



HAL
open science

Asymptotic spectrum of weighted sample covariance: a Marcenko-Pastur generalization

Benoit Oriol

► **To cite this version:**

Benoit Oriol. Asymptotic spectrum of weighted sample covariance: a Marcenko-Pastur generalization. 2024. hal-04746732

HAL Id: hal-04746732

<https://hal.science/hal-04746732v1>

Preprint submitted on 21 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Asymptotic spectrum of weighted sample covariance: a Marcenko-Pastur generalization

Benoît Oriol

CEREMADE, Université Paris-Dauphine, PSL, Paris, France
Société Générale Corporate and Investment Banking, Puteaux, France
benoit.oriol@dauphine.eu

Abstract

We propose an extension of the high dimensional spectrum analysis of sample covariance in the setting of the weighted sample covariance. We derive an asymptotic equation characterizing the limit density of the weighted sample eigenvalues generalizing for weighted sample covariance matrices the Marcenko-Pastur theorem.

Keywords: weighted covariance, asymptotic spectrum, Random Matrix Theory, Cauchy-Stieltjes transform

1 Introduction and related work

The spectrum of sample covariance and its behavior in Kolmogorov asymptotics, when the dimension grows approximately linearly with the number of samples, is a central topic in high dimensional covariance and precision matrix estimation. Among the large literature on covariance estimation, the high dimensional framework has known several recent improvements thanks to Random Matrix Theory (RMT). Main high dimensional approaches are gathered in recent surveys of the field [1, 2, 3, 4, 5].

The applications of those estimators are diverse, and an extensive list was made by Ledoit and Wolf [6] in their literature review. We can include new works in this list, as climatology [7], neuroscience [8], or sensor monitoring [9].

A main topic concerns the spectrum deformation due to the lack of samples with regard to the dimension. To address this challenge, an idea is to shrink each sample eigenvalue in order to correct the deformation. This key idea of shrinkage goes back to Stein in 1956 [10] for Gaussian mean estimation.

Covariance estimation successfully used this idea firstly with a linear formulation. Some linear shrinkage based estimators are: Ledoit and Wolf linear shrinkage [11], Oracle Approximating Shrinkage Estimator for Gaussian observations [12], linear shrinkage under factor models [13] and [14], analytical setting of linear shrinkage [15], shrinkage of Tyler M-estimator [16], linear shrinkage of elliptical distributions [17], or Ledoit-Wolf shrinkage with unknown mean [18].

RMT gave the mathematical tools to understand the deformation induced by a number of samples and a dimension of the same order of magnitude. The asymptotic sample covariance spectrum is described by a fundamental equation discovered by Marcenko and Pastur in a work of 1967 [19]. Later, Bai, Choi and Silverstein wrote a series of articles between 1995 and 1998 on the generalization of Marcenko-Pastur equation to Hermitian matrices and on characterizations of the spectrum support [20, 21, 22, 23].

These works on the deformation of the spectrum of the true covariance in the sample covariance led to Ledoit and Pécché equations for optimal non-linear shrinkage in the class of rotation-invariant estimators [24], with important consequences in covariance estimation [6, 25, 26]

These works focus on the sample covariance. However, applications in dynamic environments such as neurosciences [27], finance [28], suggest that the sample covariance suffers from the non-stationarity of the data. Weighting schemes, such as the exponential weighted moving average (EWMA), are a model-free approach, with possibly very few parameters depending on the chosen weighting schemes, and represent a transparent candidate for complex and hard-to-model dynamics. Their use in covariance estimation is not new: Pafka et al. in 2004 [29] worked on shrinkage for exponentially-weighted sample covariance in a Gaussian setting for portfolio management, which was further studied for covariance filtering [30], Svensson in 2007 [31] used Marcenko-Pastur to study the spectrum of exponentially-weighted sample covariance when the covariance is the identity, and recently Tan et al. in 2023 [32] developed a NERCOME-like approach to use non-linear shrinkage on exponentially-weighted sample covariance with applications for dynamic brain connectivity.

As they explain in the latter, no tool from Random Matrix Theory exists at our knowledge for weighted sample covariance to replicate the methods used for the sample covariance to extract its asymptotic spectrum.

This work aims at filling this gap by giving an asymptotic description of the spectrum of weighted sample covariance, as a **generalization of Marcenko-Pastur theorem for weighted sample covariance matrices**. Inspired from the works of Marcenko and Pastur [19], Bai, Choi and Silverstein [20, 21, 22], we show that under the standard, and very minimal, hypotheses of RMT, the spectrum of a weighted sample covariance converges almost surely to a distribution solving a Marcenko-Pastur-like functional equation on its Cauchy-Stieltjes transform.

2 Notation and assumptions

Let us introduce the following notation. Notation is not constant across major works on the spectrum of sample covariances. In our work, we chose to follow mostly Silverstein one [20].

Notation 1 (The data matrix). *There are N samples of dimension n . We have:*

- $c_n = \frac{n}{N}$ the concentration ratio,
- Z_n is the noise $n \times N$ matrix composed of i.i.d centered complex entries of variance 1,
- T_n is the true covariance, a non-negative definite Hermitian matrix of size $n \times n$,
- W_n is the weight matrix, a $N \times N$ diagonal non-negative real matrix,
- $Y_n = T_n^{1/2} Z_n$ is the observed data matrix.

In the following, the subscripts n are omitted when no confusion is possible.

In this work, the object of interest is the weighted sample covariance B_n , particularly its spectrum of eigenvalues which are introduced below.

Notation 2 (Weighted sample covariance). *For $n \in \mathbb{N}^*$, the weighted sample covariance is defined by:*

$$B_n := \frac{1}{N} Y_n W_n Y_n^*. \quad (1)$$

We note $(\tau_1^{(n)}, \dots, \tau_n^{(n)})$ the eigenvalues of T_n in decreasing order, and $(\lambda_1^{(n)}, \dots, \lambda_n^{(n)})$ the eigenvalues of B_n in decreasing order.

Additionally, we will need matrix in the proof:

$$\underline{B}_n := \frac{1}{N} W_n^{1/2} Y_n^* Y_n W_n^{1/2}. \quad (2)$$

\underline{B}_n and B_n have the same eigenvalues, except for $|n - N|$ zero eigenvalues.

Example 1 (Standard weighting). *The most common choice of weighting is of course the constant weighting: $W_n = I_N$. In this situation, B_n is the standard sample covariance, and its asymptotic spectrum is covered in RMT by the work of Marcenko and Pastur [19].*

Example 2 (Exponentially weighted scheme). *Another common choice in time series analysis is the exponentially weighted scheme. Parametrized by some $\alpha \in \mathbb{R}_+^*$, we define the weights as:*

$$\begin{aligned} \forall i \in \llbracket 1, N \rrbracket, (W_n)_{ii} &= \beta e^{-\alpha i/N}, \\ \beta &= e^{-\alpha/N} \frac{1 - e^{-\alpha/N}}{1 - e^{-\alpha}}. \end{aligned} \quad (3)$$

Notation 3 (Empirical spectrum distribution). *We consider a Hermitian matrix A of size $n \times n$ with real eigenvalues (μ_1, \dots, μ_n) . We define the empirical spectrum distribution of A , F^A , as:*

$$F^A := \frac{1}{n} \sum_{i=1}^n 1_{[\mu_i, +\infty[}. \quad (4)$$

In particular, we have:

$$\begin{aligned} F_n &:= F^{B_n} = \frac{1}{n} \sum_{i=1}^n 1_{[\lambda_i^{(n)}, +\infty[}, \\ \underline{F}_n &:= F^{\underline{B}_n} = (1 - c_n) 1_{[0, +\infty[} + c_n F_n. \end{aligned} \quad (5)$$

A key tool to manipulate and understand the asymptotic behavior of F_n is the Cauchy-Stieltjes transform: a transform of a real finite measure into a complex function defined on $\mathbb{C}_+ = \{z \in \mathbb{C} \mid \text{Im}(z) > 0\}$.

Notation 4 (Cauchy-Stieltjes transform). *For a real finite measure H , we denote m_H its Cauchy-Stieltjes transform:*

$$\forall z \in \mathbb{C}_+, m_H(z) = \int \frac{1}{\tau - z} dH(\tau). \quad (6)$$

In particular, for all $z \in \mathbb{C}_+$, we denote:

$$\begin{aligned} m_n(z) &:= m_{F_n}(z) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\lambda_i^{(n)} - z} \\ &= \frac{1}{n} \text{tr}((B_n - zI)^{-1}), \\ \underline{m}_n(z) &:= m_{\underline{F}_n}(z) = \frac{1}{N} \text{tr}((\underline{B}_n - zI)^{-1}) \\ &= -\frac{1 - c_n}{z} + c_n m_n(z). \end{aligned} \quad (7)$$

We handle the weakly convergence of F_n through the pointwise almost sure convergence of its Cauchy-Stieltjes transform m_n using the following fundamental theorem.

Theorem 1 (Cauchy-Stieltjes convergence, Theorem 5.8.3 [33]). *Let $(\mu_n)_n$ be random probability measures, μ be a deterministic probability measure, and $Z \subset \mathbb{C}_+$ that has an accumulation point in \mathbb{C}_+ , then with $\xrightarrow[n \rightarrow \infty]{\implies}$ denoting the weak convergence:*

$$\begin{aligned} \mu_n &\xrightarrow[n \rightarrow \infty]{\implies} \mu \text{ almost surely} \iff \\ \forall z \in Z, m_{\mu_n}(z) &\xrightarrow[n \rightarrow \infty]{\implies} m_\mu(z) \text{ almost surely.} \end{aligned} \quad (8)$$

We consider the following set of mild assumptions necessary to the proof of the main result. Except for Assumption (d) which concerns specifically the weights, those assumptions were introduced by Silverstein [22] in his proof of Marcenko-Pastur theorem for Hermitian matrices.

Assumption 1. *Assume that:*

- (a) *For all $n \in \mathbb{N}^*$ and $i \in \llbracket 1, N \rrbracket, j \in \llbracket 1, n \rrbracket$, the $(X_n)_{i,j} \in \mathbb{C}$ are i.i.d random variables with $\mathbb{E}[(X_n)_{i,j}] = 0$ and $\mathbb{E}[|(X_n)_{i,j}|^2] = 1$.*
- (b) *$N = N(n)$ and $c_n = \frac{n}{N(n)} \xrightarrow[n \rightarrow \infty]{} c \in \mathbb{R}_+^*$.*
- (c) *$F^{T_n} \xrightarrow[n \rightarrow \infty]{\implies} H$ almost surely where H is a probability distribution function - p.d.f. - on \mathbb{R}_+ .*

(d) $F^{W_n} \xrightarrow[n \rightarrow \infty]{\implies} D$ almost surely where $D \in \mathcal{L}^1(\mathbb{R}_+)$ is a p.d.f., and almost surely $\int x dF^{W_n}(x) \xrightarrow[n \rightarrow \infty]{} \int x dD(x) \in \mathbb{R}_+$.

(e) X_n, T_n and W_n are independent.

Assumption (a) makes sure that the covariance is well-defined and that $\mathbb{E}[\frac{1}{N} Y_n Y_n^*] = T_n$. Assumption (b) is the standard asymptotic setting we consider in RMT: the dimension and the number of samples are of the same order of magnitude. Assumption (c) and (d) ensure that the population covariance spectrum and the weight distribution converge: those two objects are essential to describe asymptotically the spectrum of B_n . The last assumption makes sure, for example, that the weights have no dependence on the drawn samples.

3 Main result: convergence of m_n and F_n

The main result of this work is the following theorem. It characterizes the asymptotic non random spectrum F of B_n as $n \rightarrow \infty$ through its Cauchy-Stieltjes transform m_F .

Theorem 2 (Asymptotic spectrum). *Assume (a) to (e). Then, almost surely, F_n converges weakly to a nonrandom p.d.f F , whose Cauchy-Stieltjes transform $m := m_F$ satisfies for all $z \in \mathbb{C}_+$:*

$$m(z) = \int \frac{1}{\tau \tilde{m}(z) - z} dH(\tau), \quad (9)$$

where for all $z \in \mathbb{C}_+$, $\tilde{m}(z)$ is the unique solution in $\mathbb{C} \setminus \mathbb{C}_+$ of the following equation:

$$\tilde{m}(z) = \int \frac{\delta}{1 + \delta c \int \frac{\tau}{\tau \tilde{m}(z) - z} dH(\tau)} dD(\delta). \quad (10)$$

This theorem describes the asymptotic non-random behavior of F_n , through its Cauchy-Stieltjes transform, as Marcenko-Pastur theorem does for the standard sample covariance.

Remark 1 (Role of \tilde{m}). *In Equation (9), we note that if we fix $\tilde{m}(z) = 1$, then the right side of the equation becomes $m_H(z)$, so $F = H$. An interpretation of this equation is that $\tilde{m}(z)$ carries out the deformation of H in its Cauchy-Stieltjes transform.*

We only give here the main idea of the proof which is given in the Appendix. We follow the procedure introduced in the proof of [22], and we make it work in our setting using recent developments of Lebesgue theorem [34]. We firstly truncate the noise, the population eigenvalues and the weights by roughly $\log(n)$. This transform does not affect the asymptotic behavior of F_n . Doing that, we can apply a result from concentration theory that leads to the almost sure pointwise convergence of m_n towards m . Through careful use of Lebesgue theorem for varying measures, we can extract an integral description of this limiting m that uses \tilde{m} . Finally, we check that m is indeed the Cauchy-Stieltjes transform of some p.d.f., that we denote F .

Examples of asymptotic densities are shown in figure 1 and 2.

To understand better what is behind \tilde{m} , we can go back in the setting of the standard non-weighted covariance matrix.

Remark 2 (Role of \tilde{m} when $D = 1_{[1,+\infty[}$). *When $D = 1_{[1,+\infty[}$, we have that for all $z \in \mathbb{C}_+$, $\tilde{m}(z) = 1 - c(1 + zm(z)) = -zm(z)$.*

The conclusion of this remark is that \tilde{m} in the setting of weighted sample covariance has a similar role than \underline{m} has in the uniformly weighted setting.

4 Experimental results

4.1 Impact of weight distribution on the asymptotic spectrum and on its support

We show the behavior of the asymptotic density when we vary the weight distribution. The reference distribution of true eigenvalues is $H = 0.2 \times 1_{[1,\infty[} + 0.4 \times 1_{[3,\infty[} + 0.4 \times 1_{[10,\infty[}$, introduced by

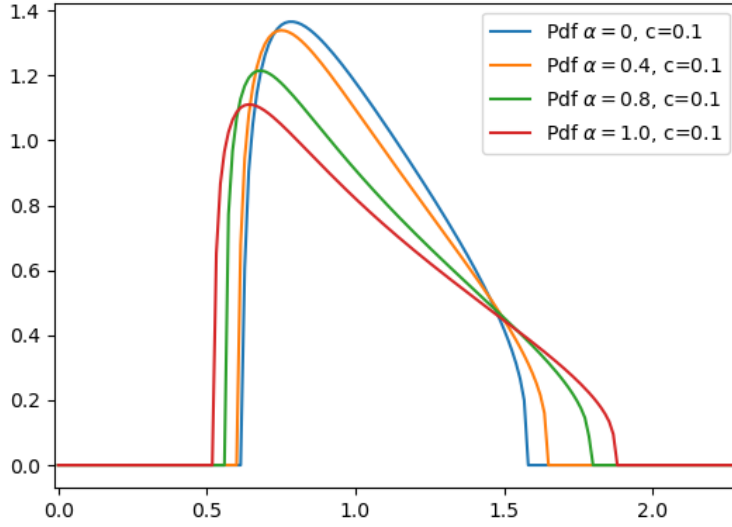


Figure 1: Different approximations regarding parameter α of the asymptotic density of weighted sample eigenvalues for $c = 0.1$, true covariance spectrum distribution $H = 1_{[1, \infty[}$ and weight distribution $D_\alpha = \frac{1}{2}1_{[1-\alpha, \infty[} + \frac{1}{2}1_{[1+\alpha, \infty[}$. $\alpha = 0$ corresponds to the classic Marcenko-Pastur density for $c = 0.1$.

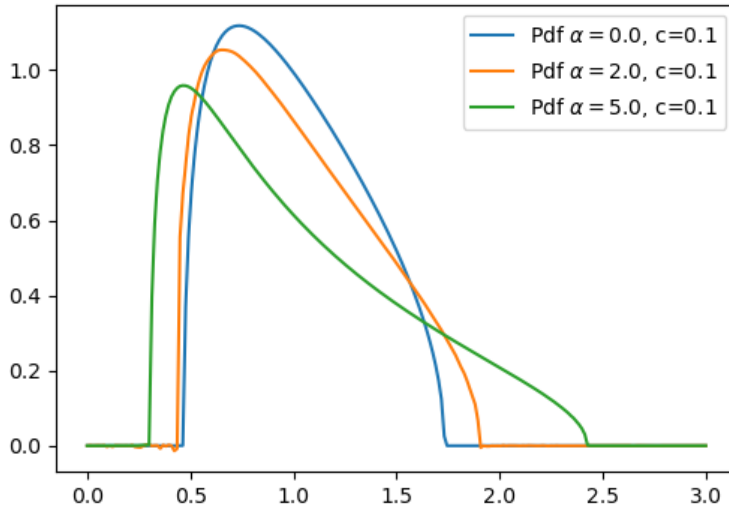


Figure 2: Different approximations regarding parameter α of the asymptotic density of weighted sample eigenvalues for $c = 0.1$, true covariance spectrum distribution $H = 1_{[1, \infty[}$ and weight distribution $D_\alpha : x \in [\beta e^{-\alpha}, \beta] \mapsto 1 + \frac{1}{\alpha} \log\left(\frac{x}{\beta}\right)$ with $\beta = \frac{\alpha}{1-e^{-\alpha}}$, defined in Definition 1. $\alpha = 0$ corresponds to the classic Marcenko-Pastur density for $c = 0.1$.

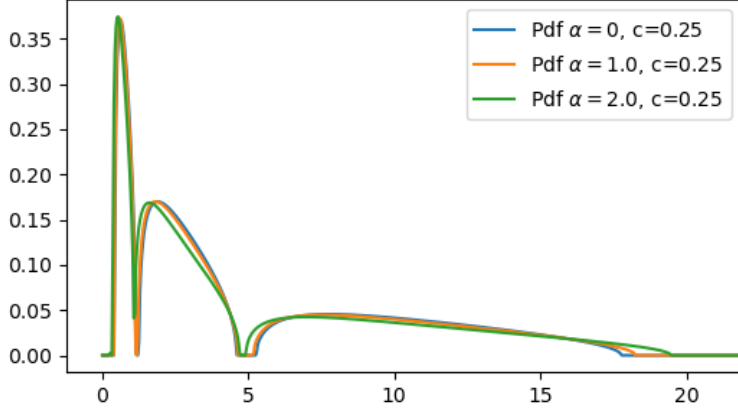


Figure 3: Approximations of the asymptotic density of weighted sample eigenvalues for $c = 0.25$, with $H = 0.2 \times 1_{[1, \infty[} + 0.4 \times 1_{[3, \infty[} + 0.4 \times 1_{[10, \infty[}$ and uniform weight distribution of parameter α .

Bai and Silverstein [23] and used by Ledoit and Péché [24]. The code is given in supplementary material. All experiments used a Lenovo laptop, Intel i7, 32Go RAM.

This distribution helps visualizing the phenomenon of spectral separation: areas of exclusion between the true eigenvalues where the asymptotic density is null. Those areas vanish when c is increasing, as was shown in the classic setting of Marcenko-Pastur theory, but we show here that the same phenomenon appears when the weight distribution is smoothly spreading.

For that matter, we are considering different weight distribution parametrized by $\alpha \in \mathbb{R}_+$ where α controls how much we are spreading the weights. We define the exponentially-weighted distribution, that corresponds to the distribution of weights in an Exponentially-Weighted Moving Average - EWMA -, used in time series analysis.

Definition 1 (α -exponentially weighted distribution). We fix $\alpha \in \mathbb{R}_+$. We define this law such as its cdf D follows: $D(\beta e^{-\alpha t}) = 1 - t$ for $t \in [0, 1]$. Moreover, we impose that $\int \delta dD(\delta) = 1$. We finally have: $\forall x \in [\beta e^{-\alpha}, \beta]$, $D(x) = 1 + \frac{1}{\alpha} \log\left(\frac{x}{\beta}\right)$, with $\beta = \frac{\alpha}{1 - e^{-\alpha}}$.

We consider three weight distributions:

- A uniform distribution with $[1 - \alpha/2, 1 + \alpha/2]$ for $\alpha \in [0, 2]$. The experiment is shown in figure 3. There are 2 spectral separation for $\alpha = 0$ and $\alpha = 1$. For $\alpha = 2$ there is only one left.
- A mixture of 2 diracs, with $D = \frac{1}{2}1_{[1-\alpha, \infty[} + \frac{1}{2}1_{[1+\alpha, \infty[}$ for $\alpha \in [0, 1]$. The experiment is shown in figure 4. There are 2 spectral separation for $\alpha = 0$. For $\alpha = 0.7$ and $\alpha = 1$ there is no more spectral separation.
- An exponentially-weighted distribution. The experiment is shown in figure 5. There are 2 spectral separation for $\alpha = 0$, only 1 for $\alpha = 2$ and no more for $\alpha = 5$.

4.2 New spectral gaps due to weight gaps

Silverstein and Choi [22] proved, in the case of the non-weighted sample covariance, that when H is a finite sum of K diracs, then there are at most $K - 1$ spectral gaps in the asymptotic spectrum F .

In the case of the weighted covariance spectrum, this result is not true anymore, as gaps in D can lead to new spectral gaps in F .

To illustrate this new phenomenon, we consider:

- $c = 0.05$ and $c = 0.01$,
- $H = 0.2 \times 1_{[1, \infty[} + 0.4 \times 1_{[3, \infty[} + 0.4 \times 1_{[10, \infty[}$,

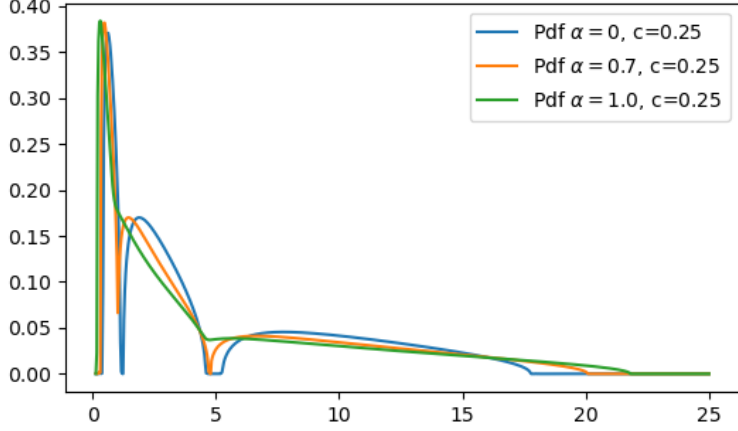


Figure 4: Approximations of the asymptotic density of weighted sample eigenvalues for $c = 0.25$, with $H = 0.2 \times 1_{[1,\infty[} + 0.4 \times 1_{[3,\infty[} + 0.4 \times 1_{[10,\infty[}$ and weight distribution from a mixture of 2 diracs of parameter α .

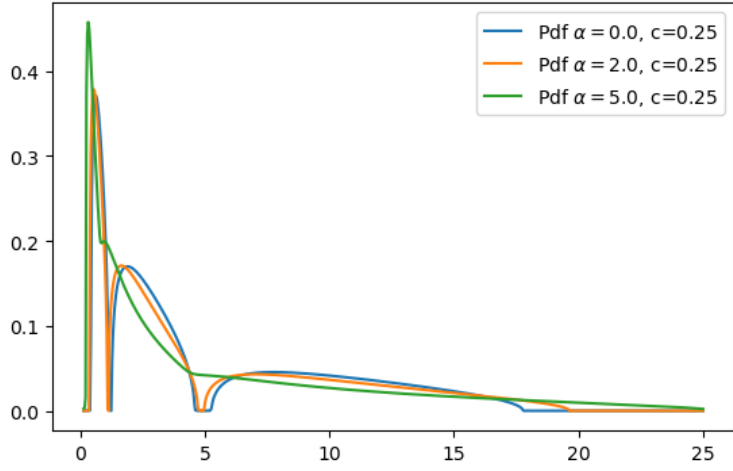


Figure 5: Approximations of the asymptotic density of weighted sample eigenvalues for $c = 0.25$, with $H = 0.2 \times 1_{[1,\infty[} + 0.4 \times 1_{[3,\infty[} + 0.4 \times 1_{[10,\infty[}$ and exponentially-weighted distribution of parameter α .

- $D = (1 - \frac{1}{2\alpha}) \times 1_{[\frac{\alpha}{2\alpha-1}, \infty[} + \frac{1}{2\alpha} 1_{[\alpha, \infty[}$ for $\alpha \in \{1, 50\}$ with $c = 0.05$ and $\alpha \in \{1, 200\}$ with $c = 0.01$.

In the standard non-weighted sample covariance case, we expect to find two spectral gaps roughly when $c < 0.34$, one spectral gap for $c \in [0.34, 0.40]$ and no spectral gaps for $c > 0.40$. These values result from the theoretical study of Silverstein and Choi [22].

In our setting, we indeed have at most two spectral gaps with $\alpha = 1$, which coincides with the non-weighted situation. However, for $\alpha = 50$ and $\alpha = 200$, a third spectral gap appears at smaller values of c , respectively numerically for $c \leq 0.05$ and $c \leq 0.01$. Results are shown in figure 6.

4.3 Speed of convergence with heavy tails

This last experimental part aims at empirically showing the speed of convergence to the asymptotic density we described in Theorem 2, depending on the heaviness of tails.

To visualize this point, we are plotting the histogram of the empirical spectrum of the weighted sample covariance along with the theoretical asymptotic density.

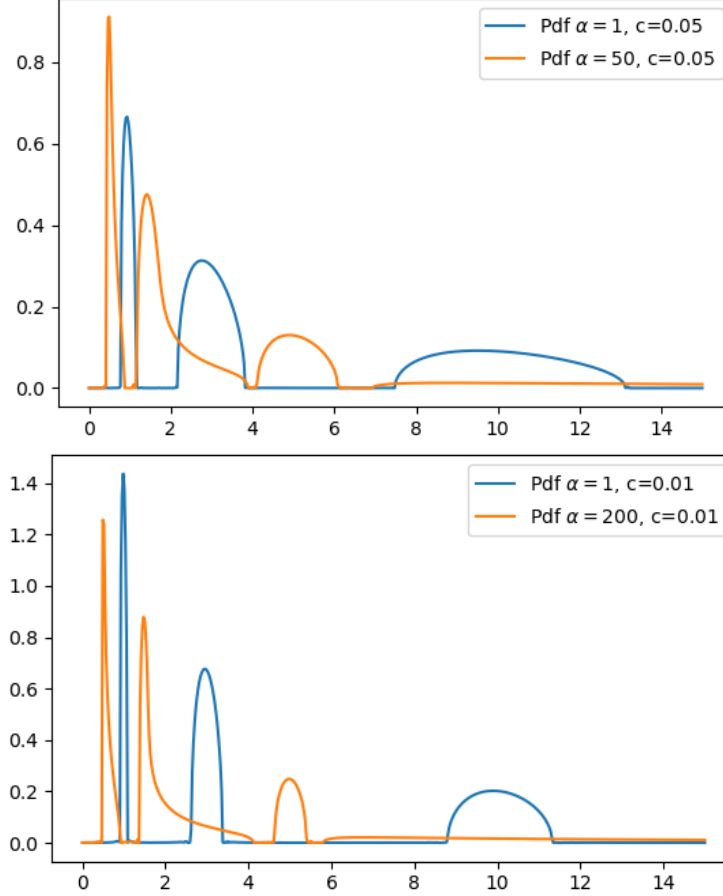


Figure 6: Approximations of the asymptotic density of weighted sample eigenvalues for $c = 0.05$ and $c = 0.01$, with $H = 0.2 \times 1_{[1,\infty[} + 0.4 \times 1_{[3,\infty[} + 0.4 \times 1_{[10,\infty[}$ and weight distribution from a mixture of 2 diracs of parameter α for spectral gaps.

We draw the (X_{ij}) from a normalized t-distribution with varying degree of freedom $\nu \in \{2.5, 3.5, 4\}$.

As previously, we chose $H = 0.2 \times 1_{[1,\infty[} + 0.4 \times 1_{[3,\infty[} + 0.4 \times 1_{[10,\infty[}$ and an exponentially-weighted distribution of parameter $\alpha = 1$ for the weights. The concentration ratio is set at $c = 0.25$ and the dimension at $n = 3000$. The convergence is meant to be almost sure, so we did only one draw of spectrum per plot. The results are shown in figure 7.

The conclusion of this experiment is that heavy tails tend to create rare but high eigenvalues, making the convergence slower as the tail grows. We recall that the convergence of the distributions is shown under weak convergence. So, the presence of very high eigenvalues far from the support of F is not a contradiction as long as their frequency of appearance converges to zero, which experimentally is the case. This type of behavior is clearly visible for $\nu = 2.5$ in Figure 7 where eigenvalues higher than 40 are sampled, but in very low proportion.

5 Limitations and extensions

The assumptions needed to prove the main theoretical result - Theorem 2 - are quite light, particularly on the properties of the underlying distributions. However, one can remark the disymmetry between H and D : we require, in Assumption (d), that $D \in \mathcal{L}^1(\mathbb{R}_+)$ and that $\int x dF^{W_n}(x) \rightarrow \int x dD(x)$ a.s. This assumption is essential in our proof, basically to remain in a compact and find a converging subsequence. But we experimentally failed to find an example that proves the minimality of this condition, so we conjecture that a weaker form should suffice.

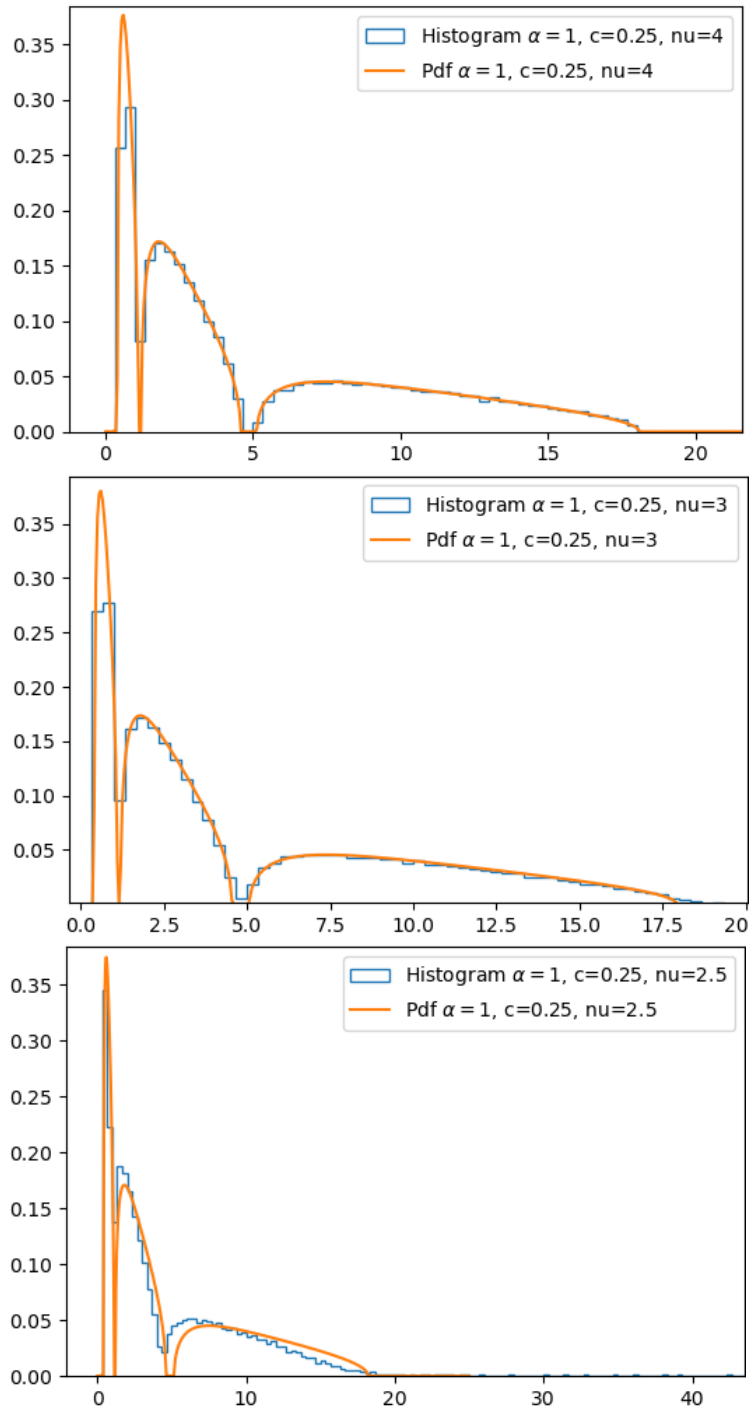


Figure 7: Histogram from a t-distribution with $\nu = 4$, $\nu = 3$ and $\nu = 2.5$ from top to bottom, and asymptotic theoretical density, for $n = 3000$, $c = 0.25$, $H = 0.2 \times 1_{[1, \infty[} + 0.4 \times 1_{[3, \infty[} + 0.4 \times 1_{[10, \infty[}$ and exponentially-weighted distribution of parameter $\alpha = 1$.

Secondly, W_n is supposed to be diagonal. This is our practical interest in multivariate time series analysis, and this is important in the proof to properly use the independence assumption. Out of curiosity, we tried to rotate W_n under Student noise in order to make it non-diagonal, and our experiments - given in Appendix - suggest that the conclusions of Theorem 2 hold in this situation, but a theoretical and empirical study needs to be made to ensure it.

6 Conclusion

In this work, we characterize the asymptotic spectrum of weighted sample covariance, extending Random Matrix Theory tools developed for the equally weighted sample covariance.

Experimentally, we show the behavior of the weighted sample covariance spectrum under different concentration ratios and weight distributions, and expose a new phenomenon on the spectral gaps. Finally, we confirm empirically the robustness of the theory and speed of convergence when the underlying distribution has heavy tails.

This theory paves a way for further works and extensions, as the optimal asymptotic non-linear shrinkage in the spirit of [24].

Acknowledgment

We would like to thank Alexandre Miot and Gabriel Turinici for their insights and advices all along the work.

References

- [1] Mohsen Pourahmadi. Covariance estimation: The glm and regularization perspectives. *Statistical Science*, 26(3):369 – 387, 2011.
- [2] Jianqing Fan, Yuan Liao, and Han Liu. An overview of the estimation of large covariance and precision matrices. *The Econometrics Journal*, 19(1):C1–C32, 2016.
- [3] Yuan Ke, Stanislav Minsker, Zhao Ren, Qiang Sun, and Wen-Xin Zhou. User-friendly covariance estimation for heavy-tailed distributions, 2019.
- [4] Olivier Ledoit and Michael Wolf. The power of (non-)linear shrinking: A review and guide to covariance matrix estimation. *Journal of Financial Econometrics*, 20(1):187–218, 06 2020.
- [5] Clifford Lam. High-dimensional covariance matrix estimation. *WIREs Computational Statistics*, 12(2):e1485, 2020.
- [6] Olivier Ledoit and Michael Wolf. Numerical implementation of the quest function, 2016.
- [7] Andrey A Popov, Adrian Sandu, Elias D. Nino-Ruiz, and Geir Evensen. A stochastic covariance shrinkage approach in ensemble transform kalman filtering, 2022.
- [8] Nicolasa Honnorat, Ehsan Adeli, Qingyu Zhao, Adolf Pfefferbaum, Edith V. Sullivan, and Kilian Pohl. Covariance shrinkage for dynamic functional connectivity. In Markus D. Schirmer, Archana Venkataraman, Islem Rekik, Minjeong Kim, and Ai Wern Chung, editors, *Connectomics in NeuroImaging*, pages 32–41, Cham, 2019. Springer International Publishing.
- [9] Ansgar Steland. Shrinkage for covariance estimation: Asymptotics, confidence intervals, bounds and applications in sensor monitoring and finance, 2018.
- [10] Charles Stein. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. pages 197–206, 1956.
- [11] Olivier Ledoit and Michael Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411, 2004.
- [12] Yilun Chen, Ami Wiesel, Yonina C. Eldar, and Alfred O. Hero. Shrinkage algorithms for mmse covariance estimation. *IEEE Transactions on Signal Processing*, 58(10):5016–5029, oct 2010.

- [13] Jianqing Fan, Yuan Liao, and Martina Mincheva. High-dimensional covariance matrix estimation in approximate factor models. *The Annals of Statistics*, 39(6):3320–3356, 2011.
- [14] Yuki Ikeda and Tatsuya Kubokawa. Linear shrinkage estimation of large covariance matrices using factor models. *Journal of Multivariate Analysis*, 152:61–81, 2016.
- [15] Zohar Halbe, Maria Bortman, and Mayer Aladjem. Regularized mixture density estimation with an analytical setting of shrinkage intensities. *Neural Networks and Learning Systems, IEEE Transactions on*, 24:460–470, 03 2013.
- [16] Arnaud Breloy, Esa Ollila, and Frédéric Pascal. Spectral shrinkage of tyler’s m -estimator of covariance matrix. In *2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 535–538, 2019.
- [17] Karina Ashurbekova, Antoine Usseglio-Carleve, Florence Forbes, and Sophie Achard. Optimal shrinkage for robust covariance matrix estimators in a small sample size setting. working paper or preprint, Mar 2021.
- [18] Benoit Oriol and Alexandre Miot. Ledoit-Wolf linear shrinkage with unknown mean, 2023.
- [19] V A Marčenko and L A Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457, apr 1967.
- [20] J. W. Silverstein. Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *Journal of Multivariate Analysis*, 55(2):331–339, 1995.
- [21] J.W. Silverstein and Z.D. Bai. On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *Journal of Multivariate Analysis*, 54(2):175–192, 1995.
- [22] Jack W. Silverstein and Sang Il Choi. Analysis of the limiting spectral distribution of large dimensional random matrices. *Journal of Multivariate Analysis*, 54:295–309, 1995.
- [23] Z. D. Bai and Jack W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *The Annals of Probability*, 26(1):316 – 345, 1998.
- [24] Olivier Ledoit and Sandrine Péché. Eigenvectors of some large sample covariance matrix ensembles, 2009.
- [25] Olivier Ledoit and Michael Wolf. Quadratic shrinkage for large covariance matrices. *SSRN Electronic Journal*, 28, 11 2019.
- [26] Olivier Ledoit and Michael Wolf. Analytical nonlinear shrinkage of large-dimensional covariance matrices. *Annals of Statistics*, 48:3043–3065, 10 2020.
- [27] Nicolas Honnorat and Mohamad Habes. Covariance shrinkage can assess and improve functional connectomes. *NeuroImage*, 256:119229, 2022.
- [28] Christian Bongiorno and Damien Challet. Covariance matrix filtering and portfolio optimisation: the average oracle vs non-linear shrinkage and all the variants of dcc-nls, 2023.
- [29] Szilard Pafka, Marc Potters, and Imre Kondor. Exponential weighting and random-matrix-theory-based filtering of financial covariance matrices for portfolio optimization, 2004.
- [30] J Daly, M Crane, and H J Ruskin. Random matrix theory filters and currency portfolio optimisation. *Journal of Physics: Conference Series*, 221(1):012003, apr 2010.
- [31] Jens Svensson. The asymptotic spectrum of the ewma covariance estimator. *Physica A: Statistical Mechanics and its Applications*, 385(2):621–630, 2007.
- [32] Vincent Tan and Stefan Zohren. Estimation of large financial covariances: A cross-validation approach, 2023.
- [33] Michael Fleermann and Werner Kirsch. Proof methods in random matrix theory, 2022.

- [34] Eugene A. Feinberg, Pavlo O. Kasyanov, and Yan Liang. Fatou's lemma in its classic form and lebesgue's convergence theorems for varying measures with applications to mdps, 2019.
- [35] Y.Q. Yin. Limiting spectral distribution for a class of random matrices. *Journal of Multivariate Analysis*, 20(1):50–68, 1986.
- [36] Jamal Najim. *Introduction à la théorie des grandes matrices aléatoires*, volume 33. 2016.

7 Appendix A - Experiment with non-diagonal W_n

We tried to rotate W_n under Student noise in order to make it non-diagonal, and check if the conclusions of Theorem 2 hold in this situation. For that matter, we compare two situations:

- in figure 8 (left), we examine the histograms of the sample spectrum for a diagonal weight matrix $W = D, T = Q' \Lambda Q'^*$.
- in figure 8 (right), we examine the histograms of the sample spectrum for a non-diagonal weight matrix $W = QDQ^*, T = Q' \Lambda Q'^*$.

where

- Λ is a diagonal matrix with a spectrum following $H = 0.2 \times 1_{[1, \infty[} + 0.4 \times 1_{[3, \infty[} + 0.4 \times 1_{[10, \infty[}$,
- D is a diagonal weight matrix where weights are sampled in an exponentially-weighted distribution of parameter $\alpha = 1$,
- Q and Q' are independent sampled unitary matrix: we sampled a noise matrix A of size (N, N) , with iid normal entries, and taking the QR decomposition, we extracted a unitary matrix Q .

The underlying noise matrix Z follows a Student distribution with $\nu \in \{4, 3, 2.5\}$ degrees of freedom.

The comparison does not clearly show any particular difference. That experiment, and the fact that the population covariance Σ does not need to be diagonal, suggests that an extension of Theorem 2 where W_n is not diagonal could be possible. However, this assumption is necessary in our proof, and further work need to be done to take this point into account.

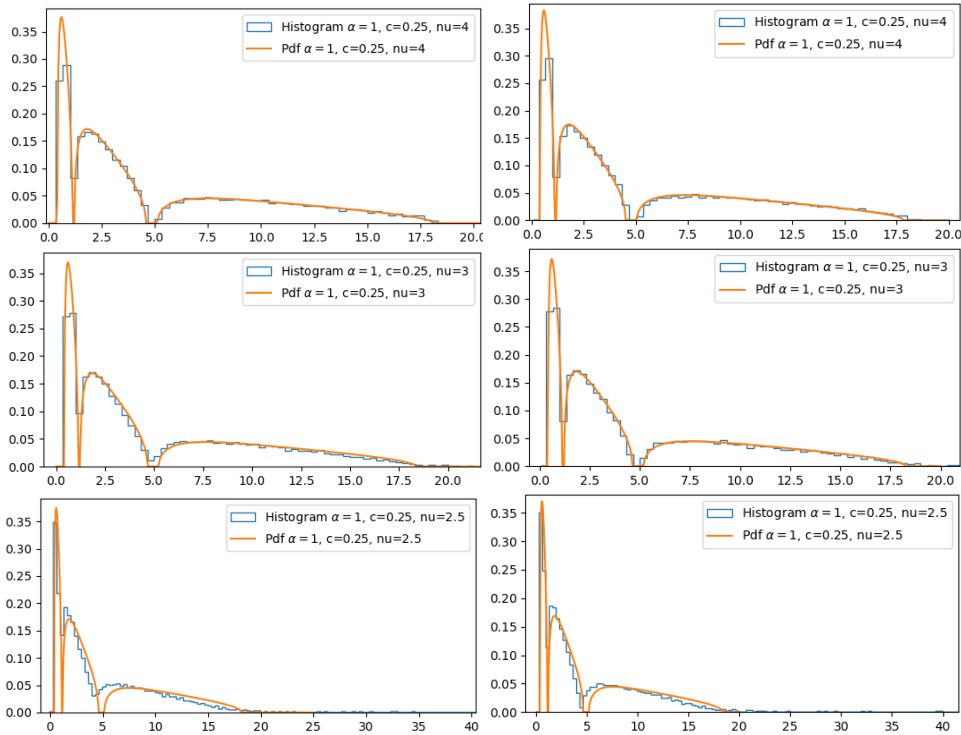


Figure 8: Histogram for W_n diagonal (left)/non-diagonal (right), from a t-distribution with $\nu = 4, \nu = 3$ and $\nu = 2.5$ from top to bottom, and asymptotic theoretical density, for $n = 2000, c = 0.25, H = 0.2 \times 1_{[1, \infty[} + 0.4 \times 1_{[3, \infty[} + 0.4 \times 1_{[10, \infty[}$ and exponentially-weighted distribution of parameter $\alpha = 1$.

8 Appendix B - Proof of Theorem 2

8.1 Truncation and centralization

In this first part of the proof, we use a similar approach as [21], aiming at truncating and centering Z , T and W while conserving the same asymptotic spectrum for F_n .

8.1.1 Truncation of Z by \sqrt{N}

Let $\hat{Z}_{ij} = Z_{ij}1_{|Z_{ij}| < \sqrt{N}}$ and $\hat{B}_n = \frac{1}{N}W^{1/2}\hat{Z}^*T\hat{Z}W^{1/2}$. We use Lemmas 2.5.a and 2.1.d [21], and with $\|\cdot\|$ denoting the ∞ -norm on bounded real functions, we have:

$$\begin{aligned} \|F^{B_n} - F^{\hat{B}_n}\| &\leq \frac{2}{N} \text{rank} \left(W^{1/2}Z - W^{1/2}\hat{Z} \right) \\ &\leq \frac{2}{N} \text{rank} \left(Z - \hat{Z} \right) \\ \|F^{B_n} - F^{\hat{B}_n}\| &\leq \frac{2}{N} \underbrace{\left\{ i \in \llbracket 1, n \rrbracket, j \in \llbracket 1, N \rrbracket \mid Z_{ij} - \hat{Z}_{ij} \neq 0 \right\}}_{\xi_n} \end{aligned} \quad (11)$$

At this point, we follow the idea of the proof p.58-59 of [35]. We have that:

$$\xi_n = \sum_{i,j} Y_{i,j} \text{ with } Y_{i,j} = 1_{|Z_{ij}| \geq \sqrt{N}}. \quad (12)$$

Remark that the Y_{ij} are i.i.d Bernoulli $\mathcal{B}(\eta)$ with $\eta = \mathbb{P} \left(|Z_{11}| \geq \sqrt{N} \right)$. As $\mathbb{E}[|Z_{11}|^2] = 1$, $\eta = o\left(\frac{1}{n}\right)$.

Let $\epsilon > 0$. For n large enough, we have:

$$\mathbb{P} \left(\frac{2}{N} \xi_n \geq \epsilon \right) = \mathbb{P} \left(\frac{1}{nN} \sum_{ij} Y_{ij} \geq \frac{\epsilon}{2n} \right) \leq \mathbb{P} \left(\frac{1}{nN} \sum_{ij} Y_{ij} \geq \eta + \frac{\epsilon}{4n} \right). \quad (13)$$

As $\mathbb{E}[Y_{ij}] = \eta$, we have by Chernoff-Hoeffding's theorem, for n large enough:

$$\begin{aligned} \mathbb{P} \left(\frac{2}{N} \xi_n \geq \epsilon \right) &\leq \exp \left[nN \left(\left(\eta + \frac{\epsilon}{4n} \right) \ln \frac{\eta}{\eta + \frac{\epsilon}{4n}} + \left(1 - \eta - \frac{\epsilon}{4n} \right) \ln \frac{1 - \eta}{1 - \eta - \frac{\epsilon}{4n}} \right) \right] \\ &\leq \exp \left[nN \left(-\frac{\left(\frac{\epsilon}{4n}\right)^2}{2\left(\eta + \frac{\epsilon}{4n}\right)} - \frac{\left(\frac{\epsilon}{4n}\right)^3}{3\left(\eta + \frac{\epsilon}{4n}\right)^2} - \frac{\left(\frac{\epsilon}{4n}\right)^2}{2\left(1 - \eta - \frac{\epsilon}{4n}\right)} + \frac{\left(\frac{\epsilon}{4n}\right)^3}{3\left(1 - \eta - \frac{\epsilon}{4n}\right)^2} \right) \right] \\ &= \exp \left[-\frac{\epsilon^2 N}{32n} \times \frac{1}{\left(\eta + \frac{\epsilon}{4n}\right)\left(1 - \eta - \frac{\epsilon}{4n}\right)} - \frac{\epsilon^3 N}{3 \times 4^3 n^2} \times \frac{1 - 2\left(\eta + \frac{\epsilon}{4n}\right)}{\left(\eta + \frac{\epsilon}{4n}\right)^2\left(1 - \eta - \frac{\epsilon}{4n}\right)^2} \right] \\ \mathbb{P} \left(\frac{2}{N} \xi_n \geq \epsilon \right) &\leq \exp \left[-\epsilon N \times \left(\frac{1}{16} + \frac{2}{3 \times 4^3} \right) \right]. \end{aligned} \quad (14)$$

Consequently,

$$\|F^{B_n} - F^{\hat{B}_n}\| \xrightarrow[n \rightarrow \infty]{} 0 \text{ almost surely.} \quad (15)$$

8.1.2 Centralization of \hat{Z}

Let $\tilde{Z}_{ij} = \hat{Z}_{ij} - \mathbb{E}[\hat{Z}_{ij}]$ and $\tilde{B}_n = \frac{1}{N}W^{1/2}\tilde{Z}^*T\tilde{Z}W^{1/2}$. The \hat{Z}_{ij} are i.i.d, so $\forall i, j, \mathbb{E}[\hat{Z}_{ij}] = \mathbb{E}[\hat{Z}_{11}]$. Consequently,

$$\text{rank} \left(W^{1/2}\tilde{Z} - W^{1/2}\hat{Z} \right) = \text{rank} \left(W^{1/2}\mathbb{E}[\hat{Z}] \right) \leq 1. \quad (16)$$

Using Lemma 2.5.a [21], we have:

$$\|F^{\tilde{B}_n} - F^{\hat{B}_n}\| \leq \frac{2}{N} \text{rank} \left(W^{1/2}\tilde{Z} - W^{1/2}\hat{Z} \right) \leq \frac{2}{N} \xrightarrow[n \rightarrow \infty]{} 0. \quad (17)$$

8.1.3 Truncation of T and W by α_n and β_n respectively

We denote U a matrix of eigenvector of T with associated eigenvalues (τ_1, \dots, τ_n) , so that $T = U \text{Diag}((\tau_i)) U^*$. For $\alpha > 0$, we define:

$$T_\alpha = U \text{Diag}((\tau_i 1_{|\tau_i| \leq \alpha})) U^*. \quad (18)$$

If α and $-\alpha$ are continuity points of H , using Lemma 2.5.b [21], we have that for any $n \times N$ Hermitian matrix Q :

$$\|F^{Q^* T Q} - F^{Q^* T_\alpha Q}\| \leq \frac{1}{N} \text{rank}(T - T_\alpha) = \frac{1}{N} \sum_{i=1}^n 1_{|\tau_i| > \alpha}. \quad (19)$$

Using Assumptions (b) and (c), we have then:

$$\frac{1}{N} \text{rank}(T - T_\alpha) \xrightarrow[n \rightarrow \infty]{} cH([-\alpha, \alpha]^c) \text{ almost surely.} \quad (20)$$

Consequently,

$$\alpha := \alpha_n \xrightarrow[n \rightarrow \infty]{} +\infty \implies \|F^{Q^* T Q} - F^{Q^* T_\alpha Q}\| \xrightarrow[n \rightarrow \infty]{} 0 \text{ almost surely.} \quad (21)$$

Similarly, for $\beta > 0$, we define:

$$W_\beta = \text{Diag}((W_{ii} 1_{|W_{ii}| \leq \beta})). \quad (22)$$

Using Lemma 2.5.b [21] and Assumptions (b) and (c), we have also for any $n \times N$ Hermitian matrix Q :

$$\|F^{W^{1/2} Q^* T_\alpha Q W^{1/2}} - F^{W_\beta^{1/2} Q^* T_\alpha Q W_\beta^{1/2}}\| \leq \frac{1}{n} \text{rank}(W - W_\beta) \xrightarrow[n \rightarrow \infty]{} \frac{1}{c} D([-\beta, \beta]^c) \text{ almost surely.} \quad (23)$$

Consequently,

$$\beta := \beta_n \xrightarrow[n \rightarrow \infty]{} +\infty \implies \|F^{W^{1/2} Q^* T_\alpha Q W^{1/2}} - F^{W_\beta^{1/2} Q^* T_\alpha Q W_\beta^{1/2}}\| \xrightarrow[n \rightarrow \infty]{} 0 \text{ almost surely.} \quad (24)$$

In the following, we choose $\alpha_n \uparrow +\infty$ and $\beta_n \uparrow +\infty$ so that:

- $(\alpha_n \beta_n)^4 (\mathbb{E}[|Z_{11}|^2 1_{|Z_{11}| \geq \ln N}] + \frac{1}{N}) \xrightarrow[n \rightarrow \infty]{} 0$,
- $\sum_{n=1}^{\infty} \frac{(\beta_n \alpha_n)^8}{N^2} (\mathbb{E}[|Z_{11}|^4 1_{|Z_{11}| \geq \sqrt{N}}] + 1) < +\infty$.

We denote $\tilde{E}_{\alpha, \beta} = \frac{1}{N} W_\beta^{1/2} \tilde{Z}^* T_\alpha \tilde{Z} W_\beta^{1/2}$. With α, β chosen following the rules above, we have:

$$\|F^{\tilde{E}_{\alpha, \beta}} - F^{\tilde{E}_n}\| \xrightarrow[n \rightarrow \infty]{} 0 \text{ almost surely.} \quad (25)$$

8.1.4 Truncation of \tilde{Z} by $\ln N$

Let $\bar{Z}_{ij} = \tilde{Z}_{ij} 1_{|Z_{ij}| < \ln N} - \mathbb{E}[\tilde{Z}_{ij} 1_{|Z_{ij}| < \ln N}]$ and $\bar{\bar{Z}}_{ij} = \tilde{Z}_{ij} - \bar{Z}_{ij}$. Remark that $\mathbb{E}[\bar{\bar{Z}}] = \mathbb{E}[\bar{Z}] = 0$. Moreover, since $|\tilde{Z}_{ij}| \leq \ln N + \mathbb{E}[|Z_{11}|]$, we have for n sufficiently large and some $a > 2$:

$$\frac{|\bar{\bar{Z}}_{ij}|}{\sqrt{\mathbb{E}[|\bar{\bar{Z}}_{11}|^2]}} \leq \frac{\ln N + \mathbb{E}[|Z_{11}|]}{\sqrt{\mathbb{E}[|Z_{11}|^2]}} \leq a \ln n := \log n. \quad (26)$$

We denote $\log n$ the logarithm of n in base $\exp(1/a)$.

In this part, we use a metric D on $\mathcal{M}(\mathbb{R})$, the set of all subprobability distribution functions on \mathbb{R} , defined in [21] p.191.

Definition 2 (D -metric, [21]). *Let $\{f_i\}$ be an enumeration of all continuous functions that take a constant $1/m$ value (m a positive integer) on $[a, b]$, where a, b are rational, $0 \text{ on }]-\infty, a - 1/m] \cup [b + 1/m, +\infty[$, and linear on each $[a - 1/m, a]$, $[b, b + 1/m]$. For $F_1, F_2 \in \mathcal{M}(\mathbb{R})$, we define:*

$$D(F_1, F_2) := \sum_{i=1}^{\infty} \left| \int f_i dF_1 - \int f_i dF_2 \right| 2^{-i}. \quad (27)$$

D is a metric on $\mathcal{M}(\mathbb{R})$ inducing the topology of vague convergence.

We denote $\bar{B}_{\alpha,\beta} = \frac{1}{N} W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2}$. Using Lemma 2.1.c and Equation (2.4) [21], we have:

$$\begin{aligned}
D^2 \left(F^{\bar{B}_{\alpha,\beta}}, F^{\bar{B}_{\alpha,\beta}} \right) &\leq \frac{1}{N} \operatorname{tr} \left[\left(W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} - W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} \right)^2 \right] \\
&= \frac{1}{N} \operatorname{tr} \left[\left(W_\beta^{1/2} (\bar{Z} + \bar{Z})^* T_\alpha (\bar{Z} + \bar{Z}) W_\beta^{1/2} - W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} \right)^2 \right] \\
D^2 \left(F^{\bar{B}_{\alpha,\beta}}, F^{\bar{B}_{\alpha,\beta}} \right) &= \frac{1}{N} \left(\operatorname{tr} \left[\left(W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} \right)^2 \right] + \operatorname{tr} \left[\left(W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} + W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} \right)^2 \right] \right. \\
&\quad \left. + 2 \operatorname{tr} \left[\left(W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} \right) \left(W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} + W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} \right) \right] \right). \tag{28}
\end{aligned}$$

Using Cauchy-Schwarz inequality, we have:

$$\begin{aligned}
D^2 \left(F^{\bar{B}_{\alpha,\beta}}, F^{\bar{B}_{\alpha,\beta}} \right) &= \frac{1}{N} \left(\operatorname{tr} \left[\left(W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} \right)^2 \right] + 4 \operatorname{tr} \left[W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} \right] \right. \\
&\quad \left. + 4 \sqrt{\operatorname{tr} \left[\left(W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} \right)^2 \right] \operatorname{tr} \left[W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} \right]} \right). \tag{29}
\end{aligned}$$

We show, through Von Neumann's trace inequality, that:

$$\begin{aligned}
\operatorname{tr} \left[\left(W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} \right)^2 \right] &\leq \alpha^2 \beta^2 \operatorname{tr} \left[\left(\bar{Z}^* \bar{Z} \right)^2 \right], \tag{30} \\
\text{and } \operatorname{tr} \left[W_\beta^{1/2} \bar{Z}^* T_\alpha \bar{Z} W_\beta \bar{Z}^* T_\alpha \bar{Z} W_\beta^{1/2} \right] &\leq \alpha^2 \beta^2 \sqrt{\operatorname{tr} \left[\left(\bar{Z}^* \bar{Z} \right)^2 \right] \operatorname{tr} \left[\left(\bar{Z}^* \bar{Z} \right)^2 \right]}.
\end{aligned}$$

And, from [21] p.185 and Assumption (b), we have:

$$\frac{1}{N} \operatorname{tr} \left[\left(\bar{Z}^* \bar{Z} \right)^2 \right] \xrightarrow{n \rightarrow \infty} c(1+c) \text{ almost surely.} \tag{31}$$

We have also from [21] p.185, with K and K' constant independent of n ,

$$\begin{aligned}
\mathbb{E} \left[\frac{1}{N} \operatorname{tr} \left[\left(\bar{Z}^* \bar{Z} \right)^2 \right] \right] &\leq K' \left(\mathbb{E} \left[|Z_{11}|^2 1_{|Z_{11}| \geq \ln N} \right] + \frac{1}{N} \right), \tag{32} \\
\mathbb{V} \left[\frac{1}{N} \operatorname{tr} \left[\left(\bar{Z}^* \bar{Z} \right)^2 \right] \right] &\leq \frac{K}{N^2} \left(\mathbb{E} \left[|Z_{11}|^4 1_{|Z_{11}| \geq \sqrt{N}} \right] + 1 \right).
\end{aligned}$$

So, $\mathbb{E} \left[\frac{\alpha^4 \beta^4}{N} \operatorname{tr} \left[\left(\bar{Z}^* \bar{Z} \right)^2 \right] \right] \xrightarrow{n \rightarrow \infty} 0$ and $\mathbb{V} \left[\frac{\alpha^4 \beta^4}{N} \operatorname{tr} \left[\left(\bar{Z}^* \bar{Z} \right)^2 \right] \right]$ is summable, so:

$$\frac{\alpha^4 \beta^4}{N} \operatorname{tr} \left[\left(\bar{Z}^* \bar{Z} \right)^2 \right] \xrightarrow{n \rightarrow \infty} 0 \text{ almost surely.} \tag{33}$$

Backing up, we have then:

$$D^2 \left(F^{\bar{B}_{\alpha,\beta}}, F^{\bar{B}_{\alpha,\beta}} \right) \xrightarrow{n \rightarrow \infty} 0 \text{ almost surely.} \tag{34}$$

So, $F^{\bar{B}_{\alpha,\beta}} - F^{\bar{B}_{\alpha,\beta}} \xrightarrow{v} 0$, \xrightarrow{v} denotes the vague convergence.

8.1.5 Conclusion of the truncations and centralizations

The conclusion of that first part of the proof is that it is sufficient to show that $F^{\bar{B}_{\alpha,\beta}} \xrightarrow{v} \underline{F}$ for some $\underline{F} \in \mathcal{M}(\mathbb{R})$ (we will in fact prove the weak convergence) in order to prove that $F^{B_n} \xrightarrow{v} \underline{F}$, which is equivalent to show that $F^{B_n} \xrightarrow{v} F := \frac{1}{c} \underline{F} + \frac{1-c}{c} 1_{[0,\infty[}$.

Moreover, as by definition $F^{B_n}(\mathbb{R}) = 1$, if $F(\mathbb{R}) = 1$, it is equivalent to prove that $F^{B_n} \xrightarrow[n \rightarrow \infty]{=} F$.

In the following, we focus on the truncated and centralized variables. In fact, we swap:

- Z_{ij} by $\bar{Z}_{ij}/\sqrt{\mathbb{E}[|\bar{Z}_{ij}|^2]}$,
- T by $\mathbb{E}[|\bar{Z}_{11}|^2]T_\alpha$,
- W by W_β .

Remark that we can impose $\alpha \leq \log n$ and $\beta \leq \log n$ without compromising any of the properties on α and β defined in Section 6.1.3.

In addition to (a)-(e), we can now use the following Assumptions.

Assumption 2. *Assume that:*

- (f) $\forall i, j, |Z_{ij}| \leq \log n$,
- (g) $\|T\| \leq \log n$,
- (h) $\|W\| \leq \log n$.

From now, $B_n = \frac{1}{N}T^{1/2}Z^*WZT^{1/2}$ and $\underline{B}_n = \frac{1}{N}W^{1/2}ZTZ^*W^{1/2}$ use the truncated variables Z , T and W , and so are F_n , \underline{F}_n , m_n and \underline{m}_n .

The following aims at proving, with those truncated variables, that for some p.d.f F , $F^{B_n} \xrightarrow{v} F$ a.s., which proves the theorem. For that, we prove that for some p.d.f F , $F^{B_n} \xrightarrow[n \rightarrow \infty]{\Rightarrow} F$ a.s. It is done using the Cauchy-Stieltjes transform: we prove that for all $z \in \mathbb{C}_+$, a.s. $m_n(z) \xrightarrow[n \rightarrow \infty]{\rightarrow} m(z)$ where m is the Cauchy-Stieltjes transform of a p.d.f.

We introduce new objects of interest in the analysis, namely q_j , r_j , and $B_{(j)}$.

Notation 5. *For $j \in \llbracket 1, N \rrbracket$, we denote:*

- $q_j = \frac{1}{\sqrt{n}}Z_{\cdot j}$,
- $r_j = \frac{1}{\sqrt{N}}T^{1/2}Z_{\cdot j}W_{jj}^{1/2}$,
- $B_{(j)} = B_n - r_j r_j^*$,

8.2 Concentration

8.2.1 Preliminary derivations

Let's introduce a technical result.

Lemma 1 (Eq (2.1) [22]). *For B a (n, n) matrix, $q \in \mathbb{C}^n$ for which B and $B + qq^*$ is invertible, we have:*

$$q^*(B + qq^*)^{-1} = \frac{1}{1 + q^*B^{-1}q}q^*B^{-1}. \quad (35)$$

We have:

$$\begin{aligned} (B_n - zI) + zI &= \sum_{j=1}^N r_j r_j^* \\ I + z(B_n - zI)^{-1} &= \sum_{j=1}^N r_j r_j^* (B_n - zI)^{-1} \\ I + z(B_n - zI)^{-1} &= \sum_{j=1}^N \frac{1}{1 + r_j^*(B_{(j)} - zI)^{-1}r_j} r_j r_j^* (B_{(j)} - zI)^{-1} \\ c_n + zc_n m_n(z) &= 1 - \frac{1}{N} \sum_{j=1}^N \frac{1}{1 + r_j^*(B_{(j)} - zI)^{-1}r_j} \\ \underline{m}_n(z) &= -\frac{1}{N} \sum_{j=1}^N \frac{1}{z(1 + r_j^*(B_{(j)} - zI)^{-1}r_j)}. \end{aligned} \quad (36)$$

Here is the crucial difference between the proof in [20] in the evenly weighted case and the weighted case. We denote:

$$\alpha_n = -\frac{1}{z\underline{m}_n(z)} \frac{1}{N} \sum_{j=1}^N W_{jj} (r_j^*(B_n - zI)^{-1} r_j - 1), \quad (37)$$

while in [22], $\alpha_n = 1$ is used.

We have equivalently:

$$\begin{aligned} \alpha_n &= -\frac{1}{z\underline{m}_n(z)} \frac{1}{N} \sum_{j=1}^N \frac{W_{jj}}{1 + r_j^*(B_{(j)} - zI)^{-1} r_j} \\ \text{and } \alpha_n &= \frac{\frac{1}{N} \sum_{j=1}^N \frac{W_{jj}}{1 + r_j^*(B_{(j)} - zI)^{-1} r_j}}{\frac{1}{N} \sum_{j=1}^N \frac{1}{1 + r_j^*(B_{(j)} - zI)^{-1} r_j}}, \end{aligned} \quad (38)$$

Then:

$$\begin{aligned} &(-z\alpha_n \underline{m}_n(z) T_n - zI)^{-1} - (B_n - zI)^{-1} = \\ &(-z\alpha_n \underline{m}_n(z) T_n - zI)^{-1} \left(z\alpha_n \underline{m}_n(z) T_n + \sum_{j=1}^N r_j r_j^* \right) (B_n - zI)^{-1} = \\ &(-z\alpha_n \underline{m}_n(z) T_n - zI)^{-1} \sum_{j=1}^N \frac{r_j r_j^* (B_{(j)} - zI)^{-1} - \frac{1}{N} \alpha_n T_n (B_n - zI)^{-1}}{1 + r_j^*(B_{(j)} - zI)^{-1} r_j}. \end{aligned} \quad (39)$$

Applying the trace and dividing by n , we obtain:

$$\begin{aligned} &\frac{1}{n} \operatorname{tr} \left((-z\alpha_n \underline{m}_n(z) T_n - zI)^{-1} \right) - m_n(z) = \\ &-\frac{1}{N} \sum_{j=1}^N \frac{1}{z(1 + r_j^*(B_{(j)} - zI)^{-1} r_j)} \left(W_{jj} q_j^* T_n^{1/2} (B_{(j)} - zI)^{-1} (\alpha_n \underline{m}_n(z) T_n + I)^{-1} T_n^{1/2} q_j \right. \\ &\quad \left. - \frac{1}{n} \operatorname{tr} \left((\alpha_n \underline{m}_n(z) T_n + I)^{-1} \alpha_n T_n (B_n - zI)^{-1} \right) \right) \end{aligned} \quad (40)$$

So,

$$\begin{aligned} &\frac{1}{n} \operatorname{tr} \left((-z\alpha_n \underline{m}_n(z) T_n - zI)^{-1} \right) - m_n(z) = \\ &-\frac{1}{zN} \sum_{j=1}^N q_j^* \left(\frac{W_{jj} T_n^{1/2} (B_{(j)} - zI)^{-1} (\alpha_n \underline{m}_n(z) T_n + I)^{-1} T_n^{1/2}}{1 + r_j^*(B_{(j)} - zI)^{-1} r_j} \right) q_j \\ &+ \frac{1}{zNn} \operatorname{tr} \left((\alpha_n \underline{m}_n(z) T_n + I)^{-1} \sum_{j=1}^N \frac{\alpha_n T_n}{1 + r_j^*(B_{(j)} - zI)^{-1} r_j} (B_n - zI)^{-1} \right) = \\ &-\frac{1}{N} \sum_{j=1}^N \frac{1}{z(1 + r_j^*(B_{(j)} - zI)^{-1} r_j)} d_j, \end{aligned} \quad (41)$$

where:

$$\begin{aligned} d_j &= W_{jj} q_j^* T_n^{1/2} (B_{(j)} - zI)^{-1} (\alpha_n \underline{m}_n(z) T_n + I)^{-1} T_n^{1/2} q_j \\ &\quad - \frac{1}{n} \operatorname{tr} \left(W_{jj} (\alpha_n \underline{m}_n(z) T_n + I)^{-1} T_n (B_n - zI)^{-1} \right). \end{aligned} \quad (42)$$

The strategy of the proof is the following:

- prove that $\max_{j \leq N} d_j \rightarrow 0$ a.s., and that $\frac{1}{n} \operatorname{tr} \left((-z\alpha_n \underline{m}_n(z) T_n - zI)^{-1} \right) - m_n(z) \rightarrow 0$ a.s.,

- show that a.s. it exists $\tilde{m}(z) \in \mathbb{C} \setminus \mathbb{C}_+$ and a subsequence n_i so that $\tilde{m}_{n_i}(z) := -z\alpha_{n_i}\underline{m}_{n_i}(z) \rightarrow \tilde{m}(z)$ a.s.,
- prove that $\tilde{m}(z)$ is the unique solution in \mathbb{C}_- of a functional equation, and deduce $\tilde{m}_n(z) \rightarrow \tilde{m}(z)$ a.s.,
- deduce that a.s. it exists $m(z) \in \mathbb{C}_+$, uniquely defined in function of $\tilde{m}(z)$ so that $m_n(z) \rightarrow m(z)$ a.s.,
- similarly deduce that a.s. it exists $\Theta^{(1)}(z) \in \mathbb{C}_+$, uniquely defined in function of $\tilde{m}(z)$ so that $\Theta_n^{(1)}(z) \rightarrow \Theta^{(1)}(z)$ a.s.,
- conclude proving that m is the Cauchy-Stieltjes transform of a p.d.f.

8.2.2 Decomposition of d_j

Much of the truth of this proof relies upon the following lemma from [21].

Lemma 2 (Lemma 3.1 [21]). *Let C a Hermitian $n \times n$ matrix so that $\|C\| \leq 1$, and $Y = (Z_1, \dots, Z_n)^T$, $Z_i \in \mathbb{C}$ where the Z_i 's are i.i.d., $\mathbb{E}[Z_1] = 0$, $\mathbb{E}[|Z_1|^2] = 1$ and $|Z_1| \leq \log n$. Then,*

$$\mathbb{E} \left[|Y^* C Y - \text{tr } C|^6 \right] \leq K n^3 \log(n)^{12}, \quad (43)$$

where the constant K does not depend on n , C , nor the distribution of Z_1 .

In order to use this Lemma on d_j , we decompose it into negligible terms and a term of the form $q_j^* C q_j - \frac{1}{n} \text{tr}(C)$ where C is independent of q_j .

For that, we denote:

$$\begin{aligned} \tilde{m}_n(z) &:= -z\alpha_n \underline{m}_n(z) \\ \text{and } \tilde{m}_{(j)}(z) &:= \frac{1}{N} \sum_{i \neq j} W_{ii} (r_i^*(B_{(j)} - zI)^{-1} r_i - 1). \end{aligned} \quad (44)$$

We have the following decomposition of d_j :

$$d_j = d_j^{(1)} + d_j^{(2)} + d_j^{(3)} + d_j^{(4)}, \quad (45)$$

with:

$$\begin{aligned} d_j^{(1)} &= W_{jj} q_j^* T_n^{1/2} (B_{(j)} - zI)^{-1} \left[\left(-\frac{\tilde{m}_n(z)}{z} T_n + I \right)^{-1} - \left(-\frac{\tilde{m}_{(j)}(z)}{z} T_n + I \right)^{-1} \right] T_n^{1/2} q_j, \\ d_j^{(2)} &= W_{jj} q_j^* T_n^{1/2} (B_{(j)} - zI)^{-1} \left(-\frac{\tilde{m}_{(j)}(z)}{z} T_n + I \right)^{-1} T_n^{1/2} q_j \\ &\quad - \frac{W_{jj}}{n} \text{tr} \left(\left(-\frac{\tilde{m}_{(j)}(z)}{z} T_n + I \right)^{-1} T_n (B_{(j)} - zI)^{-1} \right), \\ d_j^{(3)} &= \frac{W_{jj}}{n} \text{tr} \left(\left[\left(-\frac{\tilde{m}_{(j)}(z)}{z} T_n + I \right)^{-1} - \left(-\frac{\tilde{m}_n(z)}{z} T_n + I \right)^{-1} \right] T_n (B_{(j)} - zI)^{-1} \right), \\ d_j^{(4)} &= \frac{W_{jj}}{n} \text{tr} \left(\left(-\frac{\tilde{m}_n(z)}{z} T_n + I \right)^{-1} T_n [(B_{(j)} - zI)^{-1} - (B_n - zI)^{-1}] \right). \end{aligned} \quad (46)$$

In order to prove that for each $k \in \llbracket 1, 4 \rrbracket$, $\max_{j \leq N} d_j^{(k)} \rightarrow 0$ a.s., we need some technical lemmas. They essentially provide the necessary inequalities to prove that $d_j^{(1)}$, $d_j^{(3)}$ and $d_j^{(4)}$ are indeed negligible, to finally use Lemma 2 on $d_j^{(2)}$ and prove that $\max_{j \leq N} |d_j| \rightarrow 0$ a.s.

8.2.3 Technical lemmas

Lemma 3. *We have the following inequalities, for $j \in \llbracket 1, N \rrbracket$:*

$$\begin{aligned} \|(B_n - zI)^{-1}\| &\leq \frac{1}{v}, \\ \|(B_{(j)} - zI)^{-1}\| &\leq \frac{1}{v}, \\ \frac{1}{|z(1 + r_j^*(B_{(j)} - zI)^{-1}r_j)|} &\leq \frac{1}{v}. \end{aligned} \quad (47)$$

Proof. Let $j \in \llbracket 1, N \rrbracket$. The two first inequalities comes from the fact that B_n and $B_{(j)}$ are Hermitian, so for any eigenvalue λ of $B_n - zI$ or $B_{(j)} - zI$, we have that $|\lambda| \geq |\operatorname{Im}[\lambda]| = v$.

For the third inequality, we remark that:

$$\begin{aligned} \operatorname{Im} r_j^*(B_{(j)}/z - I)^{-1}r_j &= \frac{1}{2i} r_j^* [(B_{(j)}/z - I)^{-1} - (B_{(j)}/z^* - I)^{-1}] r_j \\ &= \frac{v}{|z|^2} r_j^*(B_{(j)}/z - I)^{-1} B_{(j)} (B_{(j)}/z^* - I)^{-1} r_j \end{aligned} \quad (48)$$

$$\operatorname{Im} r_j^*(B_{(j)}/z - I)^{-1}r_j \geq 0.$$

So, we deduce that:

$$\frac{1}{|z(1 + r_j^*(B_{(j)} - zI)^{-1}r_j)|} \leq \frac{1}{v}. \quad (49)$$

□

Lemma 4. *We denote $\bar{W} = \frac{1}{N} \sum_{i=1}^N W_{ii}$. For $z = u + iv$, $v > 0$ and $j \in \llbracket 1, N \rrbracket$, we have for any nonnegative Hermitian matrix A :*

$$\left\| \left(-\frac{\tilde{m}_n(z)}{z} A + I \right)^{-1} \right\| \leq f(z, \|A\|), \text{ and } \left\| \left(-\frac{\tilde{m}_{(j)}(z)}{z} A + I \right)^{-1} \right\| \leq f(z, \|A\|), \quad (50)$$

where:

$$f(z, \|A\|) = \begin{cases} \max\left(2, \frac{4}{v} \bar{W} \|A\|\right), & \text{if } u = 0, \\ 16 \left(\frac{|z|^2}{4v^2|u|} \bar{W} \|A\| + 1 \right) \times \max\left(\frac{1}{3}, \frac{|u|}{v}\right), & \text{otherwise.} \end{cases} \quad (51)$$

Proof. Let $z = u + iv$, $v > 0$, $u \in \mathbb{R}$. For $j \in \llbracket 1, N \rrbracket$, we denote by $(u_{ij})_{i=1}^n$ a set of eigenvectors of $B_{(j)}$ with associated eigenvalues $(\lambda_{ij})_{i=1}^n$. We then derive the following formulation:

$$\begin{aligned} R := \operatorname{Re} \left[-\frac{\tilde{m}_n(z)}{z} \right] &= -\frac{1}{N} \sum_{j=1}^N \frac{W_{jj}}{\left| z + r_j^* \left(\frac{B_{(j)}}{z} - I \right)^{-1} r_j \right|^2} \left(u + \sum_{i=1}^n \frac{|r_j^* u_{ij}|^2}{|\lambda_{ij} - z|^2} (\lambda_{ij} u - |z|^2) \right), \\ I := \operatorname{Im} \left[-\frac{\tilde{m}_n(z)}{z} \right] &= \frac{1}{N} \sum_{j=1}^N \frac{W_{jj}}{\left| z + r_j^* \left(\frac{B_{(j)}}{z} - I \right)^{-1} r_j \right|^2} \left(v + \sum_{i=1}^n \frac{|r_j^* u_{ij}|^2}{|\lambda_{ij} - z|^2} \lambda_{ij} v \right) \geq 0. \end{aligned} \quad (52)$$

Using Cauchy-Schwarz inequality, we deduce:

$$\begin{aligned} \frac{1}{N} \sum_{j=1}^N \frac{W_{jj} |z|^2 \sum_{i=1}^n \frac{|r_j^* u_{ij}|^2}{|\lambda_{ij} - z|^2}}{\left| z + r_j^* \left(\frac{B_{(j)}}{z} - I \right)^{-1} r_j \right|^2} &\leq \sqrt{\frac{1}{N} \sum_{j=1}^N \frac{W_{jj} \left(\sum_{i=1}^n \frac{|r_j^* u_{ij}|^2}{|\lambda_{ij} - z|^2} \right)^2}{\left| 1 + r_j^* (B_{(j)} - zI)^{-1} r_j \right|^2}} \sqrt{\frac{1}{N} \sum_{j=1}^N \frac{W_{jj}}{\left| z + r_j^* \left(\frac{B_{(j)}}{z} - I \right)^{-1} r_j \right|^2}} \\ \frac{1}{N} \sum_{j=1}^N \frac{W_{jj} |z|^2 \sum_{i=1}^n \frac{|r_j^* u_{ij}|^2}{|\lambda_{ij} - z|^2}}{\left| z + r_j^* \left(\frac{B_{(j)}}{z} - I \right)^{-1} r_j \right|^2} &\leq \frac{|z| \sqrt{\bar{W}}}{v \sqrt{v}} \sqrt{I}. \end{aligned} \quad (53)$$

So, combining both previous points, we have:

$$|R| \leq \frac{|u|}{v} I + \frac{|z|\sqrt{\bar{W}}}{v\sqrt{|u|}} \sqrt{I}. \quad (54)$$

Suppose $u \neq 0$. We denote $K := \frac{|z|\sqrt{\bar{W}}}{2v\sqrt{|u|}}$. We have then:

$$\sqrt{\frac{v}{|u|}} \left(-K + \sqrt{K^2 + |R|} \right) \leq \sqrt{I}. \quad (55)$$

Now, let $x \geq 0$. Then:

$$\begin{aligned} \left| -\frac{\tilde{m}_n(z)}{z} x + 1 \right|^2 &= (Rx + 1)^2 + I^2 x^2 \\ \left| -\frac{\tilde{m}_n(z)}{z} x + 1 \right|^2 &\geq (-|R|x + 1)^2 + \frac{v^2}{|u|^2} \left(-K\sqrt{x} + \sqrt{K^2 x + |R|x} \right)^4. \end{aligned} \quad (56)$$

We denote: $t := \sqrt{K^2 x + |R|x} \in \mathbb{R}_+$. We have then:

$$\left| -\frac{\tilde{m}_n(z)}{z} x + 1 \right|^2 \geq (t^2 - K^2 x - 1)^2 + \frac{v^2}{|u|^2} (t - K\sqrt{x})^4. \quad (57)$$

We denote $a := K\sqrt{x}$, $b := \sqrt{K^2 x + 1}$ and we split the study of the right part of the previous equation between $[0, (a+b)/2]$ and $[(a+b)/2, +\infty[$. The lower bounds rely mainly on the fact that $b - a \geq \frac{1}{2b}$:

- Let $t \in [0, (a+b)/2]$. Then,

$$\begin{aligned} (t^2 - K^2 x - 1)^2 + \frac{v^2}{|u|^2} (t - K\sqrt{x})^4 &\geq \left(\left(\frac{a+b}{2} \right)^2 - b^2 \right)^2 \\ &= \frac{(a+3b)^2 (b-a)^2}{16} \\ &\geq \frac{9(b-a)^4}{16} \\ (t^2 - K^2 x - 1)^2 + \frac{v^2}{|u|^2} (t - K\sqrt{x})^4 &\geq \frac{9}{16^2 b^4}. \end{aligned} \quad (58)$$

- Let $t \in [(a+b)/2, +\infty[$. Then,

$$\begin{aligned} (t^2 - K^2 x - 1)^2 + \frac{v^2}{|u|^2} (t - K\sqrt{x})^4 &\geq \frac{v^2}{16|u|^2} (b-a)^4 \\ (t^2 - K^2 x - 1)^2 + \frac{v^2}{|u|^2} (t - K\sqrt{x})^4 &\geq \frac{v^2}{16^2 |u|^2 b^4}. \end{aligned} \quad (59)$$

Backing up, we have that:

$$\left| -\frac{\tilde{m}_n(z)}{z} x + 1 \right|^2 \geq \frac{1}{16^2 (K^2 x + 1)^2} \times \min \left(\frac{v^2}{|u|^2}, 9 \right). \quad (60)$$

It finally leads to:

$$\left\| \left(-\frac{\tilde{m}_n(z)}{z} A + I \right)^{-1} \right\| \leq 16 \left(\frac{|z|^2}{4v^2|u|} \bar{W} \|A\| + 1 \right) \times \max \left(\frac{1}{3}, \frac{|u|}{v} \right), \quad (61)$$

which proves the first inequality when $u \neq 0$.

Now suppose $u = 0$. From Equation (62), we have:

$$|R| \leq \sqrt{\frac{\bar{W}}{v}} \sqrt{I}. \quad (62)$$

If $\bar{W} = 0$, then:

$$\left| -\frac{\tilde{m}_n(z)}{z}x + 1 \right|^2 = 1 + x^2 I^2 \geq 1 \geq \frac{1}{4}. \quad (63)$$

Otherwise, with $y = Rx$ and $a = \frac{v^2}{W^4 x^2}$:

$$\left| -\frac{\tilde{m}_n(z)}{z}x + 1 \right|^2 \geq (Rx + 1)^2 + x^2 \frac{R^4 v^2}{W^4} = ay^4 + (y + 1)^2. \quad (64)$$

Splitting the study between $]-\infty, 1/2]$ and $[1/2, +\infty[$, we have that $ay^4 + (y + 1)^2 \geq \min\left(\frac{a}{16}, \frac{1}{4}\right)$. So, we deduce the first inequality when $u = 0$:

$$\left\| \left(-\frac{\tilde{m}_n(z)}{z}A + I \right)^{-1} \right\| \leq \max\left(2, \frac{4\bar{W}\|A\|}{v}\right). \quad (65)$$

Let $j \in \llbracket 1, N \rrbracket$. Using the same method for $\left\| \left(-\frac{\tilde{m}_{(j)}(z)}{z}A + I \right)^{-1} \right\|$, we have with $\bar{W}_{(j)} := \frac{1}{N} \sum_{i \neq j} W_{ii}$:

$$\left| \operatorname{Re} \left[-\frac{\tilde{m}_{(j)}(z)}{z} \right] \right| \leq \frac{|u|}{v} \operatorname{Im} \left[-\frac{\tilde{m}_{(j)}(z)}{z} \right] + \frac{|z| \sqrt{\bar{W}_{(j)}}}{v \sqrt{v}} \sqrt{\operatorname{Im} \left[-\frac{\tilde{m}_{(j)}(z)}{z} \right]}. \quad (66)$$

As $W_{jj} \geq 0$, $\bar{W}_{(j)} \leq \bar{W}$, so we find the same equation as Equation (62):

$$\left| \operatorname{Re} \left[-\frac{\tilde{m}_{(j)}(z)}{z} \right] \right| \leq \frac{|u|}{v} \operatorname{Im} \left[-\frac{\tilde{m}_{(j)}(z)}{z} \right] + \frac{|z| \sqrt{\bar{W}}}{v \sqrt{v}} \sqrt{\operatorname{Im} \left[-\frac{\tilde{m}_{(j)}(z)}{z} \right]}. \quad (67)$$

So, from the previous proof of the first inequality, we deduce immediatly the second one, which complete the proof of this lemma:

$$\begin{aligned} u \neq 0 &\implies \left\| \left(-\frac{\tilde{m}_{(j)}(z)}{z}A + I \right)^{-1} \right\| \leq 16 \left(\frac{|z|^2}{4v^2|u|} \bar{W} \|A\| + 1 \right) \times \max\left(\frac{1}{3}, \frac{|u|}{v}\right), \\ u = 0 &\implies \left\| \left(-\frac{\tilde{m}_{(j)}(z)}{z}A + I \right)^{-1} \right\| \leq \max\left(2, \frac{4\bar{W}\|A\|}{v}\right). \end{aligned} \quad (68)$$

□

Corollary 1. *Let $j \in \llbracket 1, N \rrbracket$ and $z = u + iv$, $v > 0$. Then, for any matrices A and B of same size $n \times n$, A Hermitian non-negative, we have:*

$$\left| \operatorname{tr} \left[B \left(\left(-\frac{\tilde{m}_n(z)}{z}A + I \right)^{-1} - \left(-\frac{\tilde{m}_{(j)}(z)}{z}A + I \right)^{-1} \right) \right] \right| \leq \left| \frac{\tilde{m}_n(z)}{z} - \frac{\tilde{m}_{(j)}(z)}{z} \right| n \|B\| \|A\| f(z, \|A\|)^2. \quad (69)$$

Proof. For any invertible matrices C_1, C_2 of the same size than B , we have:

$$\begin{aligned} |\operatorname{tr} [B(C_1^{-1} - C_2^{-1})]| &= \operatorname{tr} [BC_1^{-1}(C_2 - C_1)C_2^{-1}] \\ |\operatorname{tr} [B(C_1^{-1} - C_2^{-1})]| &\leq \|B\| \times \|C_1^{-1}\| \times \|C_2^{-1}\| \times n \|C_2 - C_1\|. \end{aligned} \quad (70)$$

From that point, the result comes immediatly from Lemma 4. □

Corollary 2. Let $j \in \llbracket 1, N \rrbracket$ and $z = u + iv, v > 0$. Then, for any $n \times n$ matrix A Hermitian non-negative, and $r \in \mathbb{C}^n$, $\|r\|$ denoting its euclidean norm, we have:

$$\left| r^* \left(\left(-\frac{\tilde{m}_n(z)}{z} A + I \right)^{-1} - \left(-\frac{\tilde{m}_{(j)}(z)}{z} A + I \right)^{-1} \right) r \right| \leq \left| \frac{\tilde{m}_n(z)}{z} - \frac{\tilde{m}_{(j)}(z)}{z} \right| \|r\|^2 \|A\| f(z, \|A\|)^2. \quad (71)$$

Proof. For any invertible matrices C_1, C_2 of size $n \times n$, we have:

$$\begin{aligned} r^*(C_1^{-1} - C_2^{-1})r &= r^*C_1^{-1}(C_2 - C_1)C_2^{-1}r \\ r^*(C_1^{-1} - C_2^{-1})r &\leq \|r\|^2 \times \|C_1^{-1}\| \times \|C_2^{-1}\| \times \|C_2 - C_1\|. \end{aligned} \quad (72)$$

From that point, the result comes immediately from Lemma 4. \square

Lemma 5. Let $j \in \llbracket 1, N \rrbracket$. We denote: $A_{(j)} = \sum_{i \neq j} W_{ii} r_i r_i^*$. Then,

$$\left| \frac{\tilde{m}_n(z)}{z} - \frac{\tilde{m}_{(j)}(z)}{z} \right| \leq \frac{2 \log n}{Nv} + \frac{\|A_{(j)}\|}{|z|vN}. \quad (73)$$

Proof. Let $j \in \llbracket 1, N \rrbracket$. We have, using Lemma 1:

$$\begin{aligned} \left| \frac{\tilde{m}_n(z)}{z} - \frac{\tilde{m}_{(j)}(z)}{z} \right| &= \left| \frac{W_{jj} r_j^* (B_n - zI)^{-1} r_j}{zN} + \sum_{i \neq j} \frac{W_{ii}}{zN} r_i^* [(B_n - zI)^{-1} - (B_{(j)} - zI)^{-1}] r_i \right| \\ \left| \frac{\tilde{m}_n(z)}{z} - \frac{\tilde{m}_{(j)}(z)}{z} \right| &\leq \left| \frac{W_{jj}}{zN} \left(1 - \frac{1}{1 + r_j^* (B_{(j)} - zI)^{-1} r_j} \right) \right| \\ &\quad + \left| \frac{1}{zN} \frac{r_j^* (B_{(j)} - zI)^{-1} A_{(j)} (B_{(j)} - zI)^{-1} r_j}{1 + r_j^* (B_{(j)} - zI)^{-1} r_j} \right|. \end{aligned} \quad (74)$$

For the first term, using Lemma 3, we have:

$$\left| \frac{W_{jj}}{zN} \left(1 - \frac{1}{1 + r_j^* (B_{(j)} - zI)^{-1} r_j} \right) \right| \leq \frac{\log n}{N|z|} + \frac{\log n}{Nv} \leq \frac{2 \log n}{Nv}. \quad (75)$$

For the second term, we have:

$$\begin{aligned} \left| \frac{1}{zN} \frac{r_j^* (B_{(j)} - zI)^{-1} A_{(j)} (B_{(j)} - zI)^{-1} r_j}{1 + r_j^* (B_{(j)} - zI)^{-1} r_j} \right| &\leq \frac{\|A_{(j)}\|}{|z|N} \frac{\|(B_{(j)} - zI)^{-1} r_j\|^2}{|1 + r_j^* (B_{(j)} - zI)^{-1} r_j|} \\ \left| \frac{1}{zN} \frac{r_j^* (B_{(j)} - zI)^{-1} A_{(j)} (B_{(j)} - zI)^{-1} r_j}{1 + r_j^* (B_{(j)} - zI)^{-1} r_j} \right| &\leq \frac{\|A_{(j)}\|}{|z|vN}. \end{aligned} \quad (76)$$

Using the proof of Lemma 2.6 [21], we have additionally that:

$$\frac{\|(B_{(j)} - zI)^{-1} r_j\|^2}{|1 + r_j^* (B_{(j)} - zI)^{-1} r_j|} \leq \frac{1}{v}. \quad (77)$$

So,

$$\left| \frac{1}{zN} \frac{r_j^* (B_{(j)} - zI)^{-1} A_{(j)} (B_{(j)} - zI)^{-1} r_j}{1 + r_j^* (B_{(j)} - zI)^{-1} r_j} \right| \leq \frac{\|A_{(j)}\|}{|z|vN}, \quad (78)$$

which concludes the proof. \square

Lemma 6. We have:

$$\begin{aligned} \max_{j \leq N} \|q_j\|^2 &\longrightarrow 1 \text{ a.s.}, \\ \forall p \in \mathbb{N}, \max_{j \leq N} \frac{\log(n)^p}{N} \|A_{(j)}\| &\longrightarrow 0 \text{ a.s.} \end{aligned} \quad (79)$$

Proof. The first convergence is a direct use of Lemma 2 as remarked in [21] p.338.

Let $p \in \mathbb{N}$. The second convergence comes from the following derivations for $j \in \llbracket 1, N \rrbracket$:

$$\begin{aligned}
\frac{\log(n)^{2p}}{N^2} \|A_{(j)}\|^2 &\leq \frac{1}{N^2} \operatorname{tr} \left[\left(\sum_{i \neq j} W_{ii} r_i r_i^* \right) \right] \\
&\leq \frac{c_n^2 \log(n)^{2p}}{N^2} \left(\sum_{i \neq j} W_{ii}^4 \|T\|^2 \|q_i\|^4 + \sum_{i \neq j} \sum_{i' \neq i, i' \neq j} W_{ii}^2 W_{i'i'}^2 |q_{i'}^* T q_i|^2 \right) \\
&\leq \frac{c_n^2 \log(n)^{2p}}{N^2} \left(\sum_{i=1}^N W_{ii}^4 \|T\|^2 \|q_i\|^4 + \sum_{i=1}^N \sum_{i' \neq i} W_{ii}^2 W_{i'i'}^2 |q_{i'}^* T q_i|^2 \right) \quad (80) \\
&\leq \frac{c_n^2 \log(n)^{2p+4}}{N^2} \left(\sum_{i=1}^N \log(n)^2 \|q_i\|^4 + \sum_{i=1}^N \sum_{i' \neq i} |q_{i'}^* T q_i|^2 \right) \\
\frac{1}{N^2} \|A_{(j)}\|^2 &\leq \frac{c_n^2 \log(n)^{2p+6}}{N} \max_{i \leq N} \|q_i\|^4 + \frac{c_n^2 \log(n)^{2p+4}}{N} \max_{i \leq N} \sum_{i' \neq i} |q_{i'}^* T q_i|^2.
\end{aligned}$$

The upper bound does not depend on j anymore, so we have:

$$\max_{j \leq N} \frac{1}{N^2} \|A_{(j)}\|^2 \leq \frac{c_n^2 \log(n)^{2p+6}}{N} \max_{i \leq N} \|q_i\|^4 + \frac{c_n^2 \log(n)^{2p+4}}{N} \max_{i \leq N} \sum_{i' \neq i} |q_{i'}^* T q_i|^2. \quad (81)$$

From the first part of the proof, we have that $\max_{i \leq N} \|q_i\|^4 \rightarrow 1$ a.s., so:

$$\frac{c_n^2 \log(n)^{2p+6}}{N} \max_{i \leq N} \|q_i\|^4 \rightarrow 0 \text{ a.s.} \quad (82)$$

For the second term, we have:

$$\frac{c_n^2 \log(n)^{2p+4}}{N} \max_{i \leq N} \sum_{i' \neq i} |q_{i'}^* T q_i|^2 = \frac{c_n^2 \log(n)^{2p+4}}{N} \max_{i \leq N} \sum_{i' \neq i} q_{i'}^* (T q_i q_i^* T) q_{i'}. \quad (83)$$

Let $i \in \llbracket 1, N \rrbracket$. We use Lemma 2 in dimension $n(N-1)$ with $Y = (Z_{ki'})_{k, i' \neq i} \in \mathbb{R}^{n \times (N-1)}$ (in

vectorized form) and $C = \begin{pmatrix} T q_i q_i^* T & & (0) \\ & \ddots & \\ (0) & & T q_i q_i^* T \end{pmatrix}$ of size $n(N-1) \times n(N-1)$. We have then:

$$\begin{aligned}
\mathbb{E} \left[\left| \frac{\log(n)^{2p+4}}{N} \sum_{i' \neq i} q_{i'}^* (T q_i q_i^* T) q_{i'} - \frac{(N-1) \log(n)^{2p+4}}{nN} \|T q_i\|^2 \right|^6 \right] &\leq \\
\frac{K n^3 (N-2)^3 \log(n(N-2))^6 \log(n)^{6(2p+4)}}{N^6 n^6} \|C\|^6 &\leq \\
\frac{K \log(nN)^6}{N^3 n^3} \log(n)^{6(2p+8)}. &
\end{aligned} \quad (84)$$

So, for all $\epsilon > 0$:

$$\begin{aligned}
\mathbb{P} \left[\left| \max_{i \leq N} \frac{1}{N} \sum_{i' \neq i} q_{i'}^* (T q_i q_i^* T) q_{i'} - \max_{i \leq N} \frac{N-1}{nN} \|T q_i\|^2 \right| \geq \epsilon \right] &\leq \\
N \times \mathbb{P} \left[\left| \frac{1}{N} \sum_{i' \neq 1} q_{i'}^* (T q_1 q_1^* T) q_{i'} - \frac{N-1}{nN} \|T q_1\|^2 \right| \geq \epsilon \right] &\leq \\
\leq \frac{K \log(nN)^6}{\epsilon^6 N^2 n^3} \log(n)^{6(2p+8)}, &
\end{aligned} \quad (85)$$

which is summable. So,

$$\max_{i \leq N} \frac{\log(n)^{2p+4}}{N} \sum_{i' \neq i} q_{i'}^* (Tq_i q_i^* T) q_{i'} - \max_{i \leq N} \frac{(N-1) \log(n)^{2p+4}}{nN} \|Tq_i\|^2 \rightarrow 0 \text{ a.s.} \quad (86)$$

And, from the first part of the proof:

$$\left| \max_{i \leq N} \frac{c_n^2 (N-1) \log(n)^{2p+4}}{nN} \|Tq_i\|^2 \right| \leq \frac{c_n^2 \log(n)^{2p+6}}{nN} \max_{i \leq N} \|q_i\|^2 \rightarrow 0 \text{ a.s.} \quad (87)$$

We can conclude the proof:

$$\max_{j \leq N} \frac{\log(n)^p}{N} \|A_{(j)}\| \rightarrow 0 \text{ a.s.} \quad (88)$$

□

8.2.4 Proof that $\max_{j \leq N} |d_j| \rightarrow 0$ a.s.

Proof that $\max_{j \leq N} |d_j^{(1)}| \rightarrow 0$ a.s. Let $j \in \llbracket 1, N \rrbracket$. We recall that:

$$d_j^{(1)} = W_{jj} q_j^* T_n^{1/2} (B_{(j)} - zI)^{-1} \left[\left(-\frac{\tilde{m}_n(z)}{z} T_n + I \right)^{-1} - \left(-\frac{\tilde{m}_{(j)}(z)}{z} T_n + I \right)^{-1} \right] T_n^{1/2} q_j. \quad (89)$$

Then, using Corollary 2 and Lemma 3, we have:

$$\begin{aligned} |d_j^{(1)}| &\leq W_{jj} \|(B_{(j)} - zI)^{-1}\| \|T\|^2 \|q_j\|^2 \left| \frac{\tilde{m}_n(z)}{z} - \frac{\tilde{m}_{(j)}(z)}{z} \right| f(z, \log(n))^2 \\ |d_j^{(1)}| &\leq \frac{\log(n)^3}{v} f(z, \log(n))^2 \|q_j\|^2 \left| \frac{\tilde{m}_n(z)}{z} - \frac{\tilde{m}_{(j)}(z)}{z} \right|. \end{aligned} \quad (90)$$

Using Lemma 5, we have:

$$|d_j^{(1)}| \leq \frac{\log(n)^3}{v} f(z, \log(n))^2 \left(\frac{2 \log n}{Nv} + \frac{\|A_{(j)}\|}{|z|vN} \right) \|q_j\|^2. \quad (91)$$

By Assumption (d), $\int x dF^{W_n}(x) \rightarrow \int x dD(x) < \infty$ a.s., i.e. $\bar{W} \rightarrow \int x dD(x) \leq \infty$ a.s., so \bar{W} is bounded a.s. As a consequence, $f(z, \log(n))^2 = O(\log(n)^2)$ a.s.. Finally, using Lemma 6, we can conclude:

$$\max_{j \leq N} |d_j^{(1)}| \rightarrow 0 \text{ a.s.} \quad (92)$$

Proof that $\max_{j \leq N} |d_j^{(2)}| \rightarrow 0$ a.s. Let $j \in \llbracket 1, N \rrbracket$. We recall that:

$$\begin{aligned} d_j^{(2)} &= W_{jj} q_j^* T_n^{1/2} (B_{(j)} - zI)^{-1} \left(-\frac{\tilde{m}_{(j)}(z)}{z} T_n + I \right)^{-1} T_n^{1/2} q_j \\ &\quad - \frac{W_{jj}}{n} \text{tr} \left(\left(-\frac{\tilde{m}_{(j)}(z)}{z} T_n + I \right)^{-1} T_n (B_{(j)} - zI)^{-1} \right). \end{aligned} \quad (93)$$

Using Lemma 2, we have:

$$\mathbb{E} \left[\left| d_j^{(2)} \right|^6 \right] \leq \frac{K \log(n)^{12}}{n^3} \left\| W_{jj} \left(-\frac{\tilde{m}_{(j)}(z)}{z} T_n + I \right)^{-1} T_n (B_{(j)} - zI)^{-1} \right\|^6. \quad (94)$$

Using Lemmas 3 and 4, we have:

$$\mathbb{E} \left[\left| d_j^{(2)} \right|^6 \right] \leq \frac{K \log(n)^{24}}{v^6 n^3} f(z, \log(n))^6. \quad (95)$$

As argued above, $f(z, \log(n)) = O(\log(n))$ a.s., so $\sum_{j=1}^N \mathbb{E} \left[\left| d_j^{(2)} \right|^6 \right]$ is a.s. summable. So, we can conclude:

$$\max_{j \leq N} |d_j^{(2)}| \rightarrow 0 \text{ a.s.} \quad (96)$$

Proof that $\max_{j \leq N} |d_j^{(3)}| \rightarrow 0$ **a.s.** Let $j \in \llbracket 1, N \rrbracket$. We recall that:

$$d_j^{(3)} = \frac{W_{jj}}{n} \operatorname{tr} \left(\left[\left(-\frac{\tilde{m}_{(j)}(z)}{z} T_n + I \right)^{-1} - \left(-\frac{\tilde{m}_n(z)}{z} T_n + I \right)^{-1} \right] T_n (B_{(j)} - zI)^{-1} \right). \quad (97)$$

Using Corollary 1 and Lemma 3, we have:

$$|d_j^{(3)}| \leq \frac{\log(n)^3}{nv} \left| \frac{\tilde{m}_n(z)}{z} - \frac{\tilde{m}_{(j)}(z)}{z} \right| n f(z, \log(n))^2. \quad (98)$$

Using Lemma 5, we have:

$$|d_j^{(3)}| \leq \frac{\log(n)^3}{v} \left(\frac{2 \log n}{nv} + \frac{\|A_{(j)}\|}{|z|vn} \right) f(z, \log(n))^2. \quad (99)$$

As we argued previously, $f(z, \log(n))^2 = O(\log(n)^2)$ a.s., so from Lemma 6 we can conclude:

$$\max_{j \leq N} |d_j^{(3)}| \rightarrow 0 \text{ a.s.} \quad (100)$$

Proof that $\max_{j \leq N} |d_j^{(4)}| \rightarrow 0$ **a.s.** Let $j \in \llbracket 1, N \rrbracket$. We recall that:

$$d_j^{(4)} = \frac{W_{jj}}{n} \operatorname{tr} \left(\left(-\frac{\tilde{m}_n(z)}{z} T_n + I \right)^{-1} T_n [(B_{(j)} - zI)^{-1} - (B_n - zI)^{-1}] \right) \quad (101)$$

Using Lemma 1, we have:

$$d_j^{(4)} = \frac{W_{jj}}{n} \frac{r_j^* (B_{(j)} - zI)^{-1} \left(-\frac{\tilde{m}_n(z)}{z} T_n + I \right)^{-1} T (B_{(j)} - zI)^{-1} r_j}{1 + r_j^* (B_{(j)} - zI)^{-1} r_j} \quad (102)$$

So,

$$|d_j^{(4)}| = \frac{\log(n)}{n} \left\| \left(-\frac{\tilde{m}_n(z)}{z} T_n + I \right)^{-1} T \right\| \frac{\|(B_{(j)} - zI)^{-1} r_j\|^2}{|1 + r_j^* (B_{(j)} - zI)^{-1} r_j|} \quad (103)$$

Using the proof of Lemma 2.6 [21], we have that:

$$\frac{\|(B_{(j)} - zI)^{-1} r_j\|^2}{|1 + r_j^* (B_{(j)} - zI)^{-1} r_j|} \leq \frac{1}{v}. \quad (104)$$

So, using Lemma 5, we have:

$$|d_j^{(4)}| = \frac{\log(n)^2 f(z, \log(n))}{vn} \quad (105)$$

As argued before, $f(z, \log(n)) = O(\log(n))$ a.s., so we can conclude:

$$\max_{j \leq N} |d_j^{(4)}| \rightarrow 0 \text{ a.s.} \quad (106)$$

We can now conclude this section. The last four points prove that:

$$\max_{j \leq N} |d_j| \rightarrow 0 \text{ a.s.} \quad (107)$$

And from Lemma 3, we have for $j \in \llbracket 1, N \rrbracket$:

$$\frac{1}{|z(1 + r_j^* (B_{(j)} - zI)^{-1} r_j)|} \leq \frac{1}{v}. \quad (108)$$

So,

$$\frac{1}{N} \sum_{j=1}^N \frac{-1}{z(1 + r_j^* (B_{(j)} - zI)^{-1} r_j)} d_j \xrightarrow[n \rightarrow \infty]{} 0 \text{ a.s.} \quad (109)$$

Using Equation (41), we can now conclude that:

$$\frac{1}{n} \operatorname{tr} \left((\tilde{m}_n(z) T_n - zI)^{-1} \right) - m_n(z) \xrightarrow[n \rightarrow \infty]{} 0 \text{ a.s.} \quad (110)$$

8.3 Convergences and functional equation

For this section, we introduced an object used in [24], we define:

$$\Theta_n^{(1)}(z) = \frac{1}{n} \operatorname{tr} \left((B_n - zI)^{-1} T_n \right). \quad (111)$$

8.3.1 Proof that a.s., $\forall p \in \mathbb{N}$, $\log(n)^p \left| \tilde{m}_n(z) \Theta_n^{(1)}(z) - (1 + zm_n(z)) \right| \xrightarrow{n \rightarrow \infty} 0$

Let $p \in \mathbb{N}$. We have:

$$\begin{aligned} \left| \tilde{m}_n(z) \Theta_n^{(1)}(z) - (1 + zm_n(z)) \right| &= \left| \tilde{m}_n(z) \Theta_n^{(1)}(z) - \frac{1}{n} \sum_{j=1}^N \frac{r_j^*(B_{(j)} - zI)^{-1} r_j}{1 + r_j^*(B_{(j)} - zI)^{-1} r_j} \right| \\ &= \left| \frac{1}{N} \sum_{j=1}^N \frac{W_{jj} \left(\Theta_n^{(1)}(z) - q_j^*(B_{(j)} - zI)^{-1} q_j \right)}{1 + r_j^*(B_{(j)} - zI)^{-1} r_j} \right| \\ \left| \tilde{m}_n(z) \Theta_n^{(1)}(z) - (1 + zm_n(z)) \right| &\leq \frac{|z|}{v} \max_{j \leq N} \log(n) \left| \Theta_n^{(1)}(z) - q_j^* T^{1/2} (B_{(j)} - zI)^{-1} T^{1/2} q_j \right|. \end{aligned} \quad (112)$$

Using Lemma 2, we have that:

$$\mathbb{E} \left[\left| \log(n)^p \Theta_n^{(1)}(z) - \log(n)^p q_j^* T^{1/2} (B_{(j)} - zI)^{-1} T^{1/2} q_j \right|^6 \right] \leq K \frac{\log(n)^{6(p+3)}}{n^3 v^6}. \quad (113)$$

So,

$$\max_{j \leq N} \left| \log(n)^p \Theta_n^{(1)}(z) - \log(n)^p q_j^* T^{1/2} (B_{(j)} - zI)^{-1} T^{1/2} q_j \right| \xrightarrow{n \rightarrow \infty} 0 \text{ a.s.} \quad (114)$$

As \mathbb{N} is countable, we have also:

$$\text{a.s., } \forall p \in \mathbb{N}, \max_{j \leq N} \log(n)^p \left| \Theta_n^{(1)}(z) - q_j^* T^{1/2} (B_{(j)} - zI)^{-1} T^{1/2} q_j \right| \xrightarrow{n \rightarrow \infty} 0 \text{ a.s.} \quad (115)$$

We can conclude:

$$\text{a.s., } \forall p \in \mathbb{N}, \log(n)^p \left| \tilde{m}_n(z) \Theta_n^{(1)}(z) - (1 + zm_n(z)) \right| \xrightarrow{n \rightarrow \infty} 0. \quad (116)$$

From Assumption (d), $\bar{W} \rightarrow \int x dD(x) \in \mathbb{R}_+$ a.s. We focus now on trajectories where $\bar{W} \rightarrow \int x dD(x) \in \mathbb{R}_+$, Equations (110) and (115) hold, $F^{T_n} \implies H$ and $F^{W_n} \implies D$.

Then, $(\tilde{m}_n(z))$ is bounded. Indeed, from Lemma 3, $|\tilde{m}_n(z)| \leq \bar{W} \frac{|z|}{v}$. Then, as $\operatorname{Im}[\tilde{m}_n(z)] \leq 0$, it exists a subsequence $\{n_i\}$ of \mathbb{N} and $\tilde{m}(z) \in \mathbb{C} \setminus \mathbb{C}_+$ such that $\tilde{m}_{n_i}(z) \xrightarrow{i \rightarrow \infty} \tilde{m}(z) \in \mathbb{C} \setminus \mathbb{C}_+$.

8.3.2 Proof that $m_{n_i}(z) \xrightarrow{i \rightarrow \infty} \int \frac{1}{\tau \tilde{m}(z) - z} dH(\tau)$.

We want to prove that:

$$m_{n_i}(z) - \int \frac{1}{\tau \tilde{m}(z) - z} dH(\tau) \xrightarrow{i \rightarrow \infty} 0. \quad (117)$$

From Equation (110), it is equivalent to prove that:

$$\int \frac{1}{\tau \tilde{m}_{n_i}(z) - z} dF^{T_{n_i}}(\tau) - \int \frac{1}{\tau \tilde{m}(z) - z} dH(\tau) \xrightarrow{i \rightarrow \infty} 0. \quad (118)$$

We prove that $\int \frac{1}{\tau \tilde{m}_{n_i}(z) - z} dF^{T_{n_i}}(\tau) \xrightarrow{i \rightarrow \infty} -\frac{1}{z}$ using the Lebesgue's convergence theorem for weakly converging measures, as detailed in Corollary 5.1 [34]. We denote: $f : \tau \in \mathbb{R} \rightarrow \frac{1}{\tau \tilde{m}(z) - z}$ and $f_i : \tau \in \mathbb{R} \rightarrow \frac{1}{\tau \tilde{m}_{n_i}(z) - z}$. Regarding the hypotheses of the theorem, we have:

- $(f_i)_i$ is a.u.i. w.r.t. $(F^{T_{n_i}})_i$ (see (2.4) [?] for a definition). Indeed, $\forall \tau \in \mathbb{R}_+, \forall i \in \mathbb{R}_+, |f_i(\tau)| \leq \frac{1}{v}$, so $\lim_{K \rightarrow +\infty} \limsup_{i \rightarrow +\infty} \int |f_i(\tau)| 1_{[K, +\infty[}(|f_i(\tau)|) dF^{T_{n_i}} = 0$.
- $F^{T_{n_i}} \implies H$ by assumption.
- Let $\tau \in \mathbb{R}_+$ and $\epsilon > 0$. By assumption, \bar{W}_{n_i} is bounded and we denote by κ one of its finite upper bound. By (3), we have that $\forall i, |\tilde{m}_{n_i}(z)| \leq \frac{\kappa}{|z|v}$, so $\tilde{m}(z) \leq \frac{\kappa}{|z|v}$. We define:

$$\delta := \min \left(1, \frac{v^2}{\frac{\kappa}{|z|v} + \tau + 1} \right) > 0. \quad (119)$$

Then, there exists $i_0 \in \mathbb{N}$ such that $\forall i \geq i_0, |\tilde{m}_{n_i}(z) - \tilde{m}(z)| \leq \delta$. Now, let $\tau' \in]\tau - \delta, \tau + \delta[\cap \mathbb{R}_+$ and $i \geq i_0$. Then,

$$\begin{aligned} |f_i(\tau') - f(\tau)| &= \frac{|\tau' \tilde{m}_{n_i}(z) - \tau \tilde{m}(z)|}{|\tau' \tilde{m}_{n_i}(z) - z| \times |\tau \tilde{m}(z) - z|} \\ &\leq \frac{1}{v^2} (|(\tau' - \tau) \tilde{m}(z)| + |\tau' (\tilde{m}_{n_i}(z) - \tilde{m}(z))|) \\ &\leq \frac{\delta}{v^2} \left(\frac{\kappa}{|z|v} + \tau + 1 \right) \\ |f_i(\tau') - f(\tau)| &\leq \epsilon. \end{aligned} \quad (120)$$

So $\lim_{i \rightarrow \infty, \tau' \rightarrow \tau} f_i(\tau')$ exists.

So, using Corollary 5.1 [34] on the real and imaginary parts, we deduce that $\lim_{i \rightarrow \infty} \int f_i(\tau) dF^{T_{n_i}}(\tau)$ exists and:

$$\lim_{i \rightarrow \infty} \int f_i(\tau) dF^{T_{n_i}}(\tau) = \int f(\tau) dH(\tau). \quad (121)$$

It immediatly leads to:

$$m_{n_i}(z) - \int \frac{1}{\tau \tilde{m}(z) - z} dH(\tau) \xrightarrow{i \rightarrow \infty} 0. \quad (122)$$

8.3.3 Case $H = 1_{[0, +\infty[}$.

Suppose $H = 1_{[0, +\infty[}$. We have then that a.s. $m_n(z) \xrightarrow{n \rightarrow +\infty} -\frac{1}{z}$. Moreover, the equation $m = \int \frac{\delta}{1 + \delta c \int \frac{\delta}{\tau m - z} dH(\tau)} dD(\delta) = \int \delta dD(\delta)$ has trivially a unique solution in $\mathbb{C} \setminus \mathbb{C}_+$ that we denote $\tilde{m}^{(0)}(z) := \int \delta dD(\delta)$. As $-\frac{1}{z} = \int \frac{1}{\tau \tilde{m}^{(0)}(z) - z} dH(\tau)$, we indeed have that $m_n(z) \xrightarrow{n \rightarrow +\infty} \int \frac{1}{\tau \tilde{m}^{(0)}(z) - z} dH(\tau)$. And $z \mapsto -\frac{1}{z}$ is the Cauchy-Stieltjes transform of the p.d.f $1_{[0, +\infty[}$ which complete the proof of the Theorem in the case $H = 1_{[0, +\infty[}$.

Until the end of the proof, we now suppose that $H(]0, +\infty]) > 0$.

8.3.4 Case $D = 1_{[0, +\infty[}$.

Suppose $D = 1_{[0, +\infty[}$. Then, $\bar{W}_n \xrightarrow{n \rightarrow +\infty} 0$. By (3), $\tilde{m}_n(z) \leq \frac{\bar{W}_n}{|z|v}$. So the complete sequence (not the subsequence) $\tilde{m}_n(z) \xrightarrow{n \rightarrow +\infty} 0$. From the previous section on Equation (122), we proved that if $\tilde{m}_n(z)$ converges to some $\tilde{m}(z)$, then $m_n(z) - \int \frac{1}{\tau \tilde{m}(z) - z} dH(\tau) \xrightarrow{n \rightarrow \infty} 0$. So, with $\tilde{m}^{(1)}(z) := 0$, we have that $m_n(z) \xrightarrow{n \rightarrow \infty} \int \frac{1}{\tau \tilde{m}^{(1)}(z) - z} dH(\tau) = -\frac{1}{z}$. Moreover, we remark that $\tilde{m}^{(1)}(z) = 0$ is the unique solution to the equation $m = \int \frac{\delta}{1 + \delta c \int \frac{\delta}{\tau m - z} dH(\tau)} dD(\delta) = \int \delta dD(\delta)$ in $\mathbb{C} \setminus \mathbb{C}_+$. And $z \mapsto -\frac{1}{z}$ is the Cauchy-Stieltjes transform of the p.d.f $1_{[0, \infty[}$ which complete the proof of the Theorem in the case $D = 1_{[0, +\infty[}$.

Until the end of the proof, we now suppose that $D(]0, +\infty]) > 0$.

8.3.5 Case $\tilde{m}(z) = 0$.

Suppose $\tilde{m}(z) = 0$. Then, $m_{n_i}(z) \xrightarrow{i \rightarrow \infty} -\frac{1}{z}$ and by Cauchy-Stieltjes transform property $F^{B_{n_i}} \implies 1_{[0, +\infty[}$. Consequently, $\mathbb{E} \left[\frac{1}{n_i} \text{tr}(B_{n_i}) \right] \xrightarrow{i \rightarrow +\infty} 0$. As $\mathbb{E} \left[\frac{1}{n_i} \text{tr}(B_{n_i}) \right] = \frac{1}{n_i} \text{tr}(T_{n_i}) \times \bar{W}_{n_i}$. \bar{W}_{n_i} converges by assumption to $\int \delta dD(\delta)$, and we supposed $D(]0, +\infty]) > 0$ so $\int \delta dD(\delta) > 0$. We deduce that $\frac{1}{n_i} \text{tr}(T_{n_i}) \xrightarrow{i \rightarrow +\infty} 0$. So, $H = 1_{]0, +\infty[}$, which is absurd. So $\tilde{m}(z) \neq 0$.

8.3.6 Proof that $1 + zm_{n_i}(z) \xrightarrow{i \rightarrow \infty} \int \frac{\delta \Theta^{(1)}(z)}{1 + \delta c \Theta^{(1)}(z)} dD(\delta)$ with $\Theta^{(1)}(z) := \frac{1 + zm(z)}{\tilde{m}(z)}$.

We remind that we suppose now: $H(]0, +\infty]) > 0$, $D(]0, +\infty]) > 0$ and $\tilde{m}_{n_i}(z) \xrightarrow{i \rightarrow +\infty} \tilde{m}(z) \neq 0$.

Firstly, we denote: $M_n := \log(n) \max_{j \leq N} \left| \Theta_n^{(1)}(z) - \log(n)^p q_j^* T^{1/2} (B_{(j)} - zI)^{-1} T^{1/2} q_j \right|$. There is n_0 large enough so that $\forall n \geq n_0, M_n \leq \frac{v}{2c_n}$, due to Equation (115). Then, we have for $n_i \geq n_0$:

$$\begin{aligned} & \left| (-c_n(1 + zm_{n_i}(z)) + 1) - \frac{1}{N} \sum_{j=1}^N \frac{1}{1 + W_{jj} c_{n_i} \Theta_{n_i}^{(1)}} \right| = \\ & \left| \frac{1}{N} \sum_{j=1}^N \frac{1}{1 + r_j^* (B_{(j)} - zI)^{-1} r_j} - \frac{1}{1 + W_{jj} c_{n_i} \Theta_{n_i}^{(1)}} \right| \leq \\ & \frac{c_{n_i} \bar{W} M_{n_i}}{v(v - c_{n_i} M_{n_i})} \xrightarrow{i \rightarrow \infty} 0. \end{aligned} \quad (123)$$

Then, we denote $\Theta^{(1)}(z) := \frac{1 + zm(z)}{\tilde{m}(z)} = \int \frac{\tau}{\tau \tilde{m}(z) - z} dH(\tau)$. From Equation (116), $\Theta_{n_i}^{(1)} \xrightarrow{i \rightarrow \infty} \Theta^{(1)}(z)$. We define:

$$g : \delta \in \mathbb{R} \rightarrow \frac{1}{1 + \delta c \Theta^{(1)}(z)}. \quad (124)$$

We want to prove that:

$$\frac{1}{N} \sum_{j=1}^N \frac{1}{1 + W_{jj} c_{n_i} \Theta_{n_i}^{(1)}} \xrightarrow{i \rightarrow \infty} \int g(\delta) dD(\delta). \quad (125)$$

Let $\delta \in \mathbb{R}$. Remark that, as $\text{Im}[\tilde{m}(z)] \leq 0$:

$$\text{Im}[\Theta^{(1)}(z)] = \text{Im} \left[\frac{1 + zm(z)}{\tilde{m}(z)} \right] = \text{Im} \left[\int \frac{\tau}{\tau \tilde{m}(z) - z} dH(\tau) \right] \geq v \int \frac{\tau}{|\tau \tilde{m}(z) - z|^2} dH(\tau) > 0. \quad (126)$$

So, from [22] p.338, we have:

$$|g(\delta)| \leq \frac{|\Theta^{(1)}(z)|}{\text{Im}[\Theta^{(1)}(z)]}. \quad (127)$$

And, for n_i large enough so that $\forall n_i \geq n_1, \text{Im}[\Theta_{n_i}^{(1)}] > 0$:

$$\left| \frac{1}{1 + \delta c_{n_i} \Theta_{n_i}^{(1)}} - g(\delta) \right| \leq \frac{1}{c_{n_i} \text{Im}[\Theta_{n_i}^{(1)}(z)]} \times \frac{|\Theta^{(1)}(z)|}{\text{Im}[\Theta^{(1)}(z)]} \left| c \Theta^{(1)}(z) - c_{n_i} \Theta_{n_i}^{(1)} \right|. \quad (128)$$

Using (123), (127) and (128), we have that:

$$c_{n_i} (1 + zm_{n_i}(z)) - 1 \xrightarrow{i \rightarrow \infty} \int \frac{1}{1 + \delta c \Theta^{(1)}(z)} dD(\delta). \quad (129)$$

We conclude that:

$$1 + zm_{n_i}(z) \xrightarrow{i \rightarrow \infty} \int \frac{\delta \Theta^{(1)}(z)}{1 + \delta c \Theta^{(1)}(z)} dD(\delta). \quad (130)$$

Finally, using (122) and (130), we have that:

$$\tilde{m}(z) = \int \frac{\delta}{1 + \delta c \int \frac{\tau}{\tau \tilde{m}(z) - z} dH(\tau)} dD(\delta). \quad (131)$$

8.4 Unicity

In this section, we show that there is at most one solution $m \in \mathbb{C} \setminus \mathbb{C}_+$ so that:

$$m = \int \frac{\delta}{1 + \delta c \int \frac{\tau}{\tau m - z} dH(\tau)} dD(\delta). \quad (132)$$

Let $m \in \mathbb{C} \setminus \mathbb{C}_+$, verifying (132). Then, as $H([0, +\infty]) > 0$:

$$\operatorname{Im}[m] = - \int \frac{\delta^2 c \int \frac{\tau \operatorname{Im}[-m] + v}{|\tau m - z|^2} dH(\tau)}{\left|1 + \delta c \int \frac{\tau}{\tau m - z} dH(\tau)\right|^2} dD(\delta) < 0. \quad (133)$$

So, with $\mathbb{C}_- = \{y \in \mathbb{C} \mid \operatorname{Im}[y] < 0\}$:

$$[m \in \mathbb{C} \setminus \mathbb{C}_+ \text{ verifies (132) and } H([0, +\infty]) > 0] \iff [m \in \mathbb{C}_- \text{ verifies 132 and } H([0, +\infty]) > 0]. \quad (134)$$

Let $m_1 = u_1 - iv_1 \in \mathbb{C}_-$ and $m_2 = u_2 - iv_2 \in \mathbb{C}_-$ solving (132) at $z = u + iv \in \mathbb{C}_+$. Then,

$$\begin{aligned} m_1 - m_2 &= \int \frac{\delta^2 c \left(\int \frac{\tau}{\tau m_2 - z} dH(\tau) - \int \frac{\tau}{\tau m_1 - z} dH(\tau) \right)}{\left(1 + \delta c \int \frac{\tau}{\tau m_1 - z} dH(\tau)\right) \left(1 + \delta c \int \frac{\tau}{\tau m_2 - z} dH(\tau)\right)} dD(\delta) \\ m_1 - m_2 &= (m_1 - m_2) \times \int \frac{\delta^2 c \int \frac{\tau^2}{(\tau m_1 - z)(\tau m_2 - z)} dH(\tau)}{\left(1 + \delta c \int \frac{\tau}{\tau m_1 - z} dH(\tau)\right) \left(1 + \delta c \int \frac{\tau}{\tau m_2 - z} dH(\tau)\right)} dD(\delta) \end{aligned} \quad (135)$$

Using Hölder inequality on the last term and using (133) at the end, we have:

$$\begin{aligned} &\left| \int \frac{\delta^2 c \int \frac{\tau^2}{(\tau m_1 - z)(\tau m_2 - z)} dH(\tau)}{\left(1 + \delta c \int \frac{\tau}{\tau m_1 - z} dH(\tau)\right) \left(1 + \delta c \int \frac{\tau}{\tau m_2 - z} dH(\tau)\right)} dD(\delta) \right|^2 \leq \\ &\left| \int \frac{\delta^2 c \int \frac{\tau^2}{|\tau m_1 - z|^2} dH(\tau)}{\left|1 + \delta c \int \frac{\tau}{\tau m_1 - z} dH(\tau)\right|^2} dD(\delta) \right| \times \left| \int \frac{\delta^2 c \int \frac{\tau^2}{|\tau m_2 - z|^2} dH(\tau)}{\left|1 + \delta c \int \frac{\tau}{\tau m_2 - z} dH(\tau)\right|^2} dD(\delta) \right| < \\ &\left| \int \frac{\delta^2 c \int \frac{\tau^2 + \tau \frac{v}{v_1}}{|\tau m_1 - z|^2} dH(\tau)}{\left|1 + \delta c \int \frac{\tau}{\tau m_1 - z} dH(\tau)\right|^2} dD(\delta) \right| \times \left| \int \frac{\delta^2 c \int \frac{\tau^2 + \tau \frac{v}{v_2}}{|\tau m_2 - z|^2} dH(\tau)}{\left|1 + \delta c \int \frac{\tau}{\tau m_2 - z} dH(\tau)\right|^2} dD(\delta) \right| = \\ &\left| \frac{v_1}{v_1} \right| \times \left| \frac{v_2}{v_2} \right| = 1. \end{aligned} \quad (136)$$

Remark that the inequality is strict because we supposed $H([0, +\infty]) > 0$. Injecting this inequation in (138), we find:

$$|m_1 - m_2| \neq 0 \implies |m_1 - m_2| < |m_1 - m_2|. \quad (137)$$

So there is at most one solution $m \in \mathbb{C} \setminus \mathbb{C}_+$ verifying (132).

8.5 Convergences of $\tilde{m}_n(z)$, $m_n(z)$ and $\Theta_n^{(1)}(z)$.

Backing up, we proved that almost surely, $\tilde{m}_n(z)$ is bounded and every convergent subsequence of $\tilde{m}_n(z)$ converge towards the unique $\tilde{m}(z) \in \mathbb{C} \setminus \mathbb{C}_+$ verifying (131). So, a.s., $\tilde{m}_n(z) \xrightarrow[n \rightarrow \infty]{} \tilde{m}(z)$.

We also proved that, almost surely, if $\tilde{m}_n(z) \xrightarrow[n \rightarrow \infty]{} \tilde{m}(z)$ then:

- $m_n(z) \xrightarrow[n \rightarrow \infty]{} m(z) = \int \frac{1}{\tau \tilde{m}(z) - z} dH(\tau)$,
- $\Theta_n^{(1)}(z) \xrightarrow[n \rightarrow \infty]{} \Theta^{(1)}(z) = \int \frac{\tau}{\tau \tilde{m}(z) - z} dH(\tau)$.

So, almost surely, $m_n(z) \xrightarrow[n \rightarrow \infty]{} m(z) = \int \frac{1}{\tau \tilde{m}(z) - z} dH(\tau)$ where $\tilde{m}(z)$ is the unique solution in $\mathbb{C} \setminus \mathbb{C}_+$ to (131).

8.6 m is Stieltjes transform of a p.d.f.

The last point to prove is that m is Stieltjes transform of a p.d.f. As pointwise limit of Stieltjes transform, it is enough to prove that, using Theorem 7 [36], $iy m(iy) \xrightarrow{y \rightarrow \infty} -1$ to show that m is a Stieltjes transform of a p.d.f.

Consider a trajectory where \bar{W} is bounded, say by κ , and $\tilde{m}_n(z) \rightarrow \tilde{m}(z)$. Then, $\forall z \in \mathbb{C}_+$, $|\tilde{m}_n(z)| \leq \kappa \frac{|z|}{v}$. So, $\forall y \in \mathbb{R}_+$, $|\tilde{m}(iy)| \leq \kappa$. Consequently,

$$|-iy m(iy) - 1| = \left| \int \frac{-\tau \tilde{m}(iy)}{\tau \tilde{m}(iy) - iy} dH(\tau) \right| \leq \frac{\kappa}{y}. \quad (138)$$

So, $iy m(iy) \xrightarrow{y \rightarrow \infty} -1$, which proves that m is a Stieltjes transform a p.d.f. and concludes the proof.