ON THE BEHAVIOUR OF LARGE EMPIRICAL AUTOCOVARIANCE MATRICES BETWEEN THE PAST AND THE FUTURE

P. Loubaton

Laboratoire d'Informatique Gaspard Monge, UMR 8049, Université Paris-Est Marne la Vallée
5 Bd. Descartes, Cite Descartes, 77454 Marne la Vallée Cedex 2, France

D. Tieplova

Laboratoire d'Informatique Gaspard Monge, UMR 8049, Université Paris-Est Marne la Vallée and Dept. of Theoretical Physics, Institute for Low Temperature Physics and Engineering
47 Nauky Ave., Kharkiv, 61103, Ukraine

The asymptotic behaviour of the distribution of the squared singular values of the sample autocovariance matrix between the past and the future of a high-dimensional complex Gaussian uncorrelated sequence is studied. Using Gaussian tools, it is established that the distribution behaves as a deterministic probability measure whose support $\mathcal{S}$ is characterized. It is also established that the squared singular values are almost surely located in a neighbourhood of $\mathcal{S}$.

Keywords: Large Gaussian random matrices; Autocovariance matrices; Stieltjes transform

1. Introduction.

1.1. *The addressed problem and the results.*

In this paper, we consider a sequence of integer $(M(N))_{N\geq 1}$, and positive definite $M(N)\times M(N)$ hermitian matrices $(R_N)_{N\geq 1}$. For each N, we define an independent identically distributed sequence $(y_n)_{n\geq 1}$ (depending on N) of zero mean complex Gaussian $M(N)$–dimensional random vectors such that $y_n = R_N^{1/2}\xi_n$ where the components of the M–dimensional vector ξ_n are complex Gaussian standard i.i.d. random variables (i.e. their real and imaginary parts are i.i.d. and $\mathcal{N}(0,1/2)$ distributed). If L is a fixed integer, we consider the 2 block-Hankel $ML\times N$ matrices $W_{p,N}$ and $W_{f,N}$ defined by

$$W_{p,N} = \frac{1}{\sqrt{N}}\, Y_{p,N} = \frac{1}{\sqrt{N}} \begin{pmatrix} y_1 & y_2 & \cdots & y_{N-1} & y_N \\ y_2 & y_3 & \cdots & y_N & y_{N+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_L & y_{L+1} & \cdots & y_{N+L-2} & y_{N+L-1} \end{pmatrix} \tag{1.1}$$

and

$$W_{f,N} = \frac{1}{\sqrt{N}} Y_{f,N} = \frac{1}{\sqrt{N}} \begin{pmatrix} y_{L+1} & y_{L+2} & \cdots & y_{N-1+L} & y_{N+L} \\ y_{L+2} & y_{L+3} & \cdots & y_{N+L} & y_{N+L+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{2L} & y_{2L+1} & \cdots & y_{N+2L-2} & y_{N+2L-1} \end{pmatrix} \tag{1.2}$$

If $(\hat{\lambda}_{k,N})_{k=1,\ldots,ML}$ are the eigenvalues of the $ML \times ML$ matrix $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$, we study the behaviour of the empirical eigenvalue distribution

$$\hat{\nu}_N = \frac{1}{ML} \sum_{k=1}^{ML} \delta_{\hat{\lambda}_{k,N}}$$

of $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ in the asymptotic regime where M and N converge towards $+\infty$ in such a way that

$$c_N = \frac{ML}{N} \to c_*, \, c_* > 0.$$

Using Gaussian tools, we evaluate the asymptotic behaviour of the resolvent $Q_N(z) = (W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^* - zI)^{-1}$, and establish that the sequence $(\hat{\nu}_N)_{N \geq 1}$ has the same almost sure asymptotic behaviour than a sequence $(\nu_N)_{N \geq 1}$ of deterministic probability measures. In the following, ν_N will be referred to as the deterministic equivalent of $\hat{\nu}_N$. We evaluate the Stieltjes transform of ν_N, characterize its support, study the properties of its density, and eventually establish that almost surely, for N large enough, all the eigenvalues of $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ are located in a neighbourhood of the support of ν_N.

1.2. *Motivation*

Matrix $W_{f,N} W_{p,N}^* = \frac{Y_{f,N} Y_{p,N}^*}{N}$ represents the traditional empirical estimate of the autocovariance matrix $R_{f|p,y}^L$ between the past and the future of y defined as

$$R_{f|p,y}^L = \mathbb{E} \left[\begin{pmatrix} y_{n+L} \\ y_{n+L+1} \\ \vdots \\ y_{n+2L-1} \end{pmatrix} \left(y_n^*, y_{n+1}^*, \ldots, y_{n+L-1}^* \right) \right].$$

This matrix plays a key role in statistical inference problems related to multivariate time series with rational spectrum. In order to explain this, we consider a $M-$dimensional multivariate time series $(v_n)_{n \in \mathbb{Z}}$ generated as

$$v_n = u_n + y_n, \tag{1.3}$$

where $(y_n)_{n\in\mathbb{Z}}$ is as above a Gaussian "noise" term such that $\mathbb{E}(y_{n+k}y_k^*) = R\delta_n$ for some unknown positive definite matrix R, and where $(u_n)_{n\in\mathbb{Z}}$ is a "useful" non observable Gaussian signal with rational spectrum. u_n can thus be represented as

$$x_{n+1} = Ax_n + B\omega_n, \ u_n = Cx_n + D\omega_n, \tag{1.4}$$

where $(\omega_n)_{n\in\mathbb{Z}}$ is a $K \leq M$–dimensional white noise sequence (i.e. $\mathbb{E}(\omega_{n+k}\omega_k^*) = I_K\, \delta_n$), A is a deterministic $P \times P$ matrix whose spectral radius $\rho(A)$ is strictly less than 1, and where B, C, D are deterministic matrices. The P-dimensional Markovian sequence $(x_n)_{n\in\mathbb{Z}}$ is called the state-space sequence associated to (1.4). The state space representation (1.4) is said to be minimal if the dimension P of the state space sequence is minimal. Given the autocovariance sequence $(R_{u,n})_{n\in\mathbb{Z}}$ of u (i.e. $R_{u,n} = \mathbb{E}(u_{k+n}u_k^*)$ for each n), the so-called stochastic realization problem of $(u_n)_{n\in\mathbb{Z}}$ consists in characterizing all the minimal state space representations (1.4) of u, or equivalently in identifying all the minimum Mac-Millan degree[a] matrix-valued function $\Phi(z) = D + C(zI - A)^{-1}B$ such that $\rho(A) < 1$ and

$$S_u(e^{2i\pi f}) = \sum_{n\in\mathbb{Z}} R_{u,n}e^{-2i\pi nf} = \Phi(e^{2i\pi f})\Phi(e^{2i\pi f})^* \tag{1.5}$$

for each f. Such a function Φ is called a minimal degree causal spectral factorization of S_u. We refer the reader to [24] or [37] for more details.

The identification of P and of matrices C and A is based on the observation that the autocovariance sequence of u can be represented as

$$R_{u,n} = \mathbb{E}(u_{n+k}u_k^*) = CA^{n-1}G \tag{1.6}$$

for each $n \geq 1$, where the 3 matrices (A, C, G) are unique up to similarity transforms, thus showing that the matrices C and A associated to a minimal realization are uniquely defined (up to a similarity). Moreover, the autocovariance matrix $R^L_{f|p,u}$ between the past and the future of u can be written as

$$R^{(L)}_{f|p,u} = \mathcal{O}^{(L)}\, \mathcal{C}^{(L)}, \tag{1.7}$$

where matrix $\mathcal{O}^{(L)}$ is the $ML \times P$ "observability" matrix

$$\mathcal{O}^{(L)} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{L-1} \end{pmatrix} \tag{1.8}$$

and matrix $\mathcal{C}^{(L)}$ is the $P \times ML$ "controllability" matrix

$$\mathcal{C}^{(L)} = \left(A^{L-1}G, A^{L-2}G, \ldots, G\right). \tag{1.9}$$

[a]The Mac-Millan degree of a rational matrix-valued function Φ is defined as the minimal dimension of the matrices A for which $\Phi(z)$ can be represented as $D + C(zI - A)^{-1}B$

For each $L \geq P$, the rank of $R^{(L)}_{f|p,u}$ remains equal to P, and each minimal rank factorization of $R^{(L)}_{f|p,u}$ can be written as (1.7) for some particular triple (A, C, G). In particular, if $R^{(L)}_{f|p,u} = \Theta\Gamma\tilde{\Theta}^*$ is the singular value decomposition of $R^{(L)}_{f|p,u}$, matrix $\Theta\Gamma^{1/2}$ coincides with the observability matrix $\mathcal{O}^{(L)}$ of a pair (C, A). C and A are immediately obtained from the knowledge of the structured matrix $\mathcal{O}^{(L)}$. This discussion shows that the evaluation of P, C and A from the autocovariance sequence of u is an easy problem. We mention that, while C and A are essentially unique, there exist in general more than one pair (B, D) for which (1.4) holds because the minimal degree spectral factorization problem (1.5) has more than 1 solution. We refer the reader to [24] or [37].

We notice that as $(y_n)_{n\in\mathbb{Z}}$ in (1.3) is an uncorrelated sequence, it holds that $R_{v,n} = \mathbb{E}(v_{n+k}v_k^*)$ coincides with $R_{u,n}$ for each $n \geq 1$. Therefore, P and matrices C and A can still be identified from the autocovariance sequence of the noisy version v of u. In practice, however, the exact autocovariance sequence $(R_{v,n})_{n\geq 1}$ is in general unknown, and it is necessary to estimate P and (C, A) from the sole knowledge of N samples $v_1 = u_1 + y_1, v_2 = u_2 + y_2, \ldots, v_N = u_N + y_N$. For this, P is first estimated as the number of significant singular values of the empirical estimate $\hat{R}^L_{f|p,v}$ of the true matrix $R^L_{f|p,v} = R^L_{f|p,u}$ defined by

$$\hat{R}^L_{f|p,v} = \frac{V_{f,N}V^*_{p,N}}{N},$$

where $V_{f,N}$ and $V_{p,N}$ are defined in the same way than $Y_{f,N}$ and $Y_{p,N}$. If $(\hat{\gamma}_p)_{p=1,\ldots,P}$ and $\hat{\Theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_P)$ are the P largest singular values and corresponding left singular vectors of matrix $\hat{R}^{(L)}_{f|p,v}$, and if $\hat{\Gamma}$ is the $P \times P$ diagonal matrix with diagonal entries $(\hat{\gamma}_p)_{p=1,\ldots,P}$, the $ML \times P$ matrix $\hat{\mathcal{O}}^{(L)} = \hat{\Theta}\hat{\Gamma}^{1/2}$ is an estimator of an observability matrix $\mathcal{O}^{(L)}$. $\hat{\mathcal{O}}^{(L)}$ has not necessarily the structure of an observability matrix, but C and A can be estimated respectively by the top $M \times P$ block $\hat{C}$ of $\hat{\mathcal{O}}^{(L)}$ and by the argument $\hat{A}$ of the minimum of the quadratic fuction

$$\mathrm{Tr}\left(\left(\hat{\mathcal{O}}^{(L)}_{\mathrm{down}}A - \hat{\mathcal{O}}^{(L)}_{\mathrm{up}}\right)\left(\hat{\mathcal{O}}^{(L)}_{\mathrm{down}}A - \hat{\mathcal{O}}^{(L)}_{\mathrm{up}}\right)^*\right),$$

where the operator "down" (resp. "up") suppresses the last (resp. the first) M rows from $ML \times P$ matrix $\hat{\mathcal{O}}^{(L)}$. This approach provides consistent estimates of P, C, A when $N \to +\infty$ while M, L and P are fixed parameters. We refer the reader to [11] for a detailed analysis of this statistical inference scheme.

If M is large and that the sample size N cannot be arbitrarily larger than M, the ratio ML/N may not be small enough to make reliable the above statistical analysis. It is thus relevant to study the behaviour of the above estimators in asymptotic regimes where M and N both converge towards $+\infty$ in such a way that $\frac{ML}{N}$ converges towards a non zero constant. In this context, the truncated singular value decomposition of $\hat{R}^{(L)}_{f|p,v}$ does not provide a consistent estimate of an observability matrix $\mathcal{O}^{(L)}$, and it appears relevant to study the largest singular values and corresponding singular vectors of $\hat{R}^{(L)}_{f|p,v}$ when M and N both converge towards $+\infty$, and to precise how they are related to an observability matrix $\mathcal{O}^{(L)}$.

Without formulating specific assumptions on u, this problem seems very complicated. In the past, a number of works addressed high-dimensional inference schemes based on the eigenvalues and eigenvectors of the empirical covariance matrix of the observation (see e.g. [30], [28], [31], [17], [38], [39], [12], [36]) when the useful signal lives in a low-dimensional deterministic subspace. Using results related to spiked large random matrix models (see e.g. [4] [5], [33]), based on perturbation technics, a number of important statistical problems could be addressed using large random matrix theory technics. Our ambition is to follow the same kind of approach to address the estimation problem of P, A, C when u satisfies some low rank assumptions. The first part of this program is to study the asymptotic behaviour of the singular values of the empirical autocovariance matrix in the absence of signal $W_{f,N}W^*_{p,N} = \frac{Y_{f,N}Y^*_{p,N}}{N}$. As the singular values of $W_{f,N}W^*_{p,N}$ are the square roots of the eigenvalues of $W_{f,N}W^*_{p,N}W_{p,N}W^*_{f,N}$, this is precisely the topic of the present paper. Using the obtained results, it should be possible to use a perturbation approach in order to evaluate the behaviour of the largest singular values and corresponding left singular vectors in the presence of a useful signal, and to deduce from this some improved performance schemes for estimating P, C, A.

1.3. *On the literature.*

The large sample behaviour of high-dimensional autocovariance matrices was comparatively less studied than the high-dimensional covariance matrices. We first mention [20] which studied the asymptotic behaviour of the eigenvalue distribution of the hermitian matrix $\hat{R}_\tau + \hat{R}^*_\tau$ where $\hat{R}_\tau$ is defined as $\hat{R}_\tau = \frac{1}{N}\sum_{n=1}^{N} x_{n+\tau}x^*_n$ where $(x_n)_{n\in\mathbb{Z}}$ represents a M dimensional non Gaussian i.i.d. sequence, the components of each vector x_n being moreover i.i.d. with zero means and unit variances. In particular, $\mathbb{E}(x_n x^*_n) = I$. It is proved that the empirical eigenvalue distribution of $\hat{R}_\tau + \hat{R}^*_\tau$ converges towards a limit distribution independent from $\tau \geq 1$. Using finite rank perturbation technics of the resolvent of the matrix under consideration, the Stieltjes transform of this distribution was shown to satisfy a polynomial degree 3 equation. Solving this equation led to an explicit expression of the probability density of the limit distribution. [25] extended these results to the case where $(x_n)_{n\in\mathbb{Z}}$ is a non Gaussian linear process $x_n = \sum_{l=0}^{+\infty} A_l z_{n-l}$ where $(z_n)_{n\in\mathbb{Z}}$ is i.i.d., and where matrices $(A_l)_{l\geq 0}$ are simultaneously diagonalizable. The limit eigenvalue distribution was characterized through its Stieltjes transform that is obtained by integration of a certain kernel, itself solution of an integral equation. The proof was based on the observation that in the Gaussian case, the correlated vectors $(x_n)_{n\in\mathbb{Z}}$ can be replaced by independent ones using a classical frequency domain decorrelation procedure. The results were generalized in the non Gaussian case using the generalized Lindeberg principle. We also mention [2] (see also the book [3]) where the existence of a limit distribution of any symmetric polynomial of $(\hat{R}_\tau, \hat{R}_\tau^*)_{\tau \in T}$ for some finite set T was proved using the moment method when x is a linear non Gaussian process. [22] studied the asymptotic behaviour of matrix $\hat{R}_\tau \hat{R}_\tau^*$ when $(x_n)_{n \in \mathbb{Z}}$ represents a M dimensional non Gaussian i.i.d. sequence, the components of each vector x_n being moreover i.i.d. Using finite rank perturbation technics, it was shown that the empirical eigenvalue distribution converges towards a limit distribution whose Stieltjes transform is solution of a degree 3 polynomial equation. As in [20], this allowed to obtain the expression of the corresponding probability density function. Using combinatorial technics, [22] also established that almost surely, for large enough dimensions, all the eigenvalues of $\hat{R}_\tau \hat{R}_\tau^*$ are located in a neighbourhood of the support of the limit eigenvalue distribution. We finally mention that [23] used the results in [22] in order to study the largest eigenvalues and corresponding eigenvectors of $\hat{R}_\tau \hat{R}_\tau^*$ when the observation contains a certain spiked useful signal that is more specific than the signals $(u_n)_{n \in \mathbb{Z}}$ that motivated the present paper.

We now compare the results of the present paper with the content of the above previous works. We first study a matrix that is more general than $\hat{R}_\tau \hat{R}_\tau^*$. While we do not consider linear processes here, we do not assume that the covariance matrix of the i.i.d. sequence $(y_n)_{n \in \mathbb{Z}}$ is reduced to I as in [22]. This in particular implies that the Stieltjes transform of the deterministic equivalent ν_N of the empirical eigenvalue distribution $\hat{\nu}_N$ of $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ cannot be evaluated in closed from. Therefore, a dedicated analysis of the support and of the properties of ν_N is provided here. We also mention that in contrast with the above papers, we characterize the asymptotic behaviour of the resolvent of matrix $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ while the mentionned previous works only studied the normalized trace of the resolvent of the matrices under consideration. Studying the full resolvent matrix is necessary to address the case where a useful spiked signal u is added to the noise y. We notice that the above papers addressed the non Gaussian case while we consider the case where y is a complex Gaussian i.i.d. sequence. This situation is of course simpler in that various Gaussian tools are available, but appears to be relevant because in the context of the present paper, y is indeed supposed to represent some additive noise, which, in a number of contexts, is Gaussian.

We finally mention that some of the results of this paper may be obtained by adapting general recent results devoted to the study of the spectrum of hermitian polynomials of GUE matrices and deterministic matrices (see [6] and [27]). If we denote by Z_N the $M \times (N + 2L - 1)$ matrix $Z_N = (y_1, \ldots, y_{N+2L-1})$, then Z_N can be written as $Z_N = R_N^{1/2} X_N$ where the entries of X_N are i.i.d. complex Gaussian standard variables. Each $M \times M$ block $\Sigma_{N,k,l}$ $(1 \leq k, l \leq L)$ of $\Sigma_N = W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ is clearly a polynomial of X_N, X_N^* and various $M \times M$

and $M \times (N+2L-1)$ deterministic matrices. Assume that $M < N+2L-1$. In order to be back to a polynomial of GUE matrices, it is possible to consider the $L(N+2L-1) \times L(N+2L-1)$ matrix $\tilde{\Sigma}_N$ whose $(N+2L-1) \times (N+2L-1)$ blocks are defined by

$$\tilde{\Sigma}_{N,k,l} = \begin{pmatrix} \Sigma_{N,k,l} & 0 \\ 0 & 0 \end{pmatrix}.$$

It is clear that apart 0, the eigenvalues of $\tilde{\Sigma}_N$ coincide with those of Σ_N. If $\tilde{X}_N$ is any $(N+2L-1) \times (N+2L-1)$ matrix with i.i.d. complex Gaussian standard entries whose M first rows coincide with X_N, then, it is easily seen that each block of $\tilde{\Sigma}_N$ coincides with a hermitian polynomial of $\tilde{X}_N, \tilde{X}_N^*$ and deterministic $(N+2L-1) \times (N+2L-1)$ matrices such as

$$\tilde{R}_N = \begin{pmatrix} R_N & 0 \\ 0 & 0 \end{pmatrix}.$$

Expressing $\tilde{X}_N$ as the sum of its hermitian and anti-hermitian parts, we are back to study the behaviour of the eigenvalues of a matrix whose blocks are hermitian polynomials of 2 independent GUE matrices and of $(N+2L-1) \times (N+2L-1)$ deterministics matrices. Extending Proposition 2.2 and Theorem 1.1 in [6] to block matrices (as in Corollary 2.3 in [27]) would lead to the conclusion that $\hat{\nu}_N$ has a deterministic equivalent ν_N and that the eigenvalues of $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ are located in the neighbourhood of the support of ν_N. While this last consequence would avoid the use of the specific approach used in section 8 of the present paper, the existence of ν_N is not a sufficient information. ν_N should of course be characterized through its Stieltjes transform, and we believe that the adaptation of Proposition 2.2 and Theorem 1.1 in [6] is not the most efficient approach.

1.4. *Overview of the paper.*

As the entries of matrices $W_{p,N}$ and $W_{f,N}$ are correlated, approaches based on finite rank perturbation of the resolvent $Q_N(z)$ of matrix $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$, usually used when independence assumptions hold, are not the most efficient in our context. We rather propose to use Gaussian tools, i.e. integration by parts formula in conjunction with the Poincaré-Nash inequality (see e.g. [32]), because they are robust to correlation of the matrix entries. Moreover, as the entries of $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ are biquadratic functions of $y_1, \ldots, y_{N+2L-1}$, we rather use the well-known linearization trick that consists in studying the resolvent $\mathbf{Q}_N(z)$ of the $2ML \times 2ML$ hermitized version

$$\begin{pmatrix} 0 & W_{f,N} W_{p,N}^* \\ W_{p,N} W_{f,N}^* & 0 \end{pmatrix}$$

of matrix $W_{f,N} W_{p,N}^*$. As is well known, the first $ML \times ML$ diagonal block of $\mathbf{Q}_N(z)$ coincides with $z Q_N(z^2)$. Therefore, we characterize the asymptotic behaviour of

$\mathbf{Q}_N(z)$, and deduce from this the results concerning $Q_N(z)$. The hermitized version is this time a quadratic function of $y_1, \ldots, y_{N+2L-1}$, and the Gaussian calculus that is needed in order to study $\mathbf{Q}_N(z)$ appears much simpler than if $Q_N(z)$ was evaluated directly.

In section 3, we evaluate the variance of useful functionals for $\mathbf{Q}_N(z)$ using the Poincaré-Nash inequality. In section 4, we establish some useful lemmas related to certain Stieltjes transforms. In section 5, we use the integration by parts formula to establish that $\mathbb{E}(\mathbf{Q}_N(z))$ behaves as $I_{2L} \otimes \mathbf{S}_N(z)$ where $\mathbf{S}_N(z)$ is defined by

$$\mathbf{S}_N(z) = -\left(\frac{c_N \boldsymbol{\alpha}_N(z)}{1 - c_N^2 \boldsymbol{\alpha}_N(z)^2} R_N + z I_M\right)^{-1},$$

where $\boldsymbol{\alpha}_N(z)$ is defined by $\boldsymbol{\alpha}_N(z) = \frac{1}{ML}\mathrm{Tr}\mathbb{E}(\mathbf{Q}_{N,pp}(z))(I_L \otimes R_N)$ where $\mathbf{Q}_{N,pp}(z)$ represents the first $ML \times ML$ diagonal block of $\mathbf{Q}_N(z)$. As usual, if A and B are two $n_A \times m_A$ and $n_B \times m_B$ matrices, $A \otimes B$ represents the $n_A n_B \times m_A m_B$ matrix whose $n_B \times m_B$ blocks are the matrices $A_{i,j}B$ for $i = 1, \ldots, n_A$ and $j = 1, \ldots, m_A$. We deduce from this that

$$\mathbb{E}(Q_N(z)) = I_L \otimes S_N(z) + \Delta_N(z),$$

where $S_N(z) = -\left(zI_M + \frac{c_N z \alpha_N(z)}{1 - c_N^2 \alpha_N(z)^2} R_N\right)^{-1}$, $\alpha_N(z) = \frac{1}{ML}\mathrm{Tr}\mathbb{E}(Q_N(z))(I_L \otimes R_N)$, and where $\Delta_N(z)$ is an error term such that

$$\left|\frac{1}{ML}\mathrm{Tr}\,\Delta_N(z)\right| \leq \frac{1}{N^2} P_1(|z|) P_2\left(\frac{1}{\mathrm{Im}(z)}\right)$$

for each $z \in \mathbb{C}^+$, where P_1 and P_2 are 2 polynomials whose degrees and coefficients do not depend on N. Using this, we prove in section 6 that for each $z \in \mathbb{C}^+$,

$$\frac{1}{ML}\mathrm{Tr}\mathbb{E}\left[Q_N(z) - I_L \otimes T_N(z)\right] F_N \to 0,$$

where $(F_N)_{N \geq 1}$ is any deterministic sequence of matrices such that $\sup_N \|F_N\| < +\infty$, and where $T_N(z)$ is defined by

$$T_N(z) = -\left(zI_M + \frac{z c_N t_N(z)}{1 - z c_N^2 t_N^2(z)} R_N\right)^{-1},$$

$t_N(z)$ being the unique solution of the equation

$$t_N(z) = \frac{1}{M}\mathrm{Tr} R_N \left(-zI_M - \frac{z c_N t_N(z)}{1 - z c_N^2 t_N^2(z)} R_N\right)^{-1}$$

such that $t_N(z)$ and $z t_N(z)$ belong to $\mathbb{C}^+$ when $z \in \mathbb{C}^+$. $t_N(z)$ and $T_N(z)$ are shown to coincide with the Stieltjes transforms of a scalar measure μ_N and of a $M \times M$ positive matrix valued measure ν_N^T respectively (see Section 4 for a formal definition of a $M \times M$ positive matrix valued measure). Recalling that $\hat{\nu}_N$ denotes the empirical eigenvalue distribution of $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$, it is proved that the measure ν_N defined by $\nu_N = \frac{1}{M}\mathrm{Tr}(\nu_N^T)$ is a probability measure, and that $\hat{\nu}_N - \nu_N \to 0$ weakly almost surely. ν_N is referred to as the deterministic equivalent of $\hat{\nu}_N$. In section 7, we study the properties and the support of ν_N, or equivalently of μ_N because the 2 measures are absolutely continuous one with respect to each other. For this, we study the behaviour of $t_N(z)$ when z converges towards the real axis. For each $x > 0$, the limit of $t_N(z)$ when $z \in \mathbb{C}^+$ converges towards x exists and is finite. If $c_N \leq 1$, we deduce from this that ν_N is absolutely continuous w.r.t. the Lebesgue measure. The corresponding density $g_N(x)$ is real analytic on $\mathbb{R}^+$, and converges towards $+\infty$ when $x \to 0, x > 0$. If $c_N < 1$, it holds that $g_N(x) = \mathcal{O}(\frac{1}{\sqrt{x}})$ while $g_N(x) = \mathcal{O}(\frac{1}{x^{2/3}})$ if $c_N = 1$. If $c_N > 1$, ν_N contains a Dirac mass at 0 with weight $1 - \frac{1}{c_N}$ and an absolutely continuous component. In order to analyse the support of μ_N and ν_N, we establish that the function $w_N(z)$ defined by

$$w_N(z) = z c_N t_N(z) - \frac{1}{c_N t_N(z)}$$

is solution of the equation $\phi_N(w_N(z)) = z$ for each $z \in \mathbb{C} - \mathbb{R}^+$ where $\phi_N(w)$ is the function defined by

$$\phi_N(w) = c_N w^2 \frac{1}{M}\mathrm{Tr} R_N \left(R_N - wI\right)^{-1} \left(c_N \frac{1}{M}\mathrm{Tr} R_N \left(R_N - wI\right)^{-1} - 1\right).$$

Moreover, if we define $t_N(x)$ and $w_N(x)$ for $x > 0$ by the limit of $t_N(z)$ and $w_N(z)$ when $z \to x, z \in \mathbb{C}^+$, the equality $\phi_N(w_N(z)) = z$ is also valid on $\mathbb{R}^+$. We establish that if x is outside the support of μ_N, then, it holds that

$$\phi_N(w_N(x)) = x, \ \phi^{'}(w_N(x)) > 0, \ w_N(x) \frac{1}{M}\mathrm{Tr} R_N \left(R_N - w_N(x) I\right)^{-1} < 0.$$

This property allows to prove that apart $\{0\}$ when $c_N > 1$, the support of μ_N is a union of intervals whose end points are the extrema of ϕ_N whose arguments verify $w \frac{1}{M}\mathrm{Tr} R_N \left(R_N - wI\right)^{-1} < 0$. A sufficient condition on the eigenvalues of R_N ensuring that the support of μ_N is reduced to a single interval is formulated. Using the Haagerup-Thorbjornsen approach ([15]), it is moreover proved in section 8 that for each N large enough, all the eigenvalues of $W_{f,N} \, W_{p,N}^* \, W_{p,N} \, W_{f,N}^*$ lie in a neighbourhood of the support of the deterministic equivalent ν_N. The above results do not imply that $\hat{\nu}_N$ converges towards a limit distribution. In order to obtain this kind of result, some extra assumptions have to be formulated, such as the existence of a limit empirical eigenvalue distribution for R_N when $N \to +\infty$. If the relevant conditions are met, ν_N, and therefore $\hat{\nu}_N$, will converge towards a limit distribution whose Stieltjes transform can be obtained by replacing in the above results the empirical eigenvalue of R_N by its limit. We do not present the corresponding results here because we believe that results that characterize the behaviour of ν_N for each N large enough are more informative than the convergence towards a limit.

In section 9, we finally indicate that the use of free probability tools is an alternative approach to characterize the asymptotic behaviour of $\hat{\nu}_N$. The results of section 9 are based on the following observations:

- Up to the zero eigenvalue, the eigenvalues of $W_{f,N}W_{p,N}^*W_{p,N}W_{f,N}^*$ coincide with the eigenvalues of $W_{f,N}^*W_{f,N}W_{p,N}^*W_{p,N}$
- While the matrices $W_{f,N}^*W_{f,N}$ and $W_{p,N}^*W_{p,N}$ do not satisfy the conditions of the usual asymptotic freeness results, it turns out that they are almost surely asymptotically free. Therefore, the eigenvalue distribution of $W_{f,N}^*W_{f,N}W_{p,N}^*W_{p,N}$ converges towards the free multiplicative convolution product of the limit distributions of $W_{f,N}^*W_{f,N}$ and $W_{p,N}^*W_{p,N}$. These two distributions appear to coincide both with the limit distribution of the well known random matrix model $\frac{1}{N}X_N^*(I_L \otimes R_N)X_N$ where X_N is a $ML \times N$ complex Gaussian random matrix with standard i.i.d. entries.

The asymptotic freeness of $W_{f,N}^*W_{f,N}$ and $W_{p,N}^*W_{p,N}$ appear to be a consequence of Lemma 6 in [14]. While this approach seems to be simpler than the use of the Gaussian tools proposed in the present paper, we mention that the above free probability theory arguments do not allow to study the asymptotic behaviour of the resolvent of $W_{f,N}W_{p,N}^*W_{p,N}W_{f,N}^*$. We recall that in order to evaluate the largest eigenvalues and corresponding eigenvectors of $W_{f,N}W_{p,N}^*W_{p,N}W_{f,N}^*$ in the presence of a useful signal, the asymptotic behaviour of the full resolvent in the absence of signal has to be available.

2. Some notations, assumptions, and useful results.

In the following, it is assumed that L is a fixed integer, and that M and N converge towards $+\infty$ in such a way that

$$c_N = \frac{ML}{N} \to c_*,\, c_* > 0. \tag{2.1}$$

This regime will be referred to as $N \to +\infty$ in the following. In the regime (2.1), M should be interpreted as an integer $M = M(N)$ depending on N. The various matrices we have introduced above thus depend on N and will be denoted $R_N, Y_{f,N}, Y_{p,N}, \ldots$. In order to simplify the notations, the dependency w.r.t. N will sometimes be omitted.

We recall that the resolvent $Q_N(z)$ of $W_{f,N}W_{p,N}^*W_{p,N}W_{f,N}^*$ is defined by

$$Q_N(z) = \left(W_{f,N}W_{p,N}^*W_{p,N}W_{f,N}^* - zI\right)^{-1}. \tag{2.2}$$

As the direct study of $Q_N(z)$ is not obvious, we rather introduce the resolvent $\mathbf{Q}_N(z)$ of the $2ML \times 2ML$ block matrix

$$\mathbf{M}_N = \begin{pmatrix} 0 & W_{f,N}W_{p,N}^* \\ W_{p,N}W_{f,N}^* & 0 \end{pmatrix}. \tag{2.3}$$

It is well known that $\mathbf{Q}_N(z)$ can be expressed as

$$\mathbf{Q}_N(z) = \begin{pmatrix} zQ_N(z^2) & Q_N(z^2)W_{f,N}W_{p,N}^* \\ W_{p,N}W_{f,N}^*Q_N(z^2) & z\tilde{Q}_N(z^2) \end{pmatrix}, \tag{2.4}$$

where $\tilde{Q}_N(z)$ is the resolvent of matrix $W_{p,N}W_{f,N}^*W_{f,N}W_{p,N}^*$. As shown below, it is rather easy to evaluate the asymptotic behaviour of $\mathbf{Q}_N(z)$ using the Poincaré-Nash inequality and the integration by part formula (see Propositions 2.2 and 2.1 below). Formula (2.4) will then provide all the necessary information on $Q_N(z)$.

In the following, every $2ML \times 2ML$ matrix $\mathbf{G}$ will be written as

$$\mathbf{G} = \begin{pmatrix} \mathbf{G_{pp}} & \mathbf{G_{pf}} \\ \mathbf{G_{fp}} & \mathbf{G_{ff}} \end{pmatrix},$$

where the 4 matrices $(\mathbf{G_{ij}})_{i,j\in\{p,f\}}$ are $ML \times ML$. Sometimes, the blocks will be denoted $\mathbf{G}(pp)$, $\mathbf{G}(pf)$, ...

We denote by W_N the $2ML \times N$ matrix defined by

$$W_N = \begin{pmatrix} W_{p,N} \\ W_{f,N} \end{pmatrix}. \tag{2.5}$$

Its elements $(W_{i,j}^m)_{i\le 2L, j\le N, m\le M}$ satisfy

$$\mathbb{E}\{W_{i,j}^m (W_{i',j'}^{m'})^*\} = \frac{1}{N} R_{mm',N}\delta_{i+j,i'+j'},$$

where $W_{i,j}^m$ represents the element which lies on the $(m + M(i-1))$-th line and j-th column for $1 \le m \le M$, $1 \le i \le 2L$ and $1 \le j \le N$. Similarly, $\mathbf{Q}_{i_1 i_2}^{m_1 m_2}$, where $1 \le m_1, m_2 \le M$ and $1 \le i_1, i_2 \le 2L$, represents the entry $(m_1 + M(i_1 - 1)), (m_2 + M(i_2 - 1))$ of $\mathbf{Q}$. For each $j = 1, \ldots, N$, $\{w_j\}_{j=1}^N, \{w_{p,j}\}_{j=1}^N$ and $\{w_{f,j}\}_{j=1}^N$ are the column of matrices W, W_p and W_f respectively. For each $1 \le i \le 2L$ and $1 \le m \le M$, $\mathbf{f}_i^m$ represents the vector of the canonical basis of $\mathbb{C}^{2ML}$ with 1 at the index $m + (i-1)M$ and zeros elsewhere. In order to simplify the notations, we mention that if $i \le L$, vector $\mathbf{f}_i^m$ may also represent the vector of the canonical basis of $\mathbb{C}^{ML}$ with 1 at the index $m + (i-1)M$ and zeros elsewhere. Vector $\mathbf{e}_j$ with $1 \le j \le N$ represents the j –th vector of the canonical basis of $\mathbb{C}^N$. Also for any integer K, J_K is the $K \times K$ "shift" matrix defined by

$$(J_K)_{ij} = \delta_{j-i,1}. \tag{2.6}$$

In order to short the notations, for each integer $l \in \mathbb{Z}$, we define the symbol $\epsilon(l)$ by $\epsilon(l) = l$ if $l \ge 0$ and $\epsilon(l) = *|l|$ if $l \le 0$. Consequently, for each l, matrix $J_K^{\epsilon(l)}$ is equal to J_K^l if $l \ge 0$ and to $J_K^{*|l|}$ if $l \le 0$.

If A is a matrix, then $\|A\|$ and $\|A\|_F$ represent its spectral norm and Frobenius norm respectively. If morever A is a square matrix, $\mathrm{Im}(A)$ is the Hermitian matrix defined by $\mathrm{Im}(A) = \frac{A-A^*}{2i}$. We recall that if A and B are two $n_A \times m_A$ and $n_B \times m_B$ matrices, $A \otimes B$ represents the $n_A n_B \times m_A m_B$ matrix whose $n_B \times m_B$ blocks are the matrices $A_{i,j}B$ for $i = 1, \ldots, n_A$ and $j = 1, \ldots, m_A$.

The sequence of covariance matrices $(R_N)_{N\geq 1}$ of M–dimensional vectors $(y_n)_{n=1,\ldots,N}$ is supposed to verify

$$a\, I \leq R_N \leq b\, I \tag{2.7}$$

for each N, where $a > 0$ and $b > 0$ are two constants. $\lambda_{1,N} \geq \lambda_{2,N} \geq \ldots \geq \lambda_{M,N}$ represent the eigenvalues of R_N arranged in the decreasing order and $f_{1,N}, \ldots, f_{M,N}$ denote the corresponding eigenvectors. Hypothesis (2.7) is obviously equivalent to $\lambda_{M,N} \geq a$ and $\lambda_{1,N} \leq b$ for each N.

The eigenvalues and eigenvectors of matrix $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ are denoted $\hat{\lambda}_{1,N} \geq \ldots \geq \hat{\lambda}_{ML,N}$ and $\hat{f}_{1,N}, \ldots, \hat{f}_{ML,N}$ respectively.

By a nice constant, we mean a positive deterministic constant which does not depend on the dimensions M and N nor of the complex variable z. In the following, κ will represent a generic nice constant whose value may change from one line to the other. A nice polynomial $P(z)$ is a polynomial whose degree and coefficients are nice constants. Finally, if $(\alpha_N)_{N\geq 1}$ is a sequence of positive real numbers and if Ω is a domain of $\mathbb{C}$, we will say that a sequence of functions $(f_N(z))_{N\geq 1}$ verifies $f_N(z) = \mathcal{O}_z(\alpha_N)$ for $z \in \Omega$ if there exists two nice polynomials P_1 and P_2 such that $|f_N(z)| \leq \alpha_N P_1(|z|) P_2(\frac{1}{|\mathrm{Im} z|})$ for each $z \in \Omega$. If $\Omega = \mathbb{C}^+$, we will just write $f_N(z) = \mathcal{O}_z(\alpha_N)$ without mentioning the domain. We notice that if P_1, P_2 and Q_1, Q_2 are nice polynomials, then $P_1(|z|)P_2(\frac{1}{|\mathrm{Im} z|}) + Q_1(|z|)Q_2(\frac{1}{|\mathrm{Im} z|}) \leq (P_1 + Q_1)(|z|)(P_2 + Q_2)(\frac{1}{|\mathrm{Im} z|})$, from which we conclude that if the sequences $(f_{1,N})_{N\geq 1}$ and $(f_{2,N})_{N\geq 1}$ are $\mathcal{O}_z(\alpha_N)$ on Ω, then it also holds $f_{1,N}(z) + f_{2,N}(z) = \mathcal{O}_z(\alpha_N)$ on Ω.

$\mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R})$ represents the set of all $\mathcal{C}^\infty$ real valued compactly supported functions defined on $\mathbb{R}$.

If ξ is a random variable, we denote by ξ° the zero mean random variable defined by

$$\xi^\circ = \xi - \mathbb{E}\xi. \tag{2.8}$$

We finally recall the two Gaussian tools that will be used in the sequel in order to evaluate the asymptotic behaviour of $\mathbf{Q}_N(z)$ and $Q_N(z)$.

Proposition 2.1. ***(Integration by parts formula.)*** *Let $\xi = [\xi_1, \ldots, \xi_K]^T$ be a complex Gaussian random vector such that $\mathbb{E}\{\xi\} = 0$, $\mathbb{E}\{\xi\xi^T\} = 0$ and $\mathbb{E}\{\xi\xi^*\} = \Omega$. If $\Gamma : (\xi) \mapsto \Gamma(\xi, \bar{\xi})$ is a $\mathcal{C}^1$ complex function polynomially bounded together with its derivatives, then*

$$\mathbb{E}\{\xi_i \Gamma(\xi)\} = \sum_{k=1}^{K} \Omega_{ik} \mathbb{E}\left\{ \frac{\partial \Gamma(\xi)}{\partial \bar{\xi}_k} \right\}. \tag{2.9}$$

Proposition 2.2. ***(Poincaré-Nash inequality.)*** *Let $\xi = [\xi_1, \ldots, \xi_K]^T$ be a complex Gaussian random vector such that $\mathbb{E}\{\xi\} = 0$, $\mathbb{E}\{\xi\xi^T\} = 0$ and $\mathbb{E}\{\xi\xi^*\} = \Omega$. If $\Gamma : (\xi) \mapsto \Gamma(\xi, \bar{\xi})$ is a $\mathcal{C}^1$ complex function polynomially bounded together with its derivatives, then, noting $\nabla_\xi \Gamma = [\frac{\partial \Gamma}{\partial \xi_1}, \ldots, \frac{\partial \Gamma}{\partial \xi_K}]^T$ and $\nabla_{\bar{\xi}} \Gamma = [\frac{\partial \Gamma}{\partial \bar{\xi}_1}, \ldots, \frac{\partial \Gamma}{\partial \bar{\xi}_K}]^T$*

$$\mathbf{Var}\{\Gamma(\xi)\} \leq \mathbb{E}\left\{\nabla_\xi \Gamma(\xi)^T \Omega \overline{\nabla_\xi \Gamma(\xi)}\right\} + \mathbb{E}\left\{\nabla_{\bar{\xi}} \Gamma(\xi)^* \Omega \nabla_{\bar{\xi}} \Gamma(\xi)\right\}. \tag{2.10}$$

3. Use of the Poincaré-Nash inequality.

In this paragraph, we control the variance of various functionals of $\mathbf{Q}_N(z)$ using the Poincaré-Nash inequality. For this, it appears useful to evaluate the moments of $\|W_N\|$. The following result holds.

Lemma 3.1. *For any $l \in \mathbb{N}$, it holds that $\sup_{N \geq 1} \mathbb{E}\{\|W_N\|^{2l}\} < +\infty$.*

Proof. We first remark that it is possible to be back to the case where matrix $R_N = I_M$. Due to the Gaussianity of the i.i.d. vectors $(y_n)_{n \geq 1}$, it exists i.i.d. $\mathcal{N}_c(0, I_M)$ distributed vectors $(y_{iid,n})_{n \geq 1}$ such that $\mathbb{E}(y_{iid,n} y_{iid,n}^*) = I_M$ verifying $y_n = R_N^{1/2} y_{iid,n}$. From this, we obtain immediately that the $2ML \times N$ block Hankel matrix $W_{iid,N}$ built from $(y_{n,iid})_{n=1,\ldots,N}$ satisfies

$$W_N = \begin{pmatrix} R_N^{1/2} & & \\ & \ddots & \\ & & R_N^{1/2} \end{pmatrix} W_{iid,N}. \tag{3.1}$$

As the spectral norm of R_N is assumed uniformly bounded when N increases, the statement of the lemma is equivalent to $\sup_N \mathbb{E}\{\|W_{iid}\|^{2l}\} < +\infty$. It is shown in [26] that the empirical eigenvalue distribution of $W_{iid,N} W_{iid,N}^*$ converges towards the Marcenko-Pastur distribution $\mu_{MP,*}$ with parameter c_*, i.e. $\mu_{MP,*}$ is the limit of the empirical eigenvalue distribution of matrices such as $\frac{1}{K_2} X X^*$ where X is a $K_1 \times K_2$ random matrix with i.i.d. zero mean and unit variance entries when K_1 and K_2 converge towards $+\infty$ in such a way that $\frac{K_1}{K_2} \to c_*$. The smallest non zero eigenvalue and the largest eigenvalue of $W_{iid,N} W_{iid,N}^*$ (which coincides with $\|W_{iid,N}\|^2$) converge almost surely towards $(1 - \sqrt{c_*})^2$ and $(1 + \sqrt{c_*})^2$ respectively. We express $\mathbb{E}\{\|W_{iid}\|^{2l}\}$ as

$$\mathbb{E}\{\|W_{iid}\|^{2l}\} = \mathbb{E}\{\|W_{iid}\|^{2l} \mathbf{1}_{\|W_{iid}\|^2 \leq (1+\sqrt{c_*})^2 + \delta}\} + \mathbb{E}\{\|W_{iid}\|^{2l} \mathbf{1}_{\|W_{iid}\|^2 > (1+\sqrt{c_*})^2 + \delta}\}$$
$$\leq \kappa + \mathbb{E}\{\|W_{iid}\|_F^{2l} \mathbf{1}_{\|W_{iid}\|^2 > (1+\sqrt{c_*})^2 + \delta}\} \leq \kappa + \mathbb{E}\{\|W_{iid}\|_F^{4l}\}^{1/2} \mathbb{E}\{\mathbf{1}_{\|W_{iid}\|^2 > (1+\sqrt{c_*})^2 + \delta}\}^{1/2}$$

where $\kappa > 0$ is a nice constant. As $\mathbb{E}\{\|W_{i.i.d.}\|_F^{4l}\} = \mathcal{O}(N^{2l})$, it is sufficient to prove that $\mathbb{E}\{\mathbf{1}_{\|W_{iid}\|^2 > (1+\sqrt{c_*})^2 + \delta}\}$ is less than any power of N^{-1}. We introduce a smooth function ϕ_0 defined on $\mathbb{R}$ by

$$\phi_0(\lambda) = \begin{cases} 1, \text{ for } \lambda \in [-\infty, -\delta] \cup [(1 + \sqrt{c_*})^2 + \delta, +\infty], \\ 0, \text{ for } \lambda \in [-\delta/2, (1 + \sqrt{c_*})^2 + \delta/2] \end{cases}$$

and $\phi_0(\lambda) \in (0,\, 1)$ elsewhere. Then, it holds that

$$\mathbb{E}\{\mathbf{1}_{\|W_{iid}\|^2 > (1+\sqrt{c_*})^2+\delta}\} = \mathbb{E}\{\mathbf{1}_{\lambda_{max}(W_{iid}W_{iid}^*) > (1+\sqrt{c_*})^2+\delta}\} \leq \mathbf{P}[\mathrm{Tr}\phi_0(W_{iid}W_{iid}^*) \geq 1]$$
$$\leq \mathbb{E}\{(\mathrm{Tr}\phi_0(W_{iid}W_{iid}^*))^{2k}\}$$

for any $k \in \mathbb{N}$. Lemma 3.1 thus appears as an immediate consequence of the following lemma.

Lemma 3.2. *For each smooth function ϕ such that $\phi(\lambda) = 0$ if $\lambda \in [-\delta/2,\, (1+\sqrt{c_*})^2 + \delta/2]$ and $\phi(\lambda)$ constant on $[-\infty,\, -\delta] \cup [(1+\sqrt{c_*})^2 + \delta,\, +\infty]$, it holds that $\forall k \in \mathbb{N}$, $\mathbb{E}\left\{(\mathrm{Tr}\phi(W_{iid}W_{iid}^*))^{2k}\right\} \leq \frac{\kappa}{N^{2k}}$.*

Proof. We prove the Lemma by induction. We first consider the case $k = 1$. For more convenience we will write W instead of W_{iid} in the course of the proof. Here and below we take sum for all possible values of indexes, if not specified. From (2.10) we have

$$\mathbf{Var}\{\mathrm{Tr}\phi(WW^*)\} \leq \sum \mathbb{E}\left\{\left(\frac{\partial \mathrm{Tr}\phi(WW^*)}{\partial \overline{W}_{i_1,j_1}^{m_1}}\right)^* \mathbb{E}\{W_{i_1,j_1}^{m_1}\overline{W}_{i_2,j_2}^{m_2}\} \frac{\partial \mathrm{Tr}\phi(WW^*)}{\partial \overline{W}_{i_2,j_2}^{m_2}}\right\}$$
$$+ \sum \mathbb{E}\left\{\frac{\partial \mathrm{Tr}\phi(WW^*)}{\partial W_{i_1,j_1}^{m_1}} \mathbb{E}\{W_{i_1,j_1}^{m_1}\overline{W}_{i_2,j_2}^{m_2}\}\left(\frac{\partial \mathrm{Tr}\phi(WW^*)}{\partial W_{i_2,j_2}^{m_2}}\right)^*\right\}. \tag{3.2}$$

We only evaluate the first term, denoted by ψ, of the right handside of (3.2), because the second one can be addressed similarly. For this, we first remark that

$$\frac{\partial \mathrm{Tr}\phi(WW^*)}{\partial \overline{W}_{i_1,j_1}^{m_1}} = \mathrm{Tr}\left(\phi'(WW^*)\frac{\partial WW^*}{\partial \overline{W}_{i_1,j_1}^{m_1}}\right) = (\phi'(WW^*)W)_{i_1,j_1}^{m_1}.$$

Plugging this into (3.2) we obtain

$$\psi = \sum \frac{1}{N}\mathbb{E}\left\{(\phi'(WW^*)W)_{j_1,i_1}^{*m_1}\, \delta_{m_1,m_2}\delta_{i_1+j_1,i_2+j_2}\, (\phi'(WW^*)W)_{i_2,j_2}^{m_2}\right\}.$$

Denoting $l = i_1 - i_2$, it is easy to verify that ψ can be written as

$$\psi = \frac{1}{N}\sum_{l=-(L-1)}^{L-1} \mathbb{E}\{\mathrm{Tr}\,(\phi'(WW^*)W)^*\,(J_L^{*\epsilon(l)} \otimes I_M)\,(\phi'(WW^*)W)\,J_N^{*\epsilon(l)}\} \tag{3.3}$$

where we recall that matrix J_L is defined by (2.6) and that $\epsilon(l) = l$ if $l \geq 0$ and $\epsilon(l) = *|l|$ if $l \leq 0$. For each $ML \times N$ matrices A and B, the Schwartz inequality and the inequality between arithmetic and geometric means lead to

$$\left|\frac{1}{N}\mathrm{Tr}A^*(J_L^{*\epsilon(u)} \otimes I_M)BJ_N^{*\epsilon(u)}\right| \leq \frac{1}{2N}\mathrm{Tr}A^*(J_L^{*\epsilon(u)}J_L^{\epsilon(u)} \otimes I_M)A + \frac{1}{2N}\mathrm{Tr}BJ_N^{*\epsilon(u)}J_N^{\epsilon(u)}B^*.$$

Therefore, since $J_L^{*\epsilon(u)}J_L^{\epsilon(u)} \otimes I_M \leq I_{ML}$ and $J_N^{*\epsilon(u)}J_N^{\epsilon(u)} \leq I_N$

$$\left|\frac{1}{ML}\mathrm{Tr}A^*(J_L^{*\epsilon(u)} \otimes I_M)BJ_N^{*\epsilon(u)}\right| \leq \frac{\kappa}{N}(\mathrm{Tr}A^*A + \mathrm{Tr}BB^*). \tag{3.4}$$

By taking here $A = B = \phi'(WW^*)W$, we obtain from (3.2) and (3.3)

$$\mathbf{Var}\{\mathrm{Tr}\phi(WW^*)\} \leq \frac{\kappa}{N}\mathbb{E}\left\{\mathrm{Tr}\left(\phi'(WW^*)\right)^2 WW^*\right\}. \tag{3.5}$$

Consider the function $\eta(\lambda) = (\phi'(\lambda))^2\lambda$. It is clear that $\eta(\lambda)$ is a compactly supported smooth function. Therefore (see e.g. [26]), it holds that

$$\mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}\left((\phi'(WW^*))^2 WW^*\right)\right\} = \int_{\mathcal{S}_{MP,N}} \eta(\lambda) d\mu_{MP,N}(\lambda) + \mathcal{O}\left(\frac{1}{N^2}\right),$$

where $\mu_{MP,N}$ is the measure associated to Marcenko-Pastur distribution with parameter c_N and where $\mathcal{S}_{MP,N} \subset [0, (1+\sqrt{c_N})^2]$ represents the support of $\mu_{MP,N}$. It is clear that for N large enough, the support of ϕ' and $\mathcal{S}_{MP,N}$ do not intersect, so that $\int_{\mathcal{S}_{MP,N}} \eta(\lambda) d\mu_{MP,N}(\lambda) = 0$. Therefore, we obtain that

$$\mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}\left((\phi'(WW^*))^2 WW^*\right)\right\} = \mathcal{O}\left(\frac{1}{N^2}\right).$$

This and (3.5) lead to the conclusion that $\mathbf{Var}\{\mathrm{Tr}\phi(WW^*)\} = \mathcal{O}\left(N^{-2}\right)$. To finalize the case $k = 1$, we express $\mathbb{E}\{(\mathrm{Tr}\phi(WW^*))^2\}$ as $\mathbb{E}\{(\mathrm{Tr}\phi(WW^*))^2\} = \mathbf{Var}\{\mathrm{Tr}\phi(WW^*)\} + \mathbb{E}\{\mathrm{Tr}\phi(WW^*)\}^2$. [26, Lemma 10.1] implies that $\mathbb{E}\{\mathrm{Tr}\phi(WW^*)\} = \mathcal{O}(N^{-1})$, which completes the proof for $k = 1$.

Now we suppose that for any $n \leq k$ we have $\mathbb{E}\{(\mathrm{Tr}\phi(WW^*))^{2n}\} = \mathcal{O}(N^{-2n})$ and are about to prove that it holds for $n = k+1$. As in the previous case we write

$$\mathbb{E}\{(\mathrm{Tr}\phi(WW^*))^{2(k+1)}\} = \mathbf{Var}\{(\mathrm{Tr}\phi(WW^*))^{k+1}\} + \left(\mathbb{E}\{(\mathrm{Tr}\phi(WW^*))^{k+1}\}\right)^2. \tag{3.6}$$

To evaluate the second term of the r.h.s. of (3.6), we use the Schwartz inequality and the induction assumption

$$\mathbb{E}\{(\mathrm{Tr}\phi(WW^*))^{k+1}\} \leq \left(\mathbb{E}\{(\mathrm{Tr}\phi(WW^*))^{2k}\}\mathbb{E}\{(\mathrm{Tr}\phi(WW^*))^2\}\right)^{1/2} = \mathcal{O}\left(\frac{1}{N^{k+1}}\right), \tag{3.7}$$

We follow the same steps as in the case $k = 1$ to study the first term of the r.h.s. of (3.6). Using again the Poincaré-Nash inequality, we obtain that

$$\mathbf{Var}\{(\mathrm{Tr}\phi(WW^*))^{k+1}\} \leq \frac{\kappa}{N}\mathbb{E}\left\{(\mathrm{Tr}\phi(WW^*))^{2k}\,\mathrm{Tr}\left(\phi'(WW^*)^2 WW^*\right)\right\}.$$

Using Holder's inequality, we obtain

$$\mathbf{Var}\{(\mathrm{Tr}\phi(WW^*))^{k+1}\} \leq \frac{\kappa}{N}\mathbb{E}\left\{(\mathrm{Tr}\phi(WW^*))^{2k+2}\right\}^{\frac{k}{k+1}}\mathbb{E}\left\{\left(\mathrm{Tr}(\phi'(WW^*)^2 WW^*)\right)^{k+1}\right\}^{\frac{1}{k+1}}. \tag{3.8}$$

The properties of function $\eta(\lambda) = \phi'(\lambda)^2\lambda$ imply that it satisfies the induction hypothesis and that it verifies (3.7), i.e. $\mathbb{E}\left\{(\mathrm{Tr}(\phi'(WW^*)^2 WW^*))^{k+1}\right\} = \mathcal{O}(\frac{1}{N^{k+1}})$. Plugging this into (3.8), we get that

$$\mathbf{Var}\{(\mathrm{Tr}\phi(WW^*))^{k+1}\} \leq \frac{\kappa}{N^2}\mathbb{E}\left\{(\mathrm{Tr}\phi(WW^*))^{2k+2}\right\}^{\frac{k}{k+1}}.$$

From this, (3.7) and (3.6), we immediately obtain

$$\mathbb{E}\{(\mathrm{Tr}\phi(WW^*))^{2k+2}\} \leq \frac{\kappa_1}{N^2}\mathbb{E}\{(\mathrm{Tr}\phi(WW^*))^{2k+2}\}^{\frac{k}{k+1}} + \frac{\kappa_2}{N^{2k+2}}. \tag{3.9}$$

We denote by $z_{k,N}$ the term $z_{k,N} = N^{2k+2}\,\mathbb{E}\{(\mathrm{Tr}\phi(WW^*))^{2k+2}\}$. Then, (3.9) implies that

$$z_{k,N} \leq \kappa_1\,(z_{k,N})^{k/(k+1)} + \kappa_2.$$

This inequality leads to the conclusion that sequence $(z_{k,N})_{N\geq 1}$ is bounded, or equivalently that $\mathbb{E}\{(\mathrm{Tr}\phi(WW^*))^{2k+2}\} \leq \frac{\kappa}{N^{2k+2}}$ as expected. This completes the proof of Lemmas 3.2 and 3.1. ■

We now evaluate the variance of useful functionals of the resolvent $\mathbf{Q}_N(z)$.

Lemma 3.3. *Let $(F_N)_{N\geq1}$, $(G_N)_{N\geq1}$ be sequences of deterministic $2ML \times 2ML$ matrices and $(H_N)_{N\geq1}$ a sequence of deterministic $N \times N$ matrices such that $\max\{\sup_N \|F_N\|, \sup_N \|G_N\|, \sup_N \|H_N\|\} \leq \kappa$. Then, for each $z \in \mathbb{C}^+$, it holds that*

$$\mathbf{Var}\left\{\frac{1}{ML}\mathrm{Tr}F\mathbf{Q}\right\} \leq \frac{C(z)\kappa^2}{N^2}, \tag{3.10}$$

$$\mathbf{Var}\left\{\frac{1}{ML}\mathrm{Tr}F\mathbf{Q}GWHW^*\right\} \leq \frac{C(z)\kappa^6}{N^2}. \tag{3.11}$$

where $C(z)$ can be written as $C(z) = P_1(|z|)P_2\left(\frac{1}{\mathrm{Im}z}\right)$ for some nice polynomials P_1 and P_2.

Proof. We first prove (3.10) and denote by ξ the term $\xi = \frac{1}{ML}\mathrm{Tr}F\mathbf{Q}$. The Poincare-Nash inequality leads to

$$\mathbf{Var}\{\xi\} \leq \sum_{\substack{i_1,j_1,m_1\\ i_2,j_2,m_2}} \mathbb{E}\left\{\left(\frac{\partial\xi}{\partial\overline{W}^{m_1}_{i_1,j_1}}\right)^* \mathbb{E}\{W^{m_1}_{i_1,j_1}\overline{W}^{m_2}_{i_2,j_2}\}\frac{\partial\xi}{\partial\overline{W}^{m_2}_{i_2,j_2}}\right\}$$
$$+ \sum_{\substack{i_1,j_1,m_1\\ i_2,j_2,m_2}} \mathbb{E}\left\{\frac{\partial\xi}{\partial W^{m_1}_{i_1,j_1}}\mathbb{E}\{W^{m_1}_{i_1,j_1}\overline{W}^{m_2}_{i_2,j_2}\}\left(\frac{\partial\xi}{\partial W^{m_2}_{i_2,j_2}}\right)^*\right\}.$$

We just evaluate the first term of the r.h.s. that we denote by ϕ. For this, we need the expression of the derivative of $\mathbf{Q}$ with respect to the complex conjugates of the entries of W. We denote by Π_{pf} and Π_{fp} the $2ML \times 2ML$ matrices defined by $\Pi_{pf} = \begin{pmatrix} 0 & I_{ML} \\ 0 & 0\end{pmatrix}$ and $\Pi_{fp} = \begin{pmatrix} 0 & 0 \\ I_{ML} & 0\end{pmatrix}$. Then, after some algebra, we obtain that

$$\frac{\partial\mathbf{Q}}{\partial\overline{W}^m_{i,j}} = -\mathbf{Q}\begin{pmatrix} w_{j,f} \\ 0\end{pmatrix}(\mathbf{f}^m_{i+L})^T\mathbf{Q}\,\mathbf{1}_{i\leq L} - \mathbf{Q}\begin{pmatrix} 0 \\ w_{j,p}\end{pmatrix}(\mathbf{f}^m_{i-L})^T\mathbf{Q}\,\mathbf{1}_{i>L}$$
$$= -\mathbf{Q}\Pi_{pf}W\mathbf{e}_j\,(\mathbf{f}^m_i)^T\Pi_{pf}\mathbf{Q} - \mathbf{Q}\Pi_{fp}W\mathbf{e}_j\,(\mathbf{f}^m_i)^T\Pi_{fp}\mathbf{Q}. \tag{3.12}$$

From this, we deduce immediately that

$$\frac{\partial \xi}{\partial \overline{W}^{m_1}_{i_1,j_1}} = -\frac{1}{ML}\Big(\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W + \Pi_{fp}\mathbf{Q}F\mathbf{Q}\Pi_{fp}W\Big)^{m_1}_{i_1,j_1}.$$

Using that $\mathbb{E}\{W^{m_1}_{i_1,j_1}\overline{W}^{m_2}_{i_2,j_2}\} = \frac{1}{N}R_{m_1m_2}\delta_{i_1+j_1,i_2+j_2}$, we obtain that ϕ is given by

$$\phi = \frac{1}{N(ML)^2}\sum_{\substack{i_1,j_1,m_1\\ i_2,j_2,m_2}} (\mathbf{e}_{j_1})^T(\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W + \Pi_{fp}\mathbf{Q}F\mathbf{Q}\Pi_{fp}W)^*\mathbf{f}^{m_1}_{i_1}R_{m_1m_2}$$
$$\times\, \delta_{i_1+j_1,i_2+j_2}(\mathbf{f}^{m_2}_{i_2})^T(\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W + \Pi_{fp}\mathbf{Q}F\mathbf{Q}\Pi_{fp}W)\mathbf{e}_{j_2}.$$

We put $u = i_1 - i_2$ and remark that $\sum_{m_1,m_2,i_1-i_2=u}\mathbf{f}^{m_1}_{i_1}R_{m_1m_2}(\mathbf{f}^{m_2}_{i_2})^T = J_L^{*\epsilon(u)} \otimes R$ and that $\sum_{j_2-j_1=u}\mathbf{e}_{j_2}\mathbf{e}^T_{j_1} = J_N^{*\epsilon(u)}$. Therefore, ϕ can be written as

$$\phi = \frac{1}{MLN}\mathbb{E}\Big\{\sum_{u=-(L-1)}^{L-1}\frac{1}{ML}\mathrm{Tr}(\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W + \Pi_{fp}\mathbf{Q}F\mathbf{Q}\Pi_{fp}W)^*(J_L^{*\epsilon(u)} \otimes R)$$
$$\times\,(\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W + \Pi_{fp}\mathbf{Q}F\mathbf{Q}\Pi_{fp}W)J_N^{*\epsilon(u)}\Big\}. \tag{3.13}$$

Each term inside the sum over u can be written as $\dfrac{1}{ML}\mathrm{Tr}A^*(I_L \otimes R^{1/2})(J_L^{*\epsilon(u)} \otimes I)(I_L \otimes R^{1/2})AJ_N^{*\epsilon(u)}$, where the expression of the $ML \times N$ matrix A is omitted. As $\|R\|$ is bounded by the nice constant b (see (2.7)), (3.4) and (3.13) lead to the conclusion that we just need to evaluate $\frac{1}{ML}\mathbb{E}\{\mathrm{Tr}A^*A\}$. Using the Schwartz inequality, we obtain immediately that

$$\mathbb{E}\{\mathrm{Tr}A^*A\} \le 2\mathbb{E}\{\mathrm{Tr}\,((\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W)^*\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W)\} \tag{3.14}$$
$$+\, 2\mathbb{E}\{\mathrm{Tr}\,((\Pi_{fp}\mathbf{Q}F\mathbf{Q}\Pi_{fp}W)^*\Pi_{fp}\mathbf{Q}F\mathbf{Q}\Pi_{fp}W)\}.$$

Since $(\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf})^*\,\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf} \le \|\mathbf{Q}\|^4\|F\|^2\, I$ and $\|\mathbf{Q}\| \le \frac{1}{\mathrm{Im}z}$, we get that

$$\frac{1}{ML}\mathbb{E}\{\mathrm{Tr}\,((\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W)^*\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W)\} \le \frac{1}{(\mathrm{Im}z)^4}\,\|F\|^2\,\frac{1}{ML}\mathbb{E}\{\mathrm{Tr}W^*W\}$$
$$\le \frac{1}{(\mathrm{Im}z)^4}\,\|F\|^2\,\mathbb{E}\{\|W\|^2\}$$

Lemma 3.1 thus implies that

$$\frac{1}{ML}\mathbb{E}\{\mathrm{Tr}\,((\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W)^*\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W)\} \le \kappa^2 P\left(\frac{1}{\mathrm{Im}z}\right)$$

for some nice polynomial P. The term $\frac{1}{ML}\mathbb{E}\{\mathrm{Tr}\,(\Pi_{fp}\mathbf{Q}F\mathbf{Q}\Pi_{fp}W)^*\Pi_{fp}\mathbf{Q}F\mathbf{Q}\Pi_{fp}W)\}$ can be handled similarly. Therefore, (3.13) leads to $\phi \le \kappa^2\,\frac{1}{N^2}P\left(\frac{1}{\mathrm{Im}z}\right)$. This establishes (3.10).

To prove (3.11) one can also use Poincaré-Nash inequality for $\xi = \frac{1}{ML}\mathrm{Tr}F\mathbf{Q}GWHW^*$. After some calculations, we get that the variance of ξ is upperbounded by a term given by

$$\frac{\kappa_1}{N^2}\,\mathbb{E}\left(\frac{1}{ML}\mathrm{Tr}(F\mathbf{Q}GWH)^*(F\mathbf{Q}GWH) + \frac{1}{ML}\mathrm{Tr}(F\mathbf{Q}WH)^*(F\mathbf{Q}WH) + \eta_1 + \eta_2\right), \tag{3.15}$$

where κ_1 is some nice constant, and where η_1 and η_2 are defined by

$$\eta_1 = \frac{1}{ML}\mathrm{Tr}(\Pi_{pf}\mathbf{Q}GWHW^*F\mathbf{Q}\Pi_{pf}W)^*(\Pi_{pf}\mathbf{Q}GWHW^*F\mathbf{Q}\Pi_{pf}W),$$
$$\eta_2 = \frac{1}{ML}\mathrm{Tr}(\Pi_{fp}\mathbf{Q}GWHW^*F\mathbf{Q}\Pi_{fp}W)^*(\Pi_{fp}\mathbf{Q}GWHW^*F\mathbf{Q}\Pi_{fp}W).$$

Using Lemma 3.1 as well as the inequality $\mathbf{Q}\mathbf{Q}^* \le \frac{1}{\mathrm{Im}^2 z}I$, we obtain immediately (3.11). This completes the proof of Lemma 3.3. ■

In the following, we also need to evaluate the variance of more specific terms. The following result appears to be a consequence of Lemma 3.3 and of the particular structure (2.4) of matrix $\mathbf{Q}(z)$.

Corollary 3.1. *Let $(F_{1,N})_{N\ge 1}$ be a sequence of deterministic $ML \times ML$ matrices such that $\sup_N \|F_{1,N}\| \le \kappa$, and $(H_N)_{N\ge1}$ a sequence of deterministic $N \times N$ matrices satisfying $\sup_N \|H_N\| \le 1$. Then, if $z \in \mathbb{C}^+$ and $\mathrm{Im}z^2 > 0$, the following evaluations hold:*

$$\mathbf{Var}\left\{\frac{1}{ML}\mathrm{Tr}F_1\mathbf{Q}_{\mathbf{ij}}(z)\right\} \le \kappa^2\frac{1}{N^2}P_1(|z^2|)P_2\left(\frac{1}{\mathrm{Im}z^2}\right), \tag{3.16}$$

where i and j belong to $\{p, f\}$;

$$\mathbf{Var}\left\{\frac{1}{ML}\mathrm{Tr}\left[HW^*\Pi_{i_1j_1}\begin{pmatrix}F_1 & 0\\ 0 & 0\end{pmatrix}\mathbf{Q}(z)\Pi_{i_2j_2}W\right]\right\} \le \kappa^2\frac{1}{N^2}P_1(|z^2|)P_2\left(\frac{1}{\mathrm{Im}z^2}\right), \tag{3.17}$$

where i_1, j_1, i_2, j_2 still belong to $\{p, f\}$, but verify $i_1 \ne j_1$ and $i_2 \ne j_2$.

Proof. We first prove (3.16), and first consider the case where $i = j = p$. We define the $2ML\times 2ML$ matrix F by $F = \begin{pmatrix}F_1 & 0\\ 0 & 0\end{pmatrix}$, and remark that $\frac{1}{ML}\mathrm{Tr}F_1\mathbf{Q}_{\mathbf{pp}}(z)$ coincides with $\xi = \frac{1}{ML}\mathrm{Tr}F\mathbf{Q}(z)$. We follow the proof of (3.10), and evaluate the right hand side of (3.14) in a more accurate manner by taking into account the particular structure of the present matrix F. It is easy to check that

$$\frac{1}{ML}\mathbb{E}\{\mathrm{Tr}\,(\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W)^*\Pi_{pf}\mathbf{Q}F\mathbf{Q}\Pi_{pf}W)\}$$
$$= \frac{1}{ML}\mathbb{E}\{\mathrm{Tr}\left(W_f^*\mathbf{Q}_{\mathbf{pp}}^*F_1^*\mathbf{Q}_{\mathbf{fp}}^*\mathbf{Q}_{\mathbf{fp}}F_1\mathbf{Q}_{\mathbf{pp}}W_f\right)\}.$$

As $\mathbf{Q}_{\mathbf{fp}}(z) = W_pW_f^*Q(z^2)$, we obtain that

$$\mathbf{Q}_{\mathbf{fp}}^*(z)\mathbf{Q}_{\mathbf{fp}}(z) = (Q(z^2))^*W_fW_p^*W_pW_f^*Q(z^2) \le \|W\|^4\frac{1}{(\mathrm{Im}z^2)^2}\,I$$

if $\mathrm{Im}(z^2) > 0$. Therefore, it holds that

$$F_1^* \mathbf{Q}_{\mathbf{fp}}^* \mathbf{Q}_{\mathbf{fp}} F_1 \leq \kappa^2 \|W\|^4 \frac{1}{(\mathrm{Im} z^2)^2}\, I.$$

From this, using the expression of $\mathbf{Q}_{pp} = zQ(z^2)$, we obtain similarly that

$$W_f^* \mathbf{Q}_{\mathbf{pp}}^* F_1^* \mathbf{Q}_{\mathbf{fp}}^* \mathbf{Q}_{\mathbf{fp}} F_1 \mathbf{Q}_{\mathbf{pp}} W_f \leq \kappa^2 \|W\|^6 \frac{|z|^2}{(\mathrm{Im} z^2)^4}.$$

Lemma 3.1 thus leads to the conclusion that

$$\frac{1}{ML} \mathbb{E}\{\mathrm{Tr}\left(W_f^* \mathbf{Q}_{\mathbf{pp}}^* F_1^* \mathbf{Q}_{\mathbf{fp}}^* \mathbf{Q}_{\mathbf{fp}} F_1 \mathbf{Q}_{\mathbf{pp}} W_f\right)\} \leq \kappa^2 \frac{\kappa_1 |z|^2}{(\mathrm{Im} z^2)^4},$$

where κ_1 is a nice constant such that $\mathbb{E}(\|W_N\|^6) \leq \kappa_1$ for each N. Using similar arguments, we obtain that

$$\frac{1}{ML} \mathbb{E}\{\mathrm{Tr}\,(\Pi_{fp} \mathbf{Q} F \mathbf{Q} \Pi_{fp} W)^* \Pi_{fp} \mathbf{Q} F \mathbf{Q} \Pi_{fp} W)\} \leq \kappa^2 \frac{\kappa_1 |z^2|^2}{(\mathrm{Im} z^2)^4}.$$

This, in turn, implies (3.16) for $i = j = p$. As the arguments are essentially the same for the other values of i and j, we do not provide the corresponding proofs.

In order to establish (3.17), we follow the proof (3.11) for $F = \Pi_{i_1 j_1} \begin{pmatrix} F_1 & 0 \\ 0 & 0 \end{pmatrix}$, $G = \Pi_{i_2 j_2}$. It is necessary to check that the 4 terms inside the bracket of (3.15) can be upperbounded by $\kappa^2 P_1(|z^2|) P_2(\frac{1}{\mathrm{Im} z^2})$ for nice polynomials P_1 and P_2. As above, the use of the particular expression of matrices $(\mathbf{Q}_{\mathbf{ij}})_{i,j \in \{f,p\}}$ allows to establish this property. The corresponding easy calculations are omitted. ■

4. Various lemmas on Stieltjes transform

In this paragraph, we provide a number of useful results on certain Stieltjes transforms. We recall that if K is a positive integer, then a $K \times K$ matrix-valued positive measure ω is a σ–additive function from the Borel sets of $\mathbb{R}$ onto the set of all positive $K \times K$ matrices (see e.g. [34], Chapter 1 for more details). ω is said to be finite if the scalar positive measure $\mathrm{Tr}(\omega)$ is finite. In the following, if A is a Borel set of $\mathbb{R}$, we denote by $\mathcal{S}_M(A)$ the set of all Stieltjes transforms of $M \times M$ matrix valued positive finite measures carried by A. $\mathcal{S}_1(A)$ is denoted $\mathcal{S}(A)$. We first begin by stating well known properties of Stieltjes transforms (see e.g. the Appendix of [21], the Appendix A of [16], and the references therein).

Proposition 4.1. *The following properties hold true:*

1. Let f be the Stieltjes transform of a positive finite measure μ, then

– the function f is analytic over $\mathbb{C}^+$,

– if $z \in \mathbb{C}^+$ then $f(z) \in \mathbb{C}^+$,

– the function f satisfies: $|f(z)| \leq \frac{\mu(\mathbb{R})}{\mathrm{Im} z}$, for $z \in \mathbb{C}^+$

– if $\mu(-\infty, 0) = 0$ then its Stieltjes transform f is analytic over $\mathbb{C}/\mathbb{R}^+$. Moreover, $z \in \mathbb{C}^+$ implies $zf(z) \in \mathbb{C}^+$.

– *for all* $\phi \in \mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R})$ *we have*

$$\int_{\mathbb{R}} \phi(\lambda) d\mu(\lambda) = \frac{1}{\pi} \lim_{y \downarrow 0} \mathrm{Im} \left\{ \int_{\mathbb{R}} \phi(x) f(x+iy) dx \right\}.$$

2. Conversely, let f *be a function analytic over* $\mathbb{C}^+$ *such that* $f(z) \in \mathbb{C}^+$ *if* $z \in \mathbb{C}^+$ *and for which* $\sup_{y \geq \epsilon} |iyf(iy)| < +\infty$ *for some* $\epsilon > 0$*. Then,* f *is the Stieltjes transform of a unique positive finite measure* μ *such that* $\mu(\mathbb{R}) = \lim_{y \to +\infty} -iyf(iy)$*. If moreover* $zf(z) \in \mathbb{C}^+$ *for* z *in* $\mathbb{C}^+$ *then,* $\mu(\mathbb{R}^-) = 0$*. In particular,* f *is given by*

$$f(z) = \int_0^{+\infty} \frac{\mu(d\lambda)}{\lambda - z}$$

and has an analytic continuation on $\mathbb{C}/\mathbb{R}^+$*.*

3. Let F *be an* $P \times P$ *matrix-valued function analytic on* $\mathbb{C}^+$ *verifying*

– $\mathrm{Im}(F(z)) > 0$ *if* $z \in \mathbb{C}^+$

– $\sup_{y > \epsilon} \|iyF(iy)\| < +\infty$ *for some* $\epsilon > 0$*.*

Then, $F \in \mathcal{S}_P(\mathbb{R})$*, and if* μ^F *is the corresponding* $P \times P$ *associated positive measure, it holds that*

$$\mu^F(\mathbb{R}) = \lim_{y \to +\infty} -iyF(iy). \tag{4.1}$$

If moreover $\mathrm{Im}(zF(z)) > 0$*, then,* $F \in \mathcal{S}_P(\mathbb{R}^+)$*.*

We now state a quite useful Lemma.

Lemma 4.1. *Let* $\beta(z) \in \mathcal{S}(\mathbb{R}^+)$*, and consider function* $\boldsymbol{\beta}(z)$ *defined by* $\boldsymbol{\beta}(z) = z\beta(z^2)$*. Then* $\boldsymbol{\beta} \in \mathcal{S}(\mathbb{R})$*. Moroever, it holds that*

$$\begin{aligned} \mathbf{G}(z) &= \left(-zI_M - \frac{c\boldsymbol{\beta}(z)}{1 - c^2\boldsymbol{\beta}^2(z)} R \right)^{-1} \in \mathcal{S}_M(\mathbb{R}) \\ G(z) &= \left(-zI_M - \frac{cz\beta(z)}{1 - zc^2\beta^2(z)} R \right)^{-1} \in \mathcal{S}_M(\mathbb{R}^+) \end{aligned} \tag{4.2}$$

and that

$$\mathbf{G}(z)\left(\mathbf{G}(z)\right)^* \leq \frac{I_M}{(\mathrm{Im}z)^2},\ G(z)\left(G(z)\right)^* \leq \frac{I_M}{(\mathrm{Im}z)^2}. \tag{4.3}$$

Finally, matrices $\mathbf{G}(z)$ *and* $G(z)$ *are linked by the relation*

$$\mathbf{G}(z) = zG(z^2) \tag{4.4}$$

for each $z \in \mathbb{C}^+$*.*

Proof. Let τ be the measure carried by $\mathbb{R}^+$ corresponding to the Stieltjes transform $\beta(z)$. We first prove that $\boldsymbol{\beta}(z)$ is a Stieltjes transform. We first remark that if $z \in \mathbb{C}^+$, then $z^2 \in \mathbb{C} - \mathbb{R}^+$. β analytic on $\mathbb{C} - \mathbb{R}^+$ thus implies that $\boldsymbol{\beta}(z)$ is analytic on $\mathbb{C}^+$. Moreover, it is clear that

$$\mathrm{Im}\boldsymbol{\beta}(z) = \mathrm{Im} \int_{\mathbb{R}^+} \frac{z\, d\,\tau(\lambda)}{\lambda - z^2} = \int_{\mathbb{R}^+} \frac{\mathrm{Im}z(\lambda + |z|^2) d\,\tau(\lambda)}{|\lambda - z^2|^2} > 0, \text{ when } \mathrm{Im}z > 0.$$

To evaluate $\boldsymbol{\beta}(z)$ for $z \in \mathbb{C}^+$, we write

$$\left| \int_{\mathbb{R}^+} \frac{z d\,\tau(\lambda)}{\lambda - z^2} \right| \leq \int_{\mathbb{R}^+} \frac{d\,\tau(\lambda)}{\left|\frac{\lambda}{z} - z\right|}.$$

Using that $\left|\frac{\lambda}{z} - z\right| \geq \left|\mathrm{Im}(\frac{\lambda}{z} - z)\right| \geq \mathrm{Im} z$ for $z \in \mathbb{C}^+$ and $\lambda \geq 0$, we get that

$$|\boldsymbol{\beta}(z)| \leq \int_{\mathbb{R}^+} \frac{d\,\tau(\lambda)}{\mathrm{Im} z} = \frac{\tau(\mathbb{R}^+)}{\mathrm{Im} z}.$$

From this and Proposition 4.1, we obtain that $\boldsymbol{\beta}(z) \in \mathcal{S}(\mathbb{R})$.

To prove (4.2), it is first necessary to show that $\mathbf{G}$ is analytic on $\mathbb{C}^+$. For this, we first check that $m(z) = 1 - c^2\boldsymbol{\beta}^2(z) \neq 0$ for $z \in \mathbb{C}^+$. Indeed, write $\boldsymbol{\beta}(z) = x + iy$ with $y > 0$, then $m(z) = 1 - c^2x^2 + c^2y^2 - 2cxyi$. Hence, if $x = 0$ we have $m(z) = 1 + c^2y^2 > 0$, and if $x \neq 0$ then $2xy \neq 0$ since $y > 0$. In order to establish that matrix $\left(-zI_M - \frac{c\boldsymbol{\beta}(z)}{1 - c^2\boldsymbol{\beta}^2(z)} R\right)$ is invertible on $\mathbb{C}^+$, we verify that

$$\mathrm{Im}\left(-zI_M - \frac{c\boldsymbol{\beta}(z)}{1 - c^2\boldsymbol{\beta}^2(z)} R\right) < 0 \tag{4.5}$$

on $\mathbb{C}^+$. It is easy to check that

$$\mathrm{Im}\left(-zI_M - \frac{c\boldsymbol{\beta}(z)}{1 - c^2\boldsymbol{\beta}^2(z)} R\right) = -\mathrm{Im} z\, I_M - \frac{c\mathrm{Im}\boldsymbol{\beta}(z)(1 + c^2|\boldsymbol{\beta}(z)|^2)}{|1 - c^2\boldsymbol{\beta}^2(z)|^2} R < -\mathrm{Im} z\, I_M.$$

Therefore, $\mathrm{Im} z > 0$ and $\mathrm{Im}\boldsymbol{\beta}(z) > 0$ imply (4.5). The imaginary part of $\mathbf{G}(z)$ is given by

$$\mathrm{Im}(\mathbf{G}(z)) = -\mathbf{G}(z)\mathrm{Im}\left(-zI_M - \frac{c\boldsymbol{\beta}(z)}{1 - c^2\boldsymbol{\beta}^2(z)} R\right) (\mathbf{G}(z))^* > \mathrm{Im} z\, \left(\mathbf{G}(z)\, (\mathbf{G}(z))^*\right) > 0.$$

Therefore, $\mathrm{Im}\mathbf{G}(z) > 0$ if $z \in \mathbb{C}^+$. We finally remark that $\lim_{y \to +\infty} -iy\mathbf{G}(iy) = I_M$, which implies that $\sup_{y > \epsilon} \|iy\mathbf{G}(iy)\| < +\infty$ for each $\epsilon > 0$. Proposition 4.1 eventually implies that $\mathbf{G} \in \mathcal{S}_M(\mathbb{R})$. Moreover, if $\boldsymbol{\tau}^{\mathbf{G}}$ is the underlying $M \times M$ positive matrix valued measure, (4.1) leads to $\boldsymbol{\tau}^{\mathbf{G}}(\mathbb{R}) = I_M$.

We prove similarly the analyticity of $G(z)$ on $\mathbb{C}^+$. We first check that $1 - zc^2\beta^2(z) \neq 0$ if $z \in \mathbb{C}^+$, or equivalently that $|1 - zc^2\beta^2(z)| \neq 0$ if $z \in \mathbb{C}^+$. We remark that

$$|1 - zc^2\beta^2(z)| = |z\beta(z)||c^2\beta(z) - \frac{1}{z\beta(z)}| > \mathrm{Im} z\, \mathrm{Im}\beta(z)\, \mathrm{Im}\left(c^2\beta(z) - \frac{1}{z\beta(z)}\right). \tag{4.6}$$

As $\beta \in \mathcal{S}(\mathbb{R}^+)$, it holds that $\mathrm{Im}\left(c^2\beta(z) - \frac{1}{z\beta(z)}\right) > 0$ if $z \in \mathbb{C}^+$. Therefore, $1 - zc^2\beta^2(z) \neq 0$ if $z \in \mathbb{C}^+$. As above, we verify that

$$\mathrm{Im}\left(-zI_M - \frac{cz\beta(z)}{1 - z(c\beta(z))^2} R\right) = -\mathrm{Im} z\, I_M - \mathrm{Im}\left(\frac{cz\beta(z)}{1 - z(c\beta(z))^2}\right) R < -\mathrm{Im} z\, I_M. \tag{4.7}$$

For this, we remark that

$$\mathrm{Im}\left(\frac{cz\beta(z)}{1-z(c\beta(z))^2}\right) = \frac{c}{|1-z(c\beta(z))^2|^2}\,\left(\mathrm{Im}(z\beta(z)) + |zc\beta(z)|^2\mathrm{Im}\beta(z)\right) > 0$$

if $z \in \mathbb{C}^+$, which, of course, leads to (4.7). Therefore, matrix $\left(-zI_M - \frac{cz\beta(z)}{1-z(c\beta(z))^2}R\right)$ is invertible if $z \in \mathbb{C}^+$, and G is analytic on $\mathbb{C}^+$. Moreover, we obtain immediately that

$$\mathrm{Im}(G(z)) = G(z)\left(\mathrm{Im}z\, I_M + \mathrm{Im}\left(\frac{cz\beta(z)}{1-z(c\beta(z))^2}\right)\, R\right)(G(z))^* > \mathrm{Im}z\;(G(z)G(z)^*) > 0 \tag{4.8}$$

$$\mathrm{Im}(zG(z)) = G(z)\mathrm{Im}\left(\frac{cz\beta(z)}{1-z(c\beta(z))^2}\right)\, R(G(z))^* > 0$$

for $z \in \mathbb{C}^+$. As above, it holds that $\lim_{y\to+\infty} -iyG(iy) = I$ and that $\sup_{y>\epsilon}\|iyG(iy)\| < +\infty$ for each $\epsilon > 0$. This implies that $G \in \mathcal{S}_M(\mathbb{R}^+)$, and that if τ^G represents the associated $M\times M$ matrix-valued measure, then $\tau^G(\mathbb{R}^+) = I$.

In order to establish (4.3), we follow [15, Lemma 3.1]. More precisely, we remark that

$$\mathrm{Im}G(z) = \mathrm{Im}z\int_{\mathbb{R}^+}\frac{d\tau^G(\lambda)}{|\lambda - z|^2} < \frac{\tau^G(\mathbb{R}^+)}{\mathrm{Im}z} = \frac{I}{\mathrm{Im}z}.$$

Therefore, (4.8) leads to $(G(z)G(z)^*) \le \frac{I}{(\mathrm{Im}z)^2}$. The other statement of (4.3) is proved similarly and this completes the proof. ■

Lemma 4.2. *We consider a sequence $(\beta_N)_{N\ge1}$ of elements of $\mathcal{S}(\mathbb{R}^+)$ whose associated positive measures $(\tau_N)_{N\ge1}$ satisfy for each $N \ge 1$*

$$\tau_N(\mathbb{R}^+) = \frac{1}{M}\mathrm{Tr}R_N \tag{4.9}$$

as well as

$$\int_{\mathbb{R}^+}\lambda\, d\,\tau_N(\lambda) = c_N\,\frac{1}{M}\mathrm{Tr}R_N\,\frac{1}{M}\mathrm{Tr}R_N^2. \tag{4.10}$$

Then, it exist nice constants ω, κ such that

$$\mathrm{Im}\beta_N(z) \ge \frac{\kappa\,\mathrm{Im}z}{(\omega^2 + |z|^2)} \tag{4.11}$$

and

$$\left|1 - z\,(c_N\beta_N(z))^2\right| \ge \frac{\kappa\,(\mathrm{Im}z)^3}{(\omega^2+|z|^2)^2} \tag{4.12}$$

for each $z \in \mathbb{C}^+$ and for each $N \ge 1$. Moreover, if $\boldsymbol{\beta}_N(z)$ is defined by $\boldsymbol{\beta}_N(z) = z\,\beta_N(z^2)$, then, we also have

$$\mathrm{Im}\boldsymbol{\beta}_N(z) \ge \frac{\kappa\,(\mathrm{Im}z)^3}{(\omega^2+|z|^4)} \tag{4.13}$$

and

$$\left|1 - (c_N \boldsymbol{\beta}_N(z))^2\right| \geq \frac{\kappa\,(\mathrm{Im} z)^6}{(\omega^2 + |z|^4)^2} \tag{4.14}$$

for each $z \in \mathbb{C}^+$ *and for each* $N \geq 1$.

Proof. We first establish (4.11). $\mathrm{Im}\beta_N(z)$ is given by

$$\mathrm{Im}\beta_N(z) = \mathrm{Im} z \int_{\mathbb{R}^+} \frac{d\,\tau_N(\lambda)}{|\lambda - z|^2}.$$

For each $\omega > 0$, it is clear that

$$\int_{\mathbb{R}^+} \frac{d\,\tau_N(\lambda)}{|\lambda - z|^2} \geq \int_0^\omega \frac{d\,\tau_N(\lambda)}{|\lambda - z|^2} \geq \frac{\tau_N([0,\omega])}{2(\lambda^2 + |z|^2)}.$$

Assumption (2.7) and (4.10) imply that the sequence $(\tau_N)_{N\geq 1}$ is tight. For each $\epsilon > 0$, it thus exists $\omega > 0$ for which $\tau_N(]\omega, +\infty[) < \epsilon$ for each N, or equivalently, $\tau_N([0,\omega]) > \tau_N(\mathbb{R}^+) - \epsilon$. As $\tau_N(\mathbb{R}^+) = \frac{1}{M}\mathrm{Tr}(R_N) > a$, we choose $\epsilon = a/2$, and obtain that the corresponding ω verifies $\tau_N([0,\omega]) > a/2$ for each N. This completes the proof of (4.11). We now verify (4.12). For this, we use (4.6). As $\mathrm{Im}\left(\frac{1}{z\beta_N(z)}\right) < 0$, it holds that $\mathrm{Im}\left(c_N^2\beta_N(z) - \frac{1}{z\beta_N(z)}\right) \geq c_N^2 \mathrm{Im}\beta_N(z)$. Therefore, we obtain that

$$\left|1 - z\,(c_N\beta_N(z))^2\right| \geq c_N^2\,\mathrm{Im} z\,\left(\mathrm{Im}\beta_N(z)\right)^2 \tag{4.15}$$

which implies (4.12).

We finally verify (4.13) and (4.14). For this, we first express $\boldsymbol{\beta}_N(z)$ as

$$\boldsymbol{\beta}_N(z) = z\beta_N(z^2) = \int_{\mathbb{R}^+} \frac{z}{\lambda - z^2}\,d\,\tau_N(\lambda)$$

which leads immediately to

$$\begin{aligned}\mathrm{Im}\boldsymbol{\beta}_N(z) = \mathrm{Im} z \int_{\mathbb{R}^+} \frac{\lambda + |z|^2}{|\lambda - z^2|^2}\,d\,\tau_N(\lambda) &\geq \mathrm{Im} z\,|z|^2 \int_{\mathbb{R}^+} \frac{1}{|\lambda - z^2|^2}\,d\,\tau_N(\lambda) \\ &\geq (\mathrm{Im} z)^3 \int_{\mathbb{R}^+} \frac{1}{|\lambda - z^2|^2}\,d\,\tau_N(\lambda).\end{aligned}$$

We observe that for $\omega > 0$, then,

$$\int_{\mathbb{R}^+} \frac{1}{|\lambda - z^2|^2}\,d\,\tau_N(\lambda) \geq \int_0^\omega \frac{1}{|\lambda - z^2|^2}\,d\,\tau_N(\lambda) \geq \frac{1}{2(\omega^2 + |z|^4)}\,\tau_N([0,\omega]).$$

As justified above, it is possible to choose ω for which $\tau_N([0,\omega]) \geq \frac{a}{2}$ for each N. This leads to (4.13).

We now remark that $|1 - c_N^2\boldsymbol{\beta}_N^2| = |\boldsymbol{\beta}_N|\,|\frac{1}{\boldsymbol{\beta}_N} - c_N^2\boldsymbol{\beta}_N|$. As $\mathrm{Im}\boldsymbol{\beta}_N > 0$ on $\mathbb{C}^+$, it holds that

$$\left|\frac{1}{\boldsymbol{\beta}_N} - c_N^2\boldsymbol{\beta}_N\right| \geq \left|\mathrm{Im}\left(\frac{1}{\boldsymbol{\beta}_N} - c_N^2\boldsymbol{\beta}_N\right)\right| \geq c_N^2\mathrm{Im}\boldsymbol{\beta}_N.$$

Using that $|\boldsymbol{\beta}_N| \geq \mathrm{Im}\boldsymbol{\beta}_N$, we eventually obtain that

$$|1 - c_N^2 \boldsymbol{\beta}_N^2| \geq c_N^2 \left(\mathrm{Im}\boldsymbol{\beta}_N\right)^2$$

which, in turn, implies (4.14). ■

5. Expression of matrix $\mathbb{E}\{\mathbf{Q}\}$ obtained using the integration by parts formula

We now express $\mathbb{E}\{\mathbf{Q}(z)\}$ using the integration by parts formula and deduce from this an approximate expression of $\mathbb{E}(Q(z))$. For this, we have first to establish some useful properties of $\mathbb{E}\{\mathbf{Q}(z)\}$ that follow from the invariance properties of the probability distribution of the observations $(y_n)_{n=1,\ldots,N}$. In the following, for $k, l \in \{1, 2, \ldots, L\}$, we denote by $\mathbf{Q}_{\mathbf{pp}}^{k,l}$ and $\mathbf{Q}_{\mathbf{ff}}^{k,l}$ the $M \times M$ matrices whose entries are given by $\left(\mathbf{Q}_{\mathbf{pp}}^{k,l}\right)_{m,n} = (\mathbf{Q}_{\mathbf{pp}})_{(k-1)M+m,(l-1)M+n}$ and $\left(\mathbf{Q}_{\mathbf{ff}}^{k,l}\right)_{m,n} = (\mathbf{Q}_{\mathbf{ff}})_{(k-1)M+m,(l-1)M+n}$for each $m, n \in \{1, 2, \ldots, M\}$.

Lemma 5.1. *The matrices $\mathbb{E}\{\mathbf{Q}_{\mathbf{pp}}\}$ and $\mathbb{E}\{\mathbf{Q}_{\mathbf{ff}}\}$ are block diagonal, i.e. $\mathbb{E}\left(\mathbf{Q}_{\mathbf{pp}}^{k,l}\right) = \mathbb{E}\{\mathbf{Q}_{\mathbf{ff}}^{k,l}\} = 0$ if $k \neq l$, and*

$$\mathrm{Tr}\mathbb{E}\{\mathbf{Q}_{\mathbf{pp}}\}(I_L \otimes R) = \mathrm{Tr}\mathbb{E}\{\mathbf{Q}_{\mathbf{ff}}\}(I_L \otimes R), \tag{5.1}$$

$$\mathbb{E}\{\mathbf{Q}_{\mathbf{pf}}\} = \mathbb{E}\{\mathbf{Q}_{\mathbf{fp}}\} = 0. \tag{5.2}$$

Proof. To prove (5.2), we consider the new set of vectors $z_k = e^{-ik\theta} y_k$ and construct the matrices Z_p, Z_f in the same way as Y_p and Y_f. It is clear that sequence $(z_n)_{n\in\mathbb{Z}}$ has the same probability distribution that $(y_n)_{n\in\mathbb{Z}}$. Z_p and Z_f can be expressed as

$$Z_p = \begin{pmatrix} e^{-i\theta} I_M & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & e^{-Li\theta} I_M \end{pmatrix} Y_p \begin{pmatrix} 1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & e^{-(N-1)i\theta} \end{pmatrix},$$

$$Z_f = e^{-Li\theta} \begin{pmatrix} e^{-i\theta} I_M & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & e^{-Li\theta} I_M \end{pmatrix} Y_f \begin{pmatrix} 1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & e^{-(N-1)i\theta} \end{pmatrix}.$$

Therefore, it holds that

$$Z_f Z_p^* Z_p Z_f^* = \begin{pmatrix} e^{-i\theta} I_M & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & e^{-Li\theta} I_M \end{pmatrix} Y_f Y_p^* Y_p Y_f^* \begin{pmatrix} e^{i\theta} I_M & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & e^{Li\theta} I_M \end{pmatrix}.$$

Similarly to $\mathbf{Q}$ we define matrix $\mathbf{Q}^{\mathbf{Z}} = \begin{pmatrix} -zI_{ML} & \frac{1}{N} Z_f Z_p^* \\ \frac{1}{N} Z_p Z_f^* & -zI_{ML} \end{pmatrix}^{-1}$ and obtain immediately that

$$\mathbb{E}\{\mathbf{Q}^{\mathbf{Z}}_{\mathbf{pp}}\} = \begin{pmatrix} e^{-i\theta} I_M & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{-Li\theta} I_M \end{pmatrix} \mathbb{E}\{\mathbf{Q}_{\mathbf{pp}}\} \begin{pmatrix} e^{i\theta} I_M & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{Li\theta} I_M \end{pmatrix}.$$

Since $\mathbb{E}\{\mathbf{Q}^{\mathbf{Z}}_{\mathbf{pp}}\} = \mathbb{E}\{\mathbf{Q}_{\mathbf{pp}}\}$, then for any $M \times M$ block $\mathbb{E}\{\mathbf{Q}_{\mathbf{pp}}{}^{j,k}\}$, we have

$$\mathbb{E}\{\mathbf{Q}_{\mathbf{pp}}{}^{j,k}\} = e^{-ji\theta}\mathbb{E}\{\mathbf{Q}_{\mathbf{pp}}{}^{j,k}\}e^{ki\theta} = e^{(k-j)i\theta}\mathbb{E}\{\mathbf{Q}_{\mathbf{pp}}{}^{j,k}\}.$$

This proves that $\mathbb{E}\{\mathbf{Q}_{\mathbf{pp}}{}^{j,k}\} = 0$ if $k \neq j$ as expected. A similar proof leads to the conclusion that $\mathbb{E}\{\mathbf{Q}_{\mathbf{ff}}\}$ is block diagonal. Moroever, the equality $\mathbb{E}\{\mathbf{Q}^{\mathbf{Z}}_{\mathbf{fp}}\} = \mathbb{E}\{\mathbf{Q}_{\mathbf{fp}}\}$ implies that

$$\mathbb{E}\{\mathbf{Q}^{\mathbf{Z}}_{\mathbf{fp}}\} = e^{-Li\theta} \begin{pmatrix} e^{-i\theta} I_M & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{-Li\theta} I_M \end{pmatrix} \mathbb{E}\{\mathbf{Q}_{\mathbf{fp}}\} \begin{pmatrix} e^{i\theta} I_M & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{Li\theta} I_M \end{pmatrix}.$$

Therefore, each $M \times M$ block $\mathbf{Q}_{\mathbf{fp}}{}^{j,k}$ of $\mathbf{Q}_{\mathbf{fp}}$ verifies $\mathbb{E}\{\mathbf{Q}_{\mathbf{fp}}{}^{j,k}\} = e^{-(L+j-k)i\theta}\mathbb{E}\{\mathbf{Q}_{\mathbf{fp}}{}^{j,k}\}$. As $j-k \in \{-(L-1), \dots, L-1\}$, this implies that $\mathbb{E}\{\mathbf{Q}_{\mathbf{fp}}{}^{j,k}\} = 0$. This leads immediately to $\mathbb{E}\{\mathbf{Q}_{\mathbf{fp}}\} = 0$. We obtain similarly that $\mathbb{E}\{\mathbf{Q}_{\mathbf{pf}}\} = 0$.

To prove (5.1) we consider the sequence z defined by $z_n = y_{-n+N+2L}$ for each n. Again, the distribution of z_n will remain the same and it is easy to see that Z_p and Z_f are given by

$$Z_f = \begin{pmatrix} 0 & \dots & I_M \\ \vdots & & \vdots \\ I_M & \dots & 0 \end{pmatrix} Y_p \begin{pmatrix} 0 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 0 \end{pmatrix},$$

$$Z_p = \begin{pmatrix} 0 & \dots & I_M \\ \vdots & & \vdots \\ I_M & \dots & 0 \end{pmatrix} Y_f \begin{pmatrix} 0 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 0 \end{pmatrix}.$$

From this, we obtain that

$$\mathbb{E}\{\mathbf{Q}^{\mathbf{Z}}_{\mathbf{pp}}\} = \begin{pmatrix} 0 & \dots & I_M \\ \vdots & & \vdots \\ I_M & \dots & 0 \end{pmatrix} \mathbb{E}\{\mathbf{Q}_{\mathbf{ff}}\} \begin{pmatrix} 0 & \dots & I_M \\ \vdots & & \vdots \\ I_M & \dots & 0 \end{pmatrix}.$$

As $\mathbb{E}\{\mathbf{Q}^{\mathbf{Z}}_{\mathbf{pp}}\} = \mathbb{E}\{\mathbf{Q}_{\mathbf{pp}}\}$, this immediately implies that $\mathbb{E}\{\mathbf{Q}_{\mathbf{ff}}{}^{j,j}\} = \mathbb{E}\{\mathbf{Q}_{\mathbf{pp}}{}^{L-j,L-j}\}$, and, as a consequence, that $\mathbb{E}\{\mathrm{Tr}\mathbf{Q}_{\mathbf{pp}}(I_L \otimes R)\} = \mathbb{E}\{\mathrm{Tr}\mathbf{Q}_{\mathbf{ff}}(I_L \otimes R)\}$, as expected. ■

In order to present the following approximation of $\mathbb{E}(Q_N(z))$, we introduce some useful notations. $\alpha_N(z)$ is the function defined by

$$\alpha_N(z) = \frac{1}{ML}\mathrm{Tr}\left(\mathbb{E}\{Q_N(z)(I_L \times R_N)\}\right). \tag{5.3}$$

α_N is clearly an element of $\mathcal{S}(\mathbb{R}^+)$. In order to evaluate its associated positive measure $\overline{\mu}_N$, we denote by $\hat{\mu}_N$ the positive measure defined by

$$d\hat{\mu}_N(\lambda) = \frac{1}{ML} \sum_{i=1}^{ML} \hat{f}_i^* (I_L \otimes R) \hat{f}_i \, \delta_{\hat{\lambda}_i}, \tag{5.4}$$

where we recall that $(\hat{\lambda}_i)_{i=1,\ldots,ML}$ and $(\hat{f}_i)_{i=1,\ldots,ML}$ represent the eigenvalues and eigenvectors of $W_f W_p^* W_p W_f^*$. We remark that $\hat{\mu}_N$ is carried by $\mathbb{R}^+$ and that its mass $\hat{\mu}_N(\mathbb{R}^+)$ coincides with $\frac{1}{M}\mathrm{Tr} R_N$. Then, measure $\overline{\mu}_N$ is defined by

$$\int_{\mathbb{R}^+} \phi(\lambda) \, d\overline{\mu}_N(\lambda) = \mathbb{E}\left(\int_{\mathbb{R}^+} \phi(\lambda) \, d\hat{\mu}_N(\lambda) \right) \tag{5.5}$$

and satisfies $\overline{\mu}_N(\mathbb{R}^+) = \frac{1}{M}\mathrm{Tr} R_N$. We also define $\boldsymbol{\alpha}_N(z)$ as the function

$$\boldsymbol{\alpha}_N(z) = z\alpha_N(z^2) \tag{5.6}$$

which, due to the identity $\mathbf{Q}_{\mathbf{pp}}(z) = zQ(z^2)$, is also given by

$$\boldsymbol{\alpha}_N(z) = \frac{1}{ML} \mathbb{E}\left\{ \mathrm{Tr} \mathbf{Q}_{\mathbf{N},\mathbf{pp}}(z)(I_L \otimes R_N) \right\}. \tag{5.7}$$

Lemma 4.1 implies that $\boldsymbol{\alpha}_N \in \mathcal{S}(\mathbb{R})$ and that the $M \times M$ matrix-valued functions $S_N(z)$ and $\mathbf{S}_N(z)$ defined by

$$S_N(z) = -\left(zI_M + \frac{c_N z \alpha_N(z)}{1 - c_N^2 z \alpha_N(z)^2} R_N \right)^{-1} \tag{5.8}$$

and

$$\mathbf{S}_N(z) = -\left(\frac{c_N \boldsymbol{\alpha}(z)}{1 - c_N^2 \boldsymbol{\alpha}^2(z)} R + z \right)^{-1} = z S_N(z^2) \tag{5.9}$$

belong to $\mathcal{S}_M(\mathbb{R}^+)$ and $\mathcal{S}_M(\mathbb{R})$ respectively. We are now in position to introduce the main result of this section.

Theorem 5.1. *The matrix* $\mathbb{E}(Q_N(z))$ *can be written as*

$$\mathbb{E}\{Q_N(z)\} = I_L \otimes S_N(z) - E_N(z) \, (I_L \otimes S_N(z)), \tag{5.10}$$

where $E_N(z)$ *is an error term such that*

$$\left| \frac{1}{ML} \mathrm{Tr} E_N(z) F_N \right| \leq \kappa \, \frac{1}{N^2} \, P_1(|z|) P_2\left(\frac{1}{\mathrm{Im} z}\right) \tag{5.11}$$

for each $z \in \mathbb{C}^+$ *and for each deterministic* $ML \times ML$ *sequence of matrices* $(F_N)_{N \geq 1}$ *such that* $\sup_{N \geq 1} \|F_N\| \leq \kappa$.

In order to establish Theorem 5.1, we express $\mathbb{E}\{\mathbf{Q}(z)\}$ for $z \in \mathbb{C}^+$ by using the integration by parts formula (see Proposition 2.1), and deduce from that the expression (5.10) of $\mathbb{E}\{Q(z)\}$. The properties of the error term $E_N(z)$ is finally deduced from the results of section 3.

We recall that matrix $\mathbf{M}$ is defined by (2.3). In order to express $\mathbb{E}\{\mathbf{Q}(z)\}$ for $z \in \mathbb{C}^+$, we use the identity

$$z\mathbf{Q}(z) = -I_{2ML} + \mathbf{Q}(z)\mathbf{M} = -I_{2ML} + \sum_{j=1}^{N} \mathbf{Q}(z) \begin{pmatrix} 0 & w_{f,j} w_{p,j}^* \\ w_{p,j} w_{f,j}^* & 0 \end{pmatrix}. \tag{5.12}$$

For every $m_1, m_2 = 1, \ldots, M$, $i_1 = 1, \ldots, 2L$ and $i_2 = 1, \ldots, L$ we denote by $\hat{\mathbf{A}}_{i_1 i_2}^{m_1 m_2}$ the $2N \times 2N$ matrix defined by

$$\hat{\mathbf{A}}_{i_1 i_2}^{m_1 m_2} = \begin{pmatrix} \hat{\mathbf{A}}_{i_1 i_2}^{m_1 m_2}(pp) & \hat{\mathbf{A}}_{i_1 i_2}^{m_1 m_2}(pf) \\ \hat{\mathbf{A}}_{i_1 i_2}^{m_1 m_2}(fp) & \hat{\mathbf{A}}_{i_1 i_2}^{m_1 m_2}(ff) \end{pmatrix}, \tag{5.13}$$

where the 4 $N \times N$ blocks are given by

$$\begin{aligned} (\hat{\mathbf{A}}_{i_1 i_2}^{m_1 m_2}(pf))_{jk} &= \left(\mathbf{Q} \left(\begin{smallmatrix} 0 \\ w_{p,j} \end{smallmatrix} \right)\right)_{i_1}^{m_1} (w_{f,k}^*)_{i_2}^{m_2}, \\ (\hat{\mathbf{A}}_{i_1 i_2}^{m_1 m_2}(pp))_{jk} &= \left(\mathbf{Q} \left(\begin{smallmatrix} 0 \\ w_{p,j} \end{smallmatrix} \right)\right)_{i_1}^{m_1} (w_{p,k}^*)_{i_2}^{m_2}, \\ (\hat{\mathbf{A}}_{i_1 i_2}^{m_1 m_2}(ff))_{jk} &= \left(\mathbf{Q} \left(\begin{smallmatrix} w_{f,j} \\ 0 \end{smallmatrix} \right)\right)_{i_1}^{m_1} (w_{f,k}^*)_{i_2}^{m_2}, \\ (\hat{\mathbf{A}}_{i_1 i_2}^{m_1 m_2}(fp))_{jk} &= \left(\mathbf{Q} \left(\begin{smallmatrix} w_{f,j} \\ 0 \end{smallmatrix} \right)\right)_{i_1}^{m_1} (w_{p,k}^*)_{i_2}^{m_2}. \end{aligned} \tag{5.14}$$

We also define matrix $\mathbf{A}_{i_1 i_2}^{m_1 m_2}$ by $\mathbf{A}_{i_1 i_2}^{m_1 m_2} = \mathbb{E}\{\hat{\mathbf{A}}_{i_1 i_2}^{m_1 m_2}\}$. (5.12) implies that

$$z\mathbb{E}\{\mathbf{Q}_{i_1 i_2}^{m_1 m_2}(z)\} = -\delta_{i_1,i_2}\delta_{m_1,m_2} + \mathrm{Tr}\mathbf{A}_{i_1 i_2}^{m_1 m_2}(pf)\mathbf{1}_{i_2 \leq L} + \mathrm{Tr}\mathbf{A}_{i_1 i_2 - L}^{m_1 m_2}(fp)\mathbf{1}_{i_2 > L}. \tag{5.15}$$

In the reminder of this paragraph, we evaluate for each i_1, i_2, m_1, m_2 the elements of matrix $\mathbf{A}_{i_1 i_2}^{m_1 m_2}$ using (2.9) and (3.12). As we shall see, each element of $\mathbf{A}_{i_1 i_2}^{m_1 m_2}$ can be written as a functional of matrix $\mathbb{E}\{\mathbf{Q}\}$ plus an error term whose contribution vanishes when $N \to +\infty$. Plugging these expressions of $\mathbf{A}_{i_1 i_2}^{m_1 m_2}$ into (5.15) will establish an approximate expression of $\mathbb{E}\{\mathbf{Q}\}$. As the calculations are very tedious, we just indicate how each element $(\mathbf{A}_{i_1 i_2}^{m_1 m_2}(ff))_{j,k}$ of matrix $\mathbf{A}_{i_1 i_2}^{m_1 m_2}(ff)$ can be evaluated. By using integration by parts formula (2.9) and (3.12) we obtain

$$\mathbb{E}\left\{\left(\mathbf{Q}\begin{pmatrix} w_{f,j} \\ 0 \end{pmatrix}\right)_{i_1}^{m_1}(w^*_{f,k})^{m_2}_{i_2}\right\} = \sum_{i_3=1}^{L}\sum_{m_3}\mathbb{E}\{\mathbf{Q}^{m_1m_3}_{i_1i_3}W^{m_3}_{i_3+L,j}\overline{W}^{m_2}_{i_2+L,k}\}$$

$$= \sum_{i_3=1}^{L}\sum_{\substack{i',j' \\ m',m_3}}\mathbb{E}\{W^{m_3}_{i_3+L,j}\overline{W}^{m'}_{i',j'}\}\times\mathbb{E}\left\{\frac{\partial\left(\mathbf{Q}^{m_1m_3}_{i_1i_3}\overline{W}^{m_2}_{i_2+L,k}\right)}{\partial\overline{W}^{m'}_{i',j'}}\right\} = \frac{1}{N}\sum_{i_3=1}^{L}\sum_{\substack{i',j' \\ m',m_3}}R_{m_3m'}$$

$$\times\,\delta_{i_3+L+j,i'+j'}\mathbb{E}\left\{\mathbf{Q}^{m_1m_3}_{i_1i_3}\delta_{m_2,m'}\delta_{i_2+L,i'}\delta_{k,j'}+\overline{W}^{m_2}_{i_2+L,k}\frac{\partial\mathbf{Q}^{m_1m_3}_{i_1i_3}}{\partial\overline{W}^{m'}_{i',j'}}\right\}$$

$$= \frac{1}{N}\sum_{i_3=1}^{L}\sum_{m_3=1}^{M}\mathbb{E}\left\{\mathbf{Q}^{m_1m_3}_{i_1i_3}R_{m_3m_2}\delta_{i_3,i_2-(j-k)}\right\} - \frac{1}{N}\sum_{\substack{i_3,j' \\ m_3,m'}}\sum_{i'=1}^{L}R_{m_3m'}\delta_{i_3+L+j,i'+j'}$$

$$\times\,\mathbb{E}\left\{\overline{W}^{(f)m_2}_{i_2,k}\left(\mathbf{Q}\left(\begin{smallmatrix} w_{f,j'} \\ 0 \end{smallmatrix}\right)\right)^{m_1}_{i_1}\mathbf{Q}^{m'm_3}_{i'+Li_3}\right\} - \frac{1}{N}\sum_{\substack{i_3,j' \\ m_3,m'}}\sum_{i'=L+1}^{2L}R_{m_3m'}\delta_{i_3+L+j,i'+j'}$$

$$\times\,\mathbb{E}\left\{\overline{W}^{(f)m_2}_{i_2,k}\left(\mathbf{Q}\left(\begin{smallmatrix} 0 \\ w_{p,j'} \end{smallmatrix}\right)\right)^{m_1}_{i_1}\mathbf{Q}^{m'm_3}_{i'-Li_3}\right\} = \frac{1}{N}\sum_{i_3=1}^{L}\mathbb{E}\Big\{\left(\left(\begin{smallmatrix} \mathbf{Q_{pp}} \\ \mathbf{Q_{fp}} \end{smallmatrix}\right)(I_L\otimes R)\right)^{m_1m_2}_{i_1i_3}$$

$$\times\delta_{i_3,i_2-(j-k)}\Big\} - \frac{1}{N}\sum_{m',j'}\sum_{i_3,i'=1}^{L}\delta_{i_3+L+j,i'+j'}\mathbb{E}\left\{\left(\hat{A}^{m_1m_2}_{i_1i_2}(ff)\right)_{j',k}(\mathbf{Q_{fp}}(I_L\otimes R))^{m'm'}_{i'i_3}\right\}$$

$$-\frac{1}{N}\sum_{m',j'}\sum_{i_3,i'=1}^{L}\delta_{i_3+j,i'+j'}\mathbb{E}\left\{\left(\hat{\mathbf{A}}^{m_1m_2}_{i_1i_2}(pf)\right)_{j',k}(\mathbf{Q_{pp}}(I_L\otimes R))^{m'm'}_{i'i_3}\right\}.$$

Now we define for every $i_1 = 1,\dots,2L$, $i_2 = 1,\dots,L$ and $m_1,m_2 = 1,\dots,M$ the $2N\times 2N$ matrix $\mathbf{B}^{m_1m_2}_{i_1i_2}$ with $N\times N$ blocks

$$\left(\mathbf{B}^{m_1m_2}_{i_1i_2}(fp)\right)_{j,k} = \frac{1}{N}\mathbb{E}\left\{\left(\begin{smallmatrix} \mathbf{Q_{pp}} \\ \mathbf{Q_{fp}} \end{smallmatrix}\right)(I_L\otimes R)\right\}^{m_1,m_2}_{i_1,i_2-(j-k)-L}\mathbf{1}_{1\le i_2-(j-k)-L\le L},$$

$$\left(\mathbf{B}^{m_1m_2}_{i_1i_2}(ff)\right)_{j,k} = \frac{1}{N}\mathbb{E}\left\{\left(\begin{smallmatrix} \mathbf{Q_{pp}} \\ \mathbf{Q_{fp}} \end{smallmatrix}\right)(I_L\otimes R)\right\}^{m_1,m_2}_{i_1,i_2-(j-k)}\mathbf{1}_{1\le i_2-(j-k)\le L},$$

$$\left(\mathbf{B}^{m_1m_2}_{i_1i_2}(pp)\right)_{j,k} = \frac{1}{N}\mathbb{E}\left\{\left(\begin{smallmatrix} \mathbf{Q_{pf}} \\ \mathbf{Q_{ff}} \end{smallmatrix}\right)(I_L\otimes R)\right\}^{m_1,m_2}_{i_1,i_2-(j-k)}\mathbf{1}_{1\le i_2-(j-k)\le L},$$

$$\left(\mathbf{B}^{m_1m_2}_{i_1i_2}(pf)\right)_{j,k} = \frac{1}{N}\mathbb{E}\left\{\left(\begin{smallmatrix} \mathbf{Q_{pf}} \\ \mathbf{Q_{ff}} \end{smallmatrix}\right)(I_L\otimes R)\right\}^{m_1,m_2}_{i_1,i_2-(j-k)+L}\mathbf{1}_{1\le i_2-(j-k)+L\le L}.$$

For every $ML\times ML$ block matrix $\mathbf{D}$, we define the sequence $(\tau^{(M)}(\mathbf{D})(l))_{l=-L+1,\dots,L-1}$ as

$$\tau^{(M)}(\mathbf{D})(l) = \frac{1}{ML}\mathrm{Tr}\mathbf{D}(J_L^{\epsilon(l)}\otimes I_M) = \frac{1}{ML}\sum_{m=1}^{M}\sum_{i-i'=l}\mathbf{D}^{m,m}_{i,i'} \qquad (5.16)$$

and the $N \times N$ Toeplitz matrix $\mathcal{T}_{N,L}^{(M)}(\mathbf{D})$ given by

$$\mathcal{T}_{N,L}^{(M)}(\mathbf{D}) = \sum_{l=-L+1}^{L-1} \tau^{(M)}(\mathbf{D})(l) J_N^{*\epsilon(l)}. \tag{5.17}$$

In other words, the entries of $\mathcal{T}_{N,L}^{(M)}(\mathbf{D})$ are defined by the relation

$$\left[\mathcal{T}_{N,L}^{(M)}(\mathbf{D})\right]_{j_1,j_2} = \tau^{(M)}(\mathbf{D})(j_1 - j_2)\, \mathbf{1}_{-(L-1)\leq j_1-j_2 \leq L-1}. \tag{5.18}$$

We observe that if $\mathbf{D}$ is block diagonal, i.e. if $\mathbf{D}_{i_1,i_2}^{m_1,m_2} = 0$ for each m_1, m_2 when $i_1 \neq i_2$, then, matrix $\mathcal{T}_{N,L}^{(M)}(\mathbf{D})$ coincides with the diagonal matrix $\mathcal{T}_{N,L}^{(M)}(\mathbf{D}) = \left(\frac{1}{ML}\mathrm{Tr}\mathbf{D}\right) I_N$. It clear that

$$\frac{1}{N}\sum_{i_3=1}^{L} \mathbb{E}\left\{ \left(\begin{pmatrix} \mathbf{Q_{pp}} \\ \mathbf{Q_{fp}} \end{pmatrix} (I_L \otimes R) \right)_{i_1 i_3}^{m_1 m_2} \delta_{i_3, i_2-(j-k)} \right\} = \left(\mathbf{B}_{i_1 i_2}^{m_1 m_2}(ff)\right)_{j,k}.$$

In order to rewrite the term

$$\frac{1}{N}\sum_{m',j'} \sum_{i_3,i'=1}^{L} \delta_{i_3+L+j, i'+j'} \times \mathbb{E}\left\{ \left(\hat{A}_{i_1 i_2}^{m_1 m_2}(ff)\right)_{j',k} (\mathbf{Q_{fp}}(I_L \otimes R))_{i' i_3}^{m' m'} \right\}$$

in a more convenient way, we put $l = i' - i_3$, and remark that

$$\frac{1}{N}\sum_{m',j'} \sum_{i_3,i'=1}^{L} \delta_{i_3+L+j, i'+j'} \times \mathbb{E}\left\{ \left(\hat{A}_{i_1 i_2}^{m_1 m_2}(ff)\right)_{j',k} (\mathbf{Q_{fp}}(I_L \otimes R))_{i' i_3}^{m' m'} \right\} =$$
$$\frac{ML}{N}\sum_{m'} \sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{ \left(\hat{A}_{i_1 i_2}^{m_1 m_2}(ff)\right)_{L+j-l,k} \frac{1}{ML} \sum_{i'-i_3=l} (\mathbf{Q_{fp}}(I_L \otimes R))_{i' i_3}^{m' m'} \right\}.$$

Using the definition (5.16), this can be rewritten as

$$\frac{1}{N}\sum_{m',j'} \sum_{i_3,i'=1}^{L} \delta_{i_3+L+j, i'+j'} \times \mathbb{E}\left\{ \left(\hat{A}_{i_1 i_2}^{m_1 m_2}(ff)\right)_{j',k} (\mathbf{Q_{fp}}(I_L \otimes R))_{i' i_3}^{m' m'} \right\} =$$
$$c_N \sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{ \left(\hat{A}_{i_1 i_2}^{m_1 m_2}(ff)\right)_{L+j-l,k} \tau^M \left(\mathbf{Q_{fp}}(I_L \otimes R)\right)(l) \right\}.$$

We introduce $j' = L + j - l$, and using (5.18), we notice that

$$\frac{1}{N}\sum_{m',j'} \sum_{i_3,i'=1}^{L} \delta_{i_3+L+j, i'+j'} \times \mathbb{E}\left\{ \left(\hat{A}_{i_1 i_2}^{m_1 m_2}(ff)\right)_{j',k} (\mathbf{Q_{fp}}(I_L \otimes R))_{i' i_3}^{m' m'} \right\} =$$
$$c_N\, \mathbb{E}\left\{ \sum_{j'=1}^{N} \left[\mathcal{T}_{N,L}^{(M)}\left(\mathbf{Q_{fp}}(I_L \otimes R)\right)\right]_{L+j,j'} \left(\hat{A}_{i_1 i_2}^{m_1 m_2}(ff)\right)_{j',k} \right\} =$$
$$c_N \mathbb{E}\left\{ \left(J_N^L \mathcal{T}_{N,L}^{(M)}(\mathbf{Q_{fp}}(I_L \otimes R)) \hat{\mathbf{A}}_{i_1 i_2}^{m_1 m_2}(ff) \right)_{j,k} \right\}.$$

We obtain similarly that

$$\frac{1}{N}\sum_{m',j'}\sum_{i_3,i'=1}^{L}\delta_{i_3+j,i'+j'}\mathbb{E}\left\{\left(\hat{\mathbf{A}}_{i_1i_2}^{m_1m_2}(pf)\right)_{j',k}(\mathbf{Q_{pp}}(I_L\otimes R))_{i'i_3}^{m'm'}\right\}=$$
$$c_N\mathbb{E}\left\{\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q_{pp}}(I_L\otimes R))\hat{\mathbf{A}}_{i_1i_2}^{m_1m_2}(pf)\right)_{j,k}\right\}.$$

Therefore, matrix $\mathbf{A}_{i_1i_2}^{m_1m_2}(ff)$ is also given by

$$\left(\mathbf{A}_{i_1i_2}^{m_1m_2}(ff)\right)_{j,k}=\left(\mathbf{B}_{i_1i_2}^{m_1m_2}(ff)\right)_{j,k}-c_N\mathbb{E}\left\{\left(J_N^L\mathcal{T}_{N,L}^{(M)}(\mathbf{Q_{fp}}(I_L\otimes R))\hat{\mathbf{A}}_{i_1i_2}^{m_1m_2}(ff)\right)_{j,k}\right\}$$
$$-c_N\mathbb{E}\left\{\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q_{pp}}(I_L\otimes R))\hat{\mathbf{A}}_{i_1i_2}^{m_1m_2}(pf)\right)_{j,k}\right\}.$$

Writing $\mathbf{Q_{fp}}$ and $\mathbf{Q_{pp}}$ as $\mathbf{Q_{fp}}=\mathbb{E}\{\mathbf{Q_{fp}}\}+\mathbf{Q}_{\mathbf{fp}}^{\circ}=\mathbf{Q}_{\mathbf{fp}}^{\circ}$ (see (5.2)) and $\mathbf{Q_{pp}}=\mathbb{E}\{\mathbf{Q_{pp}}\}+\mathbf{Q}_{\mathbf{pp}}^{\circ}$, we obtain that

$$\left(\mathbf{A}_{i_1i_2}^{m_1m_2}(ff)\right)_{j,k}=\left(\mathbf{B}_{i_1i_2}^{m_1m_2}(ff)\right)_{j,k}-c_N\mathbb{E}\left\{\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q_{pp}}(I_L\otimes R))\mathbf{A}_{i_1i_2}^{m_1m_2}(pf)\right)_{j,k}\right\}$$
$$-c_N\mathbb{E}\left\{\left(J_N^L\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}_{\mathbf{fp}}^{\circ}(I_L\otimes R))\hat{\mathbf{A}}_{i_1i_2}^{m_1m_2}(ff)\right)_{j,k}\right\}$$
$$-c_N\mathbb{E}\left\{\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}_{\mathbf{pp}}^{\circ}(I_L\otimes R))\hat{\mathbf{A}}_{i_1i_2}^{m_1m_2}(pf)\right)_{j,k}\right\}.$$

We define the $N\times N$ matrix $\mathbf{\Delta}_{i_1i_2}^{m_1m_2}(ff)$ by

$$\mathbf{\Delta}_{i_1i_2}^{m_1m_2}(ff)=-c_N\mathbb{E}\left\{J_N^L\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}_{\mathbf{fp}}^{\circ}(I_L\otimes R))\hat{\mathbf{A}}_{i_1i_2}^{m_1m_2}(ff)\right\}$$
$$-c_N\mathbb{E}\left\{\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}_{\mathbf{pp}}^{\circ}(I_L\otimes R))\hat{\mathbf{A}}_{i_1i_2}^{m_1m_2}(pf)\right\}.$$

Dropping the indices i_1, i_2, m_1, m_2, we eventually obtain that

$$\mathbf{A_{ff}}=\mathbf{B_{ff}}-c_N\mathbb{E}\left\{\mathcal{T}_{N,L}^{(M)}(\mathbf{Q_{pp}}(I_L\otimes R))\right\}\mathbf{A_{pf}}+\mathbf{\Delta_{ff}}.$$

Using similar calculations, it is possible to establish that:

$$\mathbf{A_{pf}}=\mathbf{B_{pf}}-c_N\mathbb{E}\left\{\mathcal{T}_{N,L}^{(M)}(\mathbf{Q_{ff}}(I_L\otimes R))\right\}\mathbf{A_{ff}}+\mathbf{\Delta_{pf}},$$
$$\mathbf{A_{fp}}=\mathbf{B_{fp}}-c_N\mathbb{E}\left\{\mathcal{T}_{N,L}^{(M)}(\mathbf{Q_{pp}}(I_L\otimes R))\right\}\mathbf{A_{pp}}+\mathbf{\Delta_{fp}},$$
$$\mathbf{A_{pp}}=\mathbf{B_{pp}}-c_N\mathbb{E}\left\{\mathcal{T}_{N,L}^{(M)}(\mathbf{Q_{ff}}(I_L\otimes R))\right\}\mathbf{A_{fp}}+\mathbf{\Delta_{pp}},$$

where $\mathbf{\Delta_{pf}}$, $\mathbf{\Delta_{fp}}$, and $\mathbf{\Delta_{pp}}$ are defined as

$$\mathbf{\Delta_{pf}}=-c_N\mathbb{E}\left\{\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}_{\mathbf{pf}}^{\circ}(I_L\otimes R))J_N^{*L}\hat{\mathbf{A}}_{\mathbf{pf}}\right\}-c_N\mathbb{E}\left\{\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}_{\mathbf{ff}}^{\circ}(I_L\otimes R))\hat{\mathbf{A}}_{\mathbf{ff}}\right\},$$
$$\mathbf{\Delta_{fp}}=-c_N\mathbb{E}\left\{J_N^L\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}_{\mathbf{fp}}^{\circ}(I_L\otimes R))\hat{\mathbf{A}}_{\mathbf{fp}}\right\}-c_N\mathbb{E}\left\{\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}_{\mathbf{pp}}^{\circ}(I_L\otimes R))\hat{\mathbf{A}}_{\mathbf{pp}}\right\},$$
$$\mathbf{\Delta_{pp}}=-c_N\mathbb{E}\left\{\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}_{\mathbf{pf}}^{\circ}(I_L\otimes R))J_N^{*L}\hat{\mathbf{A}}_{\mathbf{pp}}\right\}-c_N\mathbb{E}\left\{\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}_{\mathbf{ff}}^{\circ}(I_L\otimes R))\hat{\mathbf{A}}_{\mathbf{fp}}\right\}.$$

By Lemma 5.1, matrices $\mathbb{E}\{\mathbf{Q_{ff}}\}$ and $\mathbb{E}\{\mathbf{Q_{pp}}\}$ are block diagonal. Therefore, matrices $\mathbb{E}\{\mathcal{T}_{N,L}^{(M)}(\mathbf{Q_{ff}}(I_L \otimes R))\}$ and $\mathbb{E}\{\mathcal{T}_{N,L}^{(M)}(\mathbf{Q_{pp}}(I_L \otimes R))\}$ reduce to $\frac{1}{ML}\mathbb{E}\{\mathrm{Tr}\mathbf{Q_{ff}}(I_L \otimes R)\}\, I_N$ and $\frac{1}{ML}\mathbb{E}\{\mathrm{Tr}\mathbf{Q_{pp}}(I_L \otimes R)\}\, I_N$ respectively. As $\mathbb{E}\{\mathrm{Tr}\mathbf{Q_{ff}}(I_L \otimes R)\} = \mathbb{E}\{\mathrm{Tr}\mathbf{Q_{pp}}(I_L \otimes R)\}$ (see (5.1)), we eventually obtain that

$$\begin{pmatrix} I_N & \frac{c_N}{ML}\mathbb{E}\left\{\mathrm{Tr}\mathbf{Q_{pp}}(I_L \otimes R)\right\} I_N \\ \frac{c_N}{ML}\mathbb{E}\left\{\mathrm{Tr}\mathbf{Q_{pp}}(I_L \otimes R)\right\} I_N & I_N \end{pmatrix} \mathbf{A} = \mathbf{B} + \boldsymbol{\Delta}. \quad (5.19)$$

Using (5.7), this can be written as

$$\begin{pmatrix} I_N & c_N\, \boldsymbol{\alpha}_N\, I_N \\ c_N\, \boldsymbol{\alpha}_N\, I_N & I_N \end{pmatrix} \mathbf{A} = \mathbf{B} + \boldsymbol{\Delta}.$$

Lemma 4.1 implies that

$$1 - (c_N \boldsymbol{\alpha}(z))^2 \neq 0$$

if $z \in \mathbb{C}^+$. This implies that the matrix governing the linear system (5.19) is invertible for $z \in \mathbb{C}^+$. Matrix $\mathbf{H}$ given by

$$\mathbf{H} = \begin{pmatrix} I_N & c_N \boldsymbol{\alpha}(z)\, I_N \\ c_N \boldsymbol{\alpha}(z)\, I_N & I_N \end{pmatrix}^{-1}.$$

is thus well defined for each $z \in \mathbb{C}^+$. The blocks of $\mathbf{H}$ are of course given by

$$\mathbf{H_{pp}} = \mathbf{H_{ff}} = \frac{1}{1 - c_N^2 \boldsymbol{\alpha}(z)^2}\, I_N,$$
$$\mathbf{H_{pf}} = \mathbf{H_{fp}} = -\frac{c_N \boldsymbol{\alpha}(z)}{1 - c_N^2 \boldsymbol{\alpha}(z)^2}\, I_N.$$

(5.19) implies that $\mathbf{A} = \mathbf{HB} + \mathbf{H\Delta}$. (5.15) implies that we only need to evaluate matrices $\mathbf{A_{pf}}$ and $\mathbf{A_{fp}}$. We obtain that these matrices are given by

$$\mathbf{A_{pf}} = \mathbf{H_{pp}B_{pf}} + \mathbf{H_{pf}B_{ff}} + \mathbf{H_{pp}}\boldsymbol{\Delta}_{\mathbf{pf}} + \mathbf{H_{pf}}\boldsymbol{\Delta}_{\mathbf{ff}},$$
$$\mathbf{A_{fp}} = \mathbf{H_{fp}B_{pp}} + \mathbf{H_{ff}B_{fp}} + \mathbf{H_{fp}}\boldsymbol{\Delta}_{\mathbf{pp}} + \mathbf{H_{ff}}\boldsymbol{\Delta}_{\mathbf{fp}}.$$

This and definition (5.14) of matrix $\mathbf{A}_{i_1 i_2}^{m_1 m_2}$ lead immediately to

$$\left(\mathbb{E}\left\{\mathbf{Q}\begin{pmatrix} 0 & W_f W_p^* \\ W_p W_f^* & 0 \end{pmatrix}\right\}\right)_{i_1 i_2}^{m_1 m_2} = \mathrm{Tr}\mathbf{A}_{i_1 i_2}^{m_1 m_2}(pf)\mathbf{1}_{i_2 \leq L} + \mathrm{Tr}\mathbf{A}_{i_1 i_2 - L}^{m_1 m_2}(fp)\mathbf{1}_{i_2 > L} =$$
$$\frac{1}{1 - c_N^2 \boldsymbol{\alpha}^2}\mathrm{Tr}\Big(\mathbf{B_{pf}} - c_N \boldsymbol{\alpha}\mathbf{B_{ff}} + \boldsymbol{\Delta}_{\mathbf{pf}} - c_N \boldsymbol{\alpha}\boldsymbol{\Delta}_{\mathbf{ff}}\Big)_{i_1 i_2}^{m_1 m_2}\mathbf{1}_{i_2 \leq L}$$
$$+ \frac{1}{1 - c_N^2 \boldsymbol{\alpha}^2}\mathrm{Tr}\Big(\mathbf{B_{fp}} - c_N \boldsymbol{\alpha}\mathbf{B_{pp}} + \boldsymbol{\Delta}_{\mathbf{fp}} - c_N \boldsymbol{\alpha}\boldsymbol{\Delta}_{\mathbf{pp}}\Big)_{i_1 i_2 - L}^{m_1 m_2}\mathbf{1}_{i_2 > L}.$$

It is easy to notice that $\mathrm{Tr}\,(\mathbf{B_{fp}})_{i_1 i_2}^{m_1 m_2} = \mathrm{Tr}\,(\mathbf{B_{pf}})_{i_1 i_2}^{m_1 m_2} = 0$, and $\mathrm{Tr}\,(\mathbf{B_{pp}})_{i_1 i_2}^{m_1 m_2} = \mathbb{E}\{(\mathbf{Q}\Pi_{ff}(I_2 L \otimes R))_{i_1 i_2 + L}^{m_1 m_2}\}$, $\mathrm{Tr}\,(\mathbf{B_{ff}})_{i_1 i_2}^{m_1 m_2} = \mathbb{E}\{(\mathbf{Q}\Pi_{pp}(I_2 L \otimes R))_{i_1 i_2}^{m_1 m_2}\}$, where

$\Pi_{ff} = \left(\begin{smallmatrix} 0 & 0 \\ 0 & I_{ML} \end{smallmatrix}\right)$ and $\Pi_{pp} = \left(\begin{smallmatrix} I_{ML} & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Hence, using that $\mathbb{E}\{\mathbf{Q_{pf}}\} = \mathbb{E}\{\mathbf{Q_{fp}}\} = 0$, we obtain that

$$\left(\mathbb{E}\left\{\mathbf{Q}\left(\begin{smallmatrix} 0 & W_f W_p^* \\ W_p W_f^* & 0 \end{smallmatrix}\right)\right\}\right)_{i_1 i_2}^{m_1 m_2} = -\frac{c_N \boldsymbol{\alpha}}{1 - c_N^2 \boldsymbol{\alpha}^2}\Big(\mathbb{E}\{\mathbf{Q}\Pi_{pp}(I_{2L} \otimes R)\}$$

$$+\mathbb{E}\{\mathbf{Q}\Pi_{ff}(I_{2L} \otimes R)\}\Big)_{i_1 i_2}^{m_1 m_2} + \mathcal{E}_{i_1 i_2}^{m_1 m_2} = -\frac{c_N \boldsymbol{\alpha}}{1 - c_N^2 \boldsymbol{\alpha}^2}\Big(\mathbb{E}\{\mathbf{Q}(I_{2L} \otimes R)\}\Big)_{i_1 i_2}^{m_1 m_2} + \mathcal{E}_{i_1 i_2}^{m_1 m_2},$$

where $\mathcal{E}_{i_1 i_2}^{m_1 m_2}$ represents the remaining terms depending on the entries of matrix $\boldsymbol{\Delta}_{i_1 i_2}^{m_1 m_2}$. Using the identity (5.12), we obtain that

$$z\mathbb{E}\{\mathbf{Q}\} + I_{2ML} = \mathbb{E}\left\{\mathbf{Q}\left(\begin{smallmatrix} 0 & W_f W_p^* \\ W_p W_f^* & 0 \end{smallmatrix}\right)\right\} = -\frac{c_N \boldsymbol{\alpha}}{1 - c_N^2 \boldsymbol{\alpha}^2}\mathbb{E}\{\mathbf{Q}\}(I_{2L} \otimes R) + \mathcal{E}, \tag{5.20}$$

which immediately leads to

$$-\mathbb{E}\{\mathbf{Q}\}\left(\frac{c_N \boldsymbol{\alpha}}{1 - c_N^2 \boldsymbol{\alpha}^2}(I_{2L} \otimes R) + z\right) = I_{2ML} - \mathcal{E}$$

or, equivalently,

$$\mathbb{E}\{\mathbf{Q}\}\ (I_{2L} \otimes \mathbf{S})^{-1} = I_{2ML} - \mathcal{E},$$

where we recall that $\mathbf{S}$ is defined by (5.9). As $\mathbb{E}\{\mathbf{Q}\}$ is block diagonal, (5.20) implies that matrix $\mathcal{E}$ is also block diagonal, i.e. $\mathcal{E}_{\mathbf{fp}} = \mathcal{E}_{\mathbf{pf}} = 0$. Moreover, it holds that

$$\mathbb{E}\{\mathbf{Q}(z)\} = I_{2L} \otimes \mathbf{S}(z) - \mathcal{E}(z)\ (I_{2L} \otimes \mathbf{S}(z))\,. \tag{5.21}$$

This allows to evaluate $\mathbb{E}\{Q(z)\}$ by identification of the first diagonal blocks of the left and right hand sides of (5.21). We thus obtain immediately that

$$\mathbb{E}\{Q(z^2)\} = I_L \otimes S(z^2) - \mathcal{E}_{\mathbf{pp}}(z)\ \left(I_L \otimes S(z^2)\right) \tag{5.22}$$

for each $z \in \mathbb{C}^+$, where we recall that $S(z)$ is given by(5.8). Therefore, $\mathcal{E}_{\mathbf{pp}}(z)$ only depends on z^2. As the image of $\mathbb{C}^+$ by the transformation $z \to z^2$ is $\mathbb{C} - \mathbb{R}^+$, we obtain that $\mathcal{E}_{\mathbf{pp}}(z) = E(z^2)$ for some function E analytic in $\mathbb{C} - \mathbb{R}^+$. This discussion leads to

$$\mathbb{E}\{Q(z)\} = I_L \otimes S(z) - E(z)\ (I_L \otimes S(z)) \tag{5.23}$$

for each $z \in \mathbb{C} - \mathbb{R}^+$.

In the following, we prove (5.11). For this, we establish following result.

Proposition 5.1. *For each deterministic sequence of $ML \times ML$ matrices $(F_{1,N})_{N \geq 1}$ such that $\sup_{N \geq 1} \|F_{1,N}\| \leq \kappa$, then*

$$\left|\frac{1}{ML}\mathrm{Tr}(\mathcal{E}_{pp}(z)\, F_{1,N})\right| \leq \kappa\, \frac{1}{N^2}\, P_1(|z^2|) P_2\Big(\frac{1}{\mathrm{Im} z^2}\Big) \tag{5.24}$$

holds for each $z \in \mathbb{C}^+$ for which $\mathrm{Im} z^2 > 0$, where P_1 and P_2 are 2 nice polynomials.

Proof. We define F_N as the $2ML \times 2ML$ matrix $F_N = \begin{pmatrix} F_{1,N} & 0 \\ 0 & 0 \end{pmatrix}$ and remark that $\frac{1}{ML}\mathrm{Tr}\mathcal{E}F = \frac{1}{ML}\mathrm{Tr}(\mathcal{E}_{\mathbf{pp}}(z)\, F_{1,N})$ can be written as

$$\frac{1}{ML}\mathrm{Tr}\mathcal{E}F = \frac{1}{1-c^2\boldsymbol{\alpha}^2} \sum_{\substack{i_1,i_2 \\ m_1,m_2}} \Big(\big(\mathrm{Tr}\boldsymbol{\Delta}_{i_1 i_2}^{m_1 m_2}(pf) - c\boldsymbol{\alpha}\mathrm{Tr}\boldsymbol{\Delta}_{i_1 i_2}^{m_1 m_2}(ff)\big)\, \mathbf{1}_{i_2 \leq L} + \big(\mathrm{Tr}\boldsymbol{\Delta}_{i_1 i_2 - L}^{m_1 m_2}(fp) - c\boldsymbol{\alpha}\mathrm{Tr}\boldsymbol{\Delta}_{i_1 i_2 - L}^{m_1 m_2}(pp)\big)\, \mathbf{1}_{i_2 > L} \Big) F_{i_2 i_1}^{m_2 m_1}. \quad (5.25)$$

As matrix F verifies $F_{i_2,i_1}^{m_2,m_1} = 0$ if $i_2 > L$, $\frac{1}{ML}\mathrm{Tr}\mathcal{E}F$ is reduced to the first term of the right hand side of (5.25) that we now evaluate.

$$\begin{aligned}
&\sum_{\substack{i_1,i_2 \\ m_1,m_2}} \mathrm{Tr}\boldsymbol{\Delta}_{i_1 i_2}^{m_1 m_2}(pf) F_{i_2 i_1}^{m_2 m_1} \mathbf{1}_{i_2 \leq L} = c \sum_{\substack{i_1,i_2 \\ m_1,m_2}} \sum_{j,k} \mathbb{E}\Big\{ \mathcal{T}_{N,L}^{M}(\mathbf{Q}_{\mathbf{ff}}^{\circ}(I_L \otimes R))_{jk} \Big(\mathbf{Q} \left(\begin{smallmatrix} w_{f,k} \\ 0 \end{smallmatrix}\right)\Big)_{i_1}^{m_1} \\
&\times \Big(w_{f,j}^*\Big)_{i_2}^{m_2} F_{i_2 i_1}^{m_2 m_1} + (\mathcal{T}_{N,L}^{M}(\mathbf{Q}_{\mathbf{pf}}^{\circ}(I_L \otimes R)) J_N^{*L})_{jk} \Big(\mathbf{Q} \left(\begin{smallmatrix} 0 \\ w_{p,k} \end{smallmatrix}\right)\Big)_{i_1}^{m_1} \Big(w_{f,j}^*\Big)_{i_2}^{m_2} F_{i_2 i_1}^{m_2 m_1} \Big\} \mathbf{1}_{i_2 \leq L} \\
&= c\,\mathrm{Tr}\mathbb{E}\Big\{ \mathcal{T}_{N,L}^{M}(\mathbf{Q}_{\mathbf{ff}}^{\circ}(I_L \otimes R)) \left(\begin{smallmatrix} W_f \\ 0 \end{smallmatrix}\right)^* F\mathbf{Q} \left(\begin{smallmatrix} W_f \\ 0 \end{smallmatrix}\right) + \mathcal{T}_{N,L}^{M}(\mathbf{Q}_{\mathbf{pf}}^{\circ}(I_L \otimes R)) J_N^{*L} \left(\begin{smallmatrix} W_f \\ 0 \end{smallmatrix}\right)^* F\mathbf{Q} \left(\begin{smallmatrix} 0 \\ W_p \end{smallmatrix}\right) \Big\} \\
&= c\,\mathrm{Tr}\mathbb{E}\Big\{ \mathcal{T}_{N,L}^{M}(\mathbf{Q}_{\mathbf{ff}}^{\circ}(I_L \otimes R))\, (\Pi_{pf} W)^* F\mathbf{Q}\, (\Pi_{pf} W) \\
&\quad + \mathcal{T}_{N,L}^{M}(\mathbf{Q}_{\mathbf{pf}}^{\circ}(I_L \otimes R)) J_N^{*L}\, (\Pi_{pf} W)^* F\mathbf{Q}\, (\Pi_{fp} W) \Big\}.
\end{aligned}$$

Similar calculations lead to the following expression of $\frac{1}{ML}\mathrm{Tr}\mathcal{E}F$:

$$\begin{aligned}
\frac{1}{ML}\mathrm{Tr}\mathcal{E}F &= \frac{c}{(1-c_N^2\boldsymbol{\alpha}^2)} \frac{1}{ML}\mathrm{Tr}\mathbb{E}\Big\{ \mathcal{T}_{N,L}^{M}(\mathbf{Q}_{\mathbf{ff}}^{\circ}(I_L \otimes R))\, (\Pi_{pf} W)^* F\mathbf{Q}\, (\Pi_{pf} W) \\
&+ \mathcal{T}_{N,L}^{M}(\mathbf{Q}_{\mathbf{pf}}^{\circ}(I_L \otimes R)) J_N^{*L}\, (\Pi_{pf} W)^* F\mathbf{Q}\, (\Pi_{fp} W) - c\boldsymbol{\alpha}\mathcal{T}_{N,L}^{M}(\mathbf{Q}_{\mathbf{pp}}^{\circ}(I_L \otimes R))\, (\Pi_{pf} W)^* F\mathbf{Q}\, (\Pi_{fp} W) \\
&- c\boldsymbol{\alpha} J_N^{L} \mathcal{T}_{N,L}^{M}(\mathbf{Q}_{\mathbf{fp}}^{\circ}(I_L \otimes R))\, (\Pi_{pf} W)^* F\mathbf{Q}\, (\Pi_{pf} W) \Big\}. \quad (5.26)
\end{aligned}$$

We now evaluate the right hand side of (5.26). The Schwartz inequality leads to

$$\left|\frac{1}{ML}\mathrm{Tr}\mathbb{E}\Big\{\mathcal{T}_{N,L}^{M}(\mathbf{Q}_{\mathbf{ff}}^{\circ}(I_L\otimes R))\,(\Pi_{pf}W)^*\,F\mathbf{Q}\,(\Pi_{pf}W)\Big\}\right|$$
$$=\left|\sum_{l=-L+1}^{L-1}\mathbb{E}\Big\{\tau^{(M)}(\mathbf{Q}_{\mathbf{ff}}^{\circ}(I_L\otimes R))(l)\frac{1}{ML}\mathrm{Tr}\Big(J_N^{*\epsilon(l)}\,(\Pi_{pf}W)^*\,F\mathbf{Q}\,(\Pi_{pf}W)\Big)\Big\}\right|$$
$$=\left|\sum_{l=-L+1}^{L-1}\mathbb{E}\Big\{\frac{1}{ML}\mathrm{Tr}(\mathbf{Q}_{\mathbf{ff}}^{\circ}(I_L\otimes R)(J_L^{\epsilon(l)}\otimes I_M))\frac{1}{ML}\mathrm{Tr}\Big(J_N^{*\epsilon(l)}\,(\Pi_{pf}W)^*\,F\mathbf{Q}\,(\Pi_{pf}W)\Big)^{\circ}\Big\}\right|$$
$$\leq\sum_{l=-L+1}^{L-1}\mathbf{Var}\left\{\frac{1}{ML}\mathrm{Tr}(\mathbf{Q}_{\mathbf{ff}}(I_L\otimes R)(J_L^{\epsilon(l)}\otimes I_M))\right\}^{1/2}$$
$$\times\,\mathbf{Var}\left\{\frac{1}{ML}\mathrm{Tr}\Big(J_N^{*\epsilon(l)}\,(\Pi_{pf}W)^*\,F\mathbf{Q}\,(\Pi_{pf}W)\Big)\right\}^{1/2}.$$

Using Corollary 3.1, we obtain that

$$\mathbf{Var}\left\{\frac{1}{ML}\mathrm{Tr}(\mathbf{Q}_{\mathbf{ff}}(I_L\otimes R)(J_L^{\epsilon(l)}\otimes I_M))\right\}\leq\frac{1}{N^2}P_1(|z^2|)P_2\left(\frac{1}{\mathrm{Im}z^2}\right)$$

and that

$$\mathbf{Var}\left\{\frac{1}{ML}\mathrm{Tr}\Big(J_N^{*\epsilon(l)}\,(\Pi_{pf}W)^*\,F\mathbf{Q}\,(\Pi_{pf}W)\Big)\right\}\leq\kappa^2\frac{1}{N^2}P_1(|z^2|)P_2\left(\frac{1}{\mathrm{Im}z^2}\right).$$

Since L does not grow with N, this implies immediately that

$$\left|\frac{1}{ML}\mathrm{Tr}\mathbb{E}\Big\{\mathcal{T}_{N,L}^{M}(\mathbf{Q}_{\mathbf{ff}}^{\circ}(I_L\otimes R))\,(\Pi_{pf}W)^*\,F\mathbf{Q}\,(\Pi_{pf}W)\Big\}\right|\leq\kappa\frac{1}{N^2}P_1(|z^2|)P_2\left(\frac{1}{\mathrm{Im}z^2}\right)$$

holds. It can be shown similarly that the 3 other normalized traces can be upper bounded by the same kind of term. It remains to control the terms $\frac{1}{1-(c_N\,\boldsymbol{\alpha}_N)^2}$ and $\frac{\boldsymbol{\alpha}_N}{1-(c_N\,\boldsymbol{\alpha}_N)^2}$. For this, we use Lemma 4.2 for the choice $\beta_N(z)=\alpha_N(z)$. It is sufficient to verify that the measures $(\overline{\mu}_N)_{N\geq 1}$ associated to functions $(\alpha_N(z))_{N\geq 1}$ verify (4.9) and (4.10). For each N, it holds that

$$\int_0^{+\infty}d\,\overline{\mu}_N(\lambda)=\mathbb{E}\Big\{\int_0^{+\infty}d\,\hat{\mu}_N(\lambda)\Big\}=\frac{1}{M}\mathrm{Tr}R_N$$

and

$$\int_0^{+\infty}\lambda\,d\,\overline{\mu}_N(\lambda)=\mathbb{E}\left(\int_0^{+\infty}\lambda\,d\,\hat{\mu}_N(\lambda)\right)=\mathbb{E}\left(\frac{1}{ML}\mathrm{Tr}((I_L\otimes R)W_fW_p^*W_pW_f^*)\right).$$

A straightforward calculation leads to $\mathbb{E}\Big\{\frac{1}{ML}\mathrm{Tr}(W_fW_p^*W_pW_f^*)\Big\}=\frac{c_N}{M^2}\mathrm{Tr}R_N\mathrm{Tr}R_N^2$. Therefore, (4.12) implies that

$$\frac{1}{|1-z(c_N\alpha_N(z))^2|}\leq P_1(|z|)P_2\left(\frac{1}{\mathrm{Im}z}\right)$$

for each $z \in \mathbb{C}^+$, and if $z^2 \in \mathbb{C}^+$, it holds that

$$\frac{1}{|1 - z^2(c_N \alpha_N(z^2))^2|} \leq P_1(|z^2|) P_2 \left(\frac{1}{\mathrm{Im} z^2} \right).$$

As $\boldsymbol{\alpha}_N(z) = z\alpha_N(z^2)$, this is equivalent to

$$\frac{1}{1 - (c_N\, \boldsymbol{\alpha}_N)^2} \leq P_1(|z^2|) P_2 \left(\frac{1}{\mathrm{Im} z^2} \right).$$

Finally, we remark that $|\alpha_N(z)| \leq \frac{1}{M} \mathrm{Tr} R_N \, \frac{1}{\mathrm{Im} z} \leq b \, \frac{1}{\mathrm{Im} z}$ for each $z \in \mathbb{C}^+$. Therefore, if $z^2 \in \mathbb{C}^+$, it holds that $|\alpha_N(z^2)| \leq b \, \frac{1}{\mathrm{Im} z^2}$ and that $|\boldsymbol{\alpha}_N(z)| = |z||\alpha_N(z^2)|$ verifies

$$|\boldsymbol{\alpha}_N(z)| \leq b|z| \, \frac{1}{\mathrm{Im} z^2} \leq b(1 + |z|^2) \, \frac{1}{\mathrm{Im} z^2}.$$

This completes the proof of Proposition 5.1. ■

Proposition 5.1 immediately leads to the following Corollary.

Corollary 5.1. *For each sequence* $(F_N)_{N\geq 1}$ *of deterministic* $ML \times ML$ *matrices such that* $\sup_{N\geq 1} \|F_N\| \leq \kappa$ *we have*

$$\left| \frac{1}{ML} \mathrm{Tr} \left[\left(\mathbb{E}\{Q_N(z)\} - I_L \otimes S_N(z) \right) F_N \right] \right| \leq \kappa \, \frac{1}{N^2} P_1(|z|) P_2 \left(\frac{1}{\mathrm{Im} z^2} \right) \tag{5.27}$$

for each $z \in \mathbb{C}^+$. *In particular, it holds that*

$$\left| \frac{1}{ML} \mathrm{Tr} \left[\left(\mathbb{E}\{Q_N(z)\} - I_L \otimes S_N(z) \right) \right] \right| \leq \kappa \, \frac{1}{N^2} P_1(|z|) P_2 \left(\frac{1}{\mathrm{Im} z^2} \right). \tag{5.28}$$

Proof. (5.22) implies that

$$\left| \frac{1}{ML} \mathrm{Tr} \left[\left(\mathbb{E}\{Q_N(z^2)\} - I_L \otimes S_N(z^2) \right) F_N \right] \right| = \left| \frac{1}{ML} \mathrm{Tr} \mathcal{E}_{\mathbf{pp}}(z) \left(I_L \otimes S_N(z^2) \right) F_N \right|$$

As $\mathcal{E}_{\mathbf{pp}}(z) = E(z^2)$ and $\|S_N(z^2)\| \leq \frac{1}{\mathrm{Im} z^2}$ if $z^2 \in \mathbb{C}^+$, the application of Proposition 5.1 to matrix $F_{1,N} = S_N(z^2) F_N$ implies that

$$\left| \frac{1}{ML} \mathrm{Tr} \left[\left(\mathbb{E}\{Q_N(z^2)\} - I_L \otimes S_N(z^2) \right) F_N \right] \right| \leq \kappa \frac{1}{N^2} P_1(|z^2|) P_2 \left(\frac{1}{\mathrm{Im} z^2} \right)$$

for each z such that $z^2 \in \mathbb{C}^+$. Exchanging z^2 by z eventually establishes (5.27). This, in turn, completes the proof of Theorem 5.1.

6. Deterministic equivalent of $\mathbb{E}\{Q\}$

6.1. *The canonical equation*

Proposition 6.1. *If* $z \in \mathbb{C}^+$, *there exists a unique solution of the equation*

$$t_N(z) = \frac{1}{M} \mathrm{Tr} R_N \left(-z I_M - \frac{z c_N t_N(z)}{1 - z c_N^2 t_N^2(z)} R_N \right)^{-1} \tag{6.1}$$

satisfying $t_N(z) \in \mathbb{C}^+$ *and* $zt_N(z) \in \mathbb{C}^+$. *Function* $z \to t_N(z)$ *is an element of* $\mathcal{S}(\mathbb{R}^+)$, *and the associated positive measure, denoted by* μ_N, *verifies*

$$\mu_N(\mathbb{R}^+) = \frac{1}{M}\mathrm{Tr}R_N, \int_{\mathbb{R}^+} \lambda\, d\,\mu_N(\lambda) = c_N\, \frac{1}{M}\mathrm{Tr}R_N\, \frac{1}{M}\mathrm{Tr}R_N^2. \tag{6.2}$$

Moreover, it exists nice constants β *and* κ *such that*

$$\frac{1}{\left|1 - z\,(c_N\, t_N(z))^2\right|} \leq \frac{\kappa\,(\beta^2 + |z|^2)^2}{(\mathrm{Im}z)^3} \tag{6.3}$$

for each N. *Finally, the* $M \times M$ *valued function* $T_N(z)$ *defined by*

$$T_N(z) = -\left(zI_M + \frac{zc_N t_N(z)}{1 - zc_N^2 t_N^2(z)} R_N\right)^{-1} \tag{6.4}$$

belongs to $\mathcal{S}_M(\mathbb{R}^+)$. *The associated* $M \times M$ *positive matrix-valued measure, denoted* ν_N^T, *verifies*

$$\nu_N^T(\mathbb{R}^+) = I_M \tag{6.5}$$

as well as

$$\mu_N = \frac{1}{M}\mathrm{Tr}R_N \nu_N^T. \tag{6.6}$$

Proof. As N is assumed to be fixed in the statement of the Proposition, we omit to mention that $t_N, T_N, \mu_N, \ldots$ depend on N in the course of the proof. We first prove the existence of a solution such that $z \to t(z)$ is an element of $\mathcal{S}(\mathbb{R}^+)$. For this, we use the classical fixed point equation scheme. We define $t_0(z) = -\frac{1}{z}$, which is of course an element of $\mathcal{S}(\mathbb{R}^+)$, and generate sequence $(t_n(z))_{n\geq 1}$ by the formula

$$t_{n+1}(z) = \frac{1}{M}\mathrm{Tr}R\left(-zI_M - \frac{zct_n(z)}{1 - zc^2t_n^2(z)}R\right)^{-1}.$$

We establish by induction that for each n, $t_n \in \mathcal{S}(\mathbb{R}^+)$, and that its associated measure μ_n verifies $\mu_n(\mathbb{R}^+) = \frac{1}{M}\mathrm{Tr}R$ and

$$\int_0^{+\infty} \lambda\mu_n(d\lambda) = c\frac{1}{M}\mathrm{Tr}(R)\frac{1}{M}\mathrm{Tr}(R^2). \tag{6.7}$$

Thanks to (2.7), this last property will imply that sequence $(\mu_n)_{n\geq 1}$ is tight. We assume that t_n indeed satisfies the above conditions, and prove that $t_{n+1}(z)$ also meets these requirements. Lemma 4.1 implies that function $T_n(z) = \left(-zI_M - \frac{zct_n(z)}{1 - zc^2t_n^2(z)}R\right)^{-1}$ is an element of $\mathcal{S}_M(\mathbb{R}^+)$. According to Proposition 4.1, to prove that $t_{n+1}(z) \in \mathcal{S}(\mathbb{R}^+)$, we need to check that $\mathrm{Im}t_{n+1}(z)$, $\mathrm{Im}zt_{n+1}(z) > 0$ if $z \in \mathbb{C}^+$, as well as that $\lim_{y\to+\infty} iyt_{n+1}(iy)$ exists. As

$T_n \in \mathcal{S}_M(\mathbb{R}^+)$ and $t_{n+1}(z) = \frac{1}{M}\mathrm{Tr}RT_n(z)$, it is clear that $\mathrm{Im}t_{n+1}(z)$, $\mathrm{Im}zt_{n+1}(z) > 0$. Finally, it holds that

$$-iyt_{n+1}(iy) = \frac{1}{M}\mathrm{Tr}R\left(I_M + \frac{ciyt_n(iy)}{iy - (ciyt_n(iy))^2}R\right)^{-1}.$$

Since $t_n(z)$ is a Stieltjes transform we have $-iyt_n(iy) \to \mu_n(\mathbb{R}^+)$, which implies that $-iyt_{n+1}(iy) \to \frac{1}{M}\mathrm{Tr}R$, i.e. that $\mu_{n+1}(\mathbb{R}^+) = \frac{1}{M}\mathrm{Tr}R$.

We finally check that μ_{n+1} satisfies (6.7). For this, we follow [16].

$$\int_0^{+\infty} \lambda\mu_{n+1}(d\lambda) = \lim_{y\to+\infty} \Re\left(-iy(iy\frac{1}{M}\mathrm{Tr}RT_n(iy) + \frac{1}{M}\mathrm{Tr}R)\right).$$

We can express T_n as

$$T_n = -\frac{1}{z}\left(I_M + \frac{ct_n}{1-zc^2t_n^2}R\right)^{-1} = -\frac{1}{z} + \frac{R}{z}\frac{ct_n}{1-zc^2t_n^2} - \left(\frac{ct_n}{1-zc^2t_n^2}\right)^2 R^2T_n,$$

from which it follows that

$$-z\left(\frac{1}{M}\mathrm{Tr}(zRT_n(z)) + \frac{1}{M}\mathrm{Tr}R)\right) = -\frac{czt_n}{1-zc^2t_n^2}\frac{1}{M}\mathrm{Tr}R^2 + \left(\frac{czt_n}{1-zc^2t_n^2}\right)^2\frac{1}{M}\mathrm{Tr}R^3T_n.$$

Since $-iyt_n(iy) \to \frac{1}{M}\mathrm{Tr}R$ and $t_n(iy) \to 0$ we can conclude that $-iy(iy\frac{1}{M}\mathrm{Tr}RT_n(iy) + \frac{1}{M}\mathrm{Tr}R) \to \frac{c}{M^2}\mathrm{Tr}R\mathrm{Tr}R^2$ as expected.

We now prove that sequence t_n converges towards a function $t \in \mathcal{S}(\mathbb{R}^+)$ verifying equation (6.1). For this we evaluate $\theta_n = t_{n+1} - t_n$

$$\theta_n = \frac{1}{M}\mathrm{Tr}R(T_n - T_{n-1}) = \frac{1}{M}\mathrm{Tr}RT_n\frac{zc(t_n - t_{n-1})(1 + zc^2t_nt_{n-1})}{(1-zc^2t_n^2)(1-zc^2t_{n-1}^2)}RT_{n-1}$$
$$= \theta_{n-1}\frac{zc(1+zc^2t_nt_{n-1})}{(1-zc^2t_n^2)(1-zc^2t_{n-1}^2)}\frac{1}{M}\mathrm{Tr}RT_nRT_{n-1}.$$

We denote by $f_n(z)$ the term defined by

$$f_n(z) = \frac{zc(1+zc^2t_nt_{n-1})}{(1-zc^2t_n^2)(1-zc^2t_{n-1}^2)}\frac{1}{M}\mathrm{Tr}RT_nRT_{n-1}. \tag{6.8}$$

Lemma 4.1 implies that $\|T_k\| \le \frac{1}{\mathrm{Im}z}$ and that $|t_k| \le \frac{b}{\mathrm{Im}z}$ for each $k \ge 1$ and each $z \in \mathbb{C}^+$. Therefore, it holds that

$$\left|zc(1+zc^2t_nt_{n-1})\frac{1}{M}\mathrm{Tr}RT_nRT_{n-1}\right| \le \kappa\left(\frac{|z|}{(\mathrm{Im}z)^2}\left(1 + \frac{|z|}{(\mathrm{Im}z)^2}\right)\right).$$

Moreover, it is clear that for each k, $|1 - zc^2t_k^2| \ge (1 - c^2\frac{|z|}{(\mathrm{Im}z)^2})$. For each $\epsilon > 0$ small enough, we consider the domain $\mathcal{D}_\epsilon$ defined by

$$\mathcal{D}_\epsilon = \{z \in \mathbb{C}^+, \frac{|z|}{(\mathrm{Im}z)^2} < \epsilon\}. \tag{6.9}$$

Then, for $z \in \mathcal{D}_\epsilon$, it holds that

$$\frac{1}{|1-zc^2t_n^2|}\frac{1}{|1-zc^2t_{n-1}^2|} \leq \frac{1}{(1-c^2\epsilon)^2}$$

and that

$$|f_n(z)| \leq \frac{\kappa}{(1-c^2\epsilon)^2}\left(\epsilon+\epsilon^2\right).$$

We choose ϵ in such a way that $\frac{\kappa}{(1-c^2\epsilon)^2}\left(\epsilon+\epsilon^2\right) < 1/2$. Then, for each $z \in \mathcal{D}_\epsilon$, it holds that

$$|\theta_n| \leq \frac{1}{2}|\theta_{n-1}|.$$

Therefore, for each z in $\mathcal{D}_\epsilon$, $(t_n(z))_{n\geq 1}$ is a Cauchy sequence. We denote by $t(z)$ its limit. $(t_n(z))_{n\geq 1}$ is uniformly bounded on every compact set of $\mathbb{C}-\mathbb{R}^+$. This implies that $(t_n(z))_{n\geq 1}$ is a normal family on $\mathbb{C}-\mathbb{R}^+$. We consider a converging subsequence extracted from $(t_n(z))_{n\geq 1}$. The corresponding limit $t_*(z)$ is analytic over $\mathbb{C}-\mathbb{R}^+$. If $z \in \mathcal{D}_\epsilon$, $t_*(z)$ must be equal to $t(z)$. Therefore, the limits of all converging subsequences extracted from $(t_n(z))_{n\geq 1}$ must coincide on $\mathcal{D}_\epsilon$, and therefore on $\mathbb{C}-\mathbb{R}^+$. This implies that $t_n(z)$ converges uniformly on each compact subset towards a function which is analytic $\mathbb{C}-\mathbb{R}^+$, and that we also denote by $t(z)$. It is clear that $t(z)$ verifies (6.1) and that $t \in \mathcal{S}(\mathbb{R}^+)$ and verifies (6.2). Moroever, Lemma 4.1 implies that $T \in \mathcal{S}_M(\mathbb{R}^+)$, while (6.6) and (6.5) are obtained immediately.

As (6.2) holds, (6.3) is a consequence of the application of Lemma 4.2 to the function $\beta_N(z) = t_N(z)$.

We now prove that if $z \in \mathbb{C}^+$ and $t_1(z)$ and $t_2(z)$ are 2 solutions of (6.1) such that $t_i(z)$ and $zt_i(z)$ belong to $\mathbb{C}^+$, $i=1,2$, then $t_1(z)=t_2(z)$. In order to prove this, we first establish the following useful Lemma.

Lemma 6.1. *If $z \in \mathbb{C}^+$ and if $t(z)$ verifies the conditions of Proposition 6.1, then, it holds that*

$$1-u(z) > 0 \tag{6.10}$$

and

$$\det(\mathbf{I}-\mathbf{D}) > 0, \tag{6.11}$$

where

$$\mathbf{D} = \begin{pmatrix} u(z) & v(z) \\ |z|^2 v(z) & u(z) \end{pmatrix}, \tag{6.12}$$

$$u(z) = c\,\frac{|czt(z)|^2\,\frac{1}{M}\mathrm{Tr}(RT(z)(T(z))^*R)}{|1-z(ct(z))^2|^2}, \tag{6.13}$$

$$v(z) = c\,\frac{\frac{1}{M}\mathrm{Tr}(RT(z)(T(z))^*R)}{|1-z(ct(z))^2|^2}. \tag{6.14}$$

Proof. Using the equation $t(z) = \frac{1}{M}\mathrm{Tr} RT(z)$, we obtain immediately after some algebra that

$$\begin{pmatrix} \frac{\mathrm{Im}(t(z))}{\mathrm{Im}(z)} \\ \frac{\mathrm{Im}(zt(z))}{\mathrm{Im}(z)} \end{pmatrix} = \mathbf{D} \begin{pmatrix} \frac{\mathrm{Im}(t(z))}{\mathrm{Im}(z)} \\ \frac{\mathrm{Im}(zt(z))}{\mathrm{Im}(z)} \end{pmatrix} + \begin{pmatrix} \frac{1}{M}\mathrm{Tr}(RT(z)(T(z))^*) \\ 0 \end{pmatrix}. \tag{6.15}$$

The first component of (6.15) implies that

$$(1-u(z))\,\frac{\mathrm{Im}(t(z))}{\mathrm{Im}(z)} = v(z)\,\frac{\mathrm{Im}(zt(z))}{\mathrm{Im}(z)} + \frac{1}{M}\mathrm{Tr}(RT(z)(T(z))^*).$$

Therefore, it holds that $(1-u(z)) > 0$. Plugging the equality

$$\frac{\mathrm{Im}(t(z))}{\mathrm{Im}(z)} = \frac{v(z)}{1-u(z)}\,\frac{\mathrm{Im}(zt(z))}{\mathrm{Im}(z)} + \frac{1}{1-u(z)}\,\frac{1}{M}\mathrm{Tr}(RT(z)(T(z))^*)$$

into the second component of (6.15) leads to

$$\left(1-u(z)-\frac{|z|^2v^2(z)}{1-u(z)}\right)\frac{\mathrm{Im}(zt(z))}{\mathrm{Im}(z)} = \frac{|z|^2v(z)}{1-u(z)}\,\frac{1}{M}\mathrm{Tr}(RT(z)(T(z))^*) > 0$$

and to (6.11).

To complete the proof of the uniqueness, we assume that equation (6.1) has 2 solutions $t_1(z)$ and $t_2(z)$ such that $t_i(z)$ and $zt_i(z)$ belong to $\mathbb{C}^+$ for $i = 1, 2$. The proof of Lemma 4.1 (see in particular (4.6)) implies that for $i = 1, 2$, then $1 - z(ct_i(z))^2 \neq 0$ and matrix $-zI - \dfrac{zct_i(z)}{1-zc^2t_i^2(z)}R$ is invertible. We denote by $T_1(z)$ and $T_2(z)$ the matrices defined by (6.4) when $t(z) = t_1(z)$ and $t(z) = t_2(z)$ respectively. $u_i(z)$ and $v_i(z)$, $i = 1, 2$, are defined similarly from (6.13) and (6.14) when $t(z) = t_1(z)$ and $t(z) = t_2(z)$. Using that $t_i(z) = \frac{1}{M}\mathrm{Tr}(RT_i(z))$ for $i = 1, 2$, we obtain immediately that

$$t_1(z) - t_2(z) = (u_{1,2}(z) + zv_{1,2}(z))\,(t_1(z) - t_2(z)),$$

where

$$u_{1,2}(z) = c\,\frac{czt_1(z)czt_2(z)\,\frac{1}{M}\mathrm{Tr}(RT_1(z)RT_2(z))}{(1-z(ct_1(z))^2)\,(1-z(ct_2(z))^2)} \tag{6.16}$$

and

$$v_{1,2}(z) = c\,\frac{\frac{1}{M}\mathrm{Tr}(RT_1(z)RT_2(z))}{(1-z(ct_1(z))^2)\,(1-z(ct_2(z))^2)}. \tag{6.17}$$

In order to prove that $t_1(z) = t_2(z)$, it is sufficient establish that $1 - u_{1,2}(z) - zv_{1,2}(z) \neq 0$. For this, we prove the following inequality:

$$|1-u_{1,2}(z)-zv_{1,2}(z)| > \sqrt{(1-u_1(z))-|z|v_1(z)}\,\sqrt{(1-u_2(z))-|z|v_2(z)} \tag{6.18}$$

which, by Lemma 6.1, implies $1 - u_{1,2}(z) - zv_{1,2}(z)) \neq 0$. For this, we remark that the Schwartz inequality leads to $|u_{1,2}(z)| \leq \sqrt{u_1(z)}\sqrt{u_2(z)}$ and $|v_{1,2}(z)| \leq \sqrt{v_1(z)}\sqrt{v_2(z)}$. Therefore,

$$|1-u_{1,2}(z)-zv_{1,2}(z)| \geq 1 - \sqrt{u_1(z)}\sqrt{u_2(z)} - \sqrt{|z|v_1(z)}\sqrt{|z|v_2(z)}.$$

We now use the inequality

$$\sqrt{ab} - \sqrt{cd} \geq \sqrt{a-c}\,\sqrt{b-d}, \tag{6.19}$$

where a, b, c, d are positive real numbers such that $a \geq c$ and $b \geq d$. (6.19) for $a = b = 1$ and $c = u_1(z)$, $d = u_2(z)$ implies that $1 - \sqrt{u_1(z)}\sqrt{u_2(z)} \geq \sqrt{1-u_1(z)}\sqrt{1-u_2(z)}$. Therefore, it holds that

$$|1 - u_{1,2}(z) - zv_{1,2}(z)| \geq \sqrt{1-u_1(z)}\sqrt{1-u_2(z)} - \sqrt{|z|v_1(z)}\sqrt{|z|v_2(z)}.$$

(6.19) for $a = 1 - u_1(z)$, $b = 1 - u_2(z)$, $c = |z|v_1(z)$ and $d = |z|v_2(z)$ eventually leads to (6.18). This completes the proof of the uniqueness of the solution of (6.1) and Proposition 6.1. ■

Remark 6.1. (6.10) and (6.11) are still valid if z belongs to $\mathbb{R}^{-*}$. To check this, it is sufficient to remark if $z = x \in \mathbb{R}^{-*}$, the fundamental equation (6.15) is still valid, but $\frac{\mathrm{Im}(t(z))}{\mathrm{Im}(z)}$ and $\frac{\mathrm{Im}(zt(z))}{\mathrm{Im}(z)}$ have to be replaced by $t^{'}(x)$ and $(xt(x))^{'}$ where $^{'}$ denotes the differentiation operator w.r.t. x. The same conclusions are obtained because $t^{'}(x) > 0$ and $(xt(x))^{'} > 0$ if $x \in \mathbb{R}^{-*}$.

6.2. *Convergence*

In this paragraph, we establish that the empirical eigenvalue distribution $\hat{\nu}_N$ of matrix $W_{f,N}W_{p,N}^*W_{p,N}W_{f,N}^*$ has almost surely the same deterministic behaviour than the probability measure ν_N defined by

$$\nu_N = \frac{1}{M}\mathrm{Tr}\nu_N^T, \tag{6.20}$$

where we recall that ν_N^T represents the positive matrix valued measure associated to $T_N(z)$. For this, we first establish the following Proposition.

Proposition 6.2. *For each sequence $(F_N)_{N\geq 1}$ of deterministic $ML \times ML$ matrices such that $\sup_{N\geq 1}\|F_N\| \leq \kappa$, then,*

$$\frac{1}{ML}\mathrm{Tr}\left[\left(\mathbb{E}\{Q_N(z)\} - I_L \otimes T_N(z)\right)F_N\right] \to 0 \tag{6.21}$$

holds for each $z \in \mathbb{C} - \mathbb{R}^+$.

Proof. Corollary 5.1 implies that

$$\frac{1}{ML}\mathrm{Tr}(\mathbb{E}\{Q_N\} - (I_L \otimes S_N))F_N = \mathcal{O}\left(\frac{1}{N^2}\right).$$

We have therefore to show that $\frac{1}{ML}\mathrm{Tr}\left(I_L \otimes (S_N - T_N)\right)F_N \to 0$. It is easy to check that

$$\frac{1}{ML}\mathrm{Tr}\left(I_L \otimes (S-T)\right)F = \frac{1}{ML}\mathrm{Tr}(I_L \otimes S)\left(\frac{zc_N\alpha}{1 - zc_N^2\alpha^2} - \frac{zc_N t}{1 - zc_N^2 t^2}\right)(I_L \otimes RT)F$$
$$= \frac{zc_N(\alpha - t)(1 + zc_N^2\alpha t)}{(1 - zc_N^2\alpha^2)(1 - zc_N^2 t^2)}\frac{1}{ML}\mathrm{Tr}(I_L \otimes SRT)F. \tag{6.22}$$

We express $\alpha - t$ as $\alpha - \frac{1}{M}\text{Tr}RS + \frac{1}{M}\text{Tr}R(S-T)$, and deduce from (6.22) that

$$\frac{1}{ML}\text{Tr}\left(I_L \otimes (S-T)\right)F = \left(\alpha - \frac{1}{M}\text{Tr}RS\right)\frac{zc_N(1+zc_N^2\alpha t)}{(1-zc_N^2\alpha^2)(1-zc_N^2 t^2)}$$
$$\times\frac{1}{ML}\text{Tr}(I_L \otimes SRT)F + \frac{1}{M}\text{Tr}R(S-T)\,\frac{zc_N(1+zc_N^2\alpha t)}{(1-zc_N^2\alpha^2)(1-zc_N^2 t^2)}\frac{1}{ML}\text{Tr}(I_L \otimes SRT)F. \tag{6.23}$$

(5.27) implies that $\alpha - \frac{1}{M}\text{Tr}RS = \mathcal{O}_z(\frac{1}{N^2})$. Therefore, in order to establish (6.21), it is sufficient to prove that $\frac{1}{M}\text{Tr}R(S-T) \to 0$. For this, we take $F = I_L \otimes R$ in (6.23) and get that

$$\frac{1}{M}\text{Tr}R(S(z)-T(z)) = f_N(z)\,\frac{1}{M}\text{Tr}R(S(z)-T(z)) + \mathcal{O}_z(\frac{1}{N^2}) \tag{6.24}$$

where $f_N(z)$ is defined by

$$f_N(z) = \frac{zc_N(1+zc_N^2\alpha t)}{(1-zc_N^2\alpha^2)(1-zc_N^2 t^2)}\frac{1}{M}\text{Tr}(RS(z)RT(z)).$$

$f_N(z)$ is similar to the term defined in (6.8). Using the arguments of the proof of Proposition 6.1, we obtain that it is possible to find $\epsilon > 0$ for which, $\sup_{N \geq N_0} |f_N(z)| < \frac{1}{2}$ for each $z \in \mathcal{D}_\epsilon$ for some large enough integer N_0. We recall that $\mathcal{D}_\epsilon$ is defined by (6.9). We therefore deduce from (6.24) that $\frac{1}{M}\text{Tr}R(S(z)-T(z)) \to 0$ and $\dfrac{1}{ML}\text{Tr}\left(I_L \otimes (S(z)-T(z))\right)F$ converge towards 0 for each $z \in \mathcal{D}_\epsilon$. As functions $z \to \dfrac{1}{ML}\text{Tr}\left(I_L \otimes (S_N(z)-T_N(z))\right)F_N$ are holomorphic on $\mathbb{C}-\mathbb{R}^+$ and are uniformly bounded on each compact subset of $\mathbb{C}-\mathbb{R}^+$, we deduce from Montel's theorem that $\dfrac{1}{ML}\text{Tr}\left(I_L \otimes (S_N(z)-T_N(z))\right)F_N$ converges towards 0 for each $z \in \mathbb{C}-\mathbb{R}^+$. ■

We deduce the following Corollary.

Corollary 6.1. *The empirical eigenvalue distribution $\hat{\nu}_N$ of $W_{f,N}W_{p,N}^* W_{p,N}W_{f,N}^*$ verifies*

$$\hat{\nu}_N - \nu_N \to 0 \tag{6.25}$$

weakly almost surely.

Proof. Proposition 6.2 implies that $\mathbb{E}\{\frac{1}{ML}\text{Tr}Q_N(z)\} - \frac{1}{M}\text{Tr}(T_N(z)) \to 0$ for each $z \in \mathbb{C}-\mathbb{R}^+$. The Poincaré-Nash inequality and the Borel Cantelli Lemma imply that $\frac{1}{ML}\text{Tr}(Q_N(z)) - \mathbb{E}\{\frac{1}{ML}\text{Tr}Q_N(z)\} \to 0$ a.s. for each $z \in \mathbb{C}-\mathbb{R}^+$. Therefore, it holds that

$$\frac{1}{ML}\text{Tr}(Q_N(z)) - \frac{1}{M}\text{Tr}(T_N(z)) \to a.s. \tag{6.26}$$

for each $z \in \mathbb{C} - \mathbb{R}^+$. Corollary 2.7 of [16] implies that $\hat{\nu}_N - \nu_N \to 0$ weakly almost surely provided we verify that $(\hat{\nu}_N)_{N \geq 1}$ is almost surely tight and that $(\nu_N)_{N \geq 1}$ is tight. It is clear that

$$\int_{\mathbb{R}^+} \lambda \, d\, \hat{\nu}_N(\lambda) = \frac{1}{ML} \mathrm{Tr} W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^* \leq \|W_N\|^4,$$

where we recall that

$$W_N = \begin{pmatrix} W_{p,N} \\ W_{f,N} \end{pmatrix}.$$

It holds that $\|W_N\| \leq \sqrt{b}\, \|W_{iid,N}\|$ where $W_{iid,N}$ is defined by (3.1). As $\|W_{iid,N}\| \to (1 + \sqrt{c_*})$ almost surely (see [26]), we obtain that $\frac{1}{ML} \mathrm{Tr} W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ is almost surely bounded for N large enough. This implies that $(\hat{\nu}_N)_{N \geq 1}$ is almost surely tight. As for sequence $(\nu_N)_{N \geq 1}$, we have shown that $\sup_N \int_{\mathbb{R}^+} \lambda \, d\, \mu_N(\lambda) < +\infty$. As $\mu_N = \frac{1}{M} \mathrm{Tr} R_N \nu_N^T$, the condition $R_N > aI$ for each N leads to

$$\int_{\mathbb{R}^+} \lambda \, d\, \mu_N(\lambda) \geq a \int_{\mathbb{R}^+} \lambda \, d\, \nu_N(\lambda).$$

Therefore, it holds that $\sup_N \int_{\mathbb{R}^+} \lambda \, d\, \nu_N(\lambda) < +\infty$, a condition which implies that $(\nu_N)_{N \geq 1}$ is tight. ■

7. Detailed study of ν_N.

In this section, we study the properties of ν_N. (2.7) implies that μ_N and ν_N are absolutely continuous one with respect each other. Hence, they share the same properties, and the same support denoted $\mathcal{S}_N$ in the following. We thus study μ_N and deduce the corresponding results related to ν_N. As in the context of other models, μ_N can be characterized by studying the Stieltjes transform $t_N(z)$ near the real axis. In the following, we denote by $\overline{M}$ the number of distinct eigenvalues $(\overline{\lambda}_{l,N})_{l=1,\ldots,\overline{M}}$ of R_N arranged in the decreasing order, and by $(m_{l,N})_{l=1,\ldots,\overline{M}}$ their multiplicities. It of course holds that $\sum_{l=1}^{\bar{M}} m_{l,N} = M$.

7.1. *Properties of $t(z)$ near the real axis.*

In this paragraph, we establish that if $x_0 \in \mathbb{R}^{+*}$, then, $\lim_{z \to x_0, z \in \mathbb{C}^+} t(z)$ exists and is finite. It will be denoted by $t(x_0)$ in order to simplify the notations. Moreover, when $c \leq 1$, $\lim_{z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*} |t(z)| = +\infty$, and $\lim_{z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*} z t(z) = 0$. The results of [35] will imply that measure μ_N is absolutely continuous w.r.t. the Lebesgue measure, and that the corresponding density is equal to $\frac{1}{\pi} \mathrm{Im}(t(x))$ for each $x \in \mathbb{R}^{+*}$. When $c > 1$, a Dirac mass appears at 0.

We first address the case where $x_0 \neq 0$, and, in order to establish the existence of $\lim_{z \to x_0, z \in \mathbb{C}^+} t(z)$, we prove the following properties:

- If $(z_n)_{n \geq 1}$ is a sequence of $\mathbb{C}^+$ converging towards x_0, then $|t(z_n)|_{n \geq 1}$ is bounded
- If $(z_{1,n})_{n \geq 1}$ and $(z_{2,n})_{n \geq 1}$ are two sequences of $\mathbb{C}^+$ converging towards x_0 and verifying $\lim_{z_{i,n} \to x_0} = t_i$ for $i = 1, 2$, then $t_1 = t_2$.

Lemma 7.1. *If $x_0 \in \mathbb{R}^{+*}$, and if $(z_n)_{n \geq 1}$ is a sequence of $\mathbb{C}^+$ such that $\lim_{n \to +\infty} z_n = x_0$, then the set $|t(z_n)|_{n \geq 1}$ is bounded.*

Proof. We assume that $|t(z_n)| \to +\infty$. Equation (6.1) can be written as

$$t(z_n) = \frac{1}{M} \sum_{l=1}^{\overline{M}} \frac{m_l \, \overline{\lambda}_l}{-z_n (1 + \frac{ct(z_n)\overline{\lambda}_l}{1 - z\, (ct(z_n))^2})}. \tag{7.1}$$

As $x_0 \neq 0$, the condition $|t(z_n)| \to +\infty$ implies that it exists l_0 for which

$$1 + \frac{ct(z_n)\overline{\lambda}_{l_0}}{1 - z\, (ct(z_n))^2} \to 0$$

or equivalently

$$z_n ct(z_n) - \frac{1}{ct(z_n)} \to \overline{\lambda}_{l_0}.$$

As $|t(z_n)| \to +\infty$, it holds that $z_n ct(z_n) \to \overline{\lambda}_{l_0}$, a contradiction because $|z_n ct(z_n)| \to +\infty$. ■

Lemma 7.2. *Consider $(z_{1,n})_{n \geq 1}$ and $(z_{2,n})_{n \geq 1}$ two sequences of $\mathbb{C}^+$ converging towards $x_0 \in \mathbb{R}^{+*}$ and verifying $\lim_{z_{i,n} \to x_0} t(z_{i,n}) = t_i$ for $i = 1, 2$. Then, it holds that $t_1 = t_2$.*

Proof. The statement of the Lemma is obvious if x_0 does not belong to $\mathcal{S}$. Therefore, we assume that $x_0 \in \mathcal{S} - \{0\}$. We first observe that if $\lim_{n \to +\infty} z_n = x_0$ ($z_n \in \mathbb{C}^+$) and $t(z_n) \to t_0$, then

$$1 - x_0\, (ct_0)^2 \neq 0, \tag{7.2}$$

$$1 + \frac{ct_0 \, \overline{\lambda}_l}{1 - x_0\, (ct_0)^2} \neq 0, \; l = 1, \ldots, \overline{M}. \tag{7.3}$$

Indeed, if (7.2) does not hold, Eq. (7.1) leads to $t_0 = 0$, a contradiction because $1 - x_0\, (ct_0)^2$ was assumed equal to 0. Similarly, if (7.3) does not hold, the limit of $t(z_n)$ cannot be finite. Therefore, matrix T_0 defined by

$$T_0 = -\left(x_0 \left[I + \frac{ct_0}{1 - x_0\, (ct_0)^2} \, R \right] \right)^{-1} \tag{7.4}$$

is well defined, and it holds that $T(z_n) \to T_0$ and that $t_0 = \frac{1}{M} \mathrm{Tr} R T_0$. In particular, for $i = 1, 2$, $T(z_{i,n}) \to T_i$ where T_i is defined by (7.4) when $t_0 = t_i$, $i = 1, 2$, and

$t_i = \frac{1}{M}\mathrm{Tr}RT_i$. Using the equation (6.1) for $z = z_{i,n}$, we obtain immediately that

$$\begin{pmatrix} t(z_{1,n}) - t(z_{2,n}) \\ z_{1,n}t(z_{1,n}) - z_{2,n}t(z_{2,n}) \end{pmatrix} = \begin{pmatrix} u_0(z_{1,n}, z_{2,n}) & v_0(z_{1,n}, z_{2,n}) \\ z_{1,n}z_{2,n}v_0(z_{1,n}, z_{2,n}) & u_0(z_{1,n}, z_{2,n}) \end{pmatrix}$$
$$\times \begin{pmatrix} t(z_{1,n}) - t(z_{2,n}) \\ z_{1,n}t(z_{1,n}) - z_{2,n}t(z_{2,n}) \end{pmatrix} + \begin{pmatrix} (z_{1,n} - z_{2,n}) \frac{1}{M}\mathrm{Tr}T(z_{1,n})RT(z_{2,n}) \\ 0 \end{pmatrix}, \quad (7.5)$$

where $u_0(z_1, z_2)$ and $v_0(z_1, z_2)$ are defined by

$$u_0(z_1, z_2) = c\, \frac{cz_1t(z_1)cz_2t(z_2)\, \frac{1}{M}\mathrm{Tr}(RT(z_1)RT(z_2))}{(1 - z_1(ct(z_1))^2)\,(1 - z_2(ct(z_2))^2)} \quad (7.6)$$

and

$$v_0(z_1, z_2) = c\, \frac{\frac{1}{M}\mathrm{Tr}(RT(z_1)RT(z_2))}{(1 - z_1(ct(z_1))^2)\,(1 - z_2(ct(z_2))^2)} \quad (7.7)$$

for $z_i \in \mathbb{C}^+$, $i = 1, 2$. Taking the limit, we obtain that

$$\begin{pmatrix} t_1 - t_2 \\ x_0(t_1 - t_2) \end{pmatrix} = \begin{pmatrix} u_0(x_0, x_0) & v_0(x_0, x_0) \\ x_0^2 v_0(x_0, x_0) & u_0(x_0, x_0) \end{pmatrix} \begin{pmatrix} t_1 - t_2 \\ x_0(t_1 - t_2) \end{pmatrix},$$

where $u_0(x_0, x_0)$ and $v_0(x_0, x_0)$ are defined by replacing $z_i, t(z_i), T(z_i)$ by x_0, t_i, T_i in (7.6, 7.7) for $i = 1, 2$. If the determinant $(1 - u_0(x_0, x_0))^2 - x_0^2 v_0(x_0, x_0)^2 \neq 0$ of the above linear system is non zero, it of course holds that $t_1 = t_2$.

We now consider the case where $(1 - u_0(x_0, x_0))^2 - x_0^2 v_0(x_0, x_0)^2 = 0$. We denote by $u_i(x_0)$ and $v_i(x_0)$, $i = 1, 2$ the limits of $u(z_{i,n})$ and $v(z_{i,n})$, $i = 1, 2$ when $n \to +\infty$. We recall that $u(z)$ and $v(z)$ are defined by (6.13) and (6.14) respectively. It is clear that $u_i(x_0)$ and $v_i(x_0)$ coincide with (6.13) and (6.14) when $(z, t(z), T(z))$ are replaced by (x_0, t_i, T_i) respectively. (6.11) thus implies that

$$(1 - u_i(x_0))^2 - x_0^2 v_i(x_0)^2 \geq 0 \quad (7.8)$$

for $i = 1, 2$. Using the Schwartz inequality and (6.19) as in the uniqueness proof of the solutions of Eq. (6.1) (see Proposition 6.1), it is easily seen that

$$\begin{aligned} |(1 - u_0(x_0, x_0))^2 - x_0^2 (v_0(x_0, x_0))^2| &\geq (1 - \sqrt{u_1(x_0)}\sqrt{u_2(x_0)})^2 - x_0^2 v_1(x_0)v_2(x_0) \\ &\geq (1 - u_1(x_0))(1 - u_2(x_0)) - x_0^2 v_1(x_0)v_2(x_0) \\ &\geq \sqrt{(1 - u_1(x_0))^2 - x_0^2 v_1(x_0)^2}\sqrt{(1 - u_2(x_0))^2 - x_0^2 v_2(x_0)^2} \geq 0. \quad (7.9) \end{aligned}$$

Therefore, $(1 - u_0(x_0, x_0))^2 - x_0^2 v_0(x_0, x_0)^2 = 0$ implies that the Schwartz inequalities and the inequalities (6.19) used to establish (7.9) are equalities. Hence, it holds that $|u_0(x_0, x_0)|^2 = u_1(x_0)u_2(x_0)$, or equivalently $|\frac{1}{M}\mathrm{Tr}(RT_1RT_2)| = (\frac{1}{M}\mathrm{Tr}(RT_1T_1^*R))^{1/2}(\frac{1}{M}\mathrm{Tr}(RT_2T_2^*R))^{1/2}$. This implies that $T_1 = aT_2^*$ for some constant $a \in \mathbb{C}$. Moreover, as $t_i = \frac{1}{M}\mathrm{Tr}(RT_i)$ for $i = 1, 2$, it must hold that $t_1 = at_2^*$. (7.9) follows from (6.19) $\{a = b = 1,\ c = u_1(x_0),\ d = u_2(x_0)\}$ and $\{a = (1 - u_1(x_0))^2,\ b = (1 - u_2(x_0))^2,\ c = x_0^2 v_1^2,\ d = x_0^2 v_2^2\}$. Since all these terms are positive real numbers, $\sqrt{ab} - \sqrt{cd} = \sqrt{a-c}\sqrt{b-d}$ if and only if $ad = bc$. It gives us

$$\begin{aligned} &u_1(x_0) = u_2(x_0), \\ &(1 - u_1(x_0))^2 x_0^2 v_2(x_0)^2 = (1 - u_2(x_0))^2 x_0^2 v_1(x_0)^2. \end{aligned} \tag{7.10}$$

Since $x_0 \neq 0$ and $v_1(x_0) > 0$, the inequality $(1 - u_1(x_0))^2 - x_0^2 v_1(x_0)^2 \geq 0$ implies that $u_1(x_0) \neq 1$. Hence, $u_1(x_0) < 1$ and (7.10) implies that $v_1(x_0) = v_2(x_0)$. From the definition of u_i and v_i one can notice that $u_i(x_0) = c^2 x_0^2 |t_i|^2 v_i(x_0)$. Which gives us immediately $|t_1|^2 = |t_2|^2$ and, as a consequence, $|a| = 1$. Using once again the fact that $v_1(x_0) = v_2(x_0)$ and $T_1 = aT_2^*$, we obtain that

$$\frac{|a|^2 \frac{1}{M} \mathrm{Tr}(T_2^* R R T_2)}{|1 - x_0 c^2 a^2 (t_2^*)^2|^2} = \frac{\frac{1}{M} \mathrm{Tr}(R T_2 T_2^* R)}{|1 - x_0 c^2 t_2^2|^2}.$$

The numerators of both sides are equal and non zero, from what follows that the denominators are also equal, i.e.

$$|1 - x_0 c^2 a^2 (t_2^*)^2| = |1 - x_0 c^2 t_2^2|.$$

We remark that if w and z satisfy $|1-w| = |1-z|$ and $|w| = |z|$, then, either $w = z$, either $w = \bar{z}$. We use this remark for $w = x_0 c^2 t_2^2$ and $z = x_0 c^2 a^2 (t_2^*)^2$. If $w = z$, it holds that $a^2 (t_2^*)^2 = t_2^2 \Rightarrow t_1^2 = t_2^2$ and since $\mathrm{Im} t_i \geq 0$ we conclude $t_1 = t_2$. If $w = \bar{z}$, we have $a^2 (t_2^*)^2 = (t_2^*)^2$. If $t_2 = 0$ then it also holds that $t_1 = 0$. Otherwise, we have $a = \pm 1$. If $a = 1$, the condition $\mathrm{Im} t_i \geq 0$, leads to the conclusion that t_1 and t_2 are real and coincide. We finally consider the case $a = -1$. We recall $T_1 = aT_2^* = -T_2^*$. Therefore, it holds that

$$x_0 I_M - \frac{x_0 t_2^*}{1 - x_0 c^2 (t_2^*)^2} R = -x_0 I_M - \frac{x_0 t_2^*}{1 - x_0 c^2 (t_2^*)^2} R,$$

which is impossible, since $x_0 \neq 0$. This completes the proof of Lemma (7.2). ∎

Lemmas 7.2 and 7.1, and their corresponding proofs imply the following result.

Proposition 7.1. *For each $x > 0$, $\lim_{z \to x, z \in \mathbb{C}^+} t(z) = t(x)$ exists. Moreover, $1 - x(ct(x))^2 \neq 0$, and matrix $(I + \frac{ct(x)}{1 - x(ct(x))^2} R)$ is invertible. Therefore, $\lim_{z \to x, z \in \mathbb{C}^+} T(z) = T(x)$ where $T(x)$ represents matrix $T(x) = \left(-x(I + \frac{ct(x)}{1-x(ct(x))^2} R)\right)^{-1}$. Moreover, $t(x)$ is solution of the equation*

$$t(x) = \frac{1}{M} \mathrm{Tr}(RT(x)). \tag{7.11}$$

If $u(x)$ and $v(x)$ represent the terms defined by (6.13) and (6.14) for $z = x$, then it holds that

$$1 - u(x) > 0 \tag{7.12}$$

and

$$(1 - u(x))^2 - x^2 (v(x))^2 \geq 0 \tag{7.13}$$

for each $x \neq 0$*. Moreover, the inequality (7.13) is strict if* $x \in \mathbb{R}^+ - \mathcal{S}$*. If moreover* $\mathrm{Im}(t(x)) > 0$*, then, we have*

$$1 - u(x) - xv(x) = 0. \tag{7.14}$$

Proof. It just remains to justify (7.12), (7.13), and (7.14). As function $z \to t(z)$ is analytic on $\mathbb{C} - \mathcal{S}$, $x \to t(x)$ is differentiable on $\mathbb{R}^+ - \mathcal{S}$. As $(t(x))^{'} > 0$ and $(xt(x))^{'} > 0$ hold on $\mathbb{R}^+ - \mathcal{S}$, the arguments used in the context of Remark 6.1 are also valid on $\mathbb{R}^+ - \mathcal{S}$, thus justifying (7.12) and the strict inequality in (7.13). $1 - u(x) \geq 0$ and inequality (7.13) also hold on $\mathcal{S} - \{0\}$ by letting $z \to x$, $z \in \mathbb{C}^+$ in Proposition 6.1. As $v(x) > 0$ for each $x \neq 0$, the strict inequality (7.12) is a consequence of (7.13).

In order to prove (7.14), we use the second component of (6.15), and remark that it implies that

$$\mathrm{Im}(t(x)) = (u(x) + xv(x))\, \mathrm{Im}(t(x)).$$

Therefore, $\mathrm{Im}(t(x)) > 0$ leads to (7.14). ■

We also add the following useful result which shows that the real part of $t(x)$ is negative for each $x > 0$.

Proposition 7.2. *For each* $x \in \mathbb{R}^{+*}$*, it holds that* $\mathrm{Re}(t(x)) < 0$.

Proof. It is easily checked that

$$\begin{pmatrix} \mathrm{Re}(t(z) \\ \mathrm{Re}(zt(z)) \end{pmatrix} = \begin{pmatrix} u(z) & -v(z) \\ -|z|^2 v(z) & u(z) \end{pmatrix} \begin{pmatrix} \mathrm{Re}(t(z) \\ \mathrm{Re}(zt(z)) \end{pmatrix} + \begin{pmatrix} -\mathrm{Re}(z)\frac{1}{M}\mathrm{Tr}(RT(z)(T(z))^*) \\ -|z|^2\frac{1}{M}\mathrm{Tr}(RT(z)(T(z))^* \end{pmatrix} \tag{7.15}$$

for each $z \in \mathbb{C} - \mathcal{S}$. Moreover, as all the terms coming into play in (7.15) have a finite limit when $z \to x$ when $x \neq 0$, (7.15) remains valid on $\mathbb{R}^*$. For $z = x$, the first component of (7.15) leads to

$$\mathrm{Re}(t(x))(1 - u(x) + xv(x)) = -x\frac{1}{M}\mathrm{Tr}(RT(x)T(x)^*). \tag{7.16}$$

Proposition 7.1 implies that $1 - u(x) > 0$, when $x \in \mathbb{R}^*$. Therefore, $1 - u(x) + xv(x)$ is strictly positive as well, and it holds that

$$\mathrm{Re}(t(x)) = -x\,\frac{1}{1 - u(x) + xv(x)}\,\frac{1}{M}\mathrm{Tr}(RT(x)T(x)^*). \tag{7.17}$$

Therefore, $x > 0$ implies that $\mathrm{Re}(t(x)) < 0$ as expected. ■

We now study the behaviour of $t(z)$ when $z \to 0$. We first establish that $\lim_{z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*} |t(z)| = +\infty$, and then that $\lim_{z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*} zt(z) = 0$ if $c \leq 1$ and is strictly negative if $c > 1$. We recall that $t(x)$ for $x > 0$ is defined by $t(x) = \lim_{z \to x, z \in \mathbb{C}^+} t(z)$. For this, we establish various lemmas.

Lemma 7.3. *It holds that* $\lim_{z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*} |t(z)| = +\infty$.

Proof. We assume that the statement of the Lemma does not hold, i.e. that it exists a sequence of elements of $\mathbb{C}^+ \cup \mathbb{R}^*$ $(z_n)_{n \geq 1}$ such that $\lim_{n \to +\infty} z_n = 0$ and $t(z_n) \to t_0$. (6.1) and (7.11) imply that

$$z_n t(z_n) = -\frac{1}{M} \sum_{l=1}^{\overline{M}} \frac{m_l \overline{\lambda}_l}{1 + \frac{ct(z_n)\overline{\lambda}_l}{1 - z_n (ct(z_n))^2}}. \tag{7.18}$$

$1 + \frac{ct(z_n)\overline{\lambda}_l}{1 - z_n(ct(z_n))^2}$ clearly converges towards $1 + ct_0 \overline{\lambda}_l$. As the left hand side of (7.18) converges towards 0, for each l, $1 + ct_0 \overline{\lambda}_l$ cannot vanish. Therefore, matrix $I + ct_0 R$ is invertible, and taking the limit of (7.18) gives

$$\frac{1}{M} \mathrm{Tr} R (I + ct_0 R)^{-1} = 0.$$

As $\mathrm{Im} \frac{1}{M} \mathrm{Tr} R(I + ct_0 R)^{-1}$ cannot be zero if t_0 is not real, t_0 must be real. We now use the observation that $|z_n| v(z_n) \leq 1$ for each n (see Lemma 6.1 and Proposition 7.1 if $z_n \in \mathbb{C}^+ \cup \mathbb{R}^{+*}$, and Remark 6.1 if $z_n \in \mathbb{R}^{-*}$). As $|1 - z_n(ct(z_n))^2|^2 \to 1$, $|z_n| v(z_n)$ bounded implies that $|z_n| \frac{1}{M} \mathrm{Tr}(RT(z_n)RT(z_n)^*)$ is bounded. It is easy to check that

$$|z_n| \frac{1}{M} \mathrm{Tr}(RT(z_n)RT(z_n)^*) = \frac{1}{|z_n|} \frac{1}{M} \mathrm{Tr}(R(I + ct_0 R)^{-1} R (I + ct_0 R)^{-1}) + \mathcal{O}(1).$$

Therefore, the boundedness of $|z_n| \frac{1}{M} \mathrm{Tr}(RT(z_n)RT(z_n)^*)$ implies that $\frac{1}{M} \mathrm{Tr}(R(I + ct_0 R)^{-1} R (I + ct_0 R)^{-1}) = 0$ which is of course impossible. ■

Lemma 7.4. *Consider a sequence* $(z_n)_{n \geq 1}$ *of elements of* $\mathbb{C}^+ \cup \mathbb{R}^*$ *such that* $\lim_{n \to +\infty} z_n = 0$. *Then, the set* $(z_n t(z_n))_{n \geq 1}$ *is bounded.*

Proof. We assume that $(z_n t(z_n))_{n \geq 1}$ is not bounded. Therefore, one can extract from $(z_n)_{n \geq 1}$ a subsequence, still denoted $(z_n)_{n \geq 1}$, such that $\lim_{n \to +\infty} |z_n t(z_n)| = +\infty$. Then,

$$\frac{ct(z_n)}{1 - z_n (ct(z_n))^2} = \frac{1}{\frac{1}{ct(z_n)} - z_n t(z_n)} \to 0.$$

Therefore,

$$-\frac{1}{M} \mathrm{Tr} R \left(I + \frac{ct(z_n)}{1 - z_n (ct(z_n))^2} R \right)^{-1} \to -\frac{1}{M} \mathrm{Tr} R.$$

This is a contradiction because the above term coincides with $z_n t(z_n)$ which cannot converge towards a finite limit. ■

Lemma 7.5. *Assume that* $(z_{1,n})_{n \geq 1}$ *and* $(z_{2,n})_{n \geq 1}$ *are sequences of elements of* $\mathbb{C}^+ \cup \mathbb{R}^*$ *such that* $\lim_{n \to +\infty} z_{i,n} = 0$ *and* $\lim_{n \to +\infty} z_{i,n} t(z_{i,n}) = \delta_i$ *for* $i = 1, 2$. *Then,* $\delta_1 = \delta_2$.

Proof. We first remark that $|t(z_{i,n})| \to +\infty$ for $i = 1, 2$. Equation (6.1) implies immediately that

$$zt(z) = \left(zct(z) - \frac{1}{ct(z)}\right) \frac{1}{M} \mathrm{Tr} R \left(R + \frac{1}{ct(z)} - zct(z)\right)^{-1}. \tag{7.19}$$

As $\frac{1}{ct(z_{i,n})} \to 0$, $z_{i,n} ct(z_{i,n}) - \frac{1}{ct(z_{i,n})} \to c\delta_i$ for $i = 1, 2$. If $\delta_i \neq 0$, Eq. (7.19) thus implies that $c\frac{1}{M}\mathrm{Tr}R\left(R + \frac{1}{ct(z_{i,n})} - z_{i,n}ct(z_{i,n})\right)^{-1}$ converges towards 1, which implies that matrix $R - c\delta_i I$ is invertible. Therefore, either $\delta_i = 0$, either δ_i is a solution of the equation

$$1 = c\frac{1}{M}\mathrm{Tr}R(R - c\delta_i I)^{-1} \tag{7.20}$$

or equivalently, δ_i verifies

$$\delta_i = c\delta_i \frac{1}{M}\mathrm{Tr}R(R - c\delta_i I)^{-1}. \tag{7.21}$$

We note that the solutions of this equation are real, so that $\delta_i \in \mathbb{R}$ for $i = 1, 2$. Eq. (7.5) leads to

$$\begin{aligned} z_{1,n}t(z_{1,n}) - z_{2,n}t(z_{2,n}) = z_{1,n}z_{2,n}v_0(z_{1,n}, z_{2,n})(t(z_{1,n}) - t(z_{2,n})) \\ + u_0(z_{1,n}, z_{2,n})(z_{1,n}t(z_{1,n}) - z_{2,n}t(z_{2,n})). \end{aligned}$$

It is straightforward to check that $z_{1,n}z_{2,n}v_0(z_{1,n}, z_{2,n})(t(z_{1,n}) - t(z_{2,n})) \to 0$ and that $u_0(z_{1,n}, z_{2,n}) \to u_0(0,0) = c\frac{1}{M}\mathrm{Tr}R(R - c\delta_1 I)^{-1}R(R - c\delta_2 I)^{-1}$. Therefore, we obtain that

$$\delta_1 - \delta_2 = u_0(0,0)(\delta_1 - \delta_2). \tag{7.22}$$

We recall that $|u_0(z_{1,n}, z_{2,n})| \leq \sqrt{u(z_{1,n})}\sqrt{u(z_{2,n})} \leq 1$. Moreover, we observe that $u(z_{i,n}) \to u_i(0) = c\frac{1}{M}\mathrm{Tr}R(R - c\delta_i I)^{-1}R(R - c\delta_i I)^{-1}$ and that $0 < u_i(0) \leq 1$. The Schwartz inequality leads to

$$|u_0(0,0)| \leq \sqrt{u_1(0)}\sqrt{u_2(0)} \leq 1. \tag{7.23}$$

If the Schwartz inequality (7.23) is strict, $|u_0(0,0)| < 1$, and $\delta_1 = \delta_2$. We now assume that $u_0(0,0) = \sqrt{u_1(0)}\sqrt{u_2(0)} = 1$. This implies that

$$R - c\delta_1 I = \kappa(R - c\delta_2 I)$$

for some real constant κ, or equivalently, $\overline{\lambda}_l - c\delta_1 = \kappa(\overline{\lambda}_l - c\delta_2)$ for each $l = 1, \ldots, \overline{M}$. If R is not a multiple of I, κ must be equal to 1, since otherwise, we would have $\overline{\lambda}_l = \overline{\lambda}_{l'}$ for each l, l'. $\kappa = 1$ implies immediately that $\delta_1 = \delta_2$. We finally consider the case where $R = \sigma^2 I$. Then, (7.21) implies that δ_i is solution of $\delta_i \frac{\sigma^2 c}{\sigma^2 - c\delta_i} = \delta_i$, i.e. $\delta_i = 0$ or

$$\delta_i = \sigma^2\left(\frac{1}{c} - 1\right). \tag{7.24}$$

We now check that $\delta_1 = 0, \delta_2 = \sigma^2\left(\frac{1}{c} - 1\right)$ or $\delta_2 = 0, \delta_1 = \sigma^2\left(\frac{1}{c} - 1\right)$ is impossible. If this holds, $u_1(0)$ and $u_2(0)$ cannot be both equal to 1, and $|u_0(0,0)| < 1$. Therefore, (7.22) leads to a contradiction, and $\delta_1 = \delta_2$ is equal either to 0, either to $\sigma^2\left(\frac{1}{c} - 1\right)$. ■

Lemmas 7.4 and 7.5 imply the following corollary.

Corollary 7.1. *If $c \leq 1$, it holds that*

$$\lim_{z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*} zt(z) = 0 \tag{7.25}$$

and that

$$\mu(\{0\}) = 0. \tag{7.26}$$

Proof. Lemmas 7.4 and 7.5 lead to the conclusion that $\lim_{z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*} zt(z) = \delta$ where δ is either equal to 0, either coincides with a solution of the equation (7.21). In order to precise this, we remark that $t(x) > 0$ if $x < 0$ implies that $\delta \leq 0$. Therefore, δ coincides with a non positive solution of equation (7.21). If $c \leq 1$, it is clear that (7.21) has no strictly negative solutions. Therefore, (7.25) is established. (7.26) is a direct consequence of the identity

$$\mu(\{0\}) = \lim_{z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*} -zt(z).$$

■

In order to address the case where $c > 1$ and to precise the behaviour of $\text{Im}(t(z))$ when $z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*$ if $c \leq 1$, we have to evaluate $z(t(z))^2$ when $z \to 0$. The following Lemma holds.

Lemma 7.6.

- *If $c = 1$, it holds that* $\lim_{z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*} |z(t(z))^2| = +\infty$.
- *If $c < 1$,*

$$\lim_{z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*} z(t(z))^2 = -\frac{1}{c(1-c)}. \tag{7.27}$$

- *If $c > 1$, the assumption* $\lim_{z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*} zt(z) = \delta = 0$ *implies that* $\lim_{z \to 0, z \in \mathbb{C}^+ \cup \mathbb{R}^*} z(t(z))^2 = -\frac{1}{c(1-c)}$, *a contradiction because the above limit is necessarily negative. Hence, δ is non zero and coincides with the strictly negative solution of Eq. (7.21), and $\mu(\{0\}) = -\delta$.*

Proof. (6.1) implies that

$$z(t(z))^2 = -\frac{1}{M}\text{Tr}R\left(\frac{I}{t(z)} + \frac{c}{1 - z(ct(z))^2}R\right)^{-1}. \tag{7.28}$$

We assume in the course of this proof that $\delta = 0$ (if $c \leq 1$, this property holds). We first establish the first item of Lemma 7.6. We assume that $c = 1$ and that there exists a sequence $(z_n)_{n \in \mathbb{C}^+ \cup \mathbb{R}^*}$ such that $z_n \to 0$ and $z_n t(z_n)^2 \to \alpha$. As

$|t(z_n)| \to +\infty$, (7.28) leads to $\alpha = \alpha - 1$, a contradiction. Therefore, if $c = 1$, $\lim_{z\to 0, \to \mathbb{C}^+\cup\mathbb{R}^*} |zt(z)^2| = +\infty$ as expected.

We now establish the 2 last items. For this, we establish that if $c \neq 1$, then, $|zt(z)^2|$ is bounded when $z \in \mathbb{C}^+ \cup \mathbb{R}^*$ and z is close from 0. For this, we assume the existence of a sequence $(z_n)_{n\geq 1}$ of elements of $\mathbb{C}^+ \cup \mathbb{R}^*$ such that $z_n \to 0$ and $|z_n t(z_n)^2| \to +\infty$. Then, it holds that

$$1 = -\frac{1}{M}\mathrm{Tr}R\left(z_n t(z_n) I + \frac{c z_n t(z_n)^2}{1 - z_n (c t(z_n))^2} R\right)^{-1}.$$

As $|z_n t(z_n)^2| \to +\infty$, $\frac{c z_n t(z_n)^2}{1 - z_n (c t(z_n))^2} \to -\frac{1}{c}$. Condition $z_n t(z_n) \to 0$ thus implies that $c = 1$, a contradiction. Using again (7.28), we obtain immediately that if $z_n (t(z_n))^2 \to \alpha$, then $\alpha = -\frac{1}{c(c-1)}$. As $|zt(z)^2|$ remains bounded when $z \in \mathbb{C}^+ \cup \mathbb{R}^*$ is close from 0, this implies that $\lim_{z\to 0, z\in\mathbb{C}^+\cup\mathbb{R}^*} z(t(z))^2 = -\frac{1}{c(1-c)}$ as expected. Taking $z \in \mathbb{R}^{-*}$ leads to the conclusion that the above limit is negative. When $c > 1$, this is a contradiction because $-\frac{1}{c(1-c)}$ is positive. Therefore, if $c > 1$, δ, the limit of $zt(z)$, cannot be equal to 0. Hence, δ coincides with the strictly negative solution of (7.21) and $\mu(\{0\}) = -\delta > 0$. This completes the proof of the Lemma. ∎

Putting all the pieces together, we obtain the following characterization of μ_N.

Theorem 7.1. *The density $f_N(x)$ of μ_N w.r.t. the Lebesgue measure is a continuous function on $\mathbb{R}^{+*}$, and is given by $f_N(x) = \frac{1}{\pi}\mathrm{Im}(t_N(x))$ for each $x > 0$. If $c_N \leq 1$, μ_N is absolutely continuous, and if $c_N > 1$, then $d\mu_N(x) = f_N(x)dx + \mu_N(\{0\})\delta_0$. 0 belongs to $\mathcal{S}_N$, and the interior $\mathcal{S}_N^\circ$ of $\mathcal{S}_N$ is given by*

$$\mathcal{S}_N^\circ = \{x \in \mathbb{R}^+, \mathrm{Im}(t(x)) > 0\}. \tag{7.29}$$

If moreover $c_N < 1$, it holds that

$$f_N(x) \simeq \frac{1}{\pi}\frac{1}{\sqrt{x\, c_N(1-c_N)}} \tag{7.30}$$

when $x \to 0^+$, while if $c_N = 1$,

$$f_N(x) \simeq \frac{1}{\pi}\frac{\sqrt{3}}{2}\left(\frac{1}{M}\mathrm{Tr}\, R^{-1}\right)^{-1/3}\frac{1}{x^{2/3}}. \tag{7.31}$$

Proof. $t(z)$ is not analytic in a neighbourhood of 0, hence, $0 \in \mathcal{S}$. As $\lim_{z\to x, z\in\mathbb{C}^+} t(z) = t(x)$ exists for $x \neq 0$, Theorem 2.1 of [35] implies that if $\mathcal{A} \subset \mathbb{R}^{+*}$ is a Borel set of zero Lebesgue measure, then $\mu(\mathcal{A}) = \int_{\mathcal{A}} f(x)dx = 0$. The continuity of f on $\mathbb{R}^{+*}$ is a also a consequence of [35].

We now prove (7.30). For this, we remark that (7.27) implies that

$$\lim_{x\to 0, x>0} x(t(x))^2 = -\frac{1}{c(1-c)}. \tag{7.32}$$

As $\mathrm{Im}(t(x)) \geq 0$ for each $x \neq 0$, (7.32) implies that $t(x) \simeq \frac{i}{\sqrt{x}\sqrt{c(1-c}}$ when $x \to 0^+$, or equivalently that $\frac{1}{\pi}\mathrm{Im}(t(x)) \simeq \frac{1}{\pi}\frac{1}{\sqrt{x\,c(1-c)}}$.

It remains to establish (7.31). For this, we first prove that

$$\lim_{x \to 0, x>0} x^2 (t(x))^3 = \left(\frac{1}{M}\mathrm{Tr}\, R_N^{-1}\right)^{-1}. \tag{7.33}$$

For this, we write (7.11) as

$$\frac{1}{M}\mathrm{Tr}R\left(-xt(x)I + \frac{1}{1-\frac{1}{x(t(x))^2}}R\right)^{-1} = 1. \tag{7.34}$$

As $c = 1$, $xt(x) \to 0$ and $|x(t(x))^2| \to +\infty$ when $x \to 0, x > 0$. The left hand side of (7.34) can be expanded as

$$\frac{1}{M}\mathrm{Tr}R\left(-xt(x)I + \frac{1}{1-\frac{1}{x(t(x))^2}}R\right)^{-1} = 1 - \frac{1}{x(t(x))^2} + \frac{1}{M}\mathrm{Tr}\, R^{-1}\, xt(x) + xt(x)\epsilon_1(x) + \frac{1}{x(t(x))^2}\epsilon_2(x),$$

where $\epsilon_1(x)$ and $\epsilon_2(x)$ converge towards 0 when $x \to 0, x > 0$. Therefore, (7.34) implies that

$$\frac{1}{M}\mathrm{Tr}\, R^{-1}\, xt(x) - \frac{1}{x(t(x))^2} = xt(x)\tilde{\epsilon}_1(x) + \frac{1}{x(t(x))^2}\tilde{\epsilon}_2(x),$$

where $\tilde{\epsilon}_1(x)$ and $\tilde{\epsilon}_2(x)$ converge towards 0 when $x \to 0, x > 0$. This leads immediately to (7.33). As function $x \to x^2(t(x))^3$ is continuous on $\mathbb{R}^{+*}$, it holds that

$$\lim_{x \to 0, x>0} x^{2/3} t(x) = e^{2ik\pi/3}\left(\frac{1}{M}\mathrm{Tr}\, R^{-1}\right)^{-1/3},$$

where k is equal to $0, 1$ or 2. If $k = 0$, the real part of $t(x)$ must be positive if x is close enough from 0. Lemma 7.2 thus leads to a contradiction. If $k = 2$, $\mathrm{Im}(t(x)) < 0$ for x small enough, a contradiction as well. Hence, k is equal to 1. Therefore,

$$\lim_{x \to 0, x>0} x^{2/3}\mathrm{Im}(t(x)) = \sin 2\pi/3 \left(\frac{1}{M}\mathrm{Tr}\, R^{-1}\right)^{-1/3}. \tag{7.35}$$

This completes the proof of (7.31). ■

We now show that function $x \to t(x)$ and $x \to f(x)$ possess a power series expansion in a neighbourhood of each point of $\mathcal{S}_N^\circ$. More precisely:

Proposition 7.3. *If $x_0 > 0$ and $\mathrm{Im}(t(x_0)) > 0$, then, t and f can be expanded as*

$$t(x) = \sum_{k=0}^{+\infty} a_k (x - x_0)^k, f(x) = \sum_{k=0}^{+\infty} b_k (x - x_0)^k$$

when $|x - x_0|$ *is small enough.*

Proof. As in [35] and [13], the proof is based on the holomorphic implicit function theorem (see [9]). We denote $t(x_0)$ by t_0. Then, Eq. (7.11) at point x_0 can be written as $h(x_0, t_0) = 0$ where function $h(z, t)$ is defined by

$$h(z,t) = t - \frac{1}{M}\mathrm{Tr}\left(R\left(-z(I + \frac{ct}{1-z(ct)^2}\,R)\right)^{-1}\right).$$

As $x_0 > 0$ and $\mathrm{Im}(t_0) > 0$, function $(z, t) \to h(z, t)$ is holomorphic in a neighbourhood of (x_0, t_0). It is easy to check that

$$\left(\frac{\partial h}{\partial t}\right)_{x_0,t_0} = 1 - u_0(x_0, x_0) - x_0^2 v_0(x_0, x_0), \tag{7.36}$$

where we recall that functions u_0 and v_0 are given by (7.6) and (7.7). Following the proof of Lemma 7.2, we obtain immediately that $1 - u_0(x_0, x_0) - x_0^2 v_0(x_0, x_0) = 0$ implies that $T(x_0) = aT(x_0)^*$, and that $t_0 = at_0^*$ for some $a \in \mathbb{C}$. The arguments of the above proof then lead to the conclusion that $t_0 = t_0^*$, a contradiction because $\mathrm{Im}(t(x_0)) > 0$. Hence, $\left(\frac{\partial h}{\partial t}\right)_{x_0,t_0} \neq 0$. The holomorphic implicit function theorem thus implies that it exists a function $z \to \tilde{t}(z)$, holomorphic in a neigbourhood N of x_0, verifying $\tilde{t}(x_0) = t_0$ and $h(z, \tilde{t}(z)) = 0$ for each $z \in N$. Moreover, condition $\mathrm{Im}(t_0) = \mathrm{Im}(\tilde{t}(x_0)) > 0$ implies that $\mathrm{Im}(\tilde{t}(z)) > 0$ and $\mathrm{Im}(z\tilde{t}(z)) > 0$ if $|z - x_0| < \epsilon$ for ϵ small enough. Therefore, if $z \in \mathbb{C}^+$ and $|z - x_0| < \epsilon$, it must hold that $\tilde{t}(z) = t(z)$ (see Proposition 6.1). Hence, $t(x) = \lim_{z\to x, z\in\mathbb{C}^+} t(z)$ must coincide with $\tilde{t}(x)$ when $|x - x_0| < \epsilon$. As $\tilde{t}(z)$ is holomorphic in a neighbourhood of x_0, function $x \to t(x)$ can be expanded as

$$t(x) = \sum_{k=0}^{+\infty} a_k (x - x_0)^k$$

when $|x - x_0| < \epsilon$. This immediately implies that f possesses a power series expansion in the interval $(x_0 - \epsilon, x_0 + \epsilon)$. ■

We finally use the above results in order the study measure ν_N associated to the Stieltjes transform

$$t_{N,\nu}(z) = \frac{1}{M}\mathrm{Tr}T_N(z).$$

As ν_N and μ_N are absolutely continuous one with respect each other, $d\nu_N(x)$ can also be written as $d\nu_N(x) = g_N(x)dx + \nu_N(\{0\})\delta_0$. Using the identity

$$\frac{1}{M}\mathrm{Tr}\left[-z\left(I + \frac{ct(z)}{1-z(ct(z))^2}R\right)T(z)\right] = 1.$$

we obtain immediately that

$$t_\nu(z) = -\frac{1}{z} - \frac{c(t(z))^2}{1-z(ct(z))^2}. \tag{7.37}$$

If $x > 0$, $t_\nu(x) = \lim_{z \to x, z \in \mathbb{C}^+} t_\nu(z)$ exists, and is given by the righthandside of (7.37) when $z = x$. Hence, for $x > 0$, $g(x) = \frac{1}{\pi} \mathrm{Im}(t_\nu(x))$, i.e.

$$g(x) = -\frac{1}{\pi} \frac{c\,\mathrm{Im}((t(x))^2)}{|1 - x(ct(x))^2|^2}. \tag{7.38}$$

If $c > 1$, $|zt(z)^2| \to +\infty$ if $z \to 0$. (7.37) thus implies that $\nu_N(\{0\}) = \lim_{z \to 0} -zt_\nu(z)$ coincides with $1 - \frac{1}{c}$, which, of course, is not surprising. We now evaluate the behaviour of g when $x \to 0, x > 0$ and $c \leq 1$.

Proposition 7.4. *If $c < 1$, it holds that*

$$g(x) \simeq_{x \to 0} \frac{1}{\pi} \frac{1}{\sqrt{c\,(1-c)}} \frac{1}{M} \mathrm{Tr}(R^{-1}) \frac{1}{\sqrt{x}} \tag{7.39}$$

while if $c = 1$, it holds that

$$g(x) \simeq_{x \to 0} \frac{1}{\pi} \frac{\sqrt{3}}{2} \left(\frac{1}{M} \mathrm{Tr}(R^{-1}) \right)^{2/3} \frac{1}{x^{2/3}}. \tag{7.40}$$

Proof. Using Eq. (7.28), we obtain after some algebra that

$$z(t(z))^2 + \frac{1}{c(1-c)} \simeq_{z \to 0} \frac{1}{M} \mathrm{Tr} R^{-1} \frac{1}{c^2(1-c)^3} \frac{1}{t(z)}.$$

As $t(x) \simeq_{x \to 0, x > 0} \frac{i}{\sqrt{x}\sqrt{c(1-c)}}$, we get that

$$\mathrm{Im}((t(x))^2) \simeq -i \frac{1}{M} \mathrm{Tr} R^{-1} \frac{1}{1-c} \frac{1}{(c(1-c))^{3/2}} \frac{1}{\sqrt{x}}.$$

Therefore, (7.38) immediately leads to (7.39). (7.40) is an immediate consequence of (7.35). ∎

Proposition 7.4 means in practice that if $c_N \leq 1$, a number of eigenvalues of matrix $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ are close from 0. Moreover, the rate of convergence of g_N towards $+\infty$ is higher if $c_N = 1$, showing that in this case, the proportion of eigenvalues close to 0 is even larger than if $c_N < 1$.

We finally mention that $t_\nu(x)$ and $g(x)$ possess a power expansion around each-point $x_0 \in \mathcal{S}^\circ$. This is an obvious consequence of Proposition 7.3 and of the above expressions of $t_\nu(x)$ and of $g(x)$ in terms of $t(x)$.

7.2. *Characterization of $\mathcal{S}_N$.*

We denote by $w_N(z)$ the function defined by

$$w_N(z) = -\frac{(1 - z(c_N t_N(z))^2)}{c_N t_N(z)} = z c_N t_N(z) - \frac{1}{c_N t_N(z)}. \tag{7.41}$$

It is clear that w is analytic on $\mathbb{C} - \mathcal{S}$, that $\mathrm{Im}(w(z)) > 0$ if $z \in \mathbb{C}^+$, that $w(x) = \lim_{z \to x, z \in \mathbb{C}^+} w(z)$ exists for each $x \in \mathbb{R}^*$, and that the limit still exists if $x = 0$. If we denote this limit by $w(0)$, then, it holds that $w(0) = 0$ if $c \leq 1$ and that $w(0) = c\delta$ if $c > 1$, where we recall that δ is defined as the solution of (7.20). Moreover, $w(x)$ is real if and only if $t(x)$ is real. Therefore, the interior $\mathcal{S}^o$ of $\mathcal{S}$ is also given by

$$\mathcal{S}^o = \{x \in \mathbb{R}^+, \mathrm{Im}(w(x)) > 0\}. \tag{7.42}$$

Moreover, as $t(x)'$ and $(xt(x))'$ are strictly positive if $x \in \mathbb{R} - \mathcal{S}$, the derivative $w'(x)$ of $w(x)$ w.r.t. x is also strictly positive on $\mathbb{R} - \mathcal{S}$. Using the equation $t(z) = \frac{1}{M}\mathrm{Tr}RT(z)$, we obtain immediately that $t(z)$ can be expressed in terms of $w(z)$ as

$$t(z) = \frac{1}{z}\, w(z)\, \frac{1}{M}\mathrm{Tr}R\, (R - w(z)I)^{-1}. \tag{7.43}$$

(7.41) implies that

$$1 + ct(z)w(z) - z(ct(z))^2 = 0. \tag{7.44}$$

Plugging (7.43) into (7.44), we obtain immediately that $w_N(z)$ verifies the equation

$$\phi_N(w_N(z)) = z, \tag{7.45}$$

where $\phi_N(w)$ is defined by

$$\phi_N(w) = c_N w^2\, \frac{1}{M}\mathrm{Tr}R_N\, (R_N - wI)^{-1} \left(c_N\, \frac{1}{M}\mathrm{Tr}R_N\, (R_N - wI)^{-1} - 1 \right). \tag{7.46}$$

Observe that (7.45) holds not only on $\mathbb{C} - \mathcal{S}$, but also for each $x \in \mathcal{S}$. Therefore, it holds that $\phi(w(x)) = x$ for each $x \in \mathbb{R}$. For each $x \in \mathbb{R} - \mathcal{S}$, it thus holds that $\phi^{'}(w(x))\, w^{'}(x) = 1$. Therefore, as $w^{'}(x) > 0$ if $x \in \mathbb{R} - \mathcal{S}$, $w(x)$ satisfies $\phi^{'}(w(x)) > 0$ for each $x \in \mathbb{R} - \mathcal{S}$. This implies that if $x \in \mathbb{R} - \mathcal{S}$, then $w(x)$ is a real solution of the polynomial equation $\phi(w) = x$ for which $\phi^{'}(w) > 0$. Moreover, Proposition 7.2 implies that if $x \in \mathbb{R}^+ - \mathcal{S}$, then, $t(x) = \mathrm{Re}(t(x))$ is strictly negative. Eq. (7.43) for $z = x$ thus leads to the conclusion that if $x > 0$ does not belong to $\mathcal{S}$, then $w(x)$ also verifies $w(x)\frac{1}{M}\mathrm{Tr}R\, (R - w(x)I)^{-1} < 0$. If $x < 0$, then, $t(x)$ is this time strictly positive and $w(x)$ still verifies $w(x)\frac{1}{M}\mathrm{Tr}R\, (R - w(x)I)^{-1} < 0$. This discussion leads to the following Proposition.

Proposition 7.5. *If $x \in \mathbb{R} - \mathcal{S}$, then $w(x)$ verifies the following properties:*

$$\phi(w(x)) = x,\ \phi^{'}(w(x)) > 0,\ w(x)\, \frac{1}{M}\mathrm{Tr}R\, (R - w(x)I)^{-1} < 0. \tag{7.47}$$

As shown below, if $x \in \mathbb{R} - \mathcal{S}$, the properties (7.47) characterize $w(x)$ among the set of all solutions of the equation $\phi(w) = x$ and allow to identify the support as the subset of $\mathbb{R}^+$ for which the equation $\phi(w) = x$ has no real solution satisfying the conditions (7.47). These results follow directly from an elementary study of function $w \to \phi(w)$.

We first consider the case $c \leq 1$, and identify the values of $x > 0$ for which the equation $\phi(w(x)) = x$ has a real solution verifying (7.47), and those for which such a solution does not exist. It is easily seen that if $x > 0$, all the real solutions of the equation $\phi(w) = x$ are strictly positive. Therefore, the third condition in (7.47) is equivalent to $\frac{1}{M}\text{Tr}R\,(R - w(x)I)^{-1} < 0$. We denote $\omega_{1,N} < \omega_{2,N} < \ldots < \omega_{\overline{M},N}$ the (necessarily real) $\overline{M}$ roots of $\frac{1}{M}\text{Tr}R_N(R_N - wI)^{-1} = \frac{1}{c_N}$ and by $\mu_{1,N} < \mu_{2,N} < \ldots < \mu_{\overline{M}-1,N}$ the roots of $\frac{1}{M}\text{Tr}R_N(R_N - wI)^{-1} = 0$. As $c \leq 1$, it is easily seen that $\omega_1 \geq 0$, and that $\omega_1 < \overline{\lambda}_{\overline{M}} < \mu_1 < \omega_2 < \overline{\lambda}_{\overline{M}-1} < \ldots < \mu_{\overline{M}-1} < \omega_{\overline{M}} < \overline{\lambda}_1$. It is clear that $\frac{1}{M}\text{Tr}R(R - wI)^{-1} < 0$ if and only if $w \in (\overline{\lambda}_{\overline{M}}, \mu_1) \cup \ldots \cup (\overline{\lambda}_2, \mu_{\overline{M}-1}) \cup (\overline{\lambda}_1, +\infty)$.

For $x > 0$, the equation $\phi(w) = x$ is easily seen to be a polynomial equation of degree $2\overline{M} + 1$. Therefore, $\phi(w) = x$ has $2\overline{M} + 1$ solutions. For each $x > 0$, this equation has at least $2\overline{M} - 1$ real solutions that cannot coincide with $w(x)$ if $x \in (\mathcal{S}^\circ)^c$:

- $\overline{M}$ solutions belong to $]\omega_1, \overline{\lambda}_{\overline{M}}[, \ldots,]\omega_{\overline{M}}, \overline{\lambda}_1[$. None of these solutions may correspond to $w(x)$ if $x \in (\mathcal{S}^\circ)^c$ because $\frac{1}{M}\text{Tr}R(R - wI)^{-1} > 0$ at these points.
- On each interval $]\overline{\lambda}_{\overline{M}}, \mu_1[, \ldots,]\overline{\lambda}_2, \mu_{\overline{M}-1}[$, the equation $\phi(w) = x$ has a real solution at which ϕ' is negative. Therefore, $\phi(w) = x$ has $\overline{M} - 1$ extra real solutions that are not equal to $w(x)$ if $x \in (\mathcal{S}^\circ)^c$.

As $\phi_N(w) \to +\infty$ if $w \to \overline{\lambda}_{1,N}, w > \overline{\lambda}_{1,N}$ and that $\phi_N(w) \to +\infty$ if $w \to +\infty$, it exists at least one point in $]\overline{\lambda}_{1,N}, +\infty[$ at which $\phi_N^{'}$ vanishes. This point is moreover unique because otherwise, $\phi_N(w) = x$ would have more than $2\overline{M} + 1$ solutions for certain values of x. We denote by $w_{+,N}$ this point, and remark that if $x > x_{+,N} = \phi_N(w_{+,N})$, $\phi_N(w) = x$ has $2\overline{M} + 1$ real solutions: the $2\overline{M} - 1$ solutions that were introduced below, and 2 extra solutions that belong to $]\overline{\lambda}_1, w_+[$ and $]w_+, +\infty[$ respectively. Therefore, $w(x)$ is real, and it is easily seen that $w(x)$ coincides with the solution that belongs to $]w_+, +\infty[$. This implies that $]x_+, +\infty[\subset \mathbb{R} - \mathcal{S}$.

If $\phi^{'}(w)$ does not vanish on $]\overline{\lambda}_{\overline{M}}, \mu_1[\cup \ldots \cup]\overline{\lambda}_2, \mu_{\overline{M}-1}[$, for each $x \in]0, x_+[$, ϕ is decreasing on these intervals. Therefore, none of the real solutions of $\phi(w) = x$ match with the properties of $w(x)$ when $x \in \mathbb{R}^+ - \mathcal{S}$. Therefore, $w(x)$ must be a complex number: $\phi(w) = x$ has thus $2\overline{M} - 1$ real solutions, and a pair of complex conjugate roots: $w(x)$ is the positive imaginary part solution. In this case, $x \in \mathcal{S}^\circ$, and the support $\mathcal{S}$ coincides with $[0, x_+]$.

We illustrate such a behaviour when $\overline{M} = 3$. In the context of Fig. 1, the support is reduced to the single interval $[0, x_+]$ because $\phi^{'}(w) \neq 0$ for $w \in [\overline{\lambda}_3, \mu_1] \cup [\overline{\lambda}_2, \mu_2]$.

In order to precise the support when $\phi^{'}$ vanishes in $]\overline{\lambda}_{\overline{M}}, \mu_1[\cup \ldots \cup]\overline{\lambda}_2, \mu_{\overline{M}-1}[$, we need to characterize the corresponding zeros. For this, we first justify that $\phi^{'}$ cannot have a multiplicity 2 zero. Assume for example that $\phi^{'}$ has a multiplicity 2 zero in $]\overline{\lambda}_{\overline{M}+1-l}, \mu_l[$, and denote by w_l this zero. Then, if $x_l = \phi(w_l)$, the equation $\phi(w) = x_l$ has $2\overline{M} - 1$ simple real roots, and the multiplicity 3 root w_l. Therefore,

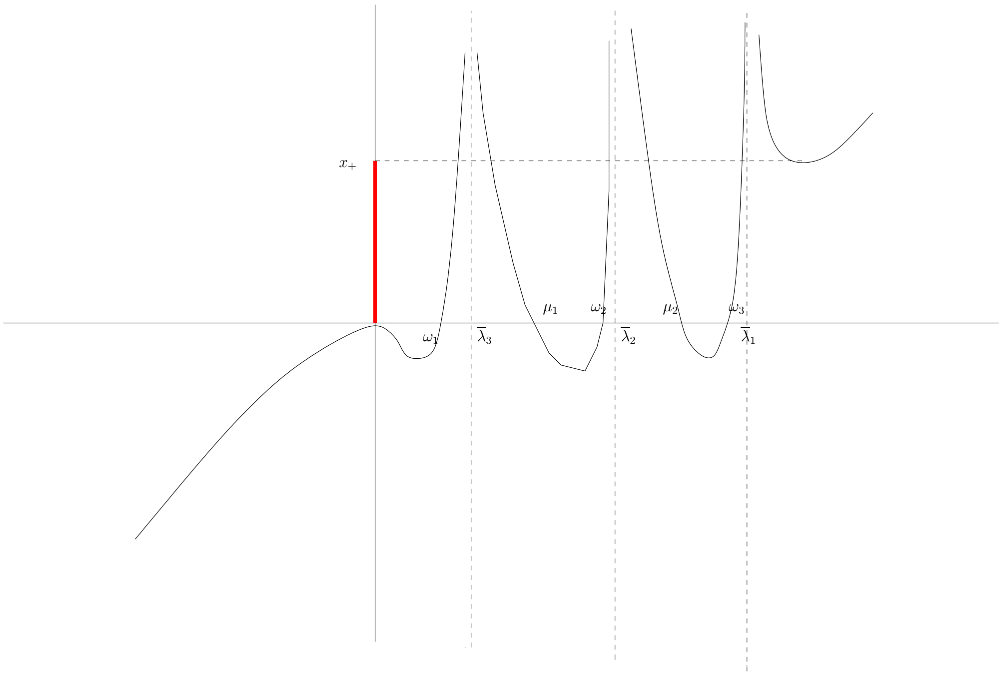

Fig. 1. Typical representation of $\phi\,(w)$ as a function of w for $\overline{M} = 3$. There is no local maximum on $[\overline{\lambda}_3, \mu_1]$ and on $[\overline{\lambda}_2, \mu_2]$, so that $\mathcal{S} = [0, x_+]$.

the equation $\phi(w) = x_l$ has $2\overline{M} + 2$ roots (counting multiplicities), a contradiction. We now establish the following useful result.

Proposition 7.6. *The number of local extrema of* ϕ_N *in* $]\overline{\lambda}_{\overline{M}}, \mu_1[\cup \ldots \cup]\overline{\lambda}_2, \mu_{\overline{M}-1}[$ *is an even number, say* $2q$*, with* $0 \le q \le \overline{M} - 1$*. If* $q \ge 1$*, we denote the arguments of these extrema by* $w^+_{1,N} < w^-_{2,N} < w^+_{2,N} < \ldots < w^+_{q-1,N} < w^-_{q,N}$*, then* $x^+_{1,N} = \phi_N(w^+_{1,N}), x^-_{2,N} = \phi_N(w^-_{2,N}), \ldots, x^+_{q-1,N} = \phi_N(w^+_{q-1,N}), x^-_{q,N} = \phi_N(w^-_{q,N})$ *verify*

$$x^+_{1,N} < x^-_{2,N} < x^+_{2,N} < \ldots < x^+_{q-1,N} < x^-_{q,N}. \tag{7.48}$$

Moreover, for each l*, the interval* $]\overline{\lambda}_{\overline{M}-(l-1)}, \mu_l[$ *contains at most one interval* $[w^+_{p,N}, w^-_{p+1,N}]$*, and* $x^+_{p,N}$ *(resp.* $x^-_{p+1,N}$*) is a local minimum (resp. local maximum) of* ϕ_N*.*

Proof. We establish that if $w_1,\, w_2 \in \{w^+_1,\, w^-_2,\, \ldots,\, w^+_{q-1},\, w^-_q\}$ such that $w_1 > w_2$, the images $x_1 = \phi(w_1)$ and $x_2 = \phi(w_2)$ are also satisfy $x_1 > x_2$. The goal is to show that ratio $(x_1 - x_2)/(w_1 - w_2)$ is always positive. For more convenience we put $f_n = \frac{c_N}{M} \mathrm{Tr} R_N (R_N - w_n I_M)^{-1} = \frac{c_N}{M} \sum_1^{\bar{M}} \frac{\overline{\lambda}_i m_i}{\overline{\lambda}_i - w_n}$ for $n = 1, 2$. With this and (7.46) we can rewrite

$$x_n = \phi(w_n) = w_n^2 f_n (f_n - 1) = w_n^2 p_n (p_n - 1), \tag{7.49}$$

where $p_n = 1 - f_n$. Let us notice that extremes w_1 and w_2 are by definition such that f_1 and f_2 are negative. Using directly (7.49) for x_1 and x_2 we can write

$$\frac{x_1 - x_2}{w_1 - w_2} = \frac{(w_1^2 p_1^2 - w_2^2 p_2^2) - (w_1^2 p_1 - w_2^2 p_2)}{w_1 - w_2}$$
$$= (w_1 p_1 + w_2 p_2)\frac{w_1 p_1 - w_2 p_2}{w_1 - w_2} - \frac{w_1^2 p_1 - w_2^2 p_2}{w_1 - w_2}. \quad (7.50)$$

With the definition of $f_{1,2}$ the first term of (7.50) can be expanded as

$$\frac{w_1 p_1 - w_2 p_2}{w_1 - w_2} = 1 + \frac{c}{M}\sum_{l=1}^{\bar{M}} \frac{\overline{\lambda}_i m_i}{w_1 - w_2}\left(\frac{w_2}{\overline{\lambda}_i - w_2} - \frac{w_1}{\overline{\lambda}_i - w_1}\right)$$
$$= 1 - \frac{c}{M}\sum_{1}^{\bar{M}} \frac{\overline{\lambda}_i^2 m_i}{(\overline{\lambda}_i - w_1)(\overline{\lambda}_i - w_2)}.$$

And similarly the second one as

$$\frac{w_1^2 p_1 - w_2^2 p_2}{w_1 - w_2} = (w_1 + w_2) + \frac{c}{M}\sum_{1}^{\bar{M}} \frac{\overline{\lambda}_i m_i}{w_1 - w_2}\left(\frac{w_2^2}{\overline{\lambda}_i - w_2} - \frac{w_1^2}{\overline{\lambda}_i - w_1}\right)$$
$$= (w_1+w_2)\left(1 - \frac{c}{M}\sum_{1}^{\bar{M}} \frac{\overline{\lambda}_i^2 m_i}{(\overline{\lambda}_i - w_1)(\overline{\lambda}_i - w_2)}\right) + w_1 w_2 \frac{c}{M}\sum_{1}^{\bar{M}} \frac{\overline{\lambda}_i m_i}{(\overline{\lambda}_i - w_1)(\overline{\lambda}_i - w_2)}.$$

Putting the last two equation in (7.50) we obtain

$$\frac{x_1 - x_2}{w_1 - w_2} = (w_1 p_1 + w_2 p_2 - w_1 - w_2)\left(1 - \frac{c}{M}\sum_{1}^{\bar{M}} \frac{\overline{\lambda}_i^2 m_i}{(\overline{\lambda}_i - w_1)(\overline{\lambda}_i - w_2)}\right)$$
$$- w_1 w_2 \frac{c}{M}\sum_{1}^{\bar{M}} \frac{\overline{\lambda}_i m_i}{(\overline{\lambda}_i - w_1)(\overline{\lambda}_i - w_2)} = -(w_1 f_1 + w_2 f_2)$$
$$\times \left(1 - \frac{c}{M}\sum_{1}^{\bar{M}} \frac{\overline{\lambda}_i^2 m_i}{(\overline{\lambda}_i - w_1)(\overline{\lambda}_i - w_2)}\right) - w_1 w_2 \frac{c}{M}\sum_{1}^{\bar{M}} \frac{\overline{\lambda}_i m_i}{(\overline{\lambda}_i - w_1)(\overline{\lambda}_i - w_2)}.$$

Now we recall that $-f_n$ is positive as well as $w_1, w_2 > 0$ from what we have $-(w_1 f_1 + w_2 f_2) > 0$. That allows us to use the inequality

$$\frac{1}{(\overline{\lambda}_i - w_1)(\overline{\lambda}_i - w_2)} \leq \frac{1}{2}\left(\frac{1}{(\overline{\lambda}_i - w_1)^2} + \frac{1}{(\overline{\lambda}_i - w_2)^2}\right)$$

and to write

$$\frac{x_1 - x_2}{w_1 - w_2} \geq -(w_1 f_1 + w_2 f_2)\left(1 - \frac{c}{2M}\sum_{1}^{\bar{M}} \frac{\overline{\lambda}_i^2 m_i}{(\overline{\lambda}_i - w_1)^2} - \frac{c}{2M}\sum_{1}^{\bar{M}} \frac{\overline{\lambda}_i^2 m_i}{(\overline{\lambda}_i - w_2)^2}\right)$$
$$- w_1 w_2 \frac{c}{M}\sum_{1}^{\bar{M}} \frac{\overline{\lambda}_i m_i}{(\overline{\lambda}_i - w_1)(\overline{\lambda}_i - w_2)}.$$

It is easy to check that $\frac{c}{M}\sum\frac{\overline{\lambda}_i^2 m_i}{(\overline{\lambda}_i-w)^2}=f(w)+wf'(w)$. Using this we can rewrite last inequality as

$$\frac{x_1-x_2}{w_1-w_2}\geq-\frac{1}{2}(w_1f_1+w_2f_2)\left(2-f_1-w_1f_1'-f_2-w_2f_2'\right) - w_1w_2\frac{c}{M}\sum_1^{\bar{M}}\frac{\overline{\lambda}_i m_i}{(\overline{\lambda}_i-w_1)(\overline{\lambda}_i-w_2)}. \quad (7.51)$$

Taking the derivatives of the expression (7.49), we obtain that $\phi'(w_n)=2w_nf_n^2-2w_nf_n+2w_n^2f_nf_n'-w_n^2f_n'$. By definition, $w_{1,2}$ are extremes of function $\phi(w)$, i.e. $\phi'(w_{1,2})=0$. This gives immediately $f_n+w_nf_n'-1=\frac{w_nf_n'}{2f_n}$. After putting this into (7.51) and regrouping terms we obtain

$$\frac{x_1-x_2}{w_1-w_2}\geq\frac{1}{4}(w_1f_1+w_2f_2)\left(\frac{w_1f_1'}{f_1}+\frac{w_2f_2'}{f_2}\right)-w_1w_2\frac{c}{M}\sum_1^{\bar{M}}\frac{\overline{\lambda}_i m_i}{(\overline{\lambda}_i-w_1)(\overline{\lambda}_i-w_2)}$$
$$=\frac{1}{4}(w_1^2f_1'+w_2^2f_2')+\frac{1}{4}w_1w_2\left(f_1'\frac{f_2}{f_1}+f_2'\frac{f_1}{f_2}\right)-w_1w_2\frac{c}{M}\sum_1^{\bar{M}}\frac{\overline{\lambda}_i m_i}{(\overline{\lambda}_i-w_1)(\overline{\lambda}_i-w_2)}.$$

Finally, we denote by I_1, I_2, I_3 the three terms of the r.h.s and show that $I_1+\frac{1}{2}I_3$ and $I_2+\frac{1}{2}I_3$ can be presented as the sum of positive terms. Using again the definition of $f_{1,2}$ we expand $I_1+\frac{1}{2}I_3$ as

$$\frac{1}{4}\left(w_1^2f_1'+w_2^2f_2'-2w_1w_2\frac{c}{M}\sum_1^{\bar{M}}\frac{\overline{\lambda}_i m_i}{(\overline{\lambda}_i-w_1)(\overline{\lambda}_i-w_2)}\right)$$
$$=\frac{c}{4M}\sum\overline{\lambda}_i m_i\Big(\frac{w_1^2}{(\overline{\lambda}_i-w_1)^2}+\frac{w_2^2}{(\overline{\lambda}_i-w_2)^2}-\frac{2w_1w_2}{(\overline{\lambda}_i-w_1)(\overline{\lambda}_i-w_2)}\Big)$$
$$=\frac{c}{4M}\sum\overline{\lambda}_i m_i\left(\frac{w_1}{\overline{\lambda}_i-w_1}-\frac{w_2}{\overline{\lambda}_i-w_2}\right)^2.$$

Similarly, $I_2+\frac{1}{2}I_3$ can be written as

$$\frac{1}{4}w_1w_2\left(f_1'\frac{f_2}{f_1}+f_2'\frac{f_1}{f_2}-2\frac{c}{M}\sum_1^{\bar{M}}\frac{\overline{\lambda}_i m_i}{(\overline{\lambda}_i-w_1)(\overline{\lambda}_i-w_2)}\right)$$
$$=w_1w_2\frac{c}{4M}\sum\overline{\lambda}_i m_i\left(\frac{f_2/f_1}{(\overline{\lambda}_i-w_1)^2}+\frac{f_1/f_2}{(\overline{\lambda}_i-w_2)^2}-\frac{2}{(\overline{\lambda}_i-w_1)(\overline{\lambda}_i-w_2)}\right)$$
$$=w_1w_2\frac{c}{4M}\sum\overline{\lambda}_i m_i\left(\frac{\sqrt{f_2/f_1}}{\overline{\lambda}_i-w_1}-\frac{\sqrt{f_1/f_2}}{\overline{\lambda}_i-w_2}\right)^2.$$

This shows that $x_1-x_2>0$, and that (7.48) holds. It remains to justify that each interval $(]\overline{\lambda}_{\overline{M}-(l-1)},\mu_l[)_{l=1,\ldots,\overline{M}-1}$ contains at most one interval $[w_{p,N}^+,w_{p+1,N}^-]$. Assume that the interval $]\overline{\lambda}_{\overline{M}-(l-1)},\mu_l[$ contains 2 intervals $[w_{p_1,N}^+,w_{p_1+1,N}^-]$ and

$[w^+_{p_2,N}, w^-_{p_2+1,N}]$ with $p_1 < p_2$. Then, it also holds that $[w^+_{p_1+1,N}, w^-_{p_1+2,N}] \subset]\overline{\lambda}_{\overline{M}-(l-1)}, \mu_l[$. $x^+_{p_1,N}$ is necessarily a local minimum because $x^+_{p_1,N} < x^-_{p_1+1,N}$ while $x^-_{p_1+1,N}$ must be a local maximum. The same property holds for $x^+_{p_1+1,N}$ and $x^-_{p_1+2,N}$. However, this contradicts the property $x^-_{p_1+1,N} < x^+_{p_1+1,N}$. This completes the proof of Proposition 7.6. ■

Proposition 7.6 allows to identify the support $\mathcal{S}_N$.

Corollary 7.2. *When $c_N \leq 1$, the support $\mathcal{S}_N$ is given by*

$$\mathcal{S}_N = [0, x^+_{1,N}] \cup [x^-_{2,N}, x^+_{2,N}] \cup \ldots [x^-_{q,N}, x_{+,N}]. \tag{7.52}$$

Proof. If x belongs to the interior of the righthandside of (7.52), $\phi(w) = x$ has only $2\overline{M} - 1$ real solutions. This implies that the 2 remaining roots are complex valued, i.e. that $x \in \mathcal{S}^\circ$. This leads to the conclusion that

$$]0, x^+_{1,N}[\cup]x^-_{2,N}, x^+_{2,N}[\cup \ldots]x^-_{q,N}, x_{+,N}[\subset \mathcal{S}^\circ$$

and that

$$[0, x^+_{1,N}] \cup [x^-_{2,N}, x^+_{2,N}] \cup \ldots [x^-_{q,N}, x_{+,N}] \subset \mathcal{S}.$$

Conversely, if $x \in \mathbb{R}^+ - \Big([0, x^+_{1,N}] \cup [x^-_{2,N}, x^+_{2,N}] \cup \ldots [x^-_{q,N}, x_{+,N}]\Big)$, the equation $\phi(w) = x$ has $2\overline{M} + 1$ real solutions, which implies that $w(x)$ is real. Therefore,

$$\mathbb{R}^+ - \Big([0, x^+_{1,N}] \cup [x^-_{2,N}, x^+_{2,N}] \cup \ldots [x^-_{q,N}, x_{+,N}]\Big) \subset \mathbb{R}^+ - \mathcal{S}$$

or equivalently,

$$\mathcal{S} \subset [0, x^+_{1,N}] \cup [x^-_{2,N}, x^+_{2,N}] \cup \ldots [x^-_{q,N}, x_{+,N}].$$

This completes the proof of Corollary (7.2). ■

We illustrate the above behaviour when $\overline{M} = 3$. In the context of Fig. 2, $\phi^{'}$ vanishes on $[\overline{\lambda}_3, \mu_1]$ and not on $[\overline{\lambda}_2, \mu_2]$. The support thus coincides with $\mathcal{S} = [0, x^+_1] \cup [x^-_2, x_+]$.

When matrix R_N is reduced to $R_N = \sigma^2 I$, i.e. $\overline{M} = 1$ and $\overline{\lambda}_1 = \sigma^2$, the support of course coincides with $\mathcal{S}_N = [0, x_{+,N}]$, and $x_{+,N}$ is given by

$$x_{+,N} = \sigma^4 c_N \left(1 + \frac{1}{\frac{1+\sqrt{1+8c_N}}{2}}\right)^2 \left(c_N + \frac{1+\sqrt{1+8c_N}}{2}\right). \tag{7.53}$$

Moreover, $w_{+,N}$ is equal to

$$w_{+,N} = \sigma^2 \left(1 + \frac{1+\sqrt{1+8c_N}}{2}\right). \tag{7.54}$$

(7.53) and (7.54) are in accordance with the results of [22].

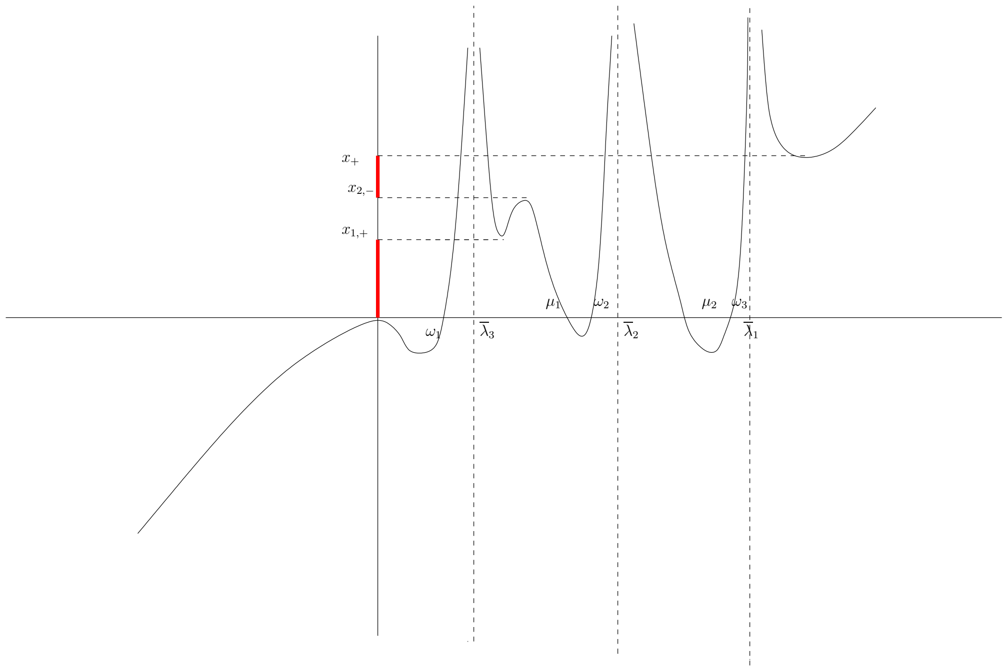

Fig. 2. Typical representation of $\phi\,(w)$ as a function of w for $\overline{M} = 3$. There are 2 local extrema on $[\overline{\lambda}_3, \mu_1]$ and no local maximum on $[\overline{\lambda}_2, \mu_2]$, so that $\mathcal{S} = [0, x_1^+] \cup [x_2^-, x_+]$.

We now briefly address the case $c_N > 1$. The behaviour of ϕ_N is essentially the same as if $c_N \leq 1$, except that the first root $\omega_{1,N}$ of the equation $\frac{1}{M}\mathrm{Tr}R_N(R_N - wI)^{-1} = \frac{1}{c_N}$ is now strictly negative. As $\phi_N(0) = 0$, this implies that it exists $w_{-,N} \in (\omega_{1,N}, 0)$ for which $\phi_N^{'}(w_{-,N}) = 0$. Moreover, this point is unique, otherwise, the equation $\phi_N(w) = x$ would have more than $2\overline{M} + 1$ roots for certain values of $x > 0$. $x_{-,N} = \phi_N(w_{-,N}) > 0$ is thus a local maximum of ϕ_N whose argument is strictly negative. We also notice that $\phi_N(w) > 0$ if $0 < w < \overline{\lambda}_{\overline{M}}$. Apart these differences, the behaviour of ϕ_N for $w > \overline{\lambda}_{\overline{M}}$ remains the same as if $c_N \leq 1$. In particular, Proposition 7.6 still holds true. However, we remark that if $0 < x < x_{-,N}$, the equation $\phi_N(w) = x$ has still $2\overline{M} - 1$ real solutions that are strictly positive, and 2 extra real roots, the smallest one being less than $w_{-,N}$ and the other one being negative and largest that $w_{-,N}$. This implies that $w_N(x)$ is real. We also notice that $w_N(x)$ coincides with the smallest extra negative root because it satisfies conditions (7.47). Hence, the interval $]0, x_{-,N}[$ is included into $\mathbb{R}^+ - \mathcal{S}_N$. If $\phi_N^{'}$ does not vanish on $]\overline{\lambda}_{\overline{M}}, \mu_1[\cup \ldots \cup]\overline{\lambda}_2, \mu_{\overline{M}-1}[$, for $x \in]x_{-,N}, x_{+,N}[$, the equation $\phi_N(w) = x$ has only $2\overline{M} - 1$ real solutions that do not satisfy conditions (7.47) and 2 extra complex conjugates solutions. Therefore, $]x_{-,N}, x_{+,N}[\subset \mathcal{S}_N^\circ$ and $[x_{-,N}, x_{+,N}] \subset \mathcal{S}_N$. Conversely, $]0, x_{-,N}[\cup]x_{+,N}, +\infty[\subset \mathbb{R}^+ - \mathcal{S}_N$, which implies that $\mathcal{S}_N \subset \{0\} \cup [x_{-,N}, x_{+,N}]$. As it was established above that $\{0\} \subset \mathcal{S}_N$, we deduce that $\mathcal{S}_N = \{0\} \cup [x_{-,N}, x_{+,N}]$ if ϕ'_N does not vanish on $]\overline{\lambda}_{\overline{M}}, \mu_1[\cup \ldots \cup]\overline{\lambda}_2, \mu_{\overline{M}-1}[$. If ϕ'_N vanishes on $]\overline{\lambda}_{\overline{M}}, \mu_1[\cup \ldots \cup]\overline{\lambda}_2, \mu_{\overline{M}-1}[$, i.e. if $q \geq 1$ (we recall that q is defined in Proposition 7.6), the support is given by

$$\mathcal{S}_N = \{0\} \cup [x_{-,N}, x^+_{1,N}] \cup [x^-_{2,N}, x^+_{2,N}] \cup \ldots [x^-_{q,N}, x_{+,N}]. \tag{7.55}$$

To justify this, we just need to establish that $x_{-,N} < x^+_{1,N}$, and to use the same arguments as in the proof of Corollary 7.2. To justify $x_{-,N} < x^+_{1,N}$, we put $w_1 = w_{-,N}, w_2 = w^+_{1,N}$, and follow step by step the arguments used to evaluate $\phi(w_2) - \phi(w_1) > 0$. We notice that in contrast with the context of the proof of Corollary 7.2, $w_1 < 0$ and $f_1 > 0$. However, $f_1 w_1$ is still negative, so that $-(w_1 f_1 + w_2 f_2)$ is still positive. This allows to conclude that all the inequalities used in the course of the proof of Corollary 7.2 remain valid, except the evaluation of the term $I_2 + I_3/2$ that needs the following simple modification: we express $I_2 + I_3/2$ as

$$-w_1 w_2 \frac{c}{4M} \sum \lambda_i m_i \times \left(\frac{-f_2/f_1}{(\lambda_i - w_1)^2} + \frac{-f_1/f_2}{(\lambda_i - w_2)^2} + \frac{2}{(\lambda_i - w_1)(\lambda_i - w_2)} \right).$$

As $-f_2/f_1$ and $-f_1/f_2$ are positive, it holds that

$$I_2 + I_3/2 = -w_1 w_2 \frac{c}{4M} \sum \lambda_i m_i \left(\frac{\sqrt{-f_2/f_1}}{\lambda_i - w_1} + \frac{\sqrt{-f_1/f_2}}{\lambda_i - w_2} \right)^2.$$

Therefore, $I_2 + I_3/2 > 0$, and $\phi(w_2) - \phi(w_1) > 0$ holds.

In order to unify the cases $c_N \leq 1$ and $c_N > 1$, we define $x_{-,N}$ for $c_N \leq 1$ by $x_{-,N} = 0$, and summarize the above discussion by the following result.

Theorem 7.2. *The support $\mathcal{S}_N$ is given by*

$$\mathcal{S}_N = \{0\} \mathbb{I}_{c_N > 1} \cup [x_{-,N}, x^+_{1,N}] \cup [x^-_{2,N}, x^+_{2,N}] \cup \ldots [x^-_{q,N}, x_{+,N}]. \tag{7.56}$$

We now establish that sequences $(w_{+,N})_{N \geq 1}$ and $(x_{+,N})_{N \geq 1}$ are bounded. In other words, for each N, the support $\mathcal{S}_N$ is included into a compact interval that does not depend on N.

Lemma 7.7.

$$\sup_{N \geq 1} w_{+,N} < +\infty, \ \sup_{N \geq 1} x_{+,N} < +\infty. \tag{7.57}$$

Proof. In order to prove this lemma, we use that $w_{+,N} > \lambda_{1,N}$ and that $\phi'_N(w_{+,N}) = 0$. It is easy to check that

$$\phi'_N(w) = 2c_N^2 w \frac{1}{M} \mathrm{Tr} R(wI - R)^{-1} - (c_N w)^2 \frac{1}{M} \mathrm{Tr} R(wI - R)^{-2}$$
$$- 2c_N^2 w \left(\frac{1}{M} \mathrm{Tr} R(wI - R)^{-1} \right)^2 - 2(c_N w)^2 \frac{1}{M} \mathrm{Tr} R(wI - R)^{-2} \frac{1}{M} \mathrm{Tr} R(wI - R)^{-1}.$$

For $w > b > \lambda_{1,N}$, it is clear that $\|(wI - R)^{-1}\| \leq \frac{1}{w-b}$. Writing that $w \frac{1}{M} \mathrm{Tr} R(wI - R)^{-1} = \frac{1}{M} \mathrm{Tr} R + \frac{1}{M} \mathrm{Tr} R^2 (wI - R)^{-1}$ and $w^2 \frac{1}{M} \mathrm{Tr} R(wI - R)^{-2} =$

$\frac{1}{M}\mathrm{Tr}R + w\left(\frac{1}{M}\mathrm{Tr}R(wI-R)^{-2}\right) - \frac{1}{M}\mathrm{Tr}R^2(wI-R)^{-1}$, we obtain immediately that $\phi_N'(w)$ can be written as

$$\phi_N'(w) = c_N^2 \frac{1}{M}\mathrm{Tr}R + \delta_N(w),$$

where $\delta_N(w)$ verifies $|\delta_N(w)| \leq \delta(w)$ and $w \to \delta(w)$ is a rational function of w that does not depend on N and which converges towards 0 when $w \to +\infty$. Therefore, for each $\eta > 0$, it exists $w_1 > b$ such that $\phi_N'(w) > c_N^2 \frac{1}{M}\mathrm{Tr}R - \eta$ for each $w \geq w_1$. As $c_N \to c_*$ and that $\frac{1}{M}\mathrm{Tr}R \geq a$, we obtain that $\phi_N'(w) > \frac{c_*^2}{2}a$ for $w \geq w_1$. As $\phi_N'(w_{+,N}) = 0$, we deduce from this that $w_{+,N} < w_1$. As w_1 does not depend on N, this establishes that $\sup_{N\geq 1} w_{+,N} < +\infty$. To prove that $x_{+,N}$ is bounded, we observe that $x_{+,N} = \phi_N(w_{+,N}) < \phi_N(w_1)$. As $w_1 > b$, it is easily seen that

$$\phi_N(w_1) < 2c_N^2 w_1^2 \left(\frac{b}{(w_1-b)^2} + \frac{b}{(w_1-b)}\right).$$

Therefore, sequences $(\phi_N(w_1))_{N\geq 1}$ and $(x_{+,N})_{N\geq 1}$ are bounded. This completes the proof of Lemma 7.7. ■

We finally provide a sufficient condition under which the support is reduced to $\mathcal{S}_N = [0, x_{+,N}]$ if $c_N < 1$ and to $\mathcal{S}_N = \{0\} \cup [x_{-,N}, x_{+,N}]$ if $c_N > 1$. More precisely, the following result holds.

Proposition 7.7. *Assume that there exists $\kappa > 0$ such that for each M large enough, the following condition holds:*

$$|\lambda_{k,N} - \lambda_{l,N}| \leq \kappa \left(\frac{|k-l|}{M}\right)^{1/2} \tag{7.58}$$

for each pair (k,l), $1 \leq k \leq l \leq M$. Then, for each M large enough, $\mathcal{S}_N = [0, x_{+,N}]$ if $c_N \leq 1$ and to $\mathcal{S}_N = \{0\} \cup [x_{-,N}, x_{+,N}]$ if $c_N > 1$.

Proof. We assume that (7.58) holds, and that $\mathcal{S}$ does not coincide with $[0, x_+]$ or $\mathcal{S} = \{0\} \cup [x_-, x_+]$, i.e. $\phi'(w)$ vanishes at a point w_0 such that $\lambda_1 < w_0 < \lambda_M$ and $\frac{1}{M}\mathrm{Tr}R(R - w_0 I)^{-1} < 0$. After some algebra, we obtain that w_0 satisfies:

$$\frac{1}{M}\mathrm{Tr}\left(R(R-w_0I)^{-1}\right)^2 = \frac{-\frac{1}{M}\mathrm{Tr}R(R-w_0I)^{-1}}{1 - 2c\frac{1}{M}\mathrm{Tr}R(R-w_0I)^{-1}}.$$

As $\frac{1}{M}\mathrm{Tr}R(R-w_0I)^{-1} < 0$, this implies that

$$\frac{1}{M}\mathrm{Tr}\left(R(R-w_0I)^{-1}\right)^2 = \frac{1}{M}\sum_{k=1}^{M}\left(\frac{\lambda_k}{\lambda_k - w_0}\right)^2 < -\frac{1}{M}\mathrm{Tr}R(R-w_0I)^{-1}$$

$$\leq \frac{1}{M}\sum_{k=1}^{M}\frac{\lambda_k}{|\lambda_k - w_0|}.$$

Jensen's inequality leads to $\left(\frac{1}{M}\sum_{k=1}^{M}\frac{\lambda_k}{|\lambda_k - w_0|}\right)^2 \leq \frac{1}{M}\sum_{k=1}^{M}\left(\frac{\lambda_k}{\lambda_k - w_0}\right)^2$. Therefore, we obtain that $\frac{1}{M}\sum_{k=1}^{M}\frac{\lambda_k}{|\lambda_k - w_0|} < 1$, and that

$$\frac{1}{M}\sum_{k=1}^{M}\left(\frac{\lambda_k}{\lambda_k - w_0}\right)^2 < 1. \tag{7.59}$$

We assume that $\lambda_{j_0} < w_0 < \lambda_{j_0+1}$. Then, hypothesis (2.7) and condition (7.58) imply that

$$\left(\frac{\lambda_k}{\lambda_k - w_0}\right)^2 > \frac{a^2}{\kappa^2}\frac{M}{(|k - j_0| + 1)}.$$

Hence, it must hold that

$$\frac{a^2}{\kappa^2}\sum_{k=1}^{M}\frac{1}{(|k - j_0| + 1)} < 1$$

for each M large enough, a contradiction because $\sum_{k=1}^{M}\frac{1}{(|k-j_0|+1)}$ is easily seen to be an unbounded term. ■

8. No eigenvalues outside the support.

In this paragraph, we establish the following result:

Theorem 8.1. *Assume that there exists $\epsilon > 0$, $\kappa_1 \in \mathbb{R}$, $\kappa_2 \in \mathbb{R} \cup \{+\infty\}$, $\kappa_2 > \kappa_1$ and an integer N_0 such that*

$$(\kappa_1 - \epsilon, \kappa_2 + \epsilon) \cap \mathcal{S}_N = \varnothing \qquad \forall\, N \geq N_0. \tag{8.1}$$

Then with probability one, no eigenvalues of $W_{f,N}W_{p,N}^*W_{p,N}W_{f,N}^*$ appears in $[\kappa_1, \kappa_2]$ for all N large enough.

We first remark that it is sufficient to consider the case where $\kappa_2 < +\infty$. To justify this claim, we recall that $\cup_{N\geq 1}\mathcal{S}_N$ is a compact subset (see Lemma 7.7), and notice that $\|W_{f,N}W_{p,N}^*W_{p,N}W_{f,N}^*\| \leq \|W_N\|^4$ where matrix W_N is defined by (2.5). Moreover, (3.1) implies that almost surely, for N large enough, $\|W_N\|^2 \leq b\,(1 + \delta + \sqrt{c_*})^2$ where $\delta > 0$. Therefore, almost surely, the largest eigenvalue of $W_{f,N}W_{p,N}^*W_{p,N}W_{f,N}^*$ is, for each N large enough, upperbounded by the nice constant $b^2\,(1 + \delta + \sqrt{c_*})^4$. This justifies that it is sufficient to assume that $\kappa_2 < +\infty$ in the following.

In order to establish Theorem 8.1, we use the Haagerup-Thornbjornsen approach ([15], see also [7]). The crucial step of the proof is the following Proposition.

Proposition 8.1. $\forall z \in \mathbb{C}^+$, *we have for N large enough,*

$$\mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}Q_N(z)\right\} = \frac{1}{M}\mathrm{Tr}T_N(z) + \frac{1}{N^2}r_N(z), \tag{8.2}$$

where r_N is holomorphic in $\mathbb{C}^+$ and satisfies

$$|r_N(z)| \leq P_1(|z|)P_2\left(\frac{1}{\text{Im}z}\right) \tag{8.3}$$

for each $z \in \mathbb{C}^+$, where P_1 and P_2 are nice polynomials.

Proof. To prove (8.2) we write

$$\mathbb{E}\left\{\frac{1}{ML}\text{Tr}Q_N(z)\right\} - \frac{1}{M}\text{Tr}T_N(z) = \frac{1}{ML}\text{Tr}\left[\mathbb{E}\left\{Q_N(z)\right\} - I_L \otimes S_N(z)\right] + \frac{1}{M}\text{Tr}\left[S_N(z) - T_N(z)\right].$$

As (5.28) holds, it is sufficient to establish that

$$\left|\frac{1}{M}\text{Tr}[S_N(z) - T_N(z)]\right| \leq \frac{1}{N^2}P_1(|z|)P_2(\text{Im}^{-1}z) \tag{8.4}$$

for some nice polynomial P_1 and P_2. In the following, we denote by $s_N(z)$ the function defined by

$$s_N(z) = \frac{1}{M}\text{Tr}R_N S_N(z). \tag{8.5}$$

It is clear that $s_N \in \mathcal{S}(\mathbb{R}^+)$. Moreover, if $\mu_{N,s}$ represents the associated positive measure, then we have

$$\mu_{N,s}(\mathbb{R}^+) = \frac{1}{M}\text{Tr}R_N, \ \int_{\mathbb{R}^+} \lambda \, d\mu_{N,s}(\lambda) = c_N \, \frac{1}{M}\text{Tr}R_N \, \frac{1}{M}\text{Tr}R_N^2 \tag{8.6}$$

(8.6) can be proved using the arguments of the proof of Proposition 6.1.

As $\frac{1}{M}\text{Tr}[S_N(z) - T_N(z)]$ is given by (6.23) for $F = I$, (8.4) appears equivalent to the property

$$\left|\frac{1}{M}\text{Tr}[R_N(S_N(z) - T_N(z))]\right| = |s_N(z) - t_N(z)| \leq \frac{1}{N^2}P_1(|z|)P_2(\text{Im}^{-1}z). \tag{8.7}$$

In order to prove (8.7), we define the following functions that appear formally similar to functions $u(z)$ and $v(z)$ defined by (6.13) and (6.14):

$$u_\alpha(z) = c\frac{|cz\alpha(z)|^2 \frac{1}{M}\text{Tr}(RS(z)S^*(z)R)}{|1 - z(c\alpha(z))^2|^2},$$

$$v_\alpha(z) = c\frac{\frac{1}{M}\text{Tr}(RS(z)S^*(z)R)}{|1 - z(c\alpha(z))^2|^2},$$

$$u_{t,\alpha}(z) = c\frac{|cz|^2 t(z)\alpha(z)\frac{1}{M}\text{Tr}(RS(z)T(z)R)}{(1 - z(c\alpha(z))^2)(1 - z(ct(z))^2)}, \tag{8.8}$$

$$v_{t,\alpha}(z) = c\frac{\frac{1}{M}\text{Tr}(RS(z)T(z)R)}{(1 - z(c\alpha(z))^2)(1 - z(ct(z))^2)}. \tag{8.9}$$

Using equation $t(z) = \frac{1}{M}\text{Tr}RT(z)$ and the definition of $s(z)$ and $S(z)$, we obtain easily that

$$\begin{pmatrix} (s(z)-t(z)) \\ z(s(z)-t(z)) \end{pmatrix} = \mathbf{D}_{t,\alpha}(z) \begin{pmatrix} (s(z)-t(z)) \\ z(s(z)-t(z)) \end{pmatrix} + \begin{pmatrix} \epsilon_1(z) \\ \epsilon_2(z) \end{pmatrix}$$

holds, where

$$\begin{aligned} \epsilon_1(z) &= (\alpha(z) - s(z))(zv_{t,\alpha}(z) + u_{t,\alpha}(z)), \\ \epsilon_2(z) &= z(\alpha(z) - s(z))(zv_{t,\alpha}(z) + u_{t,\alpha}(z)), \\ \mathbf{D}_{t,\alpha}(z) &= \begin{pmatrix} u_{t,\alpha}(z) & v_{t,\alpha}(z) \\ z^2 v_{t,\alpha}(z) & u_{t,\alpha}(z) \end{pmatrix}. \end{aligned}$$

This can also be written as

$$(\mathbf{I} - \mathbf{D}_{t,\alpha}(z)) \begin{pmatrix} (s(z)-t(z)) \\ z(s(z)-t(z)) \end{pmatrix} = \begin{pmatrix} \epsilon_1(z) \\ \epsilon_2(z) \end{pmatrix}. \tag{8.10}$$

The application of (5.27) to $F = I_L \otimes R$ leads to $\alpha(z) - s(z) = \mathcal{O}_z(N^{-2})$. In order to verify that $(\epsilon_i(z))_{i=1,2}$ are $\mathcal{O}_z(N^{-2})$ as well, we have to control $u_{t,\alpha}$ and $v_{t,\alpha}$. As $t(z), \alpha(z), \|T(z)\|$ and $\|S(z)\|$ are $\mathcal{O}_z(1)$ terms, it is sufficient to evaluate the denominator of the right handside of (8.8). As the mass and the first moment of μ and $\overline{\mu}$ (the measure associated to $\alpha(z)$) both verify the conditions of Lemma 4.2, this Lemma implies that $(1 - z(ct(z))^2)^{-1} = \mathcal{O}_z(1)$ and $(1 - z(c\alpha(z))^2)^{-1} = \mathcal{O}_z(1)$. Therefore, we have checked that $(\epsilon_i(z))_{i=1,2}$ are $\mathcal{O}_z(N^{-2})$ terms.

In order to evaluate $s(z) - t(z)$, it is of course necessary to show that matrix $I - \mathbf{D}_{t,\alpha}(z)$ is invertible on $\mathbb{C}^+$, and to control the action of its inverse on the vector $(\epsilon_1(z), \epsilon_1(z))^T$. We define matrix $\mathbf{D}_\alpha$ by

$$\mathbf{D}_\alpha(z) = \begin{pmatrix} u_\alpha(z) & v_\alpha(z) \\ z^2 v_\alpha(z) & u_\alpha(z) \end{pmatrix}$$

and establish the following result.

Lemma 8.1. *For each $z \in \mathbb{C}^+$, it exist nice constants κ and β such that*

$$\det(I - \mathbf{D}(z)) \geq \frac{\kappa \, (\text{Im}z)^8}{(|\beta|^2 + |z|^2)^4}. \tag{8.11}$$

Moreover, it exist 2 nice polynomials P_1 and P_2 for which

$$1 - u_\alpha(z) > 0 \tag{8.12}$$

and

$$\det(I - \mathbf{D}_\alpha(z)) \geq \frac{\kappa \, (\text{Im}z)^8}{(|\beta|^2 + |z|^2)^4} \tag{8.13}$$

for each $z \in \mathcal{B}_N$, where $\mathcal{B}_N$ is defined as

$$\mathcal{B}_N = \left\{ z \in \mathbb{C}^+, \ \frac{1}{MN} P_1(|z|) P_2\left(\frac{1}{\text{Im}z}\right) \leq 1 \right\}. \tag{8.14}$$

Finally, for each $z \in \mathcal{B}_N$, it holds that

$$\det(I - \mathbf{D}_{t,\alpha}(z)) \geq \frac{\kappa\, (\mathrm{Im} z)^8}{(|\beta|^2 + |z|^2)^4}. \tag{8.15}$$

Proof. To evaluate $\det(I - \mathbf{D}(z))$, we use the calculations of the proof of Lemma 6.1. In particular, we have

$$(I - \mathbf{D}(z)) \begin{pmatrix} \mathrm{Im} t(z) \\ \mathrm{Im} z t(z) \end{pmatrix} = \mathrm{Im} z \begin{pmatrix} \frac{1}{M} \mathrm{Tr} R T(z) T^*(z) \\ 0 \end{pmatrix}. \tag{8.16}$$

This implies that

$$1 - u(z) = \frac{\mathrm{Im} z}{\mathrm{Im} t(z)} \cdot \frac{1}{M} \mathrm{Tr} R T(z) T^*(z) + \frac{\mathrm{Im} z t(z)}{\mathrm{Im} t(z)} v(z) \geq \frac{\mathrm{Im} z}{\mathrm{Im} t(z)} \cdot \frac{1}{M} \mathrm{Tr} R T(z) T^*(z).$$

By applying Cramer's rule to (8.16), we obtain that

$$\det(I - \mathbf{D}(z)) = \frac{\mathrm{Im} z}{\mathrm{Im} t(z)} \cdot \frac{1}{M} \mathrm{Tr} R T(z) T^*(z) (1 - u(z)) \geq \left(\frac{\mathrm{Im} z}{\mathrm{Im} t(z)} \cdot \frac{1}{M} \mathrm{Tr} R T(z) T^*(z) \right)^2. \tag{8.17}$$

It is clear that $\mathrm{Im} t(z) \leq |t(z)| \leq \frac{1}{M} \mathrm{Tr} R \, (\mathrm{Im} z)^{-1} \leq b \, (\mathrm{Im} z)^{-1}$. Therefore, it holds that $\frac{\mathrm{Im} z}{\mathrm{Im} t(z)} \geq \frac{1}{b} \, (\mathrm{Im} t(z))^2$. We now evaluate $\frac{1}{M} \mathrm{Tr} R T(z) T^*(z)$. For this, we remark that

$$\frac{1}{M} \mathrm{Tr} R T(z) T^*(z) = \frac{1}{M} \mathrm{Tr} R T(z) T^*(z) R R^{-1} \geq \frac{1}{b} \frac{1}{M} \mathrm{Tr}(R T(z) T^*(z) R). \tag{8.18}$$

Jensen's inequality implies that $\frac{1}{M} \mathrm{Tr}(R T(z) T^*(z) R) \geq \left| \frac{1}{M} \mathrm{Tr} R T(z) \right|^2 = |t(z)|^2 \geq (\mathrm{Im}\, t(z))^2$. Therefore, the application of Lemma 4.2 to $\beta(z) = t(z)$ implies that

$$\left(\frac{\mathrm{Im} z}{\mathrm{Im} t(z)} \cdot \frac{1}{M} \mathrm{Tr} R T(z) T^*(z) \right)^2 \geq \frac{\kappa \, (\mathrm{Im} z)^8}{(|\beta|^2 + |z|^2)^4}$$

for some nice constants κ and β. (8.11) thus follows from (8.17).

We now establish (8.12) and (8.13), and denote by $\epsilon(z)$ the function $\epsilon(z) = \alpha(z) - s(z)$. Using the equation $s(z) = \frac{1}{M} \mathrm{Tr} R S(z)$, and calculating $\mathrm{Im}\, s(z)$ and $\mathrm{Im}\, z s(z)$, we obtain immediately that

$$(\mathbf{I} - \mathbf{D}_\alpha(z)) \begin{pmatrix} \mathrm{Im} \alpha(z) \\ \mathrm{Im} z \alpha(z) \end{pmatrix} = \mathrm{Im} z \begin{pmatrix} \frac{1}{M} \mathrm{Tr} R S(z) S^*(z) \\ 0 \end{pmatrix} + \begin{pmatrix} \mathrm{Im} \epsilon(z) \\ \mathrm{Im} z \epsilon(z) \end{pmatrix}. \tag{8.19}$$

The first component of (8.19) leads to

$$1 - u_\alpha = \frac{\mathrm{Im} z}{\mathrm{Im} \alpha} \cdot \frac{1}{M} \mathrm{Tr} R S S^* + \frac{\mathrm{Im} \epsilon}{\mathrm{Im} \alpha} + \frac{\mathrm{Im} z \alpha}{\mathrm{Im} \alpha} v_\alpha \geq \frac{\mathrm{Im} z}{\mathrm{Im} \alpha} \cdot \frac{1}{M} \mathrm{Tr} R S S^* + \frac{\mathrm{Im} \epsilon}{\mathrm{Im} \alpha}. \tag{8.20}$$

Using the same arguments as above, we obtain that $\frac{1}{M}\mathrm{Tr}RSS^* \geq \frac{1}{b}|s(z)|^2 \geq \frac{1}{b}(\mathrm{Im}s(z))^2$. As (8.6) holds, we can apply Lemma 4.2 to $\beta(z) = s(z)$ and obtain as above that

$$\frac{\mathrm{Im}z}{\mathrm{Im}s(z)} \cdot \frac{1}{M}\mathrm{Tr}RS(z)S^*(z) \geq \frac{\kappa\,(\mathrm{Im}z)^4}{(|\beta|^2+|z|^2)^2}$$

for some nice constants β and κ. We remark that $\frac{\mathrm{Im}\epsilon}{\mathrm{Im}\alpha} \geq -\frac{|\epsilon|}{\mathrm{Im}\alpha}$. Therefore, by Lemma 4.2 applied to $\beta(z) = \alpha(z)$, it holds that $\frac{\mathrm{Im}\epsilon}{\mathrm{Im}\alpha} \geq -\kappa_1|\epsilon|\frac{\beta_1^2+|z|^2}{\mathrm{Im}z}$ for some nice constants κ_1 and β_1. As $|\epsilon(z)| \leq \frac{1}{N^2}Q_1(|z|)Q_2(\frac{1}{\mathrm{Im}z})$ for some nice polynomials Q_1 and Q_2,we obtain that

$$1-u_\alpha \geq \frac{\mathrm{Im}z}{\mathrm{Im}\alpha}\cdot\frac{1}{M}\mathrm{Tr}RSS^* + \frac{\mathrm{Im}\epsilon}{\mathrm{Im}\alpha} \geq \frac{\mathrm{Im}z}{\mathrm{Im}\alpha}\cdot\frac{1}{M}\mathrm{Tr}RSS^* - \frac{|\epsilon|}{\mathrm{Im}\alpha} \geq \frac{1}{2}\frac{\kappa\,(\mathrm{Im}z)^4}{(|\beta|^2+|z|^2)^2} \tag{8.21}$$

if z belongs to the set $\mathcal{B}_{1,N}$ defined by

$$\frac{\kappa\,(\mathrm{Im}z)^4}{(|\beta|^2+|z|^2)^2} - \frac{1}{N^2}Q_1(|z|)Q_2(\frac{1}{\mathrm{Im}z})\,\kappa_1\,\frac{\beta_1^2+|z|^2}{\mathrm{Im}z} \geq \frac{1}{2}\frac{\kappa\,(\mathrm{Im}z)^4}{(|\beta|^2+|z|^2)^2}.$$

The set $\mathcal{B}_{1,N}$ is clearly defined in the same way than $\mathcal{B}_N$, but from 2 other nice polynomials $P_{1,1}$ and $P_{2,1}$.

Using the Cramer rule, we obtain that $\det(\mathbf{I}-\mathbf{D}_\alpha)$ can be written as

$$\det(\mathbf{I}-\mathbf{D}_\alpha) = \left(\frac{\mathrm{Im}z}{\mathrm{Im}\alpha}\cdot\frac{1}{M}\mathrm{Tr}RSS^* + \frac{\mathrm{Im}\epsilon}{\mathrm{Im}\alpha}\right)(1-u_\alpha) + \frac{\mathrm{Im}z\epsilon}{\mathrm{Im}\alpha}v_\alpha.$$

Plugging (8.21) in the last equation, we get that the inequality

$$\det(\mathbf{I}-\mathbf{D}_\alpha) \geq \left(\frac{1}{2}\frac{\kappa\,(\mathrm{Im}z)^4}{(|\beta|^2+|z|^2)^2}\right)^2 - \frac{|z|\,|\epsilon|}{\mathrm{Im}\alpha}v_\alpha$$

holds for each $z \in \mathcal{B}_{1,N}$. As $v_\alpha = \mathcal{O}_z(1)$, we obtain that

$$\left(\frac{\kappa\,(\mathrm{Im}z)^4}{(|\beta|^2+|z|^2)^2}\right)^2 - \frac{|z|\,|\epsilon|}{\mathrm{Im}\alpha}v_\alpha \geq \left(\frac{1}{4}\frac{\kappa\,(\mathrm{Im}z)^4}{(|\beta|^2+|z|^2)^2}\right)^2$$

for each $z \in \mathcal{B}_{2,N}$, where $\mathcal{B}_{2,N}$ is defined as $\mathcal{B}_N$ from 2 nice polynomials $P_{1,2}$ and $P_{2,2}$. We put $P_1(|z|) = P_{1,1}(|z|) + P_{1,2}(|z|)$ and $P_2(1/\mathrm{Im}z) = P_{2,1}(1/\mathrm{Im}z) + P_{2,2}(1/\mathrm{Im}z)$, and consider the set $\mathcal{B}_N$ defined by (8.14). It is clear that $\mathcal{B}_N \subset \mathcal{B}_{1,N} \cap \mathcal{B}_{2,N}$, and that (8.12) and (8.13) hold if $z \in \mathcal{B}_N$.

It remains to establish (8.15). For this, we remark that the inequalities

$$
\begin{aligned}
|\det(\mathbf{I}-\mathbf{D}_{t,\alpha}(z))| &\geq |1-u_{t,\alpha}(z)|^2 - |z|^2|v_{t,\alpha}(z)|^2 \geq (1-|u_{t,\alpha}(z)|)^2 \\
-|z|v_\alpha(z)\cdot|z|v(z) &\geq (1-\sqrt{u(z)u_\alpha(z)})^2 - |z|v_\alpha(z)\cdot|z|v(z) \geq (1-u(z))(1-u_\alpha(z)) \\
-|z|v_\alpha(z)\cdot|z|v(z) &\geq \sqrt{((1-u(z))^2-|z|^2v(z))((1-u_\alpha(z))^2-|z|^2v_\alpha(z))} \\
&= \sqrt{\det(\mathbf{I}-\mathbf{D}(z))\det(\mathbf{I}-\mathbf{D}_\alpha(z))}
\end{aligned}
$$

hold for each $z \in \mathcal{B}_N$. Therefore, (8.15) follows from (8.11) and (8.13). This completes the proof of Lemma 8.1. ■

Solving (8.10), we obtain immediately that it exists 2 nice polynomials Q_1 and Q_2 such that,

$$|s_N(z) - t_N(z)| \leq \frac{1}{MN} Q_1(|z|) Q_2\left(\frac{1}{\mathrm{Im} z}\right)$$

holds for each $z \in \mathcal{B}_N$. If $z \in \mathcal{B}_N^c$, we use the argument in [15]. More precisely, if $z \in \mathcal{B}_N^c$, the inequality $1 < \frac{1}{MN} P_1(|z|) P_2(1/\mathrm{Im} z)$ holds. As $|s_N(z) - t_N(z)| \leq 2\,\frac{1}{M}\mathrm{Tr} R_N\,\frac{1}{\mathrm{Im} z}$ on $\mathbb{C}^+$, we deduce that

$$|s_N(z) - t_N(z)| \leq 2b \frac{1}{MN} P_1(|z|) \frac{P_2(1/\mathrm{Im} z)}{\mathrm{Im} z}$$

for each $z \in \mathcal{B}_N^c$. This, in turn, leads to the conclusion that $s_N(z) - t_N(z) = \mathcal{O}_z(\frac{1}{N^2})$ for each $z \in \mathbb{C}^+$. This establishes (8.7) and $\frac{1}{M}\mathrm{Tr}(T_N(z) - S_N(z)) = \mathcal{O}_z(\frac{1}{N^2})$ as expected. This completes the proof of Proposition 8.1. ■

We now follow [8] and [15] and use the following Lemma.

Lemma 8.2. *Let ϕ be a compactly supported real valued smooth function defined on $\mathbb{R}^+$, i. e. $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^+, \mathbb{R}^+)$. Then,*

$$\mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)\right\} - \int_{\mathcal{S}_N} \phi(\lambda)\, d\mu_N(\lambda) = \mathcal{O}\left(\frac{1}{N^2}\right).$$

Proof. Due to Proposition 4.1 we can write

$$\mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)\right\} = \frac{1}{\pi}\lim_{y\downarrow 0}\mathrm{Im}\left\{\int_{\mathbb{R}^+}\phi(x)\mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}Q(x+iy)\right\}dx\right\}$$

as well as

$$\int_{\mathcal{S}_N} \phi(\lambda)\, d\mu_N(\lambda) = \frac{1}{\pi}\lim_{y\downarrow 0}\mathrm{Im}\left\{\int_{\mathbb{R}^+}\phi(x)\mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}T(x+iy)\right\}dx\right\}$$

Using Proposition 8.1, we obtain

$$
\begin{aligned}
&\mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)\right\} - \int_{\mathcal{S}_N} \phi(\lambda)\, d\mu_N(\lambda) \\
&\qquad = \frac{1}{N^2}\frac{1}{\pi}\lim_{y\downarrow 0}\mathrm{Im}\left\{\int_{\mathbb{R}^+}\phi(x) r_N(x+iy)\, dx\right\}. \qquad (8.22)
\end{aligned}
$$

Since the function $r_N(z) = \mathcal{O}_z(1)$, we can use the result which was proved in [7, Section 3.3] and obtain

$$\limsup_{y\downarrow 0}\left|\int_{\mathbb{R}+}\phi(x)r_N(x+iy)dx\right| \le \kappa$$

for some nice constant κ. This and (8.22) complete the proof. ■

In order to establish Theorem 8.1, we introduce a function $\phi \in \mathcal{C}_c^\infty$ such that $0 \le \phi(\lambda) \le 1$ and

$$\phi(\lambda) = \begin{cases} 1, \text{ for } \lambda \in [\kappa_1, \kappa_2], \\ 0, \text{ for } \lambda \in \mathbb{R} - (\kappa_1 - \epsilon, \kappa_2 + \epsilon). \end{cases}$$

Since for N large enough $(\kappa_1 - \epsilon, \kappa_2 + \epsilon) \cap \mathcal{S}_N = \varnothing$ then $\int_{\mathcal{S}_N} \phi(\lambda) d\mu_N(\lambda) = 0$ and according to Lemma 8.2

$$\mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)\right\} = \mathcal{O}\left(\frac{1}{N^2}\right).$$

Now we show that

$$\mathbf{Var}\left\{\frac{1}{ML}\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)\right\} = \mathcal{O}\left(\frac{1}{N^4}\right).$$

For this we use again the Poincare-Nash inequality

$$\mathbf{Var}\{\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)\} \le \sum \mathbb{E}\Big\{ \left(\frac{\partial \mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)}{\partial \overline{W}_{i_1,j_1}^{m_1}}\right)^* \mathbb{E}\{W_{i_1,j_1}^{m_1}\overline{W}_{i_2,j_2}^{m_2}\}$$

$$\times \frac{\partial \mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)}{\partial \overline{W}_{i_2,j_2}^{m_2}}\Big\} + \sum \mathbb{E}\left\{\frac{\partial \mathrm{Tr}\phi(WW^*)}{\partial W_{i_1,j_1}^{m_1}}\mathbb{E}\{W_{i_1,j_1}^{m_1}\overline{W}_{i_2,j_2}^{m_2}\}\left(\frac{\partial \mathrm{Tr}\phi(WW^*)}{\partial W_{i_2,j_2}^{m_2}}\right)^*\right\}.$$

We only evaluate the first term of the r.h.s. of the inequality, denoted by ψ, because the second is similar. For this we write first

$$\frac{\partial \mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)}{\partial \overline{W}_{i_1,j_1}^{m_1}} = \mathrm{Tr}\left(\phi'(W_f W_p^* W_p W_f^*)\frac{\partial W_f W_p^* W_p W_f^*}{\partial \overline{W}_{i_1,j_1}^{m_1}}\right)$$

$$= \begin{cases} 1 \le i_1 \le L,\ (W_p W_f^* \phi'(W_f W_p^* W_p W_f^*) W_f)_{i_1 j_1}^{m_1}, \\ L+1 \le i_1 \le 2L,\ (\phi'(W_f W_p^* W_p W_f^*) W_f^* W_f W_p)_{i_1 - L, j_1}^{m_1}. \end{cases}$$

Plugging this into (3.2) we obtain

$$\psi = \sum_{i_1,i_2=1}^{L}\sum_{j_1,j_2,m_1,m_2}\Big(\frac{1}{N}\mathbb{E}\Big\{\left(W_p W_f^* \phi'(W_f W_p^* W_p W_f^*) W_f\right)_{i_1 j_1}^{*m_1} R_{m_1 m_2}\delta_{i_1+j_1,i_2+j_2}$$

$$\times \left(W_p W_f^* \phi'(W_f W_p^* W_p W_f^*) W_f\right)_{i_2,j_2}^{m_2}\Big\} + \frac{1}{N}\mathbb{E}\Big\{\left(\phi'(W_f W_p^* W_p W_f^*) W_f W_p^* W_p\right)_{i_1 j_1}^{*m_1}$$

$$\times R_{m_1 m_2}\delta_{i_1+j_1,i_2+j_2}\left(\phi'(W_f W_p^* W_p W_f^*) W_f W_p^* W_p\right)_{i_2,j_2}^{m_2}\Big\}\Big).$$

Following the proof of Lemma 3.1, we obtain

$$\mathbf{Var}\{\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)\} \le \frac{C}{N}\mathbb{E}\{\mathrm{Tr}W_f^*\phi'(W_f W_p^* W_p W_f^*)W_f W_p^* W_p W_f^* \\ \times \phi'(W_f W_p^* W_p W_f^*)W_f\} + \frac{C}{N}\mathbb{E}\{\mathrm{Tr}W_f W_p^* W_p W_p^* W_p W_f^* \left(\phi'(W_f W_p^* W_p W_f^*)\right)^2\}. \tag{8.23}$$

To evaluate the first term ψ_1 of the r.h.s of (8.23) we denote $\eta(\lambda) = (\phi'(\lambda))^2\lambda$ and write

$$\frac{1}{N}\mathbb{E}\left\{\mathrm{Tr}W_f^*\phi'(W_f W_p^* W_p W_f^*)W_f W_p^* W_p W_f^*\phi'(W_f W_p^* W_p W_f^*)W_f\right\} \\ \le \frac{1}{N}\mathbb{E}\left\{\|W_f\|^2\mathrm{Tr}(\eta(W_f W_p^* W_p W_f^*))\right\}.$$

We recall that (3.1) implies that $\|W_f\|^2 \le b\|W_{iid}\|^2$. Therefore, it holds that

$$\psi_1 \le \frac{\kappa}{N}\mathbb{E}\{\|W_{iid}\|^2\mathbf{1}_{\|W_{iid}\|\le(1+\sqrt{c_*})^2+\delta}\mathrm{Tr}(\eta(W_f W_p^* W_p W_f^*))\} \\ + \frac{\kappa}{N}\mathbb{E}\{\|W_{iid}\|^2\mathbf{1}_{\|W_{iid}\|>(1+\sqrt{c_*})^2+\delta}\mathrm{Tr}(\eta(W_f W_p^* W_p W_f^*))\} \\ \le \frac{\kappa}{N}\mathbb{E}\{\mathrm{Tr}(\eta(W_f W_p^* W_p W_f^*))\} + \kappa\mathbb{E}^{1/2}\{\|W_{iid}\|^4\mathbf{1}_{\|W_{iid}\|>(1+\sqrt{c_*})^2+\delta}\} \\ \times \mathbb{E}^{1/2}\left\{\left(\frac{1}{N}\mathrm{Tr}(\eta(W_f W_p^* W_p W_f^*))\right)^2\right\}.$$

Lemma 8.2 implies that $\frac{1}{N}\mathbb{E}\{\mathrm{Tr}(\eta(W_f W_p^* W_p W_f^*))\} = \mathcal{O}(N^{-2})$. Throughout the proof of Lemma 3.1, we get that $\mathbb{E}\|W_{iid}\|^4\mathbf{1}_{\|W_{iid}\|>(1+\sqrt{c_*})^2+\delta} = \mathcal{O}(N^{-k})$ for all k. Since function $\phi' \in \mathcal{C}_c^\infty$, there exists a nice constant κ such that $|\phi'(\lambda)| < \kappa$ for all λ and $\phi'(\lambda) = 0$ for all $\lambda > b + 2\epsilon$. We deduce from this that it exists a nice constant κ such that $\|\eta(W_{f,N}W_{p,N}^*W_{p,N}W_{f,N}^*)\| < \kappa$ for each N. From what about we conclude that $\psi_1 = \mathcal{O}(N^{-2})$.

As for the second term (ψ_2) of the r.h.s of (8.23), we write

$$\psi_2 = \frac{\kappa}{N}\mathbb{E}\left\{\mathrm{Tr}W_p^* W_p W_p^* W_p W_f^* \left(\phi'(W_f W_p^* W_p W_f^*)\right)^2 W_f\right\} \\ \le \kappa\mathbb{E}\left\{\|W_p\|^2\frac{1}{N}\mathrm{Tr}\left(\phi'(W_f W_p^* W_p W_f^*)\right)^2 W_f W_p^* W_p W_f^*\right\}.$$

It is easy to see that ψ_2 can be evaluated as ψ_1, leading to the conclusion that $\psi_2 = \mathcal{O}(N^{-2})$. Therefore, we have checked that

$$\mathbf{Var}\{\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)\} = \mathcal{O}\left(\frac{1}{N^2}\right).$$

Now we can complete the proof of Theorem 8.1 as in [8]. For this we apply the classical Markov inequality and combine what above

$$\mathbf{P}\left\{\frac{1}{ML}\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*) > \frac{1}{N^{4/3}}\right\} \leq N^{8/3}\mathbb{E}\left\{\left(\frac{1}{ML}\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)\right)^2\right\}$$
$$= N^{8/3}\left(\mathbf{Var}\left\{\frac{1}{ML}\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)\right\} + \left(\mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*)\right\}\right)^2\right)$$
$$= \mathcal{O}\left(\frac{1}{N^{4/3}}\right).$$

Applying Borel-Cantelli lemma, we obtain that almost surely, the inequality

$$\frac{1}{ML}\mathrm{Tr}\phi(W_f W_p^* W_p W_f^* *) \leq \frac{1}{N^{4/3}}$$

holds for each N large enough. By the very definition of function ϕ, the number of eigenvalues of matrix $W_f W_p^* W_p W_f^*$ lying in the interval $[\kappa_1, \kappa_2]$ is upper bounded by $\mathrm{Tr}\phi(W_f W_p^* W_p W_f^*) \leq \frac{1}{N^{1/3}}$. Since this number of eigenvalues is an integer, we conclude that with probability one there is no eigenvalues in the interval $[\kappa_1, \kappa_2]$ for each N large enough. ■

We finally illustrate the above results by the following numerical experiment. M, N, L are given by $M = 500$, $N = 1500$ and $L = 2$ so that $c_N = 2/3$. The eigenvalues of matrix R_N are defined by $\lambda_{k,N} = 1/2 + \frac{\pi}{4}\cos\left(\frac{\pi(k-1)}{2M}\right)$ for $k = 1, \ldots, M$. Matrix R_N verifies $\frac{1}{M}\mathrm{Tr}(R_N) \simeq 1$. Fig. 3 represents the histogram of the eigenvalues of a realization of $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ as well as the graph of the density $g_N(x)$. We notice that the histogram and the graph of g_N are in accordance, and that, as expected, no eigenvalue of $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ lies outside the support of g_N.

9. Recovering the behaviour of the empirical eigenvalue distribution $\hat{\nu}_N$ using free probability tools

The purpose of this paragraph is to show that it is possible to use free probability tools in order to characterize the limiting behaviour of the empirical eigenvalue distribution $\hat{\nu}_N$ of matrix $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$. As the present paper is not focused on these kind of approach, we present briefly the following results and leave the details to the reader.

The free probability approach is based on the following observations:

- Up to the zero eigenvalue, the eigenvalues of $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ coincide with the eigenvalues of $W_{f,N}^* W_{f,N} W_{p,N}^* W_{p,N}$
- The matrices $W_{f,N}^* W_{f,N}$ and $W_{p,N}^* W_{p,N}$ are almost surely asymptotically free. Therefore, the eigenvalue distribution of $W_{f,N}^* W_{f,N} W_{p,N}^* W_{p,N}$ converges towards the free multiplicative convolution product of the limit distributions of $W_{f,N}^* W_{f,N}$ and $W_{p,N}^* W_{p,N}$. These two distributions appear

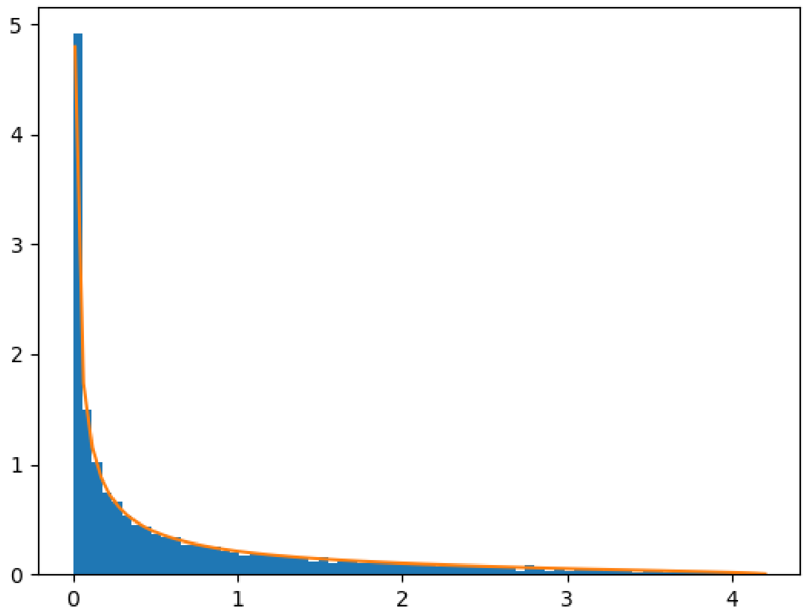

Fig. 3. Histogram of the eigenvalues and graph of $g_N(x)$ for $M = 500, N = 1500, L = 2$.

to coincide both with the limit distribution of the well known random matrix model $\frac{1}{N} X_N^* (I_L \times R_N) X_N$ where X_N is a $ML \times N$ complex Gaussian random matrix with unit variance i.i.d. entries.

In the following, we follow the definitions of asymptotic freeness provided in [18] (see in particular section 4.3) which need the existence of certain limit distributions. This is in contrast with the approach developed in the previous sections more focused on the behaviour of deterministic equivalents. We however mention that more recent free probability works (see e.g. [29] and the references therein, [6]) allow to avoid the introduction of limit distributions, and would allow to recover the previous results on the deterministic equivalent ν_N of $\hat{\nu}_N$.

In order to be in accordance with [18], we thus formulate in this section the following assumption:

Assumption 9.1. The empirical eigenvalue distribution $\omega_N = \frac{1}{M} \sum_{k=1}^{M} \delta_{\lambda_{k,N}}$ of matrix R_N converges towards a limit distribution ω.

We remark that hypothesis 2.7 implies that ω is compactly supported. Moreover, it can be shown that measures $(\mu_N)_{N \geq 1}$ and $(\nu_N)_{N \geq 1}$ both converge weakly towards limits denoted μ and ν in this section. We also notice that Lemma 7.7 implies that μ and ν are compactly supported. It is also easily checked that the Stieltjes transform $t(z)$ of μ verifies the equation

$$t(z) = -\frac{1}{z} \int_{\mathbb{R}^+} \frac{\tau \, d\omega(\tau)}{1 + \dfrac{c_* \tau t(z)}{1 - z c_*^2 t^2(z)}}, \tag{9.1}$$

while the Stieltjes transform t_ν of ν is given by

$$t_\nu(z) = -\frac{1}{z} - \frac{c_* t(z)^2}{1 - z(c_* t(z))^2}. \tag{9.2}$$

We recall that c_* represents the limit of $c_N = \frac{ML}{N}$. In the following, we establish that (9.1) and (9.2) can be obtained using free probability technics.

Before going further, we first recall the main useful definitions introduced in [18].

Definition 9.1. Consider a finite family of sequences of $N \times N$ possibly random matrices $((X_{i,N})_{N\geq 1})_{i=1,\dots,r}$. Then $(X_{i,N})_{i=1,\dots,r}$ is said to have an almost sure joint limit if for each non commutative polynomial $P(x_1, \dots, x_r)$ in r indeterminates, then $\frac{1}{N}\mathrm{Tr}P(X_{1,N}, \dots, X_{r,N})$ converges almost surely towards $\gamma(P)$ where γ is a deterministic distribution defined on the set of all non commutative polynomials in r indeterminates (i.e. γ is a linear form such that $\gamma(1) = 1$).

We remark that if $r = 1$ and $(X_{1,N})_{N\geq 1}$ are Hermitian matrices, the above condition is equivalent to the existence of a limit empirical eigenvalue distribution.

Definition 9.2. Consider p families $(X^{(1)}_{i,N})_{i=1,\dots,r_1}, \dots, (X^{(p)}_{i,N})_{i=1,\dots,r_p}$ of $N \times N$ possibly random matrices. Then, $X^{(1)}, \dots, X^{(p)}$ are said to be almost surely asymptotically free if the 2 following conditions hold:

- For each $q = 1, \dots, p$, $(X^{(q)}_{i,N})_{i=1,\dots,r_q}$ has an almost sure joint limit
- $\forall m$, $i_1, \cdots, i_m \in \{1, 2, \dots, p\}$ with $i_1 \neq i_2 \neq \cdots \neq i_m$, and for each non commutative polynomials $(P_j)_{j=1,\dots,m}$ in $(r_{i_j})_{j=1,\dots,m}$ indeterminates such that $\frac{1}{N}\mathrm{Tr}(P_j(X^{i_j}_{1,N}, \dots, X^{i_j}_{r_{i_j},N})) \to 0$ a.s. it holds that

$$\frac{1}{N}\mathrm{Tr}(P_1(X^{i_1}_{1,N}, \dots, X^{i_1}_{r_{i_1},N}) \cdots P_m(X^{i_m}_{1,N}, \dots, X^{i_m}_{r_{i_m},N})) \to 0 \quad a.s.$$

We remark that when each family $X^{(q)}$ is reduced to a single sequence $(X^{(q)}_N)_{N\geq 1}$ of $N \times N$ Hermitian, or similar to hermitian matrices [b], the almost sure freeness of $X^{(1)}, \dots, X^{(p)}$ holds if

Definition 9.3.

[b] in the sense that $X^{(q)}_N = U^{(q)}_N H^{(q)}_N (U^{(q)}_N)^{-1}$ for some $N \times N$ Hermitian matrix $H^{(q)}_N$

- For each $q = 1, \dots, p$, $(X_N^{(q)})_{N \geq 1}$ has a limit eigenvalue distribution
- $\forall m,\ i_1, \cdots, i_m \in \{1, 2, \dots, p\}$ with $i_1 \neq i_2 \neq \cdots \neq i_m$, and for each polynomials $(P_j)_{j=1 \cdots m}$ in one indeterminate such that $\frac{1}{N}\mathrm{Tr}(P_j(X_N^{i_j})) \to 0$ a.s. it holds that

$$\frac{1}{N}\mathrm{Tr}(P_1(X_N^{(i_1)})P_2(X_N^{(i_2)}) \cdots P_m(X_N^{(i_m)})) \to 0 \quad a.s. \tag{9.3}$$

We also recall the definition of the S transform of a probability measure, and recall that the S transform of the free multiplicative convolution product of two probability measures is the product of their S transforms.

Definition 9.4. Given a compactly supported probability measure μ carried by $\mathbb{R}^+$, we define $\psi_\mu(z)$ as the formal power series defined by

$$\psi_\mu(z) = \sum_{k \geq 1} z^k \int t^k d\mu(t) = \int \frac{zt}{1 - zt} d\mu(t) \tag{9.4}$$

Let χ_μ be the unique function analytic in a neighbourhood of zero, satisfying

$$\chi_\mu(\psi_\mu(z)) = z \tag{9.5}$$

for $|z|$ small enough. Then, we define the S transform of μ as the function $S_\mu(z)$ defined in a neighbourhood of zero by

$$S_\mu(z) = \chi_\mu(z)\frac{1+z}{z}. \tag{9.6}$$

Moreover, if μ_1 and μ_2 are two compactly supported probability measures carried by $\mathbb{R}^+$, the S-transform $S_{\mu_1 \boxtimes \mu_2}$ of $\mu_1 \boxtimes \mu_2$ satisfies

$$S_{\mu_1 \boxtimes \mu_2} = S_{\mu_1} S_{\mu_2}. \tag{9.7}$$

We are now in position to state the main result of this section.

Proposition 9.1. *Matrices $W_{f,N}^* W_{f,N}$ and $W_{p,N}^* W_{p,N}$ are almost surely asymptotically free.*

Proof. We first notice that it possible to replace matrices W_f and W_p by finite rank perturbations because the very definition of almost sure asymptotic freeness is not affected by finite rank perturbations. We thus exchange W_p and W_f by $\tilde{W}_p = \frac{1}{\sqrt{N}}\tilde{Y}_p$ and $\tilde{W}_f = \frac{1}{\sqrt{N}}\tilde{Y}_f$ where $\tilde{Y}_p$ and $\tilde{Y}_f$ are defined by

$$\tilde{Y}_p = \begin{pmatrix} y_1 & \cdots & \cdots & \cdots & y_N \\ y_2 & \cdots & \cdots & y_N & y_1 \\ y_3 & \cdots & y_N & y_1 & y_2 \\ \vdots & \cdots & \iddots & \iddots & \vdots \\ y_L & \cdots y_N & y_1 & y_2 \cdots & y_{L-1} \end{pmatrix},$$

$$\tilde{Y}_f = \begin{pmatrix} y_{L+1} & \cdots\cdots\cdots\cdots\cdots & y_N & y_1 & \cdots & y_L \\ y_{L+2} & \cdots\cdots\cdots\cdots & y_N & y_1 & \cdots & y_L & y_{L+1} \\ y_{L+3} & \cdots\cdots\cdots & y_N & y_1 & \cdots & y_L & y_{L+1} & y_{L+2} \\ \vdots & \cdots\cdots\cdots \iddots & \iddots & \cdots & \iddots & \iddots & \iddots & \vdots \\ y_{2L} & \cdots y_N & y_1 & \cdots & y_L & y_{L+1} & y_{L+2} & \cdots & y_{2L-1} \end{pmatrix}.$$

In other words, vectors $y_{N+1}, \dots, y_{N+L-1}, \dots, y_{N+2L-1}$ are replaced by vectors $y_1, \dots, y_{L-1}, \dots, y_{2L-1}$. In order to simplify the notations, we still denote the above finite rank modifications by Y_p, Y_f, W_p, W_f. We define the $N \times N$ matrix Π and the $M \times N$ matrix Y by

$$\Pi = \begin{pmatrix} 0 & \dots & 0 & 1 \\ 1 & \ddots & & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 1 & 0 \end{pmatrix}, \text{ and } Y = (y_1, y_2, \dots, y_N)$$

and rewrite Y_p (and Y_f respectively) as

$$Y_p = \begin{pmatrix} Y \\ Y\Pi \\ \vdots \\ Y\Pi^{L-1} \end{pmatrix}, \quad Y_f = \begin{pmatrix} Y\Pi^L \\ Y\Pi^{L+1} \\ \vdots \\ Y\Pi^{2L-1} \end{pmatrix}.$$

This allows us to obtain the useful expression for $W_p^* W_p$ and $W_f^* W_f$

$$W_p^* W_p = \textstyle\sum_{k=0}^{L-1} \Pi^{*k} \left(\frac{Y^* Y}{N} \right) \Pi^k, \tag{9.8}$$

$$W_f^* W_f = \textstyle\sum_{k=L}^{2L-1} \Pi^{*k} \left(\frac{Y^* Y}{N} \right) \Pi^k. \tag{9.9}$$

Since $N^{-1} Y^* Y$ can be written as $N^{-1} Y_{iid}^* R_N Y_{iid}$, where Y_{iid} has i.i.d. Gaussian entries, the Hermitian matrix $N^{-1} Y^* Y$ is unitarily invariant. Moreover, Assumption 9.1 implies that $N^{-1} Y^* Y$ has a limit distribution while it is easily checked that the family $\{I, \Pi^*, \Pi, \dots, \Pi^{*2L-1}, \Pi^{2L-1}\}$ has the same property. This and Theorem 4.3.5 in [18] leads to the conclusion that $Y^* Y / N$ and $\{I, \Pi^*, \Pi, \dots, \Pi^{*2L-1}, \Pi^{2L-1}\}$ are almost surely asymptotically free. Proposition 9.1 thus appears to be an immediate consequence of the following Lemma adapted from Lemma 6 in [14]. In order to make the connections between Lemma 9.1 and Lemma 6 in [14], we use nearly the same notations than in [14] in the following statement.

Lemma 9.1. *We consider a sequence of $N \times N$ Hermitian random matrices $(X^N)_{N \geq 1}$ and $N \times N$ deterministic matrices $U_1^N, W_1^N, \dots, U_m^N, W_m^N$ such that X_N and $\{U_1^N, W_1^N, \dots, U_m^N, W_m^N\}$ are almost surely asymptotically free. Then, if*

$U_1^N, W_1^N, \dots, U_m^N, W_m^N$ *satisfy*

$$U_i^N W_i^N = W_i^N U_i^N = I_N \tag{9.10}$$

for each $i = 1, \dots, m$ *as well as* $\frac{1}{N}\mathrm{Tr}(U_i^N W_j^N) = \delta_{i-j}$ *for all* $i, j = 1 \dots m$, *then the random matrices* $U_1^N X^N W_1^N, \dots, U_m^N X^N W_m^N$ *are almost surely asymptotically free.*

Proof. We prove Lemma 9.1 by following step by step the proof from [14]. For simplicity we omit index N below. Due to (9.10) we have $W_i = U_i^{-1}$ so that matrices $(U_i X W_i)_{i=1,\dots,m}$ are similar to the Hermitian matrix X. We have thus to verify the 2 items of Definition 9.3. The first item is obvious. To check condition (9.3), we consider any k, indexes $i_1, \cdots, i_k$ with $i_1 \neq \cdots \neq i_k$ and polynomials P_j such that $\frac{1}{n}\mathrm{Tr}(P_j(U_{i_j} X W_{i_j})) \to 0$ a.s. Using again (9.10) it is clear that $P_j(U_{i_j} X W_{i_j}) = U_{i_j} P_j(X) W_{i_j}$ and, as a consequence, $\frac{1}{n}\mathrm{Tr}(P_j(X)) \to 0$ a.s. We define η_N as

$$\eta_N = \frac{1}{N}\mathrm{Tr}(P_1(U_{i_1} X W_{i_1}) P_2(U_{i_2} X W_{i_2}) \cdots (U_{i_k} X W_{i_k})) =$$

$$\frac{1}{N}\mathrm{Tr}(U_{i_1} P_1(X) W_{i_1} U_{i_2} P_2(X) W_{i_2} \cdots U_{i_k} P_k(X) W_{i_k}) = \frac{1}{N}\mathrm{Tr}\left(\prod_{j=1}^{k} W_{i_{j-1}} U_{i_j} P_j(X)\right),$$

where $i_0 = i_k$. If $i_1 \neq i_k$ then by assumption $\frac{1}{n}\mathrm{Tr}(W_{i_{j-1}} U_{i_j}) = 0$ for $j = 1, \dots, m$. As we also have $\frac{1}{n}\mathrm{Tr}(P_j(X)) \to 0$ a.s, the almost sure asymptotic freeness of X and $\{U_1, W_1, \cdots, U_m, W_m\}$ leads to the conclusion that $\eta_N \to 0$ a.s. In the case when $i_1 = i_k$ we have $W_{i_k} U_{i_1} = I_N$ and the same conclusion holds. □

By taking $X = \frac{YY^*}{N}$, $U_i = \Pi^{*i-1}$ and $W_i = \Pi^{i-1}$, Lemma 9.1 gives us immediately that $\frac{Y^*Y}{N}, \Pi^*(\frac{Y^*Y}{N})\Pi, \dots, \Pi^{*2L-1}(\frac{Y^*Y}{N})\Pi^{2L-1}$ are almost surely asymptotically free. Using the expression (9.8, 9.9) of $W_p^* W_p$ and $W_f^* W_f$, we obtain that $W_p^* W_p$ and $W_f^* W_f$ are almost surely asymptotically free. ■

We also deduce that the limit distributions of $W_p^* W_p$ and $W_f^* W_f$ both coincide with the additive free convolution product of L copies of the well known limit distribution of $\frac{Y^*Y}{N}$. It is easily seen that the Stieljes transform, denoted $t_{BS}(z)$ in the following, of this free addditive convolution product is solution of the familiar equation (see e.g. [1], p. 113)

$$t_{BS}(z) = -\frac{1}{z - c_* \int \frac{\tau d\omega(\tau)}{1 + \tau t_{BS}(z)}}. \tag{9.11}$$

In the following, we denote by μ_{BS} the corresponding probability measure. It is clear that (9.11) coincides with the equation verified by the Stieltjes transform of the limit eigenvalue distribution of the random matrix $\frac{1}{N} X_N^* (I_L \times R_N) X_N$ where X_N is a $ML \times N$ complex Gaussian random matrix with unit variance i.i.d. entries. We note that this result could also be easily obtained using the Gaussian technics developed in [26] in the case where R_N is reduced to a multiple of I_M.

According to Proposition 9.1, the limit eigenvalue distribution of $W_{f,N}^* W_{f,N} W_{p,N}^* W_{p,N}$ is $\mu_{BS} \boxtimes \mu_{BS}$. In the following, we denote by $\tilde{\nu}$ this measure and by $\tilde{f}(z)$ its Stieltjes transform. To find an equation satisfied by $\tilde{f}(z)$, we use (9.7). (9.6) and (9.7) give us immediately

$$\chi_{\tilde{\nu}}(z) = \frac{1+z}{z} \chi_{BS}^2(z).$$

By replacing here z with $\psi_{\tilde{\nu}}(z)$ and taking into account (9.5) we obtain

$$z = \frac{1+\psi_{\tilde{\nu}}(z)}{\psi_{\tilde{\nu}}(z)} \chi_{BS}^2(\psi_{\tilde{\nu}}(z)). \tag{9.12}$$

We notice that by definition (9.4), we have

$$\psi_{\tilde{\nu}}(z) = \int \frac{zt}{1-zt} d\tilde{\nu}(t) = \int \frac{d\tilde{\nu}(t)}{1-zt} - 1 = -\frac{1}{z} \tilde{f}\left(\frac{1}{z}\right) - 1. \tag{9.13}$$

Putting this into (9.12) and replacing z with $\frac{1}{z}$ give us

$$\frac{z^2 \tilde{f}(z)}{1+z\tilde{f}(z)} \chi_{BS}^2 \left(\psi_{\tilde{\nu}}\left(\frac{1}{z}\right)\right) = 1.$$

From this, it is straightforward to obtain the expression of $\tilde{f}(z)$. For more convenience, we introduce the function $g(z) = \chi_{BS}(\psi_{\tilde{\nu}}(z^{-1}))$ which is analytic in the neighbourhood of infinity. It holds that

$$\tilde{f}(z) = \left(z^2 g^2(z) - z\right)^{-1}. \tag{9.14}$$

It remains to determine $g(z)$. For this we use (9.13) for ψ_{BS}, t_{BS} and replace z with $\chi_{BS}(z)$. Then (9.5) gives

$$z = -1 - \frac{1}{\chi_{BS}(z)} t_{BS}\left(\frac{1}{\chi_{BS}(z)}\right) \Rightarrow t_{BS}(\chi_{BS}^{-1}(z)) = -(1+z)\chi_{BS}(z).$$

To obtain the equation for χ_{BS} it is sufficient to use the above expression of $t_{BS}(\chi_{BS}^{-1}(z))$, and to plug it in (9.11) with $z = \chi_{BS}^{-1}(z)$. Therefore, we obtain that

$$(1+z)\chi_{BS}(z) = \frac{1}{\frac{1}{\chi_{BS}(z)} - c_* \int \frac{\tau d\omega(\tau)}{1-\tau(1+z)\chi_{BS}(z)}}.$$

After simple algebra we get that

$$\frac{z}{(1+z)\chi_{BS}(z)} = c_* \int \frac{\tau d\omega(\tau)}{1-\tau(1+z)\chi_{BS}(z)}.$$

We finally replace z by $\psi_{\tilde{\nu}}(z^{-1})$ in the above equation. Using (9.12), it is easy to see that the l.h.s. is equal to $zg(z)$. To evaluate the r.h.s., we use again (9.12) and obtain that $\psi_{\tilde{\nu}}(z^{-1}) = zg^2(z)(1 - zg^2(z))^{-1}$, and that

$$g(z) = \frac{1}{z} \int_{\mathbb{R}^+} \frac{c_* \tau \, d\omega(\tau)}{1 - \dfrac{\tau g(z)}{1 - zg^2(z)}}. \tag{9.15}$$

We recall that $t(z)$ is solution of the equation

$$t(z) = -\frac{1}{z} \int \frac{\tau d\omega(\tau)}{1 + \dfrac{c_* \tau t(z)}{1 - zc_*^2 t^2(z)}}. \tag{9.16}$$

The equations (9.15) and (9.16) are identical up to factor $-c_*$. Since it can be shown that Eq. (9.16) has a unique solution on the set of Stieltjes transforms, we obtain that $g(z) = -c_* t(z)$. Therefore, (9.14) leads to the equation

$$\tilde{f}(z) = -\frac{1}{z\,[1 - z(c_* t(z))^2]}.$$

The Stieltjes transform of the limit eigenvalue distribution of $W_f W_p^* W_p W_f^*$ is clearly equal to $\frac{1}{c_*}\left(\tilde{f}(z) + \frac{1-c_*}{z}\right)$. Using the expression (9.2) of $t_\nu(z)$, we obtain immediately that

$$\frac{1}{c_*}\left(\tilde{f}(z) + \frac{1 - c_*}{z}\right) = t_\nu(z).$$

We have thus proved that the limit eigenvalue distribution of $W_f W_p^* W_p W_f^*$ can be evaluated using free probability technics.

Acknowledgments

The authors thank warmly Prof. Leonid Pastur for fruitful discussions. This work is partially funded by the Bézout Labex, funded by ANR, reference ANR-10-LABX-58, and by the ANR Project HIDITSA, reference ANR-17-CE40-0003.

References

[1] Z. Bai, J.W. Silverstein, "Spectral Analysis of Large Dimensional Random Matrices", Mathematics Monograph Series 2, Science Press, Beijing, 2006.

[2] M. Bhattachargee, A. Bose, "Large sample behaviour of high-dimensional autocovariance matrices", Ann. of Stat., vol. 44, no. 2, pp. 598-628, 2016a.

[3] A. Bose, M. Bhattachargee, "Large Covariance and Autocovariance Matrices", Monographs on Statistics and Applied Probability 162, CRC Press, 2019.

[4] F. Benaych-Georges, R.R. Nadakuditi, "The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices", Adv. in Math., 227(1):494-521, 2011

[5] F. Benaych-Georges, R.R. Nadakuditi,"The singular values and vectors of low rank perturbations of large rectangular random matrices", J. Multivariate Anal., vol. 111 (2012), pp. 120-135.

[6] S. Belinschi, M. Capitaine, “Spectral properties of polynomials in independent Wigner and deterministic matrices”, Journal of Functional Analysis, Volume 273, Issue 12, pp 3901-3963 (2017).

[7] M. Capitaine, C. Donati-Martin, “Strong asymptotic freeness of Wigner and Wishart matrices,” Indiana Univ. Math. Journal, vol. 56, pp. 295–309, 2007.

[8] M. Capitaine, C. Donati-Martin, and D. Féral, “The largest eigenvalue of finite rank deformation of large Wigner matrices: convergence and non-universality of the fluctuations,” Ann. of Proba., vol. 37, no. 1, pp. 1–47, 2009.

[9] H. Cartan, “Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes”, Hermann, 1978.

[10] F. Chapon, R. Couillet, W. Hachem, X. Mestre, ”The outliers among the singular values of large rectangular random matrices with additive fixed rank deformation” , Markov Processes and Related Fields, 20, 183–228 (2014).

[11] A. Ciuso, G. Picci, ”Asymptotic variance of subspace estimates”, J. of Econometrics 118 (1-2):257-291, February 2004.

[12] R. Couillet, ”Robust spiked random matrices and a robust G-MUSIC estimator”, Journal of Multivariate Analysis, vol. 140, pp. 139-161, 2015.

[13] B. Dozier, J. Silverstein, ”On the Empirical Distribution of Eigenvalues of Large Dimensional Information-Plus-Noise Type Matrices”, Journal of Multivariate Analysis, 98(4) (2007), pp. 678-694.

[14] J. Evans, D. N. C. Tse, “Large System Performance of Linear Multiuser Receivers in Multipath Fading Channels,” IEEE Transactions on Information Theory, Vol. 46, No. 6, 2000.

[15] U. Haagerup, S. Thorbjornsen, ”A new application of random matrices: $Ext(C^*_{red}(F2))$ is not a group”, Annals of Mathematics, 162(2):711, 2005.

[16] W. Hachem, P. Loubaton, and J. Najim, ”Deterministic equivalents for certain functionals of large random matrices”, Annals of Applied Probability, 17(3):875–930, 2007.

[17] W. Hachem, P. Loubaton, J. Najim, X. Mestre, P. Vallet ”A subspace estimator for fixed rank perturbations of large random matrices”, Journal of Multivariate Analysis, 114 (2013), pp. 427-447

[18] F. Hiai, D. Petz, ”Asymptotic freeness almost everywhere for random matrices”, Acta Sci. Math. (Szeged), 66/3–4 (2000), 809–834

[19] R.A. Horn, C.R. Johnson, “Matrix Analysis”, Cambridge University Press, 1985.

[20] B. Jin, C. Wang, Z.D. Bai, K.K. Nair, M. Harding, “Limited spectral distribution of a symmetrised auto-cross covariance matrices”, Ann. of Proba., vol. 24, no. 3, pp. 1199-1225, 2014.

[21] M.G. Krein, A.A. Nudelman, ”The Markov Moment Problem and Extremal Problems”, Translations of Mathematical Monographs, vol. 50, American Mathematical Society, Providence, Rhode Island, 1977.

[22] Z. Li, G. Pan, J. Yao, ”On singular value distribution of large-dimensional autocovariance matrices”, J. Multivariate Analysis, vol. 137, pp. 119-140, May 2015.

[23] Z. Li, Q. Wang, J. Yao, ”Identifying the number of factors from singular values of a large sample auto-covariance matrix”, Ann. of Stat., vol. 45, no. 1, pp. 257-288, 2017.

[24] A. Lindquist, G. Picci, “Linear Stochastic Systems”, Series in Contemporary Mathematics, Vol. 1, Springer, 2015

[25] H. Liu, A. Aue, D. Paul, “On the Marcenko-Pastur law for linear time series”, Ann. of Stat., vol. 43, no. 2, pp. 675-712, 2015.

[26] P. Loubaton, ”On the almost sure location of the singular values of certain Gaussian

block-Hankel large random matrices", J. of Theoretical Probability, vol. 29, no. 4, pp. 1339-1443, December 2016.

[27] C. Male, "The norm of polynomials in large random and deterministic matrices", Probability Theory and Related Fields, vol. 154, Issue 3-4, pp 477–532, December 2012.

[28] Mestre, X. and Lagunas, M.A., "Modified subspace algorithms for DoA estimation with large arrays", *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 598-614, February 2008.

[29] J. Mingo, R. Speicher, "Free Probability and Random Matrices", Fields Institute Monographs, Springer, 2017.

[30] R.R. Nadakuditi, A. Edelman, "Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples", IEEE Trans. Signal Processing, vol. 56, no. 7, pp. 2625-2637, Jul. 2008.

[31] R.R. Nadakuditi, J.W Silverstein, "Fundamental limit of sample generalized eigenvalue based detection of signals in noise using relatively few signal-bearing and noise-only samples", IEEE J. Sel. Topics Signal Processing, vol. 4, no. 3, pp. 468 -480, Jun. 2010.

[32] L.A. Pastur, M. Shcherbina, "Eigenvalue Distribution of Large Random Matrices", Mathematical Surveys and Monographs, Providence: American Mathematical Society, 2011.

[33] D. Paul, "Asymptotics of sample eigenstruture for a large dimensional spiked covariance model", Statistica Sinica, Vol. 17, No. 4, 1617-1642, 2007.

[34] Yu.A. Rozanov, "Stationary Random Processes", Holden Day, 1967.

[35] J. W. Silverstein, and S. I. Choi, "Analysis of the limiting spectral distribution of large dimensional random matrices", J. Multivariate Anal. 54. 295-309 (1995).

[36] G.T. Pham, P. Loubaton, P. Vallet, "Performance analysis of spatial smoothing schemes in the context of large arrays", IEEE Trans. on Signal Processing, vol. 64, no. 1, pp. 160-172, January 2016.

[37] P. Van Overschee, B. de Moor, "Subspace Identification for Linear Systems: Theory, Implementation, Applications", Kluwer Academic Publishers, 1996.

[38] J. Vinogradova, R. Couillet, and W. Hachem, "Estimation of Toeplitz covariance matrices in large dimensional regime with application to source detection", IEEE Transactions on Signal Processing, 63 (18), pages 4903-4913, September 2015.

[39] J. Vinogradova, R. Couillet and W. Hachem, "Statistical Inference in Large Antenna Arrays under Unknown Noise Pattern", IEEE Transactions on Signal Processing, 61 (22), 2013, pages 5633-5645.

[40] D. Voiculescu, K. J. Dykema, A. Nica, "Free Random Variables", ser. CRM Monograph. Providence, RI: Amer. Math. Soc., 1992, vol. 1.

[41] D. Voiculescu, "Lectures on Free Probability Theory", in unpublished Notes for a Course at the Saint-Flour Summer School on Proability, 1998.