



HAL
open science

Gaussian fluctuations for linear spectral statistics of large random covariance matrices

Jamal Najim, Jianfeng Yao

► **To cite this version:**

Jamal Najim, Jianfeng Yao. Gaussian fluctuations for linear spectral statistics of large random covariance matrices. 2013. hal-00861793v2

HAL Id: hal-00861793

<https://hal.science/hal-00861793v2>

Preprint submitted on 31 Dec 2013 (v2), last revised 26 May 2015 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GAUSSIAN FLUCTUATIONS FOR LINEAR SPECTRAL STATISTICS OF LARGE RANDOM COVARIANCE MATRICES

JAMAL NAJIM AND JIANFENG YAO

ABSTRACT. Consider a $N \times n$ matrix $\Sigma_n = \frac{1}{\sqrt{n}} R_n^{1/2} X_n$, where R_n is a nonnegative definite Hermitian matrix and X_n is a random matrix with i.i.d. real or complex standardized entries. The fluctuations of the linear statistics of the eigenvalues:

$$\text{Trace } f(\Sigma_n \Sigma_n^*) = \sum_{i=1}^N f(\lambda_i), \quad (\lambda_i) \text{ eigenvalues of } \Sigma_n \Sigma_n^*,$$

are shown to be gaussian, in the regime where both dimensions of matrix Σ_n go to infinity at the same pace and in the case where f is of class C^3 , i.e. has three continuous derivatives. The main improvements with respect to Bai and Silverstein's CLT [6] are twofold: First, we consider general entries with finite fourth moment, but whose fourth cumulant is non-null, i.e. whose fourth moment may differ from the moment of a (real or complex) Gaussian random variable. As a consequence, extra terms proportional to

$$|\mathcal{V}|^2 = |\mathbb{E}(X_{11}^n)^2|^2 \quad \text{and} \quad \kappa = \mathbb{E}|X_{11}^n|^4 - |\mathcal{V}|^2 - 2$$

appear in the limiting variance and in the limiting bias, which not only depend on the spectrum of matrix R_n but also on its eigenvectors. Second, we relax the analyticity assumption over f by representing the linear statistics with the help of Helffer-Sjöstrand's formula.

The CLT is expressed in terms of vanishing Lévy-Prohorov distance between the linear statistics' distribution and a Gaussian probability distribution, the mean and the variance of which depend upon N and n and may not converge.

AMS 2000 subject classification: Primary 15A52, Secondary 15A18, 60F15.

Key words and phrases: large random matrix, fluctuations, linear statistics of the eigenvalues, central limit theorem.

CONTENTS

1. Introduction	2
2. General background - Variance and bias formulas	5
3. Statement of the CLT for the trace of the resolvent	10
4. Statement of the CLT for non-analytic functionals	15
5. Proof of Theorem 3.1 (CLT for the trace of the resolvent)	17
6. Proof of Theorem 4.3 (CLT for non-analytic functionals)	31
7. Proof of Theorem 4.4 (bias for non-analytic functionals)	34

Date: December 31, 2013.

Appendix A. Remaining proofs	35
References	43

1. INTRODUCTION

Consider a $N \times n$ random matrix $\Sigma_n = (\xi_{ij}^n)$ given by:

$$\Sigma_n = \frac{1}{\sqrt{n}} R_n^{1/2} X_n, \quad (1.1)$$

where $N = N(n)$ and R_n is a $N \times N$ nonnegative definite hermitian matrix with spectral norm uniformly bounded in N . The entries $(X_{ij}^n; i \leq N, j \leq n, n \geq 1)$ of matrices (X_n) are real or complex, independent and identically distributed (i.i.d.) with mean 0 and variance 1. Matrix $\Sigma_n \Sigma_n^*$ models a sample covariance matrix, formed from n samples of the random vector $R_n^{1/2} X_n$, with the population covariance matrix R_n .

Since the seminal work of Marčenko and Pastur [38] in 1967, the study of the spectrum of large covariance matrices of the type $X_n X_n^*$ under the asymptotic regime where:

$$N, n \rightarrow \infty \quad \text{and} \quad 0 < \liminf \frac{N}{n} \leq \limsup \frac{N}{n} < \infty, \quad (1.2)$$

(a condition that will be simply referred as $N, n \rightarrow \infty$ in the sequel) has drawn a considerable interest.

In this article, we study the fluctuations of linear spectral statistics of the form:

$$\text{tr} f(\Sigma_n \Sigma_n^*) = \sum_{i=1}^N f(\lambda_i), \quad \text{as } N, n \rightarrow \infty \quad (1.3)$$

where $\text{tr}(A)$ refers to the trace of A and the λ_i 's are the eigenvalues of $\Sigma_n \Sigma_n^*$. This subject has a rich history with contributions by Arharov [5], Girko (see [21, 22] and the references therein), Jonsson [33], Khorunzhiy et al. [36], Johansson [32], Sinai and Soshnikov [43, 44], Cabanal-Duvillard [15], Guionnet [24], Bai and Silverstein [6], Anderson and Zeitouni [4], Pan and Zhou [39], Chatterjee [17], Lytova and Pastur [37], Bai et al. [8], Shcherbina [42], etc. There are also more recent contributions for heavytailed entries (see for instance Benaych-Georges et al. [10]).

In their '04 article [6], Bai and Silverstein established a CLT for the linear spectral statistics (1.3) as the dimensions N and n grow to infinity at the same pace ($N/n \rightarrow c \in (0, \infty)$) and under two important assumptions:

- (1) The entries (X_{ij}^n) are centered with unit variance and a finite fourth moment equal to the fourth moment of a (real or complex) gaussian standard variable.
- (2) Function f in (1.3) is analytic in a neighbourhood of the asymptotic spectrum of $\Sigma_n \Sigma_n^*$.

Such a result proved to be highly useful in probability theory, statistics and various other fields.

The purpose of this article is to establish a CLT for linear spectral statistics (1.3) for general entries X_{ij}^n with fourth moment finite and for non-analytic functions f with sufficient bounded derivatives, hence to relax both Assumptions (1) and (2) in [6].

It is well known since the paper by Khorunzhiy et al. [36] that if the fourth moment of the entries differs from the fourth moment of a Gaussian random variable, then a term appears in the variance of the trace of the resolvent, which is proportional to the fourth cumulant of the entries. This term does not appear if Assumption (1) holds true, because in this case, the fourth cumulant is zero.

In Pan and Zhou [39], Assumption (1) has been relaxed under an additional assumption on matrix R_n , which somehow enforces structural conditions on R_n (in particular, these conditions are satisfied if matrix R_n is diagonal). In Hachem et al. [35, 27], CLTs have been established for specific linear statistics of interest in information theory, with general entries and (possibly non-centered) covariance random matrices with a variance profile. In Bao et al. [9], the CLT is established for the white model with general entries with fourth finite moment, featuring terms in the covariance proportional to the square of the second non-absolute moment and to the fourth cumulant.

In Lytova and Pastur [37] and Shcherbina [42], both assumptions have been relaxed for the "white" model, when R_n is equal to the identity matrix. In this case, it has been proved that mild integrability conditions over the Fourier transform of f was enough to establish the CLT. In Bai et al. [8], fluctuations for the white model are addressed as well, for non-analytic functions f . Following Shcherbina's ideas, Guédon et al. [23] establish a CLT for linear statistics of large covariance matrices with vectors with log-concave distribution. Following Lytova and Pastur, Yao [46] relaxes the analyticity assumption in [6] by using interpolation techniques and Fourier transforms. We follow here a different approach, inspired from Bordenave [12].

Non-Gaussian entries. In the case where matrix R_n is not diagonal, interesting phenomena appear when considering entries with non-Gaussian fourth moment: terms proportionals to the fourth cumulant and to $|\mathbb{E}(X_{11}^n)^2|^2$ appear in the variance; however their convergence is not granted under usual assumptions (roughly, under the convergence of R_n 's spectrum), mainly because these extra-terms also depend on the eigenvectors of R_n . As a consequence, such terms may not converge unless some very strong structural assumption over R_n (such as R_n diagonal) is made, which would ensure a joint convergence of R_n 's spectrum and eigenvectors¹. A careful description of the asymptotic variance is provided in Section 2.3. This lack of convergence has consequences on the description of the fluctuations.

Denote by $L_n(f)$ the (approximately) centered version of the linear statistics (1.3), to be properly defined below. Instead of expressing the CLT in the usual way, i.e. ($\xrightarrow{\mathcal{D}}$ stands for the convergence in distribution):

$$L_n(f) \xrightarrow[N, n \rightarrow \infty]{\mathcal{D}} \mathcal{N}(\mathcal{B}_\infty, \Theta_\infty) , \quad (1.4)$$

for some well-defined parameters $\mathcal{B}_\infty, \Theta_\infty$, we prove that the distribution of the linear statistics $L_n(f)$ becomes close to a family of Gaussian distributions, whose parameters (mean and variance) may not converge. More precisely, we establish that there exists a family of

¹Notice that this absence of convergence is not observed in [37, 9] because in the case where R_n is equal to the identity I_N , there are no issues with its eigenvectors.

Gaussian random variables $\mathcal{N}(\mathcal{B}_n^f, \Theta_n^f)$, such that

$$d_{\mathcal{L}P}(L_n(f), \mathcal{N}(\mathcal{B}_n^f, \Theta_n^f)) \xrightarrow[N, n \rightarrow \infty]{} 0, \quad (1.5)$$

where $d_{\mathcal{L}P}$ denotes the Lévy-Prohorov distance (and in particular metrizes the convergence of laws). Details are provided in Section 2.5 and the fluctuation results are stated in Theorem 3.1 (for the resolvent $f(\lambda) = (\lambda - z)^{-1}$) and Theorem 4.3 (for f of class C^3 , the space of functions with third continuous derivative).

This framework may also prove to be useful for other interesting models such as large dimensional information-plus-noise type matrices [18, 28] and more generally mixed models combining large dimensional deterministic and random matrices.

Expressing the CLT as in (1.5) makes it possible to avoid any cumbersome assumption related to the joint convergence of R_n 's eigenvectors and eigenvalues; the technical price to pay however is the need to get various uniform (in N, n) controls over the sequence $\mathcal{N}(\mathcal{B}_n, \Theta_n)$. This is achieved by introducing a matrix meta-model in Section 2.6. The case where matrix R_n is diagonal is simpler and the fluctuations express in the usual way (1.4); it is handled in Section 3.4. Remarks on the white case ($R_n = I_N$) are provided in Section 3.5.

Non-analytic functions. In Section 3, we establish the CLT for the resolvent

$$\text{tr}(\Sigma_n \Sigma_n^* - z I_N)^{-1}.$$

In order to transfer the CLT from the resolvent to the linear statistics of the eigenvalues $\text{tr} f(\Sigma_n \Sigma_n^*)$, we will use (Dynkin-)Helffer-Sjöstrand's representation formula² for a function f of class C^{k+1} and with compact support [20, 31]. Denote by $\Phi(f) : \mathbb{C}^+ \rightarrow \mathbb{C}$ the function:

$$\Phi(f)(x + iy) = \sum_{\ell=0}^k \frac{(iy)^\ell}{\ell!} f^{(\ell)}(x) \chi(y), \quad (1.6)$$

where $\chi : \mathbb{R} \rightarrow \mathbb{R}^+$ is smooth, compactly supported, with value 1 in a neighbourhood of 0 and let $\bar{\partial} = \partial_x + i\partial_y$. Helffer-Sjöstrand's formula writes:

$$\text{tr} f(\Sigma_n \Sigma_n^*) = \frac{1}{\pi} \text{Re} \int_{\mathbb{C}^+} \bar{\partial} \Phi(f)(z) \text{tr}(\Sigma_n \Sigma_n^* - z I_N)^{-1} \lambda_2(dz), \quad (1.7)$$

where λ_2 stands for the Lebesgue measure over \mathbb{C}^+ . Formula (1.7) has already been used in the context of large random matrix theory in the book by Anderson et al. [3, Eq. (5.5.11)]; see also the nice probabilistic interpretation in [2] and an elementary proof of it in [13, Chap. 5].

We believe that representation formula (1.7) provides a very streamlined way to handle non-analytic functions and in fact enables us to state the fluctuations for the linear statistics for functions of class C^3 , a lower regularity requirement than in [8, 37, 42, 46].

²In [30, Notes of chap. 8], it is written "This formula is due to Dynkin but was popularized by Helffer and Sjöstrand in the context of spectral theory, leading many authors to call it the Helffer-Sjöstrand formula".

Bias in the CLT and asymptotic expansion for the linear spectral statistics.

In order to center the linear spectral statistics $\text{tr} f(\Sigma_n \Sigma_n^*)$, we consider the (first order) expansion of $\frac{1}{N} \mathbb{E} \text{tr} f(\Sigma_n \Sigma_n^*)$:

$$\frac{1}{N} \mathbb{E} \text{tr} f(\Sigma_n \Sigma_n^*) = \mathcal{E}_{0,n}(f) + \mathcal{O}\left(\frac{1}{N}\right),$$

where $\mathcal{E}_{0,n}(f)$ is $\mathcal{O}(1)$, and define $L_n(f)$ as:

$$L_n(f) = \text{tr} f(\Sigma_n \Sigma_n^*) - N \mathcal{E}_{0,n}(f).$$

A precise description of $L_n(f)$ is provided in Section 2.4. In order to fully characterize the fluctuations of $L_n(f)$, we must study the second order expansion of $\frac{1}{N} \mathbb{E} \text{tr} f(\Sigma_n \Sigma_n^*)$:

$$\frac{1}{N} \mathbb{E} \text{tr} f(\Sigma_n \Sigma_n^*) = \mathcal{E}_{0,n}(f) + \frac{\mathcal{E}_{1,n}(f)}{N} + o\left(\frac{1}{N}\right),$$

which will naturally yield the bias of $L_n(f)$, as $\mathbb{E} L_n(f) = \mathcal{E}_{1,n}(f) + o(1)$. Asymptotic expansions for various matrix ensembles have already been studied, see for instance Pastur et al. [1], Bai and Silverstein [6], Haagerup and Thorbjørnsen [25, 26], Schultz [41], Capitaine and Donati-Martin [16], Vallet et al. [45], Hachem et al. [29], etc.

The asymptotic bias is expressed in Theorem 3.1 for the resolvent. In order to lift asymptotic expansions from the resolvent to smooth functions, we combine gaussian interpolation with ideas from Haagerup and Thorbjørnsen [25] and Loubaton et al. [29, 45]. For smooth functions, the statement is given in Theorem 4.4. Somehow surprisingly, the conditions over function f are stronger for the asymptotic expansion to hold than for the CLT as function f needs to be of class C^{18} (cf. Remark 4.1).

Acknowledgement. We are particularly indebted to Charles Bordenave who drew our attention to Helffer-Sjöstrand's formula and related variance estimates (cf. Proposition 6.2), which substantially shorten the initial proof of fluctuations for non-analytic functions; we would also like to thank Djilil Chafaï and Walid Hachem for fruitful discussions.

2. GENERAL BACKGROUND - VARIANCE AND BIAS FORMULAS

2.1. Assumptions. Recall the asymptotic regime where $N, n \rightarrow \infty$, cf. (1.2), and denote by

$$c_n = \frac{N}{n}, \quad \ell^- \triangleq \liminf \frac{N}{n} \quad \text{and} \quad \ell^+ \triangleq \limsup \frac{N}{n}.$$

Assumption A-1. *The random variables $(X_{ij}^n; 1 \leq i \leq N(n), 1 \leq j \leq n, n \geq 1)$ are independent and identically distributed. They satisfy:*

$$\mathbb{E} X_{ij}^n = 0, \quad \mathbb{E} |X_{ij}^n|^2 = 1 \quad \text{and} \quad \mathbb{E} |X_{ij}^n|^4 < \infty.$$

Assumption A-2. *Consider a sequence (R_n) of deterministic, nonnegative definite hermitian $N \times N$ matrices, with $N = N(n)$. The sequence $(R_n, n \geq 1)$ is bounded for the spectral norm as $N, n \rightarrow \infty$:*

$$\sup_{n \geq 1} \|R_n\| < \infty.$$

In particular, we will have:

$$0 \leq \lambda_R^- \triangleq \liminf_{N, n \rightarrow \infty} \|R_n\| \leq \lambda_R^+ \triangleq \limsup_{N, n \rightarrow \infty} \|R_n\| < \infty.$$

2.2. Resolvent, canonical equation and deterministic equivalents. Denote by $Q_n(z)$ (resp. \tilde{Q}_n) the resolvent of matrix $\Sigma_n \Sigma_n^*$ (resp. of $\Sigma_n^* \Sigma_n$):

$$Q_n(z) = (\Sigma_n \Sigma_n^* - zI_N)^{-1}, \quad \tilde{Q}_n(z) = (\Sigma_n^* \Sigma_n - zI_n)^{-1}, \quad (2.1)$$

and by $f_n(z)$ and $\tilde{f}_n(z)$ their normalized traces which are the Stieltjes transforms of the empirical distribution of $\Sigma_n \Sigma_n^*$'s and $\Sigma_n^* \Sigma_n$'s eigenvalues:

$$f_n(z) = \frac{1}{N} \text{tr} Q_n(z), \quad \tilde{f}_n(z) = \frac{1}{n} \text{tr} \tilde{Q}_n(z). \quad (2.2)$$

The following canonical equation³ admits a unique solution t_n in the class of Stieltjes transforms of probability measures (see for instance [6]):

$$t_n(z) = \frac{1}{N} \text{tr} (-zI_N + (1 - c_n)R_n - zc_n t_n(z)R_n)^{-1}, \quad z \in \mathbb{C} \setminus \mathbb{R}^+, \quad (2.3)$$

where c_n stands for the ratio N/n . The function t_n being introduced, we can define the following $N \times N$ matrix

$$T_n(z) = (-zI_N + (1 - c_n)R_n - zc_n t_n(z)R_n)^{-1}. \quad (2.4)$$

Matrix $T_n(z)$ can be thought of as a *deterministic equivalent* of the resolvent $Q_n(z)$ in the sense that it approximates the resolvent in various senses. For instance,

$$\frac{1}{N} \text{tr} T_n(z) - \frac{1}{N} \text{tr} Q_n(z) \xrightarrow{N, n \rightarrow \infty} 0$$

(in probability or almost surely). Otherwise stated, $t_n(z) = N^{-1} \text{tr} T_n(z)$ is the deterministic equivalent of $f_n(z)$. As we shall see later in this paper,

$$u_n^* Q_n v_n - u_n^* T_n v_n \xrightarrow{N, n \rightarrow \infty} 0 \quad (2.5)$$

where (u_n) and (v_n) are deterministic $N \times 1$ vectors with uniformly bounded euclidian norms in N . As a consequence of (2.5), not only T_n conveys information on the limiting spectrum of the resolvent Q_n but also on the eigenvectors of Q_n .

If $R_n = I_N$, then t_n is simply the Stieltjes transform of Marčenko-Pastur distribution [38] with parameter c_n .

2.3. Entries with non-null fourth cumulant and the limiting covariance for the trace of the resolvent. In [6], an important preliminary step to establish the CLT for linear statistics is to compute the CLT for the trace of the resolvent. Let \mathcal{V} be the second moment of the random variable X_{ij} and κ its fourth cumulant:

$$\mathcal{V} = \mathbb{E}(X_{ij}^n)^2 \quad \text{and} \quad \kappa = \mathbb{E}|X_{ij}^n|^4 - |\mathcal{V}|^2 - 2.$$

If the entries are real or complex standard Gaussian, then $\kappa = 0$ and $\mathcal{V} = 0$ or 1. Otherwise the fourth cumulant is a priori no longer equal to zero. This induces extra-terms in the computation of the limiting variance, mainly due to the following (\mathcal{V}, κ) -dependent identity:

$$\mathbb{E}(X_{\cdot 1}^* A X_{\cdot 1} - \text{tr} A)(X_{\cdot 1}^* B X_{\cdot 1} - \text{tr} B) = \text{tr} AB + |\mathcal{V}|^2 \text{tr} AB^T + \kappa \sum_{i=1}^N A_{ii} B_{ii}, \quad (2.6)$$

where $X_{\cdot 1}$ stands for the first column (of dimension $N \times 1$) of matrix X_n and where A, B are deterministic $N \times N$ matrices. As a consequence, there will be three terms in the limiting

³We borrow the name "canonical equation" from V.L. Girko who established in [21, 22] canonical equations associated to various models of large random matrices.

covariance of the quantity (1.3); one will raise from the first term of the right hand side (r.h.s.) of (2.6), a second one will be proportional to $|\mathcal{V}|^2$, and a third one will be proportional to κ . As will be described in the sequel, the two last terms behave differently than the first one. In order to describe these terms, we first need to introduce more notations. Let

$$\tilde{t}_n(z) = -\frac{1-c_n}{z} + c_n t_n(z). \quad (2.7)$$

The quantity $\tilde{t}_n(z)$ is the deterministic equivalent associated to $n^{-1} \text{tr}(\Sigma_n^* \Sigma_n - z I_N)^{-1}$. Denote by R_n^T the transpose matrix of R_n (notice that since R_n is hermitian, $R_n^T = \bar{R}_n$ and we shall use this later notation) and by T_n^T , the transpose matrix⁴ of T_n :

$$T_n^T(z) = (-z I_N + (1-c_n) \bar{R}_n - z c_n t_n(z) \bar{R}_n)^{-1}; \quad (2.8)$$

notice that the definition of $t_n(z)$ in (2.3) does not change if R_n is replaced by \bar{R}_n since the spectrum of both matrices R_n and \bar{R}_n is the same. We can now describe the limiting covariance of the trace of the resolvent, which is a key step in Bai and Silverstein's approach:

$$\begin{aligned} \text{cov}(\text{tr} Q_n(z_1), \text{tr} Q_n(z_2)) &= \Theta_{0,n}(z_1, z_2) + |\mathcal{V}|^2 \Theta_{1,n}(z_1, z_2) + \kappa \Theta_{2,n}(z_1, z_2) + o(1) \\ &\triangleq \Theta_n(z_1, z_2) + o(1), \end{aligned} \quad (2.9)$$

where $o(1)$ is a term that converges to zero as $N, n \rightarrow \infty$ and

$$\Theta_{0,n}(z_1, z_2) \triangleq \left\{ \frac{\tilde{t}'_n(z_1) \tilde{t}'_n(z_2)}{(\tilde{t}_n(z_1) - \tilde{t}_n(z_2))^2} - \frac{1}{(z_1 - z_2)^2} \right\} \quad (2.10)$$

$$\Theta_{1,n}(z_1, z_2) \triangleq \frac{\partial}{\partial z_2} \left\{ \frac{\partial \mathcal{A}_n(z_1, z_2)}{\partial z_1} \frac{1}{1 - |\mathcal{V}|^2 \mathcal{A}_n(z_1, z_2)} \right\} \quad (2.11)$$

$$\Theta_{2,n}(z_1, z_2) \triangleq \frac{z_1^2 z_2^2 \tilde{t}'_n(z_1) \tilde{t}'_n(z_2)}{n} \sum_{i=1}^N \left(R_n^{1/2} T_n^2(z_1) R_n^{1/2} \right)_{ii} \left(R_n^{1/2} T_n^2(z_2) R_n^{1/2} \right)_{ii} \quad (2.12)$$

with

$$\mathcal{A}_n(z_1, z_2) = \frac{z_1 z_2}{n} \tilde{t}_n(z_1) \tilde{t}_n(z_2) \text{tr} \left\{ R_n^{1/2} T_n(z_1) R_n^{1/2} \bar{R}_n^{1/2} T_n^T(z_2) \bar{R}_n^{1/2} \right\}. \quad (2.13)$$

At first sight, these formulas (established in Section 5) may seem complicated; however, much information can be inferred from them.

The term $\Theta_{0,n}$. This term is familiar as it already appears in Bai and Silverstein's CLT [6]. Notice that the quantities \tilde{t}_n and \tilde{t}'_n only depend on the spectrum of matrix R_n . Hence, under the additional assumption that:

$$c_n \xrightarrow{N, n \rightarrow \infty} c \in (0, \infty) \quad \text{and} \quad F^{R_n} \xrightarrow{N, n \rightarrow \infty, \mathcal{L}} F^{\mathbf{R}}, \quad (2.14)$$

where F^{R_n} denotes the empirical distribution of R_n 's eigenvalues and $F^{\mathbf{R}}$ is a probability measure, it can easily be proved that

$$\Theta_{0,n}(z_1, z_2) \xrightarrow{N, n \rightarrow \infty} \Theta_0(z_1, z_2) = \left\{ \frac{\tilde{t}'(z_1) \tilde{t}'(z_2)}{(\tilde{t}(z_1) - \tilde{t}(z_2))^2} - \frac{1}{(z_1 - z_2)^2} \right\}, \quad (2.15)$$

where \tilde{t}, \tilde{t}' are the limits of $\tilde{t}_n, \tilde{t}'_n$ under (2.14).

⁴Beware that T_n^T is not the entry-wise conjugate of T_n , due to the presence of z .

The term $\Theta_{1,n}$. The interesting phenomenon lies in the fact that this term involves products of matrices $R_n^{1/2}$ and its conjugate $\bar{R}_n^{1/2}$. These matrices have the same spectrum but conjugate eigenvectors. If R_n is not real, the convergence of $\Theta_{1,n}$ is not granted, even under (2.14). If however R_n and X_n 's entries are real, i.e. $\mathcal{V} = 1$, then it can be easily proved that $\Theta_{0,n} = \Theta_{1,n}$ hence the factor 2 in [6] between the complex and the real covariance.

The term $\Theta_{2,n}$. This term involves quantities of the type $(R_n^{1/2} T_n R_n^{1/2})_{ii}$ which not only depend on the spectrum of matrix R_n but also on its eigenvectors. As a consequence, the convergence of such terms does not follow from an assumption such as (2.14), except in some particular cases (for instance if R_n is diagonal) and any assumption which enforces the convergence of such terms (as for instance in [39, Theorem 1.4]) implicitly implies an asymptotic joint behaviour between R_n 's eigenvectors and eigenvalues. We shall adopt a different point of view here and will not assume the convergence of these quantities.

2.4. Representation of the linear statistics and limiting bias. Recall that $t_n(z)$ is the Stieltjes transform of a probability measure \mathcal{F}_n :

$$t_n(z) = \int_{\mathcal{S}_n} \frac{\mathcal{F}_n(d\lambda)}{\lambda - z} \quad (2.16)$$

with support \mathcal{S}_n included in a compact set. The purpose of this article is to describe the fluctuations of the linear statistics

$$L_n(f) = \sum_{i=1}^N f(\lambda_i) - N \int f(\lambda) \mathcal{F}_n(d\lambda) \quad (2.17)$$

as $N, n \rightarrow \infty$.

For a smooth enough function f with bounded support (a mild condition as we shall see in Section 4), one can rely on Helffer-Sjöstrand's formula and write:

$$\begin{aligned} L_n(f) &= \operatorname{tr} f(\Sigma_n \Sigma_n^*) - N \int f(\lambda) \mathcal{F}_n(d\lambda) \\ &= \frac{1}{\pi} \operatorname{Re} \int_{\mathbb{C}^+} \bar{\partial} \Phi(f)(z) \{ \operatorname{tr} Q_n(z) - N t_n(z) \} \lambda_2(dz) . \end{aligned} \quad (2.18)$$

where $\Phi(f)$ is defined in (1.6) and the last equality follows from the fact that

$$\int f(\lambda) \mathcal{F}_n(d\lambda) = \frac{1}{\pi} \operatorname{Re} \int_{\mathbb{C}^+} \bar{\partial} \Phi(f)(z) t_n(z) \lambda_2(dz) ,$$

Based on (2.18), we shall first study the fluctuations of:

$$\operatorname{tr} Q_n(z) - N t_n(z) = \{ \operatorname{tr} Q_n(z) - \mathbb{E} \operatorname{tr} Q_n(z) \} + \{ \mathbb{E} \operatorname{tr} Q_n(z) - N t_n(z) \}$$

for $z \in \mathbb{C}^+$. The first difference in the r.h.s. will yield the fluctuations with a covariance $\Theta_n(z_1, z_2)$ described in (2.9) while the second difference, deterministic, will yield the bias:

$$\begin{aligned} \mathbb{E} \operatorname{tr} Q_n(z) - N t_n(z) &= |\mathcal{V}|^2 \mathcal{B}_{1,n}(z) + \kappa \mathcal{B}_{2,n}(z) + o(1) \\ &\triangleq \mathcal{B}_n(z) + o(1) \end{aligned} \quad (2.19)$$

where

$$\mathcal{B}_{1,n}(z) \triangleq -z^3 \tilde{t}_n^3 \frac{\frac{1}{n} \text{tr} R_n^{1/2} T_n^2(z) R_n^{1/2} \bar{R}_n^{1/2} T_n^T(z) \bar{R}_n^{1/2}}{\left(1 - z^2 \tilde{t}_n^2 \frac{1}{n} \text{Tr} R_n^2 T_n^2\right) \left(1 - |\mathcal{V}|^2 z^2 \tilde{t}_n^2 \frac{1}{n} \text{Tr} R_n^{1/2} T_n(z) R_n^{1/2} \bar{R}_n^{1/2} T_n^T(z) \bar{R}_n^{1/2}\right)} \quad (2.20)$$

$$\mathcal{B}_{2,n}(z) \triangleq -z^3 \tilde{t}_n^3 \frac{\frac{1}{n} \sum_{i=1}^N \left(R_n^{1/2} T_n R_n^{1/2}\right)_{ii} \left(R_n^{1/2} T_n^2 R_n^{1/2}\right)_{ii}}{1 - z^2 \tilde{t}_n^2 \frac{1}{n} \text{tr} R_n^2 T_n^2} \quad (2.21)$$

The discussions on the terms $\Theta_{1,n}$ and $\Theta_{2,n}$ also apply to the terms $\mathcal{B}_{1,n}$ and $\mathcal{B}_{2,n}$ (whose expressions are established in Section 5) which are likely not to converge for similar reasons.

2.5. Gaussian processes and the central limit theorem. A priori, the mean \mathcal{B}_n and covariance Θ_n of $(\text{tr} Q_n - N t_n)$ do not converge. Hence, we shall express the Gaussian fluctuations of the linear statistics (2.17) in the following way: we first prove the existence of a family $(N_n(z), z \in \mathcal{C})_{n \in \mathbb{N}}$ of tight Gaussian processes with mean and covariance:

$$\begin{aligned} \mathbb{E} N_n(z) &= \mathcal{B}_n(z) , \\ \text{cov}(N_n(z_1), N_n(z_2)) &= \Theta_n(z_1, z_2) . \end{aligned}$$

We then express the fluctuations of the centralized trace as

$$d_{\mathcal{L}P}((\text{tr} Q_n(z) - N t_n(z)), N_n(z)) \xrightarrow[N, n \rightarrow \infty]{} 0 .$$

with $d_{\mathcal{L}P}$ the Lévy-Prohorov distance between P and Q probability measures over borel sets of $\mathbb{R}, \mathbb{R}^d, \mathbb{C}$ or \mathbb{C}^d :

$$d_{\mathcal{L}P}(P, Q) = \inf \{ \varepsilon > 0, P(A) \leq Q(A^\varepsilon) + \varepsilon \text{ for all Borel sets } A \} , \quad (2.22)$$

where A^ε is an ε -blow up of A (cf. [19, Section 11.3] for more details). If X is a random variable and $\mathcal{L}(X)$ its distribution, denote (with a slight abuse of notation) by $d_{\mathcal{L}P}(X, Y) \triangleq d_{\mathcal{L}P}(\mathcal{L}(X), \mathcal{L}(Y))$.

Similarly, we will express the fluctuations of $L_n(f)$ as:

$$d_{\mathcal{L}P}(L_n(f), \mathcal{N}_n(f)) \xrightarrow[N, n \rightarrow \infty]{} 0 ,$$

where $\mathcal{N}_n(f)$ is a well-identified gaussian random variable.

2.6. A meta-model argument. As we need to cope with a sequence of Gaussian processes (N_n) instead of a single one, it will be necessary to establish various properties uniform in n, N such as:

- (1) the tightness of the sequence (N_n) (cf. Section 5.2);
- (2) a uniform bound over the variances of $(\text{Tr} N_n(z))$ (cf. Section 6.1), needed to extend the CLT to non-analytic functionals;
- (3) a uniform bound over the biases of $(\text{Tr} N_n(z))$ (cf. Section 7.1), needed to compute the bias for non-analytic functionals.

A direct approach based on the mere definition of process N_n 's parameters seems difficult, mainly due to the definitions of Θ_n and \mathcal{B}_n which rely on quantities $(t_n$ and $\tilde{t}_n)$ defined as solutions of fixed-point equations. Since the previous properties will be established for the processes $(\text{Tr} Q_n - N t_n)$ anyway, the idea is to transfer them to N_n by means of the following matrix meta-model:

Let N , n and R_n be fixed and consider the $NM \times NM$ matrix

$$R_n(M) = \begin{pmatrix} R_n & 0 & \cdots \\ 0 & \ddots & 0 \\ \cdots & 0 & R_n \end{pmatrix}. \quad (2.23)$$

Matrix $R_n(M)$ is a block matrix with $N \times N$ diagonal blocks equal to R_n , and zero blocks elsewhere; for all $M \geq 1$ the spectral norm of $R_n(M)$ is equal to the spectral norm of R_n (which is fixed). In particular the sequence $(R_n(M); M \geq 1)$ with N, n fixed satisfies assumption (A-2) with $(R_n(M); M \geq 1)$ instead of (R_n) . Consider now the random matrix model:

$$\Sigma_n(M) = \frac{1}{\sqrt{Mn}} R_n(M)^{1/2} X_n(M) \quad (2.24)$$

where $X_n(M)$ is a $MN \times Mn$ matrix with i.i.d. random entries with the same distribution as the X_{ij} 's and satisfying (A-1). The interest of introducing matrix $\Sigma_n(M)$ lies in the fact that matrices $\Sigma_n(M)\Sigma_n(M)^*$ and $\Sigma_n\Sigma_n^*$ have loosely speaking the same deterministic equivalents. Denote by t_n , T_n and \tilde{t}_n the deterministic equivalents of $\Sigma_n\Sigma_n^*$ as defined in (2.3), (2.4) and (3.1), and by $t_n(M)$, $T_n(M)$ and $\tilde{t}_n(M)$ their counterparts for the model $\Sigma_n(M)\Sigma_n(M)^*$. Taking advantage of the block structure of $R_n(M)$, a straightforward computation yields (N, n fixed):

$$\forall M \geq 1, \quad t_n(M) = t_n, \quad \tilde{t}_n(M) = \tilde{t}_n \quad \text{and} \quad T_n(M) = \begin{pmatrix} T_n & 0 & \cdots \\ & \ddots & \\ \cdots & 0 & T_n \end{pmatrix}.$$

Similarly, denote by $\mathcal{B}_{n,M}$ and $\Theta_{n,M}$ the quantities given by formulas (2.19) and (2.9) when replacing N , t_n , T_n and \tilde{t}_n by NM , $t_n(M)$, $T_n(M)$ and $\tilde{t}_n(M)$. Straightforward computation yields:

$$\forall M \geq 1, \quad \mathcal{B}_{n,M} = \mathcal{B}_n \quad \text{and} \quad \Theta_{n,M} = \Theta_n.$$

An interesting feature of this meta-model lies in the fact that all the quantities associated to $\Sigma_n(M)\Sigma_n(M)^*$ converge as $M \rightarrow \infty$ to the deterministic equivalents t_n , \tilde{t}_n , etc. As a consequence, one can easily transfer all the estimates obtained for

$$(\text{Tr}(\Sigma_n(M)\Sigma_n(M)^* - zI_{NM})^{-1} - NMt_n)$$

to the process (N_n) .

3. STATEMENT OF THE CLT FOR THE TRACE OF THE RESOLVENT

3.1. Further notations. If A is a $N \times N$ matrix with real eigenvalues, denote by F^A the empirical distribution of the eigenvalues $(\delta_i(A), i = 1 : N)$ of A , that is:

$$F^A(dx) = \frac{1}{n} \sum_{i=1}^N \delta_{\lambda_i(A)}(dx).$$

Recall the definitions of Q_n , t_n , T_n and \tilde{t}_n (cf. (2.1), (2.3), (2.4) and (2.7)). The following relations hold true (see for instance [6]):

$$T_n(z) = -\frac{1}{z} (I_N + \tilde{t}_n(z)R_n)^{-1} \quad \text{and} \quad \tilde{t}_n(z) = -\frac{1}{z(1 + \frac{1}{n}\text{tr} R_n T_n(z))}. \quad (3.1)$$

Recall the definition of \mathcal{F}_n in (2.16) and let similarly $\tilde{\mathcal{F}}_n$ be the probability distribution associated to \tilde{t}_n . The central object of study is the signed measure:

$$G_n = N\left(F^{\Sigma_n \Sigma_n^*} - \mathcal{F}_n\right) = n\left(F^{\Sigma_n^* \Sigma_n} - \tilde{\mathcal{F}}_n\right), \quad (3.2)$$

and its Stieltjes transform

$$M_n(z) = N(f_n(z) - t_n(z)) = n\left(\tilde{f}_n(z) - \tilde{t}_n(z)\right). \quad (3.3)$$

Denote by $o_P(1)$ any random variable which converges to zero in probability.

3.2. Truncation. In this section, we closely follow Bai and Silverstein [6]. We recall the framework developed there and introduce some additional notations.

3.2.1. Truncation of random variables. Consider a sequence of positive numbers (δ_n) which satisfies:

$$\delta_n \rightarrow 0, \quad \delta_n n^{1/4} \rightarrow \infty \quad \text{and} \quad \delta_n^{-4} \int_{\{|X_{11}| \geq \delta_n \sqrt{N}\}} |X_{11}|^4 \rightarrow 0$$

as $N, n \rightarrow \infty$. Let $\hat{\Sigma}_n = n^{-1/2} R_n^{1/2} \hat{X}_n$ where \hat{X}_n is a $N \times n$ matrix having (i, j) th entry $X_{ij} 1_{\{|X_{ij}| < \delta_n \sqrt{N}\}}$. This truncation step yields:

$$\mathbb{P}\left(\Sigma_n \Sigma_n^* \neq \hat{\Sigma}_n \hat{\Sigma}_n^*\right) \xrightarrow[N, n \rightarrow \infty]{} 0. \quad (3.4)$$

Define $\tilde{\Sigma}_n = n^{-1/2} R_n^{1/2} \tilde{X}_n$ where \tilde{X}_n is a $N \times n$ matrix having (i, j) th entry $(\hat{X}_{ij} - \mathbb{E}\hat{X}_{ij})/\sigma_n$, where $\sigma_n^2 = \mathbb{E}|\hat{X}_{ij} - \mathbb{E}\hat{X}_{ij}|^2$. Let $\hat{G}_n = N(F^{\hat{\Sigma}_n \hat{\Sigma}_n^*} - \mathcal{F}_n)$ and $\tilde{G}_n = N(F^{\tilde{\Sigma}_n \tilde{\Sigma}_n^*} - \mathcal{F}_n)$, then whenever f is of class C^1 , the following holds true (cf. [6]):

$$\forall \eta > 0, \quad \mathbb{P}\left\{\left|\int f d\hat{G}_n - \int f d\tilde{G}_n\right| > \eta\right\} \xrightarrow[N, n \rightarrow \infty]{} 0. \quad (3.5)$$

Combining (3.4) and (3.5), we obtain $\int f dG_n - \int f d\tilde{G}_n \rightarrow 0$ in probability. Moreover, the moments are asymptotically not affected by these different steps:

$$\max\left(\left|\mathbb{E}\tilde{X}_{ij}^2 - \mathbb{E}X_{ij}^2\right|; \left(\mathbb{E}|\tilde{X}_{ij}|^2 - 1\right); \left(\mathbb{E}|\tilde{X}_{ij}|^4 - \mathbb{E}|X_{ij}|^4\right)\right) \xrightarrow[N, n \rightarrow \infty]{} 0. \quad (3.6)$$

Note in particular that the fourth cumulant of \tilde{X}_{ij} converges to that of X_{ij} . Hence, it is sufficient to consider variables truncated at $\delta_n \sqrt{n}$, centralized and renormalized. This will be assumed in the sequel (we shall simply write X_{ij} and all related quantities with X_{ij} 's truncated, centralized, renormalized with no superscript any more).

3.3. The Central Limit Theorem for the resolvent. We extend below Bai and Silverstein's master lemma [6, Lemma 1.1]. Let A be large enough, say

$$A > \lambda_R^+ \left(1 + \sqrt{\ell^+}\right)^2.$$

Denote by D , D^+ and D_ε the domains:

$$\begin{aligned} D &= [0, A] + \mathbf{i}[0, 1], \\ D^+ &= [0, A] + \mathbf{i}(0, 1], \\ D_\varepsilon &= [0, A] + \mathbf{i}[\varepsilon, 1] \quad (\varepsilon > 0). \end{aligned} \quad (3.7)$$

Theorem 3.1. *Assume that (A-1) and (A-2) hold true, then*

(1) The process $\{M_n(\cdot)\}$ as defined in (3.3) forms a tight sequence on D_ε , more precisely:

$$\sup_{z_1, z_2 \in D_\varepsilon, n \geq 1} \frac{\mathbb{E}|M_n(z_1) - M_n(z_2)|^2}{|z_1 - z_2|^2} < \infty$$

(2) There exists a sequence $(N_n(z), z \in D^+)$ of two-dimensional Gaussian processes with mean

$$\mathbb{E}N_n(z) = |\mathcal{V}|^2 \mathcal{B}_{1,n}(z) + \kappa \mathcal{B}_{2,n}(z) \quad (3.8)$$

where $\mathcal{B}_{1,n}(z)$ and $\mathcal{B}_{2,n}(z)$ are defined in (2.20) and (2.21), and covariance:

$$\begin{aligned} \text{cov}(N_n(z_1), N_n(z_2)) &= \mathbb{E}(N_n(z_1) - \mathbb{E}N_n(z_1))(N_n(z_2) - \mathbb{E}N_n(z_2)) \\ &= \Theta_{0,n}(z_1, z_2) + |\mathcal{V}|^2 \Theta_{1,n}(z_1, z_2) + \kappa \Theta_{2,n}(z_1, z_2) \end{aligned}$$

where $\Theta_{0,n}$, $\Theta_{1,n}$ and $\Theta_{2,n}$ are defined in (2.9), (2.10)-(2.12). Moreover $(N_n(z), z \in D_\varepsilon)$ is tight.

(3) For any functional F from $C(D_\varepsilon; \mathbb{C})$ to \mathbb{C} ,

$$\mathbb{E}F(M_n) - \mathbb{E}F(N_n) \xrightarrow{N, n \rightarrow \infty} 0$$

- Remark 3.1.* (1) The tightness of the process $\{M_n\}$ immediately follows from Bai and Silverstein's lemma as this result has been proved in [6, Lemma 1.1] under Assumption (A-1) with no extra conditions on the moments of the entries.
- (2) Differences between Theorem 3.1 and [6, Lemma 1.1] appear in the bias and in the covariance where there are respectively two terms instead of one and three terms instead of one in [6, Lemma 1.1].
- (3) Since the extra terms do not converge, we need to consider a sequence of Gaussian processes instead of a single Gaussian process as in [6, Lemma 1.1].
- (4) In order to prove that the sequence of Gaussian processes is tight, we introduce a meta-matrix model to transfer the tightness of $\{M_n\}$ to $\{N_n\}$ (see for instance Section 5.2.1).
- (5) Following Bai and Silverstein [6], it is relatively straightforward with the help of Cauchy's formula to describe the fluctuations of $L_n(f)$ for f analytic with Theorem 3.1 at hand. We skip this step since we will directly extend the CLT to non-analytic functions f in Section 4.

Proof of Theorem 3.1 is postponed to Section 5.

The end of the section is devoted to various specializations of Theorem 3.1 in the case where matrix R_n is diagonal. In this case, the results are simpler to express and comparisons can easily be made with related works.

3.4. Covariance and bias in the special case of diagonal matrices (R_n) . This case partially falls into the framework developed in Pan and Zhou [39] (note that the case $\mathcal{V} \neq 0$ and 1 is not handled there). Matrix R_n being nonnegative definite hermitian, its entries are real positive if R_n is assumed to be diagonal. In this case, matrix T_n is diagonal as well (cf.

(2.4)), $T_n = T_n^T$ and simplifications occur for the following terms:

$$\begin{aligned} \mathcal{A}_n &= \frac{z_1 z_2}{n} \tilde{t}_n(z_1) \tilde{t}_n(z_2) \text{tr} R_n T_n(z_1) R_n T_n(z_2) , \\ \Theta_{2,n}(z_1, z_2) &= \frac{z_1^2 z_2^2 \tilde{t}'_n(z_1) \tilde{t}'_n(z_2)}{n} \text{tr} (R_n^2 T_n^2(z_1) T_n^2(z_2)) , \\ \mathcal{B}_{1,n}(z) &= -z^3 \tilde{t}_n^3 \frac{\frac{1}{n} \text{tr} R_n^2 T_n^3}{\left(1 - z^2 \tilde{t}_n^2 \frac{1}{n} \text{Tr} R_n^2 T_n^2\right) \left(1 - |\mathcal{V}|^2 z^2 \tilde{t}_n^2 \frac{1}{n} \text{Tr} R_n^2 T_n^2\right)} , \\ \mathcal{B}_{2,n}(z) &= -z^3 \tilde{t}_n^3 \frac{\frac{1}{n} \text{tr} R_n^2 T_n^3}{1 - z^2 \tilde{t}_n^2 \frac{1}{n} \text{tr} R_n^2 T_n^2} . \end{aligned}$$

As one may notice, all the terms in the variance and the bias now only depend on the spectrum of R_n . Hence, the following convergence holds true under the extra assumption (2.14):

$$\begin{aligned} \mathcal{A}_n(z_1, z_2) &\xrightarrow{N, n \rightarrow \infty} \mathcal{A}(z_1, z_2) = c \tilde{t}(z_1) \tilde{t}(z_2) \int \frac{\lambda^2 F^{\mathbf{R}}(d\lambda)}{(1 + \lambda \tilde{t}(z_1))(1 + \lambda \tilde{t}(z_2))} , \\ \Theta_{1,n}(z_1, z_2) &\xrightarrow{N, n \rightarrow \infty} \Theta_1(z_1, z_2) = \frac{\partial}{\partial z_2} \left\{ \frac{\partial \mathcal{A}(z_1, z_2)}{\partial z_1} \frac{1}{1 - |\mathcal{V}|^2 \mathcal{A}(z_1, z_2)} \right\} , \\ \Theta_{2,n}(z_1, z_2) &\xrightarrow{N, n \rightarrow \infty} \Theta_2(z_1, z_2) = c \tilde{t}'(z_1) \tilde{t}'(z_2) \int \frac{\lambda^2 F^{\mathbf{R}}(d\lambda)}{(1 + \lambda \tilde{t}(z_1))^2 (1 + \lambda \tilde{t}(z_2))^2} , \\ \mathcal{B}_{1,n}(z) &\xrightarrow{N, n \rightarrow \infty} \mathcal{B}_1(z) = -\frac{c z^3 \tilde{t}^3(z)}{(1 - \mathcal{A}(z, z))(1 - |\mathcal{V}|^2 \mathcal{A}(z, z))} \int \frac{\lambda^2 F^{\mathbf{R}}(d\lambda)}{(1 + \lambda \tilde{t}(z))^3} , \\ \mathcal{B}_{2,n}(z) &\xrightarrow{N, n \rightarrow \infty} \mathcal{B}_2(z) = -\frac{c z^3 \tilde{t}^3(z)}{1 - \mathcal{A}(z, z)} \int \frac{\lambda^2 F^{\mathbf{R}}(d\lambda)}{(1 + \lambda \tilde{t}(z))^3} . \end{aligned}$$

where \tilde{t}, \tilde{t}' are the limits of $\tilde{t}_n, \tilde{t}'_n$ under (2.14). This can be packaged into the following result:

Corollary 3.2. *Assume that (A-1) and (A-2) hold true. Assume moreover that R_n is diagonal and that the convergence assumption (2.14) holds true. Then $M_n(\cdot)$ converges weakly on D_ε to a two-dimensional Gaussian process $N(\cdot)$ satisfying:*

$$\mathbb{E}N(z) = \mathcal{B}(z) \quad \text{where} \quad \mathcal{B} = |\mathcal{V}|^2 \mathcal{B}_1 + \kappa \mathcal{B}_2 , \quad z \in D_\varepsilon$$

and \mathcal{B}_1 and \mathcal{B}_2 are defined above and covariance

$$\text{cov}(N(z_1), N(z_2)) = \Theta(z_1, z_2) \quad \text{where} \quad \Theta = \Theta_0 + |\mathcal{V}|^2 \Theta_1 + \kappa \Theta_2 , \quad z_1, z_2 \in D_\varepsilon \cup \overline{D_\varepsilon}$$

and Θ_0 defined in (2.15) and Θ_1, Θ_2 defined above.

3.5. Additional computations in the case where R_n is the identity. In this section, we assume that $R_n = I_N$.

The term proportional to $|\mathcal{V}|^2$. In this case, the quantity $\mathcal{A}(z_1, z_2)$ takes the simplified form

$$\mathcal{A}(z_1, z_2) = \frac{c \tilde{t}_1 \tilde{t}_2}{(1 + \tilde{t}_1)(1 + \tilde{t}_2)} .$$

where we denote $\tilde{t}_i = \tilde{t}(z_i)$, $i = 1, 2$. Straightforward computations yield:

$$\frac{\partial}{\partial z_i} \mathcal{A}(z_1, z_2) = \frac{\tilde{t}'_i}{(1 + \tilde{t}_i) \tilde{t}_i} \mathcal{A}(z_1, z_2) , \quad i = 1, 2 .$$

and

$$\Theta_1(z_1, z_2) = \frac{c \tilde{t}'_1 \tilde{t}'_2}{(1 + \tilde{t}_1)^2 (1 + \tilde{t}_2)^2 (1 - |\mathcal{V}|^2 \mathcal{A}(z_1, z_2))^2} = \frac{c \tilde{t}'_1 \tilde{t}'_2}{((1 + \tilde{t}_1)(1 + \tilde{t}_2) - |\mathcal{V}|^2 c \tilde{t}_1 \tilde{t}_2)^2}.$$

This formula is in accordance with [9, Formula (2.2)] (use [9, (3.4)] to equate both). If needed, one can then use the explicit expression of the Stieltjes transform of Marčenko-Pastur distribution (see below).

The cumulant term. In the particular case where $R_n = I_N$, Lytova and Pastur [37] (see also [40]) provided an explicit formula for the cumulant term of the covariance based on the Stieltjes transform of Marčenko-Pastur distribution. We recover this formula hereafter and prove that

$$-\frac{\kappa}{4\pi^2} \oint \oint \varphi(z_1) \varphi(z_2) \Theta_2(z_1, z_2) dz_1 dz_2 = \frac{\kappa}{4c\pi^2} \left(\int_{\lambda^-}^{\lambda^+} \varphi(\lambda) \frac{\lambda - (1+c)}{\sqrt{(\lambda^+ - \lambda)(\lambda - \lambda^-)}} d\lambda \right)^2, \quad (3.9)$$

where $\lambda^\pm = (1 \pm \sqrt{c})^2$ i.e. $[\lambda^-, \lambda^+]$ is the support of Marčenko-Pastur distribution. This is in accordance⁵ with Lytova and Pastur's result [37, Eq. (4.65)].

Specifying Θ_2 in the case $R_n = I_N$, we obtain:

$$\begin{aligned} & -\frac{\kappa}{4\pi^2} \oint \oint \varphi(z_1) \varphi(z_2) \Theta_2(z_1, z_2) dz_1 dz_2 \quad (3.10) \\ &= -\frac{\kappa c}{4\pi^2} \left(\oint \varphi(z) \frac{\tilde{t}'(z)}{(1 + \tilde{t}(z))^2} dz \right)^2 = -\frac{\kappa c}{4\pi^2} \left(\oint \varphi'(z) \frac{1}{1 + \tilde{t}(z)} dz \right)^2 = -\frac{\kappa c}{4\pi^2} \left(\oint \varphi'(z) z t(z) dz \right)^2. \end{aligned}$$

Now, in this case, $t(z)$ is the Stieltjes transform of Marčenko-Pastur distribution and has an explicit form (see for instance [40, Chapter 7]):

$$t(z) = \frac{1}{2cz} \left\{ \sqrt{(z - (1+c))^2 - 4c} - (z - (1-c)) \right\}$$

where the branch of the square root is fixed by its asymptotics: $z - (1+c) + o(1)$ as $z \rightarrow \infty$. We have:

$$(3.10) = -\frac{\kappa c}{4\pi^2} \left(\oint \varphi(z) (z t(z))' dz \right)^2 = -\frac{\kappa c}{4\pi^2} \left(\oint \varphi(z) \frac{z - (1+c)}{2c\sqrt{(z - (1+c))^2 - 4c}} dz \right)^2$$

It remains to deform the contour into the cuts $[\lambda^-, \lambda^+]$ and to use the relations

$$\sqrt{(z - (1+c))^2 - 4c} \Big|_{z=\lambda^\pm i0} = \pm i \sqrt{(\lambda^+ - \lambda)(\lambda - \lambda^-)}$$

for $\lambda \in [(1 - \sqrt{c})^2, (1 + \sqrt{c})^2]$ to obtain

$$\begin{aligned} (3.10) &= -\frac{\kappa c}{4\pi^2} \left(\oint \varphi(z) \frac{z - (1+c)}{2c\sqrt{(z - (1+c))^2 - 4c}} dz \right)^2 = -\frac{\kappa c}{4\pi^2} \left(\int_{\lambda^-}^{\lambda^+} \varphi(\lambda) \frac{\lambda - (1+c)}{i c \sqrt{(\lambda^+ - \lambda)(\lambda - \lambda^-)}} d\lambda \right)^2 \\ &= \frac{\kappa}{4c\pi^2} \left(\int_{\lambda^-}^{\lambda^+} \varphi(\lambda) \frac{\lambda - (1+c)}{\sqrt{(\lambda^+ - \lambda)(\lambda - \lambda^-)}} d\lambda \right)^2, \end{aligned}$$

⁵Denote by the superscript ^{LP} the quantities in [37] and use the correspondance $c^{\text{LP}} \leftrightarrow 1/c$, $a^{\text{LP}} \leftrightarrow c$ and $\kappa_4^{\text{LP}} \leftrightarrow (a^{\text{LP}})^4 \kappa = c^2 \kappa$ to check that the r.h.s. of (3.9) equates the formula provided in [37].

which yields (3.9).

4. STATEMENT OF THE CLT FOR NON-ANALYTIC FUNCTIONALS

In order to lift the CLT from the trace of the resolvent to a smooth function f , the key ingredient is Helffer-Sjöstrand's formula (1.7). We first reduce the problem under study to the case of functions with compact support in Section 4.1. We then describe the fluctuations of $L_n^1(f)$ for non-analytic functions f in Section 4.2 and study the bias $L_n^2(f)$ in Section 4.3.

Let

$$\begin{aligned} L_n(f) &\stackrel{(a)}{=} \operatorname{Tr} f(\Sigma_n \Sigma_n^*) - N \int f(\lambda) \mathcal{F}_n(d\lambda) \\ &= \left(\operatorname{Tr} f(\Sigma_n \Sigma_n^*) - \mathbb{E} \operatorname{Tr} f(\Sigma_n \Sigma_n^*) \right) + \left(\mathbb{E} \operatorname{Tr} f(\Sigma_n \Sigma_n^*) - N \int f(\lambda) \mathcal{F}_n(d\lambda) \right) \\ &\triangleq L_n^1(f) + L_n^2(f), \end{aligned}$$

where \mathcal{F}_n in (a) is defined in (2.16).

4.1. Useful properties. Recall that \mathcal{S}_n is the support of the probability measure \mathcal{F}_n . Due to Assumption (A-2), it is clear that

$$\mathcal{S}_n \subset \mathcal{S}_\infty \triangleq \left[0, \lambda_R^+ \left(1 + \sqrt{\ell^+} \right)^2 \right], \quad (4.1)$$

uniformly in n .

Proposition 4.1. *Assume that (A-1) and (A-2) hold true and let the random variables be truncated as in Section 3.2. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function of class C^p ($p \geq 1$) on a η -neighborhood \mathcal{S}_∞^η of \mathcal{S}_∞ . Denote by \check{f} a function with compact support which coincides with f on \mathcal{S}_∞^η , then*

$$L_n^1(f) - L_n^1(\check{f}) \xrightarrow[N, n \rightarrow \infty]{\mathcal{P}} 0 \quad \text{and} \quad L_n^2(f) - L_n^2(\check{f}) \xrightarrow[N, n \rightarrow \infty]{} 0.$$

Proof of Proposition 4.1 is straightforward. It is based on the fact that f and \check{f} coincide on \mathcal{S}_n (and thus $\int f d\mathcal{F}_n = \int \check{f} d\mathcal{F}_n$) and on the deviation probability

$$\mathbb{P} \left(\|\Sigma_n \Sigma_n^*\| > \lambda_R^+ \left(1 + \sqrt{\ell^+} \right)^2 + \eta \right) = o(n^{-\ell})$$

for every $\ell \geq 1$ (see for instance [6, Eq. (1.9a)]). Details are left to the reader.

The following proposition underlines how a sufficient regularity of function f compensates the singularity in $\operatorname{Im}(z)^{-1}$ near the real axis.

Proposition 4.2. *Let μ, ν be two probability measures on \mathbb{R} and g_μ and g_ν their associated Stieltjes transforms. Assume that*

$$|g_\mu(z) - g_\nu(z)| \leq \frac{|h(z)|}{\operatorname{Im}(z)^k}, \quad z \in \mathbb{C}^+,$$

where h is a continuous function over $\operatorname{cl}(\mathbb{C}^+)$, the closure of \mathbb{C}^+ .

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function of order C^{k+1} with bounded support; recall the definition of $\Phi(f)$ in (1.6) and denote by

$$\|f\|_{k+1,\infty} = \sup_{0 \leq \ell \leq k+1} \|f^{(\ell)}\|_{\infty} .$$

Then

$$\begin{aligned} \left| \int f d\mu - \int f d\nu \right| &\leq \frac{1}{\pi} \left| \int_{\mathbb{C}^+} \bar{\partial}\Phi(f)(z) \{g_{\mu}(z) - g_{\nu}(z)\} \lambda_2(dz) \right| , \\ &\leq K \|f\|_{k+1,\infty} \int_{\text{supp}(f) \times \text{supp}(\chi)} |h(z)| \lambda_2(dz) , \\ &\leq K' \|f\|_{k+1,\infty} . \end{aligned} \tag{4.2}$$

Proof. Write

$$\begin{aligned} \bar{\partial}\Phi(f)(x + iy) &= \partial_x \Phi(f)(x + iy) + i \partial_y \Phi(f)(x + iy) \\ &= \frac{(iy)^k f^{(k+1)}(x)}{k!} \chi(y) + i \sum_{\ell=0}^k \frac{(iy)^{\ell} f^{(\ell)}(x)}{\ell!} \chi'(y) . \end{aligned}$$

From this and the fact that χ is equal to 1 for y small enough, we deduce that

$$\bar{\partial}\Phi(f)(x + iy) = \frac{(iy)^k f^{(k+1)}(x)}{k!}$$

near the real axis and that $|\bar{\partial}\Phi(f)(x + iy)| \leq 1_{\text{supp}(f) \times \text{supp}(\chi)}(z) K \|f\|_{k+1,\infty} \text{Im}(z)^k$, hence (4.2). \square

4.2. Fluctuations for the linear spectral statistics.

Assumption A-3. Function $f : \mathbb{R} \rightarrow \mathbb{R}$ is measurable and C^3 in a η -neighborhood of S_{∞} defined in (4.1).

Theorem 4.3. Assume that (A-1) and (A-2) hold true, and let f_1, \dots, f_k satisfy (A-3). Consider the centered Gaussian random vector $Z_n^1(\mathbf{f}) \triangleq (Z_n^1(f_1), \dots, Z_n^1(f_k))$ with covariance

$$\begin{aligned} \text{cov}(Z_n^1(f), Z_n^1(g)) &= \\ &= \frac{1}{2\pi^2} \text{Re} \int_{D^2} \bar{\partial}\Phi(f)(z_1) \overline{\bar{\partial}\Phi(g)(z_2)} \Theta_n(z_1, \bar{z}_2) \lambda_2(dz_1) \lambda_2(dz_2) \\ &\quad + \frac{1}{2\pi^2} \text{Re} \int_{D^2} \bar{\partial}\Phi(f)(z_2) \bar{\partial}\Phi(g)(z_2) \Theta_n(z_1, z_2) \lambda_2(dz_1) \lambda_2(dz_2) , \end{aligned} \tag{4.3}$$

where $\Phi(f)$ and $\Phi(g)$ are defined as in (1.6). Then, the sequence of \mathbb{R}^k -valued random vectors $Z_n^1(\mathbf{f})$ is tight and the following convergence holds true:

$$d_{\mathcal{L}P}(L_n^1(\mathbf{f}), Z_n^1(\mathbf{f})) \xrightarrow{N, n \rightarrow \infty} 0 ,$$

or equivalently for every continuous bounded function $h : \mathbb{R}^k \rightarrow \mathbb{C}$,

$$\mathbb{E} h(L_n^1(\mathbf{f})) - \mathbb{E} h(Z_n^1(\mathbf{f})) \xrightarrow{N, n \rightarrow \infty} 0 .$$

Proof of Theorem 4.3 is postponed to Section 6.

4.3. First-order expansions for the bias in the case of non-analytic functionals.

Assumption A-4. Function $f : \mathbb{R} \rightarrow \mathbb{R}$ is measurable and C^{18} in a η -neighborhood of \mathcal{S}_∞ defined in (4.1).

Theorem 4.4. Assume (A-1) and (A-2) hold true and let function f satisfy (A-4). Denote by

$$Z_n^2(f) = \frac{1}{\pi} \operatorname{Re} \int_D \bar{\partial} \Phi(f)(z) \mathcal{B}_n(z) \lambda_2(dz),$$

where \mathcal{B}_n is defined in (2.19). Then

$$\mathbb{E} \operatorname{Tr} f(\Sigma_n \Sigma_n^*) - N \int f(\lambda) \mathcal{F}_n(d\lambda) - Z_n^2(f) \xrightarrow{N, n \rightarrow \infty} 0.$$

Proof of Theorem 4.4 is postponed to Section 7.

Remark 4.1 (Why eighteen?). A quick sketch of the proof of Theorem 4.4 provides some hints. By gaussian interpolation (whose cost is $f \in C^8$), we only need to prove:

$$\mathbb{E} \operatorname{Tr} f(\Sigma_n^{\mathbb{C}}(\Sigma_n^{\mathbb{C}})^*) - N \int f(\lambda) \mathcal{F}_n(d\lambda) \rightarrow 0,$$

where $\Sigma_n^{\mathbb{C}}$ is the counterpart of Σ_n with $\mathcal{N}_{\mathbb{C}}(0, 1)$ i.i.d. entries. The proof of the latter is based on Helffer-Sjöstrand's formula:

$$\mathbb{E} \operatorname{Tr} f(\Sigma_n^{\mathbb{C}}(\Sigma_n^{\mathbb{C}})^*) - N \int f(\lambda) \mathcal{F}_n(d\lambda) = \frac{1}{\pi} \operatorname{Re} \int_{\mathbb{C}^+} \bar{\partial} \Phi(f) \{ \operatorname{Tr} \mathbb{E} Q_n^{\mathbb{C}} - N t_n \} d\lambda_2$$

where $Q_n^{\mathbb{C}} = (\Sigma_n^{\mathbb{C}}(\Sigma_n^{\mathbb{C}})^* - zI_N)^{-1}$, and on the following estimate, stated in Proposition 7.1:

$$| \mathbb{E} \operatorname{Tr} (\Sigma_n^{\mathbb{C}}(\Sigma_n^{\mathbb{C}})^* - zI_N)^{-1} - N t_n(z) | \leq \frac{1}{n} P_{12}(|z|) P_{17}(|\operatorname{Im}(z)|^{-1}), \quad (4.4)$$

where P_k denotes a polynomial with degree k and positive coefficients. In view of Proposition 4.2, f needs to be of class C^{18} . If one can improve estimate (4.4) and decrease the powers of $|\operatorname{Im}(z)|^{-1}$, then one will automatically lower the regularity assumption over f . Notice that in the case of the Gaussian Unitary Ensemble, counterpart of (4.4) features $|\operatorname{Im}(z)|^{-7}$ on its r.h.s. (cf. [25, Lemma 6.1]) hence the needed regularity is $f \in C^8$ in this case.

5. PROOF OF THEOREM 3.1 (CLT FOR THE TRACE OF THE RESOLVENT)

Recall that $M_n(z) = \operatorname{tr} Q_n(z) - N t_n(z)$. It will be convenient to decompose $M_n(z)$ as:

$$M_n(z) = M_n^1(z) + M_n^2(z) \quad \text{where} \quad \begin{cases} M_n^1(z) &= \operatorname{tr} Q_n(z) - \operatorname{tr} \mathbb{E} Q_n(z) \\ M_n^2(z) &= N (\mathbb{E} f_n(z) - t_n(z)) \end{cases}. \quad (5.1)$$

Denote by ξ_j the $N \times 1$ vector

$$\xi_j = \Sigma_{\cdot j} = \frac{1}{\sqrt{n}} R^{1/2} X_{\cdot j}$$

and by \mathbb{E}_j the conditional expectation with respect to \mathcal{G}_j , the σ -field generated by ξ_1, \dots, ξ_j ; by convention, $\mathbb{E}_0 = \mathbb{E}$. We split Theorem 3.1 into intermediate results. Recall the definitions of D_ε, D^+ and D in (3.7). Let

$$\Gamma = D^+ \cup \overline{D^+} \quad \text{where} \quad \overline{D^+} = \{ \bar{z}, \quad z \in D^+ \}.$$

Proposition 5.1. *Assume that (A-1) and (A-2) hold true; let $z_1, z_2 \in \Gamma$, then:*

$$M_n^1(z_1) = \sum_{j=1}^n Z_j^n(z_1) + o_P(1) ,$$

where the Z_j^n 's are martingale increments with respect to the σ -field \mathcal{G}_i and

$$\sum_{j=1}^n \mathbb{E}_{j-1} Z_j^n(z_1) Z_j^n(z_2) - \Theta_n(z_1, z_2) \xrightarrow[N, n \rightarrow \infty]{\mathcal{P}} 0 , \quad (5.2)$$

$$\sum_{j=1}^n \mathbb{E}_{j-1} Z_j^n(z_1) \overline{Z_j^n(z_2)} - \Theta_n(z_1, \overline{z_2}) \xrightarrow[N, n \rightarrow \infty]{\mathcal{P}} 0 , \quad (5.3)$$

where Θ_n is defined in (2.9). Moreover,

$$M_n^2(z_1) - \mathcal{B}_n(z_1) \xrightarrow[N, n \rightarrow \infty]{} 0 ,$$

where \mathcal{B}_n is defined in (2.19).

Proposition 5.2. *There exists a sequence $(N_n(z), z \in \Gamma)$ of two-dimensional Gaussian processes with mean $\mathbb{E}N_n(z) = \mathcal{B}_n(z)$ and covariance*

$$\begin{aligned} \text{cov}(N_n(z_1), N_n(z_2)) &= \mathbb{E}(N_n(z_1) - \mathbb{E}N_n(z_1))(N_n(z_2) - \mathbb{E}N_n(z_2)) \\ &= \Theta_n(z_1, z_2) . \end{aligned}$$

Moreover, $(N_n(z), z \in D_\varepsilon)$ is tight.

5.1. Proof for Proposition 5.1. The fact that (M_n) is a tight sequence has already been established in [6] (regardless of the assumption $\kappa = 0$ and $|\mathcal{V}| = 0/1$). In order to proceed, we shall heavily rely on the proof of [6, Lemma 1.1] which is the crux of Bai and Silverstein's paper. In Section 5.1.1 we recall the main steps of Bai and Silverstein's computations of the variance/covariance. In Sections 5.1.2 and 5.1.3, we compute the extra terms in the limiting variance. In Section 5.1.4, we compute the limiting bias (some details are postponed to Appendix A.1). In Section 5.3, we finally conclude the proof of Theorem 3.1 and address various subtleties which appear due to the existence of a sequence of Gaussian limiting processes.

In the sequel, we shall drop subscript n in the sequel and write Q and R instead of Q_n and R_n . Denote by $Q_j(z)$ the resolvent of matrix $\Sigma\Sigma^* - \xi_j\xi_j^*$, i.e.

$$Q_j(z) = (-zI + \Sigma\Sigma^* - \xi_j\xi_j^*)^{-1} .$$

The following quantities will be needed:

$$\begin{aligned} \beta_j(z) &= \frac{1}{1 + \xi_j^* Q_j(z) \xi_j} , \\ \bar{\beta}_j(z) &= \frac{1}{1 + \frac{1}{n} \text{tr} R_n Q_j(z)} , \\ b_n(z) &= \frac{1}{1 + \frac{1}{n} \mathbb{E} \text{tr} R_n Q_1(z)} , \\ \varepsilon_j(z) &= \xi_j^* Q_j(z) \xi_j - \frac{1}{n} \text{tr} R_n Q_j(z) , \\ \delta_j(z) &= \xi_j^* Q_j^2(z) \xi_j - \frac{1}{N} \text{tr} R_n Q_j^2(z) = \frac{d}{dz} \varepsilon_j(z) . \end{aligned}$$

5.1.1. *Preliminary variance computations.* We briefly review in this section the main steps related to the computation of the variance/covariance as presented in [6]. These standard steps will finally lead to Eq. (5.7) which will be the starting point of the computations associated to the $|\mathcal{V}|^2$ - and κ -terms of the variance.

Let $z \in \Gamma$.

$$\begin{aligned} N(f_n(z) - \mathbb{E}f_n(z)) &= -\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1}) \beta_j(z) \xi_j^* Q_j^2(z) \xi_j \\ &= -\sum_{j=1}^n \mathbb{E}_j \left(\bar{\beta}_j(z) \delta_j(z) - \bar{\beta}_j^2(z) \varepsilon_j(z) \frac{1}{n} \text{tr} R Q_j^2(z) \right) + o_P(1). \end{aligned}$$

Denote by

$$Z_j^n(z) = -\mathbb{E}_j \left(\bar{\beta}_j(z) \delta_j(z) - \bar{\beta}_j^2(z) \varepsilon_j(z) \frac{1}{n} \text{tr} R Q_j^2(z) \right) = -\mathbb{E}_j \frac{d}{dz} (\bar{\beta}_j(z) \varepsilon_j(z)).$$

Hence,

$$M_n^1(z) = N(f_n(z) - \mathbb{E}f_n(z)) = \sum_{j=1}^n Z_j^n(z) + o_P(1)$$

The r.h.s. appears as a sum of martingale increments. Such a decomposition is important since it will enable us to rely on powerful CLTs for martingales (see [11, Theorem 35.12], and the variations below in Lemmas 5.6 and 5.7). These CLTs rely on the study of the terms:

$$\sum_{j=1}^n \mathbb{E}_{j-1} Z_j^n(z_1) Z_j^n(z_2) \quad \text{and} \quad \sum_{j=1}^n \mathbb{E}_{j-1} Z_j^n(z_1) \overline{Z_j^n(z_2)}.$$

Notice that since $\overline{Z_j^n(z)} = Z_j^n(\bar{z})$, we have $\mathbb{E}_{j-1} Z_j^n(z_1) \overline{Z_j^n(z_2)} = \mathbb{E}_{j-1} Z_j^n(z_1) Z_j^n(\bar{z}_2)$. Since the set Γ is stable by complex conjugation, it is sufficient to study the limiting behavior of:

$$\sum_{j=1}^n \mathbb{E}_{j-1} Z_j^n(z_1) Z_j^n(z_2), \quad z_1, z_2 \in \Gamma$$

in order to prove (5.2) and (5.3). Now,

$$\sum_{j=1}^n \mathbb{E}_{j-1} Z_j^n(z_1) Z_j^n(z_2) = \frac{\partial^2}{\partial z_1 \partial z_2} \left\{ \sum_{j=1}^n \mathbb{E}_{j-1} [\mathbb{E}_j (\bar{\beta}_j(z_1) \varepsilon_j(z_1)) \mathbb{E}_j (\bar{\beta}_j(z_2) \varepsilon_j(z_2))] \right\}. \quad (5.4)$$

Following the same arguments as in [6, pp. 571], one can prove that it is sufficient to study the convergence in probability of

$$\sum_{j=1}^n \mathbb{E}_{j-1} [\mathbb{E}_j (\bar{\beta}_j(z_1) \varepsilon_j(z_1)) \mathbb{E}_j (\bar{\beta}_j(z_2) \varepsilon_j(z_2))] .$$

Moreover,

$$\begin{aligned}
& \sum_{j=1}^n \mathbb{E}_{j-1} \left[\mathbb{E}_j \left(\bar{\beta}_j(z_1) \varepsilon_j(z_1) \right) \mathbb{E}_j \left(\bar{\beta}_j(z_2) \varepsilon_j(z_2) \right) \right] \\
&= \sum_{j=1}^n b_n(z_1) b_n(z_2) \mathbb{E}_{j-1} \left[\mathbb{E}_j \varepsilon_j(z_1) \mathbb{E}_j \varepsilon_j(z_2) \right] + o_P(1) , \\
&= \sum_{j=1}^n z_1 \tilde{t}_n(z_1) z_2 \tilde{t}_n(z_2) \mathbb{E}_{j-1} \left[\mathbb{E}_j \varepsilon_j(z_1) \mathbb{E}_j \varepsilon_j(z_2) \right] + o_P(1) . \tag{5.5}
\end{aligned}$$

Hence, it is finally sufficient to study the limiting behaviour (in terms of convergence in probability) of the quantity:

$$\sum_{j=1}^n \mathbb{E}_{j-1} \left(\mathbb{E}_j \varepsilon_j(z_1) \mathbb{E}_j \varepsilon_j(z_2) \right) , \quad z_1, z_2 \in \Gamma . \tag{5.6}$$

Denote by A^T the transpose matrix of A . Applying (2.6) yields :

$$\begin{aligned}
\sum_{j=1}^n \mathbb{E}_{j-1} \left(\mathbb{E}_j \varepsilon_j(z_1) \mathbb{E}_j \varepsilon_j(z_2) \right) &= \frac{1}{n^2} \sum_{j=1}^n \text{tr} \left(R^{1/2} \mathbb{E}_j Q_j(z_1) R \mathbb{E}_j Q_j(z_2) R^{1/2} \right) \\
&\quad + \frac{|\mathcal{V}|^2}{n^2} \sum_{j=1}^n \text{tr} \left(R^{1/2} \mathbb{E}_j Q_j(z_1) R^{1/2} \left(R^{1/2} \mathbb{E}_j Q_j(z_2) R^{1/2} \right)^T \right) \\
&\quad + \frac{\kappa}{n^2} \sum_{j=1}^n \sum_{i=1}^N \left(R^{1/2} \mathbb{E}_j Q_j(z_1) R^{1/2} \right)_{ii} \left(R^{1/2} \mathbb{E}_j Q_j(z_2) R^{1/2} \right)_{ii} . \tag{5.7}
\end{aligned}$$

The limiting behaviour of the first term of the r.h.s. has been completely described in [6] where it has been shown that:

$$\frac{\partial^2}{\partial z_1 \partial z_2} \left\{ z_1 z_2 \tilde{t}_n(z_1) \tilde{t}_n(z_2) \frac{1}{n^2} \sum_{j=1}^n \text{tr} \left(R^{1/2} \mathbb{E}_j Q_j(z_1) R \mathbb{E}_j Q_j(z_2) R^{1/2} \right) \right\} = \Theta_{0,n}(z_1, z_2) + o_P(1) , \tag{5.8}$$

with $\Theta_{0,n}(z_1, z_2)$ defined in (2.10).

We shall focus on the second and third terms.

5.1.2. *The term proportional to $|\mathcal{V}|^2$ in the variance.* Notice first that the value of t_n and \tilde{t}_n is the same wether R is replaced by \bar{R} in (2.3) and (3.1) since t_n and \tilde{t}_n only depend on the spectrum of R (which is the same as the spectrum of \bar{R}). Notice also that $(R^{1/2})^T = \bar{R}^{1/2}$, hence:

$$\left(R^{1/2} \mathbb{E}_j Q_j(z_2) R^{1/2} \right)^T = \bar{R}^{1/2} \mathbb{E}_j Q_j^T(z_2) \bar{R}^{1/2} .$$

Recall the definition of $T_n^T(z)$ given by (2.8). Taking into account the fact that for a deterministic matrix A ,

$$\mathbb{E} \xi_j^T A \xi_j = \frac{\mathcal{V}}{n} \text{tr} \left(\bar{R}^{1/2} A R^{1/2} \right) \quad \text{and} \quad \mathbb{E} \xi_j^* A \bar{\xi}_j = \frac{\bar{\mathcal{V}}}{n} \text{tr} \left(R^{1/2} A \bar{R}^{1/2} \right) , \tag{5.9}$$

and following closely [6, Section 2], it is a matter of bookkeeping⁶ to establish that:

$$\begin{aligned} & \frac{|\mathcal{V}|^2 z_1 z_2}{n^2} \tilde{t}_n(z_1) \tilde{t}_n(z_2) \sum_{j=1}^n \operatorname{tr} \left(R^{1/2} \mathbb{E}_j Q_j(z_1) R^{1/2} \left(R^{1/2} \mathbb{E}_j Q_j(z_2) R^{1/2} \right)^T \right) \quad (5.10) \\ &= |\mathcal{V}|^2 \mathcal{A}_n(z_1, z_2) \times \frac{1}{n} \sum_{j=1}^n \frac{1}{1 - \left(\frac{j-1}{n}\right) |\mathcal{V}|^2 \mathcal{A}_n(z_1, z_2)} + o_P(1) \\ &= \int_0^{|\mathcal{V}|^2 \mathcal{A}_n(z_1, z_2)} \frac{dz}{1-z} + o_P(1) \end{aligned}$$

where

$$\mathcal{A}_n(z_1, z_2) = \frac{z_1 z_2}{n} \tilde{t}_n(z_1) \tilde{t}_n(z_2) \operatorname{tr} \left\{ R^{1/2} T_n(z_1) R^{1/2} \bar{R}^{1/2} T_n^T(z_2) \bar{R}^{1/2} \right\} .$$

Finally,

$$\frac{\partial^2}{\partial z_1 \partial z_2} (5.10) = |\mathcal{V}|^2 \Theta_{1,n}(z_1, z_2) + o_P(1) = |\mathcal{V}|^2 \frac{\partial}{\partial z_2} \left\{ \frac{\partial \mathcal{A}_n(z_1, z_2) / \partial z_1}{1 - |\mathcal{V}|^2 \mathcal{A}_n(z_1, z_2)} \right\} + o_P(1) . \quad (5.11)$$

5.1.3. *The cumulant term in the variance.* We now handle the term proportional to κ in (5.7):

$$\frac{1}{n^2} \sum_{j=1}^n \sum_{i=1}^N \left(R^{1/2} \mathbb{E}_j Q_j(z_1) R^{1/2} \right)_{ii} \left(R^{1/2} \mathbb{E}_j Q_j(z_2) R^{1/2} \right)_{ii} . \quad (5.12)$$

The objective is to prove that $\mathbb{E}_j Q_j(z)$ can be replaced by $T_n(z)$ in the formula above, which boils down to prove a convergence of quadratic forms of the type (2.5). Such a convergence has already been established in [29] for large covariance matrices based on a non-centered matrix model with separable variance profile.

By interpolating between the quantity (5.12) and its counterpart when the entries are complex i.i.d. standard Gaussian, we will be able to rely on the results in [29] by using the unitary invariance of a Gaussian matrix (see Proposition 5.4 and Eq. (5.29) below).

Let δ_z be the distance between the point $z \in \mathbb{C}$ and the real nonnegative axis \mathbb{R}^+ :

$$\delta_z = \operatorname{dist}(z, \mathbb{R}^+) . \quad (5.13)$$

Proposition 5.3. *Assume that (A-1) and (A-2) hold true and let u_n be a deterministic $N \times 1$ vector, then:*

$$\mathbb{E} |u_n^* Q(z) u_n - u_n^* \mathbb{E} Q(z) u_n|^2 \leq \frac{1}{n} \Phi(|z|) \Psi \left(\frac{1}{\delta_z} \right) \|u_n\|^2 ,$$

where Φ and Ψ are fixed polynomials with coefficients independent from N, n, z and (u_n) .

Proof of Proposition 5.3 is an easy adaptation⁷ of [29, Prop. 2.7].

Denote by $X_n^{\mathbb{C}}$ a $N \times n$ matrix whose entries are independent standard complex circular Gaussian r.v. (i.e. $X_{ij}^{\mathbb{C}} = U + iV$ where U, V are independent $\mathcal{N}(0, 2^{-1})$ random variables);

⁶Similar computations for the term proportional to $|\mathcal{V}|^2$ in the bias are outlined in Appendix A.1.

⁷Notice in particular all the cancellations that appear when adapting the proof of [29, Prop. 2.7], due to the fact that Σ_n is centered here; notice also the fact that R not being diagonal has virtually no impact.

denote accordingly $\Sigma_n^{\mathbb{C}} = n^{-1/2} R^{1/2} X_n^{\mathbb{C}}$, $\xi_j^{\mathbb{C}} = (\Sigma_n^{\mathbb{C}}) \cdot j$ and

$$Q_n^{\mathbb{C}}(z) = \left(-zI_N + \Sigma_n^{\mathbb{C}} (\Sigma_n^{\mathbb{C}})^* \right)^{-1}.$$

We now drop subscripts N and n .

Proposition 5.4. *Assume that (A-1) and (A-2) hold true and let , then:*

$$|u_n^* \mathbb{E} Q(z) u_n - u_n^* \mathbb{E} Q^{\mathbb{C}}(z) u_n| \leq \frac{1}{\sqrt{n}} \Phi(|z|) \Psi \left(\frac{1}{\delta_z} \right) \|u_n\|^2, \quad (5.14)$$

where u_n is a deterministic $N \times 1$ vector and Φ and Ψ are fixed polynomials with coefficients independent from N, n, z . Moreover,

$$|\mathbb{E} \text{Tr} Q(z) - \mathbb{E} \text{Tr} Q^{\mathbb{C}}(z)| \leq K \frac{|z|^3}{\text{Im}(z)^7}, \quad (5.15)$$

where K is independent from N, n, z .

Notice that (5.14) is of direct use in this section while (5.15) will be used in Section 7.

Proof. We first prove (5.14). Consider the resolvent

$$Q^{(i)}(z) = \left(\sum_{\ell=1}^i \xi_{\ell}^{\mathbb{C}} \xi_{\ell}^{\mathbb{C}*} + \sum_{\ell=i+1}^n \xi_{\ell} \xi_{\ell}^* - zI_N \right)^{-1}$$

defined for $1 \leq i \leq n-1$. Denote by $Q^{(0)} = Q$ and by $Q^{(n)} = Q^{\mathbb{C}}$ and write

$$u^* \mathbb{E}(Q - Q^{\mathbb{C}}) u = \sum_{i=1}^n u^* \mathbb{E}(Q^{(i-1)} - Q^{(i)}) u. \quad (5.16)$$

We shall evaluate the difference $u^* \mathbb{E}(Q^{(0)} - Q^{(1)}) u$, the other ones being handled similarly. Denote by $\check{Q}(z) = (\sum_{i=2}^n \xi_i \xi_i^* - zI_N)^{-1}$, then:

$$Q^{(0)} = \check{Q} - \frac{\check{Q} \xi_1 \xi_1^* \check{Q}}{1 + \xi_1^* \check{Q} \xi_1} \quad \text{and} \quad Q^{(1)} = \check{Q} - \frac{\check{Q} \xi_1^{\mathbb{C}} \xi_1^{\mathbb{C}*} \check{Q}}{1 + \xi_1^{\mathbb{C}*} \check{Q} \xi_1^{\mathbb{C}}}. \quad (5.17)$$

Dropping the subscript 1 to lighten the notations, we get:

$$\begin{aligned} u^* \mathbb{E} \left(Q^{(0)} - Q^{(1)} \right) u &= u^* \mathbb{E} \left(\frac{\check{Q} \xi^{\mathbb{C}} \xi^{\mathbb{C}*} \check{Q}}{1 + \xi^{\mathbb{C}*} \check{Q} \xi^{\mathbb{C}}} - \frac{\check{Q} \xi \xi^* \check{Q}}{1 + \xi^* \check{Q} \xi} \right) u \\ &= u^* \mathbb{E} \left(\frac{\check{Q} \xi^{\mathbb{C}} \xi^{\mathbb{C}*} \check{Q}}{1 + \xi^{\mathbb{C}*} \check{Q} \xi^{\mathbb{C}}} - \frac{\check{Q} \xi^{\mathbb{C}} \xi^{\mathbb{C}*} \check{Q}}{1 + \frac{1}{n} \text{tr} R \check{Q}} \right) u \\ &\quad + u^* \mathbb{E} \left(\frac{\check{Q} \xi^{\mathbb{C}} \xi^{\mathbb{C}*} \check{Q}}{1 + \frac{1}{n} \text{tr} R \check{Q}} - \frac{\check{Q} \xi \xi^* \check{Q}}{1 + \frac{1}{n} \text{tr} R \check{Q}} \right) u \\ &\quad + u^* \mathbb{E} \left(\frac{\check{Q} \xi \xi^* \check{Q}}{1 + \frac{1}{n} \text{tr} R \check{Q}} - \frac{\check{Q} \xi \xi^* \check{Q}}{1 + \xi^* \check{Q} \xi} \right) u. \end{aligned}$$

The second term in the r.h.s. above is zero (simply compute the conditional expectation with respect to \check{Q}), the first and third term are of a similar nature; we therefore only estimate the third one denoted by Δ_3 below.

$$\begin{aligned}
|\Delta_3| &= \left| u^* \mathbb{E} \left(\frac{\check{Q} \xi \xi^* \check{Q}}{1 + \frac{1}{n} \text{tr} R \check{Q}} - \frac{\check{Q} \xi \xi^* \check{Q}}{1 + \xi^* \check{Q} \xi} \right) u \right| \\
&= \left| \mathbb{E} \frac{\xi^* \check{Q} u u^* \check{Q} \xi}{(1 + \xi^* \check{Q} \xi)(1 + \frac{1}{n} \text{tr} R \check{Q})} \left(\xi^* \check{Q} \xi - \frac{1}{n} \text{tr} R \check{Q} \right) \right| \\
&\leq \frac{|z|^2}{|\text{Im}(z)|^2} \left\{ \mathbb{E} \left| \xi^* \check{Q} \xi - \frac{1}{n} \text{tr} R \check{Q} \right|^2 \mathbb{E} |\xi^* \check{Q} u u^* \check{Q} \xi|^2 \right\}^{1/2}, \tag{5.18}
\end{aligned}$$

where the last inequality follows from Cauchy-Schwarz inequality plus the fact that both $(-z(1 + \xi^* \check{Q} \xi))^{-1}$ and $(-z(1 + n^{-1} \text{tr} R \check{Q}))^{-1}$ are Stieltjes transforms and hence upper-bounded in module by $|\text{Im}(z)|^{-1}$. A control for the first expectation in the above inequality directly follows from classical estimates (see for instance [7, Lemma B.26]):

$$\mathbb{E} \left| \xi^* \check{Q} \xi - \frac{1}{n} \text{tr} R \check{Q} \right|^2 \leq \frac{K}{n^2} \mathbb{E} |X_{11}|^4 \mathbb{E} (\text{tr} R \check{Q} R \check{Q}^*) \leq \frac{K}{n} \frac{\|R\|^2}{|\text{Im}(z)|^2} c_n \mathbb{E} |X_{11}|^4, \tag{5.19}$$

where K is a constant whose value may change from line to line but which remains independent from N, n . The second expectation can be handled in the following way:

$$\begin{aligned}
\mathbb{E} |\xi^* \check{Q} u u^* \check{Q} \xi|^2 &= \mathbb{E} \left| \xi^* \check{Q} u u^* \check{Q} \xi - \frac{1}{n} \text{tr} R \check{Q} u u^* \check{Q} + \frac{1}{n} \text{tr} R \check{Q} u u^* \check{Q} \right|^2 \\
&\leq 2 \mathbb{E} \left| \xi^* \check{Q} u u^* \check{Q} \xi - \frac{1}{n} \text{tr} R \check{Q} u u^* \check{Q} \right|^2 + 2 \mathbb{E} \left| \frac{1}{n} \text{tr} R \check{Q} u u^* \check{Q} \right|^2 \\
&\leq \frac{K}{n^2} \mathbb{E} |X_{11}|^4 \mathbb{E} \text{tr} (R^{1/2} \check{Q} u u^* \check{Q} R^{1/2}) (R^{1/2} \check{Q} u u^* \check{Q} R^{1/2})^* \\
&\quad + \frac{2}{n^2} \mathbb{E} |u^* \check{Q} R \check{Q} u|^2 \\
&\leq \frac{K}{n^2} \frac{\|R\|^2 \|u\|^4}{|\text{Im}(z)|^4}. \tag{5.20}
\end{aligned}$$

It now remains to gather (5.19) and (5.20) to get:

$$\left| u^* \mathbb{E} (Q^{(0)} - Q^{(1)}) u \right| \leq \frac{1}{n\sqrt{n}} \Phi(|z|) \Psi \left(\frac{1}{\delta_z} \right) \|u\|^2.$$

Finally, the result follows by upper-bounding each term of the sum in (5.16) and (5.14) is proved.

We now establish (5.15). Using a similar decomposition as in (5.16), we get:

$$\mathbb{E} \text{Tr} (Q - Q^C) = \sum_{i=1}^n \mathbb{E} \text{Tr} (Q^{(i-1)} - Q^{(i)}).$$

We focus on the first term, use (5.17) and follow a similar notational convention by dropping subscript 1.

$$\begin{aligned} \mathbb{E}\text{Tr} \left(Q^{(0)} - Q^{(1)} \right) &= \mathbb{E}\text{Tr} \left(\frac{\check{Q}\xi^{\mathbb{C}}\xi^{\mathbb{C}*}\check{Q}}{1 + \xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}}} - \frac{\check{Q}\xi\xi^*\check{Q}}{1 + \xi^*\check{Q}\xi} \right) = \mathbb{E} \left(\frac{\xi^{\mathbb{C}*}\check{Q}^2\xi^{\mathbb{C}}}{1 + \xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}}} - \frac{\xi^*\check{Q}^2\xi}{1 + \xi^*\check{Q}\xi} \right) \\ &= \mathbb{E} \left(\frac{\xi^{\mathbb{C}*}\check{Q}^2\xi^{\mathbb{C}}\xi^*\check{Q}\xi - \xi^*\check{Q}^2\xi\xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}}}{(1 + \xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}})(1 + \xi^*\check{Q}\xi)} \right) + \mathbb{E} \left(\frac{\xi^{\mathbb{C}*}\check{Q}^2\xi^{\mathbb{C}} - \xi^*\check{Q}^2\xi}{(1 + \xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}})(1 + \xi^*\check{Q}\xi)} \right). \end{aligned} \quad (5.21)$$

Denote by

$$\begin{aligned} A_1 &= (\xi^{\mathbb{C}*}\check{Q}^2\xi^{\mathbb{C}})(\xi^*\check{Q}\xi) - (\xi^*\check{Q}^2\xi)(\xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}}), \\ A_2 &= \xi^{\mathbb{C}*}\check{Q}^2\xi^{\mathbb{C}} - \xi^*\check{Q}^2\xi, \\ B_1 &= \frac{1}{(1 + \xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}})(1 + \xi^*\check{Q}\xi)} - \frac{1}{(1 + \xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}})(1 + \frac{1}{n}\text{tr} R\check{Q})}, \\ B_2 &= \frac{1}{(1 + \xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}})(1 + \frac{1}{n}\text{tr} R\check{Q})} - \frac{1}{(1 + \frac{1}{n}\text{tr} R\check{Q})^2}, \\ B_3 &= \frac{1}{(1 + \frac{1}{n}\text{tr} R\check{Q})^2}. \end{aligned}$$

With these notations at hand, we have:

$$B_1 + B_2 + B_3 = \frac{1}{(1 + \xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}})(1 + \xi^*\check{Q}\xi)}$$

and

$$\mathbb{E}\text{Tr} \left(Q^{(0)} - Q^{(1)} \right) = \mathbb{E}A_1(B_1 + B_2 + B_3) + \mathbb{E}A_2(B_1 + B_2 + B_3).$$

Notice that $\mathbb{E}A_1B_3 = \mathbb{E}A_2B_3 = 0$. By Cauchy-Schwarz inequality, Proof of Prop. 5.4 will be completed as long as we establish the following estimates:

$$\{\mathbb{E}|A_1|^2\}^{1/2} \leq \frac{K}{\sqrt{n}} \frac{1}{\text{Im}(z)^3}, \quad (5.22)$$

$$\{\mathbb{E}|A_2|^2\}^{1/2} \leq \frac{K}{\sqrt{n}} \frac{1}{\text{Im}(z)^2}, \quad (5.23)$$

$$\{\mathbb{E}|B_1|^2\}^{1/2} \leq \frac{K}{\sqrt{n}} \frac{|z|^3}{\text{Im}(z)^4}, \quad (5.24)$$

$$\{\mathbb{E}|B_2|^2\}^{1/2} \leq \frac{K}{\sqrt{n}} \frac{|z|^3}{\text{Im}(z)^4}. \quad (5.25)$$

Estimates (5.24)-(5.25) can be established as (5.18). In order to establish (5.22), we compute exactly the expectation $\mathbb{E}|A_1|^2$ writing

$$\begin{aligned} &\mathbb{E} \left| \xi^{\mathbb{C}*}\check{Q}^2\xi^{\mathbb{C}}\xi^*\check{Q}\xi - \xi^*\check{Q}^2\xi\xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}} \right|^2 \\ &= \mathbb{E} \left(\xi^{\mathbb{C}*}\check{Q}^2\xi^{\mathbb{C}}\xi^*\check{Q}\xi - \xi^*\check{Q}^2\xi\xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}} \right) \overline{\left(\xi^{\mathbb{C}*}\check{Q}^2\xi^{\mathbb{C}}\xi^*\check{Q}\xi - \xi^*\check{Q}^2\xi\xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}} \right)}, \end{aligned} \quad (5.26)$$

which splits into 4 terms:

$$\begin{aligned} (5.26) &= \mathbb{E} \left\{ \left| \xi^{\mathbb{C}*}\check{Q}^2\xi^{\mathbb{C}} \right|^2 \left| \xi^*\check{Q}\xi \right|^2 \right\} - \mathbb{E} \left\{ (\xi^*\check{Q}^2\xi) \overline{(\xi^*\check{Q}\xi)} (\xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}}) \overline{(\xi^{\mathbb{C}*}\check{Q}^2\xi^{\mathbb{C}})} \right\} \\ &\quad + \mathbb{E} \left\{ \left| \xi^*\check{Q}^2\xi \right|^2 \left| \xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}} \right|^2 \right\} - \mathbb{E} \left\{ \overline{(\xi^*\check{Q}^2\xi)} (\xi^*\check{Q}\xi) \overline{(\xi^{\mathbb{C}*}\check{Q}\xi^{\mathbb{C}})} (\xi^{\mathbb{C}*}\check{Q}^2\xi^{\mathbb{C}}) \right\}. \end{aligned}$$

Using the independence of ξ , $\xi^{\mathbb{C}}$ and \check{Q} together with formula (2.6), lengthy but straightforward computations yield the estimate

$$\mathbb{E} \left| \xi^{\mathbb{C}*} \check{Q}^2 \xi^{\mathbb{C}} \xi^* \check{Q} \xi - \xi^* \check{Q}^2 \xi \xi^{\mathbb{C}*} \check{Q} \xi^{\mathbb{C}} \right|^2 \leq \frac{K}{n \operatorname{Im}(z)^6}.$$

Similar computations yield

$$\mathbb{E} \left| \xi^{\mathbb{C}*} \check{Q}^2 \xi^{\mathbb{C}} - \xi^* \check{Q}^2 \xi \right|^2 \leq \frac{K}{n \operatorname{Im}(z)^4}, \quad (5.27)$$

and (5.22)-(5.23) are established. Estimate (5.15) is established and proof of Prop. 5.4 is completed. \square

Corollary 5.5. *Assume that (A-1) and (A-2) hold true, then the following convergence holds true:*

$$\begin{aligned} & \frac{1}{n^2} \sum_{j=1}^n \sum_{i=1}^N \left(R^{1/2} \mathbb{E}_j Q_j(z_1) R^{1/2} \right)_{ii} \left(R^{1/2} \mathbb{E}_j Q_j(z_2) R^{1/2} \right)_{ii} \\ & - \frac{1}{n} \sum_{i=1}^N \left(R^{1/2} T(z_1) R^{1/2} \right)_{ii} \left(R^{1/2} T(z_2) R^{1/2} \right)_{ii} \xrightarrow[n, N \rightarrow \infty]{\mathcal{P}} 0. \end{aligned}$$

Proof. We first transform the sum to be calculated:

$$\frac{1}{n^2} \sum_{j=1}^n \sum_{i=1}^N \left(R^{1/2} \mathbb{E}_j Q_j(z_1) R^{1/2} \right)_{ii} \left(R^{1/2} \mathbb{E}_j Q_j(z_2) R^{1/2} \right)_{ii}. \quad (5.28)$$

Using Proposition 5.3 enables us to replace the conditional expectation \mathbb{E}_i by the true expectation in every term $(R^{1/2} \mathbb{E}_j Q_j(z) R^{1/2})_{ii}$. Now using the fact that

$$Q = Q_j - \frac{Q_j \xi_j \xi_j^* Q_j}{1 + \xi_j^* Q_j \xi_j}$$

and computations similar to those made in Proposition 5.4, one can replace $\mathbb{E} Q_j$ by $\mathbb{E} Q$. Finally, by Proposition 5.4, $\mathbb{E} Q$ can be replaced by $\mathbb{E} Q^{\mathbb{C}}$. We are led to study the sum:

$$\frac{1}{n} \sum_{i=1}^N \left(R^{1/2} \mathbb{E} Q^{\mathbb{C}}(z_1) R^{1/2} \right)_{ii} \left(R^{1/2} \mathbb{E} Q^{\mathbb{C}}(z_2) R^{1/2} \right)_{ii}.$$

Denote by $R_n = U_n \Delta U_n^*$ the spectral decomposition of covariance matrix R_n . Since matrix U_n is unitary, then $Y_n = U_n^* X_n^{\mathbb{C}}$ has i.i.d. standard complex Gaussian entries and the resolvent writes:

$$\begin{aligned} Q^{\mathbb{C}}(z) &= \left(R_n^{1/2} X_n^{\mathbb{C}} (X_n^{\mathbb{C}})^* R_n^{1/2} - z I_N \right)^{-1} = U_n \left(\Delta^{1/2} Y_n Y_n^* \Delta^{1/2} - z I_N \right)^{-1} U_n^* \\ &\triangleq U_n Q_{\Delta}(z) U_n^*. \end{aligned} \quad (5.29)$$

Denote by $T_{\Delta}(z)$ the matrix

$$T_{\Delta}(z) = (-z I_N + (1 - c_n) \Delta - z c_n t_n(z) \Delta)^{-1},$$

where $t_n(z)$ is defined in (2.3); notice that the definition of $t_n(z)$ only depends on the spectrum of R_n (or equivalently Δ); notice also that

$$T_n(z) = U_n T_{\Delta}(z) U_n^*. \quad (5.30)$$

It has been proved in [29, Theorem 1.1] that for every deterministic $N \times 1$ vector v_n :

$$\mathbb{E} |v_n^* (Q_\Delta(z) - T_\Delta(z)) v_n|^2 \leq \frac{1}{n} \Phi_2(|z|) \Psi_2 \left(\frac{1}{\delta_z} \right) \|v_n\|^4 .$$

Hence,

$$\begin{aligned} |v_n^* \mathbb{E} Q_\Delta(z) v_n - v_n^* T_\Delta(z) v_n| &\leq \left(\mathbb{E} |v_n^* (Q_\Delta(z) - T_\Delta(z)) v_n|^2 \right)^{1/2} \\ &\leq \frac{\|v_n\|^2}{\sqrt{n}} \sqrt{\Phi_2(|z|) \Psi_2(\delta_z^{-1})} \leq \frac{\|v_n\|^2}{\sqrt{n}} \left(\frac{1 + \Phi_2(|z|)}{2} \right) \left(\frac{1 + \Psi_2(\delta_z^{-1})}{2} \right) . \end{aligned}$$

In particular, let e_i be the i^{th} coordinate vector, then

$$\begin{aligned} &\left| \left(R^{1/2} \mathbb{E} Q(z) R^{1/2} \right)_{ii} - \left(R^{1/2} T(z) R^{1/2} \right)_{ii} \right| \\ &= \left| \left(R^{1/2} U \mathbb{E} Q_\Delta(z) U^* R^{1/2} \right)_{ii} - \left(R^{1/2} U T_\Delta(z) U^* R^{1/2} \right)_{ii} \right| \\ &\leq \frac{\|R^{1/2} U^* e_i\|^2}{\sqrt{n}} \left(\frac{1 + \Phi_2(|z|)}{2} \right) \left(\frac{1 + \Psi_2(\delta_z^{-1})}{2} \right) , \end{aligned}$$

which completes the proof. \square

Combining the result in Corollary 5.5 together with (5.5) and (5.7), we have proved so far that:

$$\begin{aligned} &\frac{\partial^2}{\partial z_1 \partial z_2} \left\{ \frac{z_1 z_2 \tilde{t}_n(z_1) \tilde{t}_n(z_2)}{n^2} \sum_{j=1}^n \sum_{i=1}^N \left(R^{1/2} \mathbb{E}_j Q_j(z_1) R^{1/2} \right)_{ii} \left(R^{1/2} \mathbb{E}_j Q_j(z_2) R^{1/2} \right)_{ii} \right\} \\ &= \frac{1}{n} \sum_{i=1}^N \frac{\partial^2}{\partial z_1 \partial z_2} \left\{ z_1 z_2 \tilde{t}_n(z_1) \tilde{t}_n(z_2) \left(R_n^{1/2} T_n(z_1) R_n^{1/2} \right)_{ii} \left(R_n^{1/2} T_n(z_2) R_n^{1/2} \right)_{ii} \right\} + o_P(1) . \end{aligned}$$

Taking into account (3.1) and the matrix identity $U(I + VU)^{-1}V = 1 - (I + UV)^{-1}$, we obtain:

$$\begin{aligned} &\frac{\partial^2}{\partial z_1 \partial z_2} \left\{ \frac{z_1 z_2 \tilde{t}_n(z_1) \tilde{t}_n(z_2)}{n^2} \sum_{j=1}^n \sum_{i=1}^N \left(R^{1/2} \mathbb{E}_j Q_j(z_1) R^{1/2} \right)_{ii} \left(R^{1/2} \mathbb{E}_j Q_j(z_2) R^{1/2} \right)_{ii} \right\} \\ &= \frac{1}{n} \sum_{i=1}^N \frac{\partial^2}{\partial z_1 \partial z_2} \left(I_N - (I_N + \tilde{t}_n(z_1) R_n)^{-1} \right)_{ii} \left(I_N - (I_N + \tilde{t}_n(z_2) R_n)^{-1} \right)_{ii} + o_P(1) , \\ &= \frac{z_1^2 z_2^2 \tilde{t}'_n(z_1) \tilde{t}'_n(z_2)}{n} \sum_{i=1}^N \left(R_n^{1/2} T_n^2(z_1) R_n^{1/2} \right)_{ii} \left(R_n^{1/2} T_n^2(z_2) R_n^{1/2} \right)_{ii} + o_P(1) , \\ &= \Theta_{2,n}(z_1, z_2) + o_P(1) , \end{aligned} \tag{5.31}$$

where $\Theta_{2,n}$ is given by formula (2.12).

Now gathering (5.8), (5.11) and (5.31), we have established so far:

$$\sum_{j=1}^n \mathbb{E}_{j-1} Z_j^n(z_1) Z_j^n(z_2) = \Theta_n(z_1, z_2) + o_P(1)$$

which is the first part of Proposition 5.1.

5.1.4. *Computations for the bias.* In this section, we are interested in the computation of $N(\mathbb{E}f_n(z) - t_n(z))$. As

$$\tilde{f}_n(z) = -\frac{(1-c_n)}{z} + c_n f_n(z) \quad \text{and} \quad \tilde{t}_n(z) = -\frac{(1-c_n)}{z} + c_n t_n(z),$$

we immediately obtain $N(\mathbb{E}f_n(z) - t_n(z)) = n(\mathbb{E}\tilde{f}_n(z) - \tilde{t}_n(z))$. Combining (2.7) and (3.1) yields:

$$-z - \frac{1}{\tilde{t}_n(z)} + \frac{1}{n} \text{tr} R_n (I_N + \tilde{t}_n(z) R_n)^{-1} = 0. \quad (5.32)$$

Following Bai and Silverstein [6, Section 4], we introduce the quantity $A_n(z)$ defined as:

$$\begin{aligned} A_n(z) &= z\mathbb{E}\tilde{f}_n(z) + 1 + \frac{1}{n} \text{tr} \left(I_N + \mathbb{E}\tilde{f}_n(z) R_n \right)^{-1} - c_n \\ &= z\mathbb{E}\tilde{f}_n(z) + 1 + \frac{1}{n} \text{tr} \left(I_N + \mathbb{E}\tilde{f}_n(z) R_n \right)^{-1} - \frac{1}{n} \text{tr} I_N^{-1} \\ &= -\mathbb{E}\tilde{f}_n(z) \left(-z - \frac{1}{\mathbb{E}\tilde{f}_n(z)} + \frac{1}{n} \text{tr} R_n (I_N + \mathbb{E}\tilde{f}_n(z) R_n)^{-1} \right), \end{aligned}$$

hence

$$-\frac{A_n(z)}{\mathbb{E}\tilde{f}_n(z)} = -z - \frac{1}{\mathbb{E}\tilde{f}_n(z)} + \frac{1}{n} \text{tr} R_n (I_N + \mathbb{E}\tilde{f}_n(z) R_n)^{-1}. \quad (5.33)$$

Subtracting (5.32) to (5.33) finally yields:

$$\mathbb{E}\tilde{f}_n(z) - \tilde{t}_n(z) = -A_n(z) \tilde{t}_n(z) \left[1 - \frac{\tilde{t}_n(z) \mathbb{E}\tilde{f}_n(z)}{n} \text{tr} R_n^2 \left(I_N + \mathbb{E}\tilde{f}_n(z) R_n \right)^{-1} \left(I_N + \tilde{t}_n(z) R_n \right)^{-1} \right]^{-1},$$

which is the counterpart of [6, Eq. (4.12)]. The same arguments as in [6] now yields:

$$n \left(\mathbb{E}\tilde{f}_n(z) - \tilde{t}_n(z) \right) = -n A_n(z) \tilde{t}_n(z) \left[1 - \frac{\tilde{t}_n^2(z)}{n} \text{tr} R_n^2 \left(I_N + \tilde{t}_n(z) R_n \right)^{-2} \right]^{-1} + o(1). \quad (5.34)$$

It remains to study the behaviour of $nA_n(z)$. Following [6, Eq. (4.10)], we obtain:

$$\begin{aligned} nA_n(z) &= \\ &= \frac{b_n^2}{n} \mathbb{E} \text{tr} Q_1 \left(\mathbb{E}\tilde{f}_n R_n + I_N \right)^{-1} R_n Q_1 R_n - b_n^2 n \mathbb{E} \left[\left(\xi_1^* Q_1 \xi_1 - \frac{1}{n} \text{tr} Q_1 R_n \right) \right. \\ &\quad \left. \times \left(\xi_1^* Q_1 \left(\mathbb{E}\tilde{f}_n R_n + I_N \right)^{-1} \xi_1 - \frac{1}{n} \text{tr} Q_1 \left(\mathbb{E}\tilde{f}_n R_n + I_N \right)^{-1} R_n \right) \right] + o(1). \end{aligned}$$

Applying (2.6) to the right term to the r.h.s. of the previous equation (recall that $R^T = \bar{R}$), we obtain:

$$\begin{aligned} nA_n(z) &= -|\mathcal{V}|^2 \frac{b_n^2}{n} \mathbb{E} \text{tr} R_n^{1/2} Q_1 \left(\mathbb{E}\tilde{f}_n R_n + I_N \right)^{-1} R_n^{1/2} \bar{R}_n^{1/2} Q_1^T \bar{R}_n^{1/2} \\ &\quad - \kappa \frac{b_n^2}{n} \sum_{i=1}^N \left(R_n^{1/2} Q_1 R_n^{1/2} \right)_{ii} \left(R_n^{1/2} Q_1 \left(\mathbb{E}\tilde{f}_n R_n + I_N \right)^{-1} R_n^{1/2} \right)_{ii} + o(1). \quad (5.35) \end{aligned}$$

The first term of the r.h.s. has been fully analyzed in [6] in the case where R_n and X_n are real matrices. We adapt these computations to the general case and outline in Appendix

A.1 the proof of the identity:

$$\begin{aligned} & -|\mathcal{V}|^2 \frac{b_n^2}{n} \mathbb{E} \operatorname{tr} R_n^{1/2} Q_1 \left(\mathbb{E} \tilde{f}_n R_n + I_N \right)^{-1} R_n^{1/2} \bar{R}_n^{1/2} Q_1^T \bar{R}_n^{1/2} \\ & = |\mathcal{V}|^2 \frac{\frac{z^3 \tilde{t}_n^2}{n} \operatorname{tr} R_n^{1/2} T_n^2(z) R_n^{1/2} \bar{R}_n^{1/2} T_n^T(z) \bar{R}_n^{1/2}}{1 - \frac{|\mathcal{V}|^2 z^2 \tilde{t}_n^2}{n} \operatorname{tr} R_n^{1/2} T_n(z) R_n^{1/2} \bar{R}_n^{1/2} T_n^T(z) \bar{R}_n^{1/2}} + o(1), \end{aligned} \quad (5.36)$$

where $T_n^T(z)$ is defined in (2.8). The term proportional to the cumulant in (5.35) can be analyzed as in Section 5.1.3, and one can prove that:

$$\begin{aligned} & -\kappa \frac{b_n^2}{n} \sum_{i=1}^N \left(R_n^{1/2} Q_1 R_n^{1/2} \right)_{ii} \left(R_n^{1/2} Q_1 \left(\mathbb{E} \tilde{f}_n R_n + I_N \right)^{-1} R_n^{1/2} \right)_{ii} \\ & = -\kappa \frac{z^2 \tilde{t}_n^2}{n} \sum_{i=1}^N \left(R_n^{1/2} T_n R_n^{1/2} \right)_{ii} \left(R_n^{1/2} T_n \left(\tilde{t}_n R_n + I_N \right)^{-1} R_n^{1/2} \right)_{ii} + o(1). \end{aligned} \quad (5.37)$$

We now plug (5.36) and (5.37) into (5.34) to conclude.

$$\begin{aligned} n \left(\mathbb{E} \tilde{f}_n(z) - \tilde{t}_n(z) \right) & = -|\mathcal{V}|^2 \frac{z^3 \tilde{t}_n^3}{n} \frac{\operatorname{tr} R_n^{1/2} T_n^2(z) R_n^{1/2} \bar{R}_n^{1/2} T_n^T(z) \bar{R}_n^{1/2}}{\left(1 - \frac{|\mathcal{V}|^2 z^2 \tilde{t}_n^2}{n} \operatorname{tr} R_n^{1/2} T_n(z) R_n^{1/2} \bar{R}_n^{1/2} T_n^T(z) \bar{R}_n^{1/2} \right) \left(1 - \frac{\tilde{t}_n^2}{n} \operatorname{tr} R_n^2 T_n^2 \right)} \\ & \quad - \kappa \frac{z^3 \tilde{t}_n^3}{n} \sum_{i=1}^N \frac{\left(R_n^{1/2} T_n R_n^{1/2} \right)_{ii} \left(R_n^{1/2} T_n^2 R_n^{1/2} \right)_{ii}}{1 - \frac{z^2 \tilde{t}_n^2}{n} \operatorname{tr} R_n^2 T_n^2} + o(1). \end{aligned}$$

Proof of Proposition 5.1 is completed.

5.2. Properties of the Gaussian process N_n : A meta-model argument. We prove hereafter Proposition 5.2. Recall the meta-model introduced in Section 2.6.

5.2.1. *The Gaussian process N_n .* Let

$$M_{n,M}(z) = \operatorname{tr} (\Sigma_n(M) \Sigma_n(M)^* - z I_{NM})^{-1} - M N t_n(z).$$

Applying Proposition 5.1 to the matrix model $\Sigma_n(M) \Sigma_n(M)^*$ yields:

$$\forall z \in \Gamma, \quad M_{n,M}^1(z) = \sum_{j=1}^{nM} Z_j^M(z) + o_P(1),$$

where the Z_j^M 's are martingale increments and

$$\begin{aligned} \sum_{j=1}^{nM} \mathbb{E}_{j-1} Z_j^M(z_1) Z_j^M(z_2) & \xrightarrow[N, n \text{ fixed}, M \rightarrow \infty]{\mathcal{P}} \Theta_n(z_1, z_2), \\ M_{n,M}^2(z) & \xrightarrow[N, n \text{ fixed}, M \rightarrow \infty]{} \mathcal{B}_n(z). \end{aligned}$$

Notice that there is a genuine limit in the previous convergence. Applying the central limit theorem for martingales [11, Theorem 35.12] plus the tightness argument for $(M_{n,M}(z), z \in \Gamma)$ provided by Proposition 5.1 immediately yields the fact that $M_{n,M}$ converges in distribution to a Gaussian process $(N_n(z), z \in \Gamma)$ with mean $\mathcal{B}_n(z)$ and covariance function $\Theta_n(z_1, z_2)$.

5.2.2. *Tightness of the sequence of Gaussian processes (N_n) .* In order to prove that the sequence of Gaussian processes (N_n) is tight, we shall prove, according to Prohorov's theorem, that it is relatively compact in distribution. Consider the set of matrices:

$$\{(R_n(M), M \geq 1); R_n \text{ is a } N \times n \text{ matrix, } N = N(n); n \geq 1\}$$

where $R_n(M)$ is defined in (2.23). Since $\|R_n(M)\| = \|R_n\|$ for all $M \geq 1$, we have

$$\sup_{M \geq 1, N, n \rightarrow \infty} \|R_n(M)\| = \sup_{N, n \rightarrow \infty} \|R_n\| < \infty$$

by Assumption (A-2). Hence, by Proposition 5.1, the family $\{M_{n,M}; M \geq 1\}_{N, n \rightarrow \infty}$ is tight, hence relatively compact in distribution. As the distribution $\mathcal{L}(N_n)$ of the Gaussian process N_n is the limit (in M) of the distribution $\mathcal{L}(M_{n,M})$ of $M_{n,M}$, $\mathcal{L}(N_n)$ belongs to the closure of $\{\mathcal{L}(M_{n,M})\}$, which is compact. Finally, $\{\mathcal{L}(N_n)\}$ is included in a compact set, hence is relatively compact. In particular, the family of Gaussian processes (N_n) is tight.

5.3. **Proof of Theorem 3.1.** The two propositions below are minor variations of known results. They will be helpful to conclude the proof of Theorem 3.1.

Lemma 5.6 (CLT for martingales I). *Suppose that for each n $Y_{n1}, Y_{n2}, \dots, Y_{nr_n}$ is a real martingale difference sequence with respect to the increasing σ -field $\{\mathcal{G}_{n,j}\}$ having second moments. Assume moreover that (Θ_n^2) is a sequence of nonnegative real numbers, uniformly bounded. If*

$$\sum_{j=1}^{r_n} \mathbb{E}(Y_{nj}^2 | \mathcal{G}_{n,j-1}) - \Theta_n^2 \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 0,$$

and for each $\varepsilon > 0$,

$$\sum_{j=1}^{r_n} \mathbb{E}(Y_{nj}^2 1_{|Y_{nj}| > \varepsilon}) \xrightarrow[n \rightarrow \infty]{} 0,$$

then, for every bounded continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$

$$\mathbb{E}f\left(\sum_{j=1}^{r_n} Y_{nj}\right) - \mathbb{E}f(Z_n) \xrightarrow[n \rightarrow \infty]{} 0, \quad (5.38)$$

where Z_n is a centered Gaussian random variable with variance Θ_n^2 .

Hereafter is the multidimensional and complex extension of Lemma 5.6 we shall rely on in the sequel:

Lemma 5.7 (CLT for martingales II). *Suppose that for each n $(Y_{nj}; 1 \leq j \leq r_n)$ is a \mathbb{C}^d -valued martingale difference sequence with respect to the increasing σ -field $\{\mathcal{G}_{n,j}; 1 \leq j \leq r_n\}$ having second moments. Write:*

$$Y_{nj}^T = (Y_{nj}^1, \dots, Y_{nj}^d).$$

Assume moreover that $(\Theta_n(k, \ell))_n$ and $(\tilde{\Theta}_n(k, \ell))_n$ are uniformly bounded sequences of complex numbers, for $1 \leq k, \ell \leq d$. If

$$\sum_{j=1}^{r_n} \mathbb{E}(Y_{nj}^k \bar{Y}_{nj}^d | \mathcal{G}_{n,j-1}) - \Theta_n(k, \ell) \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 0, \quad (5.39)$$

$$\sum_{j=1}^{r_n} \mathbb{E}(Y_{nj}^k Y_{nj}^\ell | \mathcal{G}_{n,j-1}) - \tilde{\Theta}_n(k, \ell) \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 0, \quad (5.40)$$

and for each $\varepsilon > 0$,

$$\sum_{j=1}^{r_n} \mathbb{E} (|Y_{nj}|^2 \mathbf{1}_{|Y_{nj}| > \varepsilon}) \xrightarrow{n \rightarrow \infty} 0, \quad (5.41)$$

then, for every bounded continuous function $f : \mathbb{C}^d \rightarrow \mathbb{R}$

$$\mathbb{E} f \left(\sum_{j=1}^{r_n} Y_{nj} \right) - \mathbb{E} f(Z_n) \xrightarrow{n \rightarrow \infty} 0, \quad (5.42)$$

where Z_n is a \mathbb{C}^d -valued centered Gaussian random vector with parameters

$$\mathbb{E} Z_n Z_n^* = (\Theta_n(k, \ell))_{k, \ell} \quad \text{and} \quad \mathbb{E} Z_n Z_n^T = (\tilde{\Theta}_n(k, \ell))_{k, \ell}.$$

Lemmas 5.6 and 5.7 are variations around the Central Limit Theorem for martingales (see Billingsley [11, Theorem 35.12]) which enables us to prove (in the real case):

$$\forall t \in \mathbb{R}, \quad \mathbb{E} e^{it \sum_{j=1}^{r_n} Y_{nj}} - e^{-\frac{t^2 \Theta_n^2}{2}} \rightarrow 0$$

and Lévy theorem for the weak convergence criterion via characteristic functions (see Kallenberg [34, Theorem 5.3 and Theorem 5.5]) which yields (5.42) from the above convergence. Details of the proof are omitted.

Lemma 5.8 (Tightness and weak convergence). *Let K be a compact set in \mathbb{C} ; let X_1, X_2, \dots and Y_1, Y_2, \dots be random elements in $C(K, \mathbb{C})$. Assume that for all $d \geq 1$, for all $z_1, \dots, z_d \in K$, for all $f \in C(\mathbb{C}^d, \mathbb{C})$ we have:*

$$\mathbb{E} f(X_n(z_1), \dots, X_n(z_d)) - \mathbb{E} f(Y_n(z_1), \dots, Y_n(z_d)) \xrightarrow{n \rightarrow \infty} 0.$$

Assume moreover that (X_n) and (Y_n) are tight, then for every continuous and bounded functional $F : C(K, \mathbb{C}) \rightarrow \mathbb{C}$, we have:

$$\mathbb{E} F(X_n) - \mathbb{E} F(Y_n) \xrightarrow{n \rightarrow \infty} 0.$$

Lemma 5.8 can be proved as [34, Lemma 16.2]; the proof is therefore omitted.

We are now in position to conclude.

In order to apply Lemma 5.7, it remains to check that Θ_n as defined in (2.9) is uniformly bounded for $z_1, z_2 \in \Gamma$ fixed but this is an easy byproduct of Proposition 5.2.

Proposition 5.1 together with Lemma 5.7 (notice that condition (5.41) can be proved as in [6]) yield the fact that for every $z_1, \dots, z_d \in \Gamma$ and for every bounded continuous function $f : \Gamma^d \rightarrow \mathbb{C}$:

$$\mathbb{E} f(M_n(z_1), \dots, M_n(z_d)) - \mathbb{E} f(N_n(z_1), \dots, N_n(z_d)) \xrightarrow{N, n \rightarrow \infty} 0,$$

where N_n is well-defined by Proposition 5.2. Now the tightness of M_n and N_n together with Lemma 5.8 yield the last statement of Theorem 3.1.

5.4. Proof of Theorem 3.1. Theorem 3.1 is now almost completely proved. It remains to prove that $L_n(\mathbf{f})$ and $Z_n(\mathbf{f})$ being tight, the following equivalence holds true:

$$d_{\mathcal{L}P}(L_n(\mathbf{f}), Z_n(\mathbf{f})) \xrightarrow{N, n \rightarrow \infty} 0 \iff \mathbb{E} h(L_n(\mathbf{f})) - \mathbb{E} h(Z_n(\mathbf{f})) \xrightarrow{N, n \rightarrow \infty} 0$$

but this immediately follows from:

Proposition 5.9. *Let (X_n) and (Y_n) be \mathbb{C}^d -valued random variables and assume that both sequences are tight, then the following are equivalent:*

- (i) *the following convergence holds true: $d_{\mathcal{L}P}(X_n, Y_n) \xrightarrow[n \rightarrow \infty]{} 0$.*
- (ii) *for every continuous bounded function $f : \mathbb{C}^d \rightarrow \mathbb{C}$, $\mathbb{E} f(X_n) - \mathbb{E} f(Y_n) \xrightarrow[n \rightarrow \infty]{} 0$,*

Proposition 5.9 can be proved easily by contradiction using the fact that $d_{\mathcal{L}P}$ metrizes the convergence of laws; its proof is hence omitted.

6. PROOF OF THEOREM 4.3 (CLT FOR NON-ANALYTIC FUNCTIONALS)

In this section, we will assume that the random variables (X_{ij}^n) are truncated, centered and normalized, following Section 3.2.

6.1. Proof of the fluctuations. Recall the definition of the sets D , D^+ and D_ε given in (3.7).

Lemma 6.1. *Let $(\varphi_n(z), z \in D^+)_{n \in \mathbb{N}}$ and $(\psi_n(z), z \in D^+)_{n \in \mathbb{N}}$ be centered complex-valued continuous random processes. Assume that:*

- (i) *The following convergence in distribution holds true: for all $d \geq 1$ and $(z_1, \dots, z_d) \in D^+$,*

$$d_{\mathcal{L}P}\left((\varphi_n(z_1), \dots, \varphi_n(z_d)), (\psi_n(z_1), \dots, \psi_n(z_d))\right) \xrightarrow[n \rightarrow \infty]{} 0$$

- (ii) *For all $\varepsilon > 0$, $\varphi_n(z)$ and $\psi_n(z)$ are tight on D_ε .*

- (iii) *The process $(\psi_n(z); z \in D^+)$ is gaussian with covariance matrix $\kappa_n(z_1, z_2)$, $(z_1, z_2 \in D^+)$.*

- (iv) *The following estimates hold true*

$$\forall n \in \mathbb{N}, \forall z \in D^+, \quad \text{var } \varphi_n(z) \leq \frac{1}{\text{Im}(z)^{2k}} \quad \text{and} \quad \text{var } \psi_n(z) \leq \frac{1}{\text{Im}(z)^{2k}}.$$

- (v) *Functions $g_\ell : \mathbb{R} \rightarrow \mathbb{R}$ ($1 \leq \ell \leq L$) are C^{k+1} in a η -neighbourhood of $[0, A]$.*

Then,

$$d_{\mathcal{L}P}\left(\frac{1}{\pi} \text{Re} \int_D \bar{\partial} \Phi(\mathbf{g})(z) \varphi_n(z) \lambda_2(dz), \frac{1}{\pi} \text{Re} \int_D \bar{\partial} \Phi(\mathbf{g})(z) \psi_n(z) \lambda_2(dz)\right) \xrightarrow[n \rightarrow \infty]{} 0,$$

where

$$\bar{\partial} \Phi(g_j)(z) = (\partial_x + i\partial_y) \sum_{\ell=0}^k \frac{(iy)^\ell}{\ell!} g_j^{(\ell)}(x) \chi(y) \quad \text{and} \quad \bar{\partial} \Phi(\mathbf{g}) = (\bar{\partial} \Phi(g_j); 1 \leq j \leq L)$$

with χ being smooth, compactly supported with value 1 in a neighbourhood of 0. Moreover,

$$\frac{1}{\pi} \text{Re} \int_D \bar{\partial} \Phi(\mathbf{g})(z) \psi_n(z) \lambda_2(dz)$$

is centered gaussian with covariance matrix:

$$\begin{aligned} \text{cov} \left(\frac{1}{\pi} \text{Re} \int_D \bar{\partial} \Phi(g_k)(z) \psi_n(z) \lambda_2(dz), \frac{1}{\pi} \text{Re} \int_D \bar{\partial} \Phi(g_\ell)(z) \psi_n(z) \lambda_2(dz) \right) \\ = \frac{1}{2\pi^2} \text{Re} \int_{D^2} \bar{\partial} \Phi(g_k)(z_1) \overline{\bar{\partial} \Phi(g_\ell)(z_2)} \kappa_n(z_1, \bar{z}_2) \lambda_2(dz_1) \lambda_2(dz_2) \\ + \frac{1}{2\pi^2} \text{Re} \int_{D^2} \bar{\partial} \Phi(g_k)(z_2) \bar{\partial} \Phi(g_\ell)(z_2) \kappa_n(z_1, z_2) \lambda_2(dz_1) \lambda_2(dz_2), \end{aligned} \quad (6.1)$$

for $1 \leq k, \ell \leq L$.

Proof of Lemma 6.1 is provided in Appendix A.2.

The strategy to prove Theorem 4.3 closely follows this lemma. Denote by

$$\varphi_n(z) = \text{tr} Q_n(z) - \mathbb{E} \text{tr} Q_n(z) \quad \text{and} \quad \psi_n(z) = N_n(z) - \mathbb{E} N_n(z),$$

the process N_n being defined in Theorem 3.1, then conditions (i), (ii) and (iii) are immediate consequences of Theorem 3.1. In order to check condition (iv), we establish the following proposition:

Proposition 6.2. *Assume that (A-1) and (A-2) hold true, then:*

(i) (Bordenave [12], Shcherbina [42]) For all $z \in \mathbb{C}^+$,

$$\text{var} Q_n(z) \leq \frac{C}{\text{Im}(z)^4},$$

(ii) For all $z \in \mathbb{C}^+$,

$$\text{var} N_n(z) \leq \frac{C}{\text{Im}(z)^4}.$$

In view of this proposition, it suffices to take functions g_ℓ 's satisfying (v) with $k = 2$ in Lemma 6.1. The fluctuation part of Theorem 4.3 immediately follows from Lemma 6.1.

Proof of Proposition 6.2. Recall the notations ξ_i, Q_i introduced in Section 5. Denote by $\Sigma^{\{i\}} = (\xi_1, \dots, \xi_{i-1}, \xi_i, \xi_{i+1}, \dots, x_n)$, where ξ_i is an independent copy of ξ_i . Let $Q^{\{i\}}$ be the associated resolvent:

$$Q^{\{i\}} = \left(\Sigma^{\{i\}} (\Sigma^{\{i\}})^{-1} - z I_N \right)^{-1}.$$

Rank-one perturbation formulas yield:

$$Q = Q_i - \frac{Q_i \xi_i \xi_i^* Q_i}{1 + \xi_i^* Q_i \xi_i} \quad \text{and} \quad Q^{\{i\}} = Q_i - \frac{Q_i \check{\xi}_i \check{\xi}_i^* Q_i}{1 + \check{\xi}_i^* Q_i \check{\xi}_i}.$$

We are now in position to apply Efron-Stein's inequality (cf. [14, Theorem 3.1]):

$$\begin{aligned} \text{var} \text{Tr} Q(z) &\leq \frac{1}{2} \sum_{i=1}^n \mathbb{E} \left| \text{Tr} Q(z) - \text{Tr} Q^{\{i\}}(z) \right|^2 \\ &= \frac{1}{2} \sum_{i=1}^n \mathbb{E} \left| \frac{\xi_i^* Q_i^2 \xi_i}{1 + \xi_i^* Q_i \xi_i} - \frac{\check{\xi}_i^* Q_i^2 \check{\xi}_i}{1 + \check{\xi}_i^* Q_i \check{\xi}_i} \right|^2 \\ &= \sum_{i=1}^n \mathbb{E} \left[\text{var}_i \left(\frac{\xi_i^* Q_i^2 \xi_i}{1 + \xi_i^* Q_i \xi_i} \right) \right] \stackrel{(a)}{\leq} \sum_{i=1}^n \mathbb{E} \left[\mathbb{E}_{\{i\}} \left| \frac{\xi_i^* Q_i^2 \xi_i}{1 + \xi_i^* Q_i \xi_i} - \frac{\mathbb{E}_{\{i\}} \xi_i^* Q_i^2 \xi_i}{1 + \mathbb{E}_{\{i\}} \xi_i^* Q_i \xi_i} \right|^2 \right] \end{aligned}$$

where var_i is the variance under the expectation $\mathbb{E}_{\{i\}}$ with respect to ξ_i , and (a) follows from the fact that $\text{var}(X) = \inf_a \mathbb{E}|X - a|^2$. Denote by

$$\overset{\circ}{X} = X - \mathbb{E}_{\{i\}}X, \quad B = \xi_i^* Q_i^2 \xi_i, \quad A = 1 + \xi_i^* Q_i \xi_i.$$

We have

$$\frac{B}{A} - \frac{\mathbb{E}_{\{i\}}B}{\mathbb{E}_{\{i\}}A} = \frac{\overset{\circ}{B}}{\mathbb{E}_{\{i\}}A} - \frac{B}{A} \frac{\overset{\circ}{A}}{\mathbb{E}_{\{i\}}A}.$$

Notice that $\text{Im}(A) = \text{Im}(z) (\xi_i^* Q_i Q_i^* \xi_i)$, hence

$$\left| \frac{B}{A} \right| \leq \frac{|\xi_i^* Q_i^2 \xi_i|}{\text{Im}(A)} \leq \frac{\xi_i^* Q_i Q_i^* \xi_i}{\text{Im}(A)} \leq \frac{1}{\text{Im}(z)}.$$

Now

$$\begin{aligned} \mathbb{E}_{\{i\}} \left(\frac{\xi_i^* Q_i^2 \xi_i}{1 + \xi_i^* Q_i \xi_i} - \frac{\mathbb{E}_{\{i\}} \xi_i^* Q_i^2 \xi_i}{1 + \mathbb{E}_{\{i\}} \xi_i^* Q_i \xi_i} \right)^2 &\leq K \mathbb{E}_{\{i\}} \left| \frac{\overset{\circ}{B}}{\mathbb{E}_{\{i\}}A} \right|^2 + K \mathbb{E}_{\{i\}} \left| \frac{B}{A} \frac{\overset{\circ}{A}}{\mathbb{E}_{\{i\}}A} \right|^2, \\ &\leq K \mathbb{E}_{\{i\}} \left| \frac{\overset{\circ}{B}}{\mathbb{E}_{\{i\}}A} \right|^2 + \frac{K}{\text{Im}(z)^2} \mathbb{E}_{\{i\}} \left| \frac{\overset{\circ}{A}}{\mathbb{E}_{\{i\}}A} \right|^2. \end{aligned}$$

We first focus on the estimation of $\mathbb{E}_{\{i\}} |\overset{\circ}{A}|^2$. Let $\delta \in (0, 1)$; we have

$$\begin{aligned} \mathbb{E}_{\{i\}} |\overset{\circ}{A}|^2 &\leq \frac{K}{n^2} \text{Tr}(Q_i R Q_i^* R), \\ &\leq \frac{K}{n^2} \|R\| \text{Tr}(Q_i R Q_i^*) \leq \frac{K}{\text{Im}(z)^{2\delta} n} \left[\frac{1}{n} \text{Tr}(Q_i R Q_i^*) \right]^{1-\delta}. \end{aligned}$$

Dividing by $|\mathbb{E}_{\{i\}}A|^2$ and noticing that $\text{Im}(E_i(A)) = \text{Im}(z) \frac{1}{n} \text{Tr}(Q_i R Q_i^*)$, we get:

$$\begin{aligned} \frac{\mathbb{E}_{\{i\}} |\overset{\circ}{A}|^2}{|\mathbb{E}_{\{i\}}A|^2} &\leq \frac{K}{\text{Im}(z)^{2\delta} n} \left[\frac{\frac{1}{n} \text{Tr}(Q_i R Q_i^*)}{\text{Im}(\mathbb{E}_{\{i\}}A)} \right]^{1-\delta} \frac{1}{|\mathbb{E}_{\{i\}}(A)|^{1+\delta}}, \\ &\leq \frac{K}{n} \frac{1}{\text{Im}(z)^{2\delta}} \frac{1}{\text{Im}(z)^{1-\delta}} \frac{1}{\text{Im}(z)^{1+\delta}} \leq \frac{K}{n} \frac{1}{\text{Im}(z)^{2+2\delta}}. \end{aligned}$$

Similarly, one can prove that

$$\frac{\mathbb{E}_{\{i\}} |\overset{\circ}{B}|^2}{|\mathbb{E}_{\{i\}}A|^2} \leq \frac{K}{n} \frac{1}{\text{Im}(z)^{4+2\delta}}.$$

From this we deduce

$$\text{var} R(z) \leq \frac{K}{\text{Im}(z)^{4+2\delta}}$$

and letting $\delta \rightarrow 0$, we get the desired estimate:

$$\text{var} R(z) \leq \frac{K}{\text{Im}(z)^4}.$$

In order to prove the second part of condition (iii), we rely on a meta-model argument (cf. Section 2.6). Denote by

$$M_{n,M}^1(z) = \text{Tr} (\Sigma_n(M) \Sigma_n^*(M) - z I_N)^{-1} - \mathbb{E} \text{Tr} (\Sigma_n(M) \Sigma_n^*(M) - z I_N)^{-1},$$

then

$$\text{var} \left\{ \text{tr} (\Sigma_n(M) \Sigma_n^*(M) - zI_N)^{-1} \right\} \leq \frac{C}{\text{Im}(z)^4},$$

moreover $M_{n,M}^1(z)$ converges in distribution to $\psi_n(z)$ as $M \rightarrow \infty$, N and n being fixed (see for instance the details in Section 5.2). Consider the continuous bounded function $h_K(x) = |x|^2 \wedge K$, then

$$\mathbb{E}h_K(\psi_n(z)) = \lim_{M \rightarrow \infty} \mathbb{E}h_K(M_{n,M}^1(z)) \leq \limsup_{M \rightarrow \infty} \mathbb{E}|M_{n,M}^1(z)|^2 \leq \frac{C}{\text{Im}(z)^4}.$$

Now letting $K \rightarrow \infty$ yields the desired bound by monotone convergence theorem:

$$\text{var}(\psi_n(z)) \leq \frac{C}{\text{Im}(z)^4}.$$

□

7. PROOF OF THEOREM 4.4 (BIAS FOR NON-ANALYTIC FUNCTIONALS)

Recall the notations $\Sigma_n^{\mathbb{C}}, Q_n^{\mathbb{C}}$, etc. introduced in Section 5. We split the bias into two terms

$$\begin{aligned} \mathbb{E}\text{Tr} f(\Sigma_n \Sigma_n^*) - N \int f(\lambda) \mathcal{F}_n(d\lambda) &= \mathbb{E}\text{Tr} f(\Sigma_n \Sigma_n^*) - \mathbb{E}\text{Tr} f(\Sigma_n^{\mathbb{C}} (\Sigma_n^{\mathbb{C}})^*) \\ &\quad + \mathbb{E}\text{Tr} f(\Sigma_n^{\mathbb{C}} (\Sigma_n^{\mathbb{C}})^*) - N \int f(\lambda) \mathcal{F}_n(d\lambda), \\ &\triangleq T_1 + T_2. \end{aligned}$$

We will prove the following: Provided that function f is of class C^8 ,

$$\mathbb{E}\text{Tr} f(\Sigma_n \Sigma_n^*) - \mathbb{E}\text{Tr} f(\Sigma_n^{\mathbb{C}} (\Sigma_n^{\mathbb{C}})^*) - \frac{1}{\pi} \text{Re} \int_{\mathbb{C}^+} \bar{\partial} \Phi(f)(z) \mathcal{B}_n(z) \lambda_2(dz) \xrightarrow{N, n \rightarrow \infty} 0. \quad (7.1)$$

Provided that function f satisfies (A-4), then:

$$\mathbb{E}\text{Tr} f(\Sigma_n^{\mathbb{C}} (\Sigma_n^{\mathbb{C}})^*) - N \int f(\lambda) \mathcal{F}_n(d\lambda) \xrightarrow{N, n \rightarrow \infty} 0. \quad (7.2)$$

As one can check, it is much more demanding in terms of assumptions to prove (7.2) than (7.1). Convergence in (7.2) should be compared to the results in Haagerup and Thørbjørnsen [25] (counterpart in the GUE case), Schultz [41] (GOE), Capitaine and Donati-Martin [16], Loubaton et al. [45] ('signal plus noise' model), etc.

7.1. Proof of (7.1). The heart of the proof lies in Helffer-Sjöstrand's formula, in Theorem 3.1 (bias part) and in a dominated convergence argument. By Theorem 3.1,

$$\mathbb{E}\text{Tr} (\Sigma_n \Sigma_n^* - zI_N)^{-1} - N t_n(z) - \mathcal{B}_n(z) \xrightarrow{N, n \rightarrow \infty} 0.$$

The same argument yields:

$$\mathbb{E}\text{Tr} (\Sigma_n^{\mathbb{C}} (\Sigma_n^{\mathbb{C}})^* - zI_N)^{-1} - N t_n(z) \xrightarrow{N, n \rightarrow \infty} 0,$$

because in the later case $\mathcal{V} = \kappa = 0$, hence the bias is zero for the matrix model $\Sigma_n^{\mathbb{C}} (\Sigma_n^{\mathbb{C}})^*$. Subtracting yields:

$$\mathbb{E}\text{Tr} Q_n(z) - \mathbb{E}\text{Tr} Q_n^{\mathbb{C}}(z) - \mathcal{B}_n(z) \xrightarrow{N, n \rightarrow \infty} 0.$$

Recall that by Proposition 5.4,

$$|\mathbb{E}\mathrm{Tr} Q_n(z) - \mathbb{E}\mathrm{Tr} Q_n^{\mathbb{C}}(z)| \leq K \frac{|z|^3}{\mathrm{Im}(z)^7} . \quad (7.3)$$

In order to transfer this bound to $\mathcal{B}_n(z)$, we invoke a meta-model argument (cf. Section 2.6): Consider matrix $\Sigma_n(M)$ and its counterpart $\Sigma_n^{\mathbb{C}}(M)$ as defined in (2.24) and recall that in this case, we have a genuine limit:

$$\mathbb{E}\mathrm{Tr} (\Sigma_n(M)\Sigma_n^*(M) - zI_{NM})^{-1} - \mathbb{E}\mathrm{Tr} (\Sigma_n^{\mathbb{C}}(M)(\Sigma_n^{\mathbb{C}}(M))^* - zI_{NM})^{-1} \xrightarrow[\substack{M \rightarrow \infty \\ N, n \text{ fixed}}]{} \mathcal{B}_n(z) .$$

Since the estimate (7.3) remains true for all $M \geq 1$, we obtain:

$$\begin{aligned} |\mathcal{B}_n(z)| &= \lim_{M \rightarrow \infty} \left| \mathbb{E}\mathrm{Tr} (\Sigma_n(M)\Sigma_n(M)^* - zI_{MN})^{-1} - \mathbb{E}\mathrm{Tr} (\Sigma_n^{\mathbb{C}}(M)_n(\Sigma_n^{\mathbb{C}}(M))^* - zI_{NM})^{-1} \right| \\ &\leq K \frac{|z|^3}{\mathrm{Im}(z)^7} . \end{aligned} \quad (7.4)$$

Write

$$\begin{aligned} \mathbb{E}\mathrm{Tr} f(\Sigma_n \Sigma_n^*) - \mathbb{E}\mathrm{Tr} f(\Sigma_n^{\mathbb{C}} (\Sigma_n^{\mathbb{C}})^*) &- \frac{1}{\pi} \mathrm{Re} \int_{\mathbb{C}_+} \bar{\partial} \Phi(f)(z) \mathcal{B}_n(z) \lambda_2(dz) \\ &= \frac{1}{\pi} \mathrm{Re} \int_{\mathbb{C}_+} \bar{\partial} \Phi(f)(z) \{ \mathbb{E}\mathrm{Tr} Q_n(z) - \mathbb{E}\mathrm{Tr} Q_n^{\mathbb{C}}(z) - \mathcal{B}_n(z) \} \lambda_2(dz) , \end{aligned} \quad (7.5)$$

In view of (7.5), we need a dominated convergence argument in order to prove (7.1); such an argument follows from Proposition 4.2, (7.3) and (7.4) as long as f is of class C^8 with large but bounded support. This concludes the proof of (7.1).

7.2. Proof of (7.2). The gist of the proof lies in the following proposition whose proof is postponed to Appendix A.3:

Proposition 7.1. *Denote by $P_\ell(X)$ a polynomial in X with degree ℓ and positive coefficients, then:*

$$\left| \mathbb{E}\mathrm{Tr} \left(\Sigma_n^{\mathbb{C}} (\Sigma_n^{\mathbb{C}})^* - zI_N \right)^{-1} - N t_n(z) \right| \leq \frac{1}{n} P_{12}(|z|) P_{17}(|\mathrm{Im}(z)|^{-1}) .$$

Using Helffer-Sjöstrand's formula, Proposition 7.1 together with Proposition 4.2 immediately yield (7.2) for any f of class C^{18} with large but bounded support.

APPENDIX A. REMAINING PROOFS

A.1. Proof of Proposition 5.1: remaining computations for the bias. In this section, we outline the proof of identity (5.36) which we recall below:

$$\begin{aligned} &- |\mathcal{V}|^2 \frac{b_n^2}{n} \mathbb{E}\mathrm{tr} R_n^{1/2} Q_1 \left(\mathbb{E} \tilde{f}_n R_n + I_N \right)^{-1} R_n^{1/2} \bar{R}_n^{1/2} Q_1^T \bar{R}_n^{1/2} \\ &= |\mathcal{V}|^2 \frac{\frac{z^3 \bar{t}_n^2}{n} \mathrm{tr} R_n^{1/2} T_n^2(z) R_n^{1/2} \bar{R}_n^{1/2} T_n^T(z) \bar{R}_n^{1/2}}{1 - \frac{|\mathcal{V}|^2 z^2 \bar{t}_n^2}{n} \mathrm{tr} R_n^{1/2} T_n(z) R_n^{1/2} \bar{R}_n^{1/2} T_n^T(z) \bar{R}_n^{1/2}} + o(1) . \end{aligned} \quad (\text{A.1})$$

The proof closely follows computations in [6, Section 4] and is essentially a matter of book-keeping; in particular, all the estimates established there remain valid in the context where R_n and X_n are not real. We shall focus here on the algebraic identities.

We first replace Q_1 by Q and approximate Q by (cf. [6, Eq. 4.13]):

$$Q(z) = -(zI_N - b_n(z)R_n)^{-1} + b_n(z)A(z) + B(z) + C(z) \quad (\text{A.2})$$

where

$$A(z) = \sum_{j=1}^n (zI_N - b_n(z)R_n)^{-1} (\xi_j \xi_j^* - n^{-1}R_n) Q_j(z) .$$

The terms $B(z)$ and $C(z)$ will not contribute in the sequel. Denote by

$$\begin{aligned} M &= (\mathbb{E}\tilde{f}_n R_n + I_N)^{-1} R_n^{1/2} \bar{R}_n^{1/2}, \\ \mathcal{T} &= \frac{1}{n} \mathbb{E} \text{tr} R_n^{1/2} Q_1 \left(\mathbb{E}\tilde{f}_n R_n + I_N \right)^{-1} R_n^{1/2} \bar{R}_n^{1/2} Q_1^T \bar{R}_n^{1/2} . \end{aligned}$$

We have:

$$\begin{aligned} \mathcal{T} &= \frac{1}{n} \mathbb{E} \text{tr} R_n^{1/2} Q_1 M Q_1^T \bar{R}_n^{1/2} \\ &= \frac{1}{n} \mathbb{E} \text{tr} R_n^{1/2} Q M Q^T \bar{R}_n^{1/2} + o(1) \\ &= -\frac{1}{n} \mathbb{E} \text{tr} R_n^{1/2} (zI_N - b_n(z)R_n)^{-1} M Q^T R_n^{1/2} + \frac{b_n(z)}{n} \mathbb{E} \text{tr} R_n^{1/2} A(z) M Q^T \bar{R}_n^{1/2} + o(1) \\ &\triangleq \mathcal{T}_1 + \mathcal{T}_2 + o(1) \end{aligned} \quad (\text{A.3})$$

In order to compute \mathcal{T}_1 , we approximate Q^T in the same way as in (A.2); we take into account the fact that for some deterministic matrix Γ , $\mathbb{E} \text{tr} (\Gamma A) = 0$; we also use the approximation $b_n(z) = -z\tilde{t}_n(z) + o(1)$ and equation (3.1). The computation of \mathcal{T}_1 then easily follows:

$$\begin{aligned} \mathcal{T}_1 &= -\frac{1}{n} \mathbb{E} \text{tr} R_n^{1/2} (zI_N - b_n(z)R_n)^{-1} M Q^T \bar{R}_n^{1/2} \\ &= \frac{1}{n} \text{tr} R_n^{1/2} (zI_N - b_n(z)R_n)^{-1} M (zI_N - b_n(z)\bar{R}_n)^{-1} \bar{R}_n^{1/2} + o(1) \\ &= -\frac{z}{n} \text{tr} R_n^{1/2} T_n^2(z) R_n^{1/2} \bar{R}_n^{1/2} T_n^T(z) \bar{R}_n^{1/2} + o(1) . \end{aligned}$$

We now focus on the term

$$\begin{aligned} \mathcal{T}_2 &= \frac{b_n(z)}{n} \mathbb{E} \text{tr} R_n^{1/2} A(z) M Q^T \bar{R}_n^{1/2} \\ &= \frac{b_n(z)}{n} \mathbb{E} \text{tr} R_n^{1/2} \sum_{j=1}^n (zI_N - b_n(z)R_n)^{-1} (\xi_j \xi_j^* - n^{-1}R_n) Q_j(z) M Q^T \bar{R}_n^{1/2} \\ &= \frac{b_n(z)}{n} \mathbb{E} \text{tr} R_n^{1/2} \sum_{j=1}^n (zI_N - b_n(z)R_n)^{-1} \{ \xi_j \xi_j^* Q_j(z) M (Q^T - Q_j^T) + D(z) + E(z) \} \bar{R}_n^{1/2} \end{aligned}$$

where

$$\begin{aligned} D(z) &= \xi_j \xi_j^* Q_j M Q_j^T - n^{-1} R_n M Q_j M Q_j^T \\ E(z) &= n^{-1} R_n M (Q_j^T - Q^T) \end{aligned}$$

will not contribute. Using the rank-one perturbation identity for $Q^T - Q_j^T$, we obtain:

$$\begin{aligned}
\mathcal{T}_2 &= \frac{b_n(z)}{n} \mathbb{E} \text{tr} R_n^{1/2} \sum_{j=1}^n (zI_N - b_n(z)R_n)^{-1} \xi_j \xi_j^* Q_j(z) M (Q^T - Q_j^T) \bar{R}^{1/2} + o(1) \\
&= -\frac{b_n(z)}{n} \mathbb{E} \text{tr} R_n^{1/2} \sum_{i=1}^n (zI_N - b_n(z)R_n)^{-1} \xi_j \xi_j^* Q_j(z) M \frac{Q_j^T \bar{\xi}_j \bar{\xi}_j^* Q_j^T}{1 + \bar{\xi}_j^* Q_j^T \bar{\xi}_j} \bar{R}^{1/2} + o(1) \\
&= -\frac{b_n(z)}{n} \sum_{j=1}^n \mathbb{E} \frac{1}{1 + \bar{\xi}_j^* Q_j^T \bar{\xi}_j} \left(\bar{\xi}_j^* Q_j^T \bar{R}^{1/2} R_n^{1/2} (zI_N - b_n(z)R_n)^{-1} \xi_j \right) \left(\xi_j^* Q_j(z) M Q_j^T \bar{\xi}_j \right) + o(1).
\end{aligned}$$

In order to pursue the computation of \mathcal{T}_2 , we shall perform the following approximations: The quantity $(1 + \bar{\xi}_j^* Q_j^T \bar{\xi}_j)^{-1}$ can be replaced by b_n and the two remaining quadratic forms in the expectation can be decorrelated. Now, using formulas (5.9), we obtain:

$$\begin{aligned}
\mathcal{T}_2 &= -\frac{b_n^2(z)}{n} \sum_{j=1}^n \mathbb{E} \left(\bar{\xi}_j^* Q_j^T \bar{R}^{1/2} R_n^{1/2} (zI_N - b_n(z)R_n)^{-1} \xi_j \right) \mathbb{E} \left(\xi_j^* Q_j(z) M Q_j^T \bar{\xi}_j \right) + o(1), \\
&= -\frac{|\mathcal{V}|^2 b_n^2(z)}{n} \sum_{j=1}^n \mathbb{E} \text{tr} \left(\bar{R}^{1/2} Q_j^T \bar{R}^{1/2} R_n^{1/2} (zI_N - b_n(z)R_n)^{-1} R^{1/2} \right) \mathbb{E} \text{tr} \left(R^{1/2} Q_j(z) M Q_j^T \bar{R}^{1/2} \right) + o(1).
\end{aligned}$$

We can now replace Q_j by Q with no loss and use equation (A.2) to obtain:

$$\begin{aligned}
\mathcal{T}_2 &= -\frac{|\mathcal{V}|^2 b_n^2(z)}{n^2} \mathbb{E} \text{tr} \left(\bar{R}^{1/2} Q^T(z) \bar{R}^{1/2} R_n^{1/2} (zI_N - b_n(z)R_n)^{-1} R^{1/2} \right) \\
&\quad \times \mathbb{E} \text{tr} \left(R^{1/2} (-zI_N + b_n(z)R_n)^{-1} + b_n(z)A(z) \right) M Q^T \bar{R}^{1/2} + o(1), \\
&= |\mathcal{V}|^2 b_n^2(z) \frac{1}{n} \text{tr} R^{1/2} T(z) R^{1/2} \bar{R}^{1/2} T^T(z) \bar{R}^{1/2} (\mathcal{T}_1 + \mathcal{T}_2) + o(1). \tag{A.4}
\end{aligned}$$

Denote by

$$\mathcal{T}_3 = \frac{1}{n} \text{tr} R^{1/2} T(z) R^{1/2} \bar{R}^{1/2} T^T(z) \bar{R}^{1/2}.$$

We now extract \mathcal{T}_2 from (A.4) and plug it into (A.3). We finally obtain:

$$\mathcal{T} = \mathcal{T}_1 + |\mathcal{V}|^2 b_n^2(z) \frac{\mathcal{T}_1 \mathcal{T}_3}{1 - |\mathcal{V}|^2 b_n^2(z) \mathcal{T}_3} + o(1) = \frac{\mathcal{T}_1}{1 - |\mathcal{V}|^2 b_n^2(z) \mathcal{T}_3} + o(1).$$

Multiplying \mathcal{T} by $-|\mathcal{V}|^2 b_n^2(z) = -|\mathcal{V}|^2 z^2 \bar{t}_n^2(z)$ finally yields (A.1).

A.2. Proof of Lemma 6.1. By Proposition 4.2,

$$\begin{aligned}
\mathbb{E} \left| \int_D \bar{\partial} \Phi(g)(z) \varphi_n(z) \lambda_2(dz) \right| &\leq \int_D |\bar{\partial} \Phi(g)(z)| \mathbb{E} |\varphi_n(z)| \lambda_2(dz) \\
&\leq \|g\|_{k+1, \infty} \int_D \text{Im}(z)^k \{\text{var} \varphi_n(z)\}^{1/2} \lambda_2(dz) < \infty,
\end{aligned}$$

by (iii) and (iv). Hence $\frac{1}{\pi} \text{Re} \int_D \bar{\partial} \Phi(g)(z) \varphi_n(z) \lambda_2(dz)$ is a well-defined a.s. finite random variable. This estimate, uniform in n , readily implies the tightness of

$$\left(\frac{1}{\pi} \text{Re} \int_D \bar{\partial} \Phi(g)(z) \varphi_n(z) \lambda_2(dz); n \in \mathbb{N} \right).$$

Notice that the integrals with ψ_n instead of φ_n are similarly well-defined and tight.

By conditions (i) and (ii), we obtain:

$$d_{\mathcal{L}P} \left(\frac{1}{\pi} \operatorname{Re} \int_{D_\varepsilon} \bar{\partial} \Phi(\mathbf{g})(z) \varphi_n(z) \lambda_2(dz), \frac{1}{\pi} \operatorname{Re} \int_{D_\varepsilon} \bar{\partial} \Phi(\mathbf{g})(z) \psi_n(z) \lambda_2(dz) \right) \xrightarrow{N, n \rightarrow \infty} 0 \quad (\text{A.5})$$

(apply Lemma 5.8).

Let $\mathbf{g} = (g_\ell; 1 \leq \ell \leq L)$ and $f : \mathbb{C}^L \rightarrow \mathbb{C}$ be bounded and continuous. Consider the following notations:

$$\begin{aligned} \xi_n &= \frac{1}{\pi} \operatorname{Re} \int_D \bar{\partial} \Phi(\mathbf{g})(z) \varphi_n(z) \lambda_2(dz), & \xi_n^\varepsilon &= \frac{1}{\pi} \operatorname{Re} \int_{D_\varepsilon} \bar{\partial} \Phi(\mathbf{g})(z) \varphi_n(z) \lambda_2(dz), \\ \eta_n &= \frac{1}{\pi} \operatorname{Re} \int_D \bar{\partial} \Phi(\mathbf{g})(z) \psi_n(z) \lambda_2(dz), & \eta_n^\varepsilon &= \frac{1}{\pi} \operatorname{Re} \int_{D_\varepsilon} \bar{\partial} \Phi(\mathbf{g})(z) \psi_n(z) \lambda_2(dz). \end{aligned}$$

We have

$$\begin{aligned} |\mathbb{E}f(\xi_n) - \mathbb{E}f(\eta_n)| \\ \leq |\mathbb{E}f(\xi_n) - \mathbb{E}f(\xi_n^\varepsilon)| + |\mathbb{E}f(\xi_n^\varepsilon) - \mathbb{E}f(\eta_n^\varepsilon)| + |\mathbb{E}f(\eta_n^\varepsilon) - \mathbb{E}f(\eta_n)|. \end{aligned} \quad (\text{A.6})$$

Given $\rho > 0$, we first prove that for all $n \geq 1$,

$$|\mathbb{E}f(\xi_n) - \mathbb{E}f(\xi_n^\varepsilon)| \leq (4\|f\|_\infty + 1)\rho \quad (\text{A.7})$$

for ε small enough.

We have

$$\mathbb{P}\{|\xi_n - \xi_n^\varepsilon| > \delta\} \leq \frac{1}{\delta} \left(\int_{[0, A] + i[0, \varepsilon]} |\bar{\partial} \Phi(\mathbf{g})(z)| \mathbb{E}|\varphi_n(z)| \lambda_2(dz) \right) \quad (\text{A.8})$$

which can be made arbitrarily small if ε is small enough, independently from n . Now,

$$\begin{aligned} |\mathbb{E}f(\xi_n) - \mathbb{E}f(\xi_n^\varepsilon)| &\leq |\mathbb{E}f(\xi_n) - \mathbb{E}f(\xi_n^\varepsilon)| \mathbf{1}_{\{|\xi_n - \xi_n^\varepsilon| > \eta\}} \\ &\quad + |\mathbb{E}f(\xi_n) - \mathbb{E}f(\xi_n^\varepsilon)| \mathbf{1}_{\{|\xi_n - \xi_n^\varepsilon| \leq \eta, |\xi_n| \vee |\xi_n^\varepsilon| > K\}} \\ &\quad + |\mathbb{E}f(\xi_n) - \mathbb{E}f(\xi_n^\varepsilon)| \mathbf{1}_{\{|\xi_n - \xi_n^\varepsilon| \leq \eta, |\xi_n| \vee |\xi_n^\varepsilon| \leq K\}}. \end{aligned}$$

First invoke the tightness of $|\xi_n| \vee |\xi_n^\varepsilon|$ and choose K large enough so that the second term of the r.h.s. is lower than $2\|f\|_\infty \rho$; then choose $\eta > 0$ small enough so that f being absolutely continuous over $\{z \in \mathbb{C}^+, |z| \leq K\}$, the third term of the r.h.s. is lower than ρ ; finally for such K and η , take advantage of (A.8) and choose ε small enough so that the first term of the r.h.s. is lower than $2\|f\|_\infty \rho$. Eq. (A.7) is proved.

One can similarly prove that $|\mathbb{E}f(\eta_n) - \mathbb{E}f(\eta_n^\varepsilon)| \leq (4\|f\|_\infty + 1)\rho$ for $\varepsilon > 0$ small enough. Such ε being fixed, it remains to control the second term of the r.h.s. of (A.6), but this immediately follows from (A.5).

In order to prove that η_n is multivariate gaussian with prescribed covariance (6.1), we first consider η_n^ε . Approximating the integral in η_n^ε by Riemann sums and using the fact that weak limits of gaussian vectors are gaussian immediately yields that η_n^ε is a gaussian vector with covariance matrix:

$$[\operatorname{cov}(\eta_n^\varepsilon)]_{k\ell} = \frac{1}{\pi^2} \mathbb{E} \left\{ \operatorname{Re} \int_{D^\varepsilon} \bar{\partial} \Phi(g_k)(z) \psi_n(z) \lambda_2(dz) \operatorname{Re} \int_{D^\varepsilon} \bar{\partial} \Phi(g_\ell)(z) \psi_n(z) \lambda_2(dz) \right\}$$

for $1 \leq k, \ell \leq L$. Using the elementary identity:

$$\operatorname{Re}(z)\operatorname{Re}(z') = \frac{\operatorname{Re}(z\bar{z}') + \operatorname{Re}(zz')}{2},$$

we obtain:

$$\begin{aligned} [\operatorname{cov}(\eta_n^\varepsilon)]_{k\ell} &= \frac{1}{2\pi^2} \operatorname{Re} \int_{(D^\varepsilon)^2} \bar{\partial}\Phi(g_k)(z_1) \overline{\bar{\partial}\Phi(g_\ell)(z_2)} \mathbb{E}\psi_n(z_1) \overline{\psi_n(z_2)} \lambda_2(dz_1) \lambda_2(dz_2) \\ &\quad + \frac{1}{2\pi^2} \operatorname{Re} \int_{(D^\varepsilon)^2} \bar{\partial}\Phi(g_k)(z_1) \bar{\partial}\Phi(g_\ell)(z_2) \mathbb{E}\psi_n(z_1) \psi_n(z_2) \lambda_2(dz_1) \lambda_2(dz_2). \end{aligned}$$

Using the fact that $\overline{\psi_n(z_2)} = \psi_n(\bar{z}_2)$ yields:

$$\begin{aligned} [\operatorname{cov}(\eta_n^\varepsilon)]_{k\ell} &= \frac{1}{2\pi^2} \operatorname{Re} \int_{(D^\varepsilon)^2} \bar{\partial}\Phi(g_k)(z_1) \overline{\bar{\partial}\Phi(g_\ell)(z_2)} \kappa_n(z_1, \bar{z}_2) \lambda_2(dz_1) \lambda_2(dz_2) \\ &\quad + \frac{1}{2\pi^2} \operatorname{Re} \int_{(D^\varepsilon)^2} \bar{\partial}\Phi(g_k)(z_1) \bar{\partial}\Phi(g_\ell)(z_2) \kappa_n(z_1, z_2) \lambda_2(dz_1) \lambda_2(dz_2). \end{aligned}$$

In order to lift the gaussianity from η_n^ε to η_n and to extend the covariance formula from the one above to formula (6.1), we rely on the approximation theorem [34, Theorem 4.28] and on assumptions (iv) and (v) on the variance estimates and on the regularity of functions g_k, g_ℓ in Lemma 6.1.

Proof of Lemma 6.1 is completed.

A.3. Proof of Proposition 7.1. In the whole section, we consider the matrix model $\Sigma_n^{\mathbb{C}}(\Sigma_n^{\mathbb{C}})^*$; we however simply write f_n, \tilde{f}_n, Q_n while the underlying random variables are $\mathcal{N}_{\mathbb{C}}(0, 1)$.

Recall that $\tilde{f}_n(z) = -\frac{1-c_n}{z} + f_n(z)$, where $f_n(z) = \frac{1}{n} \operatorname{Tr} Q_n(z)$. Denote by

$$S_n(z) = \left(-z \left(I_N + \mathbb{E}\tilde{f}_n(z) R_n \right) \right)^{-1} \quad \text{and} \quad \tilde{\tau}_n(z) = -\frac{1}{z \left(1 + \frac{1}{n} \mathbb{E}\operatorname{Tr} R_n Q_n(z) \right)}.$$

Let

$$\varepsilon_n(z) = n(\mathbb{E}\tilde{f}_n(z) - \tilde{\tau}_n(z)), \quad (\text{A.9})$$

$$\tilde{\varepsilon}_n(z) = \mathbb{E}\operatorname{Tr} R_n Q_n(z) - \operatorname{Tr} R_n S_n(z). \quad (\text{A.10})$$

Taking into account the definition of $\tilde{\tau}_n$ and \tilde{t}_n , we get:

$$\begin{aligned} n(\mathbb{E}\tilde{f}_n - \tilde{t}_n) &= n(\tilde{\tau}_n - \tilde{t}_n) + \varepsilon_n \\ &= \frac{1}{z} \frac{\operatorname{Tr} R_n T_n - \operatorname{Tr} R_n \mathbb{E}Q_n}{\left(1 + \frac{1}{n} \mathbb{E}\operatorname{Tr} R_n Q_n \right) \left(1 + \frac{1}{n} \operatorname{Tr} R_n T_n \right)} + \varepsilon_n. \end{aligned}$$

Similarly,

$$\begin{aligned} \mathbb{E}\operatorname{Tr} R_n Q_n - \operatorname{Tr} R_n T_n &= \operatorname{Tr} R_n S_n - \operatorname{Tr} R_n T_n + \tilde{\varepsilon}_n \\ &= -\frac{1}{z} \operatorname{Tr} R_n \left[(I_N + \mathbb{E}\tilde{f}_n R_n)^{-1} - (I_N + \tilde{t}_n R_n)^{-1} \right] + \tilde{\varepsilon}_n \\ &= \frac{1}{z} \operatorname{Tr} R_n (I_N + \mathbb{E}\tilde{f}_n R_n)^{-1} R_n (I_N + \tilde{t}_n R_n)^{-1} (\mathbb{E}\tilde{f}_n - \tilde{t}_n) + \tilde{\varepsilon}_n. \end{aligned}$$

Gathering the two previous equations, we finally end up with the 2×2 linear system:

$$\begin{pmatrix} n(\mathbb{E}\tilde{f}_n - \tilde{t}_n) \\ \mathbb{E}\text{Tr} R_n Q_n - \text{Tr} R_n T_n \end{pmatrix} = D_0(z) \begin{pmatrix} n(\mathbb{E}\tilde{f}_n - \tilde{t}_n) \\ \mathbb{E}\text{Tr} R_n Q_n - \text{Tr} R_n T_n \end{pmatrix} + \begin{pmatrix} \varepsilon_n \\ \tilde{\varepsilon}_n \end{pmatrix}$$

where

$$D_0(z) = \begin{pmatrix} 0 & z\tilde{\tau}_n\tilde{t}_n \\ \frac{z}{n}\text{Tr} R_n S_n R_n T_n & 0 \end{pmatrix},$$

from which we extract

$$n \left(\mathbb{E}\tilde{f}_n - \tilde{t}_n \right) = \frac{1}{\det(I_2 - D_0(z))} \left(\varepsilon_n(z) + z\tilde{\tau}_n\tilde{t}_n\tilde{\varepsilon}_n(z) \right). \quad (\text{A.11})$$

It remains to bound ε_n and $\tilde{\varepsilon}_n$ and to lowerbound $|\det(I_2 - D_0)|$. The first task relies on standard gaussian calculus for random matrices (Poincaré-Nash inequality, integration by part formula, etc. - see for instance [40]) and yields:

$$|\varepsilon_n(z)| \leq \frac{1}{n} P_2(|z|) P_5(|\text{Im}(z)|^{-1}), \quad (\text{A.12})$$

$$|\tilde{\varepsilon}_n(z)| \leq \frac{1}{n} P_3(|z|) P_7(|\text{Im}(z)|^{-1}). \quad (\text{A.13})$$

Details are omitted and can be found in [46, Chapter 3].

The second task is more involved and goes along the lines developed in [25] and [45]; it relies on the following proposition:

Proposition A.1. *There exist $\eta > 0$, polynomials P_{12}^\dagger and P_{16}^\dagger and an integer N_0 such that for every z in the set*

$$\mathcal{E}_n = \left\{ z \in \mathbb{C}^+, 1 - \frac{1}{n^2} P_{12}^\dagger(|z|) P_{16}^\dagger(|\text{Im}(z)|^{-1}) > 0 \right\}$$

and for $N \geq N_0$

$$|\det(I_2 - D_0(z))| > \frac{K |\text{Im}(z)|^4}{(\eta^2 + |z|^2)^2},$$

where K is some constant independent from N, n .

Assume for a while that Proposition A.1 holds true, and let $z \in \mathcal{E}_n$; then taking into account (A.12)-(A.13) and the fact that $|z\tilde{\tau}_n\tilde{t}_n| \leq |z| |\text{Im}(z)|^{-2}$, (A.11) yields:

$$|n(\mathbb{E}\tilde{f}_n(z) - \tilde{t}_n(z))| \leq P_8(|z|) P_{13}(|\text{Im}(z)|^{-1}).$$

If $z \notin \mathcal{E}_n$, then $1 \leq \frac{1}{n^2} P_{12}^\dagger(|z|) P_{16}^\dagger(|\text{Im}(z)|^{-1})$ and

$$\begin{aligned} |n(\mathbb{E}\tilde{f}_n(z) - \tilde{t}_n(z))| &\leq n \left(|\mathbb{E}\tilde{f}_n(z)| + |\tilde{t}_n(z)| \right) \leq \frac{2n}{|\text{Im}(z)|^{-1}} \times 1 \\ &\leq \frac{2n}{|\text{Im}(z)|^{-1}} \times \frac{1}{n^2} P_{12}^\dagger(|z|) P_{16}^\dagger(|\text{Im}(z)|^{-1}) \\ &\leq \frac{1}{n} P_{12}^\dagger(|z|) P_{17}(|\text{Im}(z)|^{-1}). \end{aligned}$$

Proof of Proposition 7.1 is completed as long as Proposition A.1 holds true.

Proof of Proposition A.1. In order to lowerbound $\det(I_2 - D_0(z))$, we introduce two auxiliary systems.

Consider matrix

$$D(z) = \begin{pmatrix} 0 & |\tilde{t}_n(z)|^2 \\ \frac{|z|^2}{n} \text{Tr} R_n T_n R_n T_n^* & 0 \end{pmatrix},$$

In order to evaluate the determinant of matrix $I_2 - D$, we need to find an equation where such a matrix appears. One can easily prove that:

$$\begin{pmatrix} \text{Im}(\tilde{t}_n) \\ \text{Im}\left(\frac{z}{n} \text{Tr} R_n T_n\right) \end{pmatrix} = D(z) \begin{pmatrix} \text{Im}(\tilde{t}_n) \\ \text{Im}\left(\frac{z}{n} \text{Tr} R_n T_n\right) \end{pmatrix} + \begin{pmatrix} |\tilde{t}_n|^2 \\ 0 \end{pmatrix} \text{Im}(z), \quad (\text{A.14})$$

from which we extract the determinant:

$$\det(I_2 - D(z)) = 1 - |z|^2 |\tilde{t}_n|^2 \frac{1}{n} \text{Tr} R_n T_n R_n T_n^* = |\tilde{t}|^2 \frac{\text{Im}(z)}{\text{Im}(\tilde{t}_n)}. \quad (\text{A.15})$$

Recall that $|\tilde{t}_n(z)| \leq (\text{Im}(z))^{-1}$; in order to lowerbound $\text{Im}(\tilde{t}_n(z))$, recall that the associated probability measures $\tilde{\mathcal{F}}_n$ form a tight family and in particular there exists $\eta > 0$ such that $\tilde{\mathcal{F}}_n([0, \eta]) > 1/2$ for every $n \geq 1$. Write:

$$\text{Im}(\tilde{t}_n(z)) = \text{Im}(z) \int_{\mathbb{R}^+} \frac{\tilde{\mathcal{F}}_n(d\lambda)}{|\lambda - z|^2} \geq \int_0^\eta \frac{\tilde{\mathcal{F}}_n(d\lambda)}{|\lambda - z|^2} \geq \frac{1}{4(\eta^2 + |z|^2)}. \quad (\text{A.16})$$

Plugging these estimates into (A.15) yields the bound:

$$\det(I_2 - D(z)) = 1 - |z|^2 |\tilde{t}_n|^2 \frac{1}{n} \text{Tr} R_n T_n R_n T_n^* \geq \frac{\text{Im}(z)^4}{16(\eta^2 + |z|^2)^2}. \quad (\text{A.17})$$

Consider now matrix

$$D'(z) = \begin{pmatrix} 0 & |\tilde{\tau}_n(z)|^2 \\ \frac{|z|^2}{n} \text{Tr} R_n S_n R_n S_n^* & 0 \end{pmatrix}.$$

In order to evaluate the determinant of matrix $I_2 - D'$, we need to find an equation where such a matrix appears. Recall the definition of ε_n and $\tilde{\varepsilon}_n$ in (A.9)-(A.10). Taking their imaginary parts, we obtain:

$$\begin{pmatrix} \text{Im}(\mathbb{E}\tilde{f}_n) \\ \text{Im}\left(\frac{z}{n} \text{Tr} \mathbb{E}R_n Q_n\right) \end{pmatrix} = \begin{pmatrix} \text{Im}(\tilde{\tau}_n) \\ \text{Im}\left(\frac{z}{n} \text{Tr} R_n S_n\right) \end{pmatrix} + \begin{pmatrix} \frac{1}{n} \text{Im}(\varepsilon_n) \\ \frac{1}{n} \text{Im}(\tilde{\varepsilon}_n) \end{pmatrix}.$$

Computing the imaginary part of $\tilde{\tau}_n$ and $n^{-1} \text{Tr} R_n S_n$, we get:

$$\begin{aligned} \text{Im}(\tilde{\tau}_n) &= |\tilde{\tau}_n|^2 \text{Im}(z) + |\tilde{\tau}_n|^2 \text{Im}\left(\frac{z}{n} \text{Tr} \mathbb{E}R_n Q_n\right), \\ \text{Im}\left(\frac{z}{n} \text{Tr} R_n S_n\right) &= \frac{|z|^2}{n} \text{Tr}(R_n S_n R_n S_n^*) \text{Im}(\mathbb{E}\tilde{f}_n). \end{aligned}$$

Hence the system:

$$\begin{pmatrix} \text{Im}(\mathbb{E}\tilde{f}_n) \\ \text{Im}\left(\frac{z}{n} \text{Tr} \mathbb{E}R_n Q_n\right) \end{pmatrix} = D'(z) \begin{pmatrix} \text{Im}(\mathbb{E}\tilde{f}_n) \\ \text{Im}\left(\frac{z}{n} \text{Tr} \mathbb{E}R_n Q_n\right) \end{pmatrix} + \begin{pmatrix} |\tilde{\tau}_n|^2 \\ 0 \end{pmatrix} \text{Im}(z) + \begin{pmatrix} \frac{1}{n} \text{Im}(\varepsilon_n) \\ \frac{1}{n} \text{Im}(\tilde{\varepsilon}_n) \end{pmatrix}. \quad (\text{A.18})$$

This system is similar to (A.14) with two extra error terms $\frac{1}{n} \text{Im}(\varepsilon_n)$ and $\frac{1}{n} \text{Im}(\tilde{\varepsilon}_n)$; but these terms are controlled by (A.12)-(A.13).

It is well-known that for all $z \in \mathbb{C}^+$, $\mathbb{E}\tilde{f}_n(z) - \tilde{t}_n(z) \rightarrow 0$ as $N, n \rightarrow \infty$. The mere definition of ε_n together with estimate (A.12) yields $\tilde{\tau}_n - \mathbb{E}\tilde{f}_n \rightarrow 0$; therefore $\tilde{\tau}_n(z) - \tilde{t}_n(z) \rightarrow 0$ for

$z \in \mathbb{C}^+$ as $N, n \rightarrow 0$. This, together with (A.16) yields:

$$\operatorname{Im}(\tilde{\tau}_n(z)) \geq \frac{\operatorname{Im}(z)}{8(\eta^2 + |z|^2)}$$

for n large enough. Taking into account the fact that $\operatorname{Im}(\frac{z}{n}\operatorname{Tr}(\mathbb{E}R_nQ_n)) \geq 0$, the first equation in (A.18) yields:

$$\begin{aligned} \operatorname{Im}(\mathbb{E}\tilde{f}_n(z)) &\geq |\tilde{\tau}_n|^2 \operatorname{Im}(z) + \frac{1}{n} \operatorname{Im}(\varepsilon_n(z)), \\ &\stackrel{(a)}{\geq} \frac{|\operatorname{Im}(z)|^3}{64(\eta^2 + |z|^2)^2} - \frac{1}{n^2} P_2(|z|) P_5(|\operatorname{Im}(z)|^{-1}), \\ &\geq \frac{|\operatorname{Im}(z)|^3}{64(\eta^2 + |z|^2)^2} \left(1 - \frac{1}{n^2} P_6(|z|) P_8(|\operatorname{Im}(z)|^{-1}) \right) \end{aligned}$$

for n large enough, where (a) follows from the previous estimate over $\tilde{\tau}_n$ and (A.12). Polynomials P_6 and P_8 being fixed as in the previous estimate, consider the set:

$$\mathcal{E}_n^1 = \left\{ z \in \mathbb{C}^+ ; 1 - \frac{1}{n^2} P_6(|z|) P_8(|\operatorname{Im}(z)|^{-1}) \geq \frac{1}{2} \right\}.$$

Then for n large enough and $z \in \mathcal{E}_n^1$,

$$\operatorname{Im}(\mathbb{E}\tilde{f}_n(z)) \geq \frac{|\operatorname{Im}(z)|^3}{128(\eta^2 + |z|^2)^2}.$$

From (A.18), we have:

$$\begin{aligned} \det(I_2 - D'(z)) &= |\tilde{\tau}_n|^2 \frac{\operatorname{Im}(z)}{\operatorname{Im}(\mathbb{E}\tilde{f}_n(z))} + |\tilde{\tau}_n|^2 \frac{\operatorname{Im}(\frac{z}{n}\tilde{\varepsilon}_n(z))}{\operatorname{Im}(\mathbb{E}\tilde{f}_n(z))} + \frac{\operatorname{Im}(\frac{1}{n}\varepsilon_n(z))}{\operatorname{Im}(\mathbb{E}\tilde{f}_n(z))} \\ &\geq |\tilde{\tau}_n|^2 |\operatorname{Im}(z)|^2 - |\tilde{\tau}_n|^2 \frac{|z\tilde{\varepsilon}_n(z)|}{n\operatorname{Im}(\mathbb{E}\tilde{f}_n(z))} - \frac{|\varepsilon_n(z)|}{n\operatorname{Im}(\mathbb{E}\tilde{f}_n(z))} \end{aligned}$$

Taking into account the various estimates previously established (recall that $|\tilde{\tau}_n| \leq |\operatorname{Im}(z)|^{-1}$), we obtain:

$$\begin{aligned} \det(I_2 - D'(z)) &\geq \frac{|\operatorname{Im}(z)|^4}{64(\eta^2 + |z|^2)^2} - \frac{1}{|\operatorname{Im}(z)|^2} \frac{|z\tilde{\varepsilon}_n(z)|}{n\operatorname{Im}(\mathbb{E}\tilde{f}_n(z))} - \frac{|\varepsilon_n(z)|}{n\operatorname{Im}(\mathbb{E}\tilde{f}_n(z))} \\ &\geq \frac{|\operatorname{Im}(z)|^4}{64(\eta^2 + |z|^2)^2} - \frac{128|z|(\eta^2 + |z|^2)^2 P_3(|z|) P_7(|\operatorname{Im}(z)|^{-1})}{n^2 |\operatorname{Im}(z)|^5} \\ &\quad - \frac{128(\eta^2 + |z|^2)^2 P_2(|z|) P_5(|\operatorname{Im}(z)|^{-1})}{n^2 |\operatorname{Im}(z)|^3} \\ &\geq \frac{|\operatorname{Im}(z)|^4}{64(\eta^2 + |z|^2)^2} \left(1 - \frac{1}{n^2} P_{12}(|z|) P_{16}(|\operatorname{Im}(z)|^{-1}) \right), \end{aligned}$$

valid for n large enough and $z \in \mathcal{E}_n^1$. Polynomials P_{12} and P_{16} being fixed as in the previous estimates, consider the set:

$$\mathcal{E}_n = \mathcal{E}_n^1 \cap \mathcal{E}_n^2 \quad \text{where} \quad \mathcal{E}_n^2 = \left\{ z \in \mathbb{C}^+ ; 1 - \frac{1}{n^2} P_{12}(|z|) P_{16}(|\operatorname{Im}(z)|^{-1}) \geq \frac{1}{2} \right\}.$$

Then for n large enough and $z \in \mathcal{E}_n$, we have:

$$\det(I_2 - D'(z)) \geq \frac{|\operatorname{Im}(z)|^4}{128(\eta^2 + |z|^2)^2}. \quad (\text{A.19})$$

We are now in position to lowerbound $\det(I_2 - D_0(z))$. Notice that

$$\begin{aligned} |\det(I_2 - D_0(z))| &= \left| 1 - \frac{z^2}{n} \tilde{\tau}_n \tilde{t}_n \operatorname{Tr} R_n S_n R_n T_n \right|, \\ &\geq 1 - \frac{|z|^2}{n} |\tilde{\tau}_n \tilde{t}_n| (\operatorname{Tr} R_n S_n R_n S_n^*)^{1/2} (\operatorname{Tr} R_n T_n R_n T_n^*)^{1/2}, \\ &\geq (\det(I_2 - D(z)))^{1/2} (\det(I_2 - D'(z)))^{1/2} \end{aligned}$$

by [29, Proposition 6.1]. Proof of Proposition A.1 is completed. \square

REFERENCES

- [1] S. Alberverio, L. Pastur, and M. Shcherbina. On the $1/n$ expansion for some unitary invariant ensembles of random matrices. *Comm. Math. Phys.*, 224(1):271–305, 2001. Dedicated to Joel L. Lebowitz.
- [2] G. W. Anderson. Convergence of the largest singular value of a polynomial in independent Wigner matrices. *Ann. Probab.*, 41(3B):2103–2181, 2013.
- [3] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.
- [4] G. W. Anderson and O. Zeitouni. A CLT for a band matrix model. *Probab. Theory Related Fields*, 134(2):283–338, 2006.
- [5] L. V. Arharov. Limit theorems for the characteristic roots of a sample covariance matrix. *Dokl. Akad. Nauk SSSR*, 199:994–997, 1971.
- [6] Z. D. Bai and J. W. Silverstein. CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.*, 32(1A):553–605, 2004.
- [7] Z.D. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer Series in Statistics. Springer, New York, second edition, 2010.
- [8] Z.D. Bai, X. Wang, and W. Zhou. Functional CLT for sample covariance matrices. *Bernoulli*, 16(4):1086–1113, 2010.
- [9] Z. Bao, G. M. Pan, and W. Zhou. On the MIMO channel capacity for the general channels. preprint, 2013.
- [10] F. Benaych-George, A. Guionnet, and C. Male. Central limit theorems for linear statistics of heavy tailed random matrices. *preprint*, 2013.
- [11] P. Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.
- [12] C. Bordenave. Personal communication, nov. 2013.
- [13] C. Bordenave. *A short course on random matrices (preliminary draft)*. available on <http://www.math.univ-toulouse.fr/~bordenave/coursRMT.pdf>, 2013.
- [14] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. OUP Oxford, 2013.
- [15] T. Cabanal-Duvillard. Fluctuations de la loi empirique de grandes matrices aléatoires. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(3):373–402, 2001.
- [16] M. Capitaine and C. Donati-Martin. Strong asymptotic freeness for Wigner and Wishart matrices. *Indiana Univ. Math. J.*, 56(2):767–803, 2007.
- [17] S. Chatterjee. Fluctuations of eigenvalues and second order Poincaré inequalities. *Probab. Theory Related Fields*, 143(1-2):1–40, 2009.
- [18] R. Brent Dozier and J. W. Silverstein. On the empirical distribution of eigenvalues of large dimensional information-plus-noise-type matrices. *J. Multivariate Anal.*, 98(4):678–694, 2007.
- [19] R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.
- [20] E. M. Dynkin. An operator calculus based on the Cauchy-Green formula. *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 30:33–39, 1972. Investigations on linear operators and the theory of functions, III.
- [21] V. L. Girko. *Theory of stochastic canonical equations. Vol. I*, volume 535 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2001.

- [22] V. L. Girko. *Theory of stochastic canonical equations. Vol. II*, volume 535 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2001.
- [23] O. Guedon, A. Lytova, A. Pajor, and L. A. Pastur. The central limit theorem for linear eigenvalue statistics of the sum of rank one projections on vectors with log-concave distribution. preprint, 2013.
- [24] A. Guionnet. Large deviations upper bounds and central limit theorems for non-commutative functionals of Gaussian large random matrices. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(3):341–384, 2002.
- [25] U. Haagerup and S. Thorbjørnsen. A new application of random matrices: $\text{Ext}(C_{\text{red}}^*(F_2))$ is not a group. *Ann. of Math. (2)*, 162(2):711–775, 2005.
- [26] U. Haagerup and S. Thorbjørnsen. Asymptotic expansions for the Gaussian unitary ensemble. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 15(1):1250003, 41, 2012.
- [27] W. Hachem, M. Kharouf, J. Najim, and J. W. Silverstein. A CLT for information-theoretic statistics of non-centered Gram random matrices. *Random Matrices Theory Appl.*, 1(2):1150010, 50, 2012.
- [28] W. Hachem, P. Loubaton, and J. Najim. Deterministic equivalents for certain functionals of large random matrices. *Ann. Appl. Probab.*, 17(3):875–930, 2007.
- [29] W. Hachem, P. Loubaton, J. Najim, and P. Vallet. On bilinear forms based on the resolvent of large random matrices. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(1):36–63, 2013.
- [30] B. Helffer. *Spectral theory and its applications*, volume 139 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2013.
- [31] B. Helffer and J. Sjöstrand. Équation de Schrödinger avec champ magnétique et équation de Harper. In *Schrödinger operators (Sønderborg, 1988)*, volume 345 of *Lecture Notes in Phys.*, pages 118–197. Springer, Berlin, 1989.
- [32] K. Johansson. On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.*, 91(1):151–204, 1998.
- [33] D. Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.*, 12(1):1–38, 1982.
- [34] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.
- [35] A. Kammoun, M. Kharouf, W. Hachem, and J. Najim. A central limit theorem for the SINR at the MMSE estimator output for large-dimensional signals. *Information Theory, IEEE Transactions on*, 55(11):5048–5063, nov. 2009.
- [36] A. M. Khorunzhy, B. A. Khoruzhenko, and L. A. Pastur. Asymptotic properties of large random matrices with independent entries. *J. Math. Phys.*, 37(10):5033–5060, 1996.
- [37] A. Lytova and L. Pastur. Central limit theorem for linear eigenvalue statistics of random matrices with independent entries. *Ann. Probab.*, 37(5):1778–1840, 2009.
- [38] V. A. Marcenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.
- [39] G. M. Pan and W. Zhou. Central limit theorem for signal-to-interference ratio of reduced rank linear receiver. *Ann. Appl. Probab.*, 18(3):1232–1270, 2008.
- [40] L. Pastur and M. Shcherbina. *Eigenvalue distribution of large random matrices*, volume 171 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2011.
- [41] H. Schultz. Non-commutative polynomials of independent Gaussian random matrices. The real and symplectic cases. *Probab. Theory Related Fields*, 131(2):261–309, 2005.
- [42] M. Shcherbina. Central limit theorem for linear eigenvalue statistics of the Wigner and sample covariance random matrices. *Zh. Mat. Fiz. Anal. Geom.*, 7(2):176–192, 197, 199, 2011.
- [43] Ya. Sinai and A. Soshnikov. Central limit theorem for traces of large random symmetric matrices with independent matrix elements. *Bol. Soc. Brasil. Mat. (N.S.)*, 29(1):1–24, 1998.
- [44] A. Soshnikov. The central limit theorem for local linear statistics in classical compact groups and related combinatorial identities. *Ann. Probab.*, 28(3):1353–1370, 2000.
- [45] P. Vallet, P. Loubaton, and X. Mestre. Improved subspace estimation for multivariate observations of high dimension: the deterministic signals case. *IEEE Trans. Inform. Theory*, 58(2):1043–1068, 2012.
- [46] J. Yao. *Estimation et fluctuations de fonctionnelles de grandes matrices aléatoires*. PhD thesis, Télécom ParisTech, 2013.

JAMAL NAJIM,
 Institut Gaspard Monge LabInfo, UMR 8049
 Université Paris Est Marne-la-Vallée
 5, Boulevard Descartes,

Champs sur Marne, 77454 Marne-la-Vallée Cedex 2, France
e-mail: najim@univ-mlv.fr

JIANFENG YAO,
Télécom Paristech,
46, rue Barrault,
75634 Paris cedex, France
e-mail: jianfeng.yao.sh@gmail.com