

# THE EMPIRICAL DISTRIBUTION OF THE EIGENVALUES OF A GRAM MATRIX WITH A GIVEN VARIANCE PROFILE

W. HACHEM, P. LOUBATON AND J. NAJIM

ABSTRACT. Consider a  $N \times n$  random matrix  $Y_n = (Y_{ij}^n)$  where the entries are given by  $Y_{ij}^n = \frac{\sigma(i/N, j/n)}{\sqrt{n}} X_{ij}^n$ , the  $X_{ij}^n$  being centered i.i.d. and  $\sigma : [0, 1]^2 \rightarrow [0, \infty)$  being a continuous function called a variance profile. Consider now a deterministic  $N \times n$  matrix  $\Lambda_n = (\Lambda_{ij}^n)$  whose off-diagonal entries are zero. Denote by  $\Sigma_n$  the non-centered matrix  $Y_n + \Lambda_n$  and by  $N \wedge n = \min(N, n)$ . Then under the assumption that  $\lim_{n \rightarrow \infty} \frac{N}{n} = c > 0$  and

$$\frac{1}{N \wedge n} \sum_{i=1}^{N \wedge n} \delta_{\left(\frac{i}{N \wedge n}, (\Lambda_{ii}^n)^2\right)} \xrightarrow[n \rightarrow \infty]{} H(dx, d\lambda),$$

where  $H$  is a probability measure, it is proven that the empirical distribution of the eigenvalues of  $\Sigma_n \Sigma_n^T$  converges almost surely in distribution to a non-random probability measure. This measure is characterized in terms of its Stieltjes transform, which is obtained with the help of an auxiliary system of equations. This kind of results is of interest in the field of wireless communication.

RÉSUMÉ. Soit  $Y_n = (Y_{ij}^n)$  une matrice  $N \times n$  dont les entrées sont données par  $Y_{ij}^n = \frac{\sigma(i/N, j/n)}{\sqrt{n}} X_{ij}^n$ , les  $X_{ij}^n$  étant des variables aléatoires centrées, i.i.d. et où  $\sigma : [0, 1]^2 \rightarrow [0, \infty)$  est une fonction continue qu'on appellera profil de variance. Considérons une matrice déterministe  $\Lambda_n = (\Lambda_{ij}^n)$  de dimensions  $N \times n$  dont les éléments non diagonaux sont nuls. Appelons  $\Sigma_n$  la matrice non centrée définie par  $\Sigma_n = Y_n + \Lambda_n$  et notons  $N \wedge n = \min(N, n)$ . Sous les hypothèses que  $\lim_{n \rightarrow \infty} \frac{N}{n} = c > 0$  et que

$$\frac{1}{N \wedge n} \sum_{i=1}^{N \wedge n} \delta_{\left(\frac{i}{N \wedge n}, (\Lambda_{ii}^n)^2\right)} \xrightarrow[n \rightarrow \infty]{} H(dx, d\lambda),$$

où  $H$  est une probabilité, on démontre que la mesure empirique des valeurs propres de  $\Sigma_n \Sigma_n^T$  converge presque sûrement vers une mesure de probabilité déterministe. Cette mesure est caractérisée par sa transformée de Stieltjes, qui s'obtient à l'aide d'un système auxiliaire d'équations. Ce type de résultats présente un intérêt dans le domaine des communications numériques sans fil.

**Key words and phrases:** Random Matrix, empirical distribution of the eigenvalues, Stieltjes transform.

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

## 1. INTRODUCTION

Consider a  $N \times n$  random matrix  $Y_n = (Y_{ij}^n)$  where the entries are given by

$$Y_{ij}^n = \frac{\sigma(i/N, j/n)}{\sqrt{n}} X_{ij}^n \tag{1.1}$$

---

Date: September 12, 2005.

where  $\sigma : [0, 1] \times [0, 1] \rightarrow [0, \infty)$  is a continuous function called a variance profile and the random variables  $X_{ij}^n$  are real, centered, independent and identically distributed (i.i.d.) with finite  $4 + \epsilon$  moment. Consider a real deterministic  $N \times n$  matrix  $\Lambda_n = (\Lambda_{ij}^n)$  whose off-diagonal entries are zero and let  $\Sigma_n$  be the matrix  $\Sigma_n = Y_n + \Lambda_n$ . This model has two interesting features: The random variables are independent but not i.i.d. since the variance may vary and  $\Lambda_n$ , the centering perturbation of  $Y_n$ , though (pseudo) diagonal can be of full rank. The purpose of this article is to study the convergence of the empirical distribution of the eigenvalues of the Gram random matrix  $\Sigma_n \Sigma_n^T$  ( $\Sigma_n^T$  being the transpose of  $\Sigma_n$ ) when  $n \rightarrow +\infty$  and  $N \rightarrow +\infty$  in such a way that  $\frac{N}{n} \rightarrow c$ ,  $0 < c < +\infty$ .

The asymptotics of the spectrum of  $N \times N$  Gram random matrices  $Z_n Z_n^T$  have been widely studied in the case where  $Z_n$  is centered (see Marčenko and Pastur [15], Yin [23], Silverstein et al. [17, 18], Girko [7, 8], Khorunzhy et al. [13], Boutet de Monvel et al. [3], etc.). For an overview on asymptotic spectral properties of random matrices, see Bai [1]. The case of a Gram matrix  $Z_n Z_n^T$  where  $Z_n$  is non-centered has comparatively received less attention. Let us mention Girko ([8], chapter 7) where a general study is carried out for the matrix  $Z_n = (W_n + A_n)$  where  $W_n$  has a given variance profile and  $A_n$  is deterministic. In [8], it is proved that the entries of the resolvent  $(Z_n Z_n^T - zI)^{-1}$  have the same asymptotic behaviour as the entries of a certain deterministic holomorphic  $N \times N$  matrix valued function  $T_n(z)$ . This matrix-valued function is characterized by a nonlinear system of  $(n + N)$  coupled functional equations (see also [10]). Using different methods, Brent Dozier and Silverstein [6] study the eigenvalue asymptotics of the matrix  $(R_n + X_n)(R_n + X_n)^T$  in the case where the matrices  $X_n$  and  $R_n$  are independent random matrices,  $X_n$  has i.i.d. entries and the empirical distribution of  $R_n R_n^T$  converges to a non-random distribution. It is proved there that the eigenvalue distribution of  $(R_n + X_n)(R_n + X_n)^T$  converges almost surely towards a deterministic distribution whose Stieltjes transform is uniquely defined by a certain functional equation.

As in [6], the model studied in this article, i.e.  $\Sigma_n = Y_n + \Lambda_n$ , is a particular case of the general case studied in ([8], chapter 7, equation  $K_7$ ) for which there exists a limiting distribution for the empirical distribution of the eigenvalues. Since the centering term  $\Lambda_n$  is pseudo-diagonal, the proof of the convergence of the empirical distribution of the eigenvalues is based on a direct analysis of the diagonal entries of the resolvent  $(\Sigma_n \Sigma_n^T - zI)^{-1}$ . This analysis leads in a natural way to the equations characterizing the Stieltjes transform of the limiting probability distribution of the eigenvalues.

In the Wigner case with a variance profile, that is when matrix  $Y_n$  and the variance profile are symmetric (such matrices are also called band matrices), the limiting behaviour of the empirical distribution of the eigenvalues has been studied by Shlyakhtenko [16] in the gaussian case (see Section 3.4 for more details).

Recently, many of these results have been applied to the field of Signal Processing and Communication Systems and some new ones have been developed for that purpose (Silverstein and Combettes [19], Tse et al. [20, 21], Debbah et al. [5], Li et al. [14], etc.). The issue addressed in this paper is mainly motivated by the performance analysis of multiple-input multiple-output (MIMO) digital communication systems. In MIMO systems with  $n$  transmit antennas and  $N$  receive antennas, one can model the communication channel by a  $N \times n$  matrix  $H_n = (H_{ij}^n)$  where the entries  $H_{ij}^n$  represent the complex gain between transmit antenna  $i$  and receive antenna  $j$ . The statistics  $C_n = \frac{1}{n} \log \det(I_n + \frac{H_n H_n^*}{\sigma^2})$  (where  $H_n^*$  is the hermitian adjoint and  $\sigma^2$  represents the variance of an additive noise corrupting the received

signals) is a popular performance analysis index since it has been shown in information theory that  $C_n$  is the mutual information, that is the maximum number of bits per channel use and per antenna that can be transmitted reliably in a MIMO system with channel matrix  $H_n$ . Since

$$C_n = \frac{1}{n} \sum_{k=1}^N \log \left( 1 + \frac{\mu_k}{\sigma^2} \right),$$

where  $(\mu_k)_{1 \leq k \leq N}$  are the eigenvalues of  $H_n H_n^*$ , the empirical distribution of the eigenvalues of  $H_n H_n^*$  gives direct information on  $C_n$  (see Tulino and Verdu [22] for an exhaustive review of recent results). For wireless systems, matrix  $H_n$  is often modelled as a zero-mean Gaussian random matrix and several articles have recently been devoted to the study of the impact of the channel statistics (via the eigenvalues of  $H_n H_n^*$ ) on the probability distribution of  $C_n$  (Chuah et al. [4], Goldsmith et al. [9], see also [22] and the references therein). Of particular interest is also the channel matrix  $H_n = F_N(Y_n + \Lambda_n)F_n^T$  where  $F_k = (F_{pq}^k)_{1 \leq p, q \leq k}$  is the Fourier matrix (i.e.  $F_{pq}^k = k^{-\frac{1}{2}} \exp(2i\pi \frac{(p-1)(q-1)}{k})$ ) and the matrix  $Y_n$  is given by (1.1) (see [22], p. 139 for more details). The matrices  $H_n$  and  $\Sigma_n$  having the same singular values, we will focus on the study of the empirical distribution of the singular values of  $\Sigma_n$ . Moreover, we will focus on matrices with real entries since the complex case is a straightforward extension.

In the sequel, we will study simultaneously quantities (Stieltjes kernels) related to the Stieltjes transforms of  $\Sigma_n \Sigma_n^T$  and  $\Sigma_n^T \Sigma_n$ . Even if the Stieltjes transforms of  $\Sigma_n \Sigma_n^T$  and  $\Sigma_n^T \Sigma_n$  are related in an obvious way, the corresponding Stieltjes kernels are not, as we shall see. We will prove that if  $N/n \xrightarrow[n \rightarrow \infty]{} c > 0$  (since we study at the same time  $\Sigma_n \Sigma_n^T$  and  $\Sigma_n^T \Sigma_n$ , we assume without loss of generality that  $c \leq 1$ ) and if there exists a probability measure  $H$  on  $[0, 1] \times \mathbb{R}$  with compact support  $\mathcal{H}$  such that

$$\frac{1}{N} \sum_{i=1}^N \delta_{\left(\frac{i}{N}, (\Lambda_{ii}^n)^2\right)} \xrightarrow[n \rightarrow \infty]{\mathcal{D}} H(dx, d\lambda)$$

where  $\mathcal{D}$  stands for convergence in distribution, then almost surely, the empirical distribution of the eigenvalues of the random matrix  $\Sigma_n \Sigma_n^T$  (resp.  $\Sigma_n^T \Sigma_n$ ) converges in distribution to a deterministic probability distribution  $\mathbb{P}$  (resp.  $\tilde{\mathbb{P}}$ ). The probability distributions  $\mathbb{P}$  and  $\tilde{\mathbb{P}}$  are characterized in terms of their Stieltjes transform

$$f(z) = \int_{\mathbb{R}^+} \frac{\mathbb{P}(dx)}{x-z} \quad \text{and} \quad \tilde{f}(z) = \int_{\mathbb{R}^+} \frac{\tilde{\mathbb{P}}(dx)}{x-z}, \quad \text{Im}(z) \neq 0.$$

as follows. Consider the following system of equations

$$\begin{aligned} \int g d\pi_z &= \int \frac{g(u, \lambda)}{-z(1 + \int \sigma^2(u, \cdot) d\tilde{\pi}_z) + \frac{\lambda}{1+c \int \sigma^2(\cdot, cu) d\pi_z}} H(du, d\lambda) \\ \int g d\tilde{\pi}_z &= c \int \frac{g(cu, \lambda)}{-z(1 + c \int \sigma^2(\cdot, cu) d\pi_z) + \frac{\lambda}{1+\int \sigma^2(u, \cdot) d\tilde{\pi}_z}} H(du, d\lambda) \\ &\quad + (1-c) \int_c^1 \frac{g(u, 0)}{-z(1 + c \int \sigma^2(\cdot, u) d\pi_z)} du \end{aligned}$$

where the unknown parameters are the complex measures  $\pi_z$  and  $\tilde{\pi}_z$  and where  $g : \mathcal{H} \rightarrow \mathbb{R}$  is a continuous test function. Then, this system admits a unique pair of solutions  $(\pi_z(dx, d\lambda), \tilde{\pi}_z(dx, d\lambda))$ . In particular,  $\pi_z$  is absolutely continuous with respect to  $H$  while

$\tilde{\pi}_z$  is not (see Section 2 for more details). The Stieltjes transforms  $f$  and  $\tilde{f}$  are then given by

$$f(z) = \int_{[0,1] \times \mathbb{R}} \pi_z(dx, d\lambda) \quad \text{and} \quad \tilde{f}(z) = \int_{[0,1] \times \mathbb{R}} \tilde{\pi}_z(dx, d\lambda).$$

The article is organized as follows. In Section 2, the notations and the assumptions are introduced and the main result (Theorem 2.3) is stated. Section 3 is devoted to corollaries and remarks. Section 4 is devoted to the proof of Theorem 2.3.

**Acknowledgment.** We thank Alice Guionnet who first drew our attention to Shlyakhtenko's work. We also thank Florent Benaych-Georges for useful discussions and the referee for several remarks which improved the presentation of this article.

## 2. CONVERGENCE OF THE STIELTJES TRANSFORM

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

$$\lim_{n \rightarrow \infty} \frac{N(n)}{n} = c.$$

Consider a  $N \times n$  random matrix  $Y_n$  where the entries are given by

$$Y_{ij}^n = \frac{\sigma(i/N, j/n)}{\sqrt{n}} X_{ij}^n$$

where  $X_{ij}^n$  and  $\sigma$  are defined below.

**Assumption A-1.** *The random variables  $(X_{ij}^n ; 1 \leq i \leq N, 1 \leq j \leq n, n \geq 1)$  are real, independent and identically distributed. They are centered with  $\mathbb{E}(X_{ij}^n)^2 = 1$  and satisfy:*

$$\exists \epsilon > 0, \quad \mathbb{E}|X_{ij}^n|^{4+\epsilon} < \infty.$$

where  $\mathbb{E}$  denotes the expectation.

*Remark 2.1.* Using truncation arguments à la Bai and Silverstein [2, 17, 18], one may improve Assumption (A-1).

**Assumption A-2.** *The real function  $\sigma : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$  is such that  $\sigma^2$  is continuous. Therefore there exists a nonnegative constant  $\sigma_{\max}$  such that*

$$\forall (x, y) \in [0, 1]^2, \quad 0 \leq \sigma^2(x, y) \leq \sigma_{\max}^2 < \infty. \quad (2.1)$$

*Remark 2.2.* The function  $\sigma$  can vanish on portions of the domain  $[0, 1]^2$ .

Denote by  $\text{var}(Z)$  the variance of the random variable  $Z$ . Since  $\text{var}(Y_{ij}^n) = \sigma^2(i/N, j/n)/n$ , the function  $\sigma$  will be called a variance profile. Denote by  $\delta_{z_0}(dz)$  the dirac measure at point  $z_0$ . The indicator function of  $A$  is denoted by  $1_A(x)$ . Denote by  $C_b(\mathcal{X})$  (resp.  $C_b(\mathcal{X}; \mathbb{C})$ ) the set of real (resp. complex) continuous and bounded functions over the topological set  $\mathcal{X}$  and by  $\|f\|_\infty = \sup_{x \in \mathcal{X}} |f(x)|$ , the supremum norm. If  $\mathcal{X}$  is compact, we simply write  $C(\mathcal{X})$  (resp.  $C(\mathcal{X}; \mathbb{C})$ ) instead of  $C_b(\mathcal{X})$  (resp.  $C_b(\mathcal{X}; \mathbb{C})$ ). We denote by  $\xrightarrow{\mathcal{D}}$  the convergence in distribution for probability measures and by  $\xrightarrow{w}$  the weak convergence for bounded complex measures.

Consider a real deterministic  $N \times n$  matrix  $\Lambda_n = (\Lambda_{ij}^n)$  whose off-diagonal entries are zero. We often write  $\Lambda_{ij}$  instead of  $\Lambda_{ij}^n$ . We introduce the  $N \times n$  matrix  $\Sigma_n = Y_n + \Lambda_n$ .

For every matrix  $A$ , we denote by  $A^T$  its transpose, by  $\text{Tr}(A)$  its trace (if  $A$  is square) and by  $F^A A^T$ , the empirical distribution function of the eigenvalues of  $A A^T$ . Denote by  $\text{diag}(a_i; 1 \leq i \leq k)$  the  $k \times k$  diagonal matrix whose diagonal entries are the  $a_i$ 's. Since we will study at the same time the limiting spectrum of the matrices  $\Sigma_n \Sigma_n^T$  and  $\Sigma_n^T \Sigma_n$ , we can assume without loss of generality that  $c \leq 1$ . We also assume for simplicity that  $N \leq n$ .

We assume that:

**Assumption A-3.** *There exists a probability measure  $H(du, d\lambda)$  over the set  $[0, 1] \times \mathbb{R}$  with compact support  $\mathcal{H}$  such that*

$$H_n(du, d\lambda) \equiv \frac{1}{N} \sum_{i=1}^N \delta_{\left(\frac{i}{N}, (\Lambda_{ii}^n)^2\right)}(du, d\lambda) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} H(du, d\lambda). \quad (2.2)$$

*Remark 2.3* (The probability measure  $H$ ). Assumption (A-3) accounts for the presence of probability measure  $H(du, d\lambda)$  in forthcoming equations (2.6) and (2.7). If  $\Lambda_{ii}^2 = f\left(\frac{i}{n}\right)$ , then (A-3) is automatically fulfilled with  $H(du, d\lambda) = \delta_{f(x)}(du)dx$ . This is in particular the case in Boutet de Monvel et al. [3], Schlyakhtenko [16], Hachem et al. ([11], Theorem 4.2).

*Remark 2.4* (The complex case). Assumptions (A-1), (A-2) and (A-3) must be slightly modified in the complex setting. Related modifications are stated in Section 3.3.

When dealing with vectors, the norm  $\|\cdot\|$  will denote the Euclidean norm. In the case of matrices, the norm  $\|\cdot\|$  will refer to the spectral norm.

*Remark 2.5* (Boundedness of the  $\Lambda_{ii}$ 's). Due to (A-3), we can assume without loss of generality that the  $\Lambda_{ii}^n$ 's are bounded for  $n$  large enough. In fact, suppose not, then by (A-3),  $\frac{1}{N} \sum_{i=1}^N \delta_{\Lambda_{ii}^2} \rightarrow H_\Lambda(d\lambda)$  whose support is compact and, say, included in  $[0, K]$ . Then Portmanteau's theorem yields  $\frac{1}{N} \sum_{i=1}^N 1_{[0, K+\delta]}(\Lambda_{ii}^2) \rightarrow 1$  thus

$$\frac{\#\{\Lambda_{ii}, \Lambda_{ii}^2 \notin [0, K+\delta]\}}{N} = 1 - \frac{1}{N} \sum_{i=1}^N 1_{[0, K+\delta]}(\Lambda_{ii}^2) \xrightarrow[n \rightarrow \infty]{} 0. \quad (2.3)$$

Denote by  $\check{\Lambda}_n = (\check{\Lambda}_{ij}^n)$  the matrix whose off-diagonal entries are zero and set  $\check{\Lambda}_{ii}^n = \Lambda_{ii}^n 1_{\{(\Lambda_{ii}^n)^2 \leq K+\delta\}}$ . Then it is straightforward to check that  $\frac{1}{N} \sum_{i=1}^N \delta_{\left(\frac{i}{N}, \check{\Lambda}_{ii}^2\right)}(du, d\lambda) \rightarrow H(du, d\lambda)$ . Moreover, if  $\check{\Sigma}_n = Y_n + \check{\Lambda}_n$  then

$$\|F^{\Sigma \Sigma^T} - F^{\check{\Sigma} \check{\Sigma}^T}\|_\infty \stackrel{(a)}{\leq} \frac{\text{rank}(\check{\Sigma} - \Sigma)}{N} \stackrel{(b)}{\leq} \frac{\#\{\Lambda_{ii}, \Lambda_{ii}^2 \notin [0, K+\delta]\}}{N} \stackrel{(c)}{\xrightarrow[n \rightarrow \infty]{} 0}.$$

where (a) follows from Lemma 3.5 in [23] (see also [18], Section 2), (b) follows from the fact that for a rectangular matrix  $A$ ,  $\text{rank}(A) \leq$  the number of non-zero entries of  $A$  and (c) follows from (2.3). Therefore,  $F^{\Sigma \Sigma^T}$  converges iff  $F^{\check{\Sigma} \check{\Sigma}^T}$  converges. In this case they share the same limit. Remark 2.5 is proved.

*Remark 2.6.* [Compacity of the support of  $H_n$ ] Due to Remark 2.5, we will assume in the sequel that for all  $n$ , the support of  $\frac{1}{N} \sum \delta_{\left(\frac{i}{N}, \Lambda_{ii}^2\right)}$  is included in a compact set  $\mathcal{K} \subset [0, 1] \times \mathbb{R}$ .

Let  $\mathbb{C}^+ = \{z \in \mathbb{C}, \text{Im}(z) > 0\}$  and  $\mathbb{C}^\nabla = \{z \in \mathbb{C}^+, |\text{Re}(z)| \leq \text{Im}(z)\}$ .

**2.2. Stieltjes transforms and Stieltjes kernels.** Let  $\nu$  be a bounded nonnegative measure over  $\mathbb{R}$ . Its Stieltjes transform  $f$  is defined by:

$$f(z) = \int_{\mathbb{R}} \frac{\nu(d\lambda)}{\lambda - z}, \quad z \in \mathbb{C}^+.$$

We list below the main properties of the Stieltjes transforms that will be needed in the sequel.

**Proposition 2.1.** *The following properties hold true:*

- (1) *Let  $f$  be the Stieltjes transform of  $\nu$ , then*
  - *the function  $f$  is analytic over  $\mathbb{C}^+$ ,*
  - *the function  $f$  satisfies:  $|f(z)| \leq \frac{\nu(\mathbb{R})}{\text{Im}(z)}$ ,*
  - *if  $z \in \mathbb{C}^+$  then  $f(z) \in \mathbb{C}^+$ ,*
  - *if  $\nu(-\infty, 0) = 0$  then  $z \in \mathbb{C}^+$  implies  $z f(z) \in \mathbb{C}^+$ .*
- (2) *Conversely, let  $f$  be a function analytic over  $\mathbb{C}^+$  such that  $f(z) \in \mathbb{C}^+$  if  $z \in \mathbb{C}^+$  and  $|f(z)||\text{Im}(z)|$  bounded on  $\mathbb{C}^+$ . Then,  $f$  is the Stieltjes transform of a bounded positive measure  $\mu$  and  $\mu(\mathbb{R})$  is given by*

$$\mu(\mathbb{R}) = \lim_{y \rightarrow +\infty} -iy f(iy).$$

*If moreover  $z f(z) \in \mathbb{C}^+$  for  $z \in \mathbb{C}^+$  then,  $\mu(\mathbb{R}^-) = 0$ .*

- (3) *Let  $\mathbb{P}_n$  and  $\mathbb{P}$  be probability measures over  $\mathbb{R}$  and denote by  $f_n$  and  $f$  their Stieltjes transforms. Then*

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

Let  $A$  be an  $n \times p$  matrix and let  $I_n$  be the  $n \times n$  identity. The resolvent of  $AA^T$  is defined by

$$Q(z) = (AA^T - zI_n)^{-1} = (q_{ij}(z))_{1 \leq i, j \leq n}, \quad z \in \mathbb{C} \setminus \mathbb{R}.$$

The following properties are straightforward.

**Proposition 2.2.** *Let  $Q(z)$  be the resolvent of  $AA^T$ , then:*

- (1) *For all  $z \in \mathbb{C}^+$ ,  $\|Q(z)\| \leq (\text{Im}(z))^{-1}$ . Similarly,  $|q_{ij}(z)| \leq (\text{Im}(z))^{-1}$ .*
- (2) *The function  $h_n(z) = \frac{1}{n} \text{Tr} Q(z)$  is the Stieltjes transform of the empirical distribution probability associated to the eigenvalues of  $AA^T$ . Since these eigenvalues are nonnegative,  $z h_n(z) \in \mathbb{C}^+$  for  $z \in \mathbb{C}^+$ .*
- (3) *Let  $\vec{\xi}$  be a  $n \times 1$  vector, then  $\text{Im} \left( z \vec{\xi} Q(z) \vec{\xi}^T \right) \in \mathbb{C}^+$  for  $z \in \mathbb{C}^+$ .*

Denote by  $\mathcal{M}_{\mathbb{C}}(\mathcal{X})$  the set of complex measures over the topological set  $\mathcal{X}$ . In the sequel, we will call Stieltjes kernel every application

$$\mu : \mathbb{C}^+ \rightarrow \mathcal{M}_{\mathbb{C}}(\mathcal{X})$$

either denoted  $\mu(z, dx)$  or  $\mu_z(dx)$  and satisfying:

- (1)  $\forall g \in C_b(\mathcal{X}), \int g d\mu_z$  is analytic over  $\mathbb{C}^+$ ,

(2)  $\forall z \in \mathbb{C}^+, \forall g \in C_b(\mathcal{X})$ ,

$$\left| \int g d\mu_z \right| \leq \frac{\|g\|_\infty}{\text{Im}(z)}$$

(3)  $\forall z \in \mathbb{C}^+, \forall g \in C_b(\mathcal{X})$  and  $g \geq 0$  then  $\text{Im} \left( \int g d\mu_z \right) \geq 0$ ,

(4)  $\forall z \in \mathbb{C}^+, \forall g \in C_b(\mathcal{X})$  and  $g \geq 0$  then  $\text{Im} \left( z \int g d\mu_z \right) \geq 0$ .

Let us introduce the following resolvents:

$$\begin{aligned} Q_n(z) &= (\Sigma_n \Sigma_n^T - zI_N)^{-1} = (q_{ij}(z))_{1 \leq i, j \leq N}, \quad z \in \mathbb{C}^+, \\ \tilde{Q}_n(z) &= (\Sigma_n^T \Sigma_n - zI_n)^{-1} = (\tilde{q}_{ij}(z))_{1 \leq i, j \leq n}, \quad z \in \mathbb{C}^+, \end{aligned}$$

and the following empirical measures defined for  $z \in \mathbb{C}^+$

$$L_z^n(du, d\lambda) = \frac{1}{N} \sum_{i=1}^N q_{ii}(z) \delta_{(\frac{i}{N}, \Lambda_{ii}^2)}(du, d\lambda), \quad (2.4)$$

$$\begin{aligned} \tilde{L}_z^n(du, d\lambda) &= \frac{1}{n} \sum_{i=1}^N \tilde{q}_{ii}(z) \delta_{(\frac{i}{n}, \Lambda_{ii}^2)}(du, d\lambda) \\ &\quad + \left( \frac{1}{n} \sum_{i=N+1}^n \tilde{q}_{ii}(z) \delta_{\frac{i}{n}}(du) \otimes \delta_0(d\lambda) \right) 1_{\{N < n\}}, \end{aligned} \quad (2.5)$$

where  $\otimes$  denotes the product of measures. Since  $q_{ii}(z)$  (resp.  $\tilde{q}_{ii}(z)$ ) is analytic over  $\mathbb{C}^+$ , satisfies  $|q_{ii}(z)| \leq (\text{Im}(z))^{-1}$  and  $\min(\text{Im}(q_{ii}(z)), \text{Im}(zq_{ii}(z))) > 0$ ,  $L^n$  (resp.  $\tilde{L}^n$ ) is a Stieltjes kernel. Recall that due to Remark 2.6,  $L^n$  and  $\tilde{L}^n$  have supports included in the compact set  $\mathcal{K}$ .

*Remark 2.7* (on the limiting support of  $L^n$ ). Consider a converging subsequence of  $L_z^n$ , then its limiting support is necessarily included in  $\mathcal{H}$ .

*Remark 2.8* (on the limiting support of  $\tilde{L}^n$ ). Denote by  $H_c$  the image of the probability measure  $H$  under the application  $(u, \lambda) \mapsto (cu, \lambda)$ , by  $\mathcal{H}_c$  its support, by  $\mathcal{R}$  the support of the measure  $1_{[c, 1]}(u) du \otimes \delta_0(d\lambda)$ . Let  $\tilde{\mathcal{H}} = \mathcal{H}_c \cup \mathcal{R}$ . Notice that  $\tilde{\mathcal{H}}$  is obviously compact. Consider a converging subsequence of  $\tilde{L}_z^n$ , then its limiting support is necessarily included in  $\tilde{\mathcal{H}}$ .

### 2.3. Convergence of the empirical measures $L_z^n$ and $\tilde{L}_z^n$ .

**Theorem 2.3.** *Assume that Assumptions (A-1), (A-2) and (A-3) hold and consider the following system of equations*

$$\int g d\pi_z = \int \frac{g(u, \lambda)}{-z(1 + \int \sigma^2(u, t) \tilde{\pi}(z, dt, d\zeta)) + \frac{\lambda}{1+c \int \sigma^2(t, cu) \pi(z, dt, d\zeta)}} H(du, d\lambda) \quad (2.6)$$

$$\begin{aligned} \int g d\tilde{\pi}_z &= c \int \frac{g(cu, \lambda)}{-z(1 + c \int \sigma^2(t, cu) \pi(z, dt, d\zeta)) + \frac{\lambda}{1+\int \sigma^2(u, t) \tilde{\pi}(z, dt, d\zeta)}} H(du, d\lambda) \\ &\quad + (1-c) \int_c^1 \frac{g(u, 0)}{-z(1 + c \int \sigma^2(t, u) \pi(z, dt, d\zeta))} du \end{aligned} \quad (2.7)$$

where (2.6) and (2.7) hold for every  $g \in C(\mathcal{H})$ . Then,

- (1) *this system admits a unique couple of solutions  $(\pi(z, dt, d\lambda), \tilde{\pi}(z, dt, d\lambda))$  among the set of Stieltjes kernels for which the support of measure  $\pi_z$  is included in  $\mathcal{H}$  and the support of measure  $\tilde{\pi}_z$  is included in  $\tilde{\mathcal{H}}$ ,*
- (2) *the functions  $f(z) = \int d\pi_z$  and  $\tilde{f}(z) = \int d\tilde{\pi}_z$  are the Stieltjes transforms of probability measures,*
- (3) *the following convergences hold true:*

$$\begin{aligned} \text{a.s. } \forall z \in \mathbb{C}^+, \quad L_z^n &\xrightarrow[n \rightarrow \infty]{w} \pi_z, \\ \text{a.s. } \forall z \in \mathbb{C}^+, \quad \tilde{L}_z^n &\xrightarrow[n \rightarrow \infty]{w} \tilde{\pi}_z, \end{aligned}$$

where  $L^n$  and  $\tilde{L}^n$  are defined by (2.4) and (2.5).

*Remark 2.9* (on the absolute continuity of  $\pi_z$  and  $\tilde{\pi}_z$ ). Due to (2.6), the complex measure  $\pi_z$  is absolutely continuous with respect to  $H$ . However, it is clear from (2.7) that  $\tilde{\pi}_z$  has an absolutely continuous part with respect to  $H_c$  (recall that  $H_c$  is the image of  $H$  under  $(u, \lambda) \mapsto (cu, \lambda)$ ) and an absolutely continuous part with respect to  $1_{[c,1]}(u) du \otimes \delta_0(d\lambda)$  (which is in general singular with respect to  $H_c$ ). Therefore, it is much more convenient to work with Stieltjes kernels  $\pi$  and  $\tilde{\pi}$  rather than with measure densities indexed by  $z$ .

Proof of Theorem 2.3 is postponed to Section 4.

**Corollary 2.4.** *Assume that (A-1), (A-2) and (A-3) hold and denote by  $\pi$  and  $\tilde{\pi}$  the two Stieltjes kernels solutions of the coupled equations (2.6) and (2.7). Then the empirical distribution of the eigenvalues of the matrix  $\Sigma_n \Sigma_n^T$  converges almost surely to a non-random probability measure  $\mathbb{P}$  whose Stieltjes transform  $f(z) = \int_{\mathbb{R}^+} \frac{\mathbb{P}(dx)}{x-z}$  is given by:*

$$f(z) = \int_{\mathcal{H}} \pi_z(dx, d\lambda).$$

*Similarly, the empirical distribution of the eigenvalues of the matrix  $\Sigma_n^T \Sigma_n$  converges almost surely to a non-random probability measure  $\tilde{\mathbb{P}}$  whose Stieltjes transform  $\tilde{f}(z)$  is given by:*

$$\tilde{f}(z) = \int_{\tilde{\mathcal{H}}} \tilde{\pi}_z(dx, d\lambda).$$

*Proof of Corollary 2.4.* The Stieltjes transform of  $\Sigma_n \Sigma_n^T$  is equal to  $\frac{1}{N} \sum_{i=1}^N q_{ii}(z) = \int dL_z^n$ . By Theorem 2.3-(3),

$$\text{a.s. } \forall z \in \mathbb{C}^+, \quad \int dL_z^n \xrightarrow[n \rightarrow \infty]{} \int d\pi_z. \quad (2.8)$$

Since  $\int d\pi_z$  is the Stieltjes transform of a probability measure  $\mathbb{P}$  by Theorem 2.3-(2), eq. (2.8) implies that  $F^{\Sigma_n \Sigma_n^T}$  converges weakly to  $\mathbb{P}$ . One can similarly prove that  $F^{\Sigma_n^T \Sigma_n}$  converges weakly to a probability measure  $\tilde{\mathbb{P}}$ .  $\square$

### 3. FURTHER RESULTS AND REMARKS

In this section, we present two corollaries of Theorem 2.3. We will discuss the case where  $\Lambda_n = 0$  and the case where the variance profile  $\sigma(x, y)$  is constant. These results are already well-known ([3, 6, 7, 8]). We also show how Assumptions (A-1), (A-2) and (A-3) must be modified in the complex case. Finally, we give some comments on Shlyakhtenko's result [16].



### 3.1. The Centered case.

**Corollary 3.1.** *Assume that (A-1) and (A-2) hold. Then the empirical distribution of the eigenvalues of the matrix  $Y_n Y_n^T$  converges a.s. to a non-random probability measure  $\mathbb{P}$  whose Stieltjes transform  $f$  is given by*

$$f(z) = \int_{[0,1]} \pi_z(dx),$$

where  $\pi_z$  is the unique Stieltjes kernel with support included in  $[0, 1]$  and satisfying

$$\forall g \in C([0, 1]), \quad \int g d\pi_z = \int_0^1 \frac{g(u)}{-z + \int_0^1 \frac{\sigma^2(u,t)}{1+c \int_0^1 \sigma^2(x,t) \pi_z(dx)} dt} du. \quad (3.1)$$

*Remark 3.1* (Absolute continuity of the Stieltjes kernel). In this case, one can prove that  $\pi_z$  is absolutely continuous with respect to  $du$ , i.e.  $\pi_z(du) = k(z, u)du$  where  $z \mapsto k(z, u)$  is analytic and  $u \mapsto k(z, u)$  is continuous. Eq. (3.1) becomes

$$\forall u \in [0, 1], \quad \forall z \in \mathbb{C}^+, \quad k(u, z) = \frac{1}{-z + \int_0^1 \frac{\sigma^2(u,t)}{1+c \int_0^1 \sigma^2(x,t) k(x,z) dx} dt}. \quad (3.2)$$

Eq. (3.2) appears (up to notational differences) in [7] and in [3] in the setting of Gram matrices based on Gaussian fields. Links between Gram matrices with a variance profile and Gram matrices based on Gaussian fields are studied in [11].

*Proof.* Assumption (A-3) is satisfied with  $\Lambda_{ii}^n = 0$  and  $H(du, d\lambda) = du \otimes \delta_0(\lambda)$  where  $du$  denotes Lebesgue measure on  $[0, 1]$ . Therefore Theorem 2.3 yields the existence of kernels  $\pi_z$  and  $\tilde{\pi}_z$  satisfying (2.6) and (2.7). It is straightforward to check that in this case  $\pi_z$  and  $\tilde{\pi}_z$  do not depend on variable  $\lambda$ . Therefore (2.6) and (2.7) become:

$$\int g d\pi_z = \int \frac{g(u)}{-z(1 + \int \sigma^2(u,t) \tilde{\pi}(z, dt))} du \quad (3.3)$$

and

$$\begin{aligned} \int g d\tilde{\pi}_z &= c \int_{[0,1]} \frac{g(cu)}{-z(1 + c \int \sigma^2(t, cu) \pi(z, dt))} du \\ &\quad + (1-c) \int_{[c,1]} \frac{g(u)}{-z(1 + c \int \sigma^2(t, u) \pi(z, dt))} du \\ &= \int_{[0,1]} \frac{g(u)}{-z(1 + c \int \sigma^2(t, u) \pi(z, dt))} du, \end{aligned} \quad (3.4)$$

where  $g \in C([0, 1])$ . Replacing  $\int \sigma^2(u, t) \tilde{\pi}(z, dt)$  in (3.3) by the expression given by (3.4), one gets the following equation satisfied by  $\pi_z(du)$ :

$$\int g d\pi_z = \int \frac{g(u)}{-z + \int \frac{\sigma^2(u,t)}{1+c \int \sigma^2(s,t) \pi(z, ds)} dt} du$$

□

### 3.2. The non-centered case with i.i.d. entries.

**Corollary 3.2.** *Assume that (A-1) and (A-2) hold where  $\sigma(x, y) = \sigma$  is a constant function. Assume moreover that  $\frac{1}{N} \sum_{i=1}^N \delta_{\Lambda_{ii}^2} \rightarrow H_\Lambda(d\lambda)$  weakly, where  $H_\Lambda$  has a compact support. Then the empirical distribution of the eigenvalues of the matrix  $\Sigma_n \Sigma_n^T$  converges a.s. to a non-random probability measure  $\mathbb{P}$  whose Stieltjes transform is given by*

$$f(z) = \int \frac{H_\Lambda(d\lambda)}{-z(1 + c\sigma^2 f(z)) + (1-c)\sigma^2 + \frac{\lambda}{1+c\sigma^2 f(z)}}. \quad (3.5)$$

*Remark 3.2.* Eq. (3.5) appears in [6] in the case where  $\Sigma_n = \sigma Z_n + R_n$  where  $Z_n$  and  $R_n$  are assumed to be independent,  $Z_{ij}^n = \frac{X_{ij}}{\sqrt{n}}$ , the  $X_{ij}$  being i.i.d. and the empirical distribution of the eigenvalues of  $R_n R_n^T$  converging to a given probability distribution. Since  $R_n$  is not assumed to be diagonal in [6], the results in [6] do not follow from Corollary 3.2.

*Proof.* We first sort the  $\Lambda_{ii}^2$ 's. Denote by  $O_n$  the  $N \times N$  permutation matrix such that  $O_n \text{diag}(\Lambda_{ii}; 1 \leq i \leq N) O_n^T = \text{diag}(\tilde{\Lambda}_{ii}; 1 \leq i \leq N)$  where  $\tilde{\Lambda}_{11}^2 \leq \dots \leq \tilde{\Lambda}_{NN}^2$ . Then  $O_n \Sigma_n \Sigma_n^T O_n^T$  and  $\Sigma_n \Sigma_n^T$  have the same eigenvalues. Denote by  $\tilde{\Lambda}_n$  the pseudo-diagonal  $N \times n$  matrix whose diagonal entries are the  $\tilde{\Lambda}_{ii}$  and by  $\check{O}_n$  the block matrix:

$$\check{O}_n = \begin{bmatrix} O_n & 0 \\ 0 & I_{N-n} \end{bmatrix}.$$

Then

$$O_n \Sigma_n \check{O}_n^T = O_n (Y_n + \Lambda_n) \check{O}_n^T = O_n Y_n \check{O}_n^T + \tilde{\Lambda}_n.$$

In particular  $O_n Y_n \check{O}_n^T$  remains a matrix with i.i.d. entries. Therefore, we can assume without loss of generality that:

$$\Lambda_{11}^2 \leq \dots \leq \Lambda_{NN}^2.$$

In this case, one can prove that the empirical distribution  $\frac{1}{N} \sum_{i=1}^N \delta_{i/N, \Lambda_{ii}^2}$  converges. In fact, denote by  $F_\Lambda(y) = H_\Lambda([0, y])$  and consider the function

$$F(x, y) = x \wedge F_\Lambda(y), \quad (x, y) \in [0, 1] \times \mathbb{R}^+.$$

where  $\wedge$  denotes the infimum. Assume that  $F_\Lambda$  is continuous at  $y$ , then:

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \delta_{i/N, \Lambda_{ii}^2}([0, x] \times [0, y]) \\ &= \frac{\#\{i, i/N \leq x \text{ and } \Lambda_{ii}^2 \leq y\}}{N} \\ &= \frac{\text{card}(A_n \cap B_n)}{N} \quad \text{where } A_n = \{i, i/N \leq x\} \text{ and } B_n = \{i, \Lambda_{ii}^2 \leq y\} \\ &\stackrel{(a)}{=} \frac{\text{card}(A_n)}{N} \wedge \frac{\text{card}(B_n)}{N} \xrightarrow{n \rightarrow \infty} x \wedge F_\Lambda(y) \end{aligned}$$

where (a) follows from the fact that  $A_n \cap B_n$  is either equal to  $A_n$  if  $A_n \subset B_n$  or to  $B_n$  if  $B_n \subset A_n$  due to the fact that the  $\Lambda_{ii}^2$  are sorted. The probability measure  $H$  associated to the cumulative distribution function  $F$  readily satisfies (A-3). Theorem 2.3 yields the existence of kernels  $\pi_z$  and  $\tilde{\pi}_z$  satisfying (2.6) and (2.7). It is straightforward to check that in this case  $\pi_z$  and  $\tilde{\pi}_z$  do not depend on variable  $u$ . Equation (2.6) becomes

$$\int g d\pi_z = \int \frac{g(u, \lambda)}{-z(1 + \sigma^2 \int \tilde{\pi}(z, dt, d\zeta)) + \frac{\lambda}{1+c\sigma^2 \int \pi(z, dt, d\zeta)}} H(du, d\lambda)$$

Let  $g(u, \lambda) = 1$  and denote by  $f = \int d\pi$ , then (2.6) becomes

$$f(z) = \int \frac{1}{-z(1 + \sigma^2 \tilde{f}(z)) + \frac{\lambda}{1 + c\sigma^2 f(z)}} H_\Lambda(d\lambda) \quad (3.6)$$

Denote by  $\tilde{f}_n(z) = \frac{1}{n} \sum_{i=1}^N \tilde{q}_{ii}(z) = \frac{1}{n} \text{Tr}(\Sigma_n^T \Sigma_n - zI)^{-1}$ . Since  $f_n(z) = \frac{1}{N} \text{Tr}(\Sigma_n \Sigma_n^T - zI)^{-1}$  and  $\tilde{f}_n(z) = \frac{1}{n} \text{Tr}(\Sigma_n^T \Sigma_n - zI)^{-1}$  (recall that  $N \leq n$ ), we have  $\tilde{f}_n(z) = \frac{N}{n} f_n(z) + (1 - \frac{N}{n}) (-\frac{1}{z})$ . This yields  $\tilde{f}(z) = cf(z) - \frac{1-c}{z}$ . Replacing  $\tilde{f}(z)$  in (3.6) by this expression, we get (3.5).  $\square$

**3.3. Statement of the results in the complex case.** In the complex setting, Assumptions (A-1)-(A-3) must be modified in the following way:

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

$$\exists \epsilon > 0, \quad \mathbb{E}|X_{ij}^n|^{4+\epsilon} < \infty.$$

**Assumption A-2.** *The complex function  $\sigma : [0, 1] \times [0, 1] \rightarrow \mathbb{C}$  is such that  $|\sigma|^2$  is continuous and therefore there exists a nonnegative constant  $\sigma_{\max}$  such that*

$$\forall (x, y) \in [0, 1]^2, \quad 0 \leq |\sigma(x, y)|^2 \leq \sigma_{\max}^2 < \infty. \quad (3.7)$$

If  $\Lambda_n$  is a complex deterministic  $N \times n$  matrix whose off-diagonal entries are zero, assume that:

**Assumption A-3.** *There exists a probability measure  $H(du, d\lambda)$  over the set  $[0, 1] \times \mathbb{R}$  with compact support  $\mathcal{H}$  such that*

$$\frac{1}{N} \sum_{i=1}^N \delta_{\left(\frac{i}{N}, |\Lambda_{ii}^n|^2\right)}(du, d\lambda) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} H(du, d\lambda). \quad (3.8)$$

In Eq. (2.6) and (2.7), one must replace  $\sigma$  by its module  $|\sigma|$ . The statements of Theorem 2.3 and Corollary 2.4 are not modified.

**3.4. Further remarks.** The problem of studying the convergence of the empirical distribution of  $\Sigma_n \Sigma_n^T$  could have been addressed differently. In particular, one could have more relied on Shlyakhtenko's ideas [16]. We give some details in this section. We also take this opportunity to thank the referee whose remarks led to this section.

*An extension of Shlyakhtenko's results.* In [16], Shlyakhtenko describes the spectral distribution of the  $n \times n$  matrix  $M_n = \Lambda_n + A_n$  where  $\Lambda_n$  is a diagonal matrix whose entries are approximately samples of a bounded function  $f$  on a regular grid, i.e.  $\Lambda_n$  is close to  $\text{diag}(f(\frac{i}{n}); 1 \leq i \leq n)$  and  $A_n$  is a Wigner matrix with a variance profile:  $A_{ij}^n = \frac{\sigma(i/n, j/n)}{\sqrt{n}} X_{ij}$ , the  $X_{ij}$ 's being i.i.d. (apart from the symmetry constraint) standard gaussian random variables and  $\sigma$  being symmetric, i.e.  $\sigma(x, y) = \sigma(y, x)$ . If instead of  $M_n$ , one can describe the limiting spectral distribution of  $\tilde{M}_n = \begin{pmatrix} 0 & \Sigma_n \\ \Sigma_n^T & 0 \end{pmatrix} = \begin{pmatrix} 0 & Y_n + \Lambda_n \\ Y_n^T + \Lambda_n^T & 0 \end{pmatrix}$ , then one can also describe the limiting spectral distribution of  $\Sigma_n \Sigma_n^T$  since  $\tilde{M}_n^2 = \begin{pmatrix} \Sigma_n \Sigma_n^T & 0 \\ 0 & \Sigma_n^T \Sigma_n \end{pmatrix}$  as noticed by Shlyakhtenko. It is not clear however how to adapt the concept of freeness with amalgamation to the study of  $\tilde{M}_n$  since the deterministic elements of  $\tilde{M}_n$  are not diagonal entries any more.

Apart from this, two issues remain with this approach: The extension of the result to the non-gaussian case and to the case where the diagonal deterministic entries are not samples of a bounded function (which is a case covered by Assumption (A-3)).

*The direct study of a Wigner matrix with a variance profile.* Another approach is based on Shlyakhtenko's remark together with the technique developed in this paper: Since it is sufficient to study the spectral measure of  $\tilde{M}_n$  in order to get the result for the spectral measure of  $\tilde{M}_n^2$ , one can directly study the diagonal elements of the resolvent  $(\tilde{M}_n - zI_n)^{-1}$ . This approach yields to results expressed in terms of the limiting Stieltjes transform of  $\frac{1}{n}\text{Tr}(\tilde{M}_n - zI_n)^{-1}$ .

#### 4. PROOF OF THEOREM 2.3

We first give an outline of the proof.

**4.1. Outline of the proof, more notations.** The proof is carried out following four steps:

- (1) Recall the definitions of the supports  $\mathcal{H}$  and  $\tilde{\mathcal{H}}$  (cf. Assumption (3) and Remark 2.8). We first prove that the system of equations (2.6) and (2.7) admits at most a unique couple of solutions  $(\pi(z, dt, d\lambda), \tilde{\pi}(z, dt, d\lambda))$  among the set of Stieltjes kernels for which the support of measure  $\pi_z$  is included in  $\mathcal{H}$  and the support of measure  $\tilde{\pi}_z$  is included in  $\tilde{\mathcal{H}}$  (Section 4.2). We also prove that if such solutions exist, then necessarily,  $f(z) = \int d\pi_z$  and  $\tilde{f}(z) = \int d\tilde{\pi}_z$  are Stieltjes transforms of probability measures.

- (2) We then prove that for each subsequence  $M(n)$  of  $n$  there exists a subsequence  $M_{\text{sub}} = M_{\text{sub}}(n)$  such that for all  $z \in \mathbb{C}^+$ ,

$$L_z^{M_{\text{sub}}} \xrightarrow[n \rightarrow \infty]{w} \mu_z \quad \text{and} \quad \tilde{L}_z^{M_{\text{sub}}} \xrightarrow[n \rightarrow \infty]{w} \tilde{\mu}_z, \quad (4.1)$$

where  $\mu_z$  and  $\tilde{\mu}_z$  are complex measures, a priori depending on  $\omega \in \Omega$ , (if  $\Omega$  denotes the underlying probability space), with support respectively included in  $\mathcal{H}$  and  $\tilde{\mathcal{H}}$  (Section 4.3). Note that the convergence of the subsequences stands everywhere (for every  $\omega \in \Omega$ ).

- (3) We then prove that  $z \mapsto \mu_z$  and  $z \mapsto \tilde{\mu}_z$  are Stieltjes kernels (Section 4.4). As in the previous step, this result holds everywhere.

- (4) We finally prove that for a countable collection of  $z \in \mathbb{C}^+$  with a limit point, say  $\mathcal{C}$ , the measures  $\mu_z$  and  $\tilde{\mu}_z$  satisfy equations (2.6) and (2.7) almost surely for all  $z \in \mathcal{C}$ . This is based on a precise description of the diagonal entry  $q_{ii}(z)$  of the resolvent  $(\Sigma_n \Sigma_n^T - zI)^{-1}$  (Lemma 4.1). Analyticity arguments yield then that almost surely,  $\forall z \in \mathbb{C}^+$ ,  $\mu_z$  and  $\tilde{\mu}_z$  satisfy equations (2.6) and (2.7). Otherwise stated:

- Almost surely, the system (2.6)-(2.7) admits a solution  $(\mu_z, \tilde{\mu}_z)$ ,
- This solution being unique, there exists a couple of deterministic Stieltjes kernels  $(\pi_z, \tilde{\pi}_z)$  such that almost surely  $(\mu_z, \tilde{\mu}_z) = (\pi_z, \tilde{\pi}_z)$ .
- Finally since almost surely the limit is the same for every subsequence in (4.1), the convergence holds for the whole sequence:

$$L_z^n \xrightarrow[n \rightarrow \infty]{w} \pi_z \quad \text{and} \quad \tilde{L}_z^n \xrightarrow[n \rightarrow \infty]{w} \tilde{\pi}_z.$$

This will conclude the proof.

We introduce some notations. Denote by

$$\begin{aligned} D(\tilde{\pi}_z, \pi_z)(u, \lambda) &= -z \left( 1 + \int \sigma^2(u, t) \tilde{\pi}(z, dt, d\zeta) \right) + \frac{\lambda}{1 + c \int \sigma^2(t, cu) \pi(z, dt, d\zeta)}, \\ d(\pi_z)(u) &= 1 + c \int \sigma^2(t, cu) \pi(z, dt, d\zeta), \\ \tilde{D}(\pi_z, \tilde{\pi}_z)(u, \lambda) &= -z \left( 1 + c \int \sigma^2(t, cu) \pi(z, dt, d\zeta) \right) + \frac{\lambda}{1 + \int \sigma^2(u, t) \tilde{\pi}(z, dt, d\zeta)}, \\ \tilde{d}(\tilde{\pi}_z)(u) &= 1 + \int \sigma^2(u, t) \tilde{\pi}(z, dt, d\zeta), \\ \kappa(\pi_z)(u) &= -z \left( 1 + c \int \sigma^2(t, u) \pi(z, dt, d\zeta) \right). \end{aligned}$$

Let  $\nu$  be a complex measure over the set  $\mathcal{H}$  (recall that  $\mathcal{H}$  is compact by (A-3)) then we denote by  $\|\nu\|_{\text{tv}}$  the total variation norm of  $\nu$ , that is

$$\begin{aligned} \|\nu\|_{\text{tv}} &= |\nu|(\mathcal{H}) \\ &= \sup \left\{ \left| \int f d\nu \right|, f \in C(\mathcal{H}; \mathbb{C}), \|f\|_{\infty} \leq 1 \right\}. \end{aligned}$$

**4.2. Step 1: Proof of the unicity of the solutions.** Consider two Stieltjes kernels  $\pi$  and  $\tilde{\pi}$  such that the support of measure  $\pi_z$  (resp. measure  $\tilde{\pi}_z$ ) is included in  $\mathcal{H}$  (resp.  $\tilde{\mathcal{H}}$ ). We prove in this section that if  $\pi$  and  $\tilde{\pi}$  satisfy equations (2.6) and (2.7), then they are unique.

Notice that the system of equations (2.6) and (2.7) remains true for every  $g \in C(\mathcal{H}; \mathbb{C})$  (consider  $g = h + ik$ ) and assume that both  $(\pi_z, \tilde{\pi}_z)$  and  $(\rho_z, \tilde{\rho}_z)$  are pairs of solutions of the given system. Let  $g \in C(\mathcal{H}; \mathbb{C})$ , then (2.6) yields:

$$\begin{aligned} \int g d\pi_z - \int g d\rho_z &= \int \frac{z g(u, \lambda) \int \sigma^2(u, t) (\tilde{\pi}(z, dt, d\zeta) - \tilde{\rho}(z, dt, d\zeta))}{D(\tilde{\pi}_z, \pi_z) \times D(\tilde{\rho}_z, \rho_z)} H(du, d\lambda) \\ &\quad + \int \frac{c \lambda g(u, \lambda) \int \sigma^2(t, cu) (\rho(z, dt, d\zeta) - \pi(z, dt, d\zeta))}{D(\tilde{\pi}_z, \pi_z) \times D(\tilde{\rho}_z, \rho_z) \times d(\pi_z) \times d(\rho_z)} H(du, d\lambda) \end{aligned}$$

and

$$\begin{aligned} \left| \int g d\pi_z - \int g d\rho_z \right| &\leq |z| \sigma_{\max}^2 \|g\|_{\infty} \|\tilde{\pi}_z - \tilde{\rho}_z\|_{\text{tv}} \int \frac{dH}{|D(\tilde{\pi}_z, \pi_z) \times D(\tilde{\rho}_z, \rho_z)|} \\ &\quad + c \sigma_{\max}^2 \|g\|_{\infty} \|\pi_z - \rho_z\|_{\text{tv}} \int \frac{\lambda H(du, d\lambda)}{|D(\tilde{\pi}_z, \pi_z) \times D(\tilde{\rho}_z, \rho_z) \times d(\pi_z) \times d(\rho_z)|} \end{aligned}$$

If one takes the supremum over the functions  $g \in C(\mathcal{H}; \mathbb{C})$ ,  $\|g\|_{\infty} \leq 1$ , one gets :

$$\|\pi_z - \rho_z\|_{\text{tv}} \leq \alpha \|\pi_z - \rho_z\|_{\text{tv}} + \beta \|\tilde{\pi}_z - \tilde{\rho}_z\|_{\text{tv}}$$

where

$$\begin{aligned} \alpha &= \alpha(\pi, \rho, \tilde{\pi}, \tilde{\rho}) = c \sigma_{\max}^2 \int \frac{\lambda H(du, d\lambda)}{|D(\tilde{\pi}_z, \pi_z) \times D(\tilde{\rho}_z, \rho_z) \times d(\pi_z) \times d(\rho_z)|} \\ \beta &= \beta(\pi, \rho, \tilde{\pi}, \tilde{\rho}) = |z| \sigma_{\max}^2 \int \frac{dH}{|D(\tilde{\pi}_z, \pi_z) \times D(\tilde{\rho}_z, \rho_z)|} \end{aligned}$$

Similarly, (2.7) yields:

$$\begin{aligned} \int g d\tilde{\pi}_z - \int g d\tilde{\rho}_z &= c \int \frac{czg(u, \lambda) \int \sigma^2(t, cu)(\pi(z, dt, d\zeta) - \rho(z, dt, d\zeta))}{\tilde{D}(\tilde{\pi}_z, \tilde{\pi}_z) \times \tilde{D}(\tilde{\rho}_z, \tilde{\rho}_z)} H(du, d\lambda) \\ &+ c \int \frac{\lambda g(u, \lambda) \int \sigma^2(u, t)(\tilde{\pi}(z, dt, d\zeta) - \tilde{\rho}(z, dt, d\zeta))}{\tilde{D}(\tilde{\pi}_z, \tilde{\pi}_z) \times \tilde{D}(\tilde{\rho}_z, \tilde{\rho}_z) \times \tilde{d}(\tilde{\pi}_z) \times \tilde{d}(\tilde{\rho}_z)} H(du, d\lambda) \\ &+ (1-c) \int_c^1 \frac{czg(u, 0) \int \sigma^2(t, u)(\pi(z, dt, d\zeta) - \rho(z, dt, d\zeta))}{\kappa(\pi_z) \times \kappa(\rho_z)} du \end{aligned}$$

and

$$\begin{aligned} \left| \int g d\tilde{\pi}_z - \int g d\tilde{\rho}_z \right| &\leq c^2 \sigma_{\max}^2 |z| \|g\|_{\infty} \|\pi_z - \rho_z\|_{\text{tv}} \int \frac{dH}{\left| \tilde{D}(\tilde{\pi}_z, \tilde{\pi}_z) \times \tilde{D}(\tilde{\rho}_z, \tilde{\rho}_z) \right|} \\ &+ c \sigma_{\max}^2 \|g\|_{\infty} \|\tilde{\pi}_z - \tilde{\rho}_z\|_{\text{tv}} \int \frac{\lambda H(du, d\lambda)}{\left| \tilde{D}(\tilde{\pi}_z, \tilde{\pi}_z) \times \tilde{D}(\tilde{\rho}_z, \tilde{\rho}_z) \times \tilde{d}(\tilde{\pi}_z) \times \tilde{d}(\tilde{\rho}_z) \right|} \\ &+ (1-c) c \sigma_{\max}^2 |z| \|g\|_{\infty} \|\pi_z - \rho_z\|_{\text{tv}} \int_c^1 \frac{du}{|\kappa(\pi_z) \times \kappa(\rho_z)|} \end{aligned}$$

As previously, by taking the supremum over  $g \in C(\mathcal{H}; \mathbb{C})$ ,  $\|g\|_{\infty} \leq 1$ , we get:

$$\|\tilde{\pi}_z - \tilde{\rho}_z\|_{\text{tv}} \leq \tilde{\alpha} \|\pi_z - \rho_z\|_{\text{tv}} + \tilde{\beta} \|\tilde{\pi}_z - \tilde{\rho}_z\|_{\text{tv}}$$

where

$$\begin{aligned} \tilde{\alpha} = \tilde{\alpha}(\pi, \rho, \tilde{\pi}, \tilde{\rho}) &= c \sigma_{\max}^2 |z| \left( c \int \frac{dH}{\left| \tilde{D}(\tilde{\pi}_z, \tilde{\pi}_z) \times \tilde{D}(\tilde{\rho}_z, \tilde{\rho}_z) \right|} + (1-c) \int_c^1 \frac{du}{|\kappa(\pi_z) \times \kappa(\rho_z)|} \right) \\ \tilde{\beta} = \tilde{\beta}(\pi, \rho, \tilde{\pi}, \tilde{\rho}) &= c \sigma_{\max}^2 \int \frac{\lambda H(du, d\lambda)}{\left| \tilde{D}(\tilde{\pi}_z, \tilde{\pi}_z) \times \tilde{D}(\tilde{\rho}_z, \tilde{\rho}_z) \times \tilde{d}(\tilde{\pi}_z) \times \tilde{d}(\tilde{\rho}_z) \right|} \end{aligned}$$

We end up with the following inequations:

$$\begin{cases} \|\pi_z - \rho_z\|_{\text{tv}} &\leq \alpha \|\pi_z - \rho_z\|_{\text{tv}} + \beta \|\tilde{\pi}_z - \tilde{\rho}_z\|_{\text{tv}} \\ \|\tilde{\pi}_z - \tilde{\rho}_z\|_{\text{tv}} &\leq \tilde{\alpha} \|\pi_z - \rho_z\|_{\text{tv}} + \tilde{\beta} \|\tilde{\pi}_z - \tilde{\rho}_z\|_{\text{tv}} \end{cases} \quad (4.2)$$

Let us prove now that for  $z \in \mathbb{C}^{\nabla}$  with  $\text{Im}(z)$  large enough, then  $\alpha < \frac{1}{2}$ .

Since  $\pi$  and  $\tilde{\pi}$  are assumed to be Stieltjes kernels,  $\text{Im}(z \int \sigma^2(u, t) \tilde{\pi}(z, dt, d\zeta)) \geq 0$  and  $\text{Im}(\int \sigma^2(t, cu) \pi(z, dt, d\zeta)) \geq 0$ . Therefore,  $\text{Im}(D(\tilde{\pi}_z, \pi_z)) \leq -\text{Im}(z)$  and hence  $|\text{Im}(D(\tilde{\pi}_z, \pi_z))| \geq \text{Im}(z)$ . Similarly,  $|\text{Im}(D(\tilde{\rho}_z, \rho_z))| \geq \text{Im}(z)$ . Thus,

$$\frac{1}{|D(\tilde{\pi}_z, \pi_z) \times D(\tilde{\rho}_z, \rho_z)|} \leq \frac{1}{\text{Im}^2(z)}.$$

Now consider  $zd(\pi_z)$ . As previously,  $\text{Im}(zd(\pi_z)) \geq \text{Im}(z)$ . As  $|zd(\pi_z)| \geq |\text{Im}(zd(\pi_z))|$ , this implies that  $\frac{1}{|zd(\pi_z)|} \leq \frac{1}{\text{Im}(z)}$  and  $\frac{1}{|d(\pi_z)|} \leq \frac{|z|}{\text{Im}(z)}$ . Since  $z \in \mathbb{C}^{\nabla}$ ,  $\frac{|z|}{\text{Im}(z)} \leq \sqrt{2}$ . The same argument holds for  $d(\rho_z)$  thus we get

$$\alpha \leq \frac{2c \sigma_{\max}^2 \int \lambda H(du, d\lambda)}{\text{Im}^2(z)} < \frac{1}{2} \quad \text{for } \text{Im}(z) \text{ large enough.}$$

With similar arguments, one can prove that

$$\beta \leq \frac{\sqrt{2}\sigma_{\max}^2}{\operatorname{Im}(z)}, \quad \tilde{\alpha} \leq \frac{3\sigma_{\max}^2}{\operatorname{Im}(z)}, \quad \tilde{\beta} \leq \frac{2\sigma_{\max}^2}{\operatorname{Im}^2(z)} \int \lambda H(dud\lambda). \quad (4.3)$$

Therefore  $\max(\alpha, \beta, \tilde{\alpha}, \tilde{\beta}) \leq \theta < \frac{1}{2}$  for  $z \in \mathbb{C}^\nabla$  and  $\operatorname{Im}(z)$  large enough where  $\theta$  does not depend on  $(\pi, \tilde{\pi}, \rho, \tilde{\rho})$ . Thus, the system (4.2) yields

$$\|\pi_z - \rho_z\|_{\text{tv}} = \|\tilde{\pi}_z - \tilde{\rho}_z\|_{\text{tv}} = 0 \quad \text{for } z \in \mathbb{C}^\nabla \text{ and } \operatorname{Im}(z) \text{ large enough.}$$

Now take  $z \in \mathbb{C}^+$  and  $g \in C(\mathcal{H})$ . Since  $\int g d\pi_z$  and  $\int g d\rho_z$  (resp.  $\int g d\tilde{\pi}_z$  and  $\int g d\tilde{\rho}_z$ ) are analytic over  $\mathbb{C}^+$  and are equal in  $\mathbb{C}^\nabla$  for  $\operatorname{Im}(z)$  large enough, they are equal everywhere. Since this is true for all  $g \in C(\mathcal{H})$ ,  $\pi_z$  and  $\rho_z$  (resp.  $\tilde{\pi}_z$  and  $\tilde{\rho}_z$ ) are identical on  $\mathbb{C}^+$ . This proves the unicity.

It remains to prove that if a solution  $(\pi_z, \tilde{\pi}_z)$  to the system exists, then  $f(z) = \int d\pi_z$  is the Stieltjes transform of a probability measure (one will prove similarly the corresponding result for  $\tilde{f}(z) = \int d\tilde{\pi}_z$ ). Recall that

$$\operatorname{Im}(f(z)) = \operatorname{Im} \left( \int d\pi_z \right) \geq \int \frac{\operatorname{Im}(z)}{|D(\tilde{\pi}_z, \pi_z)(u, \lambda)|^2} H(du, d\lambda) > 0$$

by (2.6). Moreover, since  $|f(z)| \leq \frac{1}{\operatorname{Im}(z)}$ ,  $f(z)$  is the Stieltjes transform of a subprobability measure. It remains to check that  $\lim_{y \rightarrow +\infty} iyf(iy) = -1$ . Since

$$\left| \int \sigma^2(u, t) \tilde{\pi}(iy, dt, d\zeta) \right| \leq \frac{\sigma_{\max}^2}{y} \quad \text{and} \quad \left| \int \sigma^2(t, cu) \pi(iy, dt, d\zeta) \right| \leq \frac{\sigma_{\max}^2}{y},$$

and

$$iyf(iy) = \int \frac{iy H(du, d\lambda)}{-iy(1 + \int \sigma^2(u, t) \tilde{\pi}(iy, dt, d\zeta)) + \frac{\lambda}{1+c \int \sigma^2(t, cu) \pi(iy, dt, d\zeta)}},$$

the Dominated convergence theorem yields the desired result.

**4.3. Step 2: convergence of subsequences  $L_z^{M_{\text{sub}}}$  and  $\tilde{L}_z^{M_{\text{sub}}}$ .** Let  $z_0 \in \mathbb{C}^+$ . Due to Assumption (A-3) and to the fact that  $|q_{ii}(z)| \leq \operatorname{Im}^{-1}(z)$ , Helly's theorem implies that for each subsequence of  $n$  there exists a subsequence  $M = M(n)$  and a complex measure  $\mu_{z_0}$  such that

$$L_{z_0}^M \xrightarrow[n \rightarrow \infty]{w} \mu_{z_0}.$$

Since  $L^n$  is random,  $\mu_{z_0}$  also depends on  $\omega$ . but due to (A-3), its support is included in  $\mathcal{H}$ . Let  $(z_k, k \geq 1)$  be a sequence of complex numbers dense in  $\mathbb{C}^+$ , then by Cantor diagonalization argument, one can extract a subsequence from  $M$ , say  $M_{\text{sub}}$ , such that

$$\forall k \in \mathbb{N}, \quad L_{z_k}^{M_{\text{sub}}} \xrightarrow[n \rightarrow \infty]{w} \mu_{z_k} \quad \text{and} \quad \tilde{L}_{z_k}^{M_{\text{sub}}} \xrightarrow[n \rightarrow \infty]{w} \tilde{\mu}_{z_k},$$

where  $\mu_{z_k}$  and  $\tilde{\mu}_{z_k}$  are complex measures, a priori random. Let  $g \in C(\mathcal{K})$  and let  $z \in \mathbb{C}^+$ . There exists  $z_k$  such that  $|z - z_k| \leq \epsilon$  and

$$\left| \int g dL_z^{M_{\text{sub}}(n)} - \int g dL_z^{M_{\text{sub}}(m)} \right| \leq \overbrace{\left| \int g dL_z^{M_{\text{sub}}(n)} - \int g dL_{z_k}^{M_{\text{sub}}(n)} \right|}^{(a)} + \underbrace{\left| \int g dL_{z_k}^{M_{\text{sub}}(n)} - \int g dL_{z_k}^{M_{\text{sub}}(m)} \right|}_{(b)} + \underbrace{\left| \int g dL_{z_k}^{M_{\text{sub}}(m)} - \int g dL_z^{M_{\text{sub}}(m)} \right|}_{(c)}$$

Let  $n$  and  $m$  be large enough. Since  $L_{z_k}^{M_{\text{sub}}}$  converges, (b) goes to zero. Since  $q_{ii}(z)$  is analytic and since  $|q_{ii}(z)| \leq \text{Im}^{-1}(z)$ , there exists  $K > 0$ , such that

$$\forall i \geq 1, \forall z, z' \text{ close enough, } |q_{ii}(z) - q_{ii}(z')| \leq K|z - z'|.$$

Thus  $\max\{(a), (c)\} \leq K\|g\|_{\infty}|z - z_k|$ . Therefore,  $(\int g dL_z^{M_{\text{sub}}})$  is a Cauchy sequence and converges to  $\Theta(g)(z)$ . Since  $g \mapsto \Theta(g)(z)$  is linear and since  $|\Theta(g)(z)| \leq \text{Im}^{-1}(z)\|g\|_{\infty}$ , Riesz representation's theorem yields the existence of  $\mu_z$  such that  $\Theta(g)(z) = \int g d\mu_z$  (recall that the support of  $\mu_z$  is included in  $\mathcal{H}$  which is compact). The convergence of  $\tilde{L}_z^{M_{\text{sub}}}$  can be proved similarly and (4.1) is satisfied. The second step is proved.

**4.4. Step 3: the kernels  $\mu_z$  and  $\tilde{\mu}_z$  are Stieltjes kernels.** Let us now prove that  $z \mapsto \int g d\mu_z$  is analytic over  $\mathbb{C}^+$ . Since  $|\int g dL_z^{M_{\text{sub}}}| \leq \text{Im}^{-1}(z)\|g\|_{\infty}$ , from each subsequence of  $(\int g dL_z^{M_{\text{sub}}})$ , one can extract a subsequence that converges to an analytic function. Since this limit is equal to  $\int g d\mu_z$ , the analyticity of  $z \mapsto \int g d\mu_z$  over  $\mathbb{C}^+$  is proved. Since properties (3) and (4) defining the Stieltjes kernels are satisfied by  $L_z^n$ , the kernel  $\mu_z$  inherits them. Therefore,  $\mu_z$  is a Stieltjes kernel. Similarly, one can prove that  $\tilde{\mu}_z$  is a Stieltjes kernel. The third step is proved.

**4.5. Step 4: the kernels  $\mu_z$  and  $\tilde{\mu}_z$  are almost surely equal to  $\pi_z$  and  $\tilde{\pi}_z$ .** We will now prove that almost surely for every  $z \in \mathbb{C}^+$ ,  $\mu_z$  and  $\tilde{\mu}_z$  satisfy equations (2.6) and (2.7).

In the sequel we will drop the subscript  $n$  from the notations relative to matrices, and the superscript  $n$  from  $\Lambda_{ii}^n$ . Let  $\vec{e}_i = (\delta_i(k))_{1 \leq k \leq n}$  and  $\vec{f}_i = (\delta_i(k))_{1 \leq k \leq N}$ . For the sake of simplicity,  $\Sigma^T$  will be denoted  $\Xi$ . Consider the following notations:

Matrix	$Y$	$\Lambda$	$\Sigma$	$\Sigma_{(i)}^T$	$Y^T$	$\Lambda^T$	$\Xi$	$\Xi_{(i)}^T$
$i$ th row	$\vec{y}_i$	$\Lambda_{ii}\vec{e}_i$	$\vec{\xi}_i$	$\vec{\eta}_i$	$\vec{y}_i$	$\Lambda_{ii}\vec{f}_i$	$\vec{\xi}_i$	$\vec{\eta}_i$
Matrix when $i$ th row is deleted	-	-	$\Sigma_{(i)}$	$\Sigma_{(i,i)}^T$	-	-	$\Xi_{(i)}$	$\Xi_{(i,i)}^T$

In particular,  $\vec{\xi}_i = \vec{y}_i + \Lambda_{ii}\vec{e}_i$  and  $\vec{\xi}_i = \vec{y}_i + \Lambda_{ii}\vec{f}_i$  for  $1 \leq i \leq N$ . We will denote by  $D_i$  and  $\Delta_j$  the respectively  $n \times n$  and  $N \times N$  diagonal matrices defined by

$$D_i = \text{diag} \left( \frac{\sigma\left(\frac{i}{N}, \frac{1}{n}\right)}{\sqrt{n}}, \dots, \frac{\sigma\left(\frac{i}{N}, 1\right)}{\sqrt{n}} \right), \quad \Delta_j = \text{diag} \left( \frac{\sigma\left(\frac{1}{N}, \frac{j}{n}\right)}{\sqrt{n}}, \dots, \frac{\sigma\left(1, \frac{j}{n}\right)}{\sqrt{n}} \right).$$

Finally, for  $1 \leq i \leq N$ , we denote by  $D_{(i,i)}$  and  $\Delta_{(i,i)}$  the matrices that remain after deleting row  $i$  and column  $i$  from  $D_i$  and  $\Delta_i$  respectively.

We can state our first lemma:



**Lemma 4.1.** *Assume that (A-1), (A-2) and (A-3) hold and let  $z \in \mathbb{C}^+$  be fixed.*

(1) *The  $i$ th diagonal element  $q_{ii}(z)$  of the matrix  $(\Sigma\Sigma^T - zI_N)^{-1}$  can be written :*

$$q_{ii}(z) = \frac{1}{-z - \frac{z}{n} \sum_{k=1}^n \sigma^2 \left( \frac{i}{N}, \frac{k}{n} \right) \tilde{q}_{kk}(z) + \frac{\Lambda_{ii}^2}{1 + \frac{1}{n} \sum_{k=1}^N \sigma^2 \left( \frac{k}{N}, \frac{i}{n} \right) q_{ii}(z) + \varepsilon_{i,n}^{(4)} + \varepsilon_{i,n}^{(5)}} + \Lambda_{ii} \varepsilon_{i,n}^{(1)} + \varepsilon_{i,n}^{(2)} + \varepsilon_{i,n}^{(3)}} \quad (4.4)$$

where  $1 \leq i \leq N$  and

$$\begin{aligned} \varepsilon_{i,n}^{(1)} &= -z \bar{y}_i \cdot \left( \Sigma_{(i)}^T \Sigma_{(i)} - zI \right)^{-1} \bar{e}_i^T - z \bar{e}_i \left( \Sigma_{(i)}^T \Sigma_{(i)} - zI \right)^{-1} \bar{y}_i^T \\ \varepsilon_{i,n}^{(2)} &= -z \bar{y}_i \cdot \left( \Sigma_{(i)}^T \Sigma_{(i)} - zI \right)^{-1} \bar{y}_i^T + z \operatorname{Tr} \left( D_i^2 \left( \Sigma_{(i)}^T \Sigma_{(i)} - zI \right)^{-1} \right) \\ \varepsilon_{i,n}^{(3)} &= z \operatorname{Tr} \left( D_i^2 \left( \Sigma^T \Sigma - zI \right)^{-1} \right) - z \operatorname{Tr} \left( D_i^2 \left( \Sigma_{(i)}^T \Sigma_{(i)} - zI \right)^{-1} \right) \\ \varepsilon_{i,n}^{(4)} &= \bar{\eta}_i \cdot \left( \Sigma_{(i,i)} \Sigma_{(i,i)}^T - zI \right)^{-1} \bar{\eta}_i^T - \operatorname{Tr} \left( \Delta_{(i,i)}^2 \left( \Sigma_{(i,i)} \Sigma_{(i,i)}^T - zI \right)^{-1} \right) \\ \varepsilon_{i,n}^{(5)} &= \operatorname{Tr} \left( \Delta_{(i,i)}^2 \left( \Sigma_{(i,i)} \Sigma_{(i,i)}^T - zI \right)^{-1} \right) - \operatorname{Tr} \left( \Delta_i^2 \left( \Sigma \Sigma^T - zI \right)^{-1} \right) \end{aligned}$$

Moreover, almost surely

$$\forall k, 1 \leq k \leq 5, \quad \lim_{n \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \left| \varepsilon_{i,n}^{(k)} \right| = 0. \quad (4.5)$$

(2) *If  $1 \leq i \leq N$  then the  $i$ th diagonal element  $\tilde{q}_{ii}(z)$  of the matrix  $(\Sigma^T \Sigma - zI_n)^{-1} = (\Xi \Xi^T - zI_n)^{-1}$  can be written:*

$$\tilde{q}_{ii}(z) = \frac{1}{-z - \frac{z}{n} \sum_{k=1}^N \sigma^2 \left( \frac{k}{N}, \frac{i}{n} \right) q_{kk}(z) + \frac{\Lambda_{ii}^2}{1 + \frac{1}{n} \sum_{k=1}^n \sigma^2 \left( \frac{i}{N}, \frac{k}{n} \right) \tilde{q}_{ii}(z) + \tilde{\varepsilon}_{i,n}^{(4)} + \tilde{\varepsilon}_{i,n}^{(5)}} + \Lambda_{ii} \tilde{\varepsilon}_{i,n}^{(1)} + \tilde{\varepsilon}_{i,n}^{(2)} + \tilde{\varepsilon}_{i,n}^{(3)}}. \quad (4.6)$$

If  $N+1 \leq i \leq n$ , then  $\tilde{q}_{ii}$  can be written:

$$\tilde{q}_{ii}(z) = \frac{1}{-z - \frac{z}{n} \sum_{k=1}^N \sigma^2 \left( \frac{k}{N}, \frac{i}{n} \right) q_{kk}(z) + \tilde{\varepsilon}_{i,n}^{(2)} + \tilde{\varepsilon}_{i,n}^{(3)}}. \quad (4.7)$$

where

$$\begin{aligned} \tilde{\varepsilon}_{i,n}^{(1)} &= -z \bar{y}_i \cdot \left( \Xi_{(i)}^T \Xi_{(i)} - zI \right)^{-1} \bar{f}_i^T - z \bar{f}_i \left( \Xi_{(i)}^T \Xi_{(i)} - zI \right)^{-1} \bar{y}_i^T \\ \tilde{\varepsilon}_{i,n}^{(2)} &= -z \bar{y}_i \cdot \left( \Xi_{(i)}^T \Xi_{(i)} - zI \right)^{-1} \bar{y}_i^T + z \operatorname{Tr} \left( \Delta_i^2 \left( \Xi_{(i)}^T \Xi_{(i)} - zI \right)^{-1} \right) \\ \tilde{\varepsilon}_{i,n}^{(3)} &= z \operatorname{Tr} \left( \Delta_i^2 \left( \Xi^T \Xi - zI \right)^{-1} \right) - z \operatorname{Tr} \left( \Delta_i^2 \left( \Xi_{(i)}^T \Xi_{(i)} - zI \right)^{-1} \right) \\ \tilde{\varepsilon}_{i,n}^{(4)} &= \bar{\eta}_i \cdot \left( \Xi_{(i,i)} \Xi_{(i,i)}^T - zI \right)^{-1} \bar{\eta}_i^T - \operatorname{Tr} \left( D_{(i,i)}^2 \left( \Xi_{(i,i)} \Xi_{(i,i)}^T - zI \right)^{-1} \right) \\ \tilde{\varepsilon}_{i,n}^{(5)} &= \operatorname{Tr} \left( D_{(i,i)}^2 \left( \Xi_{(i,i)} \Xi_{(i,i)}^T - zI \right)^{-1} \right) - \operatorname{Tr} \left( D_i^2 \left( \Xi \Xi^T - zI \right)^{-1} \right) \end{aligned}$$

Moreover, almost surely

$$\begin{cases} \text{for } k = 1, 4, 5 & \lim_{n \rightarrow \infty} \frac{1}{N} \sum_{1 \leq i \leq N} |\hat{\varepsilon}_{i,n}^{(k)}| = 0, \\ \text{for } k = 2, 3 & \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{1 \leq i \leq n} |\hat{\varepsilon}_{i,n}^{(k)}| = 0. \end{cases} \quad (4.8)$$

*Proof of Lemma 4.1.* Since  $q_{ii}(z) = (\Sigma \Sigma^T - zI)_{ii}^{-1}$ , this element is the inverse of the Schur complement of  $(\Sigma_{(i)} \Sigma_{(i)}^T - zI)$  in  $(\Sigma \Sigma^T - zI)$  (see for instance [12], Appendix A). In other words

$$q_{ii}(z) = \left( \|\vec{\xi}_i\|^2 - z - \vec{\xi}_i \cdot \Sigma_{(i)}^T (\Sigma_{(i)} \Sigma_{(i)}^T - zI)^{-1} \Sigma_{(i)} \vec{\xi}_i^T \right)^{-1}.$$

Using the identity

$$I - \Sigma_{(i)}^T (\Sigma_{(i)} \Sigma_{(i)}^T - zI)^{-1} \Sigma_{(i)} = -z (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1},$$

we have

$$\begin{aligned} q_{ii}(z) &= \frac{1}{-z - z \vec{\xi}_i \cdot (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \vec{\xi}_i^T} \\ &= \frac{1}{-z - z \vec{y}_i \cdot (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \vec{y}_i^T - z \Lambda_{ii}^2 \vec{e}_i (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \vec{e}_i^T + \Lambda_{ii} \varepsilon_{i,n}^{(1)}} \\ &= \frac{1}{-z - \frac{z}{n} \sum_{k=1}^n \sigma^2 \left( \frac{i}{N}, \frac{k}{n} \right) \tilde{q}_{kk}(z) - z \Lambda_{ii}^2 \vec{e}_i (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \vec{e}_i^T + \Lambda_{ii} \varepsilon_{i,n}^{(1)} + \varepsilon_{i,n}^{(2)} + \varepsilon_{i,n}^{(3)}}. \end{aligned}$$

Similarly, we have

$$\begin{aligned} \vec{e}_i (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \vec{e}_i^T &= (\Sigma_{(i)}^T \Sigma_{(i)} - zI)_{ii}^{-1} \\ &= \frac{1}{-z - z \vec{\eta}_i \cdot (\Sigma_{(ii)} \Sigma_{(ii)}^T - zI)^{-1} \vec{\eta}_i^T} \\ &= \frac{1}{-z \left( 1 + \frac{1}{n} \sum_{k=1}^N \sigma^2 \left( \frac{k}{N}, \frac{i}{n} \right) q_{kk}(z) + \varepsilon_{i,n}^{(4)} + \varepsilon_{i,n}^{(5)} \right)} \end{aligned} \quad (4.9)$$

And (4.4) is established. It is important to already note that since  $\vec{\eta}_i$  is the  $i$ th row of  $\Sigma_{(i)}^T$ ,  $\mathbb{E} \vec{\eta}_i = 0$  (while  $\mathbb{E} \vec{y}_i = (0, \dots, \Lambda_{ii}, \dots, 0)$ ). If  $i \leq N$ , (4.6) can be established in the same way. If  $i \geq N + 1$ , then  $\vec{\xi}_i$  is centered: There are no more  $\Lambda_{ii}$  and all the terms involving  $\Lambda_{ii}$  disappear in (4.6), which yields (4.7).

We now prove that

$$\frac{1}{N} \sum_{i=1}^N |\varepsilon_{i,n}^{(1)}| \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0. \quad (4.10)$$

One will prove similarly that  $\frac{1}{N} \sum_{i=1}^N |\varepsilon_{i,n}^{(1)}| \rightarrow 0$  a.s. Denote by  $R_n = (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} = (\rho_{ij})$ . Since  $R_n$  is symmetric,  $\varepsilon_{i,n}^{(1)} = -z 2 \vec{y}_i \cdot R_n \vec{e}_i^T$  and

$$|\vec{y}_i \cdot R_n \vec{e}_i^T|^4 = \left| \sum_{k=1}^n Y_{ik} \rho_{ki} \right|^4 = \sum_{k_1, k_2, l_1, l_2} Y_{ik_1} Y_{ik_2} Y_{il_1} Y_{il_2} \rho_{k_1 i} \bar{\rho}_{k_2 i} \rho_{l_1 i} \bar{\rho}_{l_2 i}.$$

Denote by  $\mathbb{E}_{R_n}$  the conditional expectation with respect to the  $\sigma$ -algebra generated by  $R_n$ . Since  $\vec{y}_i$  and  $R_n$  are independent and since  $\mathbb{E}Y_{ik} = 0$ , we get:

$$\begin{aligned} \mathbb{E}_{R_n} |\vec{y}_i \cdot R_n \vec{e}_i^T|^4 &= 2\mathbb{E}_{R_n} \sum_{k,l;k \neq l} Y_{ik}^2 |\rho_{ki}|^2 Y_{il}^2 |\rho_{li}|^2 \\ &\quad + \mathbb{E}_{R_n} \sum_{k,l;k \neq l} Y_{ik}^2 \rho_{ki}^2 Y_{il}^2 \rho_{li}^2 + \mathbb{E}_{R_n} \sum_k Y_{ik}^4 |\rho_{ki}|^4 \\ &\leq 4\mathbb{E}(X_{ij}^n)^4 \frac{\sigma_{\max}^4}{n^2} \sum_{k,l} |\rho_{ki}|^2 |\rho_{li}|^2 = 4\mathbb{E}(X_{ij}^n)^4 \frac{\sigma_{\max}^4}{n^2} \left( \sum_k |\rho_{ki}|^2 \right)^2 \end{aligned}$$

but  $\sum_k |\rho_{ki}|^2 = \|R_n e_i\|^2 \leq \|R_n\|^2 \leq \frac{1}{\text{Im}^2(z)}$ . Therefore,

$$\mathbb{E} \left| \varepsilon_{i,n}^{(1)} \right|^4 \leq \frac{|2z|^4 4\mathbb{E}(X_{ij}^n)^4 \sigma_{\max}^4}{n^2 \text{Im}^4(z)} \propto \frac{1}{n^2}. \quad (4.11)$$

Finally,

$$\begin{aligned} \mathbb{P} \left\{ \frac{1}{N} \sum_{i=1}^N \left| \varepsilon_{i,n}^{(1)} \right| > \delta \right\} &\leq \frac{1}{\delta^4 N^4} \mathbb{E} \left( \sum_{i=1}^N \left| \varepsilon_{i,n}^{(1)} \right| \right)^4 \\ &\stackrel{(a)}{\leq} \frac{1}{\delta^4 N^4} \left( \sum_{i=1}^N \left( \mathbb{E} \left| \varepsilon_{i,n}^{(1)} \right|^4 \right)^{\frac{1}{4}} \right)^4 \leq \frac{1}{\delta^4} \sup_{1 \leq i \leq N} \mathbb{E} \left| \varepsilon_{i,n}^{(1)} \right|^4 \stackrel{(b)}{\propto} \frac{1}{n^2} \end{aligned}$$

where (a) follows from Minkowski's inequality and (b) from (4.11) and Borel-Cantelli's lemma yields Eq. (4.10).

Let us now prove that

$$\frac{1}{N} \sum_{i=1}^N \left| \varepsilon_{i,n}^{(2)} \right| \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0. \quad (4.12)$$

One will prove similarly that  $\frac{1}{n} \sum |\tilde{\varepsilon}_{i,n}^{(2)}|$ ,  $\frac{1}{N} \sum |\varepsilon_{i,n}^{(4)}|$  and  $\frac{1}{N} \sum |\tilde{\varepsilon}_{i,n}^{(4)}|$  go to zero a.s. Denote by  $\vec{x}_i = (X_{i1}, \dots, X_{in})$  and write  $\vec{y}_i = \vec{x}_i \cdot D_i$ . In particular,

$$\vec{y}_i \cdot (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \vec{y}_i^T = \vec{x}_i \cdot D_i (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} D_i^T \vec{x}_i^T$$

where  $\vec{x}_i$  and  $D_i (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} D_i^T$  are independent. Lemma 2.7 in [2] states that

$$\begin{aligned} \mathbb{E} |\vec{x}_i \cdot C \vec{x}_i^T - \text{Tr} C|^p &\leq K_p \left( (\mathbb{E}(X_{i1})^4 \text{Tr} C C^T)^{p/2} + \mathbb{E}(X_{i1})^{2p} \text{Tr}(C C^T)^{p/2} \right) \quad (4.13) \end{aligned}$$

for all  $p \geq 2$ . Take  $p = 2 + \epsilon/2$  where  $\epsilon$  is given by (A-1) and let  $C = D_i (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} D_i^T$ . Then

$$\forall q \geq 1, \quad \text{Tr}(C C^T)^q \leq \frac{\sigma_{\max}^{4q}}{n^{2q-1}} \times \frac{1}{\text{Im}^{2q}(z)}. \quad (4.14)$$

Therefore, (4.13) and (4.14) yield

$$\mathbb{E} |\vec{x}_i \cdot C \vec{x}_i^T - \text{Tr} C|^{2+\epsilon/2} \leq \frac{K_1}{n^{1+\epsilon/4}} + \frac{K_2}{n^{1+\epsilon}} \leq \frac{K}{n^{1+\epsilon/4}}$$

where the constants  $K$ ,  $K_1$  and  $K_2$  depend on the moments of  $X_{i1}$ , on  $\sigma_{\max}$  and on  $\text{Im}(z)$ . Thus

$$\mathbb{E} \left| \varepsilon_{i,n}^{(2)} \right|^p \leq \frac{K |z|^p}{n^{1+\epsilon/4}}. \quad (4.15)$$

Finally,

$$\begin{aligned}
\mathbb{P} \left\{ \frac{1}{N} \sum_{i=1}^N |\varepsilon_{i,n}^{(2)}| > \delta \right\} &\leq \frac{1}{\delta^p N^p} \mathbb{E} \left( \sum_{i=1}^N |\varepsilon_{i,n}^{(2)}| \right)^p \\
&\stackrel{(a)}{\leq} \frac{1}{\delta^p N^p} \left( \sum_{i=1}^N \left( \mathbb{E} |\varepsilon_{i,n}^{(2)}|^p \right)^{\frac{1}{p}} \right)^p \\
&\leq \frac{1}{\delta^p} \sup_{1 \leq i \leq N} \mathbb{E} |\varepsilon_{i,n}^{(2)}|^p \stackrel{(b)}{\asymp} \frac{1}{n^{1+\epsilon/4}}
\end{aligned}$$

where (a) follows from Minkowski's inequality and (b) from (4.15), and Borel-Cantelli's lemma yields (4.12).

We now prove that

$$\frac{1}{N} \sum_{i=1}^N |\varepsilon_{i,n}^{(3)}| \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0. \quad (4.16)$$

One will prove similarly that  $\frac{1}{n} \sum |\tilde{\varepsilon}_{i,n}^{(3)}|$  goes to zero. Since  $\Sigma^T \Sigma = \Sigma_{(i)}^T \Sigma_{(i)} + \bar{\xi}_i^T \bar{\xi}_i$ , Lemma 2.6 in [18] yields:

$$\left| \text{Tr} \left( (\Sigma^T \Sigma - zI)^{-1} - (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \right) D_i^2 \right| \leq \frac{\sigma_{\max}^2}{n \text{Im}(z)},$$

In particular,

$$|\varepsilon_{i,n}^{(3)}| \leq \frac{|z| \sigma_{\max}^2}{n \text{Im}(z)} \quad (4.17)$$

which immediatly yields (4.16).

We finally prove that

$$\frac{1}{N} \sum_{i=1}^N |\varepsilon_{i,n}^{(5)}| \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0. \quad (4.18)$$

One will prove similarly that  $\frac{1}{N} \sum |\tilde{\varepsilon}_{i,n}^{(5)}|$  goes to zero. Write

$$\begin{aligned}
\varepsilon_{i,n}^{(5)} &= \text{Tr} \Delta_{(i,i)}^2 \left( \Sigma_{(i,i)} \Sigma_{(i,i)}^T - zI \right)^{-1} - \text{Tr} \Delta_{(i,i)}^2 \left( \Sigma_{(i)} \Sigma_{(i)}^T - zI \right)^{-1} \\
&\quad + \text{Tr} \Delta_{(i,i)}^2 \left( \Sigma_{(i)} \Sigma_{(i)}^T - zI \right)^{-1} - \text{Tr} \Delta_i^2 \left( \Sigma \Sigma^T - zI \right)^{-1}
\end{aligned}$$

As for  $\varepsilon_{i,n}^{(3)}$ , one can prove that

$$\left| \text{Tr} \Delta_{(i,i)}^2 \left( \Sigma_{(i,i)} \Sigma_{(i,i)}^T - zI \right)^{-1} - \text{Tr} \Delta_{(i,i)}^2 \left( \Sigma_{(i)} \Sigma_{(i)}^T - zI \right)^{-1} \right| \leq \frac{\sigma_{\max}^2}{n \text{Im}(z)}$$

by applying Lemma 2.6 in [18]. Let

$$\kappa_{i,n} = \text{Tr} \Delta_{(i,i)}^2 \left( \Sigma_{(i)} \Sigma_{(i)}^T - zI \right)^{-1} - \text{Tr} \Delta_i^2 \left( \Sigma \Sigma^T - zI \right)^{-1}.$$

By applying to  $\Sigma\Sigma^T - zI$  the identities relative to the inverse of a partitioned matrix (see [12], Appendix A), we obtain:  $\text{Tr} \Delta_i^2 (\Sigma\Sigma^T - zI)^{-1} = \Psi_1 + \Psi_2 + \Psi_3$  where

$$\begin{aligned}\Psi_1 &= \text{Tr} \Delta_{(i,i)}^2 \left( \Sigma_{(i)} \Sigma_{(i)}^T - zI \right)^{-1} \\ \Psi_2 &= \frac{\text{Tr} \Delta_{(i,i)}^2 \left( \Sigma_{(i)} \Sigma_{(i)}^T - zI \right)^{-1} \Sigma_{(i)} \vec{\xi}_i^T \vec{\xi}_i \Sigma_{(i)}^T \left( \Sigma_{(i)} \Sigma_{(i)}^T - zI \right)^{-1}}{-z - z \vec{\xi}_i^T (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \vec{\xi}_i^T} \\ \Psi_3 &= \frac{1}{n} \frac{\sigma^2 \left( \frac{i}{N}, \frac{i}{n} \right)}{-z - z \vec{\xi}_i^T (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \vec{\xi}_i^T}.\end{aligned}$$

In particular,  $\kappa_{i,n} = -\Psi_2 - \Psi_3$ . We have

$$\begin{aligned}|\Psi_2| &= \left| \frac{\vec{\xi}_i^T \Sigma_{(i)}^T \left( \Sigma_{(i)} \Sigma_{(i)}^T - zI \right)^{-1} \Delta_{(i,i)}^2 \left( \Sigma_{(i)} \Sigma_{(i)}^T - zI \right)^{-1} \Sigma_{(i)} \vec{\xi}_i^T}{-z - z \vec{\xi}_i^T (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \vec{\xi}_i^T} \right| \\ &\leq \|\Delta_{(i,i)}\|^2 \frac{\left\| \left( \Sigma_{(i)} \Sigma_{(i)}^T - zI \right)^{-1} \Sigma_{(i)} \vec{\xi}_i^T \right\|^2}{\left| z + z \vec{\xi}_i^T (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \vec{\xi}_i^T \right|}\end{aligned}$$

Let  $\Sigma_{(i)} = \sum_{l=1}^{N-1} \nu_l u_l v_l^T$  be a singular value decomposition of  $\Sigma_{(i)}$  where  $\nu_l$ ,  $u_l$ , and  $v_l$  are respectively the singular values, left singular vectors, and right singular vectors of  $\Sigma_{(i)}$ . Then

$$\left\| \left( \Sigma_{(i)} \Sigma_{(i)}^T - zI \right)^{-1} \Sigma_{(i)} \vec{\xi}_i^T \right\|^2 = \sum_{l=1}^{N-1} \frac{\nu_l^2 \left| v_l^T \vec{\xi}_i^T \right|^2}{|\nu_l^2 - z|^2}$$

and

$$\text{Im} \left( z + z \vec{\xi}_i^T (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \vec{\xi}_i^T \right) = \text{Im}(z) \left( 1 + \sum_{l=1}^{N-1} \frac{\nu_l^2 \left| v_l^T \vec{\xi}_i^T \right|^2}{|\nu_l^2 - z|^2} \right).$$

As a consequence,  $|\Psi_2| \leq \frac{\sigma_{\max}^2}{n} \frac{1}{\text{Im}(z)}$ . Furthermore, since  $\text{Im} \left( z \vec{\xi}_i^T (\Sigma_{(i)}^T \Sigma_{(i)} - zI)^{-1} \vec{\xi}_i^T \right) \geq 0$  by Proposition 2.2-(3), we have  $|\Psi_3| \leq \frac{\sigma_{\max}^2}{n} \frac{1}{\text{Im}(z)}$ . Thus,  $\left| \varepsilon_{i,n}^{(5)} \right| \leq \frac{\sigma_{\max}^2}{n} \frac{3}{\text{Im}(z)}$ , which immediately yields (4.18). Lemma 4.1 is proved.  $\square$

Recall notation  $D$  introduced at the beginning of Section 3:

$$D(\tilde{\pi}_z, \pi_z)(u, \lambda) = -z \left( 1 + \int \sigma^2(u, t) \tilde{\pi}(z, dt, d\zeta) \right) + \frac{\lambda}{1 + c \int \sigma^2(t, cu) \pi(z, dt, d\zeta)}.$$

We also remind the notations  $M_{\text{sub}}$ ,  $\mu_z$  and  $\tilde{\mu}_z$  (for simplicity, denote by  $n^* = M_{\text{sub}}(n)$  and  $N^* = N(n^*)$ ). It has been shown in Step 2 that from each subsequence of  $(n)$  one can extract a subsequence  $n^*$  such that

$$\forall z \in \mathbb{C}^+, \quad L_z^{n^*} \xrightarrow{n \rightarrow \infty} \mu_z \quad \text{and} \quad \tilde{L}_z^{n^*} \xrightarrow{n \rightarrow \infty} \tilde{\mu}_z. \quad (4.19)$$

Note that  $\mu$  and  $\tilde{\mu}$  depend on both  $\omega \in \Omega$  and  $M_{\text{sub}}$ .

**Corollary 4.2.** *Let the assumptions be as in Lemma 4.1 and let  $z \in \mathbb{C}^+$  be fixed. Then there exists a measurable set  $\mathcal{A}_z$  with probability one such that for every  $\omega \in \mathcal{A}_z$ :*

$$\forall g \in C(\mathcal{K}), \quad \lim_{n \rightarrow \infty} \left| \frac{1}{N^*} \sum_{i=1}^{N^*} q_{ii}^{n^*} g(\Lambda_{ii}^2, i/N^*) - \frac{1}{N^*} \sum_{i=1}^{N^*} \frac{g(\Lambda_{ii}^2, i/N^*)}{D(\tilde{\mu}_z, \mu_z)(\frac{i}{N^*})} \right| = 0. \quad (4.20)$$

where  $n^*$ ,  $\mu$  and  $\tilde{\mu}$  are given in (4.19) and  $N^* = N(n^*)$ .

*Proof of Corollary 4.2.* Let  $\mathcal{A}_z$  be the intersection of the sets where (4.5) and (4.8) hold. Following the notations  $D$  and  $d$  introduced at the beginning of Section 4, we introduce their empirical counterparts along the given subsequences:

$$\begin{aligned} d^{n^*}(u) &= 1 + \frac{1}{n^*} \sum_{k=1}^{N^*} \sigma^2 \left( \frac{k}{N^*}, \frac{N^*}{n^*} u \right) q_{ii}(z) + \varepsilon_{i,n^*}^{(4)} + \varepsilon_{i,n^*}^{(5)} \\ D^{n^*}(u) &= -z - \frac{z}{n^*} \sum_{k=1}^{n^*} \sigma^2 \left( u, \frac{k}{n^*} \right) \tilde{q}_{kk}(z) + \frac{\Lambda_{ii}^2}{d^{n^*}(u)} \\ &\quad + \Lambda_{ii} \varepsilon_{i,n^*}^{(1)} + \varepsilon_{i,n^*}^{(2)} + \varepsilon_{i,n^*}^{(3)} \end{aligned}$$

Since  $q_{ii} = (D^{n^*}(i/N^*))^{-1}$  by (4.4) and  $(\Sigma_{(i)}^T \Sigma_{(i)} - zI)_{ii}^{-1} = (-z d^{n^*}(i/N^*))^{-1}$  by (4.9), Proposition 2.2-(1) yields:

$$\frac{1}{|D^{n^*}(i/N^*)|} \leq \frac{1}{\text{Im}(z)} \quad \text{and} \quad \frac{1}{|d^{n^*}(i/N^*)|} \leq \frac{|z|}{\text{Im}(z)}. \quad (4.21)$$

On the other hand, since  $\mu_z$  and  $\tilde{\mu}_z$  are Stieltjes kernels, we have:

$$\frac{1}{|D(\tilde{\mu}_z, \mu_z)(i/N^*, \Lambda_{ii}^2)|} \leq \frac{1}{\text{Im}(z)} \quad \text{and} \quad \frac{1}{|d(\mu_z)(i/N^*)|} \leq \frac{|z|}{\text{Im}(z)}. \quad (4.22)$$

Therefore,

$$\begin{aligned} q_{ii}^{n^*} - \frac{1}{D(\tilde{\mu}_z, \mu_z)(\frac{i}{N^*}, \Lambda_{ii}^2)} &= \frac{-z \left( \int \sigma^2(i/N^*, \cdot) d\tilde{L}_z^{n^*} - \int \sigma^2(i/N^*, \cdot) d\tilde{\mu}_z \right)}{D^{n^*}(\frac{i}{N^*}) \times D(\tilde{\mu}_z, \mu_z)(\frac{i}{N^*}, \Lambda_{ii}^2)} \\ &+ \frac{\Lambda_{ii} \varepsilon_{i,n^*}^{(1)} + \varepsilon_{i,n^*}^{(2)} + \varepsilon_{i,n^*}^{(3)}}{D^{n^*}(\frac{i}{N^*}) \times D(\tilde{\mu}_z, \mu_z)(\frac{i}{N^*}, \Lambda_{ii}^2)} \\ &+ \frac{\Lambda_{ii}^2 \left( \frac{N^*}{n^*} \int \sigma^2(\cdot, i/n^*) dL_z^{n^*} - c \int \sigma^2(\cdot, i/n^*) d\mu_z \right)}{d(\mu_z)(\frac{i}{N^*}) \times d^{n^*}(\frac{i}{N^*}) \times D^{n^*}(\frac{i}{N^*}) \times D(\tilde{\mu}_z, \mu_z)(\frac{i}{N^*}, \Lambda_{ii}^2)} \\ &+ \frac{\Lambda_{ii}^2 \left( \varepsilon_{i,n^*}^{(4)} + \varepsilon_{i,n^*}^{(5)} \right)}{d(\mu_z)(\frac{i}{N^*}) \times d^{n^*}(\frac{i}{N^*}) \times D^{n^*}(\frac{i}{N^*}) \times D(\tilde{\mu}_z, \mu_z)(\frac{i}{N^*}, \Lambda_{ii}^2)} \end{aligned}$$

Recall that the  $\Lambda_{ii}$ 's are assumed to be bounded (say  $|\Lambda_{ii}| \leq K$ ). Due to (4.21) and (4.22), we get:

$$\begin{aligned} & \left| q_{ii}^{n^*} - \frac{1}{D(\tilde{\mu}_z, \mu_z) \left( \frac{i}{N^*}, \Lambda_{ii}^2 \right)} \right| \leq \frac{|z|}{\text{Im}^2(z)} \underbrace{\left| \int \sigma^2(i/N^*, \cdot) d\tilde{L}_z^{n^*} - \int \sigma^2(i/N^*, \cdot) d\tilde{\mu}_z \right|}_{I(i, n^*)} \\ & + \frac{K \left| \varepsilon_{i, n^*}^{(1)} \right| + \left| \varepsilon_{i, n^*}^{(2)} \right| + \left| \varepsilon_{i, n^*}^{(3)} \right|}{\text{Im}^2(z)} \\ & + \frac{|z|^2 K^2}{\text{Im}^4(z)} \left( \underbrace{\left| \frac{N^*}{n^*} \int \sigma^2(\cdot, i/n^*) dL_z^{n^*} - c \int \sigma^2(\cdot, i/n^*) d\mu_z \right|}_{J(i, n^*)} + \left| \varepsilon_{i, n^*}^{(4)} \right| + \left| \varepsilon_{i, n^*}^{(5)} \right| \right) \end{aligned}$$

In order to prove  $\sup_{i \leq N^*} I(i, n^*) \rightarrow 0$ , recall that  $C([0, 1]^2) = C([0, 1]) \otimes C([0, 1])$ . In particular,  $\forall \epsilon > 0$ , there exists  $k \in \mathbb{N}$ ,  $g_l \in C([0, 1])$  and  $h_l \in C([0, 1])$  for  $l \leq k$  such that  $\sup_{x, t} \left| \sigma^2(x, t) - \sum_{l=1}^k g_l(x) h_l(t) \right| \leq \epsilon$ . Therefore,

$$\left| \int \sigma^2(i/N^*, \cdot) d\tilde{L}_z^{n^*} - \int \sigma^2(i/N^*, \cdot) d\tilde{\mu}_z \right| \leq \sup_x \left| \int \sigma^2(x, \cdot) d\tilde{L}_z^{n^*} - \int \sigma^2(x, \cdot) d\tilde{\mu}_z \right| \xrightarrow{n \rightarrow \infty} 0$$

which implies that  $\sup_{i \leq N^*} |I(i, n^*)|$  goes to zero. One can prove similarly that  $\sup_{i \leq N^*} J(i, n^*)$  goes to zero. Therefore,

$$\begin{aligned} & \left| \frac{1}{N^*} \sum_{i=1}^{N^*} q_{ii}^{n^*} g(\Lambda_{ii}^2, i/N^*) - \frac{1}{N^*} \sum_{i=1}^{N^*} \frac{g(\Lambda_{ii}^2, i/N^*)}{D(\tilde{\mu}_z, \mu_z) \left( \frac{i}{N^*}, \Lambda_{ii} \right)} \right| \\ & \leq \frac{|z| \|g\|_\infty}{\text{Im}^2(z)} \sup_{i \leq N^*} I(i, n^*) + \frac{\|g\|_\infty}{\text{Im}^2(z)} \left( K \frac{1}{N^*} \sum_{i=1}^{N^*} |\varepsilon_{i, n^*}^{(1)}| + \frac{1}{N^*} \sum_{i=1}^{N^*} |\varepsilon_{i, n^*}^{(2)}| + \frac{1}{N^*} \sum_{i=1}^{N^*} |\varepsilon_{i, n^*}^{(3)}| \right) \\ & \quad + \frac{|z|^2 K^2 \|g\|_\infty}{\text{Im}^4(z)} \left( \sup_{i \leq N^*} J(i, n^*) + \frac{1}{N^*} \sum_{i=1}^{N^*} |\varepsilon_{i, n^*}^{(4)}| + \frac{1}{N^*} \sum_{i=1}^{N^*} |\varepsilon_{i, n^*}^{(5)}| \right). \end{aligned}$$

and (4.20) is proved with the help of Lemma 4.1.  $\square$

We now come back to the proof of the fourth step of Theorem 2.3.

A direct application of the Dominated convergence theorem yields that  $(\lambda, u) \mapsto \frac{g(\lambda, u)}{D(\tilde{\mu}_z, \mu_z)(u, \lambda)}$  is bounded and continuous therefore (A-3) yields

$$\frac{1}{N^*} \sum_{i=1}^{N^*} \frac{g(\Lambda_{ii}^2, i/N^*)}{D(\tilde{\mu}_z, \mu_z)(i/N^*, \Lambda_{ii}^2)} \xrightarrow{n \rightarrow \infty} \int \frac{g(\lambda, u)}{D(\tilde{\mu}_z, \mu_z)(u, \lambda)} H(d\lambda, du). \quad (4.23)$$

Moreover,

$$\frac{1}{N^*} \sum_{i=1}^{N^*} g(\Lambda_{ii}^2, i/N^*) q_{ii} \xrightarrow{n \rightarrow \infty} \int g d\mu_z. \quad (4.24)$$

Consider now a countable set  $\mathcal{C}$  with a limit point. Since  $\mathcal{C}$  is countable, (4.20) holds almost surely for every  $z \in \mathcal{C}$  and for every  $g \in C(\mathcal{K})$ . Thus (4.23) and (4.24) yield that  $\mu_z$  and  $\tilde{\mu}_z$

satisfy (2.6) (and similarly (2.7)) almost surely for all  $z \in \mathcal{C}$ .

Since  $\mu_z$  and  $\tilde{\mu}_z$  are Stieltjes kernels, one can easily prove that  $z \mapsto \int \frac{g}{D(\tilde{\mu}_z, \mu_z)} dH$  is analytic over  $\mathbb{C}^+$ . Therefore, by (2.6), the two analytic functions  $z \mapsto \int g d\mu_z$  and  $z \mapsto \int \frac{g}{D(\tilde{\mu}_z, \mu_z)} dH$  coincide almost surely over  $\mathcal{C}$  which contains a limit point. They must be equal almost surely over  $\mathbb{C}^+$ . Therefore  $\mu_z$  and  $\tilde{\mu}_z$  satisfy (2.6) (and similarly (2.7)) almost surely for all  $z \in \mathbb{C}^+$ .

Since  $\mu$  and  $\tilde{\mu}$  are Stieltjes kernels satisfying almost surely (2.6) and (2.7), they must be almost surely equal to the unique pair of solutions  $(\pi, \tilde{\pi})$  by the unicity argument developed in Step 1 (cf. Section 4.2). In particular,  $\mu$  and  $\tilde{\mu}$  are almost surely independent of  $\omega \in \Omega$ . Thus for every subsequence  $M = M(n)$ ,

$$\text{a.s., } \forall z \in \mathbb{C}^+, \quad L_z^M \xrightarrow[n \rightarrow \infty]{w} \pi_z \quad \text{and} \quad \tilde{L}_z^M \xrightarrow[n \rightarrow \infty]{w} \tilde{\pi}_z.$$

Therefore, the convergence remains true for the whole sequences  $L_z^n$  and  $\tilde{L}_z^n$ . Theorem 2.3 is proved.

## REFERENCES

- [1] Z. D. Bai. Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica*, 9(3):611–677, 1999.
- [2] Z. D. Bai and J. W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.*, 26(1):316–345, 1998.
- [3] A. Boutet de Monvel, A. Khorunzhy, and V. Vasilchuk. Limiting eigenvalue distribution of random matrices with correlated entries. *Markov Process. Related Fields*, 2(4):607–636, 1996.
- [4] C.N. Chuah, J.M. Kahn, D.N.C. Tse, and R.A. Valenzuela. Capacity scaling in mimo wireless systems under correlated fading. *IEEE Trans. Inf. Theo.*, 48(3):637–650, 2002.
- [5] M. Debbah, W. Hachem, P. Loubaton, and M. de Courville. MMSE analysis of certain large isometric random precoded systems. *IEEE Trans. Inform. Theory*, 49(5):1293–1311, 2003.
- [6] R. B. Dozier and Silverstein J. W. On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices. preprint 2004.
- [7] V. L. Girko. *Theory of random determinants*, volume 45 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1990.
- [8] V. L. Girko. *Theory of stochastic canonical equations. Vol. I*, volume 535 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2001.
- [9] A. Goldsmith, S.A. Jafar, N. Jindal, and S. Vishwanath. Capacity limits of mimo channels. *IEEE J. Sel. Areas in Comm.*, 21(5), 2003.
- [10] W. Hachem, P. Loubaton, and J. Najim. Deterministic equivalents for certain functionals of large random matrices. available at <http://arxiv.org/math.PR/0507172>, 2005.
- [11] W. Hachem, P. Loubaton, and J. Najim. The empirical eigenvalue distribution of a gram matrix: From independence to stationarity. available at <http://arxiv.org/abs/math.PR/0502535>, 2005.
- [12] T. Kailath, Sayed A.H., and Hassibi B. *Linear Estimation*. Prentice Hall, NJ, 2000.
- [13] A. Khorunzhy, B. Khoruzhenko, and L. Pastur. Asymptotic properties of large random matrices with independent entries. *J. Math. Phys.*, 37(10):5033–5060, 1996.
- [14] L. Li, A.M. Tulino, and S. Verdu. Design of reduced-rank mmse multiuser detectors using random matrix methods. *IEEE Trans. Inform. Theory*, 50(6):986–1008, june 2004.
- [15] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.
- [16] D. Shlyakhtenko. Random Gaussian band matrices and freeness with amalgamation. *Internat. Math. Res. Notices*, (20):1013–1025, 1996.
- [17] J. W. Silverstein. Strong convergence of the empirical distribution of eigenvalues of large-dimensional random matrices. *J. Multivariate Anal.*, 55(2):331–339, 1995.
- [18] 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(2):175–192, 1995.
- [19] J. W. Silverstein and P.L. Combettes. Signal detection via spectral theory of large dimensional random matrices. *IEEE Trans. Signal Proces.*, 40(8):2100–2105, 1992.



- [20] D. Tse and S. Hanly. Linear multiuser receivers: effective interference, effective bandwidth and user capacity. *IEEE Trans. Inform. Theory*, 45(2):641–657, 1999.
- [21] D. Tse and O. Zeitouni. Linear multiuser receivers in random environments. *IEEE Trans. Inform. Theory*, 46(1):171–188, 2000.
- [22] A. Tulino and S. Verdú. *Random matrix theory and wireless communications*, volume 1 of *Foundations and Trends in Communications and Information Theory*. Now Publishers, Delft, 2004.
- [23] Y. Q. Yin. Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.*, 20(1):50–68, 1986.

WALID HACHEM,  
Supélec (Ecole Supérieure d’Electricité)  
Plateau de Moulon, 3 rue Joliot-Curie  
91192 Gif Sur Yvette Cedex, France.  
e-mail: walid.hachem@supelec.fr

PHILIPPE LOUBATON,  
IGM LabInfo, UMR 8049, Institut Gaspard Monge,  
Université de Marne La Vallée, France.  
5, Bd Descartes, Champs sur Marne,  
77454 Marne La Vallée Cedex 2, France.  
e-mail: loubaton@univ-mlv.fr

JAMAL NAJIM,  
CNRS, Télécom Paris  
46, rue Barrault, 75013 Paris, France.  
e-mail: najim@tsi.enst.fr