Electronic Journal of Statistics
Vol. 17 (2023) 291–390
ISSN: 1935-7524
https://doi.org/10.1214/23-EJS2107

# Covariance discriminative power of kernel clustering methods*

**Abla Kammoun**[1] **and Romain Couillet**[2]

[1]*King Abdullah University of Science and Technology (Saudi Arabia)*
*e-mail:* abla.kammoun@kaust.edu.sa

[2]*CentraleSupélec – LSS – Université ParisSud (Gif-sur-Yvette, France)*
*e-mail:* romain.couillet@centralesupelec.fr

**Abstract:** 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 the 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 like 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 of $\tilde{\Phi}$ and their associated eigenvectors and illustrate how these results can help understand spectral clustering methods that use $\tilde{\Phi}$ as a kernel matrix.

**MSC2020 subject classifications:** Primary 15B52, 15A18; secondary 60F15.
**Keywords and phrases:** Kernel random matrices, clustering, machine learning, random matrix theory, multivariate statistical analysis, Gaussian calculus, Stieltjes transform.

Received November 2021.

## Contents



*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.

## 1. Introduction

**Motivation.** Kernel methods play a central role in statistical machine learning. They have extensively been used in many problems such as classification, clustering, regression, as well as principal component analysis and have shown to exhibit better performances than traditional statistical techniques [18, 20]. At the core of these methods is the notion of kernel matrices, constructed 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 type of these kernel matrices include inner-product kernel random matrices obtained by selecting function $k$ as $k(x_i, x_j) = f(x_i^T x_j)$ where $f$ is a given real-valued function. Kernel methods operate exclusively with the kernel matrix, be it by computing its principal eigenvectors like in kernel clustering [17] or by solving a convex problem involving it, as in support vector machine algorithms [14]. A recent line of research works has been concerned with studying the properties of large random kernel matrices when the dimension of the data and the sample size get large and are commensurable. It follows on the important wave of research focusing on the study of large sample covariance matrices given by $\frac{1}{n}\sum_{i=1}^p x_i x_i^T$ which have been the interest of several generations of mathematicians. For more details, the reader can refer to [8] and the references therein.

First works analyzing the spectrum of kernel random matrices were due to El Karoui [11, 10], who was interested in the kernel random matrices of the form $\left\{f(x_i^T x_j)\right\}_{i,j=1}^n$ where $\{x_i\}_{i=1}^n$ are zero-mean independent and identically distributed random vectors with covariance $\frac{1}{p}C$, $C$ being a $p \times p$ matrix of bounded spectral norm. More precisely, it was proven under the asymptotic setting in which the number of samples $n$ and that of features $p$ grow large and converge to a constant, that kernel random matrices are equivalent (up to some additive deterministic matrix and proper scaling) to the standard large sample covariance matrix, extensively studied in the literature. The key idea in the work of [11] lies in the observation that the kernel function is applied entry-wise to the random variable $x_i^T x_j$, which converges fast to its mean. The behavior of the kernel random matrix is then characterized by applying a Taylor-expansion of each element around its mean. However, practical machine learning applications like clustering call for more involved models, among which is the Gaussian mixture model. Based on this model, and following the same Taylor-expansion approach of [11], it was proven in [7] and subsequently in [1] that the underlying kernel random matrix behaves as a "spiked random model", that is a finite rank perturbation of (a matrix similar) to a Wishart random matrix model [4]. The work in [7] provides valuable insights into the impact of the kernel function and the data model on the clustering performance, unveiling the sufficient growth rates in the distances between covariances and means to ensure non-trivial clustering performance. More particularly, assume that data are drawn from a mixture of $c$ Gaussian distributions associated with class $\mathcal{C}_1, \cdots, \mathcal{C}_c$ such that data samples from class $\mathcal{C}_a$ follow a Gaussian distribution with zero mean and covariance $\frac{1}{p}C_a$. It can be easily seen that, under the assumption that all covariance matrices have uniformly bounded spectral norms, the off-diagonal elements of the kernel random matrix $K := \left\{f(x_i^T x_j)\right\}_{i,j=1}^n$ can be

Taylor-expanded as:

$$f(x_i^T x_j) = f(0)+f'(0)x_i^T x_j+\frac{f''(0)}{2}(x_i^T x_j)^2+O(p^{-\frac{3}{2}}) \ \ i,j=1,\cdots,n \ \ i\neq j \quad (2)$$

while the diagonal elements can be expressed as:

$$f(x_i^T x_i) = f(\frac{1}{p}\mathrm{Tr}(C_a)) + O(p^{-\frac{1}{2}}) \quad (3)$$

Clearly, if for $a \neq b$, $\frac{1}{p}\mathrm{Tr}(C_a) - \frac{1}{p}\mathrm{Tr}(C_b) = O(1)$, perfect clustering can be performed without invoking spectral clustering. Indeed, it suffices, in this case, to investigate the diagonal elements of $K$ as they would tend to different limits reflecting the class to which each observation belongs. From this, it is clear that the use of spectral clustering becomes relevant only when for all $a = 1,\cdots,c$, $\frac{1}{p}\mathrm{Tr}(C_a)$ are asymptotically the same. In this case, all diagonal elements tend to the same limits, and as such, they cannot be used to perform clustering. Let $C^\circ = \sum_{i=1}^c \frac{n_i}{n}C_i$ where for $i=1,\cdots,c$, $n_i$ refers to the number of observations in class $\mathcal{C}_i$ and define for $a=1,\cdots,c$, matrix $C_a^\circ$ as: $C_a^\circ = C_a - C^\circ$. Based on the above discussion and in order to not alter the convergence rate in (3), it is thus sensible to assume that $\frac{1}{p}\mathrm{Tr}(C_a) - \frac{1}{p}\mathrm{Tr}(C^\circ) = \frac{1}{p}\mathrm{Tr}(C_a^\circ) = O(p^{-\frac{1}{2}})$. Combining (2) and (3) together, and letting $\tau = \frac{1}{p}\mathrm{Tr}(C^\circ)$, we can easily deduce that the kernel random matrix $K$ is asymptotically equivalent to $\overline{K}_g$ given by [1]:

$$\overline{K}_g = f(0)11^T+(f(\tau)-f(0)-\tau f'(0))I_n+f'(0)X^TX+\frac{f''(0)}{2}\left\{\boldsymbol{\delta}_{i\neq j}(x_i^Tx_j)^2\right\}_{i,j=1}^n \quad (4)$$

where $X=[x_1,\cdots,x_n]$. In [8], noticing that the spectral norm

$$\left\{\boldsymbol{\delta}_{i\neq j}(x_i^Tx_j)^2\right\}_{i,j=1}^n - \left\{\frac{1}{p^2}\mathrm{Tr}(C_aC_b)1_{n_a}1_{n_b}\right\}_{a,b=1}^c$$

is $O(p^{-\frac{1}{2}})$, the authors argued that the matrix $\left\{\boldsymbol{\delta}_{i\neq j}(x_i^Tx_j)^2\right\}_{i,j=1}^n$ contains the necessary information to perform clustering, while the matrix $f'(0)X^TX$ represents the noise part. On the lookout for better performances, one is tempted to cancel the noise by choosing $f$ such that $f'(0)=0$ and $f''(0)\neq 0$. In this case, the equivalent matrix $\overline{K}_g$ satisfies

$$\overline{K}_g = f(0)11^T+(f(\tau)-f(0))I_n+\frac{f''(0)}{2}\left\{\frac{1}{p^2}\mathrm{Tr}(C_aC_b)1_{n_a}1_{n_b}\right\}_{a,b=1}^c+O_{\|.\|}(p^{-\frac{1}{2}}), \quad (5)$$

where $O_{\|.\|}(p^{-\frac{1}{2}})$ represents a matrix with a spectral norm of order $O(p^{-\frac{1}{2}})$. Equation (5) reveals that, under this particular choice of function $f$, matrix $\overline{K}_g$ is up to a vanishing matrix deterministic, which suggests the possibility

[1] The asymptotic equivalence herein is in the sense that $\|K-\overline{K}_g\|$ tends to zero as $n$ and $p$ grow large with $\frac{n}{p}\to c_0$

of perfectly clustering observations in the asymptotic regime. However, this conclusion is not guaranteed to hold if we further assume that for all $a, b \in \{1, \cdots, c\}$, $\frac{1}{p}\mathrm{Tr}(C_a^\circ C_b) = O(p^{-\frac{1}{2}})$. To see this, it suffices to expand $\frac{1}{p}\mathrm{Tr}(C_a C_b)$ as:

$$\frac{1}{p}\mathrm{Tr}(C_a C_b) = \frac{1}{p}\mathrm{Tr}(C_a^\circ C^\circ) + \frac{1}{p}\mathrm{Tr}(C_b^\circ C^\circ) + \frac{1}{p}\mathrm{Tr}(C_a^\circ C_b^\circ) + \frac{1}{p}\mathrm{Tr}((C^\circ)^2)$$

One can easily check that under the assumption $\frac{1}{p}\mathrm{Tr}(C_a^\circ C_b) = O(p^{-\frac{1}{2}})$

$$\left\| \frac{f''(0)}{2} \left\{ \boldsymbol{\delta}_{a\neq b} \left( \frac{1}{p^2}\mathrm{Tr}(C_a C_b) - \frac{1}{p^2}\mathrm{Tr}((C^\circ)^2) \right) 1_{n_a} 1_{n_b}^T \right\}_{a,b=1}^{c} \right\| = O(p^{-\frac{1}{2}}) \quad (6)$$

where for a square matrix $A$, $\|A\|$ stands for its spectral norm. Going back to (4), it becomes clear that in case $\frac{1}{p}\mathrm{Tr}(C_a^\circ C_b) = O(p^{-\frac{1}{2}})$, spectral clustering based on $K$ does not perform better than random guess clustering if $f$ is such that $f'(0) \neq 0$, since $K$ would be equivalent (up to a non-informative matrix) to the noise part $f'(0)X^T X$. On the other hand, selecting $f'(0) = 0$ cancels out the noise term, but $\overline{K}_g$ becomes equivalent (up to a vanishing matrix with spectral norm $O(p^{-\frac{1}{2}})$) to a deterministic non-informative matrix. The findings of [7] could not inform of whether clustering can be performed based on the inner-product kernel random matrix $K$. The present work aims to fill this gap. In view of (4), the answer to this question boils down to analyzing the clustering task using the following random matrix:

$$\tilde{\Phi} = \sqrt{p} \left\{ (x_k^T x_l)^2 \boldsymbol{\delta}_{k\neq l} \right\}_{k,l=1}^{n} \quad (7)$$

where the factor $\sqrt{p}$ aims to apply a kind of a "zoom" on the vanishing perturbation matrix expected to carry information about clustering.

**Contributions and summary of the obtained results.** This paper is concerned with the problem of clustering $n$ observations $x_1, \cdots, x_n$ drawn from a Gaussian mixture model with $c$ classes, in which observations from cluster $k, k = 1, \cdots, c$ follow a Gaussian distribution with zero mean and covariance $\frac{1}{p}C_k$. Of interest is the asymptotic setting in which the total number of samples $n$, that of samples in each class $n_a, a = 1, \cdots c$ and the number of features $p$ grow large while their ratios $\frac{n_a}{n}$ and $\frac{n}{p}$ are equal to constants $c_a > 0$ and $c_0 > 0$, respectively. We further assume that the covariance matrices satisfy $\frac{1}{p}\mathrm{Tr}(C_a^\circ C_b) = O(p^{-\frac{1}{2}})$, which as earlier mentioned, requires selecting the kernel function $f$ such that $f'(0) = 0$. Under this growth regime, and as shown above, the eigenvectors of the kernel matrix $K = \left\{ f(x_i^T x_j) \right\}_{i,j=1}^{n}$ are informative for clustering tasks if those of matrix $\tilde{\Phi}$ are also informative. Based on this observation, we redefine our task to that of analyzing the leading eigenvectors and eigenvalues of matrix $\tilde{\Phi}$ in (7). Decomposing $\tilde{\Phi}$ as:

$$\tilde{\Phi} = \Phi + \sqrt{p} \left\{ \mathbb{E}\left[ (x_i^T x_j)^2 \boldsymbol{\delta}_{i\neq j} \right] \right\}_{i,j=1}^{n}$$

where

$$\Phi = \sqrt{p}\left\{(x_i^T x_j)^2 \boldsymbol{\delta}_{i\neq j}\right\} - \sqrt{p}\left\{\mathbb{E}\left[(x_i^T x_j)^2 \boldsymbol{\delta}_{i\neq j}\right]\right\}_{i,j=1}^n$$

we prove that $\tilde{\Phi}$ is a sort of a spiked random matrix with matrix

$$\sqrt{p}\left\{\mathbb{E}\left[(x_i^T x_j)^2 \boldsymbol{\delta}_{i\neq j}\right]\right\}_{i,j=1}^n$$

playing the role of the finite-rank perturbation while $\Phi$ stands for the high-rank random matrix. A major result of the present work is to show that $\Phi$ behaves as a standard Wigner matrix presenting possibly isolated eigenvalues that escape from the bulk. However, these isolated eigenvalues do not carry information about clustering. The clustering information is indeed carried by the isolated eigenvalues of matrix $\tilde{\Phi}$ and their associated eigenvectors that are produced by the presence of the finite-rank perturbation matrix $\sqrt{p}\left\{\mathbb{E}\left[(x_i^T x_j)^2 \boldsymbol{\delta}_{i\neq j}\right]\right\}_{i,j=1}^n$. In a nutshell, our contributions can be summarized as follows:

- We show (Theorem 1) that in the asymptotic regime wherein $n, p$ grow to infinity with $\frac{n}{p} = c_0$ and $\frac{n_a}{n} = c_a$, the matrix $\Phi$ 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 [6], which asserts that the asymptotic behavior will involve only the contribution of a Wigner matrix once $f'(0) = 0$. Note that the result in [6] 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 [6] and mainly relies on Gaussian calculus tools as the basic instruments.
- We analyze the asymptotic behavior of bilinear forms involving the resolvent of $\Phi$ (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 $\tilde{\Phi}$.

**Related works.** This work, initially motivated by the previous work of [7], fits in the recent line of research aiming at analyzing kernel random matrices with elements

$$K_{ij} = \frac{1}{\sqrt{p}} f(\sqrt{p} x_i^T x_j)\boldsymbol{\delta}_{i\neq j}\ . \tag{8}$$

Indeed, matrix $\tilde{\Phi}$ can be thought of as a specific instance of the class of kernel matrices in the form of (8) with $f = x^2$. Compared to the kernel random matrices studied in [11], the multiplication by $\sqrt{p}$ inside function $f$ produces fluctuating off-diagonal elements. As a consequence, the Taylor-expansion method originally developed in [11] and later generalized in [15] for uniform distribution over balls and in [7] for Gaussian mixture models is not applicable. In a series of recent works in [6], [12], and [9], new tools have been developed to study the behavior of kernel random matrices in the form of (8). Contrary to the kernel random matrices studied in [11], a completely different behavior has been unveiled, according to which kernel random matrices following the model (8) behave as deformed Wigner-like matrices. However, all these results concern observations with isotropic covariance structure and hence could not be applied to understand the performance of kernel clustering methods. Compared to these works, our contribution differs as follows. $(i)$ First, we study the behavior of kernel random matrices in (8) for $f = x^2$ and when data are drawn from a Gaussian mixture model with $c$ classes. As discussed earlier, the choice $f = x^2$ aims at examining the "covariance discriminative power" of inner-product kernel random matrices $K = \left\{f(x_i^T x_j)\right\}_{i,j=1}^n$, for which, to cancel the noise, $f$ is selected such that $f'(0) = 0$. $(ii)$ Second, we characterize both the eigenvalues and the leading eigenvectors of $\tilde{\Phi}$, which allows us to gain a deeper understanding of the clustering performance. On a technical level, we develop a new approach that combines both Gaussian calculus and the Stieltjes transform tool. We believe that this approach can provide the underpinning for a unified theoretical framework to analyze general inner-product kernel random matrices in the form of (8).

*Notations:* In the remainder of the article, uppercase characters will stand for matrices, lowercase for scalars or vectors. The transpose and hermitian operation will be denoted by $(.)^T$ or $(.)^H$. The multivariate Gaussian distribution of mean $\mu$ and covariance $C$ will be denoted $\mathcal{N}(\mu, C)$. The notation $V = \{V_{ij}\}_{i=1,j=1}^{n,T}$ 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. The $i$-th element of vector $v$ is denoted by $[v]_i$, while the $(i,j)$-th entry of matrix $A$ may be denoted by either $A_{ij}$ or $[A]_{ij}$. The operator $\mathcal{D}(v) = \mathcal{D}\left\{v_a\right\}_{a=1}^k$ is the $k \times k$ diagonal matrix with $v_1, \ldots, v_k$ as its diagonal elements. When $A$ is a matrix, the operator $\mathcal{D}(A)$ refers to the diagonal matrix formed by the diagonal elements of $A$. The identity matrix of size $p$ is denoted by $I_p$ while the vector in $n \times 1$ of all ones is denoted by $1_n$. The notation $\|.\|$ refers to either the Euclidean norm of vectors or the operator norm of matrices while the notation $\|.\|_\infty$ refers to the $\ell_\infty$ norm for vectors. For $A$ and $B$ matrices with the same size, we denote the Hadamard product of $A$ and $B$ by $A \odot B$. 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 a sequence of random variables $x_p$, the notation $x_p = O(r_p)$ means that for every $\eta$ and $D$ strictly positives, $p^D \mathbb{P}\left[x_p \geq p^\eta r_p\right] \to 0$. We also define $\boldsymbol{\delta}_A$ as the indicator function of set $A$. For deterministic scalars $x_p$ and $v_p$ the notation $x_p = O_z(v_p)$ means that $|x_p| \leq v_p P(|z|) R(|\Im z|^{-1})$ for some polynomials $P$ and $R$ with non-negative coefficients and whose parameters are independent of the dimensions $n$ and $p$. Finally, we denote by $\mathbb{E}_k$, the expectation with respect to vector $x_k$ and by $\mathbf{var}_k$ its corresponding variance.

## 2. Assumptions and main results

Consider $p$-dimensional independent real Gaussian vectors $x_1, \dots, x_n$. For $n_1, \dots, n_c$ such that $n_1 + \dots + n_c = n$, we assume that

$$x_{n_1+\dots+n_{j-1}+1}, \dots, x_{n_1+\dots+n_j} \sim \mathcal{N}(0, p^{-1}C_j)$$

for $C_1, \dots, C_c \in \mathbb{R}^{p\times p}$.

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 $C_{[k]}$ as $C_{[k]} = C_j$. In other words, $C_{[k]}$ denotes the covariance matrix of observation $x_k$. Hence, observations $x_k$ can be written as:

$$x_k = \frac{1}{\sqrt{p}} C_{[k]}^{\frac{1}{2}} z_k$$

where $z_k \sim \mathcal{N}(0, I_p)$

Further, define $C^\circ = \sum_{i=1}^{c} \frac{n_i}{n} C_i$ and, for each $i$, $C_i^\circ = C_i - C^\circ$. The matrices $C_1, \dots, C_c$ 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* $a \in \{1, \dots, c\}$, $n_a/n = c_a \in (0, \infty)$
*(iii)* $\forall\ a, b \in \{1, \cdots, c\}$, $\frac{1}{p} \operatorname{tr} C_a^\circ C_b = O(p^{-\frac{1}{2}})$
*(iv) all matrices* $C_k$, $k = 1, \cdots, c$ *have bounded spectral norm, that is:*

$$\max_{1\le k\le c} \limsup_{p} \|C_k\| < \infty.$$

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

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

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

*Moreover, we assume that:*

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

*for some constant $K$ independent of $p$.*

As a direct consequence of Item *iii)* in Assumption 1, we have:

$$\frac{1}{p} \operatorname{tr} C_a^\circ C^\circ = O(p^{-\frac{1}{2}}) \tag{9}$$

$$\frac{1}{p} \operatorname{tr} C_a^\circ C_b^\circ = O(p^{-\frac{1}{2}}). \tag{10}$$

To prove (9), it suffices to replace $C^\circ$ by $\sum_{i=1}^{c}\frac{n_i}{n}C_i$ and then apply item *iii*) to each element of the obtained sum. Similarly, (10) can be proven by substituting $C_b^\circ$ by $C_b - C^\circ$ and using (9) together with item *iii*) in Assumption 1. Moreover, it also holds that for any sequence of $p\times p$ matrices $(A_p)$ with uniformly bounded spectral norm, we have:

$$\frac{1}{p}\operatorname{tr}(C_a^\circ A_p) = O(p^{-\frac{1}{4}}) \tag{11}$$

which can be directly shown by noting that

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

A possible choice of covariance matrices that satisfy Assumption 1-*iii*) is when for instance, $C_a - C_b$ has rank $\sqrt{p}$ for any $a \neq b$ in $\{1,\dots,c\}$ or when it has $O(\sqrt{p})$ eigenvalues of order 1 and the others converging to zero. In this case, it can be easily seen that $\frac{1}{p}\operatorname{tr} C_a^\circ C_b = O(p^{-\frac{1}{2}})$ and $\frac{1}{p}\operatorname{tr} C_a^\circ C_b^\circ = O(p^{-\frac{1}{2}})$.

This paper is concerned with the clustering task using the kernel:

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

Particularly, we aim to determine the conditions under which clustering using $\tilde{\Phi}$ leads to a non-trivial behavior in the growth rate regime of Assumption 1. As explained earlier, these conditions also imply the non-trivial behavior of the clustering performance using the kernel matrix $K = \left\{f(x_i^T x_j)\right\}_{i,j=1}^n$ with $f'(0) = 0$.

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 $\Phi$ obtained by element-wise centering of $\tilde{\Phi}$:

$$[\Phi]_{ij} = \sqrt{p}\left((x_i^T x_j)^2 - \frac{1}{p^2}\operatorname{tr} C_{[i]}C_{[j]}\right)\boldsymbol{\delta}_{i\neq j} \tag{12}$$

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

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

These results set the stage for the second part of our work (section 3), in which we study the behavior of spectral clustering using 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 $\Phi$ by:

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

and the Stieltjes transform of the expectation of the empirical measure of the eigenvalues of $\Phi$ 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, which is to prove the convergence of the empirical distribution of matrix $\Phi$ to the semi-circle distribution. The latter result is formally stated in the following Theorem, the proof of which is postponed to Section 5.1.

**Theorem 1.** *Let Assumption 1 hold true. Denote by $\lambda_1, \cdots, \lambda_n$ the eigenvalues of $\Phi$. 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 $\mu$ with density:*

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

*where $\boldsymbol{\delta}_A$ is the indicator function of set A. Moreover, the support of the limiting density is $\mathcal{S} = \left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right]$.*

Theorem 1 can be leveraged to approximate in the almost sure sense functionals of the eigenvalues of matrix $\Phi$ 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 $\Phi$. However, it cannot be used to infer that of the eigenvectors. As shall be seen next, a key step in analyzing the asymptotic behavior of the eigenvectors of $\Phi$ (or a perturbation of it) is to characterize the asymptotic behavior of bilinear forms associated with the resolvent matrix in the form of $a_n^T Q(z) b_n$ where $a_n$ and $b_n$ are two vectors in $\mathbb{C}^{n\times 1}$. This is the purpose of the following Theorem, the proof of which is deferred to Section 5.2.

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

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

*where here $1_n$ is the $n \times 1$ vector of all ones and thus $1_n 1_n^T$ is the $n \times n$ matrix of all ones.*

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

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

where the equivalence is in the sense that $\frac{1}{n}\operatorname{tr} A_n X_n - \frac{1}{n}\operatorname{tr} A_n Y_n \to 0$ and $a_n^T(X_n - Y_n) b_n \to 0$ for every sequence of deterministic matrices $A_n$ with bounded spectral norm and sequence of vectors $a_n, b_n$ having bounded Euclidean norms. In other words, by reference with this definition, matrix $Q$ is equivalent to a rank-one perturbation of a scaled identity. As far as classical random matrix models are considered, this behavior is met when the random matrix under study is modeled by a finite-rank perturbation of a high-rank random matrix. Based on this, we can anticipate that matrix $\Phi$ may possess a spike outside the support of the semi-circle law, which explains the presence of the rank one matrix $\frac{c_0 \Omega^2 m^3(z)}{p(1-\Omega^2 c_0^2 m^2(z))} 1_n 1_n^T$. 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 Theorem 4, which confirms the possible existence of spikes outside the main bulk of the semi-circle law.

But before moving to Theorem 4, we shall present the following result concerning the asymptotic limit of quadratic forms involving two resolvent matrices.

**Theorem 3.** *Consider the setting of Assumption 1. Let $\{a_n\}$ and $\{b_n\}$ be two sequences of deterministic vectors in $\mathbb{C}^{n\times 1}$ with bounded Euclidean norm. Let $D_n$ be a $n \times n$ sequence of diagonal matrices with uniformly bounded diagonal elements. Define for $z_1, z_2 \in \mathbb{C}\backslash\mathbb{R}$, $g(z_1, z_2)$ as:*

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

*Then,*

$$\begin{aligned}
&a_n^T Q(z_1) D_n Q(z_2) b_n - m(z_1) m(z_2) a_n^T D_n b_n \\
&\qquad - m(z_1) m(z_2) \omega^2 \Big(\frac{1}{p}\operatorname{tr}(D)\Big) g(z_1,z_2) - \tilde{r}(z_1,z_2) \xrightarrow{\text{a.s.}} 0.
\end{aligned}$$

*where:*

$$\tilde{r}(z_1,z_2)=\frac{c_0\Omega^2 m^3(z_2)m(z_1)a^T D\frac{1_n1_n^T}{p}b}{1-\Omega^2c_0^2m^2(z_2)}+\frac{c_0\Omega^2 m^3(z_1)m(z_2)a^T\frac{1_n1_n^T}{p}Db}{1-\Omega^2c_0^2m^2(z_1)}$$
$$+\frac{\frac{1}{p}\operatorname{tr}(D)\Omega^2m^2(z_1)m^2(z_2)(1+c_0^2\Omega^2m(z_1)m(z_2))a^T\frac{1_n1_n^T}{p}b}{\left(1-\Omega^2c_0^2m^2(z_1)\right)\left(1-\Omega^2c_0^2m^2(z_2)\right)}$$

*Proof.* See Section 5.3 □

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

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

*Then, with probability* 1 *for all large $n$:*

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

*Proof.* See Section 5.4. □

**Remark 1.** From Cauchy-Schwartz inequality, it follows that $\sqrt{\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)}\geq\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)$ or equivalently $\Omega\geq\omega$. Hence, if $c_0\geq1$, $\Omega\geq\frac{\omega}{\sqrt{c_0}}$. In such a case, we expect at least two spikes at positions $\tilde{\rho}$ and $-\tilde{\rho}$ escaping 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 [12, Theorem 1.7], which shows that under the i.i.d. case with all $C_k$'s equal to the identity matrix, polynomial kernel matrices might have two spikes outside the main bulk of the semi-circle law. Although restricted to $f(x)=x^2$, our work extends the result in [12] to Gaussian mixture models.

**Remark 3.** The convergence results in Theorem 2 and Theorem 3 can be easily extended to all $z\in\mathcal{I}:=\mathbb{C}\backslash\left([-2\sqrt{c_0}\omega,2\sqrt{c_0}\omega]\cup\{-\tilde{\rho},\tilde{\rho}\}\right)$ using standard arguments based on Montel's theorem. Particularly, for Theorem 2, the recipe is as follows. The random quantity $a_n^TQ(z)b_n$ and its deterministic equivalent in Theorem 2 are analytic and bounded on any compact support of $\mathcal{I}$. It follows from Montel's theorem that there exists a subsequence for which $a_n^TQ_n(z)b_n-a_n^Tb_nm(z)-\frac{c_0\Omega^2m^3(z)}{p(1-\Omega^2c_0^2m^2(z))}a_n^T1_n1_n^Tb_n$ converges to a holomorphic function on each compact set of $\mathbb{C}\backslash\mathcal{I}$. Since this limiting function is zero for all compacts

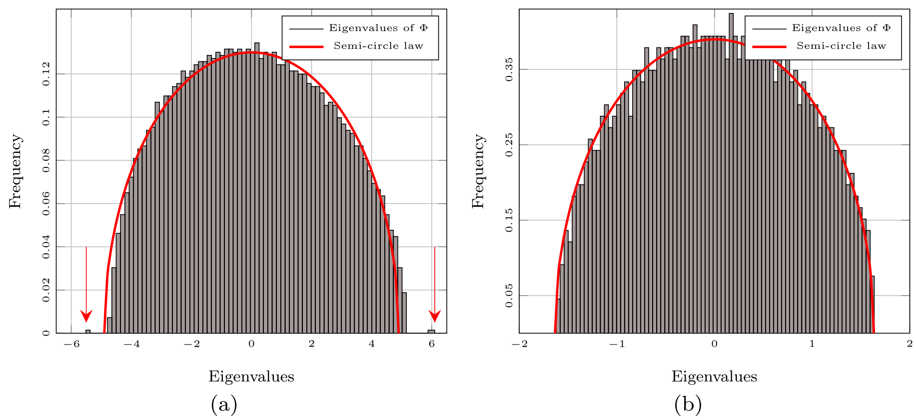


FIG 1. *Histogram of the eigenvalues of* $\Phi$ *and the semi-circle law, for (a)* $n = 4800$ *and* $p = 1600$ *and (b)* $n = 1600$ *and* $p = 4800$*. All* $C_i$*'s are equal to* $I_p$*. The semi-circle law is superposed in red, and the locations of two observed spikes is highlighted with red arrows.*

of $\mathbb{C}\backslash\mathbb{R}$, it is thus also zero for all $z \in \mathbb{C}\backslash\mathcal{I}$. Thus, the convergence of Theorem 2 holds for all $z \in \mathbb{C}\backslash\mathcal{I}$. The same argument is valid to extend the convergence result of Theorem 3 to all $z \in \mathcal{I}$.

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

In this section, we show how the previous results can be leveraged to gain a better understanding of the performance of spectral kernel clustering based on the matrix:

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

To begin with, we decompose $\tilde{\Phi}$ as:

$$\tilde{\Phi} = \Phi + \left\{\frac{1}{p^{\frac{3}{2}}} \operatorname{tr} C_{[i]} C_{[j]} \boldsymbol{\delta}_{i\neq j}\right\}_{i,j=1}^n$$

The behavior of matrix $\Phi$ is studied in the previous section, where we showed that $\Phi$ behaves like a Wigner matrix plus possibly a one-rank perturbation. It is thus not informative from a clustering perspective; the information about clustering in $\tilde{\Phi}$ is rather carried by the finite rank matrix $\left\{\frac{1}{p^{\frac{3}{2}}} \operatorname{tr} C_{[i]} C_{[j]} \boldsymbol{\delta}_{i\neq j}\right\}_{i,j=1}^n$.

To continue, we note that

$$\left\{\frac{1}{p^{\frac{3}{2}}} \operatorname{tr} C_{[i]} C_{[j]} \boldsymbol{\delta}_{i\neq j}\right\}_{i,j=1}^n = \left\{\frac{1}{p^{\frac{3}{2}}} \operatorname{tr} C_{[i]} C_{[j]}\right\}_{i,j=1}^n + O_{\|.\|}(p^{-\frac{1}{2}}) \qquad (13)$$

where $O_{\|.\|}(p^{-\frac{1}{2}})$ refers to a matrix with $O(p^{-\frac{1}{2}})$ spectral norm. Next, for $a, b = 1, \cdots, c$, decomposing $\frac{1}{p^{\frac{3}{2}}} \operatorname{tr}(C_a C_b)$ as:

$$\frac{1}{p^{\frac{3}{2}}} \operatorname{tr}(C_a C_b) = \frac{1}{p^{\frac{3}{2}}} \operatorname{tr}(C_a^\circ C_b^\circ) + \frac{1}{p^{\frac{3}{2}}} \operatorname{tr}(C_a^\circ C^\circ) + \frac{1}{p^{\frac{3}{2}}} \operatorname{tr}(C_b^\circ C^\circ) + \frac{1}{p^{\frac{3}{2}}} \operatorname{tr}((C^\circ)^2)$$

we can easily see that matrix $\mathcal{M}$ defined as:

$$\mathcal{M} := \left\{ \frac{1}{p^{\frac{3}{2}}} \operatorname{tr}(C_a C_b) 1_{n_a} 1_{n_b}^T \right\}_{a,b=1}^{c}$$

can be written in matrix form as follows:

$$\mathcal{M} = \frac{1}{p} J A J^T + \frac{1}{p} J a 1_n^T + \frac{1}{p} 1_n a^T J^T + \beta \frac{1_n 1_n^T}{p} \tag{14}$$

where $J = [j_1, \ldots, j_c] \in \mathbb{R}^{n \times c}$, with $j_i$ being the canonical vector of class $i$, taking 1 at the positions in which $x_j$ belongs to class $i$ and zero otherwise and $A$, $a$ and $\beta$ are given by:

$$\begin{aligned} a &:= \left\{ \frac{1}{\sqrt{p}} \operatorname{tr} C_i^\circ C^\circ \right\}_{i=1}^{c} \\ \beta &:= \frac{1}{\sqrt{p}} \operatorname{tr} \left( (C^\circ)^2 \right) \\ A &:= \left\{ \frac{1}{\sqrt{p}} \operatorname{tr} C_i^\circ C_j^\circ \right\}_{i,j=1}^{c}, \end{aligned} \tag{15}$$

Then, in view of (13), we obtain:

$$\tilde{\Phi} = \overline{\Phi} + O_{\|.\|}(p^{-\frac{1}{2}}) \tag{16}$$

where $\overline{\Phi}$ writes as:

$$\overline{\Phi} := \Phi + \mathcal{M} \tag{17}$$

**Behavior of the eigenvalues.** Matrix $\overline{\Phi}$ follows a spiked random model perturbed by a finite rank deterministic matrix. From standard results of random matrix theory, we expect all its eigenvalues to be located asymptotically within $\mathcal{S}_\epsilon$ except for finitely many of isolated eigenvalues escaping from $\mathcal{S}_\epsilon$. Determining the almost sure location of such eigenvalues is of fundamental importance to understand how kernel clustering works in high-dimensional settings. Since $\beta 1_n 1_n^T$ has only one non-zero eigenvalue of order $\sqrt{p}$, by the Weyl's inequalities, the $n-1$ smallest eigenvalues of $\overline{\Phi}$ are located almost surely in a compact interval, satisfying

$$\lambda_i(\overline{\Phi}) \leq \|\Phi\|_2 + \left\| \frac{1}{p} J A J^T + \frac{1}{p} J a 1_n^T + \frac{1}{p} 1_n a^T J^T \right\|_2, \quad i = 1, \cdots, n-1$$

while the largest eigenvalue cannot be bounded. The following theorem characterizes the location of the eigenvalues of $\tilde{\Phi}$ that escape from $\mathcal{S}_\epsilon$.

**Theorem 5.** *Let $\kappa$, $\lambda_1 \cdots, \lambda_{c-1}$ be the $c$ largest eigenvalues of $\Phi$ such that $\kappa \geq \lambda_1 \geq \ldots \geq \lambda_{c-1}$. Then, there exist $K_1$ and $K_2$ deterministic constants such that:*

$$K_1\sqrt{p} \leq \kappa \leq K_2\sqrt{p}$$

*Moreover,*

$$\frac{\kappa}{\sqrt{p}} - \frac{\beta c_0}{\sqrt{p}} \xrightarrow{\text{a.s.}} 0.$$

*Assume that $\frac{1}{\sqrt{p}} \operatorname{tr} C_a^\circ C_b^\circ$ converges and let:*

$$\mathcal{T} := \left\{ \sqrt{c_a}\sqrt{c_b} \lim_{n\to\infty} \frac{1}{\sqrt{p}} \operatorname{tr} C_a^\circ C_b^\circ \right\}_{a,b=1}^{c}$$

*Denote by $\nu_1 \geq \nu_2 \geq \cdots \geq \nu_{c-1}$ the $c-1$ largest eigenvalue of $\mathcal{T}$. Then, under Assumption 1, if for $i \in \{1, \cdots, c\}$, $\nu_i > \frac{\omega}{\sqrt{c_0}}$ and $\nu_i \notin \{\Omega, -\Omega\}$, then*

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

*Proof.* See Appendix C.1 □

Combining the results of Theorem 4 and Theorem 5, it can be seen that the eigenvalues of $\tilde{\Phi}$ that converge to values outside the support of the semi-circular law are informative except possibly for eigenvalues converging to $\tilde{\rho}$ and $-\tilde{\rho}$ which appear only when $\Omega \geq \frac{\omega}{\sqrt{c_0}}$. Thus, in practice, it is important to get estimates of $\Omega$ and $\omega$ in order to: 1) estimate the support of the semi-circular law, 2) determine if non-informative eigenvalues converging to $\tilde{\rho}$ and $-\tilde{\rho}$ may appear. The values of $\omega$ and $\Omega$ depend on the unknown covariances. However, as can be seen below, they can be estimated based on the observations from all classes. Indeed, we may use the Poincaré-Nash inequality stated later in (31) to prove that:

$$\operatorname{var}\left( \frac{1}{n} \sum_{i=1} \sum_{\substack{j=1 \\ j\neq i}}^{n} (x_i^T x_j)^2 \right) = O(\frac{1}{p^2}) \tag{18}$$

$$\operatorname{var}\left( \frac{1}{n} \sum_{\substack{i,j,k,m \\ i\neq j\neq k\neq m}}^{n} x_i^T x_j x_j^T x_k x_k^T x_m x_m^T x_i \right) = O(\frac{1}{p^2}) \tag{19}$$

The proof of the above results follows from very similar derivations used in the paper for variance control and is thus omitted. It follows from (18) and (19) that:

$$\frac{1}{n} \sum_{i=1} \sum_{\substack{j=1 \\ j\neq i}}^{n} (x_i^T x_j)^2 - \frac{1}{n} \sum_{i=1} \sum_{\substack{j=1 \\ j\neq i}}^{n} \mathbb{E}(x_i^T x_j)^2 \xrightarrow{\text{a.s.}} 0 \tag{20}$$

and

$$\frac{1}{n}\sum_{\substack{i,j,k,m\\ i\neq j\neq k\neq m}} x_i^T x_j x_j^T x_k x_k^T x_m x_m^T x_i - \frac{1}{n}\sum_{\substack{i,j,k,m\\ i\neq j\neq k\neq m}} \mathbb{E}[x_i^T x_j x_j^T x_k x_k^T x_m x_m^T x_i] \xrightarrow{\text{a.s.}} 0 \tag{21}$$

On the other hand, taking the expectation over the distributions of the observations, we obtain:

$$\frac{1}{n}\sum_{i=1}\sum_{\substack{j=1\\ j\neq i}}^{n} \mathbb{E}(x_i^T x_j)^2 = \frac{1}{np}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \frac{1}{p}\mathrm{tr}(C_{[i]}C_{[j]}) \stackrel{(a)}{=} \frac{(n-1)}{p}\frac{1}{p}\mathrm{tr}((C^\circ)^2) + O(p^{-\frac{1}{2}}) \tag{22}$$

$$= c_0\frac{1}{p}\mathrm{tr}((C^\circ)^2) + O(p^{-\frac{1}{2}}) \tag{23}$$

where $(a)$ follows from **Assumption 1**$-(iii)$. Similarly,

$$\begin{aligned}\frac{1}{n}\sum_{\substack{i,j,k,m\\ i\neq j\neq k\neq m}} \mathbb{E}[x_i^T x_j x_j^T x_k x_k^T x_m x_m^T x_i] &= \frac{1}{np^3}\sum_{\substack{i,j,k,m\\ i\neq j\neq k\neq m}} \frac{1}{p}\mathrm{tr}(C_{[i]}C_{[j]}C_{[k]}C_{[m]})\\ &= \frac{(n-1)(n-2)(n-3)\frac{1}{p}\mathrm{tr}((C^\circ)^4)}{p^3} + O(p^{-\frac{1}{4}})\\ &= c_0^3\frac{1}{p}\mathrm{tr}((C^\circ)^4) + O(p^{-\frac{1}{4}})\end{aligned} \tag{24}$$

where (24) follows from (11). Combining (20) and (21), consistent estimators for $\omega$ and $\Omega$ can be obtained as:

$$\hat{\omega} = \frac{\sqrt{2}p}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}(x_i^T x_j)^2 \tag{25}$$

$$\hat{\Omega} = \sqrt{\frac{2p^3}{n^4}\sum_{\substack{i,j,k,m\\ i\neq j\neq k\neq m}} x_i^T x_j x_j^T x_k x_k^T x_m x_m^T x_i} \tag{26}$$

**Behavior of the eigenvectors** The clustering performance depends on the degree of alignment between the eigenvectors of $\tilde{\Phi}$ associated with the isolated eigenvalues and hence referred to as from now on isolated eigenvectors and the columns of $J$. A perfect clustering performance would be obtained in case of perfect alignment. In our case, because of the noise matrix $\Phi$, a perfect alignment does not hold, making the eigenvectors of $\tilde{\Phi}$ fluctuate around the classes' index vectors $j_1, \cdots, j_c$, $j_i$ being the canonical vector of class $i$. Assessing the alignment of isolated eigenvectors of $\tilde{\Phi}$ to these vectors is an important step towards gauging the clustering performance. In the sequel, we will focus only on the eigenvalues of $\tilde{\Phi}$ that converge to one of the $\{\rho_i\}_{i=1}^{c-1}$ defined in Theorem 5.

Let $\hat{u}_\rho$ be an eigenvector of $\tilde{\Phi}$ associated with the isolated eigenvalue converging to $\rho$. Then, $\hat{u}_\rho$ may be decomposed as:

$$\hat{u}_\rho = \sum_{a=1}^{c} \alpha_a^\rho \frac{j_a}{\sqrt{n_a}} + \sigma_a^\rho w_a^\rho$$

where $w_a^\rho$ is a vector of unit norm supported on the indices of class $a$ and orthogonal to $j_a$ while $\alpha_a^\rho \in \mathbb{R}$ and $\sigma_a^\rho \geq 0$ are scalars to be determined. Similarly, for two isolated eigenvalues and bounded eigenvalues of $\tilde{\Phi}$ converging to $\rho_1$ and $\rho_2$, it is of interest to study the correlation:

$$\sigma^a_{\rho_1,\rho_2} = \left(\hat{u}_{\rho_1} - \alpha_a^{\rho_1} \frac{j_a}{\sqrt{n_a}}\right)^T \mathcal{D}(j_a) \left(\hat{u}_{\rho_2} - \alpha_a^{\rho_2} \frac{j_a}{\sqrt{n_a}}\right)$$

where $\mathcal{D}(j_a)$ is the diagonal matrix formed by the entries of the canonical vector $j_a$ of class $a$.

**Theorem 6.** *Consider the setting of Theorem 4. Let $i_1, i_2 \in \{1, \ldots, c-1\}$. Let $(\lambda_{\rho_{i_1}}, \hat{u}_{\rho_{i_1}})$ and $(\lambda_{\rho_{i_2}}, \hat{u}_{\rho_{i_2}})$ be the eigenpairs of $\tilde{\Phi}$ such that $\lambda_{\rho_{i_1}}$ and $\lambda_{\rho_{i_2}}$ converge respectively to $\rho_{i_1}$ and $\rho_{i_2}$. Let $(\nu_{\rho_{i_1}}, V_{\rho_{i_1}})$ and $(\nu_{\rho_{i_2}}, V_{\rho_{i_2}})$ be the eigenpairs of $\mathcal{T}$ associated with $\rho_{i_1}$ and $\rho_{i_2}$ where $\nu_{\rho_{i_1}}$ and $\nu_{\rho_{i_2}}$ have unit multiplicity. Assume that for $k = 1, 2$, $\sqrt{c_0}|\nu_{\rho_{i_k}}| > \omega$, $\nu_{\rho_{i_k}} \notin \{\Omega, -\Omega\}$ with $\tilde{\rho} = c_0\Omega + \frac{\omega^2}{\Omega}$. Then, for any $a, b \in \{1, \ldots, c\}$ and $j = 1, 2$,*

$$\alpha_a^{\rho_{i_j}} \alpha_b^{\rho_{i_j}} \xrightarrow{\text{a.s.}} (1 - \frac{1}{c_0}\frac{\omega^2}{\nu_{\rho_{i_j}}}) \left[V_{\rho_{i_j}} V_{\rho_{i_j}}^T\right]_{a,b}. \tag{27}$$

*Moreover,*

$$\sigma^a_{\rho_{i_1},\rho_{i_2}} \xrightarrow{\text{a.s.}} \boldsymbol{\delta}_{\rho_{i_1}=\rho_{i_2}} \frac{c_a}{c_0}\frac{\omega^2}{\nu^2_{\rho_{i_1}}}. \tag{28}$$

*Proof.* The proof is in Appendix C.2 □

**Remark 4.** (The Largest eigenvector is the most informative) Since $x \mapsto c_0 x + \frac{\omega^2}{x}$ is an increasing function when $x \geq \frac{\omega}{\sqrt{c_0}}$, the largest isolated eigenvalue $\rho_1$ is associated with the largest $\nu_1$ eigenvalue of $\mathcal{T}$. The fluctuations of the entries of this eigenvector around $\alpha_a^{\rho_1}$ are the smallest compared to those of other isolated eigenvectors, and is as such less prone to the noise induced by the presence of matrix $\Phi$. On the other hand, as far as the isolated eigenvalue approaches the bulk, its associated eigenvector becomes more noisy presenting the highest level of fluctuations.

**Remark 5.** (Isolated eigenvectors are asymptotically decorrelated) Since $\sigma^a_{\rho_{i_1}\rho_{i_2}}$ converges to zero whenever $\rho_{i_1} \neq \rho_{i_2}$, an interesting outcome of Theorem 6 is that the dominant eigenvectors of matrix $\tilde{\Phi}$ associated to bounded eigenvalues 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

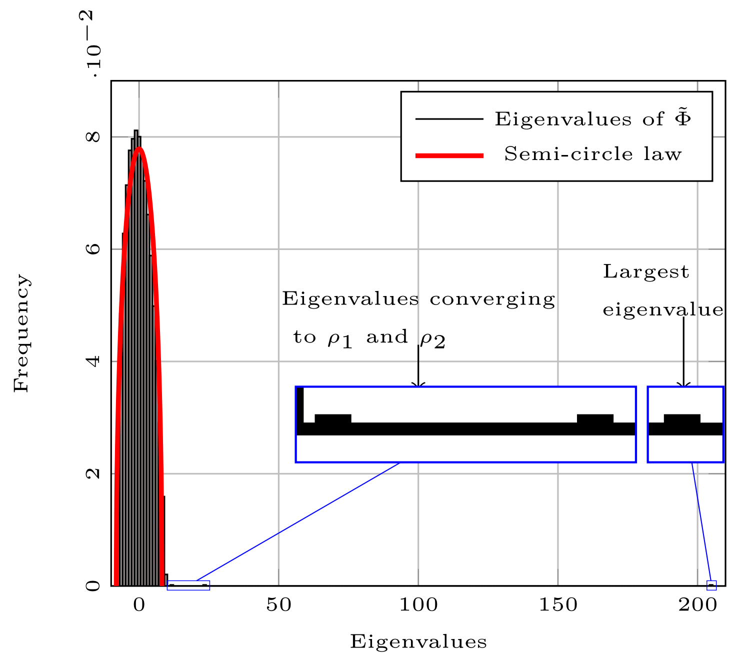


FIG 2. *Distribution of the eigenvalues of $\tilde{\Phi}$ along with the semi circle law, for $p = 1024$, $c = 3$, $c_1 = c_3 = 0.25$ and $c_2 = 0.5$, $C_i = I_p + \frac{1}{p} W_i W_i^T$ when $W_i \in \mathbb{R}^{p \times \sqrt{p}}$ with i.i.d $\mathcal{N}(0,1)$ entries.*

have different covariance matrices per class, $\Phi$ 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.

**Remark 6.** (Expression of $\alpha_a^{\rho_i}$) Since the eigenvectors are defined up to a sign, we may impose without restriction that $\alpha_s^{\rho_i} > 0$ for $s$ the smallest index $a$ for which $\alpha_a^{\rho_i} \neq 0$. Thus from Theorem 6, we find for each $a$,

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

**Numerical results** For the sake of illustration, we represent in Figure 2 the distribution of the eigenvalues of $\tilde{\Phi}$ along with the semi-circle law when observations can belong to 3 different classes with proportions $c_1 = c_3 = 0.25$ and $c_2 = 0.5$. We assume that observations are drawn from $p = 1024$ dimensional Gaussian distributions with mean zero and covariance $C_i = I_p + \theta_i \frac{1}{p} W_i W_i^T$ where $W_i$ is $p \times \sqrt{p}$ standard Gaussian matrix and $\theta_1 = 2, \theta_2 = 3$ and $\theta_3 = 4$. The total number of observations is taken to be 5000. As seen from this figure, $\tilde{\Phi}$ has two isolated eigenvalues which are bounded and one large eigenvalue scaling with $p$. We further investigate the behavior of the eigenvectors associated to the eigenvalues converging to the values $\rho_1$ and $\rho_2$ specified in Theorem 5. Figure 3

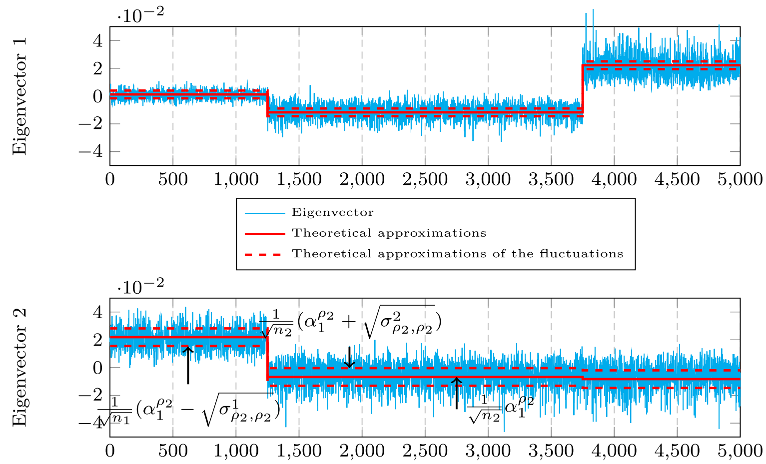


FIG 3. *Isolated eigenvectors of $\tilde{\Phi}$ converging towards $\rho_1$ and $\rho_2$ as specified in Theorem 5 versus deterministic approximations of $\alpha_a^{\rho_i} \pm \sqrt{\sigma^a_{\rho_i,\rho_i}}$ as per Theorem 6, for $n = 5000$, $p = 1024$, $c = 3$, $c_1 = c_3 = 1/4$ and $c_2 = 0.5$, $C_i = I_p + \frac{1}{p} W_i W_i^T$ with $W_i \in \mathbb{R}^{p \times \sqrt{p}}$ with i.i.d. $\mathcal{N}(0,1)$ entries.*

represents these largest eigenvectors along with the theoretical approximations provided by Theorem 6. We note a high accuracy of the provided approximations. Moreover, we note that the largest eigenvector presents the lowest variance and thus is less sensitive to the noise caused by matrix $\tilde{\Phi}$.

## 4. Mathematical tools and preliminary results

This section is dedicated to the proof of our main results in Theorems 1– 4. Throughout this section, we shall adopt the following notation. We write $x_k = C_{[k]}^{\frac{1}{2}} z_k$ and define $Z = [z_1, \cdots, z_n]$ and $X = [x_1, \ldots, x_n]$. The element $(i, j)$ of matrix $A$ will be denoted as $A_{ij}$ or $[A]_{i,j}$. Our object of interest is the resolvent matrix, which we recall is defined as

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

where $\Phi$ is defined in (12).The proof relies heavily on standard tools from Gaussian calculus as well as linear algebra relations, which we provide below for the reader's convenience.

### *4.1. Mathematical tools*

The following results will be of constant use throughout the proof of our main results.

1. Differentiation formula:
$$\frac{\partial Q_{ik}}{\partial Z_{rs}} = -2\sum_{b\neq s}(x_b^T x_s)\left[C_s^{\frac{1}{2}} x_b\right]_r (Q_{is}Q_{bk} + Q_{sk}Q_{bi}). \tag{29}$$

2. Integration by Parts formula for Gaussian functionals: Let $f$ be a $\mathcal{C}^1$ function polynomially bounded together with its derivatives. Consider $Z \in \mathbb{R}^{p\times n}$ a standard normal Gaussian matrix. Then,
$$\mathbb{E}\left[Z_{ij} f(Z)\right] = \mathbb{E}\left[\frac{\partial f(Z)}{\partial Z_{ij}}\right]. \tag{30}$$

3. Poincaré-Nash inequality: Let $Z$ and $f$ as above, then:
$$\operatorname{var}(f(Z)) \leq \sum_{i=1}^{p}\sum_{j=1}^{n}\mathbb{E}\left[\left|\frac{\partial f(Z)}{\partial Z_{ij}}\right|^2\right]. \tag{31}$$

4. Identities involving the resolvent: Define vector $\xi_k \in \mathbb{R}^n$ 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 $\xi_k$ where the $k$-th entry is replaced by zero. Let $\Phi_k$ be matrix $\Phi$ where we replace the $k$-th row and $k$-th column by zero-entry vectors. Define $Q_k = (\Phi_k - zI_n)^{-1}$. It is thus easy to notice that $Q_k$ does not depend on $x_k$ and that $[Q_k]_{kj} = 0$ for $k \neq j$. These properties will be extensively used in the proofs. Moreover, the diagonal elements of $Q$ satisfy [3, Theorem A.4]:
$$Q_{kk} = \frac{-1}{z + \xi_{(k,-k)}^T Q_k \xi_{(k,-k)}} \tag{32}$$
Furthermore, the off-diagonal element $Q_{ik}$ with $(i \neq k)$ is given by [3, page 471]:
$$Q_{ik} = \frac{e_i^T Q_k \xi_{(k,-k)}}{z + \xi_{(k,-k)}^T Q_k \xi_{(k,-k)}} = -Q_{kk} e_i^T Q_k \xi_{(k,-k)} \tag{33}$$
where $e_i$ denotes the $i$-th canonical vector of $\mathbb{C}^n$.

The integration by parts formula along with the Poincaré-Nash inequality will be extensively used to find deterministic approximations of functionals depending on the resolvent matrix and the observations $x_1, \cdots, x_n$. Let $f(Q, \{x_i\}_{i=1}^n)$ denote a scalar functional of interest. At a high level, we proceed into the following steps. First, we use the Poincaré-Nash inequality to find an upper bound of the variance. If this upper bound goes to zero with a rate $O(p^{-1-\epsilon})$ for some $\epsilon > 0$, then, from Markov inequality, the problem amounts to finding a deterministic approximation for the expectation of $f(Q, \{x_i\}_{i=1}^n)$. This is then performed by using the Integration by Parts formula (30) together with the differentiation formula (29).

## 4.2. Preliminary results

### 4.2.1. Useful inequalities

We gather in this section some matrix estimates which will be of constant use in the proof of our results.

**Lemma 1.** *Let $A$ be a $n \times n$ matrix. Then,*

$$\|\mathcal{D}(A)1_n\|_\infty \leq \|A\|$$

*Proof.* The proof follows by noticing that $\|\mathcal{D}(A)1_n\|_\infty = \max_{k=1,\cdots,n} |A_{k,k}|$ and using the fact that for all $k = 1, \cdots, n$,

$$|A_{k,k}| \leq \|A\| \qquad \square$$

**Lemma 2.** *Let $A$ and $B$ be two $n \times n$ matrices. Then,*

$$\|AB\| \leq \|A\|\|B\|.$$

*Moreover, denoting by $\odot$ the Hadamard product, we also have:*

$$\|A \odot B\| \leq \|A\|\|B\|$$

**Lemma 3.** *Let $A$ be $n \times n$ matrix. Then, the following inequality hold true,*

$$\|A\| \leq \sqrt{\operatorname{tr}(AA^T)}$$

### 4.2.2. Useful approximations of random quantities

In this section, we introduce some important results that will be extensively used in the proof of our main theorems. These results facilitate the assessment of random quantities involving the resolvent matrix. We shall for the reader's convenience, recall the notation $x_p = O(r_p)$ where $x_p$ is a sequence of random variables and $r_p$ is a rate decreasing with $p$. The notation $x_p = O(r_p)$ implies that for every $\eta$ and $D$ strictly positives $\mathbb{P}\left[x_p \geq p^\eta r_p\right] = o(p^{-D})$. As shown in [7], the notation $O(.)$ has the property that the maximum of a collection of $n^C$ random variables for any constant $C$ of order $O(r_p)$ is still $O(r_p)$. Using standard concentration inequalities, we can show that this notation holds for many functionals of Gaussian vectors. Particularly, if $z_p$ and $w_p$ are two independent standard normal vectors and $A_p$ a sequence of deterministic or random matrices with bounded spectral norm that are independent of $z_p$ and $w_p$, then $\frac{1}{p}{z_p}^T A_p w_p = O(p^{-\frac{1}{2}})$.

Let $A_{1,p}$, $A_{2,p}$ and $A_{3,p}$ be sequences of $p \times p$ deterministic matrices with spectral norm uniformly bounded in $p$. Let $k \in \{1, \cdots, n\}$. Define the $n \times n$ matrix $S_k$ such that its $(b_1, b_2)$ entry is given by:

$$[S_k]_{b_1 b_2} = \left(x_{b_1}^T A_{1,p} x_k\right)\left(x_{b_2}^T A_{2,p} x_k\right)\left(x_{b_1}^T A_{3,p} x_{b_2}\right) \boldsymbol{\delta}_{k \neq b_1} \boldsymbol{\delta}_{k \neq b_2}. \qquad (34)$$

The following result, controlling the spectral norm of this matrix is extensively required in the proofs of our results.

**Lemma 4.** *Let $S_k$ be as in (34). Then,*

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

*Proof.* We can write $S_k$ as:

$$S_k = D_k X^T A_{3,p} X \tilde{D}_k$$

where $D_k = \mathcal{D}\left\{x_b^T A_{1,p} x_k \boldsymbol{\delta}_{b\neq k}\right\}_{b=1}^n$, $\tilde{D}_k = \mathcal{D}\left\{x_b^T A_{2,p} x_k \boldsymbol{\delta}_{b\neq k}\right\}_{b=1}^n$ and $X = [x_1, \ldots, x_n]$. The result follows since $\|D_k\| = O(p^{-1/2})$, $\|\tilde{D}_k\| = O(p^{-\frac{1}{2}})$ and $\|X^T A_{3,p} X\| = O(1)$ as per [2]. □

**Lemma 5.** *Let $W_{1,p}$ and $W_{2,p}$ be two sequences of positive random variables such that there exists constant $K \geq 1$ for which*

$$\mathbb{E} W_{2,p} \leq K p^{-r} \tag{35}$$

$$\mathbb{E} W_{2,p}^4 \leq K \tag{36}$$

$$\mathbb{E} W_{1,p}^4 \leq K p^{\alpha} \tag{37}$$

*for some positive constants $\alpha$ and $r$. Assume that $W_{1,p} = O(1)$. Then, for any $\epsilon > 0$, we have*

$$\mathbb{E}\left[W_{1,p} W_{2,p}\right] \leq 2K p^{-r+\epsilon}.$$

*Proof.* Let $\epsilon > 0$. We have:

$$\begin{aligned}
\mathbb{E}\left[W_{1,p}W_{2,p}\right] &= \mathbb{E}\left[W_{1,p}W_{2,p}\boldsymbol{\delta}_{\{W_{1,p}\geq p^{\epsilon}\}}\right] + \mathbb{E}\left[W_{1,p}W_{2,p}\boldsymbol{\delta}_{\{W_{1,p}\leq p^{\epsilon}\}}\right] \\
&\leq \sqrt{\mathbb{E}W_{1,p}^2 W_{2,p}^2}\sqrt{\mathbb{P}\left[W_{1,p}\geq p^{\epsilon}\right]} + p^{\epsilon}\mathbb{E}\left[W_{2,p}\right] \\
&\leq \sqrt{K} p^{\frac{\alpha}{4}} \sqrt{\mathbb{P}\left[W_{1,p}\geq p^{\epsilon}\right]} + K p^{-r+\epsilon}
\end{aligned}$$

The result of the lemma follows by noticing that $\mathbb{P}\left[W_{1,p} \geq p^{\epsilon}\right] = o(p^{-l})$ for any $l > 0$, which follows from the definition of $O(p^{-r})$ for random variables described in the notation section. Taking $l = 2r + \frac{\alpha}{2} - 2\epsilon$ finishes the proof. □

**Remark 7.** Lemma 5 offers a practical way to control the expectation of the product of the two random variables $W_{1,p}$ and $W_{2,p}$, in which the fourth moment of one random variable, (here $W_{1,p}$) can be coarsely bounded by a constant that scales with $p^{\alpha}$. This situation occurs for instance when $W_{1,p}$ represents the maximum of random variables with bounded moments [2]. If $W_{1,p}$ additionally satisfies $W_{1,p} = O(1)$, the growth rate of the expected value of the product of these random variables will be essentially determined by that of $W_{2,p}$. For instance, when $W_{2,p} = 1$ and $W_{1,p} = O(1)$ satisfying the conditions of Lemma 5, then, for any small $\epsilon > 0$,

$$\mathbb{E}[W_{1,p}] = O(p^{\epsilon})$$

[2]Assuming $W_{1,p} = \max_{1\leq k\leq p} |Y_k|$ where $Y_k, k = 1, \cdots, p$ have all finite moments. Then, $\mathbb{E}|W_{1,p}|^4 \leq \mathbb{E}|\sum_{k=1}^p |Y_k||^4 \leq p^3 \sum_{k=1}^p \mathbb{E}[|Y_k|^4]$

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 6** (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 $\mathbb{E}_k$ the expectation operator with respect to $x_k$. Define vector $d_k$ as the $n \times 1$ vector with elements:*

$$[d_k]_i = \begin{cases} \frac{1}{\sqrt{p}} x_i^T C_{[k]} x_i - \frac{1}{p^{\frac{3}{2}}} \operatorname{Tr}(C_{[k]} C_{[i]}) & \text{for } i \neq k \\ 0, & \text{otherwise.} \end{cases} \tag{38}$$

*Define also matrix $\Sigma_k$ as the $n \times n$ matrix with elements:*

$$\Sigma_k = \begin{cases} (x_i^T C_{[k]} x_j)^2 & \text{for } i \neq k \text{ and } j \neq k \\ 0 & \text{otherwise.} \end{cases} \tag{39}$$

*Then,*

$$\mathbb{E}_k \left[ \xi_{(k,-k)}^T A \xi_{(k,-k)} \right] = d_k^T A d_k + \frac{2}{p} 1^T (\Sigma_k \odot A) \mathbf{1} \tag{40}$$

*Moreover, we also have for any $\epsilon > 0$:*

$$\mathbb{E}_k \left| \xi_{(k,-k)}^T A \xi_{(k,-k)} - \mathbb{E}_k \xi_{(k,-k)}^T A \xi_{(k,-k)} \right|^{2s} = \|A\|^{2s} O(p^{-s+\epsilon}) \tag{41}$$

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

*Proof.* See Appendix A.1 □

**Corollary 2.** *Let $k \in \{1, \ldots, n\}$. Let $A$ be a $n \times n$ symmetric matrix independent of $x_k$. Let $a$ and $b$ be $n \times 1$ vector independent of $x_k$. Denote by $\mathbb{E}_k$ the expectation operator with respect to $x_k$. Then,*

$$\left| \mathbb{E}_k \left[ \xi_{(k,-k)}^T A a b^T A \xi_{(k,-k)} \right] \right| \leq \|a\| \|b\| \|A\|^2 O(\frac{1}{p}) \tag{42}$$

*Proof.* It follows from Lemma 6 that

$$\mathbb{E}_k \left[ \xi_{(k,-k)}^T A a b^T A \xi_{(k,-k)} \right] = d_k^T A a b^T A d_k + \frac{2}{p} b^T A \Sigma_k A a \tag{43}$$

where $d_k$ and $\Sigma_k$ are defined in (38) and (39). Noting that $\|d_k\| = O(\frac{1}{\sqrt{p}})$ and that $\|\Sigma_k\| = O(1)$, we can upper-bound the first and second terms in the above equality as:

$$\left| d_k^T A a b^T A d_k \right| \leq \|A\|^2 \|a\| \|b\| O(\frac{1}{p}),$$

$$\left| \frac{2}{p} b^T A \Sigma_k A a \right| \leq 2 \|b\| \|a\| \|A\|^2 O(\frac{1}{p})$$

which proves (42). □

*4.2.3. Useful properties of the Stieltjes transform of the semi-circle distribution*

**Lemma 7.** *Let $z \in \mathbb{C} \backslash \left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right]$. 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 $\mathbb{C} \backslash \left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right]$*
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 strictly positive scalar. Then, it holds that*

$$\left(\left|1 - \alpha m^2(z)\right|\right)^{-1} \leq (|z| + 2\sqrt{c_0}\omega)^4 \left(4|\Im z|^{-4} + \frac{2}{\alpha}|\Im z|^{-2}\right). \tag{44}$$

*Moreover, if $|z| \geq 2\sqrt{2}\sqrt{c_0}\omega\sqrt{4 + \frac{2\alpha}{\omega^2 c_0}}$, then:*

$$\left|1 - \alpha m^2(z)\right| \geq \frac{|z|^4}{8(|z| + 2\sqrt{c_0}\omega)^4} \tag{45}$$

*Proof.* The proof is in Appendix A.2. □

*4.2.4. Variance evaluations of resolvent based quantities*

In this section, we leverage the Poincaré-Nash inequality to evaluate the variance of quadratic forms and weighted averages of diagonal elements of the resolvent matrix.

**Lemma 8.** *Let $z \in \mathbb{C} \backslash \mathbb{R}$. Let $\{a_n\}$ and $\{b_n\}$ be two sequences of deterministic vectors with unit norm in $\mathbb{R}^{n \times 1}$. Then, for any $\epsilon > 0$:*

$$\mathbb{E}\left[\left|a_n^T Q b_n - \mathbb{E} a_n^T Q b_n\right|^{2s}\right] = O_z(p^{-s+\epsilon}) \quad \text{for } s \in \mathbb{N}^* \tag{46}$$

$$\operatorname{var}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n Q_{ii} a_{i,n}\right) = O_z(p^{-2+\epsilon}). \tag{47}$$

*Proof.* For the sake of simplification, we shall remove the subscript $n$ from $a_n$ and $b_n$.

1. The proof of (46) is performed by induction on $s$. For $s = 1$,

$$\mathbb{E}\left|a^T Q b - \mathbb{E}\left[a^T Q b\right]\right|^2 = \operatorname{var}\left(a^T Q b\right)$$

Using Poincaré-Nash inequality, we can upper-bound the variance of $a^TQb$ as follows:

$$\operatorname{var}(a^TQb) \leq \sum_{l=1}^{p}\sum_{k=1}^{n}\mathbb{E}\Big[\Big|\frac{\partial a^TQb}{\partial Z_{lk}}\Big|^2\Big] \tag{48}$$

To prove the desired result, we shall show that:

$$\sum_{l=1}^{p}\sum_{k=1}^{n}\Big|\frac{\partial a^TQb}{\partial Z_{lk}}\Big|^2 \leq |\Im z|^{-4}O(p^{-1}) \tag{49}$$

For that, we rely on the differentiation formula in (29) to obtain:

$$\begin{aligned}\sum_{l=1}^{p}\sum_{k=1}^{n}\Big|\frac{\partial a^TQb}{\partial Z_{lk}}\Big|^2 &\leq 8\sum_{l=1}^{p}\sum_{k=1}^{n}\Big|\sum_{i=1}^{n}\sum_{j=1}^{n}a_ib_j\sum_{s\neq k}(x_s^Tx_k)\left[C_{[k]}^{\frac{1}{2}}x_s\right]_l Q_{ik}Q_{sj}\Big|^2\\ &+8\sum_{l=1}^{p}\sum_{k=1}^{n}\Big|\sum_{i=1}^{n}\sum_{j=1}^{n}a_ib_j\sum_{b\neq k}(x_s^Tx_k)\left[C_{[k]}^{\frac{1}{2}}x_s\right]_l Q_{kj}Q_{si}\Big|^2\\ &= 8\sum_{l=1}^{p}\sum_{k=1}^{n}\Big|a^TQe_k\sum_{s\neq k}(x_s^Tx_k)\left[C_{[k]}^{\frac{1}{2}}x_s\right]_l e_s^TQb\Big|^2\\ &+8\sum_{l=1}^{p}\sum_{k=1}^{n}\Big|a^TQe_s\sum_{s\neq k}(x_s^Tx_k)\left[C_{[k]}^{\frac{1}{2}}x_s\right]_l e_k^TQb\Big|^2\end{aligned} \tag{50}$$

We will only treat the first term as the second one can be handled in the same manner. Expanding the sum of the first term, we obtain:

$$8\sum_{l=1}^{p}\sum_{k=1}^{n}\left|a^TQe_k\sum_{s\neq k}(x_s^Tx_k)\left[C_{[k]}^{\frac{1}{2}}x_s\right]_l e_s^TQb\right|^2 \tag{51}$$

$$= 8\sum_{k=1}^{n}\sum_{s_1\neq k}\sum_{s_2\neq k}|a^TQe_k|^2(x_{s_1}^Tx_k)(x_{s_2}^Tx_k)x_{s_1}^TC_{[k]}x_{s_2}[Qb]_{s_1}[Q^Hb]_{s_2} \tag{52}$$

$$= 8\sum_{k=1}^{n}|a^TQe_k|^2b^TQS_kQ^Hb \tag{53}$$

where $S_k$ is the $n\times n$ matrix with entries

$$[S_k]_{s_1s_2} = (x_{s_1}^Tx_k)(x_{s_2}^Tx_k)x_{s_1}^TC_{[k]}x_{s_2}.$$

From Lemma 4, the spectral norms of matrices $S_k$ satisfy:

$$\max_{1\leq k\leq n}\|S_k\| = O(p^{-1})$$

Using the fact $\sum_{k=1}^{n}|a^TQe_k|^2$ is bounded by $|\Im z|^{-2}\|a\|^2$, we thus prove (49), and hence (46) follows by Lemma 5. Assume now that (46) holds true up to $s-1 \in \mathbb{N}^*$. Note that

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

Using the induction assumption, along with Cauchy-Schwartz inequality, the first term in the above equation can be shown to be $O_z(p^{-s+\epsilon})$. It remains thus to prove the same result for the second term. Based on the Poincaré-Nash inequality,

$$\begin{aligned}\operatorname{var}(a^T Qb - \mathbb{E}a^T Qb)^s &\leq s^2 \sum_{l=1}^p \sum_{k=1}^n \mathbb{E}\left[\left|(a^T Qb - \mathbb{E}a^T Qb)^{s-1} \frac{\partial a^T Qb}{\partial Z_{lk}}\right|^2\right] \\ &= s^2 \mathbb{E}\left[(a^T Qb - \mathbb{E}a^T Qb)^{2(s-1)} \sum_{l=1}^p \sum_{k=1}^n \left|\frac{\partial a^T Qb}{\partial Z_{lk}}\right|^2\right]\end{aligned}$$

Using (49) along with the induction assumption, we obtain:

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

2. Proof of $\operatorname{var}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n a_i Q_{ii}\right) = O_z(p^{-2+\epsilon})$. From Poincaré-Nash inequality, we have:

$$\begin{aligned}\operatorname{var}\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n a_i Q_{ii}\right) &\leq \frac{1}{n}\sum_{l=1}^p \sum_{k=1}^n \mathbb{E}\left|\sum_{i=1}^n a_i \frac{\partial Q_{ii}}{\partial Z_{lk}}\right|^2 \\ &= \frac{16}{n}\mathbb{E}\sum_{k=1}^n \left[Q\mathcal{D}\left\{a_i\right\}_{i=1}^n \left(QS_k Q^H\right)\mathcal{D}\left\{a_i\right\}_{i=1}^n Q^H\right]_{kk} \\ &= \frac{16}{n}\mathbb{E}\Big[\operatorname{tr}\left(Q\mathcal{D}\left\{a_i\right\}_{i=1}^n \left(QS_k Q^H\right)\mathcal{D}\left\{a_i\right\}_{i=1}^n Q^H\right)\Big] \\ &\leq \frac{16}{n}\mathbb{E}[\|Q\|^4 \|S_k\| \operatorname{tr}((\mathcal{D}\left\{a_i\right\}_{i=1}^n)^2)]\end{aligned}$$

where the last inequality follows from the fact that $\operatorname{tr}(AB) \leq \|A\| \operatorname{tr}(B)$ for $A$ and $B$ two $n\times n$ matrices with $B$ being hermitian non-negative. Since $\operatorname{tr}((\mathcal{D}\left\{a_i\right\}_{i=1}^n)^2) = \|a\|^2$ is bounded, we obtain: $\operatorname{var}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n a_i Q_{ii}\right) = O_z(p^{-2+\epsilon})$. □

**Lemma 9.** *Let $z \in \mathbb{C}\backslash\mathbb{R}$. For $j \in \{1,\ldots,n\}$, let $(A_{p,j})_{p\in\mathbb{N}^*}$ be a sequence of $p\times p$ matrices satisfying* $\limsup_p \max_{1\leq j\leq n}\|A_{j,p}\| < \infty$. *Then, for any $s \in \mathbb{N}^*$ and $\epsilon > 0$*

$$\max_{1\leq j\leq n} \mathbb{E}\left|\sum_{k\neq j} \frac{1}{\sqrt{p}}\left(x_k^T A_{p,j} x_k - \frac{1}{p}\operatorname{tr} C_{[k]} A_{p,j}\right) Q_{jk}\right|^{2s} = O_z(p^{-2s+\epsilon}) \qquad (54)$$

*Proof.* See Appendix A.3 □

**Lemma 10.** *Let $z \in \mathbb{C}\backslash\mathbb{R}$. Let $a$ be a unit norm deterministic vector in $\mathbb{R}^{n\times 1}$. Then, for any $\epsilon > 0$ and $s \in \mathbb{N}^*$,*

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

*Proof.* The proof is carried out by induction on $s$. For $s = 1$, the result follows by applying Corollary 2. Let $s \in \mathbb{N}$. Assume that the result holds true for all $k \leq s-1$, and let us prove it for $k = s$. To begin with, we decompose $a^T Q_k \xi_{(k,-k)}$ as:

$$a^T Q_k \xi_{(k,-k)} = a^T Q_k \xi_{(k,-k)} - \mathbb{E}_k\left[a^T Q_k \xi_{(k,-k)}\right] + \mathbb{E}_k\left[a^T Q_k \xi_{(k,-k)}\right]$$

and apply Jensen inequality to obtain:

$$\begin{aligned}\mathbb{E}\left|a^T Q_k \xi_{(k,-k)}\right|^{2s} &\leq 2^{2s-1}\mathbb{E}\Big[\left|a^T Q_k \xi_{(k,-k)} - \mathbb{E}_k\left[a^T Q_k \xi_{(k,-k)}\right]\right|^{2s}\Big] \\ &+ 2^{2s-1}\mathbb{E}\Big[\left|\mathbb{E}_k\left[a^T Q_k \xi_{(k,-k)}\right]\right|^{2s}\Big] \end{aligned} \tag{55}$$

The second term in the right-hand side of the above inequality is $O_z(p^{-s+\epsilon})$ by Corollary 2. To handle the first quantity, we use the following equality:

$$\mathbb{E}_k\Big[\left|a^T Q_k \xi_{(k,-k)} - \mathbb{E}_k\left[a^T Q_k \xi_{(k,-k)}\right]\right|^{2s}\Big] \tag{56}$$

$$\begin{aligned}&= \mathbf{var}_k\left(\left(a^T Q_k \xi_{(k,-k)} - \mathbb{E}_k\left[a^T Q_k \xi_{(k,-k)}\right]\right)^s\right) \\ &+ \left|\mathbb{E}_k\Big[\left|a^T Q_k \xi_{(k,-k)} - \mathbb{E}_k\left[a^T Q_k \xi_{(k,-k)}\right]\right|^s\Big]\right|^2\end{aligned} \tag{57}$$

where $\mathbf{var}_k$ is the variance with respect to the random vector $x_k$. Hence,

$$\begin{aligned}&\mathbb{E}\Big[\left|a^T Q_k \xi_{(k,-k)} - \mathbb{E}_k\left[a^T Q_k \xi_{(k,-k)}\right]\right|^{2s}\Big] \\ &= \mathbb{E}\left[\mathbf{var}_k\left(\left(a^T Q_k \xi_{(k,-k)} - \mathbb{E}_k\left[a^T Q_k \xi_{(k,-k)}\right]\right)^s\right)\right] \\ &+ \mathbb{E}\left[\left|\mathbb{E}_k\Big[\left|a^T Q_k \xi_{(k,-k)} - \mathbb{E}_k\left[a^T Q_k \xi_{(k,-k)}\right]\right|^s\Big]\right|^2\right]\end{aligned} \tag{58}$$

To treat the second term in (58), we apply Cauchy-Schwartz inequality to find:

$$\begin{aligned}&\left|\mathbb{E}_k\Big[\left|a^T Q_k \xi_{(k,-k)} - \mathbb{E}_k\left[a^T Q_k \xi_{(k,-k)}\right]\right|^s\Big]\right|^2 \\ &\leq \mathbb{E}_k\Big[\left|a^T Q_k \xi_{(k,-k)} - \mathbb{E}_k\left[a^T Q_k \xi_{(k,-k)}\right]\right|^{2s-2}\Big]\mathbf{var}_k\left(a^T Q_k \xi_{(k,-k)}\right)\end{aligned}$$

By Corollary 2,

$$\mathbf{var}_k(a^T Q_k \xi_{(k,-k)}) \leq \mathbb{E}_k|a^T Q_k \xi_{(k,-k)}|^2 \leq \|a\|^2|\Im z|^{-2}O(\frac{1}{p})$$

Hence using Lemma 5 along with the induction assumption, we obtain:

$$\mathbb{E}\left[\left|\mathbb{E}_k\left[\left|a^TQ_k\xi_{(k,-k)}-\mathbb{E}_k\left[a^TQ_k\xi_{(k,-k)}\right]\right|^s\right]\right|^2\right]=O_z(p^{-s+\epsilon}) \tag{59}$$

To conclude, it remains thus to handle the first term in (58). For that, we use Poincaré-Nash inequality, which leads to:

$$\begin{aligned}
&\mathbf{var}_k\left(\left(a^TQ_k\xi_{(k,-k)}-\mathbb{E}_k\left[a^TQ_k\xi_{(k,-k)}\right]\right)^s\right)\\
&\leq\sum_{l=1}^p\left|\frac{\partial\left(a^TQ_k\xi_{(k,-k)}-\mathbb{E}_k\left[a^TQ_k\xi_{(k,-k)}\right]\right)^s}{\partial Z_{lk}}\right|^2\\
&=\sum_{l=1}^p s^2\left(a^TQ_k\xi_{(k,-k)}-\mathbb{E}_k\left[a^TQ_k\xi_{(k,-k)}\right]\right)^{2(s-1)}\left|\frac{\partial a^TQ_k\xi_{k,-k}}{\partial Z_{lk}}\right|^2 &(60)\\
&=4\sum_{l=1}^p s^2\left(a^TQ_k\xi_{(k,-k)}-\mathbb{E}_k\left[a^TQ_k\xi_{(k,-k)}\right]\right)^{2(s-1)}\\
&\times\left|\sum_{j\neq k}^n[a^TQ_k]_j(x_j^Tx_k)\left[C_{[k]}^{\frac{1}{2}}x_j\right]_l\right|^2 &(61)\\
&=4s^2\left(a^TQ_k\xi_{(k,-k)}-\mathbb{E}_k\left[a^TQ_k\xi_{(k,-k)}\right]\right)^{2(s-1)}a^TQ_kS_kQ_ka &(62)
\end{aligned}$$

where $S_k$ is the $n\times n$ matrix with elements

$$[S_k]_{j_1,j_2}=\boldsymbol{\delta}_{j_1\neq k}\boldsymbol{\delta}_{j_2\neq k}(x_{j_1}^Tx_k)(x_{j_2}^Tx_k)x_{j_1}^TC_{[k]}x_{j_2}.$$

From Lemma 4, $\|S_k\|=O(p^{-1})$, hence,

$$|a^TQ_kS_kQ_ka|\leq|\Im z|^{-2}\|a\|^2O(p^{-1}) \tag{63}$$

From the induction assumption, it follows that

$$\mathbb{E}(a^TQ_k\xi_{(k,-k)}-\mathbb{E}_k\left[a^TQ_k\xi_{(k,-k)}\right]\Big)^{2(s-1)}=O_z(p^{-s+1+\epsilon}).$$

This, together with (63) and Lemma 5 leads to

$$\mathbb{E}\left[\mathbf{var}_k\left(\left(a^TQ_k\xi_{(k,-k)}-\mathbb{E}_k\left[a^TQ_k\xi_{(k,-k)}\right]\right)^s\right)\right]=O_z(p^{-s+\epsilon}) \tag{64}$$

Combining (64) with (59), we thus prove the desired result. □

A direct corollary of Lemma 10 is the following result:

**Corollary 3.** *Let $i\neq k$ with $i,k\in\{1,\cdots,n\}$. Then, for any $\epsilon>0$,*

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

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

$$\mathbb{E}[Q_{ik}]=O_z(p^{-1+\epsilon}) \tag{65}$$

*Proof.* Recalling (33), we have:

$$Q_{ik} = -Q_{kk} e_i^T Q_k \xi_{(k,-k)}$$

Hence,

$$\mathbb{E}\Big[|Q_{ik}|^s\Big] \leq |\Im z|^{-s} \mathbb{E}\Big[|e_i^T Q_k \xi_{(k,-k)}|^s\Big]$$

From Lemma 10, $\mathbb{E}\Big[|e_i^T Q_k \xi_{(k,-k)}|^s\Big] = O_z(p^{-s+\epsilon})$, and thus so is $\mathbb{E}\Big[|Q_{ik}|^s\Big]$. To prove (65), we decompose $Q_{ik}$ as:

$$Q_{ik} = -(Q_{kk} - \mathbb{E}[Q_{kk}]) e_i^T Q_k \xi_{(k,-k)} - \mathbb{E}[Q_{kk}] e_i^T Q_k \xi_{(k,-k)} \tag{66}$$

and use Lemma 10 along with Lemma 8, to prove that

$$\mathbb{E}\left[(Q_{kk} - \mathbb{E}[Q_{kk}]) e_i^T Q_k \xi_{(k,-k)}\right] = O_z(p^{-1+\epsilon}) \tag{67}$$

Computing the expectation with respect to $x_k$, we can show that:

$$\mathbb{E}\Big[e_i^T Q_k \xi_{(k,-k)}\Big] = \frac{1}{\sqrt{p}} \sum_{l \neq k} \mathbb{E}\Big[[Q_k]_{il} (x_l^T C_{[k]} x_l - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[l]}))\Big] \tag{68}$$

$$= \frac{1}{\sqrt{p}} \sum_{l \notin \{k,i\}} \mathbb{E}\Big[[Q_k]_{il} (x_l^T C_{[k]} x_l - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[l]}))\Big]$$

$$+ \frac{1}{\sqrt{p}} \mathbb{E}\Big[[Q_k]_{ii} (x_i^T C_{[k]} x_i - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[i]}))\Big] \tag{69}$$

The second term in the above inequality is obviously $O_z(p^{-1+\epsilon})$. The control of the first term can be performed using Lemma 9. To see this, we define $\tilde{X}_k$ as the $n-1 \times n-1$ matrix made up of the columns of $X$ except the $k$-th one and denote by $\tilde{x}_1, \cdots, \tilde{x}_{n-1}$ its corresponding columns. Then, we form the $n-1 \times n-1$ matrix $\hat{\Phi}_k = \left\{\sqrt{p}\left((\tilde{x}_i^T \tilde{x}_j)^2 - \mathbb{E}[(\tilde{x}_i^T \tilde{x}_j)^2]\right) \boldsymbol{\delta}_{i \neq j}\right\}_{i,j=1}^{n-1}$ and introduce its associated resolvent $\hat{Q}_k(z) = (\hat{\Phi}_k - z I_{n-1})^{-1}$. With this, it takes no much effort to notice that the first term is given by

$$\frac{1}{\sqrt{p}} \sum_{l \neq i} \mathbb{E}\Big[[\hat{Q}_k]_{il} (\tilde{x}_l^T C_{[k]} \tilde{x}_l - \mathbb{E}[\tilde{x}_l^T C_{[k]} \tilde{x}_l])\Big]$$

which is clearly $O_z(p^{-1+\epsilon})$ by Lemma 9. We thus obtain:

$$\mathbb{E}\Big[e_i^T Q_k \xi_{(k,-k)}\Big] = O_z(p^{-1+\epsilon}) \tag{70}$$

Combining (67) and (70), we thus prove that the expectation of both quantities in (66) are $O_z(p^{-1+\epsilon})$ which shows (65). ☐

*4.2.5. Other important results*

**Lemma 11.** *Let $k, b$ and $j$ integers in the set $\{1, \dots, n\}$. Let $A_{1,p}$ and $A_{2,p}$ be two sequences of $p \times p$ matrices possibly random but independent of $x_j$ and have spectral norms of order $O(1)$. Then, for any small $\epsilon > 0$,*

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

*Proof.* See Appendix A.4. □

**Lemma 12.** *Let $j, k \in \{1, \cdots, n\}$ with $j \neq k$. Let $A_{1,j,p}$, $A_{2,j,p}$, $A_{3,j,p}$ and $A_{4,j,p}$ be four sequences of $p \times p$ matrices with bounded spectral norm. Then, for any $\epsilon > 0$, we have:*

$$\max_{j \neq k} \mathbb{E} \left| \sum_{r \notin \{j,k\}} \sum_{b \notin \{j,r,k\}} x_b^T A_{1,j,p} x_k x_b^T A_{2,j,p} x_j x_k^T A_{3,j,p} x_r x_r^T A_{4,j,p} x_j Q_{br} \right|^2$$
$$= O_z(p^{-3+\epsilon})$$

*Proof.* See Appendix A.6 □

**Lemma 13.** *Let $k \in \{1, \cdots, n\}$. Let $b$ be a unit norm deterministic vector in $\mathbb{R}^n$. Let $c$ be a random vector in $\mathbb{R}^n$ independent of $x_k$ such that $\|c\|^2 = O(1)$ Then, for $z \in \mathbb{C} \backslash \mathbb{R}$ and any small $\epsilon$:*

$$\sum_{r \neq k} \sum_{l \neq k} \left( \mathbb{E}\left[ Q_{rl} b_r c_l \right] - \mathbb{E}\left[ [Q_k]_{rl} b_r c_l \right] \right) = O_z(p^{-1+\epsilon}).$$

*where $b_l$ and $c_r$ denote the $r$-th and the $l$-th entries of $a$ and $b$, respectively.*

*Proof.* See Appendix A.5 □

### 4.3. 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 $1 \le j \le n$,

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

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

$$\begin{aligned}
&\mathbb{E}\left[Q_{jk}\Phi_{kj}\right] \\
&= \frac{1}{p^{\frac{3}{2}}} \sum_{a,b=1}^{p} \sum_{l,l'=1}^{p} \sum_{m,m'=1}^{p} \big[C_{[k]}^{\frac{1}{2}}\big]_{al} \big[C_{[j]}^{\frac{1}{2}}\big]_{al'} \big[C_{[k]}^{\frac{1}{2}}\big]_{bm} \big[C_{[j]}^{\frac{1}{2}}\big]_{bm'} \mathbb{E}\big[Z_{lk} Z_{l'j} Z_{mk} Z_{m'j} Q_{jk}\big] \\
&- \frac{1}{p^{\frac{3}{2}}} \operatorname{tr} C_{[k]} C_{[j]} \mathbb{E}\left[Q_{jk}\right] \qquad (71)
\end{aligned}$$

Using the integration by parts formula in (30) along with the differentiation formula in (29), we obtain:

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

Plugging the above equation into (71), we ultimately get:

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

Hence,

$$z\mathbb{E}Q_{jj} = -1 + \alpha_j(z) + \beta_j(z) + \gamma_j(z) + \theta_j(z) \qquad (72)$$

where $\alpha_j(z), \beta_j(z), \gamma_j(z)$ and $\theta_j(z)$ write as:

$$\alpha_j(z) = \mathbb{E}\big[\sum_{k\neq j} \frac{1}{\sqrt{p}} \left(x_k^T C_{[j]} x_k - \frac{1}{p} \operatorname{tr} C_{[k]} C_{[j]}\right) Q_{jk}\big] \qquad (73)$$

$$\beta_j(z) = -2 \sum_{k\neq j} \mathbb{E}\left[\left(x_j^T x_k\right)^2 x_k^T C_{[j]} x_k Q_{jj} Q_{kk}\right] \qquad (74)$$

$$\gamma_j(z) = -2 \sum_{k\neq j} \sum_{r\notin\{j,k\}} \mathbb{E}\left[x_j^T x_k x_r^T x_j x_k^T C_{[j]} x_r Q_{jj} Q_{rk}\right] \qquad (75)$$

$$\theta_j(z) = -2 \sum_{k\neq j} \sum_{r\neq j} \mathbb{E}\left[x_j^T x_k x_r^T x_j x_k^T C_{[j]} x_r Q_{jk} Q_{rj}\right] \qquad (76)$$

The decomposition in (72) will play a key role in the proof of our main results, as will be seen in section 5.

## 5. Proof of the main results

### *5.1. Proof of Theorem 1*

The proof of Theorem 1 will rely on (72) in which quantities $\alpha_j(z), \gamma_j(z)$ and $\theta_j(z)$ constitute error terms that converge to zero in the asymptotic regime. Indeed, a direct application of Lemma 9 allows us to show that:

$$\max_{1\leq j\leq n} |\alpha_j(z)| = O_z(p^{-1+\epsilon}). \tag{77}$$

To control $\gamma_j(z)$ and $\theta_j(z)$, we will rely on Lemma 11. Indeed, by Lemma 11,

$$\max_{1\leq j\leq n} |\gamma_j(z)| = O_z(p^{-\frac{1}{2}+\epsilon}) \tag{78}$$

To handle $\theta_j(z)$, we start by decomposing it as:

$$\begin{aligned}\theta_j(z) = &-2\sum_{k\neq j}\sum_{r\notin\{j,k\}} \mathbb{E}\left[x_j^T x_k x_r^T x_j x_k^T C_{[j]} x_r Q_{jk} Q_{rj}\right] \\ &- 2\sum_{k\neq j} \mathbb{E}\left[x_j^T x_k x_k^T x_j x_k^T C_{[j]} x_k Q_{jk} Q_{kj}\right]\end{aligned}$$

Then, using Lemma 11 and the approximations in Corollary 3, it unfolds that:

$$\max_{1\leq j\leq n} |\theta_j(z)| = O_z(p^{-1+\epsilon}). \tag{79}$$

Finally, to treat $\beta_j(z)$, we use the fact that $x_k^T C_{[j]} x_k - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[j]}) = O(p^{-\frac{1}{2}})$ in combination with Lemma 5 to obtain:

$$\beta_j(z) = -2\sum_{k\neq j}\frac{1}{p}\left(\frac{1}{p}\operatorname{tr} C_{[k]}C_{[j]}\right)^2 \mathbb{E}Q_{kk}\mathbb{E}Q_{jj} + O_z(p^{-\frac{1}{2}+\epsilon}) \tag{80}$$

$$= -\frac{\omega^2 c_0}{n}\sum_{k=1}^{n} \mathbb{E}Q_{kk}\mathbb{E}Q_{jj} + O_z(p^{-\frac{1}{2}+\epsilon}) \tag{81}$$

where (81) follows from the fact that $\frac{\sqrt{2}}{p}\operatorname{tr} C_{[k]}C_{[j]} = \omega + O(p^{-\frac{1}{2}})$. Now, putting (77), (78), (79) and (81) together with (72), we obtain:

$$z\mathbb{E}Q_{jj} = -1 - \omega^2 c_0 g_n(z)\mathbb{E}Q_{jj} + O_z(p^{-\frac{1}{2}+\epsilon}) \tag{82}$$

Summing (82) over index $j$, we get

$$\omega^2 c_0 g_n^2(z) + z g_n(z) + 1 = O_z(p^{-\frac{1}{2}+\epsilon}). \tag{83}$$

where we recall that $g_n(z) = \frac{1}{n}\operatorname{tr}\mathbb{E}Q(z)$. Reaching this equation, termed as "Master equation" in [5], it can be proven by following the same steps in [5] that:

$$|g_n(z) - m(z)| = O_z(p^{-\frac{1}{2}+\epsilon}). \tag{84}$$

The weak convergence of the spectral measure of $\Phi$ 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.

### *5.2. Proof of Theorem 2*

Theorem 2 provides a deterministic equivalent for bilinear forms of the resolvent matrix $Q(z)$.Due to the almost sure convergence of $a_n^T Q(z) b_n - a_n^T \mathbb{E} Q(z) b_n$ to zero, guaranteed by Lemma 8, the problem amounts to finding an asymptotic approximation for $a_n^T \mathbb{E} Q(z) b_n$. It can be easily seen by injecting the approximation in (84) into (82) that the contribution of the diagonal elements, given by $\sum_{k=1}^n a_{k,n} b_{k,n} \mathbb{E}\left[Q(z)\right]_{kk}$ can be approximated by $m(z) a^T b$. It remains thus to study the contribution of the off-diagonal elements which we denote by:

$$\Upsilon(a_n, b_n, z) = \sum_{k=1}^n \sum_{r \neq k} a_{k,n} b_{r,n} \mathbb{E}[[Q(z)]_{kr}] \tag{85}$$

where $a_{k,n}$ and $b_{r,n}$ refers to the $k$-th and $r$-th elements of vectors $a_n$ and $b_n$ respectively. This is performed in three steps. In a first step, we establish an equation between $\Upsilon(a_n, b_n, z)$ and the quantities $\{\tilde{\alpha}_{r,j}(z)\}_{r,j=1}^n$ defined as:

$$\tilde{\alpha}_{r,j}(z) = \mathbb{E}\Big[ \sum_{k \neq j} \frac{1}{\sqrt{p}} \left( x_k^T C_{[r]} x_k - \frac{1}{p} \operatorname{tr}(C_{[r]} C_{[k]}) \right) Q_{kj} \Big] \tag{86}$$

In the second step, we establish an equation between $\tilde{\alpha}_{r,j}$ and $\Upsilon(\frac{1_n}{\sqrt{p}}, e_j, z)$ where $e_j$ is the $j$-th canonical vector of $\mathbb{R}^n$. Gathering these results, we obtain a linear equation whose solution is a deterministic equivalent for $\Upsilon(\frac{1_n}{\sqrt{p}}, b_n, z)$. Plugging this deterministic equivalent back into the relations obtained in the first and the second step, we finally derive a deterministic equivalent for $\Upsilon(a_n, b_n, z)$ and $\tilde{\alpha}_{r,j}(z)$.

#### *5.2.1. Step 1: Expression for $\Upsilon(a_n, b_n, z)$*

**Proposition 1.** *Let $\{a_n\}$ and $\{b_n\}$ be two sequences of vectors in $\mathbb{C}^{n\times 1}$ with bounded Euclidean norm. Then for any $z \in \mathbb{C} \backslash \mathbb{R}$ and any small positive $\epsilon$*

$$\Upsilon(a_n, b_n, z) = -\sum_{k=1}^n \sum_{r \neq k} a_{k,n} b_{r,n} \mathbb{E}\left[Q_{kk}\right] \tilde{\alpha}_{k,r}(z) + O_z\left(p^{-\frac{1}{2}+\epsilon}\right). \tag{87}$$

*Moreover, if $b_n$ is such that $\sum_{k=1}^n |b_{n,k}|$ is uniformly bounded in $n$, then (87) becomes:*

$$\Upsilon(a_n, b_n, z) = -\sum_{k=1}^n \sum_{r \neq k} a_{k,n} b_{r,n} \mathbb{E}\left[Q_{kk}\right] \tilde{\alpha}_{k,r}(z) + O_z\left(p^{-1+\epsilon}\right). \tag{88}$$

*Proof.* See Appendix B.1 □

*5.2.2. Step 2: Expression for $\tilde{\alpha}_{r,j}(z)$*

**Proposition 2.** *Let $r$ and $j$ be two integers in $\{1,\ldots,n\}$. The following approximation holds true:*

$$\tilde{\alpha}_{r,j}(z) = -\frac{2}{p^{\frac{3}{2}}}\sum_{k=1}^{n}\mathbb{E}(Q_{kk})\Upsilon(\frac{1_n}{\sqrt{p}}, e_j, z)\frac{1}{p}\operatorname{tr}((C^\circ)^4) - \frac{2}{p^2}\sum_{k=1}^{n}\mathbb{E}[Q_{kk}]\mathbb{E}[Q_{jj}]\frac{1}{p}\operatorname{tr}((C^\circ)^4) + O_z(p^{-\frac{5}{4}}). \tag{89}$$

*Proof.* See Appendix B.2 □

*5.2.3. Step 3: Asymptotic equivalents for $\tilde{\alpha}_{r,j}$ and $\Upsilon(a_n, b_n, z)$*

**Proposition 3.** *The following approximations hold true:*

$$\Upsilon(a_n, b_n, z) = \frac{1}{p}\frac{c_0 m^3(z)\Omega^2 a_n^T 1_n 1_n^T b_n}{1 - c_0^2\Omega^2 m^2(z)} + O_z(p^{-\frac{1}{4}}) \tag{90}$$

$$\tilde{\alpha}_{r,j}(z) = -\frac{c_0}{p}\frac{m^2(z)\Omega^2}{1 - c_0^2\Omega^2 m^2(z)} + O_z(p^{-\frac{5}{4}}), \quad r,j = 1,\cdots,n \tag{91}$$

If $b_n$ is such that $\sum_{k=1}^n |b_{n,k}|$ is uniformly bounded in $n$, then (90) becomes:

$$\Upsilon(a_n, b_n, z) = \frac{1}{p}\frac{c_0 m^3(z)\Omega^2 a_n^T 1_n 1_n^T b_n}{1 - c_0^2\Omega^2 m^2(z)} + O_z(p^{-\frac{3}{4}}) \tag{92}$$

*Proof.* Combining (89) and (87), we obtain:

$$\Upsilon(a_n, b_n, z) = \frac{2}{p^{\frac{3}{2}}}\sum_{k=1}^{n}\sum_{r\neq k} a_{k,n}b_{r,n}\mathbb{E}(Q_{kk})\sum_{\ell=1}^{n}\mathbb{E}(Q_{ll})\Upsilon(\frac{1_n}{\sqrt{p}}, e_r, z)\frac{1}{p}\operatorname{tr}((C^\circ)^4) + \frac{2}{p^2}\sum_{k=1}^{n}\sum_{r\neq k} a_{k,n}b_{r,n}\mathbb{E}[Q_{kk}]\sum_{l=1}^{n}\mathbb{E}[Q_{ll}]\mathbb{E}[Q_{rr}]\frac{1}{p}\operatorname{tr}((C^\circ)^4) + O_z(p^{-\frac{1}{4}}) \tag{93}$$

$$= \frac{2}{p^{\frac{3}{2}}}\sum_{k=1}^{n} a_{k,n}\mathbb{E}[Q_{kk}]\sum_{\ell=1}^{n}\mathbb{E}[Q_{ll}]\Upsilon(\frac{1_n}{\sqrt{p}}, b_n, z)\frac{1}{p}\operatorname{tr}((C^\circ)^4) + \frac{2}{p^2}\sum_{k=1}^{n}\sum_{r=1}^{n} a_{k,n}b_{r,n}\mathbb{E}[Q_{kk}]\sum_{l=1}^{n}\mathbb{E}[Q_{ll}]\mathbb{E}[Q_{rr}]\frac{1}{p}\operatorname{tr}((C^\circ)^4) + O_z(p^{-\frac{1}{4}}) \tag{94}$$

Particularizing (94) for $a_n = \frac{1_n}{\sqrt{p}}$, we obtain:

$$\Upsilon(\frac{1_n}{\sqrt{p}}, b_n, z) = \frac{2}{p^2}\sum_{k=1}^{n}\mathbb{E}(Q_{kk})\sum_{\ell=1}^{n}\mathbb{E}(Q_{ll})\Upsilon(\frac{1_n}{\sqrt{p}}, b_n, z)\frac{1}{p}\mathrm{tr}((C^\circ)^4)$$
$$+ \frac{2}{p^{\frac{5}{2}}}\sum_{k=1}^{n}\mathbb{E}[Q_{kk}]\sum_{l=1}^{n}\mathbb{E}[Q_{ll}]\sum_{r=1}^{n}b_{r,n}\mathbb{E}[Q_{rr}]\frac{1}{p}\mathrm{tr}((C^\circ)^4) + O_z(p^{-\frac{1}{4}}) \tag{95}$$
$$= c_0^2(g_n(z))^2\Upsilon(\frac{1_n}{\sqrt{p}}, b_n, z)\Omega^2 + c_0^2\Omega^2(g_n(z))^2\frac{1}{\sqrt{p}}\sum_{r=1}^{n}b_{r,n}\mathbb{E}[Q_{rr}]$$
$$+ O_z(p^{-\frac{1}{4}}) \tag{96}$$

Using (82) and (84), we can easily see that:

$$\mathbb{E}[Q_{rr}] - m(z) = O_z(p^{-\frac{1}{2}+\epsilon})$$

Thus,

$$\left(1 - c_0^2\Omega^2 m^2(z)\right)\Upsilon(\frac{1_n}{\sqrt{p}}, b_n, z) = c_0^2\Omega^2 m^3(z)\frac{1_n^T}{\sqrt{p}}b_n + O_z(p^{-\frac{1}{4}})$$

Invoking Lemma 7, it can be shown that:

$$\left(1 - c_0^2\Omega^2 m^2(z)\right)^{-1} = O_z(1)$$

Hence,

$$\Upsilon(\frac{1_n}{\sqrt{p}}, b_n, z) = \frac{c_0^2\Omega^2 m^3(z)}{1 - c_0^2\Omega^2 m^2(z)}\frac{1_n^T}{\sqrt{p}}b_n + O_z(p^{-\frac{1}{4}}). \tag{97}$$

Plugging (97) back into (94), we thus obtain:

$$\Upsilon(a_n, b_n, z) = \frac{1}{p}\frac{c_0\Omega^2 m^3(z) a_n^T 1_n 1_n^T b_n}{1 - c_0^2\Omega^2 m^2(z)} + O_z(p^{-\frac{1}{4}})$$

The proof of (92) follows along the same lines by using the approximation (88) of Proposition (1), while that of (91) follows by plugging the approximations in (92) into (89). □

*5.2.4. Concluding.*

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

### 5.3. Proof of Theorem 3

For the sake of simplification, we remove the subscript $n$ from the notation of $a_n$, $b_n$ and $D_n$. Following the same kind of calculations as in the proof of Lemma 8, we can show that for any $z_1, z_2 \in \mathbb{C}\backslash\mathbb{R}$, and $s \in \mathbb{N}^*$,

$$\mathbb{E}\Big[\Big|a^T Q(z_1)DQ(z_2)b - \mathbb{E}\Big[a^T Q(z_1)DQ(z_2)b\Big|^{2s}\Big] = O_z(p^{-s+\epsilon})$$

Therefore, the proof of Theorem 3 amounts to finding an asymptotically deterministic equivalent for $\mathbb{E}\Big[a^T Q(z_1)DQ(z_2)b\Big]$. To begin with, we expand it as:

$$\begin{aligned}\mathbb{E}\Big[a^T Q(z_1)DQ(z_2)b\Big] &= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \mathbb{E}\Big[a_i[Q(z_1)]_{ik}D_{kk}[Q(z_2)]_{kj}b_j\Big] \quad (98)\\ &= Z_1 + Z_2 + Z_3 + Z_4 \quad (99)\end{aligned}$$

where

$$\begin{aligned}Z_1 &= \sum_{i=1}^n \sum_{j=1}^n \sum_{k\notin\{i,j\}} \mathbb{E}\left[a_i\left[Q(z_1)\right]_{ik} D_{kk}\left[Q(z_2)\right]_{kj} b_j\right]\\ Z_2 &= \sum_{j=1}^n \sum_{i\neq j} \mathbb{E}\left[a_i\left[Q(z_1)\right]_{ii} D_{ii}\left[Q(z_2)\right]_{ij} b_j\right]\\ Z_3 &= \sum_{j=1}^n \sum_{i\neq j} \mathbb{E}\left[a_i\left[Q(z_1)\right]_{ij} D_{jj}\left[Q(z_2)\right]_{jj} b_j\right]\\ Z_4 &= \sum_{i=1}^n \mathbb{E}\left[a_i\left[Q(z_1)\right]_{ii} D_{ii}\left[Q(z_2)\right]_{ii} b_i\right]\end{aligned}$$

Using the fact that $\mathbb{E}[[Q(z)]_{ii}] - m(z) = O_z(p^{-\frac{1}{2}+\epsilon})$, we approximate $Z_4$ as:

$$Z_4 = m(z_1)m(z_2)a^T Db + O_z(p^{-\frac{1}{2}+\epsilon}) \quad (100)$$

while $Z_2$ can be treated as follows:

$$\begin{aligned}Z_2 &= \sum_{j=1}^n \sum_{i\neq j} a_i\mathbb{E}\left[\left[\left[Q(z_1)\right]_{ii} - \mathbb{E}\left[\left[Q(z_1)\right]_{ii}\right]\right] D_{ii}\left[Q(z_2)\right]_{ij} b_j\right]\\ &\quad + \sum_{j=1}^n \sum_{i\neq j} a_i\mathbb{E}\left[\left[Q(z_1)\right]_{ii}\right] D_{ii}\mathbb{E}\left[\left[Q(z_2)\right]_{ij}\right] b_j\\ &= \sum_{j=1}^n \sum_{i\neq j} a_i\mathbb{E}\left[\left[Q(z_1)\right]_{ii}\right] D_{ii}\mathbb{E}\left[\left[Q(z_2)\right]_{ij}\right] b_j + O_z(p^{-\frac{1}{2}+\epsilon})\end{aligned}$$

since, applying Cauchy-Schwartz Lemma on the first term of the above equation, we obtain:

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

Recalling the definition of $\Upsilon(a_n, b_n, z)$ in (85), we may write $Z_2$ as:

$$\begin{aligned}
Z_2 &= \sum_{i=1}^{n} a_i D_{ii}\mathbb{E}[[Q(z_1)]_{ii}]\Upsilon(e_i, b, z_2) + O_z(p^{-\frac{1}{2}+\epsilon}) \\
&= \sum_{i=1}^{n} a_i D_{ii}\Upsilon(e_i, b, z_2)m(z_1) + O_z(p^{-\frac{1}{2}+\epsilon})
\end{aligned}$$

Using (92), we thus obtain:

$$Z_2 = \frac{c_0\Omega^2 m^3(z_2)m(z_1)}{1-\Omega^2 c_0^2 m^2(z_2)} a^T D\frac{1_n 1_n^T}{p}b + O_z(p^{-\frac{1}{4}}) \tag{101}$$

Similarly, we can prove that:

$$Z_3 = \frac{c_0\Omega^2 m^3(z_1)m(z_2)}{1-c_0^2\Omega^2 m^2(z_1)} a^T\frac{1_n 1_n^T}{p}Db + O_z(p^{-\frac{1}{4}}) \tag{102}$$

It remains thus to treat the quantity $Z_1$. Using (33), we get:

$$\begin{aligned}
Z_1 &= \sum_{k=1}^{n}\mathbb{E}\Big[D_{kk}\xi_{(k,-k)}^T Q_k(z_1)ab^T Q_k(z_2)\xi_{(k,-k)}[Q(z_1)]_{kk}[Q(z_2)]_{kk}\Big] \\
&= m(z_1)m(z_2)\sum_{k=1}^{n} D_{kk}\mathbb{E}\Big[\xi_{(k,-k)}^T Q_k(z_1)ab^T Q_k(z_2)\xi_{(k,-k)}\Big] + O_z(p^{-\frac{1}{2}+\epsilon}) \\
&= m(z_1)m(z_2)\sum_{k=1}^{n} D_{kk}\mathbb{E}\Big[d_k^T Q_k(z_1)ab^T Q_k(z_2)d_k\Big] \\
&+ m(z_1)m(z_2)\frac{2}{p}\sum_{k=1}^{n} D_{kk}\mathbb{E}\Big[a^T Q_k(z_1)\Sigma_k Q_k(z_2)b\Big] + O_z(p^{-\frac{1}{2}+\epsilon})
\end{aligned}$$

where the last equality follows from Lemma 6 and $\Sigma_k$ is defined in (39). To treat the first term, we will make use of the following statements

$$\mathbb{E}\Big[\left|d_k^T Q_k(z)c\right|^2\Big] = O_z(p^{-1+\epsilon}) \tag{103}$$

$$\operatorname{var}(d_k^T Q_k(z)c) = O_z(p^{-2+\epsilon}) \tag{104}$$

where $c \in \mathbb{C}^{n\times 1}$ is deterministic with bounded norm and $z \in \mathbb{C}\backslash\mathbb{R}$. The proof of (103) follows by noting that $\|d_k\|^2 = O(p^{-1})$ whereas the proof of (104) is based on standard calculations using the Poincaré-Nash inequality and is thus omitted. Based on (103) and (104), we approximate the first term in $Z_1$ (with an error $O_z(p^{-\frac{1}{2}+\epsilon})$) as follows

$$\begin{aligned}
&m(z_1)m(z_2)\sum_{k=1}^{n} D_{kk}\mathbb{E}\Big[d_k^T Q_k(z_1)ab^T Q_k(z_2)d_k\Big] \\
&= m(z_1)m(z_2)\sum_{k=1}^{n} D_{kk}\mathbb{E}\left[d_k^T Q_k(z_1)a\right]\mathbb{E}\left[b^T Q_k(z_2)d_k\right] + O_z(p^{-\frac{1}{2}+\epsilon}) \qquad (105)\\
&= m(z_1)m(z_2)\sum_{k=1}^{n} D_{kk}\mathbb{E}\left[d_k^T Q(z_1)a\right]\mathbb{E}\left[b^T Q(z_2)d_k\right] + O_z(p^{-\frac{1}{2}+\epsilon}) \qquad (106)
\end{aligned}$$

where the last equality follows from Lemma 13. To continue, we note that the following relation holds true

$$\mathbb{E}\Big[c^T Q(z)d_k\Big] = \sum_{i=1}^{n} c_i\tilde{\alpha}_{k,i}(z) + O_z(p^{-1+\epsilon}) \tag{107}$$

with $z \in \mathbb{C}\backslash\mathbb{R}$, $c$ being a $n\times 1$ vector with bounded norm and $\tilde{\alpha}_{k,i}(z)$ defined in (86). To prove it, we shall expand $\mathbb{E}\Big[c^T Q(z)d_k\Big]$ as:

$$\begin{aligned}
&\mathbb{E}\Big[c^T Q(z)d_k\Big] = \sum_{i=1}^{n} c_i \sum_{j\neq k}\frac{1}{\sqrt{p}}\mathbb{E}\Big[\left(x_j^T C_{[k]}x_j - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[j]})\right)[Q(z)]_{ij}\Big] \qquad (108)\\
&= \sum_{i=1}^{n} c_i \sum_{j\neq i}\frac{1}{\sqrt{p}}\mathbb{E}\Big[\left(x_j^T C_{[k]}x_j - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[j]})\right)[Q(z)]_{ij}\Big] \\
&+ \sum_{i=1}^{n} c_i \frac{1}{\sqrt{p}}\mathbb{E}\Big[\left(x_i^T C_{[k]}x_i - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[i]})\right)([Q(z)]_{ii} - \mathbb{E}[[Q(z)]_{ii}])\Big] \\
&- \sum_{i=1}^{n} c_i \frac{1}{\sqrt{p}}\mathbb{E}\Big[\left(x_k^T C_{[k]}x_k - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[k]})\right)[Q(z)]_{ik}\Big] \qquad (109)
\end{aligned}$$

and note that the two last quantities in (109) are $O_z(p^{-1+\epsilon})$. With (107) at hand, the first term in $Z_1$ can be apprroximated by:

$$m(z_1)m(z_2)\sum_{k=1}^{n} D_{kk}\mathbb{E}\Big[d_k^T Q_k(z_1)ab^T Q_k(z_2)d_k\Big] \tag{110}$$

$$= m(z_1)m(z_2)\sum_{k=1}^{n} D_{kk}\sum_{i=1}^{n} a_i\tilde{\alpha}_{k,i}(z_1)\sum_{j=1}^{n} b_j\tilde{\alpha}_{k,j}(z_2) + O_z(p^{-\frac{1}{2}+\epsilon}) \tag{111}$$

We may now invoke Proposition 3 to replace $\tilde{\alpha}_{k,i}(z_1)$ and $\tilde{\alpha}_{k,j}(z_2)$ by their asymptotic equivalents. In doing so, we obtain:

$$\begin{aligned} & m(z_1)m(z_2)\sum_{k=1}^{n} D_{kk}\mathbb{E}\Big[d_k^T Q_k(z_1) ab^T Q_k(z_2) d_k\Big] \\ & = \frac{1}{p}\operatorname{tr}(D)\frac{c_0^2 m^3(z_1) m^3(z_2)\Omega^4 a^T \frac{1_n 1_n^T}{p} b}{(1-c_0^2\Omega^2 m^2(z_2))(1-c_0^2\Omega^2 m^2(z_1))} + O_z(p^{-\frac{1}{4}}) \end{aligned} \tag{112}$$

It remains now to handle the second term in $Z_1$. Adapting the calculations of [11] to our setting (Page 14-20 in [11]), we can prove that:

$$\|\Sigma_k - \frac{1}{p^2}\operatorname{tr}((C^\circ)^4)(1_n 1_n^T - I_n) - (\frac{1}{p}\operatorname{tr}((C^\circ)^2))^2 I_n\| = O(p^{-\frac{1}{4}}).$$

Hence,

$$m(z_1)m(z_2)\frac{2}{p}\sum_{k=1}^{n} D_{kk}\mathbb{E}\Big[a^T Q_k(z_1)\Sigma_k Q_k(z_2) b\Big] \tag{113}$$

$$\begin{aligned} & = m(z_1)m(z_2)\frac{2}{p}\sum_{k=1}^{n} D_{kk}\left(\frac{1}{p}\operatorname{tr}((C^\circ)^2)\right)^2 \mathbb{E}\Big[a^T Q_k(z_1) Q_k(z_2) b\Big] \\ & + m(z_1)m(z_2)\frac{2}{p}\sum_{k=1}^{n} D_{kk}\left(\frac{1}{p}\operatorname{tr}((C^\circ)^4)\right)\mathbb{E}\Big[a^T Q_k(z_1)\frac{1_n 1_n^T}{p} Q_k(z_2) b\Big] + O_z(p^{-\frac{1}{4}}) \end{aligned} \tag{114}$$

Using Lemma 13 and the fact that

$$\max(\operatorname{var}(a^T Q(z)\frac{1_n}{\sqrt{p}}), \operatorname{var}(b^T Q(z)\frac{1_n}{\sqrt{p}}) = O_z(p^{-1+\epsilon}),$$

we thus obtain:

$$m(z_1)m(z_2)\frac{2}{p}\sum_{k=1}^{n} D_{kk}\mathbb{E}\Big[a^T Q_k(z_1)\Sigma_k Q_k(z_2) b\Big] \tag{115}$$

$$\begin{aligned} & = m(z_1)m(z_2)\frac{2}{p}\sum_{k=1}^{n} D_{kk}\big(\frac{1}{p}\operatorname{tr}((C^\circ)^4)\big)\frac{1}{p}\mathbb{E}\Big[a^T Q(z_1) 1_n\Big]\mathbb{E}\Big[1_n^T Q(z_2) b\Big] \\ & + m(z_1)m(z_2)\frac{2}{p}\sum_{k=1}^{n} D_{kk}\big(\frac{1}{p}\operatorname{tr}((C^\circ)^2)\big)^2\mathbb{E}\Big[a^T Q(z_1) Q(z_2) b\Big] + O_z(p^{-\frac{1}{4}}) \\ & = m(z_1)m(z_2)\frac{2}{p}\operatorname{tr}(D)\big(\frac{1}{p}\operatorname{tr}((C^\circ)^2)\big)^2\mathbb{E}\Big[a^T Q(z_1) Q(z_2) b\Big] \\ & + \frac{m^2(z_1)m^2(z_2)\frac{1}{p}\operatorname{tr}(D)\Omega^2}{p(1-\Omega^2 c_0^2 m^2(z_1))(1-\Omega^2 c_0^2 m^2(z_2))} a^T 1_n 1_n^T b + O_z(p^{-\frac{1}{4}}) \end{aligned} \tag{116}$$

where (116) follows by using Theorem 2. Combining (112) and (116), we conclude that:

$$Z_1 = m(z_1)m(z_2)\omega^2 \frac{1}{p}\operatorname{tr}(D)\mathbb{E}\Big[a^T Q(z_1)Q(z_2)b\Big] + \frac{m^2(z_1)m^2(z_2)\frac{1}{p}\operatorname{tr}(D)\Omega^2 a^T 1_n 1_n^T b}{p(1-\Omega^2 c_0^2 m^2(z_1))(1-\Omega^2 c_0^2 m^2(z_2))} + \frac{\frac{1}{p}\operatorname{tr}(D)c_0^2 m^3(z_1)m^3(z_2)\Omega^4 a^T \frac{1_n 1_n^T}{p} b}{(1-c_0^2\Omega^2 m^2(z_1))(1-c_0^2\Omega^2 m^2(z_2))} + O_z(p^{-\frac{1}{4}}) \tag{117}$$

Combining (117), (101), (102) and (100), we obtain:

$$\mathbb{E}\Big[a^T Q(z_1)DQ(z_2)b\Big] = m(z_1)m(z_2)\frac{1}{p}\operatorname{tr}(D)\omega^2\mathbb{E}\Big[a^T Q(z_1)Q(z_2)b\Big] + \frac{m^2(z_1)m^2(z_2)\frac{1}{p}\operatorname{tr}(D)\Omega^2 a^T 1_n 1_n^T b}{p(1-\Omega^2 c_0^2 m^2(z_1))(1-\Omega^2 c_0^2 m^2(z_2))} + \frac{c_0^2\Omega^4\frac{1}{p}\operatorname{tr}(D)m^3(z_1)m^3(z_2)a^T\frac{1_n 1_n^T}{p}b}{(1-c_0^2\Omega^2 m^2(z_1))(1-c_0^2\Omega^2 m^2(z_2))} + \frac{c_0\Omega^2 m^3(z_2)m(z_1)a^T D\frac{1_n 1_n^T}{p}b}{1-\Omega^2 c_0^2 m^2(z_2)} + \frac{c_0\Omega^2 m^3(z_1)m(z_2)a^T\frac{1_n 1_n^T}{p}Db}{1-c_0^2\Omega^2 m^2(z_1)} + m(z_1)m(z_2)a^T Db + O_z(p^{-\frac{1}{4}}) \tag{118}$$

Setting $D = I_n$, we obtain:

$$\mathbb{E}\Big[a^T Q(z_1)Q(z_2)b\Big] = \omega^2 c_0 m(z_1)m(z_2)\mathbb{E}\Big[a^T Q(z_1)Q(z_2)b\Big] + m(z_1)m(z_2)a^T b \tag{119}$$

$$+ m(z_1)m(z_2)c_0\Omega^2\Big[m^2(z_1)+m^2(z_2)+m(z_1)m(z_2)-c_0^2\Omega^2 m^2(z_1)m^2(z_2)\Big] \times\Big(1-\Omega^2c_0^2m^2(z_1)\Big)^{-1}\Big(1-\Omega^2c_0^2m^2(z_2)\Big)^{-1}\frac{1}{p}a^T 1_n 1_n^T b + O_z(p^{-\frac{1}{4}}) \tag{120}$$

thus yielding:

$$\mathbb{E}\Big[a^T Q(z_1)Q(z_2)b\Big] = g(z_1,z_2) + O_z(p^{-\frac{1}{4}}) \tag{121}$$

with

$$g(z_1,z_2) = (1-\omega^2 c_0 m(z_1)m(z_2))^{-1}m(z_1)m(z_2)a^T b + m(z_1)m(z_2)c_0\Omega^2\Big[m^2(z_1)+m^2(z_2)+m(z_1)m(z_2)-c_0^2\Omega^2m^2(z_1)m^2(z_2)\Big] \times\big(1-\Omega^2c_0^2m^2(z_1)\big)^{-1}\big(1-\Omega^2c_0^2m^2(z_2)\big)^{-1}(1-\omega^2c_0m(z_1)m(z_2))^{-1}\frac{1}{p}a^T1_n1_n^Tb \tag{122}$$

Plugging (121) into (118), we get:

$$\mathbb{E}\Big[a^T Q(z_1)DQ(z_2)b\Big] = m(z_1)m(z_2)a^T Db + m(z_1)m(z_2)\omega^2\frac{1}{p}\operatorname{tr}(D)g(z_1,z_2)$$

$$+\tilde{r}(z_1,z_2)+O_z(p^{-\frac{1}{4}}) \tag{123}$$

where

$$\tilde{r}(z_1,z_2)=\frac{c_0\Omega^2 m^3(z_2)m(z_1)a^T D\frac{1_n 1_n^T}{p}b}{1-\Omega^2 c_0^2 m^2(z_2)}+\frac{c_0\Omega^2 m^3(z_1)m(z_2)a^T\frac{1_n 1_n^T}{p}Db}{1-\Omega^2 c_0^2 m^2(z_1)}$$
$$+\frac{\frac{1}{p}\operatorname{tr}(D)\Omega^2 m^2(z_1)m^2(z_2)(1+c_0^2\Omega^2 m(z_1)m(z_2))a^T\frac{1_n 1_n^T}{p}b}{\Big(1-\Omega^2 c_0^2 m^2(z_1)\Big)\Big(1-\Omega^2 c_0^2 m^2(z_2)\Big)}. \tag{124}$$

### 5.4. Almost sure location of the eigenvalues of $\Phi$ (Proof of Theorem 4)

The goal of Theorem 4 is to characterize the location of the eigenvalues of $\Phi$ in the asymptotic regime. To this end, we will resort to the tools developed in [13], which consists in analyzing the difference $g_n(z)-m(z)$. If this difference converges to zero faster than $O(p^{-1})$, then it can be proven under other mild assumptions that all the eigenvalues are almost surely located in the neighborhood of the limiting support $\mathcal{S}=\left[-2\sqrt{c_0}\omega,2\sqrt{c_0}\omega\right]$. Unfortunately, this does not hold in our case, since $g_n(z)-m(z)=O_z(p^{-\frac{1}{2}+\epsilon})$. The analysis of the location of the eigenvalues becomes thus less trivial and requires a deeper investigation of the difference $g_n(z)-m(z)$. As a matter of fact building on the ideas of [5, 16, 19], the characterization of the location of eigenvalues of $\Phi$ requires us to investigate the behavior of each term in the difference $g_n(z)-m(z)$ that converges slower than $O(p^{-1})$. More specifically, we consider showing that

$$g_n(z)-m(z)=\frac{1}{p^{\frac{1}{2}}}\tilde{f}(z)+\frac{1}{p^{\frac{3}{4}}}\tilde{h}(z)+\frac{1}{p}\tilde{k}(z)+O_z(p^{-\frac{5}{4}}) \tag{125}$$

where $\tilde{f}(z)$, $\tilde{h}(z)$ and $\tilde{k}(z)$ should be determined in terms of $m(z)$ (and not in terms of elements of $\mathbb{E}(Q(z))$. The key idea behind the technique of [5] consists in proving that $\tilde{f}(z)$, $\tilde{h}(z)$ and $\tilde{k}(z)$ are Stieltjes transforms of some distributions and characterizing their associated supports. It turns out that in our case, the supports of the distributions associated with the Stieltjes transforms $\tilde{f}(z)$, $\tilde{h}(z)$ are included in $\mathcal{S}$ while the support of that of $\tilde{k}(z)$ may present two spikes outside $\mathcal{S}$. As will be shown next, this will imply that the support of the limiting eigenvalue distribution of $\Phi$ is $\mathcal{S}$ plus possibly the two spikes that arise in the support of the distribution of Stieltjes transform $\tilde{k}(z)$. Proving (125) is the heart matter of the proof Theorem 4. To pave the way towards this, we need to derive deterministic equivalents of some quantities that will appear in our derivations. This is performed in the following next section.

#### 5.4.1. Some preliminaries

**Lemma 14.** *Let $k,b$ be two integers in $\{1,\cdots,n\}$ such that $b\neq k$. Let $A_{1,p}$ and $A_{2,p}$ be two sequences of $p\times p$ deterministic matrices with uniformly bounded* *spectral norms. Let $z \in \mathbb{C}\backslash\mathbb{R}$. Then,*

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

*Proof.* Using the relation $Q_{bk} = -e_b^T Q_k \xi_{(k,-k)} Q_{kk}$ when $b \neq k$ and invoking Lemma 10 and Lemma 8, we obtain:

$$\mathbb{E}\left[x_b^T A_{1,p} x_k x_b^T A_{2,p} x_k Q_{bk}\right] = -\mathbb{E}\left[x_b^T A_{1,p} x_k x_b^T A_{2,p} x_k e_b^T Q_k \xi_{(k,-k)} Q_{kk}\right] \quad (126)$$

$$= -\mathbb{E}\big[x_b^T A_{1,p} x_k x_b^T A_{2,p} x_k e_b^T Q_k \xi_{(k,-k)}\big]\mathbb{E}Q_{kk} + O_z(p^{-2+\epsilon}) \quad (127)$$

$$= -\mathbb{E}\big[Q_{kk}\big]\mathbb{E}\Big[x_b^T A_{1,p} x_k x_b^T A_{2,p} x_k \sum_{l \neq k} \sqrt{p}\,[Q_k]_{bl} \left((x_k^T x_l)^2 - \frac{1}{p} x_l^T C_{[k]} x_l\right)\Big]$$
$$- \mathbb{E}\big[Q_{kk}\big]\mathbb{E}\big[p^{-\frac{3}{2}} x_b^T A_{1,p} C_{[k]} A_{2,p} x_b \sum_{l \neq k} [Q_k]_{bl} \left(x_l^T C_{[k]} x_l - \frac{1}{p} \operatorname{tr} C_{[k]} C_{[l]}\right)\big]$$
$$+ O_z(p^{-2+\epsilon}) \quad (128)$$

$$= -\,\mathbb{E}\Big[Q_{kk}\Big] \frac{2}{p^{\frac{3}{2}}} \sum_{l \neq k} \mathbb{E}\big[\left[Q_k\right]_{bl} x_b^T A_{2,p} C_{[k]} x_l x_l^T C_{[k]} A_{1,p} x_b\big] + O_z(p^{-2+\epsilon}) \quad (129)$$

$$= -\,\mathbb{E}[Q_{kk}] \frac{2}{p^{\frac{3}{2}}} \sum_{l \neq k} \mathbb{E}\left[[Q]_{bl}\, x_b^T A_{2,p} C_{[k]} x_l x_l^T C_{[k]} A_{1,p} x_b\right] + O_z(p^{-2+\epsilon}) \quad (130)$$

$$= -\frac{2}{p^{\frac{3}{2}}} m^2(z) \frac{1}{p} \operatorname{tr}\left(A_{2,p} \left(C^\circ\right)^2\right) \frac{1}{p} \operatorname{tr}\left(\left(C^\circ\right)^2 A_{1,p}\right) + O_z(p^{-\frac{7}{4}}) \quad (131)$$

The second term in (128) can be proven to be $O_z(p^{-2+\epsilon})$ by applying Lemma 9 as in the proof of Corollary 3. Equality (130) in which $[Q_k]_{bl}$ is replaced by $[Q]_{bl}$ follows from Lemma 13, while (131) follows by noticing that the term obtained by taking $b = l$ is the most dominant. □

**Lemma 15.** *Let $j, k \in \{1, \cdots, n\}$ such that $j \neq k$. Let $A_{1,p}$, $A_{2,p}$ and $A_{3,p}$ be three sequences of $p \times p$ deterministic matrices with spectral norms bounded uniformly in $p$. Then,*

$$\sum_{b \notin \{j,k\}} \mathbb{E}\left[x_k^T A_{1,p} x_j x_b^T A_{2,p} x_j x_b^T A_{3,p} x_k Q_{bj}\right] = -2np^{-\frac{5}{2}} m^2(z) \frac{1}{p} \operatorname{tr}\left(\left(C^\circ\right)^2 A_{2,p}\right) \frac{1}{p} \operatorname{tr}\left(C^\circ A_{1,p} \left(C^\circ\right)^2 A_{3,p}\right) + O_z(p^{-\frac{7}{4}})$$

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

$$\sum_{b \notin \{j,k\}} \mathbb{E}\left[x_k^T A_{1,p} x_j x_b^T A_{2,p} x_j x_b^T A_{3,p} x_k Q_{bj}\right] = -\sum_{b \notin \{j,k\}} \mathbb{E}\left[x_k^T A_{1,p} x_j x_b^T A_{2,p} x_j x_b^T A_{3,p} x_k e_b^T Q_j \xi_{(j,-j)} Q_{jj}\right]$$

$$= -\sum_{b\notin\{j,k\}} \mathbb{E}\left[x_k^T A_{1,p} x_j x_b^T A_{2,p} x_j x_b^T A_{3,p} x_k e_b^T Q_j \xi_{(j,-j)}\right] \mathbb{E}Q_{jj} + O_z(p^{-2+\epsilon}) \tag{132}$$

where equation (132) follows by using the relation $e_b^T Q_j \xi_{(j,-j)} = \frac{-Q_{bj}}{Q_{jj}}$ to write:

$$\sum_{b\notin\{j,k\}} \mathbb{E}\left[x_k^T A_{1,p} x_j x_b^T A_{2,p} x_j x_b^T A_{3,p} x_k e_b^T Q_j \xi_{(j,-j)}(Q_{jj} - \mathbb{E}(Q_{jj}))\right]$$
$$= -\sum_{b\notin\{j,k\}} \mathbb{E}\left[x_k^T A_{1,p} x_j x_b^T A_{2,p} x_j x_b^T A_{3,p} x_k \frac{Q_{bj}}{Q_{jj}}(Q_{jj} - \mathbb{E}(Q_{jj}))\right] \tag{133}$$

and then applying Lemma 11 and Lemma 8 to prove the desired. Next, expanding $\xi_{(j,-j)}$, we get:

$$\sum_{b\notin\{j,k\}} \mathbb{E}\left[x_k^T A_{1,p} x_j x_b^T A_{2,p} x_j x_b^T A_{3,p} x_k Q_{bj}\right]$$
$$= \sum_{b\notin\{j,k\}} -\sqrt{p}\mathbb{E}\big[x_k^T A_{1,p} x_j x_b^T A_{2,p} x_j x_b^T A_{3,p} x_k \sum_{l\neq j} [Q_j]_{bl}\big((x_j^T x_l)^2 - \frac{1}{p} x_l^T C_{[j]} x_l\big)\big]\mathbb{E}Q_{jj}$$
$$- \sum_{b\notin\{j,k\}} \mathbb{E}\big[x_k^T A_{1,p} x_j x_b^T A_{2,p} x_j x_b^T A_{3,p} x_k \sum_{l\neq j} [Q_j]_{bl}\big(\frac{x_l^T C_{[j]} x_l}{\sqrt{p}} - \frac{1}{p^{\frac{3}{2}}} \operatorname{tr} C_{[j]} C_{[l]}\big)\big]\mathbb{E}Q_{jj}$$
$$+ O_z(p^{-2+\epsilon})$$
$$= -\sum_{b\notin\{j,k\}} \mathbb{E}\big[x_b^T A_{2,p} C_{[j]} A_{1,p} x_k x_b^T A_{3,p} x_k \sum_{l\neq j} [Q_j]_{bl}\big(\frac{x_l^T C_{[j]} x_l}{p^{\frac{3}{2}}} - \frac{1}{p^{\frac{5}{2}}} \operatorname{tr} C_{[j]} C_{[l]}\big)\big]\mathbb{E}Q_{jj}$$
$$- \sum_{b\notin\{j,k\}} \sum_{l\neq j} 2\mathbb{E}\big[[Q_j]_{bl} \frac{x_b^T A_{2,p} C_{[j]} x_l}{p^{\frac{3}{2}}} x_l^T C_{[j]} A_{1,p} x_k x_b^T A_{3,p} x_k\big]\mathbb{E}Q_{jj} + O_z(p^{-2+\epsilon}) \tag{134}$$

$$= -\sum_{b\notin\{j,k\}} 2\mathbb{E}\left[[Q_j]_{bb} \frac{x_b^T A_{3,p} x_k}{p^{\frac{3}{2}}} x_b^T A_{2,p} C_{[j]} x_b x_b^T C_{[j]} A_{1,p} x_k\right] \mathbb{E}Q_{jj} + O_z(p^{-2+\epsilon}) \tag{135}$$

$$= -\sum_{b\neq\{j,k\}} 2\mathbb{E}\,[Q_j]_{bb} \frac{1}{p^{\frac{7}{2}}} \operatorname{tr}\left((C^\circ)^2 A_{2,p}\right) \frac{1}{p} \operatorname{tr}\left(C^\circ A_{1,p} (C^\circ)^2 A_{3,p}\right) \mathbb{E}Q_{jj} + O_z(p^{-\frac{7}{4}}) \tag{136}$$

$$= -2np^{-\frac{5}{2}} m^2(z) \frac{1}{p} \operatorname{tr}\left((C^\circ)^2 A_{2,p}\right) \frac{1}{p} \operatorname{tr}\left(C^\circ A_{1,p} (C^\circ)^2 A_{3,p}\right) + O_z(p^{-\frac{7}{4}}) \tag{137}$$

Equation (134) follows by taking the expectation with respect to $x_j$. In equation (135), we used Lemma 11 to show that the summand of the first term in (134) over $b \neq l$ is $O_z(p^{-2+\epsilon})$, and handled the second term as in the proof of Corollary 3 by using Lemma 9 to show that it is $O_z(p^{-2+\epsilon})$. Next, to obtain equation (136), we use the same arguments as in the proof of Corollary 3 to interpret the diagonal elements of $Q_j$ as those of another resolvent matrix formed by discarding the observation $x_j$. This allows us to obtain $\mathrm{var}([Q_j]_{bb}) = O(p^{-\frac{1}{2}+\epsilon})$ from Lemma 8 and hence up to an error $O_z(p^{-2+\epsilon})$, $[Q_j]_{bb}$ can be replaced by its expectation. Finally, we use Lemma 13 to replace $Q_j$ by $Q$ and obtain the desired by taking the expectation with respect to the distribution of $x_b$ and $x_k$, and then using (11). $\square$

### 5.4.2. *Precise estimation of the approximation $g_n(z) - m(z)$*

With these Lemmas at hand, we are now in position to prove the estimation in (125). To this end, recall the relation involving the diagonal elements of $Q$:

$$z\mathbb{E}Q_{jj} = -1 + \alpha_j(z) + \beta_j(z) + \gamma_j(z) + \theta_j(z)$$

where $\alpha_j(z), \beta_j(z), \gamma_j(z)$ and $\theta_j(z)$ are given by (73)-(76). To prove (125), we shall first derive asymptotic equivalents that approximate all these quantities up to an error of order $O_z(p^{-\frac{5}{4}})$.

**Asymptotic equivalent for $\alpha_j(z)$.** Recall that:

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

From Proposition 3, it unfolds that:

$$\alpha_j(z) = -\frac{c_0}{p}\frac{m^2(z)\Omega^2}{1 - c_0^2\Omega^2 m^2(z)} + O_z(p^{-\frac{5}{4}})$$

**Asymptotic equivalent for $\beta_j(z)$.** Using the Integration by Parts formula, we decompose $\beta_j(z)$ as:

$$\beta_j(z) = \beta_{j,1}(z) + \beta_{j,2}(z) + \beta_{j,3}(z), \tag{138}$$

where

$$\beta_{j,1}(z) = -\frac{2}{p}\sum_{k\neq j}\mathbb{E}\left[(x_k^T C_{[j]} x_k)^2 Q_{kk} Q_{jj}\right]$$

$$\beta_{j,2}(z) = \frac{8}{\sqrt{p}}\sum_{k\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^T C_{[j]} x_k x_k^T x_j x_k^T C_{[j]} x_b x_b^T x_j Q_{jj} Q_{bj} Q_{kk}\right]$$

$$\beta_{j,3}(z) = \frac{8}{\sqrt{p}}\sum_{k\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^T C_{[j]} x_k x_k^T x_j x_k^T C_{[j]} x_b x_b^T x_j Q_{kj} Q_{bk} Q_{jj}\right]$$

By distinguishing the cases $b = k$ and $b \notin \{k, j\}$, we may decompose $\beta_{j,2}(z)$ as:

$$\beta_{j,2}(z) = \frac{8}{\sqrt{p}}\sum_{k\neq j}\sum_{b\notin\{k,j\}}\mathbb{E}\left[x_k^T C_{[j]} x_k x_k^T x_j x_k^T C_{[j]} x_b x_b^T x_j Q_{bj} Q_{jj} Q_{kk}\right]$$

$$+ \frac{8}{\sqrt{p}} \sum_{k \neq j} \mathbb{E}\left[\left(x_k^T C_{[j]} x_k\right)^2 \left(x_k^T x_j\right)^2 Q_{kj} Q_{jj} Q_{kk}\right]$$

Using the facts that

$$\begin{aligned} (x_k^T C_{[j]} x_k)^2 &= \left(\frac{1}{p}\mathrm{tr}(C_{[k]}C_{[j]})\right)^2 + O(p^{-\frac{1}{2}}) = \left(\frac{1}{p}\,\mathrm{tr}((C^\circ)^2)\right)^2 + O(p^{-\frac{1}{2}}) \\ x_k^T C_{[j]} x_k &= \frac{1}{p}\mathrm{tr}(C_{[k]}C_{[j]}) + O(p^{-\frac{1}{2}}) = \frac{1}{p}\,\mathrm{tr}((C^\circ)^2) + O(p^{-\frac{1}{2}}) \end{aligned} \tag{139}$$

together with Lemma 11, Lemma 14 and Lemma 15, we obtain:

$$\begin{aligned} \beta_{j,2}(z) &= \frac{8}{\sqrt{p}} \sum_{k \neq j} \frac{1}{p}\,\mathrm{tr}\left((C^\circ)^2\right) \mathbb{E}Q_{kk}\mathbb{E}Q_{jj} \sum_{b \notin \{k,j\}} \mathbb{E}\left[x_k^T x_j x_k^T C_{[j]} x_b x_b^T x_j Q_{bj}\right] \\ &+ \frac{8}{\sqrt{p}} \left(\frac{1}{p}\,\mathrm{tr}\,(C^\circ)^2\right)^2 \mathbb{E}Q_{jj} \sum_{k \neq j} \mathbb{E}Q_{kk}\mathbb{E}\left[(x_k^T x_j)^2 Q_{kj}\right] + O_z(p^{-\frac{3}{2}}) \\ &= -16n^2p^{-3}\left(\frac{1}{p}\,\mathrm{tr}\left(C^\circ\right)^2\right)^2 \frac{1}{p}\,\mathrm{tr}\left((C^\circ)^4\right) m^4(z) - 16np^{-2}\left(\frac{1}{p}\,\mathrm{tr}\,(C^\circ)^2\right)^4 m^4(z) \\ &+ O_z(p^{-\frac{5}{4}}) \end{aligned} \tag{140}$$

As for $\beta_{j,3}(z)$, we can see from Lemma 11 that the contribution of the summand over $b \neq \{j,k\}$ is $O_z(p^{-\frac{3}{2}+\epsilon})$. This leads to:

$$\begin{aligned} \beta_{j,3}(z) &= \frac{8}{\sqrt{p}} \sum_{k \neq j} \mathbb{E}\left[\left(x_k^T C_j x_k\right)^2 (x_k^T x_j)^2 Q_{kj} Q_{jj} Q_{kk}\right] + O_z(p^{-\frac{3}{2}+\epsilon}) \\ &= \frac{8}{\sqrt{p}} \sum_{k=1}^{n} \left(\frac{1}{p}\,\mathrm{tr}\left((C^\circ)^2\right)^2\right) \mathbb{E}Q_{kk}\mathbb{E}Q_{jj}\mathbb{E}\left[(x_k^T x_j)^2 Q_{kj}\right] + O_z(p^{-\frac{3}{2}+\epsilon}) \\ &= -16np^{-2}\left(\frac{1}{p}\,\mathrm{tr}\,(C^\circ)^2\right)^4 m^4(z) + O_z(p^{-\frac{5}{4}}) \end{aligned} \tag{141}$$

where the second equality follows from (139) and Lemma 8, while the last equality follows from Lemma 14.

It remains thus to handle the term $\beta_{j,1}(z)$. For that, we apply the Integration by Parts formula to obtain:

$$\beta_{j,2}(z) = \upsilon_1 + \upsilon_2 + \upsilon_3 + \upsilon_4$$

$$\begin{aligned} \upsilon_1 &= -\frac{2}{p} \sum_{k \neq j} \frac{1}{p}\,\mathrm{tr}\, C_{[k]} C_{[j]} \mathbb{E}\left[x_k^T C_{[j]} x_k Q_{kk} Q_{jj}\right] \\ \upsilon_2 &= -\frac{4}{p^2} \sum_{k \neq j} \mathbb{E}\left[x_k^T C_{[j]} C_{[k]} C_{[j]} x_k Q_{kk} Q_{jj}\right] \\ \upsilon_3 &= \frac{8}{p\sqrt{p}} \sum_{k \neq j} \sum_{b \neq k} \mathbb{E}\left[x_k^T C_{[j]} x_k x_k^T C_{[j]} C_{[k]} x_b x_b^T x_k Q_{jk} Q_{bj} Q_{kk}\right] \end{aligned}$$

$$\upsilon_4 = \frac{8}{p\sqrt{p}} \sum_{k\neq j} \sum_{b\neq k} \mathbb{E}\left[x_k^T C_{[j]} x_k x_k^T C_{[j]} C_{[k]} x_b x_b^T x_k Q_{bk} Q_{kk} Q_{jj}\right]$$

The term $\upsilon_1$ can be treated by applying again the Integration by Parts formula, thus leading to:

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

The first term in $\upsilon_1$ can be decomposed as:

$$\begin{aligned}&-\frac{2}{p} \sum_{k\neq j} \left(\frac{1}{p}\operatorname{tr} C_{[k]} C_{[j]}\right)^2 \mathbb{E}\left[Q_{kk} Q_{jj}\right] = -\frac{2}{p} \sum_{k=1}^{n} \left(\frac{1}{p}\operatorname{tr} C_{[k]} C_{[j]}\right)^2 \mathbb{E}\left[Q_{kk} Q_{jj}\right] \\ &+ \frac{2}{p}\left(\frac{1}{p}\operatorname{tr} C_{[j]}^2\right)^2 \mathbb{E} Q_{jj}^2 \\ &= -\frac{2}{p} \sum_{k=1}^{n} \left(\frac{1}{p}\operatorname{tr} C_{[k]} C_{[j]}\right)^2 \mathbb{E}\left[\left(Q_{kk} - \mathbb{E}Q_{kk}\right)\left(Q_{jj} - \mathbb{E}Q_{jj}\right)\right] \\ &- \frac{2}{p} \sum_{k=1}^{n} \left(\frac{1}{p}\operatorname{tr} C_{[k]} C_{[j]}\right)^2 \mathbb{E}Q_{kk}\mathbb{E}Q_{jj} + \frac{2}{p}\left(\frac{1}{p}\operatorname{tr}\left(C_{[j]}^2\right)\right)^2 \mathbb{E} Q_{jj}^2 \qquad (142)\end{aligned}$$

Using Cauchy-Schwartz inequality and the variance control in Lemma 8, we obtain:

$$\frac{2}{p} \sum_{k=1}^{n} \left(\frac{1}{p}\operatorname{tr} C_{[k]} C_{[j]}\right)^2 \mathbb{E}\left[\left(Q_{kk} - \mathbb{E}Q_{kk}\right)\left(Q_{jj} - \mathbb{E}Q_{jj}\right)\right] = O_z(p^{-\frac{3}{2}+\epsilon}) \qquad (143)$$

On the other hand, it follows from Lemma 8 that:

$$\frac{2}{p}\left(\frac{1}{p}\operatorname{tr}\left(C_{[j]}^2\right)\right)^2 \mathbb{E} Q_{jj}^2 = \frac{2}{p}\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\left(\mathbb{E}Q_{jj}\right)^2 + O_z(p^{-\frac{3}{2}+\epsilon}) \qquad (144)$$

Plugging (143) and (144) into (142), we get:

$$\begin{aligned}&-\frac{2}{p} \sum_{k\neq j} \left(\frac{1}{p}\operatorname{tr} C_{[k]} C_{[j]}\right)^2 \mathbb{E}\left[Q_{kk} Q_{jj}\right] \qquad (145) \\ &= -\omega^2 c_0 g_n(z) \mathbb{E}Q_{jj} + \frac{2}{p}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^2 \left(\mathbb{E}Q_{jj}\right)^2 \\ &- \frac{1}{p} \sum_{k=1}^{n} \left[2\left(\frac{1}{p}\operatorname{tr} C_{[k]} C_{[j]}\right)^2 - \omega^2\right] \mathbb{E}Q_{jj}\mathbb{E}Q_{kk} + O_z(p^{-\frac{3}{2}+\epsilon}) \qquad (146)\end{aligned}$$

The last term in $\upsilon_1$ can be shown $O_z(p^{-2+\epsilon})$ using the result of Lemma 11, while the second term can be treated by Lemma 14 to yield:

$$\frac{8}{p\sqrt{p}}\sum_{k\neq j}\frac{1}{p}\operatorname{tr}C_{[k]}C_{[j]}\sum_{b\neq k}\mathbb{E}\left[x_b^T x_k x_b^T C_{[k]}C_{[j]}x_k Q_{bk}Q_{kk}Q_{jj}\right]$$
$$= -16n^2p^{-3}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^2\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^4(z)+O_z(p^{-\frac{5}{4}}) \quad (147)$$

Combining (146) and (147), we thus obtain:

$$\upsilon_1 = -\omega^2 c_0 g_n(z)\mathbb{E}Q_{jj} - \frac{1}{p}\sum_{k=1}^{n}\left[2\left(\frac{1}{p}\operatorname{tr}C_{[k]}C_{[j]}\right)^2-\omega^2\right]\mathbb{E}Q_{jj}\mathbb{E}Q_{kk} \quad (148)$$
$$+\frac{2}{p}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^2m^2(z) - 16n^2p^{-3}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^2\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^4(z)$$
$$+O_z(p^{-\frac{5}{4}}) \quad (149)$$

Similarly, using Lemma 11, we can easily see that $\upsilon_3 = O_z(p^{-2+\epsilon})$, while following the same approach as before, we can prove that $\upsilon_2$ and $\upsilon_4$ can be approximated as:

$$\upsilon_2 = -\frac{4n}{p^2}\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^2(z)+O_z(p^{-\frac{5}{4}}) \quad (150)$$

$$\upsilon_4 = -16p^{-3}n^2\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^2\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^4(z)+O_z(p^{-\frac{5}{4}}) \quad (151)$$

Combining (149), (150) and (151), we thus get:

$$\beta_{j,1}(z) = -\omega^2c_0g_n(z)\mathbb{E}Q_{jj} - \frac{1}{p}\sum_{k=1}^{n}\left[2\left(\frac{1}{p}\operatorname{tr}C_{[k]}C_{[j]}\right)^2-\omega^2\right]\mathbb{E}Q_{jj}\mathbb{E}Q_{kk}$$
$$+\frac{2}{p}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^2m^2(z) - \frac{4n}{p^2}\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^2(z)$$
$$-32n^2p^{-3}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^2\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^4(z)+O_z(p^{-\frac{5}{4}}) \quad (152)$$

Plugging (152), (140), (141) into (138) leads to:

$$\beta_j(z) = -\omega^2c_0g_n(z)\mathbb{E}Q_{jj} - \frac{1}{p}\sum_{k=1}^{n}\left[2\left(\frac{1}{p}\operatorname{tr}C_{[k]}C_{[j]}\right)^2-\omega^2\right]\mathbb{E}Q_{jj}\mathbb{E}Q_{kk}$$
$$+\frac{2}{p}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^2m^2(z) - \frac{4n}{p^2}\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^2(z)$$
$$-48n^2p^{-3}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^2\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^4(z)$$

$$- 32np^{-2}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^4 m^4(z) + O_z(p^{-\frac{5}{4}})$$
$$= -\omega^2 c_0 g_n(z)\mathbb{E}Q_{jj} - \frac{1}{p}\sum_{k=1}^{n}\left[2\left(\frac{1}{p}\operatorname{tr} C_{[k]}C_{[j]}\right)^2 - \omega^2\right]\mathbb{E}Q_{jj}\mathbb{E}Q_{kk} + \frac{1}{p}\omega^2 m^2(z)$$
$$- \frac{4n}{p^2}\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^2(z) - \frac{24}{p}c_0^2\omega^2\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^4(z)$$
$$- 8c_0p^{-1}\omega^4 m^4(z) + O_z(p^{-\frac{5}{4}})$$

**Asymptotic equivalent for $\theta_j(z)$.** Using the Integration by Parts formula, we decompose $\theta_j(z)$ as:

$$\theta_j(z) = \theta_{j,1}(z) + \theta_{j,2}(z) + \theta_{j,3}(z) + \theta_{j,4}(z) \tag{153}$$

where

$$\theta_{j,1}(z) = -2\sum_{k\neq j}\sum_{r\notin\{j,k\}}\mathbb{E}\left[\frac{1}{p}\left(x_r^TC_{[j]}x_k\right)^2 Q_{rk}Q_{jj}\right]$$
$$\theta_{j,2}(z) = \frac{6}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{j,k\}}\sum_{b\neq j}\mathbb{E}\left[x_b^Tx_jx_k^TC_{[j]}x_rx_b^TC_{[j]}x_kx_r^Tx_jQ_{rk}Q_{bj}Q_{jj}\right]$$
$$\theta_{j,3}(z) = \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{j,k\}}\sum_{b\neq j}\mathbb{E}\left[x_k^TC_{[j]}x_rx_r^Tx_jx_k^TC_{[j]}x_bx_b^Tx_jQ_{rj}Q_{bk}Q_{jj}\right]$$
$$\theta_{j,4}(z) = \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{j,k\}}\sum_{b\neq j}\mathbb{E}\left[x_k^TC_{[j]}x_rx_r^Tx_jx_k^TC_{[j]}x_bx_b^Tx_jQ_{jk}Q_{br}Q_{jj}\right]$$

Based on Lemma 11, we can see that the contribution of the sum over $b\notin\{k,j\}$ in $\theta_{j,2}(z)$ is $O_z(p^{-\frac{3}{2}+\epsilon})$. Hence,

$$\theta_{j,2} = \frac{6}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{j,k\}}\mathbb{E}\left[x_k^Tx_jx_k^TC_{[j]}x_rx_r^Tx_jQ_{rk}x_k^TC_{[j]}x_kQ_{kj}Q_{jj}\right] + O_z(p^{-\frac{3}{2}+\epsilon})$$
$$= O_z(p^{-\frac{3}{2}+\epsilon})$$

where the last equality follows from Lemma 11 along with the fact that for $k\neq j$, $x_k^Tx_j = O(p^{-\frac{1}{2}})$ and $\mathbb{E}|Q_{kj}|^2 = O_z(p^{-\frac{1}{2}+\epsilon})$ by Lemma 8. Similarly, it follows from Lemma 11 that the contribution of the sum over $b\notin\{k,j\}$ in $\theta_{j,3}(z)$ is $O_z(p^{-\frac{3}{2}+\epsilon})$, which leads to:

$$\theta_{j,3}(z) = \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{k,j\}}\mathbb{E}\left[x_k^TC_{[j]}x_rx_r^Tx_jx_k^Tx_jQ_{rj}\right]\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\mathbb{E}Q_{kk}\mathbb{E}Q_{jj}$$
$$+ O_z(p^{-\frac{3}{2}+\epsilon})$$
$$= -8n^2p^{-3}m^4(z)\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^2\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right) + O_z(p^{-\frac{5}{4}}) \tag{154}$$

$$= -4c_0^2 \frac{1}{p}\omega^2 \frac{1}{p} \operatorname{tr}\left((C^\circ)^4\right) m^4(z) + O_z(p^{-\frac{5}{4}}) \tag{155}$$

where (154) follows from Lemma 15. The quantity $\theta_{j,4}$ can be shown to be $O_z(p^{-\frac{3}{2}+\epsilon})$. To see this, we first decompose it as:

$$\begin{aligned}\theta_{j,4}(z) &= \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{r\notin\{j,k\}} \sum_{b\notin\{j,r,k\}} \mathbb{E}\left[x_k^T C_{[j]} x_r x_r^T x_j x_k^T C_{[j]} x_b x_b^T x_j Q_{jk} Q_{br} Q_{jj}\right] \\ &+ \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{r\notin\{j,k\}} \mathbb{E}\left[(x_k^T C_{[j]} x_r)^2 (x_r^T x_j)^2 Q_{jk} Q_{rr} Q_{jj}\right] \\ &+ \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{r\notin\{j,k\}} \mathbb{E}\left[x_k^T C_{[j]} x_r x_r^T x_j Q_{kr} x_k^T C_{[j]} x_k x_k^T x_j Q_{jk} Q_{jj}\right]\end{aligned}$$

The first term is $O_z(p^{-\frac{3}{2}+\epsilon})$ as per Lemma 12. The second and third terms can also be shown to be $O_z(p^{-2+\epsilon})$ and $O_z(p^{-\frac{3}{2}+\epsilon})$ respectively using Lemma 11.

It remains to deal with $\theta_{j,1}(z)$. Based on the Integration by Parts formula, $\theta_{j,1}(z)$ can be decomposed as:

$$\begin{aligned}\theta_{j,1}(z) = -2 &\sum_{k\neq j} \sum_{r\notin\{k,j\}} \frac{1}{p^2} \mathbb{E}\left[Q_{rk} Q_{jj} x_r^T C_{[j]} C_{[k]} C_{[j]} x_r\right] \\ &+ 4 \sum_{k\neq j} \sum_{r\notin\{j,k\}} \sum_{b\notin\{k,r\}} \frac{1}{p^{\frac{3}{2}}} \mathbb{E}\left[x_b^T x_k x_b^T C_{[k]} C_{[j]} x_r x_r^T C_{[j]} x_k Q_{rk} Q_{bk} Q_{jj}\right] \\ &+ 4 \sum_{k\neq j} \sum_{r\notin\{k,j\}} \frac{1}{p^{\frac{3}{2}}} \mathbb{E}\left[x_r^T x_k x_r^T C_{[k]} C_{[j]} x_r x_r^T C_{[j]} x_k Q_{rk}^2 Q_{jj}\right] \\ &+ 4 \sum_{k\neq j} \sum_{r\notin\{k,j\}} \sum_{b\notin\{k,r\}} \frac{1}{p^{\frac{3}{2}}} \mathbb{E}\left[x_b^T x_k x_b^T C_{[k]} C_{[j]} x_r x_r^T C_{[j]} x_k Q_{kk} Q_{br} Q_{jj}\right] \\ &+ 4 \sum_{k\neq j} \sum_{r\notin\{k,j\}} \frac{1}{p^{\frac{3}{2}}} \mathbb{E}\left[x_r^T x_k x_r^T C_{[k]} C_{[j]} x_r x_r^T C_{[j]} x_k Q_{kk} Q_{rr} Q_{jj}\right] \\ &+ 8 \sum_{k\neq j} \sum_{r\notin\{k,j\}} \sum_{b\neq k} \frac{1}{p^{\frac{3}{2}}} \mathbb{E}\left[x_b^T x_k x_b^T C_{[k]} C_{[j]} x_r x_r^T C_{[j]} x_k Q_{rk} Q_{jk} Q_{bj}\right]\end{aligned}$$

Based on Lemma 11 and the variance evaluations in Lemma 8, we can prove that the first, fourth and fifth terms, which we denote by $\theta_{j,1,1}(z)$, $\theta_{j,1,4}(z)$ and $\theta_{j,1,5}(z)$ are the dominant ones, while all other terms are $O_z(p^{-\frac{3}{2}+\epsilon})$. To handle the first term, we note that we may replace in the first term $Q_{jj}$ by $m(z)$ and $x_r^T C_{[j]} C_{[k]} C_{[r]} x_r$ by $\frac{1}{p}\operatorname{tr}((C^\circ)^4))$ with an error $O_z(p^{-\frac{5}{4}})$. Hence, using Proposition 3, we obtain:

$$\theta_{j,1,1}(z) = -2 \sum_{k\neq j} \sum_{r\notin\{k,j\}} \frac{1}{p^2} \mathbb{E}\left[Q_{rk} Q_{jj} x_r^T C_{[j]} C_{[k]} C_{[r]} x_r\right]$$

$$= -\frac{4}{p}\frac{n^3}{p^3}\frac{\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)\right)^2 m^4(z)}{1-\frac{2n^2}{p^2}\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^2(z)} + O_z(p^{-\frac{5}{4}}) \tag{156}$$

On the other hand, we may use Lemma 11 and Lemma 15 to approximate $\theta_{j,1,4}(z)$ as:

$$\begin{aligned}
\theta_{j,1,4}(z) &= 4\sum_{k\neq j}\sum_{r\notin\{k,j\}}\sum_{b\notin\{k,r\}}\frac{1}{p^{\frac{3}{2}}}\mathbb{E}\left[x_b^T x_k x_b^T C_{[k]}C_{[j]}x_r x_r^T C_{[j]}x_k Q_{kk}Q_{br}Q_{jj}\right] \\
&= 4\sum_{k\neq j}\sum_{r\notin\{k,j\}}\sum_{b\notin\{k,r\}} m^2(z)p^{-\frac{3}{2}}\mathbb{E}\left[x_b^T x_k x_b^T C_{[k]}C_{[j]}x_r x_r^T C_{[j]}x_k Q_{br}\right] \\
&+ O_z(p^{-\frac{3}{2}+\epsilon}) \\
&= -8n^3p^{-4}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)\right)^2 m^4(z) + O_z(p^{-\frac{5}{4}})
\end{aligned} \tag{157}$$

Finally, to treat $\theta_{j,1,5}(z)$ we use the Integration by Parts formula and follow the same kind of approximation as before to obtain:

$$\begin{aligned}
\theta_{j,1,5}(z) &= 4\sum_{k\neq j}\sum_{r\notin\{k,j\}}\frac{1}{p^{\frac{3}{2}}}\mathbb{E}\left[x_r^T x_k x_r^T C_{[k]}C_{[j]}x_r x_r^T C_{[j]}x_k Q_{kk}Q_{rr}Q_{jj}\right] \\
&= 4\sum_{k\neq j}\sum_{r\notin\{k,j\}} p^{-\frac{5}{2}}\mathbb{E}\left[\left(x_r^T C_{[k]}C_{[j]}x_r\right)^2 Q_{kk}Q_{rr}Q_{jj}\right] + O_z(p^{-\frac{3}{2}+\epsilon}) \\
&= 4\sum_{k=1}^{n}\sum_{r=1}^{n} p^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr}C_{[r]}C_{[k]}C_{[j]}\right)^2\mathbb{E}\left[Q_{kk}Q_{rr}Q_{jj}\right] + O_z(p^{-\frac{3}{2}+\epsilon}) \\
&= 4\sum_{k=1}^{n}\sum_{r=1}^{n} p^{-\frac{5}{2}}\big(\frac{1}{p}\operatorname{tr}C_{[r]}C_{[k]}C_{[j]}\big)^2\mathbb{E}\left[Q_{kk}\right]\mathbb{E}\left[Q_{rr}\right]\mathbb{E}\left[Q_{jj}\right] + O_z(p^{-\frac{3}{2}+\epsilon})
\end{aligned} \tag{158}$$

where the last equality follows from the variance control in Lemma 8. Note that we cannot replace the diagonal elements of the resolvent matrix by $m(z)$ or the covariance matrices by $C^\circ$, since this would produce an error of order $O_z(p^{-\frac{3}{4}})$ which is bigger than $O(p^{-\frac{5}{4}})$. Combining (156), (157) and (158), we obtain

$$\begin{aligned}
\theta_{j,1}(z) &= -\frac{4}{p}c_0^3\frac{\big(\frac{1}{p}\operatorname{tr}\big((C^\circ)^4\big)\big)^2 m^4(z)}{1-\frac{2n^2}{p^2}\frac{1}{p}\operatorname{tr}\big((C^\circ)^4\big)m^2(z)} - 8c_0^3\frac{1}{p}\big(\frac{1}{p}\operatorname{tr}\big((C^\circ)^4\big)\big)^2 m^4(z) \\
&+ 4\sum_{k=1}^{n}\sum_{r=1}^{n} p^{-\frac{5}{2}}\big(\frac{1}{p}\operatorname{tr}C_{[r]}C_{[k]}C_{[j]}\big)^2\mathbb{E}\left[Q_{kk}\right]\mathbb{E}\left[Q_{rr}\right]\mathbb{E}\left[Q_{jj}\right] + O_z(p^{-\frac{5}{4}})
\end{aligned} \tag{159}$$

Using the fact that the dominant terms in $\theta_j(z)$ are $\theta_{j,1}(z)$ and $\theta_{j,3}(z)$, we get:

$$\theta_j(z) = -4c_0^2p^{-1}\omega^2\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^4(z) - \frac{4c_0^3}{p}\frac{\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)\right)^2 m^4(z)}{1-\frac{2n^2}{p^2}\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^2(z)}$$

$$+ 4\sum_{k=1}^{n}\sum_{r=1}^{n} p^{-\frac{5}{2}}\left(\frac{1}{p}\operatorname{tr} C_{[r]}C_{[k]}C_{[j]}\right)^2 \mathbb{E}Q_{kk}\mathbb{E}Q_{rr}\mathbb{E}Q_{jj}$$

$$- 8c_0^3 p^{-1}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)\right)^2 m^4(z) + O_z(p^{-\frac{5}{4}})$$

**Asymptotic equivalent for $\gamma_j(z)$.** Using the Integration by Parts formula, we may decompose $\gamma_j(z)$ as:

$$\gamma_j(z) = \gamma_{j,1}(z) + \gamma_{j,2}(z) + \gamma_{j,3}(z) + \gamma_{j,4}(z) + \gamma_{j,5}(z)$$

where

$$\gamma_{j,1}(z) = -2\sum_{k\neq j}\sum_{r\neq j}\mathbb{E}\left[\frac{1}{p}\left(x_r^T C_{[j]}x_k\right)^2 Q_{jk}Q_{rj}\right]$$

$$\gamma_{j,2}(z) = \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^T C_{[j]}x_r x_b^T x_j x_b^T C_{[j]}x_k x_r^T x_j Q_{jj}Q_{bk}Q_{rj}\right]$$

$$\gamma_{j,3}(z) = \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^T C_{[j]}x_r x_b^T x_j x_b^T C_{[j]}x_k x_r^T x_j Q_{jk}Q_{bj}Q_{rj}\right]$$

$$\gamma_{j,4}(z) = \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^T C_{[j]}x_r x_b^T x_j x_b^T C_{[j]}x_k x_r^T x_j Q_{rj}Q_{bj}Q_{jk}\right]$$

$$\gamma_{j,5}(z) = \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^T C_{[j]}x_r x_b^T x_j x_b^T C_{[j]}x_k x_r^T x_j Q_{jj}Q_{br}Q_{jk}\right]$$

We will start by handling $\gamma_{j,2}(z)$. Using Lemma 11, it is easy to see that summand over indexes $b \notin \{j,k\}$ and $r \notin \{j,k\}$ is $O_z(p^{-\frac{3}{2}+\epsilon})$. Hence,

$$\gamma_{j,2}(z) = \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{r\notin\{k,j\}}\mathbb{E}\left[x_k^T C_{[j]}x_r x_k^T x_j x_k^T C_{[j]}x_k x_r^T x_j Q_{jj}Q_{kk}Q_{rj}\right]$$

$$+ \frac{4}{\sqrt{p}}\sum_{k\neq j}\sum_{b\neq j}\mathbb{E}\left[x_k^T C_{[j]}x_k x_b^T x_j x_b^T C_{[j]}x_k x_k^T x_j Q_{jj}Q_{bk}Q_{kj}\right] + O_z(p^{-\frac{3}{2}+\epsilon})$$

$$= \frac{4}{\sqrt{p}}\sum_{k\neq j} m^2(z)\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\sum_{r\notin\{j,k\}}\mathbb{E}\left[x_k^T C_{[j]}x_r x_k^T x_j x_r^T x_j Q_{rj}\right]$$

$$+ \frac{4}{\sqrt{p}}\sum_{k\neq j}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^2 m^2(z)\mathbb{E}\left[(x_k^T x_j)^2 Q_{kj}\right] + O_z(p^{-\frac{3}{2}+\epsilon})$$

$$= -8n^2p^{-3}\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\right)^2\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^4(z)$$

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

where the last equality follows from Lemma 14 and Lemma 15. Using Lemma 11, we can see that $\gamma_{j,3}(z)$ and $\gamma_{j,4}(z)$ are $O_z(p^{-\frac{3}{2}+\epsilon})$. It remains thus to treat the term $\gamma_{j,5}(z)$. Again, using Lemma 11, it can be shown that the summand over indexes $r \notin \{j,k\}$ and $b \notin \{j,r,k\}$ is $O_z(p^{-\frac{3}{2}+\epsilon})$, thus leading to:

$$\begin{aligned}
\gamma_{j,5}(z) &= \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{b \neq j} \mathbb{E}\left[x_k^T C_{[j]} x_k x_b^T x_j x_b^T C_{[j]} x_k x_k^T x_j Q_{bk} Q_{jj} Q_{jk}\right] \\
&+ \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{r \notin \{j,k\}} \mathbb{E}\left[x_k^T C_{[j]} x_r x_k^T x_j x_k^T C_{[j]} x_k x_r^T x_j Q_{rk} Q_{jj} Q_{jk}\right] \\
&+ \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{r \notin \{j,k\}} \mathbb{E}\left[(x_k^T C_{[j]} x_r)^2 (x_r^T x_j)^2 Q_{jj} Q_{rr} Q_{jk}\right] + O_z(p^{-\frac{3}{2}+\epsilon}) \\
&= \frac{4}{\sqrt{p}} \sum_{k\neq j} \mathbb{E}\left[(x_k^T C_{[j]} x_k)^2 (x_k^T x_j)^2 Q_{kk} Q_{jj} Q_{jk}\right] \\
&+ \frac{4}{\sqrt{p}} \sum_{k\neq j} \sum_{r \notin \{j,k\}} \mathbb{E}\left[(x_k^T C_{[j]} x_r)^2 (x_r^T x_j)^2 Q_{jk} Q_{jj} Q_{rr}\right] + O_z(p^{-\frac{3}{2}+\epsilon}) \\
&= \frac{4}{\sqrt{p}} \sum_{k\neq j} \mathbb{E}\left[(x_k^T C_{[j]} x_k)^2 (x_k^T x_j)^2 Q_{kk} Q_{jj} Q_{jk}\right] + O_z(p^{-\frac{3}{2}+\epsilon})
\end{aligned}$$

where the last equality follows from Lemma 11. By Lemma 15,

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

Thus,

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

It remains thus to handle $\gamma_{j,1}(z)$. We may use Lemma 11 to show that the summand over index $r \notin \{j,k\}$ is $O_z(p^{-\frac{3}{2}+\epsilon})$. This leads to:

$$\begin{aligned}
\gamma_{j,1}(z) &= -2\sum_{k\neq j} \mathbb{E}\big[\frac{1}{p}\left(x_k^T C_{[j]} x_k\right)^2 Q_{jk}^2\big] + O_z(p^{-\frac{3}{2}+\epsilon}) \\
&= -\frac{2}{p}\big(\frac{1}{p}\operatorname{tr}(C^\circ)^2\big)^2 \sum_{k\neq j} \mathbb{E}\left[Q_{jk}^2\right] + O_z(p^{-\frac{3}{2}+\epsilon}) \\
&= -\frac{2}{p}\big(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\big)^2 \left(\mathbb{E}\left[[Q^2]_{jj}\right] - \mathbb{E}\left[(Q_{jj})^2\right]\right) + O_z(p^{-\frac{3}{2}+\epsilon})
\end{aligned}$$

Using Lemma 8, we can replace $Q_{jj}$ by $m(z)$ with an error $O_z(p^{-\frac{3}{2}+\epsilon})$, thus giving:

$$\gamma_{j,1}(z) = -\frac{2}{p}\big(\frac{1}{p}\operatorname{tr}\big((C^\circ)^2\big)\big)^2 \mathbb{E}\left[Q^2\right]_{jj} + \frac{2}{p} m^2(z)\big(\frac{1}{p}\operatorname{tr}\left((C^\circ)^2\right)\big)^2 + O_z(p^{-\frac{3}{2}+\epsilon})$$

It remains to find approximate equivalents 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 $\gamma_{j,1}(z)$ and using the asymptotic approximations of $\gamma_{j,2}(z)$ and $\gamma_{j,5}(z)$, we ultimately get:

$$\begin{aligned}\gamma_j(z) = &-\frac{1}{p}\frac{\omega^2 m^2(z)}{1 - c_0\omega^2 m^2(z)} + \frac{1}{p}\omega^2 m^2(z) - \frac{4}{p}c_0^2\omega^2\frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right)m^4(z)\\ &- \frac{4c_0}{p}\omega^4 m^4(z) + O_z(p^{-\frac{5}{4}})\end{aligned}$$

**Proof of** (125)**.** For $j = 1, \ldots, n$, based on the asymptotic approximations for $\alpha_j(z)$, $\beta_j(z)$, $\theta_j(z)$ and $\gamma_j(z)$, we can write $z\mathbb{E}Q_{jj}$ as:

$$z\mathbb{E}Q_{jj} = -1 - \omega^2 c_0 \mathbb{E}Q_{jj} g_n(z) + A_{p^{-\frac{1}{2}},j}(z) + A_{p^{-1}}(z) + O_z(p^{-\frac{5}{4}}),\ \ j = 1, \ldots, n. \tag{160}$$

where

$$\begin{aligned}A_{p^{-\frac{1}{2}},j}(z) &= -\frac{1}{p}\sum_{k=1}^{n}\left[2\big(\frac{1}{p}\operatorname{tr} C_{[k]}C_{[j]}\big)^2 - \omega^2\right]\mathbb{E}Q_{jj}\mathbb{E}Q_{kk}\\ &+ 4\sum_{k=1}^{n}\sum_{r=1}^{n} p^{-\frac{5}{2}}\big(\frac{1}{p}\operatorname{tr} C_{[r]}C_{[k]}C_{[j]}\big)^2\mathbb{E}Q_{kk}\mathbb{E}Q_{rr}\mathbb{E}Q_{jj}\\ A_{p^{-1}}(z) &= \frac{\omega^2 m^2(z) - 14c_0\omega^4 m^4(z) + 12c_0^2\omega^6 m^6(z)}{1 - \omega^2 c_0 m^2(z)}\\ &- 2c_0^3 p^{-1}\Omega^4 m^4(z) - 16\frac{c_0^2}{p}\omega^2\Omega^2 m^4(z) + \frac{\frac{c_0^3}{p}\Omega^4 m^4(z) - \frac{3c_0}{p}\Omega^2 m^2(z)}{1 - c_0^2\Omega^2 m^2(z)}\end{aligned}$$

We should recall that in (125), functions $\tilde{f}(z)$, $\tilde{h}(z)$ and $\tilde{k}(z)$ should be expressed solely in terms of $m(z)$ and not in terms of elements of $\mathbb{E}Q$. However, replacing diagonal elements of $\mathbb{E}Q$ by $m(z)$ could not help to identify these functions, as this would result in an error bigger than $O(p^{-1})$. In the sequel, we propose an iterative approach that allows us to compute the error we made by replacing the diagonal elements of $Q(z)$ by $m(z)$.

**Derivation of $\tilde{f}(z)$.** For $j = 1, \ldots, n$, denote by $f_j(z) = \mathbb{E}Q_{jj} - m(z)$. Then, substituting $\mathbb{E}Q_{jj}$ by $f_j(z) + m(z)$ into (160) and using the fact $zm(z) + \omega^2 c_0 m^2(z) + 1 = 0$, we obtain:

$$zf_j(z) = -\omega^2 c_0 f_j(z)m(z) - \omega^2 c_0 m(z)\frac{1}{n}\sum_{l=1}^{n} f_l(z) + 4\frac{n^2}{p^{\frac{5}{2}}}m^3(z)\big(\frac{1}{p}\operatorname{tr}\left((C^\circ)^3\right)\big)^2$$

$$-\frac{1}{p}\sum_{k=1}^{n}\left[2\left(\frac{1}{p}\operatorname{tr} C_{[k]}C_{[j]}\right)^2-\omega^2\right]m^2(z)+O_z(p^{-\frac{3}{4}}),\quad j=1,\ldots,n.$$

Again, leveraging the relation $z+\omega^2 c_0 m(z)=-\frac{1}{m(z)}$ leads to

$$f_j(z)=\omega^2 c_0 m^2(z)\frac{1}{n}\sum_{l=1}^{n}f_l(z)+\frac{1}{p}\sum_{k=1}^{n}\left[2\left(\frac{1}{p}\operatorname{tr} C_{[k]}C_{[j]}\right)^2-\omega^2\right]m^3(z)$$
$$-4n^2p^{-\frac{5}{2}}m^4(z)\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^3\right)\right)^2+O_z(p^{-\frac{3}{4}}),\quad j=1,\ldots,n.$$

The above equations define a linear system in the vector $f(z)=\left[f_1(z),\cdots,f_n(z)\right]^T$ which can be written as:

$$\left(I_n-\omega^2c_0m^2(z)\frac{1_n1_n^T}{n}\right)f=\frac{1}{p}\delta 1_n m^3(z)-4\frac{n^2}{p^{\frac{5}{2}}}m^4(z)\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^3\right)\right)^2 1_n+O_z(p^{-\frac{3}{4}})1_n$$

where $\delta$ the $n\times n$ matrix given by

$$\delta=\left\{2\left(\frac{1}{p}\operatorname{tr} C_{[k]}C_{[j]}\right)^2-\omega^2\right\}_{k,j=1}^{n}$$

Since $\left|1-\omega^2c_0m^2(z)\right|^{-1}=O_z(1)$ from Lemma 7, we have:

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

where

$$\overline{f}(z)=\frac{1}{p}m^3(z)\delta 1_n+\frac{\omega^2c_0m^5(z)1_n^T\delta 1_n}{np(1-\omega^2c_0m^2(z))}1_n-\frac{4n^2p^{-\frac{5}{2}}m^4(z)\left(\frac{1}{p}\operatorname{tr}\left((C^\circ)^3\right)\right)^2}{1-\omega^2c_0m^2(z)}1_n \tag{161}$$

Recalling that:

$$g_n(z)-m(z)=\frac{1}{n}1^Tf(z)$$

we thus obtain:

$$g_n(z)-m(z)=\frac{1}{n}1^T\overline{f}(z)+O_z(p^{-\frac{3}{4}})$$

The quantity $\frac{1}{n}1^T\overline{f}(z)$ represents thus the error of order $O_z(p^{-\frac{1}{2}})$ in the difference $g_n(z)-m(z)$. From that, we identify $\tilde{f}(z)$ as $\tilde{f}(z)=p^{\frac{1}{2}}\frac{1_n^T\overline{f}(z)}{n}$ where

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

and $\tilde{f}(z)$ simplifies as:

$$\tilde{f}(z)=\frac{1}{\sqrt{p}}\frac{m^3(z)1_n^T\delta 1_n}{n(1-\omega^2c_0m^2(z))}-\frac{4n^2p^{-2}m^4(z)\left(\frac{1}{p}\operatorname{tr}((C^\circ)^3)\right)^2}{1-\omega^2c_0m^2(z)}$$

**Derivation of $\tilde{h}(z)$.** To derive $\tilde{h}(z)$, we define for $j = 1, \dots, n$ $\quad h_j(z) = \mathbb{E}Q_{jj} - m(z) - \overline{f}_j(z)$ where $\overline{f}_j(z)$ is the $j$-th entry of vector $\overline{f}$. Following the same approach as for the derivation of $\tilde{f}(z)$, we substitute in (160) $\mathbb{E}Q_{jj}$ by $h_j(z) + m(z) + \overline{f}_j(z)$. Using (161) and the relation $zm(z) + \omega^2 c_0 m^2(z) + 1 = 0$, we obtain after simplifications:

$$zh_j(z) = -\omega^2 c_0 m(z) h_j(z) - \omega^2 c_0 m(z) \frac{1}{n} \sum_{l=1}^{n} h_l(z)$$
$$+ 4 \sum_{k=1}^{n} \sum_{r=1}^{n} p^{-\frac{5}{2}} m^4(z) \left[ \left( \frac{1}{p} \operatorname{tr} C_{[r]} C_{[k]} C_{[j]} \right)^2 - \left( \frac{1}{p} \operatorname{tr} (C^\circ)^3 \right)^2 \right] + O_z(p^{-1}),$$

which, using $(z + \omega^2 c_0 m(z))^{-1} = -m(z)$, leads to:

$$h_j(z) = \omega^2 c_0 m^2(z) \frac{1}{n} \sum_{l=1}^{n} h_l(z) \tag{162}$$
$$- 4 \sum_{k=1}^{n} \sum_{r=1}^{n} p^{-\frac{5}{2}} m^4(z) \big[ \big( \frac{1}{p} \operatorname{tr} C_{[r]} C_{[k]} C_{[j]} \big)^2 - \big( \frac{1}{p} \operatorname{tr} \big( (C^\circ)^3 \big) \big)^2 \big] + O_z(p^{-1}), \tag{163}$$

Denote by $h(z) = [h_1(z), \dots, h_n(z)]^T$. Then,

$$(I_n - \omega^2 c_0 m^2(z) \frac{1_n 1_n^T}{n}) h$$
$$= -4 \big\{ \sum_{k=1}^{n} \sum_{r=1}^{n} p^{-\frac{5}{2}} m^4(z) \big[ \big( \frac{1}{p} \operatorname{tr} C_{[r]} C_{[k]} C_{[j]} \big)^2 - \big( \frac{1}{p} \operatorname{tr} \big( (C^\circ)^3 \big) \big)^2 \big] \big\}_{j=1}^{n}$$
$$+ O_z(p^{-1}) 1_n$$

and thus:

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

where

$$\overline{h}(z) = -4 \Big\{ \sum_{k=1}^{n} \sum_{r=1}^{n} p^{-\frac{5}{2}} m^4(z) \left[ \left( \frac{1}{p} \operatorname{tr} C_{[r]} C_{[k]} C_{[j]} \right)^2 - \left( \frac{1}{p} \operatorname{tr} \left( (C^\circ)^3 \right) \right)^2 \right] \Big\}_{j=1}^{n}$$
$$- \frac{4\omega^2 c_0 m^6(z) p^{-\frac{5}{2}}}{n(1 - \omega^2 c_0 m^2(z))} \Big( \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{r=1}^{n} \big[ \big( \frac{1}{p} \operatorname{tr} C_{[r]} C_{[k]} C_{[j]} \big)^2 - \big( \frac{1}{p} \operatorname{tr} \big( (C^\circ)^3 \big) \big)^2 \big] \Big) 1_n$$

Recalling that:

$$g_n(z) - m(z) = \frac{1}{n} (1^T \overline{f}(z) + 1^T h(z))$$

we thus obtain:

$$g_n(z) - m(z) = \frac{1}{n} 1^T \overline{f}(z) + \frac{1}{n} 1^T \overline{h}(z) + O_z(p^{-1})$$

The quantity $\frac{1}{n}1^T\overline{h}(z)$ represents thus the error of order $O_z(p^{-\frac{3}{4}})$ in the difference $g_n(z)-m(z)$. From that, we identify $\tilde{h}(z)$ as $\tilde{h}(z)=p^{\frac{3}{4}}\frac{1}{n}1^T\overline{h}(z)$ where:

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

and

$$\tilde{h}(z)=-\frac{4}{n}\sum_{j=1}^n\sum_{k=1}^n\sum_{r=1}^n\frac{p^{-\frac{7}{4}}m^4(z)}{1-\omega^2c_0m^2(z)}\left[\left(\frac{1}{p}\operatorname{tr}C_{[r]}C_{[k]}C_{[j]}\right)^2-\left(\frac{1}{p}\operatorname{tr}\left(\left(C^\circ\right)^3\right)\right)^2\right]$$

**Asymptotic equivalent for $\tilde{k}(z)$.** We will now determine an asymptotic equivalent for the term vanishing at rate $O_z(p^{-1})$. To this end, we define for $j=1,\ldots,n$, $k_j(z)=\mathbb{E}Q_{jj}-m(z)-\overline{f}_j(z)-\overline{h}_j(z)$ and substitute $\mathbb{E}Q_{jj}$ by $m(z)+\overline{f}_j(z)+\overline{h}_j(z)+k_j(z)$ in (160), which yields after simplification

$$\begin{aligned}zk_j(z)&=-\omega^2c_0m(z)\frac{1}{n}\sum_{l=1}^nk_l(z)-\omega^2c_0\overline{f}_j(z)\frac{1}{n}\sum_{l=1}^n\overline{f}_l(z)-\omega^2c_0k_j(z)m(z)\\&-\frac{1}{p}\sum_{k=1}^n\delta_{kj}\left(m(z)\overline{f}_k(z)+m(z)\overline{f}_j(z)\right)+8nm^2(z)p^{-\frac{5}{2}}\sum_{r=1}^n\overline{f}_r(z)\left(\frac{1}{p}\operatorname{tr}\left(\left(C^\circ\right)^3\right)\right)^2\\&+4n^2p^{-\frac{5}{2}}m^2(z)\left(\frac{1}{p}\operatorname{tr}\left(C^\circ\right)^3\right)^2\overline{f}_j(z)+A_p^{-1}(z)+O_z(p^{-\frac{5}{4}})\end{aligned}$$

Similarly, using the relation $(z+\omega^2c_0m(z))=-\frac{1}{m(z)}$, we obtain:

$$\begin{aligned}k_j(z)&=\omega^2c_0m^2(z)\frac{1}{n}\sum_{l=1}^nk_l(z)+\omega^2c_0m(z)\overline{f}_j(z)\frac{\tilde{f}(z)}{\sqrt{p}}\\&+\frac{1}{p}\sum_{k=1}^n\delta_{kj}m^2(z)\left(\overline{f}_k(z)+\overline{f}_j(z)\right)-8m^3(z)\frac{n^2}{\sqrt{p}}p^{-\frac{5}{2}}\tilde{f}(z)\left(\frac{1}{p}\operatorname{tr}\left(C^\circ\right)^3\right)^2\\&-4n^2p^{-\frac{5}{2}}m^3(z)\left(\frac{1}{p}\operatorname{tr}\left(C^\circ\right)^3\right)^2\overline{f}_j(z)-m(z)A_{p^{-1}}(z)+O_z(p^{-\frac{5}{4}})\end{aligned}$$

Define $k(z)=[k_1(z),\ldots,k_n(z)]^T$. Then, the above equality can be equivalently written as:

$$\begin{aligned}\left(I_n-\omega^2c_0m^2(z)\frac{1_n1_n^T}{n}\right)k(z)&=\omega^2c_0m(z)\frac{\tilde{f}(z)}{\sqrt{p}}\overline{f}(z)+\frac{1}{p}m^2(z)\delta\overline{f}(z)\\&+m^2(z)\left\{\frac{1}{p}\sum_{k=1}^n\delta_{kj}\overline{f}_j(z)\right\}_{j=1}^n-8m^3(z)n^2p^{-3}\tilde{f}(z)\left(\frac{1}{p}\operatorname{tr}\left(C^\circ\right)^3\right)^21_n\\&-4n^2p^{-\frac{5}{2}}m^3(z)\left(\frac{1}{p}\operatorname{tr}\left(C^\circ\right)^3\right)^2\overline{f}(z)-m(z)A_{p^{-1}}(z)1_n+O_z(p^{-\frac{5}{4}})1_n\end{aligned}$$

or equivalently,

$$k(z)=\overline{k}(z)+O_z(p^{-\frac{5}{4}})1_n$$

with

$$\overline{k}(z) = \omega^2 c_0 m(z)\tilde{f}(z)\frac{\overline{f}(z)}{\sqrt{p}} + \frac{\omega^4 c_0^2 m^3(z)\left(\tilde{f}(z)\right)^2}{p(1-\omega^2 c_0 m^2(z))}1_n + \frac{1}{p}m^2(z)\delta\overline{f}(z)$$
$$+ \frac{\omega^2 c_0 m^4(z) 1_n^T \delta\overline{f}(z)}{pn(1-\omega^2 c_0 m^2(z))}1_n + m^2(z)\big\{\frac{1}{p}\sum_{k=1}^{n}\delta_{kj}\overline{f}_j(z)\big\}_{j=1}^{n} + \frac{\omega^2 c_0 m^4(z)1^T\delta\overline{f}(z)}{np(1-\omega^2 c_0 m^2(z))}1_n$$
$$- \frac{8m^3(z)n^2p^{-3}}{1-\omega^2 c_0 m^2(z)}\tilde{f}(z)\big(\frac{1}{p}\operatorname{tr}\left(\left(C^\circ\right)^3\right)\big)^2 1_n - 4n^2p^{-\frac{5}{2}}m^3(z)\big(\frac{1}{p}\operatorname{tr}\left(\left(C^\circ\right)^3\right)\big)^2\overline{f}(z)$$
$$- \frac{4m^5(z)n^2p^{-3}}{1-\omega^2 c_0 m^2(z)}\tilde{f}(z)\big(\frac{1}{p}\operatorname{tr}\left(\left(C^\circ\right)^3\right)\big)^2 1_n - \frac{m(z)A_{p^{-1}}(z)}{1-\omega^2 c_0 m^2(z)}1_n$$

Recalling that:

$$g_n(z) - m(z) = \frac{1}{n}(1^T\overline{f}(z) + 1^T\overline{h}(z) + 1^T k(z))$$

we thus obtain:

$$g_n(z) - m(z) = \frac{1}{n}1^T\overline{f}(z) + \frac{1}{n}1^T\overline{h}(z) + \frac{1}{n}1^T\overline{k}(z) + O_z(p^{-\frac{5}{4}})$$

from which, we identify $\tilde{k}(z)$ as $\tilde{k}(z) = p\frac{1^T\overline{k}(z)}{n}$ where

$$g_n(z) - m(z) - p^{-\frac{1}{2}}\tilde{f}(z) - p^{-\frac{3}{4}}\tilde{h}(z) - p^{-1}\tilde{k}(z) = O_z(p^{-\frac{5}{4}})$$

and

$$\tilde{k}(z) = \frac{\omega^2 c_0 m(z)\left(\tilde{f}(z)\right)^2}{1-\omega^2 c_0 m^2(z)} + \frac{2m^2(z)\overline{f}(z)^T\delta 1_n}{n(1-\omega^2 c_0 m^2(z))}$$
$$- \frac{12m^3(z)n^2p^{-2}}{1-\omega^2 c_0 m^2(z)}\tilde{f}(z)\big(\frac{1}{p}\operatorname{tr}\left(\left(C^\circ\right)^3\right)\big)^2 - \frac{m(z)pA_{p^{-1}}(z)}{1-\omega^2 c_0 m^2(z)}$$

With this, we complete the proof of (125), which will be the key for the analysis of the support of the empirical measure of $\Phi$.

### *5.4.3. Concluding.*

With the approximation in (125) at hand, we are now ready to determine the limiting support of the empirical measure of $\Phi$. We first need to prove that $\tilde{f}(z)$, $\tilde{h}(z)$ and $\tilde{k}(z)$ are Stieljtes transforms of some distributions and determine the supports thereof. To this end, we will resort to the following Lemma.

**Lemma 16.** *[16, Lemma 9.1] Let $\Lambda$ 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 $\Lambda$ 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 $\tilde{h}(z)$, we can easily see that both of them are analytic on $\mathbb{C}\backslash\left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right]$ except $\tilde{k}(z)$ which presents singularities (through the term $A_{p^{-1}}(z)$) for $z$ such that

$$m(z) = \pm\frac{1}{c_0\Omega}. \tag{164}$$

This singularity falls outside the support if $\Omega > \frac{1}{\sqrt{c_0}\omega}$, in which case the $z$'s satisfying (164) are given by the two isolated complex values $\{-\tilde{\rho}, \tilde{\rho}\}$ where $\tilde{\rho} = c_0\Omega + \frac{\omega^2}{\Omega}$.

**Proposition 4.** *$\tilde{f}(z)$ and $\tilde{h}(z)$ are the Stieltjes transforms of distributions $\Lambda_{\tilde{f}}$ and $\Lambda_{\tilde{h}}$ with support $\mathcal{S} = \left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right]$ while $\tilde{k}(z)$ is the Stieltjes transform of $\Lambda_{\tilde{k}}$ with support $\mathcal{S}_k = \mathcal{S} \cup \{-\tilde{\rho}, \tilde{\rho}\}$. Moreover, $\Lambda_{\tilde{f}}(1) = \Lambda_{\tilde{h}}(1) = \Lambda_{\tilde{k}}(1) = 0$.*

*Proof.* We will prove the result only for $\tilde{f}(z)$. The same reasoning can be applied to $\tilde{h}(z)$. For $\tilde{k}(z)$, some slight modifications should be made to account for the singularities $\{-\tilde{\rho}, \tilde{\rho}\}$. According to Lemma 16, it suffices to show that $\tilde{f}(z)$ satisfy conditions $c_1$ and $c_2$ of Lemma 16. Let $|z| \geq 4\sqrt{3c_0}\omega$, then there exist positive constants $C$ such that:

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

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 [5]. We define the interval[3]:

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

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

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

- Let $z \in D \cap \mathbb{C}\backslash\mathbb{R}$ 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:
$$\left|\tilde{f}(z)\right| \leq C_0 |\Im z|^{-8} = C_0 \mathrm{dist}(z,K)^{-8} = C_0 \max\left(\mathrm{dist}(z,K)^{-8}, 1\right)$$
- Let $z \in D \cap \mathbb{C}\backslash\mathbb{R}$ 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 $\Re z \notin K$,
$$\left|\tilde{f}(z)\right| \leq C_1 \leq C_1 \mathrm{dist}(z,K)^{-8} = C_1 \max\left(\mathrm{dist}(z,K)^{-8}, 1\right)$$
- Since $\left|\tilde{f}(z)\right| \to 0$ when $|z| \to \infty$, $\tilde{f}(z)$ is bounded on $\mathbb{C}\backslash D$. Thus, there exists some constant $C_2$ such that for any $z \in \mathbb{C}\backslash D$,
$$\left|\tilde{f}(z)\right| \leq C_2 = C_2 \max\left(\mathrm{dist}(z,K)^{-8}, 1\right)$$

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 4, 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 $\psi$ 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 17.** *Assume that $\Omega > \frac{1}{\sqrt{c_0}\omega}$. For all smooth function $\psi$ constant on the complementary of a compact interval and vanishing on $\mathcal{S} = \left[-2\sqrt{c_0}\omega, 2\sqrt{c_0}\omega\right] \cup \{-\tilde{\rho}, \tilde{\rho}\}$,*

$$\mathbb{E}\left[\frac{1}{n}\operatorname{tr}\left(\psi(\Phi)\right)\right] = O(p^{-\frac{5}{4}}) \tag{165}$$

$$\mathbb{E}\left|\frac{1}{n}\operatorname{tr}\left(\psi(\Phi)\right) - \mathbb{E}\frac{1}{n}\operatorname{tr}\left(\psi(\Phi)\right)\right|^{2l} = O(p^{-\frac{5l}{2}}) \tag{166}$$

*for each $l \geq 1$.*

*Proof.* Using the inverse Stieltjes transform, it holds that for any smooth function $\psi_c$ with compact support:

$$\begin{aligned}\frac{1}{n}\mathbb{E}\left[\operatorname{tr}\psi_c(\Phi)\right] = \int \psi_c d\mu + \frac{1}{p^{\frac{1}{2}}}\Lambda_{\tilde{f}}(\psi_c) + \frac{1}{p^{\frac{3}{4}}}\Lambda_{\tilde{h}}(\psi_c) + \frac{1}{p}\Lambda_{\tilde{k}}(\psi_c)\\ - \frac{1}{\pi}\lim_{y\to 0^+}\Im\int_{\mathbb{R}}\psi_c(x)R_n(x+iy)dx\end{aligned}$$

where $R_n = g_n(z) - m(z) - \frac{1}{p^{\frac{1}{2}}}\tilde{f}(z) - \frac{1}{p^{\frac{3}{4}}}\tilde{h}(z) - \frac{1}{p}\tilde{k}(z)$. Since, for $z \in \mathbb{C}\backslash\mathbb{R}$,

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

using the ideas of [13],

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

In order to prove (165), we follow the approach in [16]. We denote $\kappa$ 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 4, $\Lambda_{\tilde{f}}(\psi_c)=\Lambda_{\tilde{h}}(\psi_c)=\Lambda_{\tilde{k}}(\psi_c)=0$. Hence,

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

In order to prove (166), we proceed by induction on $l$. For $l=1$, using the Poincaré-Nash inequality, we have:

$$\begin{aligned}&\operatorname{var}\big(\frac{1}{n}\operatorname{tr}\psi(\Phi)\big)\\ &\leq \frac{1}{n^2}\sum_{i=1}^{p}\sum_{j=1}^{n}\mathbb{E}\big|\frac{\operatorname{tr}\psi(\Phi)}{\partial Z_{ij}}\big|^2\\ &\leq \frac{1}{n^2}\sum_{i=1}^{p}\sum_{j=1}^{n}\mathbb{E}\big|\operatorname{tr}\psi'(\Phi)\big\{\frac{\partial\Phi_{lk}}{\partial Z_{ij}}\big\}_{l,k=1}^{n}\big|^2\\ &= \frac{4}{n^2}\sum_{i=1}^{p}\sum_{j=1}^{n}\mathbb{E}\Big|\operatorname{tr}\psi'(\Phi)\big\{x_l^Tx_k\left[C_{[k]}^{\frac{1}{2}}x_l\right]_i\boldsymbol{\delta}_{j\neq k}\boldsymbol{\delta}_{l\neq k}\\ &\qquad\qquad +(x_l^Tx_k)\left[C_{[l]}^{\frac{1}{2}}x_k\right]_i\boldsymbol{\delta}_{j\neq l}\boldsymbol{\delta}_{l\neq k}\big\}_{l,k=1}^{n}\Big|^2\\ &= \frac{16}{n^2}\sum_{i=1}^{p}\sum_{j=1}^{n}\mathbb{E}\left|\sum_{a\neq j}\left[\psi'(\Phi)\right]_{aj}(x_j^Tx_a)\left[C_{[j]}^{\frac{1}{2}}x_a\right]_i\right|^2\\ &= \frac{16}{n^2}\sum_{j=1}^{n}\sum_{a_1\neq j}\sum_{a_2\neq j}\mathbb{E}\left[x_{a_2}^TC_{[j]}x_{a_1}x_j^Tx_{a_2}x_j^Tx_{a_1}\left[\psi'(\Phi)\right]_{a_1j}\left[\psi'(\Phi)\right]_{a_2j}\right]\\ &\leq \frac{16}{n^2}\sum_{j=1}^{n}\mathbb{E}\left[\left|\left[\psi'(\Phi)\right]_{a_1j}\right|\left|\left[\psi'(\Phi)\right]_{a_2j}\right|\left|x_{a_2}^TC_{[j]}x_{a_1}\right|\max_{j\neq a_1}\left|x_j^Tx_{a_1}\right|\max_{j\neq a_2}\left|x_j^Tx_{a_2}\right|\right]\end{aligned}$$

Define $R$ as:

$$[R]_{a_1a_2}=\max_{j\neq a_1}\left|x_j^Tx_{a_1}\right|\max_{j\neq a_2}\left|x_j^Tx_{a_2}\right|\max_{j\notin\{a_1,a_2\}}\left|x_{a_2}^TC_{[j]}x_{a_1}\right|$$

It is easy to see that $\|R\|=O(p^{-\frac{1}{2}+\epsilon})$ by bounding the Frobenius norm, for instance. Let $h(x)=|\psi'(x)|$ Hence:

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

From (165), $\mathbb{E}\frac{1}{n}\operatorname{tr} h^2(\Phi) = O(p^{-\frac{5}{4}})$. Using the fact that $\|R\| = O(p^{-\frac{1}{2}+\epsilon})$ along with Lemma 5 we obtain for any $\epsilon$ small and positive,

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

Assume now that (166) holds for all $l \neq k-1$. We will prove it for $l = k$. Note that:

$$\mathbb{E}\big|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\Phi\big|^{2k} = \left(\mathbb{E}\big|\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\Phi\big|^{k}\right)^2 + \operatorname{var}\big(\frac{1}{n}\operatorname{tr}\psi(\Phi) - \mathbb{E}\frac{1}{n}\operatorname{tr}\psi(\Phi)\big)^k \tag{167}$$

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

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

Using the induction assumption, it unfolds that:

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

We will now handle the second term in the right-hand side of (167). Using the Poincaré-Nash inequality, we obtain:

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

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 $\|R\|^k = O(p^{-\frac{k}{2}+\epsilon})$; From Lemma 5, 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 (165), it unfolds that:

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

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:

$$\kappa_p \leq C_1 p^{-\frac{5}{2}k} + C_2 \kappa_p^{\frac{k-1}{k}} p^{-\frac{11}{4}+\epsilon} \tag{168}$$

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 (168) in terms of $u_p$, we obtain:

$$u_p \leq C_1 + C_2 \left(u_p\right)^{1-\frac{1}{k}} p^{-\frac{1}{4}+\epsilon}$$

or equivalently:

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

thus proving that $u_p$ is a bounded sequence. This finishes the proof of Lemma 17. ☐

A direct consequence of Lemma 17 is that for any $\psi$ satisfying the condition of Lemma 17,

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

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 $\psi$ smooth such that

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

Function $\psi$ satisfies the conditions of Lemma 17. 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

$$\mathcal{S}^\epsilon := \left[-2\sqrt{c_0}\omega - \epsilon, 2\sqrt{c_0}\omega + \epsilon\right] \cup [-\tilde{\rho} - \epsilon, -\tilde{\rho} - \epsilon] \cup [\tilde{\rho} - \epsilon, \tilde{\rho} + \epsilon],$$

we conclude that almost surely for $n$ large enough, there is no eigenvalue of $\Phi$ outside $\mathcal{S}^\epsilon$.

## Appendix A Proof of the preliminary results

### *A.1 Proof of Lemma 6*

**Proof of (40).** Decomposing $\xi_{(k,-k)} = \xi_{(k,-k)} - \mathbb{E}_k\left[\xi_{(k,-k)}\right] + \mathbb{E}_k\left[\xi_{(k,-k)}\right]$ we may expand $\mathbb{E}_k\left[\xi_{(k,-k)}^T A \xi_{(k,-k)}\right]$ as:

$$\mathbb{E}\left[\xi_{(k,-k)}^T A \xi_{(k,-k)}\right] = \mathbb{E}_k\left[\left(\xi_{(k,-k)} - \mathbb{E}_k\left[\xi_{k,-k}\right]\right)^T A\left(\xi_{(k,-k)} - \mathbb{E}_k\left[\xi_{k,-k}\right]\right)\right]$$

$$+\mathbb{E}_k\left[\xi^T_{(k,-k)}\right]A\mathbb{E}_k\left[\xi_{(k,-k)}\right] \tag{169}$$

$$= \mathbb{E}_k\Big[\sum_{l\neq k}\sum_{m\neq k}\big(\sqrt{p}(x_k^Tx_l)^2-\frac{1}{\sqrt{p}}x_l^TC_{[k]}x_l\big)A_{lm}\big(\sqrt{p}(x_k^Tx_m)^2-\frac{1}{\sqrt{p}}x_m^TC_{[k]}x_m\big)\Big] \tag{170}$$

$$+\sum_{l\neq k}\sum_{m\neq k}\big(\frac{1}{\sqrt{p}}x_l^TC_{[k]}x_l-\frac{1}{p^{\frac{3}{2}}}\operatorname{tr}C_{[k]}C_{[l]}\big)A_{lm}\big(\frac{1}{\sqrt{p}}x_m^TC_{[k]}x_m-\frac{1}{p^{\frac{3}{2}}}\operatorname{tr}C_{[k]}C_{[m]}\big) \tag{171}$$

$$=\frac{2}{p}\sum_{l\neq k}\sum_{m\neq k}(x_l^TC_{[k]}x_m)^2A_{lm} \tag{172}$$

$$+\sum_{l\neq k}\sum_{m\neq k}\big(\frac{1}{\sqrt{p}}x_l^TC_{[k]}x_l-\frac{1}{p^{\frac{3}{2}}}\operatorname{tr}C_{[k]}C_{[l]}\big)A_{lm}\big(\frac{1}{\sqrt{p}}x_m^TC_{[k]}x_m-\frac{1}{p^{\frac{3}{2}}}\operatorname{tr}(C_{[k]}C_{[m]})\big) \tag{173}$$

where the last equality follows by using the fact that

$$\mathbb{E}\left[(z^TA_1z-\operatorname{tr}(A_1))(z^TA_2z-\operatorname{tr}(A_2))\right]=2\operatorname{tr}(A_1A_2)$$

for $z$ standard Gaussian random vector in $\mathbb{R}^{n\times 1}$ with $A_1$, $A_2$ $n\times n$ matrices.
**Proof of** (41)**.** The proof of (41) will be carried out by induction on $s$. For $s=1$, using Poincaré-Nash inequality, we obtain:

$$\mathbb{E}_k\left|\xi^T_{(k,-k)}A\xi_{(k,-k)}-\mathbb{E}_k\xi^T_{(k,-k)}A\xi_{(k,-k)}\right|^2\leq\sum_{j=1}^p\mathbb{E}_k\Big[\big|\frac{\partial\xi^T_{(k,-k)}A\xi_{(k,-k)}}{\partial Z_{jk}}\big|^2\Big]$$

Hence, to prove the result for $s=1$, it suffices to establish that:

$$\sum_{j=1}^p\big|\frac{\partial\xi^T_{(k,-k)}A\xi_{k,-k}}{\partial Z_{jk}}\big|^2\leq\|A\|^2O(p^{-1}) \tag{174}$$

where $O(p^{-1})$ should be understood in the sense of the convergence of random variables as described in the notation section. Moreover, as will be shown next, the inequality in (174) will also help in the proof of the result for $s>1$. Given that:

$$\begin{aligned}&\frac{\partial\xi^T_{(k,-k)}A\xi_{k,-k}}{\partial Z_{jk}}\\&=\sum_{a\neq k}\sum_{b\neq k}\frac{\partial p(x_k^Tx_a)^2A_{ab}(x_k^Tx_b)^2}{\partial Z_{jk}}-\sum_{a\neq k}\sum_{b\neq k}\frac{\partial\frac{1}{p}\operatorname{tr}C_{[k]}C_{[a]}A_{ab}(x_k^Tx_b)^2}{\partial Z_{jk}}\\&-\sum_{a\neq k}\sum_{b\neq k}\frac{\partial(x_k^Tx_a)^2A_{ab}\frac{1}{p}\operatorname{tr}C_{[k]}C_{[b]}}{\partial Z_{jk}},\end{aligned}$$

we have:

$$\sum_{j=1}^p\big|\frac{\partial\xi^T_{(k,-k)}A\xi_{(k,-k)}}{\partial Z_{jk}}\big|^2\leq\tilde{\alpha}_1+4\tilde{\alpha}_2$$

where

$$\tilde{\alpha}_1 := \sum_{j=1}^{p} \Big| \sum_{a\neq k} \sum_{b\neq k} \frac{\partial p(x_k^T x_a)^2 A_{ab} (x_k^T x_b)^2}{\partial Z_{jk}} \Big|^2$$

$$\tilde{\alpha}_2 := \sum_{j=1}^{p} \Big| \sum_{a\neq k} \sum_{b\neq k} \frac{\partial (x_k^T x_a)^2 A_{ab} \frac{1}{p} \operatorname{tr} C_{[k]} C_{[b]}}{\partial Z_{jk}} \Big|^2$$

Both $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ can be shown to satisfy:

$$\tilde{\alpha}_1 \leq \|A\|^2 O(p^{-1}), \tag{175}$$

$$\tilde{\alpha}_2 \leq \|A\|^2 O(p^{-1}). \tag{176}$$

Proving (175) and (176) implies directly the sought-for result in (174). To prove (175), we upper-bound $\tilde{\alpha}_1$ as:

$$\begin{aligned}
\tilde{\alpha}_1 &\leq p^2 \sum_{j=1}^{p} \left[ \Big| \sum_{a\neq k} \sum_{b\neq k} \frac{\partial (x_k^T x_a)^2 A_{ab} (x_k^T x_b)^2}{\partial Z_{jk}} \Big|^2 \right] \\
&\leq 16p \sum_{j=1}^{p} \left[ \Big| \sum_{a\neq k} \sum_{b\neq k} x_k^T x_a \left[ C_{[k]}^{\frac{1}{2}} x_a \right]_j A_{ab} (x_k^T x_b)^2 \Big|^2 \right] \\
&= 16p \sum_{a_1\neq k} \sum_{b_1\neq k} \sum_{a_2\neq k} \sum_{b_2\neq k} \left[ x_k^T x_{a_1} x_k^T x_{a_2} x_{a_2}^T C_{[k]} x_{a_1} A_{a_1 b_1} A^*_{a_2 b_2} (x_k^T x_{b_1})^2 \left( x_k^T x_{b_2} \right)^2 \right] \\
&= 16p 1^T D_k^2 A^H S_k A D_k^2 1 \\
&\leq 16p^2 \|D_k\|^4 \|A\|^2 \|S_k\|
\end{aligned}$$

where $D_k = \mathcal{D}\left\{ x_b^T x_k \boldsymbol{\delta}_{b\neq k} \right\}_{b=1}^n$ and $[S_k]_{a_1,a_2} = x_k^T x_{a_1} x_{a_2}^T C_{[k]} x_{a_1} x_k^T x_{a_2} \boldsymbol{\delta}_{a_1\neq k} \boldsymbol{\delta}_{a_2\neq k}$. From Lemma 4, $\|S_k\| = O(p^{-1})$. Thus, (175) follows using the fact that $\|D_k\| = O(p^{-1/2})$.

On the other hand, $\tilde{\alpha}_2$ can be treated as:

$$\begin{aligned}
\tilde{\alpha}_2 &\leq 4 \sum_{j=1}^{p} \Big| \sum_{a\neq k} \sum_{b\neq k} (x_k^T x_a) \frac{1}{\sqrt{p}} \left[ C_{[k]}^{\frac{1}{2}} x_a \right]_j A_{ab} \frac{1}{p} \operatorname{tr} C_{[k]} C_{[b]} \Big|^2 \\
&= \frac{4}{p} \sum_{a_1\neq k} \sum_{b_1\neq k} \sum_{a_2\neq k} \sum_{b_2\neq k} x_k^T x_{a_1} x_k^T x_{a_2} x_{a_2}^T C_{[k]} x_{a_1} \frac{1}{p} \operatorname{tr} C_{[k]} C_{[b_1]} \frac{1}{p} \operatorname{tr} C_{[k]} C_{[b_2]} \\
&\qquad\qquad \times A_{a_1 b_1} A^*_{a_2 b_2} \\
&= \frac{4}{p} 1^T \mathcal{D} \left\{ \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[b_2]}) \boldsymbol{\delta}_{b_2\neq k} \right\}_{b_2=1}^n A^H S_k A \mathcal{D} \left\{ \frac{1}{p} \operatorname{tr} C_{[k]} C_{[b_1]} \boldsymbol{\delta}_{b_1\neq k} \right\}_{b_1=1}^n 1 \\
&= \|A\|^2 O(p^{-1})
\end{aligned}$$

We assume that (41) holds for all integer $k = 1, \ldots, s-1$, and consider proving it for $k = s$. For that, we use the following relation

$$\mathbb{E}_k \left| \xi_{(k,-k)}^T A \xi_{(k,-k)} - \mathbb{E}_k \xi_{(k,-k)}^T A \xi_{(k,-k)} \right|^{2s}$$

$$= \left(\mathbb{E}_k\left|\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_k\xi_{(k,-k)}^T A\xi_{(k,-k)}\right|^s\right)^2$$
$$+ \mathbf{var}_k\left(\left(\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_k\xi_{(k,-k)}^T A\xi_{(k,-k)}\right)^s\right)$$

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

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

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

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

Since $\sum_{j=1}^{p}\left|\frac{\partial \xi_{(k,-k)}^T A\xi_{(k,-k)}}{\partial Z_{j,k}}\right|^2 = \|A\|^2 O(p^{-1})$ by (174) and

$$\mathbb{E}_k\left[\left|\xi_{(k,-k)}^T A\xi_{(k,-k)} - \mathbb{E}_k\xi_{(k,-k)}^T A\xi_{(k,-k)}\right|^{2(s-1)}\right] = \|A\|^{2(s-1)} O(p^{-s+1+\frac{\epsilon}{2}})$$

by the induction assumption, we have by Lemma 5,

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

### *A.2 Proof of Lemma 7*

We will only prove the inequalities (44) and (45) in the last item as the first two items have been established in [5].

**Proof of (44).** Since $\omega^2 c_0 m^2(z) = -1 - zm(z)$, we have:

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

To show (44), we start by noticing that:

$$\left|1 - \alpha m^2(z)\right| \geq \max\left(\left|\Re(1 - \alpha m^2(z))\right|, \left|\Im(1 - \alpha m^2(z))\right|\right) \tag{177}$$

In view of (177), it suffices thus to study $\left|\Re(1-\alpha m^2(z))\right|$ and $\left|\Im(1-\alpha m^2(z))\right|$.

Let $z = x + \imath y$. Then, due to the symmetry of $\mu$, $\left|\Im(1-\alpha m^2(z))\right|$ can be simplified as:

$$\begin{aligned}
&\left|\Im\left(1-\alpha m^2(z)\right)\right| \\
&= \frac{\alpha}{\omega^2 c_0}|y| \left| \int_0^{2\sqrt{c_0}\omega} \frac{\lambda}{(\lambda-x)^2+y^2}\mu(d\lambda) - \int_0^{2\sqrt{c_0}\omega} \frac{\lambda}{(\lambda+x)^2+y^2}\mu(d\lambda) \right| \\
&= \frac{\alpha|yx|}{\omega^2 c_0} \int_0^{2\sqrt{c_0}\omega} \frac{4\lambda^2}{((\lambda-x)^2+y^2)\left((\lambda+x)^2+y^2\right)}\mu(d\lambda) \qquad (178)
\end{aligned}$$

On the other hand, $\left|\Re(1-\alpha m^2(z))\right|$ can be expanded as:

$$\begin{aligned}
&\left|\Re\left(1-\alpha m^2(z)\right)\right| \\
&= 1 + \frac{\alpha}{\omega^2 c_0} + \frac{\alpha}{\omega^2 c_0}\Big( \int \frac{\lambda x}{(\lambda-x)^2+y^2}\mu(d\lambda) - \int \frac{(x^2+y^2)}{(\lambda-x)^2+y^2}\mu(d\lambda) \Big) \\
&= 1 + \frac{\alpha}{\omega^2 c_0} + \frac{\alpha}{\omega^2 c_0} \int_0^{2\sqrt{c_0}\omega} \frac{4\lambda^2 x^2}{((\lambda-x)^2+y^2)\left((\lambda+x)^2+y^2\right)}\mu(d\lambda) \\
&- \frac{\alpha}{\omega^2 c_0}(x^2+y^2) \int_0^{2\sqrt{c_0}\omega} \frac{2\lambda^2+2x^2+2y^2}{((\lambda-x)^2+y^2)\left((\lambda+x)^2+y^2\right)}\mu(d\lambda) \\
&= \Big| \int_0^{2\sqrt{c_0}\omega} \frac{P(x,y,\lambda)}{((\lambda-x)^2+y^2)\left((\lambda+x)^2+y^2\right)}\mu(d\lambda) \Big| \qquad (179)
\end{aligned}$$

where

$$\begin{aligned}
P(x,y,\lambda) &:= (2 + \frac{2\alpha}{\omega^2 c_0})\lambda^4 - \lambda^2 x^2 (4 + \frac{2\alpha}{\omega^2 c_0}) + \lambda^2 y^2 (4 + \frac{2\alpha}{\omega^2 c_0}) \\
&+ 2x^2(x^2+y^2) + 2y^2(x^2+y^2)
\end{aligned}$$

Consider the following two cases: $|x| \geq |y|$ and $|x| \leq |y|$.
Case 1: $|x| \geq |y|$. It follows from (178) that

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

Then, for $0 \leq \lambda \leq 2\sqrt{c_0}\omega$,

$$\begin{aligned}
\max\left((\lambda-x)^2+y^2, (\lambda+x)^2+y^2\right) &\leq 2\lambda^2 + 2|z|^2 \leq 2(|z|^2 + (2\sqrt{c_0}\omega)^2) \\
&\leq 2(|z| + 2\sqrt{c_0}\omega)^2
\end{aligned} \qquad (180)$$

Hence,

$$\left|1-\alpha m^2(z)\right| \geq \frac{\alpha|y^2|}{\omega^2 c_0(|z|+2\sqrt{c_0}\omega)^4} \int_0^{2\sqrt{c_0}\omega} \lambda^2 \mu(d\lambda) = \frac{\alpha|y|^2}{2(|z|+2\sqrt{c_0}\omega)^4} \qquad (181)$$

Case 2: $|x| \leq |y|$. In this case, $(-x^2+y^2)(4+\frac{2\alpha}{\omega^2 c_0}) \geq 0$. Hence, using (179), we obtain:

$$\left|1-\alpha m^2(z)\right| \geq \frac{2y^4}{4(|z|+2\sqrt{c_0}\omega)^4}\int_0^{2\sqrt{c_0\omega}} \mu(d\lambda) = \frac{y^4}{4(|z|+2\sqrt{c_0}\omega)^4} \tag{182}$$

To prove (44), we combine (181) and (182) to obtain:

$$|1-\alpha m^2(z)|^{-1} \leq (|z|+2\sqrt{c_0}\omega)^4(4|\Im z|^{-4}+\frac{2}{\alpha}|\Im z|^{-2})$$

**Proof of (45).** To show (45), we will exploit the following inequality

$$|1-\alpha m^2(z)| \geq |\Re(1-\alpha m^2(z))|$$

Note first that if $|z| \geq 2\sqrt{2}\sqrt{c_0}\omega\sqrt{4+\frac{2\alpha}{\omega^2 c_0}}$, then necessarily

$$\max(|x|,|y|) \geq 2\sqrt{c_0}\omega\sqrt{4+\frac{2\alpha}{\omega^2 c_0}} \tag{183}$$

Based on (183), we consider the following two cases:
Case 1: $|y| = \max(|x|,|y|)$. In this case, $(-x^2+y^2)(4+\frac{2\alpha}{\omega^2 c_0}) \geq 0$. Hence, using (179) along with (180), we obtain:

$$|1-\alpha m^2(z)| \geq \frac{2|z|^4\int_0^{2\sqrt{c_0}\omega}\mu(d\lambda)}{4(|z|+2\sqrt{c_0}\omega)^4} \geq \frac{|z|^4}{8(|z|+2\sqrt{c_0}\omega)^4} \tag{184}$$

Case 2: $|x| = \max(|x|,|y|)$. In this case, from (183), it holds that

$$|x| \geq 2\sqrt{c_0}\omega\sqrt{4+\frac{2\alpha}{\omega^2 c_0}}.$$

Under this condition, it can be easily checked that function $\lambda \mapsto (2+\frac{2\alpha}{\omega^2 c_0})\lambda^4 - \lambda^2 x^2(4+\frac{2\alpha}{\omega^2 c_0})$ is a decreasing function on $\left(0, 2\sqrt{c_0}\omega\right)$ and thus achieves its minimum at $\lambda = 2\sqrt{c_0}\omega$. As a result, in view of (179), $|\Re(1-\alpha m^2(z))|$ can be lower-bounded as follows:

$$\left|\Re\left(1-\alpha m^2(z)\right)\right| \geq \left|\int_0^{2\sqrt{c_0}\omega} \frac{-4c_0\omega^2x^2(4+\frac{2\alpha}{\omega^2 c_0})+x^4+(x^2+y^2)^2}{((\lambda-x)^2+y^2)\left((\lambda+x)^2+y^2\right)}\mu(d\lambda)\right| \tag{185}$$

$$\geq \frac{|z|^4}{8(|z|+2\sqrt{c_0}\omega)^4} \tag{186}$$

where (186) follows from the fact that since $|x| \geq 2\sqrt{c_0}\omega\sqrt{4+\frac{2\alpha}{\omega^2 c_0}}$,

$$x^4 - 4c_0\omega^2(4+\frac{2\alpha}{\omega^2 c_0})x^2 \geq 0.$$

Combining (184) and (186), we note that in either case $|y| = \max(|x|, |y|)$ or $|x| = \max(|x|, |y|)$,

$$|1 - \alpha m^2(z)| \geq \frac{|z|^4}{8(|z| + 2\sqrt{c_0}\omega)^4}$$

which proves the desired result.

### *A.3 Proof of Lemma 9*

To ease the notations, we denote by $d_{k,j}$ the quantity:

$$d_{k,j} = \frac{1}{\sqrt{p}}\Big(x_k^T A_{j,p} x_k - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[j]})\Big)\boldsymbol{\delta}_{k\neq j}$$

The aim of Lemma 9 is to show that:

$$\mathbb{E}\Big[\Big|\sum_{k\neq j} d_{k,j} Q_{kj}\Big|^{2s}\Big] = O_z(p^{-2s+\epsilon}) \tag{187}$$

By Lemma 5, proving that

$$\mathbb{E}_j\Big[\Big|\sum_{k\neq j} d_{k,j} Q_{kj}\Big|^{2s}\Big] \leq (|\Im z|^{-4} + |\Im z|^{-6})^s O(p^{-2s+\epsilon}) \tag{188}$$

suffices to show the desired result (187). Thus, in what follows, we consider showing (188) by induction on $s$. For $s = 1$, we decompose $\mathbb{E}_j\big[\big|\sum_{k\neq j} d_{k,j}Q_{kj}\big|^2\big]$ as:

$$\mathbb{E}_j\big[\big|\sum_{k\neq j} d_{k,j}Q_{kj}\big|^2\big] = \mathbf{var}_j\big(\frac{1}{\sqrt{p}}\sum_{k\neq j} d_{k,j}Q_{kj}\big) + \big|\mathbb{E}_j\big[\sum_{k\neq j} d_{k,j}Q_{kj}\big]\big|^2 \tag{189}$$

Using Poincaré-Nash inequality, the treatment of the first term in the right-hand side of (189) boils down to showing that:

$$\sum_{l=1}^{p}\big|\sum_{k\neq j} d_{k,j}\frac{\partial Q_{kj}}{\partial Z_{lj}}\big|^2 \leq |\Im z|^{-4} O(p^{-2}) \tag{190}$$

Indeed, using the differentiation formula in (29), we obtain:

$$\begin{aligned}
&\sum_{l=1}^{p}\big|\sum_{k\neq j} d_{k,j}\frac{\partial Q_{kj}}{\partial Z_{lj}}\big|^2 \\
&= \sum_{l=1}^{p} 4\big|\sum_{k\neq j} d_{k,j}\sum_{b\neq j}(x_b^T x_j)\Big[C_{[j]}^{\frac{1}{2}}x_b\Big]_l (Q_{kj}Q_{bj} + Q_{jj}Q_{bk})\big|^2 \\
&\leq 8\big|\big[1^T\mathcal{D}\{d_{k,j}\}_{k=1}^n Q\big]_j\big|^2 \sum_{l=1}^{p}\big|\sum_{b\neq j}(x_b^T x_j)[C_{[j]}^{\frac{1}{2}}x_b]_l Q_{bj}\big|^2
\end{aligned}$$

$$+ 8|Q_{jj}|^2 \sum_{l=1}^{p} \Big| \sum_{b \neq j} [1^T \mathcal{D}\{d_{k,j}\}_{k=1}^n Q]_b (x_b^T x_j) [C_{[j]}^{\frac{1}{2}} x_b]_l \Big|^2 \tag{191}$$

$$= 8 \left| \left[ 1^T \mathcal{D}\{d_{k,j}\}_{k=1}^n Q \right]_j \right|^2 \left[ Q^H S_j Q \right]_{jj}$$
$$+ 8|Q_{jj}|^2 1^T \mathcal{D}\{d_{k,j}\}_{k=1}^n Q S_j Q^H \mathcal{D}\{d_{k,j}\}_{k=1}^n 1. \tag{192}$$

where $S_j$ is the $n \times n$ matrix with elements

$$[S_j]_{b_1 b_2} = (x_j^T x_{b_1})(x_j^T x_{b_2}) x_{b_1}^T C_{[j]} x_{b_2} \boldsymbol{\delta}_{b_1 \neq j} \boldsymbol{\delta}_{b_2 \neq j}, \tag{193}$$

and $\mathcal{D}\{d_{k,j}\}_{k=1}^n$ is the diagonal matrix with diagonal elements $d_{1,j}, \ldots, d_{n,j}$. Obviously the spectral norm of $S_j$ is $O(p^{-1})$ by Lemma 4 and so is that of $\mathcal{D}\{d_{k,j}\}_{k=1}^n$. From this, it is easy to see that (190) holds true.

It remains thus to treat the second term in the right-hand side of (189). We consider proving that:

$$\big| \mathbb{E}_j \big[ \sum_{k \neq j} d_{k,j} Q_{kj} \big] \big|^2 \leq (|\Im z|^{-4} + |\Im z|^{-6}) O(p^{-2}) \tag{194}$$

Noticing that for $k \neq j$, $d_{k,j}$ is independent of $x_j$, we obtain:

$$\big| \mathbb{E}_j \Big[ \sum_{k \neq j} d_{k,j} Q_{kj} \Big] \big|^2$$
$$= \big| \sum_{k \neq j} d_{k,j} \mathbb{E}_j [Q_{kj}] \big|^2 \tag{195}$$
$$= \big| \sum_{k \neq j} d_{k,j} \mathbb{E}_j [Q_{jj} e_k^T Q_j \xi_{(j,-j)}] \big|^2 \tag{196}$$
$$\leq 2 \big| \sum_{k \neq j} d_{k,j} \mathbb{E}_j [(Q_{jj} - \mathbb{E}_j(Q)_{jj}) e_k^T Q_j \xi_{(j,-j)}] \big|^2$$
$$+ 2 \left| \sum_{k \neq j} d_{k,j} \mathbb{E}_j [Q_{jj}] \mathbb{E}_j \left[ e_k^T Q_j \xi_{(j,-j)} \right] \right|^2 \tag{197}$$

To establish (194) and thus complete the proof for $s = 1$, we propose to show that the following inequalities hold true:

$$\Big| \sum_{k \neq j} d_{k,j} \mathbb{E}_j [(Q_{jj} - \mathbb{E}_j(Q)_{jj}) e_k^T Q_j \xi_{(j,-j)}] \Big|^2 \leq |\Im z|^{-6} O(p^{-2}) \tag{198}$$

$$\Big| \sum_{k \neq j} d_{k,j} \mathbb{E}_j [Q_{jj}] \mathbb{E}_j \big[ e_k^T Q_j \xi_{(j,-j)} \big] \Big|^2 \leq |\Im z|^{-4} O(p^{-2}) \tag{199}$$

**Proof of** (198)**.** From Corollary 2, it holds that

$$\mathbb{E}_j [ \big| e_k^T Q_j \xi_{(j,-j)} \big|^2 ] \leq |\Im z|^{-2} O(\frac{1}{p}) \tag{200}$$

while using Poincaré-Nash inequality, we have:

$$\begin{aligned}\mathbb{E}_j\left[|Q_{jj}-\mathbb{E}_j[Q_{jj}]|^2\right] &\leq \sum_{l=1}^{p}\mathbb{E}_j\Big[\Big|\frac{\partial Q_{jj}}{\partial Z_{lj}}\Big|^2\Big] \\ &= 16\sum_{b_1\neq j}\sum_{b_2\neq j}\mathbb{E}_j\left[x_{b_1}^T x_j x_{b_2}^T x_j x_{b_1}^T C_{[j]} x_{b_2}|Q_{jj}|^2 Q_{b_1 j}Q_{b_2 j}^*\right] \\ &= 16\mathbb{E}_j\left[|Q_{jj}|^2[Q^H S_j Q]_{jj}\right] \\ &\leq |\Im z|^{-4}O(\frac{1}{p}) \qquad (201)\end{aligned}$$

Using the fact that $\max_k |d_{k,j}| = O(p^{-1})$, we obtain (198) by combining (200) and (201) and applying Cauchy-Schwartz inequality.
**Proof of** (199)**.** Computing the expectation over $x_j$ of the term $\mathbb{E}_j[e_k^T Q_j \xi_{(j,-j)}]$, we get:

$$\begin{aligned}&\left|\sum_{k\neq j} d_{k,j}\mathbb{E}_j[Q_{jj}]\mathbb{E}_j\left[e_k^T Q_j \xi_{(j,-j)}\right]\right|^2 \\ &= \left|\sum_{k\neq j} d_{k,j}\mathbb{E}_j[Q_{jj}]\sum_{m\neq j}[Q_j]_{km}\frac{1}{\sqrt{p}}\left(x_m^T C_{[j]} x_m - \frac{1}{p}\operatorname{tr}(C_{[m]}C_{[j]})\right)\right|^2 \qquad (202) \\ &= \left|\mathbb{E}_j[Q_{jj}]1^T\mathcal{D}\{d_{k,j}\}_{k=1}^n Q_j \mathcal{D}\{d_{m,j}\}_{m=1}^n 1\right|^2 \qquad (203) \\ &\leq |\Im z|^{-4}p^2\|\mathcal{D}\{d_{k,j}\}\|^4 \qquad (204)\end{aligned}$$

Hence, using the fact that $\|\mathcal{D}\{d_{k,j}\}\|^4 = O(p^{-4})$, (199) follows.

Assume that (188) holds for all integer $k = 1,\ldots,s-1$, and let us prove it for $k = s$. To begin with, we use the following relation:

$$\mathbb{E}_j\Big[\Big|\sum_{k\neq j} d_{k,j}Q_{k,j}\Big|^{2s}\Big] = \Big|\mathbb{E}_j\Big[\Big|\sum_{k\neq j} d_{k,j}Q_{kj}\Big|^s\Big]\Big|^2 + \mathbf{var}_j\Big[\Big|\sum_{k\neq j} d_{k,j}Q_{kj}\Big|^s\Big] \qquad (205)$$

and apply the Cauchy-Schwartz inequality together with the induction assumption to treat the first term of the right-hand side of (205) as follows:

$$\begin{aligned}\Big|\mathbb{E}_j\Big[\Big|\sum_{k\neq j} d_{k,j}Q_{kj}\Big|^s\Big]\Big|^2 &\leq \mathbb{E}_j\Big[\Big|\sum_{k\neq j} d_{k,j}Q_{kj}\Big|^{2s-2}\Big]\mathbb{E}_j\Big[\Big|\sum_{k\neq j} d_{k,j}Q_{kj}\Big|^2\Big] \qquad (206) \\ &\leq (|\Im z|^{-4}+|\Im z|^{-6})^s O(p^{-2s+\epsilon}) \qquad (207)\end{aligned}$$

To handle the second term in (205), we invoke the Poincaré-Nash inequality to obtain:

$$\mathbf{var}_j\Big[\Big|\sum_{k\neq j} d_{k,j}Q_{kj}\Big|^s\Big] \leq \mathbb{E}_j\Big[s^2\Big|\sum_{k\neq j} d_{k,j}Q_{kj}\Big|^{2s-2}\sum_{l=1}^{p}\Big|\sum_{k\neq j} d_{k,j}\frac{\partial Q_{kj}}{\partial Z_{lj}}\Big|^2\Big] \qquad (208)$$

and use (190) in combination with the induction assumption and Lemma 5, to ultimately get:

$$\mathbf{var}_j\Big[\Big|\sum_{k\neq j} d_{k,j}Q_{kj}\Big|^s\Big] \leq (|\Im z|^{-4} + |\Im z|^{-6})^s O(p^{-2s+\epsilon}) \tag{209}$$

From (207) and (209) and based on the decomposition in (205), we thus prove the desired result for $k = s$, which completes the proof.

### *A.4 Proof of Lemma 11*

For simplicity, we remove the subscript $p$ from the notation of $A_{1,p}$ and $A_{2,p}$. We will treat the case of $k = j$ since all other cases follow similarly. Call $\vartheta^j_{sb}$ the quantity:

$$\vartheta^j_{sb} = x_s^T A_1 x_j x_b^T A_2 x_s Q_{sj}$$

and let $\chi_{bj}$ be:

$$\chi_{bj} = \sum_{s\notin\{b,j\}} \vartheta^j_{sb}$$

With these notations, Lemma 11 aims to show that:

$$\mathbb{E}\Big[|\chi_{bj}|^2\Big] = O_z(p^{-2+\epsilon}). \tag{210}$$

Decomposing $\chi_{bj}$ as:

$$\chi_{bj} = \chi_{bj} - \mathbb{E}_j[\chi_{bj}] + \mathbb{E}_j[\chi_{bj}]$$

we obtain:

$$\mathbb{E}[|\chi_{bj}|^2] \leq 2\mathbb{E}\Big[|\chi_{bj} - \mathbb{E}_j[\chi_{bj}]|^2\Big] + 2\mathbb{E}\Big[(\mathbb{E}_j[\chi_{bj}])^2\Big] = 2\mathbb{E}[\mathbf{var}_j(\chi_{bj})] + 2\mathbb{E}\Big[(\mathbb{E}_j[\chi_{bj}])^2\Big]$$

where $\mathbf{var}_j$ is the variance with respect to the distribution of $x_j$. To prove the desired result in (210), it suffices to show that:

$$\mathbb{E}[\mathbf{var}_j(\chi_{bj})] = O_z(p^{-2+\epsilon}) \tag{211}$$

$$\mathbb{E}\Big[(\mathbb{E}_j[\chi_{bj}])^2\Big] = O_z(p^{-2+\epsilon}) \tag{212}$$

**Proof of (211).** Based on Poincaré-Nash inequality, we can upper bound $\mathbf{var}_j(\chi_{bj})$ as:

$$\mathbf{var}_j(\chi_{bj}) \leq \sum_{l=1}^{p} \Bigg|\sum_{s\notin\{b,j\}} \frac{\partial \vartheta^j_{sb}}{\partial Z_{lj}}\Bigg|^2 \tag{213}$$

Using the differentiation formula in (29), we obtain:

$$\mathbf{var}_j(\chi_{bj}) \leq 2Z_1 + 8Z_2$$

where

$$Z_1 = \sum_{l=1}^{p} \left| \sum_{s \notin \{b,j\}} \frac{1}{\sqrt{p}} \left[ C_{[j]}^{\frac{1}{2}} A_1 x_s \right]_l x_b^T A_2 x_s Q_{sj} \right|^2 \tag{214}$$

$$Z_2 = \sum_{l=1}^{p} \left| \sum_{s \notin \{b,j\}} x_s^T A_1 x_j x_b^T A_2 x_s \sum_{q \neq j} (x_q^T x_j)[C_{[j]}^{\frac{1}{2}} x_q]_l (Q_{sj} Q_{qj} + Q_{jj} Q_{qs}) \right|^2 \tag{215}$$

Using Lemma 5, the proof of (210) reduces to showing that $Z_i = O_z(p^{-2})$ for $i = 1, 2$.

**Treatment of $Z_1$.** Expanding $Z_1$, we obtain:

$$Z_1 = \sum_{s_1 \notin \{b,j\}} \sum_{s_2 \notin \{b,j\}} \frac{1}{p} x_{s_1}^T A_1 C_{[j]} A_1 x_{s_2} x_b^T A_2 x_{s_1} x_b^T A_2 x_{s_2} Q_{s_1 j} Q_{s_2 j}^* \tag{216}$$

$$= \frac{1}{p} \left[ Q \mathcal{S}_{bj} Q^H \right]_{jj} \tag{217}$$

where $\mathcal{S}_{bj}$ is the $n \times n$ matrix with elements:

$$\mathcal{S}_{bj} = x_{s_1}^T A_1 C_{[j]} A_1 x_{s_2} x_b^T A_2 x_{s_1} x_b^T A_2 x_{s_2} \boldsymbol{\delta}_{s_1 \notin \{b,j\}} \boldsymbol{\delta}_{s_2 \notin \{b,j\}}$$

By Lemma 4, $\|\mathcal{S}_{bj}\| = O(p^{-1})$. Thus $Z_1 = O(p^{-2})$.

**Treatment of $Z_2$.** Clearly, $Z_2$ can be upper-bounded as $Z_2 \leq 2Z_{21} + 2Z_{22}$ where $Z_{21}$ and $Z_{22}$ are given by:

$$Z_{21} = \left| \sum_{s \notin \{b,j\}} x_s^T A_1 x_j x_b^T A_2 x_s Q_{sj} \right|^2 \sum_{l=1}^{p} \left| \sum_{q \neq j} (x_q^T x_j)[C_{[j]}^{\frac{1}{2}} x_q]_l Q_{qj} \right|^2 \tag{218}$$

$$Z_{22} = |Q_{jj}|^2 \sum_{l=1}^{p} \left| \sum_{s \notin \{b,j\}} x_s^T A_1 x_j x_b^T A_2 x_s \sum_{q \neq j} (x_q^T x_j)[C_{[j]}^{\frac{1}{2}} x_q]_l Q_{qs} \right|^2 \tag{219}$$

Denote by $\mathcal{D}_b$ the $n \times n$ diagonal matrix with diagonal elements:

$$[\mathcal{D}_b]_{ss} = x_b^T A_2 x_s \boldsymbol{\delta}_{s \notin \{b,j\}}, \; s = 1, \dots, n$$

and by $\mathcal{S}_j$ the $n \times n$ matrix with elements:

$$[\mathcal{S}_j]_{q_1 q_2} = (x_{q_1}^T x_j)(x_{q_2}^T x_j) x_{q_1}^T C_{[j]} x_{q_2} \boldsymbol{\delta}_{q_1 \neq j} \boldsymbol{\delta}_{q_2 \neq j}$$

With these notations quantities $Z_{21}$ and $Z_{22}$ can be written in a matrix form as:

$$Z_{21} = [Q \mathcal{S}_j Q^H]_{jj} \left| \left[ X^T A_1 X \mathcal{D}_b Q \right]_{jj} \right|^2 \tag{220}$$

$$Z_{22} = |Q_{jj}|^2 x_j^T A_1 X \mathcal{D}_b Q \mathcal{S}_j Q^H \mathcal{D}_b X^T A_1 x_j \tag{221}$$

Using the fact that $\|\mathcal{S}_j\| = O(p^{-1})$ and $\|\mathcal{D}_b\| = O(p^{-\frac{1}{2}})$, we can easily deduce that $Z_{21} = O(p^{-2})$ and $Z_{22} = O(p^{-2})$. Hence $Z_2 = O(p^{-2})$, which completes the treatment of $Z_2$.

**Proof of (212).** Using (33), obtain:

$$\mathbb{E}_j\left[\chi_{bj}\right] = \mathbb{E}_j[\sum_{s\notin\{b,j\}} -x_s^T A_1 x_j x_b^T A_2 x_s Q_{jj} e_s^T Q_j \xi_{(j,-j)}]$$

Since $\mathbb{E}_j[x_s^T A_1 x_j e_s^T Q_j \xi_{(j,-j)}] = 0$, we get:

$$\mathbb{E}_j\left[\chi_{bj}\right] = \mathbb{E}_j[\sum_{s\notin\{b,j\}} -x_s^T A_1 x_j x_b^T A_2 x_s e_s^T Q_j \xi_{(j,-j)}(Q_{jj} - \mathbb{E}[Q_{jj}])]$$

Applying Cauchy-Schwartz inequality, we can bound $\mathbb{E}[|\mathbb{E}_j \chi_{bj}|^2]$ as:

$$\mathbb{E}[|\mathbb{E}_j\chi_{bj}|^2]$$

$$\leq \mathbb{E}\Big[|\sum_{s\notin\{b,j\}} \sqrt{\mathbb{E}_j\left[|e_s^T Q_j \xi_{(j,-j)}|^2\right]}\sqrt{\mathbb{E}_j[|x_s^T A_1 x_j|^2 |Q_{jj} - \mathbb{E}[Q_{jj}]|^2] |x_b^T A_2 x_s|^2}\,|^2\Big] \tag{222}$$

$$\leq \mathbb{E}\Big[\sum_{s\notin\{b,j\}} \mathbb{E}_j\left[|e_s^T Q_j \xi_{(j,-j)}|^2\right] \sum_{s\notin\{b,j\}} \mathbb{E}_j\Big[|x_s^T A_1 x_j|^2 |Q_{jj} - \mathbb{E}[Q_{jj}]|^2\Big] |x_b^T A_2 x_s|^2\Big] \tag{223}$$

Using Corollary 2, it can be shown that $\sum_{s\notin\{b,j\}} \mathbb{E}_j[|e_s^T Q_j \xi_{(j,-j)}|^2]$ can be bounded by $|\Im z|^{-2} O(1)$. On the other hand, it follows from Lemma 8 that $\mathbb{E}[|Q_{jj} - \mathbb{E}[Q_{jj}|^2]] = O_z(p^{-1+\epsilon})$. Noting that for $s \notin \{b,j\}$, both quantities $|x_s^T A_1 x_j|^2$ and $|x_b^T A_2 x_s|^2$ are $O(p^{-1})$, we show (212) by applying Lemma 5.

### *A.5 Proof of Lemma 13*

Recall that $Q_k = (\Phi_k - zI_n)^{-1}$. From the resolvent identity $Q - Q_k = Q(\Phi_k - \Phi)Q_k$, we have:

$$\begin{aligned}
Q_{lr} - [Q_k]_{lr} &= \sum_{m=1}^{n}\sum_{s=1}^{n} Q_{lm}\Big( [\Phi_k]_{ms} - \Phi_{ms}\Big) [Q_k]_{sr} \\
&= -\sum_{m\neq k} \sqrt{p} Q_{lm}\Big((x_k^T x_m)^2 - \frac{1}{p^2}\operatorname{tr} C_{[k]} C_{[m]}\Big) [Q_k]_{kr} \\
&\quad - \sum_{s\neq k} \sqrt{p} Q_{lk}\Big( \left(x_k^T x_s\right)^2 - \frac{1}{p^2}\operatorname{tr} C_{[k]} C_{[s]}\Big) [Q_k]_{sr} \\
&= -\sum_{s\neq k} \sqrt{p} Q_{lk}\Big(\left(x_k^T x_s\right)^2 - \frac{1}{p^2}\operatorname{tr} C_{[k]} C_{[s]}\Big) [Q_k]_{sr}
\end{aligned} \tag{224}$$

where the last equality follows from the fact that $[Q_k]_{kr} = 0$ when $r \neq k$. Using (224), we obtain:

$$\sum_{l\neq k}\sum_{r\neq k}\mathbb{E}\left[b_l c_r Q_{lr} - b_l c_r \left[Q_k\right]_{lr}\right]$$
$$= -\sum_{l\neq k}\sum_{r\neq k}\sum_{s\neq k}\sqrt{p}\mathbb{E}\Big[b_l c_r Q_{lk}\left[Q_k\right]_{sr}\left((x_k^T x_s)^2 - \frac{1}{p^2}\operatorname{tr}(C_{[k]}C_{[s]})\right)\Big]$$
$$= \sum_{l\neq k}\sum_{r\neq k}\sum_{s\neq k}\sqrt{p}\mathbb{E}\Big[b_l c_r e_l^T Q_k \xi_{(k,-k)} Q_{kk}\left[Q_k\right]_{sr}\left((x_k^T x_s)^2 - \frac{1}{p^2}\operatorname{tr} C_{[k]}C_{[s]}\right)\Big]$$
$$= \mathbb{E}\left[b^T Q_k \xi_{(k,-k)} c^T Q_k \xi_{(k,-k)} Q_{kk}\right] \tag{225}$$
$$= \mathbb{E}\Big[b^T Q_k \xi_{(k,-k)} c^T Q_k \xi_{(k,-k)} (Q_{kk} - \mathbb{E}[Q_{kk}])\Big] \tag{226}$$
$$+ \mathbb{E}[Q_{kk}]\mathbb{E}\Big[b^T Q_k \xi_{(k,-k)} c^T Q_k \xi_{(k,-k)}\Big] \tag{227}$$

It follows from Corollary 2 that:

$$\mathbb{E}_k\Big[b^T Q_k \xi_{(k,-k)} c^T Q_k \xi_{(k,-k)}\Big] \leq \|Q_k\|^2 \|b\| \|c\| O(\frac{1}{p})$$

Hence, using Lemma 5,

$$\mathbb{E}\Big[b^T Q_k \xi_{(k,-k)} c^T Q_k \xi_{(k,-k)}\Big] = O_z(p^{-1+\epsilon}) \tag{228}$$

On the other hand, using Lemma 10 along with Corollary 3, we can easily see that:

$$\mathbb{E}\Big[b^T Q_k \xi_{(k,-k)} c^T Q_k \xi_{(k,-k)} (Q_{kk} - \mathbb{E}[Q_{kk}])\Big] = O_z(p^{-1+\epsilon}) \tag{229}$$

Combining (228) and (229), we thus prove the sought-for result.

### *A.6 Proof of Lemma 12*

For ease of notations, we shall drop the subscript $j,p$ from matrices $A_{k,j,p}$, $k = 1,2,3,4$. Call $\hat{\theta}_{r,b}^{j,k}$ the quantity:

$$\hat{\theta}_{r,b}^{j,k} = x_b^T A_1 x_k x_b^T A_2 x_j x_k^T A_3 x_r x_r^T A_4 x_j Q_{br}$$

and let $\Theta$ be:

$$\Theta = \sum_{r\notin\{j,k\}}\sum_{b\notin\{j,r,k\}}\hat{\theta}_{r,b}^{j,k}$$

With these notations, Lemma 12 is equivalent to showing:

$$\mathbb{E}\Big[|\Theta|^2\Big] = O(p^{-3+\epsilon}) \tag{230}$$

Decomposing $\Theta$ as:

$$\Theta = \Theta - \mathbb{E}_j[\Theta] + \mathbb{E}_j[\Theta] \tag{231}$$

we obtain:

$$\mathbb{E}\Big[|\Theta|^2\Big] \leq 2\mathbb{E}\Big[|\Theta - \mathbb{E}_j[\Theta]|^2\Big] + 2\mathbb{E}\Big[\,|\mathbb{E}_j[\Theta]|^2\,\Big] = 2\mathbb{E}[\mathbf{var}_j(\Theta)] + 2\mathbb{E}\Big[\,|\mathbb{E}_j[\Theta]|^2\,\Big]$$

where $\mathbf{var}_j$ is the variance with respect to the distribution of $x_j$. To prove the desired result, we will prove that

$$\mathbb{E}[\mathbf{var}_j(\Theta)] = O_z(p^{-3+\epsilon}) \tag{232}$$

$$\mathbb{E}[|\mathbb{E}_j(\Theta)|^2] = O_z(p^{-3+\epsilon}) \tag{233}$$

Proof of (232) Based on Poincaré-Nash inequality, and using the differentiation formula in (29), $\mathbf{var}_j(\Theta)$ can be bounded as:

$$\mathbf{var}_j(\Theta) \leq \sum_{l=1}^{p} \mathbb{E}_j\Big[\big|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,j,k\}} \frac{\partial\theta_{r,b}^{j,k}}{\partial Z_{lj}}\big|^2\Big] \leq 8(\xi_1+\xi_2+\xi_3+\xi_4) \tag{234}$$

where

$$\xi_1 = \sum_{l=1}^{p}\Big|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,j,k\}} x_b^T A_1 x_k x_k^T A_3 x_r \frac{1}{\sqrt{p}}\Big[C_{[j]}^{\frac{1}{2}} A_2 x_b\Big]_l x_r^T A_4 x_j Q_{br}\Big|^2 \tag{235}$$

$$\xi_2 = \sum_{l=1}^{p}\Big|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,j,k\}} x_b^T A_1 x_k x_k^T A_3 x_r \frac{1}{\sqrt{p}}\Big[C_{[j]}^{\frac{1}{2}} A_2 x_b\Big]_l x_b^T A_2 x_j Q_{br}\Big|^2 \tag{236}$$

$$\xi_3 = \sum_{l=1}^{p}\Big|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,j,k\}}\sum_{s\neq j} x_b^T A_1 x_k x_k^T A_3 x_r x_b^T A_2 x_j x_r^T A_4 x_j (x_s^T x_j)[C_{[j]}^{\frac{1}{2}} x_s]_l Q_{bj} Q_{sr}\Big|^2 \tag{237}$$

$$\xi_4 = \sum_{l=1}^{p}\Big|\sum_{r\notin\{j,k\}}\sum_{b\notin\{r,j,k\}}\sum_{s\neq j} x_b^T A_1 x_k x_k^T A_3 x_r x_b^T A_2 x_j x_r^T A_4 x_j \times (x_s^T x_j)[C_{[j]}^{\frac{1}{2}} x_s]_l Q_{jr} Q_{sb}\Big|^2. \tag{238}$$

The treatment of $\xi_1$ and $\xi_2$ is similar. The same also holds for $\xi_3$ and $\xi_4$. We will thus only show that $\xi_1 = O(p^{-3})$ and $\xi_3 = O(p^{-3})$. Then (232) follows from Lemma 5.

**Treatment of $\xi_1$.** For $t = 1, \cdots, 4$, and $j = 1, \cdots, n$, denoting by $\mathcal{D}_{j,t}$ the diagonal matrix whose $r$-th diagonal element is given by $x_r^T A_t x_j \boldsymbol{\delta}_{r\notin\{k,j\}}$ and by $S_j$ the $n \times n$ matrix with entries:

$$[S_j]_{b_1,b_2} = x_{b_1}^T A_1 x_k x_{b_2}^T A_1 x_k x_{b_1}^T A_2 C_{[j]} A_2 x_{b_2} \boldsymbol{\delta}_{b_1\neq k} \boldsymbol{\delta}_{b_1\neq k} \boldsymbol{\delta}_{b_2\neq k}$$

we may write $\xi_1$ in a matrix form as:

$$\xi_1 = \frac{1}{p} x_k^T A_3 X \mathcal{D}_{j,4}(Q - \mathcal{D}(Q)) S_j (Q^H - \mathcal{D}(Q^H)) \mathcal{D}_{j,4} X^T A_3 x_k$$

where we recall that $\mathcal{D}(Q)$ denotes the diagonal matrix formed by the diagonal elements of $Q$ and $X = [x_1, \cdots, x_n]$. Since $\|S_j\| = O(p^{-1})$ and $\|\mathcal{D}_{j,4}\| = O(p^{-\frac{1}{2}})$, we obtain:

$$\xi_1 = O(p^{-3}).$$

**Treatment of $\xi_3$.** To treat $\xi_3$, we decompose it as the difference between two terms associated with indexes $b \notin \{j,k\}$ and $b = r$. In doing so, we obtain:

$$\xi_3 \leq 2(\xi_{31} + \xi_{32}) \tag{239}$$

where

$$\xi_{31}=\sum_{l=1}^{p} \Big| \sum_{r\notin\{j,k\}} \sum_{b\notin\{j,k\}} \sum_{s\neq j} x_b^T A_1 x_k x_k^T A_3 x_r x_b^T A_2 x_j x_r^T A_4 x_j x_s^T x_j [C_{[j]}^{\frac{1}{2}} x_s]_l Q_{bj} Q_{sr} \Big|^2 \tag{240}$$

$$\xi_{32} = \sum_{l=1}^{p} \Big| \sum_{r\notin\{j,k\}} \sum_{s\neq j} x_r^T A_1 x_k x_k^T A_3 x_r x_r^T A_2 x_j x_r^T A_4 x_j x_s^T x_j [C_{[j]}^{\frac{1}{2}} x_s]_l Q_{rj} Q_{sr} \Big|^2 \tag{241}$$

Let $\tilde{S}_j$ the $n \times n$ matrix with elements:

$$[\tilde{S}_j]_{s_1,s_2} = x_{s_1}^T x_j x_{s_1}^T C_{[j]} x_{s_2} x_{s_2}^T x_j \boldsymbol{\delta}_{s_1 \neq j} \boldsymbol{\delta}_{s_2 \neq j}$$

Quantity $\xi_{31}$ can be written in a matrix form as:

$$\begin{aligned} \xi_{31} &= \sum_{l=1}^{p} \Big| \sum_{s\neq j} \Big[ x_k^T A_3 X \mathcal{D}_{j,4} Q \Big]_s \Big[ x_j^T A_2 X \mathcal{D}_{k,1} Q \Big]_j x_s^T x_j [C_{[j]}^{\frac{1}{2}} x_s]_l \Big|^2 &(242)\\ &= \Big| \Big[ x_j^T A_2 X \mathcal{D}_{k,1} Q \Big]_j \Big|^2 x_k^T A_3 X \mathcal{D}_{j,4} Q \tilde{S}_j Q^H \mathcal{D}_{j,4} X^T A_3 x_k &(243)\\ &= O(p^{-3}) &(244)\end{aligned}$$

where the last estimate stems from the fact that $\|\mathcal{D}_{j,t}\| = O(p^{-\frac{1}{2}})$ and $\|\tilde{S}_j\| = O(p^{-1})$.

Similarly, $\xi_{32}$ can be written in a matrix form as:

$$\begin{aligned} \xi_{32} &= \sum_{l=1}^{p} \Big| \Big[ Q \mathcal{D}_{k,1} \mathcal{D}_{k,3} \mathcal{D}_{j,4} \mathcal{D}_{j,2} Q \Big]_{sj} [C_{[j]}^{\frac{1}{2}} x_s]_l (x_s^T x_j) \Big|^2 &(245)\\ &= \Big[ Q \mathcal{D}_{k,1} \mathcal{D}_{k,3} \mathcal{D}_{j,4} \mathcal{D}_{j,2} Q \tilde{S}_j Q^H \mathcal{D}_{k,1} \mathcal{D}_{k,3} \mathcal{D}_{j,4} \mathcal{D}_{j,2} Q^H \Big]_{jj} &(246)\\ &= O(p^{-5}) &(247)\end{aligned}$$

Combining (244) with (247) we get $\xi_3 = O(p^{-3})$.

Proof of (233). Using the integration by part formula, we may simplify $\mathbb{E}_j(\Theta)$ as:

$$\mathbb{E}_j(\Theta) = \Theta_{j,1} + \Theta_{j,2} + \Theta_{j,3} \tag{248}$$

where $\Theta_{j,1}, \Theta_{j,2}$ and $\Theta_{j,3}$ are given by:

$$\Theta_{j,1} = \sum_{r\notin\{j,k\}} \sum_{b\notin\{j,k,r\}} \mathbb{E}_j\left[\frac{1}{p} x_b^T A_1 x_k x_b^T A_2 C_{[j]} A_4 x_r x_k^T A_3 x_r Q_{br}\right] \tag{249}$$

$$\Theta_{j,2} = \frac{-2}{\sqrt{p}} \sum_{r\notin\{j,k\}} \sum_{b\notin\{j,k,r\}} \mathbb{E}_j\Big[\sum_{s\neq j} x_b^T A_1 x_k x_b^T A_2 x_j x_k^T A_3 x_r x_r^T A_4 C_{[j]} x_s \times x_s^T x_j Q_{jb} Q_{sr}\Big] \tag{250}$$

$$\Theta_{j,3} = \frac{-2}{\sqrt{p}} \sum_{r\notin\{j,k\}} \sum_{b\notin\{j,k,r\}} \mathbb{E}_j\Big[\sum_{s\neq j} x_b^T A_1 x_k x_b^T A_2 x_j x_k^T A_3 x_r x_r^T A_4 C_{[j]} \times x_s x_s^T x_j Q_{jr} Q_{sb}\Big] \tag{251}$$

The proof of (233) amounts to showing that for $i = 1, 2, 3$, $\mathbb{E}[|\Theta_{j,i}|^2] = O_z(p^{-3+\epsilon})$.

**Treatment of $\mathbb{E}[|\Theta_{j,1}|^2]$.** Using Cauchy-Schwartz inequality, we can bound $|\Theta_{j,1}|^2$ as:

$$\mathbb{E}[|\Theta_{j,1}|^2] \leq \mathbb{E}\Big[\sum_{r\notin\{j,k\}} \frac{|x_k^T A_3 x_r|^2}{p^2} \sum_{r\notin\{j,k\}} \Big|\sum_{b\notin\{j,k,r\}} \mathbb{E}_j\Big[x_b^T A_1 x_k x_b^T A_2 C_{[j]} A_4 x_r Q_{br}\Big]\Big|^2\Big] \tag{252}$$

$$= O_z(p^{-3+\epsilon}) \tag{253}$$

where the last estimate follows by using the fact that $\sum_{r\notin\{j,k\}} |x_k^T A_3 x_r|^2 = O(1)$ and that $\mathbb{E}[\Big|\sum_{b\notin\{j,k,r\}} \Big[x_b^T A_1 x_k x_b^T A_2 C_{[j]} A_4 x_r Q_{br}\Big]\Big|^2 = O_z(p^{-2+\epsilon})]$ from Lemma 11.

**Treatment of $\mathbb{E}[|\Theta_{j,2}|^2]$.** To begin with, we decompose $\Theta_{j,2}$ as the sum of two terms associated with index $s \notin \{j, r\}$ and with $s = r$, respectively:

$$\Theta_{j,2} = \Theta_{j,2,1} + \Theta_{j,2,2} \tag{254}$$

where

$$\Theta_{j,2,1} = \frac{-2}{\sqrt{p}} \sum_{r\notin\{j,k\}} \sum_{b\notin\{j,k,r\}} \mathbb{E}_j\Big[\sum_{s\notin\{j,r\}} x_b^T A_1 x_k x_b^T A_2 x_j x_k^T A_3 x_r x_r^T A_4 C_{[j]} x_s \times x_s^T x_j Q_{jb} Q_{sr}\Big]$$

$$\Theta_{j,2,2} = \frac{-2}{\sqrt{p}} \sum_{r\notin\{j,k\}} \sum_{b\notin\{j,k,r\}} \mathbb{E}_j\Big[x_b^T A_1 x_k x_b^T A_2 x_j x_k^T A_3 x_r x_r^T A_4 C_{[j]} x_r \times x_r^T x_j Q_{jb} Q_{rr}\Big] \tag{255}$$

The term $\mathbb{E}|\Theta_{j,2,1}|^2$ can be handled using Cauchy-Schwartz inequality in combination with Lemma 11 as follows:

$$\mathbb{E}[|\Theta_{j,2,1}|^2]$$
$$\leq \frac{4}{p}\mathbb{E}\Big[\sum_{r\notin\{j,k\}}|x_r^TA_3x_k|^2 \tag{256}$$
$$\times \sum_{r\notin\{j,k\}}\Big|\mathbb{E}_j\Big[\sum_{b\notin\{j,k,r\}}x_b^TA_1x_kx_b^TA_2x_jQ_{bj}\sum_{s\notin\{j,r\}}x_r^TA_4C_{[j]}x_s(x_s^Tx_j)Q_{sr}\Big]\Big|^2\Big] \tag{257}$$
$$\leq \frac{4}{p}\mathbb{E}\Big[\sum_{r\notin\{j,k\}}|x_r^TA_3x_k|^2\sum_{r\notin\{j,k\}}\mathbb{E}_j\Big[\Big|\sum_{b\notin\{j,k,r\}}x_b^TA_1x_kx_b^TA_2x_jQ_{bj}\Big|^2\Big]$$
$$\times\,\mathbb{E}_j\Big[\Big|\sum_{s\notin\{j,r\}}x_r^TA_4C_{[j]}x_s(x_s^Tx_j)Q_{sr}\Big|^2\Big]\Big] \tag{258}$$

Using Lemma 11 along with Lemma 5, we obtain $\mathbb{E}[|\Theta_{j,2,1}|^2]=O_z(p^{-3+\epsilon})$. On the other hand, using the fact that $\mathbb{E}|\mathbb{E}_jX|^2\leq\mathbb{E}|X|^2$, we can upper-bound $\mathbb{E}\Big[|\Theta_{j,2,2}|^2\Big]$ as:

$$\mathbb{E}\Big[|\Theta_{j,2,2}|^2\Big]\leq\frac{8}{p}\mathbb{E}\Big[|\sum_{r\notin\{j,k\}}x_k^TA_3x_r(x_r^Tx_j)Q_{rr}\Big|^2\Big|\sum_{b\notin\{j,k\}}x_b^TA_1x_kx_b^TA_2x_jQ_{jb}\Big|^2\Big] \tag{259}$$
$$+\frac{8}{p}\mathbb{E}\Big[|\sum_{r\notin\{j,k\}}x_k^TA_3x_rx_r^Tx_jQ_{rr}x_r^TA_1x_kx_r^TA_2x_jQ_{jr}\Big|^2\Big] \tag{260}$$
$$=O_z(p^{-3+\epsilon}) \tag{261}$$

where the last estimate follows by Lemma 11.

**Treatment of** $\mathbb{E}[|\Theta_{j,3}|^2]$**.** Writing $\Theta_{j,3}$ in a matrix form, we get:

$$\Theta_{j,3}$$
$$=-\frac{2}{\sqrt{p}}\sum_{r\notin\{j,k\}}\mathbb{E}_j\Big[x_k^TA_3x_rx_r^TA_4C_{[j]}X\mathcal{D}\{x_s^Tx_j\boldsymbol{\delta}_{s\neq j}\}Q\mathcal{D}\{x_b^TA_2x_j\boldsymbol{\delta}_{b\notin\{j,k,r\}}\}$$
$$\times X^TC_{[j]}x_kQ_{jr}\Big]$$

Using the facts that:

$$x_k^TA_3x_r=O(p^{-\frac{1}{2}})$$
$$\|\mathcal{D}\{x_s^Tx_j\boldsymbol{\delta}_{s\neq j}\}\|=O(p^{-\frac{1}{2}})$$
$$\mathbb{E}[|Q_{jr}|^2]=O_z(p^{-1+\epsilon})$$

we obtain $\mathbb{E}|\Theta_{j,3}|^2=O_z(p^{-3+\epsilon})$.

# Appendix B Proof of Proposition 1 and Proposition 2

## *B.1 Proof of Proposition 1*

To simplify the exposition of the proof, we remove the subscript $n$ in the notation of $a_{k,n}$ and $b_{k,n}$ and the argument $z$ from $\Upsilon(a,b,z)$. In addition, we introduce the following quantity:

$$g_k = z + \xi_{(k,-k)}^T Q_k \xi_{(k,-k)}$$

Using (33) and (32), we may decompose $\Upsilon(a,b)$ as:

$$\Upsilon(a,b) = \sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}\left[e_r^T Q_k \xi_{(k,-k)} g_k^{-1}\right] \tag{262}$$

$$= \sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}\left[e_r^T Q_k \xi_{(k,-k)} (\mathbb{E}_k(g_k))^{-1}\right] \tag{263}$$

$$- \sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}\Big[\frac{e_r^T Q_k \xi_{(k,-k)}\big(\xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}^T Q_k \xi_k]\big)}{g_k(\mathbb{E}_k(g_k))}\Big] \tag{264}$$

$$= \Upsilon_1(a,b) + \epsilon(a,b) \tag{265}$$

where

$$\Upsilon_1(a,b) = \sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}\left[e_r^T Q_k \xi_{(k,-k)} (\mathbb{E}_k(g_k))^{-1}\right]$$

$$\epsilon(a,b) = -\sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \mathbb{E}\Big[\frac{e_r^T Q_k \xi_{(k,-k)}\big(\xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}^T Q_k \xi_k]\big)}{g_k \mathbb{E}_k(g_k)}\Big]$$

Based on (265), the proof amounts to showing that:

$$\Upsilon_1(a,b) = -\sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \tilde{\alpha}_{k,r} \mathbb{E}(Q_{kk}) + O_z(p^{-\frac{1}{2}+\epsilon}) \tag{266}$$

$$\epsilon(a,b) = O_z(p^{-\frac{1}{2}+\epsilon}) \tag{267}$$

and to check that the estimates in (266) and (267) hold true with an error $O_z(p^{-1+\epsilon})$ when $\sum_{k=1}^{n}|b_k|$ is uniformly bounded in $n$.
Proof of (266) Taking the expectation with respect to $x_k$, we obtain:

$$\Upsilon_1(a,b) = \sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \sum_{l\neq k} \mathbb{E}\left[[Q_k]_{rl}\frac{1}{\sqrt{p}}(x_l^T C_{[k]} x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))(\mathbb{E}_k(g_k))^{-1}\right] \tag{268}$$

$$= \sum_{k=1}^{n} \sum_{r \neq k} a_k b_r \sum_{l \neq k} \mathbb{E}\left[ [Q_k]_{rl} \frac{1}{\sqrt{p}} (x_l^T C_{[k]} x_l - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[l]}))(g_k)^{-1} \right] \tag{269}$$

$$+ \sum_{k=1}^{n} \sum_{r \neq k} a_k b_r \sum_{l \neq k} \mathbb{E}\left[ [Q_k]_{rl} \frac{1}{\sqrt{p}} (x_l^T C_{[k]} x_l - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[l]})) \frac{(g_k - \mathbb{E}_k(g_k))}{g_k(\mathbb{E}_k(g_k))} \right] \tag{270}$$

$$= \Upsilon_{11}(a,b) + \epsilon_1(a,b) \tag{271}$$

where

$$\Upsilon_{11}(a,b) = \sum_{k=1}^{n} \sum_{r \neq k} a_k b_r \sum_{\ell \neq k} \mathbb{E}\left[ [Q_k]_{rl} \frac{1}{\sqrt{p}} (x_l^T C_{[k]} x_l - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[l]}))(g_k)^{-1} \right] \tag{272}$$

$$\epsilon_1(a,b) = \sum_{k=1}^{n} \sum_{r \neq k} \sum_{\ell \neq k} \mathbb{E}\Big[ \frac{\frac{a_k b_r [Q_k]_{rl}}{\sqrt{p}} (x_l^T C_{[k]} x_l - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[l]}))(g_k - \mathbb{E}_k(g_k))}{g_k(\mathbb{E}_k(g_k))} \Big] \tag{273}$$

To control $\epsilon_1(a,b)$, we use the fact that $Q_k$ is independent of $x_k$ to obtain:

$$\epsilon_1(a,b) = -\sum_{k=1}^{n} \sum_{r \neq k} \sum_{l \notin \{r,k\}} \mathbb{E}\Big[ \frac{\frac{a_k b_r [Q_k]_{rl}}{\sqrt{p}} (x_l^T C_{[k]} x_l - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[l]}))(g_k - \mathbb{E}_k(g_k))^2}{g_k(\mathbb{E}_k g_k)^2} \Big]$$
$$- \sum_{k=1}^{n} \sum_{r \neq k} a_k b_r \mathbb{E}\left[ [Q_k]_{rr} \frac{1}{\sqrt{p}} (x_r^T C_{[k]} x_r - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[r]})) \frac{(g_k - \mathbb{E}_k(g_k))^2}{g_k(\mathbb{E}_k g_k)^2} \right] \tag{274}$$

Both terms involved in the expression of $\epsilon_1(a,b)$ can be shown to be $O_z(p^{-1+\epsilon})$. Indeed, the first term can be upper-bounded using Cauchy-Schwartz inequality as follows:

$$\Big| \sum_{k=1}^{n} \sum_{r \neq k} a_k b_r \sum_{l \notin \{r,k\}} \mathbb{E}\left[ [Q_k]_{rl} \frac{1}{\sqrt{p}} (x_l^T C_{[k]} x_l - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[l]})) \frac{|g_k - \mathbb{E}(g_k)|^2}{g_k(\mathbb{E}_k g_k)^2} \right] \Big| \tag{275}$$

$$\leq |\Im z|^{-3} \sum_{k=1}^{n} \sum_{r \neq k} |a_k||b_r| \sqrt{\mathbb{E}\Big[ \Big| \sum_{l \notin \{r,k\}} [Q_k]_{rl} \frac{1}{\sqrt{p}} (x_l^T C_{[k]} x_l - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[l]})) \Big|^2 \Big]}$$
$$\times \sqrt{\mathbb{E}|g_k - \mathbb{E}_k(g_k)|^4} \tag{276}$$

where we used in (276) the fact that $\max(|(g_k)^{-1}|, |(\mathbb{E}_k(g_k))^{-1}|) \leq |\Im z|^{-1}$. To continue, we leverage Lemma 9 and Lemma 6 along with Lemma 5 to show that

$$\mathbb{E}\Big[ \Big| \frac{1}{\sqrt{p}} \sum_{l \neq r} [Q_k]_{rl} (x_l^T C_{[k]} x_l - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[l]})) \Big|^2 \Big] = O_z(p^{-2+\epsilon}) \tag{277}$$

and

$$\mathbb{E}|g_k - \mathbb{E}_k(g_k)|^4 = O_z(p^{-2+\epsilon}) \tag{278}$$

Using these estimates, we can easily check that the upper bound in (276) is $O_z(p^{-1+\epsilon})$. Similarly, using the fact that

$$\frac{1}{\sqrt{p}}(x_r^T C_{[k]} x_r - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[r]})) = O(p^{-1})$$

together with (278), we prove that the second term in the right-hand side of (274) is also $O_z(p^{-1+\epsilon})$.

Now, recalling that $Q_{kk} = -(g_k)^{-1}$, $\Upsilon_{11}(a,b)$ can be decomposed as:

$$\begin{aligned}\Upsilon_{11}(a,b) = &-\sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \sum_{l\neq k}\mathbb{E}\big[[Q_k]_{rl}\frac{1}{\sqrt{p}}(x_l^T C_{[k]}x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))\mathbb{E}(Q_{kk})\big] \\ &+\sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \sum_{l\neq k}\mathbb{E}\big[[Q_k]_{rl}\frac{1}{\sqrt{p}}(x_l^T C_{[k]}x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))(\mathbb{E}(Q_{kk}) - Q_{kk})\big]\end{aligned} \tag{279}$$

Obviously the second term in the right-hand side of (279) is $O_z(p^{-\frac{1}{2}+\epsilon})$ and becomes $O_z(p^{-1+\epsilon})$ when $\sum_{k=1}^n |b_k|$ is uniformly bounded in $n$. This can be shown by decomposing it as in (274) and then using Lemma 8 together with (277). We thus obtain

$$\begin{aligned}\Upsilon_{11}(a,b) = &-\sum_{k=1}^{n}\sum_{r\neq k} a_k b_r \sum_{l\neq k}\mathbb{E}\big[[Q_k]_{rl}\frac{1}{\sqrt{p}}(x_l^T C_{[k]}x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))\mathbb{E}(Q_{kk})\big] \\ &+ O_z(p^{-\frac{1}{2}+\epsilon})\end{aligned}$$

and the estimate becomes $O(p^{-1+\epsilon})$ when $\sum_{k=1}^n |b_k|$ is uniformly bounded in $n$. To complete the proof of (266), we first note that by Lemma 13[4],

$$\sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \sum_{l\neq k}\mathbb{E}\Big[[Q_k]_{rl}\frac{1}{\sqrt{p}}(x_l^T C_{[k]}x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))\Big] \tag{280}$$

$$= \sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \sum_{l\neq k}\mathbb{E}\Big[[Q]_{rl}\frac{1}{\sqrt{p}}(x_l^T C_{[k]}x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))\Big] + O_z(p^{-1+\epsilon}) \tag{281}$$

Here, it is worth mentioning that (266) is not exactly proved since, to make $\alpha_r$ appear, first, the index $l$ should be different from $r$ (and not different from $k$) and second matrix $C_{[k]}$ should be $C_{[r]}$ instead. For the moment, we focus on fixing the first problem and consider showing that:

$$\sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \sum_{l\neq k}\mathbb{E}\Big[[Q]_{rl}\frac{1}{\sqrt{p}}(x_l^T C_{[k]}x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))\Big]$$

[4]Here, we apply Lemma 13 with vector $c = \left\{x_r^T C_{[k]} x_r - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[r]})\right\}_{r=1}^n$.

$$= \sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \sum_{l\neq r} \mathbb{E}\Big[[Q]_{rl}\frac{1}{\sqrt{p}}(x_l^T C_{[k]} x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))\Big] + O_z(p^{-\frac{1}{2}+\epsilon}) \tag{282}$$

where the error in (282) becomes $O_z(p^{-1+\epsilon})$ when $\sum_{r=1}^n |b_n|$ is bounded. To show (282), we start from the following decomposition:

$$\begin{aligned}
&\sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \sum_{l\neq k} \mathbb{E}\Big[[Q]_{rl}\frac{1}{\sqrt{p}}(x_l^T C_{[k]} x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))\Big]\\
&= \sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \sum_{l\notin\{r,k\}} \mathbb{E}\Big[[Q]_{rl}\frac{1}{\sqrt{p}}(x_l^T C_{[k]} x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))\Big]\\
&+ \sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \mathbb{E}\Big[[Q]_{rr}\frac{1}{\sqrt{p}}(x_r^T C_{[k]} x_r - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[r]}))\Big] \qquad (283)\\
&= \sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \sum_{l\neq r} \mathbb{E}\Big[[Q]_{rl}\frac{1}{\sqrt{p}}(x_l^T C_{[k]} x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))\Big]\\
&- \sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \mathbb{E}\Big[[Q]_{rk}\frac{1}{\sqrt{p}}(x_k^T C_{[k]} x_k - \frac{1}{p}\operatorname{tr}(C_{[r]}C_{[k]}))\Big]\\
&+ \sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \mathbb{E}\Big[[Q]_{rr}\frac{1}{\sqrt{p}}(x_r^T C_{[k]} x_r - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[r]}))\Big] \qquad (284)\\
&= \sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \sum_{l\neq r} \mathbb{E}\Big[[Q]_{rl}\frac{1}{\sqrt{p}}(x_l^T C_{[k]} x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))\Big] + O_z(p^{-\frac{1}{2}+\epsilon}) \qquad (285)
\end{aligned}$$

and the error in (285) becomes $O_z(p^{-1+\epsilon})$ when $\sum_{r=1}^n |b_r|$ is bounded. To obtain (285), we used the fact that $\sum_{k=1}^n \sum_{r\neq k} |a_k||b_r| = O(p)$ and becomes $O(\sqrt{p})$ when $\sum_{r=1}^n |b_r|$ is bounded together with $\frac{1}{\sqrt{p}}(x_l^T C_{[r]} x_l - \frac{1}{p}\operatorname{tr}(C_{[r]}C_{[l]})) = O(p^{-1})$ and $\mathbb{E}|Q_{rk}|^2 = O_z(p^{-1+\epsilon})$. All this allows us to show that the second term in (284) satisfies:

$$\mathbb{E}\Bigg[\Bigg|\sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \mathbb{E}\left[[Q]_{rk}\frac{1}{\sqrt{p}}(x_k^T C_{[r]} x_k - \frac{1}{p}\operatorname{tr}(C_{[r]}C_{[k]}))\right]\Bigg|\Bigg] = O_z(p^{-\frac{1}{2}+\epsilon}) \tag{286}$$

and becomes $O(p^{-1+\epsilon})$ when $\sum_{r=1}^n |b_r|$ is bounded. The last term in (284) can be handled by noticing that $\mathbb{E}[\frac{1}{\sqrt{p}}(x_r^T C_{[k]} x_r - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]}))] = 0$, which allows us to write it as:

$$\sum_{k=1}^{n}\sum_{r\neq k} a_k \mathbb{E}[Q_{kk}] b_r \mathbb{E}\left[[Q]_{rr}\frac{1}{\sqrt{p}}(x_r^T C_{[k]} x_r - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[r]}))\right]$$

$$= \sum_{k=1}^{n} \sum_{r \neq k} a_k \mathbb{E}[Q_{kk}] b_r \mathbb{E}\left[ ([Q]_{rr} - \mathbb{E}[Q_{rr}]) \frac{1}{\sqrt{p}} (x_r^T C_{[k]} x_r - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[r]})) \right] \tag{287}$$

Using the facts that

$$\mathbb{E}\left[ |Q_{rr} - \mathbb{E}[Q_{rr}]|^2 \right] = O_z(p^{-1+\epsilon}) \quad \frac{1}{\sqrt{p}} (x_r^T C_{[k]} x_r - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[r]})) = O(p^{-1})$$

$$\sum_{k=1}^{n} \sum_{r \neq k} a_k \mathbb{E}[Q_{kk}] b_r \mathbb{E}\left[ [Q]_{rr} \frac{1}{\sqrt{p}} (x_r^T C_{[k]} x_r - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[r]})) \right] = O_z(p^{-\frac{1}{2}+\epsilon}) \tag{288}$$

and becomes $O_z(p^{-1+\epsilon})$ when $\sum_{r=1}^{n} |b_r|$ is bounded. Combining (286) and (288) we thus prove (285) and hence (266).

Proof of (267). First, considering the decomposition of $g_k^{-1}$ as:

$$g_k^{-1} = (\mathbb{E}_k g_k)^{-1} + (\mathbb{E}_k(g_k) - g_k)(g_k)^{-1} (\mathbb{E}_k g_k)^{-1} \tag{289}$$

$$= (\mathbb{E}_k g_k)^{-1} - (\xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}^T Q_k \xi_{(k,-k)}])(g_k)^{-1} (\mathbb{E}_k g_k)^{-1} \tag{290}$$

we can write $\epsilon(a, b)$ as:

$$\epsilon(a, b) = \epsilon_2(a, b) + \epsilon_3(a, b)$$

where

$$\epsilon_2(a,b) = -\sum_{k=1}^{n} \sum_{r \neq k} \mathbb{E}\Big[ \frac{a_k b_r e_r^T Q_k \xi_{(k,-k)} \left( \xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}^T Q_k \xi_{(k,-k)}] \right)}{(\mathbb{E}_k g_k)^2} \Big] \tag{291}$$

$$\epsilon_3(a,b) = \sum_{k=1}^{n} \sum_{r \neq k} \mathbb{E}\Big[ \frac{a_k b_r e_r^T Q_k \xi_{(k,-k)} \left( \xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}^T Q_k \xi_{(k,-k)}] \right)^2}{(\mathbb{E}_k g_k)^2 g_k} \Big] \tag{292}$$

Based on Cauchy-Schwartz inequality, we can upper-bound $\epsilon_3(a, b)$ as:

$$|\epsilon_3(a,b)| \leq \sum_{k=1}^{n} \sum_{r \neq k} |a_k||b_r||\Im z|^{-3} \sqrt{\mathbb{E}[|e_r^T Q_k \xi_{(k,-k)}|^2]} \times \sqrt{\mathbb{E}[|\xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}^T Q_k \xi_{(k,-k)}]|^4]} \tag{293}$$

and hence, using Lemma 6 and Lemma 10, we can easily check that $\epsilon_3(a, b) = O_z(p^{-\frac{1}{2}+\epsilon})$ and becomes $O_z(p^{-1+\epsilon})$ when $\sum_{k=1}^{n} |b_k|$ is uniformly bounded in $n$.

It remains thus to check that $\epsilon_2(a,b) = O_z(p^{-1+\epsilon})$. The proof relies on computing the expectation of $\epsilon_3(a,b)$ with respect to $x_k$. For that, it suffices to compute the following quantity:

$$\Gamma_r = \mathbb{E}_k \left[ e_r^T Q_k \xi_{(k,-k)} \left( \xi_{(k,-k)}^T Q_k \xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}^T Q_k \xi_{(k,-k)}] \right) \right]$$

Particularly, we prove that $\Gamma_r$ is given by:

$$\Gamma_r = 8\Gamma_{1,r} + 4\Gamma_{2,r} \tag{294}$$

with

$$\Gamma_{1,r} = \sum_{l\neq k} \sum_{m\neq k} \sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} [Q_k]_{mq} x_m^T C_{[k]} x_q x_q^T C_{[k]} x_l x_l^T C_{[k]} x_m \tag{295}$$

$$\Gamma_{2,r} = \sum_{l\neq k} \sum_{m\neq k} \sum_{q\neq k} p^{-\frac{3}{2}} [Q_k]_{rl} [Q_k]_{mq} \left( x_q^T C_{[k]} x_q - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[q]}) \right) (x_l^T C_{[k]} x_m)^2 \tag{296}$$

Before proving (294), let us see how it leads to (267). Indeed, in view of (292) and Lemma 5, it suffices to check that:

$$\sum_{r\neq k} b_r \Gamma_{1,r} = O(p^{-1}) \tag{297}$$

$$\sum_{r\neq k} b_r \Gamma_{2,r} = O(p^{-\frac{3}{2}}) \tag{298}$$

and that $\sum_{r\neq k} b_r \Gamma_{1,r} = O(p^{-\frac{3}{2}})$ when $\sum_{k=1}^n |b_{n,k}|$ is uniformly bounded in $n$. To prove (297), we write $\sum_{r\neq k} b_r \Gamma_{1,r}$ in a matrix form as:

$$\sum_{r\neq k} b_r \Gamma_{1,r} = p^{-\frac{3}{2}} b^T Q_k \mathcal{D} \left\{ X_k^T C_{[k]} X_k \left( Q_k \odot X_k^T C_{[k]} X_k \right) X_k^T C_{[k]} X_k \right\} 1_n$$

where $X = [x_1, \cdots, x_n]$ and $X_k$ is matrix $X$ with the $k$-th column and the $k$-th row replaced by zero vectors. Then (297) follows by applying Lemma 1 and using the fact that the spectral norm of $X_k^T C_{[k]} X_k$ is $O(1)$. Moreover, we can easily see that when $\sum_{k=1}^n |b_{n,k}|$ is uniformly bounded in $n$, $\sum_{r\neq k} b_r \Gamma_{1,r} = O(p^{-\frac{3}{2}})$.

Similarly, we may write $\sum_{r\neq k} b_r \Gamma_{2,r}$ in a matrix form as:

$$\sum_{r\neq k} b_r \Gamma_{2,r}$$
$$= p^{-\frac{3}{2}} b^T Q_k \big( X_k^T C_{[k]} X_l \odot X_k^T C_{[k]} X_l \big) Q_k \mathcal{D} \big\{ \big( x_q^T C_{[k]} x_q - \frac{1}{p} \operatorname{tr}(C_{[k]} C_{[q]}) \big) \boldsymbol{\delta}_{q\neq k} \big\}_{q=1}^n 1$$

and hence (298) holds true. To complete the proof, it remains thus to check (294). For that, we decompose $\xi_{(k,-k)}$ as:

$$\xi_{(k,-k)} = \xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}] + \mathbb{E}_k[\xi_{(k,-k)}]$$

and expand $\xi_{(k,-k)}^T Q_k \xi_{(k,-k)}$ as:

$$\xi_{(k,-k)}^T Q_k \xi_{(k,-k)} = \left(\xi_{(k,-k)}^T - \mathbb{E}_k[\xi_{(k,-k)}^T]\right) Q_k \left(\xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}]\right) \\ + 2\mathbb{E}_k[\xi_{k,-k}^T] Q_k \left(\xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}]\right) + \mathbb{E}_k[\xi_{k,-k}^T] Q_k \mathbb{E}_k[\xi_{k,-k}] \tag{299}$$

Using (299), $\Gamma_r$ can be further written as:

$$\Gamma_r = \mathbb{E}_k\left[e_r^T Q_k \xi_{(k,-k)} \left(\xi_{(k,-k)}^T - \mathbb{E}_k[\xi_{(k,-k)}^T]\right) Q_k \left(\xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}]\right)\right] \\ + 2\mathbb{E}_k\left[e_r^T Q_k \xi_{(k,-k)} \mathbb{E}_k[\xi_{k,-k}^T] Q_k \left(\xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}]\right)\right] \\ + \mathbb{E}_k\left[e_r^T Q_k \xi_{(k,-k)} \mathbb{E}_k[\xi_{(k,-k)}^T] Q_k \mathbb{E}_k[\xi_{(k,-k)}]\right] \\ - \mathbb{E}_k\left[e_r^T Q_k \xi_{(k,-k)} \mathbb{E}_k[\xi_{(k,-k)}^T] Q_k \xi_{(k,-k)}\right] \tag{300}$$

Using again (299), the first term in (300) is also given by:

$$\mathbb{E}_k\left[e_r^T Q_k \xi_{(k,-k)} \left(\xi_{(k,-k)}^T - \mathbb{E}_k[\xi_{(k,-k)}^T]\right) Q_k \left(\xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}]\right)\right] \\ = \mathbb{E}_k\left[e_r^T Q_k \left(\xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}]\right)\left(\xi_{(k,-k)}^T - \mathbb{E}_k[\xi_{(k,-k)}^T]\right) Q_k \left(\xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}]\right)\right] \\ + \mathbb{E}_k\left[e_r^T Q_k \mathbb{E}_k[\xi_{(k,-k)}]\left(\xi_{(k,-k)}^T - \mathbb{E}_k[\xi_{(k,-k)}^T]\right) Q_k \left(\xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}]\right)\right] \tag{301}$$

$$= p^{\frac{3}{2}} \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} [Q_k]_{rl} \mathbb{E}_k\Big[((x_l^T x_k)^2 - \frac{1}{p} x_l^T C_{[k]} x_l)((x_m^T x_k)^2 - \frac{1}{p} x_m^T C_{[k]} x_m) \\ \times [Q_k]_{mq}((x_q^T x_k)^2 - \frac{1}{p} x_q^T C_{[k]} x_q)\Big] \\ + p^{\frac{3}{2}} \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} [Q_k]_{rl} \frac{1}{p}(x_l^T C_{[k]} x_l - \frac{1}{p}\operatorname{tr}(C_{[k]} C_{[l]})) \\ \times \mathbb{E}_k\Big[((x_m^T x_k)^2 - \frac{1}{p} x_m^T C_{[k]} x_m)[Q_k]_{mq}((x_q^T x_k)^2 - \frac{1}{p} x_q^T C_{[k]} x_q)\Big] \tag{302}$$

Now, 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 $\mathbb{E}(z_1^T A_1 z_1 - \operatorname{tr} A_1)(z_1^T A_2 z_1 - \operatorname{tr} A_2) = 2 \operatorname{tr} A_1 A_2$ where $z$ is a real standard Gaussian vector, we obtain,

$$\mathbb{E}_k\left[e_r^T Q_k \xi_{(k,-k)} \left(\xi_{(k,-k)}^T - \mathbb{E}_k[\xi_{(k,-k)}^T]\right) Q_k \left(\xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}]\right)\right] \\ = 8p^{-\frac{3}{2}} \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} [Q_k]_{rl}[Q_k]_{mq}(x_l^T C_{[k]} x_m)(x_q^T C_{[k]} x_l)(x_q^T C_{[k]} x_l) \\ + 2p^{-\frac{3}{2}} \sum_{l\neq k}\sum_{m\neq k}\sum_{q\neq k} [Q_k]_{rl}[Q_k]_{mq} \frac{1}{p}(x_l^T C_{[k]} x_l - \frac{1}{p}\operatorname{tr}(C_{[k]} C_{[l]}))(x_m^T C_{[k]} x_q)^2 \tag{303}$$

In the same way, we handle the second term in (300) to obtain:

$$2\mathbb{E}_k\left[e_r^T Q_k \xi_{(k,-k)} \mathbb{E}_k[\xi_{k,-k}^T] Q_k \left(\xi_{(k,-k)} - \mathbb{E}_k[\xi_{(k,-k)}]\right)\right]$$

$$= 4p^{-\frac{3}{2}} \sum_{l\neq k} \sum_{m\neq k} \sum_{q\neq k} [Q_k]_{rl}[Q_k]_{mq} \frac{1}{p}(x_m^T C_{[k]} x_m - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[m]}))(x_q^T C_{[k]} x_l)^2 \tag{304}$$

The third term in (300) can be simplified as:

$$\begin{aligned} &\mathbb{E}_k \left[ e_r^T Q_k \xi_{(k,-k)} \mathbb{E}_k[\xi_{(k,-k)}^T] Q_k \mathbb{E}_k[\xi_{(k,-k)}] \right] \\ &= p^{-\frac{3}{2}} \sum_{l\neq k} \sum_{m\neq k} \sum_{q\neq k} [Q_k]_{rl}[Q_k]_{mq} (x_l^T C_{[k]} x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]})) \\ &\times (x_m^T C_{[k]} x_m - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[m]}))(x_q^T C_{[k]} x_q - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[q]})) \end{aligned} \tag{305}$$

Finally, to treat the last term in (300), we use Lemma 6 to get:

$$\begin{aligned} &- \mathbb{E}_k \left[ e_r^T Q_k \xi_{(k,-k)} \mathbb{E}_k[\xi_{(k,-k)}^T] Q_k \xi_{(k,-k)} \right] \\ &= -2p^{-\frac{3}{2}} \sum_{l\neq k} \sum_{m\neq k} \sum_{q\neq k} [Q_k]_{rl}[Q_k]_{mq} (x_l^T C_{[k]} x_m)^2 (x_q^T C_{[k]} x_q - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[q]})) \\ &- p^{-\frac{3}{2}} \sum_{l\neq k} \sum_{m\neq k} \sum_{q\neq k} [Q_k]_{rl}[Q_k]_{mq} (x_l^T C_{[k]} x_l - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[l]})) \qquad (306) \\ &\times (x_m^T C_{[k]} x_m - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[m]}))(x_q^T C_{[k]} x_q - \frac{1}{p}\operatorname{tr}(C_{[k]}C_{[q]})) \qquad (307) \end{aligned}$$

Taking the sum of (303)-(307), we can see that (305) and the second term in (303) cancels out with the first term in (307), thus yielding (294). This completes the proof.

### *B.2 Proof of Proposition 2*

Using the Integration by Part formula, we decompose $\tilde{\alpha}_{r,j}$ as:

$$\tilde{\alpha}_{r,j} = \chi_1 + \chi_2$$

where

$$\begin{aligned} \chi_1 &= -\frac{2}{p} \sum_{k\neq j} \sum_{b\neq k} \mathbb{E}\left[ x_b^T C_{[k]} C_{[r]} x_k x_b^T x_k Q_{kk} Q_{bj} \right] \\ \chi_2 &= -\frac{2}{p} \sum_{k\neq j} \sum_{b\neq k} \mathbb{E}\left[ x_b^T C_{[k]} C_{[r]} x_k x_b^T x_k Q_{kj} Q_{bk} \right] \end{aligned}$$

We will prove that:

$$\chi_1 = -\frac{2}{p^2} \sum_{b=1}^{n} \sum_{k=1}^{n} \mathbb{E}[Q_{kk}]\mathbb{E}[Q_{bj}] \frac{1}{p}\operatorname{tr}((C^\circ)^4) + O_z(p^{-\frac{5}{4}}) \tag{308}$$

$$\chi_2 = O_z(p^{-\frac{5}{4}}) \tag{309}$$

which obviously leads to the desired result.

**Treatment of $\chi_1$.** Using Lemma 11, we can prove that:

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

From the Integration by Parts formula, it follows that:

$$\chi_1 = \chi_{11} + \chi_{12} + \chi_{13} + O_z(p^{-\frac{3}{2}+\epsilon}) \tag{310}$$

where $\chi_{11}, \chi_{12}$ and $\chi_{13}$ are given by:

$$\chi_{11} = -\sum_{k\neq j}\sum_{b\neq k}\frac{2}{p^2}\mathbb{E}\left[x_k^T C_{[b]}C_{[k]}C_{[r]}x_k Q_{bj}\right]\mathbb{E}Q_{kk} \tag{311}$$

$$\chi_{12} = \frac{4}{p^{\frac{3}{2}}}\sum_{k\neq j}\mathbb{E}Q_{kk}\sum_{b\neq k}\sum_{s\neq b}\mathbb{E}\left[x_s^T x_b x_s^T C_{[b]}C_{[k]}C_{[r]}x_k x_k^T x_b Q_{sj}Q_{bb}\right] \tag{312}$$

$$\chi_{13} = \frac{4}{p^{\frac{3}{2}}}\sum_{k\neq j}\mathbb{E}Q_{kk}\sum_{b\neq k}\sum_{s\neq b}\mathbb{E}\left[x_s^T x_b x_s^T C_{[b]}C_{[k]}C_{[r]}x_k x_k^T x_b Q_{bj}Q_{sb}\right] \tag{313}$$

The quantity $\chi_{11}$ can be worked out as:

$$\chi_{11} = -\frac{2}{p^2}\sum_{b=1}^{n}\mathbb{E}\Big[\big(\sum_{k=1}^{n}x_k^T C_{[b]}C_{[k]}C_{[r]}x_k - \frac{1}{p}\operatorname{tr} C_{[k]}C_{[b]}C_{[k]}C_{[r]}\big)Q_{bj}\Big]\mathbb{E}Q_{kk} \tag{314}$$

$$-\frac{2}{p^2}\sum_{b=1}^{n}\sum_{k=1}^{n}\frac{1}{p}\operatorname{tr} C_{[k]}C_{[b]}C_{[k]}C_{[r]}\mathbb{E}Q_{bj}\mathbb{E}Q_{kk} + O_z(p^{-\frac{3}{2}+\epsilon}) \tag{315}$$

$$= -\frac{2}{p^2}\sum_{b=1}^{n}\sum_{k=1}^{n}\frac{1}{p}\operatorname{tr} C_{[k]}C_{[b]}C_{[k]}C_{[r]}\mathbb{E}Q_{bj}\mathbb{E}Q_{kk} + O_z(p^{-\frac{3}{2}+\epsilon}) \tag{316}$$

$$= -\frac{2}{p^2}\sum_{b=1}^{n}\sum_{k=1}^{n}\Big(\frac{1}{p}\operatorname{tr} C_{[k]}C_{[b]}C_{[k]}C_{[r]} - \frac{1}{p}\operatorname{tr}\big((C^\circ)^4\big)\Big)\mathbb{E}Q_{bj}\mathbb{E}Q_{kk}$$

$$-\frac{2}{p^2}\sum_{b=1}^{n}\sum_{k=1}^{n}\mathbb{E}Q_{kk}\mathbb{E}Q_{bj}\frac{1}{p}\operatorname{tr}\big((C^\circ)^4\big) + O_z(p^{-\frac{3}{2}+\epsilon}) \tag{317}$$

$$= -\frac{2}{p^2}\sum_{b=1}^{n}\sum_{k=1}^{n}\mathbb{E}Q_{kk}\mathbb{E}Q_{bj}\frac{1}{p}\operatorname{tr}\big((C^\circ)^4\big) + O_z(p^{-\frac{5}{4}}) \tag{318}$$

where (316) follows from the fact $\left(\sum_{k=1}^n x_k^T C_{[b]}C_{[k]}C_{[r]}x_k - \frac{1}{p}\operatorname{tr} C_{[k]}C_{[b]}C_{[k]}C_{[r]}\right)$ is $O(1)$ and (318) from $\mathbb{E}Q_{bj} = O_z(p^{-1+\epsilon})$ for $b \neq j$ and $\frac{1}{p}\operatorname{tr} C_{[k]}C_{[b]}C_{[k]}C_{[r]} - \frac{1}{p}\operatorname{tr}\left((C^\circ)^4\right) = O(p^{-\frac{1}{4}})$ by (11). It remains thus to study the two last terms in

$\chi_1$. We start by decomposing $\chi_{13}$ as:

$$\chi_{13} = \frac{4}{p^{\frac{3}{2}}} \sum_{k\neq j} \mathbb{E}[Q_{kk}] \sum_{b\neq k} \sum_{s\notin\{b,k\}} \mathbb{E}\Big[x_s^T x_b x_s^T C_{[b]} C_{[k]} C_{[r]} x_k x_k^T x_b Q_{bj} Q_{sb}\Big] \quad (319)$$

$$+ \frac{4}{p^{\frac{3}{2}}} \sum_{k\neq j} \mathbb{E}[Q_{kk}] \sum_{b\neq k} \sum_{s\notin\{b,k\}} \mathbb{E}\Big[(x_k^T x_b)^2 x_k^T C_{[b]} C_{[k]} C_{[r]} x_k Q_{bj} Q_{kb}\Big] \quad (320)$$

Both terms can be shown to be $O_z(p^{-\frac{3}{2}+\epsilon})$ by using Cauchy-Schwartz inequality and invoking Lemma 11. We will treat only the first term as the second one can be treated in a similar fashion.

$$\Big|\mathbb{E}\Big[\frac{4}{p^{\frac{3}{2}}} \sum_{k\neq j} \mathbb{E}[Q_{kk}] \sum_{b\neq k} \sum_{s\notin\{b,k\}} \mathbb{E}\Big[x_s^T x_b x_s^T C_{[b]} C_{[k]} C_{[j]} x_k x_k^T x_b Q_{bj} Q_{sb}\Big]\Big| \quad (321)$$

$$\leq \frac{4}{p^{\frac{3}{2}}} \sum_{k\neq j} |\mathbb{E}[Q_{kk}]| \sum_{b\neq k} \sqrt{\mathbb{E}[|x_k^T x_b|^2] |Q_{bj}|^2} \sqrt{\mathbb{E}\Big| \sum_{s\notin\{b,k\}} x_s^T x_b x_s^T C_{[b]} C_{[k]} C_{[j]} x_k Q_{sb}\Big|^2} \quad (322)$$

$$= O_z(p^{-\frac{3}{2}+\epsilon}) \quad (323)$$

The treatment of $\chi_{12}$ is more difficult. First, using the same calculations as for $\chi_{13}$, we can approximate $\chi_{12}$ as:

$$\chi_{12} = \frac{4}{p^{\frac{3}{2}}} \sum_{k\neq j} \mathbb{E}[Q_{kk}] \sum_{b\neq k} \mathbb{E}[Q_{bb}] \sum_{s\neq b} \mathbb{E}\Big[x_s^T x_b x_s^T C_{[b]} C_{[k]} C_{[r]} x_k x_k^T x_b Q_{sj}\Big] + O_z(p^{-\frac{3}{2}}) \quad (324)$$

It can be readily seen that the summand corresponding to $s \in \{j,k\}$ leads to a quantity that is $O_z(p^{-\frac{3}{2}})$. Indeed, for $s = j$, we have:

$$\begin{aligned} &\frac{4}{p^{\frac{3}{2}}} \sum_{k\neq j} \mathbb{E}[Q_{kk}] \sum_{b\neq k} \mathbb{E}[Q_{bb}] \mathbb{E}\Big[x_j^T x_b x_j^T C_{[b]} C_{[k]} C_{[r]} x_k x_k^T x_b Q_{jj}\Big] \\ &= \frac{4}{p^{\frac{3}{2}}} \sum_{k\neq j} \mathbb{E}[Q_{kk}] \sum_{b\notin\{k,j\}} \mathbb{E}[Q_{bb}] \mathbb{E}[Q_{jj}] \mathbb{E}\Big[x_j^T x_b x_j^T C_{[b]} C_{[k]} C_{[r]} x_k x_k^T x_b\Big] + O_z(p^{-\frac{3}{2}+\epsilon}) \\ &= \frac{4}{p^{\frac{3}{2}}} \sum_{k\neq j} \mathbb{E}[Q_{kk}] \sum_{b\notin\{k,j\}} \mathbb{E}[Q_{bb}] \mathbb{E}[Q_{jj}] \frac{1}{p^3} \operatorname{tr}(C_{[j]} C_{[b]} C_{[k]} C_{[r]} C_{[k]} C_{[b]}) + O_z(p^{-\frac{3}{2}+\epsilon}) \\ &= O_z(p^{-\frac{3}{2}+\epsilon}) \end{aligned} \quad (325)$$

Similarly, the contribution of the summand corresponding to $s = k$ can be proven to be $O_z(p^{-\frac{3}{2}+\epsilon})$. As a matter of fact,

$$\frac{4}{p^{\frac{3}{2}}} \sum_{k\neq j} \mathbb{E}[Q_{kk}] \sum_{b\neq k} \mathbb{E}[Q_{bb}] \mathbb{E}\Big[(x_k^T x_b)^2 x_k^T C_{[b]} C_{[k]} C_{[r]} x_k Q_{kj}\Big]$$

$$= \frac{4}{p^{\frac{3}{2}}} \sum_{k\neq j} \mathbb{E}[Q_{kk}] \sum_{b\notin\{k,j\}} \mathbb{E}[Q_{bb}]\mathbb{E}\Big[(x_k^T x_b)^2 x_k^T C_{[b]} C_{[k]} C_{[r]} x_k Q_{kj}\Big] + O_z(p^{-\frac{3}{2}+\epsilon}) \tag{326}$$

which is $O_z(p^{-\frac{3}{2}+\epsilon})$ by Lemma 11. Combining (325) and (326) leads to:

$$\chi_{12} \tag{327}$$

$$= \frac{4}{p^{\frac{3}{2}}} \sum_{k\neq j} \mathbb{E}[Q_{kk}] \sum_{b\neq k} \mathbb{E}[Q_{bb}] \sum_{s\notin\{b,k,j\}} \mathbb{E}\Big[x_s^T x_b x_s^T C_{[b]} C_{[k]} C_{[r]} x_k x_k^T x_b Q_{sj}\Big] + O_z(p^{-\frac{3}{2}}) \tag{328}$$

$$= \frac{4}{p^{\frac{3}{2}}} \sum_{s=1}^{n} \sum_{k\notin\{s,j\}} \sum_{b\notin\{k,s\}} \mathbb{E}[Q_{kk}]\mathbb{E}[Q_{bb}]\mathbb{E}\Big[x_s^T x_b x_s^T C_{[b]} C_{[k]} C_{[r]} x_k x_k^T x_b Q_{sj}\Big] + O_z(p^{-\frac{3}{2}}) \tag{329}$$

Denote by $\kappa_s$ the following quantity:

$$\kappa_s = \sum_{k\notin\{s,j\}} \sum_{b\notin\{k,s\}} \mathbb{E}[Q_{kk}]\mathbb{E}[Q_{bb}] x_s^T x_b x_s^T C_{[b]} C_{[k]} C_{[r]} x_k x_k^T x_b$$

Then, $\chi_{12}$ can be given by:

$$\chi_{12} = \frac{4}{p^{\frac{3}{2}}} \sum_{s=1}^{n} \mathbb{E}[\kappa_s Q_{sj}] = \frac{4}{p^{\frac{3}{2}}} \sum_{s=1}^{n} \mathbb{E}[(\kappa_s - \mathbb{E}[\kappa_s])] Q_{sj}] + \frac{4}{p^{\frac{3}{2}}} \mathbb{E}[\kappa_s]\mathbb{E}[Q_{sj}]$$

To conclude it suffices to show that:

$$\mathbb{E}[\kappa_s] = O_z(1) \tag{330}$$

$$\mathrm{var}(\kappa_s) = O_z(p^{-1}) \tag{331}$$

Indeed if (330) and (331) are satisfied, then $\chi_{12} = O_z(p^{-\frac{3}{2}+\epsilon})$ since for $s \neq j$ $\mathbb{E}[Q_{sj}] = O_z(p^{-1+\epsilon})$ from Corollary 3 and $\mathbb{E}[|Q_{sj}|^2] = O_z(p^{-1+\epsilon})$ from Lemma 10. The proof of (330) follows by computing the expectation over the variables $x_s$, $x_k$, $x_b$ while (331) follows easily by invoking Poincaré-Nash inequality. We omit the details for the sake of brevity.

**Treatment of $\chi_2$.** Using again the Integration by Part formula, we obtain

$$\chi_2 = \chi_{21} + \chi_{22} + \chi_{23} + \chi_{24} + \chi_{25}$$

where

$$\chi_{21} = -\frac{2}{p^2} \sum_{k\neq j} \sum_{b\neq k} \mathbb{E}\left[x_b^T C_{[k]} C_{[r]} C_{[k]} x_b Q_{kj} Q_{bk}\right]$$

$$\chi_{22} = \frac{4}{p\sqrt{p}} \sum_{k\neq j} \sum_{b\neq k} \sum_{s\neq k} \mathbb{E}\left[x_s^T x_k x_s^T C_{[k]} x_b x_b^T C_{[k]} C_{[r]} x_k Q_{kk} Q_{sj} Q_{bk}\right]$$

$$\chi_{23} = \frac{4}{p\sqrt{p}} \sum_{k\neq j} \sum_{b\neq k} \sum_{s\neq k} \mathbb{E}\left[x_s^T x_k x_s^T C_{[k]} x_b x_b^T C_{[k]} C_{[r]} x_k Q_{kj} Q_{sk} Q_{bk}\right]$$

$$\chi_{24} = \frac{4}{p\sqrt{p}} \sum_{k\neq j} \sum_{b\neq k} \sum_{s\neq k} \mathbb{E}\left[x_s^T x_k x_s^T C_{[k]} x_b x_b^T C_{[k]} C_{[r]} x_k Q_{bk} Q_{sk} Q_{kj}\right]$$

$$\chi_{25} = \frac{4}{p\sqrt{p}} \sum_{k\neq j} \sum_{b\neq k} \sum_{s\neq k} \mathbb{E}\left[x_s^T x_k x_s^T C_{[k]} x_b x_b^T C_{[k]} C_{[r]} x_k Q_{kk} Q_{sb} Q_{kj}\right]$$

We can prove that $\chi_{21}$ is $O_z(p^{-\frac{5}{4}})$. Indeed, $\chi_{21}$ can be further decomposed as:

$$\chi_{21} = -\frac{2}{p^2} \sum_{k\neq j} \sum_{b\neq k} \mathbb{E}\left[\left(x_b^T C_{[k]} C_{[r]} C_{[k]} x_b - \frac{1}{p} \operatorname{tr} C_{[b]} C_{[k]} C_{[r]} C_{[k]}\right) Q_{kj} Q_{bk}\right]$$
$$- \frac{2}{p^2} \sum_{k\neq j} \sum_{b\neq k} \frac{1}{p} \operatorname{tr} C_{[b]} C_{[k]} C_{[r]} C_{[k]} \mathbb{E}\left[Q_{kj} Q_{bk}\right]$$

where the first term in the right-hand side of the above equation is $O_z(p^{-\frac{3}{2}+\epsilon})$ due to Lemma 9 while the second term is $O_z(p^{-\frac{5}{4}})$ since:

$$\frac{2}{p^2} \sum_{k\neq j} \sum_{b\neq k} \frac{1}{p} \operatorname{tr} C_{[b]} C_{[k]} C_{[r]} C_{[k]} \mathbb{E}\left[Q_{kj} Q_{bk}\right]$$
$$= \frac{2}{p^2} \frac{1}{p} \operatorname{tr}\left((C^\circ)^4\right) \sum_{k\neq j} \sum_{b\neq k} \mathbb{E}\left[Q_{kj} Q_{bk}\right] + O_z(p^{-\frac{5}{4}})$$
$$= \frac{2}{p^2} \frac{1}{p} \operatorname{tr}\left((C^\circ)^4\right) \mathbb{E}\left[\left[1^T \left(Q - \mathcal{D}(Q)\right)\left(Q - \mathcal{D}(Q)\right)\right]_j\right] + O_z(p^{-\frac{5}{4}})$$
$$= O_z(p^{-\frac{5}{4}})$$

where $\mathcal{D}(Q)$ denotes the diagonal matrix whose diagonal elements are those of $Q$. To handle $\chi_{22}$, we start by decomposing it as:

$$\chi_{22} = \frac{4}{p\sqrt{p}} \sum_{k\neq j} \sum_{s\neq k} \sum_{b\notin\{k,s\}} \mathbb{E}\left[x_s^T x_k x_s^T C_{[k]} x_b x_b^T C_{[k]} C_{[r]} x_k Q_{kk} Q_{sj} Q_{bk}\right]$$
$$+ \frac{4}{p\sqrt{p}} \sum_{k\neq j} \sum_{s\neq k} \mathbb{E}\left[x_s^T x_k x_s^T C_{[k]} x_s x_s^T C_{[k]} C_{[r]} x_k Q_{kk} Q_{sj} Q_{sk}\right]$$

Using Lemma 11, the first term in $\chi_{22}$ can be treated as follows:

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

The second term in $\chi_{22}$ 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 $\chi_{22}$, we can also prove that $\chi_{23} = O_z(p^{-2+\epsilon})$ and $\chi_{24} = O_z(p^{-2+\epsilon})$. As for $\chi_{25}$, we can easily see that the contribution of the summand associated with $s \notin \{b, k\}$ is $O_z(p^{-\frac{3}{2}+\epsilon})$, leading to:

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

where the last approximation follows from the application of Lemma 11.

## Appendix C Proofs of Theorem 5 and Theorem 6

### *C.1 Proof of Theorem 5*

Let $\lambda$ be an isolated eigenvalue of $\tilde{\Phi}$. Being an isolated eigenvalue, $\lambda$ lies outside the support $\mathcal{S}^\epsilon$ defined in Theorem 4. Then, from linear algebra results, we have:

$$\det\left(\Phi + \frac{1}{p}\begin{bmatrix} J & 1_n \end{bmatrix}\begin{bmatrix} A & a \\ a^T & \beta \end{bmatrix}\begin{bmatrix} J & 1_n \end{bmatrix}^T - \lambda I_n\right) = 0$$

or equivalently:

$$\det\left(Q_\lambda^{-1} + \frac{1}{p}\begin{bmatrix} J & 1_n \end{bmatrix}\begin{bmatrix} A & a \\ a^T & \beta \end{bmatrix}\begin{bmatrix} J & 1_n \end{bmatrix}^T\right) = 0.$$

where $Q_\lambda = (\Phi - \lambda I_n)^{-1}$. Since $\lambda$ is not in the spectrum of $\Phi$, $Q_\lambda$ is well defined with probability 1 for $n$ and $p$ sufficiently large. An isolated eigenvalue of matrix $\tilde{\Phi}$ satisfies thus:

$$\det\left(I_n + Q_\lambda \frac{1}{p}\begin{bmatrix} J & 1_n \end{bmatrix}\begin{bmatrix} A & a \\ a^T & \beta \end{bmatrix}\begin{bmatrix} J & 1_n \end{bmatrix}^T\right) = 0.$$

Using Sylvester's identity, we thus obtain:

$$\det\left(I_{c+1} + \begin{bmatrix} \frac{1}{p} J^T Q_\lambda J & \frac{1}{p} J^T Q_\lambda 1_n \\ \frac{1}{p} 1_n^T Q_\lambda J & \frac{1}{p} 1_n^T Q_\lambda 1_n \end{bmatrix}\begin{bmatrix} A & a \\ a^T & \beta \end{bmatrix}\right) = 0.$$

or equivalently:

$$\det\begin{bmatrix} I_c + \frac{1}{p} J^T Q_\lambda J A + \frac{1}{p} J^T Q_\lambda 1_n a^T & \frac{1}{p} J^T Q_\lambda J a + \beta \frac{1}{p} J^T Q_\lambda 1_n \\ \frac{1}{p} 1_n^T Q_\lambda J A + \frac{1}{p} 1_n^T Q_\lambda 1_n a^T & 1 + \frac{1}{p} 1_n^T Q_\lambda J a + \beta \frac{1}{p} 1_n^T Q_\lambda 1_n \end{bmatrix} = 0.$$

Evaluating the determinant as a block-matrix determinant, we then find that an isolated eigenvalue of matrix $\tilde{\Phi}$ should satisfy:

$$\left(1+\frac{1}{p}1_n^TQ_\lambda Ja+\beta\frac{1}{p}1_n^TQ_\lambda 1_n\right)$$
$$\times\det\Big[I_c+\frac{1}{p}J^TQ_\lambda JA+\frac{1}{p}J^TQ_\lambda 1_na^T$$
$$-\frac{\left(\frac{1}{p}J^TQ_\lambda Ja+\beta\frac{1}{p}J^TQ_\lambda 1_n\right)\left(\frac{1}{p}1_n^TQ_\lambda JA+\frac{1}{p}1_n^TQ_\lambda 1_na^T\right)}{1+\frac{1}{p}1_n^TQ_\lambda Ja+\beta\frac{1}{p}1_n^TQ_\lambda 1_n}\Big]=0. \tag{332}$$

As discussed before, from Weyl's inequalities, we know that the largest eigenvalue of $\tilde{\Phi}$ is unbounded, while the $n-1$ remaining eigenvalues are bounded and thus are located asymptotically in a compact interval of the form $[-C,C]$ where $C$ is some constant greater almost surely than $c_0\Omega+\frac{\omega^2}{\Omega}+\|\frac{1}{p}JAJ^T+\frac{1}{p}Ja1_n^T+\frac{1}{p}1_na^TJ^T\|_2+\epsilon$ where $\epsilon>0$ is a small positive real. Among these eigenvalues, we focus on isolated eigenvalues that lie in $[-C,C]\backslash\mathcal{S}_\epsilon$. For such eigenvalues, and for any small $\epsilon$, there exists a positive constant $C'$ such that:

$$\left|1+\frac{1}{p}1_n^TQ_\lambda Ja+\beta\frac{1}{p}1_n^TQ_\lambda 1_n\right|\geq\sqrt{p}\,(C'-\epsilon) \tag{333}$$

To prove the above statement, it suffices to notice that:

$$\frac{1}{\sqrt{p}}\left(1+\frac{1}{p}1_n^TQ_\lambda Ja+\beta\frac{1}{p}1_n^TQ_\lambda 1_n\right)-\frac{\beta}{\sqrt{p}}\frac{c_0m(\lambda)}{1-c_0^2\Omega^2m^2(\lambda)}\xrightarrow{\text{a.s.}}0.$$

and hence for $n$ and $p$ sufficiently large:

$$\frac{1}{\sqrt{p}}\left|1+\frac{1}{p}1_n^TQ_\lambda Ja+\beta\frac{1}{p}1_n^TQ_\lambda 1_n\right|\geq\frac{\beta}{\sqrt{p}}\frac{c_0|m(\lambda)|}{|1-c_0^2\Omega^2m^2(\lambda)|}-\epsilon$$

where $\epsilon$ can be taken as small as desired. To continue, we use the fact that function $\lambda\mapsto m(\lambda)$ is analytic on $[-C,C]\backslash\mathcal{S}_\epsilon$, hence for all $\lambda\in[-C,C]\backslash\mathcal{S}_\epsilon$, $|m(\lambda)|$ is bounded by some constant $L$. From the relation $m(\lambda)=-\frac{1}{\lambda+\omega^2c_0m^2(\lambda)}$, we thus have:

$$|m(\lambda)|\geq\frac{1}{C+\omega^2c_0L},\qquad\forall\lambda\in[-C,C]\backslash\mathcal{S}_\epsilon.$$

Moreover, $\left|\frac{1}{1-\Omega^2c_0^2m^2(z)}\right|\geq\frac{1}{1+\Omega^2c_0^2L^2}$. All this together gives:

$$\frac{\beta}{\sqrt{p}}\frac{c_0|m(\lambda)|}{|1-c_0^2\Omega^2m^2(\lambda)|}\geq\frac{\beta}{\sqrt{p}}c_0\left(C+\omega^2c_0L\right)^{-1}\left(1+\Omega^2c_0^2L^2\right)^{-1}$$

which proves (333). We thus showed that an isolated eigenvalue of $\Phi$ lying in a compact interval necessarily satisfies for sufficiently large $n$ and $p$:

$$\hat{H}(\lambda):=\det\Big[I_c+\frac{1}{p}J^TQ_\lambda JA+\frac{1}{p}J^TQ_\lambda 1_na^T$$

$$-\frac{\left(\frac{1}{p}J^TQ_\lambda Ja+\beta\frac{1}{p}J^TQ_\lambda 1_n\right)\left(\frac{1}{p}1_n^TQ_\lambda JA+\frac{1}{p}1_n^TQ_\lambda 1_na^T\right)}{1+\frac{1}{p}1_n^TQ_\lambda Ja+\beta\frac{1}{p}1_n^TQ_\lambda 1_n}\Bigg]=0 \quad (334)$$

Using Theorem 2, and exploiting the fact that the fact $\underline{c}^TA=0$ where $\underline{c}=[c_1,\ldots,c_c]^T$, we can prove that $\hat{H}(\lambda)$ converges to $H(\lambda)$ given by:

$$H(\lambda)\triangleq\det\left[I_c+c_0m(\lambda)\mathcal{T}\right].$$

Let $\rho$ be such that that $H(\rho)=0$. Then $\rho$ satisfies

$$\det\left(I_c+c_0m(\rho)\mathcal{T}\right)=0 \quad (335)$$

or equivalently $m(\rho)=-\frac{1}{c_0\nu}$ where $\nu$ is one of the $c-1$ non-zero eigenvalues of $\mathcal{T}$. Since $\lambda\mapsto m(\lambda)$ is an increasing function from $(2\sqrt{c_0}\omega,\infty)$ onto $(-1/(\sqrt{c_0}\omega),0)$ and from $(-\infty,-2\sqrt{c_0}\omega)$ onto $(0,1/(\sqrt{c_0}\omega))$, the condition for existence of $\rho$ satisfying (335) is that there exists $\nu$, a non-zero eigenvalue of $\mathcal{T}$ such that $c_0\nu>\sqrt{c_0}\omega$, in which case a spike appears at the position $\rho=c_0\nu+\frac{\omega^2}{\nu}$. Since the rank of $\mathcal{T}$ is at most $c-1$, there can be no more than $c-1$ spikes $\rho_1\ldots\rho_{c-1}$ associated with $\nu_1,\cdots,\nu_{c-1}$ non-zero-eigenvalues of $\mathcal{T}$. They are eventually located at positions $\rho_i=c_0\nu_i+\frac{\omega^2}{\nu_i}$, on condition that $\nu_i>\frac{\omega}{\sqrt{c_0}}$. Going back to (332), the largest eigenvalue of $\tilde{\Phi}$, which we denote by $\kappa$ satisfies:

$$1+\frac{1}{p}1_n^TQ_\kappa Ja+\beta\frac{1}{p}1_n^TQ_\kappa 1_n=0. \quad (336)$$

We can show that there exists $C_1$ and $C_2$ positive constants such that:

$$C_1\sqrt{p}\leq\kappa\leq C_2\sqrt{p}$$

One possible way to show that is to start off from the observation that:

$$c_0\beta-\|\Phi\|_2\leq\kappa\leq c_0\beta+\|\Phi\|_2$$

and exploit the fact that for $\epsilon>0$ chosen as small as desired, and $n$ and $p$ sufficiently large,

$$\|\Phi\|\leq\max(2\sqrt{c_0}\omega,c_0\Omega+\frac{\omega^2}{\Omega})+\epsilon.$$

Since:

$$\left|\frac{1}{p}1_n^TQ_\kappa Ja\right|\leq|\frac{1}{\sqrt{p}}1_n^T|\|\frac{1}{\sqrt{p}}Ja\|\frac{1}{\kappa}$$

we have:

$$\frac{1}{p}1_n^TQ_\kappa Ja\xrightarrow{\text{a.s.}}0,$$

On the other hand,

$$\beta\frac{1}{p}1_n^TQ_\kappa 1_n+\beta\frac{c_0}{\kappa}\xrightarrow{\text{a.s.}}0.$$

Using (336), we thus obtain:

$$\frac{\kappa}{\sqrt{p}}-\frac{\beta c_0}{\sqrt{p}}\xrightarrow{\text{a.s.}}0.$$

## C.2 Proof of Theorem 6

### C.2.1 Proof of (27)

We first recall that:

$$\alpha_a^{\rho_{i_j}} \alpha_b^{\rho_{i_j}} = \frac{1}{\sqrt{n_a}\sqrt{n_b}} j_a^T \hat{u}_{\rho_{i_j}} \hat{u}_{\rho_{i_j}}^T j_b = \frac{1}{c_0\sqrt{c_a}\sqrt{c_b}} \frac{1}{p} \left[J^T \Pi_\lambda J\right]_{a,b} \tag{337}$$

where $\Pi_\lambda = \hat{u}_{\rho_{i_j}} \hat{u}_{\rho_{i_j}}^T$. Then, for any $a, b \in \{1, \dots, c\}$, From residue calculus, we have:

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

for $n$ sufficiently large, where $\mathcal{C}_\rho$ is a complex (positively oriented and with winding number one) contour circling around $\rho_{i_j}$ only. Using (16), we can easily see that:

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

where $\overline{\Phi}$ is defined in (17) as:

$$\overline{\Phi} = \Phi + \frac{1}{p} \begin{bmatrix} J & 1_n \end{bmatrix} \begin{bmatrix} A & a \\ a^T & \beta \end{bmatrix} \begin{bmatrix} J & 1_n \end{bmatrix}^T$$

Using Woodbury matrix inverse identity, we may write:

$$\begin{aligned} &\frac{1}{p} J^T \left(\overline{\Phi} - zI_n\right)^{-1} J \\ &= \frac{1}{p} J^T Q(z) J - \left[\tfrac{1}{p} J^T Q(z) JA + \tfrac{1}{p} J^T Q(z) 1_n a^T \quad \tfrac{1}{p} J^T Q(z) Ja + \beta \tfrac{1}{p} J^T Q(z) 1_n\right] \\ &\times G(z)^{-1} \begin{bmatrix} \frac{1}{p} J^T Q(z) J \\ \frac{1}{p} 1_n^T Q(z) J \end{bmatrix} \end{aligned}$$

where

$$G(z) = \begin{bmatrix} I_c + \frac{1}{p} J^T Q(z) JA + \frac{1}{p} J^T Q(z) 1_n a^T & \frac{1}{p} J^T Q(z) Ja + \beta \frac{1}{p} J^T Q(z) 1_n \\ \frac{1}{p} 1_n^T Q(z) JA + \frac{1}{p} 1_n^T Q(z) 1_n a^T & 1 + \frac{1}{p} 1_n^T Q(z) Ja + \frac{\beta}{p} 1_n^T Q(z) 1_n \end{bmatrix}$$

Define $b(z)$, $l(z)$ and $\gamma(z)$ as:

$$\begin{aligned} b(z) &= \frac{1}{p} J^T Q(z) Ja + \beta \frac{1}{p} J^T Q(z) 1_n \\ l(z) &= \frac{1}{p} A J^T Q(z) 1_n + \frac{1}{p} 1_n^T Q(z) 1_n a \\ \gamma(z) &= 1 + \frac{1}{p} 1_n^T Q(z) Ja + \frac{\beta}{p} 1_n^T Q(z) 1_n \end{aligned}$$

Let $R(z) = \left(I_c + \frac{1}{p}J^TQ(z)JA + \frac{1}{p}J^TQ(z)1_na^T - \frac{b(z)l(z)^T}{\gamma(z)}\right)^{-1}$. Then,

$$G^{-1}(z) = \begin{bmatrix} R(z) & -\frac{R(z)b(z)}{\gamma(z)} \\ -\frac{l^T(z)R(z)}{\gamma(z)} & \gamma^{-1}(z) + \gamma^{-2}(z)l^T(z)R(z)b(z) \end{bmatrix}$$

Using these notations, we thus obtain:

$$\begin{aligned}&\frac{1}{p}J^T(\overline{\Phi} - zI_n)^{-1}J \\ &= \frac{1}{p}J^TQ(z)J - \begin{bmatrix} A_{11}(z) + A_{12}(z) & A_{21}(z) + A_{22}(z) \end{bmatrix} \begin{bmatrix} \frac{1}{p}J^TQ(z)J \\ \frac{1}{p}1_n^TQ(z)J \end{bmatrix}\end{aligned}$$

where

$$A_{11}(z) = \frac{1}{p}J^TQ(z)JAR(z) + \frac{1}{p}J^TQ(z)1_na^TR(z) \tag{338}$$

$$A_{12}(z) = -\frac{1}{p}J^TQ(z)Ja\frac{l^T(z)R(z)}{\gamma(z)} - \beta\frac{1}{p}J^TQ(z)1_n\frac{l^T(z)R(z)}{\gamma(z)} \tag{339}$$

$$A_{21}(z) = -\frac{1}{p}J^TQ(z)R(z)JA\frac{b(z)}{\gamma(z)} - \frac{1}{p}J^TQ(z)1_na^TR(z)\frac{b(z)}{\gamma(z)} \tag{340}$$

$$A_{22}(z) = \left(\gamma^{-1}(z) + \gamma^{-2}(z)l^TR(z)b(z)\right)b(z) \tag{341}$$

Now exploiting Theorem 2, the following approximations are obtained for $z \in \mathcal{C}_\rho$:

$$\begin{aligned}&\frac{1}{p}J^TQ(z)J - c_0m(z)\mathcal{D}(\underline{c}) - \frac{c_0^3\Omega^2m^2(z)}{1 - c_0^2m^2(z)\Omega^2}\underline{c}\underline{c}^T \xrightarrow{\text{a.s.}} 0, \\ &\frac{1}{p}1_n^TQ(z)J - \frac{c_0m(z)\underline{c}^T}{1 - \Omega^2c_0^2m^2(z)} \xrightarrow{\text{a.s.}} 0. \\ &\frac{1}{p}1_n^TQ(z)1_n - \frac{m(z)c_0}{1 - c_0^2\Omega^2m^2(z)} \xrightarrow{\text{a.s.}} 0,\end{aligned}$$

from which we obtain the following convergence results,

$$\begin{aligned}&A_{11}(z) - c_0m(z)\mathcal{D}(\underline{c})A\mathcal{D}(\underline{c})^{\frac{1}{2}}\left(I_c + c_0m(z)\mathcal{T}\right)^{-1}\mathcal{D}(\underline{c})^{-\frac{1}{2}} \\ &- \frac{m(z)c_0\underline{c}a^T}{1 - c_0^2\Omega^2m^2(z)}\mathcal{D}(\underline{c})^{\frac{1}{2}}\left(I_c + c_0m(z)\mathcal{T}\right)^{-1}\mathcal{D}(\underline{c})^{-\frac{1}{2}} \xrightarrow{\text{a.s.}} 0, \\ &A_{12}(z) + \frac{m(z)c_0\underline{c}a^T}{1 - \Omega^2c_0^2m^2(z)}\mathcal{D}(\underline{c})^{\frac{1}{2}}\left(I_c + c_0m(z)\mathcal{T}\right)^{-1}\mathcal{D}(\underline{c})^{-\frac{1}{2}} \xrightarrow{\text{a.s.}} 0, \\ &A_{21}(z) \xrightarrow{\text{a.s.}} 0, \\ &A_{22}(z) - \underline{c} \xrightarrow{\text{a.s.}} 0\end{aligned}$$

by using the facts that $(I_c + c_0m(z)\mathcal{T})^{-1}\sqrt{\underline{c}} = \sqrt{\underline{c}}$, $a^T\underline{c} = 0$ and $A\underline{c} = 0$. It can be proven using for instance Vitali's convergence Theorem [8, Theorem 3.11] that the convergence of all the above terms is uniform on $\mathcal{C}_\rho$. Now since

$(I_c + c_0 m(z)\mathcal{T})^{-1}\sqrt{\underline{c}} = \sqrt{\underline{c}}$, the terms involving $(I_c + c_0 m(z)\mathcal{T})^{-1}\sqrt{\underline{c}}\sqrt{\underline{c}}$ will produce zero-residue when integrated over $\mathcal{C}_\rho$. We thus obtain after calculations:

$$-\frac{1}{2\pi\imath}\oint_{\mathcal{C}_\rho}\frac{1}{p}J^T\left(\tilde{\Phi}-zI_n\right)^{-1}J \\ -\frac{1}{2\pi\imath}\oint_{\mathcal{C}_\rho}c_0 m(z)\mathcal{D}(\underline{c})^{\frac{1}{2}}\left(I_c+c_0 m(z)\mathcal{T}\right)^{-1}\mathcal{D}(\underline{c})^{\frac{1}{2}}dz \xrightarrow{\text{a.s.}} 0.$$

Assuming a multiplicity 1 for $\nu_{\rho_{i_j}}$ the eigenvalue of $\mathcal{T}$ mapped with the value $\rho_{i_j}$ for which $(I_c + c_0 m(\rho_{i_j})\mathcal{T})$ is singular. Let $V_{\rho_{i_j}} \in \mathbb{R}^{c\times 1}$ be its associated eigenvector. We finally get after residue calculus,

$$\frac{1}{p}J^T\Pi_\lambda J - c_0(1-\frac{\omega^2}{c_0\nu^2_{\rho_{i_j}}})\mathcal{D}(\underline{c})^{\frac{1}{2}}V_{\rho_{i_j}}V^T_{\rho_{i_j}}\mathcal{D}(\underline{c})^{\frac{1}{2}} \xrightarrow{\text{a.s.}} 0.$$

from which it follows:

$$\alpha_a^{\rho_{i_j}}\alpha_b^{\rho_{i_j}} \xrightarrow{\text{a.s.}} (1-\frac{\omega^2}{c_0\nu^2_{\rho_{i_j}}})\left[V_{\rho_{i_j}}V^T_{\rho_{i_j}}\right]_{a,b}$$

### *C.2.2 Proof of* (28)

To evaluate $\sigma^a_{\rho_{i_1}\rho_{i_2}}$, we start by expanding $\sigma^a_{\rho_{i_1}\rho_{i_2}}$ as:

$$\sigma^a_{\rho_{i_1}\rho_{i_2}} = \hat{u}^T_{\rho_{i_1}}\mathcal{D}(j_a)\hat{u}_{\rho_{i_2}} - \alpha_a^{\rho_{i_1}}\alpha_a^{\rho_{i_2}}$$

We will prove later that it suffices to compute $\frac{1}{p}J^T\Pi_{\lambda_1}D_a\Pi_{\lambda_2}J$ and $D_a = \mathcal{D}(j_a)$. where $\Pi_{\lambda_1} = \hat{u}_{\rho_{i_1}}\hat{u}^T_{\rho_{i_1}}$ and $\Pi_{\lambda_2} = \hat{u}_{\rho_{i_2}}\hat{u}^T_{\rho_{i_2}}$, which according to the Cauchy integral relation is given by:

$$\frac{1}{p}J^T\Pi_{\lambda_1}D_a\Pi_{\lambda_2}J \\ = \frac{1}{(2\pi\imath)^2}\oint_{\mathcal{C}_{\rho_1}}\oint_{\mathcal{C}_{\rho_2}}\frac{1}{p}J^T\left(\overline{\Phi}-z_1I_n\right)^{-1}D_a\left(\overline{\Phi}-z_2I_n\right)^{-1}Jdz_1dz_2 + o(1)$$

where $\mathcal{C}_{\rho_1}$ and $\mathcal{C}_{\rho_2}$ are complex contours circling around $\rho_{i_1}$ and $\rho_{i_2}$ respectively. Based on the same approach as before, applying Woodbury's identity on each inverse $\left(\overline{\Phi}-z_1I_n\right)^{-1}$ and noticing that the generated cross-terms will have zero residue, we obtain that almost surely:

$$\frac{1}{p}J^T\Pi_{\lambda_1}D_a\Pi_{\lambda_2}J = \frac{1}{(2\pi\imath)^2}\oint_{\mathcal{C}_{\rho_1}}\oint_{\mathcal{C}_{\rho_2}}\begin{bmatrix}A_{11}(z_1)+A_{12}(z_1) & A_{21}(z_1)+A_{22}(z_1)\end{bmatrix} \\ \times\begin{bmatrix}\frac{1}{p}J^TQ(z_1)D_aQ(z_2)J & \frac{1}{p}J^TQ(z_1)D_aQ(z_2)1_n \\ \frac{1}{p}1_n^TQ(z_1)D_aQ(z_2)J & \frac{1}{p}1_n^TQ(z_1)D_aQ(z_2)1_n\end{bmatrix}\begin{bmatrix}A_{11}(z_2)^T+A_{12}(z_2)^T \\ A_{21}(z_2)^T+A_{22}(z_2)^T\end{bmatrix}dz_1dz_2 \\ + o(1)$$

Anticipating that the terms of type $A_{21}(z)+A_{22}(z)$ will not provide any residue at the end of calculus, we will thus focus on evaluating:

$$\frac{1}{(2\pi\imath)^2}\oint_{\mathcal{C}_{\rho_1}}\oint_{\mathcal{C}_{\rho_2}}\left(A_{11}(z_1)+A_{12}(z_1)\right)\frac{1}{p}J^TQ(z_1)D_aQ(z_2)J\left(A_{11}(z_2)^T+A_{12}(z_2)^T\right)$$

Now, from Theorem 3, we have:

$$\begin{aligned}&\frac{1}{p}J^TQ(z_1)D_aQ(z_2)J-\Bigg\{c_0c_am(z_1)m(z_2)\mathcal{D}(\{\boldsymbol{\delta}_{i=a}\}_{i=1}^c)\\&+c_ac_0^2\omega^2\frac{m^2(z_1)m^2(z_2)}{(1-\omega^2c_0m(z_1)m(z_2))}\mathcal{D}(\underline{c})+q_1(z_1,z_2)\underline{c}\left(\underline{c}\right)^T\\&+q_2(z_1,z_2)\mathcal{D}(\{\boldsymbol{\delta}_{i=a}\}_{i=1}^c)1_c\left(\underline{c}\right)^T+q_3(z_1,z_2)\underline{c}1_c^T\mathcal{D}(\{\boldsymbol{\delta}_{i=a}\}_{i=1}^c)\Bigg\}\xrightarrow{\text{a.s.}}0\end{aligned}$$

where $q_j(z_1,z_2)$, $j=1,\ldots,3$, are analytic functions on $\mathcal{C}_{\rho_1}$ and $\mathcal{C}_{\rho_2}$, where here $\{\boldsymbol{\delta}_{i=a}\}_{i=1}^c$ is the $c\times 1$ vector of all zeros except 1 at position $a$. Again, using the fact that $(I_c+c_0m(z)\mathcal{T})^{-1}\sqrt{\underline{c}}=\sqrt{\underline{c}}$, we deduce that the quantities $q_j(z_1,z_2)$, $j=1,\ldots,3$ will not contribute to a residue in the final expression. We thus obtain:

$$\begin{aligned}&\frac{1}{p}J^T\Pi_{\lambda_1}D_a\Pi_{\lambda_2}J\\&-\frac{1}{(2\pi\imath)^2}\oint_{\mathcal{C}_{\rho_1}}\oint_{\mathcal{C}_{\rho_2}}c_0m(z_1)m(z_2)\mathcal{D}(\underline{c})^{\frac{1}{2}}\left(I_c+c_0m(z_1)\mathcal{T}\right)^{-1}\mathcal{D}(\{\boldsymbol{\delta}_{i=a}\}_{i=1}^c)\\&\times\left(I_c+c_0m(z_2)\mathcal{T}\right)^{-1}\mathcal{D}(\underline{c})^{\frac{1}{2}}dz_1dz_2\\&-\frac{1}{(2\pi\imath)^2}\oint_{\mathcal{C}_{\rho_1}}\oint_{\mathcal{C}_{\rho_2}}c_0^2\omega^2c_a\frac{m^2(z_1)m^2(z_2)}{(1-\omega^2c_0m(z_1)m(z_2))}\mathcal{D}(\underline{c})^{\frac{1}{2}}\left(I_c+c_0m(z_1)\mathcal{T}\right)^{-1}\\&\times\left(I_c+c_0m(z_2)\mathcal{T}\right)^{-1}\mathcal{D}(\underline{c})^{\frac{1}{2}}dz_1dz_2\xrightarrow{\text{a.s.}}0.\end{aligned}$$

Then, after residue calculus, we obtain:

$$\begin{aligned}&\frac{1}{p}J^T\Pi_{\lambda_1}D_a\Pi_{\lambda_2}J\\&-\frac{1}{c_0}\frac{m(\rho_{i_1})m(\rho_{i_2})}{m'(\rho_{i_1})m'(\rho_{i_2})\nu_{i_1}\nu_{i_2}}\mathcal{D}(\underline{c})^{\frac{1}{2}}V_{\rho_{i_1}}V_{\rho_{i_2}}^T\mathcal{D}(\{\boldsymbol{\delta}_{i=a}\}_{i=1}^c)V_{\rho_{i_2}}V_{\rho_{i_2}}^T\mathcal{D}(\underline{c})^{\frac{1}{2}}\\&-\frac{\boldsymbol{\delta}_{\rho_{i_1}=\rho_{i_2}}\omega^2c_am^2(\rho_{i_1})m^2(\rho_{i_2})\mathcal{D}(\underline{c})^{\frac{1}{2}}V_{\rho_{i_1}}V_{\rho_{i_1}}^T\mathcal{D}(\underline{c})^{\frac{1}{2}}}{(1-\omega^2c_0m(\rho_{i_1})m(\rho_{i_2}))m'(\rho_{i_1})m'(\rho_{i_2})\nu_{i_1}\nu_{i_2}}\xrightarrow{\text{a.s.}}0.\end{aligned}$$

Exploiting the fact that $m(\rho_{i_j})=-\frac{1}{c_0\nu_{i_j}}$ for $i=1,2$ and the relation $m'(z)=m^2(z)\left(1-\omega^2c_0m^2(z)\right)^{-1}$, we thus obtain:

$$\frac{1}{p}J^T\Pi_{\lambda_1}D_a\Pi_{\lambda_2}J$$

$$\xrightarrow{\text{a.s.}} c_0\left(1-\frac{\omega^2}{c_0\nu_{i_1}^2}\right)\left(1-\frac{\omega^2}{c_0\nu_{i_2}^2}\right)\mathcal{D}(\underline{c})^{\frac{1}{2}}V_{\rho_{i_1}}V_{\rho_{i_1}}^T\mathcal{D}(\{\boldsymbol{\delta}_{i=a}\}_{i=1}^c)V_{\rho_{i_2}}V_{\rho_{i_2}}^T\mathcal{D}(\underline{c})^{\frac{1}{2}}$$
$$+\frac{\omega^2 c_a}{\nu_{i_1}^2}\left(1-\frac{\omega^2}{c_0\nu_{i_1}^2}\right)\mathcal{D}(\underline{c})^{\frac{1}{2}}V_{\rho_{i_1}}V_{\rho_{i_1}}^T\mathcal{D}(\underline{c})^{\frac{1}{2}}\boldsymbol{\delta}_{\rho_{i_1}=\rho_{i_2}}$$

With the above convergence at hand, we are now ready to study the convergence of $\sigma^a_{\rho_{i_1}\rho_{i_2}}$. First, we prove that if $\rho_{i_1}\neq\rho_{i_2}$, then

$$\sigma^a_{\rho_{i_1}\rho_{i_2}}\xrightarrow{\text{a.s.}}0.$$

For that, we consider two different cases.
<u>Case 1</u>. Either $[V_{\rho_{i_1}}]_a=0$ or $[V_{\rho_{i_2}}]_a=0$. In this case,

Assume that $\left[V_{\rho_{i_1}}\right]_a$ or $\left[V_{\rho_{i_2}}\right]_a$ is zero. Then, if $\rho_{i_1}\neq\rho_{i_2}$, it is easy to see that $\frac{1}{p}J^T\Pi_{\lambda_1}D_a\Pi_{\lambda_2}J\xrightarrow{\text{a.s.}}0$ and $\frac{1}{p}J^T\Pi_{\lambda_1}J\xrightarrow{\text{a.s.}}0$. Let $d_{i_1}$ and $d_{i_2}$ be such that $\liminf\left|\frac{1}{\sqrt{n_{d_{i_1}}}}\hat{u}^T_{\rho_{i_1}}j_{d_{i_1}}\right|>0$ and $\liminf\left|\frac{1}{\sqrt{n_{d_{i_2}}}}\hat{u}^T_{\rho_{i_2}}j_{d_{i_2}}\right|>0$. Such $d_{i_1}$ and $d_{i_2}$ must exist according to the asymptotic analysis in the previous section. We can thus write:

$$\sigma^a_{\rho_{i_1}\rho_{i_2}}=\frac{\left[\frac{1}{p}J^T\Pi_{\lambda_1}D_a\Pi_{\lambda_2}J\right]_{d_{i_1}d_{i_2}}}{\frac{1}{p}j^T_{d_{i_1}}\hat{u}_{\rho_{i_1}}\hat{u}^T_{\rho_{i_2}}j_{d_{i_2}}}-\alpha_a^{\rho_{i_1}}\alpha_a^{\rho_{i_2}}\tag{342}$$

and hence under the condition that $\left[V_{\rho_{i_1}}\right]_a$ or $\left[V_{\rho_{i_2}}\right]_a$ are zero and $\rho_{i_1}\neq\rho_{i_2}$, $\sigma^a_{\rho_{i_1}\rho_{i_2}}\to 0$. Now, if $\rho_{i_1}=\rho_{i_2}$ and $\left[V_{\rho_{i_1}}\right]_a$ or $\left[V_{\rho_{i_2}}\right]_a$ are zero, then $\alpha_a^{\rho_{i_1}}\alpha_a^{\rho_{i_2}}\to 0$ infinitely often. Using (342)

$$\sigma^a_{\rho_{i_1}\rho_{i_1}}\xrightarrow{\text{a.s.}}\frac{\omega^2 c_a}{c_0\nu_{i_1}^2}$$

Now assume that $\left[V_{\rho_{i_1}}\right]_a$ and $\left[V_{\rho_{i_2}}\right]_a$ are different from zero. Then, $d_{i_1}$ and $d_{i_2}$ can be chosen equal to $a$. if $\rho_{i_1}\neq\rho_{i_2}$, then:

$$\sigma^a_{\rho_{i_1}\rho_{i_2}}=\frac{\left[\frac{1}{p}J^T\Pi_{\lambda_1}D_a\Pi_{\lambda_2}J\right]_{aa}}{\sqrt{\left[\frac{1}{p}J^T\Pi_{\lambda_1}J\right]_{aa}}\sqrt{\left[\frac{1}{p}J^T\Pi_{\lambda_2}J\right]_{aa}}\operatorname{sign}\left(\alpha_a^{\rho_{i_1}}\alpha_a^{\rho_{i_2}}\right)}-\alpha_a^{\rho_{i_1}}\alpha_a^{\rho_{i_2}}$$

where $\operatorname{sign}(x)$ returns 1 if $x>0$ and $-1$ if $x<0$. If $\rho_{i_1}\neq\rho_{i_2}$, one can easily check using the above equation that

$$\sigma^a_{\rho_{i_1}\rho_{i_2}}\xrightarrow{\text{a.s.}}0.$$

while if $\rho_{i_1}=\rho_{i_2}$, we have similarly to above:

$$\sigma^a_{\rho_{i_1}\rho_{i_1}}\xrightarrow{\text{a.s.}}\frac{\omega^2 c_a}{c_0\nu_{i_1}^2}$$

In conclusion, we thus have:

$$\sigma^a_{\rho_{i_1}\rho_{i_2}}\xrightarrow{\text{a.s.}}\frac{\omega^2 c_a}{c_0\nu_{i_1}^2}\boldsymbol{\delta}_{\rho_{i_1}=\rho_{i_2}}$$

**Acknowledgments**

The authors would like to deeply thank an anonymous reviewer for his careful reading and valuable comments, which helped us to improve the quality of the manuscript.

**References**

[1] Ali, H. T., Kammoun, A., and Couillet, R. (2018). Random matrix asymptotics of inner product kernel spectral clustering. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2441–2445.

[2] 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. MR1617051

[3] Bai, Z. D. and Silverstein, J. W. (2006). *Spectral Analysis of Large Dimensional Random Matrices*. Mathematics Monograph Series 2. Science Press, Beijing. MR2567175

[4] 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. MR2944410

[5] 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. MR2489158

[6] Cheng, X. and Singer, A. (2013). The spectrum of random inner-product kernel matrices. *Random Matrices: Theory and Applications*, 02(04):1350010. MR3149440

[7] Couillet, R. and Benyach-Georges, F. (2016). Kernel Spectral Clustering of Large Dimensional Data. *Electronic Journal of Statistics*. MR3507369

[8] Couillet, R. and Debbah, M. (2011). *Random matrix methods for wireless communications*. Cambridge University Press Cambridge, MA. MR2884783

[9] Do, Y. and Vu, V. (2013). The spectrum of random kernel matrices: Universality results for rough and varying kernels. *Random matrices: Theory and application*, 2(3). MR3109422

[10] El Karoui, N. (2010a). On information plus noise kernel random matrices. *Ann. Statist.*, 38(5):3191–3216. MR2722468

[11] El Karoui, N. (2010b). The spectrum of kernel random matrices. *Ann. Statist.*, 38(1):1–50. MR2589315

[12] Fan, Z. and Montanari, A. (2015). The spectral norm of random inner-product kernel matrices. MR3109422

[13] 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. MR2183281

[14] Hofmann, T., Scholkopf, B., and Smola, A. (2008). Kernel methods in machine learning. *The Annals of Statistics*, 36(3):1171–1220. MR2418654

[15] Jiang, T. (2015). Distributions of eigenvalues of large euclidean matrices generated from lp balls and spheres. *Linear Algebra and its Applications*, 473:14–36. Special issue on Statistics. MR3338323

[16] 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. MR3571248

[17] 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.

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

[19] Schultz, H. (2005). Non-commutative polynomials of independent Gaussian random matrices. The real and symplectic cases. . *Probability theory and related fields*. MR2117954

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