SUBSPACE KERNEL SPECTRAL CLUSTERING OF LARGE DIMENSIONAL DATA

BY ABLA KAMMOUN*, ROMAIN COUILLET†

Let $x_1, \cdots, x_n$ be independent observations of size p, each of them belonging to one of c distinct classes. We assume that observations within class a are characterized by their distribution $\mathcal{N}(0, \frac{1}{p}C_a)$ where here $C_1, \cdots, C_c$ are some non-negative definite $p \times p$ matrices. This paper studies the asymptotic behavior of the symmetric matrix $\tilde{\Phi}_{kl} = \sqrt{p}\left((x_k^T x_l)^2 \boldsymbol{\delta}_{k\neq l}\right)$ when p and n grow to infinity with $\frac{n}{p} \to c_0$. Particularly, we prove that, if the class covariance matrices are sufficiently close in a certain sense, the matrix $\tilde{\Phi}$ behaves as a low-rank perturbation of a Wigner matrix, presenting possibly some isolated eigenvalues outside the bulk of the semi-circular law. We carry out a careful analysis of some of the isolated eigenvalues and eigenvectors of matrix $\tilde{\Phi}$, and illustrate how these results can help understand spectral clustering methods that use $\tilde{\Phi}$ as a kernel matrix.

1. Introduction. The past ten to fifteen years have witnessed the development of kernel methods to solve many problems in the field of machine learning. They have frequently been used thus far for such problems as classification, clustering, regression, as well as principal component analysis, and have shown to exhibit better performances than traditional statistical techniques [Schölkopf and Smola, 2001, Shawe-Taylor and Cristianini, 2004]. Central to kernel methods is the notion of kernel matrices, the construction of which is as follows. Let $x_1, \cdots, x_n$ be n observations in $\mathbb{R}^p$, the entries of the kernel matrix are given by:

$$K = \begin{cases} k(x_i, x_j), & i \neq j \\ 0, & i = j. \end{cases} \tag{1}$$

where k is a function of two variables, referred to as the kernel. A common choice of kernels comprises, for instance, $k(x_i, x_j) = f(x_i^T x_j)$ where f is a function possibly depending on p. Kernel methods operate exclusively with the kernel matrix, be it by computing its principal eigenvectors like in kernel clustering [Ng et al., 2002] or by solving a convex problem involving it, as in support vector machine algorithms [Hofmann et al., 2008]. The choice of the kernel function has a fundamental importance in the performance of kernel methods. The question of how to choose the kernel, although often posed, has thus far received no convincing answer.

It is with this ultimate goal in sight that recent works within the statistics and mathematics community have emerged. A fundamental assumption in these works is to consider the observations as samples of some random process. As problems of modern machine learning involve data sets with the number of observations n commensurable with their dimensions, the "large n large p" regime is now being considered, i.e., both n and p go to infinity at the same rate. These works can be viewed as part of a more general large random matrices framework, (see [Couillet and Debbah, 2011] and references therein).

*The research reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST).

†The work of Couillet is supported by the ANR Project RMT4GRAPH (ANR-14-CE28-0006) and the HUAWEI RMTin5G project.

When it comes to handling non-linear kernels, scaling inside or outside of the kernel function does not lead to the same asymptotic behavior. This distinction in behavior has led to two different results. The first results concern kernel matrices whose elements are of the form $\left\{f(x_i^T x_j)\right\}_{i,j=1}^n$ where $\{x_i\}_{i=1}^n$ are independent and identically distributed random vectors with covariance $\frac{1}{p}C$, C being a $p \times p$ matrix of bounded spectral norm. Under this setting, the off-diagonal elements are of order $O(\frac{1}{\sqrt{p}})$ whereas the diagonal elements are $O(1)$. If k is sufficiently smooth, it was shown in [El Karoui, 2010] that the kernel matrix behaves as

$$\overline{K} = \left(f(0) + f''(0)\frac{\operatorname{tr} C^2}{2p^2}\right) 11^T + f'(0)X^T X + \left(f\left(\frac{\operatorname{tr} C}{p}\right) - f(0) - f'(0)\frac{\operatorname{tr} C}{p}\right) I_n. \tag{2}$$

where $X = [x_1, \cdots, x_n]$ in the sense that $\left\|K - \overline{K}\right\| \to 0$ almost surely in operator norm. The limiting spectral distribution of K, thus, coincides up to an adequate shifting and rescaling, with the Marcenko-Pastur law. It appears in this case that the reason behind obtaining this equivalence is related to the convergence of the off-diagonal elements of $X^T X$ to 0. As the kernel function is applied element-wise to $X^T X$, all off-diagonal elements of the kernel matrix converge to $f(0)$. The Gram matrix $X^T X$ naturally appears upon applying a Taylor expansion on each off-diagonal element at the vicinity of zero. A more involved model, allowing observations to be drawn from different covariances, was considered in [Couillet and Benyach-Georges, 2016], and has also led to a random matrix equivalent that, up-to a deformation matrix is given by a classical random Gram matrix. From a practical standpoint, this result is somewhat disappointing, as it suggests that kernel random matrices are up to a linear transformation equivalent to standard Gram matrices. One would obtain the same results, if random equivalents (as the one in (2)) were used instead of the so-involved kernel matrix. This observation has led researchers to consider kernel matrices with different scalings [Cheng and Singer, 2013]. To avoid equivalence with a classical random matrix model, the following kernel random matrix has been considered,

$$K_{ij} = \frac{1}{\sqrt{p}} f(\sqrt{p} x_i^T x_j) \boldsymbol{\delta}_{i \neq j} \tag{3}$$

where herein the multiplication by $\sqrt{p}$ inside K serves to produce fluctuating off-diagonal elements. The asymptotic behavior of K defined in (3) has been studied in [Fan and Montanari, 2015] for i.i.d. observations with identity covariance. A completely different behavior was highlighted, by which the kernel random matrix is now asymptotically equivalent to a sort of a deformed Wigner matrix, where the deformation is non-linear with respect to the kernel in use. Studying the behavior of the kernel matrix with the second scaling should reveal new interesting aspects surrounding the effect of the kernel function, that could not be unveiled by the works in [El Karoui, 2010]. This lies behind the principal motivation of the present work, which will rather focus on the second scaling.

Differently from the aforementioned works, we consider herein the case in which $x_1, \cdots, x_n$ are Gaussian independent random vectors and belong to c different classes, where observations belonging to class a have zero mean and covariance $\frac{1}{p}C_{[a]}$ where $\frac{1}{p}\operatorname{tr} C_{[a]} = 1$. We consider the practical situation in which we perform kernel spectral clustering based on the kernel matrix

$$\tilde{\Phi} = \sqrt{p}(x_k^T x_\ell)^2 \boldsymbol{\delta}_{k \neq l}.$$

which can be easily seen to be an instance of the model in (3), obtained by choosing $f = x^2$. In the considered setting, discrimination upon data will be based on the difference in covariances, as all observations have zero means. For clustering to be asymptotically non-trivial, an assumption gauging the distance between covariance matrices across the classes need to be considered. It turns out that in our case, we should assume that for a and b indexing two different classes, $\frac{1}{p}\operatorname{tr} C^\circ_{[a]} C^\circ_{[b]} = \frac{1}{p}\operatorname{tr}(C^\circ)^2 + O(p^{-\frac{1}{2}})$ with $C^\circ = \sum_{d=1}^{c} \frac{n_d}{n} C_{[d]}$, n_d being the number of observations in class d and $C^\circ_{[d]} = C_{[d]} - C^\circ$.

To study the asymptotic behavior of $\tilde{\Phi}$, it is an essential step to start by investigating that of Φ :

$$[\Phi]_{ij} = \sqrt{p}\left((x_i^T x_j)^2 - \frac{1}{p^2}\operatorname{tr} C_i C_j\right) \boldsymbol{\delta}_{i\neq j}$$

obtained by centering the off-diagonal elements of $\tilde{\Phi}$ around their expected values. An important outcome of our analysis is that matrix Φ behaves as a standard Wigner matrix presenting possibly isolated eigenvalues that escape from the bulk of the semi-circle law. Neither these isolated eigenvalues, nor those associated with the bulk of the semi-circle law do carry information about the classes, which implies that matrix Φ act essentially as a noise matrix. It is indeed, in $\tilde{\Phi}$, the matrix formed by the means of the off-diagonal elements that carry all the information about classes. This matrix, being of low rank, can produce a finite number of eigenvalues outside the bulk of the semi-circle law, the positions of which and their associated eigenvectors are characterized and are shown to capture the necessary information surrounding classes. More precisely, our main technical contributions can be summarized as follows:

- We show (Theorem 1) that in the asymptotic regime wherein n, p grow to infinity with $\frac{n}{p} \to c_0$ and $\frac{n_a}{n} \to c_a$, the matrix Φ behaves as a real symmetric Wigner matrix, in the sense that its empirical eigenvalue distribution converges towards the semi-circle law. This result is in perfect agreement with that of [Cheng and Singer, 2013], which asserts that the asymptotic behavior will involve only the contribution of a Wigner matrix once $f'(0) = 0$. Note that the result in [Cheng and Singer, 2013] is restricted to the case of standard Gaussian random matrices, and as such could not be used to handle our specific setting. Moreover, our approach is very different from [Cheng and Singer, 2013] and mainly relies on Gaussian calculus tools as the basic instruments.
- We analyze the asymptotic behavior of bilinear forms associated with matrix Φ (Theorem 2). Particularly, we highlight a striking difference in the behavior of these quantities that, to the authors' knowledge, has never been encountered when dealing with Gram random matrices.
- We show that almost surely for n large enough, the limiting support is composed of the support of the semi-circle law plus possibly two spikes, the positions of which are derived. Moreover, almost surely, all eigenvalues lie within a neighborhood of the limiting support.
- Finally, to allow a thorough understanding of the clustering performance, we characterize the leading eigenvectors and eigenvalues of the kernel matrix that carry information.

Notations: In the remainder of the article, uppercase characters will stand for matrices, lowercase for scalars or vectors. The transpose operation will be denoted $(.)^T$. The multivariate Gussian distribution of mean μ and covariance C will be denoted $\mathcal{N}(\mu, C)$. The notation $V = \{V_{ij}\}_{i,j=1}^{n}$ denotes the matrix with $(i,j)-$ entry V_{ij} (scalar or matrix) $1 \leq i \leq n$, $1 \leq j \leq T$ while $\{V_i\}_{i=1}^{n}$ is the row-wise concatenation of the V_i's and $\{V_j\}_{j=1}^{T}$ the column-wise concatenation of the V_j's. For scalars x_p and r_p, $x_p = O(r_p)$ means that there exists a constant K independent of p and n such that $|x_p| \leq K|r_p|$. For random variables, the notation $o(1)$ refers to a random quantity that converges to 0 almost surely while the notation $O(r_p)$ refers to a random quantity that is bounded in probability at rate r_p. Finally, for deterministic scalars x_p and v_p the notation $x_p = O_z(v_p)$ means that $|x_p| \leq v_p P(|z|)Q(|\Im z|^{-1})$ for some polynomials P and Q with non-negative coefficients and whose parameters are independent of the dimensions n and p.

2. Assumptions and main results. Consider p-dimensional independent real Gaussian vectors $x_1, \ldots, x_n$. For $n_1, \ldots, n_c$ such that $n_1 + \ldots + n_c = n$, we assume that

$$x_{n_1+\ldots+n_{j-1}+1}, \ldots, x_{n_1+\ldots+n_j} \sim \mathcal{N}(0, p^{-1}C_{[j]})$$

for $C_{[1]}, \ldots, C_{[c]} \in \mathbb{R}^{p\times p}$. Let $j \in \{1, \cdots, n\}$. Then, for all integer $k \in \left[\sum_{r=1}^{j-1} n_r + 1, \sum_{r=1}^{j} n_r\right]$, we define $C_k = C_{[j]}$. When $x_i \sim \mathcal{N}(0, p^{-1}C_{[j]})$, we shall say that $x_i \in \mathcal{C}_j$.

Further define $C^\circ = \sum_{i=1}^{c} \frac{n_i}{n} C_{[i]}$ and, for each i, $C^\circ_{[i]} = C_{[i]} - C^\circ$. The matrices $C_{[1]}, \ldots, C_{[c]}$ additionally satisfy the following rate conditions.

Assumption 1 (Growth Rates). *As $p \to \infty$, we have the following assumptions:*

(i) $n/p = c_0 \in (0, \infty)$
(ii) *for each* $a \in \{1, \ldots, k\}$, $n_a/n = c_a \in (0, \infty)$
(iii) *for each* $a \in \{1, \ldots, k\}$, $\frac{1}{p} \operatorname{tr} C_{[a]} = 1$ *and* $\forall\ a, b \in \{1, \cdots, k\}$, $\frac{1}{p} \operatorname{tr} C^\circ_{[a]} C^\circ_{[b]} = O(p^{-\frac{1}{2}})$
(iv) *all matrices* $C_{[k]}$, $k = 1, \cdots, c$ *have bounded spectral norm, that is:*

$$\max_{1\leq k\leq c} \limsup_p \|C_{[k]}\| < \infty.$$

We shall further assume that $\frac{1}{p} \operatorname{tr} (C^\circ)^2$ *and* $\frac{1}{p} \operatorname{tr} (C^\circ)^4$ *converge, and define:*

$$\omega = \sqrt{2} \lim_{p\to\infty} \frac{1}{p} \operatorname{tr}(C^\circ)^2$$

$$\Omega = \sqrt{2} \sqrt{\lim_{p\to\infty} \frac{1}{p} \operatorname{tr}(C^\circ)^4}$$

Moreover, we assume that:

$$\max\left(\left| \frac{\Omega}{\sqrt{2}} - \sqrt{\frac{1}{p} \operatorname{tr} (C^\circ)^4} \right|, \left| \frac{\omega}{\sqrt{2}} - \frac{1}{p} \operatorname{tr} (C^\circ)^2 \right| \right) \leq K p^{-\frac{1}{4}}$$

for some constant K independent of p.

A direct consequence of Item *iii)* in Assumption 1 is that:

$$\frac{1}{p} \operatorname{tr} C^\circ_a C^\circ = O(p^{-\frac{1}{4}})$$

which unfolds from the Cauchy-Schwartz inequality:

$$\left| \frac{1}{p} \operatorname{tr} C^\circ_a C^\circ \right| \leq \sqrt{\frac{1}{p} \operatorname{tr} (C^\circ)^2} \sqrt{\frac{1}{p} \operatorname{tr} (C^\circ_a)^2}.$$

As shall be seen later, the above regime is mostly motivated by machine learning applications, in which we aim at clustering observations, (here denoted by $x_1, \cdots, x_n$ drawn from a c class-Gaussian mixture. Unlike the work in [Couillet and Benyach-Georges, 2016], herein we assume that all observations have zero-mean and as such clustering is made upon the distance between the covariance matrices. While [Couillet and Benyach-Georges, 2016] already covers this case, it requires a stronger assumption on the difference between the covariance matrices, that $\frac{1}{p}\operatorname{tr} C^\circ_{[a]} C^\circ_{[b]}$ be of order $O(1)$, to allow non-trivial clustering. In this paper, we perform clustering using the kernel

$$\tilde{\Phi} = \sqrt{p}\left\{\left(x_i^T x_j\right)^2 \boldsymbol{\delta}_{i\neq j}\right\}$$

For this specific kernel, we prove that non-trivial clustering is possible under the less stringent assumption in (iii). To assess the clustering performance, it is a fundamental first step to understand the asymptotic spectral behavior of the kernel matrix $\tilde{\Phi}$. This forms the main objective of the present work. We will proceed in two steps. First, we will study the asymptotic behavior of matrix Φ obtained by element-wise centering of $\tilde{\Phi}$:

$$[\Phi]_{ij} = \sqrt{p}\left((x_i^T x_j)^2 - \frac{1}{p^2}\operatorname{tr} C_i C_j\right)\boldsymbol{\delta}_{i\neq j}$$

in the growth regime defined in Assumption 1. Particularly, our main results are as follows:

1) The empirical spectral distribution of matrix Φ converges almost surely towards the semi-circle distribution (Theorem 1)
2) Bilinear forms associated with matrix Φ have deterministic equivalents in the large n, p regime which we characterize (Theorem 2)
3) Almost surely, for n large enough, all the eigenvalues of Φ are located in a neighborhood of the semi-circle distribution plus possibly two spikes at positions $c_0\Omega + \frac{1}{\sqrt{c_0}}\frac{\omega}{\Omega}$ and $-c_0\Omega - \frac{1}{\sqrt{c_0}}\frac{\omega}{\Omega}$. (Theorem 4)

These results set a stage to the second part of our work (section 3) dealing with spectral clustering applications using the kernel matrix $\tilde{\Phi}$.

For the first part of our work, the fundamental tool is the Stieltjes transform. For $z \in \mathbb{C}\backslash\mathbb{R}$, we denote the resolvent of matrix Φ by:

$$Q(z) = (\Phi - zI_n)^{-1}$$

and the Stieltjes transform of the expectation of the empirical measure of the eigenvalues of Φ by:

$$g_n(z) = \frac{1}{n}\operatorname{tr}\mathbb{E}Q(z)$$

We will prove that:

$$\frac{1}{n}\operatorname{tr} Q(z) \xrightarrow{\text{a.s.}} m(z)$$

where $m(z)$ is the unique Stieltjes transform solution of the following fixed point-equation:

$$m(z) = -\frac{1}{z + c_0\omega^2 m(z)}.$$

This allows us to achieve the first goal of the present work, namely to prove the convergence of the empirical distribution of matrix Φ to the semi-circle distribution. The latter result is formally stated in the following theorem, the proof of which is postponed to Section 5.

THEOREM 1. *Let Assumption 1 hold true. Denote by $\lambda_1, \cdots, \lambda_n$ the eigenvalues of Φ. Then, the empirical spectral distribution $\mu_n = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\delta}_{\lambda_i}$ converges almost surely (in the weak convergence of probability measures) to the probability measure μ with density:*

$$\mu(dt) = \frac{1}{2\pi c_0 \omega^2}\sqrt{(4c_0\omega^2 - t^2)^+}dt$$

having support $\mathcal{S} = \left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right]$.

Theorem 1 can be leveraged to approximate in the almost sure sense functionals of the eigenvalues of matrix Φ by virtue of the Portmanteau Lemma, thus leading to the following corollary:

COROLLARY 1. *Let Assumption 1 hold true and f be a continuous bounded function. Then,*

$$\int f(\lambda)\mu_n(d\lambda) - \int f(\lambda)\mu(d\lambda) \xrightarrow{\text{a.s.}} 0.$$

Corollary 1 prefigures the asymptotic behavior of functionals of the eigenvalues of matrix Φ but cannot be used to infer on that of the eigenvectors. As shall be seen subsequently, a key step in analyzing the asymptotic behavior of the eigenvectors of Φ (or a perturbation of it) is to characterize the asymptotic behavior of bilinear forms associated with the resolvent matrix in the form of $a^T Q(z) b$ where a and b are two vectors in $\mathbb{C}^{n\times 1}$. Such a quantity has thoroughly been studied for a large variety of random matrix models. The outcome of all existing works addressing this question is that, the resolvent matrix Q can be considered as equivalent to a matrix T depending only in the statistics of the observations in the sense that $\frac{1}{n}\operatorname{tr} Q - \frac{1}{n}\operatorname{tr} T$ and $a^T Q b - a^T T b$ converge to zero in the almost sure sense. In case the entries are centered and with the same variance, matrix $\mathbf{T}$ has been found in all studied cases to be proportional to identity. Based on these results, it might seem quite natural to expect $a^T Q(z) b$ to be well-approximated by $m(z) a^T b$. Quite surprisingly and contrary to these prior expectations, we find that the present case is, to the best of our knowledge, the first to provably make an exception to this rule. We have indeed the following result whose proof is deferred to Section 6:

THEOREM 2. *Consider the setting of Assumption 1. Let a and b be two vectors in $\mathbb{C}^{n\times 1}$ with bounded Euclidean norm. Then, for any $z \in \mathbb{C}\backslash\mathbb{R}$,*

$$a^T Q(z) b - \left(a^T b\right) m(z) - \frac{c_0 \Omega^2 m^3(z)}{1 - \Omega^2 c_0^2 m^2(z)} a^T \frac{11^T}{p} b \xrightarrow{\text{a.s.}} 0.$$

When quadratic forms involve more than one resolvent matrix, asymptotic limits are more involved as shown by the following theorem.

THEOREM 3. *Consider the setting of Assumption 1. Let a and b be two vectors in $\mathbb{C}^{n\times 1}$ with bounded Euclidean norm. Let D be a diagonal matrix with uniformly bounded diagonal* *elements. Define for* $z_1, z_2 \in \mathbb{C}\backslash\mathbb{R}$, $g(z_1, z_2)$ *as:*

$$\begin{aligned} g(z_1,z_2) &= (1-\omega^2 c_0 m(z_1)m(z_2))^{-1} m(z_1)m(z_2)a^T b \\ &+ a^T \frac{11^T}{p} b \left[1-\Omega^2 c_0^2 m^2(z_2)\right]^{-1} \left[1-\Omega^2 c_0^2 m^2(z_1)\right]^{-1} \left[1-\omega^2 c_0 m(z_1)m(z_2)\right]^{-1} \\ &\times c_0 \Omega^2 m(z_1)m(z_2) \left[m^2(z_1)+m^2(z_2)+1-c_0^2\Omega^2 m^2(z_1)m^2(z_2)\right] \end{aligned}$$

Then,

$$a^T Q(z_1) D Q(z_2) b - m(z_1)m(z_2) a^T D b - \tilde{r}(z_1,z_2) \xrightarrow{\text{a.s.}} 0.$$

where:

$$\begin{aligned} \tilde{r}(z_1,z_2) &= \left[1-\Omega^2 c_0^2 m^2(z_2)\right]^{-1} c_0 \Omega^2 m^3(z_2) m(z_1) \frac{a^T D 11^T b}{p} \\ &+ \left[1-\Omega^2 c_0^2 m^2(z_1)\right]^{-1} c_0 \Omega^2 m^3(z_1) m(z_2) \frac{a^T 11^T D b}{p} \\ &+ \frac{1}{p}\operatorname{tr} D \Omega^2 \frac{a^T 11^T b}{p} m^2(z_1)m^2(z_2) \left(1-\Omega^2 c_0^2 m^2(z_1)\right)^{-1} \left(1-\Omega^2 c_0^2 m^2(z_2)\right)^{-1} \left(1+c_0^2\Omega^2 m(z_1)m(z_2)\right) \\ &+ \left(\frac{1}{p}\operatorname{tr} D\right) \omega^2 m(z_1)m(z_2) g(z_1,z_2) \end{aligned}$$

PROOF. See Appendix G □

Theorem 1 and Theorem 2 imply that the resolvent matrix Q is equivalent to

$$\overline{Q} = m(z) I_n + \frac{\Omega^2 c_0 m^3(z)}{1-\Omega^2 c_0^2 m^2(z)} \frac{11^T}{p}$$

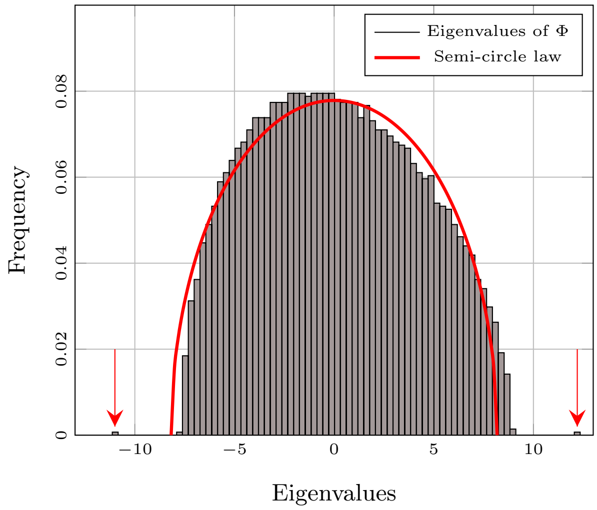
where the equivalence is in the sense that $\frac{1}{n}\operatorname{tr} AX_n - \frac{1}{n}\operatorname{tr} AY_n \to 0$ and $a_n^T(X_n - Y_n)b_n \to 0$ for every deterministic matrix A with bounded spectral norm and vectors a_n, b_n having bounded Euclidean norms. Theorem 2 suggests that Φ might possess a spike outside the support of the semi-circle law, resulting from the rank one matrix $\frac{c_0\Omega^2 m^3(z)}{1-\Omega^2 c_0^2 m^2(z)} a^T \frac{11^T}{p} b$ in $\overline{\Phi}$. Such a spike should correspond to the real values x for which $m^2(x) = \frac{1}{\Omega^2 c_0^2}$. This question is discussed in the following theorem, which confirms the possible existence of spikes outside the main bulk of the semi-circle law.

THEOREM 4 (Almost sure location of the eigenvalues of Φ). *Consider the setting of Assumption 1. Let* $\lambda_1, \cdots, \lambda_n$ *be the eigenvalues of* Φ. *Let* $\tilde{\rho} = c_0\Omega + \frac{\omega^2}{\Omega}$. *For* $\epsilon > 0$, *define* $\mathcal{S}^\epsilon$ *as*

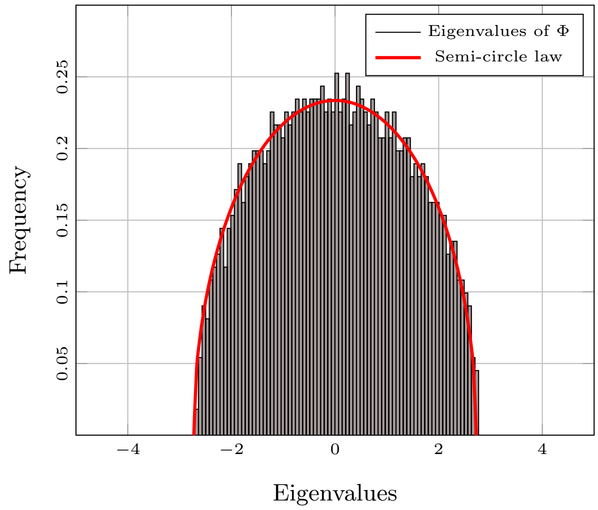
$$\mathcal{S}^\epsilon = \begin{cases} \left[-2\sqrt{c_0}\omega - \epsilon, 2\sqrt{c_0}\omega + \epsilon\right] & \text{if } \Omega \leq \frac{1}{\sqrt{c_0}}\omega \\ \left[-2\sqrt{c_0}\omega - \epsilon, 2\sqrt{c_0}\omega + \epsilon\right] \cup [\tilde{\rho}-\epsilon, \tilde{\rho}+\epsilon] \cup [-\tilde{\rho}-\epsilon, -\tilde{\rho}+\epsilon] & \text{otherwise.} \end{cases}$$

Then, for all large n *almost surely:*

$$\{\lambda_i, 1 \leq i \leq n\} \cap \mathbb{R}\backslash\mathcal{S}^\epsilon = \emptyset.$$



(a)



(b)

Fig 1: Histogram of the eigenvalues of Φ and the semi-circle law, for (a) $n = 4800$ and $p = 1600$ and (b) $n = 1600$ and $p = 4800$. All C_i's are equal to I_p. The semi-circle law is superposed in red, and the locations of two observed spikes is highlighted with red arrows.

Remark 1. From Cauchy-Shwartz inequality, it entails that $\sqrt{\frac{1}{p}\operatorname{tr}(C^\circ)^4} \geq \frac{1}{p}\operatorname{tr}(C^\circ)^2$ and hence, if $c_0 \geq 1$, $\Omega \geq \frac{1}{\sqrt{c_0\omega}}$. In such a case, we expect at least two spikes at positions $\tilde{\rho}$ and $-\tilde{\rho}$ escape from the main bulk of the semi-circle law. While in theory Theorem 4 could not infer exactly on the exact number of the spikes, simulations in Fig.1 suggest that there are exactly 2 spikes at position $\tilde{\rho}$ and $-\tilde{\rho}$.

Remark 2. The result in Theorem 4 is in agreement with [Fan and Montanari, 2015, Theorem 1.7], which shows that under the i.i.d. case with all C_p's equal to identity, polynomial kernel matrices might have two spikes outside the main bulk of the semi-circle law. Our work, in this respect, although restricting to $f(x) = x^2$, extends the result in [Fan and Montanari, 2015] to independent but not identically distributed observations.

3. Applications: Spectral clustering using $\left\{\sqrt{p}(x_i^T x_j)^2 \delta_{i\neq j}\right\}$.

In this section, we will show how the previous results can allow a better understanding of the performance of subspace clustering based on the kernel matrix:

$$\tilde{\Phi} = P\left\{(x_i^T x_j)^2 \boldsymbol{\delta}_{i\neq j}\right\}_{i,j=1}^n P.$$

where herein $x_1, \cdots, x_n$ are n zero-mean observations following Assumption 1 and $P = I_n - \frac{11^T}{n}$. The left and right multiplication by matrix P is a processing that is commonly known in the machine learning terminology as centering in the feature space.

To begin with, we first decompose $\tilde{\Phi} \triangleq \left\{\sqrt{p}(x_i^T x_j)^2 \delta_{i\neq j}\right\}$ as:

$$\tilde{\Phi} = \Phi + \left\{\frac{1}{p^{\frac{3}{2}}}\operatorname{tr} C_i^\circ C_j^\circ\right\}_{i,j=1}^n + 1\psi^T + \psi 1^T + \frac{1}{p^{\frac{3}{2}}}\operatorname{tr}(C^\circ)^2\, 11^T - \operatorname{diag}\left\{\frac{1}{p^{\frac{3}{2}}}\operatorname{tr} C_i^2\right\}_{i=1}^n \tag{4}$$

where $\psi = \left\{ \frac{1}{p^{\frac{3}{2}}} \operatorname{tr} C_i^\circ C^\circ \right\}_{i=1}^n$.

The above decomposition distinguishes four types of matrices, the spectral norms of which scale at different orders of magnitude:

- Matrix Φ, behaving like a symmetric Wigner matrix with eventually a small proportion of spikes outside the support (at most 2 by simulations), and matrix $\left\{ \frac{1}{p^{\frac{3}{2}}} \operatorname{tr} C_i^\circ C_j^\circ \right\}_{i,j=1}^n$ have $O(1)$ spectral norms,
- Matrix $\psi 1^T + 1^T \psi$ has $O(p^{\frac{1}{4}})$ spectral norm,
- Matrix $\frac{1}{p^{\frac{3}{2}}} \operatorname{tr} (C^\circ)^2 11^T$ has $O(p^{\frac{1}{2}})$ spectral norm
- Matrix $\operatorname{diag} \left\{ \frac{1}{p^{\frac{3}{2}}} \operatorname{tr} C_i^2 \right\}_{i=1}^n$ has $O(p^{-\frac{1}{2}})$ spectral norm,

Upon left and right multiplication by matrix P, we obtain:

$$P\tilde{\Phi}P = P\Phi P + \left\{ \frac{1}{p^{\frac{3}{2}}} \operatorname{tr} C_i^\circ C_j^\circ \right\}_{i,j=1}^n + O_{\|.\|}(\frac{1}{\sqrt{p}})$$

where $O_{\|.\|}(\frac{1}{\sqrt{p}})$ stands for a matrix whose spectral norm is $O(p^{-\frac{1}{2}})$. Matrix $P\tilde{\Phi}P$ follows the so-called spiked random matrix model in the sense that it is the sum of a high-rank matrix $P\Phi P$ and a deterministic matrix $\left\{ \frac{1}{p^{\frac{3}{2}}} \operatorname{tr} C_i^\circ C_j^\circ \right\}_{i,j=1}^n$ having rank at most $c-1$. As such, in the asymptotic regime, it is expected that the eigenvalues of $P\tilde{\Phi}P$ be asymptotically the same as $P\Phi P$ except for possibly finitely many of them which are allowed to escape from the main bulk of $P\Phi P$. A thorough analysis of the localization of these eigenvalues will be carried out in the next section.

3.1. *Isolated eigenvalues.*

THEOREM 5. *Let Assumption 1 hold true. Then, as $n \to \infty$, the empirical spectral distribution of $P\tilde{\Phi}P$ almost surely converges (in the weak limit of probability measures) to the probability measure μ defined in Theorem 1 with support $\mathcal{S}$. Assume further that $\frac{1}{\sqrt{p}} \operatorname{tr} C_a^\circ C_b^\circ$ converges in $[0, \infty)$, and define $\mathcal{T}$ as:*

$$\mathcal{T} = \left\{ \lim_{p\to\infty} \frac{\sqrt{c_a c_b}}{\sqrt{p}} C_{[a]}^\circ C_{[b]}^\circ \right\}_{a,b=1}^c$$

Then, for all large n almost surely, there are at most $k-1$ eigenvalues of $P\tilde{\Phi}P$ lying at a macroscopic distance of $\mathcal{S}$. They are defined as follows. Let $\eta_1 \geq \cdots \geq \eta_{c-1}$ be the strictly positive eigenvalues of $\mathcal{T}$. Then, for each $i \in \{1, \cdots, c-1\}$, if $\sqrt{c_0}|\nu_i| > \omega$ and $\eta_i \neq \Omega$, for all large n almost surely, $P\tilde{\Phi}P$ has an isolated eigenvalue λ_i satisfying:

$$\lambda_i \xrightarrow{\text{a.s.}} \rho_i \triangleq c_0 \eta_i + \frac{\omega^2}{\eta_i}$$

PROOF. See Appendix H. □

3.2. *Eigenvectors.* From the previous section, it appears that $P\tilde{\Phi}P$ has the same behavior as an additive random spiked model in which the finite rank perturbation spanned by the columns of matrix U. Noting that U is importantly constiluted of vectors j_a, it is expected that the isolated eigenvectors align to the canonical basis J, and this is all the more true that the eigenvalues of M are large. From a spectral clustering perspective, it is a fundamental question to evaluate how much the isolated eigenvectors align with the set of vector $\{j_a\}_{a=1}^c$. Let $\hat{u}_i$ be the i-th isolated eigenvector of $P\tilde{\Phi}P$. We may decompose $\hat{u}_i$ as:

$$\hat{u}_i = \sum_{a=1}^{c} \alpha_a^i \frac{j_a}{\sqrt{n_a}} + \sigma_a^i \omega_a^i$$

where ω_i^a is a vector of unit norm, supported on the indices of class a and orthogonal to j_a, while $\alpha_i^a \in \mathbb{R}$ and σ_i^a are scalars to be determined.

THEOREM 6. *Let (η_i, v_i) an eigenpair of $\mathcal{T}$ with η_i of unit multiplicity satisfying $\sqrt{c_0}|\eta_i| > \omega$. Define for each $a \in \{1, \cdots, c\}$ $\alpha_i^a = \frac{1}{\sqrt{n_a}} \hat{u}_i^T j_a$ with $(\lambda_i, \hat{u}_i)$ the eigenpair of $P\tilde{\Phi}P$ mapped to (η_i, v_i) as per Theorem 4. Assume that all the eigenvalues of $\mathcal{T}$ are different from Ω. Then, under the same setting of Theorem 4, for each $a, b \in \{1, \cdots, c\}$.*

$$\alpha_i^a \alpha_i^b \xrightarrow{\text{a.s.}} \left(1 - \frac{1}{c_0} \frac{\omega^2}{\eta_i^2}\right) \left[v_i v_i^T\right]_{ab}$$

PROOF. See Appendix I □

REMARK 3. (Expression of α_i^a)) Since the eigenvectors are defined up to a sign, we may impose without restriction that $\alpha_i^s > 0$ for s the smallest index a for which $\alpha_i^a > 0$. Thus from Theorem 6, we find for each a,

$$\alpha_i^a \xrightarrow{\text{a.s.}} \operatorname{sign}\left(\left[v_i v_i^T\right]_{sa}\right) \sqrt{\left(1 - \frac{1}{c_0} \frac{\omega^2}{\eta_i^2}\right) \left[v_i v_i^T\right]_{aa}}$$

Similarly, for two isolated eigenvalues λ_i and λ_j of $P\tilde{\Phi}P$, it shall be interesting to study the correlation:

$$\begin{aligned} \sigma_{ij}^a &\triangleq (\hat{u}_i - \alpha_i^a \frac{j_a}{\sqrt{n_a}})^T D_a (\hat{u}_j - \frac{1}{\sqrt{n_a}} \alpha_j^a j_a) \\ &= \hat{u}_i^T D_a \hat{u}_j - \alpha_i^a \alpha_j^a. \end{aligned}$$

where $D_a = \operatorname{diag}\{j_a\}$. This is the purpose of the following Theorem:

THEOREM 7. *Let $(\lambda_i, \hat{u}_i)$ and (λ_j, u_j) two eigenpairs of $P\tilde{\Phi}P$ mapped to the eigenpairs (η_i, v_i) and (η_j, v_j) of $\mathcal{T}$. Assume that $\min\left(\sqrt{c_0}|\eta_i|, \sqrt{c_0}|\eta_j|\right) > \omega$. Then, under the setting of Theorem 6,*

$$\sigma_{ij}^a \xrightarrow{\text{a.s.}} \boldsymbol{\delta}_{i=j} \frac{c_a}{c_0} \frac{\omega^2}{\eta_i^2}$$

PROOF. See Appendix J □

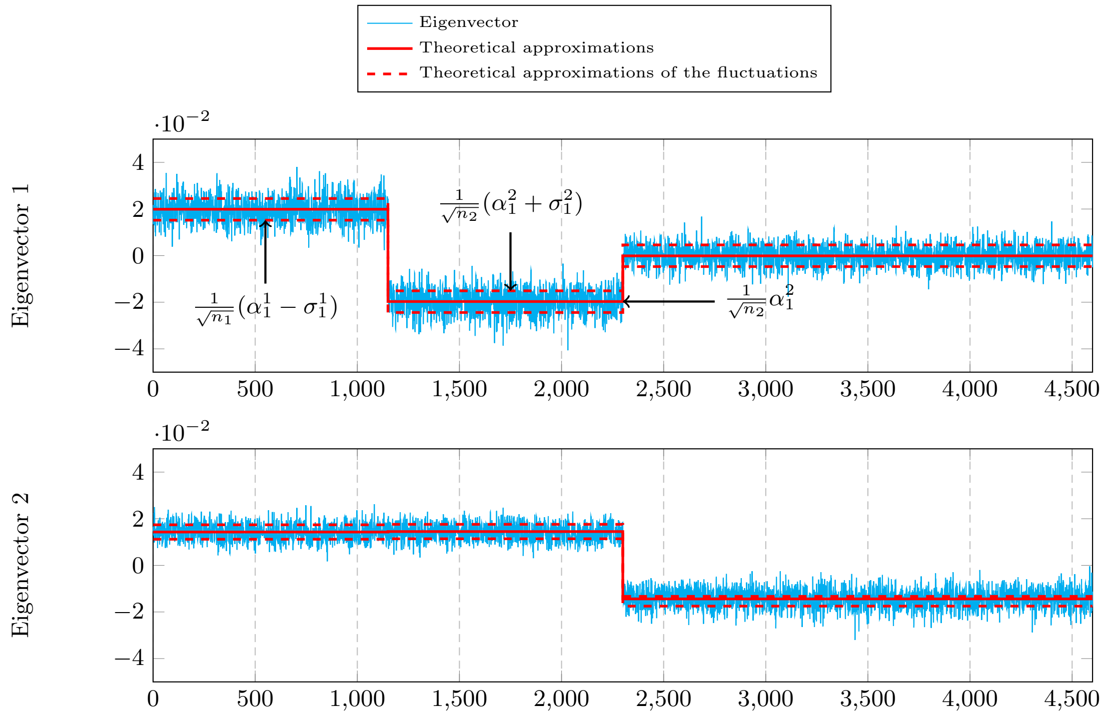


Fig 2: Leading two eigenvectors of $P\Phi P$ versus deterministic approximations of $\alpha_i^a \pm \sigma_i^a$ as per Theorem 6 and Theorem 7, for $n = 4600$, $p = 2560$, $c = 3$, $c_1 = c_2 = 1/4$ and $c_3 = 0.5$, $C_i = I_p + (p/8)^{-\frac{5}{4}} W_i W_i^T$ with $W_i \in \mathbb{R}^{p \times p/8}$ with i.i.d. $\mathcal{N}(0,1)$ entries.

Remark 4. Since $\sigma_{ij}^a \xrightarrow{\text{a.s.}} 0$ whenever $i \neq j$, an interesting outcome of Theorem 7 is that the dominant eigenvectors of matrix $P\tilde{\Phi}P$, have negligible correlation and thus can be treated independently when it comes to clustering. This behavior is a consequence of the fact that, although the x_i's have different covariance matrices per class, Φ asymptotically behaves like a matrix with i.i.d entries and thus does not asymptotically capture the difference in covariances as do the isolated eigenvectors.

For sake of illustration, we represent in Figure 2 the leading two eigenvectors of $P\tilde{\Phi}P$ against the theoretical findings of Theorem 6 and Theorem 7.

4. Mathematical tools and preliminary results.

4.1. *Mathematical tools.* The proof of our results will heavily hinge on the use of Gaussian calculus tools along with some useful identities involving the resolvent matrix, The main features of which are outlined in the sequel:

1. Differentiation formula:

$$\frac{\partial Q_{ik}}{\partial Z_{rs}} = -2 \sum_{b \neq s} (x_b^T x_s) \left[C_s^{\frac{1}{2}} x_b \right]_r (Q_{is} Q_{bk} + Q_{sk} Q_{bi}) \,. \tag{5}$$

2. Integration by Parts formula for Gaussian functionals: Let f be a $\mathcal{C}^1$ function polynomially bounded togother with its derivatives. Consider $Z \in \mathbb{R}^{p \times n}$ a standard normal

Gaussian matrix. Then,

$$\mathbb{E}\left[Z_{ij} f(Z)\right] = \mathbb{E}\left[\frac{\partial f(Z)}{\partial Z_{ij}}\right]. \tag{6}$$

3. Poincaré-Nash inequality: Let Z and f as above, then:

$$\operatorname{var}(f(Z)) \leq \sum_{i=1}^{p} \sum_{j=1}^{n} \mathbb{E}\left[\left|\frac{\partial f(Z)}{\partial Z_{ij}}\right|^2\right].$$

4. Identities involving the resolvent: Define vector $\xi_k \in \mathbb{C}^n$ with elements:

$$[\xi_k]_i = \sqrt{p}\left[(x_k^T x_i)^2 - \mathbb{E}(x_k^T x_i)^2\right]$$

We denote by $\xi_{(k,-k)}$ vector ξ_k where the k-th entry is replaced by zero. Let Φ_k be matrix Φ where we remove the k-th row and k-th column. Let $\tilde{Q}_k \in \mathbb{C}^{n-1\times n-1}$ be given by $\tilde{Q}_k = (\Phi_k - zI_{n-1})^{-1}$ and $Q_k \in \mathbb{C}^{n\times n}$ be $\tilde{Q}_k$ where a zero entries row and column entries are inserted at the k-th position. Then, the diagonal elements of Q satisfies [Bai and Silverstein, 2006, Theorem A.4]:

$$Q_{kk} = \frac{-1}{z + \xi_{(k,-k)}^T Q_k \xi_{(k,-k)}}$$

Furthermore, the off-diagonal element Q_{ik} with $(i \neq k)$ is given by [Bai and Silverstein, 2006, page 471]:

$$Q_{ik} = \frac{e_i^T Q_k \xi_{(k,-k)}}{z + \xi_{(k,-k)}^T Q_k \xi_{(k,-k)}} = -Q_{kk} e_i^T Q_k \xi_{(k,-k)} \tag{7}$$

where e_i denotes the i-th canonical vector of $\mathbb{C}^n$.

4.2. *Preliminary results.* Let A_1, A_2 and A_3 be $n \times n$ deterministic matrices with spectral norm uniformly bounded in n. Let $k \in \{1, \cdots, n\}$. Define matrix S_k as:

$$[S_k]_{b_1 b_2} = \left(x_{b_1}^T A_1 x_k\right)\left(x_{b_2}^T A_2 x_k\right)\left(x_{b_1}^T A_3 x_{b_2}\right) \boldsymbol{\delta}_{k\neq b_1} \boldsymbol{\delta}_{k \neq b_2} \tag{8}$$

The control of the spectral norm of this matrix is central to the proof of our results.

Lemma 1. *Let S_k be as in* (8). *Then,*

$$\|S_k\| = O(p^{-1}).$$

Proof. We can write S_k as:

$$S_k = D_k X^T A_3 X \tilde{D}_k$$

where $D_k = \operatorname{diag}\left\{x_b^T A_1 x_k \boldsymbol{\delta}_{b\neq k}\right\}_{b=1}^n$ and $\tilde{D}_k = \operatorname{diag}\left\{x_b^T A_2 x_k \boldsymbol{\delta}_{b\neq k}\right\}_{b=1}^n$ The result follows since $\|D_k\| = O(p^{-1/2})$, $\|\tilde{D}_k\| = O(p^{-\frac{1}{2}})$ and $\|X^T A_3 X\| = O(1)$ as per [Bai and Silverstein, 1998]. $\square$

LEMMA 2. *Let X and Y be two scalar positive random variables with bounded moments. Assume that $X = O(p^{-r})$ for some $r > 0$. Then, for any $\epsilon > 0$, there exists a constant K such that:*

$$\mathbb{E}[XY] \leq Kp^{-r+\epsilon}.$$

PROOF. Let $\epsilon > 0$. We have:

$$\begin{aligned}\mathbb{E}XY &= \mathbb{E}\left[XY1_{\{X \geq p^{-r+\epsilon}\}}\right] + \mathbb{E}\left[XY1_{X \leq p^{-r+\epsilon}}\right] \\ &\leq \sqrt{\mathbb{E}X^2Y^2}\mathbb{P}\left[X \geq p^{-r+\epsilon}\right] + p^{-r+\epsilon}\mathbb{E}Y\end{aligned}$$

The result of the lemma follows by noticing that $\mathbb{P}[X \geq p^{-r+\epsilon}] = o(p^{-l})$ for any $l > 0$. Taking $l = r - \epsilon$ ends the proof. $\square$

The asymptotic characterization of the behavior of quadratic forms has played a key role in proving many illustrative results of the field of random matrix theory. It turns out that in the currently studied case, quadratic forms of different nature involving vector $\xi_{(k,-k)}$ will arise. Studying these new kinds of quadratic forms is essential to our analysis, and is the purpose of the following lemma.

LEMMA 3 (Behavior of quadratic forms involving vector $\xi_{(k,-k)}$). *Let $k \in \{1, \cdots, n\}$. Let A be a $n \times n$ symmetric matrix independent of x_k. Denote by $\mathbb{E}_{x_k}$ the expectation operator with respect to x_k. Then,*

$$\begin{aligned}\mathbb{E}_{x_k}\left[\xi_{(k,-k)}^T A\xi_{(k,-k)}\right] &= \sum_{i\neq k}\sum_{j\neq k}\left(\frac{1}{\sqrt{p}}x_i^T C_k x_i - \frac{1}{p^{3/2}}\operatorname{tr} C_k C_i\right)\left(\frac{1}{\sqrt{p}}x_j^T C_k x_j - \frac{1}{p^{3/2}}\operatorname{tr} C_k C_j\right)A_{ij} \\ &+ \frac{2}{p}\sum_{i\neq k}\sum_{j\neq k}(x_i^T C_k x_j)^2 A_{ij}\end{aligned}$$

Moreover, we also have:

$$\mathbb{E}\left|\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T A\xi_{(k,-k)}\right|^{2s} = \|A\|^{2s}O(p^{-s+\epsilon}) \tag{9}$$

for $s \in \mathbb{N}^*$ and $s \geq 1$.

PROOF. See Appendix A $\square$

4.3. *Useful properties of the Stieltjes transform of the semi-cirle distribution.*

LEMMA 4. *Let $z \in \mathbb{C}\setminus\left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right]$. Let $m(z)$ be the unique Stieltjes transform solution of the following fixed-point equation:*

$$m(z) = -\frac{1}{z + \omega^2 c_0 m^2(z)}.$$

Then, $m(z)$ satisfies the following properties:

1. *$m(z)$ is analytic in $\mathbb{C}\setminus\left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right]$*

2. $\forall \left\{z\in\mathbb{C}, |z|>2\sqrt{c_0}\omega\right\}$,

$$|m(z)| \leq \frac{1}{|z|-2\sqrt{c_0}\omega}$$

3. Let $\alpha>0$ be a certain constant. Then, it holds that:

$$\left(\left|1-\alpha m^2(z)\right|\right)^{-1} \leq (|z|+\eta)^4 \left(|\Im z|^{-4}+\frac{1}{\frac{1}{2}c_0\omega^2}|\Im z|^{-2}\right). \tag{10}$$

for some constant η. Moreover, if $|z|\geq 2\sqrt{2}\sqrt{c_0}\omega\sqrt{4+\frac{2\alpha}{\omega^2 c_0}}$, then:

$$\left|1-\alpha m^2(z)\right| \geq \frac{|z|^4}{8(|z|+\eta)^4} \tag{11}$$

PROOF. The proof is in Appendix B. □

4.4. *Variance evaluations of resolvent based quantities.* In this section, we leverage the Poincaré-Nash inequality to evaluate the variance of quadratic forms and the normalized trace associated with the resolvent matrix.

LEMMA 5. *Let $z\in\mathbb{C}\backslash\mathbb{R}$. Let a and b be two deterministic vectors with unit norm in $\mathbb{R}^n$. Then, for any $\epsilon>0$:*

$$\mathbb{E}\left|a^TQb-\mathbb{E}a^TQb\right|^{2s} = O_z(p^{-s+\epsilon}) \quad \text{for } s\in\mathbb{N} \text{ and } s\geq 1 \tag{12}$$

$$\operatorname{var}\left(\frac{1}{p}\operatorname{tr} Q\right) = O_z(p^{-2+\epsilon}). \tag{13}$$

PROOF. 1. Proof of $\operatorname{var}(a^TQb)=O_z(p^{-1+\epsilon})$. We have:

$$\begin{aligned}
\operatorname{var}(a^TQb) &\leq \sum_{l=1}^n\sum_{k=1}^n \mathbb{E}\left|\frac{\partial a^TQb}{\partial Z_{lk}}\right|^2 \\
&\leq 8\sum_{l=1}^n\sum_{k=1}^n \mathbb{E}\left|\sum_{i=1}^n\sum_{j=1}^n a_ib_j\sum_{b\neq k}(x_b^Tx_k)\left[C_k^{\frac{1}{2}}x_b\right]_l Q_{ik}Q_{bj}\right|^2 \\
&+8\sum_{l=1}^n\sum_{k=1}^n \mathbb{E}\left|\sum_{i=1}^n\sum_{j=1}^n a_ib_j\sum_{b\neq k}(x_b^Tx_k)\left[C_k^{\frac{1}{2}}x_b\right]_l Q_{kj}Q_{bi}\right|^2 \\
&\triangleq 8\mathbb{E}X_1+8\mathbb{E}X_2
\end{aligned}$$

We need to control X_1 and X_2. We have:

$$
\begin{aligned}
X_1 &= \sum_{l=1}^{n}\sum_{k=1}^{n}\sum_{i_1=1}^{n}\sum_{j_1=1}^{n}\sum_{i_2=1}^{n}\sum_{j_2=1}^{n}\sum_{b_1\neq k}\sum_{b_2\neq k}\left[a_{i_1}b_{j_1}(x_{b_1}^T x_k)\left[C_k^{\frac{1}{2}}x_{b_1}\right]_l Q_{i_1k}Q_{b_1j_1}a_{i_2}b_{j_2}(x_{b_2}^T x_k)\left[C_k^{\frac{1}{2}}x_{b_2}\right]_l Q_{i_2k}^* Q_{b_2j_2}^*\right] \\
&= \sum_{k=1}^{n}\sum_{i_1=1}^{n}\sum_{j_1=}^{n}\sum_{i_2=1}^{n}\sum_{j_2=1}^{n}\sum_{b_1\neq k}\sum_{b_2\neq k}\left[a_{i_1}a_{i_2}b_{j_1}b_{j_2}x_{b_1}^T C_k x_{b_2}(x_{b_1}^T x_k)(x_{b_2}^T x_k)Q_{i_1k}Q_{i_2k}^* Q_{b_1j_1}Q_{b_2j_2}^*\right] \\
&= \sum_k \left[b^T\left(QS_kQ^*\right)b\left|[Qa]_k\right|^2\right] \\
&\leq \left[\max_k b^T\left(Q^T S_k Q^*\right)b\sum_k\left|[Qa]_k\right|^2\right]
\end{aligned}
$$

where $S_k = x_{b_1}^T C_k x_{b_2} x_{b_1}^T x_k x_{b_2}^T x_k \boldsymbol{\delta}_{b_1\neq k}\boldsymbol{\delta}_{b_2\neq k}$. Since $\|S_k\| = O(p^{-1})$, applying Lemma 1 and 2 we obtain $X_1 = O(p^{-1})|\Im z|^{-4}$. The treatment of X_2 follows the same arguments and is thus omitted.

2. Proof of $\mathbb{E}\left|a^TQb - a^TQb\right|^{2s} = O_z(p^{-s+\epsilon})$. The proof is performed by induction in s. Assume that the result holds true up to $s-1$. First note that:

$$
\mathbb{E}\left|a^TQb - a^TQb\right|^{2s} = \left(\mathbb{E}\left|a^TQb - \mathbb{E}a^TQb\right|^s\right)^2 + \operatorname{var}\left(a^TQb - \mathbb{E}a^TQb\right)^s
$$

Using the induction assumption, along with Cauchy-schwartz inequality, the first term in the above equation can be shown to be $O_z(p^{-s+\epsilon})$. As for the second term, we have:

$$
\operatorname{var}(a^TQb - \mathbb{E}a^TQb)^s \leq s^2\sum_{l=1}^{n}\sum_{k=1}^{n}\mathbb{E}\left|(a^TQb - \mathbb{E}a^TQb)^{s-1}\frac{\partial a^TQb}{\partial Z_{lk}}\right|^2
$$

Following the same calculations as before, we will ultimately obtain:

$$
\operatorname{var}(a^TQb - \mathbb{E}a^TQb)^s \leq 8s^2\mathbb{E}\left|(a^TQb - \mathbb{E}a^TQb)\right|^{2s-2}(X_1 + X_2)
$$

which leads to

$$
\operatorname{var}(a^TQb - \mathbb{E}a^TQb)^s = O_z(p^{-s+\epsilon}).
$$

3. Proof of $\operatorname{var}\left(\frac{1}{p}\operatorname{tr}Q\right) = O_z(p^{-2+\epsilon})$. We have:

$$
\begin{aligned}
\operatorname{var}\left(\frac{1}{p}\operatorname{tr}Q\right) &\leq \frac{1}{p^2}\sum_{l,k}\mathbb{E}\left|\sum_{i=1}^{n}\frac{\partial Q_{ii}}{\partial Z_{lk}}\right|^2 \\
&= \frac{4}{p^2}\sum_{l,k}\mathbb{E}\left|\sum_{i=1}^{n}\sum_{b\neq k}(x_b^Tx_k)\left[C_k^{\frac{1}{2}}x_b\right]_l(Q_{ik}Q_{bi} + Q_{ki}Q_{bi})\right|^2 \\
&\leq \frac{8}{p^2}\sum_{l,k}\mathbb{E}\left|\sum_{i=1}^{n}\sum_{b\neq k}(x_b^Tx_k)\left[C_k^{\frac{1}{2}}x_b\right]_l Q_{ik}Q_{bi}\right|^2 + \frac{8}{p^2}\sum_{l,k}\mathbb{E}\left|\sum_{i=1}^{n}\sum_{b\neq k}(x_b^Tx_k)\left[C_k^{\frac{1}{2}}x_b\right]_l Q_{ki}Q_{bi}\right|^2 \\
&= \frac{16}{p^2}\mathbb{E}\sum_{k=1}^{n}\left[Q^T\left(QS_kQ^H\right)Q^H\right]_{kk} = O_z(p^{-2+\epsilon})
\end{aligned}
$$

$\square$

LEMMA 6. *Let $z \in \mathbb{C}\backslash\mathbb{R}$. Let A_j be $n \times n$ matrix with bounded spectral norm. Then, for any $\epsilon > 0$:*

$$\max_{1\leq j\leq n} \mathbb{E}\left|\sum_{k\neq j} \frac{1}{\sqrt{p}}\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_k A_j\right) Q_{jk}\right|^2 = O_z(p^{-2+\epsilon})$$

PROOF. See Appendix C □

LEMMA 7. *Let $s \in \mathbb{N}^*$ and $z \in \mathbb{C}\backslash\mathbb{R}$. For $i \neq k$, and any $\epsilon > 0$, we have*

$$\mathbb{E}\left|e_i^T Q_k \xi_{(k,-k)}\right|^{2s} = O_z(p^{-s+\epsilon}).$$

PROOF. We will consider the case of $s = 1$. We can bound $\mathbb{E}\left|e_i^T Q_k \xi_{(k,-k)}\right|^2$ as:

$$\mathbb{E}\left|e_i^T Q_k \xi_{(k,-k)}\right|^2 = \mathbb{E}\left[\left|e_i^T Q_k \xi_{(k,-k)} - \mathbb{E}_{x_k} e_i^T Q_k \xi_{(k,-k)} + \mathbb{E}_{x_k} e_i^T Q_k \xi_{(k,-k)}\right|^2\right] \tag{14}$$

$$\leq 2\mathbb{E}\left[\operatorname{var}_{x_k}\left(e_i^T Q_k \xi_{(k,-k)}\right)\right] + 2\mathbb{E}\left|\mathbb{E}_{x_k} e_i^T Q_k \xi_{(k,-k)}\right|^2 \tag{15}$$

where $\operatorname{var}_{x_k}(X)$ denotes the quantity $\mathbb{E}_{x_k}\left|X - \mathbb{E}_{x_k} X\right|^2$. Note that:

$$\mathbb{E}_{x_k} e_i^T Q_k \xi_{(k,-k)} = \frac{1}{\sqrt{p}} \sum_{l\neq k} [Q_k]_{il}\left(x_l^T C_k x_l - \frac{1}{p}\operatorname{tr} C_k C_l\right)$$

Hence, from the previous lemma, we have:

$$\mathbb{E}\left|\mathbb{E}_{x_k} e_i^T Q_k \xi_{(k,-k)}\right|^2 = O_z(p^{-2+\epsilon})$$

It suffices thus to focus on $\mathbb{E}\left[\operatorname{var}_{x_k}\left(e_i^T Q_k \xi_{(k,-k)}\right)\right]$. Invoking the Poincaré-Nash inequality with respect to the elements of vector x_k, we can bound (15) as:

$$\begin{aligned}
\mathbb{E}\left|e_i^T Q_k \xi_{(k,-k)}\right|^2 &\leq p\sum_{l=1}^n \mathbb{E}\left[\left|\frac{\partial \sum_{j\neq k} [Q_k]_{ij} (x_k^T x_j)^2}{\partial Z_{lk}}\right|^2\right] \\
&= 4p\sum_{l=1}^n \mathbb{E}\left|\sum_{j\neq k} [Q_k]_{ij} (x_k^T x_j) \frac{1}{\sqrt{p}}\left[C_k^{\frac{1}{2}} x_j\right]_l\right|^2 \\
&= 4\sum_{l=1}^n \sum_{j\neq k}\sum_{r\neq k} \mathbb{E}\left[[Q_k]_{ij} [Q_k]_{ir}^* x_k^T x_j x_k^T x_r \left[C_k^{\frac{1}{2}} x_j\right]_l \left[C_k^{\frac{1}{2}} x_r\right]_l\right] \\
&= 4\sum_{j\neq k}\sum_{r\neq k} \mathbb{E}\left[x_r^T C_k x_j x_k^T x_j x_k^T x_r [Q_k]_{ij} [Q_k]_{ir}^*\right] \\
&= 4\mathbb{E}\left[Q_k S_k Q_k^H\right]_{ii} = O_z(p^{-1+\epsilon})
\end{aligned}$$

where $[S_k]_{jr} = x_k^T x_j x_r^T x_k x_r^T C_k x_j \boldsymbol{\delta}_{j\neq k} \boldsymbol{\delta}_{r\neq k}$. The inequality for $s \geq 2$ can be proved by induction in s. □

A direct corollary of Lemma 7 is the following result:

COROLLARY 2. *Let $i \neq k$ with $i,k \in \{1,\cdots,n\}$. Then, for any $\epsilon > 0$,*

$$\mathbb{E}\left|Q_{ik}\right|^{2s} = O_z(p^{-s+\epsilon})$$

for $s \in \mathbb{N}$ and $s \geq 1$.

An important outcome of Lemma 5 is that quadratic forms $a^T Q b$ and the normalized trace involving the resolvent matrix converge almost surely to their expectation. To find an asymptotic equivalent for these terms, it suffices thus to study their expectation in the asymptotic regime. As will be seen next, this can be achieved by using the Integration by Parts formula presented in section 4.1.

4.5. *Expression of matrix $\mathbb{E}Q_{jj}$ using the integration by parts formula.* The objective of this section is to develop the diagonal elements of the resolvent matrix using the integration by Parts formula. From the resolvent identity:

$$Q\Phi = I_n + zQ,$$

we have for $1 \leq j \leq n$,

$$\mathbb{E}Q_{jj} = -\frac{1}{z} + \frac{1}{z}\sum_{k\neq j}\mathbb{E}\left[Q_{ik}\Phi_{kj}\right]$$

Working on the rightmost term (with $k \neq j$) by expanding Φ_{kj} as a function of Z, we have:

$$\mathbb{E}\left[Q_{jk}\Phi_{kj}\right] = \frac{1}{p^{\frac{3}{2}}}\sum_{a,b=1}^{p}\sum_{l,l'=1}^{p}\sum_{m,m'=1}^{p}\left[C_k^{\frac{1}{2}}\right]_{al}\left[C_j^{\frac{1}{2}}\right]_{al'}\left[C_k^{\frac{1}{2}}\right]_{bm}\left[C_j^{\frac{1}{2}}\right]_{bm'}\mathbb{E}\left[Z_{lk}Z_{l'j}Z_{mk}Z_{m'j}Q_{jk}\right]$$

$$-\frac{1}{p^{\frac{3}{2}}}\operatorname{tr} C_k C_j \mathbb{E}\left[Q_{jk}\right] \tag{16}$$

Using the integration by parts formula in (6) along with the differention formula in (5), we obtain:

$$\begin{aligned}\mathbb{E}\left[Z_{lk}Z_{l'j}Z_{mk}Z_{m'j}Q_{jk}\right] &= \mathbb{E}\left[Z_{lk}\boldsymbol{\delta}_{l'm'}Z_{mk}Q_{jk}\right] + \mathbb{E}\left[Z_{lk}Z_{l'j}Z_{mk}\frac{\partial Q_{jk}}{\partial Z_{m'j}}\right] \\ &= \mathbb{E}\left[Z_{lk}\boldsymbol{\delta}_{l'm'}Z_{mk}Q_{jk}\right] - 2\mathbb{E}\left[Z_{lk}Z_{l'j}Z_{mk}\sum_{b\neq j}(x_b^T x_j)\left[C_j^{\frac{1}{2}}x_b\right]_{m'}(Q_{jj}Q_{bk} + Q_{jk}Q_{bj})\right]\end{aligned}$$

Plugging the above equation into (16), we ultimately get:

$$\begin{aligned}\sum_{k\neq j}\mathbb{E}\left[Q_{jk}\Phi_{kj}\right] &= \mathbb{E}\left[\sum_{k\neq j}\frac{1}{\sqrt{p}}\left(x_k^T C_j x_k - \frac{1}{p}\operatorname{tr} C_k C_j\right)Q_{jk}\right] \\ &\quad - 2\sum_{k\neq j}\sum_{r\neq j}\mathbb{E}\left[x_j^T x_k x_r^T x_j x_k^T C_j x_r (Q_{jj}Q_{rk} + Q_{jk}Q_{rj})\right]\end{aligned}$$

Hence,

$$z\mathbb{E}Q_{jj} = -1 + X_{1,j}(z) + X_{2,j}(z) + X_{3,j}(z) + X_{4,j}(z) \tag{17}$$

where $\{X_{l,j}(z)\}_{l=1}^4$ write as:

$$X_{1,j}(z) = \mathbb{E}\left[\sum_{k\neq j}\frac{1}{\sqrt{p}}\left(x_k^T C_j x_k - \frac{1}{p}\operatorname{tr} C_k C_j\right) Q_{jk}\right] \tag{18}$$

$$X_{2,j}(z) = -2\sum_{k\neq j}\mathbb{E}\left[\left(x_j^T x_k\right)^2 x_k^T C_j x_k Q_{jj} Q_{kk}\right] \tag{19}$$

$$X_{3,j}(z) = -2\sum_{k\neq j}\sum_{r\notin\{j,k\}}\mathbb{E}\left[x_j^T x_k x_r^T x_j x_k^T C_j x_r Q_{jj} Q_{rk}\right] \tag{20}$$

$$X_{4,j}(z) = -2\sum_{k\neq j}\sum_{r\neq j}\mathbb{E}\left[x_j^T x_k x_r^T x_j x_k^T C_j x_r Q_{jk} Q_{rj}\right] \tag{21}$$

With the decomposition in (17), we are now in position to prove Theorem 1 by studying sequentially the terms $\{X_{l,j}(z)\}_{l=1}^4$. This will be carried out in the next section.

5. Proof of Theorem 1. Quantities $X_{1,j}(z), X_{3,j}(z)$ and $X_{4,j}(z)$ constitute error terms that converge to zero in the asymptotic regime.

A direct application of Lemma 6 allows to show that:

$$\max_{1\leq j\leq n}|X_{1,j}(z)| = O_z(p^{-1+\epsilon}).$$

To control $X_{3,j}(z)$ and $X_{4,j}(z)$, the following lemma is required:

LEMMA 8. *Let A_1 and A_2 be $p\times p$ matrices independent of x_j with spectral norms of order $O(1)$. Then, for any small $\epsilon > 0$,*

$$\max_{1\leq j\leq n}\mathbb{E}\left|\sum_{s\notin\{k,b,j\}} x_s^T A_1 x_k x_s^T A_2 x_b Q_{sj}\right|^2 = O_z(p^{-2+\epsilon})$$

PROOF. See Appendix D □

A direct application of Lemma 8 implies that:

$$\max_{1\leq j\leq n}|X_{3,j}(z)| = O_z(p^{-\frac{1}{2}+\epsilon})$$

To handle $X_{4,j}(z)$, we start by decomposing it as:

$$X_{4,j}(z) = -2\sum_{k\neq j}\sum_{r\notin\{j,k\}}\mathbb{E}\left[x_j^T x_k x_r^T x_j x_k^T C_j x_r Q_{jk} Q_{rj}\right] - 2\sum_{k\neq j}\mathbb{E}\left[x_j^T x_k x_k^T x_j x_k^T C_j x_k Q_{jk} Q_{kj}\right]$$

Using Lemma 8 and the approximations in Corollary 2, it unfolds that:

$$\max_{1\leq j\leq n}|X_{4,j}(z)| = O_z(p^{-1+\epsilon}).$$

As for $X_{2,j}(z)$, we can easily see that:

$$X_{2,j}(z) = -2\sum_{k\neq j}\frac{1}{p}\left(\frac{1}{p}\operatorname{tr} C_kC_j\right)^2 \mathbb{E}Q_{kk}\mathbb{E}Q_{jj} + O_z(p^{-\frac{1}{2}+\epsilon})$$
$$= -\frac{\omega^2 c_0}{n}\sum_{k=1}^{n}\mathbb{E}Q_{kk}\mathbb{E}Q_{jj} + O_z(p^{-\frac{1}{4}})$$

which follows from the fact that $\frac{2}{p}\operatorname{tr} C_kC_j = \omega^2 + O(p^{-\frac{1}{4}})$. Hence,

$$z\mathbb{E}Q_{jj} = -1 - \omega^2 c_0 g_n(z)\mathbb{E}Q_{jj} + O_z(p^{-\frac{1}{4}}) \tag{22}$$

Summing over index j, we thus obtain:

$$\omega^2 c_0 g_n^2(z) + z g_n(z) + 1 = O_z(p^{-\frac{1}{4}}). \tag{23}$$

Reaching this equation, termed as "Master equation" in [Capitaine et al., 2009], it can be proven by following the same steps in [Capitaine et al., 2009] that:

$$|g_n(z) - m(z)| = O_z(p^{-\frac{1}{4}}). \tag{24}$$

The weak convergence of the spectral measure of Φ to the semi-circle law follows from using the fact that $\frac{1}{n}\operatorname{tr} Q - g_n(z)$ converge almost surely to zero. This ends up the proof of Theorem 1.

6. Proof of Theorem 2. Theorem 2 focuses on the study of a deterministic equivalent of bilinear forms with kernel the resolvent matrix $Q(z)$. Due to the almost sure convergence of $a^TQ(z)b - a^T\mathbb{E}Q(z)b$ guaranteed by Lemma 5, the problem amounts to finding an asymptotic approximation of $a^T\mathbb{E}Q(z)b$. It can be easily seen by injecting the approximation in (24) into (22) that the contribution of the diagonal elements $\sum_{k=1}^n a_kb_k\mathbb{E}Q(z)$ can be well-approximated by $m(z)a^Tb$. It remains thus to study the contribution of the off-diagonal elements. This is the purpose of the next proposition.

Proposition 1. *Let $\{a_k\}$ and $\{b_k\}$ be real sequences such that $\sum_{k=1}^n |a_k|^2$ and $\sum_{k=1}^n |b_k|^2$ are uniformly bounded. For $z\in\mathbb{C}\backslash\mathbb{R}$ and any small ϵ:*

$$\sum_{k=1}^{n}\sum_{r=1,r\neq k}^{n} a_kb_r\mathbb{E}Q_{kr}$$
$$= -\sum_{k=1}^{n}\sum_{r\neq k} a_kb_r\mathbb{E}\left[Q_{kk}\right]\mathbb{E}\left[\frac{1}{\sqrt{p}}\sum_{l\neq k}[Q]_{rl}\left(x_l^TC_kx_l - \frac{1}{p}\operatorname{tr} C_kC_l\right)\right] + O_z\left(p^{-1/2+\epsilon}\right). \tag{25}$$

Moreover, for $k\neq j$, $\mathbb{E}Q_{kj} = O_z(p^{-1+\epsilon})$.

Proof. See Appendix E ∎

Lemma 6 shows that $\mathbb{E}\left[\frac{1}{\sqrt{p}}\sum_{l\neq k}Q_{rl}\left(x_l^TC_kx_l - \frac{1}{p}\operatorname{tr} C_kC_l\right)\right]$ is $O_z(p^{-1+\epsilon})$, implying that the first term on the right-hand side of (25) is $O_z(1)$. To find an asymptotic approximation of this term, we need thus to refine the approximation of Lemma 6 by isolating the terms of order $O_z(p^{-1+\epsilon})$ from the other vanishing terms converging faster than $O_z(p^{-1+\epsilon})$. As shown in the proof of the following lemma, the key ingredient allowing us to carry out these derivations is the Integration by Parts formula.

PROPOSITION 2. *Let $\epsilon > 0$ be a small real. Then, the following approximation holds true:*

$$\sum_{k\neq j}\frac{1}{\sqrt{p}}\mathbb{E}\left[\left(x_k^TC_jx_k-\frac{1}{p}\operatorname{tr}C_kC_j\right)Q_{kj}\right]=-\frac{2}{p^2}\sum_{k=1}^n\mathbb{E}Q_{kk}\sum_{b=1}^n\mathbb{E}Q_{bj}\frac{1}{p}\operatorname{tr}(C^\circ)^4+O_z(p^{-\frac{5}{4}}) \tag{26}$$

PROOF. Using the Integration by Part formula, we obtain

$$\begin{aligned}
&\sum_{k\neq j}\frac{1}{\sqrt{p}}\mathbb{E}\left[\left(x_k^TC_jx_k-\frac{1}{p}\operatorname{tr}C_kC_j\right)Q_{kj}\right]=-\frac{2}{p}\sum_{k\neq j}\sum_{b\neq k}\mathbb{E}\left[x_b^TC_kC_jx_kx_b^Tx_kQ_{kk}Q_{bj}\right]\\
&-\frac{2}{p}\sum_{k\neq j}\sum_{b\neq k}\mathbb{E}\left[x_b^TC_kC_jx_kx_b^Tx_kQ_{kj}Q_{bk}\right]\\
&=\chi_1+\chi_2
\end{aligned}$$

We begin with handling χ_2. We have:

$$\begin{aligned}
\chi_2&=-\frac{2}{p^2}\sum_{k\neq j}\sum_{b\neq k}\mathbb{E}\left[x_b^TC_kC_jC_kx_bQ_{kj}Q_{bk}\right]\\
&+\frac{4}{p\sqrt{p}}\sum_{k\neq j}\sum_{b\neq k}\sum_{s\neq k}\mathbb{E}\left[x_s^Tx_kx_s^TC_kx_bx_b^TC_kC_jx_kQ_{kk}Q_{sj}Q_{bk}\right]\\
&+\frac{4}{p\sqrt{p}}\sum_{k\neq j}\sum_{b\neq k}\sum_{s\neq k}\mathbb{E}\left[x_s^Tx_kx_s^TC_kx_bx_b^TC_kC_jx_kQ_{kj}Q_{sk}Q_{bk}\right]\\
&+\frac{4}{p\sqrt{p}}\sum_{k\neq j}\sum_{b\neq k}\sum_{s\neq k}\mathbb{E}\left[x_s^Tx_kx_s^TC_kx_bx_b^TC_kC_jx_kQ_{bk}Q_{sk}Q_{kj}\right]\\
&+\frac{4}{p\sqrt{p}}\sum_{k\neq j}\sum_{b\neq k}\sum_{s\neq k}\mathbb{E}\left[x_s^Tx_kx_s^TC_kx_bx_b^TC_kC_jx_kQ_{kk}Q_{sb}Q_{kj}\right]\\
&=\chi_{21}+\chi_{22}+\chi_{23}+\chi_{24}+\chi_{25}
\end{aligned}$$

Let us treat χ_{21}. We have:

$$\begin{aligned}
\chi_{21}&=-\frac{2}{p^{\frac{3}{2}}}\mathbb{E}\left[\left[\frac{1^T}{\sqrt{p}}\left\{Q_{bk}x_b^TC_kC_jC_kx_b\boldsymbol{\delta}_{b\neq k}\boldsymbol{\delta}_{k\neq j}\right\}_{b,k=1}^nQ\right]_j\right]\\
&=O_z(p^{-\frac{3}{2}})
\end{aligned}$$

To handle χ_{22}, we start by decomposing it as:

$$\begin{aligned}
\chi_{22}&=\frac{4}{p\sqrt{p}}\sum_{k\neq j}\sum_{s\neq k}\sum_{b\notin\{k,s\}}\mathbb{E}\left[x_s^Tx_kx_s^TC_kx_bx_b^TC_kC_jx_kQ_{kk}Q_{sj}Q_{bk}\right]\\
&+\frac{4}{p\sqrt{p}}\sum_{k\neq j}\sum_{s\neq k}\mathbb{E}\left[x_s^Tx_kx_s^TC_kx_sx_s^TC_kC_jx_kQ_{kk}Q_{sj}Q_{sk}\right]
\end{aligned}$$

The first term in χ_{22} can be treated as follows:

$$\begin{aligned}
&\left|\frac{4}{p\sqrt{p}}\sum_{k\neq j}\sum_{s\neq k}\sum_{b\notin\{k,s\}}\mathbb{E}\left[x_s^Tx_kx_s^TC_kx_bx_b^TC_kC_jx_kQ_{kk}Q_{sj}Q_{bk}\right]\right| \\
&\leq \frac{4}{p\sqrt{p}}\sum_{k\neq j}\sum_{s\neq k}\sqrt{\mathbb{E}\left|x_s^Tx_kQ_{sj}\right|^2}\sqrt{\mathbb{E}\left|\sum_{b\notin\{k,s\}}x_s^TC_kx_bx_b^TC_jC_jx_kQ_{sj}\right|^2}
\end{aligned}$$

Using Lemma 8, we obtain:

$$\left|\frac{4}{p\sqrt{p}}\sum_{k\neq j}\sum_{s\neq k}\sum_{b\notin\{k,s\}}\mathbb{E}\left[x_s^Tx_kx_s^TC_kx_bx_b^TC_kC_jx_kQ_{kk}Q_{sj}Q_{bk}\right]\right| = O_z(p^{-\frac{3}{2}+\epsilon})$$

The second term in χ_{22} is obviously $O_z(p^{-\frac{3}{2}+\epsilon})$ which implies that:

$$\chi_{22} = O_z(p^{-\frac{3}{2}+\epsilon})$$

Using a similar decomposition to that used in χ_{22}, we can also prove that $\chi_{23} = O_z(p^{-2+\epsilon})$ and $\chi_{24} = O_z(p^{-2+\epsilon})$. As for χ_{25}, we can easily see that the term obtained by singling out index $b\neq s$ is $O_z(p^{-\frac{3}{2}+\epsilon})$, leading to:

$$\begin{aligned}
\chi_{25} &= \frac{4}{p\sqrt{p}}\sum_{k\neq j}\sum_{b\neq k}\mathbb{E}\left[x_b^Tx_kx_b^TC_kx_bx_b^TC_kC_jx_kQ_{kj}Q_{bb}Q_{kk}\right] + O_z(p^{-\frac{3}{2}+\epsilon}) \\
&= \frac{4}{p\sqrt{p}}\sum_{k\neq j}\sum_{b\neq k}\mathbb{E}\left[x_b^Tx_kx_b^TC_kC_jx_kQ_{kj}\right]\frac{1}{p}\operatorname{tr}\left(C^\circ\right)^2\mathbb{E}Q_{bb}\mathbb{E}Q_{kk} + O_z(p^{-\frac{5}{4}})
\end{aligned}$$

It follows from using Lemma 8 that:

$$\chi_{25} = O_z(p^{-\frac{3}{2}+\epsilon}).$$

We will now handle χ_1. Again using Lemma 8, we prove that:

$$\chi_1 = -\frac{2}{p}\sum_{k\neq j}\sum_{b\neq k}\mathbb{E}\left[x_b^TC_kC_jx_kx_b^Tx_kQ_{bj}\right]\mathbb{E}Q_{kk} + O_z(p^{-\frac{3}{2}+\epsilon})$$

Using the Integration by Parts formula, we have:

$$\begin{aligned}
\chi_1 = &-\sum_{k\neq j}\sum_{b\neq k}\frac{2}{p^2}\mathbb{E}\left[x_k^TC_bC_kC_jx_kQ_{bj}\right]\mathbb{E}Q_{kk} + \frac{4}{p^{\frac{3}{2}}}\sum_{k\neq j}\mathbb{E}Q_{kk}\sum_{b\neq k}\sum_{s\neq k}\mathbb{E}\left[x_s^Tx_bx_s^TC_bC_kC_jx_kQ_{bb}x_k^Tx_bQ_{sj}\right] \\
&+\frac{4}{p^{\frac{3}{2}}}\sum_{k\neq j}\mathbb{E}Q_{kk}\sum_{b\neq k}\sum_{s\neq k}\mathbb{E}\left[x_s^Tx_bx_s^TC_bC_kC_jx_kx_k^Tx_bQ_{bj}Q_{sb}\right] + O_z(p^{-\frac{3}{2}+\epsilon})
\end{aligned} \tag{27}$$

The first term in the right-hand side of (27) can be worked out as:

$$
\begin{aligned}
&-\sum_{k\neq j}\sum_{b\neq k}\frac{2}{p^2}\mathbb{E}\left[x_k^T C_b C_k C_j x_k Q_{bj}\right]\mathbb{E}Q_{kk}\\
&=-\frac{2}{p^2}\sum_{b=1}^{n}\mathbb{E}\left[\left(\sum_{k=1}^{n}x_k^T C_b C_k C_j x_k-\frac{1}{p}\operatorname{tr}C_k C_b C_k C_j\right)Q_{bj}\right]\mathbb{E}Q_{kk}-\frac{2}{p^2}\sum_{b=1}^{n}\sum_{k=1}^{n}\frac{1}{p}\operatorname{tr}C_k C_b C_k C_j\mathbb{E}Q_{bj}\mathbb{E}Q_{kk}\\
&+O_z(p^{-\frac{3}{2}+\epsilon})\\
&\overset{(a)}{=}-\frac{2}{p^2}\sum_{b=1}^{n}\sum_{k=1}^{n}\frac{1}{p}\operatorname{tr}C_k C_b C_k C_j\mathbb{E}Q_{bj}\mathbb{E}Q_{kk}+O_z(p^{-\frac{3}{2}+\epsilon})\\
&\overset{(b)}{=}-\frac{2}{p}1^T\left\{\mathbb{E}Q_{kk}\left(\frac{1}{p^2}\operatorname{tr}C_k C_b C_k C_j-\frac{1}{p^2}\operatorname{tr}(C^\circ)^4\right)\mathbb{E}Q_{bj}\right\}_{b,k=1}^{n}1-\frac{2}{p^2}\sum_{b=1}^{n}\sum_{k=1}^{n}\mathbb{E}Q_{kk}\mathbb{E}Q_{bj}\frac{1}{p}\operatorname{tr}(C^\circ)^4\\
&+O_z(p^{-\frac{3}{2}+\epsilon})
\end{aligned}
$$

where (a) follows by using the fact that $\left(\sum_{k=1}^{n}x_k^T C_b C_k C_j x_k-\frac{1}{p}\operatorname{tr}C_k C_b C_k C_j\right)$ is $O(1)$ and (b) from the fact that

$$
\left\|\left\{\mathbb{E}Q_{kk}\left(\frac{1}{p^2}\operatorname{tr}C_k C_b C_k C_j-\frac{1}{p^2}\operatorname{tr}(C^\circ)^4\right)\mathbb{E}Q_{bj}\right\}_{b,k=1}^{n}\right\|
$$

is $O_z(p^{-\frac{5}{4}})$. □

Plugging the approximation in (26) into (25), we obtain:

$$
\begin{aligned}
\sum_{k=1}^{n}\sum_{r\neq k}a_k b_r\mathbb{E}Q_{kr}&=-\sum_{k=1}^{n}\sum_{r\neq k}a_k b_r\mathbb{E}Q_{kk}\mathbb{E}\frac{1}{\sqrt{p}}\mathbb{E}\left[Q_{rr}\left(x_r^T C_k x_r-\frac{1}{p}\operatorname{tr}C_k C_r\right)\right]\\
&+\frac{2}{p^2}\sum_{k=1}^{n}\sum_{r\neq k}a_k b_r\mathbb{E}Q_{kk}\sum_{l=1}^{n}\mathbb{E}Q_{ll}\sum_{s=1}^{n}\mathbb{E}Q_{sr}\frac{1}{p}\operatorname{tr}(C^\circ)^4+O_z(p^{-\frac{1}{4}})\\
&=\frac{2}{p^2}\sum_{k=1}^{n}\sum_{r\neq k}a_k b_r\mathbb{E}Q_{kk}\sum_{l=1}^{n}\mathbb{E}Q_{ll}\sum_{s=1}^{n}\mathbb{E}Q_{sr}\frac{1}{p}\operatorname{tr}(C^\circ)^4+O_z(p^{-\frac{1}{4}}) \qquad (28)
\end{aligned}
$$

Taking $a_k=\frac{1}{\sqrt{p}}$, we thus have from the approximation in (24):

$$
\frac{1}{\sqrt{p}}\sum_{k=1}^{n}\sum_{r\neq k}^{n}\mathbb{E}Q_{kr}b_r=\frac{2n^2}{p^{\frac{5}{2}}}m^2(z)\sum_{r=1}^{n}\sum_{s\neq r}\mathbb{E}Q_{sr}b_r\frac{1}{p}\operatorname{tr}(C^\circ)^4+\frac{2n^2}{p^{\frac{5}{2}}}m^2(z)\sum_{r=1}^{n}\mathbb{E}Q_{rr}b_r\frac{1}{p}\operatorname{tr}(C^\circ)^4+O_z(p^{-\frac{1}{4}})
$$

From Lemma 4, we have $\frac{1}{1-\frac{2n^2}{p^2}m^2(z)\frac{1}{p}\operatorname{tr}(C^\circ)^4}=O_z(1)$. Also, it is easy to see from (22) and (23) that for all $j=1,\cdots,n$,

$$
\mathbb{E}Q_{jj}-m(z)=O_z(p^{-\frac{1}{4}})
$$

which implies that:

$$
\frac{1}{\sqrt{p}}\sum_{k=1}^{n}\sum_{r\neq k}^{n}\mathbb{E}Q_{kr}b_r=\frac{\frac{2n^2}{p^2}\frac{1}{p}\operatorname{tr}(C^\circ)^4 m^3(z)\frac{1^T}{\sqrt{p}}b}{1-\frac{2n^2}{p^2}m^2(z)\frac{1}{p}\operatorname{tr}(C^\circ)^4}+O_z(p^{-\frac{1}{4}}). \qquad (29)
$$

Plugging (29) into (28) holding for general $\{a_k\}$ and $\{b_k\}$, we finally obtain:

$$\sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}Q_{kr} = \frac{\frac{2n}{p}m^3(z)\frac{1}{p}\operatorname{tr}(C^\circ)^4\, a^T\frac{11^T}{p}b}{1-\frac{2n^2}{p^2}m^2(z)\frac{1}{p}\operatorname{tr}(C^\circ)^4} + O_z(p^{-\frac{1}{4}})$$

We end up the proof of Theorem 2 by noticing that $\sum_{k=1}^{n} a_k b_k \mathbb{E}Q_{kk} = a^T b m(z) + O_z(p^{-\frac{1}{4}})$.

7. Proof of Theorem 4. The goal of Theorem 4 is to determine the limiting support of the empirical eigenvalue measure of Φ. Referring to the work [Haagerup and Thorbjørnsen, 2005], the almost sure location of the eigenvalues of Φ around the support of the semi-circle law would hold if the difference $g_n(z) - m(z)$ converged at a rate fafter than $O_z(p^{-1-\epsilon})$. As this condition is not satisfied in our case, the difference $g_n(z) - m(z)$ being of order $O_z(p^{-\frac{1}{4}})$ as shown previously (cf.(24)), it might happen that matrix Φ presents some isolated eigenvalues. Following the ideas of [Capitaine et al., 2009] and [Loubaton, 2016], these eigenvalues can be characterized by analyzing each of the vanishing terms at rate slower than $1/p$.

Particularly, we will prove that:

$$g_n(z) - m(z) = \frac{1}{p^{\frac{1}{4}}}\tilde{f}(z) + \frac{1}{\sqrt{p}}\tilde{h}(z) + \frac{1}{p^{\frac{3}{4}}}\tilde{k}(z) + \frac{1}{p}\tilde{e}(z) + O_z(p^{-\frac{5}{4}}) \tag{30}$$

where $\tilde{f}(z)$, $\tilde{h}(z)$, $\tilde{k}(z)$ and $\tilde{e}(z)$ are Stieltjes transforms of some distributions. As will be shown next, while the support of the distributions associated with $\tilde{f}(z)$, $\tilde{h}(z)$, $\tilde{k}(z)$ is included in $\mathcal{S}$, that of $\tilde{e}(z)$ might present singularities under some conditions to be characterized. This fact will be seen to result in Φ possessing eigenvalues outside the bulk of the semi-circle law.

Proving (30) constitutes the heart of the matter and the hardest step in the proof of Theorem 4. In fact, it would not be sufficient to control the rate at which converge the slowly vanishing terms. One needs to carefully go in depth in each of these terms and determine for each an equivalent depending on the Stieltjes transform of the semi-circle law $m(z)$ with an error converging faster than $O_z(p^{-1-\epsilon})$ for some $\epsilon > 0$. In order to pave the way towards this, we will need to derive equivalents of many random quantities that will often arise in our derivations. This is performed in the following two lemmas.

LEMMA 9. *Let k, b be two integers in $\{1, \cdots, n\}$ such that $b \neq k$. Let A_1 and A_2 be $n \times n$ deterministic matrices with bounded spectral norms. Let $z \in \mathbb{C}\backslash\mathbb{R}$. Then,*

$$\mathbb{E}\left[x_b^T A_1 x_k x_b^T A_2 x_k Q_{bk}\right] = -2p^{-\frac{3}{2}} m^2(z)\frac{1}{p}\operatorname{tr}(C^\circ)^2 A_2 \frac{1}{p}\operatorname{tr}(C^\circ)^2 A_1 + O_z(p^{-\frac{7}{4}})$$

PROOF. Using the relation for $b \neq k$, $Q_{bk} = -e_b^T Q_k \xi_{(k,-k)} Q_{kk}$, we have:

$$
\begin{aligned}
&\mathbb{E}\left[x_b^T A_1 x_k x_b^T A_2 x_k Q_{bk}\right] = -\mathbb{E}\left[x_b^T A_1 x_k x_b^T A_2 x_k e_b^T Q_k \xi_{(k,-k)} Q_{kk}\right] \\
&= -\left[x_b^T A_1 x_k x_b^T A_2 x_k e_b^T Q_k \xi_{(k,-k)}\right] \mathbb{E} Q_{kk} + O_z(p^{-2+\epsilon}) \\
&= -\mathbb{E} Q_{kk} \mathbb{E}\left[x_b^T A_1 x_k x_b^T A_2 x_k \sum_{l \neq k} \sqrt{p}\, [Q_k]_{bl} \left((x_k^T x_l)^2 - \frac{1}{p} x_l^T C_k x_l\right)\right] \\
&- \mathbb{E} Q_{kk} \mathbb{E}\left[\frac{1}{p} x_b^T A_1 C_k A_2 x_b \sum_{l \neq k} p^{-\frac{1}{2}} \left(x_l^T C_k x_l - \frac{1}{p} \operatorname{tr} C_k C_l\right) [Q_k]_{bl}\right] \\
&\overset{(a)}{=} -\mathbb{E} Q_{kk} \frac{2}{p^{\frac{3}{2}}} \sum_{l \neq k} \mathbb{E}\left[[Q_k]_{bl}\, x_b^T A_2 C_k x_l x_l^T C_k A_1 x_b\right] + O_z(p^{-2+\epsilon}) \\
&= -\frac{2}{p^{\frac{3}{2}}} m^2(z) \frac{1}{p} \operatorname{tr} A_2 \left(C^\circ\right)^2 \frac{1}{p} \operatorname{tr} \left(C^\circ\right)^2 A_1 + O_z(p^{-\frac{7}{4}})
\end{aligned}
$$

where the replacement of $[Q_k]_{bl}$ by Q_{bl} can be established thanks to Lemma 14 proven in Appendix E. $\square$

LEMMA 10. *Let $j, k \in \{1, \cdots, n\}$ such that $j \neq k$. Let A_1, A_2 and A_3 be $n \times n$ deterministic matrices with bounded spectral norms. Then,*

$$
\sum_{b \notin \{j,k\}} \mathbb{E}\left[x_k^T A_1 x_j x_b^T A_2 x_j x_b^T A_3 x_k Q_{bj}\right] = -2np^{-\frac{5}{2}} m^2(z) \frac{1}{p} \operatorname{tr} \left(C^\circ\right)^2 A_2 \frac{1}{p} \operatorname{tr} C^\circ A_1 \left(C^\circ\right)^2 A_3 + O_z(p^{-\frac{7}{4}})
$$

PROOF. Again using the relation $Q_{bj} = -e_b^T Q_j \xi_{(j,-j)} Q_{jj}$ for $b \neq j$, we have:

$$\sum_{b \notin \{j,k\}} \mathbb{E}\left[x_k^T A_1 x_j x_b^T A_2 x_j x_b^T A_3 x_k Q_{bj}\right] = -\sum_{b \notin \{j,k\}} \mathbb{E}\left[x_k^T A_1 x_j x_b^T A_2 x_j x_b^T A_3 x_k e_b^T Q_j \xi_{(j,-j)} Q_{jj}\right]$$

$$= -\sum_{b \notin \{j,k\}} \mathbb{E}\left[x_k^T A_1 x_j x_b^T A_2 x_j x_b^T A_3 x_k e_b^T Q_j \xi_{(j,-j)}\right] \mathbb{E}Q_{jj} + O_z(p^{-2+\epsilon})$$

$$= -\sum_{b \notin \{j,k\}} \mathbb{E}\left[x_k^T A_1 x_j x_b^T A_2 x_j x_b^T A_3 x_k \sum_{l \neq j} [Q_j]_{bl} \sqrt{p}\left((x_j^T x_l)^2 - \frac{1}{p} x_l^T C_j x_l\right)\right] \mathbb{E}Q_{jj}$$

$$- \sum_{b \notin \{j,k\}} \mathbb{E}\left[x_k^T A_1 x_j x_b^T A_2 x_j x_b^T A_3 x_k \sum_{l \neq j} [Q_j]_{bl} p^{-\frac{1}{2}}\left(x_l^T C_j x_l - \frac{1}{p} \operatorname{tr} C_j C_l\right)\right] \mathbb{E}Q_{jj} + O_z(p^{-2+\epsilon})$$

$$= -\sum_{b \notin \{j,k\}} \sum_{l \neq j} 2p^{-\frac{3}{2}} \mathbb{E}\left[[Q_j]_{bl}\, x_b^T A_2 C_j x_l x_l^T C_j A_1 x_k x_b^T A_3 x_k\right] \mathbb{E}Q_{jj}$$

$$- \sum_{b \notin \{j,k\}} \mathbb{E}\left[\frac{1}{p} x_b^T A_2 C_j A_1 x_k x_b^T A_3 x_k p^{-\frac{1}{2}} \sum_{l \neq j} [Q_j]_{bl}\left(x_l^T C_j x_l - \frac{1}{p} \operatorname{tr} C_j C_l\right)\right] \mathbb{E}Q_{jj} + O_z(p^{-2+\epsilon})$$

$$= -\sum_{b \notin \{j,k\}} 2p^{-\frac{3}{2}} \mathbb{E}\left[[Q_j]_{bb}\, x_b^T A_3 x_k x_b^T A_2 C_j x_b x_b^T C_j A_1 x_k\right] \mathbb{E}Q_{jj} + O_z(p^{-2+\epsilon})$$

$$= -\sum_{b=1}^{n} 2p^{-\frac{5}{2}} \mathbb{E}\,[Q_j]_{bb} \frac{1}{p} \operatorname{tr} (C^\circ)^2 A_2 \frac{1}{p} \operatorname{tr} C^\circ A_1 (C^\circ)^2 A_3 \mathbb{E}Q_{jj} + O_z(p^{-2+\epsilon})$$

$$= -2np^{-\frac{5}{2}} m^2(z) \frac{1}{p} \operatorname{tr} (C^\circ)^2 A_2 \frac{1}{p} \operatorname{tr} C^\circ A_1 (C^\circ)^2 A_3 + O_z(p^{-\frac{7}{4}})$$

□

With these Lemmas at hand, we are now in position to prove the estimation in (30). To this end, recall the relation involving the diagonal elements of Q:

$$z\mathbb{E}Q_{jj} = -1 + X_{1,j}(z) + X_{2,j}(z) + X_{3,j}(z) + X_{4,j}(z)$$

where $\{X_{l,j}\}_{l=1}^4$ are given by (18)-(21). From the discussion in the beginning of this section, we need to find an equivalent for each of the $\{X_{l,j}\}_{l=1}^4$ with an error converging strictly faster than p^{-1}. These equivalents should be expressed in terms of the stieltjes transform of the semi-circle law. We start by handling $X_{1,j}$. Recall that:

$$X_{1,j} = \frac{1}{\sqrt{p}} \sum_{k \neq j} \mathbb{E}\left[(x_k^T C_j x_k - \frac{1}{p} \operatorname{tr} C_k C_j) Q_{kj}\right]$$

From Proposition 2 and Theorem 2, it unfolds that:

$$\frac{1}{\sqrt{p}} \sum_{k \neq j} \mathbb{E}\left[(x_k^T C_j x_k - \frac{1}{p} \operatorname{tr} C_k C_j) Q_{kj}\right] = -\frac{2n}{p^2} m(z) \frac{1}{p} \operatorname{tr} (C^\circ)^4 \sum_{b=1}^{n} \mathbb{E}Q_{bj} + O_z(p^{-\frac{5}{4}})$$

$$= -\frac{2np^{-2} m^2(z) \frac{1}{p} \operatorname{tr} (C^\circ)^4}{1 - \frac{2}{p} \operatorname{tr} (C^\circ)^4 \frac{n^2}{p^2} m^2(z)} + O_z(p^{-\frac{5}{4}})$$

We now deal with $X_{2,j}$. Using the Integration by Parts formula, we decompose $X_{2,j}$ as:

$$\begin{aligned}
X_{2,j} &= -\frac{2}{p}\sum_{k\neq j}\mathbb{E}\left[(x_k^T C_j x_k)^2 Q_{kk}Q_{jj}\right] \\
&+ \frac{8}{\sqrt{p}}\sum_{k\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^T C_j x_k x_k^T x_j x_k^T C_j x_b x_b^T x_j Q_{jj} Q_{bj} Q_{kk}\right] \\
&+ \frac{8}{\sqrt{p}}\sum_{k\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^T C_j x_k x_k^T x_j x_k^T C_j x_b x_b^T x_j Q_{kj} Q_{bk} Q_{jj}\right] \\
&= X_{21,j} + X_{22,j} + X_{23,j}
\end{aligned}$$

We further decompose $X_{22,j}$ as:

$$\begin{aligned}
X_{22,j} &= \frac{8}{\sqrt{p}}\sum_{k\neq j}\sum_{b\notin\{k,j\}}\mathbb{E}\left[x_k^T C_j x_k x_k^T x_j x_k^T C_j x_b x_b^T x_j Q_{bj} Q_{jj} Q_{kk}\right] \\
&+ \frac{8}{\sqrt{p}}\sum_{k\neq j}\mathbb{E}\left[\left(x_k^T C_j x_k\right)^2\left(x_k^T x_j\right)^2 Q_{kj} Q_{jj} Q_{kk}\right]
\end{aligned}$$

Using Lemma 8 along with Lemma 9 and Lemma 10, we have:

$$\begin{aligned}
X_{22,j} &= \frac{8}{\sqrt{p}}\sum_{k\neq j}\frac{1}{p}\operatorname{tr}(C^\circ)^2\,\mathbb{E}Q_{kk}\mathbb{E}Q_{jj}\sum_{b\notin\{k,j\}}\mathbb{E}\left[x_k^T x_j x_k^T C_j x_b x_b^T x_j Q_{bj}\right] \\
&+ \frac{8}{\sqrt{p}}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\mathbb{E}Q_{jj}\sum_{k\neq j}\mathbb{E}Q_{kk}\mathbb{E}\left[(x_k^T x_j)^2 Q_{kj}\right] + O_z(p^{-\frac{5}{4}}) \\
&= -\frac{16}{p^3}\sum_{k=1}^{n}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\frac{1}{p}\operatorname{tr}(C^\circ)^4\,\mathbb{E}Q_{kk}\left(\mathbb{E}Q_{jj}\right)^2\sum_{b=1}^{n}\mathbb{E}\left[Q_j\right]_{bb} \\
&- 16p^{-2}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^4\sum_{k=1}^{n}\left(\mathbb{E}Q_{kk}\right)^2\mathbb{E}Q_{jj}\mathbb{E}\left[Q_k\right]_{jj} + O_z(p^{-\frac{5}{4}}) \\
&= -16n^2p^{-3}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\frac{1}{p}\operatorname{tr}(C^\circ)^4\,m^4(z) - 16np^{-2}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^4 m^4(z) + O_z(p^{-\frac{5}{4}})
\end{aligned}$$

As for $X_{23,j}$, it follows using Lemma 8 that:

$$\begin{aligned}
X_{23,j} &= \frac{8}{\sqrt{p}}\sum_{k\neq j}\mathbb{E}\left[\left(x_k^T C_j x_k\right)^2(x_k^T x_j)^2 Q_{kj} Q_{jj} Q_{kk}\right] + O_z(p^{-\frac{3}{2}+\epsilon}) \\
&= \frac{8}{\sqrt{p}}\sum_{k=1}^{n}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\mathbb{E}Q_{kk}\mathbb{E}Q_{jj}\mathbb{E}\left[(x_k^T x_j)^2 Q_{kj}\right] + O_z(p^{-\frac{5}{4}}) \\
&= -16p^{-2}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^4\sum_{k=1}^{n}\mathbb{E}Q_{jj}\left(\mathbb{E}Q_{kk}\right)^2\mathbb{E}\left[Q_k\right]_{jj} + O_z(p^{-\frac{5}{4}}) \\
&= -16np^{-2}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^4 m^4(z) + O_z(p^{-\frac{5}{4}})
\end{aligned}$$

We will now deal with $X_{21,j}$. It unfolds from the Integration by Parts formula that:

$$\begin{aligned}X_{21,j} &= -\frac{2}{p}\sum_{k\neq j}\frac{1}{p}\operatorname{tr} C_kC_j\mathbb{E}\left[x_k^TC_jx_kQ_{kk}Q_{jj}\right] - \frac{4}{p^2}\sum_{k\neq j}\mathbb{E}\left[x_k^TC_jC_kC_jx_kQ_{kk}Q_{jj}\right]\\ &+ \frac{8}{p\sqrt{p}}\sum_{k\neq j}\sum_{b\neq k}\mathbb{E}\left[x_k^TC_jx_kx_k^TC_jC_kx_bx_b^Tx_kQ_{jk}Q_{bj}Q_{kk}\right]\\ &+ \frac{8}{p\sqrt{p}}\sum_{k\neq j}\sum_{b\neq k}\mathbb{E}\left[x_k^TC_jx_kx_k^TC_jC_kx_bx_b^Tx_kQ_{bk}Q_{kk}Q_{jj}\right]\\ &= \upsilon_1+\upsilon_2+\upsilon_3+\upsilon_4\end{aligned}$$

Again, applying the Integration by Parts formula on υ_1, we obtain:

$$\begin{aligned}\upsilon_1 = &-\frac{2}{p}\sum_{k\neq j}\left(\frac{1}{p}\operatorname{tr} C_kC_j\right)^2\mathbb{E}\left[Q_{kk}Q_{jj}\right] + \frac{8}{p\sqrt{p}}\sum_{k\neq j}\frac{1}{p}\operatorname{tr} C_kC_j\sum_{b\neq k}\mathbb{E}\left[x_b^Tx_kx_b^TC_kC_jx_kQ_{bk}Q_{kk}Q_{jj}\right]\\ &+ \frac{8}{p\sqrt{p}}\sum_{k\neq j}\frac{1}{p}\operatorname{tr} C_kC_j\sum_{b\neq k}\mathbb{E}\left[x_b^Tx_kx_b^TC_kC_jx_kQ_{jk}Q_{bj}Q_{kk}\right]\end{aligned}$$

The first term in υ_1 can be decomposed as:

$$\begin{aligned}&-\frac{2}{p}\sum_{k\neq j}\left(\frac{1}{p}\operatorname{tr} C_kC_j\right)^2\mathbb{E}\left[Q_{kk}Q_{jj}\right] = -\frac{2}{p}\sum_{k=1}^{n}\left(\frac{1}{p}\operatorname{tr} C_kC_j\right)^2\mathbb{E}\left[Q_{kk}Q_{jj}\right] + \frac{2}{p}\left(\frac{1}{p}\operatorname{tr} C_j^2\right)^2\mathbb{E}Q_{jj}^2\\ &= -\frac{2}{p}\sum_{k=1}^{n}\left(\frac{1}{p}\operatorname{tr} C_kC_j\right)^2\mathbb{E}\left[\left(Q_{kk}-\mathbb{E}Q_{kk}\right)\left(Q_{jj}-\mathbb{E}Q_{jj}\right)\right] - \frac{2}{p}\sum_{k=1}^{n}\left(\frac{1}{p}\operatorname{tr} C_kC_j\right)^2\mathbb{E}Q_{kk}\mathbb{E}Q_{jj}\\ &+ \frac{2}{p}\left(\frac{1}{p}\operatorname{tr} C_j^2\right)^2\mathbb{E}Q_{jj}^2\end{aligned}$$

Since C_k and C_j can take finite values of matrices, the first term in the right hand-side of the above equation can be treated as C_k and C_j were equal to a constant matrix, thereby implying that:

$$\frac{2}{p}\sum_{k=1}^{n}\left(\frac{1}{p}\operatorname{tr} C_kC_j\right)^2\mathbb{E}\left[\left(Q_{kk}-\mathbb{E}Q_{kk}\right)\left(Q_{jj}-\mathbb{E}Q_{jj}\right)\right] = O_z(p^{-\frac{3}{2}+\epsilon})$$

We also have:

$$\frac{2}{p}\left(\frac{1}{p}\operatorname{tr} C_j^2\right)^2\mathbb{E}Q_{jj}^2 = \frac{2}{p}\frac{1}{p}\operatorname{tr}\left(C^\circ\right)^2\left(\mathbb{E}Q_{jj}\right)^2 + O_z(p^{-\frac{5}{4}})$$

We thus have:

$$\begin{aligned}&-\frac{2}{p}\sum_{k\neq j}\left(\frac{1}{p}\operatorname{tr} C_kC_j\right)^2\mathbb{E}\left[Q_{kk}Q_{jj}\right] = -\omega^2c_0g_n(z)\mathbb{E}Q_{jj} - \frac{1}{p}\sum_{k=1}^{n}\left[2\left(\frac{1}{p}\operatorname{tr} C_kC_j\right)^2-\omega^2\right]\mathbb{E}Q_{jj}\mathbb{E}Q_{kk}\\ &+ \frac{2}{p}\left(\frac{1}{p}\operatorname{tr}\left(C^\circ\right)^2\right)^2\left(\mathbb{E}Q_{jj}\right)^2 + O_z(p^{-\frac{5}{4}})\end{aligned}$$

The last term in υ_1 can be shown $O_z(p^{-2+\epsilon})$ using the result in Lemma 8. As for the second term, it follows from Lemma 9 that:

$$
\begin{aligned}
&\frac{8}{p\sqrt{p}}\sum_{k\neq j}\frac{1}{p}\operatorname{tr}C_kC_j\sum_{b\neq k}\mathbb{E}\left[x_b^Tx_kx_b^TC_kC_jx_kQ_{bk}Q_{kk}Q_{jj}\right]\\
&=-16p^{-3}\sum_{k=1}^n\sum_{b=1}^n\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\frac{1}{p}\operatorname{tr}(C^\circ)^4(\mathbb{E}Q_{kk})^2\,\mathbb{E}Q_{jj}\mathbb{E}\left[Q_k\right]_{bb}+O_z(p^{-\frac{5}{4}})\\
&=-16n^2p^{-3}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\frac{1}{p}\operatorname{tr}(C^\circ)^4m^4(z)+O_z(p^{-\frac{5}{4}})
\end{aligned}
$$

Plugging all these results together, we thus obtain:

$$
\begin{aligned}
\upsilon_1&=-\omega^2c_0g_n(z)\mathbb{E}Q_{jj}-\frac{1}{p}\sum_{k=1}^n\left[2\left(\frac{1}{p}\operatorname{tr}C_kC_j\right)^2-\omega^2\right]\mathbb{E}Q_{jj}\mathbb{E}Q_{kk}+\frac{2}{p}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2m^2(z)\\
&-16n^2p^{-3}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\frac{1}{p}\operatorname{tr}(C^\circ)^4m^4(z)+O_z(p^{-\frac{5}{4}})
\end{aligned}
$$

Using the same arguments as above it is easy to see that $\upsilon_3=O_z(p^{-2+\epsilon})$ and υ_2 and υ_4 can be approximated as:

$$
\begin{aligned}
\upsilon_2&=-\frac{4n}{p^2}\frac{1}{p}\operatorname{tr}(C^\circ)^4m^2(z)+O_z(p^{-\frac{5}{4}})\\
\upsilon_4&=-16p^{-3}n^2\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\frac{1}{p}\operatorname{tr}(C^\circ)^4m^4(z)+O_z(p^{-\frac{5}{4}})
\end{aligned}
$$

Hence,

$$
\begin{aligned}
X_{21,j}&=-\omega^2c_0g_n(z)\mathbb{E}Q_{jj}-\frac{1}{p}\sum_{k=1}^n\left[2\left(\frac{1}{p}\operatorname{tr}C_kC_j\right)^2-\omega^2\right]\mathbb{E}Q_{jj}\mathbb{E}Q_{kk}+\frac{2}{p}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2m^2(z)\\
&-\frac{4n}{p^2}\frac{1}{p}\operatorname{tr}(C^\circ)^4m^2(z)-32n^2p^{-3}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\frac{1}{p}\operatorname{tr}(C^\circ)^4m^4(z)+O_z(p^{-\frac{5}{4}})
\end{aligned}
$$

We finally get:

$$
\begin{aligned}
X_{2,j}&=-\omega^2c_0g_n(z)\mathbb{E}Q_{jj}-\frac{1}{p}\sum_{k=1}^n\left[2\left(\frac{1}{p}\operatorname{tr}C_kC_j\right)^2-\omega^2\right]\mathbb{E}Q_{jj}\mathbb{E}Q_{kk}+\frac{2}{p}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2m^2(z)\\
&-\frac{4n}{p^2}\frac{1}{p}\operatorname{tr}(C^\circ)^4m^2(z)-48n^2p^{-3}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\frac{1}{p}\operatorname{tr}(C^\circ)^4m^4(z)-32np^{-2}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^4m^4(z)\\
&+O_z(p^{-\frac{5}{4}})\\
&=-\omega^2c_0g_n(z)\mathbb{E}Q_{jj}-\frac{1}{p}\sum_{k=1}^n\left[2\left(\frac{1}{p}\operatorname{tr}C_kC_j\right)^2-\omega^2\right]\mathbb{E}Q_{jj}\mathbb{E}Q_{kk}+\frac{1}{p}\omega^2m^2(z)\\
&-\frac{4n}{p^2}\frac{1}{p}\operatorname{tr}(C^\circ)^4m^2(z)-\frac{24}{p}c_0^2\omega^2\frac{1}{p}\operatorname{tr}(C^\circ)^4m^4(z)-8c_0p^{-1}\omega^4m^4(z)+O_z(p^{-\frac{5}{4}})
\end{aligned}
$$

We will deal now with $X_{3,j}$. Using the Intregration by Parts formula, we have:

$$\begin{aligned}
X_{3,j} &= -2\sum_{k\neq j}\sum_{r\notin\{j,k\}}\mathbb{E}\left[\frac{1}{p}\left(x_r^TC_jx_k\right)^2Q_{rk}Q_{jj}\right] \\
&+\frac{6}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{j,k\}}\sum_{b\neq j}\mathbb{E}\left[x_b^Tx_jx_k^TC_jx_rx_b^TC_jx_kx_r^Tx_jQ_{rk}Q_{bj}Q_{jj}\right] \\
&+\frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{j,k\}}\sum_{b\neq j}\mathbb{E}\left[x_k^TC_jx_rx_r^Tx_jx_k^TC_jx_bx_b^Tx_jQ_{rj}Q_{bk}Q_{jj}\right] \\
&+\frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{j,k\}}\sum_{b\neq j}\mathbb{E}\left[x_k^TC_jx_rx_r^Tx_jx_k^TC_jx_bx_b^Tx_jQ_{jk}Q_{br}Q_{jj}\right] \\
&= X_{31,j}+X_{32,j}+X_{33,j}+X_{34,j}
\end{aligned}$$

We start by handling $X_{32,j}$. From Lemma 8, we have:

$$\begin{aligned}
X_{32,j} &= \frac{6}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{j,k\}}\mathbb{E}\left[x_k^Tx_jx_k^TC_jx_rx_r^Tx_jQ_{rk}x_k^TC_jx_kQ_{kj}Q_{jj}\right]+O_z(p^{-\frac{3}{2}+\epsilon}) \\
&= O_z(p^{-\frac{3}{2}+\epsilon})
\end{aligned}$$

Again using Lemma 8, we can prove that the sum of the contributions of all terms $b\notin\{k,j\}$ in $X_{33,j}$ are $O_z(p^{-\frac{3}{2}+\epsilon})$. We thus obtain:

$$\begin{aligned}
X_{33,j} &= \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{k,j\}}\mathbb{E}\left[x_k^TC_jx_rx_r^Tx_jx_k^Tx_jQ_{rj}\right]\frac{1}{p}\operatorname{tr}\ (C^\circ)^2\,\mathbb{E}Q_{kk}\mathbb{E}Q_{jj}+O_z(p^{-\frac{3}{2}+\epsilon}) \\
&= -8n^2p^{-3}m^4(z)\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\frac{1}{p}\operatorname{tr}(C^\circ)^4+O_z(p^{-\frac{5}{4}})
\end{aligned}$$

To handle X_{34}, we need the following lemma, the proof of which is deferred to Appendix F.

Lemma 11. *Let $j,k\in\{1,\cdots,n\}$ with $j\neq k$. Let A_1, A_2, A_3 and A_4 be four $n\times n$ matrices with bounded spectral norm. Then, for any $\epsilon>0$, we have:*

$$\max_{j\neq k}\mathbb{E}\left|\sum_{r\notin\{j,k\}}\sum_{b\notin\{j,r,k\}}x_b^TA_1x_kx_b^TA_2x_jx_k^TA_3x_rx_r^TA_4x_jQ_{br}\right|^2=O_z(p^{-3+\epsilon})$$

With the above lemma at hand, we are now ready to treat $X_{34,j}$. We have:

$$\begin{aligned}
X_{34,j} &= \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{j,k\}}\sum_{b\notin\{j,r,k\}}\mathbb{E}\left[x_k^TC_jx_rx_r^Tx_jx_k^TC_jx_bx_b^Tx_jQ_{jk}Q_{br}Q_{jj}\right] \\
&+\frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{j,k\}}\mathbb{E}\left[(x_k^TC_jx_r)^2(x_r^Tx_j)^2Q_{jk}Q_{rr}Q_{jj}\right] \\
&+\frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{j,k\}}\mathbb{E}\left[x_k^TC_jx_rx_r^Tx_jQ_{kr}x_k^TC_jx_kx_k^Tx_jQ_{jk}Q_{jj}\right]
\end{aligned}$$

The first term is $O_z(p^{-\frac{3}{2}+\epsilon})$ as per Lemma 11. The second and third terms can be also shown to be $O_z(p^{-\frac{3}{2}+\epsilon})$ using respectively lemma 8.

We now deal with $X_{31,j}$. Using the Integration by Parts formula, we obtain:

$$\begin{aligned}
X_{31,j} = &-2\sum_{k\neq j}\sum_{r\notin\{k,j\}}\frac{1}{p^2}\mathbb{E}\left[Q_{rk}Q_{jj}x_r^TC_jC_kC_jx_r\right]\\
&+4\sum_{k\neq j}\sum_{r\notin\{j,k\}}\sum_{b\notin\{k,r\}}\frac{1}{p^{\frac{3}{2}}}\mathbb{E}\left[x_b^Tx_kx_b^TC_kC_jx_rx_r^TC_jx_kQ_{rk}Q_{bk}Q_{jj}\right]\\
&+4\sum_{k\neq j}\sum_{r\notin\{k,j\}}\frac{1}{p^{\frac{3}{2}}}\mathbb{E}\left[x_r^Tx_kx_r^TC_kC_jx_rx_r^TC_jx_kQ_{rk}^2Q_{jj}\right]\\
&+4\sum_{k\neq j}\sum_{r\notin\{k,j\}}\sum_{b\notin\{k,r\}}\frac{1}{p^{\frac{3}{2}}}\mathbb{E}\left[x_b^Tx_kx_b^TC_kC_jx_rx_r^TC_jx_kQ_{kk}Q_{br}Q_{jj}\right]\\
&+4\sum_{k\neq j}\sum_{r\notin\{k,j\}}\frac{1}{p^{\frac{3}{2}}}\mathbb{E}\left[x_r^Tx_kx_r^TC_kC_jx_rx_r^TC_jx_kQ_{kk}Q_{rr}Q_{jj}\right]\\
&+8\sum_{k\neq j}\sum_{r\notin\{k,j\}}\sum_{b\neq k}\frac{1}{p^{\frac{3}{2}}}\mathbb{E}\left[x_b^Tx_kx_b^TC_kC_jx_rx_r^TC_jx_kQ_{rk}Q_{jk}Q_{bj}\right]
\end{aligned}$$

It is easy to see using again lemma 8 that the second, third and last terms are $O_z(p^{-\frac{3}{2}+\epsilon})$. It remains thus to handle the other terms. We have

$$-2\sum_{k\neq j}\sum_{r\notin\{k,j\}}\frac{1}{p^2}\mathbb{E}\left[Q_{rk}Q_{jj}x_r^TC_jC_kC_rx_r\right]=-\frac{4}{p}\frac{n^3}{p^3}\frac{\left(\frac{1}{p}\operatorname{tr}(C^\circ)^4\right)^2m^4(z)}{1-\frac{2n^2}{p^2}\frac{1}{p}\operatorname{tr}(C^\circ)^4m^2(z)}+O_z(p^{-\frac{5}{4}})$$

On the other hand, using Lemma 10, we can prove that the fourth term can be approximated as:

$$\begin{aligned}
&4\sum_{k\neq j}\sum_{r\notin\{k,j\}}\sum_{b\notin\{k,r\}}\frac{1}{p^{\frac{3}{2}}}\mathbb{E}\left[x_b^Tx_kx_b^TC_kC_jx_rx_r^TC_jx_kQ_{kk}Q_{br}Q_{jj}\right]\\
&=4\sum_{k\neq j}\sum_{r\notin\{k,j\}}\sum_{b\notin\{k,r\}}m^2(z)p^{-\frac{3}{2}}\mathbb{E}\left[x_b^Tx_kx_b^TC_kC_jx_rx_r^TC_jx_kQ_{br}\right]+O_z(p^{-\frac{5}{4}})\\
&=-8n^3p^{-4}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^4\right)^2m^4(z)+O_z(p^{-\frac{5}{4}})
\end{aligned}$$

It thus remains to handle the fifth term. Using the integration by parts formula, we can prove following the same kind of calculations that:

$$\begin{aligned}
&4\sum_{k\neq j}\sum_{r\notin\{k,j\}}\frac{1}{p^{\frac{3}{2}}}\mathbb{E}\left[x_r^Tx_kx_r^TC_kC_jx_rx_r^TC_jx_kQ_{kk}Q_{rr}Q_{jj}\right]=4\sum_{k\neq j}\sum_{r\notin\{k,j\}}p^{-\frac{5}{2}}\mathbb{E}\left[\left(x_r^TC_kC_jx_r\right)^2Q_{kk}Q_{rr}Q_{jj}\right]\\
&+O_z(p^{-\frac{3}{2}+\epsilon})\\
&=4\sum_{k=1}^n\sum_{r=1}^np^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr}C_rC_kC_j\right)^2\mathbb{E}\left[Q_{kk}Q_{rr}Q_{jj}\right]+O_z(p^{-\frac{3}{2}+\epsilon})\\
&\stackrel{(a)}{=}4\sum_{k=1}^n\sum_{r=1}^np^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr}C_rC_kC_j\right)^2\mathbb{E}\left[Q_{kk}\right]\mathbb{E}\left[Q_{rr}\right]\mathbb{E}\left[Q_{jj}\right]+O_z(p^{-\frac{3}{2}+\epsilon})
\end{aligned}$$

where (a) can be proved using the control of variance of $\operatorname{tr} Q$ in Lemma 5. Note that we cannot replace the diagonal elements of the resolvent matrix by $m(z)$ or the covariance matrices by C°, since this would produce an error $O_z(p^{-\frac{3}{4}})$. The treatment of all terms involved in $X_{3,j}$ is now finished. We have the following approximation:

$$\begin{aligned}
X_{3,j} &= -8n^2p^{-3}m^4(z)\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\frac{1}{p}\operatorname{tr}(C^\circ)^4 - \frac{4}{p}\frac{n^3}{p^3}\frac{\left(\frac{1}{p}\operatorname{tr}(C^\circ)^4\right)^2 m^4(z)}{1-\frac{2n^2}{p^2}\frac{1}{p}\operatorname{tr}(C^\circ)^4 m^2(z)} \\
&\quad + 4\sum_{k=1}^n\sum_{r=1}^n p^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr} C_rC_kC_j\right)^2\mathbb{E}Q_{kk}\mathbb{E}Q_{rr}\mathbb{E}Q_{jj} + O_z(p^{-\frac{5}{4}}) - 8c_0^3p^{-1}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^4\right)m^4(z) \\
&\quad + O_z(p^{-\frac{5}{4}}) \\
&= -4c_0^2p^{-1}\omega^2\frac{1}{p}\operatorname{tr}(C^\circ)^4 m^4(z) - \frac{4c_0^3}{p}\frac{\left(\frac{1}{p}\operatorname{tr}(C^\circ)^4\right)^2 m^4(z)}{1-\frac{2n^2}{p^2}\frac{1}{p}\operatorname{tr}(C^\circ)^4 m^2(z)} \\
&\quad + 4\sum_{k=1}^n\sum_{r=1}^n p^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr} C_rC_kC_j\right)^2\mathbb{E}Q_{kk}\mathbb{E}Q_{rr}\mathbb{E}Q_{jj} - 8c_0^3p^{-1}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^4\right)m^4(z) + O_z(p^{-\frac{5}{4}})
\end{aligned}$$

We now focus on the last term $X_{4,j}$. Using the Integration by Parts formula, we have:

$$\begin{aligned}
X_{4,j} &= -2\sum_{k\neq j}\sum_{r\neq j}\mathbb{E}\left[\frac{1}{p}\left(x_r^TC_jx_k\right)^2 Q_{jk}Q_{rj}\right] \\
&\quad + \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^TC_jx_rx_b^Tx_jx_b^TC_jx_kx_r^Tx_jQ_{jj}Q_{bk}Q_{rj}\right] \\
&\quad + \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^TC_jx_rx_b^Tx_jx_b^TC_jx_kx_r^Tx_jQ_{jk}Q_{bj}Q_{rj}\right] \\
&\quad + \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^TC_jx_rx_b^Tx_jx_b^TC_jx_kx_r^Tx_jQ_{rj}Q_{bj}Q_{jk}\right] \\
&\quad + \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^TC_jx_rx_b^Tx_jx_b^TC_jx_kx_r^Tx_jQ_{jj}Q_{br}Q_{jk}\right] \\
&= X_{41,j} + X_{42,j} + X_{43,j} + X_{44,j} + X_{45,j}
\end{aligned}$$

We will start by handling $X_{42,j}$. It is easy to see that summand over indexes $b \notin \{j,k\}$ and

$r \notin \{j,k\}$ is $O_z(p^{-\frac{3}{2}+\epsilon})$. We thus have:

$$\begin{aligned}
X_{42,j} &= \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{r\notin\{k,j\}} \mathbb{E}\left[x_k^T C_j x_r x_k^T x_j x_k^T C_j x_k x_r^T x_j Q_{jj} Q_{kk} Q_{rj}\right] \\
&+ \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{b\neq j} \mathbb{E}\left[x_k^T C_j x_k x_b^T x_j x_b^T C_j x_k x_k^T x_j Q_{jj} Q_{bk} Q_{kj}\right] \\
&= \frac{4}{\sqrt{p}} \sum_{k\neq j} m^2(z) \frac{1}{p} \operatorname{tr}(C^\circ)^2 \sum_{r\notin\{j,k\}} \mathbb{E}\left[x_k^T C_j x_r x_k^T x_j x_r^T x_j Q_{rj}\right] \\
&+ \frac{4}{\sqrt{p}} \sum_{k\neq j} \left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2 m^2(z) \mathbb{E}\left[(x_k^T x_j)^2 Q_{kj}\right] + O_z(p^{-\frac{5}{4}}) \\
&= -8n^2 p^{-3} \left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2 \frac{1}{p}\operatorname{tr}(C^\circ)^4 m^4(z) - 8np^{-2}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^4 m^4(z) + O_z(p^{-\frac{5}{4}})
\end{aligned}$$

Following the same kind of calculations, it is easy to see that $X_{43,j}$ and $X_{44,j}$ are $O_z(p^{-\frac{3}{2}+\epsilon})$. We will now focus on $X_{45,j}$. From Lemma 11, it follows that the summand over indexes $r \notin \{j,k\}$ and $b \notin \{j,r,k\}$ is $O_z(p^{-\frac{3}{2}+\epsilon})$. We thus have:

$$\begin{aligned}
X_{45} &= \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{b\neq j} \mathbb{E}\left[x_k^T C_j x_k x_b^T x_j x_b^T C_j x_k x_k^T x_j Q_{bk} Q_{jj} Q_{jk}\right] \\
&+ \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{r\notin\{j,k\}} \mathbb{E}\left[x_k^T C_j x_r x_k^T x_j x_k^T C_j x_k x_r^T x_j Q_{rk} Q_{jj} Q_{jk}\right] \\
&+ \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{r\notin\{j,k\}} \mathbb{E}\left[(x_k^T C_j x_r)^2 (x_r^T x_j)^2 Q_{jj} Q_{rr} Q_{jk}\right] + O_z(p^{-\frac{3}{2}+\epsilon}) \\
&= \frac{4}{\sqrt{p}} \sum_{k\neq j} \mathbb{E}\left[(x_k^T C_j x_k)^2 (x_k^T x_j)^2 Q_{kk} Q_{jj} Q_{jk}\right] + \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{r\notin\{j,k\}} \mathbb{E}\left[(x_k^T C_j x_r)^2 (x_r^T x_j)^2 Q_{jk} Q_{jj} Q_{rr}\right] \\
&+ O_z(p^{-\frac{3}{2}+\epsilon})
\end{aligned}$$

Using Lemma 10, the first term can be approximated as:

$$\frac{4}{\sqrt{p}} \sum_{k\neq j} \mathbb{E}\left[(x_k^T C_j x_k)^2 (x_k^T x_j)^2 Q_{kk} Q_{jj} Q_{jk}\right] = -8np^{-2}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^4 m^4(z) + O_z(p^{-\frac{5}{4}})$$

The second term is $O_z(p^{-\frac{3}{2}+\epsilon})$ as per Lemma 8.We thus have:

$$X_{45,j} = -8np^{-2}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^4 m^4(z) + O_z(p^{-\frac{5}{4}})$$

We will now handle $X_{41,j}$. Again using the Integration by Parts formula, we can prove that:

$$\begin{aligned}
X_{41,j} &= -2\frac{1}{p}\operatorname{tr}(C^\circ)^4 \mathbb{E}\left[\frac{1}{p^2}\left(\sum_{k\neq j} Q_{jk}\right)^2\right] - \frac{2}{p}\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2 \mathbb{E}\left[Q^2\right]_{jj} + \frac{2}{p} m^2(z) \left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2 \\
&+ O_z(p^{-\frac{5}{4}})
\end{aligned}$$

It can be prove by decomposing $\sum_{k\neq j} Q_{jk}$ as:

$$\sum_{k\neq j} Q_{jk} = \sum_{k\neq j} Q_{jk} - \sum_{k\neq j} \mathbb{E}_j Q_{jk} + \sum_{k\neq j} \mathbb{E}_j Q_{jk}$$

and following the same steps in the proof of Lemma 8 that:

$$\left(\sum_{k\neq j} Q_{jk}\right)^2 = O_z(p^{\epsilon}).$$

We thus have:

$$X_{41,j} = -\frac{2}{p}\left(\frac{1}{p}\operatorname{tr}(C^{\circ})^2\right)^2 \mathbb{E}\left[Q^2\right]_{jj} + \frac{2}{p}m^2(z)\left(\frac{1}{p}\operatorname{tr}(C^{\circ})^2\right)^2 + O_z(p^{-\frac{5}{4}})$$

We need to find an equivalent for the diagonal elements of Q^2. From the proof of Theorem 3, we can see that:

$$\mathbb{E}\left[Q^2\right]_{ii} = \frac{m^2(z)}{1-c_0\omega^2 m^2(z)} + O_z(p^{-\frac{1}{4}})$$

Plugging the equivalent of $\mathbb{E}\left[Q^2\right]_{ii}$ into the expression of $X_{41,j}$ and using the asymptotic approximations of $X_{42,j}$ and $X_{45,j}$, we ultimately get:

$$X_{4,j} = -\frac{1}{p}\frac{\omega^2 m^2(z)}{1-c_0\omega^2 m^2(z)} + \frac{1}{p}\omega^2 m^2(z) - \frac{4}{p}c_0^2\omega^2\frac{1}{p}\operatorname{tr}(C^{\circ})^4 m^4(z) - \frac{4c_0}{p}\omega^4 m^4(z) + O_z(p^{-\frac{5}{4}})$$

Having the approximations of $X_{i,j}$, $i = 1, \cdots, 4$, we are now ready to obtain a refined estimation of the diagonal elements of $\mathbb{E}Q$. From above calculations, we have:

$$z\mathbb{E}Q_{jj} = -1 - \omega^2 c_0 \mathbb{E}Q_{jj} g_n(z) + A_{p^{-\frac{1}{4}},j}(z) + A_{p^{-\frac{1}{2}},j}(z) + A_{p^{-1}(z)} + O_z(p^{-\frac{5}{4}}) \tag{31}$$

where

$$A_{p^{-\frac{1}{4}},j}(z) = -\frac{1}{p}\sum_{k=1}^{n}\left[2\left(\frac{1}{p}\operatorname{tr} C_k C_j\right)^2 - \omega^2\right]\mathbb{E}Q_{jj}\mathbb{E}Q_{kk}$$

$$A_{p^{-\frac{1}{2}},j}(z) = 4\sum_{k=1}^{n}\sum_{r=1}^{n} p^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr} C_r C_k C_j\right)^2 \mathbb{E}Q_{kk}\mathbb{E}Q_{rr}\mathbb{E}Q_{jj}$$

$$\begin{aligned} A_{p^{-1}}(z) &= \frac{\omega^2 m^2(z) - 14c_0\omega^4 m^4(z) + 12c_0^2\omega^6 m^6(z)}{1-\omega^2 c_0 m^2(z)} - 32\frac{c_0^2}{p}\omega^2\frac{1}{p}\operatorname{tr}(C^{\circ})^4 m^4(z) \\ &+ \frac{\frac{4c_0^3}{p}\left(\frac{1}{p}\operatorname{tr}(C^{\circ})^4\right)^2 m^4(z) - \frac{6c_0}{p}\frac{1}{p}\operatorname{tr}(C^{\circ})^4 m^2(z)}{1-\frac{2c_0^2}{p}\operatorname{tr}(C^{\circ})^4 m^2(z)} - 8c_0^3 p^{-1}\left(\frac{1}{p}\operatorname{tr}(C^{\circ})^4\right)^2 m^4(z)\end{aligned}$$

Summing (31) over index j, we obtain:

$$zg_n(z) = -1 - \omega^2 c_0 g_n^2(z) + \frac{1}{n}\sum_{j=1}^{n} A_{p^{-\frac{1}{4}},j}(z) + \frac{1}{n}\sum_{j=1}^{n} A_{p^{-\frac{1}{2}},j}(z) + A_{p^{-1}}(z) + O_z(p^{-\frac{5}{4}})$$

Let $f_j(z) = \mathbb{E}Q_{jj} - m(z)$. Then,

$$zf_j(z) = -\omega^2 c_0 f_j(z)m(z) - \omega^2 c_0 m(z)\frac{1}{n}\sum_{l=1}^{n} f_l(z) - \frac{1}{p}\sum_{k=1}^{n}\left[2\left(\frac{1}{p}\operatorname{tr} C_k C_j\right)^2 - \omega^2\right] m^2(z) + O_z(p^{-\frac{1}{2}})$$

Using the fact that $z + \omega^2 c_0 m(z) = -\frac{1}{m(z)}$, we obtain:

$$f_j(z) = \omega^2 c_0 m^2(z)\frac{1}{n}\sum_{l=1}^{n} f_l(z) + \frac{1}{p}\sum_{k=1}^{n}\left[2\left(\frac{1}{p}\operatorname{tr} C_k C_j\right)^2 - \omega^2\right] m^3(z) + O_z(p^{-\frac{1}{2}})$$

Define:

$$\delta_{kj} = 2\left(\frac{1}{p}\operatorname{tr} C_k C_j\right)^2 - \omega^2$$

and $\delta \in \mathbb{R}^{n\times n}$, the matrix $\delta = \left\{\delta_{kj} 1_{n_k} 1_{n_j}^T\right\}_{k,j=1}^{c}$. Let $f(z) = [f_1(z), \cdots, f_n(z)]$. Then,

$$\left(I_n - \omega^2 c_0 m^2(z)\frac{11^T}{n}\right) f = \frac{1}{p}\delta 1 m^3(z) + O_z(p^{-\frac{1}{2}})1$$

Since $\left|1 - \omega^2 c_0 m^2(z)\right|^{-1} = O_z(1)$, we have:

$$f(z) = \overline{f}(z) + O_z(p^{-\frac{1}{2}})1$$

where

$$\overline{f}(z) = \frac{1}{p}\delta 1 + \frac{\omega^2 c_0 m^5(z) 1^T \delta 1}{np(1 - \omega^2 c_0 m^2(z))} 1.$$

Let $\tilde{f}(z) = p^{\frac{1}{4}}\frac{1^T f}{n} = p^{-\frac{1}{4}}\frac{m^3(z)1^T\delta 1}{pn(1-\omega^2 c_0 m^2(z))}$. Hence,

$$g_n(z) - m(z) - \frac{1}{p^{\frac{1}{4}}}\tilde{f}(z) = O_z(p^{-\frac{1}{2}})$$

Now, let $h_j(z) = \mathbb{E}Q_{jj} - m(z) - \overline{f}(z)$. Substituting in (31) $\mathbb{E}Q_{jj}$ by $h_j(z) + m(z) + \overline{f}_j(z)$, we obtain:

$$zh_j(z) = -\omega^2 c_0 m(z)\frac{1}{n}\sum_{l=1}^{n} h_l(z) - \omega^2 c_0 h_j(z)m(z) - \omega^2 c_0 \overline{f}_j(z)\frac{1}{n}\sum_{l=1}^{n}\overline{f}_l(z)$$
$$- \frac{1}{p}\sum_{k=1}^{n}\delta_{kj}\overline{f}_j(z)m(z) - \frac{1}{p}\sum_{k=1}^{n}\delta_{kj}\overline{f}_k(z)m(z) + 4p^{-\frac{5}{2}}\sum_{k=1}^{n}\sum_{r=1}^{n}\left(\frac{1}{p}\operatorname{tr} C_r C_k C_j\right)^2 m^3(z) + O_z(p^{-\frac{3}{4}})$$

Let $h(z) = [h_1(z), \cdots, h_n(z)]^T$. Then,

$$\left(I_n - \omega^2 c_0 m^2(z)\frac{11^T}{n}\right) h = \omega^2 c_0 \overline{f}(z)m(z)\frac{1^T\overline{f}}{n} + m^2(z)\left\{\frac{1}{p}1^T\delta e_j \overline{f}_j(z)\right\}_{j=1}^{n} + \left\{\frac{1}{p}\overline{f}^T(z)\delta e_j\right\}_{j=1}^{n} m^2(z)$$
$$- \left\{4p^{-\frac{5}{2}}\sum_{k=1}^{n}\sum_{r=1}^{n}\left(\frac{1}{p}\operatorname{tr} C_r C_k C_j\right)^2\right\}_{j=1}^{n} m^4(z) + O_z(p^{-\frac{3}{4}})1$$

We thus have:

$$h = \overline{h}(z) + O_z(p^{-\frac{3}{4}})1$$

where

$$\begin{aligned}\overline{h}(z) &= \omega^2 c_0 m(z)\overline{f}(z)\frac{1^T\overline{f}(z)}{n} + \frac{\omega^4 c_0 m^3(z)\left(\frac{1^T\overline{f}(z)}{n}\right)^2}{1-\omega^2 c_0 m^2(z)}1 + m^2(z)\left\{\frac{1}{p}1^T\delta e_j\overline{f}_j(z)\right\}_{j=1}^n \\ &+ \frac{2m^4(z)\omega^2 c_0 1^T\delta\overline{f}(z)}{np(1-\omega^2 c_0 m^2(z))}1 + m^2(z)\left\{\frac{1}{p}\overline{f}(z)^T\delta e_j\right\}_{j=1}^n - 4p^{-\frac{5}{2}}\sum_{k=1}^n\sum_{r=1}^n\left(\frac{1}{p}\operatorname{tr} C_r C_k C_j\right)^2 m^4(z) \\ &- \frac{4p^{-\frac{5}{2}}}{n(1-\omega^2 c_0 m^2(z))}\sum_{j=1}^n\sum_{k=1}^n\sum_{r=1}^n\left(\frac{1}{p}\operatorname{tr} C_r C_k C_j\right)^2 m^6(z) + O_z(p^{-\frac{3}{4}})\end{aligned}$$

Let $\tilde{h}(z) = \sqrt{p}\frac{1^T\overline{h}}{n}$. Then,

$$\begin{aligned}\tilde{h}(z) = \sqrt{p}&\frac{\omega^2 c_0 m(z)\left(\frac{1^T\overline{f}(z)}{n}\right)^2}{1-\omega^2 c_0 m^2(z)} + \sqrt{p}\frac{2m^2(z)1^T\delta\overline{f}(z)}{np\left(1-\omega^2 c_0 m^2(z)\right)} \\ &- 4p^{-2}\frac{m^4(z)}{n(1-\omega^2 c_0 m^2(z))}\sum_{j=1}^n\sum_{k=1}^n\sum_{r=1}^n\left(\frac{1}{p}\operatorname{tr} C_r C_k C_j\right)^2\end{aligned}$$

We thus have:

$$g_n(z) - m(z) - \frac{1}{p^{\frac{1}{4}}}\tilde{f}(z) - \frac{1}{\sqrt{p}}\tilde{h}(z) = O_z(p^{-\frac{3}{4}})$$

We will now determine an asymptotic equivalent for the term vanishing at rate $O_z(p^{-\frac{3}{4}})$. To this end, let $k_j(z) = \mathbb{E}Q_{jj} - m(z) - \overline{f}_j(z) - \overline{h}_j(z)$. Again, substituting $\mathbb{E}Q_{jj}$ by $m(z) + \overline{f}_j(z) + \overline{h}_j(z)$ into (31) yields:

$$\begin{aligned}zk_j(z) &= -\omega^2 c_0 m(z)\frac{1}{n}\sum_{l=1}^n k_l(z) - \omega^2 c_0\overline{f}_j(z)\frac{1^T\overline{h}(z)}{n} - \omega^2 c_0\overline{h}_j(z)\frac{1^T\overline{f}(z)}{n} \\ &- \omega^2 c_0 k_j(z)m(z) - \frac{1}{p}\sum_{k=1}^n\delta_{kj}\overline{f}_j(z)\overline{f}_k(z) - \frac{1}{p}\sum_{k=1}^n\delta_{kj}\overline{h}_j(z)m(z) - \frac{1}{p}\sum_{k=1}^n\delta_{kj}m(z)\overline{h}_k(z) \\ &+ 4\sum_{k=1}^n\sum_{r=1}^n p^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr} C_r C_k C_j\right)^2\left(m^2(z)\overline{f}_r(z) + m^2(z)\overline{f}_k(z) + m^2(z)\overline{f}_j(z)\right) + O_z(p^{-1})\end{aligned}$$

Let $k(z) = [k_1(z), \cdots, k_n(z)]^T$. Hence,

$$k(z) = \overline{k}(z) + O_z(p^{-1})1$$

with

$$\overline{k}(z) = \omega^2 c_0 m(z)\overline{f}(z)\frac{1^T\overline{h}(z)}{n} + \frac{2\omega^4 c_0^2 m^3(z)\frac{1^T\overline{f}(z)}{n}\frac{1^T\overline{h}(z)}{n}}{1-\omega^2 c_0 m^2(z)}1 + \frac{1^T\overline{f}(z)}{n}\omega^2 c_0\overline{h}(z)m(z)$$
$$+ \frac{1}{p}m(z)\left\{\overline{f}^T(z)\delta e_j\overline{f}_j(z)\right\}_{j=1}^n + \frac{\omega^2 c_0 m^3(z)\overline{f}^T(z)\delta\overline{f}(z)}{np(1-\omega^2 c_0 m^2(z))} + \frac{m^2(z)}{p}\left\{1^T\delta e_j\overline{h}_j(z)\right\}_{j=1}^n$$
$$+ \frac{2\omega^2 c_0 m^4(z)1^T\delta\overline{h}(z)}{np(1-\omega^2 c_0 m^2(z))}1 + \frac{m^2(z)}{p}\left\{\overline{h}^T(z)\delta e_j\right\}_{j=1}^n - 8p^{-\frac{5}{2}}\left\{\sum_{k=1}^n\sum_{r=1}^n\left(\frac{1}{p}\operatorname{tr} C_r C_k C_j\right)^2 m^3(z)\overline{f}_r(z)\right\}_{j=1}^n$$
$$- 4\left\{\sum_{k=1}^n\sum_{r=1}^n p^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr} C_r C_k C_j\right)^2 m^3(z)\overline{f}_j(z)\right\}_{j=1}^n$$
$$- \frac{12\omega^2 c_0 m^5(z)}{n(1-\omega^2 c_0 m^2(z))}\sum_{j=1}^n\sum_{k=1}^n\sum_{r=1}^n p^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr} C_r C_k C_j\right)^2\overline{f}_r(z)1$$

Let $\tilde{k}(z) = p^{\frac{3}{4}}\frac{1^T\overline{k}(z)}{n}$. Then:

$$\tilde{k}(z) = p^{\frac{3}{4}}\frac{2\omega^2 c_0 m(z)\frac{1^T\overline{f}(z)}{n}\frac{1^T\overline{h}(z)}{n}}{1-\omega^2 c_0 m^2(z)} + p^{\frac{3}{4}}\frac{1}{np}\frac{\overline{f}^T(z)\delta\overline{f}(z)}{1-\omega^2 c_0 m^2(z)} + p^{\frac{3}{4}}\frac{2m^2(z)1^T\delta\overline{h}(z)}{np(1-\omega^2 c_0 m^2(z))}$$
$$- \frac{12m^3(z)p^{-\frac{7}{4}}}{n(1-\omega^2 c_0 m^2(z))}\sum_{j=1}^n\sum_{k=1}^n\sum_{r=1}^n\left(\frac{1}{p}\operatorname{tr} C_r C_k C_j\right)^2\overline{f}_r(z)$$

Moreover, we have:

$$g_n(z) = m(z) + p^{-\frac{1}{4}}\tilde{f}(z) + p^{-\frac{1}{2}}\tilde{h}(z) + p^{-\frac{3}{4}}\tilde{k}(z) + O_z(p^{-1})$$

Finally, it remains to determine an asymptotic equivalent for the term vanishing at pace $O_z(p^{-1})$. Let $e_j(z) = \mathbb{E}Q_{jj} - m(z) - \overline{f}(z) - \overline{h}(z) - \overline{k}(z)$, and $e(z) = [e_1(z), \cdots, e_n(z)]^T$.

$$z\frac{1^T e(z)}{n} = -\omega^2 c_0\frac{1^T e(z)}{n}m(z) - \omega^2 c_0 m(z)\frac{1^T e(z)}{n} - 2\omega^2 c_0\frac{1^T\overline{f}(z)}{n}\frac{1^T\overline{k}(z)}{n} - \omega^2 c_0\left(\frac{1^T\overline{h}(z)}{n}\right)^2$$
$$- \frac{2}{np}m(z)\overline{k}^T(z)\delta 1 - \frac{2}{np}\overline{f}^T(z)\delta\overline{h}(z) + 12np^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr} C_0^3\right)^2\frac{1^T\overline{h}(z)}{n}$$
$$+ 12m(z)\left(1^T\overline{f}(z)\right)^2 p^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr} C_0^3\right)^2 + A_{p^{-1}}(z) + O_z(p^{-\frac{5}{4}})$$

Using the fact that $z + 2\omega^2 c_0 m(z) = -\frac{1}{m(z)}(1-\omega^2 c_0 m^2(z))$, we obtain:

$$\frac{1^T e(z)}{n} = p^{-1}\tilde{e}(z) + O_z(p^{-\frac{5}{3}})$$

where

$$\tilde{e}(z) = \frac{pm(z)}{1-\omega^2 c_0 m^2(z)}\left[2\omega^2 c_0\frac{1^T\overline{f}(z)}{n}\frac{1^T\overline{k}(z)}{n} + \omega^2 c_0\left(\frac{1^T\overline{h}(z)}{n}\right)^2 + \frac{2}{np}\overline{k}^T(z)\delta 1 m(z)\right.$$
$$\left. + \frac{2}{np}\overline{f}^T(z)\delta\overline{h}(z) - 12np^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr} C_0^3\right)^2\frac{1^T\overline{h}(z)}{n} - 12m(z)\left(1^T\overline{f}(z)\right)^2 p^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr} C_0^3\right)^2 - A_{p^{-1}}(z)\right]$$

Gathering all the above approximations togother, we obtain (30).

7.1. *Almost sure location of the eigenvalues of* Φ. With the approximation in (30) at hand, we are now ready to determine the limiting support of the empirical measure of Φ. We first need to prove that $\tilde{f}(z)$, $\tilde{h}(z)$, $\tilde{k}(z)$ and $\tilde{e}(z)$ are Stieljtes transforms of some distributions and determine the supports thereof. To this end, we will resort to the following Lemma.

LEMMA 12. *[Loubaton, 2016, Lemma 4] Let Λ be a distribution on $\mathbb{R}$ with compact support. Define its Stieltjes transform $l : \mathbb{C}\backslash\mathbb{R} \to \mathbb{C}$ by:*

$$l(z) = \Lambda\left(\frac{1}{x-z}\right)$$

Then l is analytic in $\mathbb{C}\backslash\mathbb{R}$ and has analytic continuation to $\mathbb{C}\backslash\operatorname{supp}(\Lambda)$. Moreover,

c_1) *$l(z) \to 0$ as $|z| \to \infty$,*

c_2) *There exists a constant $C > 0$, $k \in \mathbb{N}$ and a compact set $K \subset \mathbb{E}$ containing $\operatorname{supp}(\Lambda)$ such that for any $z \in \mathbb{C}\backslash\mathbb{R}$,*

$$|l(z)| \leq C \max\left\{\operatorname{dist}(z, K)^{-k}, 1\right\}$$

c_3) *for any $\phi \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ with compact support,*

$$\Lambda(\phi) = -\frac{1}{\pi} \lim_{y \to 0^+} \Im \int_{\mathbb{R}} \phi(x) l(x + \imath y) dx$$

c_4) *If $\lim_{|z|\to\infty} |zl(z)| = 0$, then it holds that:*

$$\Lambda(1) = 0.$$

Conversely if K is a compact subset of $\mathbb{R}$ and if $l : \mathbb{C}\backslash K \to \mathbb{C}$ is an analytic function satisfying c_1) and c_2) above, then l is the Stieltjes transform of a compactly supported distribution Λ on $\mathbb{R}$. Moreover, $\operatorname{supp}(\Lambda)$ is exactly the set of singular points of l in K.

From the expressions of $\tilde{f}(z)$ and $\tilde{h}(z)$, $\tilde{k}(z)$ and $\tilde{e}(z)$, we can easily see that all of them are analytic on $\mathbb{C}\backslash\left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right]$ except $\tilde{e}(z)$ which presents singularity for z such that

$$m(z) = \pm\sqrt{\frac{1}{\frac{2c_0^2}{p}\operatorname{tr}(C^\circ)^4}}. \tag{32}$$

This singularity falls outside the support if $\frac{2c_0^2}{p}\operatorname{tr}(C^\circ)^4 > \sqrt{c_0}\omega$, or equivalently $\Omega > \frac{1}{\sqrt{c_0}\omega}$, in which case the z's satisfying (32) are given by the two isolated complex values $\{-\tilde{\rho}, \tilde{\rho}\}$ where $\tilde{\rho} = c_0\Omega + \frac{\omega^2}{\Omega}$.

PROPOSITION 3. *$\tilde{f}(z)$ and $\tilde{h}(z)$, $\tilde{k}(z)$ are the Stieltjes transforms of distributions $\Lambda_{\tilde{f}}$, $\Lambda_{\tilde{h}}$, $\Lambda_{\tilde{k}}$ with support $\mathcal{S} = \left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right]$ while $\tilde{e}(z)$ is the Stieltjes transform of $\Lambda_{\tilde{e}}$ with support $\mathcal{S}_e = \mathcal{S} \cup \{-\tilde{\rho}, \tilde{\rho}\}$. Moreover, $\Lambda_{\tilde{f}}(1) = \Lambda_{\tilde{h}}(1) = \Lambda_{\tilde{k}}(1) = \Lambda_{\tilde{e}}(1) = 0$.*

PROOF. We will prove the result only for $\tilde{f}(z)$. The same reasoning can be applied to $\tilde{h}(z)$, $\tilde{k}(z)$. For $\tilde{e}(z)$, some slight modifications should be made to account for the singularities

$\{-\tilde{\rho}, \tilde{\rho}\}$. According to Lemma 12, it suffices to show that $\tilde{f}(z)$ satisfy conditions c_1 and c_2. Let $|z| \geq 4\sqrt{3c_0}\omega$, then there exists positive constants C and η such that:

$$\left|\tilde{f}(z)\right| \leq C\frac{|z+\eta|^4}{|z|^4\left||z|-2\sqrt{c_0}\omega\right|^3}$$

Hence $\tilde{f}(z)$ converges to zero as $|z|$ goes to infinity. It remains to check the condition c_2). To this end, we follow the same approach in [Capitaine et al., 2009]. We define the interval[1]:

$$K = [-1-2\sqrt{c_0}\omega, 1+2\sqrt{c_0}\omega]$$

Let $D = \{z \in \mathbb{C}, 0 < \mathrm{dist}(z,K) \leq 1\}$. We need to distinguish the following cases:

- Let $z \in D \cup \mathbb{C}\backslash\mathbb{R}$ with $\Re z \in K$. We have $\mathrm{dist}(z,K) = |\Im z| \leq 1$. Then, it is clear that there exists a constant C_0 such that:

$$\left|\tilde{f}(z)\right| \leq C_0|\Im z|^{-7} = C_0\mathrm{dist}(z,K)^{-7} = C_0 \max\left(\mathrm{dist}(z,K)^{-7}, 1\right)$$

- Let $z \in D\cup\mathbb{C}\backslash\mathbb{R}$ with $\Re z \notin K$. Since $\tilde{f}(z)$ is bounded on compact subsets of $\mathbb{C}\backslash\left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right]$, we easily deduce that there exists a constant C_1 such that for any $z \in D$ with $\Re z \notin K$,

$$\left|\tilde{f}(z)\right| \leq C_1\mathrm{dist}(z,K)^{-7} = C_1 \max\left(\mathrm{dist}(z,K)^{-7}, 1\right)$$

- Since $\left|\tilde{f}(z)\right| \to 0$ when $|z| \to \infty$, $\tilde{f}(z)$ is bounded on $\mathbb{C}\backslash D$. Thus, there eixsts some constant C_2 such that for any $z \in \mathbb{C}\backslash D$,

$$\left|\tilde{f}(z)\right| \leq C_2 = C_2 \max\left(\mathrm{dist}(z,K)^{-4}, 1\right)$$

This shows that condition c_2) is satisfied. Hence, $\tilde{f}(z)$ is the Stieltjes transform of a distribution $\Lambda_{\tilde{f}}$ whose support is in $\mathcal{S}$. Moreover, as $\lim_{|z|\to\infty} z\tilde{f}(z) = 0$, we have $\Lambda_{\tilde{f}}(1) = 0$. □

Using Proposition 3, we prove the following Lemma which evaluates the speed of convergence of the first moment as well as the central moments of $\frac{1}{n}\operatorname{tr}\psi(\Phi)$ for ψ smooth, constant on the complementary of a compact interval and vanishing on $\mathcal{S} = \left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right] \cup \{-\tilde{\rho}, \tilde{\rho}\}$:

Lemma 13. *Assume that* $\Omega > \frac{1}{\sqrt{c_0}\omega}$. *Let* $\epsilon > 0$ *be any small positive scalar. For all smooth function* ψ *constant on the complementary of a compact interval and vanishing on* $\mathcal{S} = \left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right] \cup \{-\tilde{\rho}, \tilde{\rho}\}$,

$$\mathbb{E}\left[\frac{1}{n}\operatorname{tr}\psi(\Phi)\right] = O(p^{-\frac{5}{4}}) \tag{33}$$

$$\mathbb{E}\left|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\psi(\Phi)\right|^{2l} = O(p^{-\frac{5l}{2}}) \tag{34}$$

for each $l \geq 1$.

[1]Note that if $\tilde{e}(z)$ was considered and $\Omega > \frac{1}{\sqrt{c_0}\omega}$, then the interval K should be set to $K = [-\tilde{\rho}-1, \tilde{\rho}+1]$.

PROOF. Using the inverse Stieltjes transform, it holds that for any smooth function ψ_c with compact support:

$$\frac{1}{n}\mathbb{E}\left[\operatorname{tr}\psi_c(\Phi)\right] = \int \psi_c d\mu + \frac{1}{p^{\frac{1}{4}}}\Lambda_{\tilde{f}}(\psi_c) + \frac{1}{\sqrt{p}}\Lambda_{\tilde{h}}(\psi_c) + \frac{1}{p^{\frac{3}{4}}}\Lambda_{\tilde{k}}(\psi_c) + \frac{1}{p}\Lambda_{\tilde{e}}(\psi_c) - \frac{1}{\pi}\lim_{y\to 0^+}\Im\int_{\mathbb{R}}\psi_c(x)R_n(x+iy)dx$$

where $R_n = g_n(z) - m(z) - \frac{1}{p^{\frac{1}{4}}}\tilde{f}(z) - \frac{1}{\sqrt{p}}\tilde{h}(z) - \frac{1}{p^{\frac{3}{4}}}\tilde{k}(z) - \frac{1}{p}\tilde{e}(z)$. Since, for $z \in \mathbb{C}\backslash\mathbb{R}$,

$$|R_n(z)| = O_z(p^{-\frac{5}{4}})$$

we have, using the ideas of [Haagerup and Thorbjørnsen, 2005],

$$\frac{1}{\pi}\lim_{y\to 0^+}\Im\int_{\mathbb{R}}\psi_c(x)R_n(x+iy)dx = O(p^{-\frac{5}{4}})$$

In order to prove (33), we follow the approach in [Loubaton, 2016]. We denote κ the constant for which $\psi(x) = \kappa$ for x lying outside a compact set. Function $\psi_c = \psi - \kappa$ is compactly supported and $\int \psi_c(\lambda)d\mu(\lambda) = -\kappa$. Moreover, we have from Proposition 3, $\Lambda_{\tilde{f}}(\psi_c) = \Lambda_{\tilde{h}}(\psi_c) = \Lambda_{\tilde{k}}(\psi_c) = \Lambda_{\tilde{e}}(\psi_c) = 0$. Hence,

$$\begin{aligned}\frac{1}{n}\mathbb{E}\left[\operatorname{tr}\psi(\Phi)\right] &= \frac{1}{n}\mathbb{E}\left[\operatorname{tr}\psi_c(\Phi)\right] + \kappa + O(p^{-\frac{5}{4}})\\ &= O(p^{-\frac{5}{4}})\end{aligned}$$

In order to prove (34), we proceed by induction over l. For $l = 1$, using the Poincaré-Nash inequality, we have:

$$\begin{aligned}\operatorname{var}\left(\frac{1}{n}\operatorname{tr}\psi(\Phi)\right) &\leq \frac{1}{n^2}\sum_{i=1}^{p}\sum_{j=1}^{n}\mathbb{E}\left|\frac{\operatorname{tr}\psi(\Phi)}{\partial Z_{ij}}\right|^2\\ &\leq \frac{1}{n^2}\sum_{i=1}^{p}\sum_{j=1}^{n}\mathbb{E}\left|\operatorname{tr}\psi'(\Phi)\left(\frac{\partial \Phi_{lk}}{\partial Z_{ij}}\right)\right|^2\\ &= \frac{4}{n^2}\sum_{i=1}^{p}\sum_{j=1}^{n}\mathbb{E}\left|\operatorname{tr}\psi'(\Phi)\left(x_l^T x_k\left[C_k^{\frac{1}{2}}x_l\right]_i \boldsymbol{\delta}_{jk}\boldsymbol{\delta}_{l\neq k} + (x_l^T x_k)\left[C_l^{\frac{1}{2}}x_k\right]_i \boldsymbol{\delta}_{jl}\boldsymbol{\delta}_{l\neq k}\right)_{l,k=1}^{n}\right|^2\\ &= \frac{16}{n^2}\sum_{i=1}^{p}\sum_{j=1}^{n}\mathbb{E}\left|\sum_{a\neq j}\left[\psi'(\Phi)\right]_{aj}(x_j^T x_a)\left[C_j^{\frac{1}{2}}x_a\right]_i\right|^2\end{aligned}$$

Define R as :

$$[R]_{a_1 a_2} = \max_{j\neq a_1}\left|x_j^T x_{a_1}\right|\max_{j\neq a_2}\left|x_j^T x_{a_2}\right|\max_{j\notin\{a_1,a_2\}}\left|x_{a_2}^T C_j x_{a_1}\right|$$

It is easy to see that $\|R\| = O(p^{-\frac{1}{2}})$. Let $h(x) = |\psi'(x)|$ Hence:

$$\operatorname{var}\left(\frac{1}{n}\operatorname{tr}\psi(\Phi)\right) \leq \frac{16}{n^2}\operatorname{tr}\left(h(\Phi)Rh(\Phi)\right) \leq \frac{16}{n^2}\mathbb{E}\|R\|\operatorname{tr}h^2(\Phi)$$

From (33), $\mathbb{E}\frac{1}{n}\operatorname{tr} h^2(\Phi) = O(p^{-\frac{5}{4}})$. Using the fact that $\|R\| = O(p^{-\frac{1}{2}})$, we obtain:

$$\operatorname{var}\left(\frac{1}{n}\operatorname{tr}\psi(\Phi)\right) = O(p^{-\frac{11}{4}}) = O(p^{-\frac{5}{2}})$$

Assume now that (34) holds for all $l \neq k-1$. We will prove it for $l = k$. Note that:
(35)

$$\mathbb{E}\left|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\Phi\right|^{2k} = \left(\mathbb{E}\left|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\Phi\right|^{k}\right)^2 + \operatorname{var}\left(\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\psi(\Phi)\right)^k$$

The Hölder inequality can be used to treat the first term in the right-hand side of the above equation. This leads to:

$$\mathbb{E}\left|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\Phi\right|^{k} \leq \sqrt{\mathbb{E}\left|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\Phi\right|^{2k-2}}\sqrt{\mathbb{E}\left|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\Phi\right|^{2}}$$

Using the induction assumption, it unfolds that:

$$\mathbb{E}\left|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\Phi\right|^{k} = O(p^{-\frac{5k}{4}})$$

We will now handle the second term in the right-hand side of (35). Using the Poincaré-Nash inequality, we obtain:

$$\begin{aligned}\operatorname{var}\left(\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\psi(\Phi)\right)^k &\leq \sum_{i=1}^{p}\sum_{j=1}^{n} k^2\mathbb{E}\left[\left|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\psi(\Phi)\right|\left|\frac{\partial\frac{1}{n}\operatorname{tr}\psi(\Phi)}{\partial Z_{ij}}\right|^2\right]\\ &\leq \frac{16k^2}{n^2}\mathbb{E}\left[\left|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\psi(\Phi)\right|^{2(k-1)}\operatorname{tr} h(\Phi)Rh(\Phi)\right]\\ &\leq \frac{16k^2}{n^2}\left(\mathbb{E}\left|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\psi(\Phi)\right|^{2k}\right)^{\frac{k-1}{k}}\left|\mathbb{E}\left|\operatorname{tr} h(\Phi)Rh(\Phi)\right|^k\right|^{\frac{1}{k}}\end{aligned}$$

Recall that

$$\mathbb{E}\left|\operatorname{tr} h(\Phi)Rh(\Phi)\right|^k \leq \mathbb{E}\|R\|^k\left|\operatorname{tr} h^2(\Phi)\right|^k$$

where $\|R\|^k = O(p^{-\frac{k}{2}})$; it suffices thus to treat $\mathbb{E}\left|\operatorname{tr} h^2(\Phi)\right|^k$. We have:

$$\begin{aligned}\mathbb{E}\left|\operatorname{tr} h^2(\Phi)\right|^k &\leq 2^{k-1}\mathbb{E}\left|\operatorname{tr} h^2(\Phi) - \mathbb{E}\operatorname{tr} h^2(\Phi)\right|^k + 2^{k-1}\left|\mathbb{E}\operatorname{tr} h^2(\Phi)\right|^k\\ &\leq 2^{k-1}\sqrt{\mathbb{E}\left|\operatorname{tr} h^2(\Phi) - \mathbb{E}\operatorname{tr} h^2(\Phi)\right|^{2k-2}}\sqrt{\mathbb{E}\left|\operatorname{tr} h^2(\Phi) - \mathbb{E}\operatorname{tr} h^2(\Phi)\right|^{2}} + 2^{k-1}\left|\mathbb{E}\operatorname{tr} h^2(\Phi)\right|^k\end{aligned}$$

Using the induction assumption along with (33), it unfolds that:

$$\mathbb{E}\left|\operatorname{tr} h(\Phi)Rh(\Phi)\right|^k = O(p^{-\frac{3k}{4}})$$

Let $\kappa_p = \mathbb{E}\left|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\psi(\Phi)\right|^{2k}$. From the previous derivations, it is easy to see that there exists positive constants C_1 and C_2 such that:

(36) $$\kappa_p \leq C_1 p^{-\frac{5}{2}k} + C_2 \kappa_p^{\frac{k-1}{k}} p^{-\frac{11}{4}}$$

Let $u_p = \kappa_p p^{\frac{5}{2}k}$. To conclude the proof, it suffices to check that u_p is a bounded sequence. Expressing (36) in terms of u_p, we obtain:

$$u_p \leq C_1 + C_2 u_p p^{-\frac{1}{4}}$$

thus proving that u_p is a bounded sequence. This finishes the proof of Lemma 13. □

A direct consequence of Lemma 13 is that for any ψ satisfying the condition of Lemma 13,

$$\operatorname{tr} \psi(\Phi) \xrightarrow{\text{a.s.}} 0.$$

We will now terminate the proof of Theorem 4. We will consider only the case when $\Omega > \frac{1}{\sqrt{c_0}\omega}$.

Let $\epsilon > 0$, and take ψ smooth such that

- $\psi(x) = 1$ for all $x \notin \left[-2\sqrt{c_0}\omega - \epsilon, 2\sqrt{c_0}\omega + \epsilon\right] \cup [-\tilde{\rho} - \epsilon, -\tilde{\rho} + \epsilon] \cup [\tilde{\rho} - \epsilon, \tilde{\rho} + \epsilon]$
- $\psi(x) = 0$ for $x \in \left[-2\sqrt{c_0}\omega - \frac{\epsilon}{2}, 2\sqrt{c_0}\omega + \frac{\epsilon}{2}\right] \cup [-\tilde{\rho} - \epsilon, -\tilde{\rho} - \epsilon] \cup [\tilde{\rho} - \epsilon, \tilde{\rho} + \epsilon]$
- $0 \leq \psi(x) \leq 1$ elsewhere

Function ψ satisfies the conditions of Lemma 13. Hence, we have:

$$\operatorname{tr} \psi(\Phi) \xrightarrow{\text{a.s.}} 0.$$

Since $\operatorname{tr} \psi(\Phi)$ is greater than the number of eigenvalues lying outside $\left[-2\sqrt{c_0}\omega - \epsilon, 2\sqrt{c_0}\omega + \epsilon\right] \cup [-\tilde{\rho} - \epsilon, -\tilde{\rho} - \epsilon] \cup [\tilde{\rho} - \epsilon, \tilde{\rho} + \epsilon]$, we conclude that almost surely for n large enough, there is no eigenvalue of Φ outside $\left[-2\sqrt{c_0}\omega - \epsilon, 2\sqrt{c_0}\omega + \epsilon\right] \cup [-\tilde{\rho} - \epsilon, -\tilde{\rho} - \epsilon] \cup [\tilde{\rho} - \epsilon, \tilde{\rho} + \epsilon]$.

APPENDIX A: PROOF OF LEMMA 3

Expanding $\xi_{(k,-k)}^T A \xi_{(k,-k)}$, we get:

$$\begin{aligned}\mathbb{E}_{x_k}\left[\xi_{(k,-k)}^T A \xi_{(k,-k)}\right] &= p \sum_{i \neq k} \sum_{j \neq k} \mathbb{E}_{x_k}\left[(x_i^T x_k)^2 A_{ij} (x_j^T x_k)^2\right] - \sum_{i \neq k} \sum_{j \neq k} x_i^T C_k x_i A_{ij} \frac{1}{p^2} \operatorname{tr} C_k C_j \\ &- \sum_{i \neq k} \sum_{j \neq k} x_j^T C_k x_j A_{ij} \frac{1}{p^2} \operatorname{tr} C_k C_i + \sum_{i \neq k} \sum_{j \neq k} \frac{1}{p^{3/2}} \operatorname{tr} C_k C_j \frac{1}{p^{3/2}} \operatorname{tr} C_k C_i A_{ij} \qquad (37)\end{aligned}$$

We will deal with the first term. We have:

$$\mathbb{E}_{x_k}\left[p\sum_{i\neq k}\sum_{j\neq k}(x_i^Tx_k)^2A_{ij}(x_j^Tx_k)^2\right]=\frac{1}{p^3}\sum_{i\neq k}\sum_{j\neq k}\sum_{\substack{l_1,l_2,s_1,s_2\\ m_1,m_2,n_1,n_2}}\mathbb{E}\left[Z_{l_1k}\left[C_k^{\frac12}C_i^{\frac12}\right]_{l_1l_2}Z_{l_2i}Z_{s_1k}\left[C_k^{\frac12}C_i^{\frac12}\right]_{s_1s_2}\right.$$

$$\left.\times Z_{s_2i}A_{ij}Z_{m_1k}\left[C_k^{\frac12}C_j^{\frac12}\right]_{m_1m_2}Z_{m_2j}Z_{n_1k}\left[C_k^{\frac12}C_j^{\frac12}\right]_{n_1n_2}Z_{n_2j}\right]$$

$$=\sum_{i\neq k}\sum_{j\neq k}\mathbb{E}Z_{11}^4\frac{1}{p}x_i^TC_kx_ix_j^TC_kx_jA_{ij}+\sum_{i\neq k}\sum_{j\neq k}\sum_{m\neq l}\frac{1}{p}\left(\mathbb{E}Z_{11}^2\right)^2A_{ij}\left|\left[C_k^{\frac12}x_i\right]_l\right|^2\left|\left[C_k^{\frac12}x_j\right]_m\right|^2$$

$$+\sum_{i\neq k}\sum_{j\neq k}\sum_{\ell\neq n}\frac{1}{p}\left(\mathbb{E}Z_{11}^2\right)^2\left[C_k^{\frac12}x_i\right]_l\left[C_k^{\frac12}x_i\right]_nA_{ij}\left[C_k^{\frac12}x_j\right]_l\left[C_k^{\frac12}x_j\right]_n$$

$$+\sum_{i\neq k}\sum_{j\neq k}\sum_{l\neq s}\frac{1}{p}\left(\mathbb{E}Z_{11}^2\right)^2\left[C_k^{\frac12}x_i\right]_l\left[C_k^{\frac12}x_i\right]_sA_{ij}\left[C_k^{\frac12}x_j\right]_l\left[C_k^{\frac12}x_j\right]_s$$

$$=\sum_{i\neq k}\sum_{j\neq k}(\mathbb{E}Z_{11}^4-3(\mathbb{E}Z_{11}^2)^2)\frac{1}{p}x_i^TC_kx_ix_j^TC_kx_jA_{ij}+\sum_{i\neq k}\sum_{j\neq k}\frac{1}{p}\left(\mathbb{E}Z_{11}^2\right)^2x_i^TC_kx_ix_j^TC_kx_j$$

$$+\sum_{i\neq k}\sum_{j\neq k}\frac{2}{p}\left(\mathbb{E}Z_{11}^2\right)^2A_{ij}(x_i^TC_kx_j)^2$$

Since Z_{11} is a standard Gaussian random variable, its fourth cumulant is zero. As such, the above expression can be further simplified to:

$$\begin{aligned}\mathbb{E}_{x_k}\left[p\sum_{i\neq k}\sum_{j\neq k}(x_i^Tx_k)^2A_{ij}(x_j^Tx_k)^2\right]&=\sum_{i\neq k}\sum_{j\neq k}\frac{1}{p}\left(\mathbb{E}Z_{11}^2\right)^2A_{ij}x_i^TC_kx_ix_j^TC_kx_j\\ &+\sum_{i\neq k}\sum_{j\neq k}\frac{2}{p}\left(\mathbb{E}Z_{11}^2\right)^2A_{ij}(x_i^TC_kx_j)^2\end{aligned}\tag{38}$$

Plugging (38) into (37) leads to the desired result.

We will now prove (9). For X a scalar random variable, we define the notation $\mathrm{var}_{x_k}(X)=\mathbb{E}_{x_k}\left|X-\mathbb{E}_{x_k}X\right|^2$. We have thus:

$$\begin{aligned}&\mathbb{E}\left|\xi_{(k,-k)}^TA\xi_{(k,-k)}-\mathbb{E}_{x_k}\xi_{(k,-k)}^TA\xi_{(k,-k)}\right|^2=\mathbb{E}\left[\mathrm{var}_{x_k}\left(\xi_{(k,-k)}^TA\xi_{(k,-k)}\right)\right]\\ &\leq\sum_{j=1}^n\mathbb{E}\left[\left|\frac{\partial\xi_{(k,-k)}^TA\xi_{(k,-k)}}{\partial Z_{jk}}\right|^2\right]\end{aligned}$$

Using Lemma 2, it follows that to prove (9) for $s=1$, it suffices to show that:

$$\sum_{j=1}^n\left|\frac{\partial\xi_{(k,-k)}^TA\xi_{k,-k}}{\partial Z_{jk}}\right|^2\leq\|A\|^2O(p^{-1})\tag{39}$$

Given that:

$$\begin{aligned}\frac{\partial \xi_{(k,-k)}^T A \xi_{k,-k}}{\partial Z_{jk}} &= \sum_{a\neq k}\sum_{b\neq k} \frac{\partial p(x_k^T x_a)^2 A_{ab}(x_k^T x_b)^2}{\partial Z_{jk}} - \sum_{a\neq k}\sum_{b\neq k}\frac{\partial \frac{1}{p}\operatorname{tr} C_k C_a A_{ab}(x_k^T x_b)^2}{\partial Z_{jk}} \\ &\quad - \sum_{a\neq k}\sum_{b\neq k}\frac{\partial (x_k^T x_a)^2 A_{ab}\frac{1}{p}\operatorname{tr} C_k C_b}{\partial Z_{jk}},\end{aligned}$$

one can easily see that it suffices to establish that $\max(\alpha_1,\alpha_2) = O(p^{-1})$ where

$$\begin{aligned}\alpha_1 &= \sum_{j=1}^n \left| \sum_{a\neq k}\sum_{b\neq k} \frac{\partial p(x_k^T x_a)^2 A_{ab}(x_k^T x_b)^2}{\partial Z_{jk}}\right|^2 \\ \alpha_2 &= \sum_{j=1}^n \left| \sum_{a\neq k}\sum_{b\neq k} \frac{\partial (x_k^T x_a)^2 A_{ab}\frac{1}{p}\operatorname{tr} C_k C_b}{\partial Z_{jk}}\right|^2\end{aligned}$$

We begin with α_1. We have:

$$\begin{aligned}\alpha_1 &\leq p^2 \sum_{j=1}^n \left[\left| \sum_{a\neq k}\sum_{b\neq k} \frac{\partial (x_k^T x_a)^2 A_{ab}(x_k^T x_b)^2}{\partial Z_{jk}}\right|^2\right] \\ &\leq 4p \sum_{j=1}^n \left[\left| \sum_{a\neq k}\sum_{b\neq k} x_k^T x_a \left[C_k^{\frac{1}{2}} x_a\right]_j A_{ab}(x_k^T x_b)^2\right|^2\right] \\ &= 4p \sum_{a_1\neq k}\sum_{b_1\neq k}\sum_{a_2\neq k}\sum_{b_2\neq k} \mathbb{E}\left[(x_k^T x_{a_1})(x_k^T x_{a_2}) x_{a_2}^T C_k x_{a_1} A_{a_1 b_1} A_{a_2 b_2}^* (x_k^T x_{b_1})^2 \left(x_k^T x_{b_2}\right)^2\right] \\ &= 4p 1^T D_k^2 A^H S_k A D_k^2 1\end{aligned}$$

where $D_k = \operatorname{diag}\left\{x_b^T x_k \boldsymbol{\delta}_{b\neq k}\right\}_{b=1}^n$ and $[S_k]_{a_1,a_2} = x_k^T x_{a_1} x_{a_2}^T C_k x_{a_1} x_k^T x_{a_2}$. The result follows using the fact that $\|D_k\| = O(p^{-1/2})$ and $\|S_k\| = O(p^{-1})$. We will now handle α_2.

$$\begin{aligned}\alpha_2 &\leq 4\sum_{j=1}^n \left| \sum_{a\neq k}\sum_{b\neq k} (x_k^T x_a)\frac{1}{\sqrt{p}}\left[C_k^{\frac{1}{2}} x_a\right]_j A_{ab}\frac{1}{p}\operatorname{tr} C_k C_b\right|^2 \\ &= \frac{4}{p}\sum_{a_1\neq k}\sum_{b_1\neq k}\sum_{a_2\neq k}\sum_{b_2\neq k}\left[x_k^T x_{a_1} x_k^T x_{a_2} x_{a_2}^T C_k x_{a_1}\frac{1}{p}\operatorname{tr} C_k C_{b_1}\frac{1}{p}\operatorname{tr} C_k C_{b_2} A_{a_1 b_1} A_{a_2 b_2}^*\right] \\ &= \frac{4}{p} 1^T \operatorname{diag}\left\{\frac{1}{p}\operatorname{tr} C_k C_{b_2}\right\}_{b_2=1}^n A^H S_k A \operatorname{diag}\left\{\frac{1}{p}\operatorname{tr} C_k C_{b_1}\right\}_{b_1=1}^n 1 \\ &= \|A\|^2 O(p^{-1})\end{aligned}$$

For $s\in\mathbb{N}^*$, the result can be extended by induction and using the fact that:

$$\begin{aligned}\mathbb{E}\left|\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T A\xi_{(k,-k)}\right|^{2s} &= \left(\mathbb{E}\left|\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T A\xi_{(k,-k)}\right|^s\right)^2 \\ &\quad + \operatorname{var}\left[\left(\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T A\xi_{(k,-k)}\right)^s\right]\end{aligned}$$

The first term of the above equation can be handled using the induction assumption along with the Cauchy-Schwartz inequality, leading to:

$$\begin{aligned}
&\left(\mathbb{E}\left|\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T A\xi_{(k,-k)}\right|\right)^{2s} \\
&= \mathbb{E}\left|\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T A\xi_{(k,-k)}\right|^{2(s-1)} \mathbb{E}\left|\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T A\xi_{(k,-k)}\right|^2 \\
&= \|A\|^{2s} O(p^{-s+\epsilon})
\end{aligned}$$

The second term can be treated using the Poincaré-Nash inequality as follows:

$$\begin{aligned}
&\operatorname{var}\left[\left(\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T A\xi_{(k,-k)}\right)^s\right] \\
&\leq \mathbb{E}\left[s^2\left|\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T A\xi_{(k,-k)}\right|^{2(s-1)} \sum_{j=1}^n \left|\frac{\partial \xi_{(k,-k)}^T A\xi_{(k,-k)}}{\partial Z_{j,k}}\right|^2\right]
\end{aligned}$$

Since $\sum_{j=1}^n \left|\frac{\partial \xi_{(k,-k)}^T A\xi_{(k,-k)}}{\partial Z_{j,k}}\right|^2 = \|A\|^2 O(p^{-1})$ and $\mathbb{E}\left[\left|\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T A\xi_{(k,-k)}\right|^{2(s-1)}\right] = \|A\|^{2(s-1)} O(p^{-s-1+\frac{\epsilon}{2}})$ by the induction assumption, we have by Lemma 2,

$$\operatorname{var}\left[\left(\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T A\xi_{(k,-k)}\right)^s\right] = \|A\|^{2s} O(p^{-s+\epsilon})$$

APPENDIX B: PROOF OF LEMMA 4

We will prove only the last item. The two first items can be found in [Capitaine et al., 2009]. Since $\omega^2 c_0 m^2(z) = -1 - zm(z)$, we have:

$$\begin{aligned}
1 - \alpha m^2(z) &= 1 + \frac{\alpha}{\omega^2 c_0} + \frac{\alpha z m(z)}{\omega^2 c_0} \\
&= \left(1 + \frac{\alpha}{\omega^2 c_0}\right) - \frac{\alpha}{\omega^2 c_0}|z|^2 \int \frac{1}{|\lambda - z|^2}\mu(d\lambda) + \frac{\alpha z}{\omega^2 c_0}\int \frac{1}{|\lambda - z|^2}\mu(d\lambda)
\end{aligned}$$

Since $\left|1 - \alpha m^2(z)\right| \geq \min\left(\Re 1 - \alpha m^2(z), \Im 1 - \alpha m^2(z)\right)$, it suffices to study the real and imaginary parts of $1 - \alpha m^2(z)$. Let $z = x + \imath y$. Then, due to the symmetry of μ we have:

$$\begin{aligned}
\Im\left(1 - \alpha m^2(z)\right) &= \frac{\alpha}{\omega^2 c_0}|y| \left|\int_0^{2\sqrt{c_0}\omega} \frac{\lambda}{(\lambda - x)^2 + y^2}\mu(d\lambda) - \int_0^{2\sqrt{c_0}\omega} \frac{\lambda}{(\lambda + x)^2 + y^2}\right| \\
&= \frac{\alpha|yx|}{\omega^2 c_0}\int_0^{2\sqrt{c_0}\omega} \frac{4\lambda^2}{\left((\lambda - x)^2 + y^2\right)\left((\lambda + x)^2 + y^2\right)} \qquad (40)
\end{aligned}$$

Now, we will turn to studying the real part. We have:

$$\left|\Re\left(1-\alpha m^2(z)\right)\right| = 1+\frac{\alpha}{\omega^2 c_0}+\frac{\alpha}{\omega^2 c_0}\int\frac{\lambda x}{(\lambda-x)^2+y^2}\mu(d\lambda)-\frac{\alpha}{\omega^2 c_0}(x^2+y^2)\int\frac{1}{(\lambda-x)^2+y^2}\mu(d\lambda)$$

$$= 1+\frac{\alpha}{\omega^2 c_0}+\frac{\alpha}{\omega^2 c_0}\int_0^{2\sqrt{c_0}\omega}\frac{4\lambda^2 x^2}{\left((\lambda-x)^2+y^2\right)\left((\lambda+x)^2+y^2\right)}\mu(d\lambda)$$

$$-\frac{\alpha}{\omega^2 c_0}(x^2+y^2)\int_0^{2\sqrt{c_0}\omega}\frac{2\lambda^2+2x^2+2y^2}{\left((\lambda-x)^2+y^2\right)\left((\lambda+x)^2+y^2\right)}\mu(d\lambda)$$

$$=\left|\int_0^{2\sqrt{c_0}\omega}\frac{(2+\frac{2\alpha}{\omega^2 c_0})\lambda^4-\lambda^2x^2\left(6+\frac{2\alpha}{\omega^2 c_0}\right)+\lambda^2y^2\left(2+\frac{2\alpha}{\omega^2 c_0}\right)+2(x^2+y^2)\left(\lambda^2+x^2+y^2\right)}{\left((\lambda-x)^2+y^2\right)\left((\lambda+x)^2+y^2\right)}\mu(d\lambda)\right|$$

$$=\left|\int_0^{2\sqrt{c_0}\omega}\frac{(2+\frac{2\alpha}{\omega^2 c_0})\lambda^4-\lambda^2x^2(4+\frac{2\alpha}{\omega^2 c_0})+\lambda^2y^2(4+\frac{2\alpha}{\omega^2 c_0})+2x^2(x^2+y^2)+2y^2(x^2+y^2)}{\left((\lambda-x)^2+y^2\right)\left((\lambda+x)^2+y^2\right)}\mu(d\lambda)\right| \tag{41}$$

To prove (10), we shall distinguish two different cases. If $|x|\geq|y|$, then from (40), we have:

$$\left|1-\alpha m^2(z)\right|\geq\frac{\alpha y^2}{\omega^2 c_0}\int_0^{2\sqrt{c_0}\omega}\frac{4\lambda^2}{\left((\lambda-x)^2+y^2\right)\left((\lambda+x)^2+y^2\right)}\mu(d\lambda)$$

Let $\eta>2\sqrt{c_0}\omega$. Then,

$$\max\left((\lambda-x)^2+y^2,(\lambda+x)^2+y^2\right)\leq 2\lambda^2+2|z|^2\leq 2(|z|^2+2\eta^2)$$

Hence,

$$\left|1-\alpha m^2(z)\right|\geq y^2\frac{\frac{1}{2}c_0\omega^2}{(|z|^2+\eta^2)^2}$$

If $|x|\leq|y|$, then from (41), we have:

$$\left|1-\alpha m^2(z)\right|\geq\frac{y^4}{(|z|+\eta)^4}$$

To prove (11), we first note that if $|z|\geq 2\sqrt{2}\sqrt{c_0}\omega\sqrt{4+\frac{2\alpha}{\omega^2 c_0}}$, then necessarily $\max(|x|,|y|)\geq 2\sqrt{c_0}\omega\sqrt{4+\frac{2\alpha}{\omega^2 c_0}}$. Assume that $|y|=\max(|x|,|y|)$. Then, from (41), we have:

$$|1-\alpha m^2(z)|\geq\frac{|z|^4}{4(|z|+\eta)^4}\geq\frac{|z|^4}{8(|z|+\eta)^4}$$

Now, assume that $|x|=\max(|x|,|y|)$. Hence $|x|\geq 2\sqrt{c_0}\omega\sqrt{4+\frac{2\alpha}{\omega^2 c_0}}$. Function $\lambda\mapsto(2+\frac{2\alpha}{\omega^2 c_0})\lambda^4-\lambda^2x^2(4+\frac{2\alpha}{\omega^2 c_0})$ on $\left(0,2\sqrt{c_0}\omega\right)$ is decreasing and achieves its minimum at $2\sqrt{c_0}\omega$. We thus have:

$$\left|\Re\left(1-\alpha m^2(z)\right)\right|\geq\left|\int_0^{2\sqrt{c_0}\omega}\frac{-4c_0\omega^2x^2(4+\frac{2\alpha}{\omega^2 c_0})+x^4+(x^2+y^2)^2}{\left((\lambda-x)^2+y^2\right)\left((\lambda+x)^2+y^2\right)}\mu(d\lambda)\right| \tag{42}$$

$$\geq\frac{|z|^4}{8(|z|+\eta)^4} \tag{43}$$

APPENDIX C: PROOF OF LEMMA 6

We have:

$$
\begin{aligned}
W_3 &= \mathbb{E}\left|\sum_{k\neq j}\frac{1}{\sqrt{p}}\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_k A_j\right) Q_{jk}\right|^2 \\
&\leq 2\mathbb{E}\left|\sum_{k\neq j}\frac{1}{\sqrt{p}}\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_k A_j\right) Q_{jk} - \mathbb{E}_j\left(\sum_{k\neq j}\frac{1}{\sqrt{p}}\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_k A_j\right) Q_{jk}\right)\right|^2 \\
&+ 2\mathbb{E}\left|\mathbb{E}_j\sum_{k\neq j}\frac{1}{\sqrt{p}}\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_k A_j\right) Q_{jk}\right|^2 \\
&= 2W_{31} + 2W_{32}.
\end{aligned}
$$

We have:

$$
\begin{aligned}
W_{31} &\leq \sum_{l=1}^n \mathbb{E}\left|\sum_{k\neq j}\frac{1}{\sqrt{p}}\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_k A_j\right)\frac{\partial Q_{kj}}{\partial Z_{lj}}\right|^2 \\
&\leq 8\sum_{l=1}^n \mathbb{E}\left|\sum_{k\neq j}\frac{1}{\sqrt{p}}\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_k A_j\right)\sum_{b\neq j}(x_b^T x_j)\left[C_j^{\frac{1}{2}} x_b\right]_l Q_{kj} Q_{bj}\right|^2 \\
&+ 8\sum_{l=1}^n \mathbb{E}\left|\sum_{k\neq j}\frac{1}{\sqrt{p}}\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_k A_j\right)\sum_{b\neq j}(x_b^T x_j)\left[C_j^{\frac{1}{2}} x_b\right]_l Q_{jj} Q_{bk}\right|^2 \\
&= 8W_{311} + 8W_{312}
\end{aligned}
$$

We have:

$$
\begin{aligned}
W_{311} &= \sum_{k_1\neq j}\sum_{k_2\neq j}\mathbb{E}\left[\frac{1}{p}\left(x_{k_1}^T A_j x_{k_1} - \frac{1}{p}\operatorname{tr} C_{k_1} A_j\right)\left(x_{k_2}^T A_j x_{k_2} - \frac{1}{p}\operatorname{tr} C_{k_2} A_j\right)\right. \\
&\left.\times\sum_{b_1\neq j}\sum_{b_2\neq j}(x_{b_1}^T x_j) x_{b_1}^T C_j x_{b_2}\left(x_{b_2}^T x_j\right) Q_{k_1 j} Q_{k_2 j}^* Q_{b_1 j} Q_{b_2 j}^*\right] \\
&= \mathbb{E}\left[\left|\left[\frac{1}{\sqrt{p}}^T \operatorname{diag}\left\{\left(x_{k_1}^T A_j x_{k_1} - \frac{1}{p}\operatorname{tr} C_{k_1} A_j\right)\boldsymbol{\delta}_{j\neq k}\right\}_{k_1=1}^n Q\right]_j\right|^2 \left[Q^H S_j Q\right]_{jj}\right] \\
&= O_z(p^{-2+\epsilon})
\end{aligned}
$$

where $[S_j]_{b_1b_2} = \boldsymbol{\delta}_{b_1\neq j}\boldsymbol{\delta}_{b_2\neq j}x_{b_1}^T x_j x_{b_2}^T x_j x_{b_1}^T C_j x_{b_2}$.

$$W_{312} = \sum_{k_1\neq j}\sum_{k_2\neq j}\mathbb{E}\left[\frac{1}{p}\left(x_{k_1}^T A_j x_{k_1} - \frac{1}{p}\operatorname{tr} C_{k_1}A_j\right)\left(x_{k_2}^T A_j x_{k_2} - \frac{1}{p}\operatorname{tr} C_{k_2}A_j\right)\right.$$
$$\left.\times|Q_{jj}|^2\sum_{b_1\neq j}\sum_{b_2\neq j}x_{b_1}^T C_j x_{b_2}x_{b_1}^T x_j x_{b_2}^T x_j Q_{b_1k_1}Q_{b_2k_2}^*\right]$$
$$= \mathbb{E}|Q_{jj}|^2\frac{1}{p}1^T\operatorname{diag}\left\{\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_kA_j\right)\boldsymbol{\delta}_{k\neq j}\right\}_{k=1}^n QS_jQ^H\operatorname{diag}\left\{\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_kA_j\right)\boldsymbol{\delta}_{k\neq j}\right\}_{k=1}^n 1$$
$$= O_z(p^{-2+\epsilon})$$

All this proves that:

$$\max_{1\leq j\leq n} W_{32} = O_z(p^{-2+\epsilon}).$$

Now, we move to W_{32}. We have:

$$W_{32} \leq 2\mathbb{E}\left|\mathbb{E}_j\sum_{k\neq j}\frac{1}{\sqrt{p}}\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_kA_j\right)e_k^T Q_j\xi_{(j,-j)}(Q_{jj}-\mathbb{E}Q_{jj})\right|^2$$
$$+ 2\left(\mathbb{E}Q_{jj}\right)^2\mathbb{E}\left|\sum_{k\neq j}\mathbb{E}_j\frac{1}{\sqrt{p}}\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_kA_j\right)e_k^T Q_j\xi_{(j,-j)}\right|^2$$

The first term is obviously $O_z(p^{-2+\epsilon})$ while the second term can be worked out as:

$$\mathbb{E}\left|\sum_{k\neq j}\mathbb{E}_j\frac{1}{\sqrt{p}}\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_kA_j\right)e_k^T Q_j\xi_{(j,-j)}\right|^2$$
$$= \mathbb{E}\left|\frac{1}{p}\sum_{k\neq j}\sum_{l\neq j}\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_kA_j\right)[Q_j]_{kl}\left(x_l^T A_j x_l - \frac{1}{p}\operatorname{tr} A_jC_l\right)\right|^2$$
$$= \mathbb{E}\left|\frac{1}{p}1^T\operatorname{diag}\left\{\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_kA_j\right)\boldsymbol{\delta}_{k\neq j}\right\}_{k=1^n} Q_j\operatorname{diag}\left\{\left(x_k^T A_j x_k - \frac{1}{p}\operatorname{tr} C_kA_j\right)\boldsymbol{\delta}_{k\neq j}\right\}_{k=1}^n\right|^2$$
$$= O_z(p^{-2+\epsilon})$$

APPENDIX D: PROOF OF LEMMA 8

We will treat the case of $k = j$. All other cases follow similarly. Let W_2 be given by:

$$W_2 = \mathbb{E}\left|\sum_{s\notin\{j,b\}}x_s^T A_1 x_j x_b^T A_2 x_s Q_{sj}\right|^2$$
$$\leq 2\mathbb{E}\left|\sum_{s\notin\{j,b\}}x_s^T A_1 x_j x_b^T A_2 x_s Q_{sj} - \mathbb{E}_j\sum_{s\notin\{j,b\}}x_s^T A_1 x_j x_b^T A_2 x_s Q_{sj}\right|^2 + 2\mathbb{E}\left|\mathbb{E}_j\sum_{s\notin\{j,b\}}x_s^T A_1 x_j x_b^T A_2 x_s Q_{sj}\right|^2$$
$$= 2W_{21} + 2W_{22}$$

By Poincaré-Nash inequality, we have:

$$
\begin{aligned}
W_{21} &\leq 2\sum_{l=1}^{n}\left|\sum_{s\notin\{j,b\}}\frac{1}{\sqrt{p}}\left[A_1C_j^{\frac{1}{2}}x_s\right]_l x_b^TA_2x_sQ_{sj}\right|^2 \\
&+8\sum_{l=1}^{n}\left|\sum_{s\notin\{j,b\}}x_s^TA_1x_jx_b^TA_2x_s\sum_{q\neq j}(x_q^Tx_j)\left[C_j^{\frac{1}{2}}x_q\right]_l(Q_{sj}Q_{qj}+Q_{jj}Q_{qs})\right|^2 \\
&=2Z_1+8Z_2
\end{aligned}
$$

We have:

$$
\begin{aligned}
Z_1 &= \sum_{s_1\notin\{j,b\}}\sum_{s_2\notin\{j,b\}}\frac{1}{p}\mathbb{E}\left[x_{s_1}^TA_1C_jA_1x_{s_2}x_b^TA_2x_{s_1}x_b^TA_2x_{s_2}Q_{s_1j}Q_{s_2j}\right] \\
&=\frac{1}{p}\mathbb{E}\left[Q\left\{x_{s_1}^TA_1C_jA_1x_{s_2}x_b^TA_2x_{s_1}x_b^TA_2x_{s_2}\boldsymbol{\delta}_{s_1\notin\{b,j\}}\boldsymbol{\delta}_{s_2\notin\{b,j\}}\right\}_{s_1,s_2=1}^{n}Q\right]_{jj}
\end{aligned}
$$

Using lemma 1, we thus get:

$$
\max_{j\neq k} Z_1 = O_z(p^{-2+\epsilon})
$$

We have:

$$
\begin{aligned}
Z_2 &\leq 2\sum_{s_1\notin\{j,b\}}\sum_{s_2\notin\{j,b\}}\sum_{q_1\neq j}\sum_{q_2\neq j}\mathbb{E}\left[x_{s_1}^TA_1x_jx_b^TA_2x_{s_1}x_{s_2}^TA_1x_jx_b^TA_2x_{s_2}x_{q_1}^Tx_jx_{q_2}^Tx_jx_{q_1}^TC_jx_{q_2}Q_{s_1j}Q_{s_2j}^*Q_{q_2j}^*Q_{q_1j}\right] \\
&+2\sum_{s_1\notin\{b,j\}}\sum_{s_2\notin\{b,j\}}\sum_{q_1\neq j}\sum_{q_2\neq j}\mathbb{E}\left[x_{s_1}^TA_1x_jx_b^TA_2x_{s_1}x_{s_2}^TA_1x_jx_b^TA_2x_{s_2}x_{q_1}^Tx_jx_{q_2}^Tx_jx_{q_1}^TC_jx_{q_2}|Q_{jj}|^2Q_{q_1s_1}Q_{q_2s_2}^*\right] \\
&=2\mathbb{E}\left[|Q_{jj}|^2\left[X^TA_2XD_jQS_jQ^HD_jX^TA_2X\right]_{bb}\right]+2\mathbb{E}\left[\left[X^TA_2XD_jQ\right]_{bj}\left[Q^HD_jX^TA_2X\right]_{jb}\left[Q^HS_jQ\right]_{jj}\right] \\
&=O_z(p^{-2+\epsilon})
\end{aligned}
$$

where $D_j=\operatorname{diag}\left\{x_s^TA_1x_j\boldsymbol{\delta}_{s\notin\{b,j\}}\right\}$ and $S_j=\left\{x_{q_1}^Tx_jx_{q_2}^Tx_jx_{q_1}^TC_jx_{q_2}\boldsymbol{\delta}_{q_1\neq j}\boldsymbol{\delta}_{q_2\neq j}\right\}_{q_1,q_2=1}^{n}$. We now deal with W_{22}. We have:

$$
\begin{aligned}
W_{22} &= \mathbb{E}\left|\sum_{s\notin\{j,b\}}\mathbb{E}_jx_s^TA_1x_jx_b^TA_2x_se_s^TQ_j\xi_{(j,-j)}(Q_{jj}-\mathbb{E}Q_{jj})\right|^2 \\
&=O_z(p^{-2+\epsilon})
\end{aligned}
$$

APPENDIX E: PROOF OF PROPOSITION 1

$$
\begin{aligned}
\sum_{k=1}^{n}\sum_{r=1,r\neq k}^{n} a_k b_r \mathbb{E} Q_{kr} &= \sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}\left[-e_r^T Q_k \xi_{(k,-k)} Q_{kk}\right] \\
&= \sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}\left[e_r^T Q_k \xi_{(k,-k)}\left(z+\mathbb{E}_{x_k}\xi_{(k,-k)}^T Q_k \xi_{(k-k)}\right)^{-1}\right] \\
&+ \sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}\left[-\frac{e_r^T Q_k \xi_{(k,-k)}\left(\xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T Q_k \xi_k\right)}{(z+\xi_{(k,-k)}^T Q_k \xi_{(k,-k)})(z+\mathbb{E}_{x_k}\xi_{(k,-k)}^T Q_k \xi_{(k,-k))}}\right] \\
&= \sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}\left[e_r^T Q_k \xi_{(k,-k)}\left(z+\mathbb{E}_{x_k}\xi_{(k,-k)}^T Q_k \xi_{(k-k)}\right)^{-1}\right] \\
&+ \sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}\left[-\frac{e_r^T Q_k \xi_{(k,-k)}\left(\xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T Q_k \xi_{(k,-k)}\right)}{(z+\mathbb{E}_{x_k}\xi_{(k,-k)}^T Q_k \xi_{(k,-k)})^2}\right] \\
&+ \sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}\left[\frac{e_r^T Q_k \xi_{(k,-k)}\left(\xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T Q_k \xi_{(k,-k)}\right)^2}{(z+\mathbb{E}_{x_k}\xi_{(k,-k)}^T Q_k \xi_{(k,-k)})^2(z+\xi_{(k,-k)}^T Q_k \xi_{(k,-k)})}\right] \\
&= W_1 + W_2 + W_3
\end{aligned}
$$

We start by handling W_3. We have:

$$
\begin{aligned}
|W_3| &\leq \sum_{k=1}^{n}\sum_{r\neq k} |a_k||b_r||\Im z|^{-3}\sqrt{\mathbb{E}\left|e_r^T Q_k \xi_{(k,-k)}\right|^2}\sqrt{\mathbb{E}\left(\xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T Q_k \xi_{(k,-k)}\right)^4} \\
&= O_z(p^{-\frac{1}{2}+\epsilon})
\end{aligned}
$$

As for W_2, we have:

$$
W_2 = \sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}\left[-\left(z+\mathbb{E}_{x_k}\xi_{(k,-k)}^T Q_k \xi_{(k,-k)}\right)^{-2} e_r^T Q_k \xi_{(k,-k)}\left(\xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T Q_k \xi_{(k,-k)}\right)\right]
$$

We need thus to compute:

$$
\beta_{kr} = \mathbb{E}_{x_k}\left[e_r^T Q_k \xi_{(k,-k)}\left(\xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_{x_k}\xi_{(k,-k)}^T Q_k \xi_{(k,-k)}\right)\right]
$$

We have:

$$
\begin{aligned}
\beta_{kr} &= \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}\sqrt{p}\,[Q_k]_{rl}\,\mathbb{E}_{x_k}\left[\left((x_k^Tx_l)^2-\frac{1}{p^2}\operatorname{tr}C_kC_l\right)p(x_k^Tx_m)^2\,[Q_k]_{mq}\,(x_k^Tx_q)^2\right]\\
&-2\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}\sqrt{p}\,[Q_k]_{rl}\,\mathbb{E}_{x_k}\left[\left((x_k^Tx_l)^2-\frac{1}{p^2}\operatorname{tr}C_kC_l\right)\frac{1}{p}\operatorname{tr}C_kC_m\,[Q_k]_{mq}\,(x_k^Tx_q)^2\right]\\
&+\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}\sqrt{p}\,[Q_k]_{rl}\,\mathbb{E}_{x_k}\left[\left((x_k^Tx_l)^2-\frac{1}{p^2}\operatorname{tr}C_kC_l\right)\frac{1}{p}\operatorname{tr}C_kC_m\frac{1}{p^2}\operatorname{tr}C_kC_q\right]\\
&-2\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}p^{-\frac{3}{2}}\,[Q_k]_{rl}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)(x_m^TC_kx_q)^2\,[Q_k]_{mn}\\
&-\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}p^{-\frac{3}{2}}\,[Q_k]_{rl}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)\left(x_m^TC_kx_m-\frac{1}{p}\operatorname{tr}C_kC_m\right)\left(x_q^TC_kx_q-\frac{1}{p}\operatorname{tr}C_kC_q\right)\\
&=\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}\sqrt{p}\,[Q_k]_{rl}\,[Q_k]_{mq}\,\mathbb{E}_{x_k}\left[\left((x_k^Tx_l)^2-\frac{1}{p}x_l^TC_kx_l\right)p\left((x_k^Tx_m)^2-\frac{1}{p}x_m^TC_kx_m\right)(x_k^Tx_q)^2\right]\\
&+\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}\sqrt{p}\,[Q_k]_{rl}\,[Q_k]_{mq}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)\mathbb{E}_{x_k}\left[\left((x_k^Tx_m)^2-\frac{1}{p}x_m^TC_kx_m\right)(x_k^Tx_q)^2\right]\\
&+\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}\sqrt{p}\,[Q_k]_{rl}\,[Q_k]_{mq}\left(x_m^TC_kx_m-\frac{1}{p}\operatorname{tr}C_kC_m\right)\mathbb{E}_{x_k}\left[\left((x_k^Tx_l)^2-\frac{1}{p}x_l^TC_kx_l\right)(x_k^Tx_q)^2\right]\\
&+\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}\sqrt{p}\,[Q_k]_{rl}\,[Q_k]_{mq}\left(x_m^TC_kx_m-\frac{1}{p}\operatorname{tr}C_kC_m\right)\mathbb{E}_{x_k}\left[\left(\frac{1}{p}x_l^TC_kx_l-\frac{1}{p^2}\operatorname{tr}C_kC_l\right)(x_k^Tx_q)^2\right]\\
&-\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}\sqrt{p}\,[Q_k]_{rl}\,\mathbb{E}_{x_k}\left[\left((x_k^Tx_l)^2-\frac{1}{p}x_l^TC_kx_l\right)\frac{1}{p}\operatorname{tr}C_kC_m\,[Q_k]_{mq}\,(x_k^Tx_q)^2\right]\\
&-\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}\sqrt{p}\,[Q_k]_{rl}\,\mathbb{E}_{x_k}\left[\left(\frac{1}{p}x_l^TC_kx_l-\frac{1}{p^2}\operatorname{tr}C_kC_l\right)\frac{1}{p}\operatorname{tr}C_kC_m\,[Q_k]_{mq}\,(x_k^Tx_q)^2\right]\\
&+\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}p^{-\frac{3}{2}}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)\frac{1}{p}\operatorname{tr}C_kC_m\,[Q_k]_{rl}\,[Q_k]_{mq}\,\frac{1}{p}\operatorname{tr}C_kC_q\\
&-2\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}p^{-\frac{3}{2}}\,[Q_k]_{rl}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)(x_m^TC_kx_q)^2\,[Q_k]_{mq}\\
&-\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}p^{-\frac{3}{2}}\,[Q_k]_{rl}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)\left(x_m^TC_kx_m-\frac{1}{p}\operatorname{tr}C_kC_m\right)\left(x_q^TC_kx_q-\frac{1}{p}\operatorname{tr}C_kC_q\right)
\end{aligned}
$$

Using the fact that $\mathbb{E}(z_1^T A_1 z_1 - \operatorname{tr} A_1)(z_1^T A_2 z_1 - \operatorname{tr} A_2)(z_1^T A_3 z_1 - \operatorname{tr} A_3) = 8 \operatorname{tr} A_1 A_2 A_3$ and $\mathbb{E}(z_1^T A_1 z_1 - \operatorname{tr} A_1)(z_1^T A_2 z_1 - \operatorname{tr} A_2) = 2 \operatorname{tr} A_1 A_2$, we obtain:

$$
\begin{aligned}
\beta_{kr} = 8 &\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} [Q_k]_{mq} x_m^T C_k x_q x_q^T C_k x_l x_l^T C_k x_m \\
&+ 2 \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} [Q_k]_{mq} x_q^T C_k x_q (x_l^T C_k x_m)^2 \\
&+ 2 \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} [Q_k]_{mq} \left(x_l^T C_k x_l - \frac{1}{p}\operatorname{tr} C_k C_l\right) (x_m^T C_k x_q)^2 \\
&+ 2 \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} [Q_k]_{mq} \left(x_m^T C_k x_m - \frac{1}{p}\operatorname{tr} C_k C_m\right) (x_l^T C_k x_q)^2 \\
&+ \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} [Q_k]_{mq} \left(x_m^T C_k x_m - \frac{1}{p}\operatorname{tr} C_k C_m\right) \left(x_l^T C_k x_l - \frac{1}{p}\operatorname{tr} C_k C_l\right) x_q^T C_k x_q \\
&- 2 \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} \frac{1}{p}\operatorname{tr} C_k C_m [Q_k]_{mq} (x_q^T C_k x_l)^2 \\
&- \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} [Q_k]_{mq} \frac{1}{p}\operatorname{tr} C_k C_m \left(x_l^T C_k x_l - \frac{1}{p}\operatorname{tr} C_k C_l\right) \left(x_q^T C_k x_q - \frac{1}{p}\operatorname{tr} C_k C_q\right) \\
&- 2 \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} \left(x_l^T C_k x_l - \frac{1}{p}\operatorname{tr} C_k C_l\right) (x_m^T C_k x_q)^2 [Q_k]_{mq} \\
&- \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} \left(x_l^T C_k x_l - \frac{1}{p}\operatorname{tr} C_k C_l\right) \left(x_m^T C_k x_m - \frac{1}{p}\operatorname{tr} C_k C_m\right) \left(x_q^T C_k x_q - \frac{1}{p}\operatorname{tr} C_k C_q\right) \\
&= 8 \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} [Q_k]_{mq} x_m^T C_k x_q x_q^T C_k x_l x_l^T C_k x_m \\
&+ 4 \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} [Q_k]_{mq} \left(x_q^T C_k x_q - \frac{1}{p}\operatorname{tr} C_k C_q\right) (x_l^T C_k x_m)^2 \\
&= 8\beta_{kr1} + 4\beta_{kr2}
\end{aligned}
$$

We need to show that for $j = 1, \cdots, 2$,

$$
\mathbb{E}\left|\sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \beta_{krj}\right| = O_z\left(p^{-\frac{1}{2}+\epsilon}\right)
$$

We have

$$
\begin{aligned}
&\mathbb{E}\left|\sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \beta_{kr1}\right| = \mathbb{E}\left|\sum_{k=1}^{n}\sum_{r\neq k}\sum_{l\neq k} p^{-3/2} a_k b_r [Q_k]_{rl} \left[X^T C_k X \left(Q_k \circ X^T C_k X\right) X^T C_k X\right]_{ll}\right| \\
&\leq \sum_{k=1}^{n} p^{-\frac{3}{2}} a_k \mathbb{E}\left| b^T \operatorname{diag}\left\{\alpha_{kr}\boldsymbol{\delta}_{r\neq k}\right\}_{r=1}^{n} \left[Q_k \operatorname{diag}\left(X^T C_k X \left(Q_k \circ \left\{x_m^T C_k x_q \boldsymbol{\delta}_{m\neq k}\boldsymbol{\delta}_{q\neq k}\right\}_{m,q=1}^{n}\right) X^T C_k X\right)\right] 1\right| \\
&= O_z\left(p^{-\frac{1}{2}}\right)
\end{aligned}
$$

We have:

$$\mathbb{E}\left|\sum_{k=1}^n\sum_{r\neq k}a_kb_r\beta_{kr2}\right| = \mathbb{E}\left|\sum_{k=1}^n\sum_{r\neq k}a_kb_r\sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k}p^{-\frac{3}{2}}\left[Q_k\right]_{rl}\left[Q_k\right]_{mq}\left(x_q^TC_kx_q-\frac{1}{p}\operatorname{tr}C_kC_q\right)\left(x_l^TC_kx_m\right)^2\right|$$

$$= \mathbb{E}\sum_{k=1}^n p^{-\frac{3}{2}}a_k1^T\left\{\left[Q_k\right]_{rl}\boldsymbol{\delta}_{r\neq k}\right\}_{r,l=1}^n\left\{\left(x_m^TC_kx_l\right)^2\boldsymbol{\delta}_{m\neq k}\boldsymbol{\delta}_{l\neq k}\right\}_{m,l=1}^n$$

$$\times Q_k\mathrm{diag}\left\{\left(x_q^TC_kx_q-\frac{1}{p}\operatorname{tr}C_kC_q\right)\boldsymbol{\delta}_{q\neq k}\right\}_{q=1}^n b$$

$$= O_z\left(p^{-1+\epsilon}\right)$$

We now deal with W_1. We have:

$$W_1 = \sum_{k=1}^n\sum_{r\neq k}a_kb_r\sum_{l\neq k}\mathbb{E}\left[\left[Q_k\right]_{rl}\frac{1}{\sqrt{p}}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)\left(z+\mathbb{E}_{x_k}\xi_{(k,-k)}Q_k\xi_{(k,-k)}\right)^{-1}\right]$$

$$= \sum_{k=1}^n\sum_{r\neq k}a_kb_r\sum_{l\neq k}\mathbb{E}\left[\left[Q_k\right]_{rl}\frac{1}{\sqrt{p}}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)\left(z+\xi_{(k,-k)}Q_k\xi_{(k,-k)}\right)^{-1}\right]$$

$$+ \sum_{k=1}^n\sum_{r\neq k}a_kb_r\sum_{l\neq k}\mathbb{E}\left[\left[Q_k\right]_{rl}\frac{1}{\sqrt{p}}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)\left(z+\xi_{(k,-k)}Q_k\xi_{(k,-k)}\right)^{-1}\left(z+\xi_{(k,-k)}^TQ_k\xi_{(k,-k)}\right)^{-1}\right.$$

$$\left.\times\left(\xi_{k,-k}^TQ_k\xi_{(k,-k)}-\mathbb{E}_{x_k}\xi_{(k,-k)}^TQ_k\xi_{(k,-k)}\right)\right]$$

$$= -\sum_{k=1}^n\sum_{r\neq k}a_kb_r\sum_{l\neq k}\mathbb{E}\left[\left[Q_k\right]_{rl}\frac{1}{\sqrt{p}}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)Q_{kk}\right]+O_z(p^{-\frac{1}{2}+\epsilon})$$

$$= -\sum_{k=1}^n\sum_{r\neq k}a_kb_r\sum_{l\neq k}\mathbb{E}\left[\left[Q_k\right]_{rl}\frac{1}{\sqrt{p}}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)\right]\mathbb{E}Q_{kk}+O_z(p^{-\frac{1}{2}+\epsilon})$$

To end up the proof, it suffices to check that:

$$\sum_{k=1}^n\sum_{r\neq k}a_kb_r\sum_{l\neq k}\mathbb{E}\left[\left[Q_k\right]_{rl}\frac{1}{\sqrt{p}}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)\right]\mathbb{E}Q_{kk}$$
$$= \sum_{k=1}^n\sum_{r\neq k}a_kb_r\sum_{l\neq k}\mathbb{E}\left[\left[Q\right]_{rl}\frac{1}{\sqrt{p}}\left(x_l^TC_kx_l-\frac{1}{p}\operatorname{tr}C_kC_l\right)\right]\mathbb{E}Q_{kk}+O_z(p^{-\frac{1}{2}+\epsilon})$$

This is a direct consequence of the following Lemma:

LEMMA 14. *Let $k\in\{1,\cdots,n\}$. Let $\{a_l\}_{l=1}^n$ and $\{b_l\}_{l=1}^n$ be real sequences independent of x_k such that: $\sum_{l=1}^n|a_l|^2=O(1)$ and $\sum_{l=1}^n|b_l|^2=O(1)$. Then, for $z\in\mathbb{C}\backslash\mathbb{R}$ and any small ϵ:*

$$\sum_{l=1}^n\sum_{r\neq k}\left(\mathbb{E}\left[Q_{lr}a_lb_r\right]-\mathbb{E}\left[\left[Q_k\right]_{lr}a_lb_r\right]\right)=O_z(p^{-\frac{1}{2}+\epsilon}).$$

PROOF. Let $\overline{\Phi}_k$ be $n \times n$ matrix given by:

$$\left[\overline{\Phi}_k\right]_{sl} = \begin{cases} \sqrt{p}\left((x_s^T x_l)^2 - \frac{1}{p^2}\operatorname{tr} C_s C_l\right)\boldsymbol{\delta}_{s\neq l} & \text{if } s \neq k \text{ and } l \neq k \\ 0 & \text{otherwise.} \end{cases}$$

Then, it is easy to see that:

$$Q_k = \left(\overline{\Phi}_k - zI_n\right)^{-1} + \frac{1}{z}I_n$$

Hence, from the resolvent identity $Q - (\overline{\Phi}_k - zI_n)^{-1} = Q(\overline{\Phi}_k - \Phi)\left(Q_k - \frac{1}{z}I_n\right) - \frac{1}{z}I_n$. Using the fact that $[Q_k]_{ki} = 0$ for $i \in \{1, \cdots, n\}$, we obtain

$$\begin{aligned}
&\sum_{l=1}^{n}\sum_{r\neq k}\mathbb{E}\left[a_l b_r Q_{lr} - a_l b_r \left[Q_k\right]_{lr}\right] = \sum_{l=1}^{n}\sum_{r\neq l}\mathbb{E}\left[a_l b_r \left[Q\left(\overline{\Phi}_k - \Phi\right)Q_k\right]_{lr}\right] \\
&- \frac{1}{z}\sum_{l=1}^{n}\sum_{r\neq l}\mathbb{E}\left[\left[Q\left(\overline{\Phi}_k - \Phi\right)\right]_{lr} a_l b_r\right] \\
&= \sum_{l=1}^{n}\sum_{r\neq l}\sum_{s\neq k}\mathbb{E}\left[a_l b_r Q_{lk}\left(\overline{\Phi}_{ks} - \Phi_{ks}\right)\left[Q_k\right]_{sr}\right] - \frac{1}{z}\sum_{l=1}^{n}\sum_{r\neq l}\mathbb{E}\left[a_l b_r Q_{lk}\left(\overline{\Phi}_{kr} - \Phi_{kr}\right)\right] \\
&= \eta_1 + \eta_2
\end{aligned}$$

We will first deal with η_1. We have:

$$\begin{aligned}
\eta_1 &= \sum_{l=1}^n \sum_{r\neq l}\sum_{s\neq k} \sqrt{p}\mathbb{E}\left[a_l b_r e_l^T Q_k \xi_{(k,-k)} Q_{kk}\left(x_s^T x_k\right)^2 [Q_k]_{sr}\right] \\
&= \mathbb{E}\left[a^T Q_k \xi_{(k,-k)} \left(\left\{\sqrt{p}(x_s^T x_k)^2 \boldsymbol{\delta}_{s\neq k}\right\}_{s=1}^n\right)^T Q_k b(Q_{kk}-\mathbb{E}Q_{kk})\right] \\
&+ \mathbb{E}\left[a^T Q_k \xi_{(k,-k)} \left(\left\{\sqrt{p}(x_s^T x_k)^2 \boldsymbol{\delta}_{s\neq k}\right\}_{s=1}^n\right)^T Q_k b\right]\mathbb{E}Q_{kk} \\
&= \sum_{l=1}^n \sum_{r\neq l}\sum_{s\neq k}\sum_{f\neq k} p\mathbb{E}Q_{kk}\mathbb{E}\left[a_l b_r \left((x_k^T x_f)^2 - \frac{1}{p^2}\operatorname{tr} C_k C_f\right)[Q_k]_{lf}(x_s^T x_k)^2 [Q_k]_{sr}\right] + O_z(p^{-\frac{1}{2}+\epsilon}) \\
&= \sum_{l=1}^n \sum_{r\neq l}\sum_{s\neq k}\sum_{f\neq k} \frac{2}{p}\mathbb{E}Q_{kk}\mathbb{E}\left[a_l b_r (x_f^T C_k x_s)^2 [Q_k]_{lf}[Q_k]_{sr}\right] \\
&+ \sum_{l=1}^n \sum_{r\neq l}\sum_{s\neq k}\sum_{f\neq k} \frac{1}{p}\mathbb{E}Q_{kk}\mathbb{E}\left[a_l b_r \left(x_f^T C_k x_f - \frac{1}{p}\operatorname{tr} C_k C_f\right)[Q_k]_{lf}[Q_k]_{sr} x_s^T C_k x_s\right] + O_z(p^{-\frac{1}{2}+\epsilon}) \\
&= \frac{2}{p}\mathbb{E}Q_{kk} a^T Q_k \left\{\left(x_f^T C_k x_s\right)^2 \boldsymbol{\delta}_{f\neq k}\boldsymbol{\delta}_{s\neq k}\right\}_{f,s=1}^n Q_k b - \sum_{l=1}^n\sum_{s\neq k}\sum_{f\neq k}\frac{2}{p}\mathbb{E}Q_{kk}\mathbb{E}\left[a_l b_l \left(x_f^T C_k x_s\right)^2 [Q_k]_{lf}[Q_k]_{sl}\right] \\
&+ \frac{1}{p}\mathbb{E}Q_{kk}\mathbb{E}\left[a^T Q_k \left\{\left(x_f^T C_k x_f - \frac{1}{p}\operatorname{tr} C_k C_f\right)\boldsymbol{\delta}_{f\neq k}\right\} b^T Q_k \left\{x_s^T C_k x_s \boldsymbol{\delta}_{s\neq k}\right\}_{s=1}^n\right] \\
&- \sum_{l=1}^n\sum_{s\neq k}\sum_{f\neq k}\frac{1}{p}\mathbb{E}Q_{kk}\mathbb{E}\left[a_l b_l \left(x_f^T C_k x_f - \frac{1}{p}\operatorname{tr} C_k C_f\right)[Q_k]_{lf}[Q_k]_{sl} x_s^T C_k x_s\right] + O_z(p^{-\frac{1}{2}+\epsilon}) \\
&= \frac{2}{p}\mathbb{E}Q_{kk} a^T Q_k \left\{\left(x_f^T C_k x_s\right)^2 \boldsymbol{\delta}_{f\neq k}\boldsymbol{\delta}_{s\neq k}\right\}_{f,s=1}^n Q_k b \\
&- \frac{2}{p}\mathbb{E}Q_{kk} a^T \operatorname{diag}\left\{\left[Q_k\left\{\left(x_f^T C_k x_s\right)^2 \boldsymbol{\delta}_{f\neq k}\boldsymbol{\delta}_{s\neq k}\right\} Q_k\right]_{ll}\right\}_{l=1}^n b \\
&+ \frac{1}{p}\mathbb{E}Q_{kk}\mathbb{E}\left[a^T Q_k \left\{\left(x_f^T C_k x_f - \frac{1}{p}\operatorname{tr} C_k C_f\right)\boldsymbol{\delta}_{f\neq k}\right\} b^T Q_k \left\{x_s^T C_k x_s \boldsymbol{\delta}_{s\neq k}\right\}_{s=1}^n\right] \\
&- \frac{1}{p}\mathbb{E}Q_{kk}\mathbb{E}\left[a^T \operatorname{diag}\left\{\left[Q_k\left\{\left(x_f^T C_k x_f - \frac{1}{p}\operatorname{tr} C_k C_f\right)\boldsymbol{\delta}_{f\neq k}\right\}\left\{x_s^T C_k x_s \boldsymbol{\delta}_{s\neq k}\right\}_{s=1}^n Q_k\right]_{ll}\right\} b\right] + O_z(p^{-\frac{1}{2}+\epsilon}) \\
&= O_z(p^{-\frac{1}{2}+\epsilon})
\end{aligned}$$

We will now deal with η_2. We have:

$$\begin{aligned}
\eta_2 &= \frac{\sqrt{p}}{z}\sum_{l=1}^n \sum_{r\notin\{l,k\}} \mathbb{E}\left[a_l b_r e_l^T Q_k \xi_{(k,-k)} (x_k^T x_r)^2\right]\mathbb{E}Q_{kk} + O_z(p^{-\frac{1}{2}+\epsilon}) \\
&= \frac{2}{pz}\sum_{l=1}^n \sum_{r\notin\{l,k\}}\sum_{s\neq k} \mathbb{E}\left[a_l b_r [Q_k]_{ls}\left(x_s^T C_k x_r\right)^2\right]\mathbb{E}Q_{kk} \\
&+ \frac{1}{pz}\sum_{l=1}^n \sum_{r\notin\{l,k\}}\sum_{s\neq k} \mathbb{E}\left[a_l b_r [Q_k]_{ls}\left(x_s^T C_k x_s - \frac{1}{p}\operatorname{tr} C_k C_s\right) x_r^T C_k x_r\right] + O_z(p^{-\frac{1}{2}+\epsilon}) \\
&= O_z(p^{-\frac{1}{2}+\epsilon})
\end{aligned}$$

$\square$

APPENDIX F: PROOF OF LEMMA 11

We will start by bounding the quantity in Lemma 11 by:

$$\mathbb{E}\left|\sum_{r\notin\{j,k\}}\sum_{b\notin\{j,r,k\}} x_b^T A_1 x_k x_b^T A_2 x_j x_k^T A_3 x_r x_r^T A_4 x_j Q_{br}\right|^2$$
$$\leq 2\mathbb{E}\left|\sum_{r\notin\{j,k\}}\left(\sum_{b\notin\{j,r,k\}} x_b^T A_1 x_k x_b^T A_2 x_j x_k^T A_3 x_r x_r^T A_4 x_j Q_{br} - \mathbb{E}_j x_b^T A_1 x_k x_b^T A_2 x_j x_k^T A_3 x_r x_r^T A_4 x_j Q_{br}\right)\right|^2$$
$$+ 2\mathbb{E}\left|\mathbb{E}_j \sum_{r\notin\{j,k\}}\sum_{b\notin\{j,r,k\}} x_b^T A_1 x_k x_b^T A_2 x_j x_k^T A_3 x_r x_r^T A_4 x_j Q_{br}\right|^2$$
$$= 2\xi_1 + 2\xi_2$$

By Poincaré-Nash inequality, we have for some constant K,

$$\xi_1 \leq K\sum_{l=1}^n \left|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,j,k\}} x_b^T A_1 x_k x_k^T A_3 x_r \frac{1}{\sqrt{p}}\left[C_j^{\frac{1}{2}} A_2 x_b\right]_l x_r^T A_4 x_j Q_{br}\right|^2$$
$$+ K\sum_{l=1}^n \left|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,j,k\}} x_b^T A_1 x_k x_k^T A_3 x_r \frac{1}{\sqrt{p}}\left[C_j^{\frac{1}{2}} A_4 x_r\right]_l x_b^T A_2 x_j Q_{br}\right|^2$$
$$+ K\sum_{l=1}^n \left|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,j,k\}}\sum_{s\neq j} x_b^T A_1 x_k x_k^T A_3 x_r x_b^T A_2 x_j x_r^T A_4 x_j (x_s^T x_j)\left[C_j^{\frac{1}{2}} x_s\right]_l Q_{bj} Q_{sr}\right|^2$$
$$+ K\sum_{l=1}^n \left|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,j,k\}}\sum_{s\neq j} x_b^T A_1 x_k x_k^T A_3 x_r x_b^T A_2 x_j x_r^T A_4 x_j (x_s^T x_j)\left[C_j^{\frac{1}{2}} x_s\right]_l Q_{jr} Q_{sb}\right|^2$$
$$= K\left(\xi_{11} + \xi_{12} + \xi_{13} + \xi_{14}\right)$$

We have:

$$\xi_{11} = \frac{1}{p}\mathbb{E}\Big[x_k^T A_3 X \operatorname{diag}\left\{x_r^T A_4 x_j \boldsymbol{\delta}_{r\neq k}\boldsymbol{\delta}_{r\neq j}\right\}_{r=1}^n \left\{Q_{br}\boldsymbol{\delta}_{b\neq r}\right\} S_j \left\{Q_{br}\boldsymbol{\delta}_{b\neq r}\right\}_{b,r=1}^n$$
$$\times \operatorname{diag}\left\{x_r^T A_4 x_j \boldsymbol{\delta}_{r\neq k}\boldsymbol{\delta}_{r\neq j}\right\}_{r=1}^n X^T A_3 x_k\Big]$$

where $[S_j]_{b_1 b_2} = x_{b_1}^T A_1 x_k x_{b_2}^T A_1 x_k x_{b_1}^T A_2 C_j A_2 x_{b_2} \boldsymbol{\delta}_{b_1\neq k}\boldsymbol{\delta}_{b_2\neq k}$. Since $\|S_j\| = O(p^{-1})$, we have:

$$\xi_{11} = O_z(p^{-3+\epsilon})$$

Term ξ_{12} can be treated similarly, thus leading to $\xi_{12} = O_z(p^{-3+\epsilon})$. To treat ξ_{13}, we first bound it as:

$$\begin{aligned}
\xi_{13} &\leq 2\mathbb{E}\left|\sum_{r\notin\{j,k\}}\sum_{b\notin\{j,k\}}\sum_{s\neq j} x_b^T A_1 x_k x_k^T A_3 x_r x_b^T A_2 x_j x_r^T A_4 x_j x_s^T x_j \left[C_j^{\frac{1}{2}} x_s\right]_l Q_{bj} Q_{sr}\right|^2 \\
&\quad + 2\mathbb{E}\left|\sum_{r\notin\{j,k\}}\sum_{s\neq j} x_r^T A_1 x_k x_k^T A_3 x_r x_r^T A_2 x_j x_r^T A_4 x_j x_s^T x_j \left[C_j^{\frac{1}{2}} x_s\right]_l Q_{rj} Q_{sr}\right|^2 \\
&\leq 2\mathbb{E}\left[\left|x_k^T A_1 X \operatorname{diag}\left\{\left[x_b^T A_2 x_j \boldsymbol{\delta}_{b\neq k}\boldsymbol{\delta}_{b\neq j}\right\}_{b=1}^n Q\right]_j\right|^2\right. \\
&\quad \times x_k^T A_3 X D_j Q \left\{x_{s_1}^T x_j x_{s_2}^T x_j x_{s_1}^T C_j x_{s_2} \boldsymbol{\delta}_{s_1\neq j}\boldsymbol{\delta}_{s_2\neq j}\right\}_{s_1,s_2=1}^n Q D_j X^T A_3 x_k\Big] \\
&\quad + \mathbb{E}\left[x_k^T A_1 X \tilde{D}_j Q \left\{x_{s_1}^T x_j x_{s_2}^T x_j x_{s_1}^T C_j x_{s_2} \boldsymbol{\delta}_{s_1\neq j}\boldsymbol{\delta}_{s_2\neq j}\right\}_{s_1,s_2=1}^n Q \tilde{D}_j X^T A_1 x_k\right]
\end{aligned}$$

where D_j and $\tilde{D}_j$ are $n \times n$ diagonal matrices given by:

$$\begin{aligned}
D_j &= \operatorname{diag}\left\{x_r^T A_4 x_j \boldsymbol{\delta}_{r\neq k}\boldsymbol{\delta}_{r\neq j}\right\}_{r=1}^n \\
\tilde{D}_j &= \operatorname{diag}\left\{x_r^T A_3 x_k x_r^T A_2 x_j x_r^T A_4 x_j \boldsymbol{\delta}_{r\neq k}\boldsymbol{\delta}_{r\neq j}\right\}_{r=1}^n
\end{aligned}$$

which gives:

$$\xi_{13} = O_z(p^{-3+\epsilon})$$

The treatment of ξ_{14} is similar to ξ_{13} as indexes b and r play a symmetric role. All this proves that:

$$\xi_1 = O_z(p^{-3+\epsilon})$$

We will now treat term ξ_2. Using the Integration by Parts formula, we can bound, for some constant K, ξ_2 as:

$$\begin{aligned}
\xi_2 &\leq \frac{K}{p^2}\left|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,k,j\}} x_b^T A_1 x_k x_k^T A_3 x_r x_r^T A_4 C_j A_2 x_b Q_{br}\right|^2 \\
&\quad + \frac{K}{p}\left|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,k,j\}}\sum_{s\neq j} x_b^T A_1 x_k x_k^T A_3 x_r x_s^T x_j x_s^T C_j A_4 x_r x_b^T A_2 x_j Q_{bj} Q_{sr}\right|^2 \\
&\quad + \frac{K}{p}\left|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,k,j\}}\sum_{s\neq j} x_b^T A_1 x_k x_k^T A_3 x_r x_s^T x_j x_s^T C_j A_4 x_r x_b^T A_2 x_j Q_{jr} Q_{sb}\right|^2
\end{aligned}$$

The first term in the right-hand side of the above equation is $O_z(p^{-3+\epsilon})$ as per Lemma 8. As for the second term, we have:

$$\begin{aligned}
&\frac{1}{p}\left|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,k,j\}}\sum_{s\neq j} x_b^T A_1 x_k x_k^T A_3 x_r x_s^T x_j x_s^T C_j A_4 x_r x_b^T A_2 x_j Q_{bj} Q_{sr}\right|^2 \\
&\leq \frac{1}{p}\left|\sum_{r\notin\{j,k\}}\sum_{s\notin\{j,r\}} x_s^T x_j x_s^T C_j A_4 x_r x_k^T A_3 x_r Q_{sr} \sum_{b\notin\{r,k,j\}} x_b^T A_1 x_k x_b^T A_2 x_j Q_{bj}\right|^2 \\
&+ \frac{1}{p}\left|\sum_{r\notin\{j,k\}} x_r^T x_j x_r^T C_j A_4 x_r x_k^T A_3 x_r Q_{rr} \sum_{b\notin\{r,k,j\}} x_b^T A_1 x_k x_b^T A_2 x_j Q_{bj}\right|^2 \\
&= O_z(p^{-3+\epsilon})
\end{aligned}$$

As for the last term, we have:

$$\begin{aligned}
&\frac{K}{p}\left|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,k,j\}}\sum_{s\neq j} x_b^T A_1 x_k x_k^T A_3 x_r x_s^T x_j x_s^T C_j A_4 x_r x_b^T A_2 x_j Q_{jr} Q_{sb}\right|^2 \\
&= \frac{K}{p}\left|\sum_{r\notin\{j,k\}} x_k^T A_3 x_r x_r^T A_4 C_j X \mathrm{diag}\left\{x_s^T x_j \boldsymbol{\delta}_{s\neq j}\right\}_{s=1}^n Q \mathrm{diag}\left\{x_b^T A_2 x_j \boldsymbol{\delta}_{b\neq r}\boldsymbol{\delta}_{b\neq j}\boldsymbol{\delta}_{b\neq k}\right\} X^T C_j x_k Q_{jr}\right|^2 \\
&= O_z(p^{-3+\epsilon})
\end{aligned}$$

APPENDIX G: PROOF OF THEOREM 3

To begin with, we expand $\mathbb{E}a^T Q(z_1) D Q(z_2) b$ as:

$$\begin{aligned}
\mathbb{E}a^T Q(z_1) D Q(z_2) b &= \sum_{i=1}^n\sum_{j=1}^n\sum_{k\notin\{i,j\}} \mathbb{E}\left[a_i\left[Q(z_1)\right]_{ik} D_{kk}\left[Q(z_2)\right]_{kj} b_j\right] \\
&+ \sum_{j=1}^n\sum_{i\neq j}\mathbb{E}\left[a_i\left[Q(z_1)\right]_{ii} D_{ii}\left[Q(z_2)\right]_{ij} b_j\right] + \sum_{j=1}^n\sum_{i\neq j}\mathbb{E}\left[a_i\left[Q(z_1)\right]_{ij} D_{jj}\left[Q(z_2)\right]_{jj} b_j\right] \\
&+ \sum_{i=1}^n \mathbb{E}\left[a_i\left[Q(z_1)\right]_{ii} D_{ii}\left[Q(z_2)\right]_{ii} b_i\right] \qquad (44)
\end{aligned}$$

The last term can be approximated as:

$$\sum_{i=1}^n \mathbb{E}\left[a_i\left[Q(z_1)\right]_{ii} D_{ii}\left[Q(z_2)\right]_{ii} b_i\right] = m(z_1)m(z_2)a^T D b + O_z(p^{-\frac{1}{4}})$$

The second term can be treated as follows:

$$\begin{aligned}
&\sum_{j=1}^n\sum_{i\neq j}\mathbb{E}\left[a_i\left[Q(z_1)\right]_{ii} D_{ii}\left[Q(z_2)\right]_{ij} b_j\right] = \sum_{j=1}^n\sum_{i\neq j} a_i \mathbb{E}\left[\left[\left[Q(z_1)\right]_{ii} - \mathbb{E}\left[Q(z_1)\right]_{ii}\right] D_{ii}\left[Q(z_2)\right]_{ij} b_j\right] \\
&+ \sum_{j=1}^n\sum_{i\neq j} a_i \mathbb{E}\left[Q(z_1)\right]_{ii} D_{ii} \mathbb{E}\left[Q(z_2)\right]_{ij} b_j
\end{aligned}$$

We have:

$$\left|\sum_{j=1}^{n}\sum_{i\neq j} a_i \mathbb{E}\left[\left([Q(z_1)]_{ii} - \mathbb{E}\left[Q(z_1)\right]_{ii}\right) D_{ii}\left[Q(z_2)\right]_{ij} b_j\right]\right|$$
$$\leq \sqrt{\sum_{i=1}^{n} a_i^2 D_{ii}^2 \mathbb{E}\left|[Q(z_1)]_{ii} - \mathbb{E}\left[Q(z_1)\right]_{ii}\right|^2}\sqrt{\sum_{i=1}^{n}\left|[Q(z_2)b]_i\right|^2}$$
$$= O_z(p^{-\frac{1}{2}+\epsilon})$$

Using Theorem 2, we thus obtain:

$$\sum_{j=1}^{n}\sum_{i\neq j}\mathbb{E}\left[a_i\left[Q(z_1)\right]_{ii} D_{ii}\left[Q(z_2)\right]_{ij} b_j\right] = \frac{2c_0\left(\frac{1}{p}\operatorname{tr}\left(C^\circ\right)^4\right) m^3(z_2)m(z_1)}{1-\frac{2}{p}\operatorname{tr}\left(C^\circ\right)^4 c_0^2 m^2(z_2)} a^T D \frac{11^T}{p} b + O_z(p^{-\frac{1}{4}})$$

In the same way, we can prove that:

$$\sum_{j=1}^{n}\sum_{i\neq j} a_i \mathbb{E}\left[Q(z_1)\right]_{ij} D_{jj}\mathbb{E}\left[Q(z_2)\right]_{jj} b_j = \frac{2c_0\left(\frac{1}{p}\operatorname{tr}\left(C^\circ\right)^4\right) m^3(z_1)m(z_2)}{1-\frac{2}{p}\operatorname{tr}\left(C^\circ\right)^4 c_0^2 m^2(z_1)} a^T \frac{11^T}{p} Db + O_z(p^{-\frac{1}{4}})$$

We will now treat the first term in (44). Using (7) along with Lemma 3, we have:

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k\notin\{i,j\}}\mathbb{E}\left[a_i\left[Q(z_1)\right]_{ik} D_{kk}\left[Q(z_2)\right]_{kj} b_j\right]$$
$$= \sum_{i,j=1}^{n}\sum_{k\notin\{i,j\}} a_i D_{kk} b_j \sum_{\substack{l\neq k\\ m\neq k}}\frac{1}{p}\mathbb{E}\left[(x_l^T C_k x_l - \frac{1}{p}\operatorname{tr} C_k C_l)(x_m^T C_k x_m - \frac{1}{p}\operatorname{tr} C_k C_m)\left[Q_k(z_2)\right]_{lj}\left[Q_k(z_1)\right]_{im}\right]$$
$$\times m(z_1)m(z_2) + \frac{2}{p}\sum_{i,j=1}^{n}\sum_{k\notin\{i,j\}} a_i D_{kk} b_j \sum_{l\neq k}\sum_{m\neq k}\mathbb{E}\left[\left(x_l^T C_k x_m\right)^2\left[Q_k(z_2)\right]_{lj}\left[Q_k(z_1)\right]_{im}\right] m(z_1)m(z_2)$$
$$+ O_z(p^{-\frac{1}{4}})$$

To treat the first term, we will make use of the following variance estimation,

$$\operatorname{var}\left(\sum_{j=1}^{n}\sum_{l\neq k}\frac{1}{\sqrt{p}}\left(x_l^T C_k x_l - \frac{1}{p}\operatorname{tr} C_k C_l\right)\left[Q_k(z)\right]_{lj} d_j\right) = O_z(p^{-2+\epsilon})$$

with $d \in \mathbb{C}^{n\times 1}$ with bounded norm and $z \in \mathbb{C}\backslash\mathbb{R}$, the proof of which is based on standard calculations using the Poincaré-Nash inequality and is omitted. We thus have:

$$\sum_{i,j=1}^{n}\sum_{k\notin\{i,j\}} a_i D_{kk} b_j \sum_{l\neq k}\sum_{m\neq k}\frac{1}{p}\mathbb{E}\left[(x_l^T C_k x_l - \frac{1}{p}\operatorname{tr} C_k C_l)(x_m^T C_k x_m - \frac{1}{p}\operatorname{tr} C_k C_m)\left[Q_k(z_2)\right]_{lj}\left[Q_k(z_1)\right]_{im}\right]$$
$$= \sum_{k=1}^{n} D_{kk}\mathbb{E}\left[\sum_{i\neq k} a_i \sum_{m\neq k}\frac{1}{\sqrt{p}}\left(x_m^T C_k x_m - \frac{1}{p}\operatorname{tr} C_k C_m\right)\left[Q_k(z_1)\right]_{im}\right]$$
$$\times\mathbb{E}\left[\sum_{j\neq k} b_j \sum_{l\neq k}\frac{1}{\sqrt{p}}\left(x_l^T C_k x_l - \frac{1}{p}\operatorname{tr} C_k C_l\right)\left[Q_k(z_2)\right]_{lj}\right] + O_z(p^{-1+\epsilon})$$
$$= \sum_{k=1}^{n} D_{kk}\mathbb{E}\left[\sum_{i\neq k} a_i \sum_{m\notin\{k,i\}}\frac{1}{\sqrt{p}}\left(x_m^T C_k x_m - \frac{1}{p}\operatorname{tr} C_k C_m\right)\left[Q_k(z_1)\right]_{im}\right]$$
$$\times\mathbb{E}\left[\sum_{j\neq k} b_j \sum_{l\notin\{k,j\}}\frac{1}{\sqrt{p}}\left(x_l^T C_k x_l - \frac{1}{p}\operatorname{tr} C_k C_l\right)\left[Q_k(z_2)\right]_{lj}\right] + O_z(p^{-\frac{1}{2}+\epsilon})$$

Using Proposition (2), we finally obtain:

$$\sum_{i,j=1}^{n}\sum_{k\notin\{i,j\}} a_i D_{kk} b_j \sum_{l\neq k}\sum_{m\neq k}\frac{1}{p}\mathbb{E}\left[(x_l^T C_k x_l - \frac{1}{p}\operatorname{tr} C_k C_l)(x_m^T C_k x_m - \frac{1}{p}\operatorname{tr} C_k C_m)\left[Q_k(z_2)\right]_{lj}\left[Q_k(z_1)\right]_{im}\right]$$
$$= a^T\frac{11^T}{p}b\frac{c_0^2 m^2(z_1)m^2(z_2)\Omega^4}{\left(1-\Omega^2 c_0^2 m^2(z_1)\right)\left(1-\Omega^2 c_0^2 m^2(z_2)\right)}\frac{1}{p}\operatorname{tr} D + O_z(p^{-\frac{1}{4}}) + O_z(p^{-\frac{1}{4}})$$

Finally, we can prove using standard calculations that:

$$\frac{2}{p}\sum_{i,j=1}^{n}\sum_{k\notin\{i,j\}} a_i D_{kk} b_j \sum_{l\neq k}\sum_{m\neq k}\mathbb{E}\left[\left(x_l^T C_k x_m\right)^2\left[Q_k(z_2)\right]_{lj}\left[Q_k(z_1)\right]_{im}\right] m(z_1)m(z_2)$$
$$= \frac{2}{p}\operatorname{tr} D\left(\frac{1}{p}\operatorname{tr}(C^\circ)^2\right)^2\mathbb{E}\left[a^T Q(z_1)Q(z_2)b\right] m(z_1)m(z_2)$$
$$+ \frac{2}{p}\operatorname{tr} D\frac{1}{p}\operatorname{tr}(C^\circ)^4\,\mathbb{E}\left[\frac{1}{p}a^T Q(z_1)1b^T Q(z_2)1\right] m(z_1)m(z_2) + O_z(p^{-\frac{1}{4}})$$
$$= \omega^2\frac{1}{p}\operatorname{tr} D\mathbb{E}\left[a^T Q(z_1)Q(z_2)b\right] m(z_1)m(z_2) + \frac{\frac{1}{p}\operatorname{tr} D\Omega^2 m^2(z_1)m^2(z_2)a^T\frac{11^T}{p}b}{\left(1-\Omega^2 c_0^2 m^2(z_1)\right)\left(1-\Omega^2 c_0^2 m^2(z_2)\right)} + O_z(p^{-\frac{1}{4}})$$

Gathering all the results together, we show in a first step by setting $D = I_n$ that:

$$\mathbb{E}\left[a^T Q(z_1)Q(z_2)b\right] = g(z_1,z_2) + O_z(p^{-\frac{1}{4}})$$

where

$$\begin{aligned} g(z_1,z_2) &= (1-\omega^2 c_0 m(z_1)m(z_2))^{-1}m(z_1)m(z_2)a^T b \\ &\quad + a^T\frac{11^T}{p}b\left[1-\Omega^2 c_0^2 m^2(z_2)\right]^{-1}\left[1-\Omega^2 c_0^2 m^2(z_1)\right]^{-1}\left[1-\omega^2 c_0 m(z_1)m(z_2)\right]^{-1} \\ &\quad \times c_0\Omega^2 m(z_1)m(z_2)\left[m^2(z_1)+m^2(z_2)+1-c_0^2\Omega^2 m^2(z_1)m^2(z_2)\right] \end{aligned}$$

which finally gives:

$$
\begin{aligned}
&\mathbb{E}\left[a^T Q(z_1) D Q(z_2) b\right] = m(z_1)m(z_2)a^T Db \\
&+ \left[1-\Omega^2 c_0^2 m^2(z_2)\right]^{-1} c_0\Omega^2 m^3(z_2)m(z_1)\frac{a^T D 1 1^T b}{p} \\
&+ \left[1-\Omega^2 c_0^2 m^2(z_1)\right]^{-1} c_0\Omega^2 m^3(z_1)m(z_2)\frac{a^T 1 1^T D b}{p} \\
&+ \frac{c_0^2}{p} a^T 1 1^T b \Omega^4 \left(\frac{1}{p}\operatorname{tr} D\right) m^3(z_1)m^3(z_2)\left[1-\Omega^2 c_0^2 m^2(z_1)\right]^{-1}\left[1-\Omega^2 c_0^2 m^2(z_2)\right]^{-1} \\
&+ \left(\frac{1}{p}\operatorname{tr} D\right)\Omega^2 m^2(z_1)m^2(z_2)\left[1-\Omega^2 c_0^2 m^2(z_1)\right]^{-1}\left[1-\Omega^2 c_0^2 m^2(z_2)\right]^{-1}\frac{1}{p}a^T 1 1^T b \\
&+ \left(\frac{1}{p}\operatorname{tr} D\right)\omega^2 m(z_1)m(z_2)g(z_1,z_2) + O_z(p^{-\frac{1}{4}})
\end{aligned}
$$

APPENDIX H: PROOF OF THEOREM 5

The approach to determining the isolated eigenvalues of $P\tilde{\Phi}P$ follows classical techniques from spiked random matrix models (See eig. [Benaych-Georges and Nadakuditi, 2012, Baik and Silverstein, 20 Couillet and Hachem, 2013]). The eigenvalues λ of $P\tilde{\Phi}P$ falling at macroscopic distance from $\mathcal{S}$ and different from $\{\tilde{\rho}, -\tilde{\rho}\}$ should solve the following equation:

$$
0 = \det\left(P\tilde{\Phi}P - \lambda I_n\right) = \det\left(\Phi + UMU^T + \operatorname{diag}\left\{\frac{1}{p^{\frac{3}{2}}}\operatorname{tr} C_i^2\right\}_{i=1}^n - \lambda I_n\right) + o(1) \tag{45}
$$

matrices M and U being given by

$$
\begin{aligned}
M &= \begin{bmatrix} \left(\operatorname{diag}\{c_i\}_{i=1}^k\right)^{-\frac{1}{2}} \mathcal{T} \left(\operatorname{diag}\{c_i\}_{i=1}^k\right)^{-\frac{1}{2}} & -\frac{1}{c_0}1_k \\ -\frac{1}{c_0}1_k & 0 \end{bmatrix} \\
U &= \frac{1}{\sqrt{p}}\begin{bmatrix} J & \Phi 1_n \end{bmatrix}
\end{aligned}
$$

where $J = [j_1, \cdots, j_c]$, with $j_a = \left[0_{n_1}^T, \cdots, 0_{n_a-1}^T, 1_{n_a}^T, 0_{n_a+1}^T, \cdots, 0_{n_c}^T\right]^T \in \mathbb{R}^n$ and we used the fact that $\frac{1}{n}1^T\Phi 1$ converges almost surely to zero. Define the resolvent matrix $Q_\lambda = (\Phi - \lambda I_n)^{-1}$. Then,

$$
\begin{aligned}
&\det\left(P\tilde{\Phi}P - \lambda I_n\right) \\
&= \det\left(Q_\lambda^{-1} + \operatorname{diag}\left\{\frac{1}{p^{\frac{3}{2}}}\operatorname{tr} C_i^2\right\}_{i=1}^n\right)\det\left(I_n + \left(Q_\lambda^{-1} + \operatorname{diag}\left\{\frac{1}{p^{\frac{3}{2}}}\operatorname{tr} C_i^2\right\}_{i=1}^n\right)^{-1} UMU^T\right) + o(1) \\
&= \det\left(Q_\lambda^{-1} + \operatorname{diag}\left\{\frac{1}{p^{\frac{3}{2}}}\operatorname{tr} C_i^2\right\}_{i=1}^n\right)\det\left(I_n + Q_\lambda UMU^T\right) + o(1)
\end{aligned}
$$

Using Theorem 2 along with Sylverster's identity, we can prove that:

$$
\det(I_n + Q_\lambda UMU^T) \xrightarrow{\text{a.s.}} \det H(\lambda)
$$

where $H(\lambda)$ is given by

$$H(\lambda) \triangleq \begin{bmatrix} I_c + c_0 m(\lambda) \left(\mathrm{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}} \mathcal{T} \left(\mathrm{diag}\left\{c_i\right\}_{i=1}^c\right)^{-\frac{1}{2}} & -\frac{a(\lambda)}{\lambda}\vec{c} \\ -\,\vec{c}1_c^T(1+a(\lambda)) & \\ & \\ -\lambda 1_c^T - \lambda a(\lambda) 1_c^T & -a(\lambda) \end{bmatrix}$$

with

$$a(\lambda) = \frac{\lambda m(\lambda)}{1 - \frac{2}{p}\,\mathrm{tr}\left(C^\circ\right)^4 c_0^2 m^2(\lambda)}$$

and $\vec{c} = [c_1, \cdots, c_c]^T$. Replacing and evaluating the determinant as a block-matrix determinat, we then find after simplifications:

$$\det(H(\lambda)) = (-1)^c (a(\lambda))^c \det\left(I_c + c_0 m(\lambda)\mathcal{T}\right)$$

As such, the eigenvalues of $P\tilde{\Phi}P$ at a macroscopic distance of $\mathcal{S}$ converge to those ρ's in $\mathbb{R}\backslash\mathcal{S}$ for which $\det\left(I_c + c_0 m(\lambda)\mathcal{T}\right) = 0$, or equivalently $m(\rho) = -\frac{1}{c_0\eta}$ for some non-zero eigenvalue η of $\mathcal{T}$. Since m is a growing function from $\left(2\sqrt{c_0}\omega, \infty\right)$ onto $\left(-\frac{1}{\sqrt{c_0}\omega}, 0\right)$ and from $\left(-\infty, -2\sqrt{c_0}\omega\right)$ onto $\left(0, \frac{1}{\sqrt{c_0}\omega}\right)$, the condition for the existence of such ρ is that $|c_0\eta| > \sqrt{c_0}\omega$, i.e., that $\sqrt{c_0}|\eta| > \omega$. When this condition is met, the resulting eigenvalue of $P\tilde{\Phi}P$ converges to ρ satisfying:

$$-\frac{1}{c_0\eta} = -\frac{1}{\rho + c_0\omega^2\left(-\frac{1}{c_0\eta}\right)}$$

which implies that:

$$\rho = c_0\eta + \frac{\omega^2}{\eta}$$

The proof is then concluded by an application of the argument principle that ensures that asymptotically eigenvalues of $P\tilde{\Phi}P$ and their limiting values have the same multiplicities.

APPENDIX I: PROOF OF THEOREM 6

Following again [Couillet and Benyach-Georges, 2016], let us consider an eigenvalue λ of $P\tilde{\Phi}P$ converging to ρ. Define Π_λ the projector on the eigenspace associated with the eigenvalue of $P\tilde{\Phi}P$ converging to ρ. We wish here to investigate the limit of the matrix $\frac{1}{p}J^T\Pi_\lambda J$ the entries of which are, up to a scaling, the desired quantities $\alpha_i^a\alpha_i^b$.

Note first, by residue calculus, that:

$$\frac{1}{p}J^T\Pi_\lambda J = -\frac{1}{2\pi\imath}\oint_{\mathcal{C}_\rho}\frac{1}{p}J^T\left(P\tilde{\Phi}P - zI_n\right)^{-1}Jdz$$

for all large n almost surely, where $\mathcal{C}_\rho$ is a complex (positively oriented and with winding number one) contour circling around ρ only. From the proof of Theorem 4, we have for all z over the contour C_ρ,

$$\frac{1}{p}J^T\left(P\tilde{\Phi}P - zI_n\right)^{-1}J = \frac{1}{p}J^TQ(z)J - \frac{1}{p}J^TQ(z)UM\left(I_{c+1} + U^TQ(z)UM\right)^{-1}U^TQ(z)J + o(1)$$

The first right-hand side term has asymptotically no residue when integrated over $\mathcal{C}_\rho$. Thus, we need only to focus on the second right-hand side term. Following the block-structure approach used in the proof of Theorem 4, we show, using the identity $\mathcal{T}\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{-\frac{1}{2}}\left\{c_i\right\}_{i=1}^c=0_{k\times 1}$, that:

$$\left(I_{c+1}+U^TQ(z)UM\right)^{-1}\xrightarrow{\text{a.s.}}\begin{bmatrix}\left(I_c+c_0m(z)\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}}\mathcal{T}\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{-\frac{1}{2}}\right)^{-1} & -\vec{c}\\ -\frac{z}{a(z)}1_c^T-z1_c^T & 1\end{bmatrix}$$

while:

$$\frac{1}{\sqrt{p}}J^TQ(z)UM\xrightarrow{\text{a.s.}}\begin{bmatrix}c_0m(z)\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}}\mathcal{T}\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{-\frac{1}{2}}-\vec{c}1_c^T(1+a(z)) & -\frac{a(z)}{z}\vec{c}\end{bmatrix}$$

and

$$\frac{1}{\sqrt{p}}U^TQ(z)J\xrightarrow{\text{a.s.}}\begin{bmatrix}c_0m(z)\operatorname{diag}\left(\left\{c_i\right\}_{i=1}^c\right)+\frac{2c_0^3\frac{1}{p}\operatorname{tr}(C^\circ)^4m^3(z)\vec{c}\vec{c}^T}{1-\frac{2}{p}\operatorname{tr}(C^\circ)^4c_0^2m^2(z)}\\ c_0(1+a(z))\left(\vec{c}\right)^T\end{bmatrix}$$

where

$$a(z)=\frac{zm(z)}{1-\frac{2}{p}\operatorname{tr}\left(C^\circ\right)^4c_0^2m^2(z)}$$

All calculus made, we then find:

$$\frac{1}{p}J^T\Pi_\lambda J=\frac{1}{2\pi\imath}\oint_{\mathcal{C}_\rho}c_0^2m^2(z)\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}}\mathcal{T}\mathcal{G}(z)\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}}+\frac{c_0}{z}(1+a(z))\vec{c}\left(\vec{c}\right)^T dz$$

(46)

$$+o(1)$$

where

$$\begin{aligned}\mathcal{G}(z)&\triangleq\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{-\frac{1}{2}}\left(I_c+c_0m(z)\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}}\mathcal{T}\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{-\frac{1}{2}}\right)^{-1}\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}}\\&=\left(I_c+c_0m(z)\mathcal{T}\right)^{-1}\end{aligned}$$

Note that the right-hand side term in (46) do not produce any residue when integrated over a properly chosen $\mathcal{C}_\rho$. This is because if this term has a pole, we must have:

$$m^2(\rho)=\frac{1}{\frac{2}{p}\operatorname{tr}\left(C^\circ\right)^4c_0^2}$$

or equivalently

$$\eta_i=\sqrt{\frac{2}{p}\operatorname{tr}\left(C^\circ\right)^4}$$

which leads to a contradiction as per the Assumption of Theorem 6. Computing the residue of the first term leads to the following expression:

$$\frac{1}{p}J^T\Pi J=c_0\left(1-\frac{\omega^2}{c_0}\eta_i^2\right)\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}}v_iv_i^T\left(\operatorname{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}}+o(1)$$

where v_i is the eigenvector associated with the eigenvalue η_i. We then obtain:

$$\alpha_i^a\alpha_i^b\xrightarrow{\text{a.s.}}(1-\frac{1}{c_0}\frac{\omega^2}{\eta_i^2})\left[v_iv_i^T\right]_{ab}$$

APPENDIX J: PROOF OF THEOREM 7

The evaluation of σ_{ij}^a involves that $\hat{u}_i^T D_a \hat{u}_j$ which may not be directly estimated from the application of the standard Cauchy integral method. To get this around, we express $\hat{u}_i^T D_a \hat{u}_j$ as:

$$\hat{u}_i^T D_a \hat{u}_j = \frac{\frac{1}{p} j_b^T \hat{u}_i \hat{u}_i^T D_a \hat{u}_j \hat{u}_j^T j_d}{\frac{1}{p} j_b^T \hat{u}_i \hat{u}_j^T j_d} \tag{47}$$

where b and d are some indexes for which $\frac{1}{\sqrt{p}} j_b^T \hat{u}_i$ and $\frac{1}{\sqrt{p}} j_b^T \hat{u}_i$ have non-zero asymptotic limits. The denominator in (47) can be approximated using the previous estimates of α_i^a and α_i^b, while to compute the numerator, we need to evaluate $\frac{1}{p} J^T \Pi_{\lambda_i} D_a \Pi_{\lambda_j} J$ where Π_{λ_i} and Π_{λ_j} are the projectors on the subspaces associated with λ_i and λ_j. Such a quantity can be estimated by noting that for almost sure for n large enough, we have:

$$\frac{1}{p} J^T \Pi_{\lambda_i} D_a \Pi_{\lambda_j} J = \frac{1}{(2\pi \imath)^2} \oint_{C_{\rho_i}} \oint_{\mathcal{C}_{\rho_j}} \frac{1}{p} J^T \left(P \tilde{\Phi} P - z_1 I_n \right)^{-1} D_a \left(P \tilde{\Phi} P - z_2 I_n \right)^{-1} J dz_1 dz_2$$

where ρ_i and ρ_j are the limiting values of λ_i and λ_j. With the same approach as above, applying Woodbury's identity on each inverse and noticing that the generated cross-terms will have almost surely zero-residue, we have:

$$\frac{1}{p} J^T \Pi_{\lambda_1} D_a \Pi_{\lambda_2} J = \frac{1}{(2\pi \imath)^2} \oint_{C_{\rho_1}} \oint_{\mathcal{C}_{\rho_2}} R(z_1, z_2) dz_1 dz_2 + o(1)$$

with

$$\begin{aligned} R(z_1, z_2) = {} & \frac{1}{p} J^T Q(z_1) U M \left(I_{c+1} + U^T Q(z_1) U M \right)^{-1} U^T Q(z_1) D_a Q(z_2) U \\ & \times \left(I_{c+1} + U^T Q(z_1) U M \right)^{-1} M^T U Q(z_2) J \end{aligned}$$

Denote by $G(z) = I_c + c_0 m(z) \left(\operatorname{diag} \{c_i\}_{i=1}^c \right)^{\frac{1}{2}} \mathcal{T} \left(\operatorname{diag} \{c_i\}_{i=1}^c \right)^{-\frac{1}{2}}$. Then, using the fact that $G(z)^{-1} \vec{c} = \vec{c}$ and $G(z)^{-1} 1 = 1$, it follows from the estimates derived in the proof of Theorem 6 that for $k = 1, 2$,

$$\frac{1}{\sqrt{p}} J^T Q(z_k) U M = \left[c_0 \left(\operatorname{diag} \{c_\ell\}_{\ell=1}^c \right)^{\frac{1}{2}} \mathcal{T} \left(\operatorname{diag} \{c_\ell\}_{\ell=1}^c \right)^{-\frac{1}{2}} m(z_k) G(z_k)^{-1}, \ 0 \right] + r(z_k) + o(1)$$

where $r(z_k)$ is a matrix that shall have no residue when integrated over $\mathcal{C}_{\rho_i}$. Now, from Theorem 3, we have:

$$\begin{aligned} \frac{1}{p} J^T Q(z_1) D_a Q(z_2) J = {} & c_0 c_a m(z_1) m(z_2) \left\{ \boldsymbol{\delta}_{i=a} \right\}_{i=1}^c \\ & + c_a c_0^2 \omega^2 m^2(z_1) m^2(z_2) \left(1 - \omega^2 c_0 m(z_1) m(z_2) \right)^{-1} \operatorname{diag} \{c_i\}_{i=1}^c \\ & + q_1(z_1, z_2) \vec{c} (\vec{c})^T + q_2(z_1, z_2) \operatorname{diag} \left\{ \boldsymbol{\delta}_{i=a} \right\}_{i=1}^c 1 (\vec{c})^T + q_3(z_1, z_2) \vec{c} 1^T \operatorname{diag} \left\{ \boldsymbol{\delta}_{i=a} \right\}_{i=1}^c \\ & + o(1) \end{aligned}$$

where $q_j(z_1, z_2)$, $j = 1, \cdots, 3$ are analytic functions over $\mathcal{C}_{\rho_1}$ and $\mathcal{C}_{\rho_2}$. Using again the fact that $G(z)^{-1} \vec{c} = \vec{c}$ and $G(z)^{-1} 1 = 1$, we deduce that the terms involving $q_j(z_1, z_2)$, $j = 1, \cdots, 3$

vanishes. As such, we have:

$$R(z_1,z_2) = c_0^3 m^2(z_1)m^2(z_2)\left(\mathrm{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}}\mathcal{T}\mathcal{G}^{-1}(z_1)\mathrm{diag}\left\{\boldsymbol{\delta}_{i=a}\right\}\mathcal{G}^{-1}(z_2)\mathcal{T}\left(\mathrm{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}}$$
$$+ c_0^4 c_a \omega^2 m^3(z_1)m^3(z_2)\left(1-\omega^2 c_0 m(z_1)m(z_2)\right)^{-1}\left(\mathrm{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}}\mathcal{T}\mathcal{G}^{-1}(z_1)\mathcal{G}^{-1}(z_2)\mathcal{T}\left(\mathrm{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}}$$
$$+ \tilde{q}(z_1,z_2) + o(1)$$

where $\mathcal{G}(z) = \left(\mathrm{diag}\left\{c_i\right\}_{i=1}^c\right)^{-\frac{1}{2}} G(z)\left(\mathrm{diag}\left\{c_i\right\}_{i=1}^c\right)^{\frac{1}{2}} = I_c + c_0 m(z)\mathcal{T}$ and $\tilde{q}(z_1,z_2)$ is a matrix having no asymptotic resudue when integrated over $\mathcal{C}_{\rho_1}$ and $\mathcal{C}_{\rho_2}$.

It is now interesting to note that, after residue calculus, the second right-hand side term is zero if $\rho_i \neq \rho_j$, since then the eigenspaces of $\mathcal{G}(z_1)$ and $\mathcal{G}(z_2)$ associated to distinct eigenvalues ought to be orthogonal. Developing the residues of $R(z_1,z_2)$ with the estimates obtained in the proof of Theorem 6, we then find:

$$\frac{1}{p}J^T\Pi_{\lambda_i}D_a\Pi_{\lambda_j}J = \tag{48}$$
$$c_0\left(1-\frac{\omega^2}{c_0\eta_i^2}\right)\left(1-\frac{\omega^2}{c_0\eta_j^2}\right)\left(\mathrm{diag}\left\{c_k\right\}_{k=1}^c\right)^{\frac{1}{2}} v_i v_i^T \mathrm{diag}\left\{\boldsymbol{\delta}_{k=a}\right\}_{k=1}^c v_j v_j^T\left(\mathrm{diag}\left\{c_k\right\}_{k=1}^c\right)^{\frac{1}{2}}$$
$$+\boldsymbol{\delta}_{\rho_i=\rho_j}\frac{c_a\omega^2}{\eta_i^2}\left(1-\frac{\omega^2}{c_0\eta_i^2}\right)\left(\mathrm{diag}\left\{c_k\right\}_{k=1}^c\right)^{\frac{1}{2}} v_i v_i^T\left(\mathrm{diag}\left\{c_k\right\}_{k=1}^c\right)^{\frac{1}{2}} + o(1) \tag{49}$$

Now assume that $v_i(a)$ or $v_j(a)$ are zero. If $\rho_i \neq \rho_j$, then it follows from (49) and Theorem 6

$$\frac{1}{p}J^T\Pi_{\lambda_i}D_a\Pi_{\lambda_j}J = o(1)$$

and

$$\alpha_i^a\alpha_j^a = o(1)$$

We thus have:

$$\sigma_{ij}^a = o(1)$$

Assume now that $\rho_i = \rho_j$ and the a-th element of v_i or v_j is zero. Let d be an index such that $\left[\frac{1}{p}J^T\Pi J\right]_{dd} \not\to 0$. Such index must exist because vector v_i is different from zero. Hence, using the estimates of the above section, we then find:

$$\sigma_{ii}^a \xrightarrow{\text{a.s.}} \frac{c_a\omega^2}{c_0\eta_i^2}$$

Now, we treat the case in which the a-th element of v_i and v_j are different from zero. We can easily check that if $i \neq j$,

$$\sigma_{ij}^a \xrightarrow{\text{a.s.}} 0$$

and $\sigma_{ii}^a \xrightarrow{\text{a.s.}} \frac{c_a\omega^2}{c_0\eta_i^2}$. Gathering all results together, we thus find:

$$\sigma_{ij}^a \xrightarrow{\text{a.s.}} \boldsymbol{\delta}_{i=j}\frac{c_a\omega^2}{c_0\eta_i^2}$$

REFERENCES

[Bai and Silverstein, 1998] Bai, Z. D. and Silverstein, J. W. (1998). No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Large Dimensional Sample Covariance Matrices. *Annals of Probability*, 26(1):316–345.

[Bai and Silverstein, 2006] Bai, Z. D. and Silverstein, J. W. (2006). *Spectral Analysis of Large Dimensional Random Matrices*. Mathematics Monograph Series 2. Science Press, Beijing.

[Baik and Silverstein, 2006] Baik, J. and Silverstein, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.*, 97(6):1382–1408.

[Benaych-Georges and Nadakuditi, 2012] Benaych-Georges, F. and Nadakuditi, R. R. (2012). The singular values and vectors of low rank perturbations of large rectangular random matrices. *Journal of Multivariate Analysis*, 111:120–135.

[Capitaine et al., 2009] Capitaine, M., Martin, C. D., and Féral, D. (2009). "The Largest Eigenvalues of Finite Rank Deformation of Large Wigner Matrices: Convergence and Nonuniversality of the Fluctuations". *The Annals of Probability*, 37(1):1–47.

[Cheng and Singer, 2013] Cheng, X. and Singer, A. (2013). The spectrum of random inner-product kernel matrices. *Random Matrices: Theory and Applications*, 02(04):1350010.

[Couillet and Benyach-Georges, 2016] Couillet, R. and Benyach-Georges, F. (2016). Kernel Spectral Clustering of Large Dimensional Data.

[Couillet and Debbah, 2011] Couillet, R. and Debbah, M. (2011). *Random matrix methods for wireless communications*. Cambridge University Press Cambridge, MA.

[Couillet and Hachem, 2013] Couillet, R. and Hachem, W. (2013). Fluctuations of spiked random matrix models and failure diagnosis in sensor networks. *IEEE Transactions on Information Theory*, 59(1):509–525.

[El Karoui, 2010] El Karoui, N. (2010). The spectrum of kernel random matrices. *Ann. Statist.*, 38(1):1–50.

[Fan and Montanari, 2015] Fan, Z. and Montanari, A. (2015). The spectral norm of random inner-product kernel matrices.

[Haagerup and Thorbjørnsen, 2005] Haagerup, U. and Thorbjørnsen, S. (2005). A new application of random matrices: $\mathrm{Ext}(\mathrm{C}^*_{\mathrm{red}}(\mathrm{F2}))$ is not a group. *Annals of Mathematics*, 162(2):711–775.

[Hofmann et al., 2008] Hofmann, T., Scholkopf, B., and Smola, A. (2008). Kernel methods in machine learning. *The Annals of Statistics*, 36(3):1171–1220.

[Loubaton, 2016] Loubaton, P. (2016). On the Almost Sure Location of the Singular Values of Certain Block-Hanker Large Random Matrices. *Journal of Theoretical Probability*, 29(4):1339–1443.

[Ng et al., 2002] Ng, A. Y., Jordan, M., and Weiss, Y. (2002). On spectral clustering: Analysis and an algorithm. In Dietterich, T. G., Becker, S., and Ghahramani, Z., editors, *Advances in Neural Information Processing Systems 14*, Cambridge, MA. MIT Press.

[Schölkopf and Smola, 2001] Schölkopf, B. and Smola, A. J. (2001). *Learning with Kernels : Support Vector Machines, Regularization, Optimization, and Beyond*. 1 edition.

[Shawe-Taylor and Cristianini, 2004] Shawe-Taylor, J. and Cristianini, N. (2004). *Kernel Methods for Pattern Analysis*. Cambridge University Press.

King's Abdullah University of
Science and Technology
(Saudi Arabia).
E-mail: abla.kammoun@kaust.edu.sa

CentraleSupélec – LSS –
Université ParisSud (Gif-sur-Yvette, France).
E-mail: romain.couillet@centralesupelec.fr