Supplementary Material of "Random Matrix Theory Proves that Deep Learning Representations of GAN-data Behave as Gaussian Mixtures"

Mohamed El Amine Seddik [1 2] **Cosme Louart** [1 3] **Mohamed Tamaazousti** [1] **Romain Couillet** [2 3]

Abstract

This supplementary material provides the proof of our main results, specifically a detailed proof for Theorem 3.3 and a proof for the Proposition 3.1.

1. Proof of Theorem 3.3

1.1. Setting of the proof

For simplicity, we will only suppose the case $k = 1$ and we consider the following notations that will be used subsequently.

$$\bar{\boldsymbol{x}} = \mathbb{E}\boldsymbol{x}_i, \; \boldsymbol{C} = \mathbb{E}[\boldsymbol{x}_i\boldsymbol{x}_i^\intercal], \; \boldsymbol{X}_0 = \boldsymbol{X} - \bar{\boldsymbol{x}}\mathbf{1}_n^\intercal, \; \boldsymbol{C}_0 = \mathbb{E}[\boldsymbol{X}_0\boldsymbol{X}_0^\intercal/n].$$

Let

$$\boldsymbol{X}_{-i} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}, 0, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_n)$$

the matrix X with a vector of zeros at its ith column.

Denote the resolvents

$$\boldsymbol{R} = \left(\frac{\boldsymbol{X}^\intercal\boldsymbol{X}}{p} + z\boldsymbol{I}_n\right)^{-1}, \; \boldsymbol{Q} = \left(\frac{\boldsymbol{X}\boldsymbol{X}^\intercal}{p} + z\boldsymbol{I}_p\right)^{-1}, \; \boldsymbol{Q}_{-i} = \left(\frac{\boldsymbol{X}\boldsymbol{X}^\intercal}{p} - \frac{\boldsymbol{x}_i\boldsymbol{x}_i^\intercal}{p} + z\boldsymbol{I}_p\right)^{-1} \tag{1}$$

And let

$$\tilde{\boldsymbol{Q}} = \left(\frac{1}{c}\frac{\boldsymbol{C}}{1+\delta} + z\boldsymbol{I}_p\right)^{-1}, \tag{2}$$

where δ is the solution to the fixed point equation

$$\delta = \frac{1}{p}\operatorname{tr}\left(\boldsymbol{C}\left(\frac{1}{c}\frac{\boldsymbol{C}}{1+\delta} + z\boldsymbol{I}_p\right)^{-1}\right).$$

1.2. Basic tools

Lemma 1.1 ((Ledoux, 2005))**.** *Let $\boldsymbol{z} \in \mathcal{E}_q(1 \,|\, \mathbb{R}^p, \|\cdot\|)$ and $\boldsymbol{M} \in \mathcal{E}_q(1 \,|\, \mathbb{R}^{p\times n}, \|\cdot\|_F)$. Then, for some numerical constant $C > 0$*

- $\mathbb{E}\,\|\boldsymbol{z}\| \leq \|\mathbb{E}\boldsymbol{z}\| + C\sqrt{p}, \;\; \mathbb{E}\,\|\boldsymbol{z}\|_\infty \leq \|\mathbb{E}\boldsymbol{z}\|_\infty + C\sqrt{\log p}$.
- $\mathbb{E}\,\|\boldsymbol{M}\| \leq \|\mathbb{E}\boldsymbol{M}\| + C\sqrt{p+n}, \;\; \mathbb{E}\,\|\boldsymbol{M}\|_F \leq \|\mathbb{E}\boldsymbol{M}\|_F + C\sqrt{pn}$.

[1]CEA List, Palaiseau, France [2]Centralesupélec, Gif-sur-Yvette, France [3]University Grenoble-Alpes, Grenoble, France. Correspondence to: MEA.Seddik <mohamedelamine.seddik@cea.fr>.

Proceedings of the 37^{th} International Conference on Machine Learning, Vienna, Austria, PMLR 119, 2020.

Lemma 1.2. *Denote $\boldsymbol{Q}_{\bar{\boldsymbol{x}}} = (\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^\intercal + z\boldsymbol{I}_p)^{-1}$, we have:*

$$\boldsymbol{Q}_{\bar{\boldsymbol{x}}}\bar{\boldsymbol{x}} = \frac{\bar{\boldsymbol{x}}}{\|\bar{\boldsymbol{x}}\|^2 + z} \;\; \textit{and} \;\; \|\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\|,\; \bar{\boldsymbol{x}}\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}} = \mathcal{O}(1).$$

Moreover, if $\|\bar{\boldsymbol{x}}\| \geq \sqrt{p}$, $\|\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\| = \mathcal{O}(p^{-1/2})$.

Proof. Since $z\boldsymbol{Q}_{\bar{\boldsymbol{x}}} = \boldsymbol{I}_p - \boldsymbol{Q}_{\bar{\boldsymbol{x}}}\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^\intercal$:

$$z\boldsymbol{Q}_{\bar{\boldsymbol{x}}}\bar{\boldsymbol{x}} = \bar{\boldsymbol{x}} - \|\bar{\boldsymbol{x}}\|^2\boldsymbol{Q}_{\bar{\boldsymbol{x}}}\bar{\boldsymbol{x}},$$

and we recover the first identity of the Lemma.

And since the matrix $\boldsymbol{C}_0$ is nonnegative symmetric, we have :

$$\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}} = \left(\frac{1}{c}\frac{\boldsymbol{C}_0 + \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^\intercal}{1+\delta} + z\boldsymbol{I}_p\right)^{-1}\bar{\boldsymbol{x}} \leq \frac{c(1+\delta)\bar{\boldsymbol{x}}}{\|\bar{\boldsymbol{x}}\|^2 + zc(1+\delta)}.$$

Therefore, $\bar{\boldsymbol{x}}\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}} = \frac{c(1+\delta)\|\bar{\boldsymbol{x}}\|^2}{\|\bar{\boldsymbol{x}}\|^2+zc(1+\delta)} = \mathcal{O}(1)$ and:

$$\|\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\| = \frac{c(1+\delta)\|\bar{\boldsymbol{x}}\|}{\|\bar{\boldsymbol{x}}\|^2 + zc(1+\delta)} \leq \begin{cases} \dfrac{\|\bar{\boldsymbol{x}}\|}{z} = \mathcal{O}(1) \;\text{ if } \|\bar{\boldsymbol{x}}\| \leq 1, \\ \dfrac{c(1+\delta)}{\|\bar{\boldsymbol{x}}\|} = \mathcal{O}(1) \;\text{ if } \|\bar{\boldsymbol{x}}\| \geq 1. \end{cases}$$

□

Proposition 1.3. $\bar{\boldsymbol{x}}^\intercal\mathbb{E}[\boldsymbol{Q}]\bar{\boldsymbol{x}} = \bar{\boldsymbol{x}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}} + \mathcal{O}\left(\sqrt{\frac{\log p}{p}}\right)$

Proof. Let us bound:

$$\left|\bar{\boldsymbol{x}}^\intercal\boldsymbol{Q}\bar{\boldsymbol{x}} - \bar{\boldsymbol{x}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\right| \leq \frac{c^{-1}}{1+\delta}\left|\mathbb{E}\left[\bar{\boldsymbol{x}}\boldsymbol{Q}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\left(\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i - \delta\right)\right] + \frac{1}{p}\mathbb{E}\left[\bar{\boldsymbol{x}}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\boldsymbol{Q}\boldsymbol{C}\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\right]\right|$$

Now let us consider a supplementary random vector $\boldsymbol{x}_{n+1}$ following the same low as the $\boldsymbol{x}_i$'s and independent of $\boldsymbol{X}$. We divide the set $\mathbb{I} = [n+1]$ into two sets $\mathbb{I}_{\frac{1}{2}}$ and $\mathbb{I}_{\frac{2}{2}}$ of same cardinality ($\lfloor\frac{n+1}{2}\rfloor \leq \#\mathbb{I}_{\frac{1}{2}}, \#\mathbb{I}_{\frac{2}{2}} \leq \lceil\frac{n+1}{2}\rceil$), we note $\boldsymbol{X}_{\frac{1}{2}} = (\boldsymbol{x}_i \,|\, i \in \mathbb{I}_{\frac{1}{2}})$, $\boldsymbol{X}_{\frac{2}{2}} = (\boldsymbol{x}_i \,|\, i \in \mathbb{I}_{\frac{2}{2}})$ and we introduce the diagonal matrices $\boldsymbol{\Delta} = \mathbf{diag}\left(\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i - \delta \,|\, i \in \mathbb{I}_{\frac{1}{2}}\right)$, $\boldsymbol{D} = \mathbf{diag}\left(1 + \frac{1}{p+1}\boldsymbol{x}_i^\intercal\boldsymbol{Q}\boldsymbol{x}_i \,|\, i \in \mathbb{I}_{\frac{2}{2}}\right)$. We have the bound:

$$\begin{aligned}
&\left|\mathbb{E}\left[\bar{\boldsymbol{x}}\boldsymbol{Q}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\left(\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i - \delta\right)\right]\right| \\
&\quad= \left|\mathbb{E}\left[\left(1 + \frac{1}{p}\boldsymbol{x}_{n+1}^\intercal\boldsymbol{Q}\boldsymbol{x}_{n+1}\right)\boldsymbol{x}_{n+1}\boldsymbol{Q}_{+(n+1)}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\left(\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i - \delta\right)\right]\right| \\
&\quad= \frac{1}{p^2}\left|\mathbb{E}\left[\mathbf{1}^\intercal\boldsymbol{D}\boldsymbol{X}_{\frac{2}{2}}^\intercal\boldsymbol{Q}_{+(n+1)}\boldsymbol{X}_{\frac{1}{2}}\boldsymbol{\Delta}\boldsymbol{X}_{\frac{1}{2}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\right]\right| \\
&\quad\leq \sqrt{\left|\mathbb{E}\left[\frac{1}{p^3}\mathbf{1}^\intercal\boldsymbol{D}\boldsymbol{X}_{\frac{2}{2}}^\intercal\boldsymbol{Q}_{+(n+1)}\boldsymbol{X}_{\frac{1}{2}}\boldsymbol{\Delta}^2\boldsymbol{X}_{\frac{1}{2}}^\intercal\boldsymbol{Q}_{+(n+1)}\boldsymbol{X}_{\frac{2}{2}}\boldsymbol{D}\mathbf{1}\right]\mathbb{E}\left[\frac{1}{p}\bar{\boldsymbol{x}}^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{X}_{\frac{1}{2}}\boldsymbol{X}_{\frac{1}{2}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\right]\right|} \\
&\quad\leq \sqrt{\left|\mathbb{E}\left[\left\|\frac{1}{p}\boldsymbol{X}_{\frac{2}{2}}^\intercal\boldsymbol{Q}_{+(n+1)}\boldsymbol{X}_{\frac{1}{2}}\right\|^2\|\boldsymbol{D}\|^2\|\boldsymbol{\Delta}\|^2\right]\mathbb{E}\left[\bar{\boldsymbol{x}}\tilde{\boldsymbol{Q}}\boldsymbol{C}\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\right]\right|} \;\leq\; \mathcal{O}\left(\sqrt{\frac{\log p}{p}}\right),
\end{aligned}$$

thanks to Lemma 1.1 and Lemma 1.2 (the spectral norm of Δ and D is just an infinity norm if we see them as random vectors of $\mathbb{R}^n$). We can bound $\frac{1}{p}\left|\mathbb{E}\left[\bar{\boldsymbol{x}}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\boldsymbol{Q}\boldsymbol{C}\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\right]\right|$ the same way to obtain the result of the proposition. □

Proposition 1.4. $\|\mathbb{E}[\boldsymbol{x}_i^\intercal \boldsymbol{Q}_{-i}\boldsymbol{X}_{-i}] - \frac{\bar{\boldsymbol{x}}^\intercal \tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\boldsymbol{1}^\intercal}{1+\delta}\| = \mathcal{O}(\sqrt{\log p})$

Proof. Considering $\boldsymbol{u} \in \mathbb{R}^n$ such that $\|\boldsymbol{u}\| = 1$:

$$\begin{aligned}
&\left|\mathbb{E}[\boldsymbol{x}_i^\intercal \boldsymbol{Q}_{-i}\boldsymbol{X}_{-i}\boldsymbol{u}] - \frac{\bar{\boldsymbol{x}}^\intercal \tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\boldsymbol{1}^\intercal \boldsymbol{u}}{1+\delta}\right| \\
&\qquad = \left|\sum_{\substack{j=1\\ j\neq i}}^n \boldsymbol{u}_j \mathbb{E}\left[\frac{\boldsymbol{x}_i^\intercal \boldsymbol{Q}_{\substack{-i\\-j}}\boldsymbol{x}_j}{1+\frac{1}{p}\boldsymbol{x}_j^\intercal \boldsymbol{Q}_{\substack{-j\\-i}}\boldsymbol{x}_j} - \frac{\boldsymbol{x}_i^\intercal \tilde{\boldsymbol{Q}}\boldsymbol{x}_j}{1+\delta}\right]\right| \\
&\qquad \leq \sqrt{n}\left|\mathbb{E}\left[\frac{\boldsymbol{x}_i^\intercal \boldsymbol{Q}_{\substack{-i\\-j}}\boldsymbol{x}_j}{1+\frac{1}{p}\boldsymbol{x}_j^\intercal \boldsymbol{Q}_{\substack{-j\\-i}}\boldsymbol{x}_j} - \frac{\boldsymbol{x}_i^\intercal \boldsymbol{Q}_{\substack{-i\\-j}}\boldsymbol{x}_j}{1+\delta}\right]\right| + \left|\frac{1}{1+\delta}\mathbb{E}\left[\boldsymbol{x}_i^\intercal \boldsymbol{Q}_{\substack{-i\\-j}}\boldsymbol{x}_j - \boldsymbol{x}_i^\intercal \tilde{\boldsymbol{Q}}\boldsymbol{x}_j\right]\right| \quad (\text{where } i \neq j) \\
&\qquad \leq \sqrt{n}\left|\mathbb{E}\left[\bar{\boldsymbol{x}}^\intercal \boldsymbol{Q}\boldsymbol{x}_j\left(\frac{1}{p}\boldsymbol{x}_j^\intercal \boldsymbol{Q}_{\substack{-j\\-i}}\boldsymbol{x}_j - \delta\right)\right]\right| + \sqrt{n}\left|\mathbb{E}\left[\bar{\boldsymbol{x}}^\intercal \boldsymbol{Q}_{\substack{-i\\-j}}\bar{\boldsymbol{x}} - \bar{\boldsymbol{x}}^\intercal \tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\right]\right|,
\end{aligned}$$

where the first term is treated the same way as we did in the proof of Proposition 1.3 and the second term is bounded thanks to Proposition 1.3 □

1.3. Main body of the proof

Proof of Theorem 3.3. Recall the definition of the resolvents $\boldsymbol{R}$ and $\boldsymbol{Q}$ in Equation (1). The first step of the proof is to show the concentration of $\boldsymbol{R}$. This comes from the fact that the application $\Phi : \boldsymbol{X} \mapsto (\boldsymbol{X}^\intercal \boldsymbol{X} + z\boldsymbol{I}_n)^{-1}$ is $2z^{-3/2}$-Lipschitz *w.r.t.* the Frobenius norm. Indeed, by the matrix identity $\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A}(\boldsymbol{B}^{-1} - \boldsymbol{A}^{-1})\boldsymbol{B}$, we have

$$\Phi(\boldsymbol{X}) - \Phi(\boldsymbol{X}+\boldsymbol{H}) = \Phi(\boldsymbol{X})(\boldsymbol{H}^\intercal \boldsymbol{X} + (\boldsymbol{X}+\boldsymbol{H})^\intercal \boldsymbol{H})\Phi(\boldsymbol{X}+\boldsymbol{H})$$

And by the bounds $\|\boldsymbol{A}\boldsymbol{B}\|_F \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{B}\|_F$, $\|\Phi(\boldsymbol{X})\boldsymbol{X}^\intercal\| \leq z^{-1/2}$ and $\|\Phi(\boldsymbol{X})\| \leq z^{-1}$, we have

$$\|\Phi(\boldsymbol{X}+\boldsymbol{H}) - \Phi(\boldsymbol{X})\|_F \leq \frac{2}{z^{3/2}}\|\boldsymbol{H}\|_F.$$

Therefore, given $\boldsymbol{X} \in \mathcal{E}_q(1 \,|\, \mathbb{R}^{p\times n}, \|\cdot\|_F)$ and since the application $\boldsymbol{X} \mapsto \boldsymbol{R} = \Phi(\boldsymbol{X}/\sqrt{p})$ is $2z^{-3/2}p^{-1/2}$-Lipschitz, we have by Proposition **??** that $\boldsymbol{R} \in \mathcal{E}_q(p^{-1/2} \,|\, \mathbb{R}^{n\times n}, \|\cdot\|_F)$.

The second step consists in estimating $\mathbb{E}\boldsymbol{R}(z)$ through a deterministic matrix $\tilde{\boldsymbol{R}}$. Indeed, by the identity $(\boldsymbol{M}^\intercal \boldsymbol{M} + z\boldsymbol{I})^{-1}\boldsymbol{M}^\intercal = \boldsymbol{M}^\intercal(\boldsymbol{M}\boldsymbol{M}^\intercal + z\boldsymbol{I})^{-1}$, the resolvent $\boldsymbol{R}$ can be expressed in function of $\boldsymbol{Q}$ as follows

$$\boldsymbol{R} = \frac{1}{z}\left(I_n - \frac{\boldsymbol{X}^\intercal \boldsymbol{Q}\boldsymbol{X}}{p}\right), \tag{3}$$

thus a deterministic equivalent for $\boldsymbol{R}$ can therefore be obtained through a deterministic equivalent of the matrix $\boldsymbol{X}^\intercal \boldsymbol{Q}\boldsymbol{X}$. However, as demonstrated in (Louart & Couillet, 2019), the matrix $\boldsymbol{Q}$ has as a deterministic equivalent the matrix $\tilde{\boldsymbol{Q}}$ defined in equation 2. In the following, we aim at deriving a deterministic equivalent for $\frac{1}{p}\boldsymbol{X}^\intercal \boldsymbol{Q}\boldsymbol{X}$ in function of $\tilde{\boldsymbol{Q}}$. Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be two unitary vectors in $\mathbb{R}^n$, and let us estimate

$$\Delta \equiv \mathbb{E}\left[\boldsymbol{u}^\intercal\left(\frac{\boldsymbol{X}^\intercal \boldsymbol{Q}\boldsymbol{X}}{p} - \frac{\boldsymbol{X}^\intercal \tilde{\boldsymbol{Q}}\boldsymbol{X}}{p}\right)\boldsymbol{v}\right] = \frac{1}{p}\mathbb{E}\left[\frac{\boldsymbol{u}^\intercal \boldsymbol{X}^\intercal \boldsymbol{Q}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{X}\boldsymbol{v}}{1+\delta} - \frac{1}{p}\boldsymbol{u}^\intercal \boldsymbol{X}^\intercal \boldsymbol{Q}\boldsymbol{X}\boldsymbol{X}^\intercal \tilde{\boldsymbol{Q}}\boldsymbol{X}\boldsymbol{v}\right]$$

With the following matrix identities (to explore the independence of the columns of $\boldsymbol{X}$):

$$\boldsymbol{Q} = \boldsymbol{Q}_{-i} - \frac{1}{p}\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal \boldsymbol{Q}\,, \qquad \boldsymbol{Q}\boldsymbol{x}_i = \frac{\boldsymbol{Q}_{-i}\boldsymbol{x}_i}{1+\frac{1}{p}\boldsymbol{x}_i^\intercal \boldsymbol{Q}_{-i}\boldsymbol{x}_i}, \qquad \boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A}(\boldsymbol{B}^{-1} - \boldsymbol{A}^{-1})\boldsymbol{B}$$

and the decomposition $\boldsymbol{Q}\boldsymbol{X}\boldsymbol{X}^\intercal = \sum_{i=1}^n \boldsymbol{Q}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal$, we obtain:

$$\begin{aligned}
\Delta &= \frac{1}{p^2}\mathbb{E}\left[\sum_{i=1}^n \frac{\boldsymbol{u}^\intercal\boldsymbol{X}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{X}\boldsymbol{v}}{1+\delta} - \frac{\boldsymbol{u}^\intercal\boldsymbol{X}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{X}\boldsymbol{v}}{1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i} - \frac{1}{p}\frac{\boldsymbol{u}^\intercal\boldsymbol{X}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\boldsymbol{Q}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{X}\boldsymbol{v}}{1+\delta}\right] \\
&= \frac{1}{p^2}\sum_{i=1}^n\mathbb{E}\left[\frac{\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{X}_{-i}\boldsymbol{v}}{1+\delta} - \frac{\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{X}_{-i}\boldsymbol{v}}{1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i}\right. \\
&\qquad + \frac{\boldsymbol{u}_i\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{X}_{-i}\boldsymbol{v}}{1+\delta} + \frac{\boldsymbol{v}_i\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{x}_i}{1+\delta} + \boldsymbol{u}_i\boldsymbol{v}_i\frac{\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{x}_i}{1+\delta} \\
&\qquad - \frac{\boldsymbol{u}_i\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{X}_{-i}\boldsymbol{v}}{1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i} - \frac{\boldsymbol{v}_i\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i}{1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i} - \boldsymbol{u}_i\boldsymbol{v}_i\frac{\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i}{1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i} \\
&\qquad \left. - \frac{1}{p}\frac{\boldsymbol{u}^\intercal\boldsymbol{X}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\boldsymbol{Q}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{X}\boldsymbol{v}}{1+\delta}\right]
\end{aligned}$$

We can show with Holder's inequality and the concentration bounds (mainly the fact that $\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i$ concentrates around δ) developed in (Louart & Couillet, 2019), that most of the above quantities vanish asymptotically. As a toy example, we consider the following term:

$$\begin{aligned}
&\left|\frac{1}{p^2}\sum_{i=1}^n\mathbb{E}\left[\frac{\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{X}_{-i}\boldsymbol{v}}{1+\delta} - \frac{\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{X}_{-i}\boldsymbol{v}}{1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i}\right]\right| \\
&\quad = \left|\frac{1}{p^2}\sum_{i=1}^n\mathbb{E}\left[\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{X}_{-i}\boldsymbol{v}\frac{\delta-\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i}{(1+\delta)(1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i)}\right]\right| \\
&\quad \leq \left|\frac{1}{p^2}\sum_{i=1}^n\mathbb{E}\left[(\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i)(\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{X}_{-i}\boldsymbol{v})\left(\delta-\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\right)\right]\right| \\
&\quad \leq \left|\frac{1}{p}\sum_{i=1}^n\left(\mathbb{E}\left[\left(\frac{1}{\sqrt{p}}\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\right)^3\right]\mathbb{E}\left[\left(\frac{1}{\sqrt{p}}\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{X}_{-i}\boldsymbol{v}\right)^3\right]\mathbb{E}\left[\left(\delta-\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\right)^3\right]\right)^{\frac{1}{3}}\right| \\
&\quad = \mathcal{O}\left(\frac{1}{\sqrt{p}}\right)
\end{aligned}$$

Similarly, we can show that:

$$\begin{aligned}
&\left|\frac{1}{p^2}\sum_{i=1}^n\mathbb{E}\left[\frac{\boldsymbol{u}_i\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{X}_{-i}\boldsymbol{v}}{1+\delta} + \frac{\boldsymbol{v}_i\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{x}_i}{1+\delta}\right.\right. \\
&\qquad\qquad \left.\left. + \boldsymbol{u}_i\boldsymbol{v}_i\frac{\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{x}_i}{1+\delta} - \frac{1}{p}\frac{\boldsymbol{u}^\intercal\boldsymbol{X}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\boldsymbol{Q}\boldsymbol{C}\tilde{\boldsymbol{Q}}\boldsymbol{X}\boldsymbol{v}}{1+\delta}\right]\right| = \mathcal{O}\left(\frac{1}{\sqrt{p}}\right)
\end{aligned}$$

Finally, the remaining terms in Δ can be estimated as follows:

$$\begin{aligned}
\Delta &= \frac{1}{p^2}\sum_{i=1}^n\mathbb{E}\left[-\frac{\boldsymbol{u}_i\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{X}_{-i}\boldsymbol{v}}{1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i}\right. \\
&\qquad\qquad \left. -\frac{\boldsymbol{v}_i\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i}{1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}\boldsymbol{x}_i} - \boldsymbol{u}_i\boldsymbol{v}_i\frac{\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i}{1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i}\right] + \mathcal{O}\left(\frac{1}{\sqrt{p}}\right) \\
&= -\frac{2}{p}\frac{\delta\boldsymbol{u}^\intercal\boldsymbol{1}\bar{\boldsymbol{x}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\boldsymbol{1}^\intercal\boldsymbol{v}}{1+\delta} - \frac{\delta^2\boldsymbol{u}^\intercal\boldsymbol{v}}{1+\delta} + \mathcal{O}\left(\sqrt{\frac{\log p}{p}}\right)
\end{aligned}$$

Where the last equality is obtained through the following estimation:

$$\frac{1}{p^2}\sum_{i=1}^n \mathbb{E}\left[\frac{\boldsymbol{v}_i\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i}{1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i}\right] = \frac{1}{p}\sum_{i=1}^n \mathbb{E}\left[\frac{\boldsymbol{v}_i\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\left(\frac{1}{p}\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i(1+\delta)-\delta\left(1+\frac{1}{p}\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i\right)\right)}{\left(1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\right)(1+\delta)}\right] + \frac{1}{p}\sum_{i=1}^n\frac{\boldsymbol{v}_i\delta\mathbb{E}[\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i]}{(1+\delta)}$$

With the following bound:

$$\begin{aligned}
&\left|\frac{1}{p}\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i(1+\delta)-\delta\left(1+\frac{1}{p}\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i\right)\right| \\
&\quad= \left|\frac{1}{p}\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i(1+\delta)-\delta(1+\delta)+\delta(1+\delta)-\delta\left(1+\frac{1}{p}\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i\right)\right| \\
&\quad\leq \left|\frac{1}{p}\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i-\delta\right|(1+2\delta),
\end{aligned}$$

we have again with Holder's inequality and Proposition 1.4:

$$\frac{1}{p^2}\sum_{i=1}^n \mathbb{E}\left[\frac{\boldsymbol{v}_i\boldsymbol{u}^\intercal\boldsymbol{X}_{-i}^\intercal\boldsymbol{Q}_{-i}\boldsymbol{x}_i\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_i}{1+\frac{1}{p}\boldsymbol{x}_i^\intercal\boldsymbol{Q}\boldsymbol{x}_i}\right] = \frac{1}{p}\sum_{i=1}^n\frac{\boldsymbol{v}_i\delta\boldsymbol{u}^\intercal\boldsymbol{1}\bar{\boldsymbol{x}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}}{1+\delta} + \mathcal{O}\left(\sqrt{\frac{\log p}{p}}\right)$$

Now that we estimated Δ, it remains to estimate $\mathbb{E}[\frac{1}{p}\boldsymbol{X}^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{X}]$. Indeed, given two unit norm vectors $u, v \in \mathbb{R}^n$ we have:

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{p}\boldsymbol{u}^\intercal\boldsymbol{X}^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{X}\boldsymbol{v}\right] &= \frac{1}{p}\sum_{i,j=1}^n \boldsymbol{u}_i\boldsymbol{v}_j\mathbb{E}[\boldsymbol{x}_i^\intercal\tilde{\boldsymbol{Q}}\boldsymbol{x}_j] = \frac{1}{p}\sum_{i=1}^n\sum_{\substack{j=1\\j\neq i}}^n \boldsymbol{u}_i\boldsymbol{v}_j\bar{\boldsymbol{x}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}} + \sum_{i=1}^n \boldsymbol{u}_i\boldsymbol{v}_i\delta \\
&= \frac{1}{p}\bar{\boldsymbol{x}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\boldsymbol{u}^\intercal\boldsymbol{1}\boldsymbol{1}^\intercal\boldsymbol{v} + (\delta - \frac{1}{p}\bar{\boldsymbol{x}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}})\boldsymbol{u}^\intercal\boldsymbol{v} = \frac{1}{p}\bar{\boldsymbol{x}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\boldsymbol{u}^\intercal\boldsymbol{M}_1\boldsymbol{v}^\intercal + \delta\boldsymbol{u}^\intercal\boldsymbol{v} + \mathcal{O}\left(\frac{1}{p}\right)
\end{aligned}$$

since we have $\bar{\boldsymbol{x}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}} = \mathcal{O}(1)$ by Lemma 1.2; we introduced the matrix $\boldsymbol{M}_1 = \boldsymbol{1}\boldsymbol{1}^\intercal$. Therefore we have the following estimation:

$$\frac{1}{p}\mathbb{E}\left[\boldsymbol{X}^\intercal\boldsymbol{Q}\boldsymbol{X}\right] = \frac{\delta}{1+\delta}\boldsymbol{I}_n + \frac{1}{p}\left(\frac{1-\delta}{1+\delta}\right)\bar{\boldsymbol{x}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\boldsymbol{M}_1 + \mathcal{O}_{\|\cdot\|}\left(\sqrt{\frac{\log p}{p}}\right)$$

where $\boldsymbol{A} = \boldsymbol{B} + \mathcal{O}_{\|\cdot\|}(\alpha(p))$ means that $\|\boldsymbol{A}-\boldsymbol{B}\| = \mathcal{O}(\alpha(p))$. Finally, since $\boldsymbol{R}$ concentrates around its mean, we can then conclude:

$$\boldsymbol{R} = \frac{1}{z}\left(\boldsymbol{I}_n - \frac{1}{p}\boldsymbol{X}^\intercal\boldsymbol{Q}\boldsymbol{X}\right) = \frac{1}{z}\frac{1}{1+\delta}\boldsymbol{I}_n + \frac{\delta-1}{pz(\delta+1)}\bar{\boldsymbol{x}}^\intercal\tilde{\boldsymbol{Q}}\bar{\boldsymbol{x}}\boldsymbol{M}_1 + \mathcal{O}_{\|\cdot\|}\left(\sqrt{\frac{\log p}{p}}\right).$$

□

2. Proof of Proposition 3.1

Proof. Since the Lipschitz constant of a composition of Lipschitz functions is bounded by the product of their Lipschitz constants, we consider the case $N = 1$ and a linear activation function. In this case, the Lipschitz constant corresponds to the largest singular value of the weight matrix. We consider the following notations for the proof

$$\begin{aligned}
\bar{\boldsymbol{W}}_t &= \boldsymbol{W}_t - \eta\boldsymbol{E}_t \text{ with } [\boldsymbol{E}_t]_{i,j} \sim \mathcal{N}(0,1) \\
\boldsymbol{W}_{t+1} &= \bar{\boldsymbol{W}}_t - \max(0, \bar{\sigma}_{1,t} - \sigma_*)\,\bar{\boldsymbol{u}}_{1,t}\bar{\boldsymbol{v}}_{1,t}^\intercal
\end{aligned}$$

where $\bar{\sigma}_{1,t} = \sigma_1(\bar{\boldsymbol{W}}_t)$, $\bar{\boldsymbol{u}}_{1,t} = \boldsymbol{u}_1(\bar{\boldsymbol{W}}_t)$ and $\bar{\boldsymbol{v}}_{1,t} = \boldsymbol{v}_1(\bar{\boldsymbol{W}}_t)$. The effect of spectral normalization is observed in the case where $\sigma_* > \bar{\sigma}_{1,t}$, otherwise the Lipschitz constant is bounded by σ_*. We therefore have

$$\|\bar{\boldsymbol{W}}_t\|_F^2 \leq \|\boldsymbol{W}_t\|_F^2 + \eta^2 d_1 d_0 \tag{4}$$
$$\|\boldsymbol{W}_{t+1}\|_F^2 = \|\bar{\boldsymbol{W}}_t\|_F^2 + \sigma_*^2 - \bar{\sigma}_{1,t}^2 \tag{5}$$

- If $\|\boldsymbol{W}_{t+1}\|_F \geq \|\boldsymbol{W}_t\|_F$, we have by equation 4 and equation 5

$$\|\bar{\boldsymbol{W}}_t\|_F^2 \leq \|\bar{\boldsymbol{W}}_t\|_F^2 + \sigma_*^2 - \bar{\sigma}_{1,t}^2 + \eta^2 d_1 d_0 \quad \Rightarrow \quad \|\bar{\boldsymbol{W}}_t\| = \bar{\sigma}_{1,t} \leq \sqrt{\sigma_*^2 + \eta^2 d_1 d_0} = \delta$$

  And since $\|\boldsymbol{W}_{t+1}\| \leq \|\bar{\boldsymbol{W}}_t\|$, we have $\|\boldsymbol{W}_{t+1}\| \leq \delta$.

- Otherwise, if there exits τ such that $\|\boldsymbol{W}_{\tau+1}\|_F < \|\boldsymbol{W}_\tau\|_F$, then for all $\varepsilon > 0$ there exists an iteration $\tau' \geq \tau$ such that $\|\boldsymbol{W}_{\tau'}\| \leq \delta + \varepsilon$. Indeed, otherwise we denote $\varepsilon_t = \|\boldsymbol{W}_t\|^2 - \delta^2$ and $\varepsilon_t > 0$ for all $t \geq \tau$. And if for all $t \geq \tau$, $\|\boldsymbol{W}_{t+1}\|_F \leq \|\boldsymbol{W}_t\|_F$, we have by equation 4 and equation 5

$$\|\boldsymbol{W}_t\|_F^2 - \|\boldsymbol{W}_{t+1}\|_F^2 \geq \|\bar{\boldsymbol{W}}_t\|^2 - \delta^2 \geq \|\boldsymbol{W}_{t+1}\|^2 - \delta^2 = \varepsilon_{t+1}$$

  Integrating the above expression from τ to $T-1 \geq \tau$, we end up with

$$\|\boldsymbol{W}_\tau\|_F^2 - \|\boldsymbol{W}_T\|_F^2 \geq \sum_{t=\tau}^{T-1} \varepsilon_t \quad \Rightarrow \quad 0 \leq \|\boldsymbol{W}_T\|_F^2 \leq \|\boldsymbol{W}_\tau\|_F^2 - \sum_{t=\tau}^{T-1} \varepsilon_t,$$

  therefore, when $T \to \infty$, ε_t has to tend to 0 otherwise the right hand-side of the last inequality will tend to $-\infty$ which is absurd.

□

References

Ledoux, M. *The concentration of measure phenomenon*. Number 89. American Mathematical Soc., 2005.

Louart, C. and Couillet, R. Concentration of measure and large random matrices with an application to sample covariance matrices. *submitted*, 2019.