The Annals of Applied Probability
2008, Vol. 18, No. 6, 2071–2130
DOI: 10.1214/08-AAP515

A CLT FOR INFORMATION-THEORETIC STATISTICS OF GRAM RANDOM MATRICES WITH A GIVEN VARIANCE PROFILE[1]

BY WALID HACHEM, PHILIPPE LOUBATON AND JAMAL NAJIM

Ecole Nationale Supérieure des Télécommunications, Université de Marne La Vallée and CNRS

Dedicated to Leonid A. Pastur on the occasion of his 70th birthday

Consider an $N \times n$ random matrix $Y_n = (Y_{ij}^n)$ with entries given by

$$Y_{ij}^n = \frac{\sigma_{ij}(n)}{\sqrt{n}} X_{ij}^n,$$

the X_{ij}^n being centered, independent and identically distributed random variables with unit variance and $(\sigma_{ij}(n); 1 \le i \le N, 1 \le j \le n)$ being an array of numbers we shall refer to as a variance profile. In this article, we study the fluctuations of the random variable

$$\log\det(Y_n Y_n^* + \rho I_N),$$

where Y^* is the Hermitian adjoint of Y and $\rho > 0$ is an additional parameter. We prove that, when centered and properly rescaled, this random variable satisfies a central limit theorem (CLT) and has a Gaussian limit whose parameters are identified whenever N goes to infinity and $\frac{N}{n} \to c \in (0, \infty)$. A complete description of the scaling parameter is given; in particular, it is shown that an additional term appears in this parameter in the case where the fourth moment of the X_{ij}'s differs from the fourth moment of a Gaussian random variable. Such a CLT is of interest in the field of wireless communications.

1. Introduction.

The model and the statistics. Consider an $N \times n$ random matrix $Y_n = (Y_{ij}^n)$ whose entries are given by

$$Y_{ij}^n = \frac{\sigma_{ij}(n)}{\sqrt{n}} X_{ij}^n, \tag{1.1}$$

where $(\sigma_{ij}(n), 1 \le i \le N, 1 \le j \le n)$ is a uniformly bounded sequence of real numbers and the random variables X_{ij}^n are complex, centered, independent and

Received September 2007; revised September 2007.

[1]Supported in part by the Fonds National de la Science (France) via the ACI program "Nouvelles Interfaces des Mathématiques," project MALCOM no. 205.

AMS 2000 subject classifications. Primary 15A52; secondary 15A18, 60F15.

Key words and phrases. Random matrix, empirical distribution of the eigenvalues, Stieltjes transform.

identically distributed (i.i.d.) with unit variance and finite eighth moment. Consider the following linear statistics of the eigenvalues:

$$\mathcal{I}_n(\rho) = \frac{1}{N} \log\det(Y_n Y_n^* + \rho I_N) = \frac{1}{N} \sum_{i=1}^{N} \log(\lambda_i + \rho),$$

where I_N is the $N \times N$ identity matrix, $\rho > 0$ is a given parameter and the λ_i's are the eigenvalues of matrix $Y_n Y_n^*$. This functional, known as the *mutual information* for multiple antenna radio channels, is very popular in wireless communications. Understanding its fluctuations and, in particular, being able to approximate its standard deviation is of major interest for various purposes such as, for instance, the computation of the so-called *outage probability*.

Presentation of the results. The purpose of this article is to establish a central limit theorem (CLT) for $\mathcal{I}_n(\rho)$ whenever $n \to \infty$ and $\frac{N}{n} \to c \in (0, \infty)$.

The centering procedure. In the companion paper [17], it has been proven that there exists a sequence of deterministic probability measures (π_n) such that the mathematical expectation $\mathbb{E}\mathcal{I}_n(\rho)$ satisfies

$$\mathbb{E}\mathcal{I}_n(\rho) - \int \log(\lambda + \rho)\pi_n(d\lambda) \xrightarrow[n\to\infty]{} 0.$$

Moreover, $\int \log(\lambda + \rho)\pi_n(d\lambda)$ has a closed-form formula (see Section 2.3) and is easier to compute[2] than $\mathbb{E}\mathcal{I}_n$ (whose evaluation would rely on massive Monte Carlo simulations). Accordingly, in this article, we study the fluctuations of

$$\frac{1}{N} \log\det(Y_n Y_n^* + \rho I_N) - \int \log(\rho + t)\pi_n(dt)$$

and prove that this quantity, properly rescaled, converges in distribution to a Gaussian random variable. Although phrased differently, such a centering procedure relying on a deterministic equivalent is used in [1] and [3].

In order to prove the CLT, we separately study the quantity $N(\mathcal{I}_n(\rho) - \mathbb{E}\mathcal{I}_n(\rho))$, from which the fluctuations arise, and the quantity $N(\mathbb{E}\mathcal{I}_n(\rho) - \int \log(\lambda + \rho)\pi_n(d\lambda))$, which yields a bias.

The fluctuations. We shall prove in this paper that the variance Θ_n^2 of $N(\mathcal{I}_n(\rho) - \mathbb{E}\mathcal{I}_n(\rho))$ takes a remarkably simple closed-form expression. In fact, there exists an $n \times n$ deterministic matrix A_n (described in Theorem 3.1) whose entries depend on the variance profile (σ_{ij}) such that the variance takes the form

$$\Theta_n^2 = -\log\det(I_n - A_n) + \kappa \operatorname{Tr} A_n,$$

[2]This is especially so in the important case where the variance profile is separable, that is, where $\sigma_{ij}^2(n)$ can be written as $\sigma_{ij}^2(n) = d_i(n)\tilde{d}_j(n)$.

where $\kappa = \mathbb{E}|X_{11}|^4 - 2$ is the fourth cumulant of the complex variable X_{11} and the CLT can be expressed as follows:

$$\frac{N}{\Theta_n}(\mathcal{I}_n - \mathbb{E}\mathcal{I}_n) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

The presence in the variance of a term directly dependent on the cumulant of the variable X_{11} ($\kappa = \mathbb{E}X_{11}^4 - 3\mathbb{E}X_{11}^2$ if X_{11} is real; $\kappa = \mathbb{E}|X_{11}|^4 - 2\mathbb{E}|X_{11}|^2$ if X_{11} is complex) can be traced back to the article by Khorunzhy, Khoruzhenko and Pastur [21], formula (I.15), and also appears in the recent paper by Anderson and Zeitouni [1]. In the case where $\kappa = 0$ (which happens if X_{ij} is, e.g., a complex Gaussian random variable), the variance has the log-form $\Theta_n^2 = \log\det(I_n - A_n)$. This has already been noticed for different models in the engineering literature by Moustakas, Simon and Sengupta [23] and Taricco [30]; see also Hachem et al. in [14].

The bias. It is proved in this paper that there exists a deterministic quantity $\mathcal{B}_n(\rho)$ (described in Theorem 3.3) such that

$$N\left(\mathbb{E}\mathcal{I}_n(\rho) - \int \log(\lambda + \rho)\pi_n(d\lambda)\right) - \mathcal{B}_n(\rho) \xrightarrow[n\to\infty]{} 0.$$

If $\kappa = 0$, then $\mathcal{B}_n(\rho) = 0$ and there is no bias in the CLT.

The literature. Central limit theorems have been widely studied for various models of random matrices and for various classes of linear statistics of the eigenvalues in the physics, engineering and mathematical literature.

In the mathematical literature, CLTs for Wigner matrices can be traced back to Girko [9] (see also [12]). Results for this class of matrices have also been obtained by Khorunzhy, Khoruzhenko and Pastur [21], Johansson [19], Sinai and Sochnikov [27], Soshnikov [29] and Cabanal-Duvillard [7]. For band matrices, let us mention the papers by Khorunzhy, Khoruzhenko and Pastur [21], Boutet de Monvel and Khorunzhy [5], Guionnet [13] and Anderson and Zeitouni [1]. The case of Gram matrices has been studied in Jonsson [20] and Bai and Silverstein [3]. Fluctuations for Wigner and Wishart matrices have also been studied by Mingo and Speicher in [22] with the help of free probability tools. For a more detailed overview, the reader is referred to the introduction in [1]. In the physics literature, so-called *replica methods*, as well as saddle-point methods, have long been a popular tool to compute the moments of the limiting distributions related to the fluctuations of the statistics of the eigenvalues.

Previous results and methods have recently been exploited in the engineering literature, with the growing interest in random matrix models for wireless communications (see the seminal paper by Telatar [31] and the subsequent papers of Tse and co-workers [32, 33]; see also the monograph by Tulino and Verdu [34] and the references therein). One main interest lies in the study of the convergence and the fluctuations of the mutual information $\frac{1}{N}\log\det(Y_nY_n^* + \rho I_N)$ for various models of matrices Y_n. General convergence results have been established by the authors in [15–17] while fluctuation results based on Bai and Silverstein [3] have been developed in Debbah and Müller [8] and Tulino and Verdu [35]. Other fluctuation results, based either on the replica method or on saddle-point analysis have been developed by Moustakas, Sengupta and coauthors [23, 24] and Taricco [30]. In a different fashion, and extensively based on the Gaussianity of the entries, a CLT has been proven in Hachem et al. [14].

Comparison with existing work. There are many overlaps between this work and other works in the literature, in particular, with the paper by Bai and Silverstein [3] and the paper by Anderson and Zeitouni [1] (although the latter is primarily devoted to band matrix models, i.e., symmetric matrices with a symmetric variance profile). The computation of the variance and the establishing of a closed-form formula significantly extend the results obtained in [14].

In this paper, we deal with complex variables, which are more relevant for wireless communication applications. The case of real random variables would have led to very similar computation, the cumulant $\kappa = \mathbb{E}|X|^4 - 2$ being replaced by $\tilde{\kappa} = \mathbb{E}X^4 - 3$. In [1], Anderson and Zeitouni deal with band matrices with real variables. Due to the complex nature of the variables herein, the standard trick of considering the symmetric matrix $\binom{0\ \ X}{X^*\ \ 0}$ to study the spectral distribution of XX^* does not help and one cannot rely on the CLT in [1]. Moreover, we substantially relax the moment assumptions concerning the entries with respect to [1], where the existence of moments of all orders is required.[3] In this paper, we shall only assume the finiteness of the eighth moment. Bai and Silverstein [3] consider the model $T_n^{1/2}X_nX_n^*T_n^{1/2}$, where the entries of X_n are i.i.d. and have Gaussian fourth moment. This assumption can be skipped in our framework, where a good understanding of the behavior of the individual diagonal entries of the resolvent $(-zI_n + Y_nY_n^*)^{-1}$ enables us to deal with non-Gaussian entries.

On the other hand, it must be noticed that we establish the CLT for the single functional $\log\det(Y_nY_n^* + \rho I_N)$ and do not provide results for a general class of functionals as in [1] and [3]. We do believe, however, that the computations performed in this article are a good starting point to address this issue.

Outline of the article.

Nonasymptotic vs. asymptotic results. As one may check in Theorems 3.1, 3.2 and 3.3, we have deliberately chosen to provide nonasymptotic (i.e., depending on n) deterministic formulas for the variance and the bias that appear in the

[3]However, provided one is willing to make strong moment and distribution assumptions and consider real, rather than complex, random variables, one can, in principle, get a CLT for $\mathcal{I}$ from [1], although the closed-form formula for the variance obtained here would still require a specific effort.

fluctuations of $\mathcal{I}_n(\rho)$. This approach has at least two advantages: nonasymptotic formulas exist for very general variance profiles $(\sigma_{ij}(n))$ and provide a natural discretization which can easily be implemented. In the case where the variance profile is the sampling of some continuous function, that is, $\sigma_{ij}(n) = \sigma(i/N, j/n)$ (we shall refer to this as the existence of a limiting variance profile), the deterministic formulas converge as n goes to infinity (see Section 4) and one must consider Fredholm determinants in order to express the results.

The general approach. The approach developed in this article is conceptually simple. The quantity $\mathcal{I}_n(\rho) - \mathbb{E}\mathcal{I}_n(\rho)$ is decomposed into a sum of martingale differences; we then systematically approximate random quantities such as quadratic forms $\mathbf{x}^T A\mathbf{x}$, where $\mathbf{x}$ is some random vector and A is some deterministic matrix, by their deterministic counterparts $\frac{1}{n}\,\mathrm{Trace}\, A$ (in the case where the entries of $\mathbf{x}$ are i.i.d. with variance $\frac{1}{n}$) as the size of the vectors and the matrices goes to infinity. A careful study of the deterministic quantities that arise, mainly based on (deterministic) matrix analysis, is carried out and yields the closed-form formula for the variance. The martingale method which is used to establish the fluctuations of $\mathcal{I}_n(\rho)$ can be traced back to Girko's REFORM (REsolvent, FORmula and Martingale) method (see [9, 12]) and is close to the one developed in [3].

Contents. In Section 2, we introduce the main notation, provide the main assumptions and recall all of the first-order results [deterministic approximation of $\mathbb{E}\mathcal{I}_n(\rho)$] needed in the expression of the CLT. In Section 3, we state the main results of the paper: definition of the variance Θ_n^2 (Theorem 3.1); asymptotic behavior (fluctuations) of $N(\mathcal{I}_n(\rho) - \mathbb{E}\mathcal{I}_n(\rho))$ (Theorem 3.2); asymptotic behavior (bias) of $N(\mathbb{E}\mathcal{I}_n(\rho) - \int \log(\rho + t)\pi_n(dt))$ (Theorem 3.3). Section 5 is devoted to the proof of Theorem 3.1, Section 6 to the proof of Theorem 3.2 and Section 7 to the proof of Theorem 3.3.

2. Notation, assumptions and first-order results.

2.1. *Notation and assumptions.* Let $N = N(n)$ be a sequence of integers such that

$$\lim_{n\to\infty} \frac{N(n)}{n} = c \in (0, \infty).$$

In the sequel, we shall consider an $N \times n$ random matrix Y_n with individual entries

$$Y_{ij}^n = \frac{\sigma_{ij}(n)}{\sqrt{n}} X_{ij}^n,$$

where X_{ij}^n are complex centered i.i.d random variables with unit variance and $(\sigma_{ij}(n); 1 \le i \le N, 1 \le j \le n)$ is a triangular array of real numbers. Denote by $\mathrm{var}(Z)$ the variance of the random variable Z. Since $\mathrm{var}(Y_{ij}^n) = \sigma_{ij}^2(n)/n$, the family $(\sigma_{ij}(n))$ will be referred to as a *variance profile*.

The main assumptions.

ASSUMPTION A1. The random variables $(X_{ij}^n; 1 \le i \le N, 1 \le j \le n, n \ge 1)$ are complex, independent and identically distributed. They satisfy

$$\mathbb{E}X_{ij}^n = \mathbb{E}(X_{ij}^n)^2 = 0, \qquad \mathbb{E}|X_{ij}^n|^2 = 1 \quad \text{and} \quad \mathbb{E}|X_{ij}^n|^8 < \infty.$$

ASSUMPTION A2. There exists a finite positive real number $\sigma_{\max}$ such that the family of real numbers $(\sigma_{ij}(n), 1 \le i \le N, 1 \le j \le n, n \ge 1)$ satisfies

$$\sup_{n \ge 1} \max_{\substack{1 \le i \le N \\ 1 \le j \le n}} |\sigma_{ij}(n)| \le \sigma_{\max}.$$

ASSUMPTION A3. There exists a real number $\sigma_{\min}^2 > 0$ such that

$$\liminf_{n \ge 1} \min_{1 \le j \le n} \frac{1}{n} \sum_{i=1}^N \sigma_{ij}^2(n) \ge \sigma_{\min}^2.$$

Sometimes we shall assume that the variance profile is obtained by sampling a function on the unit square of $\mathbb{R}^2$. This helps to obtain limiting expressions and limiting behaviors (cf. Theorem 2.5).

ASSUMPTION A4. There exists a continuous function $\sigma^2 : [0, 1] \times [0, 1] \to (0, \infty)$ such that $\sigma_{ij}^2(n) = \sigma^2(i/N, j/n)$.

Remarks related to the assumptions.

1. One may readily relax the assumption $\frac{N}{n} \to c \in (0, \infty)$ and assume instead that

$$0 < \liminf_n \frac{N}{n} \le \limsup_n \frac{N}{n} < \infty,$$

as done in [14]. We stick to the initial assumption in order to remain coherent with the companion paper [17].
2. Using truncation arguments à la Bai and Silverstein [2, 25, 26], one may lower the moment assumption related to the X_{ij}'s in Assumption A1.
3. Obviously, Assumption A3 holds if σ_{ij}^2 is uniformly lower bounded by some nonnegative quantity.
4. Obviously, Assumption A4 implies both Assumptions A2 and A3. When Assumption A4 holds, we shall say that there exists a *limiting variance profile*.
5. If necessary, Assumption A3 can be slightly improved by stating

$$\max\left(\liminf_{n \ge 1} \min_{1 \le j \le n} \frac{1}{n} \sum_{i=1}^N \sigma_{ij}^2(n), \liminf_{n \ge 1} \min_{1 \le i \le N} \frac{1}{n} \sum_{j=1}^n \sigma_{ij}^2(n) \right) > 0.$$

In the case where the first liminf is zero, one may note that $\log\det(Y_n Y_n^* + \rho I_N) = \log\det(Y_n^* Y_n + \rho I_n) + (n - N)\log\rho$ and consider $Y_n^* Y_n$ instead.

Notation. The indicator function of the set $\mathcal{A}$ will be denoted by $\mathbf{1}_{\mathcal{A}}(x)$, its cardinality by $\#\mathcal{A}$. As usual, $\mathbb{R}^+ = \{x \in \mathbb{R} : x \geq 0\}$ and $\mathbb{C}^+ = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$.

We denote by $\xrightarrow{\mathcal{P}}$ the convergence in probability of random variables and by $\xrightarrow{\mathcal{D}}$ the convergence in distribution of probability measures.

Denote by $\mathrm{diag}(a_i;\ 1 \leq i \leq k)$ the $k \times k$ diagonal matrix whose diagonal entries are the a_i's. Element (i, j) of matrix M will be denoted either m_{ij} or $[M]_{ij}$, depending on the notational context. Denote by M^T the matrix transpose of M, by M^* its Hermitian adjoint, by $\mathrm{Tr}(M)$ its trace, by $\det(M)$ its determinant (if M is square) and by F^{MM^*} the empirical distribution function of the eigenvalues of MM^*, that is,

$$F^{MM^*}(x) = \frac{1}{N}\#\{i : \lambda_i \leq x\},$$

where MM^* has dimensions $N \times N$ and the λ_i's are the eigenvalues of MM^*.

When dealing with vectors, $\|\cdot\|$ will refer to the Euclidean norm and $\|\cdot\|_\infty$, to the max (or ℓ_∞-) norm. In the case of matrices, $\|\cdot\|$ will refer to the spectral norm and $|||\cdot|||_\infty$ to the maximum row sum norm (referred to as the *max-row norm*), that is, $|||M|||_\infty = \max_{1 \leq i \leq N} \sum_{j=1}^N |[M]_{ij}|$ when M is a $N \times N$ matrix. We shall denote by $r(M)$ the spectral radius of matrix M.

When no confusion can occur, we shall often drop subscripts and superscripts n for readability.

2.2. *Stieltjes transforms and resolvents.* In this paper, Stieltjes transforms of probability measures play a fundamental role. Let ν be a bounded nonnegative measure over $\mathbb{R}$. Its Stieltjes transform f is defined as

$$f(z) = \int_{\mathbb{R}} \frac{\nu(d\lambda)}{\lambda - z}, \qquad z \in \mathbb{C} \setminus \mathrm{supp}(\nu),$$

where $\mathrm{supp}(\nu)$ is the support of the measure ν. We shall denote by $\mathcal{S}(\mathbb{R}^+)$ the set of Stieltjes transforms of probability measures with support in $\mathbb{R}^+$.

In the following proposition, we list the main properties of the Stieltjes transforms that will be needed in the paper.

PROPOSITION 2.1. *The following properties hold true.*

1. *If f is the Stieltjes transform of a probability measure ν on $\mathbb{R}$, then*:
   - *the function f is analytic over $\mathbb{C} \setminus \mathrm{supp}(\nu)$;*
   - *if $f(z) \in \mathcal{S}(\mathbb{R}^+)$, then $|f(z)| \leq (\mathbf{d}(z, \mathbb{R}^+))^{-1}$, where $\mathbf{d}(z, \mathbb{R}^+)$ denotes the distance from z to $\mathbb{R}^+$.*
2. *Let $\mathbb{P}_n$ and $\mathbb{P}$ be probability measures over $\mathbb{R}$ and denote by f_n and f their Stieltjes transforms. Then,*

$$\left(\forall z \in \mathbb{C}^+,\ f_n(z) \xrightarrow[n\to\infty]{} f(z)\right) \quad \Rightarrow \quad \mathbb{P}_n \xrightarrow[n\to\infty]{\mathcal{D}} \mathbb{P}.$$

There are very close ties between the Stieltjes transform of the empirical distribution of the eigenvalues of a matrix and the resolvent of this matrix. Let M be an $N \times n$ matrix. The resolvent of MM^* is defined as

$$Q(z) = (MM^* - zI_N)^{-1} = (q_{ij}(z))_{1 \le i,j, \le N}, \qquad z \in \mathbb{C} - \mathbb{R}^+.$$

The following properties are straightforward.

PROPOSITION 2.2. *If $Q(z)$ is the resolvent of MM^*, then*:

1. *The function $h_n(z) = \frac{1}{N} \operatorname{Tr} Q(z)$ is the Stieltjes transform of the empirical distribution of the eigenvalues of MM^* and, since the eigenvalues of this matrix are nonnegative, $h_n(z) \in \mathcal{S}(\mathbb{R}^+)$.*
2. *For every $z \in \mathbb{C} - \mathbb{R}^+$, $\|Q(z)\| \le (\mathbf{d}(z, \mathbb{R}^+))^{-1}$, and, in particular, if $\rho > 0$, then $\|Q(-\rho)\| \le \rho^{-1}$.*

2.3. *First order results*: *A primer.* Recall that $\mathcal{I}_n(\rho) = \frac{1}{N} \log\det(Y_n Y_n^* + \rho I)$ and let $\rho > 0$. Below, we shall recall some results related to the asymptotic behavior of $\mathbb{E}\mathcal{I}_n(\rho)$. As

$$\mathcal{I}_n(\rho) = \frac{1}{N} \sum_{i=1}^{N} \log(\lambda_i + \rho) = \int_0^\infty \log(\lambda + \rho)\, dF^{Y_n Y_n^*}(\lambda),$$

where the λ_i's are the eigenvalues of YY^*, the approximation of $\mathbb{E}\mathcal{I}_n(\rho)$ is closely related to the "first-order" approximation of $F^{Y_n Y_n^*}$ as $n \to \infty$ and $N/n \to c > 0$.

The following theorem summarizes the first-order results needed in the sequel. It is a direct consequence of [17], Sections 2 and 4 (see also [11]).

THEOREM 2.3 ([11, 17]). *Consider the family of random matrices $(Y_n Y_n^*)$ and assume that Assumptions A1 and A2 hold. Then, the following hold true*:

1. *The system of N functional equations*

$$t_i(z) = \frac{1}{-z + (1/n)\sum_{j=1}^n \frac{\sigma_{ij}^2(n)}{1 + (1/n)\sum_{\ell=1}^N \sigma_{\ell j}^2(n) t_\ell(z)}} \tag{2.1}$$

admits a unique solution $(t_1(z), \dots, t_N(z))$ in $\mathcal{S}(\mathbb{R}^+)^N$; in particular, $m_n(z) = \frac{1}{N}\sum_{i=1}^N t_i(z)$ belongs to $\mathcal{S}(\mathbb{R}^+)$ and there exists a probability measure π_n on $\mathbb{R}^+$ such that

$$m_n(z) = \int_0^\infty \frac{\pi_n(d\lambda)}{\lambda - z}.$$

2. *For every continuous and bounded function g on $\mathbb{R}^+$,*

$$\int_{\mathbb{R}^+} g(\lambda)\, dF^{Y_n Y_n^*}(\lambda) - \int_{\mathbb{R}^+} g(\lambda) \pi_n(d\lambda) \xrightarrow[n\to\infty]{} 0 \qquad a.e.$$

3. *The function* $V_n(\rho) = \int_{\mathbb{R}^+} \log(\lambda + \rho)\pi_n(d\lambda)$ *is finite for every* $\rho > 0$ *and*

$$\mathbb{E}\mathcal{I}_n(\rho) - V_n(\rho) \xrightarrow[n\to\infty]{} 0 \qquad \textit{where } \mathcal{I}_n(\rho) = \frac{1}{N}\log\det(Y_nY_n^* + \rho I_N);$$

moreover, $V_n(\rho)$ *admits the closed-form formula*

$$V_n(\rho) = -\frac{1}{N}\sum_{i=1}^N \log t_i(-\rho) + \frac{1}{N}\sum_{j=1}^n \log\left(1 + \frac{1}{n}\sum_{\ell=1}^N \sigma_{\ell j}^2(n) t_\ell(-\rho)\right) - \frac{1}{Nn}\sum_{i=1:N,j=1:n} \frac{\sigma_{ij}^2(n)t_i(-\rho)}{1 + (1/n)\sum_{\ell=1}^N \sigma_{\ell j}^2(n)t_\ell(-\rho)},$$

where the t_i*'s are defined above.*

Theorem 2.3 follows partly from the following lemma, which will be often invoked later on and whose statement emphasizes the symmetry between the study of $Y_nY_n^*$ and $Y_n^*Y_n$. Denote by $Q_n(z)$ and $\tilde{Q}_n(z)$ the resolvents of $Y_nY_n^*$ and $Y_n^*Y_n$, that is,

$$Q_n(z) = (Y_nY_n^* - zI_N)^{-1} = (q_{ij}(z))_{1\le i,j\le N}, \qquad z \in \mathbb{C} - \mathbb{R}^+,$$
$$\tilde{Q}_n(z) = (Y_n^*Y_n - zI_n)^{-1} = (\tilde{q}_{ij}(z))_{1\le i,j\le n}, \qquad z \in \mathbb{C} - \mathbb{R}^+.$$

LEMMA 2.4. *Consider the family of random matrices* $(Y_nY_n^*)$ *and assume that Assumptions* A1 *and* A2 *hold. Consider the following system of* $N+n$ *equations*:

$$\begin{cases} t_{i,n}(z) = \dfrac{-1}{z(1 + (1/n)\operatorname{Tr}\tilde{D}_{i,n}\tilde{T}_n(z))}, & \textit{for } 1 \le i \le N, \\ \tilde{t}_{j,n}(z) = \dfrac{-1}{z(1 + (1/n)\operatorname{Tr}D_{j,n}T_n(z))}, & \textit{for } 1 \le j \le n, \end{cases}$$

where

$$T_n(z) = \operatorname{diag}(t_{i,n}(z), 1 \le i \le N), \qquad \tilde{T}_n(z) = \operatorname{diag}(\tilde{t}_{j,n}(z), 1 \le j \le n),$$
$$D_{j,n} = \operatorname{diag}(\sigma_{ij}^2(n), 1 \le i \le N), \qquad \tilde{D}_{i,n} = \operatorname{diag}(\sigma_{ij}^2(n), 1 \le j \le n).$$

The following then hold true:

(a) ([17], *Theorem* 2.4) *this system admits a unique solution*

$$(t_{1,n}, \dots, t_{N,n}, \tilde{t}_{1,n}, \dots, \tilde{t}_{n,n}) \in \mathcal{S}(\mathbb{R}^+)^{N+n};$$

(b) ([17], *Lemmas* 6.1 *and* 6.6) *for every sequence* U_n *of* $N \times N$ *diagonal matrices and every sequence* $\tilde{U}_n$ *of* $n \times n$ *diagonal matrices such as*

$\sup_n \max(\|U_n\|, \|\tilde{U}_n\|) < \infty$, *the following limits hold true almost surely*:

$$\lim_{n\to\infty, N/n\to c} \frac{1}{N} \operatorname{Tr}\bigl(U_n\bigl(Q_n(z) - T_n(z)\bigr)\bigr) = 0 \qquad \forall z \in \mathbb{C} - \mathbb{R}^+;$$

$$\lim_{n\to\infty, N/n\to c} \frac{1}{n} \operatorname{Tr}\bigl(\tilde{U}_n\bigl(\tilde{Q}_n(z) - \tilde{T}_n(z)\bigr)\bigr) = 0 \qquad \forall z \in \mathbb{C} - \mathbb{R}^+.$$

In the case where there exists a limiting variance profile, the results can be expressed in the following manner.

THEOREM 2.5 ([6, 10, 16]). *Consider the family of random matrices* $(Y_n Y_n^*)$ *and assume that Assumptions* A1 *and* A4 *hold. Then*:

1. *The functional equation*

$$\tau(u, z) = \left(-z + \int_0^1 \frac{\sigma^2(u, v)}{1 + c\int_0^1 \sigma^2(x, v)\tau(x, z)\,dx}\,dv\right)^{-1} \tag{2.2}$$

   admits a unique solution among the class of functions $\Phi : [0, 1] \times \mathbb{C} \setminus \mathbb{R} \to \mathbb{C}$ *such that* $u \mapsto \Phi(u, z)$ *is continuous over* $[0, 1]$ *and* $z \mapsto \Phi(u, z)$ *belongs to* $\mathcal{S}(\mathbb{R}^+)$.
2. *The function* $f(z) = \int_0^1 \tau(u, z)\,du$, *where* $\tau(u, z)$ *is defined above, is the Stieltjes transform of a probability measure* $\mathbb{P}$; *moreover, we have*

$$F^{Y_n Y_n^*} \xrightarrow[n\to\infty]{\mathcal{D}} \mathbb{P} \qquad a.s.$$

REMARK 2.1. If one is interested in the Stieltjes function related to the limit of $F^{Y_n^* Y_n}$, then one must introduce the following function $\tilde{\tau}$, which is the counterpart of τ:

$$\tilde{\tau}(v, z) = \left(-z + c\int_0^1 \frac{\sigma^2(t, v)}{1 + \int_0^1 \sigma^2(t, s)\tilde{\tau}(s, z)\,ds}\,dt\right)^{-1}.$$

Functions τ and $\tilde{\tau}$ are related via the following equations:

$$\begin{cases} \tau(u, z) = -\left[z\left(1 + \int_0^1 \sigma^2(u, v)\tilde{\tau}(v, z)\,dv\right)\right]^{-1}, \\ \tilde{\tau}(v, z) = -\left[z\left(1 + c\int_0^1 \sigma^2(t, v)\tau(t, z)\,dt\right)\right]^{-1}. \end{cases} \tag{2.3}$$

REMARK 2.2. We briefly indicate here how Theorems 2.3 and 2.5 above can be deduced from Lemma 2.4. As $\frac{1}{N}\operatorname{Tr} Q_n(z)$ is the Stieltjes transform of $F^{Y_n Y_n^*}$, Theorem 2.4(b) with $U_n = I_N$ yields $\frac{1}{N}\operatorname{Tr} Q_n(z) - \frac{1}{N}\operatorname{Tr} T_n(z) \to 0$ almost surely. When a limiting variance profile exists, as described by Assumption A4, one can easily show that $\frac{1}{N}\operatorname{Tr} T_n(z)$ converges to the Stieltjes transform $f(z)$ given by Theorem 2.5 [equation (2.2) is the "continuous equivalent" of equations (2.1)]. Thanks to Proposition 2.1(2), we then obtain the almost sure weak convergence of $F^{Y_n Y_n^*}$ to F. In the case where Assumption A4 is not satisfied, one can similarly prove that $F^{Y_n Y_n^*}$ is approximated by π_n, as stated in Theorem 2.3(2).

3. The central limit theorem for $\mathcal{I}_n(\rho)$. When given a variance profile, one can consider the t_i's defined in Theorem 2.3(1). Recall that

$$T(z) = \operatorname{diag}(t_i(z), 1 \le i \le N) \quad \text{and} \quad D_j = \operatorname{diag}(\sigma_{ij}^2, 1 \le i \le N).$$

We shall first define, in Theorem 3.1, a nonnegative real number that will play the role of the variance in the CLT. We then state the CLT in Theorem 3.2. Theorem 3.3 deals with the bias term $N(\mathbb{E}\mathcal{I} - V)$.

THEOREM 3.1 (Definition of the variance). *Consider a variance profile (σ_{ij}) which fulfills Assumptions A2 and A3 and the related t_i's defined in Theorem 2.3(1). Let $\rho > 0$.*

1. *If $A_n = (a_{\ell,m})$ is the matrix defined by*

$$a_{\ell,m} = \frac{1}{n}\frac{(1/n)\operatorname{Tr} D_\ell D_m T(-\rho)^2}{(1 + (1/n)\operatorname{Tr} D_\ell T(-\rho))^2}, \qquad 1 \le \ell, m \le n,$$

then the quantity $\mathcal{V}_n = -\log\det(I_n - A_n)$ is well defined.

2. *Let $\mathcal{W}_n = \operatorname{Tr} A_n$ and let κ be a real number*[4] *satisfying $\kappa \ge -1$. The sequence $(\mathcal{V}_n + \kappa\mathcal{W}_n)$ satisfies*

$$0 < \liminf_n(\mathcal{V}_n + \kappa\mathcal{W}_n) \le \limsup_n(\mathcal{V}_n + \kappa\mathcal{W}_n) < \infty$$

as $n \to \infty$ and $N/n \to c > 0$. We shall write

$$\Theta_n^2 \stackrel{\triangle}{=} -\log\det(I - A_n) + \kappa \operatorname{Tr} A_n.$$

Proof of Theorem 3.1 is postponed to Section 5.

In the sequel, and for obvious reasons, we shall refer to matrix A_n as the *variance matrix*. In order to study the CLT for $N(\mathcal{I}_n(\rho) - V_n(\rho))$, we decompose it into a random term, from which the fluctuations arise,

$$N\big(\mathcal{I}_n(\rho) - \mathbb{E}\mathcal{I}_n(\rho)\big) = \log\det(Y_n Y_n^* + \rho I_N) - \mathbb{E}\log\det(Y_n Y_n^* + \rho I_N),$$

and a deterministic one, which yields a bias in the CLT,

$$N\big(\mathbb{E}\mathcal{I}_n(\rho) - V_n(\rho)\big) = \mathbb{E}\log\det(Y_n Y_n^* + \rho I_N) - N\int \log(\lambda + \rho)\pi_n(d\lambda).$$

We can now state the CLT.

[4] In the sequel, κ is defined as $\kappa = \mathbb{E}|X_{11}|^4 - 2$.

THEOREM 3.2 (The CLT). *Consider the family of random matrices* $(Y_n Y_n^*)$ *and assume that Assumptions* A1, A2 *and* A3 *hold true. Let* $\rho > 0$, *let* $\kappa = \mathbb{E}|X_{11}|^4 - 2$ *and let* Θ_n^2 *be given by Theorem* 3.1. *Then*,

$$\Theta_n^{-1}\big(\log\det(Y_n Y_n^* + \rho I_N) - \mathbb{E}\log\det(Y_n Y_n^* + \rho I_N)\big) \xrightarrow[n\to\infty, N/n\to c]{\mathcal{D}} \mathcal{N}(0,1).$$

Proof of Theorem 3.2 is postponed to Section 6.

The asymptotic bias is described in the following theorem.

THEOREM 3.3 (The bias). *Assume that the setting of Theorem* 3.2 *holds true. Then*:

1. *For every* $\omega \in [\rho, +\infty)$, *the system of* n *linear equations with unknown parameters* $(\mathbf{w}_{\ell,n}(\omega); 1 \le \ell \le n)$,

$$\mathbf{w}_{\ell,n}(\omega) = \frac{1}{n}\sum_{m=1}^{n} \frac{(1/n)\operatorname{Tr} D_\ell D_m T(-\omega)^2}{(1 + (1/n)\operatorname{Tr} D_\ell T(-\omega))^2}\mathbf{w}_{m,n}(\omega) + \mathbf{p}_{\ell,n}(\omega), \tag{3.1}$$

with

$$\begin{aligned}\mathbf{p}_{\ell,n}(\omega) = \kappa\omega^2 \tilde{t}_\ell(-\omega)^2 \Bigg(\frac{\omega}{n}\sum_{i=1}^{N}\bigg(\frac{\sigma_{i\ell}^2 t_i(-\omega)^3}{n}\operatorname{Tr}\tilde{D}_i^2\tilde{T}(-\omega)^2\bigg)& \\ - \frac{\tilde{t}_\ell(-\omega)}{n}\operatorname{Tr} D_\ell^2 T(-\omega)^2\Bigg)&,\end{aligned} \tag{3.2}$$

admits a unique solution for n *sufficiently large and, in particular, if* $\kappa = 0$, *then* $\mathbf{p}_{\ell,n} = 0$ *and* $\mathbf{w}_{\ell,n} = 0$.

2. *If we let*

$$\beta_n(\omega) = \frac{1}{n}\sum_{\ell=1}^{n}\mathbf{w}_{\ell,n}(\omega), \tag{3.3}$$

then $\mathcal{B}_n(\rho) \stackrel{\triangle}{=} \int_\rho^\infty \beta_n(\omega)\,d\omega$ *is well defined*; *moreover*,

$$\limsup_n \int_\rho^\infty |\beta_n(\omega)|\,d\omega < \infty; \tag{3.4}$$

furthermore,

$$N\big(\mathbb{E}\mathcal{I}_n(\rho) - V_n(\rho)\big) - \mathcal{B}_n(\rho) \xrightarrow[n\to\infty, N/n\to c]{} 0. \tag{3.5}$$

Proof of Theorem 3.3 is postponed to Section 7.

REMARK 3.1 (The Gaussian case). In the case where the entries X_{ij} are complex Gaussian (i.e., with independent normal real and imaginary parts, each of them centered with variance 2^{-1}), $\kappa = 0$ and the CLT can be written as

$$N[-\log\det(I - A_n)]^{-1/2}(\mathcal{I}_n(\rho) - V_n(\rho)) \xrightarrow[n\to\infty, N/n\to c]{\mathcal{D}} \mathcal{N}(0,1).$$

4. The CLT for a limiting variance profile. In this section, we shall assume that Assumption A4 holds, that is, $\sigma_{ij}^2(n) = \sigma^2(i/N, j/n)$ for some continuous nonnegative function $\sigma^2(x,y)$. Recall the definitions (2.2) of the function τ and of the t_i's [defined in Theorem 2.3(1)]. In the sequel, we take $\rho > 0$, $z = -\rho$ and let $\tau(t) \stackrel{\triangle}{=} \tau(t, -\rho)$. We first collect convergence results relating the t_i's and τ.

LEMMA 4.1. *Consider a variance profile (σ_{ij}) which fulfills Assumption A4. Recall the definitions of the t_i's and τ. Let $\rho > 0$ and let $z = -\rho$ be fixed. The following convergence results then hold true*:

1. $\frac{1}{N}\sum_{i=1}^N t_i \delta_{i/N} \xrightarrow[n\to\infty]{w} \tau(u)\,du$, *where* $\xrightarrow{w}$ *stands for the weak convergence of measures*;
2. $\sup_{i\le N} |t_i - \tau(i/N)| \underset{n\to\infty}{\to} 0$.
3. $\frac{1}{N}\sum_{i=1}^N t_i^2 \delta_{i/N} \xrightarrow[n\to\infty]{w} \tau^2(u)\,du$.

PROOF. The first item of the lemma follows from Lemma 2.4(b) together with Theorem 2.3(3) of [16].

In order to prove item (2), one must compute

$$t_i - \tau(i/N) = \left(\rho + \frac{1}{n}\sum_{j=1}^n \frac{\sigma^2(i/N, j/n)}{1 + (1/n)\sum_{\ell=1}^N \sigma^2(\ell/N, j/n) t_\ell}\right)^{-1} - \left(\rho + \int_0^1 \frac{\sigma^2(u,v)}{1 + c\int_0^1 \sigma^2(x,v)\tau(x)\,dx}\,dv\right)^{-1}$$

and use the convergence proved in the first part of the lemma. In order to prove the uniformity over $i \le N$, one may recall that $C[0,1]^2 = C[0,1] \otimes C[0,1]$, which, in particular implies that for all $\varepsilon > 0$, there exist g_ℓ and h_ℓ such that $\sup_{x,y} |\sigma^2(x,y) - \sum_{\ell=1}^L g_\ell(x)h_\ell(y)| \le \varepsilon$. Details are left to the reader.

The convergence result stated in item (3) is a direct consequence of item (2). □

4.1. *A continuous kernel and its Fredholm determinant.* Let $K : [0,1]^2 \to \mathbb{R}$ be some nonnegative continuous function, which we shall refer to as a kernel. Consider the associated operator (similarly denoted, with a slight abuse of notation)

$$K : C[0,1] \to C[0,1],$$

$$f \mapsto Kf(x) = \int_{[0,1]} K(x,y)f(y)\,dy.$$

One can then define (see, e.g., [28], Theorem 5.3.1) the Fredholm determinant $\det(1+\lambda K)$, where $1 : f \mapsto f$ is the identity operator, as

$$\det(1-\lambda K) = \sum_{k=0}^{\infty} \frac{(-1)^k \lambda^k}{k!} \int_{[0,1]^k} K\begin{pmatrix} x_1 & \cdots & x_k \\ x_1 & \cdots & x_k \end{pmatrix} \bigotimes_{i=1}^{k} dx_i, \tag{4.1}$$

where

$$K\begin{pmatrix} x_1 & \cdots & x_k \\ y_1 & \cdots & y_k \end{pmatrix} = \det(K(x_i, y_j), 1 \le i, j \le k)$$

for every $\lambda \in \mathbb{C}$. One can define the trace of the iterated kernel as

$$\operatorname{Tr} K^k = \int_{[0,1]^k} K(x_1, x_2) \cdots K(x_{k-1}, x_k) K(x_k, x_1)\, dx_1 \cdots dx_k.$$

In the sequel, we shall focus on the following kernel:

$$K_\infty(x,y) = \frac{c \int_{[0,1]} \sigma^2(u,x)\sigma^2(u,y)\tau^2(u)\,du}{(1 + c\int_{[0,1]} \sigma^2(u,x)\tau(u)\,du)^2}. \tag{4.2}$$

THEOREM 4.2 (The variance). *Suppose that Assumptions A1 and A4 hold. Let $\rho > 0$ and recall the definition of matrix A_n:*

$$a_{\ell,m} = \frac{1}{n} \frac{(1/n)\sum_{i=1}^N \sigma^2(i/N, \ell/n)\sigma^2(i/N, m/n)t_i^2}{(1 + (1/n)\sum_{i=1}^N \sigma^2(i/N, \ell/n)t_i)^2}, \qquad 1 \le \ell, m \le n.$$

Then:

1. $\operatorname{Tr} A_n \underset{n\to\infty}{\to} \operatorname{Tr} K_\infty$;
2. $\det(I_n - A_n) \underset{n\to\infty}{\to} \det(1 - K_\infty)$ *and* $\det(1-K_\infty) \neq 0$;
3. *if we let* $\kappa = \mathbb{E}|X_{11}|^4 - 2$, *then*

$$0 < -\log\det(1 - K_\infty) + \kappa \operatorname{Tr} K_\infty < \infty.$$

PROOF. The convergence of $\operatorname{Tr} A_n$ to $\operatorname{Tr} K_\infty$ follows from Lemma 4.1(1), (3). Details of the proof are left to the reader.

Let us introduce the following kernel:

$$K_n(x,y) = \frac{(1/n)\sum_{i=1}^N \sigma^2(i/N, x)\sigma^2(i/N, y)t_i^2}{(1 + (1/n)\sum_{i=1}^N \sigma^2(i/N, x)t_i)^2}.$$

One may note, in particular, that $a_{\ell,m} = \frac{1}{n} K_n(\frac{\ell}{n}, \frac{m}{n})$. Denote by $\|\cdot\|_\infty$ the supremum norm for a function over $[0,1]^2$ and by $\sigma_{\max}^2 = \|\sigma^2\|_\infty$. Then,

$$\|K_n\|_\infty \le \frac{N}{n}\frac{\sigma_{\max}^4}{\rho^2} \quad \text{and} \quad \|K_\infty\|_\infty \le c\frac{\sigma_{\max}^4}{\rho^2}. \tag{4.3}$$

The following facts (whose proof is omitted) can be established:

1. the family $(K_n)_{n\geq 1}$ is uniformly equicontinuous;
2. for every (x, y), $K_n(x, y) \to K_\infty(x, y)$ as $n \to \infty$.

In particular, Ascoli's theorem implies the uniform convergence of K_n to K_∞. It is now a matter of routine to extend these results and to get

$$K_n\begin{pmatrix} x_1 & \cdots & x_k \\ y_1 & \cdots & y_k \end{pmatrix} \xrightarrow[n\to\infty]{} K_\infty\begin{pmatrix} x_1 & \cdots & x_k \\ y_1 & \cdots & y_k \end{pmatrix} \tag{4.4}$$

uniformly over $[0, 1]^{2k}$. Using the uniform convergence (4.4) and a dominated convergence argument, we obtain

$$\frac{1}{n^k}\sum_{1\leq i_1,\dots,i_k\leq n} K_n\begin{pmatrix} i_1/n & \cdots & i_k/n \\ i_1/n & \cdots & i_k/n \end{pmatrix} \xrightarrow[n\to\infty]{} \int_{[0,1]^k} K_\infty\begin{pmatrix} x_1 & \cdots & x_k \\ x_1 & \cdots & x_k \end{pmatrix} \bigotimes_{i=1}^{k} dx_i.$$

Now, writing the determinant $\det(I_n + \lambda A_n)$ explicitly and expanding it as a polynomial in λ, we obtain

$$\det(I_n - \lambda A_n) = \sum_{k=0}^{n} \frac{(-1)^k\lambda^k}{k!}\left(\frac{1}{n^k}\sum_{1\leq i_1,\dots,i_k\leq n} K_n\begin{pmatrix} i_1/n & \cdots & i_k/n \\ i_1/n & \cdots & i_k/n \end{pmatrix}\right).$$

Applying Hadamard's inequality ([28], Theorem 5.2.1) to the determinants $K_n(\cdot)$ and $K_\infty(\cdot)$ yields

$$\frac{1}{n^k}\sum_{1\leq i_1,\dots,i_k\leq n} K_n\begin{pmatrix} i_1/n & \cdots & i_k/n \\ i_1/n & \cdots & i_k/n \end{pmatrix} \leq k^{k/2}\|K_n\|_\infty^k \overset{\text{(a)}}{\leq} k^{k/2}M^k,$$

where (a) follows from (4.3). Similarly,

$$\int_{[0,1]^k} K_\infty\begin{pmatrix} x_1 & \cdots & x_k \\ x_1 & \cdots & x_k \end{pmatrix} \bigotimes_{i=1}^{k} dx_i \leq k^{k/2}M^k.$$

Since the series $\sum_k \frac{M^k k^{k/2}}{k!}|\lambda|^k$ converges, a dominated convergence argument yields the convergence

$$\det(I_n + \lambda A_n) \xrightarrow[n\to\infty]{} \det(1 + \lambda K_\infty)$$

and item (2) of the theorem is proved. Item (3) follows from Theorem 3.1(2) and the proof of the theorem is complete. $\square$

4.2. *The CLT*: *Fluctuations and bias*.

COROLLARY 4.3 (Fluctuations). *Assume that Assumptions* A1 *and* A4 *hold. If we let*

$$\Theta_\infty^2 = -\log\det(1 - K_\infty) + \kappa \operatorname{Tr} K_\infty,$$

then

$$\frac{N}{\Theta_\infty}\big(\mathcal{I}_n(\rho) - \mathbb{E}\mathcal{I}_n(\rho)\big)$$
$$= \Theta_\infty^{-1}\big(\log\det(Y_n Y_n^* + \rho I_N) - \mathbb{E}\log\det(Y_n Y_n^* + \rho I_N)\big) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

PROOF. This follows easily from Theorem 3.2 and Theorem 4.2. □

Recall the definition of $\tilde{\tau}$ (cf. Remark 2.1).

THEOREM 4.4 (The bias). *Assume that the setting of Corollary 4.3 holds true. Let $\omega \in [\rho, \infty)$ and denote by $\mathbf{p} : [0,1] \to \mathbb{R}$ the quantity*

$$\mathbf{p}(x,\omega) = \kappa\omega^2\tilde{\tau}^2(x,-\omega)\left\{\omega c\int_0^1 \sigma^2(u,x)\tau^3(u)\left(\int_0^1 \sigma^2(s,u)\tilde{\tau}^2(s)\,ds\right)du \right.$$
$$\left. - \tilde{\tau}(x)c\int_0^1 \sigma^2(u,x)\tau^2(u)\,du\right\}.$$

The following functional equation admits a unique solution:

$$\mathbf{w}(x,\omega) = \int_0^1 \frac{c\int_0^1 \sigma^2(u,x)\sigma^2(u,y)\tau^2(u)\,du}{(1 + c\int_0^1 \sigma^2(u,x)\tau(u)\,du)^2}\mathbf{w}(y,\omega)\,dy + \mathbf{p}(x,\omega).$$

Let $\beta_\infty(\omega) = \int_0^1 \mathbf{w}(x,\omega)\,dx$. Then, $\int_\rho^\infty |\beta_\infty(\omega)|\,d\omega < \infty$. Moreover,

$$N\big(\mathbb{E}\mathcal{I}_n(\rho) - V_n(\rho)\big) \underset{n\to\infty, N/n\to c}{\longrightarrow} \mathcal{B}_\infty(\rho) \stackrel{\triangle}{=} \int_\rho^\infty \beta_\infty(\omega)\,d\omega. \tag{4.5}$$

The proof of Theorem 4.4, although technical, closely follows the classical Fredholm theory as presented in, for instance, [28], Chapter 5. We sketch it below.

SKETCH OF PROOF OF THEOREM 4.4. The existence and uniqueness of the functional equation follows from the fact that the Fredholm determinant $\det(1 - K_\infty)$ differs from zero. In order to prove the convergence (4.5), one may prove the convergence $\int_\rho^\infty \beta_n \to \int_\rho^\infty \beta_\infty$ (where β_n is defined in Theorem 3.3) by using an explicit representation for β_∞ relying on the explicit representation of the solution $\mathbf{w}$ via the resolvent kernel associated to K_∞ (see, e.g., [28], Section 5.4) and then approximating the resolvent kernel as done in the proof of Theorem 4.2. □

4.3. *The case of a separable variance profile.* We now state a consequence of Corollary 4.3 in the case where the variance profile is separable. Recall the definitions of τ and $\tilde{\tau}$ given in (2.3).

COROLLARY 4.5 (Separable variance profile). *Assume that Assumptions A1 and A4 hold. Assume, moreover, that $\rho > 0$ and that σ^2 is separable, that is, that*

$$\sigma^2(x, y) = d(x)\tilde{d}(y),$$

where both $d : [0, 1] \to (0, \infty)$ and $\tilde{d} : [0, 1] \to (0, \infty)$ are continuous functions. Let

$$\gamma = c \int_0^1 d^2(t)\tau^2(t)\,dt \quad \textit{and} \quad \tilde{\gamma} = \int_0^1 \tilde{d}^2(t)\tilde{\tau}^2(t)\,dt.$$

Then,

$$\Theta_\infty^2 = -\log(1 - \rho^2\gamma\tilde{\gamma}) + \kappa\rho^2\gamma\tilde{\gamma}. \tag{4.6}$$

REMARK 4.1. In the case where the random variables X_{ij} are standard complex circular Gaussian (i.e., $X_{ij} = U_{ij} + \mathbf{i}V_{ij}$ with U_{ij} and V_{ij} independent, real, centered Gaussian random variables with variance 2^{-1}) and where the variance profile is separable, we have

$$N\big(\mathcal{I}_n(\rho) - V_n(\rho)\big) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}\big(0, -\log(1 - \rho^2\gamma\tilde{\gamma})\big).$$

This result is in accordance with those in [23] and in [14].

PROOF OF COROLLARY 4.5. Recall the definitions of τ and $\tilde{\tau}$ given in (2.3). In the case where the variance profile is separable, the kernel K_∞ can be written as:

$$K_\infty(x, y) = \frac{c\tilde{d}(x)\tilde{d}(y)\int_{[0,1]} d^2(u)\tau^2(u)\,du}{(1 + c\tilde{d}(x)\int_{[0,1]} d(u)\tau(u)\,du)^2} = \rho^2\gamma\tilde{d}(x)\tilde{d}(y)\tilde{\tau}^2(x).$$

In particular, one can readily prove that $\operatorname{Tr} K_\infty = \rho^2\gamma\tilde{\gamma}$. Since the kernel $K_\infty(x, y)$ is itself a product of a function depending on x and a function depending on y, the determinant $K_\infty\binom{x_1 \ \cdots \ x_k}{y_1 \ \cdots \ y_k}$ is equal to zero for $k \geq 2$ and the Fredholm determinant can be written $\det(1 - K_\infty) = 1 - \int_{[0,1]} K_\infty(x, x)\,dx = 1 - \rho^2\gamma\tilde{\gamma}$. This yields

$$-\log\det(1 - A_\infty) + \kappa \operatorname{Tr} K_\infty = -\log(1 - \rho^2\gamma\tilde{\gamma}) + \kappa\rho^2\gamma\tilde{\gamma},$$

which concludes the proof. □

5. Proof of Theorem 3.1.

Recall the definition of the $n \times n$ variance matrix A_n:

$$a_{\ell,m} = \frac{1}{n^2}\frac{\operatorname{Tr} D_\ell D_m T(-\rho)^2}{(1 + (1/n)\operatorname{Tr} D_\ell T(-\rho))^2}, \qquad 1 \leq \ell, m \leq n.$$

In the course of the proof of the CLT (Theorem 3.2), the quantity that will naturally arise as a variance will turn out to be

$$\tilde{\Theta}_n^2 = \tilde{\mathcal{V}}_n + \kappa\mathcal{W}_n \tag{5.1}$$

(recall that $\mathcal{W}_n = \operatorname{Tr} A_n$), where $\tilde{\mathcal{V}}_n$ is introduced in the following lemma.

LEMMA 5.1. *Consider a variance profile (σ_{ij}) which fulfills Assumptions A2 and A3 and the related t_i's defined in Theorem 2.3(1). Let $\rho > 0$ and consider the matrix A_n defined above.*

1. *For $1 \le j \le n$, the system of $(n-j+1)$ linear equations with unknown parameters $(\mathbf{y}^{(\mathbf{j})}_{\boldsymbol{\ell},\mathbf{n}}, j \le \ell \le n)$,*

$$\mathbf{y}^{(\mathbf{j})}_{\boldsymbol{\ell},\mathbf{n}} = \sum_{m=j+1}^{n} a_{\ell,m} \mathbf{y}^{(\mathbf{j})}_{\mathbf{m},\mathbf{n}} + a_{\ell,j}, \tag{5.2}$$

admits a unique solution for n sufficiently large.

Denote by $\tilde{\mathcal{V}}_n$ the sum of the first components of vectors $(\mathbf{y}^{(\mathbf{j})}_{\boldsymbol{\ell},\mathbf{n}}, j \le \ell \le n)$, that is,

$$\tilde{\mathcal{V}}_n = \sum_{j=1}^{n} \mathbf{y}^{(\mathbf{j})}_{\mathbf{j},\mathbf{n}}.$$

2. *Let κ be a real number satisfying $\kappa \ge -1$. The sequence $(\tilde{\mathcal{V}}_n + \kappa \mathcal{W}_n)$ satisfies*

$$0 < \liminf_n (\tilde{\mathcal{V}}_n + \kappa \mathcal{W}_n) \le \limsup_n (\tilde{\mathcal{V}}_n + \kappa \mathcal{W}_n) < \infty$$

as $n \to \infty$ and $N/n \to c > 0$.

3. *The following holds true*:

$$\tilde{\mathcal{V}}_n + \log\det(I_n - A_n) \underset{n\to\infty}{\longrightarrow} 0.$$

Obviously, Theorem 3.1 is a by-product of Lemma 5.1. The remainder of the section is devoted to the proof of this lemma.

We cast the linear system (5.2) into a matrix framework and denote by $A_n^{(j)}$ the $(n-j+1)\times(n-j+1)$ submatrix $A_n^{(j)} = (a_{\ell,m})_{\ell,m=j}^n$, and by $A_n^{0,(j)}$ the $(n-j+1)\times(n-j+1)$ matrix $A_n^{(j)}$, where the first column is replaced by zeros. Denote by $\mathbf{d}^{(\mathbf{j})}_{\mathbf{n}}$ the $(n-j+1)\times 1$ vector

$$\mathbf{d}^{(\mathbf{j})}_{\mathbf{n}} = \left(\frac{1}{n}\frac{(1/n)\operatorname{Tr} D_\ell D_j T(-\rho)^2}{(1+(1/n)\operatorname{Tr} D_\ell T(-\rho))^2}\right)_{\ell=j}^n.$$

This notation being introduced, the system can be rewritten as

$$\mathbf{y}^{(\mathbf{j})}_{\mathbf{n}} = A_n^{0,(j)} \mathbf{y}^{(\mathbf{j})}_{\mathbf{n}} + \mathbf{d}^{(\mathbf{j})}_{\mathbf{n}} \quad \Leftrightarrow \quad (I - A_n^{0,(j)})\mathbf{y}^{(\mathbf{j})}_{\mathbf{n}} = \mathbf{d}^{(\mathbf{j})}_{\mathbf{n}}. \tag{5.3}$$

The key issue that appears is to study the invertibility of matrix $(I - A_n^{0,(j)})$ and one should get some bounds on its inverse.

5.1. *Results related to matrices with nonnegative entries.* The purpose of the next lemma is to state some of the properties of matrices with nonnegative entries that will appear to be satisfied by matrices $A_n^{0,(j)}$. We shall use the following notation. Assume that M is a real matrix. By $M \succ 0$ (resp. $M \succcurlyeq 0$) we mean $m_{ij} > 0$ (resp. $m_{ij} \geq 0$) for every element m_{ij} of M. We shall write $M \succ M'$ (resp. $M \succcurlyeq M'$) if $M - M' \succ 0$ (resp. $M - M' \succcurlyeq 0$). If x and y are vectors, we similarly write $x \succ 0$, $x \succcurlyeq 0$ and $x \succcurlyeq y$.

LEMMA 5.2. *Let $A = (a_{\ell,m})_{\ell,m=1}^n$ be an $n \times n$ real matrix and $u = (u_\ell, 1 \leq \ell \leq n)$, $v = (v_\ell, 1 \leq \ell \leq n)$ be two real $n \times 1$ vectors. Assume that $A \succcurlyeq 0$, $u \succ 0$ and $v \succ 0$. Assume, furthermore, that the equation*

$$u = Au + v$$

is satisfied. Then:

1. *The spectral radius $r(A)$ of A satisfies $r(A) \leq 1 - \frac{\min(v_\ell)}{\max(u_\ell)} < 1$.*
2. *The matrix $I_n - A$ is invertible and its inverse $(I_n - A)^{-1}$ satisfies*
$$(I_n - A)^{-1} \succcurlyeq 0 \quad \textit{and} \quad [(I_n - A)^{-1}]_{\ell\ell} \geq 1$$
for every $1 \leq \ell \leq n$.
3. *The max-row norm of the inverse is bounded: $|||(I_n - A)^{-1}|||_\infty \leq \frac{\max_\ell(u_\ell)}{\min_\ell(v_\ell)}$.*
4. *If we consider the $(n - j + 1) \times (n - j + 1)$ submatrix $A^{(j)} = (a_{\ell m})_{\ell,m=j}^n$ and denote by $A^{0,(j)}$ the matrix $A^{(j)}$ whenever the first column is replaced by zeros, then properties (1) and (2) are valid for $A^{0,(j)}$ and*
$$|||(I_{(n-j+1)} - A^{(j)})^{-1}|||_\infty \leq \frac{\max_{1 \leq \ell \leq n}(u_\ell)}{\min_{1 \leq \ell \leq n}(v_\ell)}.$$

PROOF. Let $\alpha = 1 - \frac{\min(v_\ell)}{\max(u_\ell)}$. Since $u \succ 0$ and $v \succ 0$, α readily satisfies $\alpha < 1$ and $\alpha u \succcurlyeq u - v = Au$ which, in turn, implies that $r(A) \leq \alpha < 1$ by [18], Corollary 8.1.29 and so (1) is proved. In order to prove (2), first note that for all $m \geq 1$, $A^m \succcurlyeq 0$. As $r(A) < 1$, the series $\sum_{m \geq 0} A^m$ converges, the matrix $I_n - A$ is invertible and $(I_n - A)^{-1} = \sum_{m \geq 0} A^m \succcurlyeq I_n \succcurlyeq 0$. This, in particular, implies that $[(I_n - A)^{-1}]_{\ell\ell} \geq 1$ for every $1 \leq \ell \leq n$ and so (2) is proved. Now, $u = (I_n - A)^{-1} v$ implies that for every $1 \leq k \leq n$,

$$u_k = \sum_{\ell=1}^n [(I_n - A)^{-1}]_{k\ell} v_\ell \geq \min(v_\ell) \sum_{\ell=1}^n [(I_n - A)^{-1}]_{k\ell},$$

hence (3).

We shall first prove (4) for matrix $A^{(j)}$, then show how $A^{0,(j)}$ inherits $A^{(j)}$'s properties. As $A \succcurlyeq 0$, one readily has $A^{(j)} \succcurlyeq 0$. In [18], matrix $A^{(j)}$ is called a *principal submatrix* of A. In particular, $r(A^{(j)}) \leq r(A)$, by [18], Corollary 8.1.20, which implies property (2) for $A^{(j)}$. Let $\tilde{A}^{(j)}$ be the matrix $A^{(j)}$ augmented with zeros to reach the size of A. The inverse $(I_{n-j+1} - A^{(j)})^{-1}$ is a principal submatrix of $(I_n - \tilde{A}^{(j)})^{-1} \succcurlyeq 0$. Therefore, $|||(I_{(n-j+1)} - A^{(j)})^{-1}|||_\infty \le |||(I_n - \tilde{A}^{(j)})^{-1}|||_\infty$. Since $A^m \succcurlyeq (\tilde{A}^{(j)})^m$ for every m, one has $\sum_{m\ge 0} A^m \succcurlyeq \sum_{m\ge 0}(\tilde{A}^{(j)})^m$; equivalently, $(I - A)^{-1} \succcurlyeq (I - \tilde{A}^{(j)})^{-1}$, which yields $|||(I - \tilde{A}^{(j)})^{-1}|||_\infty \le |||(I - A)^{-1}|||_\infty$. Finally, (4) is proved for matrix $A^{(j)}$.

We now prove (4) for $A^{0,(j)}$. By [18], Corollary 8.1.18, $r(A^{0,(j)}) \le r(A^{(j)}) < 1$, as $A^{(j)} \succcurlyeq A^{0,(j)}$. Therefore, $(I - A^{0,(j)})$ is invertible and

$$(I - A^{0,(j)})^{-1} = \sum_{k=0}^{\infty} [A^{0,(j)}]^k.$$

This yields, in particular, $(I - A^{0,(j)})^{-1} \succcurlyeq 0$ and $(I - A^{0,(j)})^{-1}_{kk} \ge 1$. Finally, as $A^{(j)} \succcurlyeq A^{0,(j)}$, one has

$$|||(I - \tilde{A}^{0,(j)})^{-1}|||_\infty \le |||(I - \tilde{A}^{(j)})^{-1}|||_\infty.$$

Item (4) is proved and so is Lemma 5.2. $\square$

5.2. *Proof of Lemma* 5.1: *Some preparation.* The following bounds will be needed.

PROPOSITION 5.3. *Let $\rho > 0$, consider a variance profile (σ_{ij}) which fulfills Assumption* A2 *and consider the related t_i's defined in Theorem* 2.3(1). *The following holds true*:

$$\frac{1}{\rho} \ge t_\ell(-\rho) \ge \frac{1}{\rho + \sigma^2_{\max}}.$$

PROOF. Recall that $t_\ell(z) \in \mathcal{S}(\mathbb{R}^+)$, by Theorem 2.3. In particular, $t_\ell(-\rho) = \int_{\mathbb{R}^+} \frac{\mu_\ell(d\lambda)}{\lambda+\rho}$ for some probability measure μ_ℓ. This yields the upper bound $t_\ell(-\rho) \le \rho^{-1}$ and the fact that $t_\ell(-\rho) \ge 0$. The lower bound now readily follows from equation (2.1). $\square$

PROPOSITION 5.4. *Let $\rho > 0$. Consider a variance profile (σ_{ij}) which fulfills Assumptions* A2 *and* A3; *consider the related t_i's defined in Theorem* 2.3(1). *Then*,

$$\liminf_{n\ge 1} \min_{1\le j\le n} \frac{1}{n} \operatorname{Tr} D_j T_n(-\rho)^2 > 0 \quad \textit{and} \quad \liminf_{n\ge 1} \min_{1\le j\le n} \frac{1}{n} \operatorname{Tr} D_j^2 T_n(-\rho)^2 > 0.$$

PROOF. Applying Proposition 5.3 yields

$$\frac{1}{N} \operatorname{Tr} D_j T(-\rho)^2 = \frac{1}{N} \sum_{i=1}^{N} \sigma^2_{ij} t^2_i(-\rho) \ge \frac{1}{(\rho + \sigma^2_{\max})^2} \frac{1}{N} \sum_{i=1}^{N} \sigma^2_{ij}, \tag{5.4}$$

which is bounded away from zero by Assumption A3. Similarly,

$$\frac{1}{N}\operatorname{Tr} D_j^2 T(-\rho)^2 \geq \frac{1}{(\rho+\sigma_{\max}^2)^2}\frac{1}{N}\sum_{i=1}^N \sigma_{ij}^4 \overset{\text{(a)}}{\geq} \frac{1}{(\rho+\sigma_{\max}^2)^2}\left(\frac{1}{N}\sum_{i=1}^N \sigma_{ij}^2\right)^2,$$

which is bounded away from zero [notice that (a) follows from the elementary inequality $(n^{-1}\sum x_i)^2 \leq n^{-1}\sum x_i^2$]. $\square$

We are now in position to study the matrix $A_n = A_n^{(1)}$.

PROPOSITION 5.5. *Let $\rho > 0$. Consider a variance profile (σ_{ij}) which fulfills Assumptions A2 and A3. Also, consider the related t_i's defined in Theorem 2.3(1) and let A_n be the variance matrix. There then exist two $n \times 1$ real vectors $u_n = (u_{\ell n}) \succ 0$ and $v_n = (v_{\ell n}) \succ 0$ such that $u_n = A_n u_n + v_n$. Moreover,*

$$\sup_n \max_{1\leq \ell\leq n}(u_{\ell n}) < \infty \quad \textit{and} \quad \liminf_n \min_{1\leq\ell\leq n}(v_{\ell n}) > 0.$$

PROOF. Let $z = -\rho + \delta\mathbf{i}$ with $\delta \in \mathbb{R} - \{0\}$. An equation involving the matrix A_n will show up by developing the expression of $\operatorname{Im}(T(z)) = (T(z) - T^*(z))/2\mathbf{i}$ and by using the expression of the $t_i(z)$'s given by Theorem 2.3(1). We first rewrite the system (2.1) as

$$T(z) = \left(-zI_N + \frac{1}{n}\sum_{m=1}^n \frac{D_m}{1+(1/n)\operatorname{Tr} D_m T}\right)^{-1}.$$

We then have

$$\begin{aligned}\operatorname{Im}(T) &= \frac{1}{2\mathbf{i}}(T - T^*) = \frac{1}{2\mathbf{i}}TT^*(T^{*-1} - T^{-1})\\ &= \frac{1}{n}\sum_{m=1}^n \frac{D_m TT^*}{|1+(1/n)\operatorname{Tr} D_m T|^2}\operatorname{Im}\left(\frac{1}{n}\operatorname{Tr} D_m T\right) + \delta TT^*.\end{aligned}$$

This yields, in particular, for any $1 \leq \ell \leq n$,

$$\begin{aligned}&\frac{1}{\delta}\operatorname{Im}\left(\frac{1}{n}\operatorname{Tr} D_\ell T\right)\\ &\quad = \frac{1}{n^2}\sum_{m=1}^n \frac{\operatorname{Tr} D_\ell D_m TT^*}{|1+(1/n)\operatorname{Tr} D_m T|^2}\frac{1}{\delta}\operatorname{Im}\left(\frac{1}{n}\operatorname{Tr} D_m T\right) + \frac{1}{n}\operatorname{Tr} D_\ell TT^*.\end{aligned} \tag{5.5}$$

Recall that for every $1 \leq i \leq N$, $t_i(z) \in \mathcal{S}(\mathbb{R}^+)$. Denote by μ_i the probability measure associated with t_i, that is, $t_i(z) = \int_{\mathbb{R}^+}\frac{\mu_i(d\lambda)}{\lambda - z}$. Then,

$$\frac{1}{\delta}\operatorname{Im}\left(\frac{1}{n}\operatorname{Tr} D_\ell T\right) = \frac{1}{n}\sum_{i=1}^N \sigma_{i\ell}^2 \int_0^\infty \frac{\mu_i(d\lambda)}{|\lambda - z|^2} \xrightarrow[\delta\to 0]{} \frac{1}{n}\sum_{i=1}^N \sigma_{i\ell}^2\int_0^\infty \frac{\mu_i(d\lambda)}{(\lambda+\rho)^2}.$$

Denote by $\tilde{u}_{\ell n}$ the right-hand side of the previous limit and let $u_{\ell n} = \frac{\tilde{u}_{\ell n}}{(1+(1/n)\operatorname{Tr} D_\ell T(-\rho))^2}$. Plugging this expression into (5.5) and letting $\delta \to 0$, we end up with the equation

$$u_n = A_n u_n + v_n,$$

where $A_n \succcurlyeq 0$ is given in the statement of the lemma and $u_n = (u_{\ell,n}; 1 \le \ell \le n)$, $v_n = (v_{\ell n}; 1 \le \ell \le n)$ are the $n \times 1$ vectors with elements

$$u_{\ell,n} = \frac{(1/n)\sum_{i=1}^N \sigma_{i\ell}^2 \int_0^\infty \mu_i(d\lambda)/(\lambda+\rho)^2}{(1+(1/n)\operatorname{Tr} D_\ell T(-\rho))^2} \quad \text{and}$$

$$v_{\ell,n} = \frac{(1/n)\operatorname{Tr} D_\ell T^2(-\rho)}{(1+(1/n)\operatorname{Tr} D_\ell T(-\rho))^2}. \tag{5.6}$$

For n large enough, the numerator of $u_{\ell,n}$ is lower than $(N\sigma_{\max}^2)/(n\rho^2)$ and its denominator is bounded away from zero (uniformly in n) due to Assumption A3 and Propositions 5.3 and 5.4. Similar arguments can be used to prove $u_n \succ 0$ and $v_n \succ 0$ and to get a uniform lower bound for $v_{\ell,n}$. This concludes the proof of Proposition 5.5. □

5.3. *Proof of Lemma* 5.1: *End of proof.*

PROOF OF LEMMA 5.1(1). Proposition 5.5 and Lemma 5.2(4) together yield that $I - A^{0,(j)}$ is invertible. Therefore, the system (5.3) admits a unique solution given by

$$\mathbf{y_n^{(j)}} = \left(I - A_n^{0,(j)}\right)^{-1} \mathbf{d_n^{(j)}}$$

and (1) is proved. □

PROOF OF LEMMA 5.1(2). Let us first prove the upper bound. Proposition 5.5 and Lemma 5.2 together yield

$$\limsup_n \max_j \left\| \left(I - A^{0,(j)}\right)^{-1} \right\|_\infty \le \limsup_{n\ge 1} \frac{\max_{1\le \ell\le n}(u_{\ell n})}{\min_{1\le\ell\le n}(v_{\ell n})} < \infty.$$

Each component of vector $\mathbf{d_n^{(j)}}$ satisfies $\mathbf{d_{\ell,n}^{(j)}} \le \frac{N\sigma_{\max}^4}{n^2\rho^2}$, that is, $\sup_{1\le j\le n} \|\mathbf{d_n^{(j)}}\|_\infty < \frac{K}{n}$. Therefore, vector $\mathbf{y_n^{(j)}}$ satisfies

$$\sup_j \|\mathbf{y_n^{(j)}}\|_\infty \le \sup_j \left\| \left(I - A_n^{0,(j)}\right)^{-1} \right\|_\infty \|\mathbf{d_n^{(j)}}\|_\infty < \frac{K}{n}.$$

Consequently,

$$0 \le \check{\mathcal{V}}_n = \sum_{j=1}^n \mathbf{y_{j,n}^{(j)}} \le \sum_{j=1}^n \|\mathbf{y_n^{(j)}}\|_\infty$$

satisfies $\limsup_n \check{\mathcal{V}}_n < \infty$. Moreover, Proposition 5.3 yields $\mathcal{W}_n \le n^{-2} \times \sum_{j=1}^n \operatorname{Tr} D_j^2 T^2 \le \sigma_{\max}^4 N(\rho^2 n)^{-1}$. In particular, $\mathcal{W}_n$ is also bounded and $\limsup_n(\check{\mathcal{V}}_n + \kappa \mathcal{W}_n) \le \limsup_n(\check{\mathcal{V}}_n + |\kappa| \mathcal{W}_n) < \infty$.

We now prove the lower bound

$$\check{\mathcal{V}}_n + \kappa \mathcal{W}_n = \sum_{j=1}^n \mathbf{y}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} + \kappa \mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} \ge \sum_{j=1}^n \mathbf{y}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} - \mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})}.$$

Recall that $\mathbf{y}_{\mathbf{n}}^{(\mathbf{j})} = (I - A^{0,(j)})^{-1} \mathbf{d}_{\mathbf{n}}^{(\mathbf{j})}$. We therefore have

$$\begin{aligned} \mathbf{y}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} - \mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} &= [\mathbf{y}_{\mathbf{n}}^{(\mathbf{j})} - \mathbf{d}_{\mathbf{n}}^{(\mathbf{j})}]_1 = [((I - A_n^{0,(j)})^{-1} - I)\mathbf{d}_{\mathbf{n}}^{(\mathbf{j})}]_1 \\ &= [(I - A_n^{0,(j)})^{-1} A_n^{0,(j)} \mathbf{d}_{\mathbf{n}}^{(\mathbf{j})}]_1. \end{aligned}$$

As $(I - A_n^{0,(j)})^{-1} \succcurlyeq I$, we have

$$\begin{aligned} \mathbf{y}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} - \mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} &\ge [A_n^{0,(j)} \mathbf{d}_{\mathbf{n}}^{(\mathbf{j})}]_1 = \sum_{\ell=j+1}^n \frac{1}{n^2} \frac{((1/n) \operatorname{Tr} D_\ell D_j T^2(-\rho))^2}{(1 + (1/n) \operatorname{Tr} D_j T(-\rho))^4} \\ &\overset{(a)}{\ge} K \sum_{\ell=j+1}^n \frac{1}{n^2} \left(\frac{1}{n} \operatorname{Tr} D_\ell D_j \right)^2, \end{aligned}$$

where (a) follows from Proposition 5.3, which is used to get both a lower bound for the numerator and an upper bound for the denominator: $(1 + \frac{1}{n} \operatorname{Tr} D_j T)^4 \le (1 + N n^{-1} \sigma_{\max}^2 \rho^{-1})^4$. Some computations remain to be carried out in order to take advantage of Assumption A3 and thereby obtain the lower bound. Recall that $\frac{1}{m} \sum_{k=1}^m x_k^2 \ge (\frac{1}{m} \sum_{k=1}^m x_k)^2$. We have

$$\begin{aligned} \sum_{j=1}^n \mathbf{y}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} - \mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} &\ge \sum_{j=1}^n \sum_{\ell=j+1}^n \frac{1}{n^2} \left(\frac{1}{n} \operatorname{Tr} D_\ell D_j \right)^2 \\ &= \frac{1}{n^2} \times \frac{n(n-1)}{2} \times \frac{2}{n(n-1)} \sum_{j<\ell} \left(\frac{1}{n} \operatorname{Tr} D_\ell D_j \right)^2 \\ &\overset{(a)}{\ge} \frac{1}{3} \left(\frac{2}{n(n-1)} \sum_{j<\ell} \frac{1}{n} \operatorname{Tr} D_\ell D_j \right)^2 \\ &\overset{(b)}{=} \frac{1}{3} \left(\frac{1}{n(n-1)} \sum_{1 \le j,\ell \le n} \frac{1}{n} \operatorname{Tr} D_\ell D_j \right)^2 + o(1) \\ &\ge \frac{1}{3} \left(\frac{1}{n^3} \sum_{i=1}^N \left(\sum_{j=1}^n \sigma_{ij}^2 \right)^2 \right)^2 + o(1) \end{aligned}$$

$$\geq \frac{1}{3}\left(\frac{N}{n^3}\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{n}\sigma_{ij}^2\right)^2\right)^2 + o(1)$$

$$\geq \frac{1}{3}\left(\frac{N}{n^3}\left(\sum_{j=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\sigma_{ij}^2\right)^2\right)^2 + o(1),$$

where (a) follows from the bound $\frac{n(n-1)}{2n^2} \geq \frac{1}{3}$, valid for n sufficiently large. The term $o(1)$ at step (b) goes to zero as $n \to \infty$ and takes into account the diagonal terms in the formula $2\sum_{j<\ell}\alpha_{j\ell} + \sum_j \alpha_{jj} = \sum_{j,\ell}\alpha_{j\ell}$. It remains now to take the $\liminf$ to obtain

$$\liminf_{n\to\infty}\left(\sum_{j=1}^{n}\mathbf{y}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} + \kappa\mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})}\right) \geq \frac{c^2\sigma_{\min}^8}{3}.$$

Item (2) is now proved. $\square$

PROOF OF LEMMA 5.1(3). We first introduce the following block-matrix notation:

$$A_n^{(j)} = \begin{pmatrix} \mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} & \bar{a}_n^{(j)} \\ \bar{\mathbf{d}}_{\mathbf{n}}^{(\mathbf{j})} & A_n^{(j+1)} \end{pmatrix} \quad \text{and} \quad A_n^{0,(j)} = \begin{pmatrix} 0 & \bar{a}_n^{(j)} \\ 0 & A_n^{(j+1)} \end{pmatrix}.$$

We can now express the inverse of

$$(I - A_n^{0,(j)}) = \begin{pmatrix} 1 & -\bar{a}_n^{(j)} \\ 0 & (I - A_n^{(j+1)}) \end{pmatrix}$$

$$\text{as } (I - A_n^{0,(j)})^{-1} = \begin{pmatrix} 1 & \bar{a}_n^{(j)}(I - A_n^{(j+1)})^{-1} \\ 0 & (I - A_n^{(j+1)})^{-1} \end{pmatrix}.$$

This, in turn, yields $\mathbf{y}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} = \mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} + \bar{a}_n^{(j)}(I - A_n^{(J+1)})^{-1}\bar{\mathbf{d}}_{\mathbf{n}}^{(\mathbf{j})}$ and one can easily check that $\mathbf{y}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} \leq \frac{K}{n}$, where K does not depend on j and n, as

$$|\mathbf{y}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})}| \leq |\mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})}| + n\|\bar{a}_n^{(j)}\|_\infty |||(I - A_n^{0,(j)})^{-1}|||_\infty \|\bar{\mathbf{d}}_{\mathbf{n}}^{(\mathbf{j})}\|_\infty.$$

Note that

$$\begin{aligned}
&\log\det(I - A_n^{(j)}) - \log\det(I - A_n^{(j+1)}) \\
&\quad = \log\det\left(\left[\begin{pmatrix} 1 - \mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} & -\bar{a}_n^{(j)} \\ -\bar{\mathbf{d}}_{\mathbf{n}}^{(\mathbf{j})} & I - A_n^{(j+1)} \end{pmatrix}\right]\left[\begin{pmatrix} 1 & 0 \\ 0 & (I - A_n^{(j+1)})^{-1} \end{pmatrix}\right]\right) \\
&\quad = \log\det\begin{bmatrix} 1 - \mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} & -\bar{a}_n^{(j)}(I - A_n^{(j+1)})^{-1} \\ -\bar{\mathbf{d}}_{\mathbf{n}}^{(\mathbf{j})} & I \end{bmatrix} \\
&\quad = \log(1 - \mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} - \bar{a}_n^{(j)}(I - A_n^{(j+1)})^{-1}\bar{\mathbf{d}}_{\mathbf{n}}^{(\mathbf{j})})
\end{aligned}$$

and write $\log\det(I - A_n)$ as

$$\begin{aligned}
\log\det(I - A_n) &= \sum_{j=1}^{n-1}\bigl(\log\det\bigl(I - A_n^{(j)}\bigr) - \log\det\bigl(I - A_n^{(j+1)}\bigr)\bigr) + \log(1 - a_{nn}) \\
&= \sum_{j=1}^{n-1} \log\bigl(1 - \mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} - \bar{a}_n^{(j)}\bigl(I - A_n^{(j+1)}\bigr)^{-1}\bar{\mathbf{d}}_{\mathbf{n}}^{(\mathbf{j})}\bigr) + \log(1 - a_{nn}) \\
&= -\sum_{j=1}^{n-1}\bigl(\mathbf{d}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} + \bar{a}_n^{(j)}\bigl(I - A_n^{(j+1)}\bigr)^{-1}\bar{\mathbf{d}}_{\mathbf{n}}^{(\mathbf{j})}\bigr) + o(1) \\
&= -\sum_{j=1}^{n-1} \mathbf{y}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} + o(1) = -\sum_{j=1}^{n} \mathbf{y}_{\mathbf{j},\mathbf{n}}^{(\mathbf{j})} + o(1) \\
&= -\check{\mathcal{V}}_n + o(1).
\end{aligned}$$

This concludes the proof of Lemma 5.1. □

6. Proof of Theorem 3.2.

6.1. *More notation*; *outline of the proof*; *key lemmas*.

More notation. Recall that $Y_n = (Y_{ij}^n)$ is an $N \times n$ matrix with $Y_{ij}^n = \frac{\sigma_{ij}}{\sqrt{n}} X_{ij}$ and that $Q_n(z) = (q_{ij}(z)) = (Y_n Y_n^* - zI_N)^{-1}$. We denote:

1. by $\tilde{Q}_n(z)$ the matrix $(\tilde{q}_{ij}(z)) = (Y_n^* Y_n - zI_n)^{-1}$;
2. by y_j the column number j of Y_n;
3. by Y_n^j the $N \times (n-1)$ matrix that remains after deleting column number j from Y_n;
4. by $Q_{j,n}(z)$ [or $Q_j(z)$ for short when there is no confusion with $Q_n(z)$] the $N \times N$ matrix
$$Q_j(z) = (Y^j Y^{j*} - zI_N)^{-1};$$
5. by ξ_i the row number i of Y_n;
6. by $Y_{i,n}$ (or Y_i for short when there is no confusion with Y_n) the $(N-1) \times n$ matrix that remains after deleting row i from Y;
7. by $\tilde{Q}_{i,n}(z)$ [or $\tilde{Q}_i(z)$] the $n \times n$ matrix
$$\tilde{Q}_i(z) = (Y_i^* Y_i - zI_n)^{-1}.$$

Recall that we use either q_{ij} or $[Q]_{ij}$ for the individual element of $Q(z)$, depending on the context (the same is true for other matrices). The following formulas are well known (see, e,g., Sections 0.7.3 and 0.7.4 in [18]):

$$Q = Q_j - \frac{Q_j y_j y_j^* Q_j}{1 + y_j^* Q_j y_j}, \qquad \tilde{Q} = \tilde{Q}_i - \frac{\tilde{Q}_i \xi_i^* \xi_i \tilde{Q}_i}{1 + \xi_i \tilde{Q}_i \xi_i^*}, \tag{6.1}$$

$$q_{ii}(z) = \frac{-1}{z(1 + \xi_i \tilde{Q}_i(z) \xi_i^*)}, \qquad \tilde{q}_{jj}(z) = \frac{-1}{z(1 + y_j^* Q_j(z) y_j)}. \tag{6.2}$$

For $1 \le j \le n$, denote by $\mathcal{F}_j$ the σ-field $\mathcal{F}_j = \sigma(y_j, \dots, y_n)$ generated by the random vectors $(y_j, \dots, y_n)$. Denote by $\mathbb{E}_j$ the conditional expectation with respect to $\mathcal{F}_j$, that is, $\mathbb{E}_j = \mathbb{E}(\cdot \mid \mathcal{F}_j)$. By convention, $\mathcal{F}_{n+1}$ is the trivial σ-field; in particular, $\mathbb{E}_{n+1} = \mathbb{E}$.

Outline of the proof. In order to prove the convergence of $\Theta_n^{-1}(\log\det(Y_n Y_n^* + \rho I_N) - \mathbb{E}\log\det(Y_n Y_n^* + \rho I_N))$ to the standard Gaussian law $\mathcal{N}(0, 1)$, we shall rely on the following CLT for martingales.

THEOREM 6.1 (CLT for martingales, Theorem 35.12 in [4]). *Let $\gamma_n^{(n)}, \gamma_{n-1}^{(n)}, \dots, \gamma_1^{(n)}$ be a martingale difference sequence with respect to the increasing filtration $\mathcal{F}_n^{(n)}, \dots, \mathcal{F}_1^{(n)}$. Assume that there exists a sequence of positive real numbers Θ_n^2 such that*

$$\frac{1}{\Theta_n^2} \sum_{j=1}^n \mathbb{E}_{j+1} \gamma_j^{(n)2} \xrightarrow[n\to\infty]{\mathcal{P}} 1. \tag{6.3}$$

Assume, further, that the Lindeberg condition holds:

$$\forall \epsilon > 0 \qquad \frac{1}{\Theta_n^2} \sum_{j=1}^n \mathbb{E}\big(\gamma_j^{(n)2} \mathbf{1}_{|\gamma_j^{(n)}| \ge \epsilon \Theta_n}\big) \xrightarrow[n\to\infty]{} 0.$$

Then, $\Theta_n^{-1} \sum_{j=1}^n \gamma_j^{(n)}$ converges in distribution to $\mathcal{N}(0, 1)$.

REMARK 6.1. The condition

$$\exists \delta > 0 \qquad \frac{1}{\Theta_n^{2(1+\delta)}} \sum_{j=1}^n \mathbb{E}|\gamma_j^{(n)}|^{2+\delta} \xrightarrow[n\to\infty]{} 0, \tag{6.4}$$

known as *Lyapunov's condition*, implies Lindeberg's condition and is easier to establish (see, e.g., [4], Section 27, page 362).

The proof of the CLT will be carried out in three steps:

1. We first show that $\log\det(Y_n Y_n^* + \rho I) - \mathbb{E}\log\det(Y_n Y_n^* + \rho I)$ can be written as $\sum_{j=1}^n \gamma_j$, where (γ_j) is a martingale difference sequence.
2. We then prove that (γ_j) satisfies Lyapunov's condition (6.4), where Θ_n^2 is given by Theorem 3.1.
3. We finally prove (6.3), which implies the CLT.

Key lemmas. The two lemmas stated below will be of constant use in the sequel. The first lemma describes the asymptotic behavior of quadratic forms related to random matrices.

LEMMA 6.2. *Let $\mathbf{x} = (x_1, \dots, x_n)$ be a $n \times 1$ vector where the x_i are centered i.i.d. complex random variables with unit variance. Let M be a $n \times n$ deterministic complex matrix.*

1. (*Bai and Silverstein*, *Lemma* 2.7 *in* [2].) *Then*, *for any $p \geq 2$*, *there exists a constant K_p for which*

$$\mathbb{E}|\mathbf{x}^* M \mathbf{x} - \operatorname{Tr} M|^p \leq K_p\big((\mathbb{E}|x_1|^4 \operatorname{Tr} MM^*)^{p/2} + \mathbb{E}|x_1|^{2p} \operatorname{Tr}(MM^*)^{p/2}\big).$$

2. (*See also equation* (1.15) *in* [3].) *If we assume*, *moreover*, *that $\mathbb{E}x_1^2 = 0$*, *then*

$$\mathbb{E}(\mathbf{x}^* M \mathbf{x} - \operatorname{Tr} M)^2 = \operatorname{Tr} M^2 + \kappa \sum_{i=1}^n m_{ii}^2,$$

where $\kappa = \mathbb{E}|x_1|^4 - 2$.

As a consequence of the first part of this lemma, there exists a constant K, independent of j and n, for which

$$\mathbb{E}\left| y_j^* Q_j(-\rho) y_j - \frac{1}{n} \operatorname{Tr} D_j Q_j(-\rho) \right|^p \leq K n^{-p/2} \tag{6.5}$$

for $p \leq 4$.

We introduce here various intermediate quantities, where $1 \leq i \leq N$ and $1 \leq j \leq n$:

$$\begin{aligned}
c_i(z) &= -\left[z\left(1 + \frac{1}{n} \operatorname{Tr} \tilde{D}_i \mathbb{E}\tilde{Q}(z)\right)\right]^{-1}, & C &= \operatorname{diag}(c_i), \\
\tilde{c}_j(z) &= -\left[z\left(1 + \frac{1}{n} \operatorname{Tr} D_j \mathbb{E}Q(z)\right)\right]^{-1}, & \tilde{C} &= \operatorname{diag}(\tilde{c}_j), \\
b_i(z) &= -\left[z\left(1 + \frac{1}{n} \operatorname{Tr} \tilde{D}_i \tilde{C}(z)\right)\right]^{-1}, & B &= \operatorname{diag}(b_i), \\
\tilde{b}_j(z) &= -\left[z\left(1 + \frac{1}{n} \operatorname{Tr} D_j C(z)\right)\right]^{-1}, & \tilde{B} &= \operatorname{diag}(\tilde{b}_j).
\end{aligned} \tag{6.6}$$

The following lemma provides various bounds and approximation results.

LEMMA 6.3. *Consider the family of random matrices $(Y_n Y_n^*)$ and assume that Assumptions* A1 *and* A2 *hold true. Let $z = -\rho$ where $\rho > 0$. Then*:

1. *Matrices* C_n *satisfy* $\|C_n\| \le \frac{1}{\rho}$ *and* $0 < c_i \le \frac{1}{\rho}$, *these inequalities remaining true when* C *is replaced with* B *or* $\tilde{C}$.
2. *If we let* U_n *and* $\tilde{U}_n$ *be two sequences of real diagonal deterministic* $N \times N$ *and* $n \times n$ *matrices and assume that* $\sup_{n\ge1} \max(\|U_n\|, \|\tilde{U}_n\|) < \infty$, *then the following hold true*:
   (a) $\frac{1}{n} \operatorname{Tr} U(\mathbb{E}Q - T) \underset{n\to\infty}{\to} 0$ *and* $\frac{1}{n} \operatorname{Tr} \tilde{U}(\mathbb{E}\tilde{Q} - \tilde{T}) \underset{n\to\infty}{\to} 0$;
   (b) $\frac{1}{n} \operatorname{Tr} U(B - T) \underset{n\to\infty}{\to} 0$;
   (c) $\sup_n \mathbb{E}(\operatorname{Tr} U(Q - \mathbb{E}Q))^2 < \infty$;
   (d) $\sup_n \frac{1}{n^2} \mathbb{E}(\operatorname{Tr} U(Q - \mathbb{E}Q))^4 < \infty$.
3. [*Rank-one perturbation inequality*] *the resolvent* Q_j *satisfies* $|\operatorname{Tr} M(Q - Q_j)| \le \frac{\|M\|}{\rho}$ *for any* $N \times N$ *matrix* M (*see Lemma* 2.6 *in* [26]).

Proof of Lemma 6.3 is postponed to Appendix A.

Finally, we shall frequently use the following identities, which are obtained from the definitions of c_i and $\tilde{c}_j$, together with equations (6.2):

$$[Q(z)]_{ii} = c_i + zc_i[Q]_{ii}\left(\xi_i \tilde{Q}_i \xi_i^* - \frac{1}{n} \operatorname{Tr} \tilde{D}_i \mathbb{E}\tilde{Q}\right); \tag{6.7}$$

$$[\tilde{Q}(z)]_{jj} = \tilde{c}_j + z\tilde{c}_j[\tilde{Q}]_{jj}\left(y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j \mathbb{E}Q\right). \tag{6.8}$$

6.2. *Proof of step* 1: *The sum of a martingale difference sequence*. Recall that $\mathbb{E}_j = \mathbb{E}(\cdot \mid \mathcal{F}_j)$, where $\mathcal{F}_j = \sigma(y_\ell, j \le \ell \le n)$. We have

$$\begin{aligned}
&\log\det(YY^* + \rho I_N) - \mathbb{E}\log\det(YY^* + \rho I_N)\\
&\quad = \sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j+1}) \log\det(YY^* + \rho I_N)\\
&\quad \overset{\text{(a)}}{=} -\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j+1}) \log\left(\frac{\det(Y^j Y^{j*} + \rho I_N)}{\det(YY^* + \rho I_N)}\right)\\
&\quad \overset{\text{(b)}}{=} -\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j+1}) \log\left(\frac{\det(Y^{j*} Y^j + \rho I_{n-1})}{\det(Y^*Y + \rho I_n)}\right)\\
&\quad \overset{\text{(c)}}{=} -\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j+1}) \log[\tilde{Q}(-\rho)]_{jj}\\
&\quad \overset{\text{(d)}}{=} \sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j+1}) \log(1 + y_j^* Q_j(-\rho) y_j),
\end{aligned}$$

where (a) follows from the fact that Y^j does not depend upon y_j, in particular, $\mathbb{E}_j \log\det(Y^j Y^{j*} + \rho I_N) = \mathbb{E}_{j+1} \log\det(Y^j Y^{j*} + \rho I_N)$; (b) follows from the fact that $\det(Y^{j*}Y^j + \rho I_{n-1}) = \det(Y^j Y^{j*} + \rho I_N) \times \rho^{n-1-N}$ [and a similar expression for $\det(Y^*Y + \rho I_n)$]; (c) follows from the equality

$$[\tilde{Q}(-\rho)]_{jj} = \frac{\det(Y^{j*}Y^j + \rho I_{n-1})}{\det(Y^*Y + \rho I_n)},$$

which is a consequence of the general inverse formula $A^{-1} = \frac{1}{\det(A)}\mathrm{adj}(A)$, where $\mathrm{adj}(A)$ is the transposed matrix of cofactors of A (see [18], Section 0.8.2); (d) follows from (6.2). We therefore have

$$\begin{aligned}&\log\det(YY^* + \rho I_N) - \mathbb{E}\log\det(YY^* + \rho I_N)\\ &\quad = \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j+1}) \log\bigl(1 + y_j^* Q_j(-\rho) y_j\bigr) \stackrel{\triangle}{=} \sum_{j=1}^{n} \gamma_j.\end{aligned}$$

As we have

$$\mathbb{E}_j \log\biggl(1 + \frac{1}{n}\operatorname{Tr} D_j Q_j\biggr) = \mathbb{E}_{j+1} \log\biggl(1 + \frac{1}{n}\operatorname{Tr} D_j Q_j\biggr),$$

γ_j can be expressed as

$$\begin{aligned}\gamma_j &= (\mathbb{E}_j - \mathbb{E}_{j+1}) \log\biggl(1 + \frac{y_j^* Q_j y_j - (1/n)\operatorname{Tr} D_j Q_j}{1 + (1/n)\operatorname{Tr} D_j Q_j}\biggr)\\ &= (\mathbb{E}_j - \mathbb{E}_{j+1}) \log(1 + \Gamma_j).\end{aligned} \tag{6.9}$$

where $\Gamma_j = \frac{y_j^* Q_j y_j - (1/n)\operatorname{Tr} D_j Q_j}{1 + (1/n)\operatorname{Tr} D_j Q_j}$. The sequence $\gamma_n, \dots, \gamma_1$ is a martingale difference sequence with respect to the increasing filtration $\mathcal{F}_n, \dots, \mathcal{F}_1$ and so step 1 is established.

6.3. *Proof of step* 2: *Validation of Lyapunov's condition* (6.4). In the remainder of this section, $z = -\rho$. Let $\delta > 0$ be a fixed positive number that will be specified below. As $\liminf \Theta_n^2 > 0$ by Theorem 3.1, we only need to prove that $\sum_{j=1}^n \mathbb{E}|\gamma_j|^{2+\delta} \to_n 0$. We have $\mathbb{E}|\gamma_j|^{2+\delta} = \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j+1}) \log(1 + \Gamma_j)|^{2+\delta}$; the Minkowski and Jensen inequalities yield

$$\begin{aligned}&(\mathbb{E}|\gamma_j|^{2+\delta})^{1/(2+\delta)}\\ &\quad \le \bigl(\mathbb{E}|\mathbb{E}_j \log(1 + \Gamma_j)|^{2+\delta}\bigr)^{1/(2+\delta)} + \bigl(\mathbb{E}|\mathbb{E}_{j+1} \log(1 + \Gamma_j)|^{2+\delta}\bigr)^{1/(2+\delta)}\\ &\quad \le 2\bigl(\mathbb{E}|\log(1 + \Gamma_j)|^{2+\delta}\bigr)^{1/(2+\delta)}.\end{aligned}$$

Otherwise stated,

$$\mathbb{E}|\gamma_j|^{2+\delta} \le K_0\, \mathbb{E}|\log(1 + \Gamma_j)|^{2+\delta}, \tag{6.10}$$

where $K_0 = 2^{2+\delta}$. Since $y_j^* Q_j y_j \ge 0$, Γ_j [defined in (6.9)] is lower bounded:

$$\Gamma_j \ge \frac{-(1/n)\operatorname{Tr} D_j Q_j}{1 + (1/n)\operatorname{Tr} D_j Q_j}.$$

Now, since

$$0 \le \frac{1}{n} \operatorname{Tr} D_j Q_j(-\rho) \le \frac{\|D_j\|}{n} \operatorname{Tr} Q_j(-\rho) \le K_1 \stackrel{\triangle}{=} \frac{\sigma_{\max}^2}{\rho} \sup_n \left(\frac{N}{n}\right)$$

and since $x \mapsto \frac{x}{1+x}$ is nondecreasing, we have

$$\frac{(1/n) \operatorname{Tr} D_j Q_j}{1 + (1/n) \operatorname{Tr} D_j Q_j} \le K_2 \stackrel{\triangle}{=} \frac{K_1}{1 + K_1} < 1. \tag{6.11}$$

In particular, $\Gamma_j \ge -K_2 > -1$. The function $(-1, \infty) \ni x \mapsto \frac{\log(1+x)}{x}$ is nonnegative, nonincreasing. Therefore, $\frac{\log(1+x)}{x} \le \frac{\log(1-K_2)}{K_2}$ for $x \in [-K_2, \infty)$. Plugging this into (6.10) yields

$$\begin{aligned} \mathbb{E}|\gamma_j|^{2+\delta} &\le K_0 K_2^{2+\delta} \mathbb{E}|\Gamma_j|^{2+\delta} \\ &\stackrel{\triangle}{=} K_3 \mathbb{E}|\Gamma_j|^{2+\delta} \le K_3 \mathbb{E}\left| y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j Q_j \right|^{2+\delta}. \end{aligned}$$

By Lemma 6.2(1), the right-hand side of the last inequality is less than $K_4 n^{-(1+\delta/2)}$ as soon as $\mathbb{E}|X_{11}|^{2+\delta} < \infty$. This is ensured by Assumption A1 for $\delta \le 6$. Therefore, Lyapunov's condition (6.4) holds and step 2 is proved.

6.4. *Proof of step* 3: *Convergence of the normalized sum of conditional variances*. This section, by far the most involved in this article, is devoted to establishing the convergence result (6.3), hence the CLT. In an attempt to guide the reader, we divide it into five stages. Recall that $z = -\rho$ and

$$\gamma_j = (\mathbb{E}_j - \mathbb{E}_{j+1}) \log(1 + \Gamma_j), \qquad \text{where } \Gamma_j = \frac{y_j^* Q_j y_j - (1/n) \operatorname{Tr} D_j Q_j}{1 + (1/n) \operatorname{Tr} D_j Q_j}.$$

In order to apply Theorem 6.1, we shall prove that $\Theta_n^{-2} \sum_{j=1}^n \mathbb{E}_{j+1} \gamma_j^2 \xrightarrow{\mathcal{P}} 1$, where Θ_n^2 is given by Theorem 3.1. Since $\liminf \Theta_n^2 > 0$, it is sufficient to establish the following convergence result:

$$\sum_{j=1}^n \mathbb{E}_{j+1} \gamma_j^2 - \Theta_n^2 \xrightarrow[n\to\infty]{\mathcal{P}} 0. \tag{6.12}$$

Instead of working with Θ_n, we shall work with $\tilde{\Theta}_n$ [introduced in Section 5, see equation (5.1)] and prove

$$\sum_{j=1}^n \mathbb{E}_{j+1} \gamma_j^2 - \tilde{\Theta}_n^2 \xrightarrow[n\to\infty]{\mathcal{P}} 0. \tag{6.13}$$

In the sequel, K will denote a constant whose value may change from line to line, but which will depend neither on n nor on $j \le n$.

Here are the main steps of the proof:

1. The following convergence result holds true:

$$\sum_{j=1}^{n} \mathbb{E}_{j+1}\gamma_j^2 - \sum_{j=1}^{n} \mathbb{E}_{j+1}(\mathbb{E}_j\Gamma_j)^2 \xrightarrow[n\to\infty]{\mathcal{P}} 0. \tag{6.14}$$

This convergence result roughly follows from a first-order approximation, as we shall informally discuss. Recall that $\gamma_i = (\mathbb{E}_j - \mathbb{E}_{j+1})\log(1+\Gamma_j)$ and that $\Gamma_j \to 0$, by Lemma 6.2(1). A first-order approximation of $\log(1+x)$ yields $\gamma_j \approx (\mathbb{E}_j - \mathbb{E}_{j+1})\Gamma_j$. As $\mathbb{E}(y_j^* Q_j y_j \mid Q_j) = \frac{1}{n}\operatorname{Tr} D_j Q_j$, one can prove that $\mathbb{E}_{j+1}\Gamma_j = 0$, hence $\gamma_j \approx \mathbb{E}_j\Gamma_j$, and one may expect $\mathbb{E}_{j+1}\gamma_j^2 \approx \mathbb{E}_{j+1}(\mathbb{E}_j\Gamma_j)^2$ and even (6.14), as we shall demonstrate.

2. Recall that $\kappa = \mathbb{E}|X_{11}|^4 - 2$. The following equality holds true:

$$\begin{aligned}
&\mathbb{E}_{j+1}(\mathbb{E}_j\Gamma_j)^2 \\
&\quad = \frac{1}{n^2(1+(1/n)\operatorname{Tr} D_j \mathbb{E} Q)^2} \\
&\qquad \times \left(\operatorname{Tr} D_j(\mathbb{E}_{j+1}Q_j)D_j(\mathbb{E}_{j+1}Q_j) + \kappa \sum_{i=1}^{N} \sigma_{ij}^4 (\mathbb{E}_{j+1}[Q_j]_{ii})^2\right) \\
&\qquad + \boldsymbol{\varepsilon}_{\mathbf{2},\mathbf{j}},
\end{aligned} \tag{6.15}$$

where

$$\max_{j\le n} \mathbb{E}|\boldsymbol{\varepsilon}_{\mathbf{2},\mathbf{j}}| \le \frac{K}{n^{3/2}}$$

for some given K.

A closer look at the right-hand side of (6.15) leads to the following observations. By Lemma 6.3(2a), the denominator $(1+\frac{1}{n}\operatorname{Tr} D_j\mathbb{E}Q)^2$ can be approximated by $(1+\frac{1}{n}\operatorname{Tr} D_j T)^2$; moreover, it is possible to prove that $[Q_j]_{ii} \approx [T]_{ii}$ [some details are given in the course of the proof of step (5) below]. Hence,

$$\frac{\kappa}{n}\sum_{i=1}^{N} \sigma_{ij}^4 (\mathbb{E}_{j+1}[Q_j]_{ii})^2 \approx \frac{\kappa}{n}\operatorname{Tr} D_j^2 T^2.$$

Therefore, it remains to study the asymptotic behavior of the term $\frac{1}{n} \times \operatorname{Tr} D_j(\mathbb{E}_{j+1}Q_j)D_j(\mathbb{E}_{j+1}Q_j)$ in order to understand (6.15). This is the purpose of step (3) below.

(3) In order to evaluate $\frac{1}{n}\operatorname{Tr} D_j(\mathbb{E}_{j+1}Q_j)D_j(\mathbb{E}_{j+1}Q_j)$ for large n, we introduce the random variables

$$\boldsymbol{\chi}_{\boldsymbol{\ell},\mathbf{n}}^{(\mathbf{j})} = \frac{1}{n}\operatorname{Tr} D_\ell(\mathbb{E}_{j+1}Q)D_j Q, \qquad j \le \ell \le n. \tag{6.16}$$

Note that, up to rank one perturbations, $\mathbb{E}_j \boldsymbol{\chi}^{(\mathbf{j})}_{\mathbf{j},\mathbf{n}}$ is very close to the term of interest. We prove here that $\boldsymbol{\chi}^{(\mathbf{j})}_{\boldsymbol{\ell},\mathbf{n}}$ satisfies the following equation:

$$\begin{aligned}\boldsymbol{\chi}^{(\mathbf{j})}_{\boldsymbol{\ell},\mathbf{n}} &= \frac{1}{n}\sum_{m=j+1}^{n} \frac{(1/n)\operatorname{Tr}(D_\ell B D_m \mathbb{E}Q)}{(1+(1/n)\operatorname{Tr} D_m \mathbb{E}Q)^2}\boldsymbol{\chi}^{(\mathbf{j})}_{\mathbf{m},\mathbf{n}} \\ &\quad + \frac{1}{n}\operatorname{Tr} D_\ell B D_j \mathbb{E}Q + \boldsymbol{\varepsilon}_{\mathbf{3},\boldsymbol{\ell}\mathbf{j}}, \qquad j \le \ell \le n,\end{aligned} \tag{6.17}$$

where B is defined in Section 6.1 and where

$$\max_{\ell, j \le n} \mathbb{E}|\boldsymbol{\varepsilon}_{\mathbf{3},\boldsymbol{\ell}\mathbf{j}}| \le \frac{K}{\sqrt{n}}.$$

(4) Recall that we have proven in Section 5 (Lemma 5.1) that the (deterministic) system

$$\mathbf{y}^{(\mathbf{j})}_{\boldsymbol{\ell},\mathbf{n}} = \sum_{m=j+1}^{n} a_{\ell,m}\mathbf{y}^{(\mathbf{j})}_{\mathbf{m},\mathbf{n}} + a_{\ell,j}, \qquad \text{for } j \le \ell \le n,$$

where $a_{\ell,m} = \frac{1}{n^2}\frac{\operatorname{Tr} D_\ell D_m T^2}{(1+\frac{1}{n}\operatorname{Tr} D_\ell T)^2}$, admits a unique solution. If we write $\mathbf{x}^{(\mathbf{j})}_{\boldsymbol{\ell},\mathbf{n}} = n(1+\frac{1}{n}\operatorname{Tr} D_\ell T)^2 \mathbf{y}^{(\mathbf{j})}_{\boldsymbol{\ell},\mathbf{n}}$, then $(\mathbf{x}^{(\mathbf{j})}_{\boldsymbol{\ell},\mathbf{n}}, j \le \ell \le n)$ readily satisfies the following system:

$$\mathbf{x}^{(\mathbf{j})}_{\boldsymbol{\ell},\mathbf{n}} = \frac{1}{n}\sum_{m=j+1}^{n}\frac{(1/n)\operatorname{Tr} D_\ell D_m T^2}{(1+(1/n)\operatorname{Tr} D_m T)^2}\mathbf{x}^{(\mathbf{j})}_{\mathbf{m},\mathbf{n}} + \frac{1}{n}\operatorname{Tr} D_\ell D_j T^2, \qquad j \le \ell \le n.$$

As one may notice, (6.17) is a perturbed version of the system above and we shall indeed prove that

$$\begin{gathered}\boldsymbol{\chi}^{(\mathbf{j})}_{\mathbf{j},\mathbf{n}} = \mathbf{x}^{(\mathbf{j})}_{\mathbf{j},\mathbf{n}} + \boldsymbol{\varepsilon}_{\mathbf{41},\mathbf{j}} + \boldsymbol{\varepsilon}_{\mathbf{42},\mathbf{j}}, \\ \text{where } \max_{j\le n}\mathbb{E}|\boldsymbol{\varepsilon}_{\mathbf{41},\mathbf{j}}| \le \frac{K}{\sqrt{n}} \text{ and } \max_{j \le n}|\boldsymbol{\varepsilon}_{\mathbf{42},\mathbf{j}}| \le \delta_n,\end{gathered} \tag{6.18}$$

the sequence (δ_n) being deterministic with $\delta_n \to 0$ as $n \to \infty$.

(5) Combining the previous results, we finally prove that

$$\sum_{j=1}^{n}\mathbb{E}_{j+1}(\mathbb{E}_j\Gamma_j)^2 - \tilde{\Theta}_n^2 \xrightarrow[n\to\infty]{\mathcal{P}} 0. \tag{6.19}$$

This, together with (6.14), yields convergence results (6.13) and (6.12) which, in turn, proves (6.3), completing the proof of Theorem 3.2.

PROOF OF (6.14). Recall that $\frac{(1/n)\operatorname{Tr} D_j Q_j}{1+(1/n)\operatorname{Tr} D_j Q_j} \le K_2 < 1$, by (6.11). In particular, $\Gamma_j \ge -K_2 > -1$. We first prove that

$$\mathbb{E}_j \log(1+\Gamma_j) = \mathbb{E}_j\Gamma_j + \boldsymbol{\varepsilon}_{\mathbf{11},\mathbf{j}} + \boldsymbol{\varepsilon}_{\mathbf{12},\mathbf{j}},$$

where

$$\begin{cases} \boldsymbol{\varepsilon}_{\mathbf{11},\mathbf{j}} = \mathbb{E}_j \log(1+\Gamma_j)\mathbf{1}_{|\Gamma_j|\le K_2} - \mathbb{E}_j \Gamma_j, \\ \boldsymbol{\varepsilon}_{\mathbf{12},\mathbf{j}} = \mathbb{E}_j \log(1+\Gamma_j)\mathbf{1}_{(K_2,\infty)}(\Gamma_j) \end{cases} \quad \text{and}$$

$$\begin{cases} \max_{j\le n} \mathbb{E}\boldsymbol{\varepsilon}^2_{\mathbf{11},\mathbf{j}} \le \dfrac{K}{n^2}, \\ \max_{j\le n} \mathbb{E}\boldsymbol{\varepsilon}^2_{\mathbf{12},\mathbf{j}} \le \dfrac{K}{n^2}. \end{cases} \tag{6.20}$$

In the sequel, we shall often omit the subscript j while dealing with the $\boldsymbol{\varepsilon}$'s. As $0 < K_2 < 1$, we have

$$\begin{aligned} |\boldsymbol{\varepsilon}_{\mathbf{11}}| &= \left|\mathbb{E}_j\left(\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}\Gamma_j^k \mathbf{1}_{|\Gamma_j|\le K_2} - \Gamma_j\right)\right| \\ &\le \mathbb{E}_j \Gamma_j \mathbf{1}_{\Gamma_j > K_2} + \sum_{k=2}^{\infty} \mathbb{E}_j |\Gamma_j|^k \mathbf{1}_{|\Gamma_j|\le K_2} \le \mathbb{E}_j \Gamma_j \mathbf{1}_{\Gamma_j>K_2} + \frac{\mathbb{E}_j \Gamma_j^2 \mathbf{1}_{|\Gamma_j|\le K_2}}{1-K_2}. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathbb{E}\boldsymbol{\varepsilon}^2_{\mathbf{11}} &\overset{\text{(a)}}{\le} 2\left(\mathbb{E}\Gamma_j^2 \mathbf{1}_{\Gamma_j>K_2} + \frac{\mathbb{E}\Gamma_j^4 \mathbf{1}_{|\Gamma_j|\le K_2}}{(1-K_2)^2}\right) \\ &\overset{\text{(b)}}{\le} \frac{2\mathbb{E}\Gamma_j^4}{K_2^2} + \frac{2\mathbb{E}\Gamma_j^4}{(1-K_2)^2} \\ &\overset{\text{(c)}}{\le} \left(\frac{2}{K_2^2} + \frac{2}{(1-K_2)^2}\right)\mathbb{E}\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j Q_j\right)^4 \overset{\text{(d)}}{\le} \frac{K}{n^2}, \end{aligned}$$

where (a) follows from $(a+b)^2 \le 2(a^2+b^2)$, (b) from the inequality $\Gamma_j^2 \mathbf{1}_{\Gamma_j>K_2} \le \Gamma_j^2 (\frac{\Gamma_j}{K_2})^2 \mathbf{1}_{\Gamma_j>K_2}$, (c) from the fact that the denominator of Γ_j is larger than one and (d) from Lemma 6.2(1), as X_{11} has a finite eighth moment by Assumption A1.

Now, $0 \le \boldsymbol{\varepsilon}_{\mathbf{12}} \le \mathbb{E}_j \Gamma_j \mathbf{1}_{\Gamma_j>K_2}$. Thus, $\mathbb{E}\boldsymbol{\varepsilon}^2_{\mathbf{12}} \le \mathbb{E}\Gamma_j^2 \mathbf{1}_{\Gamma_j>K_2} \le K_2^{-2}\mathbb{E}\Gamma_j^4 \mathbf{1}_{\Gamma_j>K_2}$. Lemma 6.2(1) again yields

$$\mathbb{E}\boldsymbol{\varepsilon}^2_{\mathbf{12}} \le \frac{K}{n^2}$$

and (6.20) is proved. Similarly, we can prove

$$\mathbb{E}_{j+1} \log(1+\Gamma_j) = \mathbb{E}_{j+1}\Gamma_j + \boldsymbol{\varepsilon}_{\mathbf{13},\mathbf{j}} \qquad \text{with } \max_{j\le n} \mathbb{E}\boldsymbol{\varepsilon}^2_{\mathbf{13},\mathbf{j}} \le \frac{K}{n^2}.$$

Denote by $\mathcal{F}_{Q_j}$ the σ-field generated by all of the y_k's except y_j, and by $\mathbb{E}_{Q_j}$ the conditional expectation with respect to $\mathcal{F}_{Q_j}$. Then,

$$\mathbb{E}_{j+1}\Gamma_j = \mathbb{E}_{j+1}\mathbb{E}_{Q_j}\left(\frac{y_j^* Q_j y_j - (1/n)\operatorname{Tr} D_j Q_j}{1 + (1/n)\operatorname{Tr} D_j Q_j}\right) = 0$$

since y_j and $\mathcal{F}_{Q_j}$ are independent and since $\mathbb{E}_{Q_j}(y_j^* Q_j y_j) = \frac{1}{n}\operatorname{Tr} D_j Q_j$. Collecting all of the previous estimates, we obtain

$$\gamma_j = \mathbb{E}_j \Gamma_j + \boldsymbol{\varepsilon}_{\mathbf{11,j}} + \boldsymbol{\varepsilon}_{\mathbf{12,j}} - \boldsymbol{\varepsilon}_{\mathbf{13,j}} \stackrel{\triangle}{=} \mathbb{E}_j \Gamma_j + \boldsymbol{\varepsilon}_{\mathbf{14,j}},$$

where $\max_{j \le n} \mathbb{E} \boldsymbol{\varepsilon}_{\mathbf{14,j}}^2 \le K n^{-2}$, by Minkowski's inequality. We therefore have $\mathbb{E}_{j+1}(\gamma_j)^2 = \mathbb{E}_{j+1}(\mathbb{E}_j \Gamma_j + \boldsymbol{\varepsilon}_{\mathbf{14,j}})^2$. Let

$$\boldsymbol{\varepsilon}_{\mathbf{1,j}} \stackrel{\triangle}{=} \mathbb{E}_{j+1}(\gamma_j)^2 - \mathbb{E}_{j+1}(\mathbb{E}_j \Gamma_j)^2 = \mathbb{E}_{j+1} \boldsymbol{\varepsilon}_{\mathbf{14,j}}^2 + 2\mathbb{E}_{j+1}(\boldsymbol{\varepsilon}_{\mathbf{14,j}} \mathbb{E}_j \Gamma_j).$$

Then,

$$\begin{aligned}\mathbb{E}|\boldsymbol{\varepsilon}_{\mathbf{1,j}}| &\le \mathbb{E}\boldsymbol{\varepsilon}_{\mathbf{14,j}}^2 + 2\mathbb{E}|\boldsymbol{\varepsilon}_{\mathbf{14,j}} \mathbb{E}_j \Gamma_j| \\ &\stackrel{\text{(a)}}{\le} \mathbb{E}\boldsymbol{\varepsilon}_{\mathbf{14,j}}^2 + 2(\mathbb{E}\boldsymbol{\varepsilon}_{\mathbf{14,j}}^2)^{1/2} (\mathbb{E}\Gamma_j^2)^{1/2} \stackrel{\text{(b)}}{\le} \frac{K}{n^{3/2}},\end{aligned}$$

where (a) follows from the Cauchy–Schwarz inequality and $(\mathbb{E}_j \Gamma_j)^2 \le \mathbb{E}_j \Gamma_j^2$, and (b) follows from Lemma 6.2(1), which yields $\mathbb{E}\Gamma_j^2 \le K n^{-1}$. Finally, we have $\sum_{j=1}^n \mathbb{E}|\mathbb{E}_{j+1}(\gamma_j)^2 - \mathbb{E}_{j+1}(\mathbb{E}_j \Gamma_j)^2| \le K n^{-1/2}$, which implies (6.14). $\square$

PROOF OF (6.15). We have

$$\begin{aligned}\mathbb{E}_j \Gamma_j &= \mathbb{E}_j \left(\frac{y_j^* Q_j y_j - (1/n)\operatorname{Tr} D_j Q_j}{1 + (1/n)\operatorname{Tr} D_j Q_j} \right) \\ &= \frac{1}{1 + (1/n)\operatorname{Tr} D_j \mathbb{E} Q} \\ &\quad \times \left\{ \mathbb{E}_j \left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j Q_j \right) \right. \\ &\quad \left. - \mathbb{E}_j \left(\frac{y_j^* Q_j y_j - (1/n)\operatorname{Tr} D_j Q_j}{1 + (1/n)\operatorname{Tr} D_j Q_j} \left(\frac{1}{n}\operatorname{Tr} D_j Q_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E} Q \right) \right) \right\}.\end{aligned}$$

Hence,

$$\begin{aligned}&\mathbb{E}_{j+1}(\mathbb{E}_j \Gamma_j)^2 \\ &\quad = \frac{1}{(1 + (1/n)\operatorname{Tr} D_j \mathbb{E} Q)^2} \\ &\qquad \times \mathbb{E}_{j+1} \left(\left(y_j^* (\mathbb{E}_j Q_j) y_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E}_j Q_j \right)^2 + \boldsymbol{\varepsilon}_{\mathbf{21,j}} + \boldsymbol{\varepsilon}_{\mathbf{22,j}} \right) \qquad (6.21) \\ &\quad = \frac{1}{(1 + (1/n)\operatorname{Tr} D_j \mathbb{E} Q)^2} \\ &\qquad \times \mathbb{E}_{j+1} \left(y_j^* (\mathbb{E}_j Q_j) y_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E}_j Q_j \right)^2 + \boldsymbol{\varepsilon}_{\mathbf{2,j}},\end{aligned}$$

where

$$\boldsymbol{\varepsilon}_{\mathbf{21,j}} = \left[\mathbb{E}_j\left(\frac{y_j^* Q_j y_j - (1/n)\operatorname{Tr} D_j Q_j}{1+(1/n)\operatorname{Tr} D_j Q_j}\left(\frac{1}{n}\operatorname{Tr} D_j Q_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E} Q\right)\right)\right]^2,$$

$$\begin{aligned}\boldsymbol{\varepsilon}_{\mathbf{22,j}} &= -2\mathbb{E}_j\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j Q_j\right)\\ &\quad\times \mathbb{E}_j\left(\frac{y_j^* Q_j y_j - (1/n)\operatorname{Tr} D_j Q_j}{1+(1/n)\operatorname{Tr} D_j Q_j}\left(\frac{1}{n}\operatorname{Tr} D_j Q_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E} Q\right)\right),\end{aligned}$$

$$\boldsymbol{\varepsilon}_{\mathbf{2,j}} = \frac{\mathbb{E}_{j+1}(\boldsymbol{\varepsilon}_{\mathbf{21,j}} + \boldsymbol{\varepsilon}_{\mathbf{22,j}})}{(1+(1/n)\operatorname{Tr} D_j \mathbb{E} Q)^2}.$$

As $\frac{1}{n}\operatorname{Tr} D_j Q_j \geq 0$, standard inequalities yield

$$\begin{aligned}\mathbb{E}\boldsymbol{\varepsilon}_{\mathbf{21,j}} &\leq \left[\mathbb{E}\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j Q_j\right)^4\right]^{1/2}\\ &\quad\times \left[\mathbb{E}\left(\frac{1}{n}\operatorname{Tr} D_j Q_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E} Q\right)^4\right]^{1/2}.\end{aligned}$$

By Lemma 6.2(1), $\mathbb{E}(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j Q_j)^4 \leq Kn^{-2}$. Due to the convex inequality $(a+b)^4 \leq 2^3(a^4+b^4)$, we obtain

$$\begin{aligned}&\mathbb{E}\left(\frac{1}{n}\operatorname{Tr} D_j(Q_j - \mathbb{E} Q)\right)^4\\ &\quad= \mathbb{E}\left(\frac{1}{n}\operatorname{Tr} D_j(Q_j - \mathbb{E} Q_j) + \frac{1}{n}\operatorname{Tr} D_j(\mathbb{E} Q_j - \mathbb{E} Q)\right)^4\\ &\quad\leq K\left\{\mathbb{E}\left(\frac{1}{n}\operatorname{Tr} D_j(Q_j - \mathbb{E} Q_j)\right)^4 + \mathbb{E}\left(\frac{1}{n}\operatorname{Tr} D_j(\mathbb{E} Q_j - \mathbb{E} Q)\right)^4\right\},\end{aligned}$$

where the first term of the right-hand side is bounded by Kn^{-2}, by (2d) in Lemma 6.3, and the second one is bounded by Kn^{-4}, due to the rank one perturbation inequality [Lemma 6.3(3)]. Therefore, $\mathbb{E}\boldsymbol{\varepsilon}_{\mathbf{21,j}} \leq Kn^{-2}$ and similar derivations yield $\mathbb{E}|\boldsymbol{\varepsilon}_{\mathbf{22,j}}| \leq Kn^{-3/2}$. Combining these two results yields the bound $\mathbb{E}|\boldsymbol{\varepsilon}_{\mathbf{2,j}}| \leq Kn^{-3/2}$. Let us now expand the term $\mathbb{E}_{j+1}(y_j^*\mathbb{E}_j Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E}_j Q_j)^2$ in the right-hand side of (6.21).

Recall that $\mathbb{E}_j Q_j = \mathbb{E}_{j+1} Q_j$ and that $y_j = D_j^{1/2}(\frac{X_{1j}}{\sqrt{n}}, \dots, \frac{X_{Nj}}{\sqrt{n}})^T$. Note also that $\mathbb{E}_{j+1}(y_j^*\mathbb{E}_j Q_j y_j) = \frac{1}{n}\operatorname{Tr} D_j \mathbb{E}_{j+1} Q_j$. Lemma 6.2(2) then immediately yields

$$\begin{aligned}&\mathbb{E}_{j+1}\left(y_j^*\mathbb{E}_j Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E}_j Q_j\right)^2\\ &\quad= \frac{1}{n^2}\left(\operatorname{Tr} D_j(\mathbb{E}_{j+1} Q_j) D_j(\mathbb{E}_{j+1} Q_j) + \kappa \sum_{\ell=1}^{N} \sigma_{\ell j}^4 (\mathbb{E}_{j+1}[Q_j]_{\ell\ell})^2\right).\end{aligned}$$

Equation (6.15) is thus proved. $\square$

PROOF OF (6.17). Recall that $\boldsymbol{\chi}_{\boldsymbol{\ell},\mathbf{n}}^{(\mathbf{j})} = \frac{1}{n}\operatorname{Tr} D_\ell(\mathbb{E}_{j+1}Q)D_jQ$. The outline of the proof of (6.17) is given by the following set of equations, the $\boldsymbol{\chi}$'s and $\boldsymbol{\varepsilon}$'s being introduced as and when required:

$$\boldsymbol{\chi}_{\boldsymbol{\ell},\mathbf{n}}^{(\mathbf{j})} = \boldsymbol{\chi}_{\mathbf{1}} + \boldsymbol{\chi}_{\mathbf{2}} - \boldsymbol{\chi}_{\mathbf{3}}; \tag{6.22}$$

$$\boldsymbol{\chi}_{\mathbf{3}} = \boldsymbol{\chi}'_{\mathbf{3}} + \boldsymbol{\varepsilon}_3; \tag{6.23}$$

$$\boldsymbol{\chi}'_{\mathbf{3}} = \boldsymbol{\chi}_{\mathbf{4}} + \boldsymbol{\chi}_{\mathbf{5}} + \boldsymbol{\varepsilon}'_3; \tag{6.24}$$

$$\boldsymbol{\chi}_{\mathbf{5}} = \boldsymbol{\chi}_{\mathbf{6}} - \boldsymbol{\chi}_{\mathbf{7}} + \boldsymbol{\varepsilon}_6 - \boldsymbol{\varepsilon}_7. \tag{6.25}$$

Collecting the previous equations, we will end up with

$$\boldsymbol{\chi}_{\boldsymbol{\ell},\mathbf{n}}^{(\mathbf{j})} = \boldsymbol{\chi}_{\mathbf{1}} + \boldsymbol{\chi}_{\mathbf{2}} - \boldsymbol{\chi}_{\mathbf{4}} - \boldsymbol{\chi}_{\mathbf{6}} + \boldsymbol{\chi}_{\mathbf{7}} + \boldsymbol{\varepsilon}, \qquad \text{where } \boldsymbol{\varepsilon} = -\boldsymbol{\varepsilon}_3 + \boldsymbol{\varepsilon}'_3 - \boldsymbol{\varepsilon}_6 + \boldsymbol{\varepsilon}_7. \tag{6.26}$$

Let us first give decomposition (6.22) and introduce $\boldsymbol{\chi}_{\mathbf{1}}$, $\boldsymbol{\chi}_{\mathbf{2}}$ and $\boldsymbol{\chi}_{\mathbf{3}}$. Recall that B (defined in Section 6.1) is the $N \times N$ diagonal matrix $B = \operatorname{diag}(b_i)$, where $b_i = (\rho(1 + \frac{1}{n}\operatorname{Tr}\tilde{D}_i\tilde{C}))^{-1}$. This yields

$$\begin{aligned} Q &= B + B(B^{-1} - Q^{-1})Q \\ &= B + B\left(\rho \operatorname{diag}\left(\frac{1}{n}\operatorname{Tr}\tilde{D}_i\tilde{C}\right) - YY^*\right)Q. \end{aligned}$$

Therefore,

$$\begin{aligned} \boldsymbol{\chi}_{\boldsymbol{\ell},\mathbf{n}}^{(\mathbf{j})} &= \frac{1}{n}\operatorname{Tr} D_\ell(\mathbb{E}_{j+1}Q)D_jQ \\ &= \frac{1}{n}\operatorname{Tr} D_\ell B D_j Q + \frac{\rho}{n}\operatorname{Tr} D_\ell B \operatorname{diag}\left(\frac{1}{n}\operatorname{Tr}\tilde{D}_i\tilde{C}\right)(\mathbb{E}_{j+1}Q)D_jQ \\ &\quad - \frac{1}{n}\operatorname{Tr} D_\ell B\left(\sum_{m=1}^n \mathbb{E}_{j+1} y_m y_m^* Q\right)D_jQ \\ &\stackrel{\triangle}{=} \boldsymbol{\chi}_{\mathbf{1}} + \boldsymbol{\chi}_{\mathbf{2}} - \boldsymbol{\chi}_{\mathbf{3}} \end{aligned}$$

and (6.22) is established. We now turn to decomposition (6.23). Identities (6.1) and (6.2) yield

$$y_m^*Q = y_m^*Q_m - y_m^*\frac{Q_m y_m y_m^* Q_m}{1 + y_m^* Q_m y_m} = \frac{y_m^* Q_m}{1 + y_m^* Q_m y_m} = \rho[\tilde{Q}]_{mm} y_m^* Q_m.$$

Using this equation, we have

$$
\begin{aligned}
\boldsymbol{\chi_3} &= \frac{1}{n}\operatorname{Tr} D_\ell B\left(\sum_{m=1}^n \mathbb{E}_{j+1} y_m y_m^* Q\right) D_j Q \\
&= \frac{\rho}{n}\operatorname{Tr} D_\ell B\left(\sum_{m=1}^n \mathbb{E}_{j+1}([\tilde{Q}]_{mm} y_m y_m^* Q_m)\right) D_j Q \\
&\overset{\text{(a)}}{=} \frac{\rho}{n}\operatorname{Tr} D_\ell B\left(\sum_{m=1}^n \tilde{c}_m \mathbb{E}_{j+1}(y_m y_m^* Q_m)\right) D_j Q \\
&\quad - \frac{\rho^2}{n}\operatorname{Tr} D_\ell B\left(\sum_{m=1}^n \tilde{c}_m \mathbb{E}_{j+1}\left([\tilde{Q}]_{mm}\left(y_m^* Q_m y_m \right.\right.\right. \\
&\qquad\qquad\qquad \left.\left.\left. - \frac{1}{n}\operatorname{Tr} D_m \mathbb{E} Q\right) y_m y_m^* Q_m\right)\right) D_j Q \\
&\overset{\triangle}{=} \boldsymbol{\chi_3'} + \boldsymbol{\varepsilon}_3,
\end{aligned}
$$

where (a) follows directly from (6.8). We are now in a position to prove that

$$
\max_{\ell, j \le n} \mathbb{E}|\boldsymbol{\varepsilon}_3| \le \frac{K}{\sqrt{n}}. \tag{6.27}
$$

Using the fact that $|\operatorname{Tr} Ayy^*B| = |y^*BAy| \le \|AB\|\|y\|^2$ together with the norm inequality $\|AB\| \le \|A\|\|B\|$, we obtain

$$
\begin{aligned}
|\boldsymbol{\varepsilon}_3| &\le \frac{\rho^2}{n}\sum_{m=1}^n \|D_j Q D_\ell B\| \tilde{c}_m \\
&\qquad \times \mathbb{E}_{j+1}\left([\tilde{Q}]_{mm}\left|y_m^* Q_m y_m - \frac{1}{n}\operatorname{Tr} D_m \mathbb{E} Q\right| \|y_m\|^2 \|Q_m\|\right) \\
&\overset{\text{(a)}}{\le} \frac{\sigma_{\max}^4}{\rho^3}\frac{1}{n}\sum_{m=1}^n \mathbb{E}_{j+1}\left(\left|y_m^* Q_m y_m - \frac{1}{n}\operatorname{Tr} D_m \mathbb{E} Q\right| \|y_m\|^2\right),
\end{aligned}
$$

where (a) follows from the fact that $\|D_j Q D_\ell B\| \tilde{c}_m \le \sigma_{\max}^4 \rho^{-3}$ and $[\tilde{Q}]_{mm} \|Q_m\| \le \rho^{-2}$. Writing $\frac{1}{n}\operatorname{Tr} D_m \mathbb{E} Q = \frac{1}{n}\operatorname{Tr} D_m Q + \frac{1}{n}\operatorname{Tr} D_m(\mathbb{E}Q - Q)$ and replacing it in the previous inequality, we obtain:

$$
\begin{aligned}
&\mathbb{E}\left(\left|y_m^* Q_m y_m - \frac{1}{n}\operatorname{Tr} D_m \mathbb{E} Q\right| \|y_m\|^2\right) \\
&\quad \le \left(\left[\mathbb{E}\left|y_m^* Q_m y_m - \frac{1}{n}\operatorname{Tr} D_m Q\right|^2\right]^{1/2} + \left[\mathbb{E}\left|\frac{1}{n}\operatorname{Tr} D_m (Q - \mathbb{E}Q)\right|^2\right]^{1/2}\right) \\
&\qquad \times (\mathbb{E}\|y_m\|^4)^{1/2},
\end{aligned}
$$

where $[\mathbb{E}|y_m^* Q_m y_m - \frac{1}{n}\operatorname{Tr} D_m Q|^2]^{1/2} \le K n^{-1/2}$ by Lemma 6.2(1) combined with Lemma 6.3(3), $[\mathbb{E}|\frac{1}{n}\operatorname{Tr} D_m(Q - \mathbb{E}Q)|^2]^{1/2} \le K n^{-1}$ by Lemma 6.3(2c) and $\mathbb{E}\|y_m\|^4 \le \sigma_{\max}^4 \mathbb{E}|X_{11}|^4 (Nn^{-1})^2$. This yields, in particular, $\max_{\ell, j \le n} \mathbb{E}|\boldsymbol{\varepsilon}_3| \le K n^{-1/2}$ and proves (6.27).

Recall that if $m \le j$, then

$$\mathbb{E}_{j+1}(y_m y_m^* Q_m) = \mathbb{E}_{j+1}(y_m y_m^*)\mathbb{E}_{j+1}(Q_m) = \frac{D_m}{n}\mathbb{E}_{j+1}(Q_m).$$

We now turn to equation (6.24) and introduce $\boldsymbol{\chi}_4$, $\boldsymbol{\chi}_5$ and $\boldsymbol{\varepsilon}_3'$.

$$\begin{aligned}
\boldsymbol{\chi}_3' &= \frac{\rho}{n}\operatorname{Tr} D_\ell B\left(\sum_{m=1}^{n} \tilde{c}_m \mathbb{E}_{j+1}(y_m y_m^* Q_m)\right) D_j Q \\
&= \frac{\rho}{n^2}\operatorname{Tr} D_\ell B\left(\sum_{m=1}^{j} \tilde{c}_m D_m \mathbb{E}_{j+1} Q_m\right) D_j Q \\
&\quad + \frac{\rho}{n}\operatorname{Tr} D_\ell B\left(\sum_{m=j+1}^{n} \tilde{c}_m \mathbb{E}_{j+1}(y_m y_m^* Q_m)\right) D_j Q \\
&= \frac{\rho}{n^2}\operatorname{Tr} D_\ell B\left(\sum_{m=1}^{j} \tilde{c}_m D_m \mathbb{E}_{j+1} Q\right) D_j Q \\
&\quad + \frac{\rho}{n}\operatorname{Tr} D_\ell B\left(\sum_{m=j+1}^{n} \tilde{c}_m \mathbb{E}_{j+1}(y_m y_m^* Q_m)\right) D_j Q \\
&\quad + \frac{\rho}{n^2}\operatorname{Tr} D_\ell B\left(\sum_{m=1}^{j} \tilde{c}_m D_m \mathbb{E}_{j+1}(Q_m - Q)\right) D_j Q \stackrel{\triangle}{=} \boldsymbol{\chi}_4 + \boldsymbol{\chi}_5 + \boldsymbol{\varepsilon}_3'
\end{aligned}$$

and decomposition (6.24) is introduced. In order to estimate $\boldsymbol{\varepsilon}_3'$, recall that, given two square matrices R and S, one has $|\operatorname{Tr} RS| \le \|R\| \operatorname{Tr} S$ for S nonnegative and Hermitian. As matrix $Q_m - Q$ is nonnegative and Hermitian by (6.1), we obtain

$$|\boldsymbol{\varepsilon}_3'| \le \frac{1}{n^2}\|D_\ell B D_j Q\| \sum_{m=1}^{j} \mathbb{E}\big(\operatorname{Tr} D_m(Q_m - Q)\big) \le \frac{\sigma_{\max}^6}{n\rho^3} \tag{6.28}$$

by Lemma 6.3(3).

We now turn to $\boldsymbol{\chi}_5$ and provide decomposition (6.25). Recall that $\operatorname{Tr} Ayy^*B = y^*BAy$. Combining (6.1) and (6.2), we get $Q = Q_m - \rho[\tilde{Q}]_{mm} Q_m y_m y_m^* Q_m$. Plugging this expression into the definition of $\boldsymbol{\chi}_5$ and using the fact that y_m is measurable with respect to $\mathcal{F}_{j+1}$ (since $m \ge j+1$), we obtain

$$\boldsymbol{\chi}_5 = \frac{\rho}{n}\operatorname{Tr} D_\ell B\left(\sum_{m=j+1}^{n} \tilde{c}_m \mathbb{E}_{j+1}(y_m y_m^* Q_m)\right) D_j Q$$

$$= \frac{\rho}{n} \sum_{m=j+1}^{n} \tilde{c}_m y_m^* (\mathbb{E}_{j+1} Q_m) D_j Q_m D_\ell B y_m$$
$$- \frac{\rho^2}{n} \sum_{m=j+1}^{n} \tilde{c}_m [\tilde{Q}]_{mm} y_m^* (\mathbb{E}_{j+1} Q_m) D_j Q_m y_m y_m^* Q_m D_\ell B y_m.$$

In order to understand the forthcoming decomposition, recall that, asymptotically, $y_m^* A_m y_m \sim \frac{1}{n} \operatorname{Tr} D_m A_m$ as long as y_m and A_m are independent, and that $\frac{1}{n} \operatorname{Tr} D_m A_m \sim \frac{1}{n} \operatorname{Tr} D_m A$ if A_m is a rank one perturbation of A. We can now introduce $\boldsymbol{\chi}_6$ and $\boldsymbol{\chi}_7$:

$$\begin{aligned}
\boldsymbol{\chi}_5 &= \frac{\rho}{n} \sum_{m=j+1}^{n} \frac{\tilde{c}_m}{n} \operatorname{Tr} D_\ell B D_m (\mathbb{E}_{j+1} Q) D_j Q \\
&\quad - \frac{\rho^2}{n} \sum_{m=j+1}^{n} \frac{\tilde{c}_m^2}{n} \operatorname{Tr} D_m (\mathbb{E}_{j+1} Q) D_j Q \times \frac{1}{n} \operatorname{Tr}(D_m Q D_\ell B) + \boldsymbol{\varepsilon}_6 - \boldsymbol{\varepsilon}_7 \\
&\stackrel{\triangle}{=} \boldsymbol{\chi}_6 - \boldsymbol{\chi}_7 + \boldsymbol{\varepsilon}_6 - \boldsymbol{\varepsilon}_7,
\end{aligned}$$

where

$$\begin{aligned}
\boldsymbol{\varepsilon}_6 &= \frac{\rho}{n} \sum_{m=j+1}^{n} \tilde{c}_m y_m^* (\mathbb{E}_{j+1} Q_m) D_j Q_m D_\ell B y_m \\
&\quad - \frac{\rho}{n} \sum_{m=j+1}^{n} \frac{\tilde{c}_m}{n} \operatorname{Tr} D_\ell B D_m (\mathbb{E}_{j+1} Q) D_j Q, \\
\boldsymbol{\varepsilon}_7 &= \frac{\rho^2}{n} \sum_{m=j+1}^{n} \tilde{c}_m [\tilde{Q}]_{mm} y_m^* (\mathbb{E}_{j+1} Q_m) D_j Q_m y_m y_m^* Q_m D_\ell B y_m \\
&\quad - \frac{\rho^2}{n} \sum_{m=j+1}^{n} \frac{\tilde{c}_m^2}{n} \operatorname{Tr} D_m (\mathbb{E}_{j+1} Q) D_j Q \times \frac{1}{n} \operatorname{Tr}(D_m Q D_\ell B).
\end{aligned}$$

It is now a matter of routine to check that

$$\mathbb{E}|\boldsymbol{\varepsilon}_6| \le \frac{K}{\sqrt{n}} \quad \text{and} \quad \mathbb{E}|\boldsymbol{\varepsilon}_7| \le \frac{K}{\sqrt{n}}. \tag{6.29}$$

Let us provide some details.

Recall that y_m is independent of $\mathbb{E}_{j+1}(Q_m)$. To obtain the bound on $\mathbb{E}|\boldsymbol{\varepsilon}_6|$, we use the facts that $\mathbb{E}(y_m^* (\mathbb{E}_{j+1} Q_m) D_j Q_m D_\ell B y_m - \frac{1}{n} \operatorname{Tr} D_\ell B D_m (\mathbb{E}_{j+1} Q_m) D_j \times Q_m)^2 \le K n^{-1}$ by Lemma 6.2(1), $|\frac{1}{n} \operatorname{Tr} D_\ell B D_m (\mathbb{E}_{j+1} Q_m) D_j (Q_m - Q)| \le K n^{-1}$ by Lemma 6.3(3), etc.

In order to prove that $\mathbb{E}|\boldsymbol{\varepsilon}_7| \le K n^{-1/2}$, we use similar arguments but we also need two additional estimates. The control $[\tilde{Q}]_{mm} - \tilde{c}_m$ (which has already been done while estimating $\boldsymbol{\varepsilon}_3$) relies on (6.8). The bounded character of $\mathbb{E}(y_m^* A_m y_m)^2$, where A_m is independent of y_m and of finite spectral norm is a by-product of Lemma 6.2(1).

We now put the pieces together and provide (6.26) satisfied by $\boldsymbol{\chi}_{\boldsymbol{\ell}\mathbf{j}}$. Recall that

$$\begin{aligned}
\boldsymbol{\chi}_{\mathbf{1}} &= \frac{1}{n} \operatorname{Tr} D_\ell B D_j Q, \\
\boldsymbol{\chi}_{\mathbf{2}} &= \frac{\rho}{n} \operatorname{Tr} D_\ell B \operatorname{diag}\left(\frac{1}{n} \operatorname{Tr} \tilde{D}_i \tilde{C}\right) (\mathbb{E}_{j+1} Q) D_j Q, \\
\boldsymbol{\chi}_{\mathbf{4}} &= \frac{\rho}{n^2} \operatorname{Tr} D_\ell B \left(\sum_{m \le j} \tilde{c}_m D_m\right) (\mathbb{E}_{j+1} Q) D_j Q, \\
\boldsymbol{\chi}_{\mathbf{6}} &= \frac{\rho}{n^2} \operatorname{Tr} D_\ell B \left(\sum_{m=j+1}^{n} \tilde{c}_m D_m\right) (\mathbb{E}_{j+1} Q) D_j Q, \\
\boldsymbol{\chi}_{\mathbf{7}} &= \frac{\rho^2}{n} \sum_{m=j+1}^{n} \frac{\tilde{c}_m^2}{n} \operatorname{Tr} D_m (\mathbb{E}_{j+1} Q) D_j Q \times \frac{1}{n} \operatorname{Tr}(D_m Q D_\ell B) \\
&= \frac{1}{n} \sum_{m=j+1}^{n} \frac{(1/n) \operatorname{Tr}(D_\ell B D_m Q)}{(1 + (1/n) \operatorname{Tr} D_m \mathbb{E} Q)^2} \boldsymbol{\chi}_{m,j}.
\end{aligned}$$

As $\frac{1}{n} \sum_{m=1}^{n} \tilde{c}_m D_m = \operatorname{diag}(\frac{1}{n} \operatorname{Tr} \tilde{D}_1 \tilde{C}, \dots, \frac{1}{n} \operatorname{Tr} \tilde{D}_N \tilde{C})$, we have $\boldsymbol{\chi}_{\mathbf{2}} - \boldsymbol{\chi}_{\mathbf{4}} - \boldsymbol{\chi}_{\mathbf{6}} = 0$ and (6.26) becomes

$$\boldsymbol{\chi}_{\boldsymbol{\ell},\mathbf{n}}^{(\mathbf{j})} = \frac{1}{n} \operatorname{Tr} D_\ell B D_j Q + \frac{1}{n} \sum_{m=j+1}^{n} \frac{(1/n) \operatorname{Tr}(D_\ell B D_m Q)}{(1 + (1/n) \operatorname{Tr} D_m \mathbb{E} Q)^2} \boldsymbol{\chi}_{\mathbf{m},\mathbf{n}}^{(\mathbf{j})} + \boldsymbol{\varepsilon},$$

where $\mathbb{E}|\boldsymbol{\varepsilon}| \le K n^{-1/2}$, by inequalities (6.27), (6.28) and (6.29). Small adjustments need to be made in order to obtain (6.17). Now, replace $\frac{1}{n} \operatorname{Tr} D_\ell B D_p Q$ by $\frac{1}{n} \operatorname{Tr} D_\ell B D_p \mathbb{E} Q$ [use Lemma 6.3(2c)]. The new error term $\boldsymbol{\varepsilon}_{\mathbf{3},\boldsymbol{\ell}\mathbf{j}}$ still satisfies $\max_{\ell, j \le n} \mathbb{E}|\boldsymbol{\varepsilon}_{\mathbf{3},\boldsymbol{\ell},\mathbf{j}}| \le K n^{-1/2}$. Equation (6.17) is thus proved. $\square$

PROOF OF (6.18). Recall that $\boldsymbol{\chi}_{\boldsymbol{\ell},\mathbf{j}}$ and $\boldsymbol{\varepsilon}_{\mathbf{3},\boldsymbol{\ell}\mathbf{j}}$ have been introduced above.

Following the matrix framework introduced to express the system satisfied by the $\mathbf{y}$'s (matrices A_n, $A_n^{(j)}$ and $A_n^{0,(j)}$), we introduce the matrix $G_n = A_n^T = (g_{\ell m})_{\ell,m=1}^n$, its $(n-j+1) \times (n-j+1)$ principal submatrix $G_n^{(j)} = (g_{\ell,m})_{\ell,m=j}^n$ and the matrix $G_n^{0,(j)}$ which differs from matrix $G_n^{(j)}$ by its first column, equal to zero. Writing $\boldsymbol{\delta}_{\mathbf{n}}^{(\mathbf{j})} = (\frac{1}{n} \operatorname{Tr} D_\ell D_j T^2; j \le \ell \le n)$, we have

$$\mathbf{x}_{\mathbf{n}}^{(\mathbf{j})} = G_n^{0,(j)} \mathbf{x}_{\mathbf{n}}^{(\mathbf{j})} + \boldsymbol{\delta}_{\mathbf{n}}^{(\mathbf{j})}.$$

Now, define the $(n-j+1)\times 1$ vector $\boldsymbol{\varepsilon}_{\mathbf{3}}^{(\mathbf{j})} = (\boldsymbol{\varepsilon}_{\mathbf{3},\boldsymbol{\ell}\mathbf{j}};\, j\le \ell\le n)$ and the $(n-j+1)\times 1$ vectors:

$$\boldsymbol{\chi}^{(\mathbf{j})} = (\boldsymbol{\chi}_{\boldsymbol{\ell},\mathbf{n}}^{(\mathbf{j})};\, j\le\ell\le n);$$

$$\breve{\boldsymbol{\delta}}^{(\mathbf{j})} = \left(\frac{1}{n}\operatorname{Tr} D_\ell B D_j \mathbb{E}Q;\, j\le\ell\le n\right).$$

Now, define the $(n-j+1)\times(n-j+1)$ matrix

$$\breve{G}^{(j)} = \left(\frac{(1/n^2)\operatorname{Tr} D_\ell B D_m \mathbb{E}Q}{(1+(1/n)\operatorname{Tr} D_m\mathbb{E}Q)^2}\right)_{\ell,m=j}^{n}$$

and $\breve{G}^{0,(j)}$, which is equal to $\breve{G}^{(j)}$ except for its first column, equal to zero. With this notation, equation (6.17), valid for $j\le\ell\le n$, can take the following matrix form:

$$\boldsymbol{\chi}^{(\mathbf{j})} = \breve{G}^{0,(j)}\boldsymbol{\chi}^{(\mathbf{j})} + \breve{\boldsymbol{\delta}}^{(\mathbf{j})} + \boldsymbol{\varepsilon}_{\mathbf{3}}^{(\mathbf{j})}.$$

We will heavily rely on the inequality

$$\limsup_n \left\|\left(I-G^{0,(j)}\right)^{-1}\right\|_\infty < \infty,$$

which can be proven as in Lemma 5.2(4) and Lemma 5.5. We drop the superscript $^{0,(j)}$ in the equation below for the sake of readability.

$$
\begin{aligned}
\boldsymbol{\chi} &= \breve{G}\boldsymbol{\chi} + \breve{\boldsymbol{\delta}} + \boldsymbol{\varepsilon}_{\mathbf{3}} \\
&\Leftrightarrow \quad \boldsymbol{\chi} = G\boldsymbol{\chi} + \boldsymbol{\delta} + \boldsymbol{\varepsilon}_{\mathbf{3}} + (\breve{G}-G)\boldsymbol{\chi} + (\breve{\boldsymbol{\delta}}-\boldsymbol{\delta}), \\
&\Leftrightarrow \quad \boldsymbol{\chi} = (I-G)^{-1}\boldsymbol{\delta} + (I-G)^{-1}\boldsymbol{\varepsilon}_{\mathbf{3}} \\
&\qquad\qquad + (I-G)^{-1}(\breve{G}-G)\boldsymbol{\chi} + (I-G)^{-1}(\breve{\boldsymbol{\delta}}-\boldsymbol{\delta}), \\
&\Leftrightarrow \quad \boldsymbol{\chi} = \mathbf{x} + (I-G)^{-1}\boldsymbol{\varepsilon}_{\mathbf{3}} \\
&\qquad\qquad + (I-G)^{-1}(\breve{G}-G)\boldsymbol{\chi} + (I-G)^{-1}(\breve{\boldsymbol{\delta}}-\boldsymbol{\delta}).
\end{aligned}
\tag{6.30}
$$

The first component of the previous equation can be written as

$$
\begin{aligned}
\boldsymbol{\chi}_{\mathbf{j}}^{(\mathbf{j})} &= \mathbf{x}_{\mathbf{j}}^{(\mathbf{j})} + \left[\left(I-G^{0,(j)}\right)^{-1}\boldsymbol{\varepsilon}_{\mathbf{3}}\right]_1 \\
&\quad + \left[\left(I-G^{0,(j)}\right)^{-1}\left(\breve{G}^{0,(j)}-G^{0,(j)}\right)\boldsymbol{\chi} + \left(I-G^{0,(j)}\right)^{-1}(\breve{\boldsymbol{\delta}}-\boldsymbol{\delta})\right]_1 \\
&\stackrel{\triangle}{=} \mathbf{x}_{\mathbf{j}}^{(\mathbf{j})} + \boldsymbol{\varepsilon}_{\mathbf{41},\mathbf{j}} + \boldsymbol{\varepsilon}_{\mathbf{42},\mathbf{j}}.
\end{aligned}
$$

Due to Lemma 5.2(4), which applies to $G^{0,(j)}$ and to the fact that $\max_{\ell,j\le n}\mathbb{E}|\boldsymbol{\varepsilon}_{\mathbf{3},\boldsymbol{\ell}\mathbf{j}}| \le Kn^{-1/2}$, we have

$$\mathbb{E}|\boldsymbol{\varepsilon}_{\mathbf{41},\mathbf{j}}| \le \sum_{m=1}^{n-j+1}\left[\left(I-G^{0,(j)}\right)^{-1}\right]_{1,m}\mathbb{E}|\boldsymbol{\varepsilon}_{\mathbf{3},\boldsymbol{\ell}\mathbf{j}}| \le \frac{K}{\sqrt{n}}.$$

The second error term $\boldsymbol{\varepsilon}_{\mathbf{42,j}}$ is the sum

$$\boldsymbol{\varepsilon}_{\mathbf{42,j}} = [(I - G^{0,(j)})^{-1}(\breve{G}^{0,(j)} - G^{0,(j)})\boldsymbol{\chi}]_1 + [(I - G^{0,(j)})^{-1}(\breve{\boldsymbol{\delta}} - \boldsymbol{\delta})]_1.$$

Let us first prove that $[(I - G^{0,(j)})^{-1}(\breve{G}^{0,(j)} - G^{0,(j)})\boldsymbol{\chi}]_1$ is dominated by a sequence independent of j that converges to zero as $n \to \infty$. The mere definition of $\boldsymbol{\chi}^{(\mathbf{j})}_{\boldsymbol{\ell},\mathbf{n}}$ [see (6.16)] yields $\|\boldsymbol{\chi}^{(\mathbf{j})}\|_\infty \le (N\sigma^4_{\max})(n\rho^2)^{-1}$, where $\|\cdot\|_\infty$ stands for the ℓ_∞-norm. Hence,

$$\begin{aligned}
&|[(I - G^{0,(j)})^{-1}(\breve{G}^{0,(j)} - G^{0,(j)})\boldsymbol{\chi}]_1| \\
&\quad\le |||(I - G^{0,(j)})^{-1}|||_\infty |||(\breve{G^{0,(j)}} - G^{0,(j)})^T|||_\infty \|\boldsymbol{\chi}\|_\infty \\
&\quad\le K|||(\breve{G}^{0,(j)} - G^{0,(j)})^T|||_\infty.
\end{aligned}$$

Let us prove that

$$|||(\breve{G}^{0,(j)} - G^{0,(j)})^T|||_\infty \xrightarrow[n\to\infty]{} 0 \tag{6.31}$$

uniformly in j. To this end, let us evaluate the (ℓ, m)-element of matrix $\breve{G}^{0,(j)} - G^{0,(j)}$ $(m > j)$:

$$\begin{aligned}
&n|[\breve{G}^{0,(j)} - G^{0,(j)}]_{\ell m}| \\
&\quad= \left|\frac{(1/n)\operatorname{Tr} D_\ell B D_m \mathbb{E}Q}{(1 + (1/n)\operatorname{Tr} D_m \mathbb{E}Q)^2} - \frac{(1/n)\operatorname{Tr} D_\ell D_m T^2}{(1 + (1/n)\operatorname{Tr} D_m T)^2}\right| \\
&\quad\le \left|\left(1 + \frac{1}{n}\operatorname{Tr} D_m T\right)^2 \frac{1}{n}\operatorname{Tr} D_\ell B D_m \mathbb{E}Q\right. \\
&\qquad\quad \left. - \left(1 + \frac{1}{n}\operatorname{Tr} D_m \mathbb{E}Q\right)^2 \frac{1}{n}\operatorname{Tr} D_\ell D_m T^2\right| \qquad (6.32) \\
&\quad\le \left|\left(1 + \frac{1}{n}\operatorname{Tr} D_m T\right)^2 \frac{1}{n}\operatorname{Tr} D_\ell B D_m(\mathbb{E}Q - T)\right| \\
&\qquad + \left|\left(1 + \frac{1}{n}\operatorname{Tr} D_m T\right)^2 \frac{1}{n}\operatorname{Tr} D_\ell T D_m(B - T)\right| \\
&\qquad + \left|\left(\left(1 + \frac{1}{n}\operatorname{Tr} D_m T\right)^2 - \left(1 + \frac{1}{n}\operatorname{Tr} D_m \mathbb{E}Q\right)^2\right)\frac{1}{n}\operatorname{Tr} D_\ell D_m T^2\right|.
\end{aligned}$$

The first term of the right-hand side of (6.32) satisfies

$$\left|\left(1 + \frac{1}{n}\operatorname{Tr} D_m T\right)^2 \frac{1}{n}\operatorname{Tr} D_\ell B D_m(\mathbb{E}Q - T)\right| \le \left(1 + \frac{\sigma^2_{\max}}{\rho}\right)^2 \frac{1}{n}\operatorname{Tr} U(\mathbb{E}Q - T),$$

where U is the $N \times N$ diagonal matrix $U = \sigma^4_{\max}\rho^{-1}\operatorname{diag}(\operatorname{sign}(\mathbb{E}[Q]_{ii} - t_i), 1 \le i \le N)$. By Lemma 6.3(2a), the right-hand side of this inequality converges to zero as $n \to \infty$.

The second and third terms of the right-hand side of (6.32) can be handled similarly with the help of Lemma 6.3 and one can prove that elements of $n(\breve{G}^{0,(j)} - G^{0,(j)})$ are dominated by a sequence independent of j that converges to zero. This implies that $|||(\breve{G}^{0,(j)} - G^{0,(j)})^T|||_\infty$ converges to zero uniformly in j and (6.31) is proved. As a consequence, $[(I - G^{0,(j)})^{-1}(\breve{G}^{0,(j)} - G^{0,(j)})\chi]_1$ is dominated by a sequence independent of j that converges to zero. The other term, $[(I - G^{0,(j)})^{-1}(\breve{\boldsymbol{\delta}} - \boldsymbol{\delta})]_1$, in the expression of $\boldsymbol{\varepsilon}_{\mathbf{42,j}}$ is handled similarly. Equation (6.18) is thus proved. □

PROOF OF (6.19). We rewrite equation (6.15) as $\mathbb{E}_{j+1}(\mathbb{E}_j\Gamma_j)^2 = \eta_{1,j} + \kappa\eta_{2,j} + \boldsymbol{\varepsilon}_{\mathbf{2,j}}$, with

$$\eta_{1,j} = \frac{1}{n^2}\frac{1}{(1+(1/n)\operatorname{Tr} D_j\mathbb{E}Q)^2}\operatorname{Tr} D_j(\mathbb{E}_{j+1}Q_j)D_j(\mathbb{E}_{j+1}Q_j),$$

$$\eta_{2,j} = \frac{1}{n^2(1+(1/n)\operatorname{Tr} D_j\mathbb{E}Q)^2}\sum_{i=1}^{N}\sigma_{ij}^4(\mathbb{E}_{j+1}[Q_j]_{ii})^2$$

and we prove that $\sum_{j=1}^n \eta_{1,j} - \tilde{\mathcal{V}}_n \xrightarrow{\mathcal{P}} 0$ and $\sum_{j=1}^n \eta_{2,j} - \mathcal{W}_n \xrightarrow{\mathcal{P}} 0$, where $\tilde{\mathcal{V}}_n$ and $\mathcal{W}_n$ are defined in Section 5. To prove the first assertion, we first notice that

$$\begin{aligned}\operatorname{Tr} D_j(\mathbb{E}_{j+1}Q_j)D_j(\mathbb{E}_{j+1}Q_j) &= \mathbb{E}_{j+1}(\operatorname{Tr} D_j(\mathbb{E}_{j+1}Q_j)D_jQ_j)\\ &= \mathbb{E}_{j+1}(\operatorname{Tr} D_j(\mathbb{E}_{j+1}Q)D_jQ) + \boldsymbol{\varepsilon},\end{aligned}$$

with $|\boldsymbol{\varepsilon}| \le 2\sigma_{\max}^4\rho^{-2}$, by Lemma 6.3(3). Therefore,

$$\eta_{1,j} = \frac{\mathbb{E}_{j+1}\boldsymbol{\chi}_{\mathbf{j,n}}^{(\mathbf{j})}}{(1+(1/n)\operatorname{Tr} D_j\mathbb{E}Q)^2} + \frac{\boldsymbol{\varepsilon}}{(1+(1/n)\operatorname{Tr} D_j\mathbb{E}Q)^2}.$$

It remains to control the difference $(1+\frac{1}{n}\operatorname{Tr} D_j\mathbb{E}Q)^{-2} - (1+\frac{1}{n}\operatorname{Tr} D_jT)^{-2}$, to use (6.18) and one easily obtains $\sum_{j=1}^n \eta_{1,j} - \tilde{\mathcal{V}}_n \xrightarrow{\mathcal{P}} 0$.

We now sketch the proof of $\sum_{j=1}^n \eta_{2,j} - \mathcal{W}_n \xrightarrow{\mathcal{P}} 0$. As in (6.2), $[Q_j]_{ii}$ satisfies $[Q_j]_{ii} = -(z(1+\xi_i^j\tilde{Q}_i^j\xi_i^{j*}))^{-1}$, where ξ_i^j is the row ξ_i without element j, and $\tilde{Q}_i^j = (Y_i^{j*}Y_i^j + \rho I_{n-1})^{-1}$, where Y_i^j is the matrix Y without row i and column j. Using this identity and Lemmas 6.2(1) and 6.3, we can show that $[Q_j]_{ii}$ is approximated by t_i, which is key to proving $\sum_{j=1}^n \eta_{2,j} - \mathcal{W}_n \xrightarrow{\mathcal{P}} 0$. □

7. Proof of Theorem 3.3. We first provide an expression of the bias that involves the Stieltjes transforms $\frac{1}{N}\operatorname{Tr} Q$ and $\frac{1}{N}\operatorname{Tr} T$. By writing $\log\det(Y_nY_n^* + \rho I_N) = N\log\rho + \log\det(\frac{1}{\rho}Y_nY_n^* + I_N)$ and taking the derivative of $\log\det(\frac{1}{\rho}Y_n \times$

$Y_n^* + I_N)$ with respect to ρ, we obtain

$$\log\det(Y_n Y_n^* + \rho I_N) = N \log \rho + N \int_\rho^\infty \left(\frac{1}{\omega} - \frac{1}{N} \operatorname{Tr} Q(-\omega) \right) d\omega.$$

Since $\frac{1}{N}\operatorname{Tr} Q(z) \in \mathcal{S}(\mathbb{R}^+)$, we have $\frac{1}{\omega} - \frac{1}{N}\operatorname{Tr} Q(-\omega) \geq 0$ for $\omega > 0$. In fact, recall that $\|Q(-\omega)\| \leq \omega^{-1}$, by Proposition 2.2. Therefore, by Fubini's theorem,

$$\mathbb{E}\log\det(Y_n Y_n^* + \rho I_N) = N \log \rho + N \int_\rho^\infty \left(\frac{1}{\omega} - \frac{1}{N} \operatorname{Tr} \mathbb{E} Q(-\omega) \right) d\omega.$$

Similarly,

$$N V_n(\rho) = N \int \log(\lambda + \rho) \pi_n(d\lambda) = N \log \rho + N \int_\rho^\infty \left(\frac{1}{\omega} - \frac{1}{N} \operatorname{Tr} T(-\omega) \right) d\omega.$$

Hence, the bias term is given by

$$\begin{aligned} \mathcal{B}_n(\rho) &\stackrel{\triangle}{=} \mathbb{E}\log\det(Y_n Y_n^* + \rho I_N) - N V_n(\rho) \\ &= \int_\rho^\infty \operatorname{Tr}\big(T(-\omega) - \mathbb{E} Q(-\omega)\big)\, d\omega. \end{aligned} \tag{7.1}$$

In Appendix B, we prove that

$$\operatorname{Tr}(T - Q) = \operatorname{Tr}(\tilde{T} - \tilde{Q}). \tag{7.2}$$

Therefore, we can also write the bias as

$$\mathcal{B}_n(\rho) = \int_\rho^\infty \operatorname{Tr}\big(\tilde{T}(-\omega) - \mathbb{E}\tilde{Q}(-\omega)\big)\, d\omega. \tag{7.3}$$

For technical reasons (and in order to rely on results of Section 5), we use representation (7.3) of the bias instead of (7.1). The proof of Theorem 3.3 will rely on the study of the asymptotic behavior of the integrand in the right-hand side of this equation.

As a by-product of Section 5, the existence and uniqueness of the solution of the system of equations (3.1) is straightforward. Indeed, define the $n \times 1$ vectors $\mathbf{w}$ and $\mathbf{p}$ as

$$\mathbf{w} = (\mathbf{w}_{j,n};\, 1 \leq j \leq n),$$
$$\mathbf{p} = (\mathbf{p}_{j,n};\, 1 \leq j \leq n).$$

The system (3.1) can then be written in matrix form as

$$\mathbf{w} = A\mathbf{w} + \mathbf{p}. \tag{7.4}$$

Since $(I - A)$ is invertible for n large enough, this proves Theorem 3.3(1).

The rest of the proof will be carried out into four steps:

1. We first introduce a perturbed version of the system (7.4). For the reader's convenience, we recall the following notation:

$$t_i = \frac{1}{\omega(1+(1/n)\operatorname{Tr}\tilde{D}_i\tilde{T})}; \qquad \tilde{t}_j = \frac{1}{\omega(1+(1/n)\operatorname{Tr}D_jT)};$$
$$c_i = \frac{1}{\omega(1+(1/n)\operatorname{Tr}\tilde{D}_i\mathbb{E}\tilde{Q})}; \qquad \tilde{c}_j = \frac{1}{\omega(1+(1/n)\operatorname{Tr}D_j\mathbb{E}Q)};$$
$$b_i = \frac{1}{\omega(1+(1/n)\operatorname{Tr}\tilde{D}_i\tilde{C})}; \qquad \tilde{b}_j = \frac{1}{\omega(1+(1/n)\operatorname{Tr}D_jC)},$$

where z is equal to $-\omega$ with $\omega \ge 0$. Write the integrand in (7.3) as

$$\operatorname{Tr}\bigl(\tilde{T}(-\omega) - \mathbb{E}\tilde{Q}(-\omega)\bigr) = \frac{1}{n}\sum_{j=1}^{n}\boldsymbol{\varphi}_j(\omega), \tag{7.5}$$
$$\text{with } \boldsymbol{\varphi}_j(\omega) \stackrel{\triangle}{=} n\bigl(\tilde{t}_j(-\omega) - \mathbb{E}[\tilde{Q}(-\omega)]_{jj}\bigr).$$

Let $\boldsymbol{\psi}^{(j)}(\omega) \stackrel{\triangle}{=} n(\tilde{b}_j(-\omega) - \mathbb{E}[\tilde{Q}(-\omega)]_{jj})$ and define the $n\times 1$ vectors $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ and the $n\times n$ matrix $\breve{A}$ as

$$\boldsymbol{\varphi} = (\boldsymbol{\varphi}_j;\, 1\le j\le n),$$
$$\boldsymbol{\psi} = (\boldsymbol{\psi}^{(j)};\, 1\le j\le n),$$
$$\breve{A} = \left(\frac{(1/n^2)\operatorname{Tr}D_jD_mCT}{(1+(1/n)\operatorname{Tr}D_jT)(1+(1/n)\operatorname{Tr}D_jC)}\right)_{j,m=1}^{n}.$$

We first prove that

$$\boldsymbol{\varphi} = \breve{A}\boldsymbol{\varphi} + \boldsymbol{\psi}. \tag{7.6}$$

2. We prove that

$$\boldsymbol{\psi}^{(j)} = \kappa\omega^2\tilde{b}_j\tilde{c}_j\left(\frac{\omega}{n}\sum_{i=1}^{N}\left(\sigma_{ij}^2c_i^3\frac{1}{n}\sum_{m=1}^{n}\sigma_{im}^4\mathbb{E}[\tilde{Q}_i]_{mm}^2\right) - \frac{\tilde{c}_j}{n}\sum_{i=1}^{N}\sigma_{ij}^4\mathbb{E}[Q_j]_{ii}^2\right) + \boldsymbol{\varepsilon}^{(j)}, \tag{7.7}$$

with $|\boldsymbol{\varepsilon}^{(j)}| \le Kn^{-1/2}$, where K is a constant that does not depend on n nor or j (but may depend on ω).

3. Matrix $\breve{A}$ readily approximates A and vector $\boldsymbol{\psi}$ approximates $\mathbf{p}$ for large n, by step 2. Therefore, by inspecting equations (7.4) and (7.6), one may expect $\boldsymbol{\varphi}$ to be close to $\mathbf{w}$. We prove here that

$$\|\boldsymbol{\varphi} - \mathbf{w}\|_\infty \underset{n\to\infty, N/n\to c}{\longrightarrow} 0. \tag{7.8}$$

4. Let $\beta_n(\omega) = \frac{1}{n}\sum_{j=1}^n \mathbf{w}_{j,n}(\omega)$. Equation (7.8) yields $\frac{1}{n}\sum_{j=1}^n \boldsymbol{\varphi}_j(\omega) - \beta_n(\omega) \to 0$. In order to prove (3.5), it remains to integrate and to provide a dominated convergence theorem argument. To this end, we shall prove that

$$|\beta_n(\omega)| \le \frac{K'}{\omega^3} \tag{7.9}$$

for n large enough. This will establish (3.4). We will also prove that

$$\left|\frac{1}{n}\sum_{j=1}^n \boldsymbol{\varphi}_j(\omega)\right| \le \frac{K'}{\omega^2} \tag{7.10}$$

for $\omega \in [\rho, +\infty)$, where K' does not depend on n or ω. This will yield (3.5) and the proof of Theorem 3.3 will be completed.

7.1. *Proof of step* 1: *equation* (7.6). Recall that $\boldsymbol{\psi}^{(j)} = n(\tilde{b}_j - \mathbb{E}[\tilde{Q}]_{jj})$. Using these expressions, we have, for $1 \le j \le n$,

$$\begin{aligned}
\boldsymbol{\varphi}_j &= n(\tilde{t}_j - \tilde{b}_j) + \boldsymbol{\psi}^{(j)} = n\tilde{b}_j\tilde{t}_j(\tilde{b}_j^{-1} - \tilde{t}_j^{-1}) + \boldsymbol{\psi}^{(j)} \\
&= \omega\tilde{b}_j\tilde{t}_j \operatorname{Tr} D_j(C - T) + \boldsymbol{\psi}^{(j)} \\
&= \omega\tilde{b}_j\tilde{t}_j \sum_{i=1}^N \sigma_{ij}^2 c_i t_i (t_i^{-1} - c_i^{-1}) + \boldsymbol{\psi}^{(j)} \\
&= \frac{\omega^2\tilde{b}_j\tilde{t}_j}{n^2} \sum_{i=1}^N \sum_{m=1}^n \sigma_{ij}^2\sigma_{im}^2 c_i t_i \boldsymbol{\varphi}_m + \boldsymbol{\psi}^{(j)} \\
&= \omega^2\tilde{b}_j\tilde{t}_j \sum_{m=1}^n \frac{1}{n^2}\operatorname{Tr}(D_j D_m C T)\boldsymbol{\varphi}_m + \boldsymbol{\psi}^{(j)},
\end{aligned}$$

which yields equation (7.6).

7.2. *Proof of step* 2: *Expression for* $\boldsymbol{\psi}^{(j)}$. We shall develop $\boldsymbol{\psi}^{(j)}$ as

$$\boldsymbol{\psi}^{(j)} = \boldsymbol{\psi}_1 + \boldsymbol{\psi}_2 - \boldsymbol{\psi}_3, \tag{7.11}$$

$$\boldsymbol{\psi}_1 = \boldsymbol{\psi}_4 + \boldsymbol{\varepsilon}_1, \tag{7.12}$$

$$\boldsymbol{\psi}_2 = -\boldsymbol{\psi}_5 + \boldsymbol{\psi}_6, \tag{7.13}$$

$$\boldsymbol{\psi}_5 = \boldsymbol{\psi}_7 + \boldsymbol{\varepsilon}_5, \tag{7.14}$$

$$\boldsymbol{\psi}_6 = \boldsymbol{\psi}_8 + \boldsymbol{\varepsilon}_6, \tag{7.15}$$

$$\boldsymbol{\psi}_3 = \boldsymbol{\psi}_9 + \boldsymbol{\varepsilon}_3, \tag{7.16}$$

where the $\boldsymbol{\psi}_k$'s and the $\boldsymbol{\varepsilon}_k$'s will be introduced when required. We shall furthermore prove that $|\boldsymbol{\varepsilon}_k| \le K n^{-1/2}$ for $k = 1, 3, 5, 6$. This will yield

$$\begin{aligned} &\boldsymbol{\psi}^{(j)} = \boldsymbol{\psi}_4 - \boldsymbol{\psi}_7 + \boldsymbol{\psi}_8 - \boldsymbol{\psi}_9 + \boldsymbol{\varepsilon}^{(j)} \\ &\qquad\qquad \text{with } |\boldsymbol{\varepsilon}^{(j)}| = |\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_3 - \boldsymbol{\varepsilon}_5 + \boldsymbol{\varepsilon}_6| \le \frac{K}{n^{1/2}}. \end{aligned} \tag{7.17}$$

Let us begin with decomposition (7.11):

$$\begin{aligned} \boldsymbol{\psi}^{(j)} &= n\tilde{b}_j \mathbb{E}([\tilde{Q}]_{jj}([\tilde{Q}]_{jj}^{-1} - \tilde{b}_j^{-1})) \\ &\overset{\text{(a)}}{=} n\omega\tilde{b}_j \mathbb{E}\Big([\tilde{Q}]_{jj}\Big(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j C\Big)\Big) \\ &\overset{\text{(b)}}{=} n\omega\tilde{b}_j\tilde{c}_j \mathbb{E}\Big(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j C\Big) \\ &\qquad - n\omega^2\tilde{b}_j\tilde{c}_j \mathbb{E}\Big([\tilde{Q}]_{jj}\Big(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E}Q\Big)\Big(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j C\Big)\Big) \\ &\overset{\text{(c)}}{=} \omega\tilde{b}_j\tilde{c}_j \operatorname{Tr} D_j \mathbb{E}(Q_j - Q) + \omega\tilde{b}_j\tilde{c}_j \operatorname{Tr} D_j(\mathbb{E}Q - C) \\ &\qquad - n\omega^2\tilde{b}_j\tilde{c}_j \mathbb{E}\Big([\tilde{Q}]_{jj}\Big(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E}Q\Big)\Big(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j C\Big)\Big) \\ &\overset{\triangle}{=} \boldsymbol{\psi}_1 + \boldsymbol{\psi}_2 - \boldsymbol{\psi}_3, \end{aligned}$$

where (a) follows from (6.2) and the definition of $\tilde{b}_j$, (b) follows from identity (6.8) and (c) follows from the equality

$$\begin{aligned} \mathbb{E}\Big(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j C\Big) &= \frac{1}{n}\operatorname{Tr} D_j(\mathbb{E}Q_j - C) \\ &= \frac{1}{n}\operatorname{Tr} D_j \mathbb{E}(Q_j - Q) + \frac{1}{n}\operatorname{Tr} D_j(\mathbb{E}Q - C). \end{aligned}$$

Equation (7.11) is thus established.

We now turn to the decomposition (7.12). Combining (6.1) and (6.2), we obtain $Q = Q_j - \omega[\tilde{Q}]_{jj} Q_j y_j y_j^* Q_j$, hence $\boldsymbol{\psi}_1 = \omega^2 \tilde{b}_j \tilde{c}_j \mathbb{E}([\tilde{Q}]_{jj} y_j^* Q_j D_j Q_j y_j)$. Using identity (6.8) and the fact that $\mathbb{E}(y_j^* Q_j D_j Q_j y_j) = \frac{1}{n}\mathbb{E}(\operatorname{Tr} D_j Q_j D_j Q_j)$, we obtain

$$\begin{aligned} \boldsymbol{\psi}_1 &= \frac{\omega^2}{n}\tilde{b}_j\tilde{c}_j^2 \mathbb{E}(\operatorname{Tr} D_j Q_j D_j Q_j) \\ &\qquad - \omega^3\tilde{b}_j\tilde{c}_j^2 \mathbb{E}\Big([\tilde{Q}]_{jj}\Big(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E}Q\Big)(y_j^* Q_j D_j Q_j y_j)\Big) \\ &\overset{\triangle}{=} \boldsymbol{\psi}_4 + \boldsymbol{\varepsilon}_1. \end{aligned}$$

We have

$$|\boldsymbol{\varepsilon}_1| \leq \frac{1}{\omega}\mathbb{E}(y_j^* Q_j D_j Q_j y_j |\boldsymbol{\varepsilon}_{11} + \boldsymbol{\varepsilon}_{12} + \boldsymbol{\varepsilon}_{13}|),$$

with $\boldsymbol{\varepsilon}_{11} = y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j Q_j$, $\boldsymbol{\varepsilon}_{12} = \frac{1}{n}\operatorname{Tr} D_j(Q_j - \mathbb{E}Q_j)$ and $\boldsymbol{\varepsilon}_{13} = \frac{1}{n} \times \operatorname{Tr} D_j \mathbb{E}(Q_j - Q)$. By Lemmas 6.2(1), 6.3(2c) and 6.3(3), we have $\mathbb{E}|\boldsymbol{\varepsilon}_{11}|^2 \leq Kn^{-1}$, $\mathbb{E}|\boldsymbol{\varepsilon}_{12}|^2 \leq Kn^{-2}$ and $|\boldsymbol{\varepsilon}_{13}|^2 \leq Kn^{-2}$, respectively. By the Cauchy–Schwarz inequality, we therefore have

$$|\boldsymbol{\varepsilon}_1| \leq \frac{K(\mathbb{E}(y_j^* Q_j D_j Q_j y_j)^2)^{1/2}}{\sqrt{n}} \leq \frac{K'}{\sqrt{n}}$$

and (7.12) is established.

We now establish decomposition (7.13):

$$\begin{aligned}
\boldsymbol{\psi}_2 &= \omega \tilde{b}_j \tilde{c}_j \operatorname{Tr} D_j(\mathbb{E}Q - C)\\
&= \omega \tilde{b}_j \tilde{c}_j \sum_{i=1}^N \sigma_{ij}^2 c_i \mathbb{E}\big([Q]_{ii}(c_i^{-1} - [Q]_{ii}^{-1})\big)\\
&= -\omega^2 \tilde{b}_j \tilde{c}_j \sum_{i=1}^N \sigma_{ij}^2 c_i \mathbb{E}\biggl([Q]_{ii}\biggl(\xi_i \tilde{Q}_i \xi_i^* - \frac{1}{n}\operatorname{Tr}\tilde{D}_i \mathbb{E}\tilde{Q}\biggr)\biggr)\\
&\overset{\text{(a)}}{=} -\omega^2 \tilde{b}_j \tilde{c}_j \sum_{i=1}^N \sigma_{ij}^2 c_i^2 \biggl(\mathbb{E}(\xi_i \tilde{Q}_i \xi_i^*) - \frac{1}{n}\operatorname{Tr}\tilde{D}_i \mathbb{E}\tilde{Q}\biggr)\\
&\qquad + \omega^3 \tilde{b}_j \tilde{c}_j \sum_{i=1}^N \sigma_{ij}^2 c_i^2 \mathbb{E}\biggl([Q]_{ii}\biggl(\xi_i \tilde{Q}_i \xi_i^* - \frac{1}{n}\operatorname{Tr}\tilde{D}_i \mathbb{E}\tilde{Q}\biggr)^2\biggr)\\
&\overset{\triangle}{=} -\boldsymbol{\psi}_5 + \boldsymbol{\psi}_6,
\end{aligned}$$

where (a) follows from (6.7). Equation (7.13) is thus established.

Let us now turn to decomposition (7.14). We have $\boldsymbol{\psi}_5 = \frac{\omega^2 \tilde{b}_j \tilde{c}_j}{n} \sum_{i=1}^N \sigma_{ij}^2 c_i^2 \times \operatorname{Tr}\tilde{D}_i \mathbb{E}(\tilde{Q}_i - \tilde{Q})$. By similar arguments as those used for $\boldsymbol{\psi}_1$, we have

$$\begin{aligned}
\boldsymbol{\psi}_5 &= \frac{\omega^3 \tilde{b}_j \tilde{c}_j}{n^2} \sum_{i=1}^N \sigma_{ij}^2 c_i^3 \mathbb{E}(\operatorname{Tr}\tilde{D}_i \tilde{Q}_i \tilde{D}_i \tilde{Q}_i)\\
&\qquad + \frac{\omega^3 \tilde{b}_j \tilde{c}_j}{n^2} \sum_{i=1}^N \sigma_{ij}^2 c_i^2 \mathbb{E}([Q]_{ii} - c_i)\operatorname{Tr}\tilde{D}_i \tilde{Q}_i \xi_i^* \xi_i \tilde{Q}_i\\
&\overset{\triangle}{=} \boldsymbol{\psi}_7 + \boldsymbol{\varepsilon}_5,
\end{aligned}$$

where $|\boldsymbol{\varepsilon}_5| \leq Kn^{-1/2}$ and (7.14) is thus established.

Turning to (7.15), we have

$$
\begin{aligned}
\boldsymbol{\psi}_6 &= \omega^3 \tilde{b}_j \tilde{c}_j \sum_{i=1}^N \sigma_{ij}^2 c_i^2 \mathbb{E}\Big([Q]_{ii}\Big(\xi_i \tilde{Q}_i \xi_i^* - \frac{1}{n}\operatorname{Tr}\tilde{D}_i \mathbb{E}\tilde{Q}\Big)^2\Big) \\
&= \omega^3 \tilde{b}_j \tilde{c}_j \sum_{i=1}^N \sigma_{ij}^2 c_i^3 \mathbb{E}\Big(\xi_i \tilde{Q}_i \xi_i^* - \frac{1}{n}\operatorname{Tr}\tilde{D}_i \mathbb{E}\tilde{Q}\Big)^2 \\
&\quad - \omega^4 \tilde{b}_j \tilde{c}_j \sum_{i=1}^N \sigma_{ij}^2 c_i^3 \mathbb{E}\Big([Q]_{ii}\Big(\xi_i \tilde{Q}_i \xi_i^* - \frac{1}{n}\operatorname{Tr}\tilde{D}_i \mathbb{E}\tilde{Q}\Big)^3\Big) \\
&\stackrel{\triangle}{=} \boldsymbol{\psi}_6' + \boldsymbol{\varepsilon}_{61},
\end{aligned}
\tag{7.18}
$$

again using (6.7). The term $\boldsymbol{\varepsilon}_{61}$ satisfies

$$
\begin{aligned}
|\boldsymbol{\varepsilon}_{61}| &\le \frac{1}{\omega^2}\sum_{i=1}^N \sigma_{ij}^2 \mathbb{E}|\boldsymbol{\varepsilon}_{611,i} + \boldsymbol{\varepsilon}_{612,i} + \boldsymbol{\varepsilon}_{613,i}|^3 \\
&\le \frac{9}{\omega^2}\sum_{i=1}^N \sigma_{ij}^2 (\mathbb{E}|\boldsymbol{\varepsilon}_{611,i}|^3 + \mathbb{E}|\boldsymbol{\varepsilon}_{612,i}|^3 + |\boldsymbol{\varepsilon}_{613,i}|^3),
\end{aligned}
$$

where $\boldsymbol{\varepsilon}_{611,i} = \xi_i \tilde{Q}_i \xi_i^* - \frac{1}{n}\operatorname{Tr}\tilde{D}_i \tilde{Q}_i$, $\boldsymbol{\varepsilon}_{612,i} = \frac{1}{n}\operatorname{Tr}\tilde{D}_i(\tilde{Q}_i - \mathbb{E}\tilde{Q}_i)$ and $\boldsymbol{\varepsilon}_{613,i} = \frac{1}{n}\operatorname{Tr}\tilde{D}_i \mathbb{E}(\tilde{Q}_i - \tilde{Q})$. By Lemma 6.2(1), $\mathbb{E}|\boldsymbol{\varepsilon}_{611,i}|^3 \le Kn^{-3/2}$. By Lemma 6.3(2d), $\mathbb{E}|\boldsymbol{\varepsilon}_{612,i}|^3 \le (\mathbb{E}|\boldsymbol{\varepsilon}_{612,i}|^4)^{3/4} \le Kn^{-3/2}$. By Lemma 6.3(3), $|\boldsymbol{\varepsilon}_{613,i}|^3 \le Kn^{-3}$, hence

$$
|\boldsymbol{\varepsilon}_{6,1}| \le \frac{K}{\sqrt{n}}.
$$

We now handle the term $\boldsymbol{\psi}_6'$ in (7.18). We have

$$
\begin{aligned}
\boldsymbol{\psi}_6' &= \omega^3 \tilde{b}_j \tilde{c}_j \sum_{i=1}^N \sigma_{ij}^2 c_i^3 \mathbb{E}\Big(\xi_i \tilde{Q}_i \xi_i^* - \frac{1}{n}\operatorname{Tr}\tilde{D}_i \tilde{Q}_i + \boldsymbol{\varepsilon}_{612,i} + \boldsymbol{\varepsilon}_{613,i}\Big)^2 \\
&= \omega^3 \tilde{b}_j \tilde{c}_j \sum_{i=1}^N \sigma_{ij}^2 c_i^3 \mathbb{E}\Big(\xi_i \tilde{Q}_i \xi_i^* - \frac{1}{n}\operatorname{Tr}\tilde{D}_i \tilde{Q}_i\Big)^2 + \boldsymbol{\varepsilon}_{62} \\
&\stackrel{\triangle}{=} \boldsymbol{\psi}_8 + \boldsymbol{\varepsilon}_{62},
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{\varepsilon}_{62} = \omega^3 \tilde{b}_j \tilde{c}_j \sum_{i=1}^N \sigma_{ij}^2 c_i^3 \Big(&\mathbb{E}(\boldsymbol{\varepsilon}_{612,i} + \boldsymbol{\varepsilon}_{613,i})^2 \\
&+ 2\mathbb{E}\Big(\Big(\xi_i \tilde{Q}_i \xi_i^* - \frac{1}{n}\operatorname{Tr}\tilde{D}_i \tilde{Q}_i\Big)(\boldsymbol{\varepsilon}_{612,i} + \boldsymbol{\varepsilon}_{613,i})\Big)\Big).
\end{aligned}
$$

Using Lemmas 6.2(1) and 6.3, it is easy to show that

$$|\boldsymbol{\varepsilon}_{62}| \le \frac{K}{\sqrt{n}}.$$

Furthermore, the terms $\mathbb{E}()^2$ in the expression of $\boldsymbol{\psi}_8$ have a more explicit form. Indeed, applying Lemma 6.2(2) yields

$$\boldsymbol{\psi}_8 = \frac{\omega^3 \tilde{b}_j \tilde{c}_j}{n^2} \sum_{i=1}^{N} \sigma_{ij}^2 c_i^3 \left(\mathbb{E}(\operatorname{Tr} \tilde{D}_i \tilde{Q}_i \tilde{D}_i \tilde{Q}_i) + \kappa \sum_{m=1}^{n} \sigma_{im}^4 \mathbb{E}([\tilde{Q}_i]_{mm}^2) \right).$$

Decomposition (7.15) is established with $\boldsymbol{\varepsilon}_6 = \boldsymbol{\varepsilon}_{61} + \boldsymbol{\varepsilon}_{62}$.

It remains to give decomposition (7.16). Using (6.8), we have

$$\begin{aligned}
\boldsymbol{\psi}_3 &= n\omega^2 \tilde{b}_j \tilde{c}_j \mathbb{E}\left([\tilde{Q}]_{jj} \left(y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j \mathbb{E} Q \right) \left(y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j C \right) \right) \\
&= n\omega^2 \tilde{b}_j \tilde{c}_j^2 \mathbb{E}\left(\left(y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j \mathbb{E} Q \right) \left(y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j C \right) \right) \\
&\quad - n\omega^3 \tilde{b}_j \tilde{c}_j^2 \mathbb{E}\left([\tilde{Q}]_{jj} \left(y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j \mathbb{E} Q \right)^2 \left(y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j C \right) \right) \\
&\stackrel{\triangle}{=} \boldsymbol{\psi}_3' + \boldsymbol{\varepsilon}_{31}.
\end{aligned}$$

The term $\boldsymbol{\varepsilon}_{31}$ satisfies

$$\begin{aligned}
|\boldsymbol{\varepsilon}_{31}| \le \frac{n}{\omega} \mathbb{E}\Bigg(&\left| y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j Q_j + \boldsymbol{\varepsilon}_{311} + \boldsymbol{\varepsilon}_{312} \right|^2 \\
&\times \left| y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j Q_j + \boldsymbol{\varepsilon}_{311} + \boldsymbol{\varepsilon}_{312} + \boldsymbol{\varepsilon}_{313} \right| \Bigg),
\end{aligned}$$

with $\boldsymbol{\varepsilon}_{311} = \frac{1}{n} \operatorname{Tr} D_j (Q_j - \mathbb{E} Q_j)$, $\boldsymbol{\varepsilon}_{312} = \frac{1}{n} \operatorname{Tr} D_j \mathbb{E}(Q_j - Q)$ and $\boldsymbol{\varepsilon}_{313} = \frac{1}{n} \times \operatorname{Tr} D_j (\mathbb{E} Q - C)$.

The terms $\boldsymbol{\varepsilon}_{311}$, $\boldsymbol{\varepsilon}_{312}$ and $y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j Q_j$ can be handled by Lemmas 6.2(1) and 6.3. The term $\boldsymbol{\varepsilon}_{313}$ coincides with $\boldsymbol{\psi}_2 (n\omega \tilde{b}_j \tilde{c}_j)^{-1}$. The derivations made on $\boldsymbol{\psi}_2$ above [decompositions (7.13)–(7.15)] show that $|\boldsymbol{\psi}_2 (\omega \tilde{b}_j \tilde{c}_j)^{-1}| \le K$, therefore $|\boldsymbol{\varepsilon}_{313}| \le K n^{-1}$.

Using these results, we obtain, after some standard manipulations,

$$|\boldsymbol{\varepsilon}_{31}| \le \frac{K}{\sqrt{n}}.$$

The term $\boldsymbol{\psi}_3'$ can be written as

$$\boldsymbol{\psi}_3' = n\omega^2 \tilde{b}_j \tilde{c}_j^2 \mathbb{E}\left(\left(y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j Q_j + \boldsymbol{\varepsilon}_{311} + \boldsymbol{\varepsilon}_{312} \right) \right.$$

$$
\begin{aligned}
&\times\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j Q_j + \boldsymbol{\varepsilon}_{311} + \boldsymbol{\varepsilon}_{312} + \boldsymbol{\varepsilon}_{313}\right)\Bigg) \\
&= n\omega^2 \tilde{b}_j \tilde{c}_j^2 \mathbb{E}\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j Q_j\right)^2 + \boldsymbol{\varepsilon}_{32} \\
&\stackrel{\triangle}{=} \boldsymbol{\psi}_9 + \boldsymbol{\varepsilon}_{32},
\end{aligned}
$$

with $|\boldsymbol{\varepsilon}_{32}| \le K n^{-1/2}$. Similarly to $\boldsymbol{\psi}_8$, we can develop $\boldsymbol{\psi}_9$ to obtain

$$
\boldsymbol{\psi}_9 = \frac{\omega^2 \tilde{b}_j \tilde{c}_j^2}{n}\left(\mathbb{E}(\operatorname{Tr} D_j Q_j D_j Q_j) + \kappa \sum_{i=1}^{N} \sigma_{ij}^4 \mathbb{E}([Q_j]_{ii}^2)\right). \tag{7.19}
$$

Decomposition (7.16) is established with $\boldsymbol{\varepsilon}_3 = \boldsymbol{\varepsilon}_{31} + \boldsymbol{\varepsilon}_{32}$.

We now put the pieces together and provide equation (7.17), satisfied by $\boldsymbol{\psi}^{(j)}$. We recall that

$$
\begin{aligned}
\boldsymbol{\psi}_4 &= \frac{\omega^2 \tilde{b}_j \tilde{c}_j^2}{n} \mathbb{E}(\operatorname{Tr} D_j Q_j D_j Q_j), \\
\boldsymbol{\psi}_7 &= \frac{\omega^3 \tilde{b}_j \tilde{c}_j}{n^2} \sum_{i=1}^{N} \sigma_{ij}^2 c_i^3 \mathbb{E}(\operatorname{Tr} \tilde{D}_i \tilde{Q}_i \tilde{D}_i \tilde{Q}_i), \\
\boldsymbol{\psi}_8 &= \frac{\omega^3 \tilde{b}_j \tilde{c}_j}{n^2} \sum_{i=1}^{N} \sigma_{ij}^2 c_i^3 \left(\mathbb{E}(\operatorname{Tr} \tilde{D}_i \tilde{Q}_i \tilde{D}_i \tilde{Q}_i) + \kappa \sum_{m=1}^{n} \sigma_{im}^4 \mathbb{E}([\tilde{Q}_i]_{mm}^2)\right), \\
\boldsymbol{\psi}_9 &= \frac{\omega^2 \tilde{b}_j \tilde{c}_j^2}{n}\left(\mathbb{E}(\operatorname{Tr} D_j Q_j D_j Q_j) + \kappa \sum_{i=1}^{N} \sigma_{ij}^4 \mathbb{E}([Q_j]_{ii}^2)\right).
\end{aligned}
$$

When computing the right-hand side of (7.17), all terms of the form $\mathbb{E}\operatorname{Tr} D_j Q_j D_j Q_j$ and $\mathbb{E}\operatorname{Tr} \tilde{D}_i \tilde{Q}_i \tilde{D}_i \tilde{Q}_i$ cancel out and we end up with equation (7.7). Step 2 is thus established.

7.3. *Proof of step* 3: $\|\boldsymbol{\varphi} - \mathbf{w}\|_\infty \to 0$. In order to prove (7.8), we need the following facts:

$$
|||(\breve{A} - A)^T|||_\infty \underset{n\to\infty}{\longrightarrow} 0; \tag{7.20}
$$

$$
\limsup_n |||(I - A)^{-1}|||_\infty < \infty; \tag{7.21}
$$

$$
I - \breve{A} \text{ is invertible for } n \text{ large enough}; \tag{7.22}
$$

$$
\limsup_n |||(I - \breve{A})^{-1}|||_\infty < \infty. \tag{7.23}
$$

The proof of (7.20) is close to the proof of (6.31) above and is therefore omitted. The bound (7.21) follows from Lemma 5.2(3). We now prove (7.22) and (7.23).

Recall that, by Lemma 5.5, there exist two vectors $u_n = (u_{\ell,n}) \succ 0$ and $v_n = (v_{\ell,n}) \succ 0$ such that $u_n = Au_n + v_n$, $\sup_n \|u_n\|_\infty < \infty$ and $\liminf_n \min_\ell(v_{\ell,n}) > 0$. Matrix $\breve{A}$ satisfies the equation $u_n = \breve{A}u_n + \breve{v}_n$ with $\breve{v}_n = (\breve{v}_{\ell n}) = v_n + (A - \breve{A})u_n$. Combining (7.20) with inequalities $\sup_n \|u_n\|_\infty < \infty$ and $\liminf_n(\min_\ell v_{\ell n}) > 0$, we have $\liminf_n(\min_\ell \breve{v}_{\ell n}) > 0$. Therefore, Lemma 5.2 applies to matrix $\breve{A}$ for n large enough, which implies (7.22) and (7.23).

We are now in a position to prove $\|\boldsymbol{\varphi} - \mathbf{w}\|_\infty \to 0$. Working out equations (7.6) and (7.4), we obtain

$$\boldsymbol{\varphi} = \mathbf{w} + (I - A)^{-1}(\breve{A} - A)\boldsymbol{\varphi} + (I - A)^{-1}(\boldsymbol{\psi} - \mathbf{p}),$$

hence

$$\|\boldsymbol{\varphi} - \mathbf{w}\|_\infty \le |||(I - A)^{-1}|||_\infty |||(\breve{A} - A)|||_\infty \|\boldsymbol{\varphi}\|_\infty + |||(I - A)^{-1}|||_\infty \|\boldsymbol{\psi} - \mathbf{p}\|_\infty.$$

Thanks to (7.22), we have $\boldsymbol{\varphi} = (I - \breve{A})^{-1}\boldsymbol{\psi}$ for n large enough. One can check from (7.7) that $\sup_n \|\boldsymbol{\psi}\|_\infty < \infty$. Therefore, by (7.23), we have $\sup_n \|\boldsymbol{\varphi}\|_\infty < \infty$. Using (7.20) and (7.21), we then have $|||(I - A)^{-1}|||_\infty |||(\breve{A} - A)|||_\infty \|\boldsymbol{\varphi}\|_\infty \to 0$.

It remains to prove that $\|\boldsymbol{\psi} - \mathbf{p}\|_\infty \to 0$. In step 3, it has been established that $\boldsymbol{\psi}$ is a perturbed version of $\mathbf{p}$, as defined in (3.2), in the sense of equation (7.7). Using the arguments developed in the course of the proof of (6.18), it is a matter of routine to check that $\|\boldsymbol{\psi} - \mathbf{p}\|_\infty \to 0$. Details are omitted. Hence,

$$|||(I - A)^{-1}|||_\infty \|\boldsymbol{\psi} - \mathbf{p}\|_\infty \to 0.$$

Consequently, $\|\boldsymbol{\varphi} - \mathbf{w}\|_\infty \to 0$ and step 3 is proved.

7.4. *Proof of step* 4: *Dominated convergence.* In this section, the constant K' does not depend on n or ω, but its value is allowed to change from line to line. We first prove (7.9). We have

$$|\beta_n| \le \|\mathbf{w}\|_\infty \le |||(I - A)^{-1}|||_\infty \|\mathbf{p}\|_\infty,$$

by (7.4). By inspecting (3.2), one obtains $\|\mathbf{p}\|_\infty \le |\kappa|(N/n)(\frac{\sigma_{\max}^6}{\omega^4} + \frac{\sigma_{\max}^4}{\omega^3}) \le K'\omega^{-3}$. We need now to bound $|||(I - A)^{-1}|||_\infty$ in terms of $\omega \in [\rho, \infty)$. Lemma 5.2(3) yields

$$|||(I - A)^{-1}|||_\infty \le \frac{\max_\ell(u_{\ell,n})}{\min_\ell(v_{\ell,n})},$$

where $u_n = (u_{\ell n})$ and $v_n = (v_{\ell n})$ are the vectors given in the statement of Lemma 5.5. We now inspect the expressions of $u_{\ell n}$ and $v_{\ell n}$. Equation (5.4) yields

$$\min_\ell(v_{\ell,n}) \ge \frac{1}{(\omega + \sigma_{\max})^2} \min_j \frac{1}{N} \operatorname{Tr} D_j$$

and $\max_\ell(u_{\ell,n}) \le (N\sigma_{\max}^2)(n\omega^2)^{-1}$ by (5.6). Collecting all of these estimates, we obtain $|\beta_n| \le K'\omega^{-3}$ and inequality (7.9) is proved.

We now prove (7.10). We have

$$\left|\frac{1}{n}\sum_{j=1}^{n}\boldsymbol{\varphi}_j\right| \le \|\boldsymbol{\varphi}\|_\infty \le |||(I-\breve{A})^{-1}|||_\infty \|\boldsymbol{\psi}\|_\infty, \tag{7.24}$$

by (7.6) and (7.22). We know that the right-hand side is bounded as $n\to\infty$. However, not much is known about the behavior of the bound with respect to ω. Using inequality (7.24) and relying on the derivations that lead to (7.6)–(7.7), one can prove that $|||(I-\breve{A})^{-1}|||_\infty$, $\|\boldsymbol{\psi}\|_\infty$ and, therefore, $\|\boldsymbol{\varphi}\|_\infty$ are bounded on the compact subsets of $[\rho,+\infty)$. Therefore, in order to establish (7.10), it is sufficient to prove that $\|\boldsymbol{\varphi}\|_\infty$ is bounded by $K'\omega^{-2}$ near infinity. To this end, we develop $|\boldsymbol{\varphi}_j(\omega)|$ as follows:

$$\begin{aligned}
|\boldsymbol{\varphi}_j(\omega)| &= n\tilde{t}_j\big|\mathbb{E}\big([\tilde{Q}]_{jj}([\tilde{Q}]_{jj}^{-1}-\tilde{t}_j^{-1})\big)\big| \\
&= n\omega\tilde{t}_j\left|\mathbb{E}\left([\tilde{Q}]_{jj}\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j T\right)\right)\right| \\
&\le \omega\tilde{t}_j\mathbb{E}[\tilde{Q}]_{jj}|\operatorname{Tr} D_j\mathbb{E}(Q-T)| \\
&\quad + n\omega\tilde{t}_j\left|\mathbb{E}\left([\tilde{Q}]_{jj}\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j\mathbb{E}Q\right)\right)\right| \\
&\overset{\text{(a)}}{\le} \omega\tilde{t}_j\mathbb{E}[\tilde{Q}]_{jj}|\operatorname{Tr} D_j\mathbb{E}(Q-T)| \\
&\quad + n\omega\tilde{t}_j\tilde{c}_j\left|\mathbb{E}\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j\mathbb{E}Q\right)\right| \\
&\quad + n\omega^2\tilde{t}_j\tilde{c}_j\left|\mathbb{E}\left([\tilde{Q}]_{jj}(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j\mathbb{E}Q)^2\right)\right| \\
&\overset{\text{(b)}}{\le} \omega\tilde{t}_j\mathbb{E}[\tilde{Q}]_{jj}|\operatorname{Tr} D_j\mathbb{E}(Q-T)| + \omega\tilde{t}_j\tilde{c}_j\big|\mathbb{E}\big(\operatorname{Tr} D_j(Q_j-Q)\big)\big| \\
&\quad + 2n\omega^2\tilde{t}_j\tilde{c}_j\mathbb{E}\left([\tilde{Q}]_{jj}\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j\mathbb{E}Q_j\right)^2\right) \\
&\quad + 2\frac{\omega^2\tilde{t}_j\tilde{c}_j}{n}\mathbb{E}[\tilde{Q}]_{jj}\big(\operatorname{Tr} D_j\mathbb{E}(Q_j-Q)\big)^2,
\end{aligned}$$

where (a) follows from (6.8) and (b) from the fact that

$$\begin{aligned}
&\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j\mathbb{E}Q\right)^2 \\
&\qquad \le 2\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j\mathbb{E}Q_j\right)^2 + 2\left(\frac{1}{n}\operatorname{Tr} D_j\mathbb{E}(Q_j-Q)\right)^2.
\end{aligned}$$

Let $\alpha(\omega) = n \max_{1\le i\le N} |t_i - \mathbb{E}[Q]_{ii}|$. Using Lemma 6.3(3), we obtain from the last inequality that

$$\|\boldsymbol{\varphi}(\omega)\|_\infty \le \frac{\sigma^2_{\max}}{\omega}\boldsymbol{\alpha}(\omega) + \frac{\sigma^2_{\max}}{\omega^2} + \frac{2n}{\omega}\mathbb{E}\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E} Q_j\right)^2 + \frac{2\sigma^4_{\max}}{n\omega^3}.$$

As in (7.19), we have

$$\begin{aligned}\mathbb{E}\left(y_j^* Q_j y_j - \frac{1}{n}\operatorname{Tr} D_j \mathbb{E} Q_j\right)^2 &= \frac{1}{n^2}\left(\mathbb{E}(\operatorname{Tr} D_j Q_j D_j Q_j) + \kappa \sum_{i=1}^N \sigma^4_{ij}\mathbb{E}[Q_j]^2_{ii}\right) \\ &\le \frac{N\sigma^4_{\max}(1+|\kappa|)}{n^2\omega^2}.\end{aligned}$$

Therefore,

$$\|\boldsymbol{\varphi}(\omega)\|_\infty \le \frac{\sigma^2_{\max}}{\omega}\boldsymbol{\alpha}(\omega) + \frac{K'}{\omega^2} \tag{7.25}$$

for $\omega \in [\rho, +\infty)$. A similar derivation yields $\boldsymbol{\alpha}(\omega) \le \frac{\sigma^2_{\max}}{\omega}\|\boldsymbol{\varphi}(\omega)\|_\infty + \frac{K'}{\omega^2}$. Plugging this inequality into (7.25), we obtain

$$(1 - \sigma^4_{\max}/\omega^2)\|\boldsymbol{\varphi}(\omega)\|_\infty \le \frac{K'}{\omega^2},$$

hence $\|\boldsymbol{\varphi}(\omega)\|_\infty \le K'\omega^{-2}$ for ω large enough.

We have proven that $\|\boldsymbol{\varphi}(\omega)\|_\infty$ is bounded on compact subsets of $[\rho, \infty)$ and, furthermore, that (7.10) is true for ω large enough. Therefore, (7.10) holds for every $\omega \in [\rho, +\infty)$. Step 4 is proved and so is Theorem 3.3.

APPENDIX A: PROOF OF LEMMA 6.3

PROOF OF LEMMA 6.3(1). This is straightforward. □

PROOF OF LEMMA 6.3(2).

PROOF OF (2A). From [17], Lemmas 6.1 and 6.6, we get

$$\frac{1}{n}\operatorname{Tr} U\bigl(Q(-\rho) - T(-\rho)\bigr) \xrightarrow[n\to\infty]{} 0 \qquad \text{a.s.}$$

Now, since

$$\left|\frac{1}{n}\operatorname{Tr} U\bigl(Q(-\rho) - T(-\rho)\bigr)\right| \le \|U\|\bigl(\|Q(-\rho)\| + \|T(-\rho)\|\bigr) \le \frac{2\|U\|}{\rho},$$

the dominated convergence theorem yields the first part of (2a). The second part is proved similarly. □

PROOF OF (2B). Recall from Theorem 2.3(1) and from the mere definitions of T and B that the matrices $T(z)$ and $B(z)$ can be written as

$$T = \left(-zI + \frac{1}{n}\sum_{j=1}^{n} \frac{1}{1+(1/n)\operatorname{Tr} D_j T} D_j\right)^{-1}$$

and

$$B = \left(-zI + \frac{1}{n}\sum_{j=1}^{n} \frac{1}{1+(1/n)\operatorname{Tr} D_j \mathbb{E}Q} D_j\right)^{-1}.$$

We therefore have

$$\begin{aligned}
&\frac{1}{n}\operatorname{Tr} U\big(B(-\rho) - T(-\rho)\big)\\
&\qquad = \frac{1}{n}\operatorname{Tr} UBT(T^{-1} - B^{-1})\\
&\qquad = \frac{1}{n^2}\operatorname{Tr}\left(UBT\sum_{j=1}^{n} \frac{(1/n)\operatorname{Tr} D_j(\mathbb{E}Q - T)}{(1+(1/n)\operatorname{Tr} D_j T)(1+(1/n)\operatorname{Tr} D_j\mathbb{E}Q)} D_j\right)\\
&\qquad = \frac{1}{n^2}\sum_{i=1}^{N}\sum_{j=1}^{n} x_{ij}^n,
\end{aligned}$$

with $x_{ij}^n = \frac{[U]_{ii} b_i t_i \sigma_{ij}^2}{(1+(1/n)\operatorname{Tr} D_j T)(1+(1/n)\operatorname{Tr} D_j \mathbb{E}Q)} \frac{1}{n}\operatorname{Tr} D_j(\mathbb{E}Q - T)$. It can be easily checked that $|x_{ij}^n| \le 2\sup_n(\|U\|)\sigma_{\max}^4/\rho^3$. Furthermore, $x_{ij}^n \to_n 0$ for every i, j, by (2a). It remains to apply the dominated convergence theorem to the integral with respect to Lebesgue measure on $[0,1]^2$ of the staircase function $f_n(x,y)$, defined as $f_n(i/N, j/n) = x_{ij}^n$, to deduce that $\frac{1}{n}\operatorname{Tr} U(B - T) \to 0$. This completes the proof of (2b). □

In the sequel, K is a constant whose value might change from line to line, but which remains independent of n.

PROOF OF (2C). We have

$$\begin{aligned}
\operatorname{Tr} U(Q - \mathbb{E}Q) &\overset{\text{(a)}}{=} \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j+1})\operatorname{Tr} UQ\\
&\overset{\text{(b)}}{=} \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j+1})\operatorname{Tr} U(Q - Q_j) \qquad \text{(A.1)}\\
&\overset{\text{(c)}}{=} -\sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j+1}) \frac{y_j^* Q_j U Q_j y_j}{1 + y_j^* Q_j y_j} \overset{\triangle}{=} \sum_{j=1}^{n} x_j,
\end{aligned}$$

where (a) follows from the fact that $\mathbb{E}_1 \operatorname{Tr} UQ = \operatorname{Tr} UQ$ and $\mathbb{E}_{n+1} \operatorname{Tr} UQ = \mathbb{E} \operatorname{Tr} UQ$, (b) follows from the fact that $\mathbb{E}_j \operatorname{Tr} UQ_j = \mathbb{E}_{j+1} \operatorname{Tr} UQ_j$ since Q_j does not depend on y_j and (c) follows from (6.1) and the fact that $\operatorname{Tr} Q_j y_j y_j^* Q_j U = y_j^* Q_j U Q_j y_j$.

Now, one can easily check that $\sum_{j=1}^n x_j (= \operatorname{Tr} U(Q - \mathbb{E}Q))$ is the sum of a martingale difference sequence with respect to the increasing filtration $\mathcal{F}_n, \dots, \mathcal{F}_1$ since $\mathbb{E}_k x_j = 0$ for $k > j$. Therefore,

$$\mathbb{E}\big(\operatorname{Tr} U(Q - \mathbb{E}Q)\big)^2 = \sum_{j=1}^n \mathbb{E} x_j^2.$$

Write $x_j = x_{j,1} + x_{j,2}$, where

$$x_{j,1} = -(\mathbb{E}_j - \mathbb{E}_{j+1})\left(\frac{y_j^* Q_j U Q_j y_j}{1 + (1/n) \operatorname{Tr} D_j Q_j}\right),$$

$$x_{j,2} = -(\mathbb{E}_j - \mathbb{E}_{j+1})\left(\frac{y_j^* Q_j U Q_j y_j}{1 + y_j^* Q_j y_j} - \frac{y_j^* Q_j U Q_j y_j}{1 + (1/n) \operatorname{Tr} D_j Q_j}\right).$$

Using the fact that y_j and $\mathcal{F}_{j+1}$ are independent and the fact that Q_j does not depend on y_j, one easily obtains

$$\mathbb{E}_{j+1}\left(\frac{y_j^* Q_j U Q_j y_j}{1 + (1/n) \operatorname{Tr} D_j Q_j}\right) = \frac{1}{n} \operatorname{Tr} D_j \mathbb{E}_{j+1}\left(\frac{Q_j U Q_j}{1 + (1/n) \operatorname{Tr} D_j Q_j}\right).$$

Thus, $x_{j,1}$ and $x_{j,2}$ can be written as

$$\begin{aligned} x_{j,1} &= -y_j^* \mathbb{E}_{j+1}\left(\frac{Q_j U Q_j}{1 + (1/n) \operatorname{Tr} D_j Q_j}\right) y_j \\ &\quad + \frac{1}{n} \operatorname{Tr} D_j \mathbb{E}_{j+1}\left(\frac{Q_j U Q_j}{1 + (1/n) \operatorname{Tr} D_j Q_j}\right), \\ x_{j,2} &= (\mathbb{E}_j - \mathbb{E}_{j+1}) \frac{y_j^* Q_j U Q_j y_j}{(1 + (1/n) \operatorname{Tr} D_j Q_j)(1 + y_j^* Q_j y_j)} \\ &\quad \times \left(y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j Q_j\right) \\ &\stackrel{\triangle}{=} (\mathbb{E}_j - \mathbb{E}_{j+1}) x_{j,3}. \end{aligned}$$

Since the matrix $\|D_j \mathbb{E}_j(\frac{Q_j U Q_j}{1+(1/n) \operatorname{Tr} D_j Q_j})\| \le K$, Lemma 6.2(1) and Assumption A1 together yield $\mathbb{E} x_{1,j}^2 \le K n^{-1}$. Furthermore, we have

$$|x_{j,3}| \le \left| y_j^* Q_j U Q_j y_j \left(y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j Q_j \right) \right|$$

since $y_j^* Q_j y_j \geq 0$ and $\frac{1}{n} \operatorname{Tr} D_j Q_j \geq 0$. The Cauchy–Schwarz inequality yields

$$\mathbb{E}x_{j,3}^2 \leq (\mathbb{E}(y_j^* Q_j U Q_j y_j)^4)^{1/2} \left(\mathbb{E}\left(y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j Q_j\right)^4\right)^{1/2},$$

which, in turn, yields $\mathbb{E}x_{j,3}^2 < \frac{K}{n}$ since

$$\mathbb{E}(y_j^* Q_j U Q_j y_j)^4 \leq K \quad \text{and} \quad \mathbb{E}\left(y_j^* Q_j y_j - \frac{1}{n} \operatorname{Tr} D_j Q_j\right)^4 \leq \frac{K}{n^2}, \tag{A.2}$$

where the first inequality in (A.2) follows from $0 \leq y_j^* Q_j U Q_j y_j \leq \|Q_j U Q_j\| \times \|y_j\|^2$ and Assumption A1, and the second from Assumption A1 and Lemma 6.2(1).

We are now in a position to conclude.

$$\begin{aligned}
\mathbb{E}x_{j,2}^2 &= \mathbb{E}\big((\mathbb{E}_j - \mathbb{E}_{j+1}) x_{j,3}\big)^2 \leq 2\mathbb{E}\big((\mathbb{E}_j x_{j,3})^2 + (\mathbb{E}_{j+1} x_{j,3})^2\big) \\
&\overset{\text{(a)}}{\leq} 2\mathbb{E}(\mathbb{E}_j x_{j,3}^2 + \mathbb{E}_{j+1} x_{j,3}^2) = 4\mathbb{E}x_{j,3}^2,
\end{aligned}$$

where (a) follows from Jensen's inequality. Now,

$$\mathbb{E}x_j^2 = \mathbb{E}(x_{j,1} + x_{j,2})^2 \leq \big((\mathbb{E}x_{j,1}^2)^{1/2} + (\mathbb{E}x_{j,2}^2)^{1/2}\big)^2 \leq \frac{K}{n}$$

and $\mathbb{E}(\operatorname{Tr} U(Q - \mathbb{E}Q))^2 = \sum_{j=1}^n \mathbb{E}x_j^2 \leq K$. Inequality (c) is thus proved. □

PROOF OF (2D). We again rely on the decomposition (A.1) and follow along the lines of the computations in [3], page 580:

$$\operatorname{Tr} U(Q - \mathbb{E}Q) = -\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j+1}) \frac{y_j^* Q_j U Q_j y_j}{1 + y_j^* Q_j y_j}.$$

Thus,

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{N} \operatorname{Tr} U(Q - \mathbb{E}Q)\right)^4 &= \frac{1}{N^4} \mathbb{E}\left(\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j+1}) \frac{y_j^* Q_j U Q_j y_j}{1 + y_j^* Q_j y_j}\right)^4 \\
&\overset{\text{(a)}}{\leq} \frac{K}{N^4} \mathbb{E}\left(\sum_{j=1}^n \left((\mathbb{E}_j - \mathbb{E}_{j+1}) \frac{y_j^* Q_j U Q_j y_j}{1 + y_j^* Q_j y_j}\right)^2\right)^2 \\
&\overset{\text{(b)}}{\leq} \frac{K}{N^4} N \sum_{j=1}^n \mathbb{E}\left((\mathbb{E}_j - \mathbb{E}_{j+1}) \frac{y_j^* Q_j U Q_j y_j}{1 + y_j^* Q_j y_j}\right)^4 \\
&\leq \frac{K}{N^2} \sup_j \mathbb{E}\left((\mathbb{E}_j - \mathbb{E}_{j+1}) \frac{y_j^* Q_j U Q_j y_j}{1 + y_j^* Q_j y_j}\right)^4,
\end{aligned}$$

where (a) follows from Burkholder's inequality and (b) from the convexity inequality $(\sum_{i=1}^n a_i)^2 \leq n \sum_{i=1}^n a_i^2$. Now, recall that $y_j^* Q_j y_j \geq 0$ and $\|Q_j(-\rho)\| \leq 1/\rho$.

Standard computations yield

$$\mathbb{E}\biggl((\mathbb{E}_j - \mathbb{E}_{j+1})\frac{y_j^* Q_j U Q_j y_j}{1 + y_j^* Q_j y_j}\biggr)^4 \le K\mathbb{E}(y_j^* Q_j U Q_j y_j)^4 \le \frac{K\|U\|^4}{\rho^8}\mathbb{E}\|y_j\|^4,$$

which is uniformly bounded by Assumptions A1 and A2. Therefore, (2d) is proved. $\square$

PROOF OF LEMMA 6.3(3). Developing the difference $Q - Q_j$ with the help of (6.1), we obtain

$$\begin{aligned} |\operatorname{Tr} M(Q - Q_j)| &= \left|\operatorname{Tr} M\biggl(\frac{Q_j y_j y_j^* Q_j}{1 + y_j^* Q_j y_j}\biggr)\right| \\ &= \frac{|y_j^* Q_j M Q_j y_j|}{1 + y_j^* Q_j y_j} \le \|M\| \frac{\|Q_j y_j\|^2}{1 + y_j^* Q_j y_j}. \end{aligned}$$

Consider a spectral representation of $Y^j Y^{j*}$, that is, $Y^j Y^{j*} = \sum_{i=1}^N \lambda_i e_i e_i^*$. We have

$$\|Q_j y_j\|^2 = \sum_{i=1}^N \frac{|e_i^* y_j|^2}{(\lambda_i + \rho)^2} \quad \text{and} \quad y_j^* Q_j y_j = \sum_{i=1}^N \frac{|e_i^* y_j|^2}{\lambda_i + \rho} \ge \rho \sum_{i=1}^N \frac{|e_i^* y_j|^2}{(\lambda_i + \rho)^2},$$

hence the result. Inequality (3) is thus proved. $\square$

APPENDIX B: PROOF OF FORMULA (7.2)

Recalling that $Q(z) = (YY^* - zI_N)^{-1}$ and $\tilde{Q}(z) = (Y^*Y - zI_n)^{-1}$, it is easy to show that $\operatorname{Tr}(Q) - \operatorname{Tr}(\tilde{Q}) = (n - N)/z$. We shall now show that $\operatorname{Tr}(T) - \operatorname{Tr}(\tilde{T}) = (n - N)/z$. Formula (7.2) is obtained by combining these two equations.

Equations (2.2) in the statement of Lemma 2.4 can be rewritten as

$$t_i + \frac{t_i}{n}\sum_{j=1}^n \sigma_{ij}^2 \tilde{t}_j = -\frac{1}{z} \qquad \text{for } 1 \le i \le N,$$

$$\tilde{t}_j + \frac{\tilde{t}_j}{n}\sum_{i=1}^N \sigma_{ij}^2 t_i = -\frac{1}{z} \qquad \text{for } 1 \le j \le n.$$

By summing the first N equations over i and the next n equations over j, and by eliminating the term $\frac{1}{n}\sum_{i=1}^N \sum_{j=1}^n \sigma_{ij}^2 t_i \tilde{t}_j$, we obtain $\sum_i t_i - \sum_j \tilde{t}_j = (n - N)/z$, which is the desired result. Equation (7.2) is thus proved.

REFERENCES

[1] ANDERSON, G. W. and ZEITOUNI, O. (2006). A CLT for a band matrix model. *Probab. Theory Related Fields* **134** 283–338. MR2222385

[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. *Ann. Probab.* **26** 316–345. MR1617051

[3] BAI, Z. D. and SILVERSTEIN, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** 553–605. MR2040792

[4] BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed. Wiley, New York. MR1324786

[5] BOUTET DE MONVEL, A. and KHORUNZHY, A. (1998). Limit theorems for random matrices. *Markov Process. Related Fields* **4** 175–197. MR1641617

[6] BOUTET DE MONVEL, A., KHORUNZHY, A. and VASILCHUK, V. (1996). Limiting eigenvalue distribution of random matrices with correlated entries. *Markov Process. Related Fields* **2** 607–636. MR1431189

[7] CABANAL-DUVILLARD, T. (2001). Fluctuations de la loi empirique de grandes matrices aléatoires. *Ann. Inst. H. Poincaré Probab. Statist.* **37** 373–402. MR1831988

[8] DEBBAH, M. and MÜLLER, R. (2003). On the limiting behavior of directional mimo models. In *Proceedings of the Sixth Baiona Workshop on Signal Processing in Communication*. Spain.

[9] GIRKO, V. L. (1975). *Sluchainye Matritsy*. Izdat. Ob″ed. "Višča Škola" pri Kiev. Gosudarstv. Univ., Kiev. MR0394812

[10] GIRKO, V. L. (1990). *Theory of Random Determinants*. *Mathematics and Its Applications* (*Soviet Series*) **45**. Kluwer Academic, Dordrecht. Translated from the Russian. MR1080966

[11] GIRKO, V. L. (2001). *Theory of Stochastic Canonical Equations*. **I**. *Mathematics and Its Applications* **535**. Kluwer Academic, Dordrecht.

[12] GIRKO, V. L. (2003). Thirty years of the central resolvent law and three laws on the $1/n$ expansion for resolvent of random matrices. *Random Oper. Stochastic Equations* **11** 167–212. MR2001000

[13] GUIONNET, A. (2002). Large deviations upper bounds and central limit theorems for non-commutative functionals of Gaussian large random matrices. *Ann. Inst. H. Poincaré Probab. Statist.* **38** 341–384. MR1899457

[14] HACHEM, W., KHORUNZHIY, O., LOUBATON, P., NAJIM, J. and PASTUR, L. (2008). A new approach for capacity analysis of large dimensional multi-antenna channels. *IEEE Trans. Inform. Theory*. **54**(9).

[15] HACHEM, W., LOUBATON, P. and NAJIM, J. (2005). The empirical eigenvalue distribution of a gram matrix: From independence to stationarity. *Markov Process. Related Fields* **11** 629–648. MR2191967

[16] HACHEM, W., LOUBATON, P. and NAJIM, J. (2006). The empirical distribution of the eigenvalues of a Gram matrix with a given variance profile. *Ann. Inst. H. Poincaré Probab. Statist.* **42** 649–670. MR2269232

[17] HACHEM, W., LOUBATON, P. and NAJIM, J. (2007). Deterministic equivalents for certain functionals of large random matrices. *Ann. Appl. Probab.* **17** 875–930. MR2326235

[18] HORN, R. A. and JOHNSON, C. R. (1994). *Topics in Matrix Analysis*. Cambridge Univ. Press, Cambridge. Corrected reprint of the 1991 original. MR1288752

[19] JOHANSSON, K. (1998). On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.* **91** 151–204. MR1487983

[20] JONSSON, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** 1–38. MR650926

[21] KHORUNZHY, A. M., KHORUZHENKO, B. A. and PASTUR, L. A. (1996). Asymptotic properties of large random matrices with independent entries. *J. Math. Phys.* **37** 5033–5060. MR1411619

[22] MINGO, J. A. and SPEICHER, R. (2006). Second order freeness and fluctuations of random matrices. I. Gaussian and Wishart matrices and cyclic Fock spaces. *J. Funct. Anal.* **235** 226–270. MR2216446

[23] MOUSTAKAS, A. L., SIMON, S. H. and SENGUPTA, A. M. (2003). MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not so) large N analysis. *IEEE Trans. Inform. Theory* **49** 2545–2561. MR2028000

[24] SENGUPTA, A. M. and MITRA, P. (2008). Capacity of multivariate channels with multiplicative noise: I. Random matrix techniques and large-n expansions for full transfer matrices. Available at http://arxiv.org/abs/physics/0010081.

[25] SILVERSTEIN, J. W. (1995). Strong convergence of the empirical distribution of eigenvalues of large-dimensional random matrices. *J. Multivariate Anal.* **55** 331–339. MR1370408

[26] SILVERSTEIN, J. W. and BAI, Z. D. (1995). On the empirical distribution of eigenvalues of a class of large-dimensional random matrices. *J. Multivariate Anal.* **54** 175–192. MR1345534

[27] SINAI, Y. and SOSHNIKOV, A. (1998). Central limit theorem for traces of large random symmetric matrices with independent matrix elements. *Bol. Soc. Brasil. Mat.* (*N.S.*) **29** 1–24. MR1620151

[28] SMITHIES, F. (1958). *Integral Equations. Cambridge Tracts in Mathematics and Mathematical Physics* **49**. Cambridge Univ. Press, New York. MR0104991

[29] SOSHNIKOV, A. (2000). The central limit theorem for local linear statistics in classical compact groups and related combinatorial identities. *Ann. Probab.* **28** 1353–1370. MR1797877

[30] TARICCO, G. (2006). On the capacity of separately-correlated MIMO Rician fading channels. In *Proc. 49th Annual IEEE Globecom Conference, 2006*. San Francisco.

[31] TELATAR, I. (1999). Capacity of multi-antenna Gaussian channel. *European Transactions on Telecommunications*.

[32] TSE, D. and ZEITOUNI, O. (2000). Linear multiuser receivers in random environments. *IEEE Trans. Inform. Theory* **46** 171–188.

[33] TSE, D. N. C. and HANLY, S. V. (1999). Linear multiuser receivers: Effective interference, effective bandwidth and user capacity. *IEEE Trans. Inform. Theory* **45** 641–657. MR1677023

[34] TULINO, A. and VERDÚ, S. (2004). Random matrix theory and wireless communications. In *Foundations and Trends in Communications and Information Theory* **1** 1–182. Now Publishers.

[35] TULINO, A. and VERDU, S. (2005). Asymptotic outage capacity of multiantenna channels. In *Proceedings of IEEE ICASSP ’05* **5** 828–835.

W. HACHEM
J. NAJIM
CNRS
TÉLÉCOM PARISTECH
46 RUE BARRAULT
75013 PARIS
FRANCE
E-MAIL: walid.hachem@enst.fr
najim@enst.fr

P. LOUBATON
IGM LABINFO
UMR 8049
INSTITUT GASPARD MONGE
UNIVERSITÉ DE MARNE LA VALLÉE
FRANCE
AND
5 BD DESCARTES
CHAMPS SUR MARNE
77454 MARNE LA VALLÉE CEDEX 2
FRANCE
E-MAIL: loubaton@univ-mlv.fr