On the Smallest Eigenvalue of General correlated Gaussian Matrices

Abla Kammoun, *Member, IEEE* and Mohamed-Slim Alouini, *Fellow, IEEE,*

***Abstract*—This paper investigates the behaviour of the spectrum of generally correlated Gaussian random matrices whose columns are zero-mean independent vectors but have different correlations, under the specific regime where the number of their columns and that of their rows grow at infinity with the same pace. This work is, in particular, motivated by applications from statistical signal processing and wireless communications, where this kind of matrices naturally arise. Following the approach proposed in [1], we prove that under some specific conditions, the smallest singular value of generally correlated Gaussian matrices is almost surely away from zero.**

I. Introduction

Let $\mathbf{\Sigma}_n$ be a rectangular random matrix of size $N\times n$. The study of the behaviour of the asymptotic spectrum of $\mathbf{\Sigma}_n$ when $N,n \to +\infty$ has been investigated in several works. As is known, when the elements of $\mathbf{\Sigma}_n$ are zero-mean and unit variance independent and identically distributed (i.i.d.) and $\frac{N}{n} \to c < 1$, the empirical measure of the eigenvalues of $\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$ converge weakly to a deterministic probability distribution which is supported by the interval $\left[(1-\sqrt{c})^2,(1+\sqrt{c})^2\right]$ [2]. A question which immediately arises in connection with this result concerns the asymptotic behaviour of the extreme singular values. At first sight, one would expect the smallest and the largest eigenvalues of $\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$ to converge to $(1-\sqrt{c})^2$ and $(1+\sqrt{c})^2$, respectively. While this statement is correct, it cannot be directly inferred from the aforementioned weak convergence result. As a matter of fact, the proof generally requires the use of more advanced techniques improving the weak convergence result. First findings related to these issues can be traced back to the works of J. Silverstein [3] and S. Geman [4], who provided a rigorous proof showing that the extreme eigenvalues of $\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$ converge in the Gaussian case to the edges of the limiting support $(1-\sqrt{c})^2$ and $(1+\sqrt{c})^2$. This result was then extended to the case of non-Gaussian matrices but with independent and identically distributed entries [5]. The characterization of the limiting support of $\mathbf{\Sigma}_n$ is much more difficult in the case where the column entries of $\mathbf{\Sigma}_n$ are correlated. Instead of determining the exact support, many works focused on establishing the almost sure absence of eigenvalues of $\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$ in any closed interval outside the support of the limiting distribution. We can cite, for sake of illustration, the work of [6] applying for the simple-correlated case where the columns of $\mathbf{\Sigma}_n$ are correlated with the same correlation matrix and that of [1] which deals with non-centered uncorrelated models.

In many applications, this result, though limited, is essential. It can be, for instance, used to efficiently handle random quantities involving the Gram matrix $\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$ or its inverse.

In this paper, we consider the generally correlated Gaussian model in which the columns of $\mathbf{\Sigma}_n$ are zero-mean independent Gaussian random vectors but with different correlations. First results related to this model are due to Wagner et al. [7] who characterize the asymptotic behaviour of the limiting distribution of $\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$. This result was in particular applied to the analysis of the performance of the regularized-zero forcing linear precoding technique [7].

Since then, this model has known an increasing popularity, mostly spurred by applications in multi-user nultiple-input-single-output (MISO) systems [8], [9] and the very recent robust signal processing applications [10]. In what follows, we provide two different applications where the general correlation Gaussian model arises.

a) Multiple Input Single Output Channel: Consider the downlink of a single-cell system in which a base station (BS) with N antennas serves n users equipped each with a single antenna each and assume that $N<n$. The downlink channel vector $\mathbf{h}_k$ between the BS and the k th user is given by [7]:

$$\mathbf{h}_k = \mathbf{R}_k^{\frac{1}{2}}\mathbf{z}_k.$$

with $\mathbf{z}_k$ is a standard complex Gaussian vector and matrix $\mathbf{R}_k$ is essentially function of the richness of the scattering between the BS and the user of interest and as such is specific for each user. To mitigate inter-user interference, the BS precodes the transmitted signal by a matrix $\mathbf{G}$ which depends on the channel conditions for all users. Among the used precoding techniques, we can cite the Zero-forcing (ZF) precoding given by [11]:

$$\mathbf{G} = \left(\frac{1}{n}\mathbf{H}\mathbf{H}^*\right)^{-1}\mathbf{H},$$

where $\mathbf{H} = [\mathbf{h}_1,\cdots,\mathbf{h}_n]$. The ZF precoding involves the inversion of the Gram matrix $\mathbf{H}\mathbf{H}^*$, a step which becomes critical in case the smallest eigenvalue is near zero. In order to analyze the performance of using the ZF precoding, the regime under which the number of antennas N and the number of users n increase with the same pace is often assumed. The performance of the ZF precoding under this regime has been studied in [7], where it has been assumed that the smallest eigenvalue of $\frac{1}{n}\mathbf{H}\mathbf{H}^*$ is bounded away from zero for all large N and n. Although this assumption holds true for specific cases where all matrices $\mathbf{R}_k$ are equal, there is no proof supporting its validity in general. This is the reason why the authors in [7] opted to add it as an assumption, which is likely to always hold true and thus is unnecessary.

b) Robust Statistics: Consider a temporal series of n vector observations $\mathbf{y}_1,\cdots,\mathbf{y}_n$ of size $N\times 1$. Assume that the contribution of each $\mathbf{y}_i$ can be decomposed as the sum of a useful signal plus an elliptical noise, i.e,

$$\mathbf{y}_i = \mathbf{s}_i + \mathbf{x}_i, \tag{1}$$

where $\mathbf{s}_1, \cdots, \mathbf{s}_n$ are Gaussian independent $N \times 1$ random Gaussian vectors with covariance $\mathbf{R}$ and $\mathbf{x}_i$ is drawn from a Compound Gaussian distribution, i.e,

$$\mathbf{x}_i = \sqrt{\tau}_i \mathbf{z}_i, \tag{2}$$

where $\mathbf{z}_i$ are standard complex Gaussian vectors and $\tau_1, \cdots, \tau_n$ are scalar positive-valued random variables. We consider the problem of estimating the covarince matrix of $\mathbf{x}_i$. In order to mitigate the impact of the heavy-tailed distributed noise, the use of robust covariance estimates known also as robust scatter estimates has been proven to be a good solution. These are given as the unique solution of the following equation:

$$\hat{\mathbf{C}}_N = \sum_{i=1}^{n} u(\mathbf{x}_i^* \hat{\mathbf{C}}_N^{-1} \mathbf{x}_i) \mathbf{x}_i \mathbf{x}_i^*, \tag{3}$$

where $x :\mapsto u(x)$ is a scalar functional satisfying certain conditions [12]. In a recent submitted work, we prove that matrix $\hat{\mathbf{C}}_N$ converges in the operator norm to $\hat{\mathbf{S}}_N$ where $\hat{\mathbf{S}}_N$ is given by:

$$\hat{\mathbf{S}}_N = \sum_{i=1}^{n} v(\delta_i) \mathbf{x}_i \mathbf{x}_i^*, \tag{4}$$

with $\delta_1, \cdots, \delta_n$ are solutions of some fixed point equations [10]. Conditioning on τ_i, matrix $\mathbf{S}_N$ follows the model of generally correlated Gaussian matrices. The proof in [10] relies on the control of the smallest eigenvalue of $\hat{\mathbf{S}}_N$.

Despite its importance, the generally correlated Gaussian model has not been extensively explored, most probably because of its recent emergence as a major practical model. Several questions related to the behaviour of the eigenvalues remain unanswered. A major question, illustrated by the two examples above, and which triggered our motivation for this work, concerns the control of the smallest eigenvalue of the Gram matrix $\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$. Knowing that the smallest eigenvalue stay away of zero in the i.i.d case when $N < n$, one can expect the same behaviour to hold for the general Gaussian correlated case under probably some mild conditions on the correlation matrices. In this paper, we provide a rigorous proof for this statement by essentially building on the techniques developed by [1].

II. Problem statement and review of some results

All along the paper, we consider integers $n, N, \overline{N}$ such that $n \geq N$ and $\overline{N} \geq N$. We denote by c_N the ratio $\frac{N}{n}$. We make the following assumptions:

Assumption A-1.

$$0 < \liminf c_N \leq \limsup c_N < 1. \tag{5}$$

The objective of this paper is to provide some interesting properties of the spectrum of generally correlated Gaussian matrices, i.e matrices whose columns are zero-mean independent random vectors but have different covariances. Throughout this paper, matrix $\mathbf{\Sigma}_n$ represents the complex-valued $N \times n$ matrix given by:

$$\mathbf{\Sigma}_n = [\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_n], \tag{6}$$

where $\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_n$ are assumed to satisfy the following assumptions:

Assumption A-2. *$(\boldsymbol{\xi}_i)_{i=1}^n$ are zero-mean complex Gaussian vectors of size $\overline{N} \times 1$ with covariance $\mathbf{\Theta}_i$ where $(\mathbf{\Theta}_i)_{i=1}^n$ is a sequence of $N \times \overline{N}$ matrices verifying:*

$$w_{\min} \triangleq \inf_N \min_{1 \leq i \leq n} \lambda_1(\mathbf{\Omega}_i) > 0, \tag{7}$$

$$w_{\max} \triangleq \sup_N \max_{1 \leq i \leq n} \lambda_N(\mathbf{\Omega}_i) < +\infty, \tag{8}$$

where $\mathbf{\Omega}_i \triangleq \mathbf{\Theta}_i\mathbf{\Theta}_i^*$ and $\lambda_1(\mathbf{\Omega}_i)$ and $\lambda_N(\mathbf{\Omega}_i)$ are the smallest and largest eigenvalues of $\mathbf{\Omega}_i$.

We denote in what follows by $\lambda_1 \leq \cdots \leq \lambda_N$ the eigenvalues of $\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$. The empirical eigenvalue distribution of $\frac{1}{n}\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ is defined as:

$$\hat{\mu}_N = \frac{1}{N}\sum_{k=1}^{N} \delta_{\lambda_k}. \tag{9}$$

In order to characterize the asymptotic behaviour of $\hat{\mu}_N$, it is in practice quite common to analyze that of its Stieltjes transform (ST). Since the ST of a positive finite measure μ is given by:

$$\Psi_\mu(z) = \int_{\mathbb{R}} \frac{d\mu(\lambda)}{\lambda - z},$$

the ST of the empirical eigenvalue distribution in (9) can be written as:

$$\hat{m}_N(z) = \frac{1}{N}\sum_{k=1}^{N} \frac{1}{\lambda_k - z}. \tag{10}$$

Denote by $\mathbf{Q}_N(z) = \left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^* - z\mathbf{I}_N\right)^{-1}$. In the parlance of random matrix theory, $\mathbf{Q}_N(z)$ is referred to as the resolvent matrix. From (10), one can easily see that:

$$\hat{m}_N(z) = \frac{1}{N}\operatorname{tr}\mathbf{Q}_N(z). \tag{11}$$

Relation (11) clearly establishes the link between the resolvent matrix and the ST of the empirical eigenvalue distribution $\hat{\mu}_N$. It is a fundamental equation that accounts for the key role played by the resolvent matrix in the theory of random matrices. As a matter of fact, the study of the asymptotic behaviour of the resolvent matrix has provided an important load of new results concerning different statistical models [13], [14]. The model of generally correlated random matrices has recently been studied in [7], where it has been proven that the ST of the empirical eigenvalue distribution converges almost surely to a deterministic function which is the ST of some probability distribution. More formally, it is well known from [7], that it exists a sequence of deterministic measures μ_N such that $\hat{\mu}_N - \mu_N$ converges weakly to zero almost surely.

Measure μ_N is characterized through its ST $m_N(z)$ which is given by:

$$m_N(z) = \frac{1}{N}\operatorname{tr}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{\Omega}_i}{1+\delta_i(z)} - zI_N\right)^{-1},$$

where $\delta_1,\cdots,\delta_n$ form the unique solutions that are ST of non-negative finite measure of the following system of equations:

$$\delta_i(z) = \frac{1}{n}\operatorname{tr}\mathbf{\Omega}_i\left(\frac{1}{n}\sum_{j=1}^{n}\frac{\mathbf{\Omega}_j}{1+\delta_j(z)} - zI_N\right)^{-1}$$

for each $z\in\mathbb{C}\backslash\mathbb{R}^+$.

In the following, we denote by $\mathbf{T}_N$, the matrix:

$$\mathbf{T}_N(z) = \left(\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{\Omega}_i}{1+\delta_i(z)} - zI_N\right)^{-1},$$

and

$$m_N(z) = \frac{1}{N}\operatorname{tr}\mathbf{T}_N(z).$$

As $\hat{\mu}_N - \mu_N$ converge to zero weakly almost surely, we have:

$$\hat{m}_N(z) - m_N(z) \overset{a.s.}{\to} 0$$

for each $z\in\mathbb{C}\backslash\mathbb{R}^+$.

III. Main results

In this paper, we prove that under Assumptions 1-2, the smallest eigenvalue of the Gram matrix $\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$ stays away zero almost surely for N large enough. This in particular implies, that for some $\epsilon>0$, $\hat{\mu}_N\,[0,\epsilon]=0$ for N large enough. Since $\hat{\mu}_N-\mu_N$ converges weakly to zero, it is not difficult to convince oneself that one needs to start by showing that the support $\mathcal{S}_N$ of μ_N does not contain 0. In particular, we prove the following result:

Theorem 1. *Under Assumption 1 and 2, $0\notin\mathcal{S}_N$. In particular, there exists $\epsilon>0$ such that:*

$$[0,\epsilon]\cap\mathcal{S}_N = \varnothing.$$

To avoid disrupting the flow of the article, the proof of Theorem 1 is deferred to Appendix B.

Theorem 1 ensures that 0 does not belong to the support of the deterministic measure μ_N. To conclude, it suffices to supplement this result with a second one, which establishes that almost surely, there is no eigenvalue of $\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$ that goes outside the support of $\mathcal{S}_N$. This kind of result has already been shown to hold for other statistical models, by either using properties of the ST and bounds on the moments of martingale difference sequences [15]–[17] or resorting to tools based on Gaussian calculus [1]. Since we assume in this paper that $\mathbf{\Sigma}_n$ has Gaussian entries, we rather build on the method of [1] which also originates from some of the ideas of [18]. In particular, we establish the following result:

Theorem 2. *Assume that there exists a positive quantity $\epsilon>0$ and two real values $a,b\in\mathbb{R}$ such that for all N large enough:*

$$]a-\epsilon, b+\epsilon[\cap\mathcal{S}_N = \varnothing$$

Then, with probability one, no eigenvalue of $\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$ appears in $[a,b]$ for all N large enough.

Proof: The following proposition will be crucial in order to prove Theorem 2. It merely quantifies the error that we incur by replacing $\mathbb{E}\frac{1}{N}\operatorname{tr}\mathbf{Q}(z)$ by $\frac{1}{N}\operatorname{tr}\mathbf{T}_N(z)$. The proof is quite demanding and heavily relies on Gaussian calculus tools. It will be detailed in the corpus of the paper, namely in section IV, since we believe that some intermediate results be of independent interest.

Proposition 3. *$\forall z\in\mathbb{C}\backslash\mathbb{R}_+$, we have for N large enough,*

$$\mathbb{E}\left[\frac{1}{N}\operatorname{tr}\mathbf{Q}(z)\right] = \frac{1}{N}\operatorname{tr}\mathbf{T}_N(z) + \frac{1}{N^2}\chi_N(z)$$

with χ is analytic on $\mathbb{C}\backslash\mathbb{R}_+$ and satisfies:

$$|\chi_N(z)| \le K\left(|z|+C\right)^k P\left(|\Im z|^{-1}\right) \tag{12}$$

for each $z\in\mathbb{C}_+$ where C,K are constants, k is an integer independent of N and P is a polynomial with positive coefficients independent of N.

Proposition 3 will essentially serve to provide asymptotic approximates of linear statistics of the eigenvalues of the Gram matrix. In fact, with the help of proposition 3, we prove the following result:

Lemma 4. *Let ϕ be a compactly supported real-valued smooth function defined on $\mathbb{R}$, i.e, $\phi\in\mathcal{C}_c^\infty(\mathbb{R},\mathbb{R})$. Then* [1]*,*

$$\mathbb{E}\left[\phi\left(\frac{1}{N}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*\right)\right] - \int_{\mathcal{S}_N}\phi(\lambda)d\mu_N(\lambda) = \mathcal{O}\left(\frac{1}{N^2}\right). \tag{13}$$

Proof: The proof is built around the use of the inversion lemma of ST. Recall that if m is the ST of some finite measure μ, then for any continuous real function ϕ with compact support in $\mathbb{R}$

$$\int_{\mathbb{R}}\phi(\lambda)\mu(d\lambda) = \frac{1}{\pi}\Im\left(\lim_{y\downarrow 0}\int_{\mathbb{R}}\phi(x)m(x+\imath y)dx\right).$$

We therefore have:

$$\mathbb{E}\left[\frac{1}{N}\operatorname{tr}\phi(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\mathrm{H}})\right] = \frac{1}{\pi}\Im\left(\lim_{y\downarrow 0}\int_{\mathbb{R}}\phi(x)\mathbb{E}\left[\frac{1}{N}\operatorname{tr}\mathbf{Q}\left(x+\imath y\right)\right]dx\right)$$

$$\int_{\mathcal{S}_N}\phi(\lambda)d\mu_N(\lambda) = \frac{1}{\pi}\Im\left(\lim_{y\downarrow 0}\int_{\mathbb{R}}\phi(x)\mathbb{E}\left[\frac{1}{N}\operatorname{tr}\mathbf{T}_N\left(x+\imath y\right)\right]dx\right).$$

By proposition 3, we get:

$$\mathbb{E}\left[\frac{1}{N}\operatorname{tr}\phi(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\mathrm{H}})\right] - \int_{\mathcal{S}_N}\phi(\lambda)d\mu_N(\lambda)$$
$$= \frac{1}{N^2}\frac{1}{\pi}\lim_{y\downarrow 0}\Im\left[\int_{\mathbb{R}_+}\phi(x)\chi_N(x+\imath y)dx\right].$$

Since the function $\chi_N(z)$ satisfies (12), Theorem 6.2 in [19] implies that:

$$\limsup_{y\downarrow 0}\left|\int_{\mathbb{R}}\phi(x)\chi_N(x+\imath y)dx\right| \le C < +\infty.$$

[1]If $\mathbf{A} = \sum_{i=1}^N\lambda_i\mathbf{u}_i\mathbf{u}_i^{\mathrm{H}}$ is an eigenvalue decomposition of $\mathbf{A}$, then $\phi(\mathbf{A}) = \sum_{i=1}^N\phi(\lambda_i)\mathbf{u}_i\mathbf{u}_i^{\mathrm{H}}$.

where C is a constant independent of N, thereby establishing (13). ■

We return now to the proof of Theorem 2. With the above results at hand, Theorem 2 can be shown along the same lines as the proof of Theorem 3 in [1]. The details are provided in the sequel for sake of completeness. Consider $\psi \in \mathcal{C}_c^\infty(\mathbb{R},\mathbb{R})$ satisfying $0 \leq \psi \leq 1$ and:

$$\psi(\lambda) = \begin{cases} 1 & \text{for} \quad \lambda \in [a,b] \\ 0 & \text{for} \quad \lambda \in \mathbb{R}\backslash\,]a-\epsilon, b+\epsilon[\,. \end{cases}$$

For N large enough, function ψ is zero in the support $\mathcal{S}_N$. Therefore,

$$\mathbb{E}\left[\frac{1}{N}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right] = \mathcal{O}\left(\frac{1}{N^2}\right).$$

We need also to prove that the variance of $\frac{1}{N}\psi(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H})$ is of order $\frac{1}{N^4}$:

$$\operatorname{var}\left[\frac{1}{N}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right] = \mathcal{O}\left(\frac{1}{N^4}\right). \tag{14}$$

To establish (18), it suffices to resort to the Nash-Poincaré inequality which is stated in Lemma 7 of the next section. Applying Lemma 7, we obtain:

$$\operatorname{var}\left(\frac{1}{N}\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right) \leq$$
$$\sum_{k=1}^n\sum_{s=1}^N\sum_{r=1}^N \frac{\partial \frac{1}{N}\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)}{\partial \xi_{s,k}}\left[\mathbf{\Omega}_k\right]_{s,r}\frac{\left[\partial\frac{1}{N}\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right]^*}{\partial\xi_{r,k}}$$
$$+\sum_{k=1}^n\sum_{s=1}^N\sum_{r=1}^N \frac{\partial \frac{1}{N}\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)}{\partial \xi_{s,k}^*}\left[\mathbf{\Omega}_k\right]_{s,r}\frac{\left[\partial\frac{1}{N}\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right]^*}{\partial\xi_{r,k}^*}. \tag{15}$$

By Lemma 4.6 in [19], we have:

$$\frac{\partial\left[\frac{1}{N}\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right]}{\partial\xi_{s,k}} = \left[\frac{1}{Nn}\mathbf{\Sigma}_n^{\rm H}\psi'\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right]_{k,s} \tag{16}$$
$$\frac{\partial\left[\frac{1}{N}\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right]}{\partial\xi_{s,k}^*} = \left[\frac{1}{Nn}\psi'\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\mathbf{\Sigma}_n\right]_{s,k}. \tag{17}$$

Plugging (16) and (17) into (15), we get:

$$\operatorname{var}\left[\frac{1}{N}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right]$$
$$\leq \sum_{k=1}^n \frac{2}{N^2n^2}\mathbb{E}\left[\operatorname{tr}\left(\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\psi'\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\mathbf{\Omega}_k\psi'\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right)\right]$$
$$\overset{(a)}{\leq} w_{\max}\sum_{k=1}^n \frac{2}{N^2n^2}\mathbb{E}\left[\operatorname{tr}\left(\psi'\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\psi'\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right)\right],$$

where (a) follows from the fact that $\operatorname{tr}\mathbf{AB} \leq \|\mathbf{A}\|\operatorname{tr}\mathbf{B}$ for $\mathbf{A}$ hermitian and $\mathbf{B}$ positive definite matrix. Consider $h:\lambda\mapsto\lambda\left|\psi'(\lambda)\right|^2$. Clearly h belongs to $\mathcal{C}_c^\infty(\mathbb{R},\mathbb{R})$. We therefore have:

$$\mathbb{E}\left[\frac{1}{n}\operatorname{tr}\left(\psi'\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\psi'\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right)\right]$$
$$= \int_{\mathcal{S}_N} h(\lambda)d\mu_N(\lambda) + \mathcal{O}\left(\frac{1}{N^2}\right).$$

It is clear that for N large enough, $\int_{\mathcal{S}_N} h(\lambda)d\mu_N(\lambda) = 0$, thus proving:

$$\operatorname{var}\left(\frac{1}{N}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right) = \mathcal{O}\left(\frac{1}{N^2}\right).$$

Applying the classical Markov inequality, we obtain:

$$\mathbb{P}\left(\frac{1}{N}\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right) \leq N^{8/3}\mathbb{E}\left[\left|\frac{1}{N}\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right|^2\right]$$
$$= N^{8/3}\left(\left|\mathbb{E}\left[\frac{1}{N}\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right]\right|^2\right.$$
$$\left.+\operatorname{var}\left(\frac{1}{N}\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)\right)\right)$$
$$= \mathcal{O}\left(\frac{1}{N^{4/3}}\right).$$

Thus, by Borel-Cantelli lemma, for N large enough,

$$\frac{1}{N}\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right) \leq \frac{1}{N^{4/3}},$$

or equivalently,

$$\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right) \leq \frac{1}{N^{1/3}}$$

By definition of function ψ, the number of eigenvalues of the Gram matrix $\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}$ that lies in the in the interval $[a,b]$ is upper-bounded by $\operatorname{tr}\psi\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^{\rm H}\right)$, and is therefore less than $\frac{1}{N^{1/3}}$ with probability 1. Since this number has to be an integer, we deduce that it is zero for N large enough. As a consequence, there is no eigenvalue in $[a,b]$ for N large enough.

■

Gathering the results of Theorem 2 and Theorem 1, we get:

Corollary 5. *Assume the setting of Theorem 1. Then, for N large enough, the smallest eigenvalue of $\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$ is bounded away from zero.*

IV. Approximation rule

This section aims at showing the approximation in proposition 3 stating that:

$$\mathbb{E}\left[\frac{1}{N}\operatorname{tr}\mathbf{Q}(z)\right] = \frac{1}{N}\operatorname{tr}\mathbf{T}_N(z) + \frac{1}{N^2}\chi_N(z)$$

for N large enough, where χ is analytic on $\mathbb{C}\backslash\mathbb{R}_+$ and satisfies inequality (12).

As far as generally correlated Gaussian matrices are concerned, the convergence of $\frac{1}{N}\operatorname{tr}\mathbf{Q}_N(z)$ to $\frac{1}{N}\operatorname{tr}\mathbf{T}_N(z)$ has been shown to hold in the almost sure sense, [7]. This result directly implies that the empirical eigenvalue distribution converges weakly to a measure μ_N which is characterized by its stieltjes transform $m_N(z) = \frac{1}{N}\operatorname{tr}\mathbf{T}_N(z)$. Its importance lies in that it gives us insights on the proportion of eigenvalues falling in any interval. But, it does not rule out the possibility of a $o(n)$ proportion of eigenvalues lying outside the limiting support of μ_N. As it has been shown above, a sufficient condition that can eliminate this possibility is constituted by the statement of proposition 3. This statement is already known to hold for other models, mainly the non-centered Gaussian model [1]. Its proof for the model of generally correlated Gaussian matrices has not been carried out, to the best of the authors' knowledge.

While the proof of proposition 3 relies on the standard use of Gaussian calculus tools, several adaptations to the specificity of the random matrix model are far from being immediate. To facilitate the understanding of the highly technical proof, we start by introducing the main key steps. In order to control the difference $\frac{1}{N}\mathbb{E}\operatorname{tr}\mathbf{Q}_N(z) - \frac{1}{N}\operatorname{tr}\mathbf{T}_N(z)$, we need to introduce, similar to previous works [14], an intermediate deterministic matrix denoted by $\mathbf{R}_N(z)$ and which writes as:

$$\mathbf{R}_N(z) = \left(\frac{1}{n}\sum_{k=1}^{n}\frac{\mathbf{\Omega}_k}{1+\alpha_k(z)} - z\mathbf{I}_N\right)^{-1},$$

where $\alpha_k(z) = \frac{1}{n}\operatorname{tr}\mathbf{\Omega}_k\mathbb{E}\mathbf{Q}(z),\ k=1,\cdots,n$. With matrix $\mathbf{R}_N(z)$ at hand, we decompose the difference $\frac{1}{N}\mathbb{E}\operatorname{tr}\mathbf{Q}_N(z) - \frac{1}{N}\operatorname{tr}\mathbf{T}_N(z)$ as:

$$\begin{aligned}\frac{1}{N}\mathbb{E}\operatorname{tr}\mathbf{Q}_N(z) - \frac{1}{N}\operatorname{tr}\mathbf{T}_N(z) &= \frac{1}{N}\mathbb{E}\operatorname{tr}\mathbf{Q}_N(z) - \frac{1}{N}\operatorname{tr}\mathbf{R}_N(z)\\ &+ \frac{1}{N}\operatorname{tr}\mathbf{R}_N(z) - \frac{1}{N}\operatorname{tr}\mathbf{T}_N(z)\\ &\triangleq \frac{1}{N^2}\chi_1(z) + \frac{1}{N^2}\chi_2(z).\end{aligned}$$

This decomposition is quite standard in random matrix theory. While the direct control of the difference $\frac{1}{N}\mathbb{E}\operatorname{tr}\mathbf{Q}_N(z) - \frac{1}{N}\operatorname{tr}\mathbf{T}_N(z)$ is complicated, much can be inferred from both differences $\frac{1}{N}\mathbb{E}\operatorname{tr}\mathbf{Q}_N(z) - \frac{1}{N}\operatorname{tr}\mathbf{R}_N(z)$ and $\frac{1}{N}\operatorname{tr}\mathbf{R}_N(z) - \frac{1}{N}\operatorname{tr}\mathbf{T}_N(z)$. In order to prove proposition 3, it suffices to show that:

$$|\chi_i(z)| \leq (|z| + C_i)^{k_i} P_i\left(|\Im z|^{-1}\right), i = 1, 2,$$

where $C_i, i=1,2$ are positive constants, $k_i, i=1,2$ are positive integers and $P_i, i=1,2$ are polynomial with positive coefficients independent of N. In addition to $\mathbf{R}_N(z)$, we will need to introduce the following deterministic quantities:

$$\begin{aligned}\tilde{r}_i &= -\frac{1}{z(1+\alpha_i(z))}, i = 1,\cdots,n\\ \tilde{\mathbf{R}}_N &= \operatorname{diag}(\tilde{r}_1,\cdots,\tilde{r}_n).\end{aligned}$$

It can be easily shown along the same lines of Proposition 5.1 of [13] that matrix valued functions $\mathbf{R}_N(z)$ and $\tilde{\mathbf{R}}_N(z)$ are holomorphic in $\mathbb{C}\backslash\mathbb{R}_+$ and coincide with the Stieltjes transforms of positive matrix valued probability measures carried by $\mathbb{R}_+$, the mass of which are equal to $\mathbf{I}$. Their spectral norms are thus bounded by $|\Im z|^{-1}$. In particular, we have:

$$\max\left(\|\tilde{\mathbf{R}}_N\|, \|\mathbf{R}_N\|\right) \leq |\Im z|^{-1}.$$

With these quantities at hand, we are now in position to sequentially control the terms $\chi_1(z)$ and $\chi_2(z)$.

A. *Control of $\chi_1(z)$*

The control of $\chi_1(z)$ will extensively rely on the use of Gaussian calculus tools, namely the Integration by Part formulae and the Nash-Poincaré inequality. Before delving into the core of the proof, we shall recall these tools.

Lemma 6 (Integration by Part Lemma)**.** *Let $\mathbf{x} = [x_1,\cdots,x_N]^T$ a complex Gaussian vector such that $\mathbb{E}[\mathbf{x}] = 0$, $\mathbb{E}[\mathbf{x}\mathbf{x}^T] = 0$ and $\mathbb{E}[\mathbf{x}\mathbf{x}^*] = \mathbf{R}$. If $\Gamma : \mathbf{x} \mapsto \Gamma(\mathbf{x})$ is a $\mathcal{C}^1$ complex function polynomially bounded together with its derivatives, then:*

$$\mathbb{E}[x_p\Gamma(x)] = \sum_{m=1}^{N}[\mathbf{R}]_{p,m}\mathbb{E}\left[\frac{\partial\Gamma(\mathbf{x})}{\partial x_m^*}\right]$$

Lemma 7 (Nash-Poincaré Inequality)**.** *Let $\mathbf{x} = [x_1,\cdots,x_N]^T$ a complex Gaussian vector such that $\mathbb{E}[\mathbf{x}] = 0$, $\mathbb{E}[\mathbf{x}\mathbf{x}^T] = 0$ and $\mathbb{E}[\mathbf{x}\mathbf{x}^*] = \mathbf{R}$. If $\Gamma : \mathbf{x} \mapsto \Gamma(\mathbf{x})$ is a $\mathcal{C}^1$ complex function polynomially bounded together with its derivatives, then, noting $\nabla_x\Gamma = \left[\frac{\partial\Gamma}{\partial x_1},\cdots,\frac{\partial\Gamma}{\partial x_M}\right]^T$ and $\nabla_{x^*}\Gamma = \left[\frac{\partial\Gamma}{\partial x_1^*},\cdots,\frac{\partial\Gamma}{\partial x_M^*}\right]^T$,*

$$\begin{aligned}\operatorname{var}(\Gamma(x)) \leq\ & \mathbb{E}\left[\nabla_{\mathbf{x}}\Gamma(x)^T\mathbf{R}\ \ (\nabla_{\mathbf{x}}\Gamma(x))^*\right]\\ &+\mathbb{E}\left[(\nabla_{\mathbf{x}^*}\Gamma(x))^*\mathbf{R}\nabla_{\mathbf{x}^*}\Gamma(x)\right].\end{aligned}$$

Applying Lemma 7, we will thus get:

$$\begin{aligned}\operatorname{var}(\Gamma(\xi_1,\cdots,\xi_n)) \leq& \sum_{k=1}^{n}\sum_{s=1}^{N}\sum_{r=1}^{N}\mathbb{E}\left[\frac{\partial\Gamma}{\partial\xi_{s,k}}[\Omega_k]_{s,r}\frac{\partial\Gamma^*}{\partial\xi_{r,k}}\right]\\ &+\sum_{k=1}^{n}\sum_{s=1}^{N}\sum_{r=1}^{N}\mathbb{E}\left[\frac{\partial\Gamma^*}{\partial\xi_{s,k}^*}[\Omega_k]_{s,r}\frac{\partial\Gamma}{\partial\xi_{r,k}^*}\right]. \quad (18)\end{aligned}$$

The application of these tools will require us to compute differentials of the resolvent matrix with respect to the entries of $\mathbf{\Sigma}_n$. In particular, we will need in the sequel, the following differentiation formulas:

$$\begin{aligned}\frac{\partial[\mathbf{Q}]_{\ell,p}}{\partial\xi_{m,k}^*} &= -\frac{1}{n}\frac{[\mathbf{Q}\partial\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*\mathbf{Q}]_{\ell,p}}{\partial\xi_{m,k}^*}\\ &= -\frac{1}{n}[\mathbf{Q}\boldsymbol{\xi}_k e_m^{\mathrm{T}}\mathbf{Q}]_{\ell,p}\\ &= -\frac{1}{n}[\mathbf{Q}\boldsymbol{\xi}_k]_\ell[\mathbf{Q}]_{m,p}. \quad (19)\end{aligned}$$

Moreover, we also have:

$$\frac{\partial[\mathbf{Q}]_{\ell,p}}{\partial\xi_{s,k}} = -\frac{1}{n}[\mathbf{Q}]_{\ell,s}[\boldsymbol{\xi}_k^*\mathbf{Q}]_p. \quad (20)$$

The use of the integration by part lemma along with the above differential formulae will allow us to establish the following lemma:

Lemma 8. *Let $\beta_i, i=1,\cdots,n$ be given by $\beta_i = \frac{1}{n}\operatorname{tr}\mathbf{\Omega}_i\mathbf{Q}(z)$. For each $z\in\mathbb{C}_+$ and any deterministic matrix $\mathbf{A}$, it holds that:*

$$\mathbb{E}\operatorname{tr}\mathbf{A}\mathbf{Q}(z) = \operatorname{tr}\mathbf{A}\mathbf{R}(z) - z\mathbb{E}\operatorname{tr}\frac{\mathbf{A}\mathbf{Q}\mathbf{\Sigma}_n\tilde{\mathbf{R}}\mathbf{B}\mathbf{\Sigma}_n^*\mathbf{R}}{n}$$

where $\mathbf{B} = \operatorname{diag}\left(\overset{o}{\beta}_1,\cdots,\overset{o}{\beta}_n\right)$ with

$$\overset{o}{\beta}_i = \beta_i - \alpha_i.$$

Proof: From the identity:

$$\mathbf{Q}\left(\frac{1}{n}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^* - z\mathbf{I}_N\right) = \mathbf{I}_N$$

we have:

$$z\mathbb{E}\left[\mathbf{Q}\right]_{p,q} = \mathbb{E}\left[\mathbf{Q}\frac{\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*}{n}\right]_{p,q} - \delta_{p,q} \tag{21}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{n}\frac{1}{n}\mathbb{E}\left[\mathbf{Q}_{p,i}\xi_{i,j}\xi_{q,j}^*\right] - \delta_{p,q}.$$

Using the integration by parts formula in Lemma 6, we have:

$$\mathbb{E}\left[\mathbf{Q}_{p,i}\xi_{i,j}\xi_{q,j}^*\right] = \sum_{m=1}^{N}\mathbb{E}\left[\left[\mathbf{\Omega}_j\right]_{i,m}\frac{\partial\xi_{q,j}^*\left[\mathbf{Q}\right]_{p,i}}{\partial\xi_{m,j}^*}\right]$$

$$= \sum_{m=1}^{N}\left[\mathbf{\Omega}_j\right]_{i,m}\delta_{m,q}\mathbb{E}\left[\mathbf{Q}\right]_{p,i}$$

$$-\sum_{m=1}^{N}\left[\mathbf{\Omega}_j\right]_{i,m}\frac{1}{n}\mathbb{E}\left[\xi_{q,j}^*\left[\mathbf{Q}\boldsymbol{\xi}_j\right]_p\left[\mathbf{Q}\right]_{m,i}\right].$$

Summing the above equality over i, we obtain:

$$\mathbb{E}\left[\left[\mathbf{Q}\boldsymbol{\xi}_j\right]_p\xi_{q,j}^*\right] = \mathbb{E}\left[\mathbf{Q}\mathbf{\Omega}_j\right]_{p,q} - \mathbb{E}\left[\beta_j\left[\mathbf{Q}\boldsymbol{\xi}_j\right]_p\xi_{q,j}^*\right]$$

Plugging $\overset{o}{\beta}_j = \beta_j - \alpha_j$ into the above equality, we get:

$$\mathbb{E}\left[\left[\mathbf{Q}\boldsymbol{\xi}_j\right]_p\xi_{q,j}^*\right] = \mathbb{E}\left[\mathbf{Q}\mathbf{\Omega}_j\right]_{p,q} - \alpha_j\mathbb{E}\left[\xi_{q,j}^*\left[\mathbf{Q}\boldsymbol{\xi}_j\right]_p\right]$$

$$-\mathbb{E}\left[\overset{o}{\beta}_j\,\xi_{q,j}^*\left[\mathbf{Q}\boldsymbol{\xi}_j\right]_p\right]$$

Hence:

$$\mathbb{E}\left[\left[\mathbf{Q}\boldsymbol{\xi}_j\right]_p\xi_{q,j}^*\right] = \mathbb{E}\left[\frac{\left[\mathbf{Q}\mathbf{\Omega}_j\right]_{p,q}}{(1+\alpha_j)}\right] - \mathbb{E}\left[\frac{\overset{o}{\beta}_j\,\xi_{q,j}^*\left[\mathbf{Q}\boldsymbol{\xi}_j\right]_p}{(1+\alpha_j)}\right]$$

Summing over j, we finally get:

$$\mathbb{E}\left[\frac{\mathbf{Q}\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*}{n}\right]_{p,q} = \mathbb{E}\left[\mathbf{Q}\frac{1}{n}\sum_{j=1}^{n}\frac{\mathbf{\Omega}_j}{(1+\alpha_j)}\right]_{p,q}$$

$$+z\mathbb{E}\left[\frac{\mathbf{Q}\mathbf{\Sigma}_n\tilde{\mathbf{R}}\mathbf{B}\mathbf{\Sigma}_n^*}{n}\right]_{p,q}$$

Plugging the above equality into (21), we thus get:

$$\mathbb{E}\left[z\mathbf{Q}\right]_{p,q} = \mathbb{E}\left[\mathbf{Q}\frac{1}{n}\sum_{j=1}^{n}\frac{\mathbf{\Omega}_j}{(1+\alpha_j)}\right]_{p,q} - \left[\mathbf{I}_N\right]_{p,q}$$

$$+z\mathbb{E}\left[\frac{\mathbf{Q}\mathbf{\Sigma}_n\tilde{\mathbf{R}}\mathbf{B}\mathbf{\Sigma}_n^*}{n}\right]_{p,q}$$

Therefore,

$$\mathbb{E}\left[\mathbf{Q}\mathbf{R}^{-1}\right]_{p,q} = \left[\mathbf{I}_N\right]_{p,q} - z\mathbb{E}\left[\frac{\mathbf{Q}\mathbf{\Sigma}_n\tilde{\mathbf{R}}\mathbf{B}\mathbf{\Sigma}_n^*}{n}\right]_{p,q}$$

thereby proving that:

$$\mathbb{E}\mathbf{Q}\mathbf{R}^{-1} = \mathbf{I}_N - z\mathbb{E}\left[\frac{\mathbf{Q}\mathbf{\Sigma}_n\tilde{\mathbf{R}}\mathbf{B}\mathbf{\Sigma}_n^*}{n}\right].$$

As a consequence:

$$\mathbb{E}\,\mathrm{tr}\,\mathbf{A}\mathbf{Q} = \mathrm{tr}\,\mathbf{A}\mathbf{R} - \frac{z}{n}\,\mathrm{tr}\,\mathbb{E}\left[\mathbf{A}\mathbf{Q}\mathbf{\Sigma}_n\tilde{\mathbf{R}}\mathbf{B}\mathbf{\Sigma}_n^*\mathbf{R}\right].$$

■

From Lemma 8, it appears that the control of χ_1 amounts to showing that:

$$z\Gamma \triangleq z\mathbb{E}\left[\mathrm{tr}\,\mathbf{Q}\mathbf{\Sigma}_n\tilde{\mathbf{R}}\mathbf{B}\mathbf{\Sigma}_n^*\mathbf{R}\right] \leq \frac{1}{n}\left(|z|+C_1\right)^{k_1}P_1\left(|\Im z|^{-1}\right)$$

with C_1, k_1 and P_1 verifying the conditions of proposition 3. The proof relies on the use of the Nash-poincaré inequality. But before that, we need to further workout quantity Γ by means of the Integration by Part formula. We first expand Γ as:

$$\Gamma = \frac{1}{n}\sum_{p,q,m=1}^{N}\sum_{\ell=1}^{n}\mathbb{E}\left[\left[\mathbf{Q}\right]_{p,q}\xi_{q,\ell}\xi_{m,\ell}^*\,\overset{o}{\beta}_\ell\right]\left[\mathbf{R}\right]_{m,p}\tilde{r}_\ell \tag{22}$$

Using the integration by part formula, we have:

$$\mathbb{E}\left[\left[\mathbf{Q}\right]_{p,q}\xi_{q,\ell}\xi_{m,\ell}^*\,\overset{o}{\beta}_\ell\right] = \sum_{s=1}^{N}\left[\mathbf{\Omega}_\ell\right]_{q,s}\mathbb{E}\left[\frac{\partial\left[\mathbf{Q}\right]_{p,q}\xi_{m,\ell}^*\,\overset{o}{\beta}_\ell}{\partial\xi_{s,\ell}^*}\right]$$

$$= \sum_{s=1}^{N}\left[\mathbf{\Omega}_\ell\right]_{q,s}\mathbb{E}\left[\left[\mathbf{Q}\right]_{p,q}\overset{o}{\beta}_\ell\right]\delta_{m,s} + \sum_{s=1}^{N}\left[\mathbf{\Omega}_\ell\right]_{q,s}\mathbb{E}\left[\left[\mathbf{Q}\right]_{p,q}\xi_{m,\ell}^*\frac{\partial\,\overset{o}{\beta}_\ell}{\partial\xi_{s,\ell}^*}\right]$$

$$-\left[\mathbf{\Omega}_\ell\right]_{q,s}\frac{1}{n}\mathbb{E}\left[\left[\mathbf{Q}\boldsymbol{\xi}_\ell\right]_p\left[\mathbf{Q}\right]_{s,q}\xi_{m,\ell}^*\,\overset{o}{\beta}_\ell\right]$$

$$= -\frac{1}{n}\mathbb{E}\left[\left[\mathbf{\Omega}_\ell\mathbf{Q}\right]_{q,q}\left[\mathbf{Q}\boldsymbol{\xi}_\ell\right]_p\xi_{m,\ell}^*\,\overset{o}{\beta}_\ell\right] + \left[\mathbf{\Omega}_\ell\right]_{q,m}\mathbb{E}\left[\left[\mathbf{Q}\right]_{p,q}\overset{o}{\beta}_\ell\right]$$

$$+\sum_{s=1}^{N}\left[\mathbf{\Omega}_\ell\right]_{q,s}\mathbb{E}\left[\left[\mathbf{Q}\right]_{p,q}\xi_{m,\ell}^*\frac{\partial\,\overset{o}{\beta}_\ell}{\partial\xi_{s,\ell}^*}\right]$$

Summing the above equation over q, we get:

$$\mathbb{E}\left[\left[\mathbf{Q}\boldsymbol{\xi}_\ell\right]_p\xi_{m,\ell}^*\,\overset{o}{\beta}_\ell\right] = -\mathbb{E}\left[\frac{1}{n}\mathrm{tr}\left(\mathbf{\Omega}_\ell\mathbf{Q}\right)\left[\mathbf{Q}\boldsymbol{\xi}_\ell\right]_p\xi_{m,\ell}^*\,\overset{o}{\beta}_\ell\right]$$

$$+\mathbb{E}\left[\left[\mathbf{Q}\mathbf{\Omega}_\ell\right]_{p,m}\overset{o}{\beta}_\ell\right] + \sum_{q=1}^{N}\sum_{s=1}^{N}\left[\mathbf{\Omega}_\ell\right]_{q,s}\mathbb{E}\left[\left[\mathbf{Q}\right]_{p,q}\xi_{m,\ell}^*\frac{\partial\,\overset{o}{\beta}_\ell}{\partial\xi_{s,\ell}^*}\right]$$

Writing $\frac{1}{n}\,\mathrm{tr}\,\mathbf{\Omega}_\ell\mathbf{Q}$ as $\overset{o}{\beta}_\ell + \alpha_\ell$ and using the same technique as in the proof of Lemma 8, we finally get:

$$\mathbb{E}\left[\left[\mathbf{Q}\boldsymbol{\xi}_\ell\right]_p\xi_{m,\ell}^*\,\overset{o}{\beta}_\ell\right] = z\tilde{r}_\ell\mathbb{E}\left[\left(\overset{o}{\beta}_\ell\right)^2\left[\mathbf{Q}\boldsymbol{\xi}_\ell\right]_p\xi_{m,\ell}^*\right]$$

$$-z\tilde{r}_\ell\mathbb{E}\left[\left[\mathbf{Q}\mathbf{\Omega}_\ell\right]_{p,m}\overset{o}{\beta}_\ell\right] - \sum_{s,q=1}^{N}z\tilde{r}_\ell\left[\mathbf{\Omega}_\ell\right]_{q,s}\mathbb{E}\left[\left[\mathbf{Q}\right]_{p,q}\xi_{m,\ell}^*\frac{\partial\,\overset{o}{\beta}_\ell}{\partial\xi_{s,\ell}^*}\right] \tag{23}$$

Plugging (23) into (22), we finally obtain:

$$\Gamma = \frac{z}{n}\mathbb{E}\left[\operatorname{tr}\left(\mathbf{Q}\boldsymbol{\Sigma}_n\tilde{\mathbf{R}}^2\mathbf{B}^2\boldsymbol{\Sigma}_n^*\mathbf{R}\right)\right] - \frac{z}{n}\sum_{\ell=1}^{n}\mathbb{E}\left[\overset{o}{\beta}_\ell \operatorname{tr}\left(\mathbf{Q}\boldsymbol{\Omega}_\ell\mathbf{R}\tilde{\mathbf{R}}^2\right)\right]$$
$$-\frac{z}{n}\sum_{\ell=1}^{n}\sum_{s=1}^{N}\tilde{r}_\ell^2\mathbb{E}\left[\left[\boldsymbol{\Sigma}_n^*\mathbf{R}\mathbf{Q}\boldsymbol{\Omega}_\ell\right]_{\ell,s}\frac{\partial\overset{o}{\beta}_\ell}{\partial\xi_{s,\ell}^*}\right]$$
$$\triangleq \Delta_1 - \Delta_2 - \Delta_3.$$

In the following we will prove that Δ_i satisfies:

$$\Delta_i \le \frac{K_i}{n}\left(|z|+\tilde{C}_i\right)^{\tilde{k}_i}\tilde{P}_i(|\Im z|^{-1})$$

for some positive constant $\tilde{C}_i, K_i$, integer k_i and polynomial $\tilde{P}_i$ independent of N. This will be sufficient to control $\chi_1(z)$ since the underlying polynomials have positive coefficients. Closer scrutiny of the expressions of $\Delta_i, i=1,2,3$, reveals that they make appear quantities of the form $\frac{1}{n}\operatorname{tr}\mathbf{A}\mathbf{Q}(z)$ with $\mathbf{A}$ is a some deterministic matrix. It is thus easy to convince oneself that controlling the variance of these terms is essential. This will be the goal of the following lemma whose proof is deferred to Appendix C:

Lemma 9. *Let* $\mathbf{A}$ *be a* $N\times N$ *deterministic matrix. Then, we have for any* $z\in\mathbb{C}_+$,

$$\operatorname{var}\left(\frac{1}{n}\operatorname{tr}\mathbf{A}\mathbf{Q}(z)\right) \le \frac{C}{n^2}\|\mathbf{A}\|^2(|z|+1)\left(\frac{1}{|\Im z|^4}+\frac{1}{|\Im z|^3}\right)$$

where C*, a positive constant and* P*, a polynomial with positive coefficients, are independent of* N.

With Lemma 9 at hand, we are now in position to handle the terms $\Delta_i, i=1,2,3$. We start by controlling Δ_1. For that, consider $\boldsymbol{\Sigma}_{(i)}$ to be the matrix $\boldsymbol{\Sigma}_n$ without its i-th column. Define $\mathbf{Q}_{(i)}$ the resolvent matrix given by:

$$\mathbf{Q}_{(i)} = \left(\frac{1}{n}\boldsymbol{\Sigma}_{(i)}\boldsymbol{\Sigma}_{(i)}^* - z\mathbf{I}_N\right)^{-1}$$

and $\beta_{i,(i)} = \frac{1}{n}\operatorname{tr}\boldsymbol{\Omega}_i\mathbf{Q}_{(i)}$. Let $\overset{o}{\beta}_{i,(i)} = \beta_{i,(i)} - \mathbb{E}\beta_{i,(i)}$ and $\mathbf{B}_{(i)} = \operatorname{diag}\left(\overset{o}{\beta}_{1,(1)},\cdots,\overset{o}{\beta}_{n,(n)}\right)$. From the rank-one perturbation Lemma [20, Lemma 2.6], we obtain:

$$\max_{1\le i\le n}\left|\overset{o}{\beta}_i - \overset{o}{\beta}_{i,(i)}\right| \le \frac{2w_{\max}}{n\,|\Im z|}$$

Decompose Δ_1 as:

$$\Delta_1 = \frac{z}{n}\sum_{i=1}^{n}\mathbb{E}\left[\left(\left|\overset{o}{\beta}_i\right|^2 - \left|\overset{o}{\beta}_{i,(i)}\right|^2\right)\left[\boldsymbol{\Sigma}_n^*\mathbf{R}_N\mathbf{Q}\boldsymbol{\Sigma}_n\tilde{\mathbf{R}}_N^2\right]_{i,i}\right]$$
$$+\frac{z}{n}\sum_{i=1}^{n}\mathbb{E}\left[\left|\overset{o}{\beta}_{i,(i)}\right|^2\left[\boldsymbol{\Sigma}_n^*\mathbf{R}_N\mathbf{Q}\boldsymbol{\Sigma}_n\tilde{\mathbf{R}}_N^2\right]_{i,i}\right]$$
$$\triangleq \Delta_{1,1} + \Delta_{1,2}.$$

We start by dealing with $\Delta_{1,1}$. First, we need to bound the quantity $\left|\overset{o}{\beta}_i\right|^2 - \left|\overset{o}{\beta}_{i,(i)}\right|^2$. We have:

$$\begin{aligned}\left|\overset{o}{\beta}_i\right|^2 - \left|\overset{o}{\beta}_{i,(i)}\right|^2 &= \left(\left|\overset{o}{\beta}_i\right| - \left|\overset{o}{\beta}_{i,(i)}\right|\right)\left(\left|\overset{o}{\beta}_i\right| + \left|\overset{o}{\beta}_{i,(i)}\right|\right)\\ &\le \frac{2Nw_{\max}}{n\,|\Im z|}\left|\overset{o}{\beta}_i - \overset{o}{\beta}_{i,(i)}\right|\\ &\le \frac{4Nw_{\max}^2}{n^2\,|\Im z|^2}. \end{aligned} \tag{24}$$

From (24), $\Delta_{1,1}$ can be bounded by:

$$\Delta_{1,1} \le \frac{|z|}{n^3}\frac{4Nw_{\max}^2}{|\Im z|^2}\sum_{i=1}^{n}\mathbb{E}\left|\left[\boldsymbol{\Sigma}_n^*\mathbf{R}_N\mathbf{Q}\boldsymbol{\Sigma}_n\tilde{\mathbf{R}}_N^2\right]_{i,i}\right|.$$

We need thus to bound $\mathbb{E}\left|\left[\frac{1}{n}\boldsymbol{\Sigma}_n^*\mathbf{R}_N\mathbf{Q}\boldsymbol{\Sigma}_n\tilde{\mathbf{R}}_N^2\right]_{i,i}\right|$. We have:

$$\begin{aligned}\mathbb{E}\left|\left[\frac{1}{n}\boldsymbol{\Sigma}_n^*\mathbf{R}_N\mathbf{Q}\boldsymbol{\Sigma}_n\tilde{\mathbf{R}}_N^2\right]_{i,i}\right| &= \mathbb{E}\left[\frac{1}{n}\boldsymbol{\xi}_i^*\mathbf{R}_N\mathbf{Q}\boldsymbol{\xi}_i\tilde{r}_i^2\right]\\ &\le |\tilde{r}|_i^2\mathbb{E}\left[\|\mathbf{R}_N\mathbf{Q}\|\frac{1}{n}\boldsymbol{\xi}_i^*\boldsymbol{\xi}_i\right]\\ &\le \frac{1}{|\Im z|^4}\frac{1}{n}\operatorname{tr}\boldsymbol{\Omega}_i\\ &\le \frac{Nw_{\max}}{n|\Im z|^4}\end{aligned}$$

and thus:

$$\Delta_{1,1} \le 4|z|\left(\limsup_N \frac{N}{n}\right)^2\frac{w_{\max}^3}{n|\Im z|^6}.$$

We now move to the control of $\Delta_{1,2}$. First, write $\Delta_{1,2}$ as:

$$\Delta_{1,2} = \frac{z}{n}\sum_{i=1}^{n}\mathbb{E}\left[\left|\overset{o}{\beta}_{i,(i)}\right|^2\left[\boldsymbol{\xi}_i^*\mathbf{R}\mathbf{Q}\boldsymbol{\xi}_i\tilde{r}_i^2\right]\right].$$

Using the relation

$$\mathbf{Q}\boldsymbol{\xi}_i = \frac{\mathbf{Q}_{(i)}\boldsymbol{\xi}_i^*}{1+\frac{1}{n}\boldsymbol{\xi}_i^*\mathbf{Q}\boldsymbol{\xi}_i^*}, \tag{25}$$

we obtain:

$$\begin{aligned}\Delta_{1,2} &\le \frac{|z|}{n}\sum_{i=1}^{n}\mathbb{E}\left[\left|\overset{o}{\beta}_{i,(i)}\right|^2\frac{\left|\boldsymbol{\xi}_i^*\mathbf{R}_N\mathbf{Q}_{(i)}\boldsymbol{\xi}_i\tilde{r}_i^2\right|}{1+\frac{1}{n}\boldsymbol{\xi}_i^*\mathbf{Q}\boldsymbol{\xi}_i}\right]\\ &\le \frac{|z|}{n|\Im z|^4}\sum_{i=1}^{n}\mathbb{E}\left[\left|\overset{o}{\beta}_{i,(i)}\right|^2\boldsymbol{\xi}_i^*\boldsymbol{\xi}_i\right].\end{aligned}$$

Since $\beta_{i,(i)}$ is independent of $\boldsymbol{\xi}_i$, and thus :

$$\begin{aligned}\Delta_{1,2} &\le \frac{|z|}{n|\Im z|^4}\sum_{i=1}^{n}\operatorname{tr}\boldsymbol{\Omega}_i\mathbb{E}|\ \overset{o}{\beta}_{i,(i)}\ |^2\\ &\le \frac{Nw_{\max}|z|}{n|\Im z|^4}\sum_{i=1}^{n}\mathbb{E}|\ \overset{o}{\beta}_{i,(i)}\ |^2\end{aligned}$$

From Lemma 9, we have:

$$\mathbb{E}\left|\overset{o}{\beta}_{i,(i)}\right|^2 \le \frac{2w_{\max}^3}{n^2}(|z|+1)\left(\frac{1}{|\Im z|^4}+\frac{1}{|\Im z|^3}\right)$$

Hence,

$$\Delta_{1,2} \leq \limsup \frac{N}{n} \frac{2w_{\max}^4}{n|\Im z|^4}(|z|+1)^2 \left(\frac{1}{|\Im z|^4}+\frac{1}{|\Im z|^3}\right)$$
$$\triangleq \frac{K}{n}(|z|+1)^2 P(|\Im z|^{-1}),$$

thereby proving the desired result. The control of Δ_2 relies on the use of the Cauchy-schwartz inequality. We have:

$$\Delta_2 = \frac{z}{n}\sum_{\ell=1}^{n}\mathbb{E}\left[\overset{o}{\beta}_\ell \operatorname{tr}\left(\mathbf{Q}\mathbf{\Omega}_\ell\mathbf{R}_N\tilde{\mathbf{R}}_N^2\right)\right]$$
$$\leq |z|\sum_{\ell=1}^{n}\sqrt{\mathbb{E}|\overset{o}{\beta}_\ell|^2}\sqrt{\operatorname{var}\frac{1}{n}\operatorname{tr}\mathbf{Q}\mathbf{\Omega}_\ell\mathbf{R}_N\tilde{\mathbf{R}}_N^2}$$

From Lemma 9, we can bound $\mathbb{E}\left|\overset{o}{\beta}_\ell\right|^2$ and $\operatorname{var}\operatorname{tr}\mathbf{Q}\mathbf{\Omega}_\ell\mathbf{R}_N\tilde{\mathbf{R}}_N^2$ as:

$$\mathbb{E}\left|\overset{o}{\beta}_\ell\right|^2 \leq \frac{2w_{\max}^3}{n^2}(|z|+1)\left(\frac{1}{|\Im z|^4}+\frac{1}{|\Im z|^3}\right)$$
$$\operatorname{var}\operatorname{tr}\frac{1}{n}\mathbf{Q}\mathbf{\Omega}_\ell\mathbf{R}_N\tilde{\mathbf{R}}_N^2 \leq \frac{2w_{\max}^3}{|\Im z|^6 n^2}(|z|+1)\left(\frac{1}{|\Im z|^4}+\frac{1}{|\Im z|^3}\right).$$

Using the fact that $\sqrt{xy}\leq\frac{x+y}{2}$ for positive scalars x,y, we finally get:

$$|\Delta_2| \leq \frac{2w_{\max}^3(|z|+1)^2}{n}\left(\frac{1}{|\Im z|^4}+\frac{1}{|\Im z|^3}+\frac{1}{|\Im z|^{10}}+\frac{1}{|\Im z|^9}\right)$$
$$\triangleq K_2(|z|+1)^2 P_2(|\Im z|^{-1})$$

Finally, we will move to the treatment of Δ_3. Recall that Δ_3 is given by:

$$\Delta_3 = \frac{z}{n}\sum_{\ell=1}^{n}\sum_{s=1}^{N}\tilde{r}_\ell^2\mathbb{E}\left[\left[\mathbf{\Sigma}_n^*\mathbf{R}_N\mathbf{Q}\mathbf{\Omega}_\ell\right]_{\ell,s}\frac{\partial\overset{o}{\beta}_\ell}{\partial\xi_{s,\ell}^*}\right].$$

Using the differentiation formulae in (19), we get:

$$\frac{\partial\overset{o}{\beta}_\ell}{\partial\xi_{s,\ell}^*} = -\frac{1}{n^2}\left[\mathbf{Q}\mathbf{\Omega}_\ell\mathbf{Q}\mathbf{\Sigma}_n\right]_{s,\ell}.$$

Hence,

$$\Delta_3 = -\frac{z}{n^3}\sum_{\ell=1}^{n}\sum_{s=1}^{N}\tilde{r}_\ell^2\mathbb{E}\left[\left[\mathbf{\Sigma}_n^*\mathbf{R}_N\mathbf{Q}\mathbf{\Omega}_\ell\right]_{\ell,s}\left[\mathbf{Q}\mathbf{\Omega}_\ell\mathbf{Q}\mathbf{\Sigma}_n\right]_{s,\ell}\right]$$
$$= -\frac{z}{n^3}\sum_{\ell=1}^{n}\tilde{r}_\ell^2\mathbb{E}\left[\boldsymbol{\xi}_\ell^*\mathbf{R}_N\mathbf{Q}\mathbf{\Omega}_\ell\mathbf{Q}\mathbf{\Omega}_\ell\mathbf{Q}\boldsymbol{\xi}_\ell\right].$$

The above relation allows us to bound Δ_3 as:

$$|\Delta_3| \leq \frac{|z|w_{\max}^2}{n^3}\sum_{\ell=1}^{n}|\tilde{r}_\ell|^2\|\mathbf{R}_N\|\mathbb{E}\left[\boldsymbol{\xi}_\ell^*\boldsymbol{\xi}_\ell\|\mathbf{Q}\|^3\right]$$
$$\leq \frac{|z|w_{\max}^3}{n|\Im z|^6}\limsup\frac{N}{n}$$
$$\triangleq \frac{K_3|z|}{n}P_3(|\Im z|^{-1}).$$

From the obtained bounds for the scalars $\Delta_i, i=1,2,3$, we can deduce that:

$$|z\Gamma| \leq \frac{1}{n}\left(|z|+C_1\right)^{k_1}P_1(|\Im z|^{-1}),$$

which is, as mentioned above, the required inequality to control χ_1.

B. Control of $\chi_2(z)$

We now move to the control of $\chi_2(z)$ given by:

$$\chi_2(z) = N\operatorname{tr}\mathbf{R}_N - N\operatorname{tr}\mathbf{T}_N.$$

To this end, we will resort to the resolvent identity : $\mathbf{A}^{-1}-\mathbf{B}^{-1}=\mathbf{B}^{-1}\left(\mathbf{A}-\mathbf{B}\right)\mathbf{A}^{-1}$ for any invertible matrices $\mathbf{B}$ and $\mathbf{A}$. We therefore obtain:

$$N\operatorname{tr}\mathbf{R}_N - N\operatorname{tr}\mathbf{T}_N = \frac{N}{n}\operatorname{tr}\mathbf{R}_N\left(\sum_{j=1}^{n}\frac{\mathbf{\Omega}_j}{1+\delta_j}-\frac{\mathbf{\Omega}_j}{1+\alpha_j}\right)\mathbf{T}$$
$$= \frac{N}{n}\sum_{j=1}^{n}\frac{\operatorname{tr}(\mathbf{R}_N\mathbf{\Omega}_j\mathbf{T})(\alpha_j-\delta_j)}{(1+\alpha_j)(1+\delta_j)}$$
$$= \frac{N}{n}\sum_{j=1}^{n}z^2\tilde{r}_j\tilde{\delta}_j\operatorname{tr}\mathbf{R}_N\mathbf{\Omega}_j\mathbf{T}(\alpha_j-\delta_j),$$

where $\tilde{\delta}_j = -\frac{1}{z(1+\delta_j)}$. Using property 6 of Lemma 1 in [1], we can easily check that $\tilde{\delta}_j, j=1,\cdots,n$ similar to $\tilde{r}_j$ are Stieltjes transforms of probability measures carried by $\mathbb{R}_+$. We therefore have:

$$\max\left(\left|\tilde{\delta}_j\right|,|\tilde{r}_j|\right) \leq \frac{1}{|\Im z|}.$$

Hence,

$$|N\operatorname{tr}\mathbf{R}_N - N\operatorname{tr}\mathbf{T}_N| \leq \frac{|z|^2N^2}{|\Im z|^4}\max_{1\leq j\leq n}|\alpha_j-\delta_j|.$$

To control χ_2, it suffices to show that there exists constants C and K, integer k and polynomial P with positive coefficients and independent of N such that:

$$\max_{1\leq j\leq n}|\alpha_j-\delta_j| \leq \frac{K}{N^2}(|z|+C)^k P(|\Im z|^{-1}).$$

This will be the objective of the next derivations in this section.

We start by decomposing $\alpha_j-\delta_j$ as:

$$\alpha_j-\delta_j = \frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbb{E}\mathbf{Q}-\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{R}+\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{R}-\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{T}$$
$$= \epsilon_j(z)+\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{R}-\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{T}.$$

The control of $\epsilon_j(z)$ is similar to that of $\chi_1(z)$, the presence of matrix $\mathbf{\Omega}_j$ instead of the identity matrix requiring only slight modifications of the proof. We can thus deduce that:

$$\max_{1\leq j\leq n}|\epsilon_j| \leq \frac{K_\epsilon}{N^2}(|z|+C_\epsilon)^{k_\epsilon}P_\epsilon(|\Im z|^{-1}), \tag{26}$$

for some constants K_ϵ and C_ϵ, integer k_ϵ and polynomial P_ϵ independent of N. Again, using the resolvent identity as above, we obtain:

$$\alpha_j-\delta_j = \epsilon_j(z)+\frac{1}{n^2}\sum_{k=1}^{n}\frac{(\alpha_k-\delta_k)\operatorname{tr}\mathbf{\Omega}_j\mathbf{R}_N\mathbf{\Omega}_k\mathbf{T}}{(1+\alpha_k)(1+\delta_k)}. \tag{27}$$

Define $\boldsymbol{\alpha}=[\alpha_1,\cdots,\alpha_n]^{\mathrm{T}}$, $\boldsymbol{\delta}=[\delta_1,\cdots,\delta_n]^{\mathrm{T}}$ and $\boldsymbol{\epsilon}=[\epsilon_1(z),\cdots,\epsilon_n(z)]$. Then (27) writes as:

$$\left(\mathbf{I}_n-\mathbf{A}\right)\left(\boldsymbol{\alpha}-\boldsymbol{\delta}\right)=\boldsymbol{\epsilon}, \tag{28}$$

where $\mathbf{A}$ is a $n\times n$ matrix with entries:

$$[\mathbf{A}]_{j,k} = \frac{1}{n^2}\frac{\operatorname{tr}\mathbf{\Omega}_j\mathbf{R}_N\mathbf{\Omega}_k\mathbf{T}}{(1+\alpha_k)(1+\delta_k)}.$$

In order to control the difference vector $\boldsymbol{\alpha}-\boldsymbol{\delta}$, we need first to check that $\mathbf{I}_n-\mathbf{A}$ is invertible. For that, notice that by Cauchy-Schwartz inequality:

$$\left|[\mathbf{A}]_{j,k}\right| \leq \sqrt{\left|[\mathbf{B}]_{j,k}\right|}\sqrt{\left|[\mathbf{C}]_{j,k}\right|}$$

where $\mathbf{B}$ and $\mathbf{C}$ are $n\times n$ matrices with entries:

$$[\mathbf{B}]_{j,k} = \frac{1}{n^2}\frac{\operatorname{tr}\boldsymbol{\Omega}_j\mathbf{R}_N\boldsymbol{\Omega}_k\mathbf{R}_N}{|1+\alpha_k|^2}$$

$$[\mathbf{C}]_{j,k} = \frac{1}{n^2}\frac{\operatorname{tr}\boldsymbol{\Omega}_j\mathbf{T}\boldsymbol{\Omega}_k\mathbf{T}}{|1+\delta_k|^2}.$$

It follows from the algebraic lemma proven in Appendix E that $\mathbf{I}_n-\mathbf{A}$ is invertible provided that $\mathbf{B}$ or $\mathbf{C}$ have spectral norms strictly less than 1, in which case:

$$\left\|(\mathbf{I}_n-\mathbf{A})^{-1}\right\|_\infty \leq \sqrt{\left\|(\mathbf{I}_n-\mathbf{B})^{-1}\right\|_\infty}\sqrt{\left\|(\mathbf{I}_n-\mathbf{C})^{-1}\right\|_\infty}. \tag{29}$$

It appears from (29) that one needs to study matrices $\mathbf{B}$ and $\mathbf{C}$, which are at first sight easier to manipulate, mainly because they either involve $\mathbf{R}_N$ or $\mathbf{T}$. This however is not trivial. We state the result in the following proposition and for sake of readability defer the proof to Appendix D.

Proposition 10. *Assume that $z\in\mathbb{C}_+$. Then,*

1) *Matrix $\mathbf{C}$ satisfies $\rho(\mathbf{C})<1$. Moreover,*

$$\left\|(\mathbf{I}_n-\mathbf{C})^{-1}\right\|_\infty \leq \frac{K(\eta^2+|z|^2)^2}{|\Im z|^4} \tag{30}$$

where K and η are some positive constants independent of N.

2) *There exists 2 polynomials Q_1 and Q_2 independent of N with positive coefficients such that for N large enough and $z\in\mathcal{E}_N$ given by*

$$\mathcal{E}_N = \left\{z\in\mathbb{C}_+, \frac{1}{N^2}Q_1(|z|)Q_2(|\Im z|^{-1}) \leq \frac{1}{2}\right\}$$

we have $\rho(\mathbf{B})\leq 1$ and:

$$\|(\mathbf{I}_n-\mathbf{B})^{-1}\| \leq \tilde{K}\frac{(\tilde{\eta}^2+|z|^2)^2}{|\Im z|^4}.$$

It follows from proposition 10 that the spectral norm of $\mathbf{A}$ is strictly less than 1. Thus, $\mathbf{I}_n-\mathbf{A}$ is invertible and for $z\in\mathcal{E}_N$,

$$\begin{aligned}\|(\mathbf{I}_n-\mathbf{A})^{-1}\|_\infty &\leq \frac{1}{2}\left\|(\mathbf{I}_n-\mathbf{B})^{-1}\right\|_\infty + \frac{1}{2}\left\|(\mathbf{I}_n-\mathbf{C})^{-1}\right\|_\infty \\ &\leq \frac{K_{\max}(\eta_{\max}^2+|z|^2)}{|\Im z|^4},\end{aligned} \tag{31}$$

where $K_{\max}=\max(K,\tilde{K})$ and $\eta_{\max}=\max(\eta,\tilde{\eta})$. Plugging (31) into (28), we obtain:

$$\|\boldsymbol{\alpha}-\boldsymbol{\delta}\|_\infty \leq \frac{K_{\max}K_\epsilon}{N^2}(|z|+C_\epsilon)^{k_\epsilon}(\eta_{\max}+|z|^2)\frac{P_\epsilon(|\Im z|^{-1})}{|\Im z|^4},$$

where the right hand side of the above inequality can be put under the form:

$$\frac{\overline{K}(\overline{C}+|z|^2)^k}{N^2}\overline{P}(|\Im z|^{-1}).$$

for $\overline{K}$ and $\overline{C}$ positive constants, $\overline{k}$ integer, and $\overline{P}$ some polynomial with positive coefficients. Consider now the case where $z\in\mathbb{C}_+\backslash\mathcal{E}_N$. We first remark that:

$$|\alpha_j-\delta_j| \leq |\alpha_j|+|\delta_j| \leq \frac{2w_{\max}}{|\Im z|}.$$

Since $z\notin\mathcal{E}_N$, we therefore have:

$$\frac{1}{N^2}Q_1(|z|)Q_2(|\Im z|^{-1}) \geq \frac{1}{2}.$$

Hence:

$$\|\boldsymbol{\alpha}-\boldsymbol{\delta}\|_\infty \leq \frac{4w_{\max}}{|\Im z|N^2}Q_1(|z|)Q_2(|\Im z|^{-1})$$

As a consequence, we can find for C,K constants, k integer and P polynomial with positive coefficients such that:

$$\|\boldsymbol{\alpha}-\boldsymbol{\delta}\|_\infty \leq \frac{K}{N^2}(|z|+C)^kP(|\Im z|^{-1}),$$

thereby ending the proof.

Appendix A
Preliminaries

Many of the results of the appendix part are based on the following key lemmas, which we recall in this section for sake of clarity.

Lemma 11. *Let $\mathbf{A}=(a_{\ell,m})_{\ell,m=1}^n$ be an $n\times n$ real matrix and $\mathbf{u}$ and $\mathbf{v}$ be two real $n\times 1$ vectors. Assume that the entries of $\mathbf{A}$ are positive and that of $\mathbf{u}$ and $\mathbf{v}$ strictly positive. Assume, furthermore, that the equation:*

$$\mathbf{u}=\mathbf{A}\mathbf{u}+\mathbf{v}$$

is satisfied. Then, the spectral radius $\rho(\mathbf{A})$ of $\mathbf{A}$ of $\mathbf{A}$ satisfies:

$$\rho(\mathbf{A}) \leq 1-\frac{\min(v_\ell)}{\max(u_\ell)} < 1.$$

Lemma 12 (Matrix Inequality)**.** *Let $\mathbf{A}$ be a $n\times n$ hermitian matrix. Then,*

$$\frac{1}{n}\operatorname{tr}\mathbf{A}\mathbf{A}^* \geq \left|\frac{1}{n}\operatorname{tr}\mathbf{A}\right|^2$$

with equality only if $\mathbf{A}$ is proportional to identity.

Proof: Let $\mathbf{A}=\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{H}}$ be an eiengevalue decomposition of $\mathbf{A}$. Consider $\lambda_1,\cdots,\lambda_n$ the eigenvalues of $\mathbf{A}$. Then, if there is $i\neq j$ such that $\lambda_i\neq\lambda_j$, we have due to the strict-convexity of $x\mapsto x^2$:

$$\begin{aligned}\frac{1}{n}\operatorname{tr}\mathbf{A}\mathbf{A}^* &= \frac{1}{n}\sum_{i=1}^n\lambda_i^2 \\ &> \left|\frac{1}{n}\sum_{i=1}^n\lambda_i\right|^2\end{aligned}$$

■

Appendix B
Proof of Theorem 1

In order to establish that 0 does not belong to the support $\mathcal{S}_N$, we show that it exists $\epsilon > 0$ for which $\mu_N([0,x]) = 0$ for each $x \in]0,\epsilon[$. To this end, define function $\phi : \mathbb{R}_+^n \times \mathbb{R}^+ \to \mathbb{R}_+^n$, with:

$$\phi(x_1,\cdots,x_n,z) = (\phi_1(x_1,\cdots,x_n,z),\cdots,\phi_n(x_1,\cdots,x_n,z))$$

where $\phi_i : \mathbb{R}_+^n \times \mathbb{R}_+ \to \mathbb{R}_+$ is given by:

$$\phi_i(x_1,\cdots,x_n,z) = \frac{1}{n}\operatorname{tr}\mathbf{\Omega}_i\left(\frac{1}{n}\sum_{k=1}^n \frac{\mathbf{\Omega}_k}{1+x_k} - zI_N\right)^{-1}.$$

We need to show that there exists $\ell_1,\cdots,\ell_n$ such that:

$$\phi_i(\ell_1,\cdots,\ell_n,0) = \ell_i.$$

Let $p \in \mathbb{N}$ and $r_p = -\frac{1}{p}$. We will first start by proving that for each p, there exists a unique $\overline{x}_1^p,\cdots,\overline{x}_n^p$ such that:

$$\phi_i(\overline{x}_1^p,\cdots,\overline{x}_n^p,r_p) = \overline{x}_i^p.$$

For that, it suffices to show that $\tilde{\phi}_p : \mathbb{R}_+^n \to \mathbb{R}_+^n, (x_1,\cdots,x_n) \mapsto \phi(x_1,\cdots,x_n,r_p)$ is a standard interference function. In particular, we need to check that ϕ satisfy the following properties:

- Nonnegativity: For each $x_1,\cdots,x_n \geq 0$ and each i and p, $\phi_i(x_1,\cdots,x_n,r_p) > 0$.
- Monotonicity: For each $x_1 \geq x_1^{'},\cdots,x_n \geq x_n^{'}$, and each i and p,
$$\phi_i(x_1,\cdots,x_n,r_p) \geq \phi_i\left(x_1^{'},\cdots,x_n^{'},r_p\right).$$
- Scalability: For each $\alpha > 1$, and each i and p, $\alpha\phi_i(x_1,\cdots,x_n,r_p) > \phi_i(\alpha x_1,\cdots,\alpha x_n,r_p)$.

The first item is obvious since $\mathbf{\Omega}_i$ are positive definite matrices, while the second one follows from the fact that for positive definite matrices, $\mathbf{A} \succeq \mathbf{B}$ implies $\mathbf{B}^{-1} \succeq \mathbf{A}^{-1}$. Finally, to prove the last item, note that for $\alpha > 1$,

$$\begin{aligned}\phi_i(\alpha x_1,\cdots,\alpha x_n,r_p) &< \frac{1}{N}\operatorname{tr}\mathbf{\Omega}_i\left(\frac{1}{n}\sum_{k=1}^n \frac{\mathbf{\Omega}_k}{\alpha(1+x_k)} - \frac{r_p}{\alpha}I_N\right)^{-1} \\ &= \alpha\phi_i(x_1,\cdots,x_n,r_p).\end{aligned}$$

Therefore,

$$\phi_i(\alpha x_1,\cdots,\alpha x_n,r_p) > \alpha\phi_i(x_1,\cdots,x_n,r_p).$$

According to [21, Theorem 2], $\tilde{\phi}_p$ is a standard interference function. To prove that there exists a unique $\overline{x}_1^p,\cdots,\overline{x}_n^p$ satisfying:

$$\overline{x}_i^p = \phi_i(\overline{x}_1^p,\cdots,\overline{x}_n^p),$$

we need to check that there exits $x_1,\cdots,x_n$ such that:

$$x_i > \phi_i(x_1,\cdots,x_n,r_p).$$

This condition holds true, since $\phi_i(x_1,\cdots,x_n) \leq \frac{1}{r_p}$, and so increasing x_i to infinity will satisfy the above inequality.

Moreover, consider the sequence:

$$x_i^{(t,p)} = \phi_i(x_1^{(t-1,p)},\cdots,x_n^{(t-1,p)}), \quad i = 1,\cdots,n$$

where $x_1^{(0,p)},\cdots,x_n^{(0,p)}$ are arbitrary positive reals. Then, $\mathbf{x}^{(t,p)} = \left(x_1^{(t,p)},\cdots,x_n^{(t,p)}\right)$ converge to $\overline{\mathbf{x}}^p = (\overline{x}_1^p,\cdots,\overline{x}_n^p)$.

From this, we can prove that for $p \geq q$, we have for each $i \in \{1,\cdots,n\}$,

$$\overline{x}_i^p \geq \overline{x}_i^q.$$

To this end, we will consider the sequence,

$$x_i^{(t,p)} = \phi_i(x_1^{(t-1,p)},\cdots,x_n^{(t-1,p)}), \quad i = 1,\cdots,n$$

where $x_i^{(0,p)} = \overline{x}_i^q$ and will show that for any t,

$$x_i^{(t,p)} \geq \overline{x}_i^q.$$

We will proceed by induction on t. For $t = 0$, the result obviously holds. Assume that the resuld holds for any $k \leq t$, i.e,

$$x_i^{(k,p)} \geq \overline{x}_i^q, \quad i = 1,\cdots,n \text{ and } k \leq t.$$

And let us prove it for $t = k+1$. We have:

$$\begin{aligned}x_i^{(t+1,p)} &= \phi_i(x_1^{(t,p)},\cdots,x_n^{(t,p)},r_p) \\ &\geq \phi_i(x_1^{(t,p)},\cdots,x_n^{(t,p)},r_q) \\ &\overset{(a)}{\geq} \phi_i(\overline{x}_1^q,\cdots,\overline{x}_n^q,r_q) \\ &= \overline{x}_i^q.\end{aligned}$$

where (a) follows since ϕ_i is increasing in each variable and $x_i^{(t,p)} \geq \overline{x}_i^{(q)}$ by the induction assumption.

We have therefore shown that for $p \geq q$,

$$\overline{x}_i^p \geq \overline{x}_i^q.$$

As p tends to infinity, $\overline{x}_i^p$ will converge to a limit $\ell_i \in \mathbb{R}_+ \cup \{+\infty\}$. Assume that for $i \in \{1,\cdots,n\}$, $\ell_i \neq +\infty$. Then, one can easily see, that necessarily, $\ell_i \neq +\infty$ for any $i \in \{1,\cdots,n\}$. We will prove now, that the case of $\ell_i = +\infty$ for all $i = 1,\cdots,n$ cannot hold. For this observe that:

$$\sum_{i=1}^n \frac{\overline{x}_i^p}{1+\overline{x}_i^p} = \sum_{i=1}^n \frac{1}{n}\operatorname{tr}\frac{\mathbf{\Omega}_i}{1+\overline{x}_i^p}\left(\frac{1}{n}\sum_{k=1}^n \frac{\mathbf{\Omega}_k}{1+x_k^p} + r_p\mathbf{I}_N\right)^{-1} \leq N.$$

Let $\overline{x}_{\min}^p = \min_{1\leq i\leq n}\overline{x}_i^p$. We have thus:

$$\frac{\overline{x}_{\min}}{1+\overline{x}_{\min}} \leq \frac{N}{n}$$

or equivalently:

$$\overline{x}_{\min} \leq \frac{\frac{N}{n}}{1-\frac{N}{n}}.$$

which is contradiction with the fact that $\ell_i = +\infty$ for all i.

Recall now that:

$$\phi_i(\overline{x}_1^p,\cdots,\overline{x}_n^p,r_p) = \overline{x}_i^p.$$

Taking the limit in p, we thus get that:

$$\phi_i(\ell_1,\cdots,\ell_n,0) = \ell_i,$$

or equivalently:

$$\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_i\left(\frac{1}{n}\sum_{k=1}^n \frac{\mathbf{\Omega}_k}{1+\ell_k}\right)^{-1} = \ell_i.$$

The Jakobian matrix corresponding to $\tilde{\phi}_\infty : \mathbb{R}_+^n \to \mathbb{R}_+^n : (x_1,\cdots,x_n) \mapsto \phi(x_1,\cdots,x_n,0)$ at $x_i=\ell_i, i=1,\cdots,n$, is given by:

$$[\mathbf{J}]_{i,m} = \frac{1}{n^2}\operatorname{tr}\mathbf{\Omega}_i\left(\frac{1}{n}\sum_{k=1}^n\frac{\mathbf{\Omega}_k}{1+\ell_k}\right)\frac{\mathbf{\Omega}_m}{(1+\ell_m)^2}\left(\frac{1}{n}\sum_{r=1}^n\frac{\mathbf{\Omega}_r}{1+\ell_r}\right)^{-1}$$

Let $\mathbf{u}=[1+\ell_1,\cdots,1+\ell_n]^{\mathrm{T}}$ and $\mathbf{v}=[\ell_1,\cdots,\ell_n]^{\mathrm{T}}$. Then, after simple calculations, one can show that:

$$\mathbf{Ju}=\mathbf{v}.$$

The entries of $\mathbf{J}$, $\mathbf{u}$ and $\mathbf{v}$ are strictly positive. A direct application of Lemma 11 in section A implies that:

$$\rho(\mathbf{J}) \le 1-\frac{\min_{1\le i\le n}\ell_i}{1+\max_{1\le i\le n}\ell_i} < 1.$$

thereby showing that $\mathbf{I}_n-\mathbf{J}$ is invertible. Hence, the implicit function theorem implies that there exists an open disk at zero with radius $\eta>0$, i.e $D(0,\eta)$ and unique analytic functions $\varphi_1,\cdots,\varphi_n$ defined in $D(0,\eta)$ such that:

$$\phi_i\left(\varphi_1(z),\cdots,\varphi_n(z),z\right)=\varphi_i(z)$$

and

$$\varphi_i(0)=\ell_i, \ \ i=1,\cdots,n.$$

On the other hand, one can show that there exists $\epsilon>0$ such that $\varphi_i(t)$ is real valued and strictly positive for any $t\in[-\epsilon,\epsilon]$. Indeed, writing $\Im\varphi_i(t)$ as:

$$\begin{aligned}\Im\varphi_i(t) &= \frac{1}{2\imath}\left(\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_i\left(\frac{1}{n}\sum_{k=1}^n\frac{\mathbf{\Omega}_k}{1+\varphi_k(t)}-t\mathbf{I}_N\right)^{-1}\right.\\ &\left.-\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_i\left(\frac{1}{n}\sum_{k=1}^n\frac{\mathbf{\Omega}_k}{1+\varphi_k^*(t)}-t\mathbf{I}_N\right)^{-1}\right)\\ &= \frac{1}{n}\operatorname{tr}\mathbf{\Omega}_i\left(\frac{1}{n}\sum_{k=1}^n\frac{\mathbf{\Omega}_k}{1+\varphi_k(t)}-t\mathbf{I}_N\right)^{-1}\\ &\times\left(\frac{1}{n}\sum_{k=1}^n\frac{\mathbf{\Omega}_k\Im(\varphi_k(t))}{|1+\varphi_k(t)|^2}\right)\left(\frac{1}{n}\sum_{k=1}^n\frac{\mathbf{\Omega}_k}{1+\varphi_k(t)}-t\mathbf{I}_N\right)^{-1}.\end{aligned}$$

Therefore, the vector $\mathbf{g}_t=[\Im(\varphi_1(t)),\cdots,\Im(\varphi_n(t))]^{\mathrm{T}}$ is solution of the following system of equations:

$$\mathbf{g}_t=\mathbf{J}_t\mathbf{g}_t.$$

As $t\mapsto\rho(\mathbf{J}_t)$ is continuous, and since for $t=0$, $\rho(\mathbf{J}_t)=\rho(\mathbf{J})<1$, there exists $\epsilon>0$ such that:

$$\rho(\mathbf{J}_t)<1$$

for every $t\in[-\epsilon,\epsilon]$. Therefore, $\mathbf{g}_t=0$. Furthermore, since at $t=0$, $\varphi_i(0)=\ell_i>0$, we can futher assume that ϵ is chosen such that $\varphi_i(t)$ is real-valued and strictly positive for any $t\in[-\epsilon,\epsilon]$. From [7, Theorem 1], we know that for $t<0$, $\delta_1(t),\cdots,\delta_n(t)$ are the unique non-negative pointwise solutions of the following system of equations

$$\delta_i(t)=\phi_i(\delta_1(t),\cdots,\delta_n(t),t),$$

thereby implying that:

$$\delta_i(t)=\varphi_i(t)$$

for any $t\in[-\epsilon,0]$. Since, the set of functionals $\delta_1(t),\cdots,\delta_n(t)$ and $\varphi_1(t),\cdots,\varphi_n(t)$ are holomorphic on $D(0,\epsilon)\backslash\{[0,\epsilon[\}$ and coincide on a set of values with an accumulation point, they must coincide on the whole domaine of analicity, namely $D(0,\epsilon)\backslash\{[0,\epsilon[\}$.

Let $\overline{m}$ be given by:

$$\overline{m}=\frac{1}{N}\operatorname{tr}\left(\frac{1}{n}\sum_{k=1}^n\frac{\mathbf{\Omega}_k}{1+\varphi_k(z)}-z\mathbf{I}_N\right)^{-1}.$$

Obviously $\overline{m}$ is analytic on $D(0,\epsilon)$ and satisfies:

$$\overline{m}(z)=m_N(z)$$

for all $z\in D(0,\epsilon)\backslash\{[0,\epsilon[\}$. We recall that for $0\le x<\epsilon$, $\mu_N([0,x])$ can be expressed as:

$$\mu_N([0,x])=\frac{1}{\pi}\lim_{y\to0,y>0}\int_0^x\Im(m_N(s+\imath y))ds.$$

Therefore,

$$\mu_N([0,x])=\frac{1}{\pi}\lim_{y\to0,y>0}\int_0^x\Im(\overline{m}(s+\imath y))ds$$

As $\overline{m}$ is holomorphic on $D(0,\epsilon)$, the dominated convergence theorem implies that:

$$\frac{1}{\pi}\lim_{y\to0,y>0}\int_0^x\Im(\overline{m}(s+\imath y))ds=\frac{1}{\pi}\int_0^x\Im(\overline{m}(s))ds=0$$

since $\overline{m}(s)\in\mathbb{R}$ for $s\in[0,x]$. Thus, we establish that $\mu_N([0,x])=0$.

Appendix C
Proof of Lemma 9

The proof follows from a direct application of the Nash-Poincaré inequality in Lemma 7. Define $\beta_{\mathbf{A}}=\frac{1}{n}\operatorname{tr}\mathbf{AQ}(z)$. We then have:

$$\begin{aligned}\operatorname{var}(\beta_{\mathbf{A}}(z)) &\le \sum_{k=1}^n\sum_{s=1}^N\sum_{r=1}^N\frac{1}{n^4}\mathbb{E}\left[[\mathbf{\Sigma}_n^*\mathbf{QAQ}]_{k,s}[\mathbf{\Omega}_k]_{s,r}[\mathbf{Q}^*\mathbf{A}^*\mathbf{Q}^*\mathbf{\Sigma}_n]_{r,k}\right]\\ &+\sum_{k=1}^n\sum_{s=1}^N\sum_{r=1}^N\frac{1}{n^4}\mathbb{E}\left[[\mathbf{Q}^*\mathbf{A}^*\mathbf{Q}^*\mathbf{\Sigma}_n]_{k,s}[\mathbf{\Omega}_k]_{s,r}[\mathbf{\Sigma}_n^*\mathbf{QAQ}]_{r,k}\right]\\ &=\sum_{k=1}^n\frac{1}{n^4}\mathbb{E}\left[[\mathbf{\Sigma}_n^*\mathbf{QAQ\Omega}_k\mathbf{Q}^*\mathbf{A}^*\mathbf{Q}^*\mathbf{\Sigma}_n]_{k,k}\right]\\ &+\sum_{k=1}^n\frac{1}{n^4}\mathbb{E}\left[[\mathbf{Q}^*\mathbf{A}^*\mathbf{Q}^*\mathbf{\Sigma}_n\mathbf{\Omega}_k\mathbf{\Sigma}_n^*\mathbf{QAQ}]_{k,k}\right].\end{aligned}$$

Since $\mathbf{\Omega}_k\preceq w_{\max}\mathbf{I}_N$ with $w_{\max}=\sup_N\max_{1\le k\le n}\|\mathbf{\Omega}_k\|$, we have:

$$\begin{aligned}\operatorname{var}(\beta_{\mathbf{A}})(z) &\le \frac{w_{\max}}{n^3}\operatorname{tr}\left(\mathbf{QAQQ}^*\mathbf{A}^*\mathbf{Q}^*\frac{\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*}{n}\right)\\ &+\frac{w_{\max}}{n^3}\operatorname{tr}\left(\mathbf{Q}^*\mathbf{A}^*\mathbf{Q}^*\frac{\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*}{n}\mathbf{QAQ}\right).\end{aligned}$$

Using the resolvent identity:

$$\mathbf{Q}(z)\frac{\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*}{n}=\frac{\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*}{n}\mathbf{Q}(z)=\mathbf{I}_N+z\mathbf{Q}(z),$$

and the inequality $\|\mathbf{Q}(z)\| \leq \frac{1}{|\Im(z)|}$, we obtain:

$$\operatorname{var}(\beta_{\mathbf{A}}(z)) \leq \frac{2w_{\max}\|\mathbf{A}\|^2}{n^2}\left(\frac{1}{|\Im(z)|^3}+\frac{|z|}{|\Im(z)|^4}\right) \leq \frac{2w_{\max}\|\mathbf{A}\|^2}{n^2}(|z|+1)\left(\frac{1}{|\Im(z)|^4}+\frac{1}{|\Im(z)|^3}\right).$$

Appendix D
Proof of proposition 10

In order to prove proposition 10, we need first to show that the sequence of measures μ_N is tight. To this end, we will follow the same steps as in [13, Lemma C1]. Observe that:

$$\int_0^{+\infty}\lambda\mu_N(d\lambda) = \lim_{y\to+\infty}\Re\left[-\imath y\left(\imath y m_N(\imath y)+1\right)\right] = \lim_{y\to+\infty}\Re\left[-\imath y\left(\imath y\frac{1}{N}\operatorname{tr}\mathbf{T}_N(\imath y)+1\right)\right]. \tag{32}$$

On the other hand:

$$\mathbf{T}_N(\imath y)\left(\frac{1}{n}\sum_{k=1}^n\frac{\mathbf{\Omega}_k}{1+\delta_k(\imath y)}-\imath y\mathbf{I}_N\right)=\mathbf{I}_N.$$

Therefore,

$$1+\frac{1}{N}\operatorname{tr}\imath y\mathbf{T}_N(\imath y) = \frac{1}{n}\sum_{k=1}^n\frac{1}{N}\frac{\operatorname{tr}\mathbf{\Omega}_k\mathbf{T}_N(\imath y)}{1+\delta_k(\imath y)} = \frac{1}{n}\sum_{k=1}^n\frac{1}{c_N}\frac{\delta_k(\imath y)}{1+\delta_k(\imath y)}. \tag{33}$$

Plugging (33) into (32), we finally get:

$$\int_0^{+\infty}\lambda\mu_N(d\lambda) = \lim_{y\to+\infty}\frac{1}{n}\frac{1}{c_N}\sum_{k=1}^n\frac{\Re\left[-\imath y\delta_k(\imath y)\right]}{|1+\delta_k(iy)|^2}.$$

Since δ_k are Stieltjes transforms of finite positive measures, we have:

$$\lim_{y\to+\infty}|\delta_k(\imath y)|=0$$

Moreover, we have $\lim_{y\to+\infty}-\imath y\delta_k(\imath y)=\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_k$, thereby establishing that:

$$\sup_N\int_0^{+\infty}\lambda\mu_N(d\lambda)<+\infty.$$

The tightness of the sequence μ_N follows directly from the above inequality. In the same way, we can also show that the sequence of measures corresponding to the Stieltjes transforms $\frac{1}{N}\operatorname{tr}\mathbf{R}$ is also tight. These two results will be of fundamental importance in the proof of proposition 10.

We now return to the proof of proposition10:

a) Proof of proposition 10-1): The proof is based on the use of Lemma 11 in section A. For that, we need to find a linear system involving matrix $\mathbf{C}$. For $z\in\mathbb{C}_+$, we have:

$$\begin{aligned}\Im(\delta_j) &= \frac{1}{2\imath n}\left(\operatorname{tr}\mathbf{\Omega}_j\mathbf{T}-\operatorname{tr}\mathbf{\Omega}_j\mathbf{T}^{\mathrm{H}}\right)\\ &= \frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{T}\left(\frac{1}{n}\sum_{k=1}^n\frac{\mathbf{\Omega}_k\Im(\delta_k)}{|1+\delta_k|^2}+\Im(z)\mathbf{I}_N\right)\mathbf{T}^{\mathrm{H}}\\ &= \frac{1}{n^2}\sum_{k=1}^n\frac{\operatorname{tr}\mathbf{\Omega}_j\mathbf{T}\mathbf{\Omega}_k\mathbf{T}^{\mathrm{H}}}{|1+\delta_k|^2}\Im(\delta_k)+\Im(z)\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{T}\mathbf{T}^{\mathrm{H}}.\end{aligned}$$

Let $\mathbf{I}_\delta$ and $\mathbf{c}$ be the $n\times1$ vectors given by:

$$\mathbf{I}_\delta = \left[\Im(\delta_1),\cdots,\Im(\delta_n)\right]^{\mathrm{T}}$$
$$\mathbf{c} = \left[\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_1\mathbf{T}\mathbf{T}^{\mathrm{H}},\cdots,\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_n\mathbf{T}\mathbf{T}^{\mathrm{H}}\right]^{\mathrm{T}},$$

Then:

$$\mathbf{I}_\delta = \mathbf{C}\,\mathbf{I}_\delta+\Im(z)\mathbf{c}.$$

Since $\Im(\delta_j)>0$ for all j and $\Im z>0$ and $\mathbf{C}$, $\mathbf{c}$ have positive entries, we get from Lemma 11,

$$\left\|(\mathbf{I}_n-\mathbf{C})^{-1}\right\|_\infty \leq \frac{\max_{1\leq j\leq n}\Im\delta_j}{\Im z\min_{1\leq j\leq n}\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{T}\mathbf{T}^{\mathrm{H}}} \leq \frac{w_{\max}}{|\Im z|^2\min_{1\leq j\leq n}\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{T}\mathbf{T}^{\mathrm{H}}},$$

where the second inequality follows from the fact that $\max_{1\leq j\leq n}\Im\delta_j\leq\max_{1\leq j\leq n}|\delta_j|\leq\frac{w_{\max}}{\Im z}$. Using the inequality $\frac{1}{n}\operatorname{tr}\mathbf{AB}\geq\lambda_1(\mathbf{A})\frac{1}{n}\operatorname{tr}\mathbf{B}$ for $\mathbf{A}$ and $\mathbf{B}$ hermitian positive definite matrices with $\lambda_1(\mathbf{A})$ the smallest eigenvalue of $\mathbf{A}$, we get:

$$\left\|(\mathbf{I}_n-\mathbf{C})^{-1}\right\|_\infty \leq \frac{w_{\max}}{|\Im z|^2 w_{\min}\frac{1}{n}\operatorname{tr}\mathbf{T}\mathbf{T}^{\mathrm{H}}}. \tag{34}$$

In order to obtain a lower bound on $\frac{1}{N}\operatorname{tr}\mathbf{T}\mathbf{T}^{\mathrm{H}}$, we first remark that by the Jensen inequality in Lemma 12: $\frac{1}{N}\operatorname{tr}\mathbf{T}\mathbf{T}^{\mathrm{H}}\geq\left|\frac{1}{N}\operatorname{tr}\mathbf{T}\right|^2=|m_N(z)|^2\geq\Im(m_N(z))^2$. As $(\mu_N)_{N\geq0}$ is tight, it exists $\eta>0$ for which $\mu_N([\eta,+\infty))\leq\frac{1}{2}$ for all N and as such:

$$\mu_N\left([0,\eta]\right)\geq\frac{1}{2}.$$

As a consequence,

$$\Im(m_N(z)) = \Im(z)\int_0^{+\infty}\frac{d\mu_N(\lambda)}{|\lambda-z|^2} > \int_0^{\eta}\frac{\Im(z)d\mu_N(\lambda)}{2(\eta^2+|z|^2)}\mu_N([0,\eta]) \geq \frac{\Im(z)}{4(\eta^2+|z|^2)}. \tag{35}$$

Plugging (35) into (34), we finally get (30).

b) Proof of proposition 10-2): The proof is similar to that of the first statement. We first decompose α_j as:

$$\alpha_j = \alpha_j-\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{R}+\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{R} = \epsilon_j+\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{R}.$$

Hence,

$$\Im(\alpha_j) = \Im(\epsilon_j(z))+\Im\left(\frac{1}{n}\operatorname{tr}\mathbf{\Omega}_j\mathbf{R}\right).$$

Using the same kind of calculations as above, we thus get:

$$\Im(\alpha_j) = \Im(\epsilon_j) + \frac{1}{n^2}\sum_{k=1}^{n}\frac{\operatorname{tr}\boldsymbol{\Omega}_j\mathbf{R}\boldsymbol{\Omega}_k\mathbf{R}^{\mathrm{H}}\Im\alpha_k}{|1+\alpha_k(z)|^2} + \Im(z)\frac{1}{n}\operatorname{tr}\boldsymbol{\Omega}_j\mathbf{R}\mathbf{R}^{\mathrm{H}}. \tag{36}$$

In order to determine a subset of $\mathbb{C}_+$ on which $\Im(z)\frac{1}{n}\operatorname{tr}\boldsymbol{\Omega}_j\mathbf{R}\mathbf{R}^{\mathrm{H}} + \Im(\epsilon_j(z)) > 0$, we evaluate a lower bound of $\frac{1}{n}\operatorname{tr}\boldsymbol{\Omega}_j\mathbf{R}\mathbf{R}^{\mathrm{H}}$. We have by the Jensen inequality in Lemma 12:

$$\frac{1}{n}\operatorname{tr}\boldsymbol{\Omega}_j\mathbf{R}\mathbf{R}^* \geq w_{\min}\left|\frac{1}{n}\operatorname{tr}\mathbf{R}\right|^2 = w_{\min}\left(\frac{N}{n}\right)^2\left|\frac{1}{N}\operatorname{tr}\mathbf{R}\right|^2.$$

From the discussion in the beginning of this section, we know that the sequence of measures corresponding to the Stieltjes transforms $\frac{1}{N}\operatorname{tr}\mathbf{R}$ is tight. Hence, there exists $\tilde{\eta}$ such that:

$$\Im\left(\frac{1}{N}\operatorname{tr}\mathbf{R}\right) \geq \frac{\Im z}{4\left(\tilde{\eta}^2+|z|^2\right)}.$$

Hence,

$$\frac{1}{n}\operatorname{tr}\boldsymbol{\Omega}_j\mathbf{R}\mathbf{R}^* \geq w_{\min}\left(\frac{N}{n}\right)^2\frac{|\Im z|^2}{16\left(\tilde{\eta}^2+|z|^2\right)^2}.$$

On the other hand, from (26), we recall that:

$$|\epsilon_j(z)| \leq \frac{K_\epsilon}{N^2}(|z|+C_\epsilon)^{k_\epsilon}P_\epsilon(|\Im z|^{-1}).$$

Consider $\mathcal{E}_{N,1}$ the set given by:

$$\mathcal{E}_{N,1} = \left\{z\in\mathbb{C}_+, \frac{w_{\min}\left(\frac{N}{n}\right)^2|\Im z|^2}{16(\tilde{\eta}^2+|z|^2)^2} - \frac{K_\epsilon}{N^2}(|z|+C_\epsilon)^{k_\epsilon}P_\epsilon(|\Im z|^{-1}) > 0\right\}$$

Then, as before, using the fact that for $z\in\mathcal{E}_{N,1}$ (36) can be cast into a linear system of equations involving positive-entries matrix and vectors, we deduce that $\rho(\mathbf{B}) < 1$ and:

$$\left\|(\mathbf{I}_n-\mathbf{B})^{-1}\right\|_\infty \leq \frac{\max_{1\leq j\leq n}\alpha_j}{\frac{w_{\min}N^2}{n^2}\frac{|\Im z|^3}{16(\tilde{\eta}^2+|z|^2)^2} - \frac{K_\epsilon}{N^2}(|z|+C_\epsilon)^{k_\epsilon}P_\epsilon(|\Im z|^{-1})}$$
$$\leq \frac{1}{\frac{w_{\min}N^2}{n^2}\frac{|\Im z|^4}{16(\tilde{\eta}^2+|z|^2)^2}\left(1-\frac{1}{N^2}Q_1(|z|)Q_2(|\Im z|^{-1})\right)},$$

where Q_1 and Q_2 are polynomials with positive coefficients.

Take $\mathcal{E}_N$ as the set defined by:

$$\mathcal{E}_N = \left\{z\in\mathbb{C}_+, \frac{1}{N^2}Q_1(|z|)\,Q_2\left(\Im z^{-1}\right) \leq \frac{1}{2}\right\}.$$

Obviously $\mathcal{E}_N \subseteq \mathcal{E}_{N,1}$, and for all $z\in\mathcal{E}_N$, we get:

$$\left\|(\mathbf{I}_n-\mathbf{B})^{-1}\right\|_\infty \leq \frac{32n^2\left(\tilde{\eta}^2+|z|^2\right)^2}{w_{\min}N^2|\Im z|^4}.$$

Appendix E
A linear algebraic result

Finally, we finish the Appendix part with a linear algebraic lemma which we need in our derivation and can be of independent interest.

Lemma 13. *Let $\mathbf{B}$ and $\mathbf{C}$ be $n\times n$ matrices with non-negative entries. Let $\mathbf{A}$ be a $n\times n$ matrix satisfying:*

$$\left|[\mathbf{A}]_{i,j}\right| \leq \sqrt{[\mathbf{B}]_{i,j}}\sqrt{[\mathbf{C}]_{i,j}}. \tag{37}$$

Then, $\rho(\mathbf{A}) \leq \sqrt{\rho(\mathbf{B})}\sqrt{\rho(\mathbf{C})}$. If furthermore $\max(\rho(\mathbf{A}),\rho(\mathbf{B})) < 1$, then $\rho(\mathbf{A}) < 1$ and:

$$\left\|(\mathbf{I}_n-\mathbf{A})^{-1}\right\|_\infty \leq \sqrt{\left\|(\mathbf{I}_n-\mathbf{B})^{-1}\right\|_\infty}\sqrt{\left\|(\mathbf{I}_n-\mathbf{C})^{-1}\right\|_\infty}$$

Proof: We start by proving that $\rho(\mathbf{A}) \leq \sqrt{\rho(\mathbf{B})}\sqrt{\rho(\mathbf{C})}$. For that, consider $\tilde{\mathbf{A}}$, the matrix given by:

$$\left[\tilde{\mathbf{A}}\right]_{i,j} = \sqrt{[\mathbf{B}]_{i,j}}\sqrt{[\mathbf{C}]_{i,j}}$$

Consider $|\mathbf{A}|$ the matrix such that $[|\mathbf{A}|]_{i,j} = \left|[\mathbf{A}]_{i,j}\right|$. Then, $\rho(|\mathbf{A}|) \leq \rho(\tilde{\mathbf{A}})$. Recall, that for any matrix $\mathbf{D}$,

$$\rho(\mathbf{D}) = \lim_{k\to+\infty}\|\mathbf{D}^k\|_\infty^{\frac{1}{k}}.$$

From the above convergence, we have:

$$\begin{aligned}\left[\tilde{\mathbf{A}}^k\right]_{i,j} &= \sum_{i_1,\cdots,i_{k-1}}^{n}\left[\tilde{\mathbf{A}}\right]_{i,i_1}\left[\tilde{\mathbf{A}}\right]_{i_2,i_3}\cdots\left[\tilde{\mathbf{A}}\right]_{i_{k-1},j}\\
&= \sum_{1\leq i_1,\cdots,i_{k-1}\leq n}\sqrt{[\mathbf{B}]_{i,i_1}[\mathbf{B}]_{i_2,i_3}\cdots[\mathbf{B}]_{i_{k-1},j}}\\
&\times\sqrt{[\mathbf{C}]_{i,i_1}[\mathbf{C}]_{i_2,i_3}\cdots[\mathbf{C}]_{i_{k-1},j}}\\
&\leq \sqrt{\sum_{1\leq i_1,\cdots,i_{k-1}\leq n}[\mathbf{B}]_{i,i_1}[\mathbf{B}]_{i_2,i_3}\cdots[\mathbf{B}]_{i_{k-1},j}}\\
&\sqrt{\sum_{1\leq i_1,\cdots,i_{k-1}\leq n}[\mathbf{C}]_{i,i_1}[\mathbf{C}]_{i_2,i_3}\cdots[\mathbf{C}]_{i_{k-1},j}}\\
&= \sqrt{[\mathbf{B}^k]_{i,j}}\sqrt{[\mathbf{C}^k]_{i,j}}.\end{aligned}$$

With this inequality at hand, we are now in position to bound $\|\tilde{\mathbf{A}}^k\|_\infty$. We have:

$$\begin{aligned}\left\|\tilde{\mathbf{A}}^k\right\|_\infty &= \max_{1\leq i\leq n}\sum_{j=1}^{n}\left[\tilde{\mathbf{A}}^k\right]_{i,j}\\
&\leq \max_{1\leq i\leq n}\sum_{j=1}^{n}\sqrt{\left[\mathbf{B}^k\right]_{i,j}}\sqrt{\left[\mathbf{C}^k\right]_{i,j}}\\
&\leq \max_{1\leq i\leq n}\sqrt{\sum_{j=1}^{n}[\mathbf{B}^k]_{i,j}}\sqrt{\sum_{j=1}^{n}[\mathbf{C}^k]_{i,j}}\\
&\leq \sqrt{\|\mathbf{B}^k\|_\infty}\sqrt{\|\mathbf{C}^k\|_\infty}.\end{aligned}$$

We therefore have:

$$\begin{aligned}\rho(\tilde{\mathbf{A}}) &= \lim_{k\to+\infty}\left\|\tilde{\mathbf{A}}^k\right\|_\infty^{\frac{1}{k}}\\ &\leq \lim_{k\to+\infty}\left\|\mathbf{B}^k\right\|_\infty^{\frac{1}{2k}}\left\|\mathbf{C}^k\right\|_\infty^{\frac{1}{2k}}\\ &= \sqrt{\rho(\mathbf{B})}\sqrt{\rho(\mathbf{C})}.\end{aligned}$$

Therefore, $\rho(\tilde{\mathbf{A}}) < 1$ and thus, $\rho(\mathbf{A}) < 1$ if $\max(\rho(\mathbf{C}),\rho(\mathbf{B})) < 1$. In this case, $\mathbf{I}_n-\mathbf{A}$ is invertible and also are $\mathbf{I}_n-\mathbf{B}$ and $\mathbf{I}_n-\mathbf{C}$. Since $(\mathbf{I}_n-\mathbf{A})^{-1} = \sum_{k=0}^{+\infty}\mathbf{A}^k$. for any $1\leq i\leq n$, we have:

$$\begin{aligned}&\sum_{j=1}^n\left|\left[(\mathbf{I}_n-\mathbf{A})^{-1}\right]_{i,j}\right| \leq \sum_{k=0}^{+\infty}\sum_{j=1}^n\left|\left[\mathbf{A}^k\right]_{i,j}\right|\\ &\leq \sum_{k=0}^{\infty}\sum_{j=1}^n\left[|\mathbf{A}|^k\right]_{i,j} \leq \sum_{k=0}^{+\infty}\sum_{j=1}^n\sqrt{[\mathbf{B}^k]_{i,j}}\sqrt{\left[|\mathbf{C}|^k\right]_{i,j}}\\ &\leq \sum_{k=0}^{+\infty}\sqrt{\sum_{j=1}^n[\mathbf{B}^k]_{i,j}}\sqrt{\sum_{j=1}^n[\mathbf{C}^k]_{i,j}}\\ &\leq \sqrt{\sum_{k=0}^{+\infty}\sum_{j=1}^n[\mathbf{B}^k]_{i,j}}\sqrt{\sum_{k=0}^{+\infty}\sum_{j=1}^n[\mathbf{C}^k]_{i,j}}\\ &\leq \sqrt{\left\|(\mathbf{I}_n-\mathbf{B})^{-1}\right\|_\infty}\sqrt{\left\|(\mathbf{I}_n-\mathbf{C})^{-1}\right\|_\infty}.\end{aligned}$$

As a consequence, we have:

$$\left\|(\mathbf{I}_n-\mathbf{A})^{-1}\right\|_\infty \leq \sqrt{\left\|(\mathbf{I}_n-\mathbf{B})^{-1}\right\|_\infty}\sqrt{\left\|(\mathbf{I}_n-\mathbf{C})^{-1}\right\|_\infty}.$$

■

REFERENCES

[1] P. Vallet, P. Loubaton, and X. Mestre, "Improved subspace estimation for multivariate observations of high dimension: the deterministic signals case," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, Feb. 2012.

[2] V. A. Marchenko and L. A. Pastur, "Distributions of eigenvalues for some sets of random matrices," *Math. USSR Sb.*, vol. 1, no. 4, 1967.

[3] J. W. Silverstein, "The smallest eigenvalue of a large dimensional wishart matrix," *The Annals of Probability*, vol. 13, no. 4, pp. 1364–1368, Nov. 1985.

[4] S. Geman, "A limit theorem for the norm of random matrices," *The Annals of Probability*, vol. 8, no. 2, pp. 252–261, Apr. 1980.

[5] Z. D. Bai and Y. Q. Yin, "Limit of the smallest eigenvalue of a large dimensional sample covariance matrix," *The Annals of Probability*, vol. 21, no. 3, pp. 1275–1294, July 1993.

[6] Z. D. Bai and J. W. Silverstein, "Exact separation of eigenvalues of large dimensional sample covariance matrices," *The Annals of Probability*, vol. 27, no. 3, pp. 1536–1555, July 1999.

[7] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in MISO broadcast channels with limited feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509–4537, July 2012.

[8] A. Adhikary and J. Nam and J. Y. Ahn and G. Caire , "Joint spatial division and multiplexing-the large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

[9] A. Kammoun, A. Müller, E. Björnson, and M. Debbah, "Linear precoding based on polynomial expansion: large-scale multi-cell MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 861–875, Oct. 2014, arXiv:1310.1799.

[10] M. S. Alouini and A. Kammoun, "The random matrix regime of Maronna's M-Estimator for observations corrupted by elliptical noises," *To be submitted to Journal of Multivariate Analysis*, 2014, http://arxiv.org/abs/1412.8344.

[11] B. Hochwald and S. Vishwanath, "Space-time multiple access: Linear growth in the sum rate," *Proc. IEEE Annual Allerton Conference on Communication, Control, and Computing (Allerton'02)*, 2002.

[12] R. Couillet, F. Pascal, and J. W. Silverstein, "The random matrix regime of Maronna's M-estimator with elliptically distributed samples," *submitted*, 2013, http://arxiv.org/abs/1311.7034.

[13] W. Hachem, Ph. Loubaton, and J. Najim, "Deterministic equivalents for certain functionals of large random matrices," *Annals of Applied Probability*, vol. 17, no. 3, pp. 875–930, June 2007.

[14] W. Hachem, O. Khorunzhy, P. Loubaton, J. Najim, and L. A. Pastur, "A new approach for capacity analysis of large dimensional multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3987–4004, Sept. 2008.

[15] Z. D. Bai and J. W. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices," *Annals of Probability*, vol. 26, no. 1, pp. 316–345, Jan. 1998.

[16] P. Debashis and J. W. Silverstein, "No eigenvalues outside the support of limiting empirical spectral distribution of a separable covariance matrix," *Journal Of Multivariate Analysis*, vol. 100, no. 1, Jan. 2009.

[17] Z. D. Bai and J. W. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of information-plus-noise type matrices," *Random Matrices: Theory and Applications*, , no. 2, pp. 1150004, Jan. 2012.

[18] M. Capitaine, C. Donati-Martin, and D. Féral, "The largest eigenvalue of finite rank deformation of large wigner matrices," *Annals of Probability*, vol. 37, no. 1, Jan. 2009.

[19] U. Haagerup and S. Thorbjørnsen, "A new application of random matrices: $\mathrm{Ext}(\mathrm{C}^*_{\mathrm{red}}(\mathrm{F2}))$ is not a group," *Annals of Mathematics*, vol. 162, no. 2, pp. 711–775, Sept. 2005.

[20] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, Aug. 1995.

[21] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sept. 1995.