Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2013, Vol. 49, No. 1, 36–63
DOI: 10.1214/11-AIHP450

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

On bilinear forms based on the resolvent of large random matrices

Walid Hachem[a], Philippe Loubaton[b], Jamal Najim[a] and Pascal Vallet[b]

[a] *CNRS, Télécom Paristech, 46 rue Barrault, 75013 Paris, France. E-mail: hachem@telecom-paristech.fr; najim@telecom-paristech.fr*
[b] *Institut Gaspard Monge LabInfo, UMR 8049, Université Paris Est Marne-la-Vallée, 5 Boulevard Descartes, Champs sur Marne, 77454 Marne-la-Vallée Cedex 2, France. E-mail: philippe.loubaton@univ-mlv.fr; pascal.vallet@univ-mlv.fr*

Received 19 April 2010; revised 16 May 2011; accepted 29 July 2011

Abstract. Consider a $N \times n$ non-centered matrix Σ_n with a separable variance profile:

$$\Sigma_n = \frac{D_n^{1/2} X_n \tilde{D}_n^{1/2}}{\sqrt{n}} + A_n.$$

Matrices D_n and $\tilde{D}_n$ are non-negative deterministic diagonal, while matrix A_n is deterministic, and X_n is a random matrix with complex independent and identically distributed random variables, each with mean zero and variance one. Denote by $Q_n(z)$ the resolvent associated to $\Sigma_n \Sigma_n^*$, i.e.

$$Q_n(z) = \left(\Sigma_n \Sigma_n^* - z I_N\right)^{-1}.$$

Given two sequences of deterministic vectors (u_n) and (v_n) with bounded Euclidean norms, we study the limiting behavior of the random bilinear form:

$$u_n^* Q_n(z) v_n \quad \forall z \in \mathbb{C} - \mathbb{R}^+,$$

as the dimensions of matrix Σ_n go to infinity at the same pace. Such quantities arise in the study of functionals of $\Sigma_n \Sigma_n^*$ which do not only depend on the eigenvalues of $\Sigma_n \Sigma_n^*$, and are pivotal in the study of problems related to non-centered Gram matrices such as central limit theorems, individual entries of the resolvent, and eigenvalue separation.

Résumé. Considérons une matrice Σ_n, non centrée, de taille $N \times n$, avec un profil de variance séparable :

$$\Sigma_n = \frac{D_n^{1/2} X_n \tilde{D}_n^{1/2}}{\sqrt{n}} + A_n.$$

Les matrices D_n et $\tilde{D}_n$ sont déterministes, diagonales et non négatives ; la matrice A_n est déterministe ; la matrice X_n est une matrice aléatoire dont les entrées complexes sont des variables aléatoires indépendantes et identiquement distribuées, de moyenne nulle et de variance unité. On note $Q_n(z)$ la résolvante associée à $\Sigma_n \Sigma_n^*$, i.e.

$$Q_n(z) = \left(\Sigma_n \Sigma_n^* - z I_N\right)^{-1}.$$

Étant données deux suites déterministes de vecteurs (u_n) et (v_n) de norme euclidienne bornée, on étudie le comportement asymptotique de la forme bilinéaire aléatoire :

$$u_n^* Q_n(z) v_n \quad \forall z \in \mathbb{C} - \mathbb{R}^+,$$

quand les dimensions de la matrice Σ_n tendent vers l'infini au même rythme. De telles quantités apparaissent dans l'étude de fonctionnelles de $\Sigma_n \Sigma_n^*$ ne dépendant pas uniquement des valeurs propres de $\Sigma_n \Sigma_n^*$, et sont centrales dans l'étude de problèmes relatifs aux matrices de Gram non centrées tels que l'établissement de théorèmes de la limite centrale, le comportement des entrées individuelles et les problèmes de séparation des valeurs propres.

MSC: Primary 15A52; secondary 15A18; 60F15

Keywords: Random matrix; Empirical distribution of the eigenvalues; Stieltjes transform

1. Introduction

The model

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

$$\Sigma_n = \frac{D_n^{1/2} X_n \tilde{D}_n^{1/2}}{\sqrt{n}} + A_n \stackrel{\triangle}{=} Y_n + A_n, \tag{1.1}$$

where D_n and $\tilde{D}_n$ are respectively $N \times N$ and $n \times n$ non-negative deterministic diagonal matrices. The entries of matrices (X_n), $(X_{ij}^n; i, j, n)$ are complex, independent and identically distributed (i.i.d.) with mean 0 and variance 1, and $A_n = (a_{ij}^n)$ is a deterministic $N \times n$ matrix whose spectral norm is bounded in n.

The purpose of this article is to study bilinear forms based on the resolvent $Q_n(z)$ of matrix $\Sigma_n \Sigma_n^*$, where Σ_n^* stands for the Hermitian adjoint of Σ_n:

$$Q_n(z) = \left(\Sigma_n \Sigma_n^* - z I_N\right)^{-1},$$

as the dimensions N and n grow to infinity at the same pace, that is:

$$0 < \liminf \frac{N}{n} \le \limsup \frac{N}{n} < \infty, \tag{1.2}$$

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

A lot of attention has been devoted to the study of quadratic forms $y^* A y$, where $y = n^{-1/2}(X_1, \ldots, X_n)^T$, the X_i's being i.i.d., and A is a matrix independent from y. It is well-known, at least since Marcenko and Pastur's seminal paper [18], Lemma 1 (see also [3], Lemma 2.7), that under fairly general conditions, $y^* A y \sim_\infty n^{-1} \operatorname{Tr} A$.

Such a result is of constant use in the study of centered random matrices, as it allows to describe the behavior of the Stieltjes transform associated to the spectral measure (empirical distribution of the eigenvalues) of the matrix under investigation, see for instance [14,15,23,24], etc. Indeed, the Stieltjes transform of the spectral measure writes:

$$f_n(z) = \frac{1}{N} \operatorname{Tr} Q_n(z) = \frac{1}{N} \sum_{i=1}^{N} \left[Q_n(z)\right]_{ii}(z),$$

where the $[Q_n(z)]_{ii}$'s denote the diagonal elements of the resolvent. Denote by $\tilde{\eta}_i$ the ith row of Σ_n and by $\Sigma_{n,i}$ matrix Σ_n when row $\tilde{\eta}_i$ has been removed, then the matrix inversion lemma yields the following expression:

$$\left[Q_n(z)\right]_{ii} = -\frac{1}{z(1 + \tilde{\eta}_i (\Sigma_{n,i}^* \Sigma_{n,i} - zI)^{-1} \tilde{\eta}_i^*)}.$$

In the case where $\Sigma_n = n^{-1/2} X_n$, the quadratic form that appears in the previous expression can be handled by the aforementioned results. However, if Σ_n is non-centered and given by (1.1), then the quadratic form writes:

$$\tilde{\eta}_i \tilde{Q}_i(z) \tilde{\eta}_i^* = \tilde{y}_i \tilde{Q}_i(z) \tilde{y}_i^* + \tilde{a}_i \tilde{Q}_i(z) \tilde{y}_i^* + \tilde{y}_i \tilde{Q}_i(z) \tilde{a}_i^* + \tilde{a}_i \tilde{Q}_i(z) \tilde{a}_i^*,$$

where $\tilde{Q}_i(z) = (\Sigma_{n,i}^* \Sigma_{n,i} - zI)^{-1}$, and $\tilde{y}_i$ and $\tilde{a}_i$ are the ith rows of matrices Y_n and A_n. The first term can be handled as in the centered case, the second and third terms go to zero; however, the fourth term involves a quadratic form $\tilde{a}_i \tilde{Q}_i(z) \tilde{a}_i^*$ based on deterministic vectors.

It is of interest to notice that, due to some fortunate cancellation, the particular study of bilinear forms of the type $u_n^* Q_n(z) v_n$ or their analogues of the type $\tilde{u}_n \tilde{Q}_n(z) \tilde{v}_n^*$ can be circumvented to establish first order results for non-centered random matrices (see for instance [7,15]). However, such a study has to be addressed for finer questions such as: Asymptotic behavior of individual entries of the resolvent (see for instance [10], Eq. (2.16), where such properties are established in the centered Wigner case to describe fine properties of the spectrum), Central Limit Theorems [13, 17], behavior of the extreme eigenvalues of $\Sigma_n \Sigma_n^*$, behavior of the eigenvalues and eigenvectors associated with finite rank perturbations of $\Sigma_n \Sigma_n^*$ [6], behavior of eigenvectors or projectors on eigenspaces of $Q(z)$ (see for instance [2] in the context of sample covariance (centered) model), etc.

In a more applied setting, functionals based on individual entries of the resolvent [1] naturally arise in the field of wireless communication (see for instance Section 2.1). Moreover, the asymptotic study of the quadratic forms $u_n^* Q_n(z) u_n$ is important in statistical inference problems. In the non-correlated case (where $D_n = I_N$ and $\tilde{D}_n = I_n$), it is proved in [25] how such quadratic forms yield consistent estimates of projectors on the subspace orthogonal to the column space of A_n in the Gaussian case (see also Section 2.2). Such projectors form the basis of MUSIC algorithm, very popular in the field of antenna array processing. A similar approach has been developed in [19,20] for sample covariance matrix models.

It is the purpose of this article to provide a quantitative description of the limiting behavior of the bilinear form $u_n^* Q_n(z) v_n$, where u_n and v_n are deterministic, as the dimensions of Σ_n go to infinity as indicated in (1.2).

Assumptions, fundamental equations, deterministic equivalents

Formal assumptions for the model are stated below, where $\|\cdot\|$ either denotes the Euclidean norm of a vector or the spectral norm of a matrix.

Assumption A-1. *The random variables $(X_{ij}^n;\ 1 \le i \le N, 1 \le j \le n, n \ge 1)$ are complex, independent and identically distributed. They satisfy $\mathbb{E} X_{ij}^n = 0$ and $\mathbb{E}|X_{ij}^n|^2 = 1$.*

Assumption A-2. *The family of deterministic $N \times n$ matrices $(A_n, n \ge 1)$ is bounded for the spectral norm as $N, n \to \infty$:*

$$\mathbf{a}_{\max} = \sup_{n \ge 1} \|A_n\| < \infty.$$

Notice that this assumption implies in particular that the Euclidean norm of any row or column of $\|A_n\|$ is uniformly bounded in N, n.

Assumption A-3. *The families of real deterministic $N \times N$ and $n \times n$ matrices (D_n) and $(\tilde{D}_n)$ are diagonal with non-negative diagonal elements, and are bounded for the spectral norm as $N, n \to \infty$:*

$$\mathbf{d}_{\max} = \sup_{n \ge 1} \|D_n\| < \infty \quad \textit{and} \quad \tilde{\mathbf{d}}_{\max} = \sup_{n \ge 1} \|\tilde{D}_n\| < \infty.$$

Moreover,

$$\mathbf{d}_{\min} = \inf_N \frac{1}{N} \operatorname{Tr} D_n > 0 \quad \textit{and} \quad \tilde{\mathbf{d}}_{\min} = \inf_n \frac{1}{n} \operatorname{Tr} \tilde{D}_n > 0.$$

We collect here results from [15].

The following system of equations:

$$\begin{cases} \delta(z) = \frac{1}{n} \operatorname{Tr} D_n \big(-z\big(I_N + \tilde{\delta}(z) D_n\big) I_N + A_n \big(I_n + \delta(z) \tilde{D}_n\big)^{-1} A_n^*\big)^{-1}, \\ \tilde{\delta}(z) = \frac{1}{n} \operatorname{Tr} \tilde{D}_n \big(-z\big(I_n + \delta(z) \tilde{D}_n\big) + A_n^* \big(I_N + \tilde{\delta}(z) D_n\big)^{-1} A_n\big)^{-1}, \end{cases} \qquad z \in \mathbb{C} - \mathbb{R}^+, \tag{1.3}$$

admits a unique solution $(\delta, \tilde{\delta})$ in the class of Stieltjes transforms of non-negative measures[1] with support in $\mathbb{R}^+$. Matrices $T_n(z)$ and $\tilde{T}_n(z)$ defined by

$$\begin{cases} T_n(z) = \left(-z\left(I_N + \tilde{\delta}(z)D_n\right) + A_n\left(I_n + \delta(z)\tilde{D}_n\right)^{-1}A_n^*\right)^{-1}, \\ \tilde{T}_n(z) = \left(-z\left(I_n + \delta(z)\tilde{D}_n\right) + A_n^*\left(I_N + \tilde{\delta}(z)D_n\right)^{-1}A_n\right)^{-1} \end{cases} \tag{1.4}$$

are approximations of the resolvent $Q_n(z)$ and the co-resolvent $\tilde{Q}_n(z) = (\Sigma_n^*\Sigma_n - zI_N)^{-1}$ in the sense that ($\xrightarrow{\text{a.s.}}$ stands for the almost sure convergence):

$$\frac{1}{N}\operatorname{Tr}\left(Q_n(z) - T_n(z)\right) \xrightarrow[N,n\to\infty]{\text{a.s.}} 0,$$

which readily gives a deterministic approximation of the Stieltjes transform $N^{-1}\operatorname{Tr} Q_n(z)$ of the spectral measure of $\Sigma_n\Sigma_n^*$ in terms of T_n (and similarly for $\tilde{Q}_n$ and $\tilde{T}_n$). Matrices T_n and $\tilde{T}_n$ will play a fundamental role in the sequel.

Nice constants and nice polynomials

By nice constants, we mean positive constants which depend upon the limiting quantities $\mathbf{d}_{\min}$, $\tilde{\mathbf{d}}_{\min}$, $\mathbf{d}_{\max}$, $\tilde{\mathbf{d}}_{\max}$, $\mathbf{a}_{\max}$, $\liminf \frac{N}{n}$ and $\limsup \frac{N}{n}$ but are independent from n and N. Nice polynomials are polynomials with fixed degree (which is a nice constant) and with non-negative coefficients, each of them being a nice constant. Further dependencies are indicated if needed.

Statement of the main result

Let $\boldsymbol{\delta}_z$ be the distance between the point $z \in \mathbb{C}$ and the real non-negative axis $\mathbb{R}^+$:

$$\boldsymbol{\delta}_z = \operatorname{dist}\left(z, \mathbb{R}^+\right). \tag{1.5}$$

Here is the main result of the paper:

Theorem 1.1. *Assume that $N, n \to \infty$ and that Assumptions A-1, A-2 and A-3 hold true. Assume moreover that there exists an integer $p \geq 1$ such that $\sup_n \mathbb{E}|X_{ij}^n|^{8p} < \infty$ and let (u_n) and (v_n) be sequences of $N \times 1$ deterministic vectors. Then, for every $z \in \mathbb{C} - \mathbb{R}^+$,*

$$\mathbb{E}\left|u_n^*\left(Q_n(z) - T_n(z)\right)v_n\right|^{2p} \leq \frac{1}{n^p}\Phi_p\left(|z|\right)\Psi_p\left(\frac{1}{\boldsymbol{\delta}_z}\right)\|u_n\|^{2p}\|v_n\|^{2p}, \tag{1.6}$$

where Φ_p and Ψ_p are nice polynomials depending on p but not on (u_n) neither on (v_n).

Remark 1.1. *Apart from providing the convergence speed $\mathcal{O}(n^{-p})$, inequality (1.6) provides a fine control of the behavior of $\mathbb{E}|u^*(Q-T)v|^{2p}$ when z is near the real axis. Such a control should be helpful for studying the behavior of the extreme eigenvalues of $\Sigma_n\Sigma_n^*$ along the lines of* [3] *and* [4].

Remark 1.2 (Influence of the eigenvectors of AA^* on the limiting behavior of u^*Qu). *Consider a matrix Σ with no variance profile ($D = I_N$, $\tilde{D} = I_n$) and let T be given by (1.4). Matrix T writes in this case:*

$$T = \left(-z(1+\tilde{\delta})I + \frac{AA^*}{1+\delta}\right)^{-1}.$$

[1] In fact, δ and $\tilde{\delta}$ are the Stieltjes transforms of measures with respective total mass $n^{-1}\operatorname{Tr} D_n$ and $n^{-1}\operatorname{Tr}\tilde{D}_n$.

Denote by $V\Delta V^*$ the spectral decomposition of AA^*, and by T_Δ:

$$T_\Delta = \left(-z(1+\tilde{\delta})I + \frac{\Delta}{1+\delta}\right)^{-1}.$$

Obviously, $T = VT_\Delta V^*$ and by Theorem 1.1, $u^*Qu - u^*VT_\Delta V^*u \to 0$. Clearly, the limiting behavior of u^*Qu not only depends on the spectrum (matrix Δ) of AA^* but also on its eigenvectors (matrix V).

Contents

In Section 2, we describe two important motivations from electrical engineering. In Section 3, we set up the notations, state intermediate results among which Lemma 3.6, which is the cornerstone of the paper. Loosely speaking, this lemma whose idea can be found in the work of Girko [11] states that quantities such as

$$\sum_{i=1}^n u^* Q_i a_i a_i^* Q_i u$$

are bounded. This control turns out to be central to take into account Assumption A-2. An intermediate deterministic matrix R_n is introduced and the proof of Theorem 1.1 is outlined. Basically, the quantity of interest $u^*(Q-T)v$ is split into three parts:

$$u^*(Q-T)v = u^*(Q-\mathbb{E}Q)v + u^*(\mathbb{E}Q-R)v + u^*(R-T)v,$$

each being studied separately.

In Section 4, the proof of estimate of $u^*(Q-\mathbb{E}Q)v$ is established, based on a decomposition of $Q-\mathbb{E}Q$ as a sum of martingale increments. Section 5 is devoted to the proof of estimate of $u^*(\mathbb{E}Q-R)v$; and Section 6, to the proof of estimate of $u^*(R-T)v$.

2. Two applications to electrical engineering

Apart from the technical motivations already mentionned in the Introduction, Theorem 1.1 has further applications in electrical engineering. In this section, we present an application to Multiple Input Multiple Output (MIMO) wireless communication systems, and an application to statistical signal processing.

2.1. *Optimal precoder in MIMO systems*

A bi-correlated MIMO wireless Ricean channel is a $N \times n$ random matrix H_n given by

$$H_n = B_n + R_n^{1/2}\frac{V_n}{\sqrt{n}}\tilde{R}_n^{1/2},$$

where B_n is a deterministic matrix, V_n is a standard complex Gaussian matrix, and where R_n and $\tilde{R}_n$ represent deterministic positive $N \times N$ and $n \times n$ matrices. An important related question is the determination of a precoder maximizing the so-called capacity after mininum mean square error detection (for more details on the application context, see [1]). Mathematically, this problem is equivalent to the evaluation of a deterministic $N \times N$ matrix K_n maximizing the function $\mathcal{I}_{\mathrm{mmse}}(K_n)$ defined on the set of all complex valued $N \times N$ matrices by

$$\mathcal{I}_{\mathrm{mmse}}(K_n) = \mathbb{E}\sum_{j=1}^N \left[\log\left(I + K_n H_n H_n^* K_n^*\right)^{-1}_{j,j}\right] \tag{2.1}$$

under the constraint $\frac{1}{N}\operatorname{Tr}(K_n K_n^*) \le a$ ($a > 0$). This optimization problem has no closed form solution and one must rely on numerical computations. However, direct numerical attempts such as optimization by steepest descent algorithms or Monte-Carlo simulations to evaluate $\mathcal{I}_{\text{mmse}}(K_n)$ before optimization, or any combination of these techniques, face major difficulties, among which: Hardly tractable expressions for $\mathcal{I}_{\text{mmse}}(K_n)$, and for its first and second derivatives, computationally intensive algorithms when relying on Monte-Carlo simulations.

If N and n are large enough, an alternative approach consists in deriving a large system approximation $\overline{\mathcal{I}}_{\text{mmse}}(K_n)$ of $\mathcal{I}_{\text{mmse}}(K_n)$ which, hopefully, is simpler to optimize w.r.t. K_n. This idea has been successfully developed in [1], in the case where $B_n = 0$, and in [9] in a slightly different context, where the functional under consideration is the Shannon capacity $\mathcal{I}_s(K_n) = \mathbb{E}\log\det(I + K_n H_n H_n^* K_n^*)$.

In the remainder of this section, we consider the case where $B_n \neq 0$ and briefly indicate how Theorem 1.1 comes into play. First remark that for every deterministic matrix K_n, the random matrix $K_n H_n$ writes:

$$K_n H_n = K_n B_n + \left(K_n R_n K_n^*\right)^{1/2} \frac{W_n}{\sqrt{n}} \tilde{R}_n^{1/2},$$

where W_n is standard Gaussian random matrix (notice that $(K_n R_n K_n^*)^{-1/2} K_n R_n^{1/2}$ is unitary).

Using the eigenvalue/eigenvector decomposition of matrices $K_n R_n K_n^*$ and $\tilde{R}_n$, the unitary invariance of the canonical equations (1.3), and Theorem 1.1, one can easily check that the diagonal entries of $(I + K_n H_n H_n^* K_n^*)^{-1}$ have the same asymptotic behaviour (when $(n, N) \to \infty$) as those of the deterministic matrix $T_n(K_n)$ defined by:

$$T_n(K_n) = \left[\left(I + \tilde{\delta}(K_n) K_n R_n K_n^*\right) + K_n B_n \left(I + \delta(K_n)\tilde{R}_n\right)^{-1} B_n^* K_n^*\right]^{-1},$$

where $\delta(K_n)$ and $\tilde{\delta}(K_n)$ are the (unique) positive solutions of the system:

$$\begin{cases} \delta(K_n) = \frac{1}{n}\operatorname{Tr} K_n R_n K_n^* \left[\left(I + \tilde{\delta}(K_n) K_n R_n K_n^*\right) + K_n B_n \left(I + \delta(K_n)\tilde{R}_n\right)^{-1} B_n^* K_n^*\right]^{-1}, \\ \tilde{\delta}(K_n) = \frac{1}{n}\operatorname{Tr} \tilde{R}_n \left[\left(I + \delta(K_n)\tilde{R}_n\right) + B_n^* K_n^* \left(I + \tilde{\delta}(K_n) K_n R_n K_n^*\right)^{-1} K_n B_n\right]^{-1}. \end{cases} \tag{2.2}$$

From this, it appears that $\mathcal{I}_{\text{mmse}}(K_n)$ can be approximated by $\overline{\mathcal{I}}_{\text{mmse}}(K_n)$ given by:

$$\overline{\mathcal{I}}_{\text{mmse}}(K_n) = \sum_{j=1}^{N} \log\left[\left(I + \tilde{\delta}(K_n) K_n R_n K_n^*\right) + K_n B_n \left(I + \delta(K_n)\tilde{R}_n\right)^{-1} B_n^* K_n^*\right]_{j,j}^{-1}.$$

Although the values taken by function $K_n \to \overline{\mathcal{I}}_{\text{mmse}}(K_n)$ are defined through the implicit equations (2.2), the first and second derivatives of $\overline{\mathcal{I}}_{\text{mmse}}$ are easy to compute, and the minimization of $\overline{\mathcal{I}}_{\text{mmse}}$ instead of $\mathcal{I}_{\text{mmse}}$ certainly leads to a computationally attractive algorithm.

A number of important related questions remain to be addressed, e.g. the accuracy of the approximation $\overline{\mathcal{I}}_{\text{mmse}}(K_n)$, its impact on the error on the optimum solution, the derivation of a more accurate approximation as in [1], the development of an efficient algorithm to compute the optimal K_n^*, etc.; however this already underlines promising applications of Theorem 1.1 in the context of wireless communication.

2.2. *Statistical signal processing applications*

There are many important applications such as source localization using antenna arrays, communication channel estimation, detection of signals corrupted by additive noise, etc. where the observations are stacked into a matrix Σ_n given by (1.1) in which A_n is a non-observable deterministic matrix modelling the information to be retrieved and where Y_n is due to an additive noise. It is therefore often relevant to estimate certain functionals of matrix A_n from Σ_n. In this section, we show how Theorem 1.1 is valuable and relevant in the context of subspace estimators when N and n are of the same order of magnitude.

Subspace estimation
Assume that $\frac{N}{n} < 1$, $D_n = I_N$ and $\tilde{D}_n = I_n$ (white noise); assume also that matrix $\mathrm{Rank}(A_n) = r < N$ where r may scale or not with the dimensions n and N. Denote by Π_n the orthogonal projection on the kernel of matrix A_n. The subspace estimation methods are based on the estimation of quadratic forms $u_n^* \Pi_n u_n$ where $(u_n)_{n\in\mathbb{N}}$ represents a sequence of unit norm deterministic N-dimensional vectors.

If N if fixed while $n \to +\infty$, it is well known that $\|\Sigma_n \Sigma_n^* - (A_n A_n^* + I)\| \to 0$. Hence, if $\check{\Pi}_n$ represents the orthogonal projection matrix on the eigenspace associated to the $N - r$ smallest eigenvalues of $\Sigma_n \Sigma_n^*$, then $\|\check{\Pi}_n - \Pi_n\| \to 0$ and thus

$$u_n^* \check{\Pi}_n u_n - u_n^* \Pi_n u_n \xrightarrow[n\to\infty, N \text{ fixed}]{} 0. \tag{2.3}$$

In order to model situations in which n and N are large and of the same order of magnitude, it is relevant to look for estimators consistent in the regime given by (1.2). Unfortunately, (2.3) is no longer valid in this context.

An estimator for large N, n
The starting point of the estimator proposed in [25], inspired by [21], is based on the observation that Π_n writes:

$$\Pi_n = \frac{1}{2\mathrm{i}\pi} \int_{\mathcal{C}^-} \left(A_n A_n^* - \lambda I\right)^{-1} \mathrm{d}\lambda,$$

where $\mathcal{C}^-$ is a clockwise oriented contour enclosing 0 but not the non-zero eigenvalues of $A_n A_n^*$. In the white noise case, matrix $T_n(z)$ writes:

$$T_n(z) = \left(1 + \delta_n(z)\right)\left(A_n A_n^* - w_n(z) I\right)^{-1},$$

where $w_n(z)$ is the function defined by $w_n(z) = z(1 + \delta_n(z))(1 + \tilde{\delta}_n(z))$. It is shown in [25] that (under additional assumptions) such a contour $\mathcal{C}^-$ is the image under w_n of the boundary $\partial\mathcal{R}_y$ of the rectangle $\mathcal{R}_y = \{z = x + \mathrm{i}v, x \in [x_-, x_+], |v| \le y\}$ for well-chosen x_- and x_+. A simple change of variable argument therefore yields the following formula for Π_n:

$$\Pi_n = \frac{1}{2\mathrm{i}\pi} \int_{\partial\mathcal{R}_y^-} \left(A_n A_n^* - w_n(z) I\right)^{-1} w_n'(z)\, \mathrm{d}z = \frac{1}{2\mathrm{i}\pi} \int_{\partial\mathcal{R}_y^-} T_n(z) \frac{w_n'(z)}{1 + \delta_n(z)}\, \mathrm{d}z.$$

Hence, $u_n^* \Pi_n u_n$ is given by:

$$u_n^* \Pi_n u_n = \frac{1}{2\mathrm{i}\pi} \int_{\partial\mathcal{R}_y^-} u_n^* T_n(z) u_n \frac{w_n'(z)}{1 + \delta_n(z)}\, \mathrm{d}z. \tag{2.4}$$

Equation (2.4) is particularly interesting because all the terms in the integrand admit consistent estimators: Quantities $\delta_n(z)$ and $\tilde{\delta}_n(z)$ can be estimated by $\hat{\delta}_n(z) = \frac{1}{n}\mathrm{Tr}(Q_n(z))$ and $\hat{\tilde{\delta}}_n(z) = \frac{1}{n}\mathrm{Tr}(\tilde{Q}_n(z))$, $w_n'(z)$ can be estimated by the derivative of $\hat{w}_n(z) = z(1 + \hat{\delta}_n(z))(1 + \hat{\tilde{\delta}}_n(z))$; finally, Theorem 1.1 implies that $u_n^* Q_n(z) u_n - u_n^* T_n(z) u_n \to 0$ for $N, n \to \infty$. A reasonnable estimator for Π_n should therefore be

$$\hat{\Pi}_n = \frac{1}{2\mathrm{i}\pi} \int_{\partial\mathcal{R}_y^-} Q_n(z) \frac{\hat{w}_n'(z)}{1 + \hat{\delta}_n(z)}\, \mathrm{d}z \tag{2.5}$$

and it should be expected that $u_n^* \hat{\Pi}_n u_n - u_n^* \Pi_n u_n \to 0$ for $N, n \to \infty$.

Remaining mathematical issues
The full definition of $\hat{\Pi}_n$ requires to prove that none of the poles of the integrand of the r.h.s. of (2.5) can be equal to x_- or x_+. Otherwise, the mere definition of $\hat{\Pi}_n$ does not make sense. This problem has been solved in the Gaussian case in [25]. In the non-Gaussian case, partial results concerning “no eigenvalue separation for the signal plus noise model” [5] together with Theorem 1.1 tend to indicate that the estimator $u_n^* \hat{\Pi}_n u_n$ is also consistent.

3. Notations, preliminary results and sketch of proof

3.1. *Notations*

The indicator function of the set $\mathcal{A}$ will be denoted by $\mathbf{1}_{\mathcal{A}}(x)$, its cardinality by $\#\mathcal{A}$. Denote by $a \wedge b = \inf(a, b)$ and by $a \vee b = \sup(a, b)$. As usual, $\mathbb{R}^+ = \{x \in \mathbb{R}: x \geq 0\}$ and $\mathbb{C}^+ = \{z \in \mathbb{C}: \operatorname{Im}(z) > 0\}$; similarly $\mathbb{C}^- = \{z \in \mathbb{C}: \operatorname{Im}(z) < 0\}$; $\mathbf{i} = \sqrt{-1}$; if $z \in \mathbb{C}$, then $\bar{z}$ stands for its complex conjugate. Denote by $\xrightarrow{\mathcal{P}}$ the convergence in probability of random variables and by $\xrightarrow{\mathcal{D}}$ the convergence in distribution of probability measures. Denote by $\operatorname{diag}(a_i; 1 \leq i \leq k)$ the $k \times k$ diagonal matrix whose diagonal entries are the a_i's. Element (i, j) of matrix M will be either denoted m_{ij} or $[M]_{ij}$ depending on the notational context. If M is a $n \times n$ square matrix, $\operatorname{diag}(M) = \operatorname{diag}(m_{ii}; 1 \leq i \leq n)$. Denote by M^T the matrix transpose of M, by M^* its Hermitian adjoint, by $\operatorname{Tr}(M)$ its trace and $\det(M)$ its determinant (if M is square). We shall use Landau's notation: By $a_n = \mathcal{O}(b_n)$, it is meant that there exists a nice constant K such that $|a_n| \leq K|b_n|$ as $N, n \to \infty$. Recall that when dealing with vectors, $\|\cdot\|$ will refer to the Euclidean norm; in the case of matrices, $\|\cdot\|$ will refer to the spectral norm.

Due to condition (1.2), we can assume (without loss of generality) that there exist $0 < \boldsymbol{\ell}^- \leq \boldsymbol{\ell}^+ < \infty$ such that

$$\forall N, n \in \mathbb{N}^*, \quad \boldsymbol{\ell}^- \leq \frac{N}{n} \leq \boldsymbol{\ell}^+.$$

We may drop occasionally subscripts and superscripts n for readability.

Denote by Y the $N \times n$ matrix $n^{-1/2} D^{1/2} X \tilde{D}^{1/2}$; by (η_j), (a_j), (x_j) and (y_j) the columns of matrices Σ, A, X and Y. Denote by Σ_j, A_j and Y_j, the matrices Σ, A and Y where column j has been removed. The associated resolvent is $Q_j(z) = (\Sigma_j \Sigma_j^* - zI_N)^{-1}$. Denote by $\mathbb{E}_j$ the conditional expectation with respect to the σ-field $\mathcal{F}_j$ generated by the vectors $(y_\ell, 1 \leq \ell \leq j)$. By convention, $\mathbb{E}_0 = \mathbb{E}$. Denote by $\mathbb{E}_{y_j}$ the conditional expectation with respect to the σ-field generated by the vectors $(y_\ell, \ell \neq j)$.

3.2. *Classical and useful results*

We remind here classical identities of constant use in the sequel. The first one expresses the diagonal elements of the co-resolvent; the other ones are based on low-rank perturbations of inverses (see for instance [16], Section 0.7.4).

Diagonal elements of the co-resolvent; *rank-one perturbation of the resolvent*

$$\tilde{q}_{jj}(z) = -\frac{1}{z(1 + \eta_j^* Q_j(z)\eta_j)}, \tag{3.1}$$

$$Q(z) = Q_j(z) - \frac{Q_j(z)\eta_j\eta_j^* Q_j(z)}{1 + \eta_j^* Q_j \eta_j}, \tag{3.2}$$

$$Q_j(z) = Q(z) + \frac{Q(z)\eta_j\eta_j^* Q(z)}{1 - \eta_j^* Q \eta_j}, \tag{3.3}$$

$$1 + \eta_j^* Q_j \eta_j = \frac{1}{1 - \eta_j^* Q \eta_j}. \tag{3.4}$$

A useful consequence of (3.2) is:

$$\eta_j^* Q(z) = \frac{\eta_j^* Q_j(z)}{1 + \eta_j^* Q_j(z)\eta_j} = -z\tilde{q}_{jj}(z)\eta_j^* Q_j(z). \tag{3.5}$$

Recall that $\boldsymbol{\delta}_z = \operatorname{dist}(z, \mathbb{R}^+)$. Considering the eigenvalues of $Q(z)$ immediately yields $\|Q(z)\| \leq \boldsymbol{\delta}_z^{-1}$. Taking into account the fact that

$$-\frac{1}{z(1 + n^{-1}\tilde{d}_j \operatorname{Tr} Q_j + a_j^* Q_j a_j)} \quad \text{and} \quad -\frac{1}{z(1 + \eta_j^* Q_j \eta_j)}$$

are Stieltjes transforms of probability measures over $\mathbb{R}^+$, and based on standard properties of Stieltjes transforms (see for instance [15], Proposition 2.2), we readily obtain the following estimates:

$$\frac{1}{|1+(\tilde{d}_j/n)\operatorname{Tr} DQ_j + a_j^* Q_j a_j|} \le \frac{|z|}{\boldsymbol{\delta}_z} \quad \text{and} \quad \frac{1}{|1+\eta_j^* Q_j \eta_j|} \le \frac{|z|}{\boldsymbol{\delta}_z} \quad \forall z \in \mathbb{C}-\mathbb{R}^+. \tag{3.6}$$

The following lemma describes the behavior of quadratic forms based on random vectors (see for instance [3], Lemma 2.7).

Lemma 3.1. *Let $\mathbf{x}=(x_1,\dots,x_n)$ be a $n\times 1$ vector where the x_i's are centered i.i.d. complex random variables with unit variance; consider $p\ge 2$ and assume that $\mathbb{E}|x_1|^{2p}<\infty$. Let $M=(m_{ij})$ be a $n\times n$ complex matrix independent of $\mathbf{x}$. Then there exists a constant K_p such that*

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

Let $\mathbf{u}\in\mathbb{C}^n$ be deterministic, then $\mathbb{E}|\mathbf{x}^*\mathbf{u}|^p=\mathcal{O}(\|\mathbf{u}\|^p)$. Moreover, $\mathbb{E}\|\mathbf{x}\|^p=\mathcal{O}(n^{p/2})$.

Note by $D=\operatorname{diag}(d_i;1\le i\le N)$ and $\tilde{D}=\operatorname{diag}(\tilde{d}_i;1\le i\le n)$. Gathering the previous estimates yields the following useful corollary:

Corollary 3.2. *Let $z\in\mathbb{C}-\mathbb{R}^+$, and let $p\ge 2$. Denote by Δ_j the quantity:*

$$\Delta_j=\eta_j^* Q_j\eta_j-\frac{\tilde{d}_j}{n}\operatorname{Tr} DQ_j - a_j^* Q_j a_j.$$

Then

$$\mathbb{E}_{y_j}|\Delta_j|^p=\mathcal{O}\left(\frac{1}{n^{p/2}\boldsymbol{\delta}_z^p}\right).$$

Theorem 3.3 (Burkholder inequality). *Let (X_k) be a complex martingale difference sequence with respect to the filtration $(\mathcal{F}_k)$. For every $p\ge 1$, there exists K_p such that:*

$$\mathbb{E}\left|\sum_{k=1}^n X_k\right|^{2p} \le K_p\left(\mathbb{E}\left(\sum_{k=1}^n \mathbb{E}\big(|X_k|^2|\mathcal{F}_{k-1}\big)\right)^p + \sum_{k=1}^n \mathbb{E}|X_k|^{2p}\right).$$

A result on holomorphic functions:

Lemma 3.4 (Part of Schwarz's lemma, Theorem 12.2 in [22]). *Let f be an holomorphic function on the open unit disc U such that $f(0)=0$ and $\sup_{z\in U}|f(z)|\le 1$. Then $|f(z)|\le|z|$ for every $z\in U$.*

Rules about nice polynomials and nice constants

Some very simple rules of calculus related to nice polynomials will be particularly helpful in the sequel:

If $(\Phi_k, 1\le k\le K)$ and $(\Psi_k, 1\le k\le K)$ are nice polynomials, then there exist nice polynomials Φ and Ψ such that:

$$\sum_{k=1}^K \Phi_k(x)\Psi_k(y)\le\Phi(x)\Psi(y) \quad \text{for } x,y>0. \tag{3.7}$$

Take for instance $\Phi(x)=\sum_{k=1}^K\Phi_k(x)$ and $\Psi(x)=\sum_{k=1}^K\Psi_k(x)$.

If Φ_1 and Ψ_1 are nice polynomials, then there exist nice polynomials Φ and Ψ such that:

$$\sqrt{\Phi_1(x)\Psi_1(y)}\le\Phi(x)\Psi(y) \quad \text{for } x,y>0. \tag{3.8}$$

Take for instance $\Phi = 2^{-1}(1+\Phi_1)$ and $\Psi = (1+\Psi_1)$ and note that:

$$\sqrt{\Phi_1(x)\Psi_1(y)} \le \frac{1}{2}\big(1+\Phi_1(x)\Psi_1(y)\big) \le \frac{(1+\Phi_1(x))}{2}\big(1+\Psi_1(y)\big).$$

The values of nice constants or nice polynomials may change from line to line within the proofs, the constant or the polynomial remaining nice.

3.3. *Important estimates*

Lemma 3.5. *Assume that the setting of Theorem 1.1 holds true. Let u be a deterministic complex $N \times 1$ vector. Then, for every $z \in \mathbb{C} - \mathbb{R}^+$, the following estimates hold true*:

$$\mathbb{E}\left(\sum_{j=1}^n \mathbb{E}_{j-1}\big(u^* Q a_j a_j^* Q^* u\big)\right)^p \le K_p \frac{\|u\|^{2p}}{\boldsymbol{\delta}_z^{2p}}, \tag{3.9}$$

$$\mathbb{E}\left(\sum_{j=1}^n \mathbb{E}_{j-1}\big(u^* Q \eta_j \eta_j^* Q^* u\big)\right)^p \le \tilde{K}_p \frac{|z|^p \|u\|^{2p}}{\boldsymbol{\delta}_z^{2p}}, \tag{3.10}$$

where K_p and $\tilde{K}_p$ are nice constants depending on p but not on $\|u\|$.

Proof of Lemma 3.5 is postponed to Appendix A.

Lemma 3.6. *Assume that the setting of Theorem 1.1 holds true. Let u be a deterministic complex $N \times 1$ vector. Then, for every $z \in \mathbb{C} - \mathbb{R}^+$, the following estimates hold true*:

$$\sum_{j=1}^n \mathbb{E}\big(u^* Q_j a_j a_j^* Q_j^* u\big)^2 \le \Phi\big(|z|\big)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right)\|u\|^4, \tag{3.11}$$

$$\mathbb{E}\left(\sum_{j=1}^n \mathbb{E}_{j-1}\big(u^* Q_j a_j a_j^* Q_j^* u\big)\right)^p \le \tilde{\Phi}\big(|z|\big)\tilde{\Psi}\left(\frac{1}{\boldsymbol{\delta}_z}\right)\|u\|^{2p}, \tag{3.12}$$

where Φ, Ψ, $\tilde{\Phi}$ and $\tilde{\Psi}$ are nice polynomials not depending on $\|u\|$.

Proof of Lemma 3.6 is postponed to Appendix A.

In order to proceed, it is convenient to introduce the following intermediate quantities ($z \in \mathbb{C} - \mathbb{R}^+$):

$$\alpha_n(z) = \frac{1}{n}\operatorname{Tr} D_n \mathbb{E} Q_n(z), \qquad \tilde{\alpha}_n(z) = \frac{1}{n}\operatorname{Tr} \tilde{D}_n \mathbb{E} \tilde{Q}_n(z), \tag{3.13}$$

$$R_n(z) = \big(-z\big(I_N + \tilde{\alpha}(z) D_n\big) I_N + A_n\big(I_n + \alpha(z)\tilde{D}_n\big)^{-1} A_n^*\big)^{-1}, \tag{3.14}$$

$$\tilde{R}_n(z) = \big(-z\big(I_n + \alpha(z)\tilde{D}_n\big) + A_n^*\big(I_N + \tilde{\alpha}(z) D_n\big)^{-1} A_n\big)^{-1}. \tag{3.15}$$

A slight modification of the proof of [15], Proposition 5.1-(3), yields the following estimates:

$$\big\|R_n(z)\big\| \le \frac{1}{\boldsymbol{\delta}_z}, \qquad \big\|\tilde{R}_n(z)\big\| \le \frac{1}{\boldsymbol{\delta}_z} \quad \text{for } z \in \mathbb{C} - \mathbb{R}^+.$$

The same estimates hold true for $\|T_n(z)\|$ and $\|\tilde{T}_n(z)\|$.

3.4. *Main steps of the proof*

In order to prove Theorem 1.1, we split the quantity of interest $u^*(Q-T)u$ into three parts:

$$u^*(Q-T)v = u^*(Q-\mathbb{E}Q)v + u^*(\mathbb{E}Q-R)v + u^*(R-T)v,$$

and handle each term separately in the following propositions:

Proposition 3.7. *Assume that the setting of Theorem 1.1 holds true. Let (u_n) and (v_n) be sequences of $N \times 1$ deterministic vectors. Then, for every $z \in \mathbb{C} - \mathbb{R}^+$,*

$$\mathbb{E}\left|u_n^*\big(Q_n(z) - \mathbb{E}Q_n(z)\big)v_n\right|^{2p} \le \frac{1}{n^p}\Phi_p\big(|z|\big)\Psi_p\left(\frac{1}{\delta_z}\right)\|u_n\|^{2p}\|v_n\|^{2p},$$

where Φ_p and Ψ_p are nice polynomials depending on p but not on (u_n) nor on (v_n).

Proposition 3.7 is proved in Section 4.

Proposition 3.8. *Assume that the setting of Theorem 1.1 holds true.*

(i) *Let (u_n) and (v_n) be sequences of $N \times 1$ deterministic vectors. Then, for every $z \in \mathbb{C} - \mathbb{R}^+$,*

$$\left|u_n^*\big(\mathbb{E}Q_n(z) - R_n(z)\big)v_n\right| \le \frac{1}{\sqrt{n}}\Phi\big(|z|\big)\Psi\left(\frac{1}{\delta_z}\right)\|u_n\|\|v_n\|,$$

where Φ and Ψ are nice polynomials, not depending on (u_n) nor on (v_n).

(ii) *Let M_n be a $N \times N$ deterministic matrix. Then, for every $z \in \mathbb{C} - \mathbb{R}^+$,*

$$\left|\frac{1}{n}\operatorname{Tr} M_n\mathbb{E}Q_n(z) - \frac{1}{n}\operatorname{Tr} M_n R_n(z)\right| \le \frac{1}{n}\Phi\big(|z|\big)\Psi\left(\frac{1}{\delta_z}\right)\|M_n\|,$$

where Φ and Ψ are nice polynomials, not depending on M_n.

Proposition 3.8(i) is proved in Section 5; proof of Proposition 3.8(ii) is very similar and thus omitted.

Proposition 3.9. *Assume that the setting of Theorem 1.1 holds true. Let (u_n) and (v_n) be sequences of $N \times 1$ deterministic vectors.*

Then, for every $z \in \mathbb{C} - \mathbb{R}^+$,

$$\left|u_n^*\big(R_n(z) - T_n(z)\big)v_n\right| \le \frac{1}{n}\Phi\big(|z|\big)\Psi\left(\frac{1}{\delta_z}\right)\|u_n\|\|v_n\|,$$

where Φ and Ψ are nice polynomials, not depending on (u_n) nor on (v_n).

Proposition 3.9 is proved in Section 6.

Theorem 1.1 is then easily proved using these three propositions together with inequality $|x+y+z|^{2p} \le K_p(|x|^{2p}+|y|^{2p}+|z|^{2p})$ and (3.7).

4. Proof of Proposition 3.7

In this section, we establish the estimate:

$$\mathbb{E}\left|u^*\big(Q(z) - \mathbb{E}Q(z)\big)v\right|^{2p} \le \frac{1}{n^p}\Phi_p\big(|z|\big)\Psi_p\left(\frac{1}{\delta_z}\right)\|u\|^{2p}\|v\|^{2p} \quad \forall z \in \mathbb{C} - \mathbb{R}^+. \tag{4.1}$$

4.1. *Reduction to unit vectors and quadratic forms*

Using a polarization identity, it is sufficient in order to establish estimate (4.1) for the bilinear form $u^*(Q-\mathbb{E}Q)v$ to establish the related estimate for the quadratic form $u^*(Q-\mathbb{E}Q)u$ and for unit vectors $\|u\|$ (just consider $u/\|u\|$ if necessary):

$$\mathbb{E}\big|u^*\big(Q(z)-\mathbb{E}Q(z)\big)u\big|^{2p}\le\frac{1}{n^p}\Phi_p\big(|z|\big)\Psi_p\left(\frac{1}{\boldsymbol{\delta}_z}\right). \tag{4.2}$$

4.2. *Martingale difference sequence and Burkholder inequality*

We first express the difference $u^*(Q-\mathbb{E}Q)u$ as the sum of martingale difference sequences:

$$\begin{aligned}u^*(Q-\mathbb{E}Q)u&=\sum_{j=1}^n(\mathbb{E}_j-\mathbb{E}_{j-1})\big(u^*Qu\big)=\sum_{j=1}^n(\mathbb{E}_j-\mathbb{E}_{j-1})\big(u^*(Q-Q_j)u\big)\\&=-\sum_{j=1}^n(\mathbb{E}_j-\mathbb{E}_{j-1})\left(\frac{u^*Q_j\eta_j^*\eta_jQ_ju}{1+\eta_j^*Q_j\eta_j}\right)\\&\stackrel{\triangle}{=}-\sum_{j=1}^n(\mathbb{E}_j-\mathbb{E}_{j-1})\Gamma_j.\end{aligned}$$

One can easily check that $((\mathbb{E}_j-\mathbb{E}_{j-1})\Gamma_j)$ is the sum of a martingale difference sequence with respect to the filtration $(\mathcal{F}_j,j\le n)$; hence Burkholder's inequality yields:

$$\begin{aligned}&\mathbb{E}\left|\sum_{j=1}^n(\mathbb{E}_j-\mathbb{E}_{j-1})\Gamma_j\right|^{2p}\\&\quad\le K\left(\mathbb{E}\left(\sum_{j=1}^n\mathbb{E}_{j-1}\big|(\mathbb{E}_j-\mathbb{E}_{j-1})\Gamma_j\big|^2\right)^p+\sum_{j=1}^n\mathbb{E}\big|(\mathbb{E}_j-\mathbb{E}_{j-1})\Gamma_j\big|^{2p}\right).\end{aligned} \tag{4.3}$$

Recall the definition of $\Delta_j=\eta_j^*Q_j\eta_j-n^{-1}\tilde{d}_j\operatorname{Tr}DQ_j-a_j^*Q_ja_j$. In order to control the right-hand side of Burkholder's inequality, we write Γ_j as:

$$\begin{aligned}\Gamma_j&=\frac{u^*Q_j\eta_j^*\eta_jQ_ju}{1+\eta_j^*Q_j\eta_j}=\frac{u^*Q_j\eta_j^*\eta_jQ_ju}{1+\eta_j^*Q_j\eta_j}\times\frac{1+(\tilde{d}_j/n)\operatorname{Tr}DQ_j+a_j^*Q_ja_j}{1+(\tilde{d}_j/n)\operatorname{Tr}DQ_j+a_j^*Q_ja_j}\\&=\frac{u^*Q_j\eta_j^*\eta_jQ_ju}{1+\eta_j^*Q_j\eta_j}\times\frac{1+\eta_j^*Q_j\eta_j-\Delta_j}{1+(\tilde{d}_j/n)\operatorname{Tr}DQ_j+a_j^*Q_ja_j}\\&\stackrel{\triangle}{=}\Gamma_{1j}-\Gamma_{2j},\end{aligned}$$

where

$$\Gamma_{1j}=\frac{u^*Q_j\eta_j\eta_j^*Q_ju}{1+(\tilde{d}_j/n)\operatorname{Tr}DQ_j+a_j^*Q_ja_j}\quad\text{and}\quad\Gamma_{2j}=\frac{\Gamma_j\Delta_j}{1+(\tilde{d}_j/n)\operatorname{Tr}DQ_j+a_j^*Q_ja_j}. \tag{4.4}$$

In the following proposition, we establish relevant estimates.

Proposition 4.1. *Assume that the setting of Theorem 1.1 holds true. There exist nice polynomials* $(\Phi_i, 1 \le i \le 4)$ *and* $(\Psi_i, 1 \le i \le 4)$ *such that the following estimates hold true*:

$$\mathbb{E}\left(\sum_{j=1}^{n} \mathbb{E}_{j-1}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})\Gamma_{1j}\right|^2\right)^p \le \frac{1}{n^p}\Phi_1(|z|)\Psi_1\left(\frac{1}{\boldsymbol{\delta}_z}\right), \tag{4.5}$$

$$\sum_{j=1}^{n} \mathbb{E}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})\Gamma_{1j}\right|^{2p} \le \frac{1}{n^p}\Phi_2(|z|)\Psi_2\left(\frac{1}{\boldsymbol{\delta}_z}\right), \tag{4.6}$$

$$\mathbb{E}\left(\sum_{j=1}^{n} \mathbb{E}_{j-1}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})\Gamma_{2j}\right|^2\right)^p \le \frac{1}{n^p}\Phi_3(|z|)\Psi_3\left(\frac{1}{\boldsymbol{\delta}_z}\right), \tag{4.7}$$

$$\sum_{j=1}^{n} \mathbb{E}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})\Gamma_{2j}\right|^{2p} = \frac{1}{n^p}\Phi_4(|z|)\Psi_4\left(\frac{1}{\boldsymbol{\delta}_z}\right). \tag{4.8}$$

It is now clear that the proof of Proposition 3.7 directly follows from Burkholder's inequality together with the estimates of Proposition 4.1. The rest of the section is devoted to the proof of Proposition 4.1.

4.3. *Proof of Proposition* 4.1

In this section, we shall establish estimates (4.5) and (4.6). Estimates (4.7) and (4.8) can be proved similarly and are therefore omitted.

We split Γ_{1j} as $\Gamma_{1j} = \chi_{1j} + \chi_{2j} + \chi_{3j}$, where:

$$\chi_{1j} = \frac{u^* Q_j y_j y_j^* Q_j u}{1 + (\tilde{d}_j/n)\operatorname{Tr} DQ_j + a_j^* Q_j a_j},$$

$$\chi_{2j} = \frac{y_j^* Q_j u u^* Q_j a_j}{1 + (\tilde{d}_j/n)\operatorname{Tr} DQ_j + a_j^* Q_j a_j} + \frac{a_j^* Q_j u u^* Q_j y_j}{1 + (\tilde{d}_j/n)\operatorname{Tr} DQ_j + a_j^* Q_j a_j},$$

$$\chi_{3j} = \frac{u^* Q_j a_j a_j^* Q_j u}{1 + (\tilde{d}_j/n)\operatorname{Tr} DQ_j + a_j^* Q_j a_j}.$$

Notice that $(\mathbb{E}_j - \mathbb{E}_{j-1})(\chi_{3j}) = 0$, hence χ_{3j} will play no further role in the sequel. As Q_j is independent from column y_j, we have:

$$(\mathbb{E}_j - \mathbb{E}_{j-1})(\chi_{1j}) = \frac{\tilde{d}_j}{n}\mathbb{E}_j\left(\frac{x_j^* D^{1/2} Q_j u u^* Q_j D^{1/2} x_j - \operatorname{Tr} DQ_j u u^* Q_j}{1 + (\tilde{d}_j/n)\operatorname{Tr} DQ_j + a_j^* Q_j a_j}\right) \tag{4.9}$$

and

$$\begin{aligned}
\mathbb{E}_{j-1}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})(\chi_{1j})\right|^2 &\overset{(a)}{\le} \frac{\tilde{\mathbf{d}}_{\max}^2}{n^2} \times \mathbb{E}_{j-1}\left|\frac{x_j^* D^{1/2} Q_j u u^* Q_j D^{1/2} x_j - \operatorname{Tr} DQ_j u u^* Q_j}{1 + (\tilde{d}_j/n)\operatorname{Tr} DQ_j + a_j^* Q_j a_j}\right|^2 \\
&\overset{(b)}{\le} \frac{\tilde{\mathbf{d}}_{\max}^2}{n^2}\frac{|z|^2}{\boldsymbol{\delta}_z^2} \times \mathbb{E}_{j-1}\left[\mathbb{E}_{y_j}\left|x_j^* D^{1/2} Q_j u u^* Q_j D^{1/2} x_j - \operatorname{Tr} DQ_j u u^* Q_j\right|^2\right] \\
&\overset{(c)}{\le} K\frac{\tilde{\mathbf{d}}_{\max}^2}{n^2}\frac{|z|^2}{\boldsymbol{\delta}_z^2} \times \mathbb{E}_{j-1}\left(\operatorname{Tr} D^{1/2} Q_j u u^* Q_j D^{1/2} D^{1/2} Q_j^* u u^* Q_j^* D^{1/2}\right) \\
&= \mathcal{O}\left(\frac{|z|^2}{n^2\boldsymbol{\delta}_z^6}\right),
\end{aligned} \tag{4.10}$$

where (a) follows from Jensen's inequality, (b) from estimate (3.6), and (c) from Lemma 3.1. Thus

$$\mathbb{E}\left(\sum_{j=1}^{n}\mathbb{E}_{j-1}\left|(\mathbb{E}_j-\mathbb{E}_{j-1})(\chi_{1j})\right|^2\right)^p=\mathcal{O}\left(\frac{|z|^{2p}}{n^p\boldsymbol{\delta}_z^{6p}}\right). \tag{4.11}$$

We now turn to the contribution of χ_{2j}. Arguments similar as previously yield:

$$\begin{aligned}\mathbb{E}_{j-1}\left|(\mathbb{E}_j-\mathbb{E}_{j-1})(\chi_{2j})\right|^2&=\mathbb{E}_{j-1}|\mathbb{E}_j\chi_{2j}|^2\le\mathbb{E}_{j-1}|\chi_{2j}|^2\\&\le\frac{2}{n}\mathbb{E}_{j-1}\left(\left|\frac{x_j^*D^{1/2}Q_juu^*Q_ja_j}{1+(\tilde{d}_j/n)\operatorname{Tr}DQ_j+a_j^*Q_ja_j}\right|^2+\left|\frac{a_j^*Q_juu^*Q_jD^{1/2}x_j}{1+(\tilde{d}_j/n)\operatorname{Tr}DQ_j+a_j^*Q_ja_j}\right|^2\right)\\&\le\frac{2}{n}\frac{|z|^2}{\boldsymbol{\delta}_z^2}\mathbb{E}_{j-1}\big(\mathbb{E}_{y_j}\big(x_j^*D^{1/2}Q_juu^*Q_j^*D^{1/2}x_j\big)\times u^*Q_ja_ja_j^*Q_j^*u\\&\quad+\mathbb{E}_{y_j}\big(x_j^*D^{1/2}Q_j^*uu^*Q_jD^{1/2}x_j\big)\times u^*Q_j^*a_ja_j^*Q_ju\big)\\&\le\frac{K}{n}\frac{|z|^2}{\boldsymbol{\delta}_z^4}\big(\mathbb{E}_{j-1}\big(u^*Q_j^*a_ja_j^*Q_ju\big)+\mathbb{E}_{j-1}\big(u^*Q_ja_ja_j^*Q_j^*u\big)\big).\end{aligned} \tag{4.12}$$

Now, using Eq. (3.12) in Lemma 3.6 yields:

$$\mathbb{E}\left(\sum_{j=1}^{n}\mathbb{E}_{j-1}\left|(\mathbb{E}_j-\mathbb{E}_{j-1})(\chi_{2j})\right|^2\right)^p\le\frac{1}{n^p}\Phi\big(|z|\big)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right). \tag{4.13}$$

Hence, gathering (4.11) and (4.13) yields estimate (4.5).

We now establish estimate (4.6). As previously, consider identity (4.9); take it this time to the power p. Using the same arguments as for (4.10), we obtain:

$$\mathbb{E}\left|(\mathbb{E}_j-\mathbb{E}_{j-1})(\chi_{1j})\right|^{2p}=\mathcal{O}\left(\frac{|z|^{2p}}{n^{2p}\boldsymbol{\delta}_z^{6p}}\right),$$

hence:

$$\mathbb{E}\sum_{j=1}^{n}\left|(\mathbb{E}_j-\mathbb{E}_{j-1})(\chi_{1j})\right|^{2p}=\mathcal{O}\left(\frac{|z|^{2p}}{n^{2p-1}\boldsymbol{\delta}_z^{6p}}\right). \tag{4.14}$$

Similarly, using the same arguments as in (4.12), together with elementary manipulations, we obtain:

$$\mathbb{E}_{j-1}\left|(\mathbb{E}_j-\mathbb{E}_{j-1})(\chi_{2j})\right|^{2p}\le\frac{K}{n^p}\frac{|z|^{2p}}{\boldsymbol{\delta}_z^{4p}}\big(\mathbb{E}_{j-1}\big(u^*Q_j^*a_ja_j^*Q_ju\big)^p+\mathbb{E}_{j-1}\big(u^*Q_ja_ja_j^*Q_j^*u\big)^p\big).$$

Due to the rough estimate (A.1), we obtain

$$\mathbb{E}\left|(\mathbb{E}_j-\mathbb{E}_{j-1})(\chi_{2j})\right|^{2p}\le\frac{K}{n^p}\frac{|z|^{2p}}{\boldsymbol{\delta}_z^{6p-4}}\big(\mathbb{E}\big(u^*Q_j^*a_ja_j^*Q_ju\big)^2+\mathbb{E}\big(u^*Q_ja_ja_j^*Q_j^*u\big)^2\big),$$

which after summation, and the estimate obtained in Lemma 3.6, yields:

$$\mathbb{E}\sum_{j=1}^{n}\left|(\mathbb{E}_j-\mathbb{E}_{j-1})(\chi_{2j})\right|^{2p}\le\frac{1}{n^p}\Phi'\big(|z|\big)\Psi'\left(\frac{1}{\boldsymbol{\delta}_z}\right), \tag{4.15}$$

where Φ' and Ψ' are nice polynomials. Gathering (4.14) and (4.15) yields estimate (4.6).

5. Proof of Proposition 3.8

The argument referred to in Section 4.1 still holds true here; therefore it is sufficient to establish, for $z \in \mathbb{C} - \mathbb{R}^+$ and for a unit vector u:

$$\left|u^*\left(\mathbb{E}Q(z) - R(z)\right)u\right| \le \frac{1}{\sqrt{n}}\Phi\left(|z|\right)\Psi\left(\frac{1}{\delta_z}\right). \tag{5.1}$$

Recalling that $R = [-z(I + \tilde{\alpha}D) + A(I + \alpha\tilde{D})^{-1}A^*]^{-1}$, the resolvent identity yields:

$$\begin{aligned}
u^*(R - Q)u &= u^*R\left(Q^{-1} - R^{-1}\right)Qu \\
&= u^*R\left(\Sigma\Sigma^* - A(I + \alpha\tilde{D})^{-1}A^*\right)Qu + z\tilde{\alpha}u^*RDQu \\
&= u^*R\left(\sum_{j=1}^{n}\eta_j\eta_j^* - \sum_{j=1}^{n}\frac{a_j a_j^*}{1 + \alpha\tilde{d}_j}\right)Qu + z\tilde{\alpha}u^*RDQu \\
&\stackrel{(a)}{=} \sum_{j=1}^{n}\frac{u^*R\eta_j\eta_j^*Q_j u}{1 + \eta_j^*Q_j\eta_j} - \sum_{j=1}^{n}\frac{u^*Ra_j a_j^*Q_j u}{1 + \alpha\tilde{d}_j} \\
&\quad + \sum_{j=1}^{n}\frac{u^*Ra_j a_j^*Q_j\eta_j\eta_j^*Q_j u}{(1 + \eta_j^*Q_j\eta_j)(1 + \alpha\tilde{d}_j)} - \sum_{j=1}^{n}\frac{\tilde{d}_j}{n}\mathbb{E}\left(\frac{1}{1 + \eta_j^*Q_j\eta_j}\right)u^*RDQu \\
&\triangleq \sum_{j=1}^{n} Z_j,
\end{aligned}$$

where (a) follows from (3.2) and (3.5), together with the mere definition of $\tilde{\alpha}$.

As usual, we now write $\eta_j = y_j + a_j$, group the terms that compensate one another and split Z_j accordingly:

$$Z_j = Z_{1j} + Z_{2j} + Z_{3j} + Z_{4j},$$

where

$$\begin{aligned}
Z_{1j} &= \frac{y_j^*Q_j uu^*Ry_j}{1 + \eta_j^*Q_j\eta_j} - \frac{\tilde{d}_j}{n}\mathbb{E}\left(\frac{1}{1 + \eta_j^*Q_j\eta_j}\right)u^*RDQu, \\
Z_{2j} &= \frac{(\alpha\tilde{d}_j - y_j^*Q_j y_j)u^*Ra_j a_j^*Q_j u}{(1 + \eta_j^*Q_j\eta_j)(1 + \alpha\tilde{d}_j)}, \\
Z_{3j} &= \frac{y_j^*Q_j ua_j^*Q_j y_j \times u^*Ra_j}{(1 + \eta_j^*Q_j\eta_j)(1 + \alpha\tilde{d}_j)}, \\
Z_{4j} &= \frac{u^*Ry_j a_j^*Q_j u + u^*Ra_j y_j^*Q_j u}{1 + \eta_j^*Q_j\eta_j} \\
&\quad - \frac{y_j^*Q_j a_j u^*Ra_j a_j^*Q_j u + a_j^*Q_j y_j u^*Ra_j a_j^*Q_j u}{(1 + \eta_j^*Q_j\eta_j)(1 + \alpha\tilde{d}_j)} \\
&\quad + \frac{u^*Ra_j a_j^*Q_j a_j y_j^*Q_j u + u^*Ra_j a_j^*Q_j y_j a_j^*Q_j u}{(1 + \eta_j^*Q_j\eta_j)(1 + \alpha\tilde{d}_j)}.
\end{aligned}$$

Now, the estimate (5.1) immediately follows from similar estimates for the terms $\mathbb{E}\sum_{j=1}^{n} Z_{\ell j}$, $1 \le \ell \le 4$.

The rest of the section is devoted to establish such an estimate for the term $\mathbb{E}\sum_{j=1}^n Z_{1j}$. Estimates for the terms $\mathbb{E}\sum_{j=1}^n Z_{\ell j}$ $(2 \le \ell \le 4)$ rely on the same arguments and are therefore omitted.

Convergence to zero of $\sum_j \mathbb{E}Z_{1j}$

We have

$$\begin{aligned}
\mathbb{E}Z_{1j} &= \mathbb{E}\left(\frac{y_j^* Q_j u u^* R y_j}{1+\eta_j^* Q_j \eta_j}\right) - \frac{\tilde{d}_j}{n}\mathbb{E}\left(\frac{1}{1+\eta_j^* Q_j \eta_j}\right)\mathbb{E}\left(u^* R D Q u\right) \\
&= \mathbb{E}\left[\left(\frac{y_j^* Q_j u u^* R y_j}{1+\eta_j^* Q_j \eta_j}\right) - \frac{\tilde{d}_j}{n}\left(\frac{u^* R D Q_j u}{1+\eta_j^* Q_j \eta_j}\right)\right] \\
&\quad + \frac{\tilde{d}_j}{n}\left[\mathbb{E}\left(\frac{u^* R D Q_j u}{1+\eta_j^* Q_j \eta_j}\right) - \mathbb{E}\left(\frac{1}{1+\eta_j^* Q_j \eta_j}\right)\mathbb{E}\left(u^* R D Q_j u\right)\right] \\
&\quad + \frac{\tilde{d}_j}{n}\mathbb{E}\left(\frac{1}{1+\eta_j^* Q_j \eta_j}\right)\mathbb{E}\left(u^* R D (Q_j - Q) u\right) \\
&\stackrel{\triangle}{=} \chi_{1j} + \chi_{2j} + \chi_{3j}.
\end{aligned}$$

We first handle χ_{ij}. Recall that $\Delta_j = \eta_j^* Q_j \eta_j - n^{-1}\tilde{d}_j \operatorname{Tr} D Q_j - a_j^* Q_j a_j$. Since $\mathbb{E}_{y_j}(y_j^* Q_j u u^* R y_j) = \tilde{d}_j n^{-1} u^* R D \times Q_j u$, we get:

$$\begin{aligned}
\chi_{1j} &= \mathbb{E}\left[\left(\frac{y_j^* Q_j u u^* R y_j}{1+\eta_j^* Q_j \eta_j}\right) - \frac{\tilde{d}_j}{n}\left(\frac{u^* R D Q_j u}{1+\eta_j^* Q_j \eta_j}\right)\right] \\
&= \mathbb{E}\left[\left(\frac{1}{1+\eta_j^* Q_j \eta_j} - \frac{1}{1+(\tilde{d}_j/n)\operatorname{Tr} D Q_j + a_j^* Q_j a_j}\right)\left(y_j^* Q_j u u^* R y_j - \frac{\tilde{d}_j}{n}\left(u^* R D Q_j u\right)\right)\right] \\
&= \mathbb{E}\left[\Delta_j \frac{y_j^* Q_j u u^* R y_j - (\tilde{d}_j/n)(u^* R D Q_j u)}{(1+\eta_j^* Q_j \eta_j)(1+(\tilde{d}_j/n)\operatorname{Tr} D Q_j + a_j^* Q_j a_j)}\right].
\end{aligned}$$

Hence,

$$\begin{aligned}
|\chi_{1j}| &\le \frac{|z|^2}{\boldsymbol{\delta}_z^2}\sqrt{\mathbb{E}|\Delta_j|^2}\left[\mathbb{E}\left|y_j^* Q_j u u^* R y_j - \frac{\tilde{d}_j}{n}\left(u^* R D Q_j u\right)\right|^2\right]^{1/2} \\
&\le \frac{|z|^2}{\boldsymbol{\delta}_z^2} \times \frac{1}{\sqrt{n}\boldsymbol{\delta}_z} \times \frac{1}{n\boldsymbol{\delta}_z^2} = \mathcal{O}\left(\frac{|z|^2}{n^{3/2}\boldsymbol{\delta}_z^5}\right).
\end{aligned}$$

Summing over j yields the estimate $\sum_j |\chi_{1j}| = \mathcal{O}(|z|^2 n^{-1/2}\boldsymbol{\delta}_z^{-5})$.

We now handle χ_{2j}. Using the inequality $\operatorname{cov}(XY) \le \sqrt{\operatorname{var}(X)\operatorname{var}(Y)}$, we get:

$$|\chi_{2j}| \le \frac{K}{n}\frac{|z|}{\boldsymbol{\delta}_z}\sqrt{\mathbb{E}\left|u^* R D (Q_j - \mathbb{E}Q_j) u\right|^2}.$$

Hence, applying Proposition 3.7 to $|u^* R D (Q_j - \mathbb{E}Q_j) u|^2$ and summing over j yields the estimate $\sum_j |\chi_{2j}| = n^{-1/2}\Phi(|z|)\Psi(\boldsymbol{\delta}_z^{-1})$.

Let us now handle the term χ_{3j}. Using the decomposition of $Q_j - Q$, Schwarz inequality and the fact that $\sqrt{ab} \le 2^{-1}(a+b)$ yields

$$\begin{aligned} |\chi_{3j}| &= \left| \frac{\tilde{d}_j}{n} \mathbb{E}\left(\frac{1}{1+\eta_j^* Q_j \eta_j} \right) \mathbb{E}\big(u^* R D (Q_j - Q) u\big) \right| \\ &\le \frac{K}{n} \frac{|z|^2}{\boldsymbol{\delta}_z^2} \left(\mathbb{E}\left|u^* R D Q_j \eta_j\right|^2 + \mathbb{E}\left|\eta_j^* Q_j u\right|^2 \right). \end{aligned} \tag{5.2}$$

Now, as:

$$\begin{aligned} \mathbb{E}\left|u^* R D Q_j \eta_j\right|^2 &= \mathbb{E} u^* R D Q_j y_j y_j^* Q_j^* D R^* u + \mathbb{E} u^* R D Q_j a_j a_j^* Q_j^* D R^* u, \\ \mathbb{E}\left|\eta_j^* Q_j u\right|^2 &= \mathbb{E} u^* Q_j^* y_j y_j^* Q_j u + \mathbb{E} u^* Q_j^* a_j a_j^* Q_j u, \end{aligned}$$

it remains to sum over j and to apply Lemma 3.6 to get the estimate $\sum_j |\chi_{3j}| = n^{-1} \Phi(|z|) \Psi(\boldsymbol{\delta}_z^{-1})$. Gathering the partial estimates yields:

$$\left| \mathbb{E} \sum_j Z_{1j} \right| \le \frac{\Phi(|z|)\Psi(\boldsymbol{\delta}_z^{-1})}{\sqrt{n}}. \tag{5.3}$$

6. Proof of Proposition 3.9

As mentioned in Section 4.1, it is sufficient to establish the estimate:

$$\left| u^* \big(R(z) - T(z) \big) u \right| \le \frac{1}{n} \Phi\big(|z|\big) \Psi\left(\frac{1}{\boldsymbol{\delta}_z} \right) \tag{6.1}$$

for $z \in \mathbb{C} - \mathbb{R}^+$ in the case where u has norm one.

6.1. *The estimate for $u^*(R-T)u$*

Recall the definitions of $\delta, \tilde{\delta}$ (1.3), $\alpha, \tilde{\alpha}$ (3.13) and $R, \tilde{R}$ (3.14)–(3.15). Using twice the resolvent identity yields:

$$u^*(R-T)u = (\tilde{\alpha} - \tilde{\delta})\kappa_1 + (\alpha - \delta)\kappa_2, \tag{6.2}$$

where

$$\begin{cases} \kappa_1 = z u^* R D T u, \\ \kappa_2 = u^* R A (I + \alpha \tilde{D})^{-1} \tilde{D} (I + \delta \tilde{D})^{-1} A^* T u. \end{cases}$$

The following bounds are straightforward:

$$|\kappa_1| \le \frac{|z| \tilde{\mathbf{d}}_{\max}}{\boldsymbol{\delta}_z^2} \quad \text{and} \quad |\kappa_2| \le \frac{\|A\|^2 \tilde{\mathbf{d}}_{\max}}{\boldsymbol{\delta}_z^2} \times \left\| (I + \alpha \tilde{D})^{-1} \right\| \times \left\| (I + \delta \tilde{D})^{-1} \right\|.$$

It remains to control the spectral norms of $(I + \alpha \tilde{D})^{-1}$ and $(I + \delta \tilde{D})^{-1}$. Recall that α is the Stieltjes transform of a positive measure with support included in $\mathbb{R}^+$. This in particular implies that $\operatorname{Im}(z\alpha) > 0$ for $z \in \mathbb{C}^+$. One can check that

$$\Upsilon_j(z) = \frac{1}{-z(1 + \alpha \tilde{d}_j)}$$

is analytic and satisfies $\mathrm{Im}(\Upsilon_j) > 0$ and $\mathrm{Im}(z\Upsilon_j) > 0$ on $\mathbb{C}^+$ and that $\lim_{y\to\infty}(-\mathbf{i}y\Upsilon_j(\mathbf{i}y)) = 1$. As a consequence, Υ_j is the Stieltjes transform of a probability measure with support included in $\mathbb{R}^+$ (see e.g. [15], Proposition 2.2(2)). In particular,

$$\left|\Upsilon_j(z)\right| \le \frac{1}{\boldsymbol{\delta}_z} \quad \text{for } 1 \le j \le n,$$

which readily implies that $\|(I + \alpha\tilde{D})^{-1}\| \le |z|\boldsymbol{\delta}_z^{-1}$. The same argument applies for $\|(I + \delta\tilde{D})^{-1}\|$. Finally,

$$|\kappa_2| \le \frac{|z|^2\|A\|^2\tilde{\mathbf{d}}_{\max}}{\boldsymbol{\delta}_z^4}.$$

In view of the estimates obtained for κ_1 and κ_2, it is sufficient, in order to establish (6.1), to obtain estimates for $\alpha - \delta$ and $\tilde{\alpha} - \tilde{\delta}$. Assume that the following estimate holds true:

$$\forall z \in \mathbb{C} - \mathbb{R}^+, \quad \max\bigl(|\alpha - \delta|, |\tilde{\alpha} - \tilde{\delta}|\bigr) \le \frac{1}{n}\Phi\bigl(|z|\bigr)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right), \tag{6.3}$$

where Φ and Ψ are nice polynomials. Then, plugging (6.3) into (6.2) immediately yields the desired result (6.1).

The rest of the section is devoted to establish (6.3).

6.2. *Auxiliary estimates over* $(\alpha - \delta)$ *and* $(\tilde{\alpha} - \tilde{\delta})$

Writing $\alpha = n^{-1}\operatorname{Tr} DR + n^{-1}\operatorname{Tr} D(\mathbb{E}Q - R)$ and $\delta = n^{-1}\operatorname{Tr} DT$, the difference $\alpha - \delta$ expresses as $n^{-1}\operatorname{Tr} D(R - T) + n^{-1}\operatorname{Tr} D(\mathbb{E}Q - R)$. Now using the resolvent identity $R - T = -R(R^{-1} - T^{-1})T$ and performing the same computation for the tilded quantities yields the following system of equations:

$$\begin{pmatrix} \alpha - \delta \\ \tilde{\alpha} - \tilde{\delta} \end{pmatrix} = C_0 \begin{pmatrix} \alpha - \delta \\ \tilde{\alpha} - \tilde{\delta} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon} \\ \tilde{\boldsymbol{\varepsilon}} \end{pmatrix}, \quad \text{where } C_0 = \begin{pmatrix} u_0 & zv_0 \\ z\tilde{v}_0 & \tilde{u}_0 \end{pmatrix}, \tag{6.4}$$

the coefficients being defined as:

$$\begin{cases} u_0 = \frac{1}{n}\operatorname{Tr} D^{1/2}RA(I + \alpha\tilde{D})^{-1}\tilde{D}(I + \delta\tilde{D})^{-1}A^*TD^{1/2}, \\ \tilde{u}_0 = \frac{1}{n}\operatorname{Tr} \tilde{D}^{1/2}\tilde{R}A^*(I + \tilde{\alpha}D)^{-1}D(I + \tilde{\delta}D)^{-1}A\tilde{T}\tilde{D}^{1/2}, \\ v_0 = \frac{1}{n}\operatorname{Tr} DRDT, \\ \tilde{v}_0 = \frac{1}{n}\operatorname{Tr} \tilde{D}\tilde{R}\tilde{D}\tilde{T} \end{cases} \tag{6.5}$$

and the quantities $\boldsymbol{\varepsilon}$ and $\tilde{\boldsymbol{\varepsilon}}$ being given by:

$$\boldsymbol{\varepsilon} = \frac{1}{n}\operatorname{Tr} D(\mathbb{E}Q - R) \quad \text{and} \quad \tilde{\boldsymbol{\varepsilon}} = \frac{1}{n}\operatorname{Tr} \tilde{D}(\mathbb{E}\tilde{Q} - \tilde{R}). \tag{6.6}$$

The general idea, in order to transfer the estimates over $\boldsymbol{\varepsilon}$ and $\tilde{\boldsymbol{\varepsilon}}$ (as provided in Proposition 3.8(ii)), to $\alpha - \delta$ and $\tilde{\alpha} - \tilde{\delta}$, is to obtain an estimate over $1/\det(I - C_0)$, and then to solve the system (6.4).

Lower bound for $\det(I - C_0)$
The mere definition of $I - C_0$ yields

$$\begin{aligned} \left|\det(I - C_0)\right| &= \left|(1 - u_0)(1 - \tilde{u}_0) - z^2 v_0\tilde{v}_0\right| \\ &\ge \bigl(1 - |u_0|\bigr) \times \bigl(1 - |\tilde{u}_0|\bigr) - |z|^2|v_0| \times |\tilde{v}_0|. \end{aligned}$$

In order to control the quantities $u_0, \tilde{u}_0, v_0$ and $\tilde{v}_0$, we shall use the following inequality:

$$\left|\operatorname{Tr} AB^*\right| \le \bigl(\operatorname{Tr} AA^*\bigr)^{1/2} \times \bigl(\operatorname{Tr} BB^*\bigr)^{1/2}, \tag{6.7}$$

together with the following quantities:

$$\begin{cases} u_1 = \frac{1}{n}\operatorname{Tr} DTA\big(I+\delta^*\tilde{D}\big)^{-1}\tilde{D}(I+\delta\tilde{D})^{-1}A^*T^*, \\ \tilde{u}_1 = \frac{1}{n}\operatorname{Tr} \tilde{D}\tilde{T}A^*(I+\tilde{\delta}D)^{-1}D\big(I+\tilde{\delta}^*D\big)^{-1}A\tilde{T}^*, \\ v_1 = \frac{1}{n}\operatorname{Tr} DTDT^*, \\ \tilde{v}_1 = \frac{1}{n}\operatorname{Tr}\tilde{D}\tilde{T}\tilde{D}\tilde{T}^* \end{cases}$$

and

$$\begin{cases} u_2 = \frac{1}{n}\operatorname{Tr} DRA\big(I+\alpha^*\tilde{D}\big)^{-1}\tilde{D}(I+\alpha\tilde{D})^{-1}A^*R^*, \\ \tilde{u}_2 = \frac{1}{n}\operatorname{Tr} \tilde{D}\tilde{R}A^*(I+\tilde{\alpha}D)^{-1}D\big(I+\tilde{\alpha}^*D\big)^{-1}A\tilde{R}^*, \\ v_2 = \frac{1}{n}\operatorname{Tr} DRDR^*, \\ \tilde{v}_2 = \frac{1}{n}\operatorname{Tr}\tilde{D}\tilde{R}\tilde{D}\tilde{R}^*. \end{cases} \tag{6.8}$$

Using (6.7) together with identity $(I+\delta\tilde{D})^{-1}A^*T = \tilde{T}A^*(I+\tilde{\delta}D)^{-1}$ (and similar ones for related quantities), we obtain:

$$|u_0| \le (\tilde{u}_1 u_2)^{1/2}, \qquad |\tilde{u}_0| \le (u_1\tilde{u}_2)^{1/2}, \qquad |v_0| \le (v_1 v_2)^{1/2}, \qquad |\tilde{v}_0| \le (\tilde{v}_1\tilde{v}_2)^{1/2},$$

hence the lower bound:

$$\big|\det(I - C_0)\big| \ge \big(1-(\tilde{u}_1 u_2)^{1/2}\big)\big(1-(u_1\tilde{u}_2)^{1/2}\big) - |z|^2(v_1 v_2 \tilde{v}_1\tilde{v}_2)^{1/2}. \tag{6.9}$$

Notice that it is not proved yet that the right-hand side of the previous inequality is non-negative.

In order to handle estimate (6.9), we shall rely on the following proposition.

Proposition 6.1. *Consider the nonnegative real numbers* x_i, y_i, s_i, t_i $(i = 1, 2)$. *Assume that*:

$$x_i \le 1, \qquad y_i \le 1 \quad \textit{and} \quad (1-x_i)(1-y_i) - s_i t_i \ge 0 \quad \textit{for } i = 1, 2.$$

Then:

$$\big(1-\sqrt{x_1x_2}\big)\big(1-\sqrt{y_1y_2}\big) - \sqrt{s_1s_2t_1t_2} \\ \ge \sqrt{(1-x_1)(1-y_1) - s_1t_1}\sqrt{(1-x_2)(1-y_2) - s_2t_2}.$$

Proof. If $a \ge c$ (≥ 0) and $b \ge d$ (≥ 0), then:

$$\sqrt{ab} - \sqrt{cd} \ge \sqrt{a-c}\sqrt{b-d}.$$

To prove this, simply take the difference of the squares. Applying once this inequality yields $1 - \sqrt{x_1x_2} \ge \sqrt{(1-x_1)(1-x_2)}$, hence:

$$\big(1-\sqrt{x_1x_2}\big)\big(1-\sqrt{y_1y_2}\big) - \sqrt{s_1s_2t_1t_2} \ge \sqrt{(1-x_1)(1-x_2)(1-y_1)(1-y_2)} - \sqrt{s_1s_2t_1t_2}.$$

Applying again the first inequality yields then the desired result. □

Our goal is to apply Proposition 6.1 to (6.9). The main idea, in order to fulfill assumptions of Proposition 6.1 (at least on some portions of $\mathbb{C} - \mathbb{R}^+$), is to consider the quantities of interest, i.e. $u_i, \tilde{u}_i, v_i, \tilde{v}_i$ $(i = 1, 2)$ as coefficients of linear systems whose determinants are the desired quantities $(1-u_i)(1-\tilde{u}_i) - |z|^2 v_i\tilde{v}_i$.

Consider the following matrices:

$$C_i(z) = \begin{pmatrix} u_i & v_i \\ |z|^2\tilde{v}_i & \tilde{u}_i \end{pmatrix}, \quad i = 1, 2.$$

The following proposition holds true:

Proposition 6.2. *Assume that* $z \in \mathbb{C} - \mathbb{R}^+$. *Then*:

(i) *The following holds true*: $1 - u_1(z) \geq 0$ *and* $1 - \tilde{u}_1(z) \geq 0$. *Moreover, there exists positive constants* K, η *such that*:

$$\det\big(I - C_1(z)\big) \geq K \frac{\boldsymbol{\delta}_z^8}{(\eta^2 + |z|^2)^4}.$$

(ii) *There exist nice polynomials* Φ *and* Ψ *and a set*

$$\mathcal{E}_n = \left\{ z \in \mathbb{C}^+,\ \frac{1}{n}\Phi\big(|z|\big)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right) \leq 1/2 \right\},$$

such that for every $z \in \mathcal{E}_n$, $1 - u_2(z) \geq 0$, $1 - \tilde{u}_2(z) \geq 0$, *and*

$$\det(I - C_2) \geq K \frac{\boldsymbol{\delta}_z^8}{(\eta^2 + |z|^2)^4},$$

where K, η *are positive constants*.

Proof of Proposition 6.2 is postponed to Appendix B.

We are now in position to establish the following estimate:

$$\forall z \in \mathcal{E}_n, \quad \max\big(|\alpha - \delta|, |\tilde{\alpha} - \tilde{\delta}|\big) \leq \frac{1}{n}\Phi\big(|z|\big)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right). \tag{6.10}$$

Assume $z \in \mathcal{E}_n$. Thanks to Proposition 6.2, assumptions of Proposition 6.1 are fulfilled by $u_i, \tilde{u}_i, v_i$ and $\tilde{v}_i$, and (6.9) yields:

$$\det(I - C_0) \geq \sqrt{\det(I - C_1)}\sqrt{\det(I - C_2)} \geq K \frac{\boldsymbol{\delta}_z^8}{(\eta^2 + |z|^2)^4}, \tag{6.11}$$

where K, η are nice constants.

Solving now the system (6.4), we obtain:

$$\begin{cases} \alpha - \delta = \big(\det(I - C_0)\big)^{-1}\big((1 - \tilde{u}_0)\boldsymbol{\varepsilon} + z v_0 \tilde{\boldsymbol{\varepsilon}}\big), \\ \tilde{\alpha} - \tilde{\delta} = \big(\det(I - C_0)\big)^{-1}\big((1 - u_0)\tilde{\boldsymbol{\varepsilon}} + z \tilde{v}_0 \boldsymbol{\varepsilon}\big). \end{cases}$$

It remains to use (6.11), Proposition 3.8(ii), and obvious bounds over $u_0, \tilde{u}_0, v_0$ and $\tilde{v}_0$ to conclude and obtain (6.10).

We turn out to the case where $z \in \mathbb{C} - \mathbb{R}^+ - \mathcal{E}_n$, and rely on the same argument as in Haagerup and Thorbjornsen [12] (see also [8]). In this case,

$$\frac{1}{n}\Phi\big(|z|\big)\Psi\big(\boldsymbol{\delta}_z^{-1}\big) \geq \frac{1}{2}.$$

As $|\alpha - \delta| = |n^{-1} \operatorname{Tr} D(\mathbb{E}Q - T)| \leq 2\boldsymbol{\ell}^+ \mathbf{d}_{\max} \boldsymbol{\delta}_z^{-1}$, we obtain:

$$\forall z \in \mathbb{C} - \mathbb{R}^+ - \mathcal{E}_n, \quad |\alpha - \delta| \leq \frac{2\boldsymbol{\ell}^+ \mathbf{d}_{\max}}{\boldsymbol{\delta}_z} \times \frac{2\Phi(|z|)\Psi(1/\boldsymbol{\delta}_z)}{n};$$

a similar estimate holds for $\tilde{\alpha} - \tilde{\delta}$ for $z \notin \mathcal{E}_n$. Gathering the cases where $z \in \mathcal{E}_n$ and $z \notin \mathcal{E}_n$ yields (6.3).

Appendix A: Remaining proofs for Section 3

Proof of Lemma 3.5. Note that it is sufficient to establish the result for a vector u with norm one (which is assumed in the sequel). The general result follows by considering $u/\|u\|$.

We proceed by induction over p. Let $p = 1$ and consider:

$$0 \le \mathbb{E}\sum_{j=1}^{n} \mathbb{E}_{j-1} u^* Q a_j a_j^* Q^* u = \mathbb{E} u^* Q A A^* Q^* u \le \mathbf{a}_{\max}^2 \mathbb{E}\|Q\|^2.$$

As $\|Q\| \le \delta_z^{-1}$, we obtain the desired bound.

Now, write

$$\begin{aligned}
\mathbb{E}\left|\sum_{j=1}^{n} \mathbb{E}_{j-1}\big(u^* Q a_j a_j^* Q^* u\big)\right|^p &= \sum_{j_1,\dots,j_p} \mathbb{E}\big[\mathbb{E}_{j_1-1}\big(u^* Q a_{j_1} a_{j_1}^* Q^* u\big)\cdots \mathbb{E}_{j_p-1}\big(u^* Q a_{j_p} a_{j_p}^* Q^* u\big)\big] \\
&\le p! \sum_{j_1\le\cdots\le j_p} \mathbb{E}\big[\mathbb{E}_{j_1-1}\big(u^* Q a_{j_1} a_{j_1}^* Q^* u\big)\cdots \mathbb{E}_{j_p-1}\big(u^* Q a_{j_p} a_{j_p}^* Q^* u\big)\big] \\
&= p! \sum_{j_1\le\cdots\le j_p} \mathbb{E}\Bigg[\mathbb{E}_{j_p-1}\big(u^* Q a_{j_p} a_{j_p}^* Q^* u\big) \underbrace{\prod_{k=1}^{p-1} \mathbb{E}_{j_k-1}\big(u^* Q a_{j_k} a_{j_k}^* Q^* u\big)}_{\mathcal{F}_{j_p-1}\ \text{measurable}}\Bigg] \\
&= p! \sum_{j_1\le\cdots\le j_{p-1}} \mathbb{E}\Bigg[\sum_{j_p=j_{p-1}}^{n} \big(u^* Q a_{j_p} a_{j_p}^* Q^* u\big) \prod_{k=1}^{p-1} \mathbb{E}_{j_k-1}\big(u^* Q a_{j_k} a_{j_k}^* Q^* u\big)\Bigg] \\
&\overset{(a)}{\le} p! \frac{\mathbf{a}_{\max}^2}{\delta_z^2} \mathbb{E}\left|\sum_{j=1}^{n} \mathbb{E}_{j-1}\big(u^* Q a_j a_j^* Q^* u\big)\right|^{p-1},
\end{aligned}$$

where (a) follows from the fact that

$$\sum_{j_p=j_{p-1}}^{n} \big(u^* Q a_{j_p} a_{j_p}^* Q^* u\big) \le \sum_{j_p=1}^{n} \big(u^* Q a_{j_p} a_{j_p}^* Q^* u\big) \le \frac{\mathbf{a}_{\max}^2}{\delta_z^2}.$$

It remains to plug the induction assumption to conclude. Hence (3.9) is established.

In order to establish (3.10), one may use the same arguments as previously together with the identity $Q\Sigma\Sigma^* = I + zQ$, which yields the factor $|z|^p$ in estimate (3.10). □

Proof of Lemma 3.6. We prove the lemma in the case where $\|u\| = 1$, the general result readily follows by considering $u/\|u\|$.

Write $u^* Q_j a_j a_j^* Q_j^* u = \chi_{1j} + \chi_{2j} + \chi_{3j} + \chi_{4j}$ with:

$$\begin{aligned}
\chi_{1j} &= u^*(Q_j - Q) a_j a_j^* (Q_j - Q)^* u, \\
\chi_{2j} &= u^* Q a_j a_j^* Q^* u, \\
\chi_{3j} &= u^*(Q_j - Q) a_j a_j^* Q^* u, \\
\chi_{4j} &= u^* Q a_j a_j^* (Q_j - Q)^* u.
\end{aligned}$$

Hence,

$$\sum_{j=1}^{n}\mathbb{E}\big(u^* Q_j a_j a_j^* Q_j^* u\big)^2 \le \sum_{j=1}^{n}\mathbb{E}\chi_{1j}^2 + \sum_{j=1}^{n}\mathbb{E}\chi_{2j}^2 + \sum_{j=1}^{n}\mathbb{E}|\chi_{3j}|^2 + \sum_{j=1}^{n}\mathbb{E}|\chi_{4j}|^2.$$

Notice that:

$$\mathbb{E}|\chi_{3j}|^2 \le \frac{1}{2}\big(\mathbb{E}\chi_{1j}^2 + \mathbb{E}\chi_{2j}^2\big) \quad \text{and} \quad \mathbb{E}|\chi_{4j}|^2 \le \frac{1}{2}\big(\mathbb{E}\chi_{1j}^2 + \mathbb{E}\chi_{2j}^2\big).$$

Note that using the facts that $a_j a_j^* \le AA^*$ and $\eta_j \eta_j^* \le \Sigma\Sigma^*$ together with the identity $Q\Sigma\Sigma^* = I + zQ$ yield the rough but useful estimates:

$$u^* Q a_j a_j^* Q^* u = \mathcal{O}\big(\boldsymbol{\delta}_z^{-2}\big) \quad \text{and} \quad u^* Q \eta_j \eta_j^* Q^* u = \mathcal{O}\left(\frac{|z|}{\boldsymbol{\delta}_z^2}\right). \tag{A.1}$$

We first begin by the contribution of $\sum_j \mathbb{E}\chi_{2j}^2$:

$$\begin{aligned}
\sum_{j=1}^{n}\chi_{2j}^2 &= \sum_{j=1}^{n} u^* Q a_j a_j^* Q^* u \times u^* Q a_j a_j^* Q^* u \\
&\le \sum_{j=1}^{n} u^* Q a_j a_j^* Q^* u \times u^* Q A A^* Q^* u \\
&\le \big(u^* Q A A^* Q^* u\big)^2 = \mathcal{O}\big(\boldsymbol{\delta}_z^{-4}\big) \le \Phi_2\big(|z|\big)\Psi_2\left(\frac{1}{\boldsymbol{\delta}_z}\right).
\end{aligned} \tag{A.2}$$

Similarly,

$$\sum_{j=1}^{n}\big(u^* Q \eta_j \eta_j^* Q^* u\big)^2 = \mathcal{O}\left(\frac{|z|^2}{\boldsymbol{\delta}_z^4}\right). \tag{A.3}$$

We now turn to the contribution of $\sum_j \mathbb{E}\chi_{1j}^2$. Using the decompositions (3.2), (3.3) and (3.4), χ_{1j} writes:

$$\begin{aligned}
\chi_{1j} &= \left|\frac{1+\eta_j^* Q_j \eta_j}{1-\eta_j^* Q \eta_j}\right| \times \big|u^* Q \eta_j \eta_j^* Q a_j a_j^* Q^* \eta_j \eta_j^* Q^* u\big| \\
&= \big|1+\eta_j^* Q_j \eta_j\big| \times \big|u^* Q \eta_j \eta_j^* Q^* u\big| \times \left|\frac{a_j^* Q^* \eta_j \eta_j^* Q a_j}{1-\eta_j^* Q \eta_j}\right|.
\end{aligned} \tag{A.4}$$

We first prove that

$$\frac{a_j^* Q^* \eta_j \eta_j^* Q a_j}{1-\eta_j^* Q \eta_j} = \mathcal{O}\left(\frac{|z|}{\boldsymbol{\delta}_z^2}\right). \tag{A.5}$$

In fact:

$$\begin{aligned}
\left|\frac{a_j^* Q^* \eta_j \eta_j^* Q a_j}{1-\eta_j^* Q \eta_j}\right| &\le \left|\frac{a_j^* Q^* \eta_j \eta_j^* Q^* a_j}{1-\eta_j^* Q \eta_j}\right| + \left|\frac{a_j^* Q^* \eta_j \eta_j^* (Q-Q^*) a_j}{1-\eta_j^* Q \eta_j}\right| \\
&\overset{(a)}{\le} \big|a_j^* (Q_j - Q)^* a_j\big| + 2\big|\mathrm{Im}(z)\big|\big|a_j^* (Q_j - Q) Q a_j\big| \\
&= \mathcal{O}\left(\frac{1}{\boldsymbol{\delta}_z}\right) + \mathcal{O}\left(\frac{|z|}{\boldsymbol{\delta}_z^2}\right) = \mathcal{O}\left(\frac{|z|}{\boldsymbol{\delta}_z^2}\right),
\end{aligned}$$

where we use the fact that $Q - Q^* = 2\mathbf{i}\,\mathrm{Im}(z) Q^* Q$ to obtain (a). Now,

$$\big|1 + \eta_j^* Q_j \eta_j\big| \le 1 + |\Delta_j| + \left|\frac{\tilde{d}_j}{n} \operatorname{Tr} D Q_j + a_j^* Q_j a_j\right|. \tag{A.6}$$

Since $|n^{-1}\tilde{d}_j \operatorname{Tr} D Q_j + a_j^* Q_j a_j| = \mathcal{O}(\boldsymbol{\delta}_z^{-1})$, we obtain:

$$\begin{aligned}
\sum_{j=1}^n \mathbb{E}\chi_{1j}^2 &= \left(\mathcal{O}\left(\frac{|z|^2}{\boldsymbol{\delta}_z^4}\right) + \mathcal{O}\left(\frac{|z|^2}{\boldsymbol{\delta}_z^6}\right)\right) \times \sum_{j=1}^n \mathbb{E}\big(u^* Q \eta_j \eta_j^* Q^* u\big)^2 \\
&\quad + \mathcal{O}\left(\frac{|z|^2}{\boldsymbol{\delta}_z^4}\right) \times \sum_{j=1}^n \mathbb{E}\big(u^* Q \eta_j \eta_j^* Q^* u\big)^2 \times |\Delta_j|^2 \\
&\overset{(a)}{=} \mathcal{O}\left(\frac{|z|^4}{\boldsymbol{\delta}_z^8}\right) + \mathcal{O}\left(\frac{|z|^4}{\boldsymbol{\delta}_z^{10}}\right) + \mathcal{O}\left(\frac{|z|^4}{\boldsymbol{\delta}_z^8}\right) \times \sum_{j=1}^n \mathbb{E}|\Delta_j|^2 \\
&\overset{(b)}{=} \mathcal{O}\left(\frac{|z|^4}{\boldsymbol{\delta}_z^8}\right) + \mathcal{O}\left(\frac{|z|^4}{\boldsymbol{\delta}_z^{10}}\right) \\
&\le \Phi_1\big(|z|\big)\Psi_1\left(\frac{1}{\boldsymbol{\delta}_z}\right),
\end{aligned}$$

where (a) follows from (A.3) and (A.1) and (b), from Corollary 3.2.

It remains to gather the contributions of χ_{1j}, χ_{2j}, χ_{3j} and χ_{4j} to get:

$$\sum_{j=1}^n \mathbb{E}\big(u^* Q_j a_j a_j^* Q_j^* u\big)^2 \le 2\Phi_1\big(|z|\big)\Psi_1\left(\frac{1}{\boldsymbol{\delta}_z}\right) + 2\Phi_2\big(|z|\big)\Psi_2\left(\frac{1}{\boldsymbol{\delta}_z}\right) \overset{(a)}{\le} \Phi\big(|z|\big)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right),$$

where (a) follows from (3.7). Equation (3.11) is proved.

In order to prove (3.12), first note that:

$$\begin{aligned}
&\mathbb{E}\left(\sum_{j=1}^n \mathbb{E}_{j-1}\big(u^* Q_j a_j a_j^* Q_j^* u\big)\right)^p \\
&\quad \le K\left(\mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\chi_{1j}\right|^p + \mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\chi_{2j}\right|^p + \mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\chi_{3j}\right|^p + \mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\chi_{4j}\right|^p\right).
\end{aligned}$$

Hence, it remains to evaluate the contributions of each term. Using decomposition (A.4) together with the estimate (A.5), we obtain:

$$\mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\chi_{1j}\right|^p = \mathcal{O}\left(\frac{|z|^p}{\boldsymbol{\delta}_z^{2p}}\right) \times \mathbb{E}\left(\sum_{j=1}^n \mathbb{E}_{j-1}\big|1 + \eta_j^* Q_j \eta_j\big| \times u^* Q \eta_j \eta_j^* Q^* u\right)^p.$$

Using (A.6) together with (3.10) yields:

$$\mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\chi_{1j}\right|^p = \mathcal{O}\left(\frac{|z|^{2p}}{\boldsymbol{\delta}_z^{4p}}\right) + \mathcal{O}\left(\frac{|z|^{2p}}{\boldsymbol{\delta}_z^{5p}}\right) + \mathcal{O}\left(\frac{|z|^p}{\boldsymbol{\delta}_z^{2p}}\right) \times \mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\big(|\Delta_j| \times u^* Q \eta_j \eta_j^* Q^* u\big)\right|^p.$$

Combining standard inequalities (Cauchy–Schwarz, $|\sum_j a_j b_j| \le (\sum_j a_j^2)^{1/2}(\sum_j b_j^2)^{1/2}$, and Cauchy–Schwarz again), we obtain:

$$\mathbb{E}\left(\sum_{j=1}^n \mathbb{E}_{j-1}\big(|\Delta_j| \times u^* Q\eta_j\eta_j^* Q^* u\big)\right)^p$$
$$\le \left[\mathbb{E}\left(\sum_{j=1}^n \mathbb{E}_{j-1}\big(u^* Q\eta_j\eta_j^* Q^* u\big)^2\right)^p \times \mathbb{E}\left(\sum_{j=1}^n \mathbb{E}_{j-1}|\Delta_j|^2\right)^p\right]^{1/2} \overset{\text{(a)}}{=} \mathcal{O}\left(\frac{|z|^p}{\boldsymbol{\delta}_z^{3p}}\right),$$

where (a) follows from (A.1), Corollary 3.2 and (3.10). Finally,

$$\mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\chi_{1j}\right|^p = \mathcal{O}\left(\frac{|z|^{2p}}{\boldsymbol{\delta}_z^{4p}}\right) + \mathcal{O}\left(\frac{|z|^{2p}}{\boldsymbol{\delta}_z^{5p}}\right) + \mathcal{O}\left(\frac{|z|^{2p}}{\boldsymbol{\delta}_z^{5p}}\right) \le \Phi_1\big(|z|\big)\Psi_1\big(\boldsymbol{\delta}_z^{-1}\big). \tag{A.7}$$

Equation (3.9) directly yields the estimate:

$$\mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\chi_{2j}\right|^p = \mathcal{O}\left(\frac{1}{\boldsymbol{\delta}_z^{2p}}\right) \le \Phi_2\big(|z|\big)\Psi_2\big(\boldsymbol{\delta}_z^{-1}\big). \tag{A.8}$$

Finally,

$$\mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\chi_{3j}\right|^p \le \left(\mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\chi_{1j}\right|^p \mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\chi_{2j}\right|^p\right)^{1/2} \le \Phi_3\big(|z|\big)\Psi_3\left(\frac{1}{\boldsymbol{\delta}_z}\right). \tag{A.9}$$

A corresponding inequality exists for $\mathbb{E}|\sum \mathbb{E}_{j-1}\chi_{4j}|^p$, obtain:

$$\mathbb{E}\left|\sum_{j=1}^n \mathbb{E}_{j-1}\chi_{4j}\right|^p \le \Phi_4\big(|z|\big)\Psi_4\left(\frac{1}{\boldsymbol{\delta}_z}\right). \tag{A.10}$$

Gathering (A.7), (A.8), (A.9) and (A.10), we end up with (3.12), and Lemma 3.6 is proved. □

Appendix B: Remaining proofs for Section 6

Proof of Proposition 6.2(i). Recall that $\delta = \frac{1}{n}\operatorname{Tr} DT$ and $\tilde\delta = \frac{1}{n}\operatorname{Tr}\tilde D\tilde T$. We consider first the case where $z \in \mathbb{C}^+ \cup \mathbb{C}^-$. We have

$$\operatorname{Im}(\delta) = \frac{1}{2\mathbf{i}n}\operatorname{Tr} DT\big(T^{-*} - T^{-1}\big)T^* \quad \text{and} \quad \operatorname{Im}(z\tilde\delta) = \frac{1}{2\mathbf{i}n}\operatorname{Tr}\tilde D(z\tilde T)\big[(z\tilde T)^{-*} - (z\tilde T)^{-1}\big](z\tilde T)^*.$$

Developing the previous identities, we end up with the system:

$$(I - C_1)\begin{pmatrix}\operatorname{Im}(\delta)\\ \operatorname{Im}(z\tilde\delta)\end{pmatrix} = \operatorname{Im}(z)\begin{pmatrix} w_1(z)\\ \tilde x_1(z)\end{pmatrix}, \tag{B.1}$$

where

$$\begin{cases} w_1(z) = \frac{1}{n}\operatorname{Tr} DTT^* & (>0),\\ \tilde x_1(z) = \frac{1}{n}\operatorname{Tr}\tilde D\tilde T A^*(I+\tilde\delta D)^{-1}\big(I+\tilde\delta^* D\big)^{-1}A\tilde T^* & (>0).\end{cases}$$

By developing the first equation of this system, and by recalling that $\delta(z)$ is the Stieltjes transform of a positive measure μ_n with support included in $\mathbb{R}^+$, we obtain

$$1-u_1 = w_1 \frac{\mathrm{Im}(z)}{\mathrm{Im}(\delta)} + v_1 \frac{\mathrm{Im}(z\tilde{\delta})}{\mathrm{Im}(\delta)} \geq w_1 \frac{\mathrm{Im}(z)}{\mathrm{Im}(\delta)} \geq 0.$$

Replacing $(\mathrm{Im}(\delta), \mathrm{Im}(z\tilde{\delta}))$ with $(\mathrm{Im}(\tilde{\delta}), \mathrm{Im}(z\delta))$ and repeating the same argument, we obtain

$$1-\tilde{u}_1 = \tilde{w}_1 \frac{\mathrm{Im}(z)}{\mathrm{Im}(\tilde{\delta})} + \tilde{v}_1 \frac{\mathrm{Im}(z\delta)}{\mathrm{Im}(\tilde{\delta})} \geq \tilde{w}_1 \frac{\mathrm{Im}(z)}{\mathrm{Im}(\tilde{\delta})} \geq 0.$$

By continuity of $u_1(z)$ and $\tilde{u}_1(z)$ at any point of the open real negative axis, we have $1-u_1 \geq 0$ and $1-\tilde{u}_1 \geq 0$ for any $z \in \mathbb{C} - \mathbb{R}^+$. The first two inequalities in the statement of Proposition 6.2(i) are proven.

By applying Cramer's rule ([16], Section 0.8.3) where the first column of $I - C_1$ is replaced with the right-hand member of (B.1), we obtain

$$\det(I-C_1) = (1-\tilde{u}_1) w_1 \frac{\mathrm{Im}(z)}{\mathrm{Im}(\delta)} + v_1 \tilde{x}_1 \frac{\mathrm{Im}(z)}{\mathrm{Im}(\delta)} \geq (1-\tilde{u}_1) w_1 \frac{\mathrm{Im}(z)}{\mathrm{Im}(\delta)} \geq w_1 \tilde{w}_1 \frac{\mathrm{Im}(z)}{\mathrm{Im}(\delta)} \frac{\mathrm{Im}(z)}{\mathrm{Im}(\tilde{\delta})}. \tag{B.2}$$

Using the fact that the positive measure μ_n is supported by $\mathbb{R}^+$ and has a total mass $n^{-1} \operatorname{Tr} D$, we have

$$0 \leq \frac{\mathrm{Im}(\delta)}{\mathrm{Im}(z)} = \int \frac{1}{|t-z|^2} \mu_n(\mathrm{d}t) \leq \frac{1}{\boldsymbol{\delta}_z^2} \frac{1}{n} \operatorname{Tr} D \leq \frac{\boldsymbol{\ell}^+ \mathbf{d}_{\max}}{\boldsymbol{\delta}_z^2} \quad \text{and} \quad 0 \leq \frac{\mathrm{Im}(\tilde{\delta})}{\mathrm{Im}(z)} \leq \frac{\tilde{\mathbf{d}}_{\max}}{\boldsymbol{\delta}_z^2}. \tag{B.3}$$

In order to find a lower bound on w_1 and $\tilde{w}_1$, we begin by finding a lower bound on $|\delta|$.

A computation similar to [15], Lemma C.1, shows that the sequence of measures (μ_n) is tight. Hence there exists $\eta > 0$ such that:

$$\mu_n[0,\eta] \geq \frac{1}{2} \frac{1}{n} \operatorname{Tr} D \geq \frac{\boldsymbol{\ell}^- \mathbf{d}_{\min}}{2}.$$

We have

$$|\delta| \geq \left|\mathrm{Im}(\delta)\right| = \left|\mathrm{Im}(z)\right| \int \frac{\mu_n(\mathrm{d}t)}{|t-z|^2} \geq \left|\mathrm{Im}(z)\right| \int_0^\eta \frac{\mu_n(\mathrm{d}t)}{2(t^2+|z|^2)} \geq \left|\mathrm{Im}(z)\right| \frac{\boldsymbol{\ell}^- \mathbf{d}_{\min}}{4(\eta^2+|z|^2)}. \tag{B.4}$$

Furthermore, when $\mathrm{Re}(z) < 0$, we have

$$|\delta| \geq \mathrm{Re}(\delta) = \int \frac{t - \mathrm{Re}(z)}{|t-z|^2} \mu_n(\mathrm{d}t) \geq -\mathrm{Re}(z) \int \frac{\mu_n(\mathrm{d}t)}{|t-z|^2} \geq -\mathrm{Re}(z) \frac{\boldsymbol{\ell}^- \mathbf{d}_{\min}}{4(\eta^2+|z|^2)},$$

which results in

$$|\delta| \geq \boldsymbol{\delta}_z \frac{\boldsymbol{\ell}^- \mathbf{d}_{\min}}{4(\eta^2+|z|^2)}.$$

We can now find a lower bound to w_1:

$$\begin{aligned}
w_1 &= \frac{1}{n} \operatorname{Tr} DTT^* = \frac{1}{n} \sum_{i=1}^N d_i \sum_{j=1}^N |T_{ij}|^2 = \frac{1}{n} \operatorname{Tr} D \sum_{i=1}^N \kappa_i \sum_{j=1}^N |T_{ij}|^2 \quad \text{with } \kappa_i = \frac{d_i}{\operatorname{Tr} D} \\
&\overset{(a)}{\geq} \frac{1}{n} \operatorname{Tr} D \left(\sum_{i=1}^N \kappa_i \left(\sum_{j=1}^N |T_{ij}|^2 \right)^{1/2} \right)^2 \geq \frac{1}{n} \operatorname{Tr} D \left(\sum_{i=1}^N \kappa_i |T_{ii}| \right)^2 \geq \frac{1}{n} \operatorname{Tr} D \left| \sum_{i=1}^N \kappa_i T_{ii} \right|^2 \\
&= \frac{|\delta|^2}{(1/n) \operatorname{Tr} D} \geq \frac{(\boldsymbol{\delta}_z \boldsymbol{\ell}^- \mathbf{d}_{\min})^2}{16 \boldsymbol{\ell}^+ \mathbf{d}_{\max} (\eta^2+|z|^2)^2},
\end{aligned}$$

where (a) follows by convexity. A similar computation yields $\tilde{w}_1 \geq (\boldsymbol{\delta}_z \tilde{\mathbf{d}}_{\min})^2/(16\tilde{\mathbf{d}}_{\max}(\tilde{\eta}^2 + |z|^2)^2)$ where $\tilde{\eta}$ is a positive constant. Grouping these estimates with those in (B.3) and plugging them into (B.2), we obtain

$$\begin{aligned}\det(I - C_1) &\geq \frac{\boldsymbol{\delta}_z^8(\boldsymbol{\ell}^- \tilde{\mathbf{d}}_{\min}\tilde{\mathbf{d}}_{\min})^2}{256(\boldsymbol{\ell}^+\mathbf{d}_{\max}\tilde{\mathbf{d}}_{\max})^2(\eta^2 + |z|^2)^2(\tilde{\eta}^2 + |z|^2)^2} \\ &\geq K \frac{\boldsymbol{\delta}_z^8}{(\max(\eta, \tilde{\eta})^2 + |z|^2)^4},\end{aligned}$$

where K is a nice constant. The same bound holds for $z \in (-\infty, 0)$ by continuity of $\det(I - C_1(z))$ at any point of the open real negative axis. □

Proof of Proposition 6.2(ii). Recall that

$$\boldsymbol{\varepsilon}_n = \frac{1}{n} \operatorname{Tr} D(\mathbb{E}Q - R).$$

We first establish useful estimates.

Lemma B.1. *There exists nice polynomials Φ and Ψ such that*:

$$\left|\frac{\operatorname{Im}(\boldsymbol{\varepsilon}_n(z))}{\operatorname{Im}(z)}\right| \leq \frac{1}{n}\Phi(|z|)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right) \quad \textit{and} \quad \left|\frac{\operatorname{Im}(z\boldsymbol{\varepsilon}_n(z))}{\operatorname{Im}(z)}\right| \leq \frac{1}{n}\Phi(|z|)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right) \quad \textit{for } z \in \mathbb{C} - \mathbb{R}^+.$$

Proof. We prove the first inequality. By Proposition 3.8(ii), the sequence of functions $(\boldsymbol{\varepsilon}_n)$ satisfies over $\mathbb{C} - \mathbb{R}_+$

$$|\boldsymbol{\varepsilon}_n(z)| \leq \frac{1}{n}\Phi(|z|)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right),$$

where Φ and Ψ are nice polynomials. Let $\mathcal{R}$ be the region of the complex plane defined as $\mathcal{R} = \{z\colon \operatorname{Re}(z) < 0, |\operatorname{Im}(z)| < -\operatorname{Re}(z)/2\}$. If $z \in \mathbb{C} - \mathbb{R}^+ - \mathcal{R}$, then $|\operatorname{Im}(z)| \geq \boldsymbol{\delta}_z/\sqrt{5}$, therefore $|\operatorname{Im}\boldsymbol{\varepsilon}(z)/\operatorname{Im} z| \leq n^{-1}\sqrt{5}\boldsymbol{\delta}_z^{-1}\Phi(|z|) \times \Psi(\boldsymbol{\delta}_z^{-1})$ and the result is proven. Assume now that $z \in \mathcal{R}$. In this case, z belongs to the open disc $\mathcal{D}_z$ centered at $\operatorname{Re}(z)$ with radius $-\operatorname{Re}(z)/2$. For any $u \in \mathcal{D}_z$, we have $|\boldsymbol{\varepsilon}(u)| \leq n^{-1}\Phi(|u|)\Psi(|u|^{-1})$. Moreover,

$$\forall u \in \mathcal{D}_z, \quad \frac{\boldsymbol{\delta}_z}{\sqrt{5}} \leq -\frac{\operatorname{Re}(z)}{2} \leq |u| \leq -\frac{3\operatorname{Re}(z)}{2} \leq \frac{3|z|}{2}.$$

As $\Phi(x)$ is increasing and $\Psi(1/x)$ is decreasing in $x > 0$, we obtain:

$$|\boldsymbol{\varepsilon}_n(u)| \leq \frac{1}{n}\Phi\left(\frac{3|z|}{2}\right)\Psi\left(\frac{\sqrt{5}}{\boldsymbol{\delta}_z}\right) \quad \text{for } u \in \mathcal{D}_z. \tag{B.5}$$

The function $\boldsymbol{\varepsilon}$ is holomorphic on $\mathcal{D}_z$. Consider the function: Applying Lemma 3.4 with

$$f(\zeta) = \frac{\boldsymbol{\varepsilon}(|\operatorname{Re}(z)/2|\zeta + \operatorname{Re}(z)) - \boldsymbol{\varepsilon}(\operatorname{Re}(z))}{\sup_{u\in\mathcal{D}_z}|\boldsymbol{\varepsilon}(u) - \boldsymbol{\varepsilon}(\operatorname{Re}(z))|}.$$

Let $\zeta = \mathbf{i}2\operatorname{Im}(z)/\operatorname{Re}(z)$, apply Lemma 3.4, and use (B.5). This yields:

$$|\boldsymbol{\varepsilon}(z) - \boldsymbol{\varepsilon}(\operatorname{Re}(z))| \leq \frac{2|\operatorname{Im}(z)|}{|\operatorname{Re}(z)|} \times \frac{1}{n}\Phi(|z|)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right) \leq \frac{\sqrt{5}|\operatorname{Im}(z)|}{\boldsymbol{\delta}_z} \times \frac{1}{n}\Phi(|z|)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right),$$

where Φ and Ψ are nice polynomials. As $\operatorname{Im}(\boldsymbol{\varepsilon}(\operatorname{Re}(z))) = 0$, we obtain

$$\left|\frac{\operatorname{Im}(\boldsymbol{\varepsilon}_n(z))}{\operatorname{Im}(z)}\right| \leq \left|\frac{\boldsymbol{\varepsilon}(z) - \boldsymbol{\varepsilon}(\operatorname{Re}(z))}{\operatorname{Im}(z)}\right| \leq \frac{\sqrt{5}}{\boldsymbol{\delta}_z n}\Phi(|z|)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right).$$

This proves the first inequality. The second one can be proved similarly. □

We now tackle the proof of Proposition 6.2(ii), following closely the line of the proof of Proposition 6.2(i). Recall that $\alpha = \frac{1}{n}\operatorname{Tr} D\mathbb{E}Q$, $\tilde{\alpha} = \frac{1}{n}\operatorname{Tr} \tilde{D}\mathbb{E}\tilde{Q}$, $\boldsymbol{\varepsilon} = \frac{1}{n}\operatorname{Tr} D(\mathbb{E}Q - R)$ and $\tilde{\boldsymbol{\varepsilon}} = \frac{1}{n}\operatorname{Tr} \tilde{D}(\mathbb{E}\tilde{Q} - \tilde{R})$. We begin by establishing the lower bound on $\det(I - C_2)$. Assume that $z \in \mathbb{C}^+ \cup \mathbb{C}^-$. Writing $\alpha = \frac{1}{n}\operatorname{Tr} DR + \boldsymbol{\varepsilon}$ and $\tilde{\alpha} = \frac{1}{n}\operatorname{Tr} \tilde{D}\tilde{R} + \tilde{\boldsymbol{\varepsilon}}$ and developing $\operatorname{Im}(\alpha)$ and $\operatorname{Im}(z\tilde{\alpha})$ with the help of the resolvent identity, we get the following system:

$$(I - C_2)\begin{pmatrix} \operatorname{Im}(\alpha) \\ \operatorname{Im}(z\tilde{\alpha}) \end{pmatrix} = \operatorname{Im}(z)\begin{pmatrix} w_2(z) \\ \tilde{x}_2(z) \end{pmatrix} + \begin{pmatrix} \operatorname{Im}(\boldsymbol{\varepsilon}) \\ \operatorname{Im}(z\tilde{\boldsymbol{\varepsilon}}) \end{pmatrix},$$

where $w_2(z) = \frac{1}{n}\operatorname{Tr} DRR^*$ and $\tilde{x}_2(z) > 0$. Let $\tilde{w}_2 = n^{-1}\operatorname{Tr} \tilde{D}\tilde{R}\tilde{R}^*$. Using the same arguments as in the proof of Proposition 6.2(i), we obtain

$$1 - u_2 = w_2 \frac{\operatorname{Im}(z)}{\operatorname{Im}(\alpha)} + v_2 \frac{\operatorname{Im}(z\tilde{\alpha})}{\operatorname{Im}(\alpha)} + \frac{\operatorname{Im}(\boldsymbol{\varepsilon})}{\operatorname{Im}(\alpha)} \ge w_2 \frac{\operatorname{Im}(z)}{\operatorname{Im}(\alpha)} + \frac{\operatorname{Im}(\boldsymbol{\varepsilon})}{\operatorname{Im}(\alpha)}, \tag{B.6}$$

$$1 - \tilde{u}_2 = \tilde{w}_2 \frac{\operatorname{Im}(z)}{\operatorname{Im}(\tilde{\alpha})} + \tilde{v}_2 \frac{\operatorname{Im}(z\alpha)}{\operatorname{Im}(\tilde{\alpha})} + \frac{\operatorname{Im}(\tilde{\boldsymbol{\varepsilon}})}{\operatorname{Im}(\tilde{\alpha})} \ge \tilde{w}_2 \frac{\operatorname{Im}(z)}{\operatorname{Im}(\tilde{\alpha})} + \frac{\operatorname{Im}(\tilde{\boldsymbol{\varepsilon}})}{\operatorname{Im}(\tilde{\alpha})}, \tag{B.7}$$

$$\begin{aligned} \det(I - C_2) &\ge w_2\tilde{w}_2 \frac{\operatorname{Im}(z)}{\operatorname{Im}(\alpha)} \frac{\operatorname{Im}(z)}{\operatorname{Im}(\tilde{\alpha})} + w_2 \frac{\operatorname{Im}(z)}{\operatorname{Im}(\alpha)} \frac{\operatorname{Im}(\tilde{\boldsymbol{\varepsilon}})}{\operatorname{Im}(\tilde{\alpha})} + (1 - \tilde{u}_2)\frac{\operatorname{Im}(\boldsymbol{\varepsilon})}{\operatorname{Im}(\alpha)} + v_2 \frac{\operatorname{Im}(z\tilde{\boldsymbol{\varepsilon}})}{\operatorname{Im}(\alpha)} \\ &\stackrel{\triangle}{=} w_2\tilde{w}_2 \frac{\operatorname{Im}(z)}{\operatorname{Im}(\alpha)} \frac{\operatorname{Im}(z)}{\operatorname{Im}(\tilde{\alpha})} + e(z). \end{aligned} \tag{B.8}$$

We now find an upper bound on the perturbation term $e(z)$. To this end, we have $0 \le w_2 \le \boldsymbol{\ell}^+ \mathbf{d}_{\max}/\boldsymbol{\delta}_z^2$ and $0 \le v_2 \le \boldsymbol{\ell}^+ \mathbf{d}_{\max}^2/\boldsymbol{\delta}_z^2$. Recalling (6.8), we also have

$$|1 - \tilde{u}_2| \le 1 + \frac{\mathbf{d}_{\max}\tilde{\mathbf{d}}_{\max}\mathbf{a}_{\max}^2 |z|^2}{\boldsymbol{\delta}_z^4}.$$

Using the same arguments as in the proof of Proposition 6.2(i) (involving this time the tightness of the measures associated with the Stieltjes transforms $\frac{1}{n}\operatorname{Tr} DR$ and $\frac{1}{n}\operatorname{Tr} \tilde{D}\tilde{R}$) yields:

$$\frac{\operatorname{Im}(z)}{\operatorname{Im}(\alpha)} \le \frac{4(\eta^2 + |z|^2)}{\boldsymbol{\ell}^- \mathbf{d}_{\min}}, \qquad \bigl|e(z)\bigr| \le \frac{1}{n}\Phi\bigl(|z|\bigr)\Psi\bigl(\boldsymbol{\delta}_z^{-1}\bigr), \qquad \frac{\operatorname{Im}(z)}{\operatorname{Im}(\alpha)} \le \frac{4(\eta^2 + |z|^2)}{\boldsymbol{\ell}^- \mathbf{d}_{\min}}$$

for every $z \in \mathbb{C}^+ \cup \mathbb{C}^-$, where η, K are positive constants, and Φ and Ψ, nice polynomials.

Finally, we can state that there exist nice polynomials Φ and Ψ such that:

$$\det(I - C_2) \ge K \frac{\boldsymbol{\delta}_z^8}{(\eta^2 + |z|^2)^4}\left(1 - \frac{1}{n}\Phi\bigl(|z|\bigr)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right)\right).$$

By continuity of $\det(I - C_2(z))$ at any point of the open real negative axis, this inequality is true for any $z \in \mathbb{C} - \mathbb{R}^+$. Denote by $\mathcal{E}_n$ the set:

$$\mathcal{E}_n = \left\{ z \in \mathbb{C} - \mathbb{R}^+,\ \frac{1}{n}\Phi\bigl(|z|\bigr)\Psi\left(\frac{1}{\boldsymbol{\delta}_z}\right) \le 1/2 \right\}.$$

If $z \in \mathcal{E}_n$, then $\det(I - C_2)$ is readily lower-bounded by the quantity stated in Proposition 6.2(ii).

By considering inequalities (B.6) and (B.7) and by possibly modifying the polynomials Φ and Ψ, we have $1 - u_2 \ge 0$ and $1 - \tilde{u}_2 \ge 0$ for $z \in \mathcal{E}_n$. The proof of Proposition 6.2(ii) is completed. □

Acknowledgment

This work was partially supported by the Agence Nationale de la Recherche (France), project SESAME n°ANR-07-MDCO-012-01.

References

[1] C. Artigue and P. Loubaton. On the precoder design of flat fading MIMO systems equipped with MMSE receivers: A large-system approach. *IEEE Trans. Inform. Theory* **57** (2011) 4138–4155.

[2] Z. D. Bai, B. Q. Miao and G. M. Pan. On asymptotics of eigenvectors of large sample covariance matrix. *Ann. Probab.* **35** (2007) 1532–1572. MR2330979

[3] 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** (1998) 316–345. MR1617051

[4] Z. D. Bai and J. W. Silverstein. Exact separation of eigenvalues of large-dimensional sample covariance matrices. *Ann. Probab.* **27** (1999) 1536–1555. MR1733159

[5] 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** (2012) 1150004. MR2930382

[6] F. Benaych-Georges and R. N. Rao. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. Preprint, 2009. Available at arXiv:0910.2120.

[7] 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** (2007) 678–694. MR2322123

[8] M. Capitaine, C. Donati-Martin and D. Féral. The largest eigenvalues of finite rank deformation of large Wigner matrices: Convergence and nonuniversality of the fluctuations. *Ann. Probab.* **37** (2009) 1–47. MR2489158

[9] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton and J. Najim. On the capacity achieving covariance matrix for Rician MIMO channels: An asymptotic approach. *IEEE Trans. Inform. Theory* **56** (2010) 1048–1069. MR2723661

[10] L. Erdös, H.-T. Yau and J. Yin. Rigidity of eigenvalues of generalized Wigner matrices. Unpublished manuscript, 2010. Available at http://arxiv.org/pdf/1007.4652.

[11] V. L. Girko. *An Introduction to Statistical Analysis of Random Arrays*. VSP, Utrecht, 1998. MR1694087

[12] U. Haagerup and S. Thorbjørnsen. A new application of random matrices: $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ is not a group. *Ann. of Math. (2)* **162** (2005) 711–775. MR2183281

[13] W. Hachem, M. Kharouf, J. Najim and J. Silverstein. A CLT for information-theoretic statistics of non-centered Gram random matrices. *Random Matrices Theory Appl.* **1** (2012) 1150010. MR2934716

[14] W. Hachem, P. Loubaton and J. Najim. The empirical distribution of the eigenvalues of a Gram matrix with a given variance profile. *Ann. Inst. Henri Poincaré Probab. Stat.* **42** (2006) 649–670. MR2269232

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

[16] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge Univ. Press, Cambridge, 1994. MR1288752

[17] A. Kammoun, M. Kharouf, W. Hachem, J. Najim and A. El Kharroubi. On the fluctuations of the mutual information for non centered MIMO channels: The non Gaussian case. In *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2010.

[18] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)* **72** (1967) 507–536. MR0208649

[19] X. Mestre. Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates. *IEEE Trans. Inform. Theory* **54** (2008) 5113–5129. MR2589886

[20] X. Mestre. On the asymptotic behavior of the sample estimates of eigenvalues and eigenvectors of covariance matrices. *IEEE Trans. Signal Process.* **56** (2008) 5353–5368. MR2472837

[21] X. Mestre and M. A. Lagunas. Modified subspace algorithms for DOA estimation with large arrays. *IEEE Trans. Signal Process.* **56** (2008) 598–614. MR2445537

[22] W. Rudin. *Real and Complex Analysis*, 3rd edition. McGraw-Hill, New York, 1986.

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

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

[25] P. Vallet, P. Loubaton and X. Mestre. Improved subspace estimation for multivariate observations of high dimension: The deterministic signals case. *IEEE Trans. Inform. Theory* **58** (2012) 1043–1068.