On the largest singular values of certain large random matrices with application to the estimation of the minimal dimension of the state-space representations of high-dimensional time series.

D. Tieplova, P. Loubaton, *Fellow, IEEE*,

Abstract

This paper is devoted to the estimation of the minimal dimension P of the state-space realizations of a high-dimensional time series y, defined as a noisy version (the noise is white and Gaussian) of a useful signal with low rank rational spectral density, in the high-dimensional asymptotic regime where the number of available samples N and the dimension of the time series M converge towards infinity at the same rate. In the classical low-dimensional regime, P is estimated as the number of significant singular values of the empirical autocovariance matrix between the past and the future of y, or as the number of significant estimated canonical correlation coefficients between the past and the future of y. Generalizing large random matrix methods developed in the past to analyze classical spiked models, the behaviour of the above singular values and canonical correlation coefficients is studied in the high-dimensional regime. It is proved that they are smaller than certain thresholds depending on the statistics of the noise, except a finite number of outliers that are due to the useful signal. The number of singular values of the sample autocovariance matrix above the threshold is evaluated, is shown to be almost independent from P in general, and cannot therefore be used to estimate P accurately. In contrast, the number s of canonical correlation coefficients larger than the corresponding threshold is shown to be less than or equal to P, and explicit conditions under which it is equal to P are provided. Under the corresponding assumptions, s is thus a consistent estimate of P in the high-dimensional regime. The core of the paper is the development of the necessary large random matrix tools.

Index Terms

Minimal state space realization of rational spectrum time series, autocovariance matrix between the past and the future, canonical correlation coefficients between the past and the future, high-dimensional regime, large Gaussian random matrix theory, low rank perturbations of large random matrices, Stieltjes transform.

I. Introduction

A. The addressed problem and the results.

Due to the spectacular development of data acquisition devices and sensor networks, it becomes very common to be faced with high-dimensional time series in various fields such as digital communications, environmental sensing, electroencephalography, analysis of financial datas, industrial monitoring, ... In this context, it is not always possible to collect a large enough number of observations to perform statistical inference because the durations of the signals are limited and/or because their statistics are not time-invariant over large enough temporal windows. As a result, fundamental inference schemes do not behave as in the classical low-dimensional regimes. This stimulated considerably in the ten past years the development of new statistical approaches aiming at mitigating the above mentioned difficulties.

Daria Tieplova was with the Department of Theoretical Physics, Institute for Low Temperature Physics and Engineering, Kharkiv, Ukraine and with Laboratoire d'Informatique Gaspard Monge, UMR 8049, Université Paris-Est Marne la Vallée, 5 Bd. Descartes, Cité Descartes, 77454 Marne la Vallée Cedex 2, France. She is now with the Department of Statistics and Actuarial Science, Run Run Shaw Building, The University of Hong Kong, Pokfulam Road, Hong Kong, daria.tieplova@u-pem.fr

Philippe Loubaton is with Laboratoire d'Informatique Gaspard Monge, UMR 8049, Université Paris-Est Marne la Vallée, 5 Bd. Descartes, Cité Descartes, 77454 Marne la Vallée Cedex 2 , France, philippe.loubaton@u-pem.fr

This work was partially supported by Labex Bézout, funded by ANR, project ANR-10-LABX-58, and by ANR project HIDITSA, ANR-17-CE40-0003

In particular, a number of works proposed to use large random matrix theory in the context of high-dimensional statistical signal processing, traditionally modelled by a double asymptotic regime in which the dimension M of the time series and the sample size N both converge towards $+\infty$. These contributions addressed, among others, detection or estimation problems in the context of the so-called narrow band array processing model, also known in the statistical community as the linear factor model. The M–dimensional observation $(y_n)_{n=1,\ldots,N}$ is a noisy version of a useful signal $(u_n)_{n\in\mathbb{Z}}$ that can be written as $u_n = Hs_n$ where $(s_n)_{n\in\mathbb{Z}}$ is a K-dimensional non observable signal and H is a $M \times K$ unknown (or partially unknown) deterministic matrix. In this context, some relevant informations have to be inferred on the useful signal $(u_n)_{n\in\mathbb{Z}}$ from the available samples $y_1, \ldots, y_N$, e.g. estimation of K, of the column space of H, the non zero eigenvalues and associated eigenvectors of the covariance matrix $R_u = \mathbb{E}(u_n u_n^*)$ when $(s_n)_{n\in\mathbb{Z}}$is assumed to be a stationary sequence,....The $M \times N$ observed matrix Y collecting the N observations appears as the sum of a full rank random matrix due to the additive noise with the $M \times N$ rank K matrix $U = HS$ where U and S are defined in the same way than Y. In this context, a number of detection and estimation schemes are based on functionals of the empirical "spatial" covariance matrix $\hat{R}_y = \frac{YY^*}{N}$. In the traditional low dimensional regime where M is fixed while $N \to +\infty$, $\hat{R}_y$ behaves as the true covariance matrix $R_y = \mathbb{E}(y_n y_n^*)$ of the observation, and this allows to evaluate quite easily the behaviour of the various algorithms. The main difficulty of the high-dimensional regime follows from the well known observation that, when M and N converge towards $+\infty$ at the same rate, then $\hat{R}_y$ is not a good estimator of R_y in the sense that the spectral norm $\|\hat{R}_y - R_y\|$ of the estimation error does not converge towards 0. However, when the rank K does not scale with M and N, an assumption which in practice means that $\frac{K}{M}$ is small enough, large random matrix theory results related to the so-called spiked models, characterizing, among others, the eigenvalue distribution and the K largest eigenvalues and related eigenvectors of $\hat{R}_y$ (see e.g. [3], [4], [6], [7], [40]), allow to evaluate the behaviour of the relevant functionals of $\hat{R}_y$, to analyze the performance of the traditional schemes, and, sometimes, to propose improved algorithms (see e.g. [8], [15], [14] [24], [29], [36], [37], [50], [46], [51]). In particular, provided the K non zero eigenvalues of R_u are larger than a certain threshold depending on the noise statistics, then, under certain extra assumptions, K can be estimated consistently as the number of "significant" eigenvalues of $\hat{R}_y$.

In this work, we consider the more general context where the useful signal $(u_n)_{n\in\mathbb{Z}}$ coincides with the output of a K inputs / M outputs filter, $K < M$, with unknown causal and causally invertible rational transfer function $H(z)$ driven by a K dimensional non observable sequence $(i_n)_{n\in\mathbb{Z}}$ verifying $\mathbb{E}(i_{n+k} i_n^*) = I_K \delta_k$, which, necessarily, coincides with a normalized version of the innovation sequence of u defined as the prediction error $u_n - u_n/\text{sp}(u_{n-k}, k \geq 1)$. Normalized version of the innovation means that for each n, the components of i_n represent an orthonormal basis of the K–dimensional space generated by the components of $u_n - u_n/\text{sp}(u_{n-k}, k \geq 1)$. We remark that, for each frequency $f \in [-1/2, 1/2]$, the spectral density of u is a rank $K < M$ matrix, except if $e^{2i\pi f}$ is a zero of $H(z)$. In the following, we denote by P the Mac-Millan degree of $H(z)$, i.e. the minimal dimension of the state-space representations of $H(z) = D + C(zI - A)^{-1}B$ where A is a $P \times P$ matrix with spectral radius $\rho(A) < 1$ and where C, B, D are $M \times P, P \times K, M \times M$ matrices respectively. It is well known (see e.g. [27], [49], [31], Appendix A) that the minimality of the state-space representation of $H(z)$ is equivalent to (C, A) observable and (A, B) commandable. We recall that (C, A) observable means that for each $L \geq P$, the $ML \times P$ observability matrix $\mathcal{O}^{(L)}$ of (C, A) defined by

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

is a rank P matrix, while, similarly, (A, B) is commandable if the $P \times ML$ commandability matrix $\mathcal{C}^{(L)}$ of (A, B) defined by

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

is rank P as well. Then, u_n can be represented in the state-space form as

$$x_{n+1} = Ax_n + Bi_n, \; u_n = Cx_n + Di_n, \tag{I.3}$$

where the P-dimensional Markovian sequence $(x_n)_{n\in\mathbb{Z}}$ is called the state-space sequence associated to (I.3). Moreover, we assume that the observed M–dimensional multivariate time series $(y_n)_{n\in\mathbb{Z}}$ is given by

$$y_n = u_n + v_n, \tag{I.4}$$

where $(v_n)_{n\in\mathbb{Z}}$ is a complex Gaussian "noise" term such that $\mathbb{E}(v_{n+k}v_n^*) = R\delta_k$ for some unknown positive definite matrix R. $(v_n)_{n\in\mathbb{Z}}$ is of course independent from the useful signal u.

The estimation of the (minimal) dimension P of the state-space representation (I.3) from N avalaible samples $y_1,\ldots,y_N$ is an important problem of multivariate time series analysis in that estimating P first allows to address the estimation of matrices C and A, as well as of matrices B, D and R, at least if the three later matrices are identifiable. In the standard asymptotic regime where $N \to +\infty$ while M remains fixed, standard estimation procedures are based on the following well known ingredients (the reader is referred e.g. to [49] and [31] and the references therein). First, as $(v_n)_{n\in\mathbb{Z}}$ is an uncorrelated sequence, the autocovariance sequence $(R_{y,k})_{k\in\mathbb{Z}}$ defined by $R_{y,k} = \mathbb{E}(y_{n+k}y_n^*)$ verifies $R_{y,k} = R_{u,k} = \mathbb{E}(u_{n+k}u_n^*)$ for each $k \neq 0$. Next, the autocovariance sequence R_u of u can be represented as

$$R_{u,k} = CA^{k-1}G \tag{I.5}$$

for each $k \geq 1$, where matrix G coincides with $G = \mathbb{E}(x_{n+1}u_n^*)$, which is also equal to $G = \mathbb{E}(x_{n+1}y_n^*)$ because signals u and v are uncorrelated. Moreover, the pair (A, G) is commandable, and every triple (A', C', G') of $P \times P, M \times P, P \times M$ matrices for which (I.5) holds can be obtained from (A, C, G) by a similarity transform. If we define the autocovariance matrix $R^L_{f|p,u}$ between the past and the future of u as

$$R^L_{f|p,u} = \mathbb{E}\left[\begin{pmatrix} u_{n+L} \\ u_{n+L+1} \\ \vdots \\ u_{n+2L-1}\end{pmatrix}\left(u_n^*, u_{n+1}^*, \ldots, u_{n+L-1}^*\right)\right] \tag{I.6}$$

then, it holds that

$$R^{(L)}_{f|p,u} = \mathcal{O}^{(L)}\,\mathcal{C}^{(L)}, \tag{I.7}$$

where $\mathcal{O}^{(L)}$ is the observability matrix of the pair (C, A) and $\mathcal{C}^{(L)}$ represents the commandability matrix of (A, G). For each $L \geq P$, matrices $\mathcal{O}^{(L)}$ and $\mathcal{C}^{(L)}$ are full rank, so that 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 (I.7) for some particular triple (A, C, G). As matrix $R^L_{f|p,y}$ defined in the same way than $R^{(L)}_{f|p,u}$ coincides with $R^{(L)}_{f|p,u}$, we deduce from the above properties that P coincides with the rank of $R^L_{f|p,y}$ for each integer $L \geq P$. Moreover, a particular pair (C, A) can be identified from any minimal rank factorisation of $R^L_{f|p,y}$.

In order to estimate P from the available samples $y_1,\ldots,y_N$, a standard approach is to estimate P as the number of "significant" singular values of the empirical estimate $\hat{R}^L_{f|p,y}$ of the true matrix $R^L_{f|p,y} = R^L_{f|p,u}$ defined by

$$\hat{R}^L_{f|p,y} = \frac{Y_{f,N}Y^*_{p,N}}{N},$$

where matrices $Y_{f,N}$ and $Y_{p,N}$ are defined as

$$Y_{p,N} = \begin{pmatrix} y_1 & y_2 & \cdots & y_{N-1} & y_N \\ y_2 & y_3 & \cdots & y_N & y_{N+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_L & y_{L+1} & \cdots & y_{N+L-2} & y_{N+L-1}\end{pmatrix} \tag{I.8}$$

and

$$Y_{f,N} = \begin{pmatrix} y_{L+1} & y_{L+2} & \cdots & y_{N-1+L} & y_{N+L} \\ y_{L+2} & y_{L+3} & \cdots & y_{N+L} & y_{N+L+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{2L} & y_{2L+1} & \cdots & y_{N+2L-2} & y_{N+2L-1}\end{pmatrix}. \tag{I.9}$$

We note that the samples $(y_{N+l})_{l=1,\ldots,2L-1}$ are supposed to be available while we have assumed that only the first N samples are observed. In order to simplify the presentation, this end effect is neglected. We also notice that a pair (C, A) can also be estimated from the truncated singular value decomposition of $\hat{R}^L_{f|p,y}$ (see [49] and [11] for a statistical analysis of the corresponding estimates). This approach provides consistent estimates of P, C, A when $N \rightarrow +\infty$ while M, K, P and L are fixed because in this context, $\|\hat{R}^L_{f|p,y} - R^L_{f|p,y}\| \rightarrow 0$.

Another way to estimate P is to resort to the canonical analysis of the observation y. In particular, P coincides with the number of non zero canonical correlation coefficients between the spaces $\mathcal{Y}_{p,L}$ and $\mathcal{Y}_{f,L}$ generated respectively by the components of $y_{n+k}, k = 0, \ldots, L-1$ and $y_{n+k}, k = L, \ldots, 2L-1$ for any $L \geq P$. We recall that these coefficients are defined as the singular values of matrix $(R^L_y)^{-1/2} R^L_{f|p,y} (R^L_y)^{-1/2}$ where R^L_y represents the covariance matrix of the ML–dimensional vector $(y_n^T, \ldots, y_{n+L-1}^T)^T$. In order to estimate P from the N avalaible observations $y_1, \ldots, y_N$, a standard solution is to estimate the canonical correlation coefficients between $\mathcal{Y}_{p,L}$ and $\mathcal{Y}_{f,L}$ by the canonical correlation coefficients between the row spaces of matrices $Y_{p,N}$ and $Y_{f,N}$ defined by (I.8) and (I.9) respectively, and to estimate P as the number of significant coefficients, i.e. as the number of significant singular values of matrix $(\hat{R}^L_{f,y})^{-1/2} \hat{R}^L_{f|p,y} (\hat{R}^L_{p,y})^{-1/2}$, or equivalently as the number of significant eigenvalues of $(\hat{R}^L_{f,y})^{-1/2} \hat{R}^L_{f|p,y} (\hat{R}^L_{p,y})^{-1} \hat{R}^{L*}_{f|p,y} (\hat{R}^L_{f,y})^{-1/2}$. Here, matrices $\hat{R}^L_{f,y}$ and $\hat{R}^L_{p,y}$ are defined by $\hat{R}^L_{f,y} = \frac{Y_{f,N} Y^*_{f,N}}{N}$ and $\hat{R}^L_{p,y} = \frac{Y_{p,N} Y^*_{p,N}}{N}$ respectively. In the standard low-dimensional regime $N \rightarrow +\infty$ and M, K, P, L are fixed, it holds that $\|\hat{R}^L_{i,y} - R^L_y\| \rightarrow 0$ for $i = p, f$ as well as $\|\hat{R}^L_{f|p,y} - R^L_{f|p,y}\| \rightarrow 0$. This immediately leads to the conclusion that this approach provides consistent estimates of P. We again refer to [49] and [31] and the references therein.

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, in the sense that it cannot be expected that $\hat{R}^L_{f|p,y}$ and $\hat{R}^L_{i,y}$, $i = p, f$ are close enough in the spectral norm sense from the true matrices $R^L_{f|p,y}$ and R^L_y respectively. It is thus relevant to study the behaviour of the above estimators of P in asymptotic regimes where M and N both converge towards $+\infty$ in such a way that $c_N = \frac{ML}{N}$ converges towards a non zero constant c_*. In this context, matrix $\hat{R}^L_{f|p,y}$ is no longer a consistent estimate of the true matrix $R^{(L)}_{f|p,y}$ in the spectral norm sense. Therefore, the singular values of $\hat{R}^{(L)}_{f|p,y}$ have no reasons to behave as those of $R^{(L)}_{f|p,y}$, and the same conclusion holds for matrices $(\hat{R}^L_{f,y})^{-1/2} \hat{R}^L_{f|p,y} (\hat{R}^L_{p,y})^{-1/2}$ and $(R^L_y)^{-1/2} R^L_{f|p,y} (R^L_y)^{-1/2}$. Thus, it appears of fundamental interest to evaluate the behaviour of the singular values of $\hat{R}^L_{f|p,y}$ and $(\hat{R}^L_{f,y})^{-1/2} \hat{R}^L_{f|p,y} (\hat{R}^L_{p,y})^{-1/2}$, and to study whether the largest singular values still allow to estimate P consistently, at least if the power of the useful signal u and the non zero singular values of $R^L_{f|p,u}$ or the non zero canonical correlation coefficients between the spaces $\mathcal{U}_{p,L}$ and $\mathcal{U}_{f,L}$ are large enough.

In this paper, we address these problems when the integers K and P do not scale with M and N, and thus remain fixed integers. This in practice means that the following results are likely to be useful when the rank K of the spectral density of u is much smaller than M, and when P is small enough compared to M and N. As P is supposed to be a fixed integer, the integer $L \geq P$ will also be assumed to remain fixed when M and N converge towards $+\infty$. As explained below, the assumption K, P, L remain fixed implies that the matrices $\hat{R}^L_{f|p,y}$ and $(\hat{R}^L_{f,y})^{-1/2} \hat{R}^L_{f|p,y} (\hat{R}^L_{p,y})^{-1/2}$ are low rank perturbations of the random matrices $\hat{R}^L_{f|p,v}$ and $(\hat{R}^L_{f,v})^{-1/2} \hat{R}^L_{f|p,v} (\hat{R}^L_{p,v})^{-1/2}$ built from the noise samples $v_1, \ldots, v_N$ instead of $y_1, \ldots, y_N$. It is thus in principle possible to use the perturbation techniques developed in [6], [7], [40]. However, the random matrix models that come into play in this paper are considerably more complicated than in [6], [7], [40]. Thus, the following results cannot be considered as direct consequences of [6], [7], [40].

We first evaluate in Section II the behaviour of the largest singular values of $\hat{R}^L_{f|p,y}$, or equivalently of the largest eigenvalues of $\hat{R}^L_{f|p,y} \hat{R}^{L*}_{f|p,y}$ and take benefit of the results in [33] in which the asymptotic behaviour of the eigenvalues of $\hat{R}^L_{f|p,v} \hat{R}^{L*}_{f|p,v}$ is characterized. Introducing some extra assumptions, we deduce from [33] that for each $\epsilon > 0$, almost surely, for N large enough, all the eigenvalues $\hat{R}^L_{f|p,v} \hat{R}^{L*}_{f|p,v}$ are less than $x_{+,*} + \epsilon$ for a certain $x_{+,*} > 0$. Using the perturbation techniques developed in [7] and [40], we obtain that the number of eigenvalues of $\hat{R}^L_{f|p,y} \hat{R}^{L*}_{f|p,y}$ that may escape from the interval $[0, x_{+,*}]$ is between 0 and $2r$ where r represents the rank of the covariance matrix $R^{(L)}_u$ of the vector $(u_n^T, \ldots, u_{n+L-1}^T)^T$. When $P = 1$ and $R = \sigma^2 I$ for some σ^2, for any $r \geq 1$, we indicate how to produce simple examples such that $2r - 1$ eigenvalues of $\hat{R}^L_{f|p,y}$ escape from $[0, x_{+,*}]$.

This behaviour leads to the conclusion that P cannot be estimated consistently as the number of eigenvalues that are larger than $x_{+,*}$ even if the useful u is powerful enough and the non zero singular values of $R^{(L)}_{f|p,u}$ are large enough. While it would be possible to address the case $c_* \geq 1$, we will assume that $c_* < 1$ to simplify the exposition. Therefore, c_N verifies $c_N < 1$ for each N large enough .

Always under the assumption $c_* < 1$, using the same approach, we then study in Section III the largest eigenvalues of $(\hat{R}^L_{f,y})^{-1/2}\hat{R}^L_{f|p,y}(\hat{R}^L_{p,y})^{-1}(\hat{R}^L_{f|p,y})^*(\hat{R}^L_{f,y})^{-1/2}$, which also coincide with those of matrix $\Pi_{p,y}\Pi_{f,y}$ where $\Pi_{p,y}$ and $\Pi_{f,y}$ represent the orthogonal projection matrices on the spaces generated by the rows of Y_p and Y_f respectively. We first study the eigenvalue distribution of $\Pi_{p,v}\Pi_{f,v}$, and establish that it converges towards the free multiplicative convolution product of $c_*\delta_1 + (1-c_*)\delta_0$ with itself. Notice that $c_N\delta_1 + (1-c_N)\delta_0$ coincides with the eigenvalue distribution of matrices $\Pi_{p,v}$ and $\Pi_{f,v}$. We also establish that almost surely, for N large enough, all the eigenvalues of $\Pi_{p,v}\Pi_{f,v}$ lie in a neighbourhood of the support $[0, 4c_*(1-c_*)] \cup \{1\}\mathbf{1}_{c_*>1/2}$ of its limit distribution. Using the above mentioned perturbation techniques, we establish that if s represents the number of eigenvalues of $\Pi_{p,y}\Pi_{f,y}$ that escape from $[0, 4c_*(1-c_*)] \cup \{1\}\mathbf{1}_{c_*>1/2}$, then, $s \leq P$, and eventually provide the explicit conditions under which $s = P$. These conditions hold if $c_* < 1/2$ and if the power of u and the non zero canonical correlation coefficients between the spaces $\mathcal{U}_{p,L}$ and $\mathcal{U}_{f,L}$ are large enough. These results allow to conclude that, under certain reasonable well defined assumptions, it is possible to estimate P consistently using the largest singular values of $(\hat{R}^L_{f,y})^{-1/2}\hat{R}^L_{f|p,y}(\hat{R}^L_{p,y})^{-1/2}$, but that the use of the largest singular values of $\hat{R}^L_{f|p,y}$ appears unreliable.

It has hard to explain intuitively why the use of the normalized matrix $(\hat{R}^L_{f,y})^{-1/2}\hat{R}^L_{f|p,y}(\hat{R}^L_{p,y})^{-1/2}$ allows to estimate P consistently under certain assumptions, while this is not the case for matrix $\hat{R}^L_{f|p,y}$. We however mention that matrix $(\hat{R}^L_{f,v})^{-1/2}\hat{R}^L_{f|p,v}(\hat{R}^L_{p,v})^{-1/2}$ defined by replacing y by v does not depend on the covariance matrix R of the random vectors $(v_n)_{n\in\mathbb{Z}}$, while it is of course not the case of matrix $\hat{R}^L_{f|p,v}$. This invariance property appears of course attractive, and plays an important role in the following. We also mention that matrix $(\hat{R}^L_{f,y})^{-1/2}\hat{R}^L_{f|p,y}(\hat{R}^L_{p,y})^{-1/2}$ is connected with the canonical analysis of the time series y. Generally speaking, this analysis has well established merits. In particular, it leads to the concept of stochastically balanced state-space realizations which are known to be useful to derive model reduction algorithms ([49] and [31] and the references therein).
We believe that the main findings of this paper are of potential interest for statistical signal processing and time series analysis researchers. However, the large random matrix models that come into play in this paper are rather complicated, and were almost not considered in previous works. Therefore, new random matrix tools have to be developed and a number of technical intermediate results have to be established. In order to improve the readability of this paper, we postpone the most technical steps in the Appendix, and sometimes provide sketches of proof rather than detailed arguments.

B. On the literature.

We first mention that the problems considered in this paper have connections with the "Generalized Dynamic Factor Models" introduced in the econometrics field, see e.g. [17], [18], [16]. In these works, the observation is still given by $y_n = u_n + v_n$ where $u_n = [H(z)]i_n$ and v_n are called the common component and the idiosyncratic component respectively. $H(z)$ is not assumed in [17] and [18] to be rational, while v is not necessarily an uncorrelated time series. These papers still address estimation problems in the asymptotic regime where M and N converge towards $+\infty$, but [17], [18], [16] assume that the eigenvalues of the spectral density matrix of v remain bounded when M and N increase, while the K non zero eigenvalues of the spectral density of u converge towards $+\infty$. In this context, it appears possible to estimate consistently from the available samples a number of parameters attached to the useful signal u. In particular, if $H(z)$ is rational, the estimation of P does not pose any problem (see [16] devoted to the case $H(z)$ rational). In contrast, the technical assumptions formulated in the present paper imply that the eigenvalues of the spectral densities of u and v are of the same order of magnitude when M and N increase. We refer the reader to [41] for a discussion on the practical relevance of the context of the present paper. Therefore, the solutions developed in [17], [18], [16] cannot be used to design consistent estimates of P under our assumptions.

We next review the existing works that are more directly related to the present paper. The behaviour of the eigenvalues of matrix $R^L_{f|p,v}R^{L*}_{f|p,v}$ was studied in [33], and we refer to this paper for the various references that addressed similar problems when $L = 1$, in the non Gaussian case, or when the time series $(v_n)_{n\in\mathbb{Z}}$ is possibly correlated in the time domain. Apart [30], we are not aware of any previous work addressing the behaviour of the largest singular values of matrices depending on estimated autocovariance matrices of y at non zero lags in the presence of a low rank useful signal u. [30] assumes that v is possibly non Gaussian with covariance matrix $R = \sigma^2 I$, and that the useful signal u is given by $u_n = Hs_n$ where H is a $M \times K$ matrix verifying $H^*H = I_K$ and where the components $(s_{k,n})_{n\in\mathbb{Z}}$ of $(s_n)_{n\in\mathbb{Z}}$ are independent times series. Using the above mentioned perturbation analysis, [30] studies the eigenvalues of $\hat{R}^1_{f|p,y}\hat{R}^{1*}_{f|p,y}$ that escape from the interval $[0, x_{+,*}]$ introduced above. We notice that if $L = 1$, matrix $\hat{R}^1_{f|p,y}$ coincides with the standard estimate of the autocovariance matrix of y at lag 1.

We finally mention that a number of previous works addressed the behaviour of the canonical correlation coefficients between the row spaces of two large random matrices. However, the underlying random matrix models are simpler than in the present paper. More specifically, the structured random matrices $Y_{p,L}$ and $Y_{f,L}$ as well as $V_{p,L}$ and $V_{f,L}$ are replaced by mutually independent matrices X_1 and X_2 with i.i.d. elements, a property that is not verified by $Y_{p,L}$, $Y_{f,L}$, $V_{p,L}$ and $V_{f,L}$. [53] addressed the case of $M \times N$ mutually independent complex Gaussian matrices X_1 and X_2 with i.i.d. entries, and derived the corresponding limit distribution of the squared canonical correlation coefficients. This is equivalent to evaluating the limit eigenvalue distribution of $\Pi_1\Pi_2$ where Π_1 and Π_2 represent the orthogonal matrices on the row spaces of X_1 and X_2. We note that the result of [53] appears as a trivial consequence of basic free probability theory results (see e.g. [52] [25], [35], as well as [47] for a more engineering oriented presentation) because under the above hypotheses, Π_1 and Π_2 are almost surely asymptotically free. More recently, [54] extended this result to the case where X_1 and X_2 are independent matrices with non Gaussian i.i.d. entries. We also note that [55] took benefit of this result to propose independence tests between 2 sets of i.i.d. high-dimensional samples. We mention that [5] extended the result of [53] to the case where X_1 and X_2 have Gaussian i.i.d. entries, but this time $\mathbb{E}\{\frac{X_1X_2^*}{N}\}$ is a non zero low rank matrix. We finally notice that in [44], we established the convergence of the eigenvalue distribution of $\Pi_{p,v}\Pi_{f,v}$ by establishing the almost sure freeness of $\Pi_{p,v}$ and $\Pi_{f,v}$. We however mention that in order to study the largest eigenvalues of $\Pi_{p,y}\Pi_{f,y}$ using perturbation techniques, it is also necessary to evaluate the asymptotic behaviour of the resolvent of $\Pi_{p,v}\Pi_{f,v}$, a more difficult issue that is solved in the present paper.

C. Assumptions, notations and basic tools.

We now introduce the main assumptions, notations and fundamental tools that will be used throughout this paper.

Assumptions

- We assume that L is a fixed parameter verifying $L \geq P$, and that M and N converge towards $+\infty$ in such a way that
$$c_N = \frac{ML}{N} \to c_*, \; 0 < c_* < 1 \tag{I.10}$$
This regime will be referred to as $N \to +\infty$ in the following. In the regime (I.10), 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 notice that the results of Section II devoted to the study of the largest eigenvalues of $\hat{R}^L_{f|p,y}\hat{R}^{L*}_{f|p,y}$ could be generalized to the case $c_* \geq 1$, but we prefer to assume $c_* < 1$ in order to simplify the presentation of the corresponding results. It therefore holds that $c_N < 1$ for each N large enough.
- The sequence of covariance matrices $(R_N)_{N\geq 1}$ of M–dimensional vectors $(v_n)_{n=1,\ldots,N}$ is supposed to verify
$$a\, I \leq R_N \leq b\, I \tag{I.11}$$
for each N, where $a > 0$ and $b > 0$ are 2 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 (I.11) is obviously equivalent to $\lambda_{M,N} \geq a$ and $\lambda_{1,N} \leq b$ for each N.

Notations

- For each $1 \leq i \leq 2L$ and $1 \leq m \leq M$, $\mathbf{f}_i^m$ 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 $i \leq L$, vector $\mathbf{f}_i^m$ may also represent, depending on the context, 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 $1 \leq j \leq N$ represents the j –th vector of the canonical basis of $\mathbb{C}^N$.

- For each integer $l \geq 1$, we define the $l \times l$ "shift" matrix J_l as

$$(J_l)_{ij} = \delta_{j-(i+1)}. \tag{I.12}$$

- $\mathbb{R}^+$ and $\mathbb{R}^-$ represent respectively the set of all non-negative numbers and non-positive numbers, and we denote $\mathbb{R}^* \equiv \mathbb{R} \setminus \{0\}$, $\mathbb{R}^{+*} \equiv \mathbb{R}^+ \setminus \{0\}$ and $\mathbb{R}^{-*} \equiv \mathbb{R}^- \setminus \{0\}$. We also define $\mathbb{C}^+ = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$. We finally denote by $\rho(z)$ the distance from $z \in \mathbb{C}$ to $\mathbb{R}^+$, i.e.

$$\rho(z) = \mathrm{dist}(z, \mathbb{R}^+) \tag{I.13}$$

- 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 that appears in the various Stieltjes transforms introduced in this paper. 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.
- 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.
- For any matrix A, $\|A\|$ and $\|A\|_F$ represent its spectral norm and Frobenius norm respectively. The transpose, conjugate, and conjugate transpose of A are respectively denoted by A^T ,$\bar{A}$ and A^*. If A is a square matrix, $\mathrm{Im}(A)$ is the Hermitian matrix defined by $\mathrm{Im}(A) = \frac{A - A^*}{2i}$. If A and B are Hermitian matrices, $A \geq B$ stands for $A - B$ non-negative definite.
- $\mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R})$ 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

$$\xi^\circ = \xi - \mathbb{E}\xi. \tag{I.14}$$

Fundamentals tools

If n is a positive integer, then a $n \times n$ matrix-valued positive measure ω is a σ–additive function from the Borel sets of $\mathbb{R}$ onto the set of all positive $n \times n$ matrices (see e.g. [42], Chapter 1 for more details). If ω is a $n \times n$ matrix-valued positive finite [1] measure, the Stieltjes transform S_ω of ω is the function defined for each $z \in \mathbb{C} \setminus \mathrm{Supp}(\omega)$ by

$$S_\omega(z) = \int \frac{d\omega(\lambda)}{\lambda - z} \tag{I.15}$$

In the following, if B is a Borel set of $\mathbb{R}$, we denote by $\mathcal{S}_n(B)$ the set of all Stieltjes transforms of $n \times n$ matrix-valued positive finite measures carried by B. $\mathcal{S}_1(B)$ is denoted $\mathcal{S}(B)$. We just mention the following useful properties of the elements of $\mathcal{S}_n(\mathbb{R})$ and $\mathcal{S}_n(\mathbb{R}^+)$: if $S \in \mathcal{S}_n(\mathbb{R})$ and if ω represents its associated $n \times n$ matrix-valued positive finite measure , then, S is analytic on $\mathbb{C} \setminus \mathbb{R}$ and verifies

$$\|S(z)\| \leq \frac{\|\omega(\mathbb{R})\|}{\mathrm{Im} z}, \ \mathrm{Im} S(z) \geq 0 \tag{I.16}$$

if $z \in \mathbb{C}^+$. Moreover, $\omega(\mathbb{R}) = \lim_{y \to +\infty} -iyS(iy)$. When the positive matrix $\omega(\mathbb{R})$ is positive definite, $\mathrm{Im} S(z) > 0$ on $\mathbb{C}^+$. If $S \in \mathcal{S}_n(\mathbb{R}^+)$, then S is analytic on $\mathbb{C} - \mathbb{R}^+$ and also satisfies

$$\mathrm{Im} z S(z) \geq 0, \ z \in \mathbb{C}^+, \ \|S(z)\| \leq \frac{\|\omega(\mathbb{R}_+)\|}{\rho(z)}, z \in \mathbb{C} - \mathbb{R}^+ \tag{I.17}$$

When $\omega(\mathbb{R}_+) > 0$, we also have $\mathrm{Im} z S(z) > 0$ on $\mathbb{C}^+$. We refer the reader to Proposition 4.1 in [33] for other useful properties, and for a converse of (I.16, I.17). We finally mention the following immediate properties:

$$S \in \mathcal{S}_n(\mathbb{R}^+) \implies \mathbf{S} \in \mathcal{S}_n(\mathbb{R}) \tag{I.18}$$

[1] finite means that $\mathrm{Tr}\,(\omega(\mathbb{R})) < +\infty$

where $\mathbf{S}(z)$ is defined for $z \in \mathbb{C}^+$ by $\mathbf{S}(z) = zS(z^2)$. Moreover, if ω and $\boldsymbol{\omega}$ are the positive matrix-valued measures associated to S and $\mathbf{S}$, the following equality holds:

$$\omega(\mathbb{R}^+) = \boldsymbol{\omega}(\mathbb{R}) \tag{I.19}$$

If A is a $n \times n$ matrix, the resolvent of A is defined as the matrix-valued function Q_A defined on $\mathbb{C} - \{\lambda_1(A), \dots, \lambda_n(A)\}$ by

$$Q_A(z) = (A - zI)^{-1} \tag{I.20}$$

If A is Hermitian, it is clear that Q_A coincides with the Stieltjes transform of the $n \times n$ positive matrix-valued measure ω_A given by

$$\omega_A = \sum_{k=1}^{n} \delta_{\lambda_k(A)} f_k(A) f_k(A)^*$$

where $(f_k(A))_{k=1,\dots,n}$ represent the eigenvectors of A. We notice that $\omega_A(\mathbb{R}) = I$, so that (I.16) leads to

$$\|Q_A(z)\| \leq \frac{1}{\mathrm{Im} z} \tag{I.21}$$

on $\mathbb{C}^+$ and

$$\|Q_A(z)\| \leq \frac{1}{\rho(z)} \tag{I.22}$$

on $\mathbb{C} - \mathbb{R}^+$ if $A \geq 0$. We also mention that Q_A satisfies the "resolvent identity"

$$I + zQ_A(z) = Q_A(z)\, A = A\, Q_A(z) \tag{I.23}$$

for each z. If $\nu_A = \frac{1}{n}\sum_{k=1}^{n} \delta_{\lambda_k(A)}$ represents the empirical eigenvalue distribution of A, $\frac{1}{n}\mathrm{Tr}\, Q_A(z)$ is the Stieltjes transform of ν_A.

We recall that if $(A_N)_{N \geq 1}$ is a sequence of $N \times N$ Hermitian (possibly random) matrices, a convenient way to study the behaviour of the sequence of probability measures $(\nu_{A_N})_{N \geq 1}$ when $N \to +\infty$ is to study the asymptotic behaviour of the corresponding Stieltjes transforms $S_{\nu_{A_N}}(z) = \frac{1}{N}\mathrm{Tr}\,(Q_{A_N}(z))$ because the weak convergence of sequence $(\nu_{A_N})_{N \geq 1}$ towards a probability measure ν_* is equivalent to the convergence of $S_{\nu_{A_N}}(z)$ towards the Stieltjes transform of ν_* for each $z \in \mathbb{C}^+$. This explains why Stieltjes transforms and resolvents play an important role in large random matrix theory. We refer the reader to e.g. [2], [39], [56]. See also [13] and [47] for more engineering oriented books.

We also recall Montel's theorem (see e.g. [12]), also called the Normal Family Theorem, which is frequently used in the large random matrix literature. If $(s_N(z))_{N \geq 1}$ is a sequence of functions that are holomorphic on a domain Ω, and such that, for each compact set $\mathcal{K} \subset \Omega$, $\sup_{N \geq 1} \sup_{z \in \mathcal{K}} |s_N(z)| < +\infty$, then it is possible to extract from $(s_N(z))_{N \geq 1}$ a subsequence converging uniformly on each compact subset of Ω towards a function $s_*(z)$ holomorphic on Ω. Note in particular that if for each $N \geq 1$, s_N is the Stieltjes transform of a probability measure, then $(s_N(z))_{N \geq 1}$ verifies the above assumptions for $\Omega = \mathbb{C}^+$ because $|s_N(z)| \leq \frac{1}{\mathrm{Im} z}$ on $\mathbb{C}^+$ for each $N \geq 1$.

In this paper, we will consider frequently $2n \times 2n$ matrices $\mathbf{A}$ given by

$$\mathbf{A} = \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix}$$

where B and C are $n \times n$ matrices. The resolvent $\mathbf{Q_A}$ of $\mathbf{A}$ is given by

$$\mathbf{Q_A}(z) = \begin{pmatrix} zQ_{BC}(z^2) & Q_{BC}(z^2)\, B \\ C\, Q_{BC}(z^2) & zQ_{CB}(z^2) \end{pmatrix} \tag{I.24}$$

If the eigenvalues of BC are real and positive, the eigenvalues of $\mathbf{A}$ are the $\pm \left(\sqrt{\lambda_k(BC)}\right)_{k=1,\dots,n}$.

We finally recall the two Gaussian tools that will be used in the sequel in order to evaluate the asymptotic behaviour of certain resolvents.

Proposition I.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*

$$\mathbb{E}\{\xi_i \Gamma(\xi)\} = \sum_{k=1}^{K} \Omega_{ik} \mathbb{E}\left\{ \frac{\partial \Gamma(\xi)}{\partial \bar{\xi}_k} \right\}. \tag{I.25}$$

Proposition I.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 $\nabla_{\bar{\xi}} \Gamma = [\frac{\partial \Gamma}{\partial \bar{\xi}_1}, \ldots, \frac{\partial \Gamma}{\partial \bar{\xi}_K}]^T$*

$$\mathbf{Var}\{\Gamma(\xi)\} \leq \mathbb{E}\left\{ \nabla_\xi \Gamma(\xi)^T \Omega \overline{\nabla_\xi \Gamma(\xi)} \right\} + \mathbb{E}\left\{ \nabla_{\bar{\xi}} \Gamma(\xi)^* \Omega \nabla_{\bar{\xi}} \Gamma(\xi) \right\}. \tag{I.26}$$

The combination of these two tools was first proposed in [38], see also [39] for an exhaustive reference. We also mention [22] in which Propositions I.1 and I.2 are used in order to study the capacity of large MIMO channels.

II. The largest singular values of the empirical autocovariance matrix.

A. *Review of the zero signal case.*

In this paragraph, we briefly present the useful results from [33] concerning the study of the singular values of matrix $\hat{R}^L_{f|p,v}$, or equivalently of the eigenvalues of $\hat{R}^L_{f|p,v}(\hat{R}^L_{f|p,v})^*$. All along Section II, we will denote by $W_{p,N}$ and $W_{f,N}$ the $ML \times N$ normalized matrices defined by

$$W_{p,N} = \frac{1}{\sqrt{N}} V_{p,N}, \; W_{f,N} = \frac{1}{\sqrt{N}} V_{f,N} \tag{II.1}$$

and by W_N the $2ML \times N$ matrix given by

$$W_N = \begin{pmatrix} W_{p,N} \\ W_{f,N} \end{pmatrix} \tag{II.2}$$

We first mention that matrices $W_{i,N}$, $i = p, f$ verify the following property.

There exists a nice constant κ such that, almost surely, for each N large enough, $\|W_{i,N}\| < \kappa$ (II.3)

If R_N was equal to I_M, this property would be an immediate consequence of Theorem 1.1 in [32]. In the general case, it is an immediate consequence of Eq. (3.1) in [33] and of (I.11).

In order to study the asymptotic behaviour of the eigenvalues of $W_{f,N} W^*_{p,N} W_{p,N} W^*_{f,N}$, [33] studied the behaviour of the resolvent, denoted $Q_{N,W}(z)$, of the $ML \times ML$ matrix $W_{f,N} W^*_{p,N} W_{p,N} W^*_{f,N}$, i.e.

$$Q_{N,W}(z) = \left(W_{f,N} W^*_{p,N} W_{p,N} W^*_{f,N} - zI \right)^{-1} \tag{II.4}$$

The entries of $Q_{N,W}$ are easily seen to concentrate almost surely around their mathematical expectations. Therefore, it is sufficient to study the behaviour of $\mathbb{E}(Q_{N,W}(z))$ using Propositions I.1 and I.2. As the entries of $W_{f,N} W^*_{p,N} W_{p,N} W^*_{f,N}$ are bi-quadratic functions of the entries of W_N, the Gaussian calculations that allow to evaluate $\mathbb{E}(Q_{N,W}(z))$ are very complicated. Therefore, [33] used the well-known linearization trick that consists in studying the resolvent $\mathbf{Q}_{N,W}(z)$ of the $2ML \times 2ML$ hermitized version

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

Formula (I.24) allows eventually to deduce $\mathbb{E}(Q_{N,W}(z))$ from the first diagonal block of $\mathbf{Q}_{N,W}(z)$. This linearization trick will also be used extensively in the present paper. In the following, every $2ML \times 2ML$ matrix $\mathbf{G}$ such as $\mathbf{Q}_N(z)$ will be written

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

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

In order to introduce the main results of [33], we recall Proposition 6.1 in [33]: for each $z \in \mathbb{C}^+$, the equation

$$t_N(z) = \frac{1}{M}\mathrm{Tr} R_N \left(-zI_M - \frac{z c_N t_N(z)}{1 - z c_N^2 t_N^2(z)} R_N\right)^{-1} \quad \text{(II.5)}$$

has a unique solution for which $t_N(z)$ and $z t_N(z)$ belongs to $\mathbb{C}^+$. Moreovoer, t_N is the Stieltjes transform of a positive measure μ_N carried by $\mathbb{R}^+$, and the $M \times M$ matrix-valued function $T_N(z)$ defined by

$$T_N(z) = -\left(zI_M + \frac{z c_N t_N(z)}{1 - z c_N^2 t_N^2(z)} R_N\right)^{-1}, \quad \text{(II.6)}$$

belongs to $\mathcal{S}_M(\mathbb{R}^+)$. Its associated positive matrix-valued measure, denoted ν_N^T, verifies $\nu_N^T(\mathbb{R}^+) = I$. We also define $\mathbf{t}_N(z)$ and $\mathbf{T}_N(z)$ by

$$\mathbf{t}_N(z) = z t_N(z^2) \quad \text{(II.7)}$$

and

$$\mathbf{T}_N(z) = z T_N(z^2) = \left(-zI_M - \frac{c_N \mathbf{t}_N(z)}{1 - c_N^2 \mathbf{t}_N^2(z)} R_N\right)^{-1} \quad \text{(II.8)}$$

which, by (I.18), belong to $\mathcal{S}(\mathbb{R})$ and $\mathcal{S}_M(\mathbb{R})$ respectively. Moreover, the positive matrix-valued measure $\boldsymbol{\nu}_N^{\mathbf{T}}$ associated to $\mathbf{T}_N$ verifies $\boldsymbol{\nu}_N^{\mathbf{T}}(\mathbb{R}) = \nu_N^T(\mathbb{R}^+) = I$. Then, the following Proposition can be deduced from the results of [33].

Proposition II.1. *We consider sequences of deterministic $ML \times ML$ and $2ML \times 2ML$ matrices $(A_N)_{N\geq 1}$ and $(\mathbf{A}_N)_{N\geq 1}$ verifying $\sup_N \|A_N\| < +\infty$ and $\sup_N \|\mathbf{A}_N\| < +\infty$. Then, we have*

$$\frac{1}{ML}\mathrm{Tr}\left((Q_N(z) - I_L \otimes T_N(z))A_N\right) \to 0 \quad \text{(II.9)}$$

and

$$\frac{1}{2ML}\mathrm{Tr}\left((\mathbf{Q}_N(z) - I_{2L} \otimes \mathbf{T}_N(z))\mathbf{A}_N\right) \to 0 \quad \text{(II.10)}$$

where the convergence holds almost surely and uniformly on the compact subsets of $\mathbb{C}\backslash\mathbb{R}^+$ and of $\mathbb{C}\backslash\mathbb{R}$ respectively. Moreover, if $(a_N)_{N\geq 1}$, and $(b_N)_{N\geq 1}$ (resp. $(\mathbf{a}_N)_{N\geq 1}$, $(\mathbf{b}_N)_{N\geq 1}$) represent sequences of ML-dimensional (resp. $2ML$-dimensional) deterministic vectors verifying $\sup_N \|a_N\| < +\infty$ and $\sup_N \|b_N\| < +\infty$ (resp. $\sup_N \|\mathbf{a}_N\| < +\infty$ and $\sup_N \|\mathbf{b}_N\| < +\infty$), we also have

$$a_N^* \left(Q_N(z) - I_M \otimes T_N(z)\right) b_N \to 0 \quad \text{(II.11)}$$

and

$$\mathbf{a}_N^* \left(\mathbf{Q}_N(z) - I_{2L} \otimes \mathbf{T}_N(z)\right) \mathbf{b}_N \to 0 \quad \text{(II.12)}$$

almost surely and uniformly on the compact subsets of $\mathbb{C} \setminus \mathbb{R}^+$ and $\mathbb{C} \setminus \mathbb{R}$ respectively.

We denote in the following $\hat{\nu}_N$ the empirical eigenvalue distribution of matrix $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$. The use of (II.9) for $A_N = I$ leads to the conclusion that if ν_N represents the probability measure defined by

$$\nu_N = \frac{1}{M}\mathrm{Tr}\, \nu_N^T \quad \text{(II.13)}$$

then $\hat{\nu}_N - \nu_N \to 0$ weakly almost surely. Therefore, the empirical eigenvalue distribution $\hat{\nu}_N$ of $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ has a deterministic behaviour when $N \to +\infty$, and measure ν_N will be referred to as the deterministic equivalent of $\hat{\nu}_N$ in the following. [33] also characterized the support of ν_N, or equivalently the support of μ_N because Assumption (I.11) implies that ν_N and μ_N are absolutely continuous one with respect to each other. For this, the behaviour of $t_N(z)$ when z converges towards the real axis is studied in [33]. It is shown that for each $x > 0$, the limit of $t_N(z)$ when $z \in \mathbb{C}^+$ converges towards x exists and is finite. This limit is still denoted $t_N(x)$ in the following. This property implies that μ_N and ν_N are absolutely continuous w.r.t. the Lebesgue measure (see e.g. Theorem 2.1 in [43]). Moreover, it is shown that the corresponding densities converge towards $+\infty$ when $x \to 0, x > 0$. In order to analyse the common support $\mathcal{S}_N$ of μ_N and ν_N, the function $w_N(z)$ defined by

$$w_N(z) = z c_N t_N(z) - \frac{1}{c_N t_N(z)} \quad \text{(II.14)}$$

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

$$\phi_N(w) = c_N w^2 \frac{1}{M} \mathrm{Tr} R_N \left(R_N - wI\right)^{-1} \left(c_N \frac{1}{M} \mathrm{Tr} R_N \left(R_N - wI\right)^{-1} - 1\right). \tag{II.15}$$

To understand the equation $\phi_N(w_N(z)) = z$, we remark that $T_N(z)$ can be written in terms of $w_N(z)$ as

$$T_N(z) = \frac{w_N(z)}{z} \left(R_N - w_N(z)I\right)^{-1} \tag{II.16}$$

so that $t_N(z) = \frac{1}{M} \mathrm{Tr}\, R_N T_N(z)$ is equal to

$$t_N(z) = \frac{w_N(z)}{z} \frac{1}{M} \mathrm{Tr} \left(R_N (R_N - w_N(z)I)^{-1}\right) \tag{II.17}$$

Plugging (II.17) into (II.14) leads to $\phi_N(w_N(z)) = z$. Moreover, if we define by $w_N(x)$ for $x > 0$ the limit of $w_N(z)$ when $z \to x, z \in \mathbb{C}^+$, the equality $\phi_N(w_N(z)) = z$ is also valid on $\mathbb{R}^+$. It is proved that $x \in \mathcal{S}_N^\circ$ if and only $\mathrm{Im}(w_N(x)) > 0$ ($\mathcal{S}_N^\circ$ represents the interior of $\mathcal{S}_N$) and that $x \in \left(\mathcal{S}_N^\circ\right)^c$ if and only if $w_N(x)$ is real. Moreover, $w_N'(x) > 0$ for each $x \in \left(\mathcal{S}_N\right)^c$. Finally, if $x \in \left(\mathcal{S}_N\right)^c$, it holds that

$$\phi_N(w_N(x)) = x, \ \phi^{'}(w_N(x)) > 0, \ w_N(x) \frac{1}{M} \mathrm{Tr} R_N \left(R_N - w_N(x)I\right)^{-1} < 0. \tag{II.18}$$

This property allows to prove that the support $\mathcal{S}_N$ of μ_N contains 0, and coincides with the union of intervals whose end points, apart 0, are the extrema of ϕ_N whose arguments verify $\frac{1}{M} \mathrm{Tr} R_N \left(R_N - wI\right)^{-1} < 0$, see Corollary 7.2 in [33]. If we denote by $x_{+,N}$ the largest element of $\mathcal{S}_N$, then, $x_{+,N} = \phi_N(w_{+,N})$ where $w_{+,N} > \lambda_{1,N} = \lambda_1(R_N)$ is the largest solution of $\phi_N'(w) = 0$. It is established that $\sup_{N \geq 1} x_{+,N} < +\infty$ and $\sup_{N \geq 1} w_{+,N} < +\infty$. A sufficient condition on the eigenvalues of R_N ensuring that the support of μ_N is reduced to the single interval $[0, x_{+,N}]$ is formulated (see Lemma 7.7 in [33]). Using the Haagerup-Thorbjornsen approach ([20]), it is finally proved that if $d > 0$ verifies $[d, +\infty) \cap \cup_{N \geq N_0} \mathcal{S}_N = \emptyset$ for some integer N_0, almost surely, for N large enough, all the eigenvalues of $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ are smaller than d. When $\mathcal{S}_N$ is reduced to $[0, x_{+,N}]$, this property implies that for each $\epsilon > 0$, for each N large enough, then all the eigenvalues of $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ are smaller than $\sup_{N \geq N_0} x_{+,N} + \epsilon$ where N_0 is a large enough integer.

B. *Signal model and first assumptions*

Now we pass to the case when signal $(u_n)_{n \in \mathbb{Z}}$ is present, and evaluate its influence on the eigenvalues of matrix $\frac{Y_f Y_p^*}{N} \left(\frac{Y_f Y_p^*}{N}\right)^*$. For this, we use a classical approach based on the observation that matrix $\frac{Y_f Y_p^*}{N}$ is a finite rank perturbation of matrix $\frac{V_f V_p^*}{N}$ due to the noise $(v_n)_{n \in \mathbb{Z}}$. It will be assumed that for each N large enough, the support $\mathcal{S}_N$ of measure μ_N associated to $t_N(z)$ is reduced to the single interval $\mathcal{S}_N = [0, x_{N,+}]$, see Assumption II.5 below.

We recall that the useful signal $(u_n)_{n \in \mathbb{Z}}$ is generated by the minimal state-space representation (I.3). As M is supposed to increase towards $+\infty$, it is first necessary to precise how the parameters of (I.3) depend on M. We formulate the following assumptions:

Assumption II.1.
- *$(i_n)_{n \in \mathbb{Z}}$ is a K–dimensional white noise sequence such that $\mathbb{E}(i_n i_n^*) = I_K$, and which is independent of M and N*
- *The dimension P of the state-space does not scale with M and N and matrices A and B are independent of M and N.*
- *Matrices $C = C_N$ and $D = D_N$ depend of M and thus on N, and are supposed to verify*

$$\sup_N \|C_N\| < +\infty, \sup_N \|D_N\| < +\infty \tag{II.19}$$

We recall that $L \geq P$. As a consequence of Assumption II.1, the P–dimensional Markovian signal $(x_n)_{n\in\mathbb{Z}}$ is independent of M and N. We define matrix $\mathcal{H}_N$ as the $ML \times KL$ block-Toeplitz matrix defined by

$$\mathcal{H}_N = \begin{pmatrix} D_N & 0 & \dots & \dots & 0 \\ C_N B & D_N & 0 & \ddots & 0 \\ \vdots & C_N B & \ddots & \ddots & \vdots \\ C_N A^{L-3} B & \ddots & \ddots & \ddots & \vdots \\ C_N A^{L-2} B & C_N A^{L-3} B & \ddots & C_N B & D_N \end{pmatrix} \tag{II.20}$$

Then, it is easy to check that the ML–dimensional vector $u_n^L = (u_n^T, \dots, u_{n+L-1}^T)^T$ can be written as

$$u_n^L = (\mathcal{O}_N, \mathcal{H}_N) \begin{pmatrix} x_n \\ i_n^L \end{pmatrix} \tag{II.21}$$

where i_n^L is defined as u_n^L and where we recall that the observability matrix $\mathcal{O}_N$ is defined by (I.1). We formulate the following assumption:

Assumption II.2. *The rank $r \leq P + KL$ of matrix $(\mathcal{O}_N, \mathcal{H}_N)$ remains constant for N large enough.*

In the following, we denote by $U_{f,N}$ and $U_{p,N}$ the $ML \times N$ matrices defined as the analogues of $Y_{f,N}$ and $Y_{p,N}$ obtained by replacing the M–dimensional vectors $(y_n)_{n=1,\dots,N+2L-1}$ by the M–dimensional vectors $(u_n)_{n=1,\dots,N+2L-1}$. We also denote by $R_{u,N}^L = \mathbb{E}(u_n^L u_n^{L*})$ the covariance matrix of u_n^L, and recall that $\mathbb{E}(u_{n+L}^L u_n^{L*})$ coincides with $R_{f|p,N}^L = \mathbb{E}(y_{n+L}^L y_n^{L*})$. We also recall that $\mathrm{Rank}(R_{f|p,N}^L) = P$ for each $L \geq P$ and claim that Assumption II.2 implies that for N large enough, $\mathrm{Rank}(R_{u,N}^L) = r$ for each $L \geq P$. This is because $R_{u,N}^L$ is given by

$$R_{u,N}^L = (\mathcal{O}_N, \mathcal{H}_N) \begin{pmatrix} R_x & 0 \\ 0 & I_{KL} \end{pmatrix} (\mathcal{O}_N, \mathcal{H}_N)^* \tag{II.22}$$

where $R_x = \mathbb{E}(x_n x_n^*)$ coincides with

$$R_x = \sum_{k=0}^{\infty} A^k B B^* A^{*k}$$

R_x is positive definite because the minimality of the state-space representation (I.3) of u implies that the pair (A, B) is commandable. Therefore, Assumption II.2 implies that $\mathrm{Rank}(R_{u,n}^L) = r$ for each N large enough. In the following, we denote by

$$R_{u,N}^L = \Theta_N \Delta_N^2 \Theta_N^* \tag{II.23}$$

the eigenvalue / eigenvector decomposition of $R_{u,N}^L$ where $\Delta_N^2 = \mathrm{Diag}(\delta_{1,N}^2, \dots, \delta_{r,N}^2)$ where $(\delta_{k,N}^2)_{k=1,\dots,r}$ are the eigenvalues of $R_{u,N}^L$ arranged in the decreasing order and where Θ_N is the $ML \times r$ orthogonal matrix corresponding to the eigenvectors.

We now take benefit of Assumptions II.1 and II.2 to evaluate the properties of matrices $\frac{U_{i,N} U_{i,N}^*}{N}$ for $i = p, f$ and $\frac{U_{f,N} U_{p,N}^*}{N}$.

Proposition II.2. *The following convergence result hold:*

$$\left\| \frac{U_{i,N} U_{i,N}^*}{N} - R_{u,N}^L \right\| \to 0 \tag{II.24}$$

$$\left\| \frac{U_{f,N} U_{p,N}^*}{N} - R_{f|p,N}^L \right\| \to 0 \tag{II.25}$$

for $i = p, f$.

Proof. In the following, we denote by $X_{1,N}$ and $X_{L+1,N}$ the $P \times N$ matrices defined by

$$X_{1,N} = (x_1, x_2, \dots, x_N), \; X_{L+1,N} = (x_{L+1}, x_{L+2}, \dots, x_{N+L}) \tag{II.26}$$

and by $I_{f,N}$ and $I_{p,N}$ the $KL \times N$ matrices defined as the analogues of $Y_{f,N}$ and $Y_{p,N}$ obtained by replacing the M–dimensional vectors $(y_n)_{n=1,\ldots,N+2L-1}$ by the K–dimensional vectors $(i_n)_{n=1,\ldots,N+2L-1}$. It is easy to check that

$$U_{p,N} = \mathcal{O}_N \, X_1 + \mathcal{H}_N \, I_{p,N}, \; U_{f,N} = \mathcal{O}_N \, X_{L+1,N} + \mathcal{H}_N \, I_{f,N} \tag{II.27}$$

As P, K, L remain fixed, matrix

$$\frac{1}{N} \begin{pmatrix} X_{1,N} \\ I_{p,N} \end{pmatrix} \begin{pmatrix} X_{1,N}^* & I_{p,N}^* \end{pmatrix}$$

converges almost surely towards the covariance matrix of vector $\begin{pmatrix} x_n \\ i_n^L \end{pmatrix}$, i.e. matrix

$$\begin{pmatrix} R_x & 0 \\ 0 & I_{KL} \end{pmatrix}$$

As the rank of this matrix is obviously $P + KL$, the same property holds for $\begin{pmatrix} X_{1,N} \\ I_{p,N} \end{pmatrix}$ for N large enough. Moreover, (II.19) implies that

$$\sup_N \|(\mathcal{O}_N, \mathcal{H}_N)\| < +\infty \tag{II.28}$$

Using the equation

$$\frac{U_{p,N} U_{p,N}^*}{N} = (\mathcal{O}_N, \mathcal{H}_N) \, \frac{1}{N} \begin{pmatrix} X_{1,N} \\ I_{p,N} \end{pmatrix} \begin{pmatrix} X_{1,N}^* & I_{p,N}^* \end{pmatrix} (\mathcal{O}_N, \mathcal{H}_N)^* ,$$

(II.22) and (II.28) imply that

$$\|R_{u,N}^L - \frac{U_{p,N} U_{p,N}^*}{N}\| \to 0 \tag{II.29}$$

It holds similarly that

$$\|R_{u,N}^L - \frac{U_{f,N} U_{f,N}^*}{N}\| \to 0 \tag{II.30}$$

Moreover, the column space of matrices $U_{p,N}$ and $U_{f,N}$ both coincide with the r–dimensional column space of $(\mathcal{O}_N, \mathcal{H}_N)$ for N large enough. Therefore, $\mathrm{Rank}\left(\frac{U_{i,N} U_{i,N}^*}{N}\right) = r$ almost surely for N large enough. We also remark that

$$\frac{1}{N} \begin{pmatrix} X_{L+1,N} \\ I_{f,N} \end{pmatrix} \left(X_{1,N}^* I_{p,N}^*\right) \to \mathbb{E}\left[\begin{pmatrix} x_{n+L} \\ i_{n+L}^L \end{pmatrix} \left(x_n^*, i_n^{L*}\right)\right] = \begin{bmatrix} \mathbb{E}\left(x_{n+L}(x_n^*, i_n^{L*})\right) \\ 0 \end{bmatrix}$$

Therefore, using (II.28), we obtain that

$$\left\| \frac{1}{N} (\mathcal{O}_N, \mathcal{H}_N) \begin{pmatrix} X_{L+1,N} \\ I_{f,N} \end{pmatrix} \left(X_{1,N}^* I_{p,N}^*\right) \begin{pmatrix} \mathcal{O}_N^* \\ \mathcal{H}_N^* \end{pmatrix} - (\mathcal{O}_N, \mathcal{H}_N) \begin{pmatrix} \mathbb{E}\left(x_{n+L} \, u_n^{L*}\right) \\ 0 \end{pmatrix} \right\| \to 0 \tag{II.31}$$

because (II.21) holds. It is easily seen that matrix $\mathbb{E}\left(x_{n+L} u_n^{L*}\right)$ coincides with $\mathcal{C}_N = (A^{L-1}G, \ldots, G)$ (we recall that $G = \mathbb{E}(x_{n+1} u_n^*)$, see Paragraph I-A). Moroever, as $R_{f|p,N}^L = \mathbb{E}(u_{n+L}^L u_n^{L*})$ is equal to $\mathcal{O}_N \mathcal{C}_N$, we obtain that

$$(\mathcal{O}_N, \mathcal{H}_N) \begin{pmatrix} \mathbb{E}\left(x_{n+L} \, u_n^{L*}\right) \\ 0 \end{pmatrix} = \mathcal{O}_N \mathcal{C}_N = R_{f|p,N}^L$$

Therefore, (II.31) implies that (II.25) holds. ■

We introduce the singular value decompositions of matrices $\frac{U_{p,N}}{\sqrt{N}}$ and $\frac{U_{f,N}}{\sqrt{N}}$:

$$\frac{U_{p,N}}{\sqrt{N}} = \Theta_{p,N} \, \Delta_{p,N} \tilde{\Theta}_{p,N}^*, \; \frac{U_{f,N}}{\sqrt{N}} = \Theta_{f,N} \, \Delta_{f,N} \tilde{\Theta}_{f,N}^* \tag{II.32}$$

where $\Theta_{i,N}, \Delta_{i,N}, \tilde{\Theta}_{i,N}$ are $ML \times r$, $r \times r$, $N \times r$ matrices that of course depend on N for $i = p, f$. We deduce from (II.25) that $\mathrm{Rank}\left(\frac{U_{f,N} U_{p,N}^*}{N}\right) = P$ for each N large enough. As $\frac{U_{f,N} U_{p,N}^*}{N}$ coincides with $\Theta_{f,N} \Delta_{f,N} \tilde{\Theta}_{f,N}^* \tilde{\Theta}_{p,N} \Delta_{p,N} \Theta_{p,N}^*$, we obtain that $\mathrm{Rank}\left(\Delta_{f,N} \tilde{\Theta}_{f,N}^* \tilde{\Theta}_{p,N} \Delta_{p,N}\right) = P$, $\mathrm{Rank}\left(\Delta_N \tilde{\Theta}_{f,N}^* \tilde{\Theta}_{p,N} \Delta_N\right) = P$ and that $\mathrm{Rank}\left(\tilde{\Theta}_{f,N}^* \tilde{\Theta}_{p,N}\right) = P$ for each N large enough. As in the previous works devoted to the study of conventional spiked models (see e.g. [7], [10]), it is necessary to introduce assumptions concerning the existence of limits of certain terms depending on the statistics of the useful signal u. In particular, we will need the following assumption.

Assumption II.3. *$r \times r$ matrices Δ_N and $\Theta_N^* R_{f|p,N}^L \Theta_N$ converge towards matrices Δ_* and Γ_* respectively. It is moreover assumed that $\Delta_* > 0$.*

We notice that $\mathrm{Rank}(\Gamma_*) = P$. As seen below, the proofs of the main results of this paper appear simpler when we assume the following condition

$$\delta_{1,*} > \ldots > \delta_{r,*} \tag{II.33}$$

where $(\delta_{k,*})_{k=1,\ldots,r}$ represent the diagonal entries of Δ_*. Therefore, in the following, we will assume that condition (II.33) holds, and discuss briefly in Sections II-F and III-C below how the results can be extended to the case where some of the diagonal entries of $\Delta_* > 0$ coincide. In order to explain why condition (II.33) allows to simplify the following arguments, we establish the following result.

Proposition II.3. *For $i = p, f$, matrices $(\Delta_{i,N})_{N \geq 1}$ verify*

$$\|\Delta_{i,N} - \Delta_N\| \to 0 \ a.s. \tag{II.34}$$

Moreover, if condition (II.33) holds, and if the r left singular vectors $(\theta_{i,N,k})_{k=1,\ldots,r}$ of $\frac{U_{i,N}}{\sqrt{N}}$ are chosen in such a way that $\theta_{N,k}^* \theta_{i,N,k}$ is real and positive, then we have

$$\|\Theta_{i,N} - \Theta_N\| \to 0 \ a.s. \tag{II.35}$$

Proof. (II.34) is a consequence of (II.24) and of the Weyl inequalities which imply that $|\delta_{i,N,k}^2 - \delta_{N,k}^2| \leq \|\frac{U_{i,N} U_{i,N}^*}{N} - R_{u,N}^L\|$. We thus notice that (II.34) holds even when Assumption (II.3) is not verified. In order to verify (II.35), we first remark that (II.24), Assumption (II.3) and (II.34) imply that

$$\|\Theta_{i,N} \Delta_*^2 \Theta_{i,N}^* - \Theta_N \Delta_*^2 \Theta_N^*\| \to 0 \tag{II.36}$$

when $N \to +\infty$. Condition (II.33) implies that the eigenvalues of matrices $\Theta_{i,N} \Delta_*^2 \Theta_{i,N}^*$ and $\Theta_N \Delta_*^2 \Theta_N^*$ have multiplicity 1. Therefore, standard results of perturbation theory of Hermitian matrices lead to the conclusion that

$$\|\theta_{i,N,k} \theta_{i,N,k}^* - \theta_{N,k} \theta_{N,k}^*\| \to 0$$

for $k = 1, \ldots, r$. This implies that $\|\theta_{i,N,k} - (\theta_{N,k}^* \theta_{i,N,k})\, \theta_{N,k}\| \to 0$ as well as $|\theta_{N,k}^* \theta_{i,N,k}|^2 \to 1$. The condition $(\theta_{N,k}^* \theta_{i,N,k})$ real positive leads to $(\theta_{N,k}^* \theta_{i,N,k}) \to 1$ and to

$$\|\theta_{i,N,k} - \theta_{N,k}\| \to 0$$

for each $k = 1, \ldots, r$. This completes the proof of (II.35). ■

Condition (II.33) allows to replace matrices $\Delta_{i,N}$ and $\Theta_{i,N}$ for $i = p, f$ by matrices Δ_N and Θ_N up to error terms that converge towards 0. In particular, $\frac{U_{f,N} U_{p,N}^*}{N} = \Theta_{f,N} \Delta_{f,N} \tilde{\Theta}_{f,N}^* \tilde{\Theta}_{p,N} \Delta_{p,N} \Theta_{p,N}^*$ verifies $\|\frac{U_{f,N} U_{p,N}^*}{N} - \Theta_N \Delta_N \tilde{\Theta}_{f,N}^* \tilde{\Theta}_{p,N} \Delta_N \Theta_N^*\| \to 0$. We introduce the rank P matrix Γ_N given by

$$\Gamma_N = \Delta_N \tilde{\Theta}_{f,N}^* \tilde{\Theta}_{p,N} \Delta_N \tag{II.37}$$

Then, (II.25) implies that

$$\|R_{f|p,N}^L - \Theta_N \Gamma_N \Theta_N^*\| \to 0 \tag{II.38}$$

and that, under condition (II.33),

$$\lim_{N \to +\infty} \Gamma_N = \Gamma_* \tag{II.39}$$

We notice that if some of the entries of Δ_* coincide, then (II.39) does no longer hold. This point will be explained in Section II-F. If we consider the singular value decomposition

$$\Gamma_N = \Upsilon_N \Xi_N \tilde{\Upsilon}_N^* \tag{II.40}$$

of matrix Γ_N, then, (II.38) implies that the P non zero singular values of $R_{f|p,N}^L$ have the same asymptotic behaviour than the P non zero singular values $(\chi_{k,N})_{k=1,\ldots,P}$ of Γ_N, and converge towards the singular values of matrix Γ_*.

We finally notice that the canonical correlation coefficients between the row spaces of $U_{p,N}$ and $U_{f,N}$, i.e. the singular values of matrix $\tilde{\Theta}_{f,N}^* \tilde{\Theta}_{p,N}$, and the canonical correlation coefficients between the spaces $\mathcal{U}_{p,L}$ and $\mathcal{U}_{f,L}$ generated by the components of $(u_{n+L})_{n=0,\ldots,L-1}$ and $(u_{n+L})_{n=L,\ldots,2L-1}$, i.e. the singular values of matrix $\Delta_N^{-1} \Theta_N^* R_{f|p,N}^L \Theta_N \Delta_N^{-1}$, have the same asymptotic behaviour. For this, we just use (II.38), (II.39) as well as the convergence of Δ_N towards $\Delta_* > 0$, and obtain that

$$\|\Delta_N^{-1} \Theta_N^* R_{f|p,N}^L \Theta_N \Delta_N^{-1} - \tilde{\Theta}_{f,N}^* \tilde{\Theta}_{p,N}\| \rightarrow 0 \tag{II.41}$$

C. General approach

We first briefly explain the general approach that will be used in the following to evaluate the behaviour of the eigenvalues of $\frac{Y_{f,N} Y_{p,N}^*}{N} \frac{Y_{p,N} Y_{f,N}^*}{N}$. In order to simplify the notations, we denote by $\Sigma_{i,N}$ and $W_{i,N}$ matrices $\Sigma_{i,N} = \frac{Y_{i,N}}{\sqrt{N}}$ and $W_{i,N} = \frac{V_{i,N}}{\sqrt{N}}$ for $i = p, f$. It is easy to check that

$$\Sigma_f \Sigma_p^* = W_f W_p^* + (\Theta_f, W_f \tilde{\Theta}_p \Delta_p) \begin{pmatrix} \Delta_f \tilde{\Theta}_f^* \tilde{\Theta}_p \Delta_p & I_r \\ I_r & 0 \end{pmatrix} \begin{pmatrix} \Theta_p^* \\ \Delta_f \tilde{\Theta}_f^* W_p^* \end{pmatrix} \tag{II.42}$$

We denote by $\mathcal{A}$ and $\mathcal{B}$ the matrices defined by

$$\mathcal{A} = \left(\Theta_f, W_f \tilde{\Theta}_p \Delta_p\right) \tag{II.43}$$

and

$$\mathcal{B} = \left(\Theta_p, W_p \tilde{\Theta}_f \Delta_f\right) \begin{pmatrix} \Delta_p \tilde{\Theta}_p^* \tilde{\Theta}_f \Delta_f & I_r \\ I_r & 0 \end{pmatrix} \tag{II.44}$$

Then, an easy calculation leads to

$$\begin{pmatrix} -z\, I & \Sigma_f \Sigma_p^* \\ \Sigma_p \Sigma_f^* & -z\, I \end{pmatrix} = \begin{pmatrix} -z\, I & W_f W_p^* \\ W_p W_f^* & -z\, I \end{pmatrix} + \begin{pmatrix} \mathcal{A} & 0 \\ 0 & \mathcal{B} \end{pmatrix} \begin{pmatrix} 0 & I_{2r} \\ I_{2r} & 0 \end{pmatrix} \begin{pmatrix} \mathcal{A}^* & 0 \\ 0 & \mathcal{B}^* \end{pmatrix} \tag{II.45}$$

We recall that $\mathbf{Q}_W(z)$ represents the resolvent of matrix $\begin{pmatrix} 0 & W_f W_p^* \\ W_p W_f^* & 0 \end{pmatrix}$. Consider a positive real number y such that y is not eigenvalue of $\begin{pmatrix} 0 & W_f W_p^* \\ W_p W_f^* & 0 \end{pmatrix}$ for each N large enough (some conditions on such an eigenvalue will be precised below). For $z = y$, the left handside of (II.45) can also be written as

$$\begin{pmatrix} -y\, I & \Sigma_f \Sigma_p^* \\ \Sigma_p \Sigma_f^* & -y\, I \end{pmatrix} = \begin{pmatrix} -y\, I & W_f W_p^* \\ W_p W_f^* & -y\, I \end{pmatrix} \left(I_{2ML} + \mathbf{Q}_W(y) \begin{pmatrix} \mathcal{A} & 0 \\ 0 & \mathcal{B} \end{pmatrix} \begin{pmatrix} 0 & I_{2r} \\ I_{2r} & 0 \end{pmatrix} \begin{pmatrix} \mathcal{A}^* & 0 \\ 0 & \mathcal{B}^* \end{pmatrix}\right) \tag{II.46}$$

Therefore, y is eigenvalue of $\begin{pmatrix} 0 & \Sigma_f \Sigma_p^* \\ \Sigma_p \Sigma_f^* & 0 \end{pmatrix}$ if and only the determinant of the second term of the right handside of (II.46) vanishes. Using the identity $\det(I + EF) = \det(I + FE)$, we obtain that y is an eigenvalue of $\begin{pmatrix} 0 & \Sigma_f \Sigma_p^* \\ \Sigma_p \Sigma_f^* & 0 \end{pmatrix}$ if and only

$$\det\left(I_{4r} + \begin{pmatrix} \mathcal{A}^* & 0 \\ 0 & \mathcal{B}^* \end{pmatrix} \mathbf{Q}_W(y) \begin{pmatrix} \mathcal{A} & 0 \\ 0 & \mathcal{B} \end{pmatrix} \begin{pmatrix} 0 & I_{2r} \\ I_{2r} & 0 \end{pmatrix}\right) = 0 \tag{II.47}$$

or equivalently if

$$\det\left(I_{4r} + F_N(y)\right) = 0 \tag{II.48}$$

where $F_N(z)$ is the $4r \times 4r$ matrix-valued function given by

$$F_N(z) = \begin{pmatrix} \mathcal{A}^* \mathbf{Q}_{W,pf}(z) \mathcal{B} & \mathcal{A}^* \mathbf{Q}_{W,pp}(z) \mathcal{A} \\ \mathcal{B}^* \mathbf{Q}_{W,ff}(z) \mathcal{B} & \mathcal{B}^* \mathbf{Q}_{W,fp}(z) \mathcal{A} \end{pmatrix} \tag{II.49}$$

We will see that under certain technical assumptions, $F_N(y)$ converges towards a deterministic matrix $F_*(y)$ and that the solutions of (II.48) converge towards the solutions of the deterministic equation $\det\left(I_{4r} + F_*(y)\right) = 0$, which, fortunately, can be analyzed.

D. New assumptions and their consequences.

We need to distinguish two kinds of extra-assumptions.

- Assumptions on the asymptotic behaviour of the eigenvalue distribution of matrix R_N.

Assumption II.4. *If $\omega_N = \frac{1}{M}\sum_{k=1}^{M}\delta_{\lambda_{k,N}}$ is the eigenvalue distribution of matrix R_N, it is assumed that*

$$\lim_{N\to+\infty}\lambda_{1,N} = \lambda_{+,*} \quad \lim_{N\to+\infty}\lambda_{M,N} = \lambda_{-,*} \tag{II.50}$$

We note that $\lambda_{-,*} \geq a > 0$ and $\lambda_{+,*} \leq b$ where a and b are defined by (I.11). Moreover, sequence $(\omega_N)_{N\geq 1}$ is assumed to converge weakly towards a probability measure ω_*, which, necessarily, is carried by $[\lambda_{-,*}, \lambda_{+,*}]$

Assumption II.5. *It is assumed that for each N large enough, it exists a nice constant $\kappa > 0$ such that the eigenvalues $(\lambda_{k,N})_{k=1,\ldots,M}$ satisfy*

$$|\lambda_{k,N} - \lambda_{l,N}| \leq \kappa\left(\frac{|k-l|}{M}\right)^{1/2} \tag{II.51}$$

for each pair (k,l), $1 \leq k \leq l \leq M$, so that the support $\mathcal{S}_N$ of μ_N is equal to $\mathcal{S}_N = [0, x_{+,N}]$ (see Lemma 7.7 in [33]). Moreover, we add the following condition: for each N large enough,

$$\lambda_{1,N} - \lambda_{k,N} \leq \kappa\,\frac{k-1}{M} \tag{II.52}$$

for some nice constant κ.

- Assumptions on the asymptotic behaviour of matrices depending both of the useful signal and the noise.

Assumption II.6. *We recall that $(f_{k,N})_{k=1,\ldots,M}$ represent the eigenvectors of matrix R_N. We consider the $M\times M$ matrix-valued function positive measure ω_N^R defined by*

$$\omega_N^R = \sum_{k=1}^{M}\delta_{\lambda_{k,N}}\, f_{k,N}f_{k,N}^*$$

and introduce the $r\times r$ positive matrix-valued measure γ_N defined by

$$d\gamma_N(\lambda) = \Theta_N^*\left(I_L \otimes d\omega_N^R(\lambda)\right)\Theta_N \tag{II.53}$$

Then it is assumed that the sequence $(\gamma_N)_{N\geq 1}$ converges weakly towards a certain measure γ_*.

It is clear that Assumptions II.4, II.5, II.6 look rather strong (notice however that the assumptions are satisfied when $R_N = \sigma^2 I$ for some $\sigma^2 > 0$). This does not limit the usefulness of the results of Section II because our goal is to establish that, despite the above strong Assumptions, the number of largest eigenvalues of $\hat{R}_{f|p,y}\hat{R}_{f|p,y}^* = \Sigma_f\Sigma_p^*\Sigma_p\Sigma_f^*$ that escape from $[0, x_{+,N}]$ is not at all related to P. Therefore, the conclusion of the results of Section II is that, even if strong Assumptions hold, the largest eigenvalues of $\hat{R}_{f|p,y}\hat{R}_{f|p,y}^*$ cannot be used to estimate P consistently.

We now state some consequences of Assumption II.4 and Assumption II.5, which, in some sense, show that $x_{+,N}, w_{+,N}$, functions $t_N(z), w_N(z)$ and measure μ_N have, when $N\to+\infty$, limits that satisfy the same properties that their finite N equivalents. We recall that $w_{+,N} > 0$ is defined by $w_{+,N} = w_N(x_{+,N})$ and verifies $x_{+,N} = \phi_N(w_{+,N})$, $\phi_N'(w_{+,N}) = 0$ and $w_{+,N} > \lambda_{1,N}$ (we recall that w_N and ϕ_N are defined by (II.14) and (II.15)). We omit the proof of the two following Propositions, and refer the reader to the proofs of Proposition 4.1 and Proposition 4.2 in the Thesis [45].

Proposition II.4. *Sequences $(w_{+,N})_{N\geq 1}$ and $(x_{+,N})_{N\geq 1}$ converge towards finite limits $w_{+,*}$ and $x_{+,*}$ respectively. Moreover, $w_{+,*}$ verifies $w_{+,*} > \lambda_{+,*}$. If $\phi_*(w)$ is the function defined on $\mathbb{C} - [\lambda_{-,*}, \lambda_{+,*}]$ by*

$$\phi_*(w) = (c_* w)^2\left(\int_{\lambda_{-,*}}^{\lambda_{+,*}}\frac{\lambda\, d\omega_*(\lambda)}{w-\lambda}\right)^2 + c_* w^2\int_{\lambda_{-,*}}^{\lambda_{+,*}}\frac{\lambda\, d\omega_*(\lambda)}{w-\lambda} \tag{II.54}$$

then, $\phi_N(w) \to \phi_*(w)$ uniformly on the compact subsets of $\mathbb{C} - [\lambda_{-,*}, \lambda_{+,*}]$. Moreover, it holds that

$$x_{+,*} = \phi_*(w_{+,*}) \tag{II.55}$$

The sequence $(\mu_N)_{N\geq 1}$ converges weakly towards a probability measure μ_*. The support $\mathcal{S}_*$ of μ_* is included into $[0, x_{+,*}]$, and the Stieltjes transform $t_*(z)$ of μ_* verifies the equation

$$t_*(z) = \int_{\lambda_{-,*}}^{\lambda_{+,*}} \frac{\lambda}{-z(1+\frac{c_* t_*(z)\,\lambda}{1-z(c_* t_*(z))^2})} \, d\omega_*(\lambda) \tag{II.56}$$

for each $z \in \mathbb{C} - [0, x_{+,*}]$. Moreover, $t_N(z)$ converges uniformly towards $t_*(z)$ on the compact subsets of $\mathbb{C} - [0, x_{+,*}]$. If $w_*(z)$ is the function defined on $\mathbb{C} - [0, x_{+,*}]$ by

$$w_*(z) = c_* z t_*(z) - \frac{1}{c_* t_*(z)} \tag{II.57}$$

then, w_* is holomorphic on $\mathbb{C} - [0, x_{+,*}]$ and $w_N(z)$ converges uniformly towards $w_*(z)$ on the compact subsets of $\mathbb{C} - [0, x_{+,*}]$. $w_*(z)$ satisfies

$$\phi_*(w_*(z)) = z \tag{II.58}$$

for each $z \in \mathbb{C} - [0, x_{+,*}]$. Finally,

$$\lim_{x\to x_{+,*}, x > x_{+,*}} t_*(x) \text{ exists, is finite, is still denoted } t_*(x_{+,*}), \text{ and Eq. (II.56) holds for } z = x_{+,*} \tag{II.59}$$

Moreover, we have

$$w_{+,*} = w_*(x_{+,*}) \tag{II.60}$$

We recall that ν_N^T is the $M \times M$ matrix-valued positive measure associated to matrix-valued Stieltjes transform $T_N(z)$, and introduce for each N the $r \times r$ matrix-valued measure β_N defined by

$$d\beta_N(\lambda) = \Theta_N^* \left(I_L \otimes d\nu_N^T(\lambda) \right) \Theta_N \tag{II.61}$$

We notice that $\nu_N^T(\mathbb{R}^+) = I$ implies that $\beta_N(\mathbb{R}^+) = I$. Using the identity (II.16), we obtain immediately that the Stieltjes transform $T_{\beta_N}(z)$ of β_N is given by

$$T_{\beta_N}(z) = \frac{w_N(z)}{z} \int \frac{d\gamma_N(\lambda)}{\lambda - w_N(z)} \tag{II.62}$$

Then, the following result is a consequence of Assumption II.6.

Proposition II.5. *The sequence of measures $(\beta_N)_{N\geq 1}$ converges weakly towards a measure β_* whose support is included into $[0, x_{+,*}]$, and which verifies $\beta_*([0, x_{+,*}]) = \beta_*(\mathbb{R}^+) = I$. The Stieltjes transform $T_{\beta_*}(z)$ of β_* is given by*

$$T_{\beta_*}(z) = \frac{w_*(z)}{z} \int_{\lambda_{-,*}}^{\lambda_{+,*}} \frac{d\gamma_*(\lambda)}{\lambda - w_*(z)} \tag{II.63}$$

for each $z \in \mathbb{C} - [0, x_{+,*}]$. Moreover, it holds that

$$T_{\beta_*}(x_{+,*}) = \lim_{x\to x_{+,*}, x > x_{+,*}} T_{\beta_*}(x) = \lim_{N\to+\infty} T_{\beta_N}(x_{+,N}) = \frac{w_{+,*}}{x_{+,*}} \int_{\lambda_{-,*}}^{\lambda_{+,*}} \frac{d\gamma_*(\lambda)}{\lambda - w_{+,*}} \tag{II.64}$$

We finally conclude this paragraph by the following result.

Proposition II.6. *Assume that $y > \sqrt{x_{+,*}}$. Then, for each N large enough, y is not eigenvalue of matrix $\begin{pmatrix} 0 & W_{f,N} W_{p,N}^* \\ W_{p,N} W_{f,N}^* & 0 \end{pmatrix}$, and y^2 is not eigenvalue of $W_f W_p^* W_p W_f^*$.*

Proof. As $y > \sqrt{x_{+,*}}$ and that $\lim_{N\to+\infty} x_{+,N} = x_{+,*}$, it exists N_0 such that $y > \sqrt{x_{+,N}}$ and $y^2 > x_{+,N}$ for each $N \geq N_0$. Therefore, y^2 does not belong to $\cup_{N\geq N_0} \mathcal{S}_N$. Theorem 8.1 in [33] thus implies that y^2 and y cannot be one of the eigenvalues of matrices $W_{f,N} W_{p,N}^* W_{p,N} W_{f,N}^*$ and $\begin{pmatrix} 0 & W_{f,N} W_{p,N}^* \\ W_{p,N} W_{f,N}^* & 0 \end{pmatrix}$ for $N \geq N_0$ respectively. ■

E. Asymptotic behaviour of the eigenvalues of $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$.

In this paragraph, we characterize the possible eigenvalues of $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$ that escape from the interval $[0, x_{+,*}]$. For this, for each $\delta > 0$ small enough, we study the positive eigenvalues of $\begin{pmatrix} 0 & \Sigma_{f,N}\Sigma_{p,N}^* \\ \Sigma_{p,N}\Sigma_{f,N}^* & 0 \end{pmatrix}$ that are almost surely, for N large enough, strictly greater than $\sqrt{x_{+,*}+\delta}$. We first mention that Theorem 8.1 in [33] implies that the resolvent $Q_W(z)$ of $W_{f,N}W_{p,N}^*W_{p,N}W_{f,N}^*$ and the resolvent $\mathbf{Q}_W(z)$ of matrix $\begin{pmatrix} 0 & W_{f,N}W_{p,N}^* \\ W_{p,N}W_{f,N}^* & 0 \end{pmatrix}$ are almost surely, for each N large enough, holomorphic in $\mathbb{C} - [0, x_{+,N}]$ and in $\mathbb{C} - [-\sqrt{x_{+,N}}, \sqrt{x_{+,N}}]$ respectively. Therefore, almost surely, for each N large enough, function $F_N(z)$ defined by (II.49) is holomorphic on $\mathbb{C} - [-\sqrt{x_{+,N}}, \sqrt{x_{+,N}}]$. As $\lim_{N\to+\infty} x_{+,N} = x_{+,*}$, $F_N(z)$ is also holomorphic on $\mathbb{C} - [-\sqrt{x_{+,*}+\delta}, \sqrt{x_{+,*}+\delta}]$ for each $\delta > 0$ for N large enough.

We first establish that the sequence of analytic functions $(F_N(z))_{N\geq 1}$ almost surely converges uniformly on each compact subset of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$ towards a deterministic function $F_*(z)$ which is analtyic in $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$. Adapting the stability results of the zeros of certain analytic functions proved in [6] and [10], we obtain that for δ small enough, the solutions of the equation $\det(I + F_N(y)) = 0$, $y > \sqrt{x_{+,*}+\delta}$, converge towards the solutions of the limit equation $\det(I + F_*(y)) = 0$, $y > \sqrt{x_{+,*}}$.

In order to study the asymptotic behaviour of F_N, we first consider the asymptotic behaviour of matrix $\mathcal{A}^*\mathbf{Q}_{W,N}(pf)\mathcal{B}$, which is given by

$$\mathcal{A}^*\mathbf{Q}_{W,N}(pf)\mathcal{B} = \begin{pmatrix} \Theta_f^* \\ \Delta_p\tilde{\Theta}_p^* W_f^* \end{pmatrix} \mathbf{Q}_{W,N}(pf) \left(\Theta_p, W_p\tilde{\Theta}_f\Delta_f\right) \begin{pmatrix} \Delta_p\tilde{\Theta}_p^*\tilde{\Theta}_f\Delta_f & I_r \\ I_r & 0 \end{pmatrix}$$

In order to study matrix $\mathcal{A}^*\mathbf{Q}_{W,N}(pf)\mathcal{B}$ when $N \to +\infty$, it is necessary to evaluate the asymptotic behaviour of sesquilinear forms of matrices $\mathbf{Q}_{W,N}(pf)$, $W_f^*\mathbf{Q}_{W,N}(pf)$, $\mathbf{Q}_{W,N}(pf)W_p$ and $W_f^*\mathbf{Q}_{W,N}(pf)W_p$. The following result holds.

Lemma II.1. *For each $z \in \mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$ and for each bounded sequences $(a_N, b_N)_{N\geq 1}$ and $(\tilde{a}_N, \tilde{b}_N)$ of ML–dimensional and N–dimensional deterministic vectors, it holds that*

- $a_N^* \,\mathbf{Q}_{W,N}(pf)\, b_N \to 0$ *almost surely*
- $\tilde{a}_N^* \,W_f^* \,\mathbf{Q}_{W,N}(pf)\, b_N \to 0$ *almost surely*
- $a_N^* \,\mathbf{Q}_{W,N}(pf)\, W_p\, \tilde{b}_N \to 0$ *almost surely*
- $\tilde{a}_N^* \,W_f^* \,\mathbf{Q}_{W,N}(pf)\, W_p\, \tilde{b}_N + \frac{(c_N \mathbf{t}_N(z))^2}{1-(c_N \mathbf{t}_N(z))^2}\, \tilde{a}_N^*\tilde{b}_N \to 0$ *almost surely.*

Moreover, the convergence is uniform over each compact subset of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$ and it holds that, almost surely

$$\mathcal{A}^* \,\mathbf{Q}_{W,N}(pf)\, \mathcal{B} - \begin{pmatrix} 0 & 0 \\ -\frac{(c_N \mathbf{t}_N(z))^2}{1-(c_N \mathbf{t}_N(z))^2}\Gamma_N^* & 0 \end{pmatrix} \to 0 \tag{II.65}$$

the convergence being uniform on compact subsets of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$. Finally, the above properties hold if $a_N, b_N, \tilde{a}_N, \tilde{b}_N$ are random bounded vectors that are independent from the noise sequence $(v_n)_{n\geq 1}$, i.e. from the entries of matrices $(W_N)_{N\geq 1}$.

Sketch of proof. The proof of this result uses ingredients that are very similar to the calculations of Section 5 and Paragraph 6.2 in [33]. We therefore only provide a sketch of proof. When $z \in \mathbb{C}^+$, the first item follows from (II.10) and from the observation that $(I_{2L} \otimes T_N(z))\,(pf) = 0$. The convergence for each $z \in \mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$ follows from the observation that almost surely, for each $\delta > 0$, functions $(a_N^* \,\mathbf{Q}_{W,N}(pf)\, b_N)$ are analytic on $\mathbb{C} - [-\sqrt{x_{+,*}+\delta}, \sqrt{x_{+,*}+\delta}]$ for N large enough. The use of Montel's theorem allows to prove the almost sure convergence for each $z \in \mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$, as well as the uniformity of the convergence on each compact subset of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$. To establish the second and the third item of Lemma II.1 when $z \in \mathbb{C}^+$, we first show that $\mathbb{E}(W_f^* \,\mathbf{Q}_{W,pf}) = 0$ and $\mathbb{E}(\mathbf{Q}_{W,pf}\, W_p) = 0$ using the invariance of the distribution of $(v_n)_{n\in\mathbb{Z}}$ under the transformation $v_n \to e^{in\theta}v_n$ for each θ, and use the Poincaré-Nash inequality. We finally prove the uniform convergence on compact subsets of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$ using Montel's theorem. We note that the sequences of functions defined in item (ii) and (iii) are almost surely bounded on each compact subsets of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$

because matrices W_f and W_p are almost surely bounded, see (II.3).

We denote by $\alpha_N(z)$ and $\boldsymbol{\alpha}_N(z)$ the functions defined by

$$\alpha_N(z) = \mathbb{E}\left(\frac{1}{ML}\mathrm{Tr}\ [(I_L \otimes R_N) Q_{W,N}(z)]\right) \tag{II.66}$$

and

$$\boldsymbol{\alpha}_N(z) = \mathbb{E}\left(\frac{1}{ML}\mathrm{Tr}\ [(I_L \otimes R_N) \mathbf{Q}_{W,N}(pp)(z)]\right) \tag{II.67}$$

We notice that $\boldsymbol{\alpha}_N(z) = z\alpha_N(z^2)$. The proof of the fourth item of Lemma II.1 needs to use the Gaussian calculations of Section 5 in [33] to establish that

$$\tilde{a}_N^* \, W_f^* \, \mathbf{Q}_{W,N}(pf) \, W_p \, \tilde{b}_N + \frac{(c_N \boldsymbol{\alpha}_N(z))^2}{1 - (c_N \boldsymbol{\alpha}_N(z))^2} \, \tilde{a}_N^* \tilde{b}_N \to 0 \ a.s.$$

for each $z \in \mathbb{C}^+$. It is proved in Paragraph 5.2 in [33] that $\alpha_N(z) - t_N(z) \to 0$ for each $z \in \mathbb{C}^+$. As $\boldsymbol{\alpha}_N(z) = z\alpha_N(z^2)$ and $\mathbf{t}_N(z) = z t_N(z^2)$, this implies that $\boldsymbol{\alpha}_N(z) - \mathbf{t}_N(z) \to 0$ if $\mathrm{Arg}(z) \in]0, \pi/2[$. This convergence domain can be extended to $\mathbb{C}^+$ using classical arguments based Montel's theorem. From this, we deduce immediately that

$$\frac{(c_N \boldsymbol{\alpha}_N(z))^2}{1 - (c_N \boldsymbol{\alpha}_N(z))^2} - \frac{(c_N \mathbf{t}_N(z))^2}{1 - (c_N \mathbf{t}_N(z))^2} \to 0$$

for each $z \in \mathbb{C}^+$, and that, for each $z \in \mathbb{C}^+$,

$$\tilde{a}_N^* \, W_f^* \, \mathbf{Q}_{W,N}(pf) \, W_p \, \tilde{b}_N + \frac{(c_N \mathbf{t}_N(z))^2}{1 - (c_N \mathbf{t}_N(z))^2} \, \tilde{a}_N^* \tilde{b}_N \to 0, \ a.s. \tag{II.68}$$

Matrices W_f and W_p are almost surely bounded. Therefore, for each $\delta > 0$, $\tilde{a}_N^* \, W_f^* \, \mathbf{Q}_{W,N}(pf) \, W_p \, \tilde{b}_N$ and $\frac{(c_N \mathbf{t}_N(z))^2}{1 - (c_N \mathbf{t}_N(z))^2}$ are analytic on $\mathbb{C} - [-\sqrt{x_{+,*} + \delta}, \sqrt{x_{+,*} + \delta}]$ and bounded on each compact subset of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$. Montel's theorem thus implies that (II.68) holds for each $z \in \mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$. Moreover, the convergence is uniform on each compact subset of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$.

We now assume that $a_N, b_N, \tilde{a}_N, \tilde{b}_N$ are random bounded vectors independent from the $(v_n)_{n \geq 1}$, and just verify that $a_N^* \mathbf{Q}_{W,N}(pf) b_N \to 0$ almost surely still holds. We denote by $(\Omega_{a,b}, \mathbb{P}_{a,b})$ and $(\Omega_v, \mathbb{P}_v)$ the probability spaces on which $(a_N, b_N)_{N \geq 1}$ and the random variables $(v_n)_{n \geq 1}$ are defined. We consider the event A on which $a_N^* \mathbf{Q}_{W,N}(pf) b_N$ does not converge towards zero, and justify that $\mathbb{P}(A) = 0$ where $\mathbb{P} = \mathbb{P}_{a,b} \otimes \mathbb{P}_v$. For each element $\omega_{a,b} \in \Omega_{a,b}$, we denote by $A_{\omega_{a,b}}$ the event

$$A_{\omega_{a,b}} = \{\omega_v \in \Omega_v, (\omega_{a,b}, \omega_v) \in A\}$$

Then, the Fubini theorem leads to

$$\mathbb{P}(A) = \int_{\Omega_{a,b}} \mathbb{P}(A_{\omega_{a,b}}) \mathbb{P}_{a,b}(d\omega_{a,b}) \tag{II.69}$$

As the sequence of realizations $(a_N(\omega_{a,b}))_{N \geq 1}$ and $(b_N(\omega_{a,b}))_{N \geq 1}$ are bounded vectors, item (i) implies that $a_N^*(\omega_{a,b}) \mathbf{Q}_{W,N}(pf) b_N(\omega_{a,b}) \to 0$ almost surely, or equivalently, $\mathbb{P}(A_{\omega_{a,b}}) = 0$. (II.69) leads to the conclusion that $\mathbb{P}(A) = 0$ as expected.

(II.65) is an immediate consequence of the statements of items (i) to (iv) and their generalization to the context of random vectors $(a_N, b_N, \tilde{a}_N, \tilde{b}_N)$ (because the columns of $\Theta_{i,N}, \tilde{\Theta}_{i,N}$ are bounded random vectors for $i = p, f$ and the entries of $\Delta_{i,N}$ are bounded random variables), as well as of Condition (II.33) which implies that $r \times r$ diagonal matrices $\Delta_{p,N}$ and $\Delta_{f,N}$ (resp. orthogonal $ML \times r$ matrices $\Theta_{f,N}$ and $\Theta_{p,N}$) have the same asymptotic behaviour than matrix Δ_N (resp. matrix Θ_N). ■

Using the same kind of arguments as in the proof of Lemma II.1, it is possible to establish the following result.

Proposition II.7. *For each $z \in \mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$, it holds that*

$$\mathcal{A}^* \, \mathbf{Q}_{W,N}(pp) \, \mathcal{A} - \begin{pmatrix} -\Theta_N^* \left(zI + \frac{c_N \mathbf{t}_N(z)}{1 - (c_N \mathbf{t}_N(z))^2} I_L \otimes R_N\right)^{-1} \Theta_N & 0 \\ 0 & \frac{c_N \mathbf{t}_N(z)}{1 - (c_N \mathbf{t}_N(z))^2} \Delta_N^2 \end{pmatrix} \to 0 \ a.s. \tag{II.70}$$

$$\mathcal{B}^* \, \mathbf{Q}_{W,N}(ff) \, \mathcal{B} - \begin{pmatrix} \Gamma_N^* & I \\ I & 0 \end{pmatrix} \begin{pmatrix} -\Theta_N^* \left(zI + \frac{c_N \mathbf{t}_N(z)}{1-(c_N \mathbf{t}_N(z))^2} I_L \otimes R_N \right)^{-1} \Theta_N & 0 \\ I & 0 \end{pmatrix} \begin{pmatrix} \Gamma_N & I \\ I & 0 \end{pmatrix} \rightarrow 0 \; a.s. \tag{II.71}$$

$$\mathcal{B}^* \, \mathbf{Q}_{W,N}(ff) \, \mathcal{A} - \begin{pmatrix} 0 & -\frac{(c_N \mathbf{t}_N(z))^2}{1-(c_N \mathbf{t}_N(z))^2} \Gamma_N \\ 0 & 0 \end{pmatrix} \rightarrow 0 \; a.s. \tag{II.72}$$

The convergence is moreover uniform on each compact subset of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$.

Lemma II.1 and Proposition II.7 imply that for each $z \in \mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$, almost surely, matrix $F_N(z)$ has the same asymptotic behaviour than the $4r \times 4r$ matrix $F_{d,N}(z)$ defined by

$$F_{d,N}(z) = \begin{pmatrix} F_{d,N}^{11}(z) & F_{d,N}^{1,2}(z) \\ F_{d,N}^{2,1}(z) & F_{d,N}^{2,2}(z) \end{pmatrix} \tag{II.73}$$

where the $2r \times 2r$ blocks of $F_{d,N}(z)$ are characterized in Lemma II.1 and in Proposition II.7. The assumptions formulated in Paragraph II-D imply that matrix $F_{d,N}(z)$ converges for each $z \in \mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$ towards a limit $F_*(z)$, the convergence being uniform on each compact subset of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$. More precisely, $t_N(z)$ converges towards $t_*(z)$ uniformly on each compact subset of $\mathbb{C} - [0, x_{+,*}]$, which implies that $\mathbf{t}_N(z) = z t_N(z^2)$ converges uniformly on each compact subset of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$ towards $\mathbf{t}_*(z) = z t_*(z^2)$. We notice that matrix $-\left(zI + \frac{c_N \mathbf{t}_N(z)}{1-(c_N \mathbf{t}_N(z))^2} I_L \otimes R_N \right)^{-1}$ coincides with matrix $I_L \otimes \mathbf{T}_N(z) = I_L \otimes z T_N(z^2) = z \int_0^{x_{+,N}} \frac{I_L \otimes d\nu_N^T(\lambda)}{\lambda - z^2}$. We denote by $\mathbf{T}_{\beta_N}(z)$ the function defined by $\mathbf{T}_{\beta_N}(z) = z T_{\beta_N}(z^2)$, which can also be written as

$$\Theta_N^* (I_L \otimes \mathbf{T}_N(z)) \Theta_N = \mathbf{T}_{\beta_N}(z)$$

and which, by (I.18), coincides with the Stieltjes transform of a positive matrix-valued measure carried by $[-\sqrt{x_{+,N}}, \sqrt{x_{+,N}}]$. Proposition II.5 implies that $\mathbf{T}_{\beta_N}(z)$ converges uniformly on each compact subset of $\mathbb{C} \setminus [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$ towards the $r \times r$ matrix $\mathbf{T}_{\beta_*}(z)$ defined by

$$\mathbf{T}_{\beta_*}(z) = z T_{\beta_*}(z^2) \tag{II.74}$$

where we recall that $T_{\beta_*}(z) = \int_0^{x_{+,*}} \frac{d\beta_*(\lambda)}{\lambda - z}$ is the Stieltjes transform of the positive matrix-valued measure β_*. $\mathbf{T}_{\beta_*}$ is an element of $\mathcal{S}_r(\mathbb{R})$, its associated positive measure, denoted $\boldsymbol{\beta}_*$, is carried by $[-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$, and verifies $\boldsymbol{\beta}_*([-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]) = \boldsymbol{\beta}_*(\mathbb{R}) = I$ because $\beta_*([0, x_{+,*}] = \beta_*(\mathbb{R}^+) = I$ (see (I.18) and (I.19)). All this imply that

$$F_{d,N}^{(1,1)}(z) = \begin{pmatrix} 0 & 0 \\ -\frac{(c_N \mathbf{t}_N(z))^2}{1-(c_N \mathbf{t}_N(z))^2} \Gamma_N^* & 0 \end{pmatrix} \rightarrow F_*^{1,1}(z) = \begin{pmatrix} 0 & 0 \\ -\frac{(c_* \mathbf{t}_*(z))^2}{1-(c_* \mathbf{t}_*(z))^2} \Gamma_*^* & 0 \end{pmatrix}$$

$$F_{d,N}^{(1,2)}(z) = \begin{pmatrix} \mathbf{T}_{\beta_N}(z) & 0 \\ 0 & \frac{c_N \mathbf{t}_N(z)}{1-(c_N \mathbf{t}_N(z))^2} \Delta_N^2 \end{pmatrix} \rightarrow F_*^{1,2}(z) = \begin{pmatrix} \mathbf{T}_{\beta_*}(z) & 0 \\ 0 & \frac{c_* \mathbf{t}_*(z)}{1-(c_* \mathbf{t}_*(z))^2} \Delta_*^2 \end{pmatrix}$$

$$F_{d,N}^{2,1}(z) = \begin{pmatrix} \Gamma_N & I \\ I & 0 \end{pmatrix} F_{d,N}^{1,2}(z) \begin{pmatrix} \Gamma_N^* & I \\ I & 0 \end{pmatrix} \rightarrow F_*^{2,1}(z) = \begin{pmatrix} \Gamma_* & I \\ I & 0 \end{pmatrix} F_*^{1,2}(z) \begin{pmatrix} \Gamma_*^* & I \\ I & 0 \end{pmatrix}$$

$$F_{d,N}^{2,2}(z) = \begin{pmatrix} 0 & -\frac{(c_N \mathbf{t}_N(z))^2}{1-(c_N \mathbf{t}_N(z))^2} \Gamma_N \\ 0 & 0 \end{pmatrix} \rightarrow F_*^{2,2}(z) = \begin{pmatrix} 0 & -\frac{(c_* \mathbf{t}_*(z))^2}{1-(c_* \mathbf{t}_*(z))^2} \Gamma_* \\ 0 & 0 \end{pmatrix}$$

where we recall that Γ_* is defined by Assumption II.3. The previous results show that $(F_N(z))_{N \geq 1}$ converge uniformly towards $F_*(z)$ over each compact subset of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$. It is thus reasonable to expect that for $\delta > 0$ small enough, the solutions of the equation $\det(I + F_N(y)) = 0$ satisfying $y > \sqrt{x_{+,*} + \delta}$ will converge towards the roots of $\det(I + F_*(y)) = 0$ satisfying $y > \sqrt{x_{+,*}}$.

We now study the solutions of $\det(I+F_*(y))=0$, $y>\sqrt{x_{+,*}}$. For $y>\sqrt{x_{+,*}}$, we express in a more convenient manner the equation $\det(I+F_*(y))=0$. This equation holds if and only

$$\det\left(\left(\begin{array}{cc} I & 0 \\ 0 & \Omega_* \end{array}\right)(I+F_*(y))\left(\begin{array}{cc} \Omega_*^* & 0 \\ 0 & I \end{array}\right)\right)=0 \tag{II.75}$$

where

$$\Omega_*=\left(\begin{array}{cc} \Gamma_* & I \\ I & 0 \end{array}\right)^{-1}=\left(\begin{array}{cc} 0 & I \\ I & -\Gamma_* \end{array}\right)$$

The matrix whose determinant vanishes in (II.75) is equal to

$$\left(\begin{array}{cccc} 0 & I & \mathbf{T}_{\beta_*}(z) & 0 \\ I & -\frac{\Gamma_*^*}{1-(c_*\mathbf{t}_*(z))^2} & 0 & \frac{c_N\mathbf{t}_*(z)}{1-(c_*\mathbf{t}_*(z))^2}\Delta_*^2 \\ \mathbf{T}_{\beta_*}(z) & 0 & 0 & I \\ 0 & \frac{c_*\mathbf{t}_*(z)}{1-(c_*\mathbf{t}_*(z))^2}\Delta_*^2 & I & -\frac{\Gamma_*}{1-(c_*\mathbf{t}_*(z))^2} \end{array}\right) \tag{II.76}$$

As the lower diagonal $2r\times 2r$ block of this matrix is invertible, its determinant is 0 if and only the determinant of its Schur complement is 0. After some calculations, we obtain that $\det(I+F_*(y))=0$ if and only if $\det(I-K_*(y))=0$ where $K_*(z)$ is the $2r\times 2r$ matrix-valued function defined for each $z\in\mathbb{C}-[-\sqrt{x_{+,*}},\sqrt{x_{+,*}}]$ by

$$K_*(z)=\left(\begin{array}{cc} \frac{c_*\mathbf{t}_*(z)}{1-(c_*\mathbf{t}_*(z))^2}\Delta_*^2\mathbf{T}_{\beta_*}(z) & \frac{\Gamma_*^*}{1-(c_*\mathbf{t}_*(z))^2} \\ \frac{\mathbf{T}_{\beta_*}(z)\,\Gamma_*\,\mathbf{T}_{\beta_*}(z)}{1-(c_*\mathbf{t}_*(z))^2} & \frac{c_*\mathbf{t}_*(z)}{1-(c_*\mathbf{t}_*(z))^2}\mathbf{T}_{\beta_*}(z)\,\Delta_*^2 \end{array}\right) \tag{II.77}$$

$K_*(z)$ can be factorized as

$$K_*(z)=\left(\begin{array}{cc} I & 0 \\ 0 & \mathbf{T}_{\beta_*}(z) \end{array}\right)\left(\begin{array}{cc} \frac{c_*\mathbf{t}_*(z)}{1-(c_*\mathbf{t}_*(z))^2}\Delta_*^2 & \frac{\Gamma_*^*}{1-(c_*\mathbf{t}_*(z))^2} \\ \frac{\Gamma_*}{1-(c_*\mathbf{t}_*(z))^2} & \frac{c_*\mathbf{t}_*(z)}{1-(c_*\mathbf{t}_*(z))^2}\Delta_*^2 \end{array}\right)\left(\begin{array}{cc} \mathbf{T}_{\beta_*}(z) & 0 \\ 0 & I \end{array}\right)$$

For each $y>\sqrt{x_{+,*}}$, $\mathbf{T}_{\beta_*}(y)$ can be written as

$$\mathbf{T}_{\beta_*}(y)=\int_{-\sqrt{x_{+,*}}}^{\sqrt{x_{+,*}}}\frac{d\boldsymbol{\beta}_*(\lambda)}{\lambda-y}$$

and verifies $\mathbf{T}_{\beta_*}(y)\leq -\frac{1}{\sqrt{x_{+,*}}+y}\boldsymbol{\beta}_*([-\sqrt{x_{+,*}},\sqrt{x_{+,*}}])=-\frac{I}{\sqrt{x_{+,*}}+y}$ because we recall that $\boldsymbol{\beta}_*([-\sqrt{x_{+,*}},\sqrt{x_{+,*}}])=\boldsymbol{\beta}_*(\mathbb{R})=I$. Therefore, $\mathbf{T}_{\beta_*}(y)$ is negative definite, and thus invertible. Hence, $\det(I-K_*(y))=0$ if and only

$$\det\left(\left(\begin{array}{cc} \frac{c_*\mathbf{t}_*(y)}{1-(c_*\mathbf{t}_*(y))^2}\Delta_*^2 & \frac{\Gamma_*^*}{1-(c_*\mathbf{t}_*(y))^2} \\ \frac{\Gamma_*}{1-(c_*\mathbf{t}_*(y))^2} & \frac{c_*\mathbf{t}_*(y)}{1-(c_*\mathbf{t}_*(y))^2}\Delta_*^2 \end{array}\right)-\left(\begin{array}{cc} (\mathbf{T}_{\beta_*}(y))^{-1} & 0 \\ 0 & (\mathbf{T}_{\beta_*}(y))^{-1} \end{array}\right)\right)=0 \tag{II.78}$$

In the following, we denote by $H_*(z)$ the $2r\times 2r$ matrix-valued function defined on $\mathbb{C}-[\sqrt{x_{+,*}},\sqrt{x_{+,*}}]$ by

$$H_*(z)=\left(\begin{array}{cc} \frac{c_*\mathbf{t}_*(z))}{1-(c_*\mathbf{t}_*(z))^2}\Delta_*^2-(\mathbf{T}_{\beta_*}(z))^{-1} & \frac{\Gamma_*^*}{1-(c_*\mathbf{t}_*(z))^2} \\ \frac{\Gamma_*}{1-(c_*\mathbf{t}_*(z))^2} & \frac{c_*\mathbf{t}_*(z)}{1-(c_*\mathbf{t}_*(z))^2}\Delta_*^2-(\mathbf{T}_{\beta_*}(z))^{-1} \end{array}\right) \tag{II.79}$$

$H_*(z)$ is of course holomorphic on $\mathbb{C}-[-\sqrt{x_{+,*}},\sqrt{x_{+,*}}]$, and the solutions of $\det(I+F_*(y))=0$, $y>\sqrt{x_{+,*}}$, coincide with the solutions of

$$\det\left(H_*(y)\right)=0 \tag{II.80}$$

where $y>\sqrt{x_{+,*}}$. In order to characterize the roots of (II.80), we first establish the following Proposition.

Proposition II.8. *For each $z\in\mathbb{C}^+$, $\mathrm{Im}(H_*(z))>0$, and function $y\to H_*(y)$ is increasing in the sense of the partial order defined on the set of all Hermitian matrices on the interval $[\sqrt{x_{+,*}},+\infty[$.*

Proof. As $\boldsymbol{\beta}_*(\mathbb{R})=I$, $\mathrm{Im}(\mathbf{T}_{\beta_*}(z))$ is positive definite for $z\in\mathbb{C}^+$ and $\mathrm{Im}\left((\mathbf{T}_{\beta_*}(z))^{-1}\right)<0$. Therefore, in order to establish that $\mathrm{Im}(H_*(z))>0$ on $\mathbb{C}^+$, it is sufficient to prove that $\mathrm{Im}(H_{*,1}(z))>0$ on $\mathbb{C}^+$ where $H_{*,1}(z)$ is the function defined by

$$H_{*,1}(z)=\left(\begin{array}{cc} \frac{c_*\mathbf{t}_*(z)}{1-(c_*\mathbf{t}_*(z))^2}\Delta_*^2 & \frac{\Gamma_*^*}{1-(c_*\mathbf{t}_*(z))^2} \\ \frac{\Gamma_*}{1-(c_*\mathbf{t}_*(z))^2} & \frac{c_*\mathbf{t}_*(z)}{1-(c_*\mathbf{t}_*(z))^2}\Delta_*^2 \end{array}\right)$$

After some calculations, we obtain that

$$\mathrm{Im}(H_{*,1}(z)) = \frac{1}{|1-(c_*\mathbf{t}_*(z))^2|^2} \begin{pmatrix} \mathrm{Im}(c_*t_*(z))(1+|c_*t_*(z)|^2)\Delta_*^2 & \mathrm{Im}\left((c_*t_*(z))^2\right)\Gamma_*^* \\ \mathrm{Im}\left((c_*t_*(z))^2\right)\Gamma_* & \mathrm{Im}(c_*t_*(z))(1+|c_*t_*(z)|^2)\Delta_*^2 \end{pmatrix}$$

It is clear that $\mathrm{Im}(c_*t_*(z))(1+|c_*t_*(z)|^2)\Delta_*^2 > 0$. Therefore, $\mathrm{Im}(H_{*,1}(z)) > 0$ if and only if

$$\mathrm{Im}(c_*t_*(z))(1+|c_*t_*(z)|^2)\Delta_*^2 - \frac{\left[\mathrm{Im}\left((c_*t_*(z))^2\right)\right]^2}{\mathrm{Im}(c_*t_*(z))(1+|c_*t_*(z)|^2)}\Gamma_*^*\Delta_*^{-2}\Gamma_* > 0$$

or equivalently, if and only if

$$I - \frac{\left[\mathrm{Im}\left((c_*t_*(z))^2\right)\right]^2}{\left[\mathrm{Im}(c_*t_*(z))(1+|c_*t_*(z)|^2)\right]^2}\Delta_*^{-1}\Gamma_*^*\Delta_*^{-2}\Gamma_*\Delta_*^{-1} > 0 \tag{II.81}$$

We first claim that $\Delta_*^{-1}\Gamma_*^*\Delta_*^{-2}\Gamma_*\Delta_*^{-1} \leq I$. To verify this, we notice that for each N, matrix $\Delta_N^{-1}\Gamma_N^*\Delta_N^{-2}\Gamma_N\Delta_N^{-1}$ coincides with $\tilde{\Theta}_{f,N}^*\tilde{\Theta}_{p,N}\tilde{\Theta}_{p,N}^*\tilde{\Theta}_{f,N}$ which is less than I. Therefore,

$$\lim_{N\to+\infty}\Delta_N^{-1}\Gamma_N^*\Delta_N^{-2}\Gamma_N\Delta_N^{-1} = \Delta_*^{-1}\Gamma_*^*\Delta_*^{-2}\Gamma_*\Delta_*^{-1} \leq I$$

$\frac{\left[\mathrm{Im}\left((c_*t_*(z))^2\right)\right]^2}{\left[\mathrm{Im}(c_*t_*(z))(1+|c_*t_*(z)|^2)\right]^2}$ is equal to

$$\frac{\left[\mathrm{Im}\left((c_*t_*(z))^2\right)\right]^2}{\left[\mathrm{Im}(c_*t_*(z))(1+|c_*t_*(z)|^2)\right]^2} = \frac{4\left[\mathrm{Re}(c_*t_*(z))\right]^2}{(1+|c_*t_*(z)|^2)^2}$$

For $z \in \mathbb{C}^+$, $\mathrm{Im}(t_*(z)) > 0$. Therefore, it holds that $\left(\mathrm{Re}(c_*t_*(z))\right)^2 < |c_*t_*(z)|^2$ and that

$$\frac{\left[\mathrm{Im}\left((c_*t_*(z))^2\right)\right]^2}{\left[\mathrm{Im}(c_*t_*(z))(1+|c_*t_*(z)|^2)\right]^2} < \frac{4|c_*t_*(z)|^2}{(1+|c_*t_*(z)|^2)^2} \leq 1$$

This establishes (II.81) and $\mathrm{Im}(H_*(z)) > 0$. ■

We now prove that $y \to H_*(y)$ is increasing on the interval $[\sqrt{x_{+,*}}, +\infty[$. For this, we use the following representation of holomorphic matrix-valued functions whose imaginary part is positive on $\mathbb{C}^+$ (see e.g. [19]):

$$H_*(z) = A + Bz + \int \frac{1+\lambda z}{\lambda - z}\frac{d\sigma(\lambda)}{1+\lambda^2} \tag{II.82}$$

where A is Hermitian, $B \geq 0$ and σ is a positive matrix-valued measure for which

$$\mathrm{Tr}\left(\frac{d\sigma(\lambda)}{1+\lambda^2}\right) < +\infty$$

$B = \lim_{y\to+\infty}\frac{H_*(iy)}{iy}$ is easily seen to be equal to

$$B = \lim_{y\to+\infty}\begin{pmatrix} -\frac{\mathbf{T}_{\beta_*}(iy)}{iy} & 0 \\ 0 & -\frac{\mathbf{T}_{\beta_*}(iy)}{iy} \end{pmatrix} = I_{2r}$$

while for any interval $[y_1, y_2]$, it holds that

$$\sigma([y_1,y_2]) = \frac{1}{\pi}\lim_{\epsilon\to 0}\int_{y_1}^{y_2}\mathrm{Im}(H_*(y+i\epsilon))dy$$

As $\mathrm{Im}(H_*(y)) = 0$ if $|y| > \sqrt{x_{+,*}}$, the support of σ is included into $[-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$. Therefore, we get immediately from (II.82) that $y \to H_*(y)$ is strictly increasing on $]\sqrt{x_{+,*}}, +\infty]$, i.e. $H_*(y_2) > H_*(y_1)$ if $y_2 > y_1$. We also notice that the last item of Proposition II.4 as well as Proposition II.5 imply that $\lim_{y\to\sqrt{x_{+,*}}} H_*(y) = H_*(\sqrt{x_{+,*}})$ exists and is finite. Moreover, it holds that $H_*(\sqrt{x_{+,*}}) < H_*(y)$ for $y > \sqrt{x_{+,*}}$.

Corollary II.1. *The eigenvalues (arranged in the decreasing order) $(\lambda_{k,*}(y))_{k=1,\dots,2r}$ of matrix $H_*(y)$ are strictly increasing functions of y on $[\sqrt{x_{+,*}}, +\infty[$, i.e., for each $k = 1, \dots, 2r$, it holds that*

$$\lambda_{k,*}(y_1) < \lambda_{k,*}(y_2) \text{ if } \sqrt{x_{+,*}} \leq y_1 < y_2 \tag{II.83}$$

Moreover, the number s of solutions of (II.80) (taking into account their multiplicities) for which $y > \sqrt{x_{+,*}}$ belongs to $\{0, 1, \dots, 2r\}$, and coincides with the number of strictly negative eigenvalues of matrix $H_*(\sqrt{x_{+,*}})$.

Proof. We have shown that if $\sqrt{x_{+,*}} \le y_1 < y_2$, then $H_*(y_1) < H_*(y_2)$. The Weyl's inequalities (see e.g. [26], Paragraph 4.3) thus imply that (II.83) holds. Moreover, as matrix B in (II.82) is equal to I_{2r}, it is clear that for each $k = 1, \dots, 2r$, $\lambda_{k,*}(y)$ converges towards $+\infty$ when $y \to +\infty$. For $k = 1, \dots, 2r$, the equation $\lambda_{k,*}(y) = 0$ has thus 1 solution $y > \sqrt{x_{+,*}}$ if $\lambda_{k,*}(x_{+,*}) < 0$ and no solution if $\lambda_{k,*}(x_{+,*}) \ge 0$. (II.80) holds if and only if one of the eigenvalues of $H_*(y)$ is equal to 0. Therefore, if we denote by $\tilde{s}$ the number of positive eigenvalues of $H_*(\sqrt{x_{+,*}})$, for $j = 1, \dots, \tilde{s}$, it must hold that $\lambda_{j,*}(y) > 0$ for $y > \sqrt{x_{+,*}}$. Moroever, $\lambda_{\tilde{s}+1,*}(\sqrt{x_{+,*}}) < 0$ implies that the equation $\lambda_{\tilde{s}+1,*}(y) = 0$ has a unique solution $y_{1,*} > \sqrt{x_{+,*}}$. Similarly, the equation $\lambda_{\tilde{s}+2,*}(y) = 0$ has a unique solution denoted $y_{2,*}$. Moreover, as $\lambda_{\tilde{s}+2,*}(y) \le \lambda_{\tilde{s}+1,*}(y)$ for each y, we deduce that $\lambda_{\tilde{s}+2,*}(y_{1,*}) \le \lambda_{\tilde{s}+1,*}(y_{1,*}) = 0$. If $\lambda_{\tilde{s}+2,*}(y_{1,*}) < 0$, $y_{2,*}$ must be strictly greater than $y_{1,*}$. As a root of (II.80), $y_{1,*}$ has thus multiplicity 1. If $\lambda_{\tilde{s}+2,*}(y_{1,*}) = 0$, the multiplicity of $y_{1,*}$ as a root of (II.80) is at least equal to 2. Iterating the process, we obtain that the number of solutions s (taking into account the multiplicities) of (II.80) is equal to $s = 2r - \tilde{s}$. Moreover, solutions $y_{1,*}, \dots, y_{s,*}$ satisfy $y_{1,*} \le y_{2,*} \le \dots \le y_{s,*}$. ■

Corollary II.1 implies that Eq. $\det(I + F_*(y)) = 0$ has s ($0 \le s \le 2r$) solutions $(y_{k,*})_{k=1,\dots,s}$ strictly greater than $\sqrt{x_{+,*}}$. We recall that, almost surely, the sequence of functions $(F_N(z))_{N \ge 1}$ converges uniformly on each compact subset of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$ towards $F_*(z)$. We now take benefit of the arguments used in [6], Lemma 6.1 and in the proof of Theorem 2.1 in [10] to derive the following result.

Corollary II.2. *For each $\delta > 0$ small enough, almost surely, for N large enough, Eq. $\det(I + F_N(y)) = 0$ has s solutions $y_{1,N} \le y_{2,N} \dots \le y_{s,N}$ such that $y_{k,N} > \sqrt{x_{+,*} + \delta}$, and for each $k = 1, \dots, s$, it holds that $\lim_{N \to +\infty} y_{k,N} = y_{k,*}$.*

Proof. We just provide a sketch of proof because we follow the arguments in [6] and [10]. In order to simplify the exposition, we assume that $y_{1,*} < \dots < y_{s,*}$, but the following arguments can be extended immediately to the case where some $(y_{k,*})_{k=1,\dots,s}$ coincide. We first justify that almost surely, for N large enough, the solutions of $\det(I + F_N(y)) = 0$, $y > \sqrt{x_{+,*} + \delta}$ are bounded by a nice constant. To verify this, we remark that (II.3) implies that it exists a nice constant κ for which, almost surely, $\|\mathcal{A}_N\| \le \kappa$ and $\|\mathcal{B}_N\| \le \kappa$ for each N large enough. We recall that $\mathcal{A}_N$ and $\mathcal{B}_N$ are defined by (II.43) and (II.44). Moreover, for $y > \sqrt{x_{+,*} + \delta}$, the inequality $\|\mathbf{Q}_W(y)\| \le \frac{1}{y - \sqrt{x_{+,*} + \delta}}$ holds. Therefore, matrix $F_N(y)$ verifies $\|F_N(y)\| < \frac{\kappa}{y - \sqrt{x_{+,*} + \delta}}$ for some nice constant κ, and all the eigenvalues of $F_N(y)$ satisfy $|\lambda_j(F_N(y))| \le \frac{\kappa}{y - \sqrt{x_{+,*} + \delta}}$, for $j = 1, \dots, 2r$. For y larger than a nice constant y_{max}, $\det(I_{4r} + F_N(y))$ cannot therefore vanish. This implies that almost surely, for N large enough, the solutions of $\det(I + F_N(y)) = 0$, $y > \sqrt{x_{+,*} + \delta}$ belong to $(\sqrt{x_{+,*} + \delta}, y_{max})$. We choose δ in such a way that $\sqrt{x_{+,*} + \delta} < y_{1,*}$. We consider any open interval (a_1, a_2) such that $(a_1, a_2) \subset (\sqrt{x_{+,*} + \delta}, \max(y_{max}, y_{s,*}) + \delta)$ and $a_i \neq y_{k,*}$ for $i = 1, 2$ and $k = 1, \dots, s$. Then, using the arguments in the proof of Theorem 2.1 in [10], we obtain that the equations $\det(I + F_N(y)) = 0$ and $\det(I + F_*(y)) = 0$ have the same number of solutions located in (a_1, a_2). Choosing $(a_1, a_2) = (\sqrt{x_{+,*} + \delta}, \max(y_{max}, y_{s,*}) + \delta)$ leads to the conclusion that $\det(I + F_N(y)) = 0$ has s solutions $y_{1,N}, \dots, y_{s,N}$ larger than $\sqrt{x_{+,*} + \delta}$. We fix $k \in \{1, 2, \dots, s\}$ and establish that $y_{k,N} \to y_{k,*}$. For this, we choose $\epsilon > 0$ arbitrarily small, and choose $(a_1, a_2) = (y_{k,*} - \epsilon, y_{k,*} + \epsilon)$. Then, [10] implies that almost surely, for N large enough, $\det(I + F_N(y)) = 0$ has 1 solution $y_{k,N}$ in $(y_{k,*} - \epsilon, y_{k,*} + \epsilon)$, and that $|y_{k,N} - y_{k,*}| < \epsilon$. This is equivalent to $\lim_{N \to +\infty} y_{k,N} = y_{k,*}$ as expected. ■.

Remark II.1. *We notice that the existence of the limits $\Delta_*, \Gamma_*, t_*, \beta_*$ introduced in the various Assumptions of Section II allows to establish that $F_N(z)$ converges towards the deterministic and independent of N function $F_*(z)$, and to prove that the solutions of $\det(I + F_N(y)) = 0$ larger than $\sqrt{x_{+,*} + \delta}$ converge towards the corresponding solutions of $\det(I + F_*(y)) = 0$. If the above limits are not supposed to exist, we can just establish that $F_N(z)$ has the same asymptotic behaviour that the term $F_{d,N}(z)$ introduced in (II.73). As $F_{d,N}(z)$ depends on N, it is not possible to adapt the arguments in the proof of Theorem 2.1 in [10] to establish rigorously that the solutions of $\det(I + F_N(y)) = 0$ larger than $\sqrt{x_{+,N}}$ have the same behaviour than the corresponding solutions of $\det(I + F_{d,N}(y)) = 0$. However, the existence of $\Delta_*, \Gamma_*, t_*, \beta_*$ can be considered as purely technical assumptions that allow to derive well founded mathematical results. In particular, even if the limits are not supposed to exist, in practice, for N large enough, the eigenvalues of $\Sigma_{f,N} \Sigma_{p,N}^* \Sigma_{p,N} \Sigma_{f,N}^*$ that escape from $[0, x_{+,N}]$ should be close* *from the solutions of* $\det(I + F_{d,N}(y)) = 0$ *larger than* $x_{+,N}$ *in a number of scenarios. However, the derivation of reasonable alternative conditions under which this behaviour holds seems difficult.*

We have thus established the Theorem:

Theorem II.1. *Almost surely, for each* N *large enough, the* s *largest eigenvalues* $\hat{\lambda}_{1,N} \geq \ldots \geq \hat{\lambda}_{s,N}$ *of matrix* $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ *escape from the interval* $[0, x_{+,*}]$*, and converge towards* $\rho_{1,*} \geq \ldots \geq \rho_{s,*} > x_{+,*}$ *defined by* $\rho_{k,*} = y_{s+1-k,*}^2$ *for* $k = 1, \ldots, s$*. Moreover, for each* $\delta > 0$*, the eigenvalues* $(\hat{\lambda}_{k,N})_{k \geq s+1}$ *belong to* $[0, x_{+,*} + \delta]$*.*

s and the limit eigenvalues $(\rho_{k,*})_{k=1,\ldots,s}$ depend on the limit distributions ω_* and β_* that are rather immaterial. It is thus more appropriate to evaluate the asymptotic behaviour of the largest eigenvalues of $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ by using a finite N equivalent of $H_*(z)$. We thus define function $H_N(z)$ by

$$H_N(z) = \begin{pmatrix} \frac{c_N \mathbf{t}_N(z))}{1-(c_N \mathbf{t}_N(z))^2}\Delta_N^2 - (\mathbf{T}_{\beta_N}(z))^{-1} & \frac{\Gamma_N^*}{(1-(c_N \mathbf{t}_N(z))^2} \\ \frac{\Gamma_N}{(1-(c_N \mathbf{t}_N(z))^2} & \frac{c_N \mathbf{t}_N(z)}{1-(c_N \mathbf{t}_N(z))^2}\Delta_N^2 - (\mathbf{T}_{\beta_N}(z))^{-1} \end{pmatrix} \tag{II.84}$$

For each N large enough and for each $\delta > 0$, $H_N(z)$ is holomorphic in $\mathbb{C} - [-\sqrt{x_{+,*} + \delta}, \sqrt{x_{+,*} + \delta}]$ and converges uniformly on each compact subset of $\mathbb{C} - [-\sqrt{x_{+,*}}, \sqrt{x_{+,*}}]$ towards function $H_*(z)$. Using again the approach of [6] and [10], we obtain that, for each N large enough, the equation $\det(H_N(y)) = 0$ has s solutions $y_{1,N,*} \leq \ldots \leq y_{s,N,*}$ strictly larger than $\sqrt{x_{+,N} + \delta}$ for some $\delta > 0$ small enough, and which satisfy $y_{k,N,*} - y_{k,*} \to 0$ when $N \to +\infty$. Moreover, the convergence of $x_{+,N}$ and $w_{+,N}$ towards $x_{+,*} = \phi_*(w_{+,*})$ and $w_{+,*} = w_*(x_{+,*})$ imply that $\mathbf{t}_N(x_{+,N})$ converge towards $\mathbf{t}_*(x_{+,*})$. Therefore, (II.64) leads to the following Corollary.

Corollary II.3. $H_N(\sqrt{x_{+,N}})$ *converges towards* $H_*(\sqrt{x_{+,*}})$*. Moreover, if* $\det(H_*(\sqrt{x_{+,*}})) \neq 0$*, for* N *large enough,* s *also coincides with the number of strictly negative eigenvalues of matrix* $H_N(\sqrt{x_{+,N}})$*. Finally, if we define* $\rho_{k,N}$ *by* $\rho_{k,N} = y_{s+1-k,N,*}^2$ *for* $k = 1, \ldots, s$ *, then it holds that* $\hat{\lambda}_{k,N} - \rho_{k,N} \to 0$ *almost surely.*

Proof. It just remains to remark that if 0 is not eigenvalue of $H_*(\sqrt{x_{+,*}})$, then, for each N large enough, s is equal to the number of strictly negative eigenvalues of matrix $H_N(\sqrt{x_{+,N}})$. ■

Matrix $H_N(\sqrt{x_{+,N}})$ can be written in a more explicit way, so that s can be evaluated using the following alternative formulation.

Corollary II.4. *Define* G_N *as the* $r \times r$ *matrix given by*

$$G_N = \frac{c_N w_{+,N}}{\sqrt{x_{+,N}}} \frac{1}{M}\mathrm{Tr}(R_N(w_{+,N}I - R_N)^{-1}) \left[\left(\Theta_N^*(I_L \otimes (w_{+,N}I - R_N)^{-1})\Theta_N\right)^{-1} - \Delta_N^2 \right] \tag{II.85}$$

Then, if $\det(H_*(\sqrt{x_{+,*}})) \neq 0$*, for each* N *large enough,* s *coincides with the number of strictly negative eigenvalues of the* $2r \times 2r$ *matrix*

$$\begin{pmatrix} G_N & \Gamma_N^* \\ \Gamma_N & G_N \end{pmatrix} \tag{II.86}$$

Proof. Writing $\mathbf{t}_N(z)$ as $\mathbf{t}_N(z) = z t_N(z^2)$, and using the expression (II.17) of $t_N(z)$ in terms of $w_N(z)$, we obtain after some algebra that matrix $H_N(\sqrt{x_{+,N}})$ is given by

$$H_N(\sqrt{x_{+,N}}) = \left(1 + c_N \frac{1}{M}\mathrm{Tr}(R_N(w_{+,N}I - R_N)^{-1})\right) \begin{pmatrix} G_N & \Gamma_N^* \\ \Gamma_N & G_N \end{pmatrix} \tag{II.87}$$

As $w_{+,N} > \lambda_{1,N}$, we have $\frac{1}{M}\mathrm{Tr}(R_N(w_{+,N}I - R_N)^{-1} > 0$ and $1 + c_N \frac{1}{M}\mathrm{Tr}(R_N(w_{+,N}I - R_N)^{-1}) > 0$. s thus coincides with the number of strictly negative eigenvalues of (II.86). ■

F. When Condition (II.33) does not hold.

We now consider the case where Condition (II.33) does not hold, and briefly indicate how the above results have to be modified. For this, we denote by k the number of different diagonal entries of Δ_* and by $m_1, \ldots, m_k$ their multiplicities, which also coincide with the multiplicities of the k different eigenvalues of matrices $\Theta_N\Delta_*\Theta_N^*$ and $\Theta_{i,N}\Delta_*\Theta_{i,N}^*$ for $i = p, f$. If $(\Theta_N(l))_{l=1,\ldots,k}$ and $(\Theta_{i,N}(l))_{l=1,\ldots,k}$ represent the $ML \times m_l$ matrices for defined by $\Theta_N = (\Theta_N(1), \ldots, \Theta_N(k))$ and $\Theta_{i,N} = (\Theta_{i,N}(1), \ldots, \Theta_{i,N}(k))$, then, (II.36) and standard results of perturbation theory imply that

$$\|\Theta_{i,N}(l)\Theta_{i,N}(l)^* - \Theta_N(l)\Theta_N(l)^*\| \to 0 \tag{II.88}$$

for each $l = 1, \ldots, k$. We denote by $X_{i,N}(l)$ the $m_l \times m_l$ random matrix defined by $X_{i,N}(l) = \Theta_N(l)^* \Theta_{i,N}(l)$, and deduce from (II.88) that

$$\|\Theta_{i,N}(l) - \Theta_N(l) X_{i,N}(l)\| \to 0 \tag{II.89}$$

as well as

$$X_{i,N}(l)^* X_{i,N}(l) - I_{m_l} \to 0, \; X_{i,N}(l) X_{i,N}(l)^* - I_{m_l} \to 0 \tag{II.90}$$

Therefore, matrix $\Theta_{i,N}$ can be replaced up to error terms by matrix $\Theta_N X_{i,N}$ where $X_{i,N}$ represents the $r \times r$ block diagonal matrix with diagonal blocks $X_{i,N}(1), \ldots, X_{i,N}(k)$. It is useful to notice that the very definition of $X_{i,N}$ implies that the equality

$$X_{i,N} \Delta_* = \Delta_* X_{i,N} \tag{II.91}$$

holds. Another consequence of (II.90) is related to the asymptotic behaviour of matrix Γ_N. In particular, (II.39) does no longer hold, and we rather have

$$\Gamma_N - X_{f,N}^{-1} \Gamma_* X_{p,N}^{-*} \to 0 \; a.s. \tag{II.92}$$

To justify (II.92), we notice that (II.25) leads to

$$\Theta_N^* R_{f|p,N}^L \Theta_N - \Theta_N^* \Theta_{f,N} \Delta_N \tilde{\Theta}_{f,N}^* \tilde{\Theta}_{p,N} \Delta_N \Theta_{p,N}^* \Theta_N \to 0$$

Therefore, (II.89) leads to

$$X_{f,N} \Gamma_N X_{p,N}^* - \Gamma_* \to 0$$

and to (II.92). Therefore, Γ_N does not converge towards a deterministic matrix, and rather behaves as the random matrix $X_{f,N}^{-1} \Gamma_* X_{p,N}^{-*}$. Moreover, the reader may check that the convergence results (II.70) and (II.71) have to be modified as follows: in (II.70), matrix

$$\Theta_N^* \left(zI + \frac{c_N \mathbf{t}_N(z)}{1 - (c_N \mathbf{t}_N(z))^2} I_L \otimes R_N \right)^{-1} \Theta_N = \mathbf{T}_{\beta_N}(z)$$

has to be replaced by $X_{f,N}^* \mathbf{T}_{\beta_N}(z) X_{f,N}$ while in (II.71), $\mathbf{T}_{\beta_N}(z)$ has to be exchanged with $X_{p,N}^* \mathbf{T}_{\beta_N}(z) X_{p,N}$. Matrix $F_{d,N}$ is thus modified. The modified matrix, still denoted $F_{d,N}(z)$, does no longer converge towards matrix $F_*(z)$ introduced after Proposition II.7, but appears to have almost surely the same asymptotic behaviour than the random matrix $F_{*,N}(z)$ obtained by replacing Γ_N by $X_{f,N}^{-1} \Gamma_* X_{p,N}^{-*}$, and $\mathbf{T}_{\beta_*}$ in the definitions of $F_*^{1,2}(z)$ and $F_*^{2,1}(z)$ by $X_{f,N}^* \mathbf{T}_{\beta_*}(z) X_{f,N}$ and $X_{p,N}^* \mathbf{T}_{\beta_*}(z) X_{p,N}$ respectively. However, after some algebra, it is easily seen that $\det(I + F_{*,N}(z)) = 0$ if and only if $\det(I - K_{*,N}(z)) = 0$, where $K_{*,N}$ is defined by

$$K_{*,N}(z) = \begin{pmatrix} \frac{c_* \mathbf{t}_*(z)}{1 - (c_* \mathbf{t}_*(z))^2} \Delta_*^2 X_{p,N}^* \mathbf{T}_{\beta_*}(z) X_{p,N} & \frac{X_{p,N}^{-1} \Gamma_*^* X_{f,N}^{-*}}{(1 - (c_* \mathbf{t}_*(z))^2} \\ \frac{X_{f,N}^* \mathbf{T}_{\beta_*}(z) \, \Gamma_* \mathbf{T}_{\beta_*}(z) X_{p,N}}{1 - (c_* \mathbf{t}_*(z))^2} & \frac{c_* \mathbf{t}_*(z)}{1 - (c_* \mathbf{t}_*(z))^2} X_{f,N}^* \mathbf{T}_{\beta_*}(z) X_{f,N} \, \Delta_*^2 \end{pmatrix} \tag{II.93}$$

Using (II.90) and (II.91), we obtain that

$$\begin{pmatrix} X_{p,N} & 0 \\ 0 & X_{f,N} \end{pmatrix} (I - K_{*,N}(z)) \begin{pmatrix} X_{p,N}^* & 0 \\ 0 & X_{f,N}^* \end{pmatrix} \to I - K_*(z) \tag{II.94}$$

where $K_*(z)$ is defined by (II.77). The solutions $y > \sqrt{x_{+,*}}$ of the equation $\det(I - K_*(y)) = 0$ are the $(y_{k,*})_{k=1,\ldots,s}$ introduced in Section II-E. Using the arguments in [6] and [10], we obtain from (II.94) that for $\delta > 0$ small enough, the equation $\det(I - K_{*,N}(y)) = 0$, or equivalently the equation $\det(I + F_{*,N}(y)) = 0$ has s solutions $(y_{k,N,*})_{k=1,\ldots,s}$ larger than $\sqrt{x_{+,*} + \delta}$ and verifying $y_{k,N,*} \to y_{k,*}$ for each $k = 1, \ldots, s$. [6] and [10] imply that the equations $\det(I + F_{d,N}(y)) = 0$ and $\det(I + F_N(y)) = 0$ have also s solutions larger than $\sqrt{x_{+,*} + \delta}$ and converging almost surely towards the $(y_{k,*})_{k=1,\ldots,s}$. This, in turn, shows that Theorem II.1 remains still valid when condition (II.33) does not hold.

G. Particular cases and examples

In order to get some insights on the number of eigenvalues s that escape from $\mathcal{S}_N = [0, x_{+,N}]$ for each N large enough, we first study informally the behaviour of s when $c_N \to 0$. Intuitively, we should recover the results corresponding to the traditional regime, i.e. that $s = P$. For this, we use Corollary II.4 and remark that $w_{+,N}$, which depends on c_N, satisfies $\phi_N^{'}(w_{+,N}) = 0$. Using $\phi_N^{'}(w_{+,N}) = 0$ and following the proof of Proposition 7.7 in [33] for $w_0 = w_{+,N}$ until Eq. (7.59), we obtain that $\frac{1}{M}\mathrm{Tr}(R_N(w_{+,N}I - R_N)^{-1}) < 1$. As $R_N > aI$ (see Assumption (I.11)), we obtain that

$$\frac{1}{M}\sum_{k=1}^{M}\frac{1}{w_{+,N} - \lambda_{k,N}} < \frac{1}{a}$$

holds for each c_N. This implies that $\liminf_{c_N \to 0} w_{+,N} - \lambda_{1,N} > 0$, and that matrix $\left(\Theta_N^*(I_L \otimes (w_{+,N}I - R_N)^{-1})\Theta_N\right)^{-1}$ remains bounded when $c_N \to 0$. As $x_{+,N} = \phi_N(w_{+,N})$, it is easy to check that $x_{+,N} = \mathcal{O}(c_N)$. Therefore, $\frac{c_N w_{+,N}}{\sqrt{x_{+,N}}} = \mathcal{O}(\sqrt{c_N})$, and $G_N \to 0$ when $c_N \to 0$. Therefore, when $c_N \to 0$,

$$H_N(\sqrt{x_{+,N}}) \to \begin{pmatrix} 0 & \Gamma_N^* \\ \Gamma_N & 0 \end{pmatrix}$$

As mentioned previously, matrix Γ_N has rank $P \leq r$. Therefore, the eigenvalues of matrix $\begin{pmatrix} 0 & \Gamma_N^* \\ \Gamma_N & 0 \end{pmatrix}$ are 0 with mutiplicity $2(r-P)$, $(\chi_k)_{k=1,\ldots,P}$ and $-(\chi_k)_{k=1,\ldots,P}$ where we recall that $(\chi_k)_{k=1,\ldots,P}$ represent the P non zero singular values of matrix Γ_N. Therefore, when $c_N \to 0$, s converges towards P. This is in accordance with the traditional asymptotic regime where $N \to +\infty$ and M is fixed. Indeed, in this context, matrix $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ converges towards the rank P matrix $R_{f|p}^L \left(R_{f|p}^L\right)^*$, i.e. for N large enough, matrix $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ has P eigenvalues that are significantly larger the $M-P$ smallest ones.

When c_N does not converge towards 0, the presence of matrix G_N in the expression (II.87) in general deeply modifies the value of s. In particular, the value of s depends on the singular values $(\chi_{k,N})_{k=1,\ldots,P}$ of matrix Γ_N, but also on the diagonal entries $(\delta_{k,N}^2)_{k=1,\ldots,r}$ of matrix Δ_N^2, or equivalently, on the non zero eigenvalues of $R_{u,N}^L = \mathbb{E}(u_n^L u_n^{*L})$. In contrast with the context of the usual spiked empirical covariance matrix models, s may be larger than the number P of non zero eigenvalues of the true matrix $R_{f|p}R_{f|p}^*$. This implies that if c_N is not small enough, then estimating the rank P of matrix $R_{f|p}R_{f|p}^*$ by the number s of eigenvalues of $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ that escape from $[0, x_{+,N}]$ does not lead to a consistent estimation scheme.

We now construct explicit examples where $R_N = \sigma^2 I_M$ for some $\sigma^2 > 0$, $P = L = 1$ and for which s does not coincide with $P = 1$. In particular, we now establish that for each $r \geq 2$, there exists useful signals u for which $\mathrm{Rank}(R_{u,N}^1) = \mathrm{Rank}(\mathbb{E}(u_n u_n^*)) = r$ and $s = 2r-1$. Moreover, we show that the non zero eigenvalues of $R_{u,N}^1$ as well as the non zero singular value of $R_{f|p,N}^1$ can be arbitrarily large. In the following, matrices $R_{u,N}^1$ and $R_{f|p,N}^1$ will be denoted $R_{u,N}$ and $R_{f|p,N}$. We define $K = r-1$ and we consider a $M \times r$ matrix $\Theta_N = (C_N, D_{1,N}, \ldots, D_{K,N})$ verifying $\Theta_N^*\Theta_N = I_r$ and a K–dimensional white noise sequence $(i_n)_{n\in\mathbb{Z}}$ verifying $\mathbb{E}(i_n i_n^*) = I_K$. If $0 \leq a < 1$ and $b_1, \ldots, b_K$ are real numbers, we define the signal $(u_n)_{n\in\mathbb{Z}}$ by

$$\begin{aligned} x_{n+1} &= a\, x_n + \sum_{k=1}^{K} b_k\, i_{k,n} \\ u_n &= C_N\, x_n + \sum_{k=1}^{K} \delta_{k+1}\, D_{k,N}\, i_{k,n} \end{aligned} \tag{II.95}$$

where $\delta_2, \ldots, \delta_r$ are strictly positive real numbers. As the state-space sequence is 1–dimensional, P coincides with 1. We denote by δ_1 the positive real number such that

$$\mathbb{E}(|x_n|^2) = \frac{\sum_{k=1}^{K} b_k^2}{1-a^2} = \delta_1^2$$

Then, if we denote by Δ^2 the $r \times r$ diagonal matrix $\mathrm{Diag}(\delta_1^2, \ldots, \delta_r^2)$, we obtain immediately that

$$R_{u,N} = \Theta_N \Delta^2 \Theta_N^* = \delta_1^2\, C_N C_N^* + \sum_{k=1}^{K} \delta_{k+1}^2\, D_{k,N} D_{k,N}^*$$

coincides with the eigenvalue/ eigenvector decomposition of $R_{u,N}$. We observe that the eigenvalues $(\delta_k^2)_{k=1,\ldots,r}$ do not depend on N. Therefore, matrix Δ_N coincides for each N with $\Delta = \Delta_*$. The matrix $R_{f|p,N} = \mathbb{E}(u_{n+1}u_n^*)$ is given by

$$R_{f|p,N} = C_N \left(a\delta_1^2 \, C_N^* + \sum_{k=1}^{K} b_k \delta_{k+1} \, D_{k,N}^* \right)$$

Therefore, $R_{f|p,N}$ can we written as

$$R_{f|p,N} = \Theta_N \Gamma_* \Theta_N^*$$

where $\Gamma_* = \Theta_N^* R_{f|p,N} \Theta_N$ is equal to

$$\Gamma_* = e_1 \left(a\delta_1^2, b_1\delta_2, \ldots, b_K \delta_{K+1}\right) = \chi \, \Upsilon \, \tilde{\Upsilon}^*$$

where e_1 is the first vector of the canonical basis of $\mathbb{C}^r$, $\Upsilon = e_1$, $\chi = \left((a\delta_1^2)^2 + \sum_{k=1}^K (b_k \delta_{k+1})^2 \right)^{1/2}$ and $\tilde{\Upsilon}$ is the unit norm vector $\tilde{\Upsilon} = \frac{1}{\chi}(a\delta_1^2, b_1\delta_2, \ldots, b_K\delta_{K+1})^T$. χ thus represents the non zero singular value of rank 1 matrix $R_{f|p,N}$. Using (II.38), we obtain immediately that matrix Γ_* coincides with $\lim_{N\to+\infty} \Gamma_N$ where Γ_N is defined by (II.37). As $R_N = \sigma^2 I_M$, and that $H_*(\sqrt{x_{+,*}}) = \lim_{N\to+\infty} H_N(\sqrt{x_{+,N}})$, it is easy to check using (II.87) that s coincides with the number of strictly negative eigenvalues of matrix $\begin{pmatrix} G_* & \Gamma_*^* \\ \Gamma_* & G_* \end{pmatrix}$ where G_* is defined by

$$G_* = \left(\frac{\sigma^2 c_*}{w_{+,*} - \sigma^2(1-c_*)} \right)^{1/2} \left((w_{+,*} - \sigma^2) \, I_r - \Delta^2 \right)^{-1}$$

Here, $w_{+,*} = \lim_{N\to+\infty} w_{+,N}$ is equal to

$$w_{+,*} = \sigma^2 \left(1 + \frac{1 + \sqrt{1 + 8c_*}}{2} \right)$$

(see the expression of $w_{+,N}$, Eq. (7-54) in [33]). It is easily checked that all the previous required Assumptions are verified by the present model. We now indicate how it is possible to choose the various parameters in order that s coincides with $2r-1$. We first fix parameters $(\delta_k)_{k=1,\ldots,r}$ in such a way that $\delta_1^2 \geq \delta_2^2 \geq \ldots \geq \delta_r^2$ and $\delta_k^2 > (w_{+,*} - \sigma^2)$ for each $k = 1, \ldots, r$ while we consider in the following $(b_k)_{k=1,\ldots,K}$ verifying

$$1 - \frac{\sum_{k=1}^K b_k^2}{\delta_1^2} > 0$$

and choose

$$a = \left(1 - \frac{\sum_{k=1}^K b_k^2}{\delta_1^2} \right)^{1/2} \tag{II.96}$$

Therefore, we of course have $\delta_1^2 = \mathbb{E}(|x_n|^2)$. We claim that with these set of parameters, $s = 2r-1$. For this, we first remark that matrix $G_* < 0$. As $\mathrm{Rank}(\Gamma_*) = 1$, the number s of stricty negative eigenvalues of $\begin{pmatrix} G_* & \Gamma_*^* \\ \Gamma_* & G_* \end{pmatrix}$ is equal to $2r-1$ or to $2r$. Using the Schur complement trick twice, we obtain after some algebra that $s = 2r-1$ if and only

$$-\chi^2 \tilde{\Upsilon}^* G_*^{-1} \tilde{\Upsilon} > -\left(\Upsilon^* G_*^{-1} \Upsilon \right)^{-1}$$

a condition equivalent to

$$a^2 > \frac{\sigma^2 c_*}{\sigma^2 c_* + w_{+,*} - \sigma^2} \left(1 - \frac{w_{+,*} - \sigma^2}{\delta_1^2} \right)^2 - \sum_{k=1}^{K} \frac{b_k^2}{\delta_1^2} \left(\frac{1 - \frac{w_{+,*}-\sigma^2}{\delta_1^2}}{1 - \frac{w_{+,*}-\sigma^2}{\delta_{k+1}^2}} \right) \tag{II.97}$$

Using that $a^2 = 1 - \frac{\sum_{k=1}^K b_k^2}{\delta_1^2}$ as well as $\delta_1^2 \geq \delta_{k+1}^2$ for each $k = 1, \ldots, K$, we obtain that the above condition holds, and, therefore, that $s = 2r-1$. We also mention that the condition $\delta_k^2 > w_{+,*} - \sigma^2$ for each $k = 1, \ldots, r$ does not induce any power limitation on the useful signal u. Moreover, using $a^2 = 1 - \frac{\sum_{k=1}^K b_k^2}{\delta_1^2}$, we obtain that the non zero singular value χ of $R_{f|p,N}$ is given by

$$\chi^2 = \delta_1^4 \left(1 - \frac{\sum_{k=1} b_k^2}{\delta_1^2} \right) + \sum_{k=1}^{K} b_k^2 \delta_{k+1}^2$$

As the $(\delta_k^2)_{k=1,\ldots,r}$ can take any large values, the same property holds for χ. In sum, even for powerful enough signals u for which the largest singular value of $R_{f|p,N}$ is large, s may be strictly larger than P.

We illustrate the above analysis by numerical simulations in which $N = 1200$ and $c_N = \frac{1}{2}$. The parameters of model (II.95) are chosen as above by replacing c_* by c_N. Figures 1 and 2 plot an histogram of the eigenvalues of a realization of matrix $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$, as well as the graph of the density g_N of the deterministic equivalent measure ν_N of the empirical eigenvalue $\hat{\nu}_N$ of $W_{f,N}W_{p,N}^*W_{p,N}W_{f,N}^*$. In the context of Fig. 1, $r = 2$ and it is seen that $s = 3$ eigenvalues of $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ escape from the support of ν_N. In the context of Fig. 2, $r = 3$ and $s = 5$ as expected. We mention that in both figures, the largest eigenvalue, which, in some sense, is due to the useful signal, appears much larger than the other spurious ones. It can be checked that, as expected, for smaller values of c_N, the spurious eigenvalues that escape from the support of ν_N tend to become closer from $x_{+,N}$. This will be confirmed in Section IV where more exhaustive Monte Carlo simulation results evaluate the behaviour of two estimates of s when $s = 5$ and $c_N = \frac{1}{4}$. It will be seen that the estimates of s belong to $\{2, 3, 4, 5, 6, 7, 8\}$, fail to detect $s = 5$ very often, but never take the value 1.

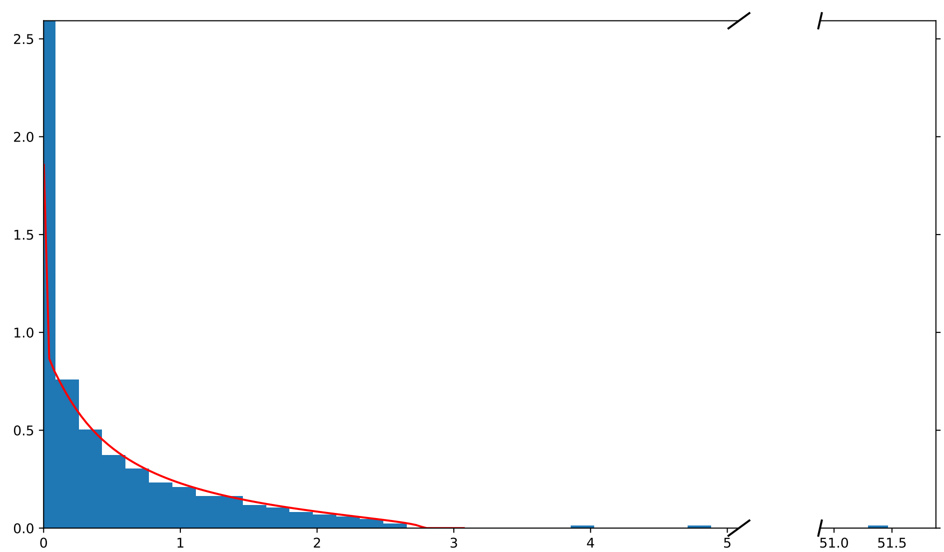

Figure 1. Histogram of the eigenvalues and graph of g_N, $r = 2$, $s = 3$

The above examples show that s can take any odd value larger than 3. We finally show that s can also be equal to 2, and consider the following simple case. We assume that $P = K = 1$, and that the scalar state-space sequence $(x_n)_{n\in\mathbb{Z}}$ is given by $x_{n+1} = ax_n + bi_n$ where $a \in]0, 1[$, $b > 0$, and $(i_n)_{n\in\mathbb{Z}}$ is scalar unit variance i.i.d. sequence. Moreover, u_n is given by

$$u_n = \theta_N x_{n+1} = a\theta_N x_n + b\theta_N i_n \tag{II.98}$$

where θ_N is a unit norm M–dimensional vector. Therefore, matrices C_N and D_N coincide with vectors $a\theta_N$ and $b\theta_N$ respectively. We also consider the case where $L = 1$. The covariance matrix $R_{u,N} = E(u_n u_n^*)$ is of course equal to $\delta^2\, \theta_N \theta_N^*$ where $\delta^2 = \mathbb{E}(|x_n|^2) = \frac{|b|^2}{1-a^2}$, so that $r = P = 1$. We also mention that in the present case, δ^2 does not depend on N. Moreover, $R_{f|p,N} = E(x_{n+1}x_n^*) = a\delta^2\theta_N\theta_N^*$. Therefore, matrix Γ_* is reduced to the scalar $a\delta^2$, which also coincides with the non zero singular value χ of $R_{f|p,N}$. As $r = P = 1$, s may take the values $0, 1, 2$. In the following, we justify that it is possible to find a and b for which $s = 2$.

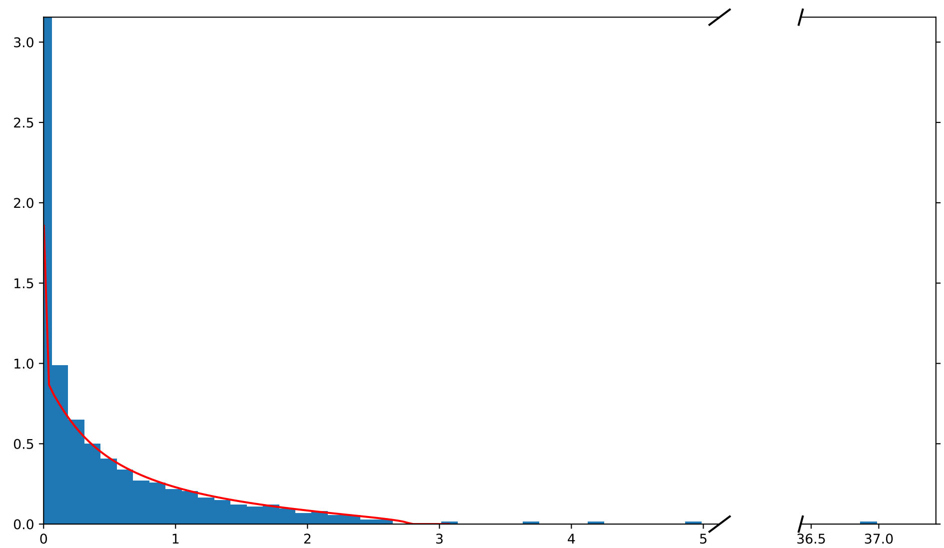

Figure 2. Histogram of the eigenvalues and graph of g_N, $r = 3$, $s = 5$

It is easily seen that $s = 2$ if $\delta^2 > w_{+,*} - \sigma^2$ and

$$a^2 < \frac{\sigma^2 c_*}{\sigma^2 c_* + (w_{+,*} - \sigma^2)} \left(1 - \frac{(w_{+,*} - \sigma^2)}{\delta^2}\right)^2 \tag{II.99}$$

In order to find $a \in]0,1[$ and b for which these conditions hold, we fix $\delta^2 > w_{+,*} - \sigma^2$, then choose $a \in]0,1[$ such that (II.99) holds, and finally select b in such a way that $|b|^2 = \delta^2(1 - a^2)$. We again mention that δ^2 and $\chi = a\delta^2$ can take arbitrarily large values, χ being however less than $\delta^2 \left(\frac{\sigma^2 c_*}{\sigma^2 c_* + (w_{+,*} - \sigma^2)}\right)^{1/2} \left(1 - \frac{(w_{+,*} - \sigma^2}{\delta^2}\right)$.

We again illustrate the above example by Fig. 3 obtained when $N = 1200$, $c_N = \frac{1}{2}$, and where it is seen that $s = 2$.

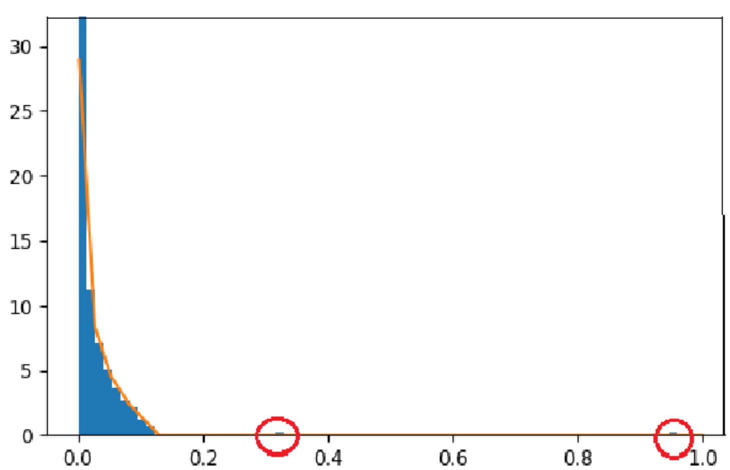

Figure 3. Histogram of the eigenvalues and graph of g_N, $r = 1$, $s = 2$

III. The canonical correlation coefficients between the past and the future

We showed in Section II that the number of eigenvalues of $\hat{R}_{f|p}\hat{R}_{f|p}^*$ that escape from the interval $[0, x_{+,N}]$ is in general not a consistent estimator of the dimension P of the minimal state space representation (I.3). In this section, we thus study the largest singular values of matrix $(\frac{Y_f Y_f^*}{N})^{-1/2}\frac{Y_f Y_p^*}{N}(\frac{Y_p Y_p^*}{N})^{-1/2}$, or equivalently the largest eigenvalues of matrix $(\frac{Y_f Y_f^*}{N})^{-1/2}\frac{Y_f Y_p^*}{N}(\frac{Y_p Y_p^*}{N})^{-1}\frac{Y_p Y_f^*}{N}(\frac{Y_f Y_f^*}{N})^{-1/2}$. It is clear that apart 0, the eigenvalues of the above matrix coincide with the eigenvalues of matrix $\Pi_{p,y}\Pi_{f,y}$ where for each $i = p, f$, $\Pi_{i,y}$ represents the orthogonal projection matrix on the row space of matrix Y_i, i.e.

$$\Pi_{i,y} = \frac{Y_i^*}{\sqrt{N}}\left(\frac{Y_i Y_i^*}{N}\right)^{-1}\frac{Y_i}{\sqrt{N}} \tag{III.1}$$

We remark that the eigenvalues of $\Pi_{p,y}\Pi_{f,y}$ of course belong to $[0,1]$. We follow the same approach than in Section II. We first study the eigenvalues of $\Pi_{p,v}\Pi_{f,v}$, where $\Pi_{i,v}$ is obtained from $\Pi_{i,y}$ by replacing y by the noise v. Under certain assumptions on the useful signal u (that appear simpler than in Section II), we study the largest eigenvalues of $\Pi_{p,y}\Pi_{f,y}$ by remarking that $\Pi_{p,y}\Pi_{f,y}$ is a low rank perturbation of $\Pi_{p,v}\Pi_{f,v}$, and use the approach developed in [6], [7], [40]. We again mention that, while this general approach appears classical, as in Section II, the complexity of the random matrix models that come into play makes the following results not obvious at all.

In the following, for the sake of simplicity, we will often use the same notations as in Section II to represent different objects. This will not introduce any confusion because Section III and Section II are independent. In particular, if $(\alpha_N)_{N\geq 1}$ is a sequence of positive numbers, we will say in this section that function $f_N(z) = \mathcal{O}_z(\alpha_N)$ on a domain $\Omega \subset \mathbb{C}\setminus\mathbb{R}^+$ if there exists two nice polynomials P_1 and P_2 such that $|f_N(z)| \leq \alpha_N P_1(|z|)P_2(\frac{1}{\rho(z)})$ for each $z \in \Omega$, where $\rho(z) = \mathrm{dist}(z, \mathbb{R}^+)$. If $\Omega = \mathbb{C}\setminus\mathbb{R}^+$, we will just write $f_N(z) = \mathcal{O}_z(\alpha_N)$ without mentioning the domain. For any diagonal $K\times K$ matrix $A(z)$, by $A(z) = \mathcal{O}_z^K(\alpha_N)$, we mean that each diagonal element of $A(z)$ is $\mathcal{O}_z(\alpha_N)$. Finally, we will use a lot the notation $f_N(z) = \mathcal{O}_{z^2}(\alpha_N)$ without mentioning the domain, which will mean that $|f_N(z)| \leq \alpha_N P_1(|z^2|)P_2(\frac{1}{\rho(z^2)})$ for some nice polynomials P_1, P_2 when $z^2 \in \mathbb{C}\setminus\mathbb{R}^+$, or equivalently, when $z \in \mathbb{C}\setminus\mathbb{R}$. We notice that if P_1, P_2 and Q_1, Q_2 are nice polynomials, then $P_1(|z|)P_2(\frac{1}{\rho(z)}) + Q_1(|z|)Q_2(\frac{1}{\rho(z)}) \leq (P_1+Q_1)(|z|)(P_2+Q_2)(\frac{1}{\rho(z)})$, from which we conclude that if functions f_1 and f_2 are $\mathcal{O}_z(\alpha_N)$ then also $f_1(z)+f_2(z) = \mathcal{O}_z(\alpha_N)$.

A. In the absence of signal

In this paragraph, we study the behaviour of the eigenvalues of $\Pi_{p,v}\Pi_{f,v}$. Due to the Gaussianity of the i.i.d. vectors $(v_n)_{n\geq 1}$, it exists i.i.d. $\mathcal{N}_c(0, I_M)$ distributed vectors $(v_{iid,n})_{n\geq 1}$ such that $\mathbb{E}(v_{iid,n}v_{iid,n}^*) = I_M$ verifying $v_n = R_N^{1/2}v_{iid,n}$. It is clear that the row spaces of V_p and V_f coincide with the row spaces of the block Hankel matrices $V_{p,iid}$ and $V_{f,iid}$ defined from vectors $(v_{n,iid})_{n=1,\ldots,N+2L-1}$. Therefore, the projection matrices $\Pi_{i,v}$ and $\Pi_{i,v_{iid}}$ coincide for $i = p, f$ and there is thus no restriction to assume in Section III-A that $R_N = I_M$.

As before, we denote by W_p, W_f the matrices defined by $W_p = \frac{1}{\sqrt{N}}V_p$ and $W_f = \frac{1}{\sqrt{N}}V_f$. In order to simplify the notations of this Section, matrices $\Pi_{i,v}$, $i = p, f$ are denoted Π_i, $i = p, f$. Therefore, we have $\Pi_p = W_p^*(W_p W_p^*)^{-1}W_p$ and $\Pi_f = W_f^*(W_f W_f^*)^{-1}W_f$. We recall that W_N is the $2ML\times N$ matrix

$$W_N = \begin{pmatrix} W_{p,N} \\ W_{f,N} \end{pmatrix}, \tag{III.2}$$

As R_N is supposed to be equal to I_M, the elements $(W_{i,j}^m)_{i\leq 2L, j\leq N, m\leq M}$ of W_N satisfy

$$\mathbb{E}\{W_{i,j}^m W_{i',j'}^{m'}\} = \frac{1}{N}\,\delta_{m-m'}\delta_{i+j-(i'+j')}. \tag{III.3}$$

where $W_{i,j}^m$ represents the element which lies on the $(m+M(i-1))$-th line and j-th column for $1\leq m\leq M$, $1\leq i\leq 2L$ and $1\leq j\leq N$. 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.

We first verify that, almost surely, for N large enough, matrices $W_{i,N}W_{i,N}^*$ are invertible, so that the orthogonal projection matrices Π_i, $i = p, f$ are well defined. For this, we mention that [32] (see Theorem 1.1) established that the empirical eigenvalue distribution of $W_{i,N}W_{i,N}^*$ for $i = \{p, f\}$ converges towards the Marcenko-Pastur distribution with parameter c_*, and that almost surely, for N greater than a random integer, its eigenvalues are located in a neighbourhood of $[(1-\sqrt{c_*})^2, (1+\sqrt{c_*})^2]$. Therefore, almost surely, for N large enough, matrices $W_{f,N}W_{f,N}^*$ and $W_{p,N}W_{p,N}^*$ are invertible. Matrices $\Pi_{i,N}$ are thus well defined for N large enough.

We next use again the results of [32] to show the following Lemma which will be useful to establish Theorem III.1 below.

Lemma III.1. *If $c_* > \frac{1}{2}$, then, almost surely, for N large enough, 1 is eigenvalue of $\Pi_{p,N}\Pi_{f,N}$ with multiplicity $2ML - N$*

Proof. It is clear that the eigenspace of $\Pi_{p,N}\Pi_{f,N}$ associated to the eigenvalue 1 coincides with $\mathrm{sp}_r(W_{p,N}) \cap \mathrm{sp}_r(W_{f,N})$, where for a matrix A, $\mathrm{sp}_r(A)$ represents the space generated by the rows of A. We have thus to verify that if $c_* > 1/2$, then almost surely, for N large enough, $\dim\left(\mathrm{sp}_r(W_{p,N}) \cap \mathrm{sp}_r(W_{f,N})\right) = 2ML - N$. For this, we use again [32]. The eigenvalue distribution of $W_N W_N^*$ converges towards the Marcenko-Pastur distribution with parameter $2c_*$, and if $c_* > \frac{1}{2}$, i.e. if $2c_* > 1$, then, for each $\epsilon > 0$, 0 is eigenvalue of $W_N W_N^*$ with multiplicity $2ML - N$ and the remaining N eigenvalues are located almost surely for each N large enough in $[(1-\sqrt{2c_*})^2 - \epsilon, (1+\sqrt{2c_*})^2 + \epsilon]$. Therefore, we obtain that $\dim(\mathrm{sp}_r(W_N)) = N$ while we already know that $\dim(\mathrm{sp}_r(W_{p,N})) + \dim(\mathrm{sp}_r(W_{f,N})) = 2ML$. As $\mathrm{sp}_r(W_N) = \mathrm{sp}_r(W_{p,N}) + \mathrm{sp}_r(W_{f,N})$, we obtain as expected that $\dim\left(\mathrm{sp}_r(W_{p,N}) \cap \mathrm{sp}_r(W_{f,N})\right) = 2ML - N$. ■.

1) Preliminary results: In order to be able to use the perturbation approach developed in [6], [7], [40], it appears necessary to evaluate the asymptotic behaviour of the resolvent of matrix $\Pi_p\Pi_f$. The corresponding results will also provide a characterization of the eigenvalues of $\Pi_p\Pi_f$. For this, we use in the following the integration by parts formula and the Poincaré-Nash inequality (see Propositions I.1, I.2). The resolvent of $\Pi_p\Pi_f$ will be interpreted as a function of the entries of matrix W_N. However, this approach needs some care because, considered as a function of the entries of W_N, matrices Π_p and Π_f are not differentiable everywhere. In particular, for $i = p, f$, Π_i is not differentiable when the rank of $W_{i,N}$ is less than ML. But, we have seen that almost surely, for N large enough, matrices $W_f W_f^*$ and $W_p W_p^*$ are invertible. In order to take benefit of this property, we use in the following a regularization term η_N already introduced in [23] in a different context. Another problem posed by the evaluation of the resolvent of $\Pi_p\Pi_f$ is due to the observation that, while matrix $\Pi_p\Pi_f$ has real eigenvalues that belong to $[0,1]$, it is not Hermitian. Some basic properties of the resolvent of $\Pi_p\Pi_f$ thus do not hold, in particular the upper bound (I.22). In this paragraph, we first present the regularization term η_N as well as some extra useful properties.

a) Regularization term: We define η_N by

$$\eta_N = \det\,[\phi(W_{f,N}W_{f,N}^*)]\det\,[\phi(W_{p,N}W_{p,N}^*)], \tag{III.4}$$

where ϕ is a smooth function such that

$$\begin{aligned} \phi(\lambda) &= 1 \text{ for } \lambda \in [(1-\sqrt{c_*})^2 - \epsilon],\ [(1+\sqrt{c_*})^2 + \epsilon], \\ \phi(\lambda) &= 0 \text{ for } \lambda \in [-\infty,\ (1-\sqrt{c_*})^2 - 2\epsilon] \cup [(1+\sqrt{c_*})^2 + 2\epsilon,\ +\infty] \end{aligned} \tag{III.5}$$

and $\phi(\lambda) \in (0,1)$ elsewhere. Here, ϵ verifies $(1-\sqrt{c_*})^2 - 2\epsilon > 0$. Taking into account the almost sure behaviour of the eigenvalues of matrices $W_p W_p^*$ and $W_f W_f^*$, $\eta_N = 1$ and

$$(W_{i,N}W_{i,N}^*)^{-1}\eta_N \leq \frac{I_{ML}}{(1-\sqrt{c_*})^2 - 2\epsilon}. \tag{III.6}$$

almost surely for each N larger than a random integer. We first mention the following useful property.

Lemma III.2. *For each $l, k \in \mathbb{N}$ it holds that*

$$\mathbb{E}\{\eta_N^l\} = 1 + \mathcal{O}\left(\frac{1}{N^k}\right) \tag{III.7}$$

Moreover, if X is a bounded random variable, we have for each integer $l \geq 1$

$$\mathbb{E}(\eta^l X) = \mathbb{E}(X) + \mathcal{O}\left(\frac{1}{N^k}\right) \tag{III.8}$$

for each integer k.

Proof. Denote

$$\mathcal{E}_N = \{\text{one of the eigenvalues of } W_pW_p^* \text{ or } W_fW_f^* \text{ escapes from the } [(1-\sqrt{c_*})^2-\epsilon,\, (1+\sqrt{c_*})^2+\epsilon]\} \quad \text{(III.9)}$$

and define another smooth function ϕ_0 as

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

and $\phi_0(\lambda) \in (0,\, 1)$ elsewhere. Then we have

$$P(\mathcal{E}_N) \leq P\left(\mathrm{Tr}\phi_0(W_pW_p^*) \geq 1\right) \leq \mathbb{E}\left\{\left(\mathrm{Tr}\phi_0(W_pW_p^*)\right)^{2k}\right\}$$

for all $k \in \mathbb{N}$. In order to evaluate $\mathbb{E}\left\{\left(\mathrm{Tr}\phi_0(W_pW_p^*)\right)^{2k}\right\}$, one can use the same steps as in the proof of Lemma 3.2 [33] and get immediately that $\mathbb{E}\left\{\left(\mathrm{Tr}\phi_0(W_pW_p^*)\right)^{2k}\right\} = \mathcal{O}\left(\frac{1}{N^{2k}}\right)$ and therefore that $P(\mathcal{E}_N) = \mathcal{O}\left(\frac{1}{N^{2k}}\right)$ for each k. To show (III.7) we write

$$\begin{aligned} |\mathbb{E}\{\eta_N^l - 1\}|^2 = |\mathbb{E}\{(\eta_N-1)(1+\ldots+\eta_N^{l-1})\}|^2 \leq \mathbb{E}\{(\eta_N-1)^2\}\mathbb{E}\{(1+\ldots+\eta_N^{l-1})^2\} \\ \leq \kappa\mathbb{E}\{(\eta_N-1)^2\mathbf{1}_{\mathcal{E}_N}\} \end{aligned}$$

because $\eta_N - 1 = 0$ on $\mathcal{E}_N^c$. Since by definition $\phi(\lambda) \in [0,1]$, we conclude that $0 \leq \eta_N \leq 1$ and $0 \leq (\eta_N-1)^2 \leq 1$. This allows us to write that $\kappa\mathbb{E}\{(\eta_N-1)^2\mathbf{1}_{\mathcal{E}_N}\} \leq \kappa\mathbb{E}\{\mathbf{1}_{\mathcal{E}_N}\} = \kappa P(\mathcal{E}_N) = \mathcal{O}\left(\frac{1}{N^{2k}}\right)$, which completes the proof. To verify (III.8), we remark that

$$|\mathbb{E}\{(\eta_N^l-1)X\}|^2 \leq \mathbb{E}\{(1-\eta_N^l)^2\}\mathbb{E}\left\{|X|^2\right\} = \kappa\left(1-2\left(1+\mathcal{O}(N^{-k})\right)\right)+1+\mathcal{O}(N^{-k})) = \mathcal{O}\left(\frac{1}{N^k}\right).$$

■

b) Linearisation: It is clear that almost surely, $\eta_N\Pi_{i,N} = \Pi_{i,N}$ for each N large enough. Therefore, in order to evaluate the almost sure behaviour of the resolvent of $\Pi_{p,N}\Pi_{f,N}$, it is sufficient to study the behaviour of the resolvent $Q_N(z)$ of $\eta_N\Pi_{p,N}\eta_N\Pi_{f,N}$ defined by

$$Q_N(z) = (\eta_N\Pi_{p,N}\eta_N\Pi_{f,N} - zI)^{-1}$$

As the direct study of $Q_N(z)$ is not obvious, we rather use, as in Section II, the linearisation trick and introduce the resolvent $\mathbf{Q}_N(z)$ of the $2N \times 2N$ block matrix

$$\begin{pmatrix} 0 & \eta_N\Pi_{p,N} \\ \eta_N\Pi_{f,N} & 0 \end{pmatrix}.$$

which can be written as

$$\mathbf{Q}_N(z) = \begin{pmatrix} (\mathbf{Q_{pp}})_N(z) & (\mathbf{Q_{pf}})_N(z) \\ (\mathbf{Q_{fp}})_N(z) & (\mathbf{Q_{ff}})_N(z) \end{pmatrix} = \begin{pmatrix} zQ_N(z^2) & Q_N(z^2)\eta_N\Pi_{p,N} \\ \eta_N\Pi_{f,N}Q_N(z^2) & z\hat{Q}_N(z^2) \end{pmatrix} \quad \text{(III.10)}$$

where $\hat{Q}_N(z)$ is the resolvent of matrix $\eta_N\Pi_{f,N}\eta_N\Pi_{p,N}$. Since $Q_N(z)$ and $\mathbf{Q}_N(z)$ are resolvents of non Hermitian matrices, the usual bound (I.22) is not necessarily verified. A more specific control is thus needed.

Lemma III.3. *If* $\mathrm{Im}z \neq 0$ *(i.e.* $z^2 \in \mathbb{C} \setminus \mathbb{R}^+$*), then* $\|\mathbf{Q}(z)\| = \mathcal{O}_{z^2}(1)$.

Proof. It is sufficient to bound each of the four blocks of $\mathbf{Q}$. We start with $\mathbf{Q_{pf}}$. For this we use expression (III.10) for $\mathbf{Q_{pf}}$, the fact that $\Pi_p = \Pi_p^2$ and that $(AB-x)^{-1}A = A(BA-x)^{-1}$ in the case $A = \eta\Pi_p, B = \eta\Pi_p\Pi_f$. This leads to

$$\mathbf{Q_{pf}} = (\eta_N^2\Pi_p\Pi_f - z^2)^{-1}\eta_N\Pi_p\Pi_p = \eta_N\Pi_p(\eta_N^2\Pi_p\Pi_f\Pi_p - z^2)^{-1}\Pi_p. \quad \text{(III.11)}$$

$(\eta_N^2\Pi_p\Pi_f\Pi_p - z^2)^{-1}$ is the resolvent of a positive Hermitian matrix evaluated at $z^2 \in \mathbb{C}\setminus\mathbb{R}^+$, so that its norm can be bounded by $(\rho(z^2))^{-1}$ (see (I.22)). Since $\|\Pi_p\| \leq 1$ and $\eta_N \leq 1$, we have

$$\|\mathbf{Q_{pf}}\| \leq \frac{1}{\rho(z^2)} \tag{III.12}$$

It is easily seen that $\|\mathbf{Q_{fp}}\|$ can be evaluated similarly. In order to address $\mathbf{Q_{pp}}$, we use again (III.10) and the resolvent identity (I.23), and observe that:

$$\mathbf{Q_{pp}} = z(\eta_N^2\Pi_p\Pi_f - z^2)^{-1} = \frac{1}{z}(-I_N + \eta_N^2\Pi_p\Pi_f(\eta_N^2\Pi_p\Pi_f - z^2)^{-1}) = \frac{1}{z}(-I_N + \eta_N\Pi_p\mathbf{Q_{fp}})$$

It obviously holds that $\|-I_N+\eta_N\Pi_p\mathbf{Q_{fp}}\| \leq 1+\frac{1}{\rho(z^2)}$. To show that $|z^{-1}| \leq P(\rho(z^2)^{-1})$ for some nice polynomial P, we write

$$\frac{1}{|z|^2} \leq \frac{1}{\rho(z^2)} \leq 1 + \frac{1}{\rho(z^2)} \leq \left(1 + \frac{1}{\rho(z^2)}\right)^2 \tag{III.13}$$

This brings us to the conclusion that $\|\mathbf{Q_{pp}}\| = \mathcal{O}_{z^2}(1)$ and so for $\mathbf{Q_{ff}}$. This completes the proof of the Lemma. ■

Remark III.1. *It is worth to remark that in the course of the proof, we obtained that* $\frac{1}{|z|}\mathcal{O}_{z^2}(1)$ *is still* $\mathcal{O}_{z^2}(1)$*. Since* $|z| \leq \frac{1}{2}(1+|z|^2)$ *holds, we also have* $|z|\mathcal{O}_{z^2}(1) = \mathcal{O}_{z^2}(1)$*.*

Remark III.2. *While* $Q_N(z)$ *is not the resolvent of an Hermitian matrix,* $\frac{1}{N}\mathrm{Tr}Q_N(z)$ *coincides with the Stieltjes transform of the empirical eigenvalue distribution* $\hat{\nu}_N$ *of matrix* $\eta^2\Pi_p\Pi_f$*, which, of course, is a probability measure carried by* $[0,1]$*, and thus by* $\mathbb{R}^+$*. Therefore, (III.10) and property (I.18) imply* $\frac{1}{N}\mathrm{Tr}\,\mathbf{Q_{pp}}(z) = \frac{1}{N}\mathrm{Tr}\,\mathbf{Q_{ff}}(z)$ *coincide with the Stieltjes transform of a probability measure carried by* $[-1,1]$ *which appears to be the eigenvalue distribution of matrix* $\begin{pmatrix} 0 & \eta_N\Pi_{p,N} \\ \eta_N\Pi_{f,N} & 0 \end{pmatrix}$

The proof of Lemma III.3 also leads to the following useful Corollary.

Corollary III.1. $N^{-1}\mathrm{Tr}\mathbf{Q_{pf}}(z)$ *and* $N^{-1}\mathrm{Tr}\mathbf{Q_{fp}}(z)$ *coincide with the value taken at* z^2 *by the Stieltjes transforms of some positive measures carried by* $\mathbb{R}^+$*. The same property holds for* $\mathbb{E}\{N^{-1}\mathrm{Tr}\mathbf{Q_{pf}}(z)\}$ *and* $\mathbb{E}\{N^{-1}\mathrm{Tr}\mathbf{Q_{fp}}(z)\}$*, and the mass of the corresponding measures can be written as* $c_N + \mathcal{O}(N^{-k})$ *for each* $k \in \mathbb{N}$*.*

Proof. We just establish the properties of $N^{-1}\mathrm{Tr}\mathbf{Q_{pf}}(z)$. $(\eta_N^2\Pi_p\Pi_f\Pi_p - z^2)^{-1}$ is the resolvent of a positive Hermitian matrix evaluated at point z^2. Therefore, $N^{-1}\mathrm{Tr}\eta_N\Pi_p(\eta_N^2\Pi_p\Pi_f\Pi_p - z^2)^{-1}\Pi_p = N^{-1}\mathrm{Tr}\mathbf{Q_{pf}}(z)$ (see Eq. III.11)) coincides with the Stieltjes transform of a positive measure carried by $\mathbb{R}^+$ of total mass $N^{-1}\mathrm{Tr}\eta_N\Pi_p^2 = N^{-1}\mathrm{Tr}\eta_N\Pi_p$ evaluated at z^2. This implies that $N^{-1}\mathbb{E}\{\mathrm{Tr}\mathbf{Q_{pf}}\}$ has the same property, and that the mass of the corresponding measure is equal to $N^{-1}\mathbb{E}\{\mathrm{Tr}\eta_N\Pi_p\}$. We claim that

$$N^{-1}\mathbb{E}\{\mathrm{Tr}\eta_N\Pi_p\} = c_N + \mathcal{O}(N^{-k}) \tag{III.14}$$

for each integer k. To justify (III.14), we first use (III.8) and obtain that $N^{-1}\mathbb{E}\{\mathrm{Tr}\eta_N\Pi_p\} = N^{-1}\mathbb{E}\{\mathrm{Tr}\Pi_p\} + \mathcal{O}(N^{-k})$. It is clear that $N^{-1}\mathrm{Tr}\,\Pi_p = c_N$ on the event $\mathcal{E}_N^c$, where we recall that $\mathcal{E}_N$ is the set defined by (III.9). Writing

$$N^{-1}\mathbb{E}\{\mathrm{Tr}\Pi_p\} = N^{-1}\mathbb{E}\{\mathrm{Tr}\Pi_p\mathbf{1}_{\mathcal{E}_N^c}\} + N^{-1}\mathbb{E}\{\mathrm{Tr}\Pi_p\mathbf{1}_{\mathcal{E}_N}\} = c_N P(\mathcal{E}_N^c) + N^{-1}\mathbb{E}\{\mathrm{Tr}\Pi_p\mathbf{1}_{\mathcal{E}_N}\}$$

and using that $P(\mathcal{E}_N) = \mathcal{O}(N^{-k})$ for each k, we obtain that $N^{-1}\mathbb{E}\{\mathrm{Tr}\Pi_p\mathbf{1}_{\mathcal{E}_N}\} = \mathcal{O}(N^{-k})$ for each k and that $N^{-1}\mathbb{E}\{\mathrm{Tr}\Pi_p\} = c_N + \mathcal{O}(N^{-k})$. This completes the proof of (III.14). ■

c) Properties based on the invariance of the complex Gaussian distribution :

Lemma III.4. *The matrix* $\mathbb{E}\{\eta_N(W_iW_i^*)^{-1}\}$ *is block diagonal and matrices* $\mathbb{E}\{\eta_N\Pi_i\}$*,* $\mathbb{E}\{\mathbf{Q_{ij}}\}$*,* $\mathbb{E}\{\eta_N\mathbf{Q_{ij}}\}$*,* $\mathbb{E}\{\eta_N\Pi_h\mathbf{Q_{ij}}\}$ *and* $\mathbb{E}\{\eta_N\mathbf{Q_{ij}}W_h^*(W_hW_h^*)^{-2}W_h\}$ *are diagonal, for* $i,j,h = \{p,f\}$*. Moreover,*

$\mathbb{E}(\eta_N\, W_h^*(W_hW_h^*)^{-1}) = \mathbb{E}(\eta_N\, \mathbf{Q_{ij}}\, W_h^*(W_hW_h^*)^{-1}) = \mathbb{E}(\eta_N\, \Pi_k\, \mathbf{Q_{ij}}\, W_h^*(W_hW_h^*)^{-1}) = 0$ *for* $i, j, h, k = \{p, f\}$. *Finally, if* $i, j, h = \{p, f\}$,*for each* $n = 1, \ldots, N$, *we have*

$$\mathbb{E}\{(\mathbf{Q_{ij}})^{n,n}\} = \mathbb{E}\{(\mathbf{Q_{\tilde{i}\tilde{j}}})^{N+1-n,N+1-n}\} \tag{III.15}$$

$$\mathbb{E}\{\eta_N(\Pi_h\mathbf{Q_{ij}})^{n,n}\} = \mathbb{E}\{\eta_N(\Pi_{\tilde{h}}\mathbf{Q_{\tilde{i}\tilde{j}}})^{N+1-n,N+1-n}\} \tag{III.16}$$

$$\mathrm{Tr}\mathbb{E}\{\mathbf{Q_{ij}}\} = \mathrm{Tr}\mathbb{E}\{\mathbf{Q_{\tilde{i}\tilde{j}}}\}, \tag{III.17}$$

$$\mathrm{Tr}\mathbb{E}\{\eta_N\Pi_h\mathbf{Q_{ij}}\} = \mathrm{Tr}\mathbb{E}\{\eta_N\Pi_{\tilde{h}}\mathbf{Q_{\tilde{i}\tilde{j}}}\}, \tag{III.18}$$

where " ˜ " changes index to opposite: $p \to f, f \to p$.

The proof is postponed to the Appendix. To establish the first statements of the Lemma, we remark that for each θ, the probability distribution of $(v_n)_{n\in\mathbb{Z}}$ coincides with the probability distribution of $(z_n)_{n\in\mathbb{Z}}$ where z is chosen as $z_n = v_n e^{-in\theta}$ for each n. We use the same trick when $z_n = v_{-n+N+2L}$ for each n to prove (III.15) –(III.18).

We now establish that the diagonal matrices $\mathbb{E}\{\eta_N(W_iW_i^*)^{-1}\}$ and $\mathbb{E}\{\eta_N\Pi_i\}$ are multiples of the identity matrix up to error terms.

Lemma III.5. *For* $i = \{p, f\}$, *we have:*

$$\mathbb{E}\{\eta_N(W_iW_i^*)^{-1}\} = \frac{1}{1-c_N}I_{ML} + \mathcal{O}^{ML}\left(\frac{1}{N^{3/2}}\right) \tag{III.19}$$

$$\mathbb{E}\{\eta_N\Pi_i\} = c_N I_N + \mathcal{O}^N\left(\frac{1}{N^{3/2}}\right). \tag{III.20}$$

Moreover, $(ML)^{-1}\mathrm{Tr}\mathbb{E}\{\eta_N(W_iW_i^*)^{-1}\} = (1-c_N)^{-1} + \mathcal{O}(\frac{1}{N^2})$.

The proof of Lemma III.5 uses the integration by parts formula and the Poincaré-Nash inequality, and is provided in the Appendix.

2) Expression of matrix $\mathbb{E}\{\mathbf{Q}\}$ *obtained using the integration by parts formula:* We now establish that matrices $\mathbf{Q_{ij}}$ are, up to error terms, multiples of I_N, and characterize the asymptotic behaviour of their common diagonal terms. For this, we state the following Proposition that is proved by using the integration by parts formula and the Poincaré-Nash inequality.

Proposition III.1. *The following equalities hold for each* $z \in \mathbb{C}^+$.

$$\mathbb{E}\Big\{\mathbf{Q_{pp}}\eta\Pi_p\Big\} = c_N\mathbb{E}\Big\{\mathbf{Q_{pp}}\Big\} - (1-c_N)\mathbb{E}\Big\{\eta\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-2}W_p\Big\}\frac{1}{N}\mathbb{E}\Big\{\mathrm{Tr}\left(\eta\Pi_p^\perp\mathbf{Q_{fp}}\right)\Big\} + \mathbf{\Delta_{pp}} \tag{III.21}$$

$$\mathbb{E}\{\mathbf{Q_{pf}}\eta\Pi_p\} = c_N\mathbb{E}\Big\{\mathbf{Q_{pf}}\Big\} - (1-c_N)\mathbb{E}\Big\{\eta\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-2}W_p\Big\}\frac{1}{N}\mathbb{E}\Big\{\mathrm{Tr}\left(\eta\Pi_p^\perp\mathbf{Q_{ff}}\right)\Big\} + \mathbf{\Delta_{pf}} \tag{III.22}$$

$$\mathbb{E}\Big\{\mathbf{Q_{pp}}\eta\Pi_f\Big\} = c_N\mathbb{E}\Big\{\mathbf{Q_{pp}}\Big\} - (1-c_N)\mathbb{E}\Big\{\eta\mathbf{Q_{pf}}W_f^*(W_fW_f^*)^{-2}W_f\Big\}\frac{1}{N}\mathbb{E}\Big\{\mathrm{Tr}\left(\eta\Pi_f^\perp\mathbf{Q_{pp}}\right)\Big\} + \mathbf{\Delta^1_{pp}} \tag{III.23}$$

$$\mathbb{E}\Big\{\mathbf{Q_{pf}}\eta\Pi_f\Big\} = c_N\mathbb{E}\Big\{\mathbf{Q_{pf}}\Big\} - (1-c_N)\mathbb{E}\Big\{\eta\mathbf{Q_{pf}}W_f^*(W_fW_f^*)^{-2}W_f\Big\}\frac{1}{N}\mathbb{E}\Big\{\mathrm{Tr}\left(\eta\Pi_f^\perp\mathbf{Q_{pf}}\right)\Big\} + \mathbf{\Delta^1_{pf}} \tag{III.24}$$

where matrices $\mathbf{\Delta_{pp}}, \mathbf{\Delta_{pf}}, \mathbf{\Delta^1_{pp}}, \mathbf{\Delta^1_{pf}}$ *are diagonal matrices whose entries are* $\mathcal{O}_{z^2}\left(N^{-3/2}\right)$ *terms, and whose normalized traces are* $\mathcal{O}_{z^2}\left(N^{-2}\right)$ *terms.*

(III.21) is proved in the Appendix. (III.22, III.23, III.24) are established similarly.

In order to introduce the next result, we denote by $w_N(z)$ the function defined by

$$w_N(z) = 1 + \frac{1}{N}\mathbb{E}\{\mathrm{Tr}\left(\eta_N\Pi_p^\perp\mathbf{Q_{fp}}\right)\} = 1 + \frac{1}{N}\mathbb{E}\{\mathrm{Tr}\left(\eta_N\Pi_f^\perp\mathbf{Q_{pf}}\right)\} \tag{III.25}$$

where the equality between the second term and the third term in (III.25) comes from (III.18). We claim that

$$\frac{1}{c_N^2 - z^2w_N^2} = \mathcal{O}_{z^2}(1) \tag{III.26}$$

To verify (III.26), we first notice that (III.11) implies that $\frac{1}{N}\mathbb{E}\{\mathrm{Tr}\,(\eta_N \Pi_f^\perp \mathbf{Q_{pf}} \Pi_f^\perp)\} = \frac{1}{N}\mathbb{E}\{\mathrm{Tr}\,(\eta_N \Pi_f^\perp \mathbf{Q_{pf}})\}$ (the equality follows from $\Pi_f^\perp = (\Pi_f^\perp)^2$) coincides with the value taken at point z^2 by the Stieltjes transform of a positive measure carried by $\mathbb{R}^+$. Proposition 5.1, item 4 in [21] thus implies that function $-\left(z^2 w_N(z)\right)^{-1}$ has the same property. Moreover, the converse of (I.16, I.17) in Proposition 4.1 in [33] leads to the conclusion that $-\left(z^2(w_N - \frac{c_N^2}{z^2 w_N})\right)^{-1}$ also coincides with the value taken at point z^2 by the Stieltjes transform of a positive measure carried by $\mathbb{R}^+$. Writing $|c_N^2 - z^2 w_N^2|^{-1}$ as

$$\left|\frac{1}{c_N^2 - z^2 w_N^2}\right| = \left|\frac{1}{z^2 w_N\left(-\frac{c_N^2}{z^2 w_N} + w_N\right)}\right| = \left|\frac{1}{z^2 w_N}\right| |z|^2 \left|\frac{1}{z^2(-\frac{c_N^2}{z^2 w_N} + w_N)}\right|$$

leads to (III.26). We are in position to precise the behaviour of the diagonal matrices $\mathbb{E}(\mathbf{Q_{ij}})$.

Proposition III.2. *For $i = p, f$, and for $i \neq j$, we have*

$$\mathbb{E}(\mathbf{Q_{ii}}(z)) = \frac{z w_N^2(z)}{c_N^2 - (z w_N(z))^2}\, I_N + \mathcal{O}_{z^2}^N\left(\frac{1}{N^{3/2}}\right) \tag{III.27}$$

$$\mathbb{E}(\mathbf{Q_{ij}}(z)) = \frac{c_N w_N(z)}{c_N^2 - (z w_N(z))^2}\, I_N + \mathcal{O}_{z^2}^N\left(\frac{1}{N^{3/2}}\right) \tag{III.28}$$

where the normalized traces of the $\mathcal{O}_{z^2}^N\left(\frac{1}{N^{3/2}}\right)$ error terms are $\mathcal{O}_{z^2}\left(\frac{1}{N^2}\right)$ terms.

Proof. We just establish (III.27) for $i = p$ and (III.28) for $i = p, j = f$ because, due to (III.15), (III.27) and (III.28) for $i = f$ and for $i = f, j = p$, respectively, are consequences of (III.27) for $i = p$ and (III.28) for $i = p, j = f$. We consider Proposition III.1, and begin by showing that the use of (III.21) and (III.22) allows to obtain the following relationship between $\mathbb{E}(\mathbf{Q_{pp}})$ and $\mathbb{E}(\mathbf{Q_{pf}})$

$$z\mathbb{E}(\mathbf{Q_{pf}}(z))\, w_N(z) = c\mathbb{E}(\mathbf{Q_{pp}}(z)) + \mathcal{O}_{z^2}^N\left(\frac{1}{N^{3/2}}\right) \tag{III.29}$$

where the normalized trace of the $\mathcal{O}_{z^2}^N\left(\frac{1}{N^{3/2}}\right)$ error term is a $\mathcal{O}_{z^2}\left(\frac{1}{N^2}\right)$ term. To check (III.29), we first notice that (III.10) and (I.23) lead to the equality

$$\mathbf{Q_{ff}}\eta\Pi_p^\perp = z(\eta^2\Pi_f\Pi_p - z^2)^{-1}\eta\Pi_p^\perp = z\left(-\frac{1}{z^2}\eta\Pi_p^\perp + \frac{1}{z^2}(\eta^2\Pi_f\Pi_p - z^2)^{-1}\eta^3\Pi_f\Pi_p\Pi_p^\perp\right) = -\frac{1}{z}\eta\Pi_p^\perp \tag{III.30}$$

(III.14) implies that $N^{-1}\mathbb{E}\left(\mathrm{Tr}\eta\Pi_p^\perp\right) = (1 - c_N) + \mathcal{O}(\frac{1}{N^k})$ for each k. Therefore, (III.30) leads to $\mathbb{E}\{N^{-1}\mathrm{Tr}\mathbf{Q_{ff}}\eta\Pi_p^\perp\} = -\frac{(1-c_N)}{z} + \mathcal{O}(\frac{1}{N^k})$ for each k. Moreover, (III.10) and $\Pi_p^2 = \Pi_p$ lead to $E(\mathbf{Q_{pf}}\eta\Pi_p) = E(\eta\mathbf{Q_{pf}})$, which, using again (III.8), can also be written as $E(\mathbf{Q_{pf}}) + \mathcal{O}_{z^2}^N(\frac{1}{N^k})$ for each k. Therefore, (III.22) implies that

$$\begin{aligned}\mathbb{E}\Big\{\eta\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-2}W_p\Big\} =& \frac{z}{1-c_N}\mathbb{E}\{\mathbf{Q_{pf}}\} - \frac{z}{(1-c_N)^2}\mathbf{\Delta_{pf}} + \\ & \mathcal{O}_{z^2}^N\left(\frac{1}{N^k}\right) + \mathbb{E}\Big\{\eta\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-2}W_p\Big\}\mathcal{O}_{z^2}\left(\frac{1}{N^k}\right).\end{aligned}$$

It is easily seen that $\mathbb{E}\{\eta\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-2}W_p\}\mathcal{O}_{z^2}(\frac{1}{N^k}) = \mathcal{O}_{z^2}^N(\frac{1}{N^k})$. We next notice that (III.10) implies that $\mathbf{Q_{pp}}\eta\Pi_p = z\mathbf{Q_{pf}}$. Plugging this and the above expression of $\mathbb{E}\Big\{\eta\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-2}W_p\Big\}$ into (III.21), we obtain easily (III.29). Moreover, the property of the normalized trace of the error term in (III.29) follows immediately from $\frac{1}{N}\mathrm{Tr}\,\mathbf{\Delta_{pp}} = \mathcal{O}_{z^2}\left(\frac{1}{N^2}\right)$ and $\frac{1}{N}\mathrm{Tr}\,\mathbf{\Delta_{pf}} = \mathcal{O}_{z^2}\left(\frac{1}{N^2}\right)$.

We now use in a similar way (III.23) and (III.24) to obtain another relationship between $\mathbb{E}(\mathbf{Q_{pp}})$ and $\mathbb{E}(\mathbf{Q_{pf}})$. We first notice that by (III.18) and (III.30),

$$\mathbb{E}\{N^{-1}\mathrm{Tr}\mathbf{Q_{pp}}\eta\Pi_f^\perp\} = \mathbb{E}\{N^{-1}\mathrm{Tr}\mathbf{Q_{ff}}\eta\Pi_p^\perp\} = -\frac{(1-c_N)}{z} + \mathcal{O}(\frac{1}{N^k}) \tag{III.31}$$

for each k. We then remark that (III.20) implies that

$$\mathbb{E}\{\mathbf{Q_{pp}}\eta\Pi_f\} = z\mathbb{E}\{(-\frac{1}{z^2} + \frac{1}{z^2}(\eta_N^2\Pi_p\Pi_f - z^2)^{-1}\eta_N^2\Pi_p\Pi_f)\eta_N\Pi_f\} = -\frac{1}{z}\mathbb{E}\{\eta_N\Pi_f\} + \frac{1}{z}\mathbb{E}\{\mathbf{Q_{pf}}\eta\Pi_f\} + \mathcal{O}_{z^2}^N(N^{-k}) = -\frac{c_N}{z}I_N + \frac{1}{z}\mathbb{E}\{\mathbf{Q_{pf}}\eta\Pi_f\} + \mathcal{O}_{z^2}^N(N^{-3/2})$$

for each k. Moreover, it holds that $\mathbb{E}\{\mathbf{Q_{pf}}\eta\Pi_f\} = \mathbb{E}\{(\eta^2\Pi_p\Pi_f - z^2)^{-1}\eta^2\Pi_p\Pi_f\} = I_N + z\mathbb{E}\{\mathbf{Q_{pp}}\}$. (III.23) thus allows to obtain that

$$(1-c_N)\mathbb{E}\{\eta\mathbf{Q_{pf}}W_f^*(W_fW_f^*)^{-2}W_f\} = I_N + z\mathbb{E}(\mathbf{Q_{pp}}) + \mathcal{O}_{z^2}^N(N^{-3/2})$$

(III.24) eventually leads to

$$(I_N + z\mathbb{E}(\mathbf{Q_{pp}}))\ w_N(z) = c_N\mathbb{E}(\mathbf{Q_{pp}}) + \mathcal{O}_{z^2}^N(N^{-3/2}) \tag{III.32}$$

where the normalized trace of the $\mathcal{O}_{z^2}^N\left(\frac{1}{N^{3/2}}\right)$ error term is a $\mathcal{O}_{z^2}\left(\frac{1}{N^2}\right)$ term. (III.27), (III.28), and the property of the normalized traces of the error terms then follow from (III.29) and (III.32). ■

Finally, to complete this paragraph, we denote

$$\tilde{\boldsymbol{\alpha}}_N = \frac{1}{N}\mathbb{E}\{\mathrm{Tr}\mathbf{Q_{pp}}\} = \frac{1}{N}\mathbb{E}\{\mathrm{Tr}\mathbf{Q_{ff}}\} \tag{III.33}$$

$$\boldsymbol{\alpha}_N = \frac{1}{N}\mathbb{E}\{\mathrm{Tr}\mathbf{Q_{pf}}\} = \frac{1}{N}\mathbb{E}\{\mathrm{Tr}\mathbf{Q_{fp}}\} \tag{III.34}$$

and remark that taking the normalized traces of (III.27) and (III.28) implies that $\tilde{\boldsymbol{\alpha}}_N(z) = \frac{zw_N^2(z)}{c_N^2-(zw_N(z))^2} + \mathcal{O}_{z^2}\left(\frac{1}{N^2}\right)$ and $\boldsymbol{\alpha}_N(z) = \frac{c_Nw_N(z)}{c_N^2-(zw_N(z))^2} + \mathcal{O}_{z^2}^N\left(\frac{1}{N^2}\right)$. We have thus shown the following Corollary.

Corollary III.2. *For $i = p, f$, and for $i \neq j$, we have*

$$\mathbb{E}(\mathbf{Q_{ii}}(z)) \quad = \quad \tilde{\boldsymbol{\alpha}}_N(z)\, I_N + \mathcal{O}_{z^2}^N\left(\frac{1}{N^{3/2}}\right) \tag{III.35}$$

$$\mathbb{E}(\mathbf{Q_{ij}}(z)) \quad = \quad \boldsymbol{\alpha}_N(z)\, I_N + \mathcal{O}_{z^2}^N\left(\frac{1}{N^{3/2}}\right) \tag{III.36}$$

where the normalized traces of the $\mathcal{O}_{z^2}^N\left(\frac{1}{N^{3/2}}\right)$ error terms are $\mathcal{O}_{z^2}\left(\frac{1}{N^2}\right)$.

We now establish a relationship between $\boldsymbol{\alpha}_N$ and $\tilde{\boldsymbol{\alpha}}_N$ and take benefit of this to show that $\tilde{\boldsymbol{\alpha}}_N$ is a solution of a perturbed degree 2 polynomial equation. We will deduce from this that $\mathbb{E}\left(\frac{1}{N}\mathrm{Tr}Q_N(z)\right)$ verifies a similar equation. This property will be useful to evaluate the limit eigenvalue distribution of $\Pi_p\Pi_f$. We notice that

$$\frac{1}{N}\mathrm{Tr}\eta\Pi_f^{\perp}\mathbf{Q_{pp}} = \frac{1}{N}\mathrm{Tr}\left(\eta\mathbf{Q_{pp}} - \eta_N\Pi_f z(\eta^2\Pi_p\Pi_f - z^2)^{-1}\right) = \frac{1}{N}\mathrm{Tr}\left(\eta\mathbf{Q_{pp}} - z\mathbf{Q_{fp}}\right)$$

Taking the expectation from the both sides, using (III.17) and replacing η by 1 in $\dfrac{1}{N}\mathrm{Tr}\left(\eta\mathbf{Q_{pp}}\right)$, we get that

$$\frac{1}{N}\mathbb{E}\{\eta\mathrm{Tr}\Pi_f^{\perp}\mathbf{Q_{pp}}\} = \tilde{\boldsymbol{\alpha}} - z\boldsymbol{\alpha} + \mathcal{O}_{z^2}\left(\frac{1}{N^k}\right) \tag{III.37}$$

for each k. (III.31) thus implies that

$$\boldsymbol{\alpha}_N(z) = \frac{\tilde{\boldsymbol{\alpha}}_N(z)}{z} + \frac{1-c_N}{z^2} + \mathcal{O}_{z^2}\left(\frac{1}{N^k}\right) \tag{III.38}$$

Taking the normalized trace of (III.29) leads to

$$c_N\tilde{\boldsymbol{\alpha}}_N(z) = z\boldsymbol{\alpha}_N(z)w_N(z) + \mathcal{O}_{z^2}\left(\frac{1}{N^2}\right) \tag{III.39}$$

We now express w_N in terms of $\boldsymbol{\alpha}_N$ and $\tilde{\boldsymbol{\alpha}}_N$. For this we use $\mathbf{Q_{fp}} = \eta\Pi_f(\eta^2\Pi_p\Pi_f - z^2)^{-1}$ and write

$$N^{-1}\mathbb{E}\{\mathrm{Tr}(\eta\Pi_p^{\perp}\mathbf{Q_{fp}})\} = N^{-1}\mathbb{E}\{\mathrm{Tr}(\eta\mathbf{Q_{fp}})\} - N^{-1}\mathbb{E}\{\mathrm{Tr}(\eta^2\Pi_p\Pi_f(\eta^2\Pi_p\Pi_f - z^2)^{-1})\} = \boldsymbol{\alpha} - 1 - zN^{-1}\mathbb{E}\{\mathrm{Tr}(\mathbf{Q_{pp}})\} + \mathcal{O}_{z^2}\left(\frac{1}{N^k}\right) = \boldsymbol{\alpha} - 1 - z\tilde{\boldsymbol{\alpha}} + \mathcal{O}_{z^2}\left(\frac{1}{N^k}\right) \tag{III.40}$$

Therefore, we obtain that $w(z) = \boldsymbol{\alpha}(z) - z\tilde{\boldsymbol{\alpha}}(z) + \mathcal{O}_{z^2}\left(\frac{1}{N^k}\right)$ for each k. Plugging this into (III.39) and using (III.38), we obtain after some algebra that

$$(1-z^2)\tilde{\boldsymbol{\alpha}}_N^2 + \left(\frac{2(1-c_N)}{z} - z\right)\tilde{\boldsymbol{\alpha}}_N + \frac{(1-c_N)^2}{z^2} = \mathcal{O}_{z^2}\left(\frac{1}{N^2}\right). \tag{III.41}$$

We define $\tilde{\alpha}_N(z)$ and $\alpha_N(z)$ by

$$\tilde{\alpha}_N(z) = \frac{1}{N}\mathbb{E}\{\mathrm{Tr}Q_N(z)\} \tag{III.42}$$

$$\alpha_N(z) = \frac{1}{N}\mathbb{E}\{\mathrm{Tr}\eta\Pi_p Q_N(z)\} \tag{III.43}$$

for each $z \in \mathbb{C}\setminus\mathbb{R}^+$. (III.10) implies that $\tilde{\boldsymbol{\alpha}}_N(z) = z\tilde{\alpha}_N(z^2)$ and $\boldsymbol{\alpha}_N(z) = \alpha_N(z^2)$ if $\mathrm{Im}z \neq 0$ or equivalently if $z^2 \in \mathbb{C}\setminus\mathbb{R}^+$. Therefore, we deduce from (III.41) that $\tilde{\alpha}_N(z)$ is a solution of the perturbed equation

$$(1-z^2)z^2\tilde{\alpha}_N^2(z^2) + \left(2(1-c_N) - z^2\right)\tilde{\alpha}_N(z^2) + \frac{(1-c_N)^2}{z^2} = \mathcal{O}_{z^2}\left(\frac{1}{N^2}\right).$$

The l.h.s of this equation is a function of $z^2 \in \mathbb{C}\setminus\mathbb{R}^+$, thus the error term at the r.h.s is also a function of z^2. By exchanging z^2 with z we have

$$(1-z)z\tilde{\alpha}_N^2(z) + (2(1-c_N) - z)\tilde{\alpha}_N(z) + \frac{(1-c_N)^2}{z} = \mathcal{O}_z\left(\frac{1}{N^2}\right) \tag{III.44}$$

on $\mathbb{C}\setminus\mathbb{R}^+$. Moreover, from (III.38), we obtain that

$$\alpha_N(z) = \tilde{\alpha}_N(z) + \frac{1-c_N}{z} + \mathcal{O}_z\left(\frac{1}{N^k}\right). \tag{III.45}$$

on $\mathbb{C}\setminus\mathbb{R}^+$ for each integer $k \geq 1$.

Remark III.3. *Corollary III.1 implies that α_N is the Stieltjes transform of a positive measure carried by $\mathbb{R}^+$ with mass $c_N + \mathcal{O}_z(N^{-k})$. This is in accordance with (III.45) because $\tilde{\alpha}_N$ is the Stieltjes transform of a probability measure carried by $\mathbb{R}^+$ (i.e. the expectation of the empirical eigenvalue distribution of $\eta^2\Pi_p\Pi_f$) and $-\frac{1-c_N}{z}$ is the Stieltjes transform of measure $(1-c_N)\delta_0$.*

3) Limiting distribution and almost sure localisation of the eigenvalues of $\Pi_p\Pi_f$: In this paragraph, we evaluate the almost sure asymptotic behaviour of the empirical eigenvalue distribution $\hat{\nu}_N$ of matrix $\Pi_p\Pi_f$. As $\eta_N = 1$ almost surely for N large enough, this can be done by evaluating the almost sure behaviour of $\frac{1}{N}\mathrm{Tr}(Q_N(z))$ where we recall that $Q_N(z)$ is the resolvent of the regularized matrix $\eta_N^2\Pi_p\Pi_f$. We first notice that, in conjunction with the Borel-Cantelli Lemma, Lemma A.2, Eq. (A.23), applied for $i = j = p$ and $F = I$, implies immediately that

$$\frac{1}{N}\mathrm{Tr}(Q_N(z)) - \mathbb{E}\left(\frac{1}{N}\mathrm{Tr}(Q_N(z))\right) \to 0 \ a.s. \tag{III.46}$$

for each $z \in \mathbb{C}\setminus\mathbb{R}^+$. We are thus back to the evaluation of the asymptotic behaviour of $\tilde{\alpha}_N(z) = \mathbb{E}\left(\frac{1}{N}\mathrm{Tr}(Q_N(z))\right)$. For this, we introduce the probability measure $\tilde{\nu}_N$ defined by

$$\tilde{\nu}_N = (c_N\delta_1 + (1-c_N)\delta_0) \boxtimes (c_N\delta_1 + (1-c_N)\delta_0) \tag{III.47}$$

where $\boxtimes$ represents the free multiplicative convolution product operator (see e.g. [52] Section 3.6). We recall that if Π_1 and Π_2 are orthogonal projection matrices onto the rows of two mutually independent random Gaussian $ML \times N$ matrices with i.i.d. standard Gaussian entries, then the results of [52] imply that the empirical eigenvalue distribution of $\Pi_1\Pi_2$ has the same asymptotic behaviour than $\tilde{\nu}_N$. In the following, we establish that, while Π_p and Π_f are not generated as Π_1 and Π_2, $\hat{\nu}_N$ behaves as $\tilde{\nu}_N$.

For this, we denote by $\tilde{t}_N$ the Stieltjes transform of $\tilde{\nu}_N$. The expression and the properties of $\tilde{t}_N$ and of $\tilde{\nu}_N$ are well-known, see for example Example 3.6.7. [52]. If $z \in \mathbb{C}^+$, $\tilde{t}_N$ is given by

$$\tilde{t}_N(z) = \frac{z - 2(1-c_N) + \sqrt{z(z - 4c_N(1-c_N))}}{2(1-z)z}, \tag{III.48}$$

where we define function $z \mapsto \sqrt{z}$ for $z = |z|e^{i\theta}$, $\theta \in [0, 2\pi)$ as $\sqrt{z} = \sqrt{|z|}e^{i\theta/2}$. In particular, if $x \in \mathbb{R}^+$ and $z = xe^{i\theta}$ then $\sqrt{z} \xrightarrow[\theta \searrow 0]{} \sqrt{x}$ and $\sqrt{z} \xrightarrow[\theta \nearrow 2\pi]{} -\sqrt{x}$. Then one can easily obtain that $\lim_{z \to x, z \in \mathbb{C}^+} \tilde{t}_N(z)$ exists for $x \in (-\infty, 0) \cap (4c_N(1-c_N), +\infty)$ and $x \neq 1$. This limit is still denoted $\tilde{t}_N(x)$, and

$$\tilde{t}_N(x) = \begin{cases} \dfrac{x - 2(1-c_N) - \sqrt{x(x - 4c_N(1-c_N))}}{2(1-x)x}, & x < 0 \\ \dfrac{x - 2(1-c_N) + \sqrt{x(x - 4c_N(1-c_N))}}{2(1-x)x}, & x > 4c_N(1-c_N), x \neq 1 \end{cases} \tag{III.49}$$

Moreover, $\tilde{\nu}_N = (c_N\delta_1 + (1-c_N)\delta_0) \boxtimes (c_N\delta_1 + (1-c_N)\delta_0)$ is given by

$$d\tilde{\nu}_N(\lambda) = \frac{\sqrt{\lambda(4c_N(1-c_N) - \lambda)}}{2\pi\lambda(1-\lambda)} \mathbf{1}_{[0, 4c_N(1-c_N)]} d\lambda + (1-c_N)\delta_0 + \max(2c_N - 1, 0)\delta_1. \tag{III.50}$$

The support of $\tilde{\nu}_N$, denoted by $\mathcal{S}_N$, is thus given by

$$\mathcal{S}_N = [0, 4c_N(1-c_N)] \cup \{1\} \, \mathbf{1}_{c_N > 1/2}. \tag{III.51}$$

Finally, $\tilde{t}_N$ satisfies the equation (III.44), but in which the term $\mathcal{O}_z(N^{-2})$ is replaced by 0, i.e.

$$z(1-z)\tilde{t}_N^2(z) + (2(1-c_N) - z)\tilde{t}_N(z) + \frac{(1-c_N)^2}{z} = 0 \tag{III.52}$$

a property which suggests that $\tilde{\alpha}_N(z) - \tilde{t}_N(z) \to 0$. In order to establish this formally, we establish the following Proposition.

Proposition III.3. *$\tilde{\alpha}_N(z)$ can be written as*

$$\tilde{\alpha}_N(z) = \tilde{t}_N(z) + \tilde{r}_N(z), \tag{III.53}$$

where $\tilde{r}_N$ is holomorphic in $\mathbb{C} \setminus \mathbb{R}^+$, and verifies

$$|r_N(z)| \leq \frac{1}{N^2} P_1(|z|) P_2\left(\frac{1}{\text{Im} z}\right) \tag{III.54}$$

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

The proof is given in the Appendix.

As $c_N \to c_*$, $\tilde{t}_N(z) \to \tilde{t}_*(z)$ where $\tilde{t}_*(z)$ is obtained from $\tilde{t}_N(z)$ by replacing c_N by c_* in Eq. (III.48). $\tilde{t}_*$ is of course the Stieljes transform of the measure $\tilde{\nu}_*$ given by

$$d\tilde{\nu}_*(\lambda) = \frac{\sqrt{\lambda(4c_*(1-c_*) - \lambda)}}{2\pi\lambda(1-\lambda)} \mathbf{1}_{[0, 4c_*(1-c_*)]} d\lambda + (1-c_*)\delta_0 + \max(2c_* - 1, 0)\delta_1 \tag{III.55}$$

and the support $\mathcal{S}_*$ of $\tilde{\nu}_*$ is obtained by replacing c_N by c_* in (III.51). Sequence $(\tilde{\nu}_N)_{N \geq 1}$ of course converges weakly towards the probability measure $\tilde{\nu}_*$. We deduce from this and from Proposition III.3 the following Theorem which states that $(\hat{\nu}_N)_{N \geq 1}$ converges weakly almost surely towards $\tilde{\nu}_*$. Moreover, all the eigenvalues of $\Pi_p \Pi_f$ are almost surely localised in a neighbourhood of $\mathcal{S}_*$.

Theorem III.1. *The empirical eigenvalue distribution $\hat{\nu}_N$ of $\Pi_{p,N}\Pi_{f,N}$ verifies*

$$\hat{\nu}_N \to \tilde{\nu}_* \tag{III.56}$$

weakly almost surely. If $c_* < \frac{1}{2}$, for each $\epsilon > 0$, almost surely, for each N large enough, all the eigenvalues of $\Pi_p \Pi_f$ belong to $[0, 4c_*(1-c_*) + \epsilon]$. If $c_* > \frac{1}{2}$, 1 is eigenvalue of $\Pi_p \Pi_f$ with multiplicity $2ML - N$, and for each $\epsilon > 0$, the $2(N - ML)$ remaining eigenvalues are almost surely located in $[0, 4c_*(1-c_*) + \epsilon]$ for N large enough.

Proof. (III.46) and Proposition III.3 imply that

$$\frac{1}{N}\text{Tr}(Q_N(z)) - \tilde{t}_N(z) \to 0 \; a.s. \tag{III.57}$$

for each $z \in \mathbb{C}^+$. As $\tilde{t}_N(z) \to \tilde{t}_*(z)$ on $\mathbb{C}^+$, we obtain that $\frac{1}{N}\text{Tr}(Q_N(z)) \to \tilde{t}_*(z) \to 0$ almost surely on $\mathbb{C}^+$, and that (III.56) holds.

We remark that if $c_* = \frac{1}{2}$, the support $\mathcal{S}_*$ of $\tilde{\nu}_*$ is equal to the whole interval $[0,1]$. As we know that the eigenvalues of $\Pi_p \Pi_f$ belong to $[0,1]$, the knowledge of $\mathcal{S}_*$ does not provide any valuable information of the almost sure location of these eigenvalues if $c_* = \frac{1}{2}$. If $c_* \neq \frac{1}{2}$, the almost sure localisation of the eigenvalues of $\Pi_p \Pi_f$ can be established using the Haagerup-Thornbjornsen approach ([20]) using decomposition (III.53) of $\tilde{\alpha}_N(z)$. As the corresponding proof is rather standard, we just provide a sketch of proof. We first mention that (III.53) implies that if ψ is a $\mathcal{C}_\infty$ function constant on the complementary of a compact subset, then, we have

$$\mathbb{E}\left(\mathrm{Tr}\left(\psi(\Pi_p \Pi_f)\right)\right) = N \int_{\mathcal{S}_N} \psi(\lambda)\, d\tilde{\nu}_N(\lambda) + \mathcal{O}(\frac{1}{N}) \tag{III.58}$$

(see Proposition 6.2 in [20] or Proposition 4.6 in [9]). If moreover ψ vanishes on $\mathcal{S}_N$, we obtain that

$$\mathbb{E}\left(\mathrm{Tr}\left(\psi(\Pi_p \Pi_f)\right)\right) = \mathcal{O}(\frac{1}{N}) \tag{III.59}$$

while if ψ' vanishes on $\mathcal{S}_N$, the Poincaré-Nash inequality allows to establish that $\mathrm{Tr}\left(\psi(\Pi_p \Pi_f) - \mathbb{E}\left(\mathrm{Tr}\left(\psi(\Pi_p \Pi_f)\right)\right) \to 0\right.$ almost surely. Therefore, (III.58) implies that $\mathrm{Tr}\left(\psi(\Pi_p \Pi_f) - N \int_{\mathcal{S}_N} \psi(\lambda)\, d\tilde{\nu}_N(\lambda) \to 0\right.$ almost surely if ψ' vanishes on $\mathcal{S}_N$. We consider $\epsilon > 0$ small enough, and a function $\psi_1 \in \mathcal{C}_\infty$ that verifies:

$$\begin{aligned}
\psi_1(\lambda) =& 1 \text{ if } \lambda \in \left([0, 4c_*(1-c_*)+\epsilon] \cup [1-\epsilon, 1+\epsilon]\mathbf{1}_{c_*>1/2}\right)^c \\
\psi_1(\lambda) =& 0 \text{ if } \lambda \in [0, 4c_*(1-c_*)+\epsilon/2] \cup [1-\epsilon/2, 1+\epsilon/2]\mathbf{1}_{c_*>1/2} \\
\psi_1(\lambda) \in& [0,1] \text{ elsewhere}
\end{aligned}$$

As $c_N \to c_*$, ψ_1 (and therefore ψ_1') vanishes on $\mathcal{S}_N$ for N large enough, so that $\mathrm{Tr}\left(\psi_1(\Pi_p \Pi_f)\right) \to 0$. The number of eigenvalues of $\Pi_p \Pi_f$ located into $\in \left([0, 4c_*(1-c_*)+\epsilon] \cup [1-\epsilon, 1+\epsilon]\mathbf{1}_{c_*>1/2}\right)^c$ is clearly less than $\mathrm{Tr}\left(\psi_1(\Pi_p \Pi_f)\right)$ which converges towards 0. Therefore, almost surely, for each N large enough, all the eigenvalues of $\Pi_p \Pi_f$ belong to $[0, 4c_*(1-c_*)+\epsilon] \cup [1-\epsilon, 1+\epsilon]\mathbf{1}_{c_*>1/2}$. This completes the proof of Theorem III.57 when $c_* < 1/2$. In order to address the case $c_* > 1/2$, we consider a function $\psi_2 \in \mathcal{C}_\infty$ satisfying

$$\begin{aligned}
\psi_2(\lambda) =& 1 \text{ if } \lambda \in [1-\epsilon, 1+\epsilon] \\
\psi_2(\lambda) =& 0 \text{ if } \lambda \in [1-2\epsilon, 1+2\epsilon]^c \\
\psi_2(\lambda) \in& [0,1] \text{ elsewhere}
\end{aligned}$$

ψ_2' vanishes of $\mathcal{S}_N$, and $\int_{\mathcal{S}_N} \psi_2(\lambda)\, d\tilde{\nu}_N(\lambda) = 2c_N - 1$. Therefore, we obtain that $\mathrm{Tr}\left(\psi_2(\Pi_p \Pi_f)\right) - (2ML - N) \to 0$ almost surely. As there is no eigenvalue of $\Pi_p \Pi_f$ in $[1-2\epsilon, 1-\epsilon)$, $\mathrm{Tr}\left(\psi_2(\Pi_p \Pi_f)\right)$ coincides with the number of eigenvalues of $\Pi_p \Pi_f$ located into $[1-\epsilon, 1]$. As $\mathrm{Tr}\left(\psi_2(\Pi_p \Pi_f)\right) - (2ML - N) \to 0$ almost surely, we obtain that for N large enough, $\Pi_p \Pi_f$ has $2ML - N$ eigenvalues located in $[1-\epsilon, 1]$. Lemma III.1 implies that 1 is eigenvalue of $\Pi_p \Pi_f$ with multiplicity $2ML - N$, from which we get that if $c_* > 1/2$, the eigenvalues of $\Pi_p \Pi_f$ belong to $[0, 4c_*(1-c_*)+\epsilon] \cup \{1\}$. ■

In the following, it will be useful to introduce the measure ν_N defined by

$$\nu_N = \frac{1}{c_N}\tilde{\nu}_N - \frac{1-c_N}{c_N}\delta_0 = \frac{\sqrt{\lambda(4c_N(1-c_N)-\lambda)}}{2\pi c_N \lambda(1-\lambda)} \mathbf{1}_{[0,4c_N(1-c_N)]} d\lambda + \max(2c_N - 1, 0)\delta_1 \tag{III.60}$$

It is easily seen that ν_N is the probability measure carried by $\mathcal{S}_N$ with Stieltjes transform $t_N(z)$ defined on $\mathbb{C} \setminus \mathcal{S}_N$ by

$$t_N(z) = \frac{\tilde{t}_N(z)}{c_N} + \frac{1-c_N}{c_N z} \tag{III.61}$$

After some algebra, we obtain that

$$\begin{aligned}
t_N(z) &= \frac{z(2c_N-1) + \sqrt{z(z-4c_N(1-c_N))}}{2c_N(1-z)z}, \quad z \in \mathbb{C}^+ \\
t_N(x) &= \begin{cases} \dfrac{x(2c_N-1) - \sqrt{x(x-4c_N(1-c_N))}}{2c_N(1-x)x}, & x < 0 \\ \dfrac{x(2c_N-1) + \sqrt{x(x-4c_N(1-c_N))}}{2c_N(1-x)x}, & x > 4c_N(1-c_N), x \neq 1 \end{cases}
\end{aligned} \tag{III.62}$$

We also define $\tilde{\mathbf{t}}_N(z) = z\tilde{t}_N(z^2)$ and $\mathbf{t}_N(z) = t_N(z^2)$ which are related by

$$\mathbf{t}_N(z) = \frac{\tilde{\mathbf{t}}_N(z)}{c_N z} + \frac{1-c_N}{c_N z^2}. \tag{III.63}$$

(I.18) implies that $\tilde{\mathbf{t}}_N$ is the Stieljes transform of a probability measure whose support is clearly the set $\boldsymbol{\mathcal{S}}_N$ defined by

$$\boldsymbol{\mathcal{S}}_N = [-\sqrt{4c_N(1-c_N)}, \sqrt{4c_N(1-c_N)}] \cup \{\pm 1\}\mathbf{1}_{c_N>1/2} \tag{III.64}$$

While $\mathbf{t}_N$ is not a Stieltjes transform, we however mention that $\mathbf{t}_N$ is also holomorphic outside $\boldsymbol{\mathcal{S}}_N$. Then, we deduce from (III.45) and (III.53) the following obvious, but useful properties.

Corollary III.3. *The sequence $(\alpha_N(z))_{N\geq 1}$ verifies*

$$\alpha_N(z) - c_N t_N(z) \to 0$$

for $z \in \mathbb{C} \setminus \mathbb{R}^+$. Moreover, we also have

$$\boldsymbol{\alpha}_N(z) - c_N \mathbf{t}_N(z) \to 0 \tag{III.65}$$

$$\tilde{\boldsymbol{\alpha}}_N(z) - \tilde{\mathbf{t}}_N(z) \to 0 \tag{III.66}$$

on $\mathbb{C}^+$.

We also denote by ν_* and $t_*(z)$ the limits of ν_N and $t_N(z)$ when $N \to +\infty$, i.e. their expressions are obtained by replacing c_N by c_* in (III.60) and (III.62). We have of course $\tilde{\nu}_* = c_N \nu_* + (1-c_*)\delta_0$. We also remark that if $\hat{\nu}'_N$ represents the eigenvalue distribution of matrix $(W_p W_p^*)^{-1/2} W_p W_f^* (W_f W_f^*)^{-1} W_f W_p^* (W_p W_p^*)^{-1/2}$, then $\hat{\nu}_N = c_N \hat{\nu}'_N + (1-c_N)\delta_0$. Therefore, the relation $\tilde{\nu}_* = c_N \nu_* + (1-c_*)\delta_0$ and the convergence result (III.56) imply that

$$\hat{\nu}'_N \to \nu_* \tag{III.67}$$

almost surely.

In the following, we also denote by $\mathbf{t}_*(z)$ and $\tilde{\mathbf{t}}_*(z)$ the functions $t_*(z^2)$ and $z\tilde{t}_*(z^2)$ respectively, that can also be seen as the limits of $\mathbf{t}_N(z)$ and $\tilde{\mathbf{t}}_N(z)$ when $N \to +\infty$. $\tilde{\mathbf{t}}_*(z)$ is of course the Stieltjes transform of a probability measure carried by the set $\boldsymbol{\mathcal{S}}_*$ obtained by replacing c_N by c_* in (III.64).

We finally conclude this section by a result which can be seen as the counterpart of Lemma II.1 derived in Section II.

Lemma III.6. *For each $z \in \mathbb{C}\backslash\boldsymbol{\mathcal{S}}_*$, $i \neq j \in \{p, f\}$ and for each bounded sequences $(a_N, b_N)_{N\geq 1}$ of N–dimensional deterministic vectors, it holds that*

$$a_N^* \, (\mathbf{Q_{ii}})_N(z) \, b_N - \tilde{\mathbf{t}}_N(z) a_N^* b_N \to 0 \text{ almost surely} \tag{III.68}$$

$$a_N^* \, (\mathbf{Q_{ij}})_N(z) \, b_N - c_N \mathbf{t}_N(z) a_N^* b_N \to 0 \text{ almost surely} \tag{III.69}$$

Moreover, these convergences hold uniformly on each compact subset of $\mathbb{C} \setminus \boldsymbol{\mathcal{S}}_*$. The properties are still valid if a_N, b_N are random bounded vectors that are independent from the noise sequence $(v_n)_{n\geq 1}$.

The proof is given in the Appendix.

B. *In the presence of signal*

In this section we assume that signal $(u_n)_{n\in\mathbb{Z}}$ is present, and evaluate its influence on the eigenvalues of matrix $\Pi_{p,y}\Pi_{f,y}$. For this, we notice that matrices $\Pi_{p,y}$ and $\Pi_{f,y}$ are finite rank perturbation of matrices $\Pi_{p,v}$ and $\Pi_{f,v}$ due to the noise $(v_n)_{n\in\mathbb{Z}}$,. Therefore, $\Pi_{p,y}\Pi_{f,y}$ is itself a finite rank perturbation of $\Pi_{p,v}\Pi_{f,v}$ We can thus use the same approach as in the previous chapter. Since the useful signal $(u_n)_{n\in\mathbb{Z}}$ is generated by the same minimal state-space representation (I.3), we keep the notations from the Section II-B. As before, we denote $\Sigma_{i,N} = \frac{Y_{i,N}}{\sqrt{N}} = W_{i,N} + \frac{U_{i,N}}{\sqrt{N}}$. $\Pi_{i,y}$ and $\Pi_{i,v}$ are denoted respectively Π_i and Π_i^W for $i = p, f$ from now on. We remind that in the presence of signal, we cannot assume that $R_N = I_M$, thus $W_i = (I_L \otimes R_N)^{1/2} W_{i,iid}$ where matrix $W_{i,iid}$ is built from i.i.d. $\mathcal{N}_c(0, I_M)$ distributed random vector $(v_{n,iid})_{n=1,\ldots,n+2L-1}$. However, we recall that $\Pi_i^W = \Pi_i^{W_{iid}}$ for $i = p, f$. In

the following, we will denote by η_N (rather than $\eta_{N,iid}$) the regularization term defined by (III.4) by replacing W by W_{iid} in order to simplify the notations.
We also keep Assumptions II.1 and II.2, as well as Assumption II.3 on the limits of Δ_N and $\Theta_N^* \mathbf{R}_{f|p,N}^L \Theta_N$ related to the signal model. As in Section II-E, we derive the following results under condition (II.33), and briefly justify that Theorem III.2 remains valid if some of the entries of matrix Δ_* coincide. Finally, it appears that the more involved Assumptions II.4, II.5, and II.6 are not needed here and can be replaced with the following milder one.

Assumption III.1. *$r \times r$ matrix $G_N = \Theta_N^*(I_L \otimes R_N^{-1})\Theta_N$ converge towards some matrix G_*.*

We now take benefit of Proposition II.2 to evaluate the behaviour of the canonical correlation coefficients between the row spaces of matrices $U_{p,N}$ and $U_{f,N}$ when $N \to +\infty$. For this, we recall that Γ_* represents the limit of $\Theta_N^* R_{f|p,N}^L \Theta_N$, as well as, under condition (II.33), the limit of of $\Gamma_N = \Delta_N \tilde{\Theta}_{f,N}^* \Theta_{p,N} \Delta_N$ (see Eq. (II.37) for the definition of Γ_N). As $\Delta_N \to \Delta_* > 0$, $\tilde{\Theta}_{f,N}^* \tilde{\Theta}_{p,N}$ converges towards the matrix Ω_* given by

$$\Omega_* = \Delta_*^{-1} \Gamma_* \Delta_*^{-1} \tag{III.70}$$

Ω_* of course verifies $\|\Omega_*\| \leq 1$ and $\mathrm{Rank}(\Omega_*) = P$.

We are now in position to formulate the main result of this Section. For this, we denote by F_* the rank P $r \times r$ matrix defined by

$$F_* = \Omega_*^*(I_r + \Delta_*^{-1} G_*^{-1} \Delta_*^{-1})^{-1} \Omega_* (I_r + \Delta_*^{-1} G_*^{-1} \Delta_*^{-1})^{-1} \tag{III.71}$$

As matrix Ω_* verifies $\|\Omega_*\| \leq 1$, matrix F_* satisfies $\|F_*\| < 1$. Moreover, the eigenvalues of F_* are real and belong to $[0, 1)$.

Theorem III.2.
- *The function $f_*(x)$ defined by*

$$f_*(x) = x \left(\frac{\tilde{t}_*(x)}{(1 - c_*) t_*(x)} \right)^2 \tag{III.72}$$

  is strictly increasing on $[4c_*(1 - c_*), 1]$, verifies $f_*(4c_*(1 - c_*)) = \frac{c_*}{1-c_*}$, $f(1) = 1$ if $c_* < \frac{1}{2}$ and $f(1) = \left(\frac{c_*}{1-c_*}\right)^2$ if $c_* > \frac{1}{2}$.
- *If $c_* \geq \frac{1}{2}$, the equation*

$$\det\left(f_*(x)\, I_r - F_*\right) = 0 \tag{III.73}$$

  has no solution in $(4c_*(1-c_*), 1)$, and for each $\delta > 0$, almost surely, for N large enough, all the eigenvalues of $\Pi_p \Pi_f$ belong to $[0, 4c_*(1-c_*)+\delta] \cup [1-\delta, 1]$. Among the eigenvalues contained in $[1-\delta, 1]$, $2ML - N + \mathcal{O}(1)$ are equal to 1, and, possibly, $o(N)$ other eigenvalues converge towards 1.
- *If $c_* < \frac{1}{2}$, the equation (III.73) has $0 \leq s \leq P$ solutions that belong to $(4c_*(1 - c_*), 1)$ where s is the number of eigenvalues (taking into account the multiplicities) of F_* that are striclty larger than $\frac{c_*}{1-c_*} < 1$. If $\rho_{1,*}, \ldots, \rho_{s,*}$ are the corresponding solutions, then the s largest eigenvalues of $\Pi_p \Pi_f$ converge almost surely towards $\rho_{1,*}, \ldots, \rho_{s,*}$, and, for each $\delta > 0$, almost surely, for N large enough, the remaining $N - s$ ones belong to $[0, 4c_*(1 - c_*) + \delta]$.*

Proof. The properties of function f_* are proved in the Appendix. $x \in (4c_*(1 - c_*), 1)$ is solution of equation (III.72) if and only $f_*(x)$ coincides with one of the eigenvalues of F_*. If $c_* \geq \frac{1}{2}$, $f_*(x) \in [\frac{c_*}{1-c_*}, \left(\frac{c_*}{1-c_*}\right)^2]$ if $x \in [4c_*(1 - c_*), 1]$. As $\frac{c_*}{1-c_*} \geq 1$ and the eigenvalues of F_* belong to $[0, 1)$, equation (III.72) has no solution in $(4c_*(1 - c_*), 1)$. If $c_* < \frac{1}{2}$, $f_*((4c_*(1 - c_*), 1)$ coincides with the interval $(\frac{c_*}{1-c_*}, 1)$. Therefore, equation (III.72) has s solutions, where s represents the number of eigenvalues of F_* strictly larger than $\frac{c_*}{1-c_*}$.

We now establish the last statements of Theorem III.2 related to the possible eigenvalues of $\Pi_p \Pi_f$ that escape from $\mathcal{S}_* = [0, 4c_*(1 - c_*)] \cup \{1\} \mathbf{1}_{c_* > 1/2}$. We first present the general approach of the proof. As before, we study the squares of the positive eigenvalues of the linearised version $\begin{pmatrix} 0 & \Pi_p \\ \Pi_f & 0 \end{pmatrix}$ that escape form $[0, 2\sqrt{c_*(1 - c_*)}] \cup \{1\} \mathbf{1}_{c_* > 1/2}$. For this, for each $\delta > 0$ small enough, we consider $y \in (\sqrt{4c_*(1 - c_*) + \delta}, 1 - \delta)$

if $c_* > \frac{1}{2}$ and $y \in (\sqrt{4c_*(1-c_*)+\delta}, 1]$ if $c_* < \frac{1}{2}$, which by Theorem III.1, cannot be, almost surely, for N large enough, an eigenvalue of matrix $\begin{pmatrix} 0 & \Pi_p^W \\ \Pi_f^W & 0 \end{pmatrix}$. We take benefit of this property to express $\det \begin{pmatrix} -yI_N & \Pi_p \\ \Pi_f & -yI_N \end{pmatrix}$ in terms of $\det \begin{pmatrix} -yI_N & \Pi_p^W \\ \Pi_f^W & -yI_N \end{pmatrix}$ and of the resolvent of matrix $\begin{pmatrix} 0 & \Pi_p^W \\ \Pi_f^W & 0 \end{pmatrix}$ evaluated at y, which is well defined. As we establish almost sure convergence results in the following, we notice that the regularisation term η_N defined by (III.4) by exchanging W_i by $W_{i,iid}$, $i = p, f$, can be considered to be equal to 1. Therefore, the later resolvent coincides with $\mathbf{Q}^{\mathbf{W}}(y) = \mathbf{Q}^{\mathbf{W}_{\text{iid}}}(y)$ defined by (III.10) for $z = y$. We then evaluate the asymptotic behaviour of $\det \begin{pmatrix} -yI_N & \Pi_p \\ \Pi_f & -yI_N \end{pmatrix}$, and deduce from this the last statements of Theorem III.2.

The key point is to use that $\begin{pmatrix} -yI_N & \Pi_p \\ \Pi_f & -yI_N \end{pmatrix}$ is a low rank perturbation of $\begin{pmatrix} -yI_N & \Pi_p^W \\ \Pi_f^W & -yI_N \end{pmatrix}$. In order to evaluate the corresponding low-rank matrix, we have first to evaluate $\Pi_i - \Pi_i^W$ for $i = p, f$. It is easy to see that $\Sigma_i \Sigma_i^*$ can be expressed as

$$\Sigma_i \Sigma_i^* = W_i W_i^* + (W_i \tilde{\Theta}_i \Delta_i, \Theta_i) \begin{pmatrix} 0 & I_r \\ I_r & \Delta_p^2 \end{pmatrix} \begin{pmatrix} \Delta_i \tilde{\Theta}_i^* W_i^* \\ \Theta^* \end{pmatrix}$$

where we recall that $\frac{U_i}{\sqrt{N}} = \Theta_i \Delta_i \tilde{\Theta}_i^*$ is the singular value decomposition of $\frac{U_i}{\sqrt{N}}$ (see Eq. (II.32)).
We first establish that, almost surely, for N large enough, matrix $\Sigma_i \Sigma_i^*$ is invertible. For this, we need the following Lemma proved in the Appendix.

Lemma III.7. *We define E_i as the $2r \times 2r$ matrix given by*

$$E_i = \begin{pmatrix} 0 & I_r \\ I_r & \Delta_i^2 \end{pmatrix}^{-1} \left(I_{2r} + \begin{pmatrix} 0 & I_r \\ I_r & \Delta_i^2 \end{pmatrix} \begin{pmatrix} \Delta_i \tilde{\Theta}_i^* \Pi_i^W \tilde{\Theta}_i \Delta_i & \Delta_i \tilde{\Theta}_i^* W_i^* (W_i W_i^*)^{-1} \Theta_i \\ \Theta_i^* (W_i W_i^*)^{-1} W_i \tilde{\Theta}_i \Delta_i & \Theta_i^* (W_i W_i^*)^{-1} \Theta_i \end{pmatrix} \right) \tag{III.74}$$

Then, we have

$$E_i - \begin{pmatrix} -(1-c_N)\Delta_N^2 & I_r \\ I_r & \frac{1}{1-c_N} \Theta_N^* (I_L \otimes R_N^{-1}) \Theta_N \end{pmatrix} \to 0 \textit{ almost surely} \tag{III.75}$$

The determinant of the second term of the left hand side of (III.75) is equal to

$$\det\left(-(1-c_N)\Delta_N^2\right) \det\left(\frac{1}{1-c_N} (\Theta_N^* (I_L \otimes R_N^{-1}) \Theta_N + \Delta_N^{-2}) \right)$$

and thus converges towards a non zero term. Therefore, almost surely, for N large enough, matrix E_i is invertible. In the following, we denote by $D_i = E_i^{-1}$ the inverse of E_i. The Woodbury's identity implies that $\Sigma_i \Sigma_i^*$ is almost surely invertible for each N large enough, and that

$$(\Sigma_i \Sigma_i^*)^{-1} = (W_i W_i^*)^{-1} - ((W_i W_i^*)^{-1} W_i \tilde{\Theta}_i \Delta_i, (W_i W_i^*)^{-1} \Theta_i) D_i \begin{pmatrix} \Delta_i \tilde{\Theta}_i^* W_i^* (W_i W_i^*)^{-1} \\ \Theta_i^* (W_i W_i^*)^{-1} \end{pmatrix},$$

After some algebra, we obtain that

$$\Pi_i - \Pi_i^W = -\mathcal{A}_i D_i \mathcal{A}_i^*,$$

where

$$\mathcal{A}_i = (-\Pi_i^{W,\perp} \tilde{\Theta}_i \Delta_i, W_i^* (W_i W_i^*)^{-1} \Theta_i) \tag{III.76}$$

From this, we immediately get that

$$\begin{pmatrix} -yI_N & \Pi_p \\ \Pi_f & -yI_N \end{pmatrix} = \begin{pmatrix} -yI_N & \Pi_p^W \\ \Pi_f^W & -yI_N \end{pmatrix} - \begin{pmatrix} \mathcal{A}_p & 0 \\ 0 & \mathcal{A}_f \end{pmatrix} \begin{pmatrix} D_p & 0 \\ 0 & D_f \end{pmatrix} \begin{pmatrix} 0 & \mathcal{A}_p^* \\ \mathcal{A}_f^* & 0 \end{pmatrix} \tag{III.77}$$

or equivalently

$$\begin{pmatrix} -yI_N & \Pi_p \\ \Pi_f & -yI_N \end{pmatrix} = \begin{pmatrix} -yI_N & \Pi_p^W \\ \Pi_f^W & -yI_N \end{pmatrix} \left(I_{2N} - \mathbf{Q}^{\mathbf{W}}(y) \begin{pmatrix} \mathcal{A}_p & 0 \\ 0 & \mathcal{A}_f \end{pmatrix} \begin{pmatrix} D_p & 0 \\ 0 & D_f \end{pmatrix} \begin{pmatrix} 0 & \mathcal{A}_p^* \\ \mathcal{A}_f^* & 0 \end{pmatrix} \right) \tag{III.78}$$

Therefore, y is an eigenvalue of $\begin{pmatrix} 0 & \Pi_p \\ \Pi_f & 0 \end{pmatrix}$ if and only if the determinant of the second term at the r.h.s. of (III.78) vanishes, or equivalently if

$$\det \left(I_{2r} - \begin{pmatrix} \mathcal{A}_p^* \mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}(y)\mathcal{A}_p & \mathcal{A}_p^* \mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}}(y)\mathcal{A}_f \\ \mathcal{A}_f^* \mathbf{Q}_{\mathbf{pp}}^{\mathbf{W}}(y)\mathcal{A}_p & \mathcal{A}_f^* \mathbf{Q}_{\mathbf{pf}}^{\mathbf{W}}(y)\mathcal{A}_f \end{pmatrix} \begin{pmatrix} D_p & 0 \\ 0 & D_f \end{pmatrix} \right) = 0 \tag{III.79}$$

or

$$\det \left(\begin{pmatrix} E_p & 0 \\ 0 & E_f \end{pmatrix} - \begin{pmatrix} \mathcal{A}_p^* \mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}(y)\mathcal{A}_p & \mathcal{A}_p^* \mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}}(y)\mathcal{A}_f \\ \mathcal{A}_f^* \mathbf{Q}_{\mathbf{pp}}^{\mathbf{W}}(y)\mathcal{A}_p & \mathcal{A}_f^* \mathbf{Q}_{\mathbf{pf}}^{\mathbf{W}}(y)\mathcal{A}_f \end{pmatrix} \right) = 0 \tag{III.80}$$

We now establish that for each $y \in (\sqrt{4c_*(1-c_*)}, 1)$, the left hand side of (III.80) converges towards a deterministic term. In particular, we have the following result.

Lemma III.8. *For each $z \in \mathbb{C} \setminus \boldsymbol{\mathcal{S}}_*$, where $\boldsymbol{\mathcal{S}}_* = (-2\sqrt{c_*(1-c_*)}, 2\sqrt{c_*(1-c_*)}) \cup \{\pm 1\}\mathbf{1}_{c_*>1/2}$ and $i \neq j \in \{p, f\}$ we have:*

- $\mathcal{A}_i^* \mathbf{Q}_{\mathbf{ji}}^{\mathbf{W}} \mathcal{A}_i - \begin{pmatrix} -\dfrac{(1-c_N)(1+z\tilde{\mathbf{t}}_N(z))}{z\tilde{\mathbf{t}}_N(z)+1-c_N}\Delta_N^2 & 0 \\ 0 & \dfrac{1+\tilde{\mathbf{t}}_N(z)z}{c_N(1-c_N)}\Theta_N^*(I_L \otimes R_N^{-1})\Theta_N \end{pmatrix} \to 0$ *almost surely*
- $\mathcal{A}_f^* \mathbf{Q}_{\mathbf{pp}}^{\mathbf{W}} \mathcal{A}_p - \begin{pmatrix} -\dfrac{(1-c_N)^2}{z^2\tilde{\mathbf{t}}_N(z)}\Gamma_N & 0 \\ 0 & 0 \end{pmatrix} \to 0$ *almost surely*
- $\mathcal{A}_p^* \mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}} \mathcal{A}_f - \begin{pmatrix} -\dfrac{(1-c_N)^2}{z^2\tilde{\mathbf{t}}_N(z)}\Gamma_N^* & 0 \\ 0 & 0 \end{pmatrix} \to 0$ *almost surely*

Moreover, almost surely, the three convergence items hold uniformly on each compact subset of $\mathbb{C} \setminus \boldsymbol{\mathcal{S}}_*$.

Proof. The proof of this Lemma is postponed to the Section H.

We remind that $\Theta_N^*(I_L \otimes R_N^{-1})\Theta_N$ is denoted by G_N. After trivial algebra, Lemma (III.8) implies that asymptotically, for $N \to \infty$, the "limiting form" of Eq. (III.80) is

$$\det \begin{pmatrix} \dfrac{(1-c_N)c_N}{z\tilde{\mathbf{t}}_N(z)+1-c_N}\Delta_N^2 & I_r & \dfrac{(1-c_N)^2}{z^2\tilde{\mathbf{t}}_N(z)}\Gamma_N^* & 0 \\ I_r & -\dfrac{1-c_N+z\tilde{\mathbf{t}}_N(z)}{c_N(1-c_N)}G_N & 0 & 0 \\ \dfrac{(1-c_N)^2}{z^2\tilde{\mathbf{t}}_N(z)}\Gamma_N & 0 & \dfrac{(1-c_N)c_N}{z\tilde{\mathbf{t}}_N(z)+1-c_N}\Delta_N^2 & I_r \\ 0 & 0 & I_r & -\dfrac{1-c_N+z\tilde{\mathbf{t}}_N(z)}{c_N(1-c_N)}G_N \end{pmatrix} = 0 \tag{III.81}$$

Replacing $z\tilde{\mathbf{t}}_N(z)+1-c_N$ by $z^2 c_N \mathbf{t}_N(z)$ (see (III.63)) and taking the limits of the various terms when $N \to +\infty$ (due to Assumptions II.2, II.3, III.1), we can expect that the solutions of equation (III.80) tend to the solutions of the limiting equation, i.e.

$$\det \begin{pmatrix} \dfrac{1-c_*}{y^2\mathbf{t}_*(y)}\Delta_*^2 & I_r & \dfrac{(1-c_*)^2}{y^2\tilde{\mathbf{t}}_*(y)}\Gamma_*^* & 0 \\ I_r & -\dfrac{y^2\mathbf{t}_*(y)}{1-c_*}G_* & 0 & 0 \\ \dfrac{(1-c_*)^2}{y^2\tilde{\mathbf{t}}_*(y)}\Gamma_* & 0 & \dfrac{1-c_*}{y^2\mathbf{t}_*(y)}\Delta_*^2 & I_r \\ 0 & 0 & I_r & -\dfrac{y^2\mathbf{t}_*(y)}{1-c_*}G_* \end{pmatrix} = 0. \tag{III.82}$$

We now study the solutions of (III.82). If we interchange the second and third row blocks and second and third column blocks, the determinant will not change and using the Schur complement formula, the l.h.s. of (III.82) becomes

$$\det \begin{pmatrix} -\frac{y^2 \mathbf{t}_*(y)}{1-c_*} G_* & 0 \\ 0 & -\frac{y^2 \mathbf{t}_*(y)}{1-c_*} G_* \end{pmatrix} \times$$

$$\det \left[\begin{pmatrix} \frac{1-c_*}{y^2 \mathbf{t}_*(y)} \Delta_*^2 & \frac{(1-c_*)^2}{y^2 \tilde{\mathbf{t}}_*(y)} \Gamma_*^* \\ \frac{(1-c_*)^2}{y^2 \tilde{\mathbf{t}}_*(y)} \Gamma_* & \frac{1-c_*}{y^2 \mathbf{t}_*(y)} \Delta_*^2 \end{pmatrix} - \begin{pmatrix} -\frac{y^2 \mathbf{t}_*(y)}{1-c_*} G_* & 0 \\ 0 & -\frac{y^2 \mathbf{t}_*(y)}{1-c_*} G_* \end{pmatrix}^{-1} \right]$$

Since $\det \begin{pmatrix} -\frac{y^2 \mathbf{t}_*(y)}{1-c_*} G_* & 0 \\ 0 & -\frac{y^2 \mathbf{t}_*(y)}{1-c_*} G_* \end{pmatrix} \neq 0$, Eq. (III.82) is equivalent to

$$\det \begin{pmatrix} \frac{1-c_*}{y^2 \mathbf{t}_*(y)} (\Delta_*^2 + G_*^{-1}) & \frac{(1-c_*)^2}{y^2 \tilde{\mathbf{t}}_*(y)} \Gamma_*^* \\ \frac{(1-c_*)^2}{y^2 \tilde{\mathbf{t}}_*(y)} \Gamma_* & \frac{1-c_*}{y^2 \mathbf{t}_*(y)} (\Delta_*^2 + G_*^{-1}) \end{pmatrix} = 0$$

Using again the Schur complement formula, we obtain that the limiting form of Eq. (III.80) is

$$\det \left(\frac{(1-c_*)^2}{y^4 \mathbf{t}_*^2(y)} (\Delta_*^2 + G_*^{-1}) - \frac{(1-c_*)^4}{y^4 \tilde{\mathbf{t}}_*^2(y)} \Gamma_*^* (\Delta_*^2 + G_*^{-1})^{-1} \Gamma_* \right) = 0,$$

or equivalently,

$$\det \left(\frac{1}{(1-c_*)^2} \frac{\tilde{\mathbf{t}}_*^2(y)}{\mathbf{t}_*^2(y)} - \Gamma_*^* (\Delta_*^2 + G_*^{-1})^{-1} \Gamma_* (\Delta_*^2 + G_*^{-1})^{-1} \right) = 0. \tag{III.83}$$

We write that $\tilde{\mathbf{t}}_*(y) = y \tilde{t}_*(y^2)$ and $\mathbf{t}_*(y) = t_*(y^2)$, and put $x = y^2 \in (4c_*(1-c_*), 1)$. Then, using (III.70), Eq. (III.83) leads to equation (III.73).

In order to complete the proof of Theorem III.2, it remains to resort to the stability arguments in [6] and [10]. For this, it is sufficient to use exactly the same arguments as in the proof of Corollary II.2. We thus omit the details. We just justify the statements related to the number of eigenvalues located into $[1-\delta, 1]$ when $c_* > \frac{1}{2}$. Lemma III.1 implies that 1 is eigenvalue of $\Pi_p^W \Pi_f^W$ with multiplicity $2ML - N$. As $\Pi_p \Pi_f$ is a finite rank perturbation of $\Pi_p^W \Pi_f^W$, 1 is eigenvalue of $\Pi_p \Pi_f$ with a multiplicity equal to $2ML - N + \mathcal{O}(1)$. The stability arguments in [6] and [10] do not preclude the existence of other eigenvalues of $\Pi_p \Pi_f$ that converge towards 1. As the eigenvalue distribution of $\Pi_p \Pi_f$ has the same limit as the eigenvalue distribution of $\Pi_p^W \Pi_f^W$, i.e. measure $\tilde{\nu}_*$, for each $\delta > 0$ small enough, $\frac{1}{N} \#\{\lambda_i(\Pi_p \Pi_f) \in [\delta, 1]\} \to \tilde{\nu}_*([\delta, 1]) = 2c_* - 1$. Therefore, the number of remaining eigenvalues converging towards 1 is a $o(N)$ term, as expected. ■

Theorem III.2 allows to derive immediately the conditions under which it is possible to estimate consistently P by the number of eigenvalues of $\Pi_p \Pi_f$ that escape from $\mathcal{S}_*$.

Corollary III.4. *P coincides with the number of eigenvalues that escape from $\mathcal{S}_*$ if and only if $c_* < \frac{1}{2}$ and if the P non zero eigenvalues of F_* are strictly larger than $\frac{c_*}{1-c_*}$*

The condition that the non zero eigenvalues of F_* are bigger than $\frac{c_*}{1-c_*}$ implies that the singular values of Ω_* and the eigenvalues of Δ_* are large enough. In practice, this means that the canonical correlation coefficients between the past and the future of u are large enough (thus making the singular values of Ω_* large) and the r eigenvalues of $R_{u,N}^L$ are also large enough (thus making matrix Δ_*^{-1} small). It is interesting to notice that if $c_* > \frac{1}{2}$, the largest eigenvalues of $\Pi_p \Pi_f$ cannot be used to estimate P.

We finally mention that, as in the context of Corollary II.3, Theorem III.2 can be formulated in terms of the finite N equivalents of matrix F_* and function $f_*(z)$ defined by

$$F_N = \Delta_N^{-1}\Gamma_N\Delta_N^{-1}(I_r + \Delta_N^{-1}G_N^{-1}\Delta_N^{-1})^{-1}\Delta_N^{-1}\Gamma_N\Delta_N^{-1}(I_r + \Delta_N^{-1}G_N^{-1}\Delta_N^{-1})^{-1} \tag{III.84}$$

and

$$f_N(x) = x\left(\frac{\tilde{t}_N(x)}{(1-c_N)t_N(x)}\right)^2 \tag{III.85}$$

It is easily seen that the properties of function f_N are similar to the properties of f_* stated in item (i) of Theorem III.2, parameter c_N replacing c_*. We thus have the following result.

Corollary III.5. *If $c_* < \frac{1}{2}$, and if $\delta > 0$ is small enough, for N large enough, s coincides with the number of solutions of the equation $\det(f_N(x) - F_N) = 0$ that belong to $(4c_N(1-c_N)+\delta, 1)$, as well as with the number of eigenvalues of F_N that are strictly larger than $\frac{c_N}{1-c_N} + \kappa < 1$ for some $\kappa > 0$ small enough. If $\rho_{1,N}, \ldots, \rho_{s,N}$ are the corresponding solutions, then $\rho_{1,N}, \ldots, \rho_{s,N}$ converge almost surely towards $\rho_{1,*}, \ldots, \rho_{s,*}$. The s largest eigenvalues of $\Pi_p\Pi_f$ have the same asymptotic behaviour than $\rho_{1,N}, \ldots, \rho_{s,N}$, and for each $\delta > 0$, almost surely, for N large enough, the remaining $N-s$ ones belong to $[0, 4c_N(1-c_N)+\delta]$.*

We illustrate the above discussion by numerical experiments showing that eigenvalues outside the bulk indeed tend to thesolutions of equation (III.73). We consider a simple case, when $P = 2$, $K = 1$ and A is diagonal with eigenvalues a_1 and a_2. Figures 4, 5 represent histograms of the eigenvalues of realizations of the matrix $(\hat{R}^L_{f,y})^{-1/2}\hat{R}^L_{f|p,y}(\hat{R}^L_{p,y})^{-1}\hat{R}^{L*}_{f|p,y}(\hat{R}^L_{f,y})^{-1/2}$, as well as the graph of the density of measure $\nu_N = \frac{1}{c_N}\tilde{\nu}_N - \frac{1-c_N}{c_N}\delta_0$ and the solutions of equation (III.73).

We take $N = 2000$, $M = 130$ and $L = 4$, so $c_N = 0.26$. 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 $k = 1, \ldots, M$, so that matrix R_N verifies $\frac{1}{M}\mathrm{Tr}(R_N) \simeq 1$. Figure 4 corresponds to a choice of (a_1, a_2) for which $s = 1$, while $s = 2$ in the context of Figure 5.

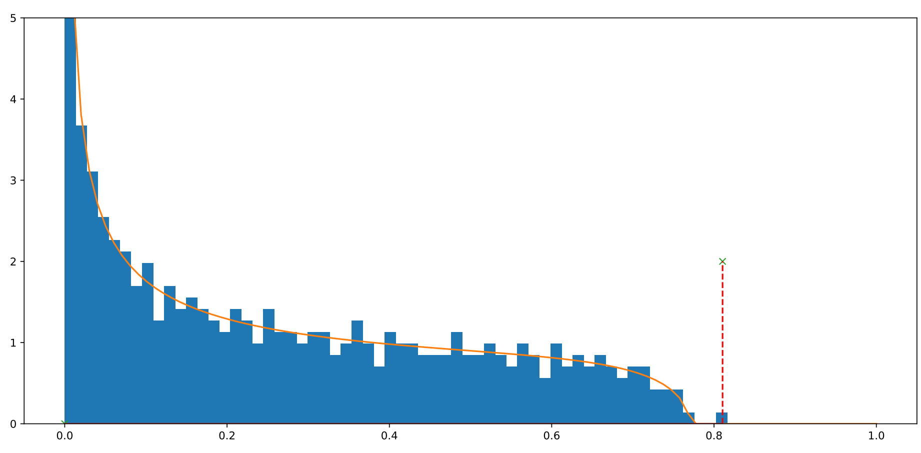

Figure 4. Histogram of the eigenvalues and graph of the density of ν_N with 1 outlier

C. When condition (II.33) does not hold.

We briefly justify that Theorem III.2 remains valid when some of the entries of Δ_* coincide. For this, we use the same notations as in Section II-F. The reader may check that when condition (II.33) does not hold, the limiting

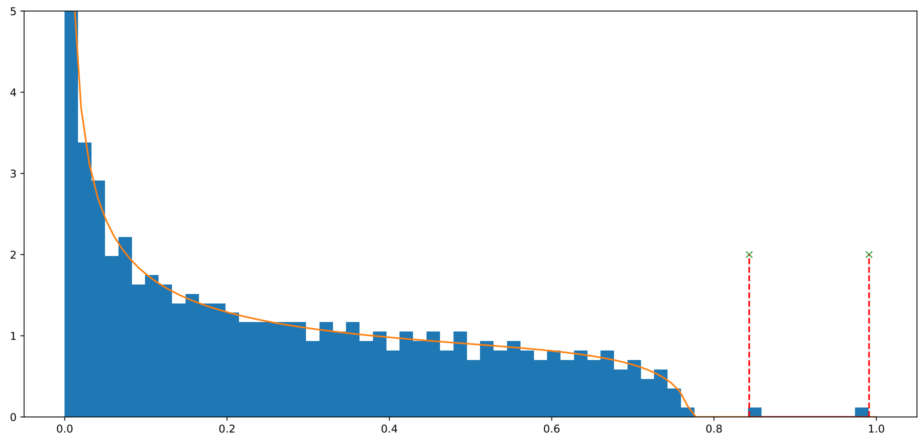

Figure 5. Histogram of the eigenvalues and graph of the density of ν_N with 2 outliers

equation (III.82) is replaced by

$$\det\begin{pmatrix} \frac{1-c_*}{y^2\mathbf{t}_*(y)}\Delta_*^2 & I_r & \frac{(1-c_*)^2}{y^2\tilde{\mathbf{t}}_*(y)}X_{p,N}^{-1}\Gamma_*^*X_{f,N}^{-*} & 0 \\ I_r & -\frac{y^2\mathbf{t}_*(y)}{1-c_*}X_{p,N}^*G_*X_{p,N} & 0 & 0 \\ \frac{(1-c_*)^2}{y^2\tilde{\mathbf{t}}_*(y)}X_{f,N}^{-1}\Gamma_*X_{p,N}^{-*} & 0 & \frac{1-c_*}{y^2\mathbf{t}_*(y)}\Delta_*^2 & I_r \\ 0 & 0 & I_r & -\frac{y^2\mathbf{t}_*(y)}{1-c_*}X_{f,N}^*G_*X_{f,N} \end{pmatrix} = 0. \tag{III.86}$$

Following the same steps as in the proof of Theorem III.2, we obtain that (III.86) is equivalent to

$$\det\left(\frac{1}{(1-c_*)^2}\frac{\tilde{\mathbf{t}}_*^2(y)}{\mathbf{t}_*^2(y)} - X_p^{-1}\Gamma_*^*X_f^{-*}(\Delta_*^2+(X_f^*G_*X_f)^{-1})^{-1}X_f^{-1}\Gamma_*X_p^{-*}(\Delta_*^2+(X_p^*G_*X_p)^{-1})^{-1}\right) = 0 \tag{III.87}$$

or to

$$\det\left(\frac{1}{(1-c_*)^2}\frac{\tilde{\mathbf{t}}_*^2(y)}{\mathbf{t}_*^2(y)} - \Gamma_*^*X_f^{-*}(\Delta_*^2+(X_f^*G_*X_f)^{-1})^{-1}X_f^{-1}\Gamma_*X_p^{-*}(\Delta_*^2+(X_p^*G_*X_p)^{-1})^{-1}X_p^{-1}\right) = 0 \tag{III.88}$$

We remark that for $i = p, f$

$$X_i^{-*}(\Delta_*^2+(X_i^*G_*X_i)^{-1})^{-1}X_i^{-1} = \left(X_i\Delta_*^2X_i^*+G_*^{-1}\right)^{-1} = \left(\Delta_*^2X_iX_i^*+G_*^{-1}\right)^{-1}$$

because we recall that $X_i\Delta_* = \Delta_*X_i$. As $X_{i,N}X_{i,N}^* \to I_r$ (see (II.90)), it appears the limiting form of (III.87) is (III.83), i.e. the final equation derived in the proof of Theorem III.2. Using again the stability arguments in [6] and [10], we deduce that Theorem III.2 remains valid.

D. Example.

We now consider the particular models defined by (II.95) and assume that $R_N = \sigma^2 I_M$. We use again the notations introduced to derive the properties of (II.95), and evaluate the conditions under which $s = P = 1$. For this, we have first to compute matrix Ω_*. We notice that

$$(R_{u,N})^{\#1/2}R_{f|p}(R_{u,N})^{\#1/2} = \Theta_N\Delta_*^{-1}\Gamma_*\Delta_*^{-1}\Theta_N^* = \Theta_N\Delta_*^{-1}\kappa\Upsilon\tilde{\Upsilon}^*\Delta_*^{-1}\Theta_N$$

where we recall that Υ coincides with the first vector e_1 of the canonical basis of $\mathbb{C}^r$. Therefore, a simple calculation leads to the conclusion that matrix $\tilde{\Theta}_{f,N}^* \tilde{\Theta}_{p,N}$ converges towards Ω_* given by

$$\Omega_* = \frac{1}{\delta_1} e_1 \left(a\delta_1^2, b_1\delta_2, \ldots, b_K\delta_{K+1}\right) \Delta_*^{-1} = e_1 \left(a, \frac{b_1}{\delta_1}, \ldots, \frac{b_K}{\delta_1}\right)$$

The non zero singular value of Ω_* is thus equal $a^2 + \frac{\|b\|^2}{\delta_1^2}$, which, by (II.96), coincides with 1. We notice that this is not surprising because it is easily seen that the intersection of the row spaces of matrices U_p and U_f is not reduced to 0, and coincides with the one dimensional space generated by $(x_2, \ldots, x_{N+1})$. As $G_* = \frac{I}{\sigma^2}$, matrix F_* is thus given by

$$F_* = \frac{1}{1 + \frac{\sigma^2}{\delta_1^2}} \begin{pmatrix} a \\ \frac{b_1}{\delta_1} \\ \vdots \\ \frac{b_K}{\delta_1} \end{pmatrix} \left(a, \frac{b_1}{\delta_1}, \ldots, \frac{b_K}{\delta_1}\right) \left(I + \sigma^2 \Delta_*^{-2}\right)^{-1} \tag{III.89}$$

and the non zero eigenvalue $\lambda_1(F_*)$ of F_* is given by

$$\lambda_1(F_*) = \left(\frac{1}{1 + \frac{\sigma^2}{\delta_1^2}}\right) \left[\left(\frac{a^2}{1 + \frac{\sigma^2}{\delta_1^2}}\right) + \sum_{k=1}^{K} \frac{b_k^2}{\delta_1^2} \frac{1}{1 + \frac{\sigma^2}{\delta_{k+1}^2}}\right]$$

We conclude that $s = P = 1$ if and only $c_* < \frac{1}{2}$ and $\lambda_1(F_*) > \frac{c_*}{1-c_*}$. In order to get more insights on this condition, we assume that the $(\delta_k)_{k=1,r}$ all coincide with δ. In this context, the ratio $\frac{\delta^2}{\delta^2}$ can be interpreted as the signal to noise ratio. Then, $\lambda_1(F_*) > \frac{c_*}{1-c_*}$ is equivalent to

$$\left(\frac{1}{1 + \frac{\sigma^2}{\delta^2}}\right)^2 > \frac{c_*}{1 - c_*} \tag{III.90}$$

or to

$$\frac{\delta^2}{\sigma^2} > \frac{1}{\left(\frac{1-c_*}{c_*}\right)^{1/2} - 1} = \frac{\sqrt{c_*}}{\sqrt{1 - c_*} - \sqrt{c_*}} = \frac{c_* + \sqrt{c_*(1 - c_*)}}{1 - 2c_*} \tag{III.91}$$

It is interesting to notice that for $i = p, f$, $Y_{i,N} = U_{i,N} + V_{i,N}$, where $\frac{U_{i,N} U_{i,N}^*}{N}$ is a rank r matrix whose r non zero eigenvalues converge towards δ^2. Therefore, usual results related to spiked models imply that the r largest eigenvalues of $\frac{Y_{i,N} Y_{i,N}^*}{N}$ escape from the support of the Marcenko-Pastur distribution $[\sigma^2(1 - \sqrt{c_*})^2, \sigma^2(1 + \sqrt{c_*})^2]$ if and only if the signal to noise ratio $\frac{\delta^2}{\sigma^2}$ is larger than the threshold $\sqrt{c_*}$. Not surprisingly, condition (III.91) appears stronger than $\frac{\delta^2}{\sigma^2} > \sqrt{c_*}$. However, if c_* is small enough, $\frac{\sqrt{c_*}}{\sqrt{1-c_*} - \sqrt{c_*}} \simeq \sqrt{c_*}$ and the 2 conditions are nearly equivalent.

IV. Monte Carlo Simulations

Our theoretical results allow to evaluate the number s of eigenvalues of $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$ and of $\Pi_p \Pi_f$ that escape from the support of the limit eigenvalue distribution of $W_f W_p^* W_p W_f^*$ and $\Pi_p^W \Pi_f^W$ respectively. In this section, using Monte Carlo simulation results, we evaluate the behaviour of two estimates of s, and check whether the true value of s is in practice well estimated. We still consider the simple model defined by (II.95), and choose the various parameters in such a way that $s = 2r - 1$ and $s = P = 1$ in the context of matrices $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$ and $\Pi_p \Pi_f$ respectively. More precisely, we take $c_N = 0.25$, $R_N = I_M$ (that is $\sigma = 1$), $K = 2$ and therefore $r = K + 1 = 3$ and $s = 5$. a is chosen equal to 0.2, and we choose $\delta_1 = \delta_2 = \delta_3 = \delta$ and $b_1 = b_2 = b = \frac{1}{\sqrt{2}} \delta (1 - a^2)^{1/2}$. δ is chosen equal to $\delta = (w_{+,N} - \sigma^2)^{1/2} + 0.3$ where $w_{+,N} = \sigma^2 \left(1 + \frac{1 + \sqrt{1 + 8c_N}}{2}\right)$, so that the signal to noise ratio $\frac{\delta^2}{\sigma^2}$ is equal to $3.3dB$. Our goal is twofold. While we know that $s = 2r - 1 = 5$, we first check that in the context of $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$, the probability of estimating s by $P = 1$ is very low, thus confirming that estimating P from the largest eigenvalues of $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$ is irrelevant both theoretically and practically. Second, in the context of matrix $\Pi_p \Pi_f$, we evaluate the empirical probability that the estimates of s take the value $s = P = 1$.

1000 realisations of matrices $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$ and $\Pi_p \Pi_f$ were generated. Table I reports the results corresponding to the estimation of s in the context of matrix $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$. The first estimate $\tilde{s}$ of s is the number of eigenvalues of $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$ that are larger than $x_{+,N}(1+\epsilon_1)$ for $\epsilon_1 = 0.01$. The second estimate, $\hat{s}$, already used in [51] and [30], is defined by

$$\hat{s} = \underset{k}{\text{argmin}} \Big\{ \frac{\lambda_{k+1}}{\lambda_k} > 1 - \epsilon_2 \Big\} - 1 \tag{IV.1}$$

for $\epsilon_2 = 0.05$. $\hat{s}$ appears to be more realistic than $\tilde{s}$ because, in practice, the noise variance σ^2, and thus $x_{+,N}$, are not necessarily known. Table I provides the empirical probabilities that $\tilde{s}$ and $\hat{s}$ equal to $0, 1, 2, 3, 4, 5, 6, 7, 8$ for various values of M and N and Figure 6 represents the ratios of eigenvalues λ_{i+1}/λ_i of a realisation of $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$ in terms of $i-1$ when $(M, N) = (600, 2400)$. Figure 6 indicates that the largest eigenvalue λ_1 is much larger

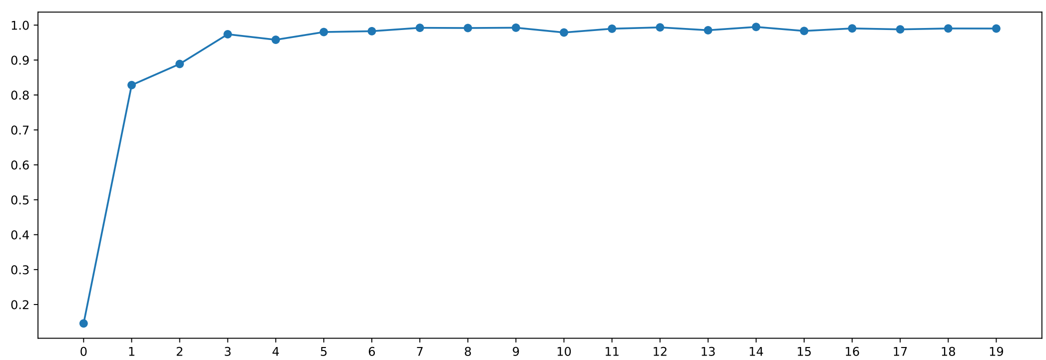

Figure 6. Ratios of eigenvalues λ_{i+1}/λ_i of $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$ w.r.t. $i-1$

than the next four ones because $\frac{\lambda_2}{\lambda_1} < 0.1$. Moreover, λ_2 is nearly equal to $1.3\,\lambda_3$, and the next eigenvalues appear to be much closer one from each others. This confirms what we already noticed in the context of the numerical experiment of Section II-G: as $c_N = \frac{1}{4}$ is rather small, the largest eigenvalue corresponds to the useful signal, and appears much larger than the other 4 spurious outliers. Table I tends to confirm that $(\lambda_i)_{i=3,4,5}$ are likely to be close from $x_{+,N}$ while λ_2 is very often significantly larger than $x_{+,N}$ thus explaining that $\tilde{s}$ and $\hat{s}$ do not take the value 1, and that $\tilde{s}$ and $\hat{s}$ take the values $2, 3, 4, 5$ (and $\hat{s}$ sometimes $6, 7, 8$) . These experiments tend to indicate that the true value of s is difficult to estimate, and more importantly, that the estimates are never equal to $P = 1$. This confirms that P cannot be estimated reliably from the largest eigenvalues of $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$.

Table I
BEHAVIOUR OF $\tilde{s}$ AND $\hat{s}$ FOR MATRIX $\Sigma_f \Sigma_p^* \Sigma_p \Sigma_f^*$

	M=100 N=400	M=200 N=800	M=400 N=1600	M=600 N=2400		M=100 N=400	M=200 N=800	M=400 N=1600	M=600 N=2400
$\tilde{s}=8$	0	0	0	0	$\hat{s}=8$	0.061	0	0	0
$\tilde{s}=7$	0	0	0	0	$\hat{s}=7$	0.128	0	0	0
$\tilde{s}=6$	0	0	0	0	$\hat{s}=6$	0.179	0.01	0	0
$\tilde{s}=5$	0	0.005	0.09	0.27	$\hat{s}=5$	0.25	0.335	0.097	0
$\tilde{s}=4$	0.235	0.56	0.86	0.72	$\hat{s}=4$	0.247	0.298	0.357	0.033
$\tilde{s}=3$	0.745	0.425	0.05	0.01	$\hat{s}=3$	0.12	0.21	0.32	0.287
$\tilde{s}=2$	0.02	0.01	0	0	$\hat{s}=2$	0.005	0.147	0.226	0.68
$\tilde{s}=1$	0	0	0	0	$\hat{s}=1$	0.01	0	0	0
$\tilde{s}$=0	0	0	0	0	$\hat{s}=0$	0	0	0	0

Table II is related to the estimation of s in the context of matrix $\Pi_p \Pi_f$. ϵ_1 and ϵ_2 being still equal to 0.01 and 0.05, $\tilde{s}$ represents this time the number of eigenvalues of $\Pi_p \Pi_f$ that are larger than $4c_N(1-c_N)(1+\epsilon_1)$, while $\hat{s}$ is defined by

$$\hat{s} = \underset{k}{\text{argmin}} \Big\{ \frac{\lambda_{k+1}}{\lambda_k} > 1 - \epsilon_2 \Big\} - 1$$

We also represent Fig. 7 the ratios of eigenvalues $\frac{\lambda_{i+1}}{\lambda_i}$ of a realisation of $\Pi_p \Pi_f$ in terms of $i-1$ when $(M,N) = (600, 2400)$. The largest eigenvalue appears significantly larger than λ_2, and the other eigenvalues are quite close one from each others. This behaviour is confirmed by the behaviour of $\tilde{s}$ and $\hat{s}$ which take the value 1 with high probability, thus confirming the relevance of the estimate of P based on the largest eigenvalues of $\Pi_p \Pi_f$. We notice that in the context of matrix $\Pi_p \Pi_f$, the estimate $\tilde{s}$ is in practice relevant because $4c_N(1-c_N)$ is of course known.

Table II
BEHAVIOUR OF $\tilde{s}$ AND $\hat{s}$ FOR MATRIX $\Pi_p \Pi_f$

	M=100 N=400	M=200 N=800	M=400 N=1600	M=600 N=2400		M=100 N=400	M=200 N=800	M=400 N=1600	M=600 N=2400
$\tilde{s}=4$	0	0	0	0	$\hat{s}=4$	0	0.007	0.007	0.008
$\tilde{s}=3$	0	0	0	0	$\hat{s}=2$	0	0.013	0.008	0.017
$\tilde{s}=2$	0.001	0	0	0	$\hat{s}=2$	0.07	0.012	0	0.01
$\tilde{s}=1$	0.999	1	1	1	$\hat{s}=1$	0.91	0.966	0.974	0.965
$\tilde{s}=0$	0	0	0	0	$\hat{s}=0$	0.02	0.002	0.011	0

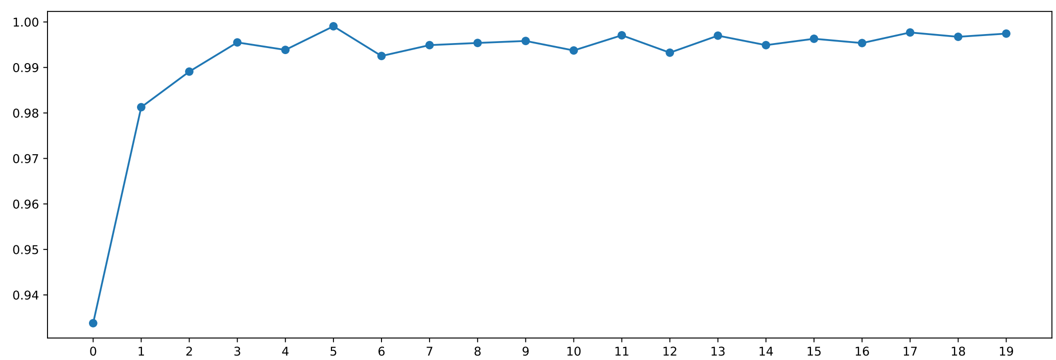

Figure 7. Ratios of eigenvalues λ_{i+1}/λ_i of $\Pi_{p,y}\Pi_{f,y}$ w.r.t. $i-1$

V. CONCLUSION

In this paper, motivated by the problem of estimating consistently the minimal dimension P of the state space realizations of the high-dimensional time series y, we have studied the behaviour of the largest singular values of the empirical autocovariance matrix $\hat{R}^L_{f|p,y}$ as well as of its normalized version $(\hat{R}^L_{f,y})^{-1/2}\hat{R}^L_{f|p,y}(\hat{R}^L_{p,y})^{-1/2}$. In the high-dimensional asymptotic regime defined in Section I-C, and under certain technical assumptions, we have shown that all the singular values of $\hat{R}^L_{f|p,y}$ are less than a certain threshold, except a finite number s of outliers. Unfortunately, s is not related to P, and, when $P=1$, we have built simple examples for which s can take any odd value. We also showed that the singular values of the normalized matrix $(\hat{R}^L_{f,y})^{-1/2}\hat{R}^L_{f|p,y}(\hat{R}^L_{p,y})^{-1/2}$ lie almost surely in a neighbourhood of the interval $[0, 2\sqrt{c_*(1-c_*)}]$, but this time, we proved that the number s of outliers belong to $\{0,1,\ldots,P\}$, and that $s=P$ if $c_* < \frac{1}{2}$ and if the P non zero eigenvalues of the rank P matrix F_* defined by (III.71) are larger than $\frac{c_*}{1-c_*}$. Under this condition, which, in practice, means that the useful signal u is powerful enough and its non zero canonical correlation coefficients between the past and the future are large enough, P can be estimated consistently by the number of singular values of $(\hat{R}^L_{f,y})^{-1/2}\hat{R}^L_{f|p,y}(\hat{R}^L_{p,y})^{-1/2}$ that are larger than $2\sqrt{c_*(1-c_*)}(1+\epsilon)$ for a certain parameter ϵ small enough. These results are established using a general approach already proposed in the literature in the context of simple large random matrix models. However, the random matrix models considered in the present paper are quite complicated, and we needed to solve a number of non obvious new technical issues. We have also provided numerical simulation results that confirm the practical relevance of our theoretical results. We finally mention that the existence of a consistent estimate of P allows to consider the problem of estimating other parameters of the state space realizations of the useful signal u. This is a topic for future research.

References

[1] K. Abed-Meraim, P. Loubaton, E. Moulines, "A subspace algorithm for certain blind identification problems", *IEEE Inf. Theory*, Vol. 43 (4), March 1997, pp. 499-511.

[2] Z.D. Bai, J.W. Silverstein, "Spectral Analysis of Large Random Matrices", Springer, 2010.

[3] J. Baik, G. Ben Arous, S. Péché, "Phase transition of the largest eigenvalue for non-null complex sample covariance matrices", Annals of Proba., vol. 33, no. 5 (2005), pp. 1643-1697.

[4] J. Baik, J.W. Silverstein, "Eigenvalues of large sample covariance matrices of spiked population models", Journal of Multivariate Analysis, vol. 97, no. 6 (2006), pp. 1382-1408.

[5] Z. Bao, J. Hu, G. Pan, W. Zhou, "Canonical correlation coefficients of high-dimensional Gaussian vectors: Finite rank case", Ann. of Stat., vol. 47, no. 1, pp. 612-640, 2019.

[6] 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.

[7] 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.

[8] P. Bianchi, M. Debbah, M. Maeda and J. Najim, "Performance of statistical tests for single source detection using random matrix theory", *IEEE Inf. Theory*, Vol. 57 (4), April 2011, pp. 2400–2419.

[9] 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.

[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] J. Conway, "Functions of One Complex Variable I", Springer-Verlag, 1978.

[13] R. Couillet, M. Debbah, "Random Matrix Methods for Wireless Communications", Cambridge University Press, 2011.

[14] R. Couillet, "Robust spiked random matrices and a robust G-MUSIC estimator", Journal of Multivariate Analysis, vol. 140, pp. 139-161, 2015.

[15] R. Couillet, F. Pascal, J. W. Silverstein, "The random matrix regime of Maronna’s M-estimator with elliptically distributed samples", Elsevier Journal of Multivariate Analysis, vol. 139, pp. 56-78, 2015.

[16] M. Deistler, B.D.O. Anderson, A. Filler, W. Chen, "Generalized linear dynamic factor models: an approach via singular autoregressions", European J. on Control, (2010), 3(211-224).

[17] M. Forni, M. Hallin, M. Lippi, L. Reichlin, "The generalized dynamic factor model: Identification and estimation", Rev. Econ. Stat., vol. 82, no. 4, pp. 540-554, 2000.

[18] M. Forni, M. Hallin, M. Lippi, L. Reichlin, "The generalized dynamic factor model, consistency and rates", Journal of Econometrics 119 (2), pp. 231-255, 2004.

[19] F. Gesztesy, E. Tsekanovskii, “On matrix-valued Herglotz functions”, Math. Nach. 218 61-138, 2000.

[20] 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.

[21] 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.

[22] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, L. Pastur, "A New Approach for Capacity Analysis of Large Dimensional Multi-Antenna Channels", IEEE Transactions on Information Theory, vol. 54, no. 9, pp. 3987-4004, September 2008.

[23] W. Hachem, P. Loubaton, X. Mestre, J. Najim, P. Vallet, "Large information plus noise random matrix models and consistent subspace estimation in large sensor networks", Random Matrices: Theory and Applications, Vol. 1, No. 2, April 2012.

[24] 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

[25] F. Hiai, D. Petz, "The Semicircle Law, Free Random Variables and Entropy”, Mathematical Survey and Monographs, vol. 77, American Mathematical Society, Providence, RI, 2000.

[26] R.A. Horn, C.R. Johnson, “Matrix Analysis”, Cambridge University Press, 1985.

[27] T. Kailath, "Linear Systems", Prenctice-Hall, 1980.

[28] T. Kato, "Perturbation Theory for Linear Operators", Second Edition, Springer Verlag 1980.

[29] S. Kritchman, B. Nadler, "Non-parametric detection of the number of signals, hypothesis testing and random matrix theory, *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3930–3941, 2009.

[30] 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.

[31] A. Lindquist, G. Picci, “Linear Stochastic Systems”, Series in Contemporary Mathematics, Vol. 1, Springer, 2015

[32] 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.

[33] P. Loubaton, D. Tieplova, "On the behaviour of large empirical autocovariance matrices between the past and the future", Random Matrices: Theory and Applications, vol. 10, no. 2, 2150021 (2021).

[34] X. Mestre, M.A. Lagunas, "Modified subspace algorithms for DoA estimation with large arrays", *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 598-614, February 2008.

[35] J. Mingo, R. Speicher, “Free Probability and Random Matrices”, Fields Institute Monographs, Springer, 2017.

[36] 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.

[37] 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.
[38] L. Pastur, "A simple approach to the global regime of Gaussian ensembles of random matrices", Ukrainian Math. J., 57, no. 6, 936-966 (2005).
[39] L.A. Pastur, M. Shcherbina, "Eigenvalue Distribution of Large Random Matrices", Mathematical Surveys and Monographs, Providence: American Mathematical Society, 2011.
[40] D. Paul, “Asymptotics of sample eigenstruture for a large dimensional spiked covariance model”, Statistica Sinica, Vol. 17, No. 4, 1617-1642, 2007.
[41] A. Rosuel, P. Vallet, P. Loubaton, X. Mestre, "On the detection of low-rank signal in the presence of spatially uncorrelated noise: a frequency domain approach", IEEE Transactions on Signal Processing, 69, art. no. 9496131, pp. 4458-4473, 2021.
[42] Yu.A. Rozanov, "Stationary Random Processes", Holden Day, 1967.
[43] 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).
[44] D. Tieplova, P. Loubaton, L. Pastur, "On the limit distribution of the canonical correlation coefficients between the past and the future of a high-dimensional white noise", in Proc. of the 2020 International Conference on Acoustic, Speech and Signal Processing, 4-8 May 2020, Barcelona, pp. 8772-8776.
[45] D. Tieplova, "Application of Large Random Matrices to Multivariate Time Series Analysis", Joint PHD Thesis Université Paris-Est Marne la Vallée and Boris Verkin Institute for Low Temperature Physics and Engineering, December 2020, can be downloaded at the URL http://www-syscom.univ-mlv.fr/ loubaton/hiditsa
[46] P. Vallet, X. Mestre, P. Loubaton, "Performance analysis of an improved MUSIC DoA estimator", IEEE Trans. on Signal Processing, vol. 63, no. 23, pp. 6407-6422, December 1 2015.
[47] A. Tulino, S. Verdu, "Random Matrix Theory and Wireless Communications", Fundations and Trends in Communications and Information Theory, vol. 1, no. 1, 2004 (1-182).
[48] 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.
[49] P. Van Overschee, B. de Moor, "Subspace Identification for Linear Systems: Theory, Implementation, Applications", Kluwer Academic Publishers, 1996.
[50] 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), pp. 4903-4913, September 2015.
[51] 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, pp. 5633-5645.
[52] Voiculescu, D. V., Dykema, K. J., Nica, A.: Free Random Variables. A Noncommutative Probability Approach to Free Products with Applications to Random Matrices, Operator Algebras and Harmonic Analysis on Free Groups, CRM Monograph Series, Vol. 1, American Mathematical Society, Providence, RI, (1992).
[53] K. W. Wachter, "The limiting empirical measure of multiple discriminant ratios", Ann. Statist., vol. 8, no. 5, pp. 937-957, 1980.
[54] Y. R. Yang, G. M. Pan, "The convergence of the empirical distribution of canonical correlation coefficients", Electron. J. Proba., 17: 1-13 (2012).
[55] Y. Yang, G. M. Pan, "Independence test for high dimensional data based on regularized canonical correlation coefficients", Ann. of Stat., 43(2), pp. 467-500, 2015.
[56] J. Yao, S. Zheng, Z.B. Bai, "Large Sample Covariance Matrices and High-Dimensional Data Analysis", Cambridge University Press, 2015.

APPENDIX

A. Proof of Lemma III.4

To prove that matrices $\mathbb{E}\{\mathbf{Q_{ij}}\}$, $i, j = p, f$ are diagonal, we consider the new set of vectors $z_k = e^{-ik\theta} v_k$ and construct the matrices Z_p, Z_f in the same way as W_p and W_f. It is clear that sequence $(z_n)_{n\in\mathbb{Z}}$ has the same probability distribution that $(v_n)_{n\in\mathbb{Z}}$. Z_p and Z_f can be expressed as

$$
\begin{aligned}
Z_p &= \begin{pmatrix} e^{-i\theta} I_M & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{-Li\theta} I_M \end{pmatrix} W_p \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{-(N-1)i\theta} \end{pmatrix}, \\
Z_f &= e^{-Li\theta} \begin{pmatrix} e^{-i\theta} I_M & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{-Li\theta} I_M \end{pmatrix} W_f \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{-(N-1)i\theta} \end{pmatrix}.
\end{aligned}
$$

Then

$$Z_i Z_i^* = \begin{pmatrix} e^{-i\theta} I_M & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{-Li\theta} I_M \end{pmatrix} W_i W_i^* \begin{pmatrix} e^{i\theta} I_M & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{Li\theta} I_M \end{pmatrix}, \tag{A.1}$$

$$(Z_i Z_i^*)^{-1} = \begin{pmatrix} e^{-i\theta} I_M & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{-Li\theta} I_M \end{pmatrix} (W_i W_i^*)^{-1} \begin{pmatrix} e^{i\theta} I_M & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{Li\theta} I_M \end{pmatrix} \tag{A.2}$$

as well as $\phi(Z_f Z_f^*) = \phi(W_f W_f^*)$ and $\phi(Z_p Z_p^*) = \phi(W_p W_p^*)$ where ϕ is defined by (III.5). Therefore, η^z coincides with η. Next we define $\Pi_i^z = Z_i^*(Z_i Z_i^*)^{-1} Z_i$, $i = \{p, f\}$. The equality

$$\Pi_i^z = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{(N-1)i\theta} \end{pmatrix} \Pi_i \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{-(N-1)i\theta} \end{pmatrix} \tag{A.3}$$

holds for $i = \{p, f\}$. We define matrix $\mathbf{Q^Z} = \begin{pmatrix} -zI_{ML} & \eta^z \Pi_p^z \\ \eta^z \Pi_f^z & -zI_{ML} \end{pmatrix}^{-1}$ and obtain immediately that

$$\mathbb{E}\{\mathbf{Q^Z}\} = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \mathbb{E}\{\mathbf{Q}\} \begin{pmatrix} A^* & 0 \\ 0 & A^* \end{pmatrix}, \mathbb{E}\{\eta_N \mathbf{Q^Z}\} = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \mathbb{E}\{\eta_N \mathbf{Q}\} \begin{pmatrix} A^* & 0 \\ 0 & A^* \end{pmatrix}$$

where the $N \times N$ matrix A is defined as

$$A = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{(N-1)i\theta} \end{pmatrix}$$

Obviously for each $N \times N$ block $\mathbb{E}\{\mathbf{Q_{ij}}^z\}$, $i, j = \{p, f\}$, we have

$$\mathbb{E}\{\mathbf{Q_{ij}}^z\} = A\, \mathbb{E}\{\mathbf{Q_{ij}}\}\, A^*, \mathbb{E}\{\eta_N \mathbf{Q_{ij}}^z\} = A\, \mathbb{E}\{\eta_N \mathbf{Q_{ij}}\}\, A^*$$

and

$$\mathbb{E}\{\eta_N \Pi_h^z \mathbf{Q_{ij}}^z\} = A\, \mathbb{E}\{\eta_N \Pi_h \mathbf{Q_{ij}}\}\, A^*$$

for $h = \{p, f\}$. Since $\mathbb{E}\{\mathbf{Q^Z}\} = \mathbb{E}\{\mathbf{Q}\}$, $\mathbb{E}\{\eta_N \mathbf{Q^Z}\} = \mathbb{E}\{\eta_N \mathbf{Q}\}$, and $\mathbb{E}\{\eta_N \Pi_h^z \mathbf{Q_{ij}}^z\} = \mathbb{E}\{\eta_N \Pi_h \mathbf{Q_{ij}}\}$, for $1 \leq k, l \leq N$ and $i, j, h = \{p, f\}$, we have

$$\begin{aligned} &\mathbb{E}\{\mathbf{Q_{ij}}^{k,l}\} = e^{(k-1)i\theta} \mathbb{E}\{\mathbf{Q_{ij}}^{k,l}\} e^{-(l-1)i\theta} = e^{(k-l)i\theta} \mathbb{E}\{\mathbf{Q_{ij}}^{k,l}\} \\ &\mathbb{E}\{\eta_N \mathbf{Q_{ij}}^{k,l}\} = e^{(k-1)i\theta} \mathbb{E}\{\eta_N \mathbf{Q_{ij}}^{k,l}\} e^{-(l-1)i\theta} = e^{(k-l)i\theta} \mathbb{E}\{\eta_N \mathbf{Q_{ij}}^{k,l}\} \\ &\mathbb{E}\{\eta_N (\Pi_h \mathbf{Q_{ij}})^{k,l}\} = e^{(k-1)i\theta} \mathbb{E}\{\eta_N (\Pi_h \mathbf{Q_{ij}})^{k,l}\} e^{-(l-1)i\theta} = e^{(k-l)i\theta} \mathbb{E}\{\eta_N (\Pi_h \mathbf{Q_{ij}})^{k,l}\} \end{aligned}$$

This proves that $\mathbb{E}\{\mathbf{Q_{ij}}^{k,l}\} = 0$, $\mathbb{E}\{\eta_N \mathbf{Q_{ij}}^{k,l}\} = 0$ and $\mathbb{E}\{\eta_N (\Pi_h \mathbf{Q_{ij}})^{k,l}\} = 0$ if $k \neq l$, as expected. We can prove similarly that matrices $\mathbb{E}\{\eta_N (W_i W_i^*)^{-1}\}$, $\mathbb{E}\{\eta_N \Pi_i\}$ and $\mathbb{E}\{\eta_N \mathbf{Q_{ij}} W_h^* (W_h W_h^*)^{-2} W_h\}$ are diagonal. We just verify that $\mathbb{E}(\eta_N W_p^* (W_p W_p)^{-1}) = 0$, and omit the proof of $\mathbb{E}(\eta_N\, \mathbf{Q_{ij}}\, W_h^* (W_h W_h^*)^{-1}) = \mathbb{E}(\eta_N\, \Pi_k\, \mathbf{Q_{ij}}\, W_h^* (W_h W_h^*)^{-1}) = 0$. It is clear that

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

The equalities $\eta^z = \eta$ and $\mathbb{E}(\eta^z Z_p^* (Z_p Z_p^*)^{-1}) = \mathbb{E}(\eta W_p^* (W_p W_p^*)^{-1})$ lead immediately to $\left(\mathbb{E}(\eta W_p^* (W_p W_p^*)^{-1})\right)_{n,l} = e^{i(n-1+l)\theta} \left(\mathbb{E}(\eta W_p^* (W_p W_p^*)^{-1})\right)_{n,l}$ for each $1 \leq n \leq N$ and $1 \leq l \leq L$. As θ can be chosen arbitrarily, we obtain that $\left(\mathbb{E}(\eta W_p^* (W_p W_p^*)^{-1})\right)_{n,l} = 0$ as expected.

To prove (III.15) to (III.18), we consider the sequence $(z_n)_{n\in\mathbb{Z}}$ defined by $z_n = v_{-n+N+2L}$ for each n. Again, the distributions of z and v coincide, and it is easy to see that for $i \in \{p, f\}$, Z_i is given by

$$Z_i = \begin{pmatrix} 0 & \dots & I_M \\ \vdots & & \vdots \\ I_M & \dots & 0 \end{pmatrix} W_{\tilde{i}} \begin{pmatrix} 0 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 0 \end{pmatrix},$$

and as consequence

$$Z_i Z_i^* = \begin{pmatrix} 0 & \dots & I_M \\ \vdots & & \vdots \\ I_M & \dots & 0 \end{pmatrix} W_{\tilde{i}} W_{\tilde{i}}^* \begin{pmatrix} 0 & \dots & I_M \\ \vdots & & \vdots \\ I_M & \dots & 0 \end{pmatrix}$$

Therefore, $Z_i Z_i^*$ and $W_{\tilde{i}} W_{\tilde{i}}^*$ have the same eigenvalues, which implies that $\phi(Z_i Z_i^*) = \phi(W_{\tilde{i}} W_{\tilde{i}}^*)$, and that the new regularization term $\eta^z = \det\, \phi(Z_p Z_p^*) \det\, \phi(Z_f Z_f^*)$ will remain the same, i.e. $\eta^z = \eta$. It is easy to see that $\Pi_p^z = A \Pi_f A$ and $\Pi_f^z = A \Pi_p A$, where this time $A = \begin{pmatrix} 0 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 0 \end{pmatrix}$. From this, we obtain that

$$\mathbb{E}\{\mathbf{Q}^{\mathbf{Z}}\} = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \mathbb{E}\Big\{ \begin{pmatrix} -zI_N & \eta \Pi_f \\ \eta \Pi_p & -zI_N \end{pmatrix}^{-1} \Big\} \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}.$$

Using the inverse block matrix formula and the fact that $\mathbb{E}\{\mathbf{Q}^{\mathbf{Z}}\} = \mathbb{E}\{\mathbf{Q}\}$, we obtain that $\mathbb{E}\{\mathbf{Q_{pp}}\} = A\mathbb{E}\{\mathbf{Q_{ff}}\}A$ and $\mathbb{E}\{\mathbf{Q_{pf}}\} = A\mathbb{E}\{\mathbf{Q_{fp}}\}A$. This immediately implies that for every $1 \le k \le N$ and $h, i, j = \{p, f\}$ we have $\mathbb{E}\{(\mathbf{Q_{ij}})^{k,k}\} = \mathbb{E}\{(\mathbf{Q_{\tilde{i}\tilde{j}}})^{N+1-k,N+1-k}\}$ and $\mathbb{E}\{\eta_N (\Pi_h \mathbf{Q_{ij}})^{k,k}\} = \mathbb{E}\{\eta_N (\Pi_{\tilde{h}} \mathbf{Q_{\tilde{i}\tilde{j}}})^{N+1-k,N+1-k}\}$. As a consequence, $\mathbb{E}\{\mathrm{Tr}\mathbf{Q_{ij}}\} = \mathbb{E}\{\mathrm{Tr}\mathbf{Q_{\tilde{i}\tilde{j}}}\}$ and $\mathbb{E}\{\eta_N \Pi_h \mathbf{Q_{ij}}\} = \mathbb{E}\{\eta_N \Pi_{\tilde{h}} \mathbf{Q_{\tilde{i}\tilde{j}}}\}$ as expected.

B. Proof of Lemma III.5

The lemma is established using the integration by parts formula and the Poincaré-Nash inequality. As the partial derivatives of η with respect to elements of W_p, W_f will appear, we first state the following useful lemma. We recall that ϕ and $\mathcal{E}_N$ are defined respectively by (III.5) and (III.9).

Lemma A.1. *Let Ω be the event defined by:*

$$\Omega = \mathcal{E}_N \cap \{\textit{all eigenvalues of } W_p W_p^* \textit{ and } W_f W_f^* \in \mathrm{Supp}(\phi)\}. \tag{A.4}$$

Then it holds that

$$\frac{\partial \eta_N}{\partial W_{i,j}^m} = 0 \textit{ on } \Omega^c \tag{A.5}$$

and

$$\mathbb{E}\left\{ \left| \frac{\partial \eta_N}{\partial W_{i,j}^m} \right|^2 \right\} = \mathcal{O}\left(\frac{1}{N^k}\right) \tag{A.6}$$

for all $1 \le m \le M$, $1 \le i \le 2L$, $1 \le j \le N$ and each k.

The proof of the lemma is an adaptation of Lemma 11 and calculations from Proposition 4 of [23].

We just prove Lemma III.5 for $i = p$. In the following, we drop index i and denote $G = (WW^*)^{-1}$. To prove the lemma, we apply the integration by parts formula (I.25) to $\eta_N G_{i_1 i_2}^{m_1 m_2} W_{i_2,j_2}^{m_2} \bar{W}_{j_1,i_3}^{m_3}$ considered as a function of the entries of the $2ML \times N$ matrix $\overline{W}_N$ whose elements are the complex conjugates of those of W_N. We recall that the correlation structure of the elements of W_N is given by (III.3).

$$\begin{aligned} \mathbb{E}\{\eta_N G_{i_1 i_2}^{m_1 m_2} W_{i_2,j_2}^{m_2} \bar{W}_{j_1,i_3}^{m_3}\} &= \sum_{m',i',j'} \mathbb{E}\{\bar{W}_{j_1,i_3}^{m_3} W_{i',j'}^{m'}\} \\ &\times \left(\mathbb{E}\left\{ \frac{\partial \eta_N}{\partial W_{i',j'}^{m'}} G_{i_1 i_2}^{m_1 m_2} W_{i_2,j_2}^{m_2} \right\} + \mathbb{E}\left\{ \eta_N \frac{\partial G_{i_1 i_2}^{m_1 m_2}}{\partial W_{i',j'}^{m'}} W_{i_2,j_2}^{m_2} \right\} + \mathbb{E}\left\{ \eta_N G_{i_1 i_2}^{m_1 m_2} \frac{\partial W_{i_2,j_2}^{m_2}}{\partial W_{i',j'}^{m'}} \right\} \right) \end{aligned} \tag{A.7}$$

Lemma A.1 implies that the first term of the r.h.s of (A.7) is of order $\mathcal{O}(N^{-k})$ for each k. Indeed,

$$\mathbb{E}\left\{\frac{\partial \eta_N}{\partial W^{m'}_{i',j'}} G^{m_1 m_2}_{i_1 i_2} W^{m_2}_{i_2,j_2}\right\} = \mathbb{E}\left\{\mathbf{1}_\Omega \frac{\partial \eta_N}{\partial W^{m'}_{i',j'}} G^{m_1 m_2}_{i_1 i_2} W^{m_2}_{i_2,j_2}\right\}$$

where we recall that Ω is defined by (A.4). The Schwartz inequality leads to

$$\left|\mathbb{E}\left\{\mathbf{1}_\Omega \frac{\partial \eta_N}{\partial W^{m'}_{i',j'}} G^{m_1 m_2}_{i_1 i_2} W^{m_2}_{i_2,j_2}\right\}\right|^2 \leq \mathbb{E}\left\{\left|\frac{\partial \eta_N}{\partial W^{m'}_{i',j'}}\right|^2\right\} \mathbb{E}\left\{\left|\mathbf{1}_\Omega G^{m_1 m_2}_{i_1 i_2} W^{m_2}_{i_2,j_2}\right|^2\right\} \tag{A.8}$$

On Ω, the eigenvalues of WW^* belong to $((1-\sqrt{c_*})^2 - 2\epsilon, (1+\sqrt{c_*})^2 + 2\epsilon)$, so that $\|G\mathbf{1}_\Omega\|$ and $\|W\mathbf{1}_\Omega\|$ are bounded by a nice constant. Therefore $\left|\mathbf{1}_\Omega G^{m_1 m_2}_{i_1 i_2} W^{m_2}_{i_2,j_2}\right|$ has the same property. After some calculations, (A.7) becomes

$$\begin{aligned}\mathbb{E}\{\eta_N G^{m_1 m_2}_{i_1 i_2} W^{m_2}_{i_2,j_2} \bar{W}^{m_3}_{j_1,i_3}\} &= \frac{1}{N} \sum_{m',i',j'} \delta_{m',m_3} \delta_{i_3+j_1,i'+j'} \\ &\times \left(-\mathbb{E}\left\{\eta_N G^{m_1 m'}_{i_1 i'} (W^* G)^{m_2}_{j',i_2} W^{m_2}_{i_2,j_2}\right\} + \mathbb{E}\{\eta_N G^{m_1 m_2}_{i_1 i_2} \delta_{m',m_2} \delta_{i_2,i'} \delta_{j_2,j'}\}\right) + \mathcal{O}\left(\frac{1}{N^k}\right)\end{aligned} \tag{A.9}$$

Defining $l = i_3 - i' = j' - j_1$ which runs from $-L+1$ to $L-1$ and taking into account (I.12), we get $\delta_{m',m_3}\delta_{i_3+j_1,i'+j'} = (J_L^{(l)} \otimes I_M)^{m' m_3}_{i' i_3} (J_N^{(l)})_{j_1 j'}$. Then, after summing over i', j' and m', (A.9) becomes

$$\begin{aligned}\mathbb{E}\{\eta_N G^{m_1 m_2}_{i_1 i_2} W^{m_2}_{i_2,j_2} \bar{W}^{m_3}_{j_1,i_3}\} &= -\frac{1}{N} \mathbb{E}\left\{\eta_N \left(G(J_L^{(l)} \otimes I_M)\right)^{m_1 m_3}_{i_1 i_3} (J_N^{(l)} W^* G)^{m_2}_{j_1,i_2} W^{m_2}_{i_2,j_2}\right\} \\ &+ \frac{1}{N} \mathbb{E}\{\eta_N G^{m_1 m_2}_{i_1 i_2} (J_L^{(l)} \otimes I_M)^{m_2 m_3}_{i_2 i_3} (J_N^{(l)})_{j_1 j_2}\} + \mathcal{O}\left(\frac{1}{N^k}\right)\end{aligned}$$

Summing both sides over i_2, m_2, we obtain that

$$\begin{aligned}\mathbb{E}\{\eta_N (GW)^{m_1}_{i_1 j_2} \bar{W}^{m_3}_{j_1,i_3}\} &= -\frac{1}{N} \sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{\eta_N (G(J_L^{(l)} \otimes I_M))^{m_1 m_3}_{i_1 i_3} (J_N^{(l)} \Pi)_{j_1 j_2}\right\} \\ &+ \frac{1}{N} \sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{\eta_N (G(J_L^{(l)} \otimes I_M))^{m_1 m_3}_{i_1 i_3} (J_N^{(l)})_{j_1 j_2}\right\} + \mathcal{O}\left(\frac{1}{N^k}\right).\end{aligned} \tag{A.10}$$

At this point, in order to prove (III.19), we take $j_2 = j_1$ and sum over this index. Since $GWW^* = I_{ML}$, we have

$$\mathbb{E}\{\eta_N\} I_{ML} = -\sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{\eta_N G(J_L^{(l)} \otimes I_M) \frac{1}{N} \mathrm{Tr}(J_N^{(l)} \Pi)\right\} + \sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{\eta_N G(J_L^{(l)} \otimes I_M) \frac{1}{N} \mathrm{Tr} J_N^{(l)}\right\} + \mathcal{O}\left(\frac{1}{N^k}\right)$$

Obviously $\frac{1}{N}\mathrm{Tr} J_N^{(l)}$ is equal to 0 for $l \neq 0$ and to 1 if $l = 0$, and, as was discussed above, we can replace $\mathbb{E}\{\eta_N\}$ by 1 on the l.h.s. and η_N by η_N^2 on the first term of the r.h.s. while adding a $\mathcal{O}(N^{-k})$ term. Then

$$\begin{aligned}I_{ML} = &-\sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{\eta_N G(J_L^{(l)} \otimes I_M)\right\} \mathbb{E}\left\{\frac{1}{N} \mathrm{Tr}(\eta_N J_N^{(l)} \Pi)\right\} - \sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{\eta_N G(J_L^{(l)} \otimes I_M) \frac{1}{N} \mathrm{Tr}(\eta_N J_N^{(l)} \Pi)^\circ\right\} \\ &+ \mathbb{E}\{\eta_N G\} + \mathcal{O}\left(\frac{1}{N^k}\right)\end{aligned} \tag{A.11}$$

Lemma III.4 implies that $\mathbb{E}\{\eta_N \Pi\}$ is diagonal, so $\mathbb{E}\left\{\frac{1}{N}\mathrm{Tr}(\eta_N J_N^{(l)} \Pi)\right\} = 0$ for all $l \neq 0$ and moreover since $\frac{1}{N}\mathrm{Tr}\Pi = c_N$ it is easy to see that $\mathbb{E}\left\{\frac{1}{N}\mathrm{Tr}(\eta_N \Pi)\right\} = c_N + \mathcal{O}(N^{-k})$ for each k. Thus, (A.11) leads to

$$\mathbb{E}\left\{\eta_N (WW^*)^{-1}\right\} = \frac{1}{1-c} I_{ML} + \frac{1}{1-c} \sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{\eta_N G(J_L^{(l)} \otimes I_M) \frac{1}{N} \mathrm{Tr}(\eta_N J_N^{(l)} \Pi)^\circ\right\} + \mathcal{O}\left(\frac{1}{N^k}\right) \tag{A.12}$$

Finally, we show that each element of matrix $\sum \mathbb{E}\left\{\eta_N G(J_L^{(l)} \otimes I_M) \frac{1}{N}\mathrm{Tr}(J_N^{(l)}(\eta_N \Pi)^\circ)\right\}$ is of order $\mathcal{O}(N^{-3/2})$. For this, we apply the Schwartz inequality:

$$\left|\mathbb{E}\left\{(\mathbf{f}_{i_1}^{m_1})^* \eta_N G(J_L^{(l)} \otimes I_M)\mathbf{f}_{i_2}^{m_2} \frac{1}{N}\mathrm{Tr}(\eta_N J_N^{(l)} \Pi^\circ)\right\}\right| \leq \left(\mathbf{Var}\left((\mathbf{f}_{i_1}^{m_1})^* \eta_N G(J_L^{(l)} \otimes I_M)\mathbf{f}_{i_2}^{m_2}\right) \mathbf{Var}\left(\frac{1}{N}\mathrm{Tr}(\eta_N J_N^{(l)} \Pi)\right)\right)^{1/2}$$

In order to evaluate these variances, one should follow the steps of the proof of Proposition 3.1 in [32]. In [32], matrix ηG is replaced by the resolvent of WW^* evaluated at $z \in \mathbb{C}^+$. The proof of Proposition 3.1 in [32] uses the fact that the norm of this resolvent is bounded by $\frac{1}{\mathrm{Im} z}$, a result that is of course not true in the present context. However, the above upper bound is replaced by $\eta_N G \leq \kappa I_N$ (see (III.6)). This allows to obtain the same estimations as in Proposition 3.1 in [32]:

$$\mathbf{Var}\left((\mathbf{f}_{i_1}^{m_1})^* \eta_N G(J_L^{(l)} \otimes I_M)\mathbf{f}_{i_2}^{m_2}\right) = \mathcal{O}\left(\frac{1}{N}\right)$$
$$\mathbf{Var}\left(\frac{1}{N}\mathrm{Tr}(\eta_N J_N^{(l)} \Pi)\right) = \mathcal{O}\left(\frac{1}{N^2}\right)$$
$$\mathbf{Var}\left(\frac{1}{ML}\mathrm{Tr}\eta_N G(J_L^{(l)} \otimes I_M)\right) = \mathcal{O}\left(\frac{1}{N^2}\right)$$

and to conclude that (III.19) holds.

To estimate the expectation of $(ML)^{-1}\mathrm{Tr}\eta_N(WW^*)^{-1}$ we take the normalized trace from both sides of (A.12) and use again the Schwartz inequality:

$$\left|\mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}(\eta_N G(J_L^{(l)} \otimes I_M))\frac{1}{N}\mathrm{Tr}(\eta_N J_N^{(l)} \Pi^\circ)\right\}\right| \leq \left(\mathbf{Var}\left(\frac{1}{ML}\mathrm{Tr}\eta_N G(J_L^{(l)} \otimes I_M)\right) \mathbf{Var}\left(\frac{1}{N}\mathrm{Tr}(\eta_N J_N^{(l)} \Pi)\right)\right)^{1/2} = \mathcal{O}\left(\frac{1}{N^2}\right)$$

Then we get immediately $(ML)^{-1}\mathrm{Tr}\mathbb{E}\{\eta_N(W_i W_i^*)^{-1}\} = (1-c_N)^{-1} + \mathcal{O}(\frac{1}{N^2})$.

Finally, to prove (III.20) we return to equation (A.10) but this time we take $m_1 = m_3$, $i_1 = i_3$ and sum both sides over these indexes:

$$\mathbb{E}\{\eta_N \Pi\} = -c_N \sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}(\eta_N G(J_L^{(l)} \otimes I_M))(J_N^{(l)} \Pi)\right\} + c_N \sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}(\eta_N G(J_L^{(l)} \otimes I_M))J_N^{(l)}\right\} + \mathcal{O}\left(\frac{1}{N^k}\right)$$

Analogous to what we have seen above, we replace η_N by η_N^2 in the first term of the r.h.s. and remark that $\mathbb{E}\{\mathrm{Tr}(\eta_N G(J_L^{(l)} \otimes I_M))\} = 0$ for all $l \neq 0$, since $\mathbb{E}\{\eta_N G\}$ is block diagonal. Moreover $\mathbb{E}\{(ML)^{-1}\mathrm{Tr}(\eta_N G)\} = (1-c_N)^{-1} + \mathcal{O}(\frac{1}{N^2})$, so that, after trivial algebra, we get

$$\mathbb{E}\{\eta_N \Pi\} = c_N I_N + \mathcal{O}\left(\frac{1}{N^2}\right) + \sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{\frac{1}{ML}\mathrm{Tr}(\eta_N G(J_L^{(l)} \otimes I_M))^\circ \eta_N J_N^{(l)} \Pi\right\}$$

The Schwartz inequality allows to obtain (III.20).

C. Proof of Proposition III.1

We just establish (III.21). For this, we evaluate each entry of $\mathbb{E}\left(\mathbf{Q}_{\mathbf{pp}}\eta\Pi_p\right)$ by using the integration by parts formula (I.25). In this formula, each entry $\mathbb{E}\{(\mathbf{Q}_{\mathbf{pp}}\eta\Pi_p)_{rs}\}$ of $\mathbb{E}\left(\mathbf{Q}_{\mathbf{pp}}\eta\Pi_p\right)$ is considered as a function of the entries of the $2ML \times N$ matrix $\overline{W}_N$ whose elements are the complex conjugate of those of W_N.

$$\mathbb{E}\{(\mathbf{Q_{pp}}\eta\Pi_p)_{rs}\} = \sum_{t=1}^{N}\sum_{i_1,i_2=1}^{L}\sum_{m_1,m_2=1}^{M}\mathbb{E}\Big\{\mathbf{Q}_{\mathbf{pp}}^{rt}\eta\bar{W}_{p,i_1t}^{m_1}\left((W_pW_p^*)^{-1}\right)_{i_1i_2}^{m_1m_2}W_{p,i_2s}^{m_2}\Big\} = \sum\mathbb{E}\{\bar{W}_{p,i_1t}^{m_1}W_{i_3u}^{m_3}\}$$
$$\times\mathbb{E}\left\{\frac{\partial\left(\mathbf{Q}_{\mathbf{pp}}^{rt}\eta\left((W_pW_p^*)^{-1}\right)_{i_1i_2}^{m_1m_2}W_{p,i_2s}^{m_2}\right)}{\partial W_{i_3u}^{m_3}}\right\} = \frac{1}{N}\sum\mathbb{E}\left\{\delta_{m_1,m_3}\delta_{i_1+t,i_3+u}\mathbf{Q}_{\mathbf{pp}}^{rt}\eta\left((W_pW_p^*)^{-1}\right)_{i_1i_2}^{m_1m_2}\frac{\partial W_{p,i_2s}^{m_2}}{\partial W_{i_3u}^{m_3}}\right.$$
$$\left.+\mathbf{Q}_{\mathbf{pp}}^{rt}\eta\frac{\partial\left((W_pW_p^*)^{-1}\right)_{i_1i_2}^{m_1m_2}}{\partial W_{i_3u}^{m_3}}W_{p,i_2s}^{m_2}+\frac{\partial\mathbf{Q}_{\mathbf{pp}}^{rt}}{\partial W_{i_3u}^{m_3}}\eta\left((W_pW_p^*)^{-1}\right)_{i_1i_2}^{m_1m_2}W_{p,i_2s}^{m_2}+\mathbf{Q}_{\mathbf{pp}}^{rt}\frac{\partial\eta}{\partial W_{i_3u}^{m_3}}\left((W_pW_p^*)^{-1}\right)_{i_1i_2}^{m_1m_2}W_{p,i_2s}^{m_2}\right\} \tag{A.13}$$

Here we take the derivative with respect to each element of $W = (W_p^T, W_f^T)^T$, so index i_3 takes values from 1 to $2L$. We denote each term of the r.h.s. of (A.13) without expectation by $(T_1)_{rs}$, $(T_2)_{rs}$, $(T_3)_{rs}$, $(T_4)_{rs}$ respectively and treat them separately. In order to simplify the notations, for $i = 1, 2, 3, 4$, we denote $(T_i)_{rs}$ by T_i in the following calculations.

$$T_1 = \frac{1}{N}\sum\delta_{m_1,m_3}\delta_{i_1+t,i_3+u}\mathbf{Q}_{\mathbf{pp}}^{rt}\eta\left((W_pW_p^*)^{-1}\right)_{i_1i_2}^{m_1m_2}\frac{\partial W_{p,i_2s}^{m_2}}{\partial W_{i_3u}^{m_3}} = \frac{1}{N}\sum\delta_{m_1,m_2}\delta_{i_1+t,i_2+s}\mathbf{Q}_{\mathbf{pp}}^{rt}\eta\left((W_pW_p^*)^{-1}\right)_{i_1i_2}^{m_1m_2}$$

We define $l = i_1 - i_2$ and rewrite $\delta_{i_1+t,i_2+s} = \delta_{i_1-i_2,l}\delta_{s-t,l} = (J_M^{(l)})_{i_2i_1}(J_N^{(l)})_{ts}$. Taking into account (I.12), we obtain

$$T_1 = \frac{1}{N}\sum(J_N^{(l)})_{ts}(J_L^{(l)}\otimes I_M)_{i_2i_1}^{m_2m_1}\mathbf{Q}_{\mathbf{pp}}^{rt}\eta\left((W_pW_p^*)^{-1}\right)_{i_1i_2}^{m_1m_2} = \sum_{l=-(L-1)}^{L-1}\left(\mathbf{Q_{pp}}J_N^{(l)}\right)_{rs}\frac{1}{N}\mathrm{Tr}\left((J_L^{(l)}\otimes I_M)\eta(W_pW_p^*)^{-1}\right) \tag{A.14}$$

We take the expectation and obtain

$$\mathbb{E}\{T_1\} = \sum_{l=-(L-1)}^{L-1}\mathbb{E}\Big\{\left(\mathbf{Q_{pp}}J_N^{(l)}\right)_{rs}\Big\}\frac{1}{N}\mathbb{E}\Big\{\mathrm{Tr}\left((J_L^{(l)}\otimes I_M)\eta(W_pW_p^*)^{-1}\right)\Big\}$$
$$+\sum_{l=-(L-1)}^{L-1}\mathbb{E}\Big\{\left(\mathbf{Q}_{\mathbf{pp}}^{\circ}J_N^{(l)}\right)_{rs}\frac{1}{N}\mathrm{Tr}\left((J_L^{(l)}\otimes I_M)\eta(W_pW_p^*)^{-1}\right)\Big\}$$

We denote the second term of the r.h.s by $T_1^{\mathcal{E}}$. According to (III.19), $\mathbb{E}\{(ML)^{-1}\mathrm{Tr}\eta(W_pW_p^*)^{-1}\} = \frac{1}{1-c_N}+\mathcal{O}(\frac{1}{N^2})$. Therefore, if $l = 0$ we have

$$\frac{1}{N}\mathbb{E}\{\mathrm{Tr}(\eta(W_pW_p^*)^{-1})\} = \frac{c_N}{(1-c_N)}+\mathcal{O}\left(\frac{1}{N^2}\right)$$

and if $l \neq 0$, we have $\frac{1}{N}\mathbb{E}\Big\{\mathrm{Tr}\left((J_L^{(l)}\otimes I_M)\eta(W_pW_p^*)^{-1}\right)\Big\} = 0$ by Lemma III.4. Lemma III.3 thus leads to

$$\mathbb{E}\{T_1\} = \frac{c_N}{1-c_N}\mathbb{E}\Big\{\left(\mathbf{Q_{pp}}\right)_{rs}\Big\}+\mathcal{O}_{z^2}\left(\frac{1}{N^2}\right)+T_1^{\mathcal{E}}. \tag{A.15}$$

For second term, we have

$$T_2 = -\frac{1}{N}\sum\delta_{m_1,m_3}\delta_{i_1+t,i_3+u}\mathbf{Q}_{\mathbf{pp}}^{rt}\eta\left((W_pW_p^*)^{-1}\right)_{i_1i_3}^{m_1m_3}\left(W_p^*(W_pW_p^*)^{-1}\right)_{ui_2}^{m_2}W_{p,i_2s}^{m_2}$$

We define $l = i_1 - i_3$. Then, $\delta_{i_1+t,i_3+u} = \delta_{i_1-i_3,l}\delta_{u-t,l} = (J_M^{(l)})_{i_3i_1}(J_N^{(l)})_{tu}$. This gives us

$$T_2 = -\sum_{l=-(L-1)}^{L-1}\left(\eta\mathbf{Q_{pp}}J_N^{(l)}\Pi_p\right)_{rs}\frac{1}{N}\mathrm{Tr}\left((J_L^{(l)}\otimes I_M)(W_pW_p^*)^{-1}\right) \tag{A.16}$$

Taking the expectation and replacing η by η^2, we have for each $k \geq 1$,

$$\mathbb{E}\{T_2\} = -\sum_{l=-(L-1)}^{L-1} \mathbb{E}\Big\{ \left(\eta \mathbf{Q_{pp}} J_N^{(l)} \Pi_p\right)_{rs} \Big\} \frac{1}{N} \mathbb{E}\Big\{ \mathrm{Tr}\left(\eta (J_L^{(l)} \otimes I_M)(W_p W_p^*)^{-1}\right) \Big\} - \sum_{l=-(L-1)}^{L-1} \mathbb{E}\Big\{ \left(\eta \mathbf{Q_{pp}} J_N^{(l)} \Pi_p\right)^{\circ}_{rs} \frac{1}{N} \mathrm{Tr}\left(\eta (J_L^{(l)} \otimes I_M)(W_p W_p^*)^{-1}\right) \Big\} + \mathcal{O}_{z^2}\left(\frac{1}{N^k}\right)$$

As previously, we denote the second term of the r.h.s. by $T_2^{\mathcal{E}}$ and notice that in the first term of the r.h.s., according to Lemma III.4, all the terms corresponding to $l \neq 0$ are zeros, and $\mathbb{E}\{(ML)^{-1}\mathrm{Tr}\eta(W_p W_p^*)^{-1}\} = \frac{1}{1-c_N} + \mathcal{O}(\frac{1}{N^2})$. Therefore, we obtain that

$$\mathbb{E}\{T_2\} = -\frac{c_N}{1-c_N} \mathbb{E}\Big\{ \left(\eta \mathbf{Q_{pp}} \Pi_p\right)_{rs} \Big\} + T_2^{\mathcal{E}} + \mathcal{O}_{z^2}\left(\frac{1}{N^2}\right) \tag{A.17}$$

To deal with the third term, T_3, we first should find the derivatives of the resolvent w.r.t. the entries of W. For this, we write

$$\partial \mathbf{Q} = -\mathbf{Q} \partial \begin{pmatrix} 0 & \eta \Pi_p \\ \eta \Pi_f & 0 \end{pmatrix} \mathbf{Q} = -\begin{pmatrix} \mathbf{Q_{pf}} \partial(\eta \Pi_f) \mathbf{Q_{pp}} + \mathbf{Q_{pp}} \partial(\eta \Pi_p) \mathbf{Q_{fp}} & \mathbf{Q_{pf}} \partial(\eta \Pi_f) \mathbf{Q_{pf}} + \mathbf{Q_{pp}} \partial(\eta \Pi_p) \mathbf{Q_{ff}} \\ \mathbf{Q_{ff}} \partial(\eta \Pi_f) \mathbf{Q_{pp}} + \mathbf{Q_{fp}} \partial(\eta \Pi_p) \mathbf{Q_{fp}} & \mathbf{Q_{ff}} \partial(\eta \Pi_f) \mathbf{Q_{pf}} + \mathbf{Q_{fp}} \partial(\eta \Pi_p) \mathbf{Q_{ff}} \end{pmatrix} \tag{A.18}$$

and evaluate the derivative with respect to the element $W_{i_3 u}^{m_3}$. As was discussed before, since $\|\mathbf{Q}\|$ and $\|\Pi_i\|$, $i = p, f$, are bounded (see Lemma III.3), the expectation of the entries of the terms containing $\frac{\partial \eta}{\partial W_{i_3 u}^{m_3}}$ are $\mathcal{O}_{z^2}(N^{-k})$ terms for each $k \geq 1$. This justifies that we gather all this terms together in a matrix, denoted $\mathcal{E}$, whose entries are $\mathcal{O}_{z^2}(N^{-k})$ for any k. We also need to evaluate the derivative of projectors Π_p and Π_f. For this, we use classical perturbation theory results ([28], see also Theorem 6 in [1] for the statement of the result), and obtain

$$\delta \Pi_p = \Pi_p^{\perp} \delta(W_p^* W_p)(W_p^* W_p)^{\#} + (W_p^* W_p)^{\#} \delta(W_p^* W_p) \Pi_p^{\perp} \tag{A.19}$$

where $(W_p^* W_p)^{\#}$ is the pseudoinverse of $W_p^* W_p$, which, in this case, is equal to $W_p^*(W_p W_p^*)^{-2} W_p$. The expression of $\delta \Pi_f$ is similar. The derivative with respect to $W_{i_3 u}^{m_3}$ is thus given by

$$\frac{\partial \Pi_p}{\partial W_{i_3 u}^{m_3}} = \left(\Pi_p^{\perp} W_p^* \mathbf{f}_{i_3}^{m_3} \mathbf{e}_u W_p^* (W_p W_p^*)^{-2} W_p + W_p^* (W_p W_p^*)^{-2} W_p W_p^* \mathbf{f}_{i_3}^{m_3} \mathbf{e}_u \Pi_p^{\perp}\right) \mathbf{1}_{i_3 \leq L}.$$

In this context, $\mathbf{f}_{i_3}^{m_3}$ is a vector of the canonical basis of $\mathbb{C}^{ML}$ rather than of $\mathbb{C}^{2ML}$. Since $\Pi_p^{\perp} W_p^* = 0$ the first term disappears and we obtain

$$\frac{\partial \Pi_p}{\partial W_{i_3 u}^{m_3}} = W_p^* (W_p W_p^*)^{-1} \mathbf{f}_{i_3}^{m_3} \mathbf{e}_u^* \Pi_p^{\perp} \mathbf{1}_{i_3 \leq L}$$

For Π_f the formula is analogous, but $\mathbf{f}_{i_3}^{m_3}$ is replaced by $\mathbf{f}_{i_3 - L}^{m_3}$:

$$\frac{\partial \Pi_f}{\partial W_{i_3 u}^{m_3}} = W_f^* (W_f W_f^*)^{-1} \mathbf{f}_{i_3 - L}^{m_3} \mathbf{e}_u^* \Pi_f^{\perp} \mathbf{1}_{i_3 > L}$$

Putting these expressions in (A.18), we get that

$$\frac{\partial \mathbf{Q}}{\partial W_{i_3 u}^{m_3}} = -\eta \mathbf{1}_{i_3 \leq L} \begin{pmatrix} \mathbf{Q_{pp}} (W_p^* (W_p W_p^*)^{-1} \mathbf{f}_{i_3}^{m_3} \mathbf{e}_u^* \Pi_p^{\perp} \mathbf{Q_{fp}} & \mathbf{Q_{pp}} (W_p^* (W_p W_p^*)^{-1} \mathbf{f}_{i_3}^{m_3} \mathbf{e}_u^* \Pi_p^{\perp} \mathbf{Q_{ff}} \\ \mathbf{Q_{fp}} (W_p^* (W_p W_p^*)^{-1} \mathbf{f}_{i_3}^{m_3} \mathbf{e}_u^* \Pi_p^{\perp} \mathbf{Q_{fp}} & \mathbf{Q_{fp}} (W_p^* (W_p W_p^*)^{-1} \mathbf{f}_{i_3}^{m_3} \mathbf{e}_u^* \Pi_p^{\perp} \mathbf{Q_{ff}} \end{pmatrix} - \eta \mathbf{1}_{i_3 > L} \begin{pmatrix} \mathbf{Q_{pf}} (W_f^* (W_f W_f^*)^{-1} \mathbf{f}_{i_3 - L}^{m_3} \mathbf{e}_u^* \Pi_f^{\perp} \mathbf{Q_{pp}} & \mathbf{Q_{pf}} (W_f^* (W_f W_f^*)^{-1} \mathbf{f}_{i_3 - L}^{m_3} \mathbf{e}_u^* \Pi_f^{\perp} \mathbf{Q_{pf}} \\ \mathbf{Q_{ff}} (W_f^* (W_f W_f^*)^{-1} \mathbf{f}_{i_3 - L}^{m_3} \mathbf{e}_u^* \Pi_f^{\perp} \mathbf{Q_{pp}} & \mathbf{Q_{ff}} (W_f^* (W_f W_f^*)^{-1} \mathbf{f}_{i_3 - L}^{m_3} \mathbf{e}_u^* \Pi_f^{\perp} \mathbf{Q_{pf}} \end{pmatrix} + \mathcal{E} \tag{A.20}$$

We are now ready to address the term T_3. We first we sum over i_2, m_2, and obtain that T_3 can be written as

$$T_3 = -\frac{1}{N} \sum \delta_{m_1, m_3} \delta_{i_1 + t, i_3 + u} \eta^2 \left(\mathbf{Q_{pp}} W_p^* (W_p W_p^*)^{-1}\right)_{r i_3}^{m_3} \left(\Pi_p^{\perp} \mathbf{Q_{fp}}\right)_{ut} \left((W_p W_p^*)^{-1} W_p\right)_{i_1 s}^{m_1} \mathbf{1}_{i_3 \leq L} - \frac{1}{N} \sum \delta_{m_1, m_3} \delta_{i_1 + t, i_3 + u} \eta^2 \left(\mathbf{Q_{pf}} W_f^* (W_f W_f^*)^{-1}\right)_{r i_3 - L}^{m_3} \left(\Pi_f^{\perp} \mathbf{Q_{pp}}\right)_{ut} \left((W_p W_p^*)^{-1} W_p\right)_{i_1 s}^{m_1} \mathbf{1}_{i_3 > L} + \mathcal{E}_3 \tag{A.21}$$

where $\mathcal{E}_3$ represents the contribution of matrix $\mathcal{E}$ to T_3. In order to express the first term of the r.h.s. of (A.21), we again define $l = i_1 - i_3$ which belongs to $\{-(L-1), \dots, L-1\}$ and notice that $\delta_{i_1+t,i_3+u} = \delta_{i_1-i_3,l}\delta_{u-t,l} = (J_L^{(l)})_{i_3 i_1}(J_N^{(l)})_{tu}$. As for second term of the r.h.s. of (A.21), we first exchange $i_3 > L$ by $i_3 - L$ which runs from 1 to L. The second term becomes

$$\frac{1}{N}\sum \delta_{m_1,m_3}\delta_{i_1+t,i_3+L+u}\eta^2(\mathbf{Q_{pf}}W_f^*(W_fW_f^*)^{-1})_{ri_3}^{m_3}(\Pi_f^\perp \mathbf{Q_{pp}})_{ut}((W_pW_p^*)^{-1}W_p)_{i_1 s}^{m_1}\mathbf{1}_{i_3<L}$$

We again put $l = i_1 - i_3$ and remark that $\delta_{i_1+t,i_3+L+u} = \delta_{i_1-i_3,l}\delta_{u-t,l-L} = (J_L^{(l)})_{i_3,i_1}(J_N^{(l-L)})_{tu}$. Therefore, we obtain that T_3 is equal to

$$T_3 = -\sum_{l=-(L-1)}^{L-1}\eta^2\left(\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-1}(J_L^{(l)}\otimes I_M)(W_pW_p^*)^{-1}W_p\right)_{rs}\frac{1}{N}\mathrm{Tr}\left(\Pi_p^\perp\mathbf{Q_{fp}}J_N^{(l)}\right)$$
$$-\sum_{l=-(L-1)}^{L-1}\eta^2\left(\mathbf{Q_{pf}}W_f^*(W_fW_f^*)^{-1}(J_L^{(l)}\otimes I_M)(W_pW_p^*)^{-1}W_p\right)_{rs}\frac{1}{N}\mathrm{Tr}\left(\Pi_f^\perp\mathbf{Q_{pp}}J_N^{(l-L)}\right)+\mathcal{E}_3$$

Taking the expectation, we obtain that

$$\mathbb{E}\{T_3\} = -\sum_{l=-(L-1)}^{L-1}\mathbb{E}\Big\{\eta\left(\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-1}(J_L^{(l)}\otimes I_M)(W_pW_p^*)^{-1}W_p\right)_{rs}\Big\}\frac{1}{N}\mathbb{E}\Big\{\mathrm{Tr}\left(\eta\Pi_p^\perp\mathbf{Q_{fp}}J_N^{(l)}\right)\Big\}$$
$$-\sum_{l=-(L-1)}^{L-1}\mathbb{E}\Big\{\eta\left(\mathbf{Q_{pf}}W_f^*(W_fW_f^*)^{-1}(J_L^{(l)}\otimes I_M)(W_pW_p^*)^{-1}W_p\right)_{rs}\Big\}\frac{1}{N}\mathbb{E}\Big\{\mathrm{Tr}\left(\eta\Pi_f^\perp\mathbf{Q_{pp}}J_N^{(l-L)}\right)\Big\}$$
$$+\mathbb{E}\{\mathcal{E}_3\}+T_3^{\mathcal{E}},$$

where, as above, $T_3^{\mathcal{E}}$ is defined by

$$T_3^{\mathcal{E}} = \sum_l\mathbb{E}\left[\left(\eta\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-1}(J_L^{(l)}\otimes I_M)(W_pW_p^*)^{-1}W_p\right)_{rs}^\circ\frac{1}{N}\mathrm{Tr}\left(\Pi_p^\perp\mathbf{Q_{fp}}J_N^{(l)}\right)^\circ\right]+$$
$$\sum_l\mathbb{E}\left[\left(\eta\mathbf{Q_{pf}}W_f^*(W_fW_f^*)^{-1}(J_L^{(l)}\otimes I_M)(W_pW_p^*)^{-1}W_p\right)_{rs}^\circ\frac{1}{N}\mathrm{Tr}\left(\Pi_f^\perp\mathbf{Q_{pp}}J_N^{(l-L)}\right)^\circ\right]$$

According to Lemma III.4, $\mathbb{E}\{\eta\Pi_p^\perp\mathbf{Q_{fp}}\}$ and $\mathbb{E}\{\eta\Pi_f^\perp\mathbf{Q_{pp}}\}$ are diagonal. Therefore, the traces of these matrices multiplied by $J_N^{(k)}$ for $k \neq 0$ are zeros. This leads to

$$\mathbb{E}\{T_3\} = -\mathbb{E}\Big\{\eta\left(\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-2}W_p\right)_{rs}\Big\}\frac{1}{N}\mathbb{E}\Big\{\mathrm{Tr}\left(\eta\Pi_p^\perp\mathbf{Q_{fp}}\right)\Big\}+\mathbb{E}\{\mathcal{E}_3\}+T_3^{\mathcal{E}}. \tag{A.22}$$

Finally, the various terms of T_4 contain the terms $(\frac{\partial\eta}{\partial W_{i_3u}^{m_3}})_{i_3=1,\dots,2L,m_3=1,\dots,M}$. Therefore, T_4 is denoted $\mathcal{E}_4$, and $\mathbb{E}(T_4) = \mathbb{E}(\mathcal{E}_4) = \mathcal{O}_{z^2}(N^{-k})$ for each k.

Combining (A.15), (A.17), (A.22), we have thus obtained that

$$\mathbb{E}\Big\{(\mathbf{Q_{pp}}\eta\Pi_p)_{rs}\Big\} = \frac{c_N}{1-c_N}\mathbb{E}\Big\{(\mathbf{Q_{pp}})_{rs}\Big\}-\frac{c_N}{1-c_N}\mathbb{E}\Big\{\eta(\mathbf{Q_{pp}}\Pi_p)_{rs}\Big\}$$
$$-\mathbb{E}\Big\{\eta(\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-2}W_p)_{rs}\Big\}\frac{1}{N}\mathbb{E}\Big\{\mathrm{Tr}\left(\eta\Pi_p^\perp\mathbf{Q_{fp}}\right)\Big\}+\frac{1}{1-c_N}(\mathbf{\Delta_{pp}})_{rs}$$

where $(\mathbf{\Delta_{pp}})_{rs}$ represents the term

$$(\mathbf{\Delta_{pp}})_{rs} = (1-c_N)\left[\mathbb{E}\{T_1^{\mathcal{E}}+T_2^{\mathcal{E}}+T_3^{\mathcal{E}}\}+\mathbb{E}\{\mathcal{E}_3\}+\mathbb{E}\{\mathcal{E}_4\}+\mathcal{O}_{z^2}(\frac{1}{N^2})\right]$$

obtained by gathering the various error terms defined in the evaluation of $(T_i)_{i=1,2,3,4}$. Therefore, we eventually get the following expression of matrix $\mathbb{E}\Big\{\mathbf{Q_{pp}}\eta\Pi_p\Big\}$:

$$\mathbb{E}\Big\{\mathbf{Q_{pp}}\eta\Pi_p\Big\} = \frac{c_N}{1-c_N}\mathbb{E}\Big\{\mathbf{Q_{pp}}\Big\} - \frac{c_N}{1-c_N}\mathbb{E}\Big\{\eta\mathbf{Q_{pp}}\Pi_p\Big\} - \mathbb{E}\Big\{\eta\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-2}W_p\Big\}\frac{1}{N}\mathbb{E}\Big\{\mathrm{Tr}\left(\eta\Pi_p^{\perp}\mathbf{Q_{fp}}\right)\Big\} + \frac{1}{1-c_N}\boldsymbol{\Delta}_{\mathbf{pp}}$$

which leads immediately to (III.21).

It remains to establish the properties of matrix $\boldsymbol{\Delta}_{\mathbf{pp}}$. According to Lemma III.4, $\mathbb{E}\{\mathbf{Q_{pp}}\eta\Pi_p\}, \mathbb{E}\{\mathbf{Q_{pp}}\}$ and $\mathbb{E}\{\eta\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-2}W_p\}$ are diagonal. Therefore, (III.21) implies that $\boldsymbol{\Delta}_{\mathbf{pp}}$ is also diagonal. In order to evaluate the order of magnitude of the entries of $\boldsymbol{\Delta}_{\mathbf{pp}}$ and of $\frac{1}{N}\mathrm{Tr}(\boldsymbol{\Delta}_{\mathbf{pp}})$, we first prove the next lemma which is based on the Poincaré-Nash inequality.

Lemma A.2. *Let $(F_N)_{N\geq 1}$ and $(G_N)_{N\geq 1}$ be sequences of deterministic $N\times N$ matrices such that $\sup_N\|F_N\|$, $\sup_N\|G_N\|\leq\kappa$, and consider sequences of deterministic N–dimensional vectors $(a_{1,N})_{N\geq1}$, $(a_{2,N})_{N\geq 1}$ such that $sup_N\|a_{i,N}\|\leq\kappa$ for $i=1,2$. Then, for each $z\in\mathbb{C}^+$ and $i,j,h=\{p,f\}$, it holds that*

$$\mathbf{Var}\left\{\frac{1}{N}\mathrm{Tr}F\mathbf{Q_{ij}}\right\} = \mathcal{O}_{z^2}\left(\frac{1}{N^2}\right), \tag{A.23}$$

$$\mathbf{Var}\left\{\frac{1}{N}\mathrm{Tr}\mathbf{Q_{ij}}F\eta_N\Pi_hG\right\} = \mathcal{O}_{z^2}\left(\frac{1}{N^2}\right), \tag{A.24}$$

$$\mathbf{Var}\left\{\frac{1}{N}\mathrm{Tr}\mathbf{Q_{ij}}F\eta_N\Pi_h^{\perp}G\right\} = \mathcal{O}_{z^2}\left(\frac{1}{N^2}\right), \tag{A.25}$$

$$\mathbf{Var}\left\{\frac{1}{N}\mathrm{Tr}\eta\mathbf{Q_{ij}}W_h^*(W_hW_h^*)^{-1}F(W_kW_k^*)^{-1}W_k\right\} = \mathcal{O}_{z^2}\left(\frac{1}{N}\right), \tag{A.26}$$

$$\mathbf{Var}\left(a_1^*\eta\mathbf{Q_{ij}}W_h^*(W_hW_h^*)^{-1}F(W_kW_k^*)^{-1}W_ka_2\right) = \mathcal{O}_{z^2}\left(\frac{1}{N}\right), \tag{A.27}$$

$$\mathbf{Var}\left\{a_1^*\mathbf{Q_{ij}}a_2\right\} = \mathcal{O}_{z^2}\left(\frac{1}{N}\right), \tag{A.28}$$

$$\mathbf{Var}\left[\left(a_1^*\mathbf{Q_{ij}}a_2 - \mathbb{E}(a_1^*\mathbf{Q_{ij}}a_2)\right)^2\right] = \mathcal{O}_{z^2}\left(\frac{1}{N^2}\right), \tag{A.29}$$

$$\mathbf{Var}\left\{a_1^*\mathbf{Q_{ij}}F\eta_N\Pi_ha_2\right\} = \mathcal{O}_{z^2}\left(\frac{1}{N}\right). \tag{A.30}$$

Proof. We just prove (A.23) for $\mathbf{Q_{pp}}$. The proofs of the other items are omitted. We denote by ξ the term $\xi = \frac{1}{N}\mathrm{Tr}F\mathbf{Q_{pp}}$. The Poincaré-Nash inequality (I.2) leads to

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

We just evaluate the second term of the r.h.s., denoted by ϕ. The derivative of ξ can be found with the help of (A.20):

$$\frac{\partial\xi}{\partial W_{i_1j_1}^{m_1}} = -\frac{\eta}{N}\mathrm{Tr}F\mathbf{Q_{pp}}W_p^*(W_pW_p^*)^{-1}\mathbf{f}_{i_1}^{m_1}\mathbf{e}_{j_1}^*\Pi_p^{\perp}\mathbf{Q_{fp}}\mathbf{1}_{i_1\leq L} - \frac{\eta}{N}\mathrm{Tr}F\mathbf{Q_{pf}}W_f^*(W_fW_f^*)^{-1}\mathbf{f}_{i_1-L}^{m_1}\mathbf{e}_{j_1}^*\Pi_f^{\perp}\mathbf{Q_{pp}}\mathbf{1}_{i_1>L} + \mathcal{O}\left(\frac{1}{N^k}\right)$$

ϕ is clearly the sum of four similar terms. We just evaluate

$$\frac{1}{N^3}\sum_{\substack{i_1,j_1,m_1\\ i_2,j_2,m_2}} \delta_{m_1,m_2}\delta_{i_1+j_1,i_2+j_2}\mathbb{E}\left\{\eta^2 \mathbf{e}_{j_1}^* \Pi_p^\perp \mathbf{Q_{fp}} F \mathbf{Q_{pp}} W_p^*(W_pW_p^*)^{-1}\mathbf{f}_{i_1}^{m_1}\mathbf{f}_{i_2}^{m_2 *}(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{pp}}^* F\mathbf{Q}_{\mathbf{fp}}^*\Pi_p^\perp \mathbf{e}_{j_2}\right\} \tag{A.31}$$

where $1 \leq i_1, i_2 \leq L$. Now we again denote $l = i_1 - i_2 = j_2 - j_1$ which lies in $(-L+1, L-1)$ and remark that $\sum_{m_1,m_2,i_1,i_2} \delta_{m_1,m_2}\delta_{i_1-i_2,l}\mathbf{f}_{i_1}^{m_1}\mathbf{f}_{i_2}^{m_2*} = (J_L^{(l)} \otimes I_M)$ as well as $\sum_{j_1,j_2}\delta_{j_2-j_1,l}\mathbf{e}_{j_2}\mathbf{e}_{j_1}^* = J_N^{(l)}$. This allows to rewrite (A.31) as

$$\frac{1}{N^3}\sum_{l=-(L-1)}^{L-1}\mathbb{E}\left\{\eta^2 \mathrm{Tr}\Pi_p^\perp \mathbf{Q_{fp}} F\mathbf{Q_{pp}} W_p^*(W_pW_p^*)^{-1}(J_L^{(l)}\otimes I_M)(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{pp}}^* F^*\mathbf{Q}_{\mathbf{fp}}^*\Pi_p^\perp J_N^{(l)}\right\} \tag{A.32}$$

For each $N \times ML$ matrices A and B, the Schwartz inequality and the inequality between arithmetic and geometric means lead to

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

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

$$\left|\frac{1}{N}\mathrm{Tr}A(I_M\otimes J_L^{*(l)})B^* J_N^{*(l)}\right| \leq \frac{C}{2N}(\mathrm{Tr}A^*A + \mathrm{Tr}B^*B). \tag{A.33}$$

holds. We take $A = B = \Pi_p^\perp \mathbf{Q_{fp}} F\mathbf{Q_{pp}} W_p^*\eta(W_pW_p^*)^{-1}$, and have to check that $N^{-1}\mathbb{E}\{\mathrm{Tr}AA^*\} = \mathcal{O}_z^2(1)$. For this, we remark that $\eta W_p^* W_p \leq ((1+\sqrt{c_*})^2 + 2\epsilon)I_N$ and $\eta^2(W_pW_p^*)^{-2} \leq ((1-\sqrt{c_*})^2 - 2\epsilon)^{-2}I_{ML}$ (see (III.6)). Therefore, $W_p^*\eta^2(W_pW_p^*)^{-2}W_p \leq \kappa I_N$, and

$$\frac{1}{N}\mathbb{E}\left\{\mathrm{Tr}\Pi_p^\perp \mathbf{Q_{fp}} F\mathbf{Q_{pp}} W_p^*\eta^2(W_pW_p^*)^{-2}W_p\mathbf{Q}_{\mathbf{pp}}^* F^*\mathbf{Q}_{\mathbf{fp}}^*\Pi_p^\perp\right\} = \mathcal{O}_z^2(1) \tag{A.34}$$

as expected. This completes the proof of (A.23). ■

We return to the evaluation of the (diagonal) entries of $\mathbf{\Delta_{pp}}$. As the terms $\mathbb{E}(\mathcal{E}_3)$ and $\mathbb{E}(\mathcal{E}_4)$ are $\mathcal{O}_z^2(\frac{1}{N^k})$ for each k, it remains to evaluate the order of magnitude of the terms $(T_i^{\mathcal{E}})_{rr}$ for $i = 1, 2, 3$ defined by (A.15,A.17,A.22) respectively when $r = s$. We start with $(T_1^{\mathcal{E}})_{rr}$ and use Schwartz inequality:

$$\begin{aligned}|(T_1^{\mathcal{E}})_{rr}| &= \left|\sum_{l=-(L-1)}^{L-1}\mathbb{E}\left\{\left(\mathbf{Q}_{\mathbf{pp}}^\circ J_N^{(l)}\right)_{rr}\frac{1}{N}\mathrm{Tr}\left((I_M\otimes J_M^{(l)})\eta(W_pW_p^*)^{-1}\right)\right\}\right| \\ &\leq \sum_{l=-(L-1)}^{L-1}\left(\mathbf{Var}\left(\left(\mathbf{Q_{pp}}J_N^{(l)}\right)_{rr}\right)\mathbf{Var}\left(\frac{1}{N}\mathrm{Tr}\left((I_M\otimes J_M^{(l)})\eta(W_pW_p^*)^{-1}\right)\right)\right)^{1/2}\end{aligned}$$

We apply (A.28) for $a_1 = \mathbf{e}_r$ and $a_2 = J_N^{(l)}\mathbf{e}_r$ and take into account that $\mathbf{Var}(\frac{1}{N}\mathrm{Tr}((I_M\otimes J_M^{(l)})\eta(W_pW_p^*)^{-1})) = \mathcal{O}(N^{-2})$. Then

$$|(T_1^{\mathcal{E}})_{rr}| \leq \mathcal{O}_z^2\left(\frac{1}{N^{3/2}}\right) \tag{A.35}$$

As for $(T_2^{\mathcal{E}})_{rr}$, we have

$$\begin{aligned}|(T_2^{\mathcal{E}})_{rr}| &= \left|\sum_{l=-(L-1)}^{L-1}\mathbb{E}\left\{\left(\eta\mathbf{Q_{pp}}J_N^{(l)}\Pi_p\right)_{rr}^\circ\frac{1}{N}\mathrm{Tr}\left((I_M\otimes J_M^{(l)})\eta(W_pW_p^*)^{-1}\right)\right\}\right| \\ &\leq \sum_{l=-(L-1)}^{L-1}\left(\mathbf{Var}\left(\left(\eta\mathbf{Q_{pp}}J_N^{(l)}\Pi_p\right)_{rr}\right)\mathbf{Var}\left(\frac{1}{N}\mathrm{Tr}\left((I_M\otimes J_M^{(l)})\eta(W_pW_p^*)^{-1}\right)\right)\right)^{1/2}\end{aligned}$$

From (A.30) we get immediately

$$|(T_2^{\mathcal{E}})_{rr}| = \mathcal{O}_z^2\left(\frac{1}{N^{3/2}}\right) \tag{A.36}$$

For $T_3^{\mathcal{E}}$ we obtain

$$\begin{aligned}
|(T_3^{\mathcal{E}})_{rr}| = & \Bigg| \sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{ \left(\eta \mathbf{Q_{pp}} W_p^*(W_pW_p^*)^{-1}(I_M \otimes J_M^{(l)})(W_pW_p^*)^{-1}W_p\right)^{\circ}_{rr} \frac{1}{N}\mathrm{Tr}\left(\eta J_N^{(l)} \Pi_p^{\perp} \mathbf{Q_{fp}}\right)\right\} \\
& + \sum_{l=-(L-1)}^{L-1} \mathbb{E}\left\{ \left(\eta \mathbf{Q_{pf}} W_f^*(W_fW_f^*)^{-1}(I_M \otimes J_M^{(l)})(W_pW_p^*)^{-1}W_p\right)^{\circ}_{rr} \frac{1}{N}\mathrm{Tr}\left(\eta J_N^{(l-L)} \Pi_f^{\perp} \mathbf{Q_{pp}}\right)\right\} \Bigg| \\
\leq & \sum_{l=-(L-1)}^{L-1} \left(\mathbf{Var}\left(\left(\eta \mathbf{Q_{pp}} W_p^*(W_pW_p^*)^{-1}(I_M \otimes J_M^{(l)})(W_pW_p^*)^{-1}W_p\right)_{rr}\right)\mathbf{Var}\left(\frac{1}{N}\mathrm{Tr}\left(\eta J_N^{(l)} \Pi_p^{\perp} \mathbf{Q_{fp}}\right)\right)\right)^{1/2} \\
+ & \sum_{l=-(L-1)}^{L-1} \left(\mathbf{Var}\left(\left(\eta \mathbf{Q_{pf}} W_f^*(W_fW_f^*)^{-1}(I_M \otimes J_M^{(l)})(W_pW_p^*)^{-1}W_p\right)_{rr}\right)\mathbf{Var}\left(\frac{1}{N}\mathrm{Tr}\left(\eta J_N^{(l)} \Pi_f^{\perp} \mathbf{Q_{pp}}\right)\right)\right)^{1/2}
\end{aligned}$$

from what, using again (A.27) and (A.25), we immediately get

$$|(T_3^{\mathcal{E}})_{rr}| = \mathcal{O}_z^2\left(\frac{1}{N^{3/2}}\right) \tag{A.37}$$

To evaluate the normalized trace of $\mathbf{\Delta_{pp}}$, we still use the Schwartz inequality, and take benefit of the estimates (A.23)-(A.26) to improve the rate of convergence of $\frac{1}{N}\mathrm{Tr}\mathbf{\Delta_{pp}}$.

D. Proof of Proposition III.3

In order to evaluate $\tilde{\alpha}_N - \tilde{t}_N$, it is natural to take the difference between equations (III.44) and (III.52):

$$(\tilde{\alpha}_N - \tilde{t}_N)\left((1-z)z(\tilde{\alpha}_N + \tilde{t}_N) + 2(1-c_N) - z\right) = \mathcal{O}_z\left(\frac{1}{N^2}\right)$$

We remind that $\tilde{\alpha}_N = \alpha_N - \frac{1-c_N}{z} + \mathcal{O}_z(N^{-k})$ (see (III.45)) and rewrite the last equation as

$$(\tilde{\alpha}_N - \tilde{t}_N)\left((1-z)z\alpha_N - (1-z)(1-c_N) + (1-z)z\tilde{t}_N + 2(1-c_N) - z + \mathcal{O}_z(N^{-k})\right) = \mathcal{O}_z\left(\frac{1}{N^2}\right)$$

or equivalently as

$$(\tilde{\alpha}_N - \tilde{t}_N)\left((1-z)z\alpha_N - (1-z)(1-c_N) + (1-z)z\tilde{t}_N + 2(1-c_N) - z\right) + (\tilde{\alpha}_N - \tilde{t}_N)\mathcal{O}_z\left(\frac{1}{N^k}\right) = \mathcal{O}_z\left(\frac{1}{N^2}\right)$$

Since $\tilde{\alpha}_N$, $\tilde{t}_N$ and α_N (see Remark III.3) are the Stieltjes transforms of a positive measures carried by $\mathbb{R}^+$, we obtain that $\alpha_N = \mathcal{O}_z(1)$, $\tilde{\alpha}_N = \mathcal{O}_z(1)$, $\tilde{t}_N = \mathcal{O}_z(1)$, and that $(\tilde{\alpha}_N - \tilde{t}_N)\mathcal{O}_z(N^{-k}) = \mathcal{O}_z(N^{-k})$. $(\tilde{\alpha}_N - \tilde{t}_N)$ can thus be written as

$$\tilde{\alpha}_N - \tilde{t}_N = \frac{\mathcal{O}_z\left(N^{-2}\right)}{(1-z)z\alpha_N - (1-z)(1-c_N) + (1-z)z\tilde{t}_N + 2(1-c_N) - z}$$

We now evaluate the denominator for $z \in \mathbb{C}^+$. For this, we return to (III.52) and write:

$$(1-z)z\tilde{t}_N + 2(1-c_N) - z = -\frac{(1-c_N)^2}{z\tilde{t}_N}$$

Moreover, since $\tilde{t}_N$ is the Stieltjes transform of a positive measure carried by $\mathbb{R}^+$, $\mathrm{Im} z\tilde{t}_N > 0$ for $z \in \mathbb{C}^+$ (see (I.17)) and $\mathrm{Im}((1-z)z\tilde{t}_N) = \mathrm{Im}z - \mathrm{Im}\dfrac{(1-c_N)^2}{z\tilde{t}_N} > \mathrm{Im}z$. We rewrite the denominator as

$$(1-z)z\alpha_N - (1-z)(1-c_N) + (1-z)z\tilde{t}_N + 2(1-c_N) - z = (1-z)\left(z\alpha_N - (1-c_N) - \frac{(1-c_N)^2}{(1-z)z\tilde{t}_N}\right)$$

$\mathrm{Im} z\alpha_N > 0$ for $z \in \mathbb{C}^+$ because α_N is the Stieltjes transform of a positive measure curried by $\mathbb{R}^+$. Thus

$$|(1-z)z\alpha_N - (1-z)(1-c_N) + (1-z)z\tilde{t}_N + 2(1-c_N) - z| \geq |1-z|\mathrm{Im}\frac{-(1-c_N)^2}{(1-z)z\tilde{t}_N} = \frac{(1-c_N)^2\mathrm{Im}((1-z)z\tilde{t}_N)}{|1-z||z|^2|\tilde{t}_N|^2}$$

As $\mathrm{Im}((1-z)z\tilde{t}_N) > \mathrm{Im} z$ and $|\tilde{t}_N(z)| \leq (\mathrm{Im} z)^{-1}$ on $\mathbb{C}^+$, and that

$$\left|\mathcal{O}_z\left(\frac{1}{N^2}\right)\right| \leq \frac{1}{N^2} P_1(|z|) P_2\left(\frac{1}{\mathrm{Im} z}\right)$$

on $\mathbb{C}^+$ (because $\frac{1}{\rho(z)} \leq \frac{1}{\mathrm{Im} z}$ on $\mathbb{C}^+$), we obtain that

$$|\tilde{\alpha}_N(z) - \tilde{t}_N(z)| \leq \frac{1}{N^2} P_1(|z|) P_2\left(\frac{1}{\mathrm{Im} z}\right) \tag{A.38}$$

for each $z \in \mathbb{C}^+$. This completes the proof of Proposition III.3.

E. Proof of Lemma III.6

We first justify (III.68) and (III.69) when $z \in \mathbb{C}^+$. For this, we mention that (A.28) and (A.29) imply that the fourth order moments of $a_N^* \, (\mathbf{Q_{ii}})_N(z) \, b_N - \mathbb{E}(a_N^* \, (\mathbf{Q_{ii}})_N(z) \, b_N)$ and $a_N^* \, (\mathbf{Q_{ij}})_N(z) \, b_N - \mathbb{E}(a_N^* \, (\mathbf{Q_{ij}})_N(z) \, b_N)$ are $\mathcal{O}_{z^2}\left(\frac{1}{N^2}\right)$ terms. Borel-Cantelli's Lemma thus implies that $a_N^* \, (\mathbf{Q_{ii}})_N(z) \, b_N - \mathbb{E}(a_N^* \, (\mathbf{Q_{ii}})_N(z) \, b_N)$ and $a_N^* \, (\mathbf{Q_{ij}})_N(z) \, b_N - \mathbb{E}(a_N^* \, (\mathbf{Q_{ij}})_N(z) \, b_N)$ converge almost surely towards 0 . (III.35) and (III.36) as well as Corollary III.3 complete the proof of (III.68) and (III.69) on $\mathbb{C}^+$. In order to extend the convergence to $\mathbb{C} \setminus \boldsymbol{\mathcal{S}}_*$, we remark that Theorem III.1 implies that almost surely, for N large enough, $a_N^* \, (\mathbf{Q_{ii}})_N(z) \, b_N - \tilde{\mathbf{t}}_N(z) a_N^* b_N$ and $a_N^* \, (\mathbf{Q_{ij}})_N(z) \, b_N - c_N \mathbf{t}_N(z) a_N^* b_N$ are holomorphic on $\mathbb{C} \setminus \boldsymbol{\mathcal{S}}_*$, and bounded on each compact subset of $\mathbb{C} \setminus \boldsymbol{\mathcal{S}}_*$. Therefore, Montel's theorem implies that (III.35) and (III.36) holds for each $z \in \mathbb{C} \setminus \boldsymbol{\mathcal{S}}_*$ and uniformly on the compact subsets of $\mathbb{C} \setminus \boldsymbol{\mathcal{S}}_*$. The extension to the context of random vectors (a_N, b_N) is justified using the arguments used in the course of the proof of Lemma II.1.

F. Proof of the properties of function f_* defined by (III.72)

(III.61) implies that $c_* x t_*(x) = x\tilde{t}_*(x) + 1 - c_*$ for each $x \in (4c_*(1-c_*), 1)$. As $\tilde{t}_*$ is the Stieltjes transform of a positive measure supported by $\mathcal{S}_*$, function $x \to x\tilde{t}_*(x)$ is increasing on $(4c_*(1-c_*), 1)$. As we also have

$$\frac{\tilde{t}_*(x)}{t_*(x)} = c_* - \frac{1-c_*}{x t_*(x)}. \tag{A.39}$$

we obtain that $x \to \frac{\tilde{t}_*(x)}{t_*(x)}$ is increasing on $(4c_*(1-c_*), 1)$. Using (III.62), we check that $x t_*(x)$ and $\dfrac{\tilde{t}_*(x)}{t_*(x)}$ are well defined at $4c_*(1-c_*)$ and that

$$(x t_*(x))\Big|_{x=4c_*(1-c_*)} = \frac{4c_*(1-c_*)(2c_*-1)}{2c_*(2c_*-1)^2} = \frac{2(1-c_*)}{2c_*-1} \Rightarrow \frac{\tilde{t}_*(x)}{t_*(x)}\Big|_{x=4c_*(1-c_*)} = c_* - \frac{2c_*-1}{2} = \frac{1}{2}$$

This shows that $\frac{\tilde{t}_*(x)}{t_*(x)}$ is positive on $[4c_*(1-c_*), 1)$ and that $x \to \left(\frac{\tilde{t}_*(x)}{t_*(x)}\right)^2$, and thus $x \to f_*(x)$, are increasing on $[4c_*(1-c_*), 1)$. Moreover, it is easily checked that $f_*(4c_*(1-c_*)) = \frac{c_*}{1-c_*}$. It remains to show that f_* is well defined at 1, and to evaluate $f_*(1)$. For this, we remark that if $c_* < 1/2$, then due to (III.49) and (III.62), we have

$$\frac{\tilde{t}_*(x)}{t_*(x)}\Big|_{x=1} = \lim_{y \to 1} \frac{c_* 4(1-c_*)^2(1-y)\left(x(2c_*-1) - \sqrt{x(x-4c_*(1-c_*))}\right)}{x(1-x)4c_*(1-c_*)\left(x - 2(1-c_*) - \sqrt{x(x-4c_*(1-c_*))}\right)} = 1 - c_*$$

and for $c_* \geq 1/2$

$$\frac{\tilde{t}_*(x)}{t_*(x)}\Big|_{x=1} = \lim_{x \to 1} \frac{c_*\left(x - 2(1-c_*) + \sqrt{x(x-4c_*(1-c_*))}\right)}{x(2c_*-1) + \sqrt{x(x-4c_*(1-c_*))}} = c_*$$

This leads to $f_*(1) = 1$ if $c_* < \frac{1}{2}$ and $f_*(1) = \left(\frac{c_*}{1-c_*}\right)^2$ if $c_* \geq 1/2$.

G. Proof of Lemma III.7

We express E_i as

$$E_i = \begin{pmatrix} -\Delta_i^2 & I_r \\ I_r & 0 \end{pmatrix} + \begin{pmatrix} \Delta_i \tilde{\Theta}_i^* \Pi_i^W \tilde{\Theta}_i \Delta_i & \Delta_i \tilde{\Theta}_i^* W_i^* (W_i W_i^*)^{-1} \Theta_i \\ \Theta_i^* (W_i W_i^*)^{-1} W_i \tilde{\Theta}_i \Delta_i & \Theta_i^* (W_i W_i^*)^{-1} \Theta_i \end{pmatrix} \tag{A.40}$$

We remind that, as $\eta_N = 1$ almost surely for N large enough, it is possible to introduce η_N whenever it is useful without modifying the almost sure behaviour of the various terms. Lemma III.4 implies that$\mathbb{E}\{\eta W_i^* (W_i W_i^*)^{-1}\} = \mathbb{E}\{\eta (W_i W_i^*)^{-1} W_i\} = 0$ for $i = p, f$ while (III.19)-(III.20) lead to $\mathbb{E}\{\eta \Pi_i^W\} = c_N I_r + \mathcal{O}(N^{-k})$ and $\mathbb{E}\{\eta (W_i W_i^*)^{-1}\} = (1 - c_N)^{-1} (I_L \otimes R_N^{-1}) + \mathcal{O}(N^{-k})$ for each $k \in \mathbb{N}$. Combining these evaluations with the Poincaré-Nash inequality, we obtain immediately (III.75).

H. Proof of Lemma III.8

We just provide a brief justification of the first item of Lemma III.8. For this, we notice that the $2r \times 2r$ matrix $\mathcal{A}_i^* \mathbf{Q}_{\mathbf{ji}}^{\mathbf{W}} \mathcal{A}_i$ is given by

$$\mathcal{A}_i^* \mathbf{Q}_{\mathbf{ji}}^{\mathbf{W}} \mathcal{A}_i = \begin{pmatrix} -\Delta_i \tilde{\Theta}_i^* \Pi_i^{W,\perp} \\ \Theta_i^* (W_i W_i^*)^{-1} W_i^* \end{pmatrix} \mathbf{Q}_{\mathbf{ji}}^{\mathbf{W}} \left(-\Pi_i^{W,\perp} \tilde{\Theta}_i \Delta_i, W_i^* (W_i W_i^*)^{-1} \Theta_i \right) \tag{A.41}$$

We recall that, as $\eta_N = 1$ for each N large enough, we can add η_N everywhere in the 4 $r \times r$ blocks of $\mathcal{A}_i^* \mathbf{Q}_{\mathbf{ji}}^{\mathbf{W}} \mathcal{A}_i$ without modifying their almost sure behaviour. We first justify that the two $r \times r$ non diagonal blocks of $\mathcal{A}_i^* \mathbf{Q}_{\mathbf{ji}}^{\mathbf{W}} \mathcal{A}_i$ converge almost surely towards 0. For this, we first notice that, using the same arguments as in Lemma III.4, it can be easily shown that $\mathbb{E}\{\Pi_i^W \mathbf{Q}_{\mathbf{ji}}^{\mathbf{W}} \eta W_i^* (W_i W_i^*)^{-1}\} = \mathbb{E}\{\mathbf{Q}_{\mathbf{ji}}^{\mathbf{W}} \eta W_i^* (W_i W_i^*)^{-1}\} = 0$. Using the Poincaré-Nash inequality, it is possible to prove that $\mathbf{Var}\{a_N^* \Pi_i^{W,\perp} \mathbf{Q}_{\mathbf{ij}}^{\mathbf{W}} \eta W_i^* (W_i W_i^*)^{-1} b_N\} = \mathcal{O}_{z^2}(\frac{1}{N})$ and $\mathbf{Var}\{\left(a_N^* \Pi_i^{W,\perp} \mathbf{Q}_{\mathbf{ij}}^{\mathbf{W}} \eta W_i^* (W_i W_i^*)^{-1} b_N\right)^2\} = \mathcal{O}_{z^2}(\frac{1}{N^2})$, where a_N (resp. b_N) is a N dimensional (resp. ML-dimensional) deterministic vector such that $\sup_N \|a_N\| < +\infty$ (resp. $\sup_N \|b_N\| < +\infty$). This immediately implies that $\mathbb{E}\left|a_N^* \Pi_i^{W,\perp} \mathbf{Q}_{\mathbf{ij}}^{\mathbf{W}} \eta W_i^* (W_i W_i^*)^{-1} b_N\right|^4 = \mathcal{O}_{z^2}(\frac{1}{N^2})$ and that $a_N^* \Pi_i^{W,\perp} \mathbf{Q}_{\mathbf{ij}}^{\mathbf{W}} \eta W_i^* (W_i W_i^*)^{-1} b_N$, and thus $a_N^* \Pi_i^{W,\perp} \mathbf{Q}_{\mathbf{ij}}^{\mathbf{W}} W_i^* (W_i W_i^*)^{-1} b_N$ converge almost surely towards 0. The extension of these properties to the context of bounded random vectors (a_N, b_N) independent from the sequence $(v_n)_{n \in \mathbb{Z}}$ (see the proof of Lemma II.1) leads to the conclusion that the two $r \times r$ non diagonal blocks of $\mathcal{A}_i^* \mathbf{Q}_{\mathbf{ji}}^{\mathbf{W}} \mathcal{A}_i$ converge almost surely towards 0 as expected.

We now evaluate the almost sure behaviour of the two $r \times r$ diagonal blocks of $\mathcal{A}_i^* \mathbf{Q}_{\mathbf{ji}}^{\mathbf{W}} \mathcal{A}_i$, and consider the case $i = p$, $j = f$ without loss of generality. In the expression of $(\mathcal{A}_p^* \mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}} \mathcal{A}_p)_{11} = \Delta_p \tilde{\Theta}_p^* \Pi_p^{W,\perp} \mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}} \Pi_p^{W\perp} \tilde{\Theta}_p \Delta_p$, it is possible to replace $\Pi_p^{W,\perp} = I - \Pi_p^W$ by $I - \eta \Pi_p^W$. Using (III.10) and the resolvent identity

$$I + z\mathbf{Q}^{\mathbf{W}} = \mathbf{Q}^{\mathbf{W}} \begin{pmatrix} 0 & \eta \Pi_p^W \\ \eta \Pi_f^W & 0 \end{pmatrix} = \begin{pmatrix} 0 & \eta \Pi_p^W \\ \eta \Pi_f^W & 0 \end{pmatrix} \mathbf{Q}^{\mathbf{W}} \tag{A.42}$$

we obtain easily that

$$(I - \eta \Pi_p^W) \mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}} (I - \eta \Pi_p^W) = \mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}} - I_N - z\mathbf{Q}_{\mathbf{pp}}^{\mathbf{W}} - I_N - z\mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}} + \eta \Pi_p^W + z^2 \mathbf{Q}_{\mathbf{pf}}^{\mathbf{W}}$$

Using the Poincaré-Nash inequality, it is easy to check that if a_N and b_N are two deterministic N–dimensional vectors for which $\sup_N \|a_N\| < +\infty$ and $\sup_N \|b_N\| < +\infty$, then, it holds that $\mathrm{Var}\{a_N^* \eta \Pi_p^W b_N\} = \mathcal{O}(\frac{1}{N})$ and $\mathrm{Var}\{\left(a_N^* \eta \Pi_p^W b_N - \mathbb{E}(a_N^* \eta \Pi_p^W b_N)\right)^2\} = \mathcal{O}(\frac{1}{N^2})$. Therefore, $\mathbb{E}\left(a_N^* \eta \Pi_p^W b_N - \mathbb{E}(a_N^* \eta \Pi_p^W b_N)\right)^4 = \mathcal{O}(\frac{1}{N^2})$, so that $a_N^* \eta \Pi_p^W b_N - \mathbb{E}(a_N^* \eta \Pi_p^W b_N) \to 0$ almost surely. (III.20) thus leads to the conclusion that

$$a_N^* \eta \Pi_p^W b_N - c_N \, a_N^* b_N \to 0 \; a.s.$$

The use of Lemma III.6 implies that $a_N^* \Pi_p^{W,\perp} \mathbf{Q}_{\mathbf{fp}} \Pi_p^{W,\perp} b_N - ((1+z^2) c_N \mathbf{t}_N(z) - 1 - 2z\tilde{\mathbf{t}}_N(z) - (1-c_N)) a_N^* b_N \to 0$. Moreover, this property also holds when (a_N, b_N) are bounded random vectors (a_N, b_N) independent from the sequence $(v_n)_{n \in \mathbb{Z}}$. We deduce from this that

$$(\mathcal{A}_p^* \mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}} \mathcal{A}_p)_{11} - \left((1+z^2) c_N \mathbf{t}_N(z) - 1 - 2z\tilde{\mathbf{t}}_N(z) - (1-c_N)\right) \Delta_N \to 0 \; a.s.$$

holds. In order to obtain the expression stated in the Lemma, we refer to (III.63) and replace $c_N\mathbf{t}_N(z)$ by $c_N\mathbf{t}_N(z) = \frac{\tilde{\mathbf{t}}_N(z)}{z} + \frac{1-c_N}{z^2}$:

$$(1+z^2)c_N\mathbf{t}_N(z) - 1 - 2z\tilde{\mathbf{t}}_N(z) - (1-c_N) = \tilde{\mathbf{t}}_N(z)\left(\frac{1}{z} - z\right) + \frac{1-c_N}{z^2} - 1 \quad \text{(A.43)}$$

Let us remind that $\tilde{\mathbf{t}}_N$ satisfies Eq. (III.41) but in which term $\mathcal{O}_{z^2}(N^{-2})$ is replaced with 0, i.e.

$$(1-z^2)\tilde{\mathbf{t}}_N^2(z) + \left(\frac{2(1-c_N)}{z} - z\right)\tilde{\mathbf{t}}_N(z) + \frac{(1-c_N)^2}{z^2} = 0 \quad \text{(A.44)}$$

In order to simplify (A.43) we rewrite Eq. (A.44) as

$$(z\tilde{\mathbf{t}}_N(z) + (1-c_N))\left(\tilde{\mathbf{t}}_N(z)\left(\frac{1}{z} - z\right) + \frac{1-c_N}{z^2} - 1\right) + (1-c_N) + z(1-c_N)\tilde{\mathbf{t}}_N(z) = 0$$

and get immediately that the r.h.s. of (A.43) is equal to $-\frac{(1-c_N)(1+z\tilde{\mathbf{t}}_N(z))}{z\tilde{\mathbf{t}}_N(z)+(1-c_N)}$. This establishes the expression stated in the Lemma.

We finally evaluate the behaviour of $(\mathcal{A}_p^*\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\mathcal{A}_p)_{22} = \Theta_p^*(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}W_p^*(W_pW_p^*)^{-1}\Theta_p$. We recall that $W_i = (I\otimes R^{1/2})W_{i,iid}$ for $i = p, f$, so that

$$\Theta_p^*(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}W_p^*(W_pW_p^*)^{-1}\Theta_p = \Theta_p^*(I\otimes R^{-1/2})(W_{p,iid}W_{p,iid}^*)^{-1}W_{p,iid}\mathbf{Q}_{\mathbf{iid},\mathbf{fp}}^{\mathbf{W}}W_{p,iid}^*(W_{p,iid}W_{p,iid}^*)^{-1}(I\otimes R^{-1/2})\Theta_p$$

because $\mathbf{Q}_{\mathbf{iid},\mathbf{fp}}^{\mathbf{W}} = \mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}$. It is thus sufficient to study the behaviour of

$$a_N^*\eta_N(W_{p,iid}W_{p,iid}^*)^{-1}W_{p,iid}\mathbf{Q}_{\mathbf{iid},\mathbf{fp}}^{\mathbf{W}}W_{p,iid}^*(W_{p,iid}W_{p,iid}^*)^{-1}b_N$$

where a_N, b_N are deterministic ML–dimensional vectors such that $\sup_N\|a_N\| < +\infty$ and $\sup_N\|b_N\| < +\infty$. We also recall that the regularization term η_N is built from matrix W_{iid}. In order to simplify the notations, we prefer to denote $W_{i,iid}$ by W_i in the following. After some calculations, the Poincaré-Nash inequality leads to

$$a_N^*\eta_N(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\eta W_p^*(W_pW_p^*)^{-1}b_N - \mathbb{E}\left(a_N^*\eta_N(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\eta W_p^*(W_pW_p^*)^{-1}b_N\right) \to 0 \ a.s.$$

It is thus sufficient to evaluate $\mathbb{E}\{(\eta W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\eta W_p^*(W_pW_p^*)^{-1}\}$ using the integration by parts formula. By Lemma III.4, this matrix is diagonal, and we therefore consider its diagonal elements. For this, we need to repeat the calculations of Section III-A2. In order to avoid to reproduce another tedious and similar calculations, we provide only the ideas and main steps. It is first necessary to apply the integration by parts formula for $\sum_{r,t,m_2,i_2}\mathbb{E}\{\eta{\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}}_{rt}\bar{W}_{p,i_1t}^{m_1}((W_pW_p^*)^{-1})_{i_1i_2}^{m_1m_2}W_{p,i_2r}^{m_2}\}$ and follow the calculations of Section III-A2 using similar arguments. We first obtain

$$\begin{aligned}\mathbb{E}\{\left(\eta(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}W_p^*\right)_{i_1i_1}^{m_1m_1}\} &= \mathbb{E}\{\eta((W_pW_p^*)^{-1})_{i_1i_1}^{m_1m_1}\}\frac{1}{N}\left(\mathbb{E}\{\mathrm{Tr}\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\} - \mathbb{E}\{\mathrm{Tr}\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\eta\Pi_p^W\}\right)\\ &- \mathbb{E}\left\{\eta\left((W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\eta W_p^*(W_pW_p^*)^{-1}\right)_{i_1i_1}^{m_1m_1}\right\}\frac{1}{N}\mathbb{E}\{\mathrm{Tr}\eta\Pi_p^{W,\perp}\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\} + \mathcal{O}_{z^2}^N(N^{-3/2})\end{aligned}$$

Since $\mathbb{E}\{\eta(W_pW_p^*)^{-1}\} = (1-c_N)^{-1}I_N + \mathcal{O}^N(N^{-3/2})$ and that the equality $\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\eta\Pi_p^W = I_N + z\mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}}$ holds (see (A.42)), we can simplify the r.h.s. of the last equation:

$$\begin{aligned}\mathbb{E}\{\eta\left((W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}W_p^*\right)_{i_1i_1}^{m_1m_1}\} &= \frac{1}{1-c_N}(\boldsymbol{\alpha}_N - 1 - z\tilde{\boldsymbol{\alpha}}_N)\\ &- \mathbb{E}\left\{\eta\left((W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\eta W_p^*(W_pW_p^*)^{-1}\right)_{i_1i_1}^{m_1m_1}\right\}(\boldsymbol{\alpha}_N - 1 - z\tilde{\boldsymbol{\alpha}}_N) + \mathcal{O}_{z^2}(N^{-3/2}) \quad \text{(A.45)}\end{aligned}$$

We express $\mathbb{E}\{\eta\left((W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}}W_p^*\right)_{i_1i_1}^{m_1m_1}\}$ similarly:

$$\begin{aligned}\mathbb{E}\{\eta\left((W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}}W_p^*\right)_{i_1i_1}^{m_1m_1}\} &= \mathbb{E}\{\eta((W_pW_p^*)^{-1})_{i_1i_1}^{m_1m_1}\}\frac{1}{N}\left(\mathbb{E}\{\mathrm{Tr}\mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}}\} - \mathbb{E}\{\mathrm{Tr}\mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}}\eta\Pi_p^W\}\right)\\ &- \mathbb{E}\left\{\eta\left((W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}}\eta W_p^*(W_pW_p^*)^{-1}\right)_{i_1i_1}^{m_1m_1}\right\}\frac{1}{N}\mathbb{E}\{\mathrm{Tr}\eta\Pi_p^{W,\perp}\mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}}\} + \mathcal{O}_{z^2}(N^{-3/2})\end{aligned}$$

We remark that $N^{-1}\mathbb{E}\{\mathrm{Tr}\mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}}\eta\Pi_p^{W,\perp}\} = -z^{-1}\mathbb{E}\{N^{-1}\eta\mathrm{Tr}\Pi_p^{W,\perp}\} = -\frac{1-c_N}{z} + \mathcal{O}_{z^2}(N^k)$ for each integer k (see (III.30)). The last equation can thus be rewritten as

$$\mathbb{E}\{\eta\left((W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}}W_p^*\right)_{i_1i_1}^{m_1m_1}\} = -\frac{1}{z} + \frac{1-c_N}{z}\mathbb{E}\left\{\eta\left((W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\eta W_p^*(W_pW_p^*)^{-1}\right)_{i_1i_1}^{m_1m_1}\right\} + \mathcal{O}_{z^2}(N^{-3/2}) \quad \text{(A.46)}$$

Moreover, using the resolvent identity, $(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{ff}}^{\mathbf{W}}W_p^*$ can be rewritten as $(W_pW_p^*)^{-1}W_p(-z^{-1}I_N + z^{-1}(\Pi_f^W\Pi_p^W - z^2)^{-1}\Pi_f^W\Pi_p^W)W_p^* = -z^{-1}I_N + z^{-1}(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\eta\Pi_p^W W_p^*$. Using the obvious identity $\Pi_p^W W_p^* = W_p^*$ and comparing (A.45) and (A.46) we that:

$$\mathbb{E}\{\left(\eta(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\eta W_p^*(W_pW_p^*)^{-1}\right)_{i_1i_1}^{m_1m_1}\} = \frac{\boldsymbol{\alpha}_N - 1 - z\tilde{\boldsymbol{\alpha}}_N}{(1-c_N)((1-c_N)+\boldsymbol{\alpha}_N - 1 - z\tilde{\boldsymbol{\alpha}}_N)} + \mathcal{O}_{z^2}(N^{-3/2})$$

As we can see, all diagonal elements of $\mathbb{E}\{\eta(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\eta W_p^*(W_pW_p^*)^{-1}\}$ are equal up to an error term. Therefore, the matrix $\mathbb{E}\{\eta(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\eta W_p^*(W_pW_p^*)^{-1}\}$ is a multiple of I_N up to an error term. We thus conclude that

$$a_N^*(W_pW_p^*)^{-1}W_p\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}W_p^*(W_pW_p^*)^{-1}b_N - \frac{c_N\mathbf{t}_N - 1 - z\tilde{\mathbf{t}}_N}{(1-c_N)((1-c_N)+c_N\mathbf{t}_N - 1 - z\tilde{\mathbf{t}}_N)}a_N^*b_N \to 0. \quad \text{(A.47)}$$

After replacing $c_N\mathbf{t}_N$ with $\frac{\tilde{\mathbf{t}}_N(z)}{z} + \frac{1-c_N}{z^2}$ we find that $c_N\mathbf{t}_N - 1 - z\tilde{\mathbf{t}}_N = \tilde{\mathbf{t}}_N(z)(\frac{1}{z} - z) + \frac{1-c_N}{z^2} - 1$ which is also the expression obtained in (A.43). We remind that

$$\tilde{\mathbf{t}}_N(z)\left(\frac{1}{z} - z\right) + \frac{1-c_N}{z^2} - 1 = -\frac{(1-c_N)(1+z\tilde{\mathbf{t}}_N(z))}{z\tilde{\mathbf{t}}_N(z) + (1-c_N)}$$

Plugging this expression into (A.47), and remarking that (A.47) still holds when (a_N, b_N) are random bounded vectors independent from $(v_n)_{n\geq 1}$, we obtain the asymptotic behaviour $(\mathcal{A}_p^*\mathbf{Q}_{\mathbf{fp}}^{\mathbf{W}}\mathcal{A}_p)_{22}$ stated in the Lemma.