A CLT for linear spectral statistics of large random information-plus-noise matrices

Marwa Banna, Jamal Najim, Jianfeng Yao

A CLT FOR LINEAR SPECTRAL STATISTICS OF LARGE RANDOM INFORMATION-PLUS-NOISE MATRICES

MARWA BANNA, JAMAL NAJIM, AND JIANFENG YAO

Abstract. Consider a matrix $\mathrm{Y}_n = \frac{\sigma}{\sqrt{n}}\mathrm{X}_n + \mathrm{A}_n$, where $\sigma > 0$ and $\mathrm{X}_n = (x_{ij}^n)$ is a $N \times n$ random matrix with i.i.d. real or complex standardized entries and A_n is a $N \times n$ deterministic matrix with bounded spectral norm. The fluctuations of the linear spectral statistics of the eigenvalues:

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

are shown to be Gaussian, in the case where f is a smooth function of class C^3 with bounded support, and in the regime where both dimensions of matrix Y_n go to infinity at the same pace.

The CLT is expressed in terms of vanishing Lévy-Prohorov distance between the linear statistics' distribution and a centered Gaussian probability distribution, the variance of which depends upon N and n and may not converge. The proof combines ideas from Bai and Silverstein [3], Hachem et al. [18] and Najim and Yao [32].

1. Introduction

The model. Consider a $N \times n$ random matrix $\mathrm{Y}_n = (y_{ij}^n)$ given by:

$$\mathrm{Y}_n = \frac{\sigma}{\sqrt{n}}\mathrm{X}_n + \mathrm{A}_n \ ,$$

where $\sigma > 0$ and X_n is a $N \times n$ matrix whose entries $(x_{ij}^n\ ; i, j, n)$ are real or complex, independent and identically distributed (i.i.d.) with mean 0 and variance 1. Matrix A_n has the same dimensions and is deterministic. Matrix Y_n is sometimes coined as "Information-plus-noise" type matrix in the literature.

The purpose of this article is to study the fluctuations of linear spectral statistics of the form:

$$\mathrm{Tr}\, f(\mathrm{Y}_n\mathrm{Y}_n^*) = \sum_{i=1}^{N} f(\lambda_i) \ , \tag{1.1}$$

where $\mathrm{Tr}\,(\mathbf{M})$ refers to the trace of $\mathbf{M}$, the λ_i's are the eigenvalues of $\mathrm{Y}_n\mathrm{Y}_n^*$, and f is a smooth function, under the regime where the dimensions n and $N = N(n)$ go to infinity at the same pace:

$$N, n \to \infty \quad \text{and} \quad 0 \ < \ \liminf \frac{N}{n} \ \le \ \limsup \frac{N}{n} \ < \ \infty \ . \tag{1.2}$$

This condition will simply be referred to as $N, n \to \infty$ in the sequel.

Large information-plus-noise matrices, and more generally large non-centered random matrices, have recently attracted a lot of attention. Under mild conditions over the moments of the entries of X_n and the spectral norm of matrix A_n the asymptotic behavior of the empirical distribution of $\mathrm{Y}_n\mathrm{Y}_n^*$'s eigenvalues (also called spectral distribution of $\mathrm{Y}_n\mathrm{Y}_n^*$) defined as:

$$F^{\mathrm{Y}_n\mathrm{Y}_n^*}(B) = \frac{\#\{i,\ \lambda_i \in B\}}{N} \qquad \text{for } B \text{ a Borel set in } \mathbb{R} \ , \tag{1.3}$$

has been studied by Girko [16, chapter 7], Dozier and Silverstein [13], Hachem et al. [21], etc. Following these results, various properties of the asymptotic spectrum were studied, see for instance [12, 29, 1, 8].

From an applied point of view, information-plus-noise matrices are versatile models in many contexts, from Rice channels in wireless communication to noisy data and small rank perturbations [31, 15, 19, 20]. From a theoretical standpoint, hermitian non-centered models of the type

$$\left(\frac{\sigma}{\sqrt{n}}\mathrm{X}_n - z\mathrm{I}_N\right)\left(\frac{\sigma}{\sqrt{n}}\mathrm{X}_n - z\mathrm{I}_N\right)^*$$

Date: April 24, 2019.

2010 *Mathematics Subject Classification.* Primary 15B52, Secondary 15A18, 60F15.

Key words and phrases. Large random matrices, linear statistics of the eigenvalues, central limit theorem.

MB is supported by the ERC Advanced Grant NCDFP 339760, held by Roland Speicher, and partially by the French Agence Nationale de la Recherche under the grant ANR-12-MONU-0003.

JN is supported by French ANR grant ANR-12-MONU-0003 and Labex Bézout.

JY is supported by program "Futur et ruptures" of Fondation Télécom.

are a key device to understand the spectrum of large $N \times N$ non-hermitian matrices $\frac{\sigma}{\sqrt{n}} \mathrm{X}_n$ via Girko's hermitization trick.

While fluctuations of functionals of large random covariance matrices have attracted a lot of attention, see for instance [26, 25, 7, 17, 3, 22, 18, 33, 5, 11, 30, 34, 32] and the references therein, there seems to be very few results (in fact one to the authors' knowledge) for large information-plus-noise type matrices. In the specific case of a non-centered matrix with a separable variance profile, i.e. $\Sigma_n = \frac{1}{\sqrt{n}} \mathrm{D}_n^{1/2} \mathrm{X}_n \widetilde{\mathrm{D}}_n^{1/2} + \mathrm{A}_n$, with $\mathrm{D}_n, \widetilde{\mathrm{D}}_n$ deterministic diagonal matrices, the fluctuations have been described for the specific functional (known as the mutual information in wireless communications)

$$\log\det\left(\mathrm{I}_n + \Sigma_n \Sigma_n^*\right) = \sum_{i=1}^{N} \log\left(1 + \lambda_i(\Sigma_n \Sigma_n^*)\right) \ , \tag{1.4}$$

first at a physical level of rigor by Moustakas et al. [31] for complex Gaussian entries, then for general entries by Hachem et al. in [18]. This shortage of results might be related to the fact that the addition of a deterministic component A_n to a large random matrix $\frac{\sigma}{\sqrt{n}} \mathrm{X}_n$ substantially increases the complexity of the computations needed to establish the CLT. Equation (3.20) in [18] illustrates this fact.

Fluctuations and representation of linear spectral statistics. We now present the main object of interest:

$$L_n(f) = \sum_{i=1}^{N} f(\lambda_i) - \mathbb{E} \sum_{i=1}^{N} f(\lambda_i)$$

as $N, n \to \infty$. In the case where f is a function of class C^{k+1} with compact support, denote by $\Phi_k(f) : \mathbb{C}^+ \to \mathbb{C}$ its so-called almost analytic extension defined as $\Phi_k(f)(x + \mathbf{i}y) = \sum_{\ell=0}^{k} \frac{(\mathbf{i}y)^\ell}{\ell!} f^{(\ell)}(x)\chi(y)$, where $\chi : \mathbb{R} \to \mathbb{R}^+$ is a smooth, compactly supported function with value 1 in the neighbourhood of 0 and $\bar{\partial} = \frac{1}{2}(\partial_x + \mathbf{i}\partial_y)$. Helffer-Sjöstrand's formula yields that:

$$\sum_{i=1}^{N} f(\lambda_i) - \sum_{i=1}^{N} \mathbb{E} f(\lambda_i) = \frac{2}{\pi} \mathrm{Re} \int_{\mathbb{C}^+} \bar{\partial} \Phi_k(f)(z) \left\{ \mathrm{Tr}\, \mathrm{Q}_n(z) - \mathrm{Tr}\, \mathbb{E}\mathrm{Q}_n(z) \right\} \ell(dz) \tag{1.5}$$

where $\mathrm{Q}_n(z) = (\mathrm{Y}_n \mathrm{Y}_n^* - z I_N)^{-1}$ stands for the resolvent of $\mathrm{Y}_n \mathrm{Y}_n^*$ and $\ell(dz) = dx\, dy$ for the Lebesgue measure over $\mathbb{C}^+$. It is clear from (1.5) that in order to describe the fluctuations of $L_n(f)$, a natural approach is to study the fluctuations of the process $(\mathrm{Tr}(\mathrm{Q}_n(z) - \mathbb{E}\mathrm{Q}_n(z))\ ; z \in \Gamma)$ where $\Gamma \subset \mathbb{C}$ is some given compact set. In order to proceed, we define

$$M_n(z) := \mathrm{Tr}\, \mathrm{Q}_n(z) - \mathbb{E}\mathrm{Tr}\, \mathrm{Q}_n(z) \tag{1.6}$$

and handle this term by martingale techniques, a strategy successfully applied in [3, 33, 22, 27, 18, 32, 5].

Entries with non-null fourth cumulant and a family of Gaussian random variables. It is well known since the paper by Khorunzhiy et al. [28] that if the fourth moment of the entries differs from its Gaussian counterpart, then other terms may appear in the variance of the trace of the resolvent, one being proportional to the fourth cumulant κ of the entries:

$$\kappa = \mathbb{E}\,|x_{11}^n|^4 - |\vartheta|^2 - 2 \quad \text{where} \quad \vartheta = \mathbb{E}(x_{11}^n)^2 \ . \tag{1.7}$$

The same phenomenon will occur here but the convergence of these additional terms may fail to happen under usual assumptions such as the convergence of the spectral distribution $F^{\mathrm{A}_n \mathrm{A}_n^*}$ of matrix $\mathrm{A}_n \mathrm{A}_n^*$ to a probability measure as $N, n \to \infty$.

As we shall see later, the reason for this lack of convergence lies in the fact that these additional terms not only depend on the spectrum of $\mathrm{A}_n \mathrm{A}_n^*$, but also on the spectrum of $\mathrm{A}_n \mathrm{A}_n^{\mathrm{T}}$ and on $\mathrm{A}_n \mathrm{A}_n^*$'s eigenvectors. In order to avoid cumbersome assumptions enforcing the joint convergence of these quantities, we shall express our fluctuation results in the same way as in [32] and prove that the distribution of the linear statistics $L_n(f)$ becomes close to a family of centered Gaussian distributions, whose variance might not converge. Namely, we shall establish that there exists a Gaussian random variable $\mathcal{N}(0, \Theta_n(f))$ such that:

$$d_{\mathcal{L}P}\left(L_n(f), \mathcal{N}(0, \Theta_n(f))\right) \xrightarrow[N,n\to\infty]{} 0 \ , \tag{1.8}$$

where $d_{\mathcal{L}P}$ denotes the Lévy-Prohorov distance (and in particular metrizes the convergence of laws).

A simple expression for the variance (for real A_n and real or circular $\mathbf{x}_{ij}$'s). We first introduce some key quantities whose properties will be recalled and studied in Section 2. The following equations admit a unique solution $(\delta_n, \tilde{\delta}_n)$ in the class of Stieltjes transforms of nonnegative measures with supports $\mathcal{S}_n$ and $\tilde{\mathcal{S}}_n$ in $\mathbb{R}^+$ (see for instance [21, 18], see also [16, Section 7.11]).

$$\left\{\begin{array}{lcl} \delta_n(z) & = & \frac{\sigma}{n}\mathrm{Tr}\left(-z(1+\sigma\tilde{\delta}_n(z))\mathrm{I}_N + \frac{\mathrm{A}_n\mathrm{A}_n^*}{1+\sigma\delta_n(z)}\right)^{-1} \\ \tilde{\delta}_n(z) & = & \frac{\sigma}{n}\mathrm{Tr}\left(-z(1+\sigma\delta_n(z))\mathrm{I}_n + \frac{\mathrm{A}_n^*\mathrm{A}_n}{1+\sigma\tilde{\delta}_n(z)}\right)^{-1} \end{array}\right. , \quad z\in\mathbb{C}^+ . \tag{1.9}$$

Associated to δ_n and $\tilde{\delta}_n$ are the $N\times N$ and $n\times n$ matrices:

$$\left\{\begin{array}{l} \mathrm{T}_n(z) := \left(-z(1+\sigma\tilde{\delta}_n(z))\mathrm{I}_N + \frac{\mathrm{A}_n\mathrm{A}_n^*}{1+\sigma\delta_n(z)}\right)^{-1} = [t_{ij}(z)] \\ \tilde{\mathrm{T}}_n(z) := \left(-z(1+\sigma\delta_n(z))\mathrm{I}_n + \frac{\mathrm{A}_n^*\mathrm{A}_n}{1+\sigma\tilde{\delta}_n(z)}\right)^{-1} = [\tilde{t}_{ij}(z)] \end{array}\right. , \tag{1.10}$$

and the quantity

$$s_n(z) := z(1+\sigma\delta_n(z))(1+\sigma\tilde{\delta}_n(z)) . \tag{1.11}$$

With these quantities at hand, the variance $\Theta_n(f)$ which appears in (1.8) takes a remarkably simple form, to be compared with [3, Eq. (1.17)] and [32, Eq. (4.7)], if matrix A_n is real and the $\mathbf{x}_{ij}$'s are real ($\vartheta = 1$) or circular[1] ($\vartheta = 0$).

$$\begin{aligned} \Theta_n(f) \quad = \quad & \frac{1+\vartheta}{2\pi^2}\int_{\mathcal{S}_n^2} f'(x)f'(y)\ln\left|\frac{s_n(x)-\overline{s_n(y)}}{s_n(x)-s_n(y)}\right| dxdy \\ & + \frac{\sigma^4\kappa}{\pi^2 n^2}\sum_{i=1}^N\sum_{j=1}^n\left(\int_{\mathcal{S}_n} f'(x)\mathrm{Im}\left(x\, t_{ii}(x)\tilde{t}_{jj}(x)\right)\, dx\right)^2 . \end{aligned} \tag{1.12}$$

The quantities $s_n(x), t_{ii}(x), \tilde{t}_{jj}(x)$ are the limits of the corresponding quantities $s_n, t_{ii}, \tilde{t}_{jj}$, evaluated at $z\in\mathbb{C}^+$, as $z\to x\in\mathbb{R}$. In the case where matrix A_n is not real or $\vartheta\notin\{0,1\}$, then the term proportional to ϑ above is substantially more complicated.

While the heart of the computations needed to establish the CLT is a (substantial) variation of those performed in [18], the identification of the variance is an important contribution of this article. In particular one may notice that the quantity s_n defined in (1.11) is central to express the variance in (1.12) while it does not appear in the formula of the variance of the mutual information (1.4).

Organization of the paper. The main results of the paper are introduced in Section 2. Central Limit Theorems are stated in Theorem 1 for the trace of the resolvent and in Theorem 2 for general linear statistics. Simplified expressions for the variance are provided in Theorem 3. Sections 3 and 4 are devoted to the proofs. Useful estimates are recalled in Appendix A.

2. Statement of the Central Limit Theorem

2.1. **Notations and assumptions.** Throughout the paper, $\boldsymbol{i} = \sqrt{-1}$, $\mathbb{R}^+ = \{x\in\mathbb{R} : x\geq 0\}$ and $\mathbb{C}^+ = \{z\in\mathbb{C} : \mathrm{Im}\, z > 0\}$. Denote by $\xrightarrow{a.s.}$ (resp. $\xrightarrow{\mathcal{P}}$, $\xrightarrow{\mathcal{D}}$) the almost sure convergence (resp. in probability, in distribution). Denote by $\mathrm{diag}(a_i; 1\leq i\leq k)$ the $k\times k$ diagonal matrix whose diagonal entries are the a_i's. Element (i,j) of the matrix M will be denoted by m_{ij} or $[\mathrm{M}]_{ij}$.

For a matrix M, denote by M^T its transpose, M^* its Hermitian adjoint, $\overline{\mathrm{M}}$ its entry-wise conjugate, $\det(\mathrm{M})$ its determinant and $\mathrm{vdiag}(\mathrm{M})$ the vector whose entries are the diagonal elements (m_{ii}). When dealing with vectors and matrices, $\|.\|$ refers to the Euclidean and the spectral norm respectively.

We shall denote by K a generic constant that does not depend on N, n but whose value may change from line to line. Function $\mathbf{1}_A$ denotes the indicator function of the set A.

Notations $u_n = \mathcal{O}(v_n)$ and $u_n = o(v_n)$ stand for the usual big $\mathcal{O}$ and little o notations when $N, n\to\infty$. We might also use $\mathcal{O}_z$ or $\mathcal{O}_\varepsilon$ to underline the dependence of the constant in $\mathcal{O}$ on z or ε. If X_n and Y_n are sequences of random variables, $X_n = o_P(Y_n)$ stands for the fact that there exists a sequence Z_n such as $X_n = Z_nY_n$ and Z_n converges to zero in probability.

Denote by $d_{\mathcal{LP}}(P,Q)$ the Lévy-Prohorov distance between two probability measures P, Q defined as:

$$d_{\mathcal{LP}}(P,Q) = \inf\{\varepsilon > 0 : P(A)\leq Q(A^\varepsilon)+\varepsilon \quad \text{for all Borel sets } A\subset\mathbb{R}^d\},$$

where A^ε is an ε-blow up of A (see [6, Chapter 1, Section 6] for more details). If X and Y are random variables with distributions $\mathcal{L}(X)$ and $\mathcal{L}(Y)$, we simply write (with a slight abuse of notations) $d_{\mathcal{LP}}(X,Y)$

[1] By circular, we mean that $\mathbf{x}_{ij}$ has decorrelated real and imaginary part, each with the same variance 1/2, i.e. $\mathbb{E}|\mathbf{x}_{ij}|^2 = 1$ and $\vartheta = \mathbb{E}(\mathbf{x}_{ij})^2 = 0$.

instead of $d_{\mathcal{L}P}(\mathcal{L}(X), \mathcal{L}(Y))$. It is well-known that the Lévy-Prohorov distance metrizes the convergence in distribution (see for instance [14, Chapter 11]).

The set $C_c^k(\mathbb{R})$ denotes the class of functions with k continuous derivatives and compact support.

We now state the main assumptions of the article. Recall the fact that $N = N(n)$ and the asymptotic regime (1.2) where $N, n \to \infty$ and denote by

$$c_n := \frac{N}{n}\ , \quad c_- := \liminf \frac{N}{n} \quad \text{and} \quad c_+ := \limsup \frac{N}{n}\ .$$

Assumption 1. *The random variables* $(x_{ij}^n\ ;\ 1 \le i \le N(n), 1 \le j \le n, n \ge 1)$ *are real or complex, independent and identically distributed (i.i.d.). They satisfy*

$$\mathbb{E}x_{ij}^n = 0\ , \qquad \mathbb{E}|x_{ij}^n|^2 = 1 \quad \text{and} \quad \mathbb{E}|x_{ij}^n|^{16} < \infty\ .$$

Remark 2.1. *The 16th moment assumption above could be relaxed to an optimal 4th moment assumption as in* [3, 32]*, with extra work involving the improvement of some estimates from* [23]*. We do not pursue in this direction here.*

Associated to these moments are the quantities introduced in (1.7). We mention two important special cases: The case where $\vartheta = 1$ corresponding to real x_{ij}'s and the case where $\vartheta = 0$, corresponding to complex x_{ij}'s with decorrelated real and imaginary part of equal variance.

Assumption 2. *The family of deterministic* $N \times n$ *complex matrices* (A_n) *is bounded for the spectral norm:*

$$a_{\max} \;=\; \sup_{n\ge 1} \|\mathrm{A}_n\| < \infty\,.$$

2.2. **Resolvent, canonical equations and deterministic equivalents.** Denote by $\mathrm{Q}_n(z)$ and $\tilde{\mathrm{Q}}_n(z)$ the resolvents of matrices $\mathrm{Y}_n\mathrm{Y}_n^*$ and $\mathrm{Y}_n^*\mathrm{Y}_n$:

$$\mathrm{Q}_n(z) = \left(\mathrm{Y}_n\mathrm{Y}_n^* - z\mathrm{I}_N\right)^{-1}\ , \qquad \tilde{\mathrm{Q}}_n(z) = \left(\mathrm{Y}_n^*\mathrm{Y}_n - z\mathrm{I}_n\right)^{-1}\ . \tag{2.1}$$

Their normalized traces $\frac{1}{N}\mathrm{Tr}\, Q_n(z)$ and $\frac{1}{n}\mathrm{Tr}\, \tilde{Q}_n(z)$ are respectively the Stieltjes transforms of the empirical distribution of $\mathrm{Y}_n\mathrm{Y}_n^*$'s eigenvalues and of $\mathrm{Y}_n^*\mathrm{Y}_n$'s eigenvalues.

Recall the definition of the Stieltjes transforms δ_n and $\tilde{\delta}_n$ as solutions of the canonical equations (1.9) and those of matrices T_n and $\tilde{\mathrm{T}}_n$:

$$\left\{\begin{array}{lllll} \delta_n(z) &=& \frac{\sigma}{n}\mathrm{Tr}\left(-z(1+\sigma\tilde{\delta}_n(z))\mathrm{I}_N + \frac{\mathrm{A}_n\mathrm{A}_n^*}{1+\sigma\delta_n(z)}\right)^{-1} &=& \frac{\sigma}{n}\mathrm{Tr}\,\mathrm{T}_n(z) \\ \tilde{\delta}_n(z) &=& \frac{\sigma}{n}\mathrm{Tr}\left(-z(1+\sigma\delta_n(z))\mathrm{I}_n + \frac{\mathrm{A}_n^*\mathrm{A}_n}{1+\sigma\tilde{\delta}_n(z)}\right)^{-1} &=& \frac{\sigma}{n}\mathrm{Tr}\,\tilde{\mathrm{T}}_n(z) \end{array}\right. \ , \quad z \in \mathbb{C}^+\,.$$

The measures associated to δ_n and $\tilde{\delta}_n$ have respective total masses given by

$$\lim_{y\to\infty}\ -\boldsymbol{i}y\,\delta_n(\boldsymbol{i}y) = \frac{N}{n}\sigma \quad \text{and} \quad \lim_{y\to\infty}\ -\boldsymbol{i}y\,\tilde{\delta}_n(\boldsymbol{i}y) = \sigma\ .$$

Matrix $\mathrm{T}_n(z)$ defined in (1.10) is a deterministic equivalent of the resolvent Q_n in the sense that for $z \in \mathbb{C}\backslash\mathbb{R}^+$:

$$\frac{1}{N}\mathrm{Tr}(\mathrm{Q}_n(z) - \mathrm{T}_n(z)) \xrightarrow[N,n\to\infty]{a.s.} 0 \quad \text{and} \quad u_n^*\mathrm{Q}_n v_n - u_n^*\mathrm{T}_n v_n \xrightarrow[N,n\to\infty]{a.s.} 0\ ,$$

where u_n and v_n are deterministic $N \times 1$ vectors with uniformly bounded euclidian norms (in n), see for instance [13, 23]. A symmetric result holds for $\tilde{\mathrm{Q}}_n$ and $\tilde{\mathrm{T}}_n$.

We will often drop the subscript n and when dealing with δ_n, $\tilde{\delta}_n$, T_n and $\tilde{\mathrm{T}}_n$, we may emphasize the z-dependence by writing δ_z, $\tilde{\delta}_z$, T_z and $\tilde{\mathrm{T}}_z$ instead.

Remark 2.2. *In the sequel, we will handle* T_z, $\overline{\mathrm{T}}_z$, $\mathrm{T}_z^{\mathsf{T}}$ *and* T_z^*. *Beware that these quantities are a priori different. Definition* (1.10) *yields*

$$\mathrm{T}_z^{\mathsf{T}} \;=\; \left(-z(1+\sigma\tilde{\delta}_z)\mathrm{I}_N + \frac{\bar{\mathrm{A}}\bar{\mathrm{A}}^*}{1+\sigma\delta_z}\right)^{-1}$$

hence the identities $\mathrm{T}_z^* = \mathrm{T}_{\bar{z}}$ *and* $\overline{\mathrm{T}}_z = \mathrm{T}_{\bar{z}}^{\mathsf{T}}$.

2.3. **Expression of the variance and statement of the main results.** In order to express the variance, we need to introduce a number of auxiliary quantities. Let

$$\nu(z_1,z_2)=\frac{\sigma^2}{n}\frac{\mathrm{Tr}\,\mathrm{T}_{z_1}\mathrm{A}\mathrm{A}^*\mathrm{T}_{z_2}}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})}\,,\quad \gamma(z_1,z_2)=\frac{\sigma^2}{n}\mathrm{Tr}\,\mathrm{T}_{z_1}\mathrm{T}_{z_2}\,,\quad \tilde\gamma(z_1,z_2)=\frac{\sigma^2}{n}\mathrm{Tr}\,\tilde{\mathrm{T}}_{z_1}\tilde{\mathrm{T}}_{z_2}\,. \tag{2.2}$$

The following quantity will be instrumental in the sequel.

$$\Delta_n(z_1,z_2)=(1-\nu)^2-z_1z_2\,\gamma\,\tilde\gamma\,. \tag{2.3}$$

Consider now

$$\begin{aligned}\gamma^\dagger(z_1,z_2)&=\frac{\sigma^2}{n}\mathrm{Tr}\,\mathrm{T}_{z_1}\mathrm{T}_{z_2}^{\mathsf T}\,, & \nu^\dagger(z_1,z_2)&=\frac{\sigma^2}{n}\frac{\mathrm{Tr}\,\mathrm{T}_{z_2}^{\mathsf T}\bar{\mathrm{A}}\mathrm{A}^*\mathrm{T}_{z_1}}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})}\,,\\ \tilde\gamma^\dagger(z_1,z_2)&=\frac{\sigma^2}{n}\mathrm{Tr}\,\tilde{\mathrm{T}}_{z_1}\tilde{\mathrm{T}}_{z_2}^{\mathsf T}\,, & \tilde\nu^\dagger(z_1,z_2)&=\frac{\sigma^2}{n}\frac{\mathrm{Tr}\mathrm{T}_{z_2}\mathrm{A}\mathrm{A}^{\mathsf T}\mathrm{T}_{z_1}^{\mathsf T}}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})}\,.\end{aligned} \tag{2.4}$$

and the following counterpart to Δ_n

$$\Delta_n^\vartheta(z_1,z_2)=\left(1-\vartheta\nu^\dagger\right)\left(1-\bar\vartheta\tilde\nu^\dagger\right)-|\vartheta|^2\,z_1z_2\,\gamma^\dagger\tilde\gamma^\dagger\,. \tag{2.5}$$

Proposition 2.3 (Properties of s_n, Δ_n and Δ_n^ϑ)**.** *The following properties hold:*

1. *Function $s_n:\mathbb{C}^+\to\mathbb{C}^+$ is analytic and if $z_1,z_2\in\mathbb{C}^+$ and $z_1\neq z_2$ then $s_n(z_1)\neq s_n(z_2)$.*
2. *Function $\Delta_n:\mathbb{C}^+\times\mathbb{C}^+\to\mathbb{C}$ never vanishes and the following identity holds:*

$$\Delta_n(z_1,z_2)=\frac{z_1-z_2}{s_n(z_1)-s_n(z_2)}\,. \tag{2.6}$$

3. *Function $\Delta_n^\vartheta:\mathbb{C}^+\times\mathbb{C}^+\to\mathbb{C}$ never vanishes.*

Proposition 2.3 follows from Proposition 3.1 below.

Remark 2.4. *Formula (2.6) is a non-trivial representation of Δ_n which seems to first appear here (to the authors' knowledge). It plays a key role in this paper (a) at a technical level to establish stability conditions when proving the CLT (see proof of Lemma 3.10-(ii)) and (b) to obtain important formulas for the variance (see (1.12) and (2.7)). The absence of such a representation for Δ_n^ϑ prevents us from obtaining a good representation formula for $\Theta_{2,n}(\kappa,f,g)$ (see Theorem 3 and Remark 2.7).*

Remark 2.5 (Simplifications)**.** *Simplifications may occur depending on the values of* A *and ϑ:*

1. *If matrix* A *has real entries, then $\mathrm{T}_z^{\mathsf T}=\mathrm{T}_z$, $\gamma=\gamma^\dagger$, $\tilde\gamma=\tilde\gamma^\dagger$. Moreover*

$$\nu^\dagger(z_1,z_2)=\tilde\nu^\dagger(z_1,z_2)=\frac{\sigma^2}{n}\frac{\mathrm{Tr}\,\mathrm{T}_{z_2}\mathrm{A}\mathrm{A}^{\mathsf T}\mathrm{T}_{z_1}}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})}\,.$$

2. *In the case where $\vartheta=1$ (real entries (x_{ij})) and A has real entries then $\Delta_n=\Delta_n^\vartheta$.*
3. *In the case where $\vartheta=0$ then $\Delta_n^\vartheta=1$.*

We are now in position to introduce the covariance function. Denote by Θ_n the quantity:

$$\Theta_n(z_1,z_2)\;:=\;\Theta_{0,n}(z_1,z_2)+\Theta_{1,n}(\vartheta,z_1,z_2)+\Theta_{2,n}(\kappa,z_1,z_2)\,,\qquad z_1,z_2\in\mathbb{C}^+ \tag{2.7}$$

where

$$\Theta_{0,n}(z_1,z_2)\;:=\;-\frac{\partial}{\partial z_2}\left(\frac{1}{\Delta_n}\frac{\partial\Delta_n}{\partial z_1}\right)\;=\;\frac{s_n'(z_1)s_n'(z_2)}{(s_n(z_1)-s_n(z_2))^2}-\frac{1}{(z_1-z_2)^2}\,, \tag{2.8}$$

$$\Theta_{1,n}(\vartheta,z_1,z_2)\;:=\;-\frac{\partial}{\partial z_2}\left(\frac{1}{\Delta_n^\vartheta}\frac{\partial\Delta_n^\vartheta}{\partial z_1}\right)\,, \tag{2.9}$$

$$\Theta_{2,n}(\kappa,z_1,z_2)\;:=\;\kappa\frac{\partial^2}{\partial z_1\partial z_2}\left\{\frac{\sigma^4z_1z_2}{n^2}\sum_{i=1}^N t_{ii}(z_1)t_{ii}(z_2)\sum_{j=1}^n\tilde t_{jj}(z_1)\tilde t_{jj}(z_2)\right\}\,. \tag{2.10}$$

Consider the following subsets of $\mathbb{C}$, with $A>0$

$$\begin{aligned}D&=[0,A]+\boldsymbol{i}[0,1]\,, & D^\pm&=\{z\in D^+\}\cup\{\bar z\in D^+\}\,,\\ D^+&=[0,A]+\boldsymbol{i}(0,1]\,, & D_\varepsilon&=[0,A]+\boldsymbol{i}[\varepsilon,1]\,,\quad(\varepsilon>0)\,.\end{aligned} \tag{2.11}$$

We first study the Gaussian fluctuations for the trace of the resolvent.

Theorem 1 (CLT for the trace of the resolvent)**.** *Recall the definition of M_n in (1.6) and let Assumptions 1 and 2 hold. Then for every $\varepsilon>0$,* *(1) There exists $z_0 \in \mathbb{C}^+$ such that*

$$\sup_{n\geq 1} \mathbb{E}|M_n(z_0)|^2 < \infty \qquad \text{and} \qquad \sup_{z_1,z_2\in D_\varepsilon, n\geq 1} \frac{\mathbb{E}|M_n(z_1)-M_n(z_2)|^2}{|z_1-z_2|^2} < \infty\,.$$

In particular, the process $(M_n(z), z\in D_\varepsilon)$ is tight.

(2) There exists a sequence $(G_n(z), z\in D^\pm)$ of centered Gaussian processes such that for any $z_1, z_2 \in D^\pm$:

$$\mathrm{cov}(G_n(z_1), G_n(z_2)) = \Theta_n(z_1,z_2) \qquad \text{and} \qquad \mathrm{cov}(G_n(z_1), \overline{G_n(z_2)}) = \mathrm{cov}(G_n(z_1), G_n(\overline{z_2}))\ ,$$

where Θ_n is defined in (2.7). Moreover, $(G_n(z), z\in D_\varepsilon)$ is tight.

(3) For any continuous and bounded functional F from $C(D_\varepsilon;\mathbb{C})$ to $\mathbb{C}$, $\mathbb{E}F(M_n) - \mathbb{E}F(G_n) \xrightarrow[N,n\to\infty]{} 0$.

Theorem 1 is an extension of Bai and Silverstein's master lemma [3, Lemma 1.1] to the non-centered case. The proof[2] is postponed to Section 3.

Having the CLT for the trace of the resolvent at hand, we can now extend it to non-analytic functions via Helffer-Sjöstrand's formula (1.5).

Theorem 2 (CLT for general linear statistics)**.** *Let Assumptions 1 and 2 hold. Let $f_1,\dots,f_k \in C^3_c(\mathbb{R})$ and let $L_n(\mathbf{f}) = (L_n(f_1),\dots,L_n(f_k))$ with*

$$L_n(f) = \mathrm{Tr}\, f(\mathrm{YY}^*) - \mathbb{E}\mathrm{Tr}\, f(\mathrm{YY}^*)\,, \quad f\in\{f_1,\cdots,f_k\}\,.$$

Then there exists an $\mathbb{R}^k$-valued sequence of centered Gaussian vectors

$$Z_n(\mathbf{f}) = (Z_n(f_1),\dots,Z_n(f_k)) \tag{2.12}$$

with covariance given by

$$\begin{aligned}\mathrm{Cov}\,(Z_n(f), Z_n(g)) = &\frac{2}{\pi^2}\mathrm{Re}\int_{(\mathbb{C}^+)^2} \bar\partial\Phi_2(f)(z_1)\overline{\bar\partial\Phi_2(g)(z_2)}\Theta_n(z_1,\bar z_2)\ell(dz_1)\ell(dz_2)\\ &+\frac{2}{\pi^2}\mathrm{Re}\int_{(\mathbb{C}^+)^2} \bar\partial\Phi_2(f)(z_2)\bar\partial\Phi_2(g)(z_1)\Theta_n(z_1,z_2)\ell(dz_1)\ell(dz_2)\,,\end{aligned}$$

for $f,g\in\{f_1,\dots,f_k\}$. Moreover, the sequence $(Z_n(\mathbf{f}), n\geq 1)$ is tight and

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

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

$$\mathbb{E}F\big(L_n(\mathbf{f})\big) - \mathbb{E}F\big(Z_n(\mathbf{f})\big) \xrightarrow[N,n\to\infty]{} 0\,.$$

Proof of Theorem 2 is skipped as it closely follows the proof of [32, Theorem 2]. It relies on [32, Lemma 6.3] and on the following estimates of the variance

$$\mathrm{Var}\,\mathrm{Tr}\,\mathrm{Q}_n(z) = \mathcal{O}\left(|\mathrm{Im} z|^{-4}\right) \quad \text{and} \quad \mathrm{Var}\,\mathrm{Tr}\, G_n(z) = \mathcal{O}\left(|\mathrm{Im} z|^{-4}\right)\,.$$

which is a variation of [32, Proposition 6.4].

Due to the decomposition of $\Theta_n(z_1,z_2)$ in (2.7), the covariance $\mathrm{Cov}\,(Z_n(f),Z_n(g))$ can be split into three terms

$$\mathrm{Cov}\,(Z_n(f),Z_n(g)) := \Theta_{0,n}(f,g) + \Theta_{1,n}(\vartheta,f,g) + \Theta_{2,n}(\kappa,f,g)$$

where (we drop the dependence in ϑ,κ), for $i=0,1,2$,

$$\begin{aligned}\Theta_{i,n}(f,g) = &\frac{2}{\pi^2}\mathrm{Re}\int_{(\mathbb{C}^+)^2} \bar\partial\Phi_2(f)(z_1)\overline{\bar\partial\Phi_2(g)(z_2)}\Theta_{i,n}(z_1,\bar z_2)\ell(dz_1)\ell(dz_2)\\ &+\frac{2}{\pi^2}\mathrm{Re}\int_{(\mathbb{C}^+)^2} \bar\partial\Phi_2(f)(z_2)\bar\partial\Phi_2(g)(z_1)\Theta_{i,n}(z_1,z_2)\ell(dz_1)\ell(dz_2)\,.\end{aligned}$$

In order to provide simplified formulas, we evaluate various quantities defined on $\mathbb{C}^+$ along the real axis.

Proposition 2.6 (cf. Theorem 2.1 in [12])**.** *Let $x\in\mathbb{R}\setminus\{0\}$, then the following limits exist*

$$s_n(x) := \lim_{\varepsilon\downarrow 0} s_n(x+\boldsymbol{i}\varepsilon)\ ,\quad t_{ii}(x) := \lim_{\varepsilon\downarrow 0} t_{ii}(x+\boldsymbol{i}\varepsilon)\ ,\quad \tilde t_{jj}(x) \ := \ \lim_{\varepsilon\downarrow 0}\tilde t_{jj}(x+\boldsymbol{i}\varepsilon)\ .$$

Recall that $\mathcal{S}_n$ denotes the support of the measure associated to the Stieltjes transform $\delta_n(z)$. Alternatively, $\mathcal{S}_n$ is the support of the probability distribution $\mathbb{P}_n$ associated to the Stieltjes transform $N^{-1}\mathrm{Tr}\,T_n(z)$.

We can now express simplified formulas.

[2]Proofs of the tightness of the processes $(M_n(z), z\in D_\varepsilon)$ and $(G_n(z), z\in D_\varepsilon)$ are left to the reader as the arguments are standard. For the latter (tightness of (G_n)), a meta-model argument can be used, see [32, Section 5.2.2].

Theorem 3 (Alternative expression for the covariance formula)**.** *Let Assumptions 1 and 2 hold and* $f, g \in C_c^3(\mathbb{R})$. *Then*

$$\Theta_{0,n}(f,g) \;=\; \frac{1}{2\pi^2}\int_{\mathcal{S}_n^2} f'(x)g'(y)\ln\left|\frac{s_n(x)-\overline{s_n(y)}}{s_n(x)-s_n(y)}\right| dxdy \;=\; \frac{1}{2\pi^2}\int_{\mathcal{S}_n^2} f'(x)g'(y)\ln\left|\frac{\Delta_n(x,y)}{\underline{\Delta}_n(x,y)}\right| dxdy\,,$$

where $\Delta_n(x,y) := \lim_{\varepsilon\downarrow 0}\Delta_n(x+\boldsymbol{i}\varepsilon, y+\boldsymbol{i}\varepsilon)$ *and* $\underline{\Delta}_n(x,y) := \lim_{\varepsilon\downarrow 0}\Delta_n(x+\boldsymbol{i}\varepsilon, y-\boldsymbol{i}\varepsilon)$, *and*

$$\Theta_{2,n}(\kappa,f,g) = \frac{\sigma^4\kappa}{\pi^2 n^2}\sum_{i=1}^N\sum_{j=1}^n \int_{\mathcal{S}_n} f'(x)\mathrm{Im}\left(x\, t_{ii}(x)\tilde{t}_{jj}(x)\right)\, dx \int_{\mathcal{S}_n} g'(y)\mathrm{Im}\left(y\, t_{ii}(y)\tilde{t}_{jj}(y)\right)\, dy\,.$$

Proof of Theorem 3 is postponed to Section 4.

Remark 2.7 (about the term $\Theta_{1,n}(\vartheta,f,g)$)**.** *We have not succeeded so far to establish the natural formula:*

$$\Theta_{1,n}(\vartheta,f,g) \overset{?}{=} \frac{\vartheta}{2\pi^2}\int_{\mathcal{S}_n^2} f'(x)g'(y)\ln\left|\frac{\Delta_n^\vartheta(x,y)}{\underline{\Delta}_n^\vartheta(x,y)}\right| dxdy\,.$$

We could only prove the following boundary value representation in Proposition 4.2:

$$\Theta_{1,n}(\vartheta,f,g) = -\frac{1}{4\pi^2}\lim_{\varepsilon\searrow 0}\sum_{\pm_1,\pm_2}(\pm_1\pm_2)\int_{\mathbb{R}^2} f(x)g(y)\Theta_{1,n}(\vartheta, x\pm_1 \boldsymbol{i}\varepsilon, y\pm_2 \boldsymbol{i}\varepsilon)dxdy\,,$$

where $\pm_1,\pm_2 \in \{+,-\}$ *and* $\pm_1\pm_2$ *is the sign resulting from the product* $\pm_1 1$ *by* $\pm_2 1$.

Remark 2.8 (more simplifications)**.** *The following simplifications occur:*

(1) If $\vartheta = 0$, *then*

$$\Theta_{1,n}(\vartheta,f,g)\mid_{\vartheta=0} = 0\ .$$

(2) If $\kappa = 0$ *(Gaussian moments of order 1, 2, 4) then*

$$\Theta_{2,n}(\kappa,f,g)\mid_{\kappa=0} = 0\ .$$

(3) For real entries (x_{ij}) *(corresponding to* $\vartheta = 1$*) and real matrix* A, *then*

$$\Theta_{1,n}(\vartheta,f,g)\mid_{\vartheta=1} \;=\; \Theta_{0,n}(f,g) \quad \text{and} \quad \mathrm{Cov}\left(Z_n(f), Z_n(g)\right) = 2\Theta_{0,n}(f,g) + \Theta_{2,n}(\kappa,f,g)\ .$$

Remark 2.9 (relaxing the support compactness of test functions)**.** *Let the framework of Remark 2.8-(1) or (3) holds, so that an explicit expression of the variance as provided in Theorem 3 is available. Then combining Theorem 2 and an argument of spectrum confinement (see for instance* [1], [9, Theorem 5.2]*), one can obtain the following fluctuation result: let* $f \in C^3(\mathbb{R})$ *(notice that* f *has no longer a bounded support) and let* $\boldsymbol{h}:\mathbb{R}\to[0,1]$ *a* $C_c^\infty(\mathbb{R})$ *function with value 1 on* $\mathcal{S}_n$, *then*

$$d_{\mathcal{L}P}\left(\sum_{i=1}^N f(\lambda_i) - \sum_{i=1}^N \mathbb{E}(f\boldsymbol{h})(\lambda_i)\ ,\ Z_n(f)\right) \xrightarrow[N,n\to\infty]{} 0\,,$$

where $Z_n(f)$ *is a Gaussian random variable with variance given by* $2\Theta_{0,n}(f,f) + \Theta_{2,n}(\kappa,f,f)$. *A similar extension for a different matrix model is available in* [32, Corollary 4.3].

2.4. **Remarks concerning the bias.** We have provided so far fluctuation results for quantities $\sum_{i=1}^N f(\lambda_i) - \mathbb{E}\sum_{i=1}^N f(\lambda_i)$. Let $\mathbb{P}_n$ be the probability distribution associated to the Stieltjes transform $\frac{1}{N}\mathrm{Tr}\, T_n(z)$.

The study of the biases

$$\mathbb{E}\mathrm{Tr}\, Q_n(z) - \mathrm{Tr}\, T_n(z) \qquad \text{and} \qquad \mathbb{E}\sum_{i=1}^N f(\lambda_i) - N\int_{\mathbb{R}^+} f(\lambda)\mathbb{P}_n(d\lambda)$$

is an interesting question, computationally challenging, that we only superficially address hereafter, for the simple case of complex standard Gaussian entries.

Proposition 2.10. *Assume that the random variables* $(\mathbf{x}_{ij}, 1\le i\le N, 1\le j\le n)$ *are i.i.d. complex standard Gaussian entries, that is* $\mathbf{x}_{ij} = 2^{-1}(U_{ij} + \boldsymbol{i}V_{ij})$ *where* U_{ij} *and* V_{ij} *are real standard Gaussian entries. Assume moreover that Assumption 2 holds. Then*

$$\mathbb{E}\mathrm{Tr}\, Q_n(z) - \mathrm{Tr}\, T_n(z) = \frac{1}{n}\Pi_1(|z|)\Pi_2\left(\frac{1}{\mathrm{Im}(z)}\right)\,,$$

where Π_1 *and* Π_2 *are polynomials with fixed degree independent from* n. *Denote by* k_0 *the degree of* Π_2. *Let* $f \in C_c^{k_0+1}(\mathbb{R})$, *then*

$$\mathbb{E}\sum_{i=1}^N f(\lambda_i) - N\int_{\mathbb{R}^+} f(\lambda)\mathbb{P}_n(d\lambda) = \mathcal{O}\left(\frac{1}{n}\right)$$

The first part of the proposition can be proved as in [15, Theorem 2], [10], [29, Lemma 4] and one can track down the minimal value of k_0 by carefully following these proofs. The second part of the proposition is a mere application of Helffer-Sjöstrand formula.

Remark 2.11 (relaxing the support compactness of test functions - continued)**.** *Combining the previous proposition and Remark 2.9, one obtains the following fluctuation result for an information-plus-noise matrix with standard complex Gaussian entries: let $f \in C^{k_0+1}(\mathbb{R})$ then*

$$d_{\mathcal{L}P}\left(\sum_{i=1}^{N} f(\lambda_i) - N\int_{\mathcal{S}_n} f(\lambda)\mathbb{P}_n(d\lambda)\ ,\ Z_n(f)\right) \xrightarrow[N,n\to\infty]{} 0\,,$$

where $Z_n(f)$ is a centered Gaussian random variable with variance given by $\Theta_{0,n}(f,f)$.

3. Proof of Theorem 1: The CLT for the trace of the resolvent

3.1. General properties of s_n, Δ_n and Δ_n^{ϑ}.

Recall the definitions of s, γ, $\tilde{\gamma}$, ν and Δ introduced in (1.11), (2.2)-(2.3). We provide hereafter various important properties from which Proposition 2.3 follows.

Proposition 3.1. *Let δ and $\tilde{\delta}$ be the Stieltjes transforms solution of (1.9) and recall that*

$$s_z = z(1+\sigma\delta_z)(1+\sigma\tilde{\delta}_z)\,, \quad z\in\mathbb{C}^+\,.$$

1. *Function $s:\mathbb{C}^+\to\mathbb{C}^+$ is analytic.*
2. *Let $z, z_1, z_2 \in \mathbb{C}^+$, then*
$$\delta_z \;=\; \tilde{\delta}_z + \frac{\sigma(1-c_n)}{z}\,.$$
In particular, if $s_{z_1} = s_{z_2}$ then $z_1 = z_2$.
3. *Let $z_1, z_2 \in \mathbb{C}^+$ with $z_1 \neq z_2$, then the following identities hold*
$$\gamma(z_1,z_2) = \sigma\frac{\delta_{z_1}-\delta_{z_2}}{s_{z_1}-s_{z_2}}\,, \qquad \tilde{\gamma}(z_1,z_2) = \sigma\frac{\tilde{\delta}_{z_1}-\tilde{\delta}_{z_2}}{s_{z_1}-s_{z_2}}\,,$$
$$1-\nu(z_1,z_2) = \frac{z_1(1+\sigma\tilde{\delta}_{z_1}) - z_2(1+\sigma\tilde{\delta}_{z_2})}{s_{z_1}-s_{z_2}}\,, \qquad \Delta(z_1,z_2) = \frac{z_1-z_2}{s_{z_1}-s_{z_2}}\,.$$
4. *Let $z\in\mathbb{C}^+$ and $A\neq 0$ then the following inequalities hold*
$$0 \;<\; \nu(z,\bar{z}) < 1\,,\qquad \Delta(z,\bar{z}) \;>\; 0\,,\qquad 0 \;<\; |z|^2\gamma(z,\bar{z})\tilde{\gamma}(z,\bar{z}) \;<\; 1\,. \tag{3.1}$$
5. *Let $z\in\mathbb{C}^+$. If $c_n\le 1$ then*
$$\gamma(z,\bar{z}) < \frac{1}{|z|} \qquad \textit{and} \qquad |\delta_z| \le \frac{\sqrt{c_n}}{\sqrt{|z|}}\,.$$
If $c_n = 1$ then $\gamma(z,\bar{z}) = \tilde{\gamma}(z,\bar{z})$ and $\tilde{\gamma}(z,\bar{z}) \le |z|^{-1}$.

Recall the definition of Δ_n^{ϑ} in (2.5).

6. *Let $z_1, z_2\in\mathbb{C}^+$ then*
$$|z_1 z_2 \gamma(z_1,z_2)\tilde{\gamma}(z_1,z_2)| < 1\,,\qquad |\Delta(z_1,z_2)| > 0\,,\qquad |\Delta^{\vartheta}(z_1,z_2)| > 0\,.$$

Proof. Function s is obviously analytic. The mere definition of δ and $\tilde{\delta}$ yields

$$\frac{\delta_z}{1+\sigma\delta_z} = \frac{\sigma}{n}\mathrm{Tr}\,(-s_z + \mathrm{AA}^*)^{-1} \qquad\text{and}\qquad \frac{\tilde{\delta}_z}{1+\sigma\tilde{\delta}_z} = \frac{\sigma}{n}\mathrm{Tr}\,(-s_z+\mathrm{A}^*\mathrm{A})^{-1} \tag{3.2}$$

from which we deduce that for all $z\in\mathbb{C}^+$, s_z does not belong to the spectrum of AA^*. Taking the conjugate and applying the resolvent identity, we obtain

$$\frac{\mathrm{Im}(\delta_z)}{|1+\sigma\delta_z|^2} = \mathrm{Im}(s_z)\frac{\sigma}{n}\mathrm{Tr}\,(-s_z+\mathrm{AA}^*)^{-1}(-\overline{s_z}+\mathrm{AA}^*)^{-1}\,,$$

that is $s_z\in\mathbb{C}^+$. Item (1) is proved.

Comparing the spectra of AA^* and $\mathrm{A}^*\mathrm{A}$ we obtain

$$\frac{\sigma}{n}\mathrm{Tr}\,(-s_z+\mathrm{A}^*\mathrm{A})^{-1} = \frac{\sigma}{n}\mathrm{Tr}\,(-s_z+\mathrm{AA}^*)^{-1} - \frac{\sigma}{s_z}(1-c_n)\,,$$

hence using (3.2) we get

$$\frac{\delta_z}{1+\sigma\delta_z} = \frac{\tilde{\delta}_z}{1+\sigma\tilde{\delta}_z} + \frac{\sigma}{s_z}(1-c_n)$$

from which we deduce the desired identity. Applying (3.2) to $z = z_1$ and $z = z_2$ and substracting yields

$$\frac{\delta_{z_1}}{1+\sigma\delta_{z_1}} - \frac{\delta_{z_2}}{1+\sigma\delta_{z_2}} = (s_{z_1} - s_{z_2})\frac{\sigma}{n}\mathrm{Tr}\,(-s_{z_1} + \mathrm{AA}^*)^{-1}(-s_{z_2} + \mathrm{AA}^*)^{-1} \ , \tag{3.3}$$

from which we deduce that $s_{z_1} = s_{z_2}$ implies that $\delta_{z_1} = \delta_{z_2}$. Assume that $\delta_{z_1} = \delta_{z_2} = \delta^*$ then

$$s_{z_{1,2}} = z_{1,2}(1+\sigma\delta_{z_{1,2}})(1+\sigma\tilde{\delta}_{z_{1,2}}) = z_{1,2}(1+\sigma\delta^*)^2 - \sigma^2(1-c_n)(1+\sigma\delta^*) \ .$$

Hence

$$s_{z_1} - s_{z_2} = (z_1 - z_2)(1+\sigma\delta^*)^2 \ .$$

Since $z \mapsto \delta(z)$ is the Stieltjes transform of a positive measure with support in $\mathbb{R}^+$, so is $[-z(1+\sigma\delta)]^{-1}$. In particular, $|z(1+\sigma\delta)|^{-1} \le \mathrm{Im}(z)^{-1}$ and $|1+\sigma\delta| \ge \mathrm{Im}(z)/|z|$, which guarantees that $(1+\sigma\delta^*)^2 \neq 0$. Necessarily, $z_1 = z_2$. Item (2) is proved.

The first formula of item (3) immediatly follows from (3.3). The second formula can be obtained similarly. We now apply the resolvent identity to $\delta_{z_1} - \delta_{z_2}$ and obtain, after simplification

$$\begin{aligned}\delta_{z_1} - \delta_{z_2} &= \frac{\sigma}{n}\left(z_1(1+\sigma\tilde{\delta}_{z_1}) - z_2(1+\sigma\tilde{\delta}_{z_2})\right)\mathrm{Tr}\,\mathrm{T}_{z_1}\mathrm{T}_{z_2} + (\delta_{z_1} - \delta_{z_2})\frac{\sigma^2}{n}\frac{\mathrm{Tr}\,\mathrm{T}_{z_1}\mathrm{AA}^*\mathrm{T}_{z_2}}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})} \ , \\ &= \frac{z_1(1+\sigma\tilde{\delta}_{z_1}) - z_2(1+\sigma\tilde{\delta}_{z_2})}{s_{z_1} - s_{z_2}}(\delta_{z_1} - \delta_{z_2}) + (\delta_{z_1} - \delta_{z_2})\nu(z_1, z_2) \ .\end{aligned}$$

If $\delta_{z_1} \neq \delta_{z_2}$ then we simply divide by $\delta_{z_1} - \delta_{z_2}$ and obtain the third formula. If $\delta_{z_1} = \delta_{z_2}$, consider $z_1^n \neq z_1$ with $z_1^n \to z_1$. Since the zeros of the analytic function $\omega \mapsto \delta_\omega - \delta_{z_2}$ are isolated, $\delta_{z_1^n} \neq \delta_{z_2}$ for n large enough and one obtains the desired formula for (z_1^n, z_2) as previously. By continuity, the formula remains true for (z_1, z_2).

Using the previously established formulas, we now express $\Delta(z_1, z_2)$.

$$\begin{aligned}\Delta(z_1, z_2) &= (1 - \nu(z_1, z_2))^2 - z_1 z_2 \gamma(z_1, z_2)\tilde{\gamma}(z_1, z_2) \ , \\ &= \frac{(z_1(1+\sigma\tilde{\delta}_{z_1}) - z_2(1+\sigma\tilde{\delta}_{z_2}))^2 - \sigma^2 z_1 z_2(\delta_{z_1} - \delta_{z_2})(\tilde{\delta}_{z_1} - \tilde{\delta}_{z_2})}{(s_{z_1} - s_{z_2})^2} \ .\end{aligned}$$

We focus on the numerator

$$\begin{aligned}&\left(z_1(1+\sigma\tilde{\delta}_{z_1}) - z_2(1+\sigma\tilde{\delta}_{z_2})\right)^2 - \sigma^2 z_1 z_2(\delta_{z_1} - \delta_{z_2})\left(\tilde{\delta}_{z_1} - \tilde{\delta}_{z_2}\right) \\ &= z_1^2(1+\sigma\tilde{\delta}_{z_1})^2 + z_2^2(1+\sigma\tilde{\delta}_{z_2})^2 - 2z_1 z_2(1+\sigma\tilde{\delta}_{z_1})(1+\sigma\tilde{\delta}_{z_2}) \\ &\quad -\sigma^2 z_1 z_2\left(\tilde{\delta}_{z_1} + (1-c_n)\frac{\sigma}{z_1} - \tilde{\delta}_{z_2} - (1-c_n)\frac{\sigma}{z_2}\right)(\tilde{\delta}_{z_1} - \tilde{\delta}_{z_2}) \ , \\ &= z_1^2(1+\sigma\tilde{\delta}_{z_1})^2 + z_2^2(1+\sigma\tilde{\delta}_{z_2})^2 - 2z_1 z_2(1+\sigma\tilde{\delta}_{z_1})(1+\sigma\tilde{\delta}_{z_2}) \\ &\quad -z_1 z_2\left(1+\sigma\tilde{\delta}_{z_1} - (1+\sigma\tilde{\delta}_{z_2})\right)^2 + \sigma^3(1-c_n)(z_1 - z_2)(\tilde{\delta}_{z_1} - \tilde{\delta}_{z_2}) \ , \\ &= z_1^2(1+\sigma\tilde{\delta}_{z_1})^2 + z_2^2(1+\sigma\tilde{\delta}_{z_2})^2 - z_1 z_2\left((1+\sigma\tilde{\delta}_{z_1})^2 + (1+\sigma\tilde{\delta}_{z_2})^2\right) \\ &\quad +\sigma^3(1-c_n)(z_1 - z_2)(\tilde{\delta}_{z_1} - \tilde{\delta}_{z_2}) \ , \\ &= (z_1 - z_2)\left(z_1(1+\sigma\tilde{\delta}_{z_1})^2 + \sigma^2(1-c_n)(1+\sigma\tilde{\delta}_{z_1})\right) \\ &\quad -(z_1 - z_2)\left(z_2(1+\sigma\tilde{\delta}_{z_2})^2 + \sigma^2(1-c_n)(1+\sigma\tilde{\delta}_{z_2})\right) \ .\end{aligned}$$

It remains to notice that

$$s_z = z(1+\sigma\delta_z)(1+\sigma\tilde{\delta}_z) = z(1+\sigma\tilde{\delta}_z)^2 + \sigma^2(1-c_n)(1+\sigma\tilde{\delta}_z)$$

to conclude that the numerator writes $(z_1 - z_2)(s_{z_1} - s_{z_2})$. The formula for $\Delta(z_1, z_2)$ immediatly follows, and item (3) is proved.

Let $z \in \mathbb{C}^+$, then the mere definition of ν yields $\nu(z, \bar{z}) > 0$ for $A \neq 0$. Recall that since $\tilde{\delta}$ is the Stieltjes transform of a measure with support in $\mathbb{R}^+$, then $\mathrm{Im}(z\tilde{\delta}_z) \ge 0$. By the formula established in (3),

$$1 - \nu(z, \bar{z}) = \frac{\mathrm{Im}(z) + \sigma\mathrm{Im}(z\tilde{\delta}_z)}{\mathrm{Im}(s_z)} > 0$$

hence $\nu(z, \bar{z}) < 1$. Similarly,

$$\Delta(z, \bar{z}) = \frac{z - \bar{z}}{s_z - \overline{s_z}} = \frac{\mathrm{Im}(z)}{\mathrm{Im}(s_z)} \ > \ 0 \,,$$

from which we deduce that

$$0 \ < \ |z|^2\gamma(z, \bar{z})\tilde{\gamma}(z, \bar{z}) \ < \ (1 - \nu(z, \bar{z}))^2 \ < \ 1 \,.$$

Proof of item (4) is completed.

Using the relation between δ_z and $\tilde{\delta}_z$, we obtain

$$\tilde{\gamma}(z,\bar{z}) = \frac{\sigma}{s_z - \overline{s_z}}\left(\delta_z - \overline{\delta_z} - \frac{\sigma(1-c_n)}{z} + \frac{\sigma(1-c_n)}{\bar{z}}\right) = \gamma(z,\bar{z}) + \frac{\sigma^2(1-c_n)}{|z|^2}\frac{\operatorname{Im}(z)}{\operatorname{Im}(s_z)}\,.$$

In particular, $\gamma(z,\bar{z}) < \tilde{\gamma}(z,\bar{z})$ if $c_n < 1$ and $\gamma(z,\bar{z}) = \tilde{\gamma}(z,\bar{z})$ if $c_n = 1$. Plugging this into the last inequality of (3.1), we get $|z|^2\gamma^2(z,\bar{z}) < 1$ which is the desired inequality. Finally, we use the elementary inequality $|\operatorname{Tr}(AB)| \le \sqrt{\operatorname{Tr} AA^*}\sqrt{\operatorname{Tr} BB^*}$ to obtain

$$|\delta_z| = \left|\frac{\sigma}{n}\operatorname{Tr} T_z\right| \le \sqrt{\gamma(z,\bar{z})}\frac{\sqrt{\operatorname{Tr} I_N}}{\sqrt{n}} \le \frac{\sqrt{c_n}}{\sqrt{|z|}}\,.$$

Item (5) is proved.

Using the mere definition of γ, we have

$$|\gamma(z_1,z_2)| = \left|\frac{\sigma^2}{n}\operatorname{Tr} T_{z_1}T_{z_2}\right| \le \left(\frac{\sigma^2}{n}\operatorname{Tr} T_{z_1}T_{z_1}^*\right)^{1/2}\left(\frac{\sigma^2}{n}\operatorname{Tr} T_{z_2}T_{z_2}^*\right)^{1/2} = \sqrt{\gamma(z_1,\overline{z_1})}\sqrt{\gamma(z_2,\overline{z_2})}\,.$$

Similarly, using the definition of $\tilde{\gamma}$, we get that $|\tilde{\gamma}(z_1,z_2)| \le \sqrt{\tilde{\gamma}(z_1,\overline{z_1})}\sqrt{\tilde{\gamma}(z_2,\overline{z_2})}$ and hence

$$|z_1z_2\gamma(z_1,z_2)\tilde{\gamma}(z_1,z_2)| \le \sqrt{|z_1|^2\gamma(z_1,\overline{z_1})\tilde{\gamma}(z_1,\overline{z_1})|}\sqrt{|z_2|^2\gamma(z_2,\overline{z_2})\tilde{\gamma}(z_2,\overline{z_2})|} < 1$$

by (3.1) and the first inequality of item (6) is proved. We now prove that

$$|\Delta(z_1,z_2)| \;\ge\; \sqrt{\Delta(z_1,\overline{z_1})}\sqrt{\Delta(z_2,\overline{z_2})} \tag{3.4}$$

where the last quantity is positive by item (4). We have

$$|1-\nu(z_1,z_2)| \;\ge\; 1-|\nu(z_1,z_2)| \;\ge\; 1-\sqrt{\nu(z_1,\overline{z_1})}\sqrt{\nu(z_2,\overline{z_2})} \;>\; 0\,.$$

Hence

$$\begin{aligned}|\Delta(z_1,z_2)| \;&>\; \left(1-\sqrt{\nu(z_1,\overline{z_1})}\sqrt{\nu(z_2,\overline{z_2})}\right)^2\\ &\quad -\sqrt{|z_1|^2\gamma(z_1,\overline{z_1})\tilde{\gamma}(z_1,\overline{z_1})|}\sqrt{|z_2|^2\gamma(z_2,\overline{z_2})\tilde{\gamma}(z_2,\overline{z_2})|}\,.\end{aligned} \tag{3.5}$$

We now rely on elementary inequalities (for a proof see [23, Proposition 6.1]) to conclude:

Proposition 3.2. *(1) Let $a_1, a_2 \ge 0$, then*

$$(1-\sqrt{a_1a_2})^2 \;>\; (1-a_1)(1-a_2)\,.$$

(2) Assume moreover that $b_i \ge 0$ and $(1-a_i)^2 - b_i > 0$ for $i=1,2$, then

$$(1-\sqrt{a_1a_2})^2 - \sqrt{b_1b_2} \;>\; \sqrt{(1-a_1)^2-b_1}\sqrt{(1-a_2)^2-b_2}\,.$$

Using the second inequality of the previous proposition in (3.5) yields (3.4).

In order to handle Δ^ϑ, notice that

$$\begin{aligned}|\gamma^\dagger(z_1,z_2)| &\le \sqrt{\gamma(z_1,\overline{z_1})}\sqrt{\gamma(z_2,\overline{z_2})}\,, & |\nu^\dagger(z_1,z_2)| &\le \sqrt{\nu(z_1,\overline{z_1})}\sqrt{\nu(z_2,\overline{z_2})}\,,\\ |\tilde{\gamma}^\dagger(z_1,z_2)| &\le \sqrt{\tilde{\gamma}(z_1,\overline{z_1})}\sqrt{\tilde{\gamma}(z_2,\overline{z_2})}\,, & |\tilde{\nu}^\dagger(z_1,z_2)| &\le \sqrt{\nu(z_1,\overline{z_1})}\sqrt{\nu(z_2,\overline{z_2})}\,.\end{aligned}$$

Hence

$$\begin{aligned}\left|\Delta^\vartheta(z_1,z_2)\right| \;&\ge\; \left(1-|\vartheta|\sqrt{\nu(z_1,\overline{z_1})}\sqrt{\nu(z_2,\overline{z_2})}\right)^2\\ &\quad -\sqrt{|z_1|^2\gamma(z_1,\overline{z_1})\tilde{\gamma}(z_1,\overline{z_1})|}\sqrt{|z_2|^2\gamma(z_2,\overline{z_2})\tilde{\gamma}(z_2,\overline{z_2})|}\,.\end{aligned} \tag{3.6}$$

Since $|\vartheta| \le 1$, we obtain the same lower bound as in (3.5), from which we can conclude as previously. Item (6) is proved. Proof of Proposition 3.1 is completed. □

3.2. **Technical means and outline of the proof.** We first prove that under Assumptions 1 and 2 $M_n(z)$ defined in (1.6) can be written as the sum of martingale increments:

$$M_n(z) = \sum_{j=1}^n P_j(z) + o_P(1)\,, \tag{3.7}$$

see (3.12)-(3.14) below. This decomposition allows to establish its Gaussian fluctuations via powerful CLTs for martingales such as [6, Th. 35.12] and [32, Lemma 5.6]. For the reader's convenience, we recall the latter.

Lemma 3.3 ([32], Lemma 5.6). *Suppose that for each n, $(Y_{nj}; 1 \le j \le r_n)$ is a $\mathbb{C}^d$-valued martingale difference sequence with respect to the increasing σ-field $\{\mathcal{F}_{n,j}; 1 \le j \le r_n\}$ having second moments. Write*

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

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

$$\sum_{j=1}^{r_n} \mathbb{E}\left(Y_{nj}^k \bar{Y}_{nj}^\ell \mid \mathcal{F}_{n,j-1}\right) - \Theta_n(k,\ell) \xrightarrow[n\to\infty]{\mathcal{P}} 0 \qquad \text{and} \qquad \sum_{j=1}^{r_n} \mathbb{E}\left(Y_{nj}^k Y_{nj}^\ell \mid \mathcal{F}_{n,j-1}\right) - \tilde{\Theta}_n(k,\ell) \xrightarrow[n\to\infty]{\mathcal{P}} 0\ , \tag{3.8}$$

and for each $\varepsilon > 0$, the following Lyapunov condition holds true:

$$\sum_{j=1}^{r_n} \mathbb{E}\left(\|Y_{nj}\|^2 \mathbf{1}_{\|Y_{nj}\|>\varepsilon}\right) \xrightarrow[n\to\infty]{} 0\ . \tag{3.9}$$

Then $d_{\mathcal{LP}}\left(\sum_{j=1}^{r_n} Y_{nj}\ , Z_n\right) \xrightarrow[n\to\infty]{} 0$, or equivalently for any continuous bounded function $f : \mathbb{C}^d \to \mathbb{R}$:

$$\mathbb{E}f\Big(\sum_{j=1}^{r_n} Y_{nj}\Big) - \mathbb{E}f(Z_n) \xrightarrow[n\to\infty]{} 0\ ,$$

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

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

Lemma 3.3 can be strengthened with the following lemma:

Lemma 3.4 (cf. Lemma 5.7 in [32]). *Let K be a compact set in $\mathbb{C}$ and let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be random elements in $C(K,\mathbb{C})$. Assume that for all $d \ge 1$, $z_1, \ldots, z_d \in K$, $f \in C(\mathbb{C}^d, \mathbb{C})$ we have:*

$$\mathbb{E}f(X_n(z_1), \ldots, X_n(z_d)) - \mathbb{E}f(Y_n(z_1), \ldots, Y_n(z_d)) \xrightarrow[n\to\infty]{} 0\,.$$

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

$$\mathbb{E}F(X_n) - \mathbb{E}F(Y_n) \xrightarrow[n\to\infty]{} 0\,.$$

We can now provide an outline of the proof:

1. The martingale decomposition $M_n(z) = \sum_{j=1}^n P_j(z) + o_P(1)$ is established in Section 3.3.
2. Lyapunov's condition (3.9) is established for $Y_{nj}^{\mathsf{T}} = (P_j(z_1), \cdots, P_j(z_d))$ in Section 3.3.
3. Convergence (3.8), the more demanding part of the proof, is performed in Sections 3.4 and 3.5.
4. The tightness for $(M_n(z),\, z \in D_\varepsilon)$ and of the Gaussian process $(G_n(z),\, z \in D_\varepsilon)$ rely on standard arguments and computations and are thus skipped.

3.3. **Sum of martingale increments and Lyapunov's condition.** We introduce now some notations. Denote by $\mathbf{y}_j$, $\mathbf{a}_j$, $\mathbf{x}_j$ and $\mathbf{r}_j$ the j^{th} columns of the matrices Y_n, A_n, X_n and $\frac{\sigma}{\sqrt{n}}\mathrm{X}_n$ respectively and let

$$\mathrm{Y}_j\mathrm{Y}_j^* := \mathrm{Y}\mathrm{Y}^* - \mathbf{y}_j\mathbf{y}_j^* = \sum_{\ell\neq j} \mathbf{y}_\ell \mathbf{y}_\ell^*\ .$$

Recalling that $\mathrm{Q}_z = (\mathrm{Y}_n\mathrm{Y}_n^* - z\mathrm{I}_N)^{-1}$, we denote by $\mathrm{Q}_{z,j} := \mathrm{Q}_j(z) = (\mathrm{Y}_j\mathrm{Y}_j^* - z\mathrm{I}_N)^{-1}$ and by $\tilde{\mathrm{Q}}_{z,j} := \tilde{\mathrm{Q}}_j(z) = (\mathrm{Y}_j^*\mathrm{Y}_j - z\mathrm{I}_n)^{-1}$ and finally note that the diagonal entries $\tilde{q}_{z,jj} = [\tilde{\mathrm{Q}}_z]_{jj}$ of the co-resolvent are given by

$$\tilde{q}_{jj} = \frac{-1}{z(1 + \mathbf{y}_j^* \mathrm{Q}_{z,j}\mathbf{y}_j)}\ ,$$

where this formula is obtained by combining formulas for the inverse of a partitioned matrix [24, Section 0.7.3] and Woodbury's formula [24, Section 0.7.4]. We now introduce several notations that will be used all along this paper. Denote by:

$$\begin{aligned}
b_{z,j} &= \frac{-1}{z(1 + \frac{\sigma^2}{n}\mathrm{Tr}\mathbb{E}\mathrm{Q}_{z,j} + \mathbf{a}_j^*\mathbb{E}\mathrm{Q}_{z,j}\mathbf{a}_j)},\\
\tilde{b}_{z,j} &= \frac{-1}{z(1 + \frac{\sigma^2}{n}\mathrm{Tr}\mathrm{Q}_{z,j} + \mathbf{a}_j^*\mathrm{Q}_{z,j}\mathbf{a}_j)}\ ,\\
\tau_{z,j} &= \mathbf{y}_j^*\mathrm{Q}_{z,j}\mathbf{y}_j - \frac{\sigma^2}{n}\mathrm{Tr}\mathbb{E}\mathrm{Q}_{z,j} - \mathbf{a}_j^*\mathbb{E}\mathrm{Q}_{z,j}\mathbf{a}_j,
\end{aligned}$$

$$\begin{aligned}\hat{\tau}_{z,j} &= \mathbf{y}_j^* \mathrm{Q}_{z,j}\mathbf{y}_j - \frac{\sigma^2}{n}\mathrm{Tr}\mathrm{Q}_{z,j} - \mathbf{a}_j^*\mathrm{Q}_{z,j}\mathbf{a}_j ,\\ \alpha_{z,j} &= \mathbf{y}_j^* \mathrm{Q}_{z,j}^2\mathbf{y}_j - \frac{\sigma^2}{n}\mathrm{Tr}\mathrm{Q}_{z,j}^2 - \mathbf{a}_j^*\mathrm{Q}_{z,j}^2\mathbf{a}_j .\end{aligned}$$

When no confusion occurs, we drop the variable z and write $\tilde{q}_{z,jj} := \tilde{q}_{jj}(z)$, $\tilde{t}_{z,jj} := \tilde{t}_{jj}(z)$, $\mathrm{T}_z := \mathrm{T}(z)$, etc.

Let $\mathbb{E}_0 = \mathbb{E}$ denote the expectation and $\mathbb{E}_j$ the conditional expectation with respect to the σ-field $\mathcal{F}_{n,j}$ generated by $\{\mathbf{x}_\ell\,,\ 1\le \ell\le j\}$. By the rank-one perturbation formula

$$\mathrm{Q} - \mathrm{Q}_j = z\tilde{q}_{jj}\mathrm{Q}_j\mathbf{y}_j\mathbf{y}_j^*\mathrm{Q}_j\,, \tag{3.10}$$

and the definition of M_n in (1.6), we have

$$M_n = \sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1})\mathrm{Tr}\,(\mathrm{Q}-\mathrm{Q}_j) = \sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1}) z\tilde{q}_{jj}\mathbf{y}_j^*\mathrm{Q}_j^2\mathbf{y}_j\,.$$

Note that

$$\tilde{q}_{jj} = \tilde{b}_j + z\tilde{q}_{jj}\tilde{b}_j\hat{\tau}_j = \tilde{b}_j + z\tilde{b}_j^2\hat{\tau}_j + z^2\tilde{b}_j^2\hat{\tau}_j^2\tilde{q}_{jj}\ , \tag{3.11}$$

and develop $M_n(z)$ as follows

$$M_n = z\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1})(\tilde{b}_j + z\tilde{b}_j^2\hat{\tau}_j + z^2\tilde{b}_j^2\hat{\tau}_j^2\tilde{q}_{jj})\left(\alpha_j + \frac{\sigma^2}{n}\mathrm{Tr}\mathrm{Q}_j^2 + \mathbf{a}_j^*\mathrm{Q}_j^2\mathbf{a}_j\right) = \sum_{j=1}^n P_j(z) + \sum_{j=1}^n P_j'(z)\,, \tag{3.12}$$

where

$$\begin{aligned} P_j(z) &:= z\mathbb{E}_j\left(\tilde{b}_j\alpha_j + z\tilde{b}_j^2\hat{\tau}_j\left(\frac{\sigma^2}{n}\mathrm{Tr}\mathrm{Q}_j^2 + \mathbf{a}_j^*\mathrm{Q}_j^2\mathbf{a}_j\right)\right)\,,\\ P_j'(z) &:= z(\mathbb{E}_j - \mathbb{E}_{j-1})\left[z\tilde{b}_j^2\hat{\tau}_j\alpha_j + z^2\tilde{b}_j^2\hat{\tau}_j^2\tilde{q}_{jj}\alpha_j + z^2\tilde{b}_j^2\hat{\tau}_j^2\tilde{q}_{jj}\left(\frac{\sigma^2}{n}\mathrm{Tr}\mathrm{Q}_j^2 + \mathbf{a}_j^*\mathrm{Q}_j^2\mathbf{a}_j\right)\right]\,. \end{aligned}$$

Since $\tilde{q}_{z,jj}$ and $\tilde{b}_{z,j}$ are Stieltjes transforms of probability measures, we have

$$\max\left(|\tilde{q}_{z,jj}|, |\tilde{b}_{z,j}|\right) \le (\mathrm{Im}(z))^{-1}\,. \tag{3.13}$$

We decompose P_j' into three termes. By orthogonality, Cauchy-Schwarz's inequality and the estimates provided in Lemma A.2, we have

$$\begin{aligned}\mathbb{E}\Big|\sum_{j=1}^n z(\mathbb{E}_j - \mathbb{E}_{j-1})\left[z\tilde{b}_j^2\hat{\tau}_j\alpha_j\right]\Big|^2 &= \sum_{j=1}^n \mathbb{E}\Big|z(\mathbb{E}_j - \mathbb{E}_{j-1})\left[z\tilde{b}_j^2\hat{\tau}_j\alpha_j\right]\Big|^2,\\ &\le \frac{4|z|^2}{\mathrm{Im}(z)^4}\sum_{j=1}^n\left(\mathbb{E}|\hat{\tau}_j|^4\mathbb{E}|\alpha_j|^4\right)^{1/2} = \mathcal{O}_z\left(\frac{1}{n}\right)\,.\end{aligned}$$

In the same way, we control the other terms and prove that

$$\mathbb{E}\left|\sum_{j=1}^n P_j'(z)\right|^2 = \mathcal{O}_z\left(\frac{1}{n}\right)\,. \tag{3.14}$$

This implies that for any $z\in D^+\cup\overline{D^+}$, $M_n(z)$ verifies (3.7).

We now prove Lyapunov's condition (3.9). First note that $\mathbb{E}|P_j(z)|^4 = \mathcal{O}_z(n^{-2})$ by Lemma A.2. Thus for any $z_1,\dots,z_d\in\mathbb{C}^+$,

$$\sum_{j=1}^n \mathbb{E}\Big[\big(\sum_{\ell=1}^d |P_j(z_\ell)|^2\big)\mathbf{1}_{\{\sum_{\ell=1}^d |P_j(z_\ell)|^2\ge\varepsilon^2\}}\Big] \le \frac{1}{\varepsilon^2}\sum_{j=1}^n\mathbb{E}\left(\sum_{\ell=1}^d |P_j(z_\ell)|^2\right)^2 \le \frac{d}{\varepsilon^2}\sum_{j=1}^n\sum_{\ell=1}^d \mathbb{E}|P_j(z_\ell)|^4 \xrightarrow[N,n\to\infty]{} 0\,.$$

Lyapunov's condition is hence verified.

3.4. **Computation of the covariance: some preparation.** Recall the decomposition of $M_n(z)$ in (3.7). We shall prove that $(P_j, 1\le j\le n)$ satisfies (3.8) with Θ_n defined in (2.7). It is sufficient to prove that:

$$\sum_{j=1}^n \mathbb{E}_{j-1}P_j(z_1)P_j(z_2) - \Theta_n(z_1,z_2) \xrightarrow[N,n\to\infty]{\mathcal{P}} 0 \qquad\text{and}\qquad \sum_{j=1}^n \mathbb{E}_{j-1}P_j(z_1)\overline{P_j(z_2)} - \Theta_n(z_1,\overline{z}_2) \xrightarrow[N,n\to\infty]{\mathcal{P}} 0\,.$$

Due to the expression

$$P_j(z) = \mathbb{E}_j\left\{z\tilde{b}_{z,j}\alpha_{z,j} + z^2\tilde{b}_{z,j}^2\hat{\tau}_{z,j}\left(\frac{\sigma^2}{n}\mathrm{Tr}\mathrm{Q}_{z,j}^2 + \mathbf{a}_j^*\mathrm{Q}_{z,j}^2\mathbf{a}_j\right)\right\}\,,$$

we have $\overline{P_j(z)} = P_j(\bar z)$. As $D^+ \cup \overline{D^+}$ is stable under conjugation, it suffices to prove the convergence of $\sum_{i=1}^n \mathbb{E}_{j-1} P_j(z_1)P_j(z_2)$ for any $z_1, z_2 \in D^+ \cup \overline{D^+}$.

We introduce

$$
\begin{aligned}
\Gamma_j(z) &= z\tilde b_{z,j}\left(\mathbf{y}_j^* \mathrm{Q}_{z,j}\mathbf{y}_j - \frac{\sigma^2}{n}\mathrm{Tr}\mathrm{Q}_{z,j} - \mathbf{a}_j^*\mathrm{Q}_{z,j}\mathbf{a}_j\right) = z\tilde b_{z,j}\hat\tau_{z,j}\,,\\
A_n(z_1,z_2) &= \sum_{j=1}^n \mathbb{E}_{j-1}\left\{\mathbb{E}_j\Gamma_j(z_1)\mathbb{E}_j\Gamma_j(z_2)\right\}\,. \qquad (3.15)
\end{aligned}
$$

Since

$$
\frac{\partial}{\partial z}(\hat\tau_{z,j}) = \alpha_{z,j} \qquad \text{and} \qquad \frac{\partial}{\partial z}(z\tilde b_{z,j}) = z^2\tilde b_{z,j}^2(\mathbf{y}_j^*\mathrm{Q}_{z,j}^2\mathbf{y}_j - \alpha_{z,j})\,,
$$

we can easily prove that

$$
\frac{\partial^2}{\partial z_1\partial z_2}A_n(z_1,z_2) = \sum_{j=1}^n \mathbb{E}_{j-1}P_j(z_1)P_j(z_2)\,. \qquad (3.16)
$$

By the same arguments as in Bai and Silverstein [3, page 571] and [4, page 273], it is sufficient to study the convergence in probability to zero of

$$
A_n(z_1,z_2) - \Upsilon_n(z_1,z_2) \quad \text{where} \quad \frac{\partial^2}{\partial z_1\partial z_2}\Upsilon_n(z_1,z_2) = \Theta_n(z_1,z_2)
$$

and the uniform boundedness (in n) of Υ_n (the latter being easy to establish by Lemma 3.10 and similar results). We now slightly simplify the study of $A_n(z_1,z_2)$ and prove that:

$$
A_n(z_1,z_2) - \sum_{j=1}^n z_1z_2\tilde t_{z_1,jj}\tilde t_{z_2,jj}\mathbb{E}_{j-1}\left\{\mathbb{E}_j\hat\tau_{z_1,j}\mathbb{E}_j\hat\tau_{z_2,j}\right\} \xrightarrow[n\to\infty]{\mathcal{P}} 0\,. \qquad (3.17)
$$

In the following computations, we shall use repeatedly the fact that t_{ii} and $\tilde t_{jj}$ are bounded, as Stieltjes transforms. Indeed, by Lemma A.4, $\mathbb{E}|\tilde b_j - \tilde t_{jj}|^2 \le 2\mathbb{E}|\tilde b_j - \tilde q_{jj}|^2 + 2\mathbb{E}|\tilde q_{jj} - \tilde t_{jj}|^2 = \mathcal{O}(n^{-1})$. Using Cauchy-Schwarz inequality and Lemma A.2, we obtain

$$
\begin{aligned}
&\mathbb{E}\left|\mathbb{E}_{j-1}\left\{\mathbb{E}_j\left[\tilde b_{z_1,j}\hat\tau_{z_1,j}\right]\mathbb{E}_j\left[\tilde b_{z_2,j}\hat\tau_{z_2,j}\right]\right\} - \tilde t_{z_1,jj}\tilde t_{z_2,jj}\mathbb{E}_{j-1}\left\{\mathbb{E}_j\hat\tau_{z_1,j}\mathbb{E}_j\hat\tau_{z_2,j}\right\}\right|\\
&\quad= \mathbb{E}\left|\mathbb{E}_{j-1}\left\{\mathbb{E}_j\left[(\tilde b_{z_1,j} - \tilde t_{z_1,jj})\hat\tau_{z_1,j}\right]\mathbb{E}_j\left[\tilde b_{z_2,j}\hat\tau_{z_2,j}\right]\right\} + \tilde t_{z_1,jj}\mathbb{E}_{j-1}\left\{\mathbb{E}_j[\hat\tau_{z_1,j}]\mathbb{E}_j[(\tilde b_{z_2,j} - \tilde t_{z_2,jj})\hat\tau_{z_2,j}]\right\}\right|\,,\\
&\quad= \mathcal{O}(n^{-3/2})\,.
\end{aligned}
$$

Summing over j, we prove the convergence (3.17). Notice that $\mathbb{E}_j(\mathrm{Q}_{z_1,j})$ is $\mathcal{F}_{n,j-1}$ measurable. By applying [18, Equation (3.20)] to $\mathrm{M} = \mathbb{E}_j(\mathrm{Q}_{z_1,j})$ and $\mathrm{P} = \mathbb{E}_j(\mathrm{Q}_{z_2,j})$ (the random vector being $(x_{1j},\dots,x_{Nj})$), we obtain

$$
\sum_{j=1}^n z_1z_2\tilde t_{z_1,jj}\tilde t_{z_2,jj}\mathbb{E}_{j-1}[\mathbb{E}_j\hat\tau_{z_1,j}\mathbb{E}_j\hat\tau_{z_2,j}] := \sum_{j=1}^n\left(\xi_{1j} + \xi_{2j} + \xi'_{2j} + \xi_{3j} + \xi_{4j}\right),
$$

where

$$
\begin{aligned}
\xi_{1j} &:= \frac{\kappa\sigma^4}{n^2}z_1z_2\tilde t_{z_1,jj}\tilde t_{z_2,jj}\sum_{i=1}^N \mathbb{E}_j[Q_{z_1,j}]_{ii}\mathbb{E}_j[Q_{z_2,j}]_{ii}\,, && (3.18)\\
\xi_{2j} &:= \frac{\sigma^3}{n}z_1z_2\tilde t_{z_1,jj}\tilde t_{z_2,jj}\mathbb{E}(|x_{11}|^2x_{11})\\
&\qquad\times\left(\frac{\mathbf{a}_j^*(\mathbb{E}_j\mathrm{Q}_{z_1,j})\mathrm{vdiag}(\mathbb{E}_j\mathrm{Q}_{z_2,j})}{\sqrt n} + \frac{\mathbf{a}_j^*(\mathbb{E}_j\mathrm{Q}_{z_2,j})\mathrm{vdiag}(\mathbb{E}_j\mathrm{Q}_{z_1,j})}{\sqrt n}\right), && (3.19)\\
\xi'_{2j} &:= \frac{\sigma^3}{n}z_1z_2\tilde t_{z_1,jj}\tilde t_{z_2,jj}\mathbb{E}(|x_{11}|^2\bar x_{11})\\
&\qquad\times\left(\frac{\mathrm{vdiag}(\mathbb{E}_j\mathrm{Q}_{z_2,j})^{\mathsf T}(\mathbb{E}_j\mathrm{Q}_{z_1,j})\mathbf{a}_j}{\sqrt n} + \frac{\mathrm{vdiag}(\mathbb{E}_j\mathrm{Q}_{z_1,j})^{\mathsf T}(\mathbb{E}_j\mathrm{Q}_{z_2,j})\mathbf{a}_j}{\sqrt n}\right), && (3.20)\\
\xi_{3j} &:= \frac{\sigma^2}{n}z_1z_2\tilde t_{z_1,jj}\tilde t_{z_2,jj}\Big(\frac{\sigma^2}{n}\mathrm{Tr}\big(\mathbb{E}_j\mathrm{Q}_{z_1,j}\mathbb{E}_j\mathrm{Q}_{z_2,j}\big)\\
&\qquad+ \mathbf{a}_j^*\mathbb{E}_j\mathrm{Q}_{z_1,j}\mathbb{E}_j\mathrm{Q}_{z_2,j}\mathbf{a}_j + \mathbf{a}_j^*\mathbb{E}_j\mathrm{Q}_{z_2,j}\mathbb{E}_j\mathrm{Q}_{z_1,j}\mathbf{a}_j\Big)\,, && (3.21)\\
\xi_{4j} &:= \frac{\sigma^2}{n}z_1z_2\tilde t_{z_1,jj}\tilde t_{z_2,jj}\Big(\frac{1}{n}|\vartheta|^2\sigma^2\mathrm{Tr}\big(\mathbb{E}_j\mathrm{Q}_{z_1,j}\mathbb{E}_j\mathrm{Q}_{z_2,j}^{\mathsf T}\big)\\
&\qquad+ \vartheta\mathbf{a}_j^*\mathbb{E}_j\mathrm{Q}_{z_2,j}\mathbb{E}_j\mathrm{Q}_{z_1,j}^{\mathsf T}\bar{\mathbf{a}}_j + \bar\vartheta\mathbf{a}_j^{\mathsf T}\mathbb{E}_j\mathrm{Q}_{z_1,j}^{\mathsf T}\mathbb{E}_j\mathrm{Q}_{z_2,j}\mathbf{a}_j\Big)\,. && (3.22)
\end{aligned}
$$

In the following series of lemmas, we describe the behaviour of each sum. Recall the formula of the covariance in (2.7), then $\sum_{j=1}^n \xi_{1j}$ is associated to the term $\Theta_{2,n}$ while $\sum_{j=1}^n \xi_{3j}$ and $\sum_{j=1}^n \xi_{4j}$ correspond to $\Theta_{0,n}$ and $\Theta_{1,n}$ respectively. The terms $\sum_{j=1}^n \xi_{2j}$ and $\sum_{j=1}^n \xi'_{2j}$ have no contribution in the final expression.

In [18], the terms above have been studied in the case where $z_1 = z_2 = -\rho \in (-\infty, 0)$ and many computations performed there can be established for general $z_1, z_2 \in \mathbb{C}^+$ by mere book keeping. A technical issue however remains: the invertibility of systems of equations that appear when studying the terms $\sum_j \xi_{3j}$ and $\sum_j \xi_{4j}$. In this case, the generalization from $z_1 = z_2 = -\rho \in (-\infty, 0)$ to general $z_1, z_2 \in \mathbb{C}^+$ is not trivial and is carefully developed hereafter, cf. Lemma 3.10-(ii).

Lemma 3.5. *Let Assumptions 1 and 2 hold, then*

$$\sum_{j=1}^n \xi_{1j} - \frac{\kappa\sigma^4 z_1 z_2}{n^2} \sum_{i=1}^N t_{z_1,ii} t_{z_2,ii} \sum_{j=1}^n \tilde{t}_{z_1,jj} \tilde{t}_{z_2,jj} \xrightarrow[N,n\to\infty]{\mathcal{P}} 0\,.$$

Proof of Lemma 3.5 is similar to the proof of [18, Lemma 4.1] and is omitted.

Lemma 3.6. *Let Assumptions 1 and 2 hold, then*

$$\sum_{j=1}^n \xi_{2j} \xrightarrow[N,n\to\infty]{\mathcal{P}} 0 \quad \text{and} \quad \sum_{j=1}^n \xi'_{2j} \xrightarrow[N,n\to\infty]{\mathcal{P}} 0\,.$$

Proof of Lemma 3.6 is similar to the proof of [18, Lemma 4.2] and is also omitted.

Lemma 3.7. *Let Assumptions 1 and 2 hold, then*

$$\frac{\partial^2}{\partial z_1 \partial z_2} \sum_{j=1}^n \xi_{3j} - \frac{s'_n(z_1) s'_n(z_2)}{(s_n(z_1) - s_n(z_2))^2} + \frac{1}{(z_1 - z_2)^2} \xrightarrow[n\to\infty]{\mathcal{P}} 0\,,$$

where $s_n(z) = z(1+\sigma\delta_z)(1+\sigma\tilde{\delta}_z)$.

Proof of Lemma 3.7 is provided in Section 3.5.

Lemma 3.8. *Let Assumptions 1 and 2 hold, then*

$$\frac{\partial^2}{\partial z_1 \partial z_2} \sum_{j=1}^n \xi_{4j} - \Theta_{1,n} \xrightarrow[n\to\infty]{\mathcal{P}} 0\,,$$

where $\Theta_{1,n}$ *is defined in (2.7).*

Proof of Lemma 3.8 is very close to the proof of 3.7 and is thus omitted.

3.5. **Proof of Lemma 3.7.** We first recall the definition of ξ_{3j} and introduce an auxiliary quantity $\tilde{\xi}_{3j}$:

$$\begin{aligned}
\xi_{3j} &= \frac{\sigma^2}{n} z_1 z_2 \tilde{t}_{z_1,jj} \tilde{t}_{z_2,jj} \Big(\frac{\sigma^2}{n} \mathrm{Tr}\big(\mathbb{E}_j Q_{z_1,j} \mathbb{E}_j Q_{z_2,j}\big) + \mathbf{a}_j^* \mathbb{E}_j Q_{z_1,j} \mathbb{E}_j Q_{z_2,j} \mathbf{a}_j + \mathbf{a}_j^* \mathbb{E}_j Q_{z_2,j} \mathbb{E}_j Q_{z_1,j} \mathbf{a}_j \Big)\,, \\
\tilde{\xi}_{3j} &= \frac{\sigma^2}{n} z_1 z_2 \tilde{t}_{z_1,jj} \tilde{t}_{z_2,jj} \Big(\frac{\sigma^2}{n} \mathrm{Tr}\mathbb{E}\left\{\mathbb{E}_j Q_{z_1} \mathbb{E}_j Q_{z_2}\right\} + \mathbf{a}_j^* \mathbb{E}\left\{\mathbb{E}_j Q_{z_1,j} \mathbb{E}_j Q_{z_2,j}\right\} \mathbf{a}_j + \mathbf{a}_j^* \mathbb{E}\left\{\mathbb{E}_j Q_{z_2,j} \mathbb{E}_j Q_{z_1,j}\right\} \mathbf{a}_j \Big)\,.
\end{aligned}$$

By rank-one perturbation and Lemma A.5, we easily prove that

$$\sum_{i=1}^n \xi_{3j} - \sum_{j=1}^n \tilde{\xi}_{3j} \xrightarrow[n\to\infty]{\mathcal{P}} 0\,.$$

Consider the following notations:

$$\begin{aligned}
\psi_j(z_1,z_2) &= \frac{\sigma^2}{n} \mathrm{Tr}\mathbb{E}\left\{(\mathbb{E}_j Q_{z_1})(\mathbb{E}_j Q_{z_2})\right\} &&= \frac{\sigma^2}{n} \mathrm{Tr}\mathbb{E}\{(\mathbb{E}_j Q_{z_1}) Q_{z_2}\}\,, \\
\zeta_{kj}(z_1,z_2) &= \sigma \mathbb{E}\{\mathbf{a}_k^* (\mathbb{E}_j Q_{z_1})(\mathbb{E}_j Q_{z_2}) \mathbf{a}_k\} &&= \sigma \mathbb{E}\{\mathbf{a}_k^* (\mathbb{E}_j Q_{z_1}) Q_{z_2} \mathbf{a}_k\}\,, \\
\theta_{kj}(z_1,z_2) &= \sigma \mathbb{E}\{\mathbf{a}_k^* (\mathbb{E}_j Q_{z_1,k})(\mathbb{E}_j Q_{z_2,k}) \mathbf{a}_k\} &&= \sigma \mathbb{E}\{\mathbf{a}_k^* (\mathbb{E}_j Q_{z_1,k}) Q_{z_2,k}\, \mathbf{a}_k\}\,, \\
\phi_j(z_1,z_2) &= \frac{\sigma}{n} \sum_{k=1}^j z_1 z_2 \tilde{t}_{z_1,kk} \tilde{t}_{z_2,kk} \theta_{kj}(z_1,z_2)\,.
\end{aligned} \tag{3.23}$$

If clear from the context, we simply write ψ_j, ζ_{kj}, θ_{kj} and ϕ_j instead of $\psi_j(z_1,z_2)$, $\zeta_{kj}(z_1,z_2)$, $\theta_{kj}(z_1,z_2)$ and $\phi_j(z_1,z_2)$ respectively. With these notations at hand, $\tilde{\xi}_{3j}$ writes

$$\tilde{\xi}_{3j}(z_1,z_2) = \frac{1}{n} z_1 z_2 \tilde{t}_{jj}(z_1) \tilde{t}_{jj}(z_2) \big(\sigma^2 \psi_j(z_1,z_2) + \sigma \theta_{jj}(z_1,z_2) + \sigma \theta_{jj}(z_2,z_1) \big)\,.$$

The following part of the proof is inspired from [18]: Since it seems difficult to obtain a direct expression for the quantities ψ_j, ζ_{kj}, θ_{kj} and ϕ_j, we establish in the following lemma a system of (perturbed) equations which describes the structural links between these quantities.

Lemma 3.9. *Let Assumptions 1 and 2 hold and recall that $\gamma = \gamma(z_1, z_2) = \frac{\sigma^2}{n}\mathrm{Tr}\mathrm{T}_{z_1}\mathrm{T}_{z_2}$. Then*

$$\begin{aligned}
\zeta_{kj} &= \psi_j\Big(\sum_{\ell=1}^{j} \frac{\sigma \mathbf{a}_k^*\mathrm{T}_{z_1}\mathbf{a}_\ell\mathbf{a}_\ell^*\mathrm{T}_{z_2}\mathbf{a}_k}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})} + \sigma^3\mathbf{a}_k^*\mathrm{T}_{z_1}\mathrm{T}_{z_2}\mathbf{a}_k\frac{1}{n}\sum_{\ell=1}^{j} z_1z_2\tilde{t}_{z_1,\ell\ell}\tilde{t}_{z_2,\ell\ell}\Big)\\
&\quad +\sigma\mathbf{a}_k^*\mathrm{T}_{z_1}\mathrm{T}_{z_2}\mathbf{a}_k + \sigma\mathbf{a}_k^*\mathrm{T}_{z_1}\mathrm{T}_{z_2}\mathbf{a}_k\phi_j + \mathcal{O}_{z_1,z_2}(n^{-1/2}),\\
\zeta_{kj} &= z_1z_2\tilde{t}_{z_1,kk}\tilde{t}_{z_2,kk}(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})\theta_{kj} + \frac{\sigma\mathbf{a}_k^*\mathrm{T}_{z_1}\mathbf{a}_k\mathbf{a}_k^*\mathrm{T}_{z_2}\mathbf{a}_k}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})}\psi_j + \mathcal{O}_{z_1,z_2}(n^{-1/2}),\\
\psi_j &= \psi_j\Big(\frac{\sigma^2}{n}\sum_{\ell=1}^{j}\frac{\mathbf{a}_\ell^*\mathrm{T}_{z_2}\mathrm{T}_{z_1}\mathbf{a}_\ell}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})} + \frac{\gamma\sigma^2}{n}\sum_{\ell=1}^{j} z_1z_2\tilde{t}_{z_1,\ell\ell}\tilde{t}_{z_2,\ell\ell}\Big) + \frac{\sigma^2}{n}\mathrm{Tr}\mathrm{T}_{z_1}\mathrm{T}_{z_2} + \gamma\phi_j + \mathcal{O}_{z_1,z_2}(n^{-1/2}).
\end{aligned}$$

Lemma 3.9 is a generalization of computations performed in [18, Section 5] for $z_1 = z_2 \in (-\infty, 0)$ to general $z_1, z_2 \in \mathbb{C}^+$. Its proof is omitted.

Combining the two first equations of Lemma 3.9, we get

$$\begin{aligned}
\sigma z_1z_2\tilde{t}_{z_1,kk}\tilde{t}_{z_2,kk}\theta_{kj} &= \frac{\sigma^2\mathbf{a}_k^*\mathrm{T}_{z_1}\mathrm{T}_{z_2}\mathbf{a}_k}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})} + \frac{\sigma^2\mathbf{a}_k^*\mathrm{T}_{z_1}\mathrm{T}_{z_2}\mathbf{a}_k}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})}\phi_j\\
&\quad + \psi_j\left(\sum_{\ell=1}^{j}\frac{\sigma^2\mathbf{a}_k^*\mathrm{T}_{z_1}\mathbf{a}_\ell\mathbf{a}_\ell^*\mathrm{T}_{z_2}\mathbf{a}_k}{(1+\sigma\delta_{z_1})^2(1+\sigma\delta_{z_2})^2} + \frac{\sigma^2\mathbf{a}_k^*\mathrm{T}_{z_1}\mathrm{T}_{z_2}\mathbf{a}_k}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})}\frac{\sigma^2}{n}\sum_{\ell=1}^{j} z_1z_2\tilde{t}_{z_1,\ell\ell}\tilde{t}_{z_2,\ell\ell}\right)\\
&\quad - \psi_j\frac{\sigma^2\mathbf{a}_k^*\mathrm{T}_{z_1}\mathbf{a}_k\mathbf{a}_k^*\mathrm{T}_{z_2}\mathbf{a}_k}{(1+\sigma\delta_{z_1})^2(1+\sigma\delta_{z_2})^2} + \mathcal{O}_{z_1,z_2}(n^{-1/2}). \qquad (3.24)
\end{aligned}$$

In order to simplify the notations, we introduce the following quantities:

$$\begin{aligned}
\nu_j &:= \nu_j(z_1,z_2) = \frac{\sigma^2}{n}\sum_{k=1}^{j}\frac{a_k^*\mathrm{T}_{z_1}\mathrm{T}_{z_2}a_k}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})},\\
\eta_j &:= \eta_j(z_1,z_2) = \frac{\sigma^2}{n}\sum_{\ell=1}^{j} z_1z_2\tilde{t}_{z_1,\ell\ell}\tilde{t}_{z_2,\ell\ell},\\
\omega_j &:= \omega_j(z_1,z_2) = \frac{\sigma^2}{n}\sum_{k=1}^{j}\sum_{\ell=1,\ell\neq k}^{j}\frac{\mathbf{a}_k^*\mathrm{T}_{z_1}\mathbf{a}_\ell\mathbf{a}_\ell^*\mathrm{T}_{z_2}\mathbf{a}_k}{(1+\sigma\delta_{z_1})^2(1+\sigma\delta_{z_2})^2}.
\end{aligned}$$

Notice that for $j = n$ the notation ν_j is consistant with the definition (2.2). Eq. (3.24) yields

$$(1-\nu_j)\phi_j - (\omega_j + \nu_j\eta_j)\psi_j = \nu_j + \mathcal{O}_{z_1,z_2}(n^{-1/2})$$

and the last equation of Lemma 3.9 writes

$$-\gamma\phi_j + (1-\nu_j-\gamma\eta_j)\psi_j = \gamma + \mathcal{O}_{z_1,z_2}(n^{-1/2}).$$

We finally end up with a system of two perturbed linear equations for ϕ_j and ψ_j:

$$\left\{\begin{array}{lcl} (1-\nu_j)\phi_j - (\omega_j+\nu_j\eta_j)\psi_j &=& \nu_j + \mathcal{O}_{z_1,z_2}(n^{-1/2}),\\ -\gamma\phi_j + (1-\nu_j-\gamma\eta_j)\psi_j &=& \gamma + \mathcal{O}_{z_1,z_2}(n^{-1/2}). \end{array}\right. \qquad (3.25)$$

We study hereafter the properties of the determinant of the system $\mathcal{D}_j$ given by

$$\mathcal{D}_j = (1-\nu_j)^2 - \gamma(\eta_j + \omega_j). \qquad (3.26)$$

Lemma 3.10. *Let Assumption 2 hold and recall the definition of Δ_n given in (2.3). The determinant $\mathcal{D}_j$ satisfies the following properties:*

(i) for any $z_1, z_2 \in \mathbb{C}^+$, we have for $j = n$

$$\mathcal{D}_n(z_1,z_2) = \Delta_n(z_1,z_2) = \left(1 - \frac{\sigma^2}{n}\frac{\mathrm{Tr}\mathrm{T}_{z_1}AA^*\mathrm{T}_{z_2}}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})}\right)^2 - z_1z_2\gamma\tilde{\gamma},$$

(ii) for any $z_1, z_2 \in \mathbb{C}^+$,

$$\liminf_n \inf_{1\le j\le n} |\mathcal{D}_j(z_1,z_2)| > 0.$$

Proof. For $1 \le j \le n$ denote by $A_{1:j}$ the $N \times n$ matrix defined by $A_{1:j} := [\mathbf{a}_1, \dots, \mathbf{a}_j, \mathbf{0}, \dots, \mathbf{0}]$ and write

$$\omega_j = \frac{\sigma^2}{n}\frac{\mathrm{Tr}\mathrm{A}_{1:j}^*\mathrm{T}_{z_1}\mathrm{A}_{1:j}\mathrm{A}_{1:j}^*\mathrm{T}_{z_2}A_{1:j}}{(1+\sigma\delta_{z_1})^2(1+\sigma\delta_{z_2})^2} - \frac{\sigma^2}{n}\sum_{k=1}^{j}\frac{\mathbf{a}_k^*\mathrm{T}_{z_1}\mathbf{a}_k\mathbf{a}_k^*\mathrm{T}_{z_2}\mathbf{a}_k}{(1+\sigma\delta_{z_1})^2(1+\sigma\delta_{z_2})^2}.$$

Using a standard identity [24, Section 0.7.4] applied to T_z and $\tilde{\mathrm{T}}_z$ yields the identity

$$\tilde{\mathrm{T}}_z = -\frac{1}{z(1+\sigma\delta_z)} + \frac{1}{z(1+\sigma\delta_z)^2}\mathrm{A}^*\mathrm{T}_z\mathrm{A} \tag{3.27}$$

from which we obtain $(1+\sigma\delta_z)^{-2}\mathbf{a}_k^*\mathrm{T}_z\mathbf{a}_k = (1+\sigma\delta_z)^{-1} + z\tilde{t}_{kk}$ and thus

$$\frac{\sigma^2}{n}\sum_{k=1}^{j}\frac{\mathbf{a}_k^*\mathrm{T}_{z_1}\mathbf{a}_k\mathbf{a}_k^*\mathrm{T}_{z_2}\mathbf{a}_k}{(1+\sigma\delta_{z_1})^2(1+\sigma\delta_{z_2})^2} = \frac{\sigma^2}{n}\sum_{k=1}^{j} z_1 z_2 \tilde{t}_{z_1,kk}\tilde{t}_{z_2,kk}$$
$$+\frac{\sigma^2}{n}\sum_{k=1}^{j}\frac{z_1\tilde{t}_{z_1,kk}}{1+\sigma\delta_{z_2}} + \frac{\sigma^2}{n}\sum_{k=1}^{j}\frac{\mathbf{a}_k^*\mathrm{T}_{z_2}\mathbf{a}_k}{(1+\sigma\delta_{z_1})(1+\sigma\delta_{z_2})^2}\,.$$

Introduce

$$-z\tilde{\mathrm{T}}_{1:j}(z) := (1+\sigma\delta_z)^{-1}\mathrm{I}_{1:j} - (1+\sigma\delta_z)^{-2}A_{1:j}^*\mathrm{T}_z A_{1:j}$$

where $\mathrm{I}_{1:j} := \sum_{k=1}^{j} e_k e_k^{\mathsf{T}}$ and (e_i) is the canonical basis of $\mathbb{R}^n$, one can check that

$$\eta_j + \omega_j = \frac{\sigma^2 z_1 z_2}{n}\mathrm{Tr}\big(\tilde{\mathrm{T}}_{1:j}(z_1)\tilde{\mathrm{T}}_{1:j}(z_2)\big)\,. \tag{3.28}$$

In particular $\eta_n + \omega_n = z_1 z_2 \tilde{\gamma}$ hence the identity $\mathcal{D}_n = (1-\nu_n)^2 - z_1 z_2\gamma\tilde{\gamma} = \Delta_n$ and (i) is established.

We now prove (ii) and start by showing that for any $z\in\mathbb{C}^+$,

$$\liminf_n \inf_{1\le j\le n} \mathcal{D}_j(z,\bar{z}) > 0\,. \tag{3.29}$$

It is straightforward to check that

$$0 < \nu_j(z,\bar{z}) \le \nu_n(z,\bar{z}) \overset{(a)}{<} 1 \qquad \text{and} \qquad 0 \le \omega_j(z,\bar{z}) + \eta_j(z,\bar{z}) \le \omega_n(z,\bar{z}) + \eta_n(z,\bar{z})\,,$$

where (a) follows from Proposition 3.1-(3). Hence $\mathcal{D}_i(z,\bar{z}) \ge \mathcal{D}_n(z,\bar{z})$. Since by (i) we have proved that

$$\mathcal{D}_n(z,\bar{z}) \;=\; \Delta_n(z,\bar{z}) \;=\; (1-\nu_n(z,\bar{z}))^2 - \gamma(z,\bar{z})(\eta_n(z,\bar{z}) + \omega_n(z,\bar{z}))\,,$$

we obtain the following estimate

$$\inf_{1\le j\le n} \mathcal{D}_j(z,\bar{z}) \;\ge\; \Delta_n(z,\bar{z})\,. \tag{3.30}$$

Recall that δ_n and $\tilde{\delta}_n$ are Stieltjes transforms associated to measures with respective total mass $\sigma N n^{-1}$ and σ hence

$$|s_n(z)| \;\le\; |z|\left(1+\frac{N\sigma^2}{n\mathrm{Im}(z)}\right)\left(1+\frac{\sigma^2}{\mathrm{Im}(z)}\right) \;\le\; K_z$$

uniformly in $n\ge 1$. By Proposition 3.1-(3),

$$|\Delta_n(z,\bar{z})| = \frac{\mathrm{Im}(z)}{\mathrm{Im}(s_n(z))} \;\ge\; \frac{\mathrm{Im}(z)}{K_z}\,.$$

Combining this estimate with (3.30) yields (3.29). To conclude the proof, we show that for $z_1, z_2 \in \mathbb{C}^+$,

$$|\mathcal{D}_j(z_1,z_2)| \ge (\mathcal{D}_j(z_1,\bar{z}_1)\mathcal{D}_j(z_2,\bar{z}_2))^{1/2}\,. \tag{3.31}$$

Starting from (3.28), we have

$$\mathcal{D}_j(z_1,z_2) \;=\; (1-\nu_j(z_1,z_2))^2 - \gamma z_1 z_2 \frac{\sigma^2}{n}\mathrm{Tr}\,\tilde{\mathrm{T}}_{1:j}(z_1)\tilde{\mathrm{T}}_{1:j}^*(z_2)\,.$$

From this identity, we can conclude as in the proof of Proposition 3.1-(6). □

We now go back to the proof of Lemma 3.7. By the above lemma, the system (3.25) has the solution

$$\begin{bmatrix}\phi_j\\ \psi_j\end{bmatrix} = \frac{1}{\mathcal{D}_j}\begin{bmatrix}\nu_j(1-\nu_j-\gamma\eta_j)+\gamma(\omega_j+\nu_j\eta_j)\\ \gamma\nu_j+\gamma(1-\nu_j)\end{bmatrix} + \mathcal{O}_{z_1,z_2}(n^{-1/2})\,. \tag{3.32}$$

Notice that since T_{z_1} and T_{z_2} commute $\nu(z_1,z_2) = \nu(z_2,z_1)$. Dividing by n and plugging the solution (3.32) in (3.24) yields:

$$\frac{\sigma}{n} z_1 z_2 \tilde{t}_{z_1,jj}\tilde{t}_{z_2,jj}\big(\theta_{jj}(z_1,z_2) + \theta_{jj}(z_2,z_1)\big)$$
$$= 2(\nu_j - \nu_{j-1}) + \frac{\gamma}{\mathcal{D}_j}\big((\omega_j - \omega_{j-1}) + 2\eta_j(\nu_j - \nu_{j-1})\big)$$
$$+\frac{1}{\mathcal{D}_j}(\nu_j - \nu_{j-1})\big(\nu_j(1-\nu_j) + \gamma\omega_j\big) + \mathcal{O}_{z_1,z_2}(n^{-3/2})\,,$$
$$= \frac{1}{\mathcal{D}_j}\big(2(\nu_j - \nu_{j-1})(1-\nu_j) + \gamma(\omega_j - \omega_{j-1})\big) + \mathcal{O}_{z_1,z_2}(n^{-3/2})\,,$$

where by convention we set $\omega_0 = \nu_0 = \eta_0 = 0$ and $\mathcal{D}_0 = 1$. Therefore, we get

$$\frac{1}{n} z_1 z_2 \tilde{t}_{z_1,jj} \tilde{t}_{z_2,jj} \big(\sigma^2 \psi_j(z_1,z_2) + \sigma\theta_{jj}(z_1,z_2) + \sigma\theta_{jj}(z_2,z_1)\big) = \frac{2(\nu_j - \nu_{j-1})(1-\nu_j) + \gamma(\eta_j - \eta_{j-1}) + \gamma(\omega_j - \omega_{j-1})}{\mathcal{D}_j} + \mathcal{O}_{z_1,z_2}(n^{-3/2}).$$

Moreover, going back to the definition (3.26) of $\mathcal{D}_j$, we have

$$\begin{aligned} \mathcal{D}_{j-1} - \mathcal{D}_j &= 2(\nu_j - \nu_{j-1})(2 - \nu_{j-1} - \nu_j) + \gamma(\eta_j - \eta_{j-1}) + \gamma(\omega_j - \omega_{j-1}), \\ &= 2(\nu_j - \nu_{j-1})(1 - \nu_j) + \gamma(\eta_j - \eta_{j-1}) + \gamma(\omega_j - \omega_{j-1}) + \mathcal{O}_{z_1,z_2}(n^{-2}). \end{aligned}$$

Then

$$\frac{1}{n}\sum_{j=1}^n z_1 z_2 \tilde{t}_{z_1,jj}\tilde{t}_{z_2,jj}\big(\sigma^2\psi_j(z_1,z_2)+\sigma\theta_{jj}(z_1,z_2)+\sigma\theta_{jj}(z_2,z_1)\big) = \sum_{j=1}^n \frac{\mathcal{D}_{j-1}-\mathcal{D}_j}{\mathcal{D}_j} + \mathcal{O}_{z_1,z-2}(n^{-1/2}), \quad (3.33)$$

and

$$\mathcal{D}_0 = 1 \qquad \text{and} \qquad |\mathcal{D}_{j-1} - \mathcal{D}_j| = \mathcal{O}_{z_1,z_2}(n^{-1}) \qquad \text{for} \quad 1 \le j \le n. \quad (3.34)$$

For a sufficiently large fixed constant K and for any $1 \le j \le n$, we denote by $\mathcal{B}_j := B(\mathcal{D}_j, K/n)$ the ball of center $\mathcal{D}_j$ and radius K/n and we let $[\mathcal{D}_j, \mathcal{D}_{j-1}] \subset \mathcal{B}_j$ be the segment joining $\mathcal{D}_j$ and $\mathcal{D}_{j-1}$. We suppose that n is large enough so that $K/n < |\mathcal{D}_j|/2$. Thus for any $z \in [\mathcal{D}_j, \mathcal{D}_{j-1}]$,

$$\frac{|\mathcal{D}_j|}{2} < |z| \le |\mathcal{D}_j| + \frac{K}{n} \quad \text{and} \quad |z - \mathcal{D}_j| \le \frac{K}{n}. \quad (3.35)$$

As $z \mapsto z^{-1}$ is analytic over $\mathcal{B} := \cup_{j=1}^n \mathcal{B}_j$, we write

$$\frac{\mathcal{D}_{j-1}-\mathcal{D}_j}{\mathcal{D}_j} - \int_{[\mathcal{D}_j,\mathcal{D}_{j-1}]} \frac{1}{z}\mathrm{d}z = \sum_{j=1}^n \int_{[\mathcal{D}_j,\mathcal{D}_{j-1}]} \Big(\frac{1}{\mathcal{D}_j} - \frac{1}{z}\Big)\mathrm{d}z = \int_{[\mathcal{D}_j,\mathcal{D}_{j-1}]} \frac{z-\mathcal{D}_j}{z\mathcal{D}_j}\mathrm{d}z = \mathcal{O}_{z_1,z_2}(n^{-2})$$

where the last equality follows from (3.34) and (3.35). We finally obtain

$$\left| \sum_{j=1}^n \frac{\mathcal{D}_{j-1}-\mathcal{D}_j}{\mathcal{D}_j} - \int_{\Delta_n}^1 \frac{1}{z}\mathrm{d}z \right| = \mathcal{O}_{z_1,z_2}(n^{-1}).$$

Using Lemma 3.10-(ii), one can prove that the r.h.s. above is uniformly bounded and apply [4, Lemma 2.14] to obtain the convergence of the derivative. Differentiating with respect to z_2, we get

$$\frac{\partial}{\partial z_2} \int_{\Delta_n}^1 \frac{1}{z}\mathrm{d}z = -\frac{1}{\Delta_n}\frac{\partial \Delta_n}{\partial z_2}.$$

Differentiating again with respect to z_1 and relying on the identity of Proposition 3.1-(4), we conclude the proof of Lemma 3.7.

4. Proof of Theorem 3: Alternative expression for the covariance

We first recall some properties of s_n useful in the sequel. Recall that $\mathcal{S}_n$ is the support of the measure whose Stieltjes transform is δ_n and denote by $\mathcal{S}_{\boldsymbol{A}}$ the support of the empirical distribution of the eigenvalues of $A_n A_n^*$.

Proposition 4.1 (Properties of s_n near the real axis). *The following properties hold*

1. *The limit $s_n(x) := \lim_{\varepsilon\downarrow 0} s_n(x + \boldsymbol{i}\varepsilon)$ exists and is continuous for all $x \in \mathbb{R} \setminus 0$.*
2. *$x \in \mathcal{S}_n^c$ implies that $s_n(x) \in \mathcal{S}_{\boldsymbol{A}}^c$.*
3. *If $c_n = \frac{N}{n} < 1$ then $0 \in \mathcal{S}_n^c$.*
4. *The quantity $s_n(z)$ is bounded for $|z| < \eta$ for some $\eta > 0$ and $z \neq 0$.*

Proof. Items (1), (2) follow from Theorems 2.1 and 3.3 in [12]. Item (3) can be found in [8, Theorem 1.3]. To prove item (4), write

$$s_n(z) = z(1 + \sigma\delta_n(z))(1 + \sigma\tilde{\delta}_n(z)) = z(1 + \sigma\delta_n(z))^2 - \sigma(1 - c_n)(1 + \sigma\delta_n(z)).$$

The first part of the r.h.s. is bounded by Proposition 3.1-(5). If $c_n = 1$, the second part vanishes; if $c_n < 1$ then $0 \notin \mathcal{S}_n$ and $\delta_n(z)$ is analytic in a small neighbourhood of zero. □

4.1. **A boundary value representation for the covariance.**

Proposition 4.2. *Let $(Z_n(f), Z_n(g))$ be the Gaussian process defined in Theorem 2, and Θ_n the covariance defined in Theorem 1, then the covariance of $(Z_n(f), Z_n(g))$ admits the following representation:*

$$\begin{aligned}
\mathrm{cov}(Z_n(f), Z_n(g)) &= -\frac{1}{4\pi^2}\lim_{\varepsilon\searrow 0}\int_{\mathbb{R}^2} f(x)g(y)\big(\Theta_n(x+\boldsymbol{i}\varepsilon, y+\boldsymbol{i}\varepsilon)+\Theta_n(x-\boldsymbol{i}\varepsilon, y-\boldsymbol{i}\varepsilon)\big)dxdy\\
&\quad +\frac{1}{4\pi^2}\lim_{\varepsilon\searrow 0}\int_{\mathbb{R}^2} f(x)g(y)\big(\Theta_n(x+\boldsymbol{i}\varepsilon, y-\boldsymbol{i}\varepsilon)+\Theta_n(x-\boldsymbol{i}\varepsilon, y+\boldsymbol{i}\varepsilon)\big)dxdy\ ,\\
&= -\frac{1}{4\pi^2}\lim_{\varepsilon\searrow 0}\sum_{\pm_1,\pm_2}(\pm_1\pm_2)\int_{\mathbb{R}^2} f(x)g(y)\Theta_n(x\pm_1\boldsymbol{i}\varepsilon, y\pm_2\boldsymbol{i}\varepsilon)dxdy\ ,
\end{aligned}$$

where $\pm_1, \pm_2 \in \{+,-\}$ and $\pm_1\pm_2$ is the sign resulting from the product $\pm_1 1$ by $\pm_2 1$.

For a proof, see [32, Proposition 4.1].

4.2. **Proof of Theorem 3.** Notice that due to the symmetry of equations (1.9), we only need to consider the case where $c \le 1$, which we now assume. Recall the definition of the quantity

$$\underline{\Delta}_n(x,y) = \lim_{\varepsilon\downarrow 0}\Delta_n(x+\boldsymbol{i}\varepsilon, y-\boldsymbol{i}\varepsilon)\,.$$

The covariance $\Theta_n(z_1,z_2)$ splits into three parts $\Theta_n(z_1,z_2) = \Theta_{0,n}(z_1,z_2)+\Theta_{1,n}(z_1,z_2)+\Theta_{2,n}(z_1,z_2)$, cf. (2.7). We first prove that

$$-\frac{1}{4\pi^2}\lim_{\varepsilon\searrow 0}\sum_{\pm_1,\pm_2}(\pm_1\pm_2)\int_{\mathbb{R}^2} f(x)g(y)\Theta_{0,n}(x\pm_1\boldsymbol{i}\varepsilon, y\pm_2\boldsymbol{i}\varepsilon)dxdy = \frac{1}{2\pi^2}\int_{\mathcal{S}_n^2} f'(x)g'(y)\ln\left|\frac{\underline{\Delta}_n(x,y)}{\Delta_n(x,y)}\right| dxdy$$

Taking advantage of formula (2.8) and performing a double integration by parts yields

$$\begin{aligned}
&-\frac{1}{4\pi^2}\lim_{\varepsilon\searrow 0}\sum_{\pm_1,\pm_2}(\pm_1\pm_2)\int_{\mathbb{R}^2} f(x)g(y)\Theta_{0,n}(x\pm_1\boldsymbol{i}\varepsilon, y\pm_2\boldsymbol{i}\varepsilon)dxdy\\
&= \frac{1}{4\pi^2}\lim_{\varepsilon\searrow 0}\sum_{\pm_1,\pm_2}(\pm_1\pm_2)\int_{\mathbb{R}^2} f(x)g(y)\frac{\partial}{\partial y}\left\{\frac{1}{\Delta_n(x\pm_1\boldsymbol{i}\varepsilon, y\pm_2\boldsymbol{i}\varepsilon)}\frac{\partial}{\partial x}\Delta_n(x\pm_1\boldsymbol{i}\varepsilon, y\pm_2\boldsymbol{i}\varepsilon)\right\}dxdy\ ,\\
&= -\frac{1}{4\pi^2}\lim_{\varepsilon\searrow 0}\sum_{\pm_1,\pm_2}(\pm_1\pm_2)\int_{\mathbb{R}^2} f(x)g'(y)\frac{1}{\Delta_n(x+\boldsymbol{i}\varepsilon, y+\boldsymbol{i}\varepsilon)}\frac{\partial}{\partial x}\Delta_n(x+\boldsymbol{i}\varepsilon, y+\boldsymbol{i}\varepsilon)dxdy\ ,\\
&\overset{(a)}{=} \frac{1}{4\pi^2}\lim_{\varepsilon\searrow 0}\sum_{\pm_1,\pm_2}(\pm_1\pm_2)\int_{\mathbb{R}^2} f'(x)g'(y)\log\left(\Delta_n(x\pm_1\boldsymbol{i}\varepsilon, y\pm_2\boldsymbol{i}\varepsilon)\right)dxdy\ ,\\
&\overset{(b)}{=} \frac{1}{4\pi^2}\lim_{\varepsilon\searrow 0}\sum_{\pm_1,\pm_2}(\pm_1\pm_2)\int_{\mathbb{R}^2} f'(x)g'(y)\ln\left|\Delta_n(x\pm_1\boldsymbol{i}\varepsilon, y\pm_2\boldsymbol{i}\varepsilon)\right|dxdy\ ,\\
&\overset{(c)}{=} \frac{1}{2\pi^2}\lim_{\varepsilon\searrow 0}\int_{\mathbb{R}^2} f'(x)g'(y)\ln\left|\frac{\Delta_n(x+\boldsymbol{i}\varepsilon, y+\boldsymbol{i}\varepsilon)}{\Delta_n(x+\boldsymbol{i}\varepsilon, y-\boldsymbol{i}\varepsilon)}\right|dxdy\ ,\\
&\overset{(d)}{=} \frac{1}{2\pi^2}\lim_{\varepsilon\searrow 0}\int_{\mathbb{R}^2} f'(x)g'(y)\ln\left|\frac{s_n(x+\boldsymbol{i}\varepsilon)-s_n(y-\boldsymbol{i}\varepsilon)}{s_n(x+\boldsymbol{i}\varepsilon)-s_n(y+\boldsymbol{i}\varepsilon)}\right|dxdy\ ,
\end{aligned}$$

where $\log(\cdot)$ is any branch of the complex logarithm in (a), where (b) follows from the fact that the covariance being real, the argument part of the complex logarithm necessarily vanishes and where (c) and (d) follow from the representation formula for Δ_n (cf. Proposition 3.1-(3)) and the fact that

$$\overline{\Delta_n(z_1,z_2)} = \frac{\overline{z_1}-\overline{z_2}}{s_n(\overline{z_1})-s_n(\overline{z_2})} = \Delta_n(\overline{z_1},\overline{z_2})\ .$$

Write now

$$\begin{aligned}
\ln\left|\frac{s_n(x+\boldsymbol{i}\varepsilon)-s_n(y-\boldsymbol{i}\varepsilon)}{s_n(x+\boldsymbol{i}\varepsilon)-s_n(y+\boldsymbol{i}\varepsilon)}\right| &= \frac{1}{2}\ln\left|\frac{s_n(x+\boldsymbol{i}\varepsilon)-s_n(y-\boldsymbol{i}\varepsilon)}{s_n(x+\boldsymbol{i}\varepsilon)-s_n(y+\boldsymbol{i}\varepsilon)}\right|^2\ ,\\
&= \frac{1}{2}\ln\left(1+\frac{4\,\mathrm{Im}\,s_n(x+\boldsymbol{i}\varepsilon)\mathrm{Im}\,s_n(y+\boldsymbol{i}\varepsilon)}{|s_n(x+\boldsymbol{i}\varepsilon)-s_n(y+\boldsymbol{i}\varepsilon)|^2}\right)\ .
\end{aligned}$$

In order to apply the dominated convergence theorem, we need to majorize the right hand side above by an integrable function of $(x,y)\in[-K,K]^2$ where K is sufficiently large to contain the supports of functions f

and g. Let $\varepsilon_0 > 0$. Function s being continuous on a rectangle $[0, K] \times [0, \varepsilon_0]$, it is bounded. In particular, $\mathrm{Im}(s(x + \boldsymbol{i}\varepsilon)) \le |s(x + \boldsymbol{i}\varepsilon)|$ is also bounded on $[0, K] \times [0, \varepsilon_0]$.

Let $z_1 = x + \boldsymbol{i}\varepsilon$ and $z_2 = y + \boldsymbol{i}\varepsilon$. Then by the definition (2.3) of Δ_n,

$$\begin{aligned} |\Delta_n(z_1, z_2)| &\le \left(1 + \sqrt{\nu(z_1, \overline{z_1})}\sqrt{\nu(z_2, \overline{z_2})}\right)^2 + \sqrt{|z_1|^2\gamma(z_1, \overline{z_1})\tilde{\gamma}(z_1, \overline{z_1})|}\sqrt{|z_2|^2\gamma(z_2, \overline{z_2})\tilde{\gamma}(z_2, \overline{z_2})|} \\ &\le 5 \end{aligned}$$

by Proposition 3.1-(4). By the representation of Δ_n provided in Proposition 3.1-(3), we have

$$|\Delta_n(z_1, z_2)|^2 = \frac{|z_1 - z_2|^2}{|s(z_1) - s(z_2)|^2} \le 25 \qquad \Rightarrow \qquad \frac{1}{|s(z_1) - s(z_2)|^2} \le \frac{25}{|z_1 - z_2|^2} = \frac{25}{|x - y|^2}\,.$$

In the end,

$$\ln\left|\frac{s_n(x + \boldsymbol{i}\varepsilon) - s_n(y - \boldsymbol{i}\varepsilon)}{s_n(x + \boldsymbol{i}\varepsilon) - s_n(y + \boldsymbol{i}\varepsilon)}\right| \le \ln\left(1 + \frac{K'}{|x - y|^2}\right)$$

which is integrable. It remains to apply the dominated convergence theorem to conclude and obtain (**??**).

We now prove that

$$\begin{aligned} &-\frac{1}{4\pi^2}\lim_{\varepsilon \searrow 0}\sum_{\pm_1, \pm_2}(\pm_1\pm_2)\int_{\mathbb{R}^2} f(x)g(y)\Theta_{2,n}(x \pm_1 \boldsymbol{i}\varepsilon, y \pm_2 \boldsymbol{i}\varepsilon)dxdy \\ &\qquad = \frac{4\sigma^4\kappa}{\pi^2 n^2}\sum_{i=1}^N\sum_{j=1}^n \int_{\mathcal{S}_n} f'(x)\mathrm{Im}\left(x\, t_{ii}(x)\tilde{t}_{jj}(x)\right)\, dx \int_{\mathcal{S}_n} g'(y)\mathrm{Im}\left(y\, t_{ii}(y)\tilde{t}_{jj}(y)\right)\, dy\,. \end{aligned}$$

By the mere definition of $\Theta_{2,n}$, we have

$$\begin{aligned} &-\frac{1}{4\pi^2}\lim_{\varepsilon \searrow 0}\sum_{\pm_1, \pm_2}(\pm_1\pm_2)\int_{\mathbb{R}^2} f(x)g(y)\Theta_{2,n}(x \pm_1 \boldsymbol{i}\varepsilon, y \pm_2 \boldsymbol{i}\varepsilon)dxdy \\ &= -\frac{\kappa\sigma^4}{4n^2\pi^2}\lim_{\varepsilon \searrow 0}\sum_{\pm_1, \pm_2}(\pm_1\pm_2)\sum_{i,j}\int_{\mathbb{R}^2} f(x)g(y)\frac{\partial}{\partial x}(x \pm_1 \boldsymbol{i}\varepsilon)t_{ii}(x \pm_1 \boldsymbol{i}\varepsilon)\tilde{t}_{jj}(x \pm_1 \boldsymbol{i}\varepsilon) \\ &\qquad\qquad \times\frac{\partial}{\partial y}(y \pm_2 \boldsymbol{i}\varepsilon)t_{ii}(y \pm_2 \boldsymbol{i}\varepsilon)\tilde{t}_{jj}(y \pm_2 \boldsymbol{i}\varepsilon)dxdy\,, \\ &= \frac{\kappa\sigma^4}{n^2\pi^2}\lim_{\varepsilon \searrow 0}\sum_{i,j}\int_{\mathbb{R}} f'(x)\mathrm{Im}\left\{(x + \boldsymbol{i}\varepsilon)t_{ii}(x + \boldsymbol{i}\varepsilon)\tilde{t}_{jj}(x + \boldsymbol{i}\varepsilon)\right\}\, dx \\ &\qquad\qquad \times\int_{\mathbb{R}} g'(y)\mathrm{Im}\left\{(y + \boldsymbol{i}\varepsilon)t_{ii}(y + \boldsymbol{i}\varepsilon)\tilde{t}_{jj}(y + \boldsymbol{i}\varepsilon)\right\} dy\,. \end{aligned}$$

It remains to prove that

$$\lim_{\varepsilon \searrow 0}\int_{\mathbb{R}} f'(x)\mathrm{Im}\left\{(x + \boldsymbol{i}\varepsilon)t_{ii}(x + \boldsymbol{i}\varepsilon)\tilde{t}_{jj}(x + \boldsymbol{i}\varepsilon)\right\}\, dx = \int_{\mathbb{R}} f'(x)\mathrm{Im}\left\{x\, t_{ii}(x)\tilde{t}_{jj}(x)\right\}\, dx\,. \tag{4.1}$$

Let $x > 0$, then by [12, Theorem 2.1]

$$(x + \boldsymbol{i}\varepsilon)t_{ii}(x + \boldsymbol{i}\varepsilon)\tilde{t}_{jj}(x + \boldsymbol{i}\varepsilon) \xrightarrow[\varepsilon \searrow 0]{} x\, t_{ii}(x)\tilde{t}_{jj}(x)\ .$$

In order to apply the dominated convergence theorem, we handle separately the cases $c_n = 1$ and $c_n < 1$.

If $c_n = 1$, then $\delta_z = \tilde{\delta}_z$ (apply Prop. 3.1-(2) for instance), $\|\mathrm{T}_z\| = \|\tilde{\mathrm{T}}_z\|$ and

$$\|\mathrm{T}_z\| \le \frac{\sqrt{n}}{\sigma\sqrt{|z|}} \qquad \text{and} \qquad \|\tilde{\mathrm{T}}_z\| \le \frac{\sqrt{n}}{\sigma\sqrt{|z|}} \quad \text{for all } z \in \mathbb{C}^+\,.$$

In fact,

$$\|\mathrm{T}_z\| = \sqrt{\lambda_{\max}(\mathrm{T}_z\mathrm{T}_z^*)} \le \sqrt{\mathrm{Tr}\,\mathrm{T}_z\mathrm{T}_z^*} = \sqrt{n\sigma^{-2}\gamma(z, \bar{z})} \le \frac{\sqrt{n}}{\sigma\sqrt{|z|}}\,,$$

where the last inequality follows from Proposition 3.1-6. Now

$$|zt_{ii}(z)\tilde{t}_{jj}(z)| \le |z|\,\|\mathrm{T}_z\|\,\|\tilde{\mathrm{T}}_z\| \le \frac{n}{\sigma^2}\,.$$

We therefore apply the dominated convergence theorem and prove (4.1) in the case where $c_n = 1$.

Assume now that $c_n < 1$ then $0 \notin \mathcal{S}_n$, where $\mathcal{S}_n$ denotes the support of the measure associated to the Stieltjes transform δ_n. In particular, there exists $\eta > 0$ such that $(-\eta, \eta) \cap \mathcal{S}_n = \emptyset$. In the sequel, we will alternatively bound $|zt_{ii}(z)\tilde{t}_{jj}(z)|$ on the sets

$$\mathcal{D} = \{|z| \le \eta/2\} \qquad \text{and} \qquad \mathcal{D}'_A = \{|z| > \eta/2\} \cap [\eta/2, A] \times [0, A]\,,$$

where $A>0$ is an arbitrary constant. This will enable us to apply the dominated convergence theorem and prove (4.1).

Let $z\in\mathcal{D}$. One has

$$\delta_z=\frac{\sigma}{n}\sum_{i=1}^N t_{ii}(z)\qquad\text{and}\qquad \tilde\delta_z=\frac{\sigma}{n}\sum_{j=1}^n \tilde t_{jj}(z)$$

hence $\mathrm{Im}(t_{ii}(z))\le\frac{n}{\sigma}\mathrm{Im}(\delta_z)$ and $\mathrm{Im}(\tilde t_{jj}(z))\le\frac{n}{\sigma}\mathrm{Im}(\tilde\delta_z)$. We deduce from these inequalities that the probability measure μ_{ii} associated to the Stieltjes transform t_{ii} has a support included in $\mathcal{S}_n$ hence

$$|t_{ii}(z)|=\left|\int_{\mathcal{S}_n}\frac{\mu_{ii}(d\lambda)}{\lambda-z}\right|\le\frac{2}{\eta}\qquad\text{for } z\in\mathcal{D}\,.$$

By Proposition 3.1-(2), the support of the measure associated to $\tilde\delta_z$ is $\{0\}\cup\mathcal{S}_n$. Let $\tilde\mu_{jj}$ be the probability distribution with Stieltjes transform $\tilde t_{jj}$, then $\mathrm{supp}(\tilde\mu_{jj})\subset\{0\}\cup\mathcal{S}_n$. Otherwise stated, $\tilde\mu_{jj}$ has a Dirac component at zero and a component $\check\mu_{jj}$ with support included in $\mathcal{S}_n$ hence

$$\tilde t_{jj}(z)=-\frac{\alpha_j}{z}+\int_{\mathcal{S}_n}\frac{\check\mu_{jj}(d\lambda)}{\lambda-z}\qquad\text{and}\qquad\left|z\tilde t_{jj}(z)\right|\le\alpha_j+\frac{2}{\eta}\quad\text{for } z\in\mathcal{D}\,.$$

combining the two previous estimates yields a bound for $|zt_{ii}(z)\tilde t_{jj}(z)|$ for $z\in\mathcal{D}$.

Let $z\in\mathcal{D}'_A$. By Proposition 3.1-(5), we have

$$\|\mathrm{T}_z\|\le\sqrt{n\sigma^{-2}\gamma(z,\bar z)}\le\sqrt{\frac{2n}{\sigma^2\eta}}\,.$$

Recall the identity (3.27):

$$\tilde{\mathrm{T}}_z=-\frac{1}{z(1+\sigma\delta_z)}+\frac{1}{z(1+\sigma\delta_z)^2}\mathrm{A}^*\mathrm{T}_z\mathrm{A}\,.$$

We now use the fact that for any set $[\eta/2,A]\times[0,A]$, the function $z\mapsto 1+\sigma\delta_z$ is continuous [12, Theorem 2.1] and does not vanish. In fact, if $z\in\mathbb{C}^+$ then $1+\sigma\delta_z\in\mathbb{C}^+$ and if $x\in[\eta/2,A]$ then $\mathrm{Re}(1+\sigma\delta_x)>0$ by combining Lemma 2.1 and Theorem 2.1 in [12]. Finally

$$\|z\tilde{\mathrm{T}}_z\|\le\frac{1}{|1+\sigma\delta_z|}+\frac{1}{|1+\sigma\delta_z|^2}\|\mathrm{A}\|^2\|\mathrm{T}_z\|$$

is bounded over $\mathcal{D}'_A$. This finally yields a bound for $|zt_{ii}(z)\tilde t_{jj}(z)|$ for $z\in\mathcal{D}'_A$. As a consequence of these bounds and the dominated convergence theorem, we obtain (4.1). Proof of Theorem 3 is completed.

Appendix A. Useful identities and estimates

By a simple application of the classical identity for the inverse of a perturbated matrix:

$$(A+XRY)^{-1}=A^{-1}-A^{-1}X(R^{-1}+YA^{-1}X)^{-1}YA^{-1}\,,$$

we obtain

$$\mathrm{Q}=\mathrm{Q}_j-\frac{\mathrm{Q}_j\mathbf{y}_j\mathbf{y}_j^*\mathrm{Q}_j}{1+\mathbf{y}_j^*\mathrm{Q}_j\mathbf{y}_j}\qquad\text{and}\qquad\mathrm{Q}_j=\mathrm{Q}+\frac{\mathrm{Q}\mathbf{y}_j\mathbf{y}_j^*\mathrm{Q}}{1-\mathbf{y}_j^*\mathrm{Q}\mathbf{y}_j}.$$

We recall a result by Bai and Silverstein [2] that allows the control of the moments of quadratic forms.

Lemma A.1. [2, Lemma 2.7] *Let $\mathbf{x}=(x_1,\cdots,x_n)$ be an $n\times 1$ vector where x_i are centered i.i.d. complex random variables with unit variance. Let M be an $n\times n$ Hermitian complex matrix. Then for any $p\ge 2$, there exists a constant K_p depending only on p for which*

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

The above lemma is the key for the following estimates:

Lemma A.2. *Let $\mathbf{x}$ be defined as in Lemma A.1. Let M_n be an $n\times n$ Hermitian complex matrix independent of $\mathbf{x}_n$ and having a uniformly bounded spectral norm. Then for $p\in[2,8]$*

$$\max\left(\mathbb{E}\left|\frac{1}{n}\mathbf{x}_n^*\mathrm{M}_n\mathbf{x}_n-\frac{1}{n}\mathrm{Tr}\mathrm{M}_n\right|^p,\ \mathbb{E}\left|\mathbf{y}_n^*\mathrm{M}_n\mathbf{y}_n-\frac{\sigma^2}{n}\mathrm{Tr}\mathrm{M}_n-\mathbf{a}_n^*\mathrm{M}_n\mathbf{a}_n\right|^p\right)=\mathcal{O}\left(\frac{\|\mathrm{M}_n\|^p}{n^{p/2}}\right).$$

In particular,

$$\mathbb{E}|\hat\tau_j(z)|^p=\mathcal{O}_z\left(n^{-p/2}\right)\quad\textit{and}\quad\mathbb{E}|\hat\alpha_j(z)|^p=\mathcal{O}_z\left(n^{-p/2}\right).$$

Recall the definition of $D_\varepsilon=[0,A]+\boldsymbol{i}[\varepsilon,1]$, for $A>0$ and $\varepsilon\in(0,1)$ fixed.

Lemma A.3. *Assume that Assumptions 1 and 2 hold. Let $k_1, k_2 \in \mathbb{N}$ and, for any $z_1, z_2 \in D_\varepsilon$, define*

$$M_j := \mathrm{M}_j(z_1, z_2) = \mathrm{Q}_j^{k_1}(z_1)\mathrm{Q}_j^{k_2}(z_2)\ .$$

Then for any $p \in [2, 8]$,

$$\begin{aligned}
\sup_{z \in D_\varepsilon} \|\mathrm{Q}_j(z)\|^p &\leq \varepsilon^{-p}\,, \\
\sup_{z \in D_\varepsilon} \mathbb{E}|\tau_j(z)|^p &= \mathcal{O}_\varepsilon(n^{-p/2})\,, \\
\sup_{z \in D_\varepsilon} \mathbb{E}|\mathbf{y}_j^* \mathrm{M}_j(z)\mathbf{y}_j|^p &= \mathcal{O}(\varepsilon^{-(k_1+k_2)p})\,, \\
\sup_{z_1, z_2 \in D_\varepsilon} \mathbb{E}\Big|\mathbf{y}_j^* \mathrm{M}_j \mathbf{y}_j - \frac{\sigma^2}{n}\mathrm{Tr}\mathrm{M}_j - \mathbf{a}_j^* \mathrm{M}_j \mathbf{a}_j\Big|^p &= \mathcal{O}_\varepsilon(n^{-p/2})\ .
\end{aligned}$$

The estimates in Lemmas A.2 and A.3 mainly follow from Lemma A.1 and thus their proof is omitted. For more details on the estimate $\sup_{z \in D_\varepsilon} \mathbb{E}|\tau_j(z)|^p = \mathcal{O}_\varepsilon(n^{-p/2})$, one can check Appendix A.2 in [18].

Lemma A.4. [18, Theorem 3.3] *Let $(\mathbf{u}_n)$ and $(\mathbf{v}_n)$ be two sequences of deterministic complex $N \times 1$ vectors bounded by*

$$\sup_{n \geq 1} \max(\|\mathbf{u}_n\|, \|\mathbf{v}_n\|) < \infty,$$

and let $(\mathbf{U}_n)$ be a sequence of deterministic $N \times N$ matrix with bounded spectral norm

$$\sup_{n \geq 1} \|\mathbf{U}_n\| < \infty.$$

Then, in the setting of Theorem 2:

(1) there exists a constant K such that

$$\sum_{j=1}^{n} \mathbb{E}|\mathbf{u}_n^* \mathrm{Q}_j \mathbf{a}_j|^2 \leq K\,,$$

(2) there exists a constant K such that

$$\left|\frac{1}{n}\mathrm{Tr}\mathbf{U}(\mathrm{T} - \mathbb{E}\mathrm{Q})\right| \leq \frac{K}{n}\,,$$

(3) there exists a constant K such that

$$\mathbb{E}|\mathrm{Tr}\mathbf{U}(\mathrm{Q} - \mathbb{E}\mathrm{Q})|^2 \leq K\,,$$

(4) for any $p \in [1, 2]$, there exists a constant K_p such that

$$\max\left\{\mathbb{E}|\mathbf{u}_n^*(\mathrm{Q} - \mathrm{T})\mathbf{v}_n|^{2p},\ \mathbb{E}|\mathbf{u}_n^*(\mathrm{Q}_j - \mathrm{T}_j)\mathbf{v}_n|^{2p}\right\} \leq \frac{K_p}{n^p}\,,$$

(5) for any $p \in [1, 4]$, there exists a constant K such that

$$\max\left\{\mathbb{E}|\tilde{q}_{jj} - \tilde{b}_j|^{2p},\ \mathbb{E}|\tilde{q}_{jj} - \tilde{t}_{jj}|^{2p}\right\} \leq \frac{K}{n^p}\,.$$

The next result is a counterpart of [18, Lemma 5.1] and is used for the computation of the asymptotic covariance.

Lemma A.5. *For any $N \times 1$ vector $\mathbf{a}$ with bounded Euclidean norm, we have*

$$\max_j \mathrm{Var}\{\mathbf{a}^*(\mathbb{E}_j \mathrm{Q}_{z_1})(\mathbb{E}_j \mathrm{Q}_{z_2})\mathbf{a}\} = \mathcal{O}_{z_1,z_2}\left(\frac{1}{n}\right) \quad \textit{and} \quad \max_j \mathrm{Var}\{\mathrm{Tr}(\mathbb{E}_j \mathrm{Q}_{z_1})(\mathbb{E}_j \mathrm{Q}_{z_2})\} = \mathcal{O}_{z_1,z_2}(1)\,.$$

References

[1] Z. Bai and J. W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of information-plus-noise type matrices. Random Matrices Theory Appl., 1(1):1150004, 44, 2012.

[2] Z. D. Bai and J. W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. Ann. Probab., 26(1):316–345, 1998.

[3] 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.

[4] Z. D. Bai and J. W. Silverstein. Spectral analysis of large dimensional random matrices. Springer Series in Statistics. Springer, New York, second edition, 2010.

[5] Z. Bao, G. M. Pan, and W. Zhou. On the MIMO channel capacity for the general channels. preprint, 2013.

[6] 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.

[7] T. Cabanal-Duvillard. Fluctuations de la loi empirique de grandes matrices aléatoires. Ann. Inst. H. Poincaré Probab. Statist., 37(3):373–402, 2001.

[8] M. Capitaine. Exact separation phenomenon for the eigenvalues of large information-plus-noise type matrices, and an application to spiked models. Indiana Univ. Math. J., 63(6):1875–1910, 2014.

[9] M. Capitaine. Limiting eigenvectors of outliers for spiked information-plus-noise type matrices. arXiv preprint arXiv:1701.08069, 2017.
[10] M. Capitaine and C. Donati-Martin. Strong asymptotic freeness for Wigner and Wishart matrices. Indiana Univ. Math. J., 56(2):767–803, 2007.
[11] S. Chatterjee. Fluctuations of eigenvalues and second order Poincaré inequalities. Probab. Theory Related Fields, 143(1-2):1–40, 2009.
[12] R. B. Dozier and J. W. Silverstein. Analysis of the limiting spectral distribution of large dimensional information-plus-noise type matrices. J. Multivariate Anal., 98(6):1099–1122, 2007.
[13] R. B. 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.
[14] 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.
[15] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton, and J. Najim. On the capacity achieving covariance matrix for rician mimo channels: An asymptotic approach. Information Theory, IEEE Transactions on, 56(3):1048 –1069, march 2010.
[16] V. L. Girko. Theory of stochastic canonical equations. Vol. I, volume 535 of Mathematics and its Applications. Kluwer Academic Publishers, Dordrecht, 2001.
[17] 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.
[18] 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.
[19] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet. Large information plus noise random matrix models and consistent subspace estimation in large sensor networks. Random Matrices Theory Appl., 1(2):1150006, 51, 2012.
[20] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet. A subspace estimator for fixed rank perturbations of large random matrices. J. Multivariate Anal., 114:427–447, 2013.
[21] W. Hachem, P. Loubaton, and J. Najim. Deterministic equivalents for certain functionals of large random matrices. Ann. Appl. Probab., 17(3):875–930, 2007.
[22] W. Hachem, P. Loubaton, and J. Najim. A CLT for information-theoretic statistics of gram random matrices with a given variance profile. Ann. Appl. Probab., 18(6):2071–2130, 2008.
[23] 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.
[24] R. A. Horn and C. R. Johnson. Matrix analysis. Cambridge University Press, Cambridge, second edition, 2013.
[25] K. Johansson. On fluctuations of eigenvalues of random Hermitian matrices. Duke Math. J., 91(1):151–204, 1998.
[26] D. Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. J. Multivariate Anal., 12(1):1–38, 1982.
[27] A. Kammoun, M. Kharouf, W. Hachem, and J. Najim. A central limit theorem for the sinr at the lmmse estimator output for large-dimensional signals. Information Theory, IEEE Transactions on, 55(11):5048 –5063, nov. 2009.
[28] 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.
[29] P. Loubaton and P. Vallet. Almost sure localization of the eigenvalues in a Gaussian information plus noise model—application to the spiked models. Electron. J. Probab., 16:no. 70, 1934–1959, 2011.
[30] 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.
[31] A. L. Moustakas, S. H. Simon, and A. M. Sengupta. MIMO capacity through correlated channels in the presence of correlated interferers and noise: a (not so) large N analysis. IEEE Trans. Inform. Theory, 49(10):2545–2561, 2003. Special issue on space-time transmission, reception, coding and signal processing.
[32] J. Najim and J. Yao. Gaussian fluctuations for linear spectral statistics of large random covariance matrices. Ann. Appl. Probab., 26(3):1837–1887, 2016.
[33] 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.
[34] 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.

Marwa Banna, Saarland University, Fachbereich Mathematik, Postfach 151150, 66041 Saarbrücken, Germany
E-mail address: banna@math.uni-sb.de

Jamal Najim, Laboratoire d'Informatique Gaspard Monge (UMR 8049) Université Paris-Est Marne-La-Vallée, 5, boulevard Descartes, Champs sur Marne, 77454 Marne-La-Vallée Cedex 2, France
E-mail address: najim@univ-mlv.fr

Jianfeng Yao, Shanghai, China
E-mail address: jianfeng_yao_sh@sina.com