World Scientific
www.worldscientific.com

LARGE INFORMATION PLUS NOISE RANDOM MATRIX MODELS AND CONSISTENT SUBSPACE ESTIMATION IN LARGE SENSOR NETWORKS

WALID HACHEM

Telecom Paristech (CNRS UMR 5141)
46 rue Barrault 75634 Paris, France
hachem@telecom-paristech.fr

PHILIPPE LOUBATON

Université Paris-Est Marne la Vallée
LIGM, CNRS UMR 8049
5 Boulevard Descartes 77454 Marne-la-Vallée, France
loubaton@univ-mlv.fr

XAVIER MESTRE

Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
Av. Carl Friedrich Gauss 08860 Castelldefels, Barcelona, Spain
xavier.mestre@cttc.cat

JAMAL NAJIM

Telecom Paristech (CNRS UMR 5141)
46 rue Barrault 75634 Paris, France
najim@telecom-paristech.fr

PASCAL VALLET

Université Paris-Est Marne la Vallée
LIGM, CNRS UMR 8049
5 Boulevard Descartes 77454 Marne-la-Vallée, France
vallet@univ-mlv.fr

In array processing, a common problem is to estimate the angles of arrival of K deterministic sources impinging on an array of M antennas, from N observations of the source signal, corrupted by Gaussian noise. In the so-called subspace methods, the problem reduces to estimate a quadratic form (called "localization function") of a certain projection matrix related to the source signal empirical covariance matrix. The estimates of the angles of arrival are then obtained by taking the K deepest local minima of the estimated localization function. Recently, a new subspace estimation method has

been proposed, in the context where the number of available samples N is of the same order of magnitude than the number of sensors M. In this context, the traditional subspace methods tend to fail because they are based on the empirical covariance matrix of the observations which is a poor estimate of the source signal covariance matrix. The new subspace method is based on a consistent estimator of the localization function in the regime where M and N tend to $+\infty$ at the same rate. However, the consistency of the angles estimator was not addressed, and the purpose of this paper is to prove this consistency in the previous asymptotic regime. For this, we prove the property that the singular values of $M \times N$ Gaussian information plus noise matrix escape from certain intervals is an event of probability decreasing at rate $\mathcal{O}(N^{-p})$ for all p. A regularization trick is also introduced, which allows to confine these singular values into certain intervals and to use standard tools as Poincaré inequality to characterize any moments of the estimator. These results are believed to be of independent interest.

1. Introduction

1.1. *Motivation*

This paper is motivated by the problem of source localization using a large sensor network. In this context, the observation is a complex valued M-variate time series $(\mathbf{y}_n)_{n \in \mathbb{Z}}$ (M represents the number of sensors of the array) given by

$$\mathbf{y}_n = \sum_{k=1}^{K} s_{k,n} \mathbf{a}(\theta_k) + \mathbf{v}_n = \mathbf{A}(\boldsymbol{\theta}) \mathbf{s}_n + \mathbf{v}_n,$$

where

- The $K < M$ scalar (in general complex-valued) time series $(s_{k,n})_{n \in \mathbb{Z}}$ for $k = 1, \ldots, K$ are non-observable, and represent the signals transmitted by K transmitters. The vector $\mathbf{s}_n$ is given by $\mathbf{s}_n = (s_{1,n}, \ldots, s_{K,n})^T$.
- For each k, θ_k is a scalar real parameter characterizing the direction of arrival of transmitter k. $\theta \to \mathbf{a}(\theta)$ is a known $\mathbb{C}^M$-valued function depending on the sensor network geometry, and matrix $\mathbf{A}(\boldsymbol{\theta})$ is defined as $\mathbf{A}(\boldsymbol{\theta}) = (\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K))$.
- $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ finally represents an additive complex Gaussian noise, i.e. $\mathbf{v}_n = (v_{1,n}, \ldots, v_{M,n})^T$ where the M time series $((v_{k,n})_{n \in \mathbb{Z}})_{k=1,\ldots,M}$ are mutually independent identically distributed (iid) sequences such that $\mathrm{Re}(v_{k,n})$ and $\mathrm{Im}(v_{k,n})$ are independent real Gaussian random variables with zero mean and variance $\sigma^2/2$.

The classical source localization problem consists in estimating vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)^T$ from N samples collected in the $M \times N$ matrix $\mathbf{Y}_N = (\mathbf{y}_1, \ldots, \mathbf{y}_N)$. This problem was extensively studied in the past (see e.g. [21] and the references therein). The so-called subspace estimator of $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)^T$ is based on the observation that if matrices $\mathbf{A}(\boldsymbol{\theta})$ and $\mathbf{S}_N = (\mathbf{s}_1, \ldots, \mathbf{s}_N)$ have both full rank K,

then the angles $(\theta_k)_{k=1,\ldots,K}$ are solutions[a] of the equation $\mathbf{a}(\theta)^* \mathbf{\Pi}_N \mathbf{a}(\theta) = 0$, where $\mathbf{\Pi}_N$ represents the orthogonal projection matrix on the kernel of matrix $\mathbf{A}(\boldsymbol{\theta})\mathbf{S}_N\mathbf{S}_N^*\mathbf{A}(\boldsymbol{\theta})^*$. The subspace method is therefore irrelevant if $K \geq M$, which explains why it is assumed that $K < M$. The existing subspace methods consist in estimating for each θ the quadratic form $\eta_N(\theta) = \mathbf{a}(\theta)^*\mathbf{\Pi}_N\mathbf{a}(\theta)$ of $\mathbf{\Pi}_N$ by a certain term $\hat{\eta}_N(\theta)$, and then to estimate the K angles as the argument of the K most significant local minima of function $\theta \to \hat{\eta}_N(\theta)$. This approach has been extensively developed when $N \to +\infty$ and M fixed. In this context, $\eta_N(\theta)$ can be estimated consistently for each θ by $\hat{\eta}_N(\theta) = \mathbf{a}(\theta)^*\hat{\mathbf{\Pi}}_N\mathbf{a}(\theta)$ with $\hat{\mathbf{\Pi}}_N$ the orthogonal projection matrix on the eigenspace associated to the $M - K$ smallest eigenvalues of the empirical covariance matrix $\frac{1}{N}\mathbf{Y}_N\mathbf{Y}_N^*$. It clearly holds that $\sup_{\theta \in [-\pi,\pi]} |\hat{\eta}_N(\theta) - \eta_N(\theta)|$ converges towards 0 almost surely, and this allows to prove that the corresponding estimators $(\hat{\theta}_k)_{k=1,\ldots,K}$ of the direction of arrivals are consistent.

If however M and N are of the same order of magnitude, a quite common situation if the number of sensors M is large, then the above estimators show poor performances because $\hat{\mathbf{\Pi}}_N$ is no longer an accurate estimator of $\mathbf{\Pi}_N$. In order to study this context, Mestre and Lagunas [18] were the first to propose consistent estimators of $\eta_N(\theta)$ when $M, N \to +\infty$ in such a way that $c_N = \frac{M}{N} \to c$, with $c > 0$. In Mestre and Lagunas [18], it is assumed that the source signals $(s_{k,n})_{k=1,\ldots,K}$ are mutually independent complex Gaussian iid time series with unit variance elements. Under this assumption, $\mathbf{y}_n$ can be written as

$$\mathbf{y}_n = \mathbf{R}_y^{1/2}\mathbf{x}_n,$$

where $\mathbf{R}_y = \mathbf{A}(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})^* + \sigma^2\mathbf{I}_M$ represents the covariance matrix of the time series $(\mathbf{y}_n)_{n\in\mathbb{Z}}$ and $\mathbf{x}_n$ is a complex standard Gaussian vector. Matrix $\mathbf{\Pi}_N$ coincides with the orthogonal projection matrix over the eigenspace of $\mathbf{R}_y$ associated to the eigenvalue σ^2, and Mestre and Lagunas addressed the problem of estimating consistently any quadratic form of $\mathbf{\Pi}_N$ from the empirical covariance matrix $\frac{1}{N}\mathbf{Y}_N\mathbf{Y}_N^*$ where $\mathbf{Y}_N = \mathbf{R}_y^{1/2}\mathbf{X}_N$ and $\mathbf{X}_N = (\mathbf{x}_1,\ldots,\mathbf{x}_N)$. Mestre and Lagunas [18] used properties (see Silverstein and Choi [20], Bai and Silverstein [1, 2]) of the empirical covariance matrix, and were able to exhibit an $M \times M$ matrix $\tilde{\mathbf{\Pi}}_{\text{iid},N}$ such that $\mathbf{a}_N^*\tilde{\mathbf{\Pi}}_{\text{iid},N}\mathbf{a}_N - \mathbf{a}_N^*\mathbf{\Pi}_N\mathbf{a}_N \to 0$ for each deterministic bounded sequence of vectors $(\mathbf{a}_N)$ when $M, N \to +\infty$ in such a way that $c_N = \frac{M}{N} \to c$, with $c > 0$. In some sense, matrix $\tilde{\mathbf{\Pi}}_{\text{iid},N}$ can be viewed as a consistent estimate of $\mathbf{\Pi}_N$ but in a weak sense because in general, it does not hold that $\|\tilde{\mathbf{\Pi}}_{\text{iid},N} - \mathbf{\Pi}_N\| \to 0$, where we have denoted by $\|\cdot\|$ the usual spectral norm. Mestre and Lagunas concluded that for each θ, $\mathbf{a}(\theta)^*\tilde{\mathbf{\Pi}}_{\text{iid},N}\mathbf{a}(\theta)$ is a consistent estimate of $\eta_N(\theta)$. However, the consistency of the angular estimates was not established. Note that these results do not require any hypothesis on K which may scale with N or not.

The estimator proposed in [18] is in principle valid only when the source signals $(s_{k,n})_{k=1,\ldots,K}$ are mutually independent complex Gaussian time series, each time

---

[a]The K angles are the unique solutions under certain assumptions on function $\theta \to \mathbf{a}(\theta)$.

series being independent and identically distributed. This is clearly a strong hypothesis which is not verified in a number of applicative contexts. In particular, the source signals may be non-Gaussian and/or correlated in the time domain. In Vallet *et al.* [23], a more general case was considered where the time series $(s_{k,n})_{k=1,\ldots,K}$ are deterministic signals for which the spectral norm of matrix $\frac{1}{\sqrt{N}}\mathbf{A}(\boldsymbol{\theta})\mathbf{S}_N$ is bounded with respect to the dimensions M, N, K. As a realization of any discrete time random process can be interpreted as a deterministic signal, the results of [23] can be used whatever the properties of the $(s_{k,n})_{k=1,\ldots,K}$. In particular, it is proved in [23] that if the source signals are realizations of independent Gaussian iid sequences, then the estimators of [18, 23] have the same behavior. In the deterministic source signals case considered in the present paper, random matrix $\mathbf{Y}_N$ is non-zero mean, and corresponds to the so-called "Information plus Noise model" investigated in various works of Girko (see e.g. [11]) and Dozier and Silverstein (see [9, 8]). Using new results on the almost sure localization of the eigenvalues of the empirical covariance matrix $\frac{1}{N}\mathbf{Y}_N\mathbf{Y}_N^*$, [23] generalized the estimator of Mestre and Lagunas [18], and derived a "weakly consistent estimator" $\tilde{\boldsymbol{\Pi}}_N$ of $\boldsymbol{\Pi}_N$, i.e. $\mathbf{a}_N^*\tilde{\boldsymbol{\Pi}}_N\mathbf{a}_N - \mathbf{a}_N^*\boldsymbol{\Pi}_N\mathbf{a}_N \to 0$ for each deterministic bounded sequences of vectors $(\mathbf{a}_N)$. Therefore, it holds that for each θ, $\tilde{\eta}_N(\theta) = \mathbf{a}(\theta)^*\tilde{\boldsymbol{\Pi}}_N\mathbf{a}(\theta)$ is a consistent estimate of $\eta_N(\theta)$ if $\mathbf{a}(\theta)$ is uniformly bounded in N.

In order to emphasize on the practical interest of doubly infinite consistent estimators of [18, 23], we refer to the various numerical experiments of [18, 23]. The possible values of (M, N) $(20, 15)$, $(20, 40)$, $(20, 75)$, $(20, 150)$, and for each case, the gain produced by the doubly consistent estimator over the traditional subspace method appeared to be quite substantial, even for the case $(20, 150)$ for which c_N is far from 1. This tends to indicate that the use of the new estimators can be recommended in practice in the context of large sensor networks whatever the values of M and N.

The goal of the present paper is to pursue the work [23], and to establish that the angle estimates defined as the K most significant local minima of function $\theta \to \mathbf{a}(\theta)^*\tilde{\boldsymbol{\Pi}}_N\mathbf{a}(\theta)$ are consistent. As it will be shown below, the consistency of the angle estimates is based on the property

$$\sup_{\theta\in[-\pi,\pi]}|\tilde{\eta}_N(\theta) - \eta_N(\theta)| \xrightarrow[N\to\infty]{\text{a.s.}} 0, \tag{1.1}$$

that we shall refer to as the *uniform consistency* of the estimate $\tilde{\eta}_N(\theta)$ of $\eta_N(\theta)$.

This paper is organized as follows. In Sec. 2, we provide some background material on the asymptotic eigenvalue distribution of the large information plus noise model, on the almost sure localization of the eigenvalues of the empirical covariance matrix, and on the consistent estimator of $\mathbf{a}_N^*\boldsymbol{\Pi}_N\mathbf{a}_N$ proposed in [23]. The new results of this paper are essentially presented in Sec. 3 devoted to the proof of the uniform consistency (1.1). As in [23], the proofs are based on Gaussian techniques introduced by Haagerup and Thorbjornsen in [12]. However, (1.1) is much more demanding than the consistency of $\tilde{\eta}_N(\theta)$ proved in [23], and needs the introduction

of new tools. Using a classical discretization scheme of the parameter space, we show that it is sufficient to prove that for each θ and each $\alpha > 0$, $\mathbb{P}(|\tilde{\eta}_N(\theta) - \eta_N(\theta)| > \alpha)$ converges towards 0 fast enough. For this, a tempting choice is to use the Markov inequality, and to evaluate the behavior of the moments of $\tilde{\eta}_N(\theta) - \eta_N(\theta)$. However, these moments do not necessarily exist for some technical reasons. We therefore introduce a well-chosen regularization term χ_N, for which $\mathbb{P}(\chi_N \neq 1)$ decreases at rate $\frac{1}{N^l}$ for each integer l, and such that the moments of $\tilde{\eta}_N(\theta)\chi_N$ exist. We establish that for each $l \in \mathbb{N}$, $\mathbb{E}|\tilde{\eta}_N(\theta)\chi_N - \eta_N(\theta)|^{2l}$ decreases at rate $\frac{1}{N^l}$. This, in turn, allows to establish that (1.1) holds. The properties of the regularized factor χ_N heavily rely on results concerning the probability that the eigenvalues of $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$ and of a related matrix escape from the intervals in which they are located almost surely for N large enough. These results are new, and are believed to be of independent interest. We also mention that the uniform consistency is proved when function $\mathbf{a}(\theta)$ is defined by

$$\mathbf{a}(\theta) = \frac{1}{\sqrt{M}}(1, e^{i\theta}, \ldots, e^{i(M-1)\theta})^T, \tag{1.2}$$

for all $\theta \in [-\pi, \pi]$. (1.1) of course holds for more general functions. In particular, if $\mathbf{a}(\theta)$ is any function satisfying the conditions $\|\mathbf{a}(\theta)\| = 1$ for all θ and

$$\|\mathbf{a}(\theta_1) - \mathbf{a}(\theta_2)\| \leq CN^r|\theta_1 - \theta_2|^s, \tag{1.3}$$

for $s > 0$ and $r \geq 0$, then, the last step of the proof of (1.1) (see Sec. 3.3) can be easily adapted to model (1.3) (the function $\mathbf{a}(\theta)$ defined by (1.2) verifies (1.3) for $r = s = 1$). Note that in (1.3), we have mentioned that function $\mathbf{a}(\theta)$ depends on the dimensions of the system. As model (1.2) is very popular, we prefer to present our results in this context for the sake of simplicity.

In Sec. 4, we finally deduce from (1.1) the consistency of the K most significant local minima of function $\theta \to \tilde{\eta}_N(\theta)$ by following the approach in [15].

1.2. *General notations and useful results*

We now introduce various notations and results used throughout the paper.

- If $E \subset \mathbb{R}$, $\text{Int}(E)$ and ∂E represent the interior and the boundary of E respectively.
- If $z \in \mathbb{C}$, the complex conjugate of z is denoted $\overline{z}$ or z^*. For a complex matrix $\mathbf{A}$, we denote its transpose by $\mathbf{A}^T$ and its Hermitian adjoint by $\mathbf{A}^*$.
- We denote by $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ (respectively $\mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R})$) the set of all smooth real-valued functions (respectively, compactly supported smooth real values functions).
- The quantity C will represent a generic positive constant whose main feature is to be deterministic and independent of M and N. The value of C may change from one line to another.
- Similarly, P_1 and P_2 will denote generic polynomials, independent of M and N, with positive coefficients. The polynomials may change from one line to another.

- Complex Gaussian distribution: A complex-valued random variable $Z = X + iY$ follows the distribution $\mathcal{CN}(\alpha+i\beta, \sigma^2)$ if X and Y are independent real Gaussian random variables $\mathcal{N}(\alpha, \frac{\sigma^2}{2})$ and $\mathcal{N}(\beta, \frac{\sigma^2}{2})$, respectively. The variance of Z, denoted as $\mathrm{Var}(Z)$ is defined as $\mathrm{Var}(Z) = \mathbb{E}|Z - \mathbb{E}[Z]|^2 = \sigma^2$.

- Poincaré inequality (see Chen [6]): Let $Z_1 = X_1 + iY_1, \ldots, Z_p = X_p + iY_p$ be p iid $\mathcal{CN}(0, \sigma^2)$ random variables and consider a function γ defined on $\mathbb{R}^{2p}$ continuously differentiable with polynomially bounded partial derivatives. Then, if $\mathbf{X} = (X_1, \ldots, X_p)^T$ and $\mathbf{Y} = (Y_1, \ldots, Y_p)^T$, the random variable $\gamma(\mathbf{X}, \mathbf{Y})$ can be written as $\gamma(\mathbf{X}, \mathbf{Y}) = \tilde{\gamma}(\mathbf{Z}, \overline{\mathbf{Z}})$ and

$$\mathrm{Var}[\gamma(\mathbf{X}, \mathbf{Y})] = \mathrm{Var}[\tilde{\gamma}(\mathbf{Z}, \overline{\mathbf{Z}})]$$

$$\leq \sigma^2 \sum_{i=1}^{p} \left(\mathbb{E} \left| \frac{\partial \tilde{\gamma}(\mathbf{Z}, \overline{\mathbf{Z}})}{\partial z_i} \right|^2 + \mathbb{E} \left| \frac{\partial \tilde{\gamma}(\mathbf{Z}, \overline{\mathbf{Z}})}{\partial \overline{z}_i} \right|^2 \right),$$

where we define as usual the differential operators $\frac{\partial}{\partial z} = \frac{1}{2}(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y})$ and $\frac{\partial}{\partial \overline{z}} = \frac{1}{2}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y})$. If γ is real-valued, it is clear that $\frac{\partial \tilde{\gamma}(\mathbf{Z}, \overline{\mathbf{Z}})}{\partial \overline{z}_i}$ coincides with the complex conjugate of $\frac{\partial \tilde{\gamma}(\mathbf{Z}, \overline{\mathbf{Z}})}{\partial z_i}$. In this case, the Poincaré inequality reduces to

$$\mathrm{Var}(\gamma(\mathbf{X}, \mathbf{Y})) \leq 2\sigma^2 \sum_i \mathbb{E} \left| \frac{\partial \tilde{\gamma}(\mathbf{Z}, \overline{\mathbf{Z}})}{\partial z_i} \right|^2.$$

- Stieltjes transform: Let μ be a positive finite measure on $\mathbb{R}$. Its Stieltjes transform m is the function defined by

$$m(z) = \int_{\mathbb{R}} \frac{\mathrm{d}\mu(\lambda)}{\lambda - z}, \quad \forall z \in \mathbb{C}\backslash\mathrm{supp}(\mu),$$

where $\mathrm{supp}(\mu)$ represents the support of measure μ. Function m is holomorphic on $\mathbb{C}\backslash\mathrm{supp}(\mu)$ and satisfies $\frac{\mathrm{Im}(m(z))}{\mathrm{Im}(z)} > 0$ for $z \in \mathbb{C}\backslash\mathbb{R}$ and $m(iy) \to 0$ when $y \to +\infty$. Moreover, $\mathrm{supp}(\mu) \subset \mathbb{R}^+$ if and only if $\frac{\mathrm{Im}(zm(z))}{\mathrm{Im}(z)} > 0$ for $z \in \mathbb{C}\backslash\mathbb{R}$. The mass of the measure μ can be evaluated through the formula

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

We also notice that if $m(z)$ is the Stieltjes transform of positive measure μ, then it holds that

$$|m(z)| \leq \frac{\mu(\mathbb{R})}{\mathrm{dist}(z, \mathrm{supp}(\mu))} \leq \frac{\mu(\mathbb{R})}{|\mathrm{Im}(z)|},$$

and that $m'(z) = \int_{\mathbb{R}} \frac{\mathrm{d}\mu(\lambda)}{(\lambda - z)^2}$ satisfies

$$|m'(z)| \leq \frac{\mu(\mathbb{R})}{\mathrm{dist}(z, \mathrm{supp}(\mu))^2} \leq \frac{\mu(\mathbb{R})}{|\mathrm{Im}(z)|^2},$$

on $\mathbb{C}\backslash\mathrm{supp}(\mu)$. We finally recall the following version of the inverse Stieltjes transform formula: for each function $\psi \in \mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R})$, we have

$$\int_{\mathbb{R}} \psi(\lambda)\mathrm{d}\mu(\lambda) = \frac{1}{\pi} \lim_{y \downarrow 0} \mathrm{Im} \left(\int_{\mathbb{R}} \psi(\lambda)m(\lambda + iy)\mathrm{d}\lambda \right). \tag{1.4}$$

2. Background on the Information Plus Noise Model and on the Estimator of [23]

All along this paper, we consider integers $M, N, K \in \mathbb{N}^*$ such that $1 \leq K < M$, $K = K(N)$ and $M = M(N)$ are functions of N with $c_N = \frac{M}{N} \to c > 0$ as $N \to \infty$. We assume that

Assumption A-1. $0 < c_N < 1$ and $0 < c < 1$.

Assumption A-1 is formulated in order to simplify the exposition. The case $c_N > 1$ and $c > 1$ is also relevant, and the consistent estimator $\tilde{\eta}_N(\theta)$ can be easily generalized. However, certain intermediate results proved in Sec. 3 have to be modified, but the uniform consistency is still valid. We refer the reader to the discussion of Sec. 3.4 for more information on the case $c_N > 1$ and $c > 1$. The case $c = 1$ is more intricate, and would need a specific treatment because the density of the asymptotic eigenvalue distribution μ_N of matrix $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ (to be defined below) may become singular at the origin when $c_N \to 1$. This complicates the description of the support of μ_N as well as certain proofs. We therefore do not consider this specific case in the following.

In this section, $\mathbf{\Sigma}_N$ represents the complex-valued $M \times N$ random matrix given by

$$\mathbf{\Sigma}_N = \frac{\mathbf{Y}_N}{\sqrt{N}} = \mathbf{B}_N + \mathbf{W}_N,$$

where $\mathbf{B}_N = \frac{\mathbf{A}(\theta)\mathbf{S}_N}{\sqrt{N}}$ and $\mathbf{W}_N = \frac{\mathbf{V}_N}{\sqrt{N}}$. Matrices $\mathbf{B}_N$ and $\mathbf{W}_N$ are assumed to satisfy the following assumptions:

Assumption A-2. Matrix $\mathbf{B}_N$ is deterministic and satisfies $\sup_N \|\mathbf{B}_N\| < +\infty$.

Assumption A-3. $\mathrm{Rank}(\mathbf{B}_N \mathbf{B}_N^*) = K < M$ where K may scale with N or not.

Assumption A-4. The entries of matrix $\mathbf{W}_N$ are iid and follow the complex normal distribution $\mathcal{CN}(0, \frac{\sigma^2}{N})$.

We assume moreover that the non-zero eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ have multiplicities 1 in order to simplify the notations. In the following, we denote by $0 = \lambda_{1,N} = \cdots = \lambda_{M-K,N} < \lambda_{M-K+1,N} < \cdots < \lambda_{M,N}$ and $(\mathbf{u}_{k,N})_{k=1,\dots,M}$ the ordered eigenvalues and associated eigenvectors of $\mathbf{B}_N \mathbf{B}_N^*$. The eigenvalues and the eigenvectors of matrix $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ are denoted $(\hat{\lambda}_{k,N})_{k=1,\dots,M}$ and $(\hat{\mathbf{u}}_{k,N})_{k=1,\dots,M}$, and $\hat{\mu}_N$ represents the empirical eigenvalue distribution of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ defined by

$$\hat{\mu}_N = \frac{1}{M} \sum_{k=1}^{M} \delta_{\hat{\lambda}_{k,N}}.$$

As we assume $c_N < 1$, the joint probability distribution of $(\hat{\lambda}_{k,N})_{k=1,\dots,M}$ is absolutely continuous (see e.g. James [16]) and it holds that the $(\hat{\lambda}_{k,N})_{k=1,\dots,M}$ have multiplicity 1 almost surely. We finally denote by $\mathbf{Q}_N(z)$ the resolvent of matrix $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$, i.e. $\mathbf{Q}_N(z) = (\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* - z\mathbf{I}_M)^{-1}$.

2.1. *The asymptotic eigenvalue distribution μ_N of $\hat{\mu}_N$*

It is well-known ([11, Theorem 7.4], [9, Theorem 1.1]) that it exists a sequence of deterministic probability measures (μ_N) such that $\hat{\mu}_N - \mu_N \to_N 0$ weakly almost surely. Measure μ_N is characterized by its Stieltjes transform $m_N(z)$ which is known to satisfy the equation

$$m_N(z) = \frac{1}{M}\text{Tr}\left[-z(1+\sigma^2 c_N m_N(z))\mathbf{I}_M + \sigma^2(1-c_N)\mathbf{I}_M + \frac{\mathbf{B}_N\mathbf{B}_N^*}{1+\sigma^2 c_N m_N(z)}\right]^{-1},$$
(2.1)

for each $z \in \mathbb{C}\backslash\mathbb{R}^+$. In the following, we denote by $\mathcal{S}_N$ the support of μ_N. As $\hat{\mu}_N - \mu_N \to_N 0$ weakly almost surely, it holds that

$$\hat{m}_N(z) - m_N(z) \to 0$$
(2.2)

almost surely for each $z \in \mathbb{C}\backslash\mathbb{R}^+$. The following result will be of help.

Lemma 2.1 ([12, 4]). *Let $\psi \in \mathcal{C}_c^\infty(\mathbb{R},\mathbb{R})$ and (r_N) a sequence of holomorphic functions on $\mathbb{C}\backslash\mathbb{R}$ such that*

$$|r_N(z)| \leq \text{P}_1(|z|)\text{P}_2\left(\frac{1}{|\text{Im}(z)|}\right),$$

with P_1 and P_2 two polynomials with positive coefficients, independent of N. Then,

$$\limsup_{y\downarrow 0}\left|\int_{\mathbb{R}}\psi(x)r_N(x+iy)\mathrm{d}x\right| \leq C < \infty,$$

with C a constant independent of N.

Taking into account the previous result, it is shown in [23] that

$$\mathbb{E}[\hat{m}_N(z)] = m_N(z) + \frac{r_N(z)}{N^2},$$
(2.3)

with r_N as in Lemma 2.1. Using the inverse Stieltjes transform formula (1.4), we obtain that for each function $\psi \in \mathcal{C}_c^\infty(\mathbb{R},\mathbb{R})$, it holds that

$$\frac{1}{M}\mathbb{E}[\text{Tr}\,\psi(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*)] = \frac{1}{M}\sum_{k=1}^M\mathbb{E}[\psi(\hat{\lambda}_{k,N})] = \int_{\mathcal{S}_N}\psi(\lambda)\mu_N(d\lambda) + \mathcal{O}\left(\frac{1}{N^2}\right).$$
(2.4)

If we denote by $\mathbf{T}_N(z)$ the matrix-valued function defined by

$$\mathbf{T}_N(z) = \left[-z(1+\sigma^2 c_N m_N(z))\mathbf{I}_M + \sigma^2(1-c_N)\mathbf{I}_M + \frac{\mathbf{B}_N\mathbf{B}_N^*}{1+\sigma^2 c_N m_N(z)}\right]^{-1},$$

then $\mathbf{T}_N$ coincides with the Stieltjes transform of a positive matrix-valued measure $\boldsymbol{\mu}_N$ with support $\mathcal{S}_N$ such that $\boldsymbol{\mu}_N(\mathcal{S}_N) = \mathbf{I}_M$ (see Hachem *et al.* [13, Theorem 2.4

and Proposition 2.2]), i.e.

$$\mathbf{T}_N(z) = \int_{\mathcal{S}_N} \frac{d\boldsymbol{\mu}_N(\lambda)}{\lambda - z}.$$

As $m_N(z)$ verifies Eq. (2.1), it is clear that $\frac{1}{M}\text{Tr}\,\boldsymbol{\mu}_N = \mu_N$.

In the remainder of the paper, we will make use of the following result proved in [23] if $\mathbf{W}_N$ is complex Gaussian and in Hachem *et al.* [14] in the non-Gaussian case.

Theorem 2.1. *Consider two sequences of deterministic vectors* $(\mathbf{b}_N), (\mathbf{d}_N)$ *such that* $\sup_N \|\mathbf{b}_N\| < +\infty$ *and* $\sup_N \|\mathbf{d}_N\| < +\infty$. *Then, under Assumptions* A-1, A-2 *and* A-4, *it holds that*

$$\mathbf{b}_N^* \mathbf{Q}_N(z)\mathbf{d}_N - \mathbf{b}_N^* \mathbf{T}_N(z)\mathbf{d}_N \underset{N}{\longrightarrow} 0, \tag{2.5}$$

almost surely for each $z \in \mathbb{C}\backslash\mathbb{R}^+$.

2.2. *The characterization of the support* $\mathcal{S}_N$ *of* μ_N

The support $\mathcal{S}_N$ of μ_N was first studied in Dozier and Silverstein [8] and a more convenient characterization was presented in [23]. We first recall (see [8]) that if $z \in \mathbb{C}^+$ converges towards $x \in \mathbb{R}$, then, $m_N(z)$ converges towards a finite limit still denoted $m_N(x)$. Function $x \to m_N(x)$ is continuous on $\mathbb{R}$, continuously differentiable on $\mathbb{R}\backslash\partial\mathcal{S}_N$, and verifies Eq. (2.1) on $\mathbb{R}\backslash\partial\mathcal{S}_N$. Moreover, μ_N is absolutely continuous and its density coincides with function $\frac{1}{\pi}\text{Im}(m_N(x))$.

In order to present the characterization of $\mathcal{S}_N$, we first introduce the following notations. We denote by f_N, ϕ_N and w_N the functions defined by

$$f_N(w) = \frac{1}{M}\text{Tr}(\mathbf{B}_N\mathbf{B}_N^* - w\mathbf{I}_M)^{-1}, \tag{2.6}$$

$$\phi_N(w) = w(1 - \sigma^2 c_N f_N(w))^2 + \sigma^2(1 - c_N)(1 - \sigma^2 c_N f_N(w)), \tag{2.7}$$

$$w_N(z) = z(1 + \sigma^2 c_N m_N(z))^2 - \sigma^2(1 - c_N)(1 + \sigma^2 c_N m_N(z)). \tag{2.8}$$

We are now in position to characterize $\mathcal{S}_N$.

Theorem 2.2. *Assume Assumptions* A-1–A-4 *hold. The function* ϕ_N *admits* $2Q$ *non-negative local extrema counting multiplicities (with* $1 \leq Q \leq K+1$*) whose preimages are denoted* $w_{1,N}^- < 0 < w_{1,N}^+ \leq w_{2,N}^- \cdots \leq w_{Q,N}^- < w_{Q,N}^+$. *Define* $x_{q,N}^- = \phi_N(w_{q,N}^-)$ *and* $x_{q,N}^+ = \phi_N(w_{q,N}^+)$ *for* $q = 1, \ldots, Q$. *Then,*

$$x_{1,N}^- < x_{1,N}^+ < x_{2,N}^- < \cdots < x_{Q,N}^- < x_{Q,N}^+,$$

and the support $\mathcal{S}_N$ *of* μ_N *is given by*

$$\mathcal{S}_N = \bigcup_{q=1}^{Q} [x_{q,N}^-, x_{q,N}^+].$$

Moreover, for $q = 1, \ldots, Q$, each interval $]w_{q,N}^-, w_{q,N}^+[$ contains at least an element of the set $\{0, \lambda_{M-K+1,N}, \ldots, \lambda_{M,N}\}$ and each eigenvalue of $\mathbf{B}_N\mathbf{B}_N^*$ belongs to one of these intervals.

The second statement of the theorem shows that each eigenvalue of $\mathbf{B}_N\mathbf{B}_N^*$ corresponds to a certain interval of $\mathcal{S}_N$. More precisely, an eigenvalue of $\mathbf{B}_N\mathbf{B}_N^*$ will be said to be associated to cluster $[x_{q,N}^-, x_{q,N}^+]$ if it belongs to the interval $(w_{q,N}^-, w_{q,N}^+)$. We note that the eigenvalue 0 is necessarily associated to the first cluster $[x_{1,N}^-, x_{1,N}^+]$.

We finally recall the useful properties of function w_N defined by (2.8) (see [23]). We still denote by $w_N(x)$ the limit of $w_N(z)$ when $z \in \mathbb{C}^+$ converges towards $x \in \mathbb{R}$.

Proposition 2.1. *Under Assumptions* A-1–A-4, *function* $w_N : \mathbb{C} \to \mathbb{C}$ *satisfies the following properties*:

- *Function* $x \to w_N(x)$ *is continuous on* $\mathbb{R}$ *and continuously differentiable on* $\mathbb{R}\backslash\partial\mathcal{S}_N$,
- $\mathrm{Im}(w_N(z)) > 0$ *if* $\mathrm{Im}(z) > 0$,
- w_N *is real and strictly increasing on* $\mathbb{R}\backslash\mathcal{S}_N$,
- $w_N(x_{q,N}^-) = w_{q,N}^-$ *and* $w_N(x_{q,N}^+) = w_{q,N}^+$ *for each* $1 \leq q \leq Q$,
- $\mathrm{Im}(w_N(x)) > 0$ *if and only if* $x \in \mathrm{Int}(\mathcal{S}_N)$.

2.3. *Some useful evaluations*

In this paragraph, we gather some useful bounds related to certain Stieltjes transforms. We first recall that the inequality

$$|1 + \sigma^2 c_N m_N(z)| \geq \mathrm{Re}(1 + \sigma^2 c_N m_N(z)) \geq 1/2 \tag{2.9}$$

holds for $z \in \mathbb{C}$ (see Loubaton and Vallet [17]). We now consider function $z \to \frac{-1}{z(1+\sigma^2 c_N m_N(z))}$. [13, Proposition 2.2] implies that it coincides with the Stieltjes transform of a probability measure carried by $\mathbb{R}^+$. Moreover, (2.9) shows that the support of this measure is included in $\mathcal{S}_N \cup \{0\}$. Therefore, we obtain that

$$\frac{1}{|1 + \sigma^2 c_N m_N(z)|} \leq \frac{|z|}{|\mathrm{Im}(z)|} \tag{2.10}$$

for each $z \in \mathbb{C}\backslash\mathbb{R}$ as well as

$$\frac{1}{|1 + \sigma^2 c_N m_N(z)|} \leq \frac{|z|}{\mathrm{dist}(z, \mathcal{S}_N)}$$

for each $z \in \mathbb{C}^*\backslash\mathcal{S}_N$. We also recall that matrix $\mathbf{T}_N(z)$ satisfies $\mathbf{T}_N(z)\mathbf{T}_N(z)^* \leq \frac{\mathbf{I}_M}{\mathrm{Im}(z)^2}$ for $z \in \mathbb{C}^+$ (see [13, Proposition 5.1]). We now claim that the inequality

$$\mathbf{T}_N(z)\mathbf{T}_N(z)^* \leq \frac{\mathbf{I}_M}{\mathrm{dist}(z, \mathcal{S}_N)^2} \tag{2.11}$$

also holds on $\mathbb{C}\backslash\mathcal{S}_N$. In order to establish (2.11), we follow the proof of [13, Proposition 5.1]. We first remark that function $\tilde{m}_N(z)$ defined by

$$\tilde{m}_N(z) = c_N m_N(z) - \frac{1 - c_N}{z}$$

is the Stieltjes transform of probability measure $\tilde{\mu}_N = c_N \mu_N + (1 - c_N)\delta_0$. The support of $\tilde{\mu}_N$ thus coincides with $\mathcal{S}_N \cup \{0\}$, and is included in $\mathbb{R}^+$. Therefore, it holds that $\frac{\text{Im}(z\tilde{m}_N(z))}{\text{Im}(z)} > 0$ if $z \in \mathbb{C}\backslash\mathbb{R}$. We remark that

$$\frac{\mathbf{T}_N(z) - \mathbf{T}_N(z)^*}{2i} = \text{Im}(z) \int_{\mathcal{S}_N} \frac{d\boldsymbol{\mu}_N(\lambda)}{|\lambda - z|^2}.$$

By using the identity, $\mathbf{T}_N(z) - \mathbf{T}_N(z)^* = \mathbf{T}_N(z)(\mathbf{T}_N(z)^{-*} - \mathbf{T}_N(z)^{-1})\mathbf{T}_N(z)^*$, we get after some algebra

$$\text{Im}(z) \int_{\mathcal{S}_N} \frac{d\boldsymbol{\mu}_N(\lambda)}{|\lambda - z|^2} = \text{Im}(z)\mathbf{T}_N(z)\mathbf{T}_N(z)^* + \sigma^2 \text{Im}(z\tilde{m}_N(z))\mathbf{T}_N(z)\mathbf{T}_N(z)^*$$

$$+ \frac{\sigma^2 c_N}{|1 + \sigma^2 c_N m_N(z)|^2}\text{Im}(m_N(z))\mathbf{T}_N(z)\mathbf{B}_N\mathbf{B}_N^*\mathbf{T}_N(z)^*,$$

for each $z \in \mathbb{C}\backslash\mathbb{R}$, or equivalently

$$\int_{\mathcal{S}_N} \frac{d\boldsymbol{\mu}_N(\lambda)}{|\lambda - z|^2} = \mathbf{T}_N(z)\mathbf{T}_N(z)^* + \sigma^2 \frac{\text{Im}(z\tilde{m}_N(z))}{\text{Im}(z)}\mathbf{T}_N(z)\mathbf{T}_N(z)^*$$

$$+ \frac{\sigma^2 c_N}{|1 + \sigma^2 c_N m_N(z)|^2}\frac{\text{Im}(m_N(z))}{\text{Im}(z)}\mathbf{T}_N(z)\mathbf{B}_N\mathbf{B}_N^*\mathbf{T}_N(z)^*.$$

Consequently, we obtain that

$$\mathbf{T}_N(z)\mathbf{T}_N(z)^* \leq \int_{\mathcal{S}_N} \frac{d\boldsymbol{\mu}_N(\lambda)}{|\lambda - z|^2}$$

for $z \in \mathbb{C}\backslash\mathbb{R}$, but also for $z \in \mathbb{C}\backslash\mathcal{S}_N$ because both members of above inequality are continuous on $\mathbb{C}\backslash\mathcal{S}_N$. This immediately leads to (2.11). This inequality also implies that for each $z \in \mathbb{C}\backslash\mathcal{S}_N$,

$$\min_{k=1,\ldots,M} |\lambda_{k,N} - w_N(z)| \geq \frac{1}{2}\text{dist}(z, \mathcal{S}_N). \tag{2.12}$$

Indeed, $\mathbf{T}_N(z)$ can be written as

$$\mathbf{T}_N(z) = (1 + \sigma^2 c_N m_N(z))(\mathbf{B}_N\mathbf{B}_N^* - w_N(z)\mathbf{I}_M)^{-1}.$$

Therefore, $\|\mathbf{T}_N(z)\|$ is equal to

$$\|\mathbf{T}_N(z)\| = \frac{|1 + \sigma^2 c_N m_N(z)|}{\min_{k=1,\ldots,M} |\lambda_{k,N} - w_N(z)|}$$

so that (2.12) follows from (2.11) and (2.9).

Since $\hat{m}_N(z)$ is the Stieltjes transform of the distribution $\frac{1}{M}\sum_{k=1}^{M}\delta_{\hat{\lambda}_{k,N}}$, it holds that

$$|\hat{m}_N(z)| \leq \frac{1}{\text{dist}(z, \{\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M,N}\})},$$

as well as

$$|\hat{m}'_N(z)| \leq \frac{1}{\text{dist}(z, \{\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M,N}\})^2}.$$

We now consider the rational function $z \mapsto \frac{1}{1+\sigma^2 c_N \hat{m}_N(z)}$ which will play an important role in the following. Its poles are solutions of the equation $1+\sigma^2 c_N \hat{m}_N(z) = 0$, and satisfy some useful properties. From now on, we denote by $\hat{\boldsymbol{\Lambda}}_N$ the diagonal matrix $\hat{\boldsymbol{\Lambda}}_N = \text{Diag}(\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M,N})$ and by $\hat{\boldsymbol{\Omega}}_N$ the matrix

$$\hat{\boldsymbol{\Omega}}_N = \hat{\boldsymbol{\Lambda}}_N + \frac{\sigma^2 c_N}{M}\mathbf{1}\mathbf{1}^T, \tag{2.13}$$

where $\mathbf{1}$ denotes vector $\mathbf{1} = (1, 1, \ldots, 1)^T$. We denote $\hat{\omega}_{1,N} \leq \cdots \leq \hat{\omega}_{M,N}$ its eigenvalues. Then we have the following straightforward properties.

- The zeros of $z \mapsto 1 + \sigma^2 c_N \hat{m}_N(z)$ are included in the set $\{\hat{\omega}_{1,N}, \ldots, \hat{\omega}_{M,N}\}$.
- If the eigenvalues $\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M,N}$ of $\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*$ have multiplicity 1, the equation $1 + \sigma^2 c_N \hat{m}_N(z) = 0$ has M multiplicity 1 solutions which coincide with the $(\hat{\omega}_{k,N})_{k=1,\ldots,M}$. Moreover, $\hat{\lambda}_{1,N} < \hat{\omega}_{1,N} < \cdots < \hat{\lambda}_{M,N} < \hat{\omega}_{M,N}$.
- If the eigenvalue $\hat{\lambda}_{k,N}$ has multiplicity $p > 1$, i.e.

$$\hat{\lambda}_{k-1,N} < \hat{\lambda}_{k,N} = \hat{\lambda}_{k+p-1,N} < \hat{\lambda}_{k+p,N},$$

then,

$$\hat{\omega}_{k-1,N} < \hat{\lambda}_{k,N} = \hat{\omega}_{k,N} = \cdots = \hat{\lambda}_{k+p-2,N} = \hat{\omega}_{k+p-2,N}$$
$$= \hat{\lambda}_{k+p-1,N} < \hat{\omega}_{k+p-1,N} < \hat{\lambda}_{k+p,N},$$

and the $\hat{\omega}_{k,N}$ that do not coincide with some eigenvalues of $\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*$ are zeros of $1 + \sigma^2 c_N \hat{m}_N(z)$.

Remark 2.1. Since $c_N < 1$, we recall that the eigenvalues $(\hat{\lambda}_{k,N})_{k=1,\ldots,M}$ have multiplicity 1 almost surely. However, in Sec. 3.2, it will be necessary to define properly the solutions of $1 + \sigma^2 c_N \hat{m}_N(z) = 0$ everywhere. This explains why the case where some of the $(\hat{\lambda}_{k,N})_{k=1,\ldots,M}$ are multiple has to be considered.

Function $z \mapsto \frac{-1}{z(1+\sigma^2 c_N \hat{m}_N(z))}$ is the Stieltjes transform of a probability measure whose support coincides with the set of all roots of the equation $z(1+\sigma^2 c_N \hat{m}_N(z)) = 0$, which is included into the set $\{0, \hat{\omega}_{1,N}, \ldots, \hat{\omega}_{M,N}\}$. Therefore, it holds that

$$\frac{1}{|1+\sigma^2 c_N \hat{m}_N(z)|} \leq \frac{|z|}{\text{dist}(z, \{0, \hat{\omega}_{1,N}, \ldots, \hat{\omega}_{M,N}\})}$$

for $z \in \mathbb{C} \backslash \{0, \hat{\omega}_{1,N}, \ldots, \hat{\omega}_{M,N}\}$ and

$$\frac{1}{|1 + \sigma^2 c_N \hat{m}_N(z)|} \leq \frac{|z|}{|\text{Im}(z)|} \tag{2.14}$$

for $z \in \mathbb{C} \backslash \mathbb{R}$. We eventually notice that

$$\|\mathbf{Q}_N(z)\| \leq \frac{1}{\text{dist}(z, \{\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M,N}\})}.$$

2.4. Almost sure localization of the eigenvalues $(\hat{\lambda}_{k,N})_{k=1,\ldots,M}$

We recall the two following useful results of [23, 17].

Theorem 2.3 ([23]). *Assume Assumptions A-1–A-4 hold. Let $a, b \in \mathbb{R}, \epsilon > 0$ and $N_0 \in \mathbb{N}$ such that*

$$]a - \epsilon, b + \epsilon[\cap \mathcal{S}_N = \emptyset,$$

for each $N > N_0$. Then, with probability one, no eigenvalue of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ belongs to $[a, b]$ for N large enough.

Theorem 2.4 ([17]). *Assume Assumptions A-1–A-4 hold. Let $a, b \in \mathbb{R}, \epsilon > 0$, $N_0 \in \mathbb{N}$ such that $]a - \epsilon, b + \epsilon[\cap \mathcal{S}_N = \emptyset$ for $N > N_0$. Then, with probability one,*

$$\text{card}\{k : \hat{\lambda}_{k,N} < a\} = \text{card}\{k : \lambda_{k,N} < w_N(a)\}, \tag{2.15}$$

$$\text{card}\{k : \hat{\lambda}_{k,N} > b\} = \text{card}\{k : \lambda_{k,N} > w_N(b)\}, \tag{2.16}$$

for N large enough.

It is useful to mention that $\sup_N x_{Q_N,N}^+ < +\infty$ and that these two theorems are still valid if $b = +\infty$ (see [17]).

2.5. The consistent estimate of quadratic forms of $\mathbf{\Pi}_N$

Let $\mathbf{\Pi}_N$ be the orthogonal projection matrix on the kernel of $\mathbf{B}_N \mathbf{B}_N^*$ and let $(\mathbf{a}_N)_{N \in \mathbb{N}}$ be a sequence of deterministic M-dimensional vectors such that $\sup_N \|\mathbf{a}_N\| < \infty$. Then, [23] proposed a consistent estimate of η_N defined by

$$\eta_N = \mathbf{a}_N^* \mathbf{\Pi}_N \mathbf{a}_N.$$

The approach of [23] is valid under the following assumptions.

Assumption A-5. For N large enough, none of the strictly positive eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ is associated to the first cluster $[x_{1,N}^-, x_{1,N}^+]$, i.e. $\lambda_{M-K+1,N} > w_N(x_{1,N}^+)$ for N large enough.

Assumption A-6. It holds that

$$0 < \liminf_{N \to +\infty} x_{1,N}^- < \limsup_{N \to +\infty} x_{1,N}^+ < \liminf_{N \to +\infty} x_{2,N}^-.$$

Using Theorems 2.3 and 2.4, we deduce that if $t_1^-, t_1^+, t_2^-, t_2^+$ are real numbers independent of N satisfying

$$0 < t_1^- < \liminf_{N \to +\infty} x_{1,N}^- < \limsup_{N \to +\infty} x_{1,N}^+ < t_1^+ < t_2^-$$

$$< \liminf_{N \to +\infty} x_{2,N}^- \leq \limsup_{N \to +\infty} x_{Q_N,N}^+ < t_2^+ \qquad (2.17)$$

then, almost surely, for N large enough, it holds that

$$0 < t_1^- < \hat{\lambda}_{1,N} < \cdots < \hat{\lambda}_{M-K,N} < t_1^+ < t_2^- < \hat{\lambda}_{M-K+1,N} < \cdots < \hat{\lambda}_{M,N} < t_2^+.$$
$$(2.18)$$

Assumptions A-5 and A-6 thus imply that, almost surely, the smallest $M - K$ eigenvalues of $\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*$ are separated from the K greatest ones for N large enough in the sense that the two sets of eigenvalues are included into two disjoint intervals that do not depend on N. It is interesting to remark that Assumptions A-5 and A-6 are "deterministic conditions" depending only on $\sigma^2, c_N = \frac{M}{N}$, and on the eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$. If K remains fixed, recent results of Benaych–Rao [3] (see also [17]) imply that Assumptions A-5 and A-6 and hold if and only if $\liminf_{N \to +\infty} \lambda_{M-K+1,N} > \sigma^2 \sqrt{c}$. If however K scales with N, the derivation of more explicit conditions equivalent to Assumptions A-5 and A-6 is still an open problem.

We are now in position to present the consistent estimator of η_N proposed in [23]. It is based on the observation that

$$\boldsymbol{\Pi}_N = \frac{1}{2i\pi} \int_{\mathcal{C}^-} (\mathbf{B}_N \mathbf{B}_N^* - \lambda \mathbf{I}_M)^{-1} \mathrm{d}\lambda,$$

where $\mathcal{C}$ represents a contour enclosing 0 and not the strictly positive eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$, and the symbol $\mathcal{C}^-$ means that the contour is oriented clockwise. The estimator of [23] is based on the observation that under Assumptions A-5 and A-6, function $w_N(z)$ provides such a contour for N large enough. In the following, for $y > 0$ and $\epsilon > 0$, $\epsilon < \frac{y}{3}$ small enough, we consider the rectangle $\mathcal{R}_y$ defined by

$$\mathcal{R}_y = \{z = x + iv, 0 < t_1^- - 3\epsilon \leq x \leq t_1^+ + 3\epsilon < t_2^- - 3\epsilon, -y \leq v \leq y\} \quad (2.19)$$

and its boundary $\partial \mathcal{R}_y$. Then, the properties of function $w_N(z)$ (see Proposition 2.1) imply that for N large enough, the set $w_N(\partial \mathcal{R}_y)$ is a contour enclosing the origin, but not the other eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$. Therefore, $\boldsymbol{\Pi}_N$ can also be written as

$$\boldsymbol{\Pi}_N = \frac{1}{2i\pi} \int_{\partial \mathcal{R}_y^-} (\mathbf{B}_N \mathbf{B}_N^* - w_N(z)\mathbf{I}_M)^{-1} w_N'(z) \mathrm{d}z$$

or equivalently

$$\boldsymbol{\Pi}_N = \frac{1}{2i\pi} \int_{\partial \mathcal{R}_y} \mathbf{T}_N(z) \frac{w_N'(z)}{1 + \sigma^2 c_N m_N(z)} \mathrm{d}z \qquad (2.20)$$

because $(\mathbf{B}_N \mathbf{B}_N^* - w_N(z)\mathbf{I}_M)^{-1} = \frac{\mathbf{T}_N(z)}{1 + \sigma^2 c_N m_N(z)}$. Using (2.2) and (2.5) as well as the following lemma

Lemma 2.2. *Under Assumptions* A-1–A-6, *almost surely, for N large enough, the M solutions* $(\hat{\omega}_{k,N})_{k=1,\ldots,M}$ *of the equation* $1 + \sigma^2 c_N m_N(z) = 0$ *satisfy*

$$t_1^- < \hat{\lambda}_{1,N} < \hat{\omega}_{1,N} < \cdots < \hat{\lambda}_{M-K,N} < \hat{\omega}_{M-K,N} < t_1^+ < t_2^-$$

$$< \hat{\lambda}_{M-K+1,N} < \hat{\omega}_{M-K+1,N} < \cdots < \hat{\lambda}_{M,N} < \hat{\omega}_{M,N} < t_2^+.$$

It is showed in [23] that matrix $\tilde{\mathbf{\Pi}}_N$ defined by

$$\tilde{\mathbf{\Pi}}_N = \frac{1}{2i\pi} \int_{\partial \mathcal{R}_y} \mathbf{Q}_N(z) \frac{\hat{w}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} \mathrm{d}z \qquad (2.21)$$

where $\hat{w}_N(z) = z(1 + \sigma^2 c_N \hat{m}_N(z))^2 - \sigma^2(1 - c_N)(1 + \sigma^2 c_N \hat{m}_N(z))$, satisfies $\mathbf{a}_N^* \tilde{\mathbf{\Pi}}_N \mathbf{a}_N - \mathbf{a}_N^* \mathbf{\Pi}_N \mathbf{a}_N \to 0$ almost surely. We note that the poles of the integrand of the right-hand side of (2.21) coincide with the set $\{\hat{\lambda}_{k,N}, \hat{\omega}_{k,N} : k = 1, \ldots, M\}$, which by (2.18) and Lemma 2.2, verifies

$$\text{dist}(\partial \mathcal{R}_y, \{(\hat{\lambda}_{k,N}, \hat{\omega}_{k,N})_{k=1,\ldots,M}\}) > 3\epsilon \qquad (2.22)$$

almost surely for N large enough. In practice, the above estimator is quite easy to implement because, as the localization of the poles of the integrand in (2.21) with respect to the contour $\partial \mathcal{R}_y$ is known (see Lemma 2.2), the contour integral in (2.21) can be solved, and expressed in closed form in terms of the $(\hat{\mathbf{u}}_{k,N}, \hat{\lambda}_{k,N}, \hat{\omega}_{k,N})_{k=1,\ldots,M}$.

3. Statement and Proof of the Uniform Consistency of Estimate $\tilde{\eta}_N(\theta)$

From now on, we assume that vector $\mathbf{a}(\theta)$ is given by (1.2) and that Assumptions A-5 and A-6 hold. We consider t_1^-, t_1^+, t_2^- and t_2^+ satisfying (2.17) as well a rectangle $\mathcal{R}_y$ defined by (2.19). We prove here the following result.

Theorem 3.1. *Under Assumptions* A-1–A-6, *it holds*

$$\sup_{\theta \in [-\pi,\pi]} |\tilde{\eta}_N(\theta) - \eta_N(\theta)| \xrightarrow[N\to\infty]{\text{a.s.}} 0.$$

In order to prove Theorem 3.1, we show that it is sufficient to establish that for each $\alpha > 0$ and for each $\theta \in [-\pi, \pi]$, $\mathbb{P}(|\mathbf{a}(\theta)^*(\tilde{\mathbf{\Pi}}_N - \mathbf{\Pi}_N)\mathbf{a}(\theta)| > \alpha)$ decreases fast enough towards 0. For this, a tempting choice is to use the Markov inequality, and to establish that the moments of $\mathbf{a}(\theta)^*(\tilde{\mathbf{\Pi}}_N - \mathbf{\Pi}_N)\mathbf{a}(\theta)$ decrease fast enough. However, the observation that (2.22) holds for N greater than a random integer does not necessarily imply the existence of the moments of $\mathbf{a}(\theta)^* \tilde{\mathbf{\Pi}}_N \mathbf{a}(\theta)$. In order to solve this technical problem, we establish that the probability that at least one element of $\{\hat{\lambda}_{k,N}, \hat{\omega}_{k,N} : k = 1, \ldots, M\}$ escapes from $[t_1^- - 2\epsilon, t_1^+ + 2\epsilon] \cup [t_2^- - 2\epsilon, t_2^+ + 2\epsilon]$ decreases at rate $\frac{1}{N^l}$ for any $l \in \mathbb{N}$, and prove that the moments of a convenient regularized version of $\mathbf{a}(\theta)^*(\tilde{\mathbf{\Pi}}_N - \mathbf{\Pi}_N)\mathbf{a}(\theta)$ converge fast enough towards 0.

In the following, we denote by $\mathcal{T}_\epsilon$ the set

$$\mathcal{T}_\epsilon = [t_1^- - \epsilon, t_1^+ + \epsilon] \cup [t_2^- - \epsilon, t_2^+ + \epsilon].$$

We first establish in Secs. 3.1 and 3.2 that the events $\mathcal{E}_{1,N}$ and $\mathcal{E}_{2,N}$ defined by

$$\mathcal{E}_{1,N} = \{\text{at least one of the } (\hat{\lambda}_{k,N})_{k=1,\dots,M} \text{ escapes from } \mathcal{T}_\epsilon\}, \tag{3.1}$$

$$\mathcal{E}_{2,N} = \{\text{at least one of the } (\hat{\omega}_{k,N})_{k=1,\dots,M} \text{ escapes from } \mathcal{T}_\epsilon\}, \tag{3.2}$$

verify $\mathbb{P}(\mathcal{E}_{i,N}) = \mathcal{O}(\frac{1}{N^l})$ for each $l \in \mathbb{N}$. Using this result, we introduce in Sec. 3.3 the regularization term, denoted χ_N, defined as follows. We consider a function $\phi \in \mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R}^+)$ satisfying

$$\phi(\lambda) = \begin{cases} 1, & \text{for } \lambda \in \mathcal{T}_\epsilon, \\ 0, & \text{for } \lambda \in \mathbb{R} \backslash ([t_1^- - 2\epsilon, t_1^+ + 2\epsilon] \cup [t_2^- - 2\epsilon, t_2^+ + 2\epsilon]), \end{cases} \tag{3.3}$$

and $\phi(\lambda) \in (0,1)$ elsewhere, and define the random variable

$$\chi_N = \det \phi(\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*) \det \phi(\hat{\boldsymbol{\Omega}}_N), \tag{3.4}$$

which verifies $\mathbb{1}_{\mathcal{E}_N^c} \leq \chi_N$ where $\mathcal{E}_N = \mathcal{E}_{1,N} \cup \mathcal{E}_{2,N}$. We will prove that, considered as a function of the real and imaginary part of the entries of $\mathbf{W}_N$, χ_N is a $\mathcal{C}^1$ function, and using Poincaré inequality, we will establish that

$$\mathbb{E}|\mathbf{a}(\theta)^*(\tilde{\boldsymbol{\Pi}}_N - \boldsymbol{\Pi}_N)\mathbf{a}(\theta)\chi_N|^{2l} = \mathcal{O}\left(\frac{1}{N^l}\right),$$

for each integer l. The above-mentioned properties eventually allow to prove the uniform consistency of estimator $\tilde{\eta}_N(\theta)$.

3.1. *Evaluation of the escape probability of $(\hat{\lambda}_{k,N})_{k=1,\dots,N}$*

The purpose of this section is to prove the following technical result.

Proposition 3.1. *Under Assumptions A-1–A-6, for each $l \in \mathbb{N}$, it holds that*

$$\mathbb{P}(\mathcal{E}_{1,N}) = \mathcal{O}\left(\frac{1}{N^l}\right).$$

To prove this result, we consider a function $\psi_0 \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^+)$ such that

$$\psi_0(\lambda) = \begin{cases} 1, & \text{for } \lambda \in \mathcal{T}_\epsilon^c, \\ 0, & \text{for } \lambda \in [t_1^-, t_1^+] \cup [t_2^-, t_2^+] \end{cases} \tag{3.5}$$

and $\psi_0(\lambda) \in (0,1)$ elsewhere. From this definition, we clearly have

$$\mathbb{P}(\mathcal{E}_{1,N}) \leq \mathbb{P}(\text{Tr}\, \psi_0(\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*) \geq 1) \leq \mathbb{E}[(\text{Tr}\, \psi_0(\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*))^{2l}]$$

for $l \in \mathbb{N}$. In order to establish Proposition 3.1, it is therefore sufficient to prove that $\mathbb{E}[(\text{Tr}\, \psi_0(\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*))^{2l}] = \mathcal{O}(\frac{1}{N^{2l}})$ for each integer l which is the object of the next lemma.

Lemma 3.1. *Assume Assumptions* A-1–A-6 *hold. Then, for all function* $\psi \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ *constant over the complementary of a compact interval and which vanishes on the support* $\mathcal{S}_N$ *of* μ_N *for all* N *large enough, it holds that*

$$\mathbb{E}[(\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^{2l}] = \mathcal{O}\left(\frac{1}{N^{2l}}\right) \tag{3.6}$$

for each $l \in \mathbb{N}$.

Proof. We prove Lemma 3.1 by induction on l. We first consider the case $l = 1$, and consider a function ψ as above, and denote by C the constant value taken by ψ over the complementary of a certain compact interval. We follow [12] and write ψ as $\psi = \tilde{\psi} + C$, where $\tilde{\psi} \in \mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R})$, and verifies $\tilde{\psi} = -C$ over $\mathcal{S}_N$ for N large enough. Using the technique developed in [12] based on (2.4) and Poincaré inequality, we have

$$\operatorname{Var}[\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)] = \operatorname{Var}[\operatorname{Tr}\tilde{\psi}(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)] = \mathcal{O}\left(\frac{1}{N^2}\right),$$

$$\mathbb{E}[\operatorname{Tr}\tilde{\psi}(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)] = M \int_{\mathbb{R}} \tilde{\psi}(\lambda)\mathrm{d}\mu_N(\lambda) + \mathcal{O}\left(\frac{1}{N}\right) = -MC + \mathcal{O}\left(\frac{1}{N}\right).$$

As $\mathbb{E}[\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)] = CM + \mathbb{E}[\operatorname{Tr}\tilde{\psi}(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)]$, this leads to $\mathbb{E}[\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)] = \mathcal{O}(\frac{1}{N})$. As

$$\mathbb{E}[(\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^2] = (\mathbb{E}[\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)])^2 + \operatorname{Var}[\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)] \tag{3.7}$$

we finally obtain that (3.6) holds for $l = 1$.

We now assume that (3.6) holds until the order $l - 1$ for each function of $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ vanishing on $\mathcal{S}_N$ for N large enough and constant over the complementary of a compact interval. We consider such a function ψ and evaluate the behavior of the $2l$th-order moment of $\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)$. We have

$$\mathbb{E}[(\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^{2l}] = (\mathbb{E}[(\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^l])^2 + \operatorname{Var}[(\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^l]. \tag{3.8}$$

The first term of the right-hand side of (3.8) can be upperbounded as follows

$$(\mathbb{E}[(\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^l])^2 \leq \mathbb{E}[(\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^2]\mathbb{E}[(\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^{2(l-1)}]$$

$$= \mathcal{O}\left(\frac{1}{N^{2l}}\right),$$

using that (3.6) holds until the order $l - 1$. The second term of the right-hand side of (3.8) can be evaluated using the Poincaré inequality. Using that the partial derivative of $\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)$ with respect to $W_{i,j,N}$ and $\overline{W}_{i,j,N}$ are equal respectively to $\mathbf{e}_j^T\boldsymbol{\Sigma}_N^*\psi'(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)\mathbf{e}_i$ and $\mathbf{e}_i^T\psi'(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)\boldsymbol{\Sigma}_N\mathbf{e}_j$, we immediately obtain that

$$\operatorname{Var}[(\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^l] \leq C\mathbb{E}\left[\frac{1}{N}\operatorname{Tr}(\psi'(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)^2\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)(\operatorname{Tr}\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^{2l-2}\right].$$

Using Hölder's inequality, we get immediately that

$$
\mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\,(\psi'(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)^2\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)(\mathrm{Tr}\,\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^{2l-2}\right]
$$

$$
\leq C\left(\mathbb{E}\left|\frac{1}{N}\mathrm{Tr}\,(\psi'(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)^2\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)\right|^l\right)^{\frac{1}{l}}(\mathbb{E}[(\mathrm{Tr}\,\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^{2l}])^{\frac{l-1}{l}}. \quad (3.9)
$$

Since the function $\lambda \to \psi'(\lambda)^2\lambda$ belongs to $\mathcal{C}_c^\infty(\mathbb{R},\mathbb{R})$ and has a support disjoint from $\mathcal{S}_N$ for N large enough, it holds that

$$
\mathbb{E}\left|\frac{1}{N}\mathrm{Tr}\,([\psi'(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)]^2\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)\right|^l
$$

$$
\leq \sqrt{\mathbb{E}\left|\frac{1}{N}\mathrm{Tr}\,([\psi'(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)]^2\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)\right|^2}\sqrt{\mathbb{E}\left|\frac{1}{N}\mathrm{Tr}\,([\psi'(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)]^2\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)\right|^{2(l-1)}}
$$

$$
= \mathcal{O}\left(\frac{1}{N^{2l}}\right).
$$

Plugging the previous estimates into (3.9), we get

$$
\mathrm{Var}[(\mathrm{Tr}\,\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^l] \leq \frac{C}{N^2}(\mathbb{E}[(\mathrm{Tr}\,\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^{2l}])^{\frac{l-1}{l}}.
$$

Define $x_N = \mathbb{E}[(\mathrm{Tr}\,\psi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*))^{2l}]$ and $u_N = N^{2l}x_N$. From (3.8), we have the inequalities $x_N \leq \frac{C_1}{N^2}x_N^{\frac{l-1}{l}} + \frac{C_2}{N^{2l}}$ and $u_N \leq C_1 u_N^{\frac{l-1}{l}} + C_2$. We claim that the sequence (u_N) is bounded. If this is not the case, it exists a subsequence u_{k_N} extracted from u_N which converges towards $+\infty$. However, the inequality $\frac{C_1}{u_{k_N}^{1/l}} + \frac{C_2}{u_{k_N}} \geq 1$ must hold for N large enough. As $u_{k_N} \to +\infty$, this leads to a contradiction. Therefore, u_N is bounded and $x_N \leq \frac{C}{N^{2l}}$ for N large enough. This proves Lemma 3.1. $\qquad\square$

3.2. Evaluation of the escape probability of the $(\hat{\omega}_{k,N})_{k=1,...,N}$

In this section, we will prove the following result.

Proposition 3.2. *Under Assumptions* A-1–A-6, *for each $l \in \mathbb{N}$, it holds that*

$$
\mathbb{P}(\mathcal{E}_{2,N}) = \mathcal{O}\left(\frac{1}{N^l}\right).
$$

We follow the same approach than in Sec. 3.1 and first prove that the $(\hat{\omega}_{k,N})_{k=1,...,M}$ satisfy a property similar to (2.4). For this, we study the behavior of the Stieltjes transform $\hat{n}_N(z)$ of the distribution $\frac{1}{M}\sum_{k=1}^M \delta_{\hat{\omega}_{k,N}}$ defined by

$$
\hat{n}_N(z) = \frac{1}{M}\mathrm{Tr}(\hat{\boldsymbol{\Omega}}_N - z\mathbf{I})^{-1}
$$

and use Lemma 2.1 as well as the inverse Stieltjes transform formula (1.4). Our starting point is the following result showing that the empirical eigenvalue distribution of $\hat{\boldsymbol{\Omega}}_N$ is very similar to the distribution of the eigenvalues of $\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*$. The following auxiliary result will be useful.

Lemma 3.2. *Under Assumptions* A-1–A-6, *it holds that*

$$\mathbb{E}[\hat{m}'_N(z)] - m'_N(z) = \frac{t_N(z)}{N^2}, \tag{3.10}$$

where t_N is analytic on $\mathbb{C}\backslash\mathbb{R}$ and can be upper bounded by $\mathrm{P}_1(|z|)\mathrm{P}_2(\frac{1}{|\mathrm{Im}(z)|})$ on $\mathbb{C}\backslash\mathbb{R}$.

Proof. The proof is given in Appendix A.1. $\qquad\square$

We now prove the fundamental following result.

Lemma 3.3. *Under Assumptions* A-1–A-6, *for each $z \in \mathbb{C}\backslash\mathbb{R}$, it holds that*

$$\mathbb{E}[\hat{n}_N(z)] = \int_{\mathcal{S}_N} \frac{d\mu_N(\lambda)}{\lambda - z} - \frac{1}{M}\int_{\mathcal{S}_N} \frac{d\kappa_N(\lambda)}{\lambda - z} + \frac{r_N(z)}{N^2},$$

with κ_N a finite signed measure carried by $\mathcal{S}_N$ such that $\kappa_N([x^-_{q,N}, x^+_{q,N}]) = 0$ for $q = 1, \ldots, Q_N$, and r_N a holomorphic function on $\mathbb{C}\backslash\mathbb{R}$ satisfying

$$|r_N(z)| \leq \mathrm{P}_1(|z|)\mathrm{P}_2\left(\frac{1}{|\mathrm{Im}(z)|}\right),$$

with P_1, P_2 two polynomials with positive coefficients independent of N.

Proof. Using that $\hat{\boldsymbol{\Omega}}_N$ is a rank 1 perturbation of $\hat{\boldsymbol{\Lambda}}_N$, we obtain immediately that

$$\hat{n}_N(z) = \hat{m}_N(z) - \frac{1}{M}\frac{\sigma^2 c_N \hat{m}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)}.$$

Therefore, for $z \in \mathbb{C}\backslash\mathbb{R}$, it holds that

$$\mathbb{E}[\hat{n}_N(z)] = \mathbb{E}[\hat{m}_N(z)] - \frac{1}{M}\mathbb{E}\left[\frac{\sigma^2 c_N \hat{m}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)}\right]. \tag{3.11}$$

We first establish that

$$\mathbb{E}\left[\frac{\sigma^2 c_N \hat{m}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)}\right] = \frac{\sigma^2 c_N m'_N(z)}{1 + \sigma^2 c_N m_N(z)} + \frac{c_N}{N}r_N(z), \tag{3.12}$$

where $r_N(z)$ is holomorphic on $\mathbb{C}\backslash\mathbb{R}$ and satisfies $|r_N(z)| \leq \mathrm{P}_1(|z|)\mathrm{P}_2(\frac{1}{|\mathrm{Im}(z)|})$. For this, we write

$$\frac{\sigma^2 c_N \hat{m}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} - \frac{\sigma^2 c_N m'_N(z)}{1 + \sigma^2 c_N m_N(z)}$$

$$= \frac{\sigma^2 c_N(\hat{m}'_N(z) - m'_N(z))}{(1 + \sigma^2 c_N \hat{m}_N(z))(1 + \sigma^2 c_N m_N(z))}$$

$$+ \frac{(\sigma^2 c_N)^2(m_N(z)(\hat{m}'_N(z) - m'_N(z)) + m'_N(z)(m_N(z) - \hat{m}_N(z)))}{(1 + \sigma^2 c_N \hat{m}_N(z))(1 + \sigma^2 c_N m_N(z))}. \tag{3.13}$$

In order to study the expectation of this expression, we use (2.10) and (2.14). Moreover, (2.3) and a straightforward application of the Poincaré inequality to

$\hat{m}_N(z)$ considered for z fixed as a function of the entries of $\mathbf{W}_N$ leads immediately to

$$\mathbb{E}|\hat{m}_N(z) - m_N(z)|^2 \leq \frac{1}{N^2}\mathrm{P}_1(|z|)\mathrm{P}_2\left(\frac{1}{|\mathrm{Im}(z)|}\right),$$

for some polynomials P_1, P_2 with positive coefficients and independent of N. Therefore,

$$\mathbb{E}|\hat{m}_N(z) - m_N(z)| \leq \frac{1}{N}\left(\mathrm{P}_1(|z|) + \mathrm{P}_2\left(\frac{1}{|\mathrm{Im}(z)|}\right)\right).$$

Applying also Poincaré inequality to bound $\mathrm{Var}[\hat{m}'_N(z)]$, together with Lemma 3.2, we get

$$\mathbb{E}|\hat{m}'_N(z) - m'_N(z)|^2 \leq \frac{1}{N^2}\mathrm{P}_1(|z|)\mathrm{P}_2\left(\frac{1}{|\mathrm{Im}(z)|}\right).$$

Therefore, it holds that

$$\mathbb{E}|\hat{m}'_N(z) - m'_N(z)| \leq \frac{1}{N}\left(\mathrm{P}_1(|z|) + \mathrm{P}_2\left(\frac{1}{|\mathrm{Im}(z)|}\right)\right).$$

Using $|m_N(z)| \leq \frac{1}{|\mathrm{Im}(z)|}$, $|m'_N(z)| \leq \frac{1}{|\mathrm{Im}(z)|^2}$, as well as (2.10) and (2.14), we eventually get from (3.13) that

$$\mathbb{E}\left|\frac{\sigma^2 c_N \hat{m}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} - \frac{\sigma^2 c_N m'_N(z)}{1 + \sigma^2 c_N m_N(z)}\right| \leq \frac{1}{N}\mathrm{P}_1(|z|)\mathrm{P}_2\left(\frac{1}{|\mathrm{Im}(z)|}\right).$$

This immediately implies (3.12). Now define the function $h_N(z)$ by

$$h_N(z) = \frac{\sigma^2 c_N m'_N(z)}{1 + \sigma^2 c_N m_N(z)}.$$

This function coincides with the Stieltjes transform of a signed measure κ_N satisfying the conditions of Lemma 3.3: Using (2.9), we obtain that $|h_N(z)| \leq 2\sigma^2 c_N |m'_N(z)|$. As $|m'_N(z)| \leq \frac{1}{\mathrm{dist}(z,\mathcal{K})^2}$ where $\mathcal{K}$ is a compact containing $\mathcal{S}_N$, it holds that $|h_N(z)| \leq C\frac{1}{\mathrm{dist}(z,\mathcal{K})^2}$. Using [5, Theorem 4.3], we obtain that $h_N(z)$ is the Stieltjes transform of a finite signed measure κ_N, the support of which is the set of singular points of $h_N(z)$, i.e. $\mathcal{S}_N$. In order to evaluate $\kappa_N([x_{q,N}^-, x_{q,N}^+])$, we use the inverse Stieltjes transform formula,

$$\kappa_N([x_{q,N}^-, x_{q,N}^+]) = \frac{1}{\pi}\lim_{y \downarrow 0}\mathrm{Im}\left(\int_{[x_{q,N}^-, x_{q,N}^+]} h_N(x + iy)\mathrm{d}x\right). \tag{3.14}$$

It is clear that $h_N(x + iy) = \frac{\partial \log(1 + \sigma^2 c_N m_N(x + iy))}{\partial x}$, where the complex logarithm corresponds to the principal determination defined on $\mathbb{C}\backslash\mathbb{R}^-$. We note that (2.9) justifies the use of the principal determination. Therefore,

$$\int_{[x_{q,N}^-, x_{q,N}^+]} h_N(x + iy)\mathrm{d}x = \log(1 + \sigma^2 c_N m_N(x_{q,N}^+ + iy))$$

$$- \log(1 + \sigma^2 c_N m_N(x_{q,N}^- + iy)). \tag{3.15}$$

As $x_{q,N}^{\pm} \in \partial \mathcal{S}_N$, then $m_N(x_{q,N}^{\pm}) \in \mathbb{R}$. Moreover, (2.9) implies that $1 + \sigma^2 c_N m_N(x_{q,N}^{\pm}) > 0$. When $y \to 0$, the right-hand side of (3.15) converges towards $\log(1 + \sigma^2 c_N m_N(x_{q,N}^{+})) - \log(1 + \sigma^2 c_N m_N(x_{q,N}^{-}))$, a real quantity. This shows that $\kappa_N([x_{q,N}^{-}, x_{q,N}^{+}]) = 0$. Consequently,

$$\mathbb{E}\left[\frac{\sigma^2 c_N \hat{m}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)}\right] = \int_{\mathcal{S}_N} \frac{\mathrm{d}\kappa_N(\lambda)}{\lambda - z} + \frac{c_N r_N(z)}{N},$$

where $r_N(z)$ is holomorphic on $\mathbb{C}\backslash\mathbb{R}$ such that $|r_N(z)| \leq \mathrm{P}_1(|z|)\mathrm{P}_2(\frac{1}{|\mathrm{Im}(z)|})$. Lemma 3.3 follows immediately from (3.11). $\square$

We now handle the proof of Proposition 3.2. Although certain steps of the present proof are similar to the proof of Proposition 3.1, more work is needed because matrix $\hat{\boldsymbol{\Omega}}_N$ considered as a function of the entries of $\mathbf{W}_N$ is more complicated than $\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*$. We still consider function $\psi_0 \in \mathcal{C}^{\infty}(\mathbb{R}, \mathbb{R}^+)$ defined by (3.5) and remark that

$$\mathbb{P}(\mathcal{E}_{2,N}) \leq \mathbb{P}(\mathrm{Tr}\,\psi_0(\hat{\boldsymbol{\Omega}}_N) \geq 1) \leq \mathbb{E}[(\mathrm{Tr}\,\psi_0(\hat{\boldsymbol{\Omega}}_N))^{2l}]$$

for $l \in \mathbb{N}$. In order to establish Proposition 3.2, it is therefore sufficient to prove that $\mathbb{E}[(\mathrm{Tr}\,\psi_0(\hat{\boldsymbol{\Omega}}_N))^{2l}] = \mathcal{O}(\frac{1}{N^{2l}})$ for each integer l. For this, we still use the Poincaré inequality. However, in contrast with the context of Proposition 3.1, the entries of $\hat{\boldsymbol{\Omega}}_N$, considered as functions of the real and imaginary parts of the entries of $\mathbf{W}_N$, are not continuously differentiable on $\mathbb{R}^{2MN}$ because function $\mathbf{W}_N \to \hat{\lambda}_{k,N}$ is not differentiable at points for which eigenvalue $\hat{\lambda}_{k,N}$ has multiplicity strictly larger than 1. The use of Poincaré inequality has therefore to be justified carefully. The following useful lemma is proved in the Appendix.

Lemma 3.4. *Assume Assumptions* A-1–A-6 *hold. Let* $\tilde{\psi}$ *be a function of* $\mathcal{C}_c^{\infty}(\mathbb{R}, \mathbb{R})$. *Then,* $\mathrm{Tr}\,\tilde{\psi}(\hat{\boldsymbol{\Omega}}_N)$, *considered as a function of the real and imaginary parts of the entries of* $\mathbf{W}_N$, *is continuously differentiable. Moreover, if the eigenvalues of* $\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*$ *have multiplicity 1, it holds that*

$$\frac{\partial}{\partial W_{i,j,N}}\left\{\frac{1}{M}\mathrm{Tr}(\tilde{\psi}(\hat{\boldsymbol{\Omega}}_N))\right\} = \frac{1}{M}\left[\boldsymbol{\Sigma}_N^* \sum_{l=1}^{M}[\tilde{\psi}'(\hat{\boldsymbol{\Omega}}_N)]_{l,l}\hat{\boldsymbol{\Pi}}_{l,N}\right]_{j,i}, \qquad (3.16)$$

where $\hat{\boldsymbol{\Pi}}_{l,N}$ *represents the orthogonal projection matrix on the one-dimensional eigenspace associated to the eigenvalue* $\hat{\lambda}_{l,N}$ *of* $\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*$.

We will also need that

Lemma 3.5. *Under Assumption* A-4, *for each integer* $p > 0$, *it holds that*

$$\sup_N \mathbb{E}[\|\mathbf{W}_N \mathbf{W}_N^*\|^p] < +\infty,$$

a property also established in the Appendix.

We now prove the following result.

Lemma 3.6. *Assume Assumptions* A-1–A-6 *hold. For all function $\psi \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ constant over the complementary of a compact interval and which vanishes on the support $\mathcal{S}_N$ of μ_N for all N large enough, it holds that*

$$\mathbb{E}[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^{2l}] = \mathcal{O}\left(\frac{1}{N^{2l}}\right) \tag{3.17}$$

for each $l \in \mathbb{N}$.

Proof. As previously, we prove Lemma 3.6 by induction on l. We first consider the case $l = 1$, and consider a function ψ as above, and denote by b the constant value taken by ψ over the complementary of a certain compact interval, and by $\tilde{\psi}$ the function of $\mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R})$ defined by $\tilde{\psi}(\lambda) = \psi(\lambda) - b$, which, of course, is equal to $-b$ on $\mathcal{S}_N$. Using Lemmas 2.1 and 3.3, we obtain

$$\mathbb{E}\left[\frac{1}{M}\operatorname{Tr}\tilde{\psi}(\hat{\boldsymbol{\Omega}}_N)\right] = \int_{\mathcal{S}_N} \tilde{\psi}(\lambda)\mathrm{d}\mu_N(\lambda) - \frac{1}{M}\int_{\mathcal{S}_N}\tilde{\psi}(\lambda)\mathrm{d}\kappa_N(\lambda) + \mathcal{O}\left(\frac{1}{N^2}\right). \tag{3.18}$$

Using that $\kappa_N([x_{q,N}^-, x_{q,N}^+]) = 0$ for each $q = 1, \ldots, Q_N$, we get that $\int_{\mathcal{S}_N} \tilde{\psi}(\lambda)\mathrm{d}\kappa_N(\lambda) = 0$ and that

$$\mathbb{E}\left[\frac{1}{M}\operatorname{Tr}\tilde{\psi}(\hat{\boldsymbol{\Omega}}_N)\right] = -b + \mathcal{O}\left(\frac{1}{N^2}\right).$$

Therefore, it holds that

$$\mathbb{E}\left[\frac{1}{M}\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N)\right] = \mathcal{O}\left(\frac{1}{N^2}\right).$$

Moreover, we prove the following lemma.

Lemma 3.7. *Under Assumptions* A-1–A-6, *it holds that*

$$\operatorname{Var}\left[\frac{1}{M}\operatorname{Tr}(\psi(\hat{\boldsymbol{\Omega}}_N))\right] = \mathcal{O}\left(\frac{1}{N^4}\right). \tag{3.19}$$

Proof. We first note that, considered as a function of $(\operatorname{Re}(W_{i,j,N}), \operatorname{Im}(W_{i,j,N}))_{1 \le i \le M, 1 \le j \le N}$, function $\frac{1}{M}\operatorname{Tr}\tilde{\psi}(\hat{\boldsymbol{\Omega}}_N)$ is continuously differentiable by Lemma 3.4. Therefore, function $\frac{1}{M}\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N)$ is continuously differentiable as well. It is thus possible to use the Poincaré inequality to evaluate the left-hand side of (3.19). Furthermore, as the probability that the eigenvalues $(\hat{\lambda}_{k,N})_{k=1,\ldots,M}$ have multiplicity 1 is equal to 1, it is sufficient to evaluate the partial derivatives of function $\frac{1}{M}\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N)$ when $\mathbf{W}_N$ is such that the $(\hat{\lambda}_{k,N})_{k=1,\ldots,M}$ have multiplicity 1. As the derivative of ψ coincides with $\tilde{\psi}'$, (3.16) and Poincaré inequality lead to

$$\operatorname{Var}\left[\frac{1}{M}\operatorname{Tr}(\psi(\hat{\boldsymbol{\Omega}}_N))\right] \le \frac{C}{N^2}\mathbb{E}\left[\frac{1}{M}\operatorname{Tr}\left(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^* \sum_{l=1}^M |[\psi'(\hat{\boldsymbol{\Omega}}_N)]_{l,l}|^2\hat{\boldsymbol{\Pi}}_{l,N}\right)\right],$$

or equivalently,

$$\text{Var}\left[\frac{1}{M}\text{Tr}(\psi(\hat{\mathbf{\Omega}}_N))\right] \le \frac{C}{N^2}\mathbb{E}\left[\frac{1}{M}\sum_{l=1}^{M}\hat{\lambda}_{l,N}|[\psi'(\hat{\mathbf{\Omega}}_N)]_{l,l}|^2\right].$$

We claim that

$$|[\psi'(\hat{\mathbf{\Omega}}_N)]_{l,l}|^2 \le ([\psi'(\hat{\mathbf{\Omega}}_N)]^2)_{l,l}. \tag{3.20}$$

Indeed, if $(\hat{\mathbf{v}}_{k,N})_{k=1,\dots,M}$ represent the eigenvectors of $\hat{\mathbf{\Omega}}$, then

$$[\psi'(\hat{\mathbf{\Omega}}_N)]_{l,l} = \sum_{k=1}^{M}\psi'(\hat{\omega}_{k,N})|\mathbf{e}_l^T\hat{\mathbf{v}}_{k,N}|^2.$$

As $\sum_{k=1}^{M}|\mathbf{e}_l^T\hat{\mathbf{v}}_{k,N}|^2 = 1$, Jensen's inequality yields to (3.20). Therefore, it holds that

$$\text{Var}\left[\frac{1}{M}\text{Tr}(\psi(\hat{\mathbf{\Omega}}_N))\right] \le \frac{C}{N^2}\mathbb{E}\left[\|\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*\|\frac{1}{M}\sum_{l=1}^{M}\psi'(\hat{\omega}_l)^2\right]. \tag{3.21}$$

As $\sup_N\|\mathbf{B}_N\mathbf{B}_N^*\| < +\infty$, we get using Lemma 3.5 that

$$\sup_N\mathbb{E}[\|\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*\|^p] < +\infty.$$

We remark that $\|\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*\| < t_2^+ + \epsilon$ on the set $\mathcal{E}_{1,N}^c$, and write the right-hand side of (3.21) as

$$\frac{C}{N^2}\mathbb{E}\left[\|\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*\|(\mathbb{1}_{\mathcal{E}_{1,N}} + \mathbb{1}_{\mathcal{E}_{1,N}^c})\frac{1}{M}\sum_{l=1}^{M}\psi'(\hat{\omega}_l)^2\right].$$

It holds that

$$\mathbb{E}\left[\|\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*\|\mathbb{1}_{\mathcal{E}_{1,N}^c}\frac{1}{M}\sum_{l=1}^{M}\psi'(\hat{\omega}_l)^2\right] \le (t_2^+ + \epsilon)\mathbb{E}\left[\frac{1}{M}\text{Tr}\left(\psi'(\hat{\mathbf{\Omega}}_N)\right)^2\right].$$

Function ψ'^2 belongs to $\mathcal{C}_c^\infty(\mathbb{R},\mathbb{R})$ and vanishes on $\mathcal{S}_N$. Therefore, Lemma 3.3 implies that $\mathbb{E}[\frac{1}{M}\text{Tr}\left(\psi'(\hat{\mathbf{\Omega}}_N)\right)^2] = \mathcal{O}(\frac{1}{N^2})$ (see Eq. (3.18)). Moreover, as $\frac{1}{M}\text{Tr}(\psi'(\hat{\mathbf{\Omega}}_N)^2) \le \sup_\lambda\psi'(\lambda)^2 < C$, we have

$$\mathbb{E}\left[\|\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*\|\mathbb{1}_{\mathcal{E}_{1,N}}\frac{1}{M}\sum_{l=1}^{M}\psi'(\hat{\omega}_l)^2\right] < C\mathbb{E}[\|\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*\|\mathbb{1}_{\mathcal{E}_{1,N}}],$$

which is itself upperbounded by

$$C(\mathbb{E}[\|\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*\|^2])^{1/2}\mathbb{P}(\mathcal{E}_{1,N})^{1/2} = \mathcal{O}\left(\frac{1}{N^p}\right),$$

for each integer p. This completes the proof of Lemma 3.7. $\qquad\square$

Assume that (3.17) holds until integer $l-1$. We write as previously that

$$\mathbb{E}[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^{2l}] = (\mathbb{E}[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^l])^2 + \operatorname{Var}[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^l]. \tag{3.22}$$

The Cauchy–Schwarz inequality leads immediately to

$$(\mathbb{E}[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^l])^2 \le \mathbb{E}[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^2]\mathbb{E}[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^{2l-2}]$$

$$= \mathcal{O}\left(\frac{1}{N^{2l}}\right). \tag{3.23}$$

For the second term of the right-hand side of (3.22), we use Poincaré inequality and Hölder's inequality to obtain

$$\operatorname{Var}[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^l] \le C\mathbb{E}\left[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^{2(l-1)}\frac{1}{M}\sum_{k=1}^{M}\hat{\lambda}_{k,N}([\psi'(\hat{\boldsymbol{\Omega}}_N)]_{k,k})^2\right],$$

$$\le C\left(\mathbb{E}\left|\frac{1}{M}\sum_{k=1}^{M}\hat{\lambda}_{k,N}([\psi'(\hat{\boldsymbol{\Omega}}_N)]_{k,k})^2\right|^l\right)^{\frac{1}{l}}(\mathbb{E}[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^{2l}])^{\frac{l-1}{l}}.$$

Jensen's inequality leads again to

$$\mathbb{E}\left|\frac{1}{M}\sum_{k=1}^{M}\hat{\lambda}_{k,N}([\psi'(\hat{\boldsymbol{\Omega}}_N)]_{k,k})^2\right|^l \le \mathbb{E}\left|\|\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\|\frac{1}{M}\operatorname{Tr}\psi'(\hat{\boldsymbol{\Omega}}_N)^2\right|^l.$$

We write again that

$$\mathbb{E}\left|\|\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\|\frac{1}{M}\operatorname{Tr}\psi'(\hat{\boldsymbol{\Omega}}_N)^2\right|^l = \mathbb{E}\left|\|\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\|\frac{1}{M}\operatorname{Tr}\psi'(\hat{\boldsymbol{\Omega}}_N)^2\mathbb{1}_{\mathcal{E}_{1,N}}\right|^l$$

$$+ \mathbb{E}\left|\|\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\|\frac{1}{M}\operatorname{Tr}\psi'(\hat{\boldsymbol{\Omega}}_N)^2\mathbb{1}_{\mathcal{E}_{1,N}^c}\right|^l,$$

and obtain as previously that

$$\mathbb{E}\left|\|\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\|\frac{1}{M}\operatorname{Tr}\psi'(\hat{\boldsymbol{\Omega}}_N)^2\right|^l \le C\left(\mathbb{E}\left|\frac{1}{M}\operatorname{Tr}\psi'(\hat{\boldsymbol{\Omega}}_N)^2\right|^l + (\mathbb{P}(\mathcal{E}_{1,N}))^{1/2}\right).$$

But, applying Cauchy–Schwarz inequality as in (3.23) to $\mathbb{E}|\operatorname{Tr}\psi'(\hat{\boldsymbol{\Omega}}_N)^2|^l$ leads to $\mathbb{E}|\frac{1}{M}\operatorname{Tr}\psi'(\hat{\boldsymbol{\Omega}}_N)^2|^l = \mathcal{O}(\frac{1}{N^{2l}})$. Gathering all the previous inequalities, we find that

$$\mathbb{E}[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^{2l}] \le \frac{C}{N^2}(\mathbb{E}[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^{2l}])^{\frac{l-1}{l}} + \mathcal{O}\left(\frac{1}{N^{2l}}\right),$$

and in the same way as in the proof of Proposition 3.1, we obtain $\mathbb{E}[(\operatorname{Tr}\psi(\hat{\boldsymbol{\Omega}}_N))^{2l}] = \mathcal{O}(\frac{1}{N^{2l}})$. This concludes the proof of Lemma 3.6.

3.3. *End of the proof of Theorem* 3.1

We now complete the proof of Theorem 3.1 when function $\theta \to \mathbf{a}(\theta)$ is given by

$$\mathbf{a}(\theta) = \frac{1}{\sqrt{M}}[1, e^{i\theta}, \ldots, e^{i(M-1)\theta}]^T,$$

for $\theta \in [-\pi, \pi]$. We recall that $\mathcal{E}_N$ is defined by

$$\mathcal{E}_N = \mathcal{E}_{1,N} \cup \mathcal{E}_{2,N},$$

where $(\mathcal{E}_{i,N})_{i=1,2}$ are defined by (3.1) and (3.2), and that $\mathbb{1}_{\mathcal{E}_N^c} \le \chi_N$ where $\chi_N = \det \phi(\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*) \det \phi(\hat{\boldsymbol{\Omega}}_N)$. We first give a useful lemma which appears as a straightforward consequence of the evaluations of Sec. 2.3.

Lemma 3.8. *Under Assumptions* A-1–A-6, *for each* N, *it holds that*

$$\sup_{z \in \partial \mathcal{R}_y} \|\mathbf{T}_N(z)\| \le C,$$

$$\sup_{z \in \partial \mathcal{R}_y} \left| \frac{1}{1 + \sigma^2 c_N m_N(z)} \right| \le C,$$

$$\sup_{z \in \partial \mathcal{R}_y} \left| \frac{w_N'(z)}{1 + \sigma^2 c_N m_N(z)} \right| \le C,$$

and for N *large enough, we have*

$$\sup_{z \in \partial \mathcal{R}_y} \|\mathbf{Q}_N(z)\| \chi_N \le C,$$

$$\sup_{z \in \partial \mathcal{R}_y} \left| \frac{\chi_N}{1 + \sigma^2 c_N \hat{m}_N(z)} \right| \le C,$$

$$\sup_{z \in \partial \mathcal{R}_y} \left| \frac{\hat{w}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} \right| \chi_N \le C.$$

We consider the set

$$\vartheta_N = \left\{ -\pi + \frac{2(k-1)\pi}{N^2} : k = 1, \ldots, N^2 \right\},$$

and remark that for each $\theta \in [-\pi, \pi]$ and for each N, there exists $\theta_N \in \vartheta_N$ such that $|\theta - \theta_N| \le \frac{2\pi}{N^2}$. For each $\theta \in [-\pi, \pi]$, it holds that

$$\tilde{\eta}_N(\theta) - \eta_N(\theta) = [\tilde{\eta}_N(\theta) - \tilde{\eta}_N(\theta_N)] + [\tilde{\eta}_N(\theta_N) - \eta_N(\theta_N)] + [\eta_N(\theta_N) - \eta_N(\theta)]. \tag{3.24}$$

It is easy to check that the third term of the right-hand side of (3.24) satisfies

$$\sup_{\theta \in [-\pi, \pi]} |\eta_N(\theta_N) - \eta_N(\theta)| \le 2 \sup_{\theta \in [-\pi, \pi]} \|\mathbf{a}(\theta) - \mathbf{a}(\theta_N)\| = \mathcal{O}\left(\frac{1}{N}\right). \tag{3.25}$$

In order to evaluate the behavior of the supremum over θ of the first term of the right-hand side of (3.24), we prove that for each $\alpha > 0$,

$$\mathbb{P}\left(\sup_{\theta \in [-\pi,\pi]} |\tilde{\eta}_N(\theta) - \tilde{\eta}_N(\theta_N)| > \alpha\right) = \mathcal{O}\left(\frac{1}{N^{1+\beta}}\right),$$

where $\beta > 0$. We first remark that for each $l \in \mathbb{N}$, it holds that

$$\mathbb{P}\left(\sup_{\theta \in [-\pi,\pi]} |\tilde{\eta}_N(\theta) - \tilde{\eta}_N(\theta_N)| > \alpha\right)$$

$$\leq \mathbb{P}\left(\sup_{\theta \in [-\pi,\pi]} |\tilde{\eta}_N(\theta) - \tilde{\eta}_N(\theta_N)| \mathbb{1}_{\mathcal{E}_N^c} > \alpha\right) + \mathbb{P}(\mathcal{E}_N)$$

$$\leq \frac{1}{\alpha^l} \mathbb{E}\left[\sup_{\theta \in [-\pi,\pi]} |\tilde{\eta}_N(\theta) - \tilde{\eta}_N(\theta_N)|^l \mathbb{1}_{\mathcal{E}_N^c}\right] + \mathcal{O}\left(\frac{1}{N^l}\right).$$

Moreover,

$$|\tilde{\eta}_N(\theta) - \tilde{\eta}_N(\theta_N)|^l \mathbb{1}_{\mathcal{E}_N^c}$$

$$\leq C \oint_{\partial \mathcal{R}_y^-} \left|(\mathbf{a}(\theta) - \mathbf{a}(\theta_N))^* \mathbf{Q}_N(z) \frac{\hat{w}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} \mathbf{a}(\theta_N)\right|^l \mathbb{1}_{\mathcal{E}_N^c} |dz|.$$

Lemma 3.8 and the inequality $\mathbb{1}_{\mathcal{E}_N^c} \leq \chi_N$ imply that

$$\sup_{z \in \partial \mathcal{R}_y} \|\mathbf{Q}_N(z)\| \left|\frac{\hat{w}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)}\right| \mathbb{1}_{\mathcal{E}_N^c} < C \tag{3.26}$$

for some constant term C. Inequality (3.25) thus implies that

$$\sup_{\theta \in [-\pi,\pi]} \left|(\mathbf{a}(\theta) - \mathbf{a}(\theta_N))^* \mathbf{Q}_N(z) \frac{\hat{w}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} \mathbf{a}(\theta_N)\right|^l \mathbb{1}_{\mathcal{E}_N^c} \leq \frac{C}{N^l}$$

thus showing that

$$\mathbb{P}\left(\sup_{\theta \in [-\pi,\pi]} |\tilde{\eta}_N(\theta) - \tilde{\eta}_N(\theta_N)| > \alpha\right) = \mathcal{O}\left(\frac{1}{N^l}\right)$$

for each integer l. Borel–Cantelli's lemma eventually implies that

$$\sup_{\theta \in [-\pi,\pi]} |\tilde{\eta}_N(\theta) - \tilde{\eta}_N(\theta_N)| \to 0$$

almost surely.

We finally study the supremum of the second term of (3.24). We denote by $\nu_{k,N}$ the elements of ϑ_N. Let $\alpha > 0$, then

$$
\mathbb{P}\left(\sup_{\theta \in [-\pi,\pi]} |\tilde{\eta}_N(\theta_N) - \eta_N(\theta_N)| > \alpha\right)
$$

$$
\leq \mathbb{P}\left(\sup_{k=1,\ldots,N^2} |\tilde{\eta}_N(\nu_{k,N}) - \eta_N(\nu_{k,N})| > \alpha\right)
$$

$$
\leq \sum_{k=1}^{N^2} \mathbb{P}(|\tilde{\eta}_N(\nu_{k,N}) - \eta_N(\nu_{k,N})| > \alpha)
$$

$$
\leq \sum_{k=1}^{N^2} [\mathbb{P}(\{|\tilde{\eta}_N(\nu_{k,N}) - \eta_N(\nu_{k,N})| > \alpha\} \cap \mathcal{E}_N^c)] + \mathcal{O}\left(\frac{1}{N^l}\right)
$$

for each integer l. We now introduce in the above term the regularization term $\chi_N = \det \phi(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*) \det \phi(\hat{\mathbf{\Omega}}_N)$ defined in (3.4). As χ_N is equal to 1 on $\mathcal{E}_N^c$, it holds that

$$
\mathbb{P}(\{|\tilde{\eta}_N(\nu_{k,N}) - \eta_N(\nu_{k,N})| > \alpha\} \cap \mathcal{E}_N^c)
$$

$$
= \mathbb{P}(\{|\tilde{\eta}_N(\nu_{k,N}) - \eta_N(\nu_{k,N})|\chi_N^2 > \alpha\} \cap \mathcal{E}_N^c)
$$

$$
\leq \mathbb{P}(|\tilde{\eta}_N(\nu_{k,N}) - \eta_N(\nu_{k,N})|\chi_N^2 > \alpha)
$$

$$
\leq \frac{1}{\alpha^{2l}} \mathbb{E}|(\tilde{\eta}_N(\nu_{k,N}) - \eta_N(\nu_{k,N}))\chi_N^2|^{2l}.
$$

The introduction of χ_N is in part motivated by the observation that the moments of $\tilde{\eta}_N(\nu_{k,N})\chi_N^2$ are finite. Moreover, it holds that

$$
\mathbb{E}|(\tilde{\eta}_N(\nu_{k,N}) - \eta_N(\nu_{k,N}))\chi_N^2|^{2l}
$$

$$
\leq C \oint_{\partial \mathcal{R}_y^-} \mathbb{E}\left[\left|\mathbf{a}(\nu_{k,N})^* \left(\mathbf{Q}_N(z)\frac{\hat{w}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)}\right.\right.\right.
$$

$$
\left.\left.\left. - \mathbf{T}_N(z)\frac{w_N'(z)}{1 + \sigma^2 c_N m_N(z)}\right) \mathbf{a}(\nu_{k,N})\chi_N^2\right|^{2l}\right] |\mathrm{d}z|.
$$

In order to complete the proof of Theorem 3.1, we establish the following proposition.

Proposition 3.3. *If $(\mathbf{a}_N)_{N \in \mathbb{N}}$ is sequence of deterministic vectors satisfying $\|\mathbf{a}_N\| = 1$, then, under Assumptions A-1–A-6, for each integer l, it holds that*

$$
\sup_{z \in \partial \mathcal{R}_y} \mathbb{E}\left[\left|\mathbf{a}_N^* \left(\mathbf{Q}_N(z)\frac{\hat{w}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} - \mathbf{T}_N(z)\frac{w_N'(z)}{1 + \sigma^2 c_N m_N(z)}\right) \mathbf{a}_N \chi_N^2\right|^{2l}\right]
$$

$$
\leq \frac{C}{N^l}, \tag{3.27}
$$

where the constant C does not depend on the sequence $(\mathbf{a}_N)$.

Proof. In order to shorten the notations, we denote by $\hat{g}_N(z)$ and $g_N(z)$ the functions defined by

$$\hat{g}_N(z) = \mathbf{a}_N^* \mathbf{Q}_N(z) \mathbf{a}_N \frac{\hat{w}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)},$$

and

$$g_N(z) = \mathbf{a}_N^* \mathbf{T}_N(z) \mathbf{a}_N \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)}.$$

In order to evaluate $\mathbb{E}|\hat{g}_N(z) - g_N(z)\chi_N^2|^{2l}$, we use the Poincaré inequality. For this, we first state the following lemma proved in the Appendix. We recall that if $\mathbf{H}$ is a Hermitian matrix with a spectral decomposition $\mathbf{H} = \sum_l \gamma_l \mathbf{x}_l \mathbf{x}_l^*$, its adjoint (i.e. the transpose of its cofactor matrix) denoted by $\mathrm{adj}(\mathbf{H})$ is given by $\mathrm{adj}(\mathbf{H}) = \sum_l (\prod_{k \neq l} \gamma_k) \mathbf{x}_l \mathbf{x}_l^*$. When $\mathbf{H}$ is invertible, $\mathrm{adj}(\mathbf{H}) = \det(\mathbf{H}) \mathbf{H}^{-1}$. Next, we state the following lemma proved in the Appendix.

Lemma 3.9. *Assume Assumptions* A-1–A-6. *Considered as functions of the real and imaginary parts of the entries of* $\mathbf{W}_N$, *functions* $\det \phi(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*)$ *and* $\det \phi(\hat{\mathbf{\Omega}}_N)$ *belong to* $\mathcal{C}^1(\mathbb{R}^{2MN})$, *and their partial derivatives with respect to* $W_{i,j,N}$ *denoted by*

$$[\mathbf{D}_1]_{i,j,N} := \frac{\partial}{\partial W_{i,j,N}} \{\det \phi(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*)\},$$

$$[\mathbf{D}_2]_{i,j,N} := \frac{\partial}{\partial W_{i,j,N}} \{\det \phi(\hat{\mathbf{\Omega}}_N)\},$$

are given almost surely by

$$[\mathbf{D}_1]_{i,j,N} = \mathbf{e}_j^* \mathbf{\Sigma}_N^* \, \mathrm{adj}(\phi(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*)) \phi'(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*) \mathbf{e}_i, \tag{3.28}$$

$$[\mathbf{D}_2]_{i,j,N} = \left[\mathbf{\Sigma}_N^* \sum_{l=1}^{M} [\mathrm{adj}(\phi(\hat{\mathbf{\Omega}}_N)) \phi'(\hat{\mathbf{\Omega}}_N)]_{ll} \hat{\mathbf{\Pi}}_{l,N} \right]_{ji}. \tag{3.29}$$

If we denote by $\mathcal{A}_{1,N}$ *and* $\mathcal{A}_{2,N}$ *the events defined by*

$$\mathcal{A}_{1,N} = \{\exists k : \hat{\lambda}_{k,N} \notin \mathcal{T}_\epsilon\} \cap \{\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M,N} \in \mathrm{supp}(\phi)\},$$

$$\mathcal{A}_{2,N} = \{\exists k : \hat{\omega}_{k,N} \notin \mathcal{T}_\epsilon\} \cap \{\hat{\omega}_{1,N}, \ldots, \hat{\omega}_{M,N} \in \mathrm{supp}(\phi)\},$$

then $[\mathbf{D}_1]_{i,j,N} = 0$ *on* $\mathcal{A}_{1,N}^c$ *and* $[\mathbf{D}_2]_{i,j,N} = 0$ *on* $\mathcal{A}_{2,N}^c$.

We now establish (3.27) by induction on l, and first consider the case $l = 1$. We write the second moment of $(\hat{g}_N(z) - g_N(z))\chi_N^2$ as

$$\mathbb{E}|(\hat{g}_N(z) - g_N(z))\chi_N^2|^2 = |\mathbb{E}((\hat{g}_N(z) - g_N(z))\chi_N^2)|^2 + \mathrm{Var}((\hat{g}_N(z) - g_N(z))\chi_N^2).$$

We evaluate $\mathrm{Var}[(\hat{g}_N(z) - g_N(z))\chi_N^2]$ using the Poincaré inequality and get

$$\mathrm{Var}[(\hat{g}_N(z) - g_N(z))\chi_N^2] \leq \frac{\sigma^2}{N} \sum_{i,j} \mathbb{E}\left[\chi_N^4 \left(\left|\frac{\partial \hat{g}_N(z)}{\partial W_{i,j,N}}\right|^2 + \left|\frac{\partial \hat{g}_N(z)}{\partial \overline{W}_{i,j,N}}\right|^2\right)\right]$$

$$+ 2\mathbb{E}\left[|\hat{g}_N(z) - g_N(z)|^2 \left|\frac{\partial \chi_N^2}{\partial W_{i,j,N}}\right|^2\right]. \qquad (3.30)$$

It is clear that

$$\frac{\partial \hat{g}_N(z)}{\partial W_{i,j,N}} = \mathbf{a}_N^* \frac{\partial \mathbf{Q}_N(z)}{\partial W_{i,j,N}} \mathbf{a}_N \frac{\hat{w}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)}$$

$$+ \mathbf{a}_N^* \mathbf{Q}_N(z) \mathbf{a}_N \frac{\partial}{\partial W_{i,j,N}} \left\{\frac{\hat{w}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)}\right\}.$$

We verify that

$$\mathbf{a}_N^* \frac{\partial \mathbf{Q}_N(z)}{\partial W_{i,j,N}} \mathbf{a}_N = -\mathbf{a}_N^* \mathbf{Q}_N(z) \mathbf{e}_i \mathbf{e}_j \mathbf{\Sigma}_N^* \mathbf{Q}_N(z) \mathbf{a}_N,$$

so that

$$\sum_{i,j} \left|\mathbf{a}_N^* \frac{\partial \mathbf{Q}_N(z)}{\partial W_{i,j,N}} \mathbf{a}_N\right|^2 = \mathbf{a}_N^* \mathbf{Q}_N(z) \mathbf{Q}_N(z)^* \mathbf{a}_N \mathbf{a}_N^* \mathbf{Q}_N(z) \mathbf{\Sigma}_N \mathbf{\Sigma}_N^* \mathbf{Q}_N(z)^* \mathbf{a}_N.$$

As $\chi_N \neq 0$ implies that $\hat{\lambda}_{M,N} = \|\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*\| \leq t_2^+ + 2\epsilon$, Lemma 3.8 implies that

$$\sup_{z \in \partial \mathcal{R}_y} \chi_N^2 \mathbf{a}_N^* \mathbf{Q}_N(z) \mathbf{Q}_N(z)^* \mathbf{a}_N \mathbf{a}_N^* \mathbf{Q}_N(z) \mathbf{\Sigma}_N \mathbf{\Sigma}_N^* \mathbf{Q}_N(z)^* \mathbf{a}_N \leq C.$$

Using again Lemma 3.8, we get that

$$\sup_{z \in \partial \mathcal{R}_y} \chi_N^4 \left|\frac{\hat{w}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)}\right|^2 \sum_{i,j} \left|\mathbf{a}_N^* \frac{\partial \mathbf{Q}_N(z)}{\partial W_{i,j,N}} \mathbf{a}_N\right|^2 \leq C. \qquad (3.31)$$

We obtain similarly that

$$\sup_{z \in \partial \mathcal{R}_y} \chi_N^4 |\mathbf{a}_N^* \mathbf{Q}_N(z) \mathbf{a}_N|^2 \sum_{i,j} \left|\frac{\partial}{\partial W_{i,j,N}} \left\{\frac{\hat{w}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)}\right\}\right|^2 \leq \frac{C}{N}. \qquad (3.32)$$

The same conclusions hold when the derivatives with respect to variables $\overline{W}_{i,j,N}$ are considered. This shows that the first term of the right-hand side of (3.30) is a $\mathcal{O}(\frac{1}{N})$ term. We now evaluate the behavior of the second term of the right-hand side of (3.30), and establish that

$$\sup_{z \in \partial \mathcal{R}_y} \mathbb{E}\left[|\hat{g}_N(z) - g_N(z)|^2 \sum_{i,j} \left|\frac{\partial \chi_N^2}{\partial W_{i,j,N}}\right|^2\right] = \mathcal{O}\left(\frac{1}{N^p}\right) \qquad (3.33)$$

for each integer p. We express $\frac{\partial \chi_N^2}{\partial W_{i,j,N}}$ as $2\chi_N \frac{\partial \chi_N}{\partial W_{i,j,N}}$. Lemma 3.8 implies that $\sup_{z \in \partial \mathcal{R}_y} \chi_N^2 |\hat{g}_N(z) - g_N(z)|^2 < C$. Therefore, it is sufficient to check that

$$\mathbb{E}\left[\sum_{i,j} \left|\frac{\partial \chi_N}{\partial W_{i,j,N}}\right|^2\right] = \mathcal{O}\left(\frac{1}{N^p}\right)$$

for each integer p. $\frac{\partial \chi_N}{\partial W_{i,j,N}}$ can be written as

$$\frac{\partial \chi_N}{\partial W_{i,j,N}} = [\mathbf{D}_1]_{i,j,N} \det \phi(\hat{\mathbf{\Omega}}_N) + [\mathbf{D}_2]_{i,j,N} \det \phi(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*).$$

It holds that

$$\mathbb{E}\left[\sum_{i,j} |[\mathbf{D}_1]_{i,j,N} \det \phi(\hat{\mathbf{\Omega}}_N)|^2\right]$$

$$= \mathbb{E}[\det \phi(\hat{\mathbf{\Omega}}_N)^2 \operatorname{Tr}(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* \phi'(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*)^2 \operatorname{adj}(\phi(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*))^2) \mathbb{1}_{\mathcal{A}_{1,N}}]$$

Moreover, we can write

$$\operatorname{Tr}(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* \phi'(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*)^2 \operatorname{adj}(\phi(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*))^2) = \sum_k \hat{\lambda}_{k,N} \phi'(\hat{\lambda}_{k,N})^2 \prod_{l \neq k} \phi(\hat{\lambda}_{l,N})^2$$

$$\leq \operatorname{Tr}(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* \phi'(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*)^2),$$

because $\phi(\lambda) \leq 1$ on $\mathbb{R}$. Therefore, it holds that

$$\mathbb{E}\left[\sum_{i,j} |[\mathbf{D}_1]_{i,j,N} \det \phi(\hat{\mathbf{\Omega}}_N)|^2\right] \leq \mathbb{E}[\det \phi(\hat{\mathbf{\Omega}}_N)^2 \operatorname{Tr}(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* \phi'(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*)^2) \mathbb{1}_{\mathcal{A}_{1,N}}]$$

$$\leq C N \mathbb{P}(\mathcal{A}_{1,N}),$$

because $\det \phi(\hat{\mathbf{\Omega}}_N) \leq 1$, and $\operatorname{Tr}(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* \phi'(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*)^2) \leq CN$ on $\mathcal{A}_{1,N}$. As $\mathcal{A}_{1,N} \subset \mathcal{E}_{1,N}$, Proposition 3.1 implies that

$$\mathbb{E}\left[\sum_{i,j} |[\mathbf{D}_1]_{i,j,N} \det \phi(\hat{\mathbf{\Omega}}_N)|^2\right] = \mathcal{O}\left(\frac{1}{N^p}\right)$$

for each integer p. Using similar calculations and Proposition 3.2, we obtain that

$$\mathbb{E}\left[\sum_{i,j} |[\mathbf{D}_2]_{i,j,N} \det \phi(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*)|^2\right] = \mathcal{O}\left(\frac{1}{N^p}\right)$$

for each integer p. This completes the proof of (3.33) and establishes that

$$\sup_{z \in \partial \mathcal{R}_y} \operatorname{Var}[(\hat{g}_N(z) - g_N(z))\chi_N^2] = \mathcal{O}\left(\frac{1}{N}\right).$$

In order to evaluate the term $|\mathbb{E}[(\hat{g}_N(z) - g_N(z))\chi_N^2]|^2$, we also need the following auxilliary lemma proved in the Appendix.

Lemma 3.10. *Under Assumptions* A-1–A-6, *it holds that*

$$\sup_{z\in\partial\mathcal{R}_y}|\mathbb{E}[\mathbf{a}_N^*\mathbf{Q}_N(z)\mathbf{a}_N\chi_N - \mathbf{a}_N^*\mathbf{T}_N(z)\mathbf{a}_N]| = \mathcal{O}\left(\frac{1}{N^{\frac{3}{2}}}\right), \qquad (3.34)$$

$$\sup_{z\in\partial\mathcal{R}_y}|\mathbb{E}[\hat{m}_N(z)\chi_N - m_N(z)]| = \mathcal{O}\left(\frac{1}{N^2}\right), \qquad (3.35)$$

$$\sup_{z\in\partial\mathcal{R}_y}|\mathbb{E}[\hat{m}'_N(z)\chi_N - m'_N(z)]| = \mathcal{O}\left(\frac{1}{N^2}\right). \qquad (3.36)$$

We express $(\hat{g}_N(z) - g_N(z))\chi_N^2$ as $\beta_{1,N}(z) + \beta_{2,N}(z)$ where

$$\beta_{1,N}(z) = \chi_N(\mathbf{a}_N^*\mathbf{Q}_N(z)\mathbf{a}_N - \mathbf{a}_N^*\mathbf{T}_N(z)\mathbf{a}_N)\frac{\hat{w}'_N(z)\chi_N}{1+\sigma^2 c_N\hat{m}_N(z)}$$

and

$$\beta_{2,N}(z) = \chi_N^2\mathbf{a}_N^*\mathbf{T}_N(z)\mathbf{a}_N\left(\frac{\hat{w}'_N(z)}{1+\sigma^2 c_N\hat{m}_N(z)} - \frac{w'_N(z)}{1+\sigma^2 c_N m_N(z)}\right),$$

and establish that

$$\sup_{z\in\partial\mathcal{R}_y}\mathbb{E}|\beta_{1,N}|^2 = \mathcal{O}\left(\frac{1}{N}\right) \quad \text{and} \quad \sup_{z\in\partial\mathcal{R}_y}\mathbb{E}|\beta_{2,N}|^2 = \mathcal{O}\left(\frac{1}{N^2}\right). \qquad (3.37)$$

Using Lemma 3.8, (3.37) for $\beta_{1,N}$ will be established if we show that

$$\sup_{z\in\partial\mathcal{R}_y}\mathbb{E}|\chi_N(\mathbf{a}_N^*\mathbf{Q}_N(z)\mathbf{a}_N - \mathbf{a}_N^*\mathbf{T}_N(z)\mathbf{a}_N)|^2 = \mathcal{O}\left(\frac{1}{N}\right).$$

For this, we write that

$$\mathbb{E}|\chi_N(\mathbf{a}_N^*\mathbf{Q}_N(z)\mathbf{a}_N - \mathbf{a}_N^*\mathbf{T}_N(z)\mathbf{a}_N)|^2$$
$$= \mathrm{Var}(\chi_N\mathbf{a}_N^*\mathbf{Q}_N(z)\mathbf{a}_N) + |\mathbb{E}(\chi_N(\mathbf{a}_N^*\mathbf{Q}_N(z)\mathbf{a}_N - \mathbf{a}_N^*\mathbf{T}_N(z)\mathbf{a}_N))|^2.$$

The above calculations prove that $\sup_{z\in\partial\mathcal{R}_y}\mathrm{Var}[\chi_N\mathbf{a}_N^*\mathbf{Q}_N(z)\mathbf{a}_N] = \mathcal{O}(\frac{1}{N})$, while (3.34) and $1 - \mathbb{E}(\chi_N) = \mathcal{O}(\frac{1}{N^p})$ for each p imply that

$$\mathbb{E}[\chi_N(\mathbf{a}_N^*\mathbf{Q}_N(z)\mathbf{a}_N - \mathbf{a}_N^*\mathbf{T}_N(z)\mathbf{a}_N)] = \mathcal{O}\left(\frac{1}{N^{\frac{3}{2}}}\right).$$

This completes the proof of (3.37) for $\beta_{1,N}$. In order to show (3.37) for $\beta_{2,N}$, we first remark that by Lemma 3.8, $|\mathbf{a}_N^*\mathbf{T}_N(z)\mathbf{a}_N|$ is uniformly bounded on $\partial\mathcal{R}_y$, and write that

$$\chi_N^2\left(\frac{\hat{w}'_N(z)}{1+\sigma^2 c_N\hat{m}_N(z)} - \frac{w'_N(z)}{1+\sigma^2 c_N m_N(z)}\right)$$
$$= \sigma^2 c_N\chi_N^2(\hat{m}_N(z) - m_N(z)) + 2z\sigma^2 c_N\chi_N^2(\hat{m}'_N(z) - m'_N(z))$$
$$- \sigma^4 c_N(1-c_N)\chi_N^2\left(\frac{\hat{m}'_N(z)}{1+\sigma^2 c_N\hat{m}_N(z)} - \frac{m'_N(z)}{1+\sigma^2 c_N m_N(z)}\right),$$

or equivalently that

$$\chi_N^2 \left(\frac{\hat{w}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} - \frac{w_N'(z)}{1 + \sigma^2 c_N m_N(z)} \right)$$

$$= \sigma^2 c_N \chi_N^2 (\hat{m}_N(z) - m_N(z)) + 2z\sigma^2 c_N \chi_N^2 (\hat{m}_N'(z) - m_N'(z))$$

$$- (\sigma^2 c_N)^2 \sigma^2 (1 - c_N)$$

$$\times \frac{m_N(z)\chi_N(\hat{m}_N'(z) - m_N'(z)) - m_N'(z)\chi_N(\hat{m}_N(z) - m_N(z))}{(1 + \sigma^2 c_N \hat{m}_N(z))(1 + \sigma^2 c_N m_N(z))} \chi_N$$

$$- \sigma^2 c_N \sigma^2 (1 - c_N) \frac{\chi_N(\hat{m}_N'(z) - m_N'(z))}{(1 + \sigma^2 c_N \hat{m}_N(z))(1 + \sigma^2 c_N m_N(z))} \chi_N.$$

The Poincaré inequality and Lemma 3.10 imply that

$$\sup_{z \in \partial \mathcal{R}_y} \mathbb{E}|\chi_N(\hat{m}_N(z) - m_N(z))|^2 = \mathcal{O}\left(\frac{1}{N^2}\right)$$

and

$$\sup_{z \in \partial \mathcal{R}_y} \mathbb{E}|\chi_N(\hat{m}_N'(z) - m_N'(z))|^2 = \mathcal{O}\left(\frac{1}{N^2}\right).$$

Equation (3.37) follows immediately from

$$\sup_{z \in \partial \mathcal{R}_y} \left| \frac{\chi_N}{(1 + \sigma^2 c_N \hat{m}_N(z))(1 + \sigma^2 c_N m_N(z))} \right| \leq C,$$

for some deterministic constant C (see Lemma 3.8). This completes the proof of (3.27) for $l = 1$.

We now assume that (3.27) holds until integer $l - 1$ and write that

$$\mathbb{E}|\chi_N^2(\hat{g}_N(z) - g_N(z))|^{2l} = |\mathbb{E}[(\chi_N^2(\hat{g}_N(z) - g_N(z)))^l]|^2$$

$$+ \mathrm{Var}[(\chi_N^2(\hat{g}_N(z) - g_N(z)))^l].$$

The Cauchy–Schwarz inequality implies that

$$|\mathbb{E}[(\chi_N^2(\hat{g}_N(z) - g_N(z)))^l]|^2$$

$$\leq \mathbb{E}|\chi_N^2(\hat{g}_N(z) - g_N(z))|^2 \mathbb{E}|\chi_N^2(\hat{g}_N(z) - g_N(z))|^{2(l-1)},$$

and shows that

$$\sup_{z \in \partial \mathcal{R}_y} |\mathbb{E}(\chi_N^2(\hat{g}_N(z) - g_N(z)))^l|^2 = \mathcal{O}\left(\frac{1}{N^l}\right).$$

The Poincaré inequality gives

$$\text{Var}[(\chi_N^2(\hat{g}_N(z) - g_N(z)))^l]$$

$$\leq \frac{\sigma^2 l^2}{N} \mathbb{E}\left[|\chi_N^2(\hat{g}_N(z) - g_N(z))|^{2(l-1)} \chi_N^4 \sum_{i,j} \left(\left|\frac{\partial \hat{g}_N(z)}{\partial W_{i,j,N}}\right|^2 + \left|\frac{\partial \hat{g}_N(z)}{\partial \overline{W}_{i,j,N}}\right|^2\right)\right]$$

$$+ \frac{8\sigma^2 l^2}{N} \mathbb{E}\left[|\hat{g}_N(z) - g_N(z)|^{2l} \chi_N^{4l-2} \sum_{i,j} \left|\frac{\partial \chi_N}{\partial W_{i,j,N}}\right|^2\right]. \tag{3.38}$$

Finally, (3.31) and (3.32) imply that

$$\sup_{z \in \partial \mathcal{R}_y} \chi_N^4 \sum_{i,j} \left(\left|\frac{\partial \hat{g}_N(z)}{\partial W_{i,j,N}}\right|^2 + \left|\frac{\partial \hat{g}_N(z)}{\partial \overline{W}_{i,j,N}}\right|^2\right) \leq C,$$

for some deterministic constant C. Therefore, the supremum over $z \in \partial \mathcal{R}_y$ of first term of the right-hand side of (3.38) is a $\mathcal{O}(\frac{1}{N^l})$. Moreover, it can be shown as in the case $l = 1$ that the supremum over $z \in \partial \mathcal{R}_y$ of the second term of the right-hand side of (3.38) is a $\mathcal{O}(\frac{1}{N^p})$ for each integer p. This completes the proof of Proposition 3.3 and of the uniform consistency of estimator $\tilde{\eta}_N(\theta)$. $\qquad\square$

3.4. *A brief discussion on the case $c > 1$*

We now briefly explain how the previous results can be adapted to the case $c > 1$ and assume without restriction in this section that $c_N = \frac{M}{N} > 1$ for each N. In this context, the smallest $M - N$ eigenvalues $\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M-N,N}$ of matrix $\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*$ are identically 0. The behavior of the empirical eigenvalue distribution $\hat{\mu}_N$ of $\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*$ is still characterized by the results of Sec. 2.1 which do not depend on the value of c_N. However, the characterization of the support $\mathcal{S}_N$ of μ_N has to be adapted because $\{0\}$ is carried by μ_N, the corresponding mass being equal to $1 - \frac{1}{c_N}$. Using the results of Sec. 2.2 to the dual model

$$\boldsymbol{\Sigma}_N^* = \mathbf{B}_N^* + \mathbf{W}_N^*, \tag{3.39}$$

it is easily seen that $\mathcal{S}_N$ is given by

$$\mathcal{S}_N = \{0\} \cup \left(\bigcup_{q=1}^{Q} [x_{q,N}^-, x_{q,N}^+]\right),$$

where the points $(x_{q,N}^-, x_{q,N}^+)_{q=1,\ldots,Q}$ are still defined by the Q positive extrema of function $\phi_N(w)$ defined by (2.7). Note that $\cup_{q=1}^{Q}[x_{q,N}^-, x_{q,N}^+]$ coincides with the support of the measure $\tilde{\mu}_N = c_N \mu_N + (1 - c_N)\delta_0$, i.e. the asymptotic eigenvalue distribution of the empirical eigenvalue distribution $\frac{1}{N}\sum_{k=M-N+1}^{M} \delta_{\hat{\lambda}_{k,N}}$ of matrix $\boldsymbol{\Sigma}_N^* \boldsymbol{\Sigma}_N$. Moreover, function $w_N(z)$ defined by (2.8) still verifies Proposition 2.1 where $(w_{q,N}^-, w_{q,N}^+)_{q=1,\ldots,Q}$ are the preimages of the $(x_{q,N}^-, x_{q,N}^+)_{q=1,\ldots,Q}$. Also, the N non-zero eigenvalues $(\hat{\lambda}_{k,N})_{k=M-N+1,\ldots,M}$ satisfy Theorems 2.3 and 2.4. Finally,

under Assumptions A-5 and A-6, matrix $\mathbf{\Pi}_N$ is still given by (2.20), and $\tilde{\eta}_N(\theta) = \mathbf{a}_N(\theta)^* \tilde{\mathbf{\Pi}}_N \mathbf{a}_N(\theta)$ where $\tilde{\mathbf{\Pi}}_N$ is defined by (2.21) is still a consistent estimator of $\eta_N(\theta)$. We note that the properties of the zeros of $1 + \sigma^2 c_N \hat{m}_N(z)$ and of the eigenvalues $(\hat{\omega}_{k,N})_{k=1,\ldots,M}$ of matrix $\hat{\mathbf{\Omega}}_N$ defined by (2.13) slightly differ from the case $c_N < 1$. Indeed, as 0 is eigenvalue of matrix $\hat{\mathbf{\Lambda}}_N$ with multiplicity $M - N$, $\hat{\omega}_{1,N} = \cdots = \hat{\omega}_{M-N-1,N} = 0 < \hat{\omega}_{M-N,N}$. The zeros of $1 + \sigma^2 c_N \hat{m}_N(z)$ are the $(\hat{\omega}_k)_{k=M-N,\ldots,M}$ and they verify

$$0 < \hat{\omega}_{M-N,N} < \hat{\lambda}_{M-N+1,N} < \hat{\omega}_{M-N+1,N} < \cdots < \hat{\omega}_{M-1,N} < \hat{\lambda}_{M,N} < \hat{\omega}_{M,N}.$$

It can be shown as in [23] that almost surely, for N large enough,

$$0 < \hat{\omega}_{M-N,N} < t_1^- - 4\epsilon,$$

$$\hat{\omega}_{M-N+1,N}, \ldots, \hat{\omega}_{M-K,N} \in (t_1^-, t_1^+),$$

$$\hat{\omega}_{M-K+1,N}, \ldots, \hat{\omega}_{M,N} \in (t_2^-, t_2^+).$$

Note that we have to assume here that $\epsilon < \frac{y}{4}$.

We now discuss on the uniform consistency of $\tilde{\eta}_N(\theta)$ in the case $c_N > 1$. It is easily seen that the probability that the non-zero eigenvalues $(\hat{\lambda}_{k,N})_{k=M-N+1,\ldots,M}$ escape from the set $\mathcal{T}_\epsilon$ still decreases at rate $\frac{1}{N^l}$ for each integer l. In order to extend property (1.1), the reader may check that it is sufficient to prove that the probability that the $(\hat{\omega}_k)_{k=M-N,\ldots,M}$ escape from $[0, t_1^- - 4\epsilon) \cup \mathcal{T}_\epsilon$ converge to 0 at rate $\frac{1}{N^l}$ for each integer l. For this, we have to prove that Lemma 3.6 is still valid. In the case $c_N < 1$, Lemma 3.6 depends on Lemma 3.3. In the case $c_N > 1$, the statement of Lemma 3.3 has to be adapted because the properties of the signed measure κ_N are slightly different. More precisely, it can be shown that for $c_N > 1$, it holds that

$$\kappa_N(\{0\}) = 1, \quad \kappa_N([x_{1,N}^-, x_{1,N}^+]) = -1, \quad \kappa_N([x_{q,N}^-, x_{q,N}^+]) = 0, \quad q \geq 2, \tag{3.40}$$

$\kappa_N([x_{1,N}^-, x_{1,N}^+]) = -1$ because in contrast with the case $c_N < 1$, it can be shown that $1 + \sigma^2 c_N m_N(x_{1,N}^-)$ is real and negative if $c_N > 1$; therefore, $\log(1 + \sigma^2 c_N m_N(x_{1,N}^-))$ is equal to $i\pi$, and (3.14) leads to $\kappa_N([x_{1,N}^-, x_{1,N}^+]) = -1$. As $\kappa_N(\{0\}) + \kappa_N([x_{1,N}^-, x_{1,N}^+]) = 0$, it is easy to check that the proof of Lemma 3.6 is still valid.

4. Consistency of the Angular Estimates

We now address the consistency of the DoA estimates defined as the local minima of function $\theta \to \tilde{\eta}_N(\theta)$. For this, we assume that the number of sources K is fixed, i.e. that K does not scale with N. In other words, model $\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N$ corresponds to a finite rank perturbation of the complex Gaussian iid matrix $\mathbf{W}_N$.

Remark 4.1. In this context, it is possible to derive in a simpler way than above an alternative consistent estimator, say $\theta \to \hat{\eta}_{N,\text{spike}}(\theta)$ of function $\theta \mapsto \eta_N(\theta)$.

This estimator is obtained by assuming from the very beginning that K is fixed, and is based on the recent work of Benaych–Nadakuditi [3]; see [22] for more details. However, as shown in [22], estimator $\tilde{\eta}_N(\theta)$ always leads in practice to the same performance as $\hat{\eta}_{N,\text{spike}}(\theta)$ if $\frac{K}{M} \ll 1$ (typical value $\frac{1}{10}$ in [22]), but outperforms $\hat{\eta}_{N,\text{spike}}(\theta)$ for greater values of $\frac{K}{M}$ (typical value $\frac{1}{4}$ in [22]). Therefore, the use of estimator $\tilde{\eta}_N(\theta)$ appears in practice more relevant than $\hat{\eta}_{N,\text{spike}}(\theta)$.

In order to define the estimators of $\theta_1, \ldots, \theta_K$ properly, we consider K disjoint intervals $I_1, \ldots, I_K$, such that $\theta_k \in I_k$, and define for each k the estimator $\tilde{\theta}_{k,N}$ of θ_k by $\tilde{\theta}_{k,N} = \arg\min_{\theta \in I_k} |\tilde{\eta}_N(\theta)|$. We prove the following result.

Proposition 4.1. *Under Assumptions* A-1–A-6, *for* $k = 1, \ldots, K$,

$$N(\tilde{\theta}_{k,N} - \theta_k) \xrightarrow[N \to \infty]{\text{a.s.}} 0.$$

In order to establish the proposition, we follow a classical approach initiated by Hannan [15] to study sinusoid frequency estimates. For this, we first recall the following useful lemma.

Lemma 4.1. *Let* (α_M) *a real-valued sequence of a compact subset of* $(-\pi, \pi]$, *and converging to* α *as* $M \to \infty$. *Define* $q_M(\alpha_M) = \frac{1}{M} \sum_{k=0}^{M-1} e^{ik\alpha_M}$. *If* $\alpha \neq 0$ *or if* $\alpha = 0$ *and* $M|\alpha_M| \to \infty$, *then* $q_M(\alpha_M) \to 0$. *If* $\alpha = 0$ *and* $M\alpha_M \xrightarrow[M \to \infty]{} \beta \in \mathbb{R}$, *then* $q_M(\alpha_M) \to e^{i\frac{\beta}{2}} \operatorname{sinc}(\beta/2)$.

We denote by $\mathbf{A}$ the matrix $\mathbf{A}(\mathbf{\Theta})$ corresponding to the true angles $\mathbf{\Theta} = (\theta_1, \ldots, \theta_K)^T$. It is clear that $\eta_N(\theta) = 1 - \mathbf{a}(\theta)^* \mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*\mathbf{a}(\theta)$. By the very definition of $\tilde{\theta}_{k,N}$, $|\tilde{\eta}_N(\tilde{\theta}_{k,N})| \leq |\tilde{\eta}_N(\theta_k)|$. From Theorem 3.1 and the equality $\eta_N(\theta_k) = 0$, we have $|\tilde{\eta}_N(\tilde{\theta}_{k,N})| \to 0$ w.p.1., as $N \to \infty$. Consequently,

$$
\begin{aligned}
|\eta_N(\tilde{\theta}_{k,N})| &\leq |\eta_N(\tilde{\theta}_{k,N}) - \tilde{\eta}_N(\tilde{\theta}_{k,N})| + |\tilde{\eta}_N(\tilde{\theta}_{k,N})| \\
&\leq \sup_{\theta \in [-\pi, \pi]} |\eta_N(\theta) - \tilde{\eta}_N(\theta)| + |\tilde{\eta}_N(\tilde{\theta}_{k,N})| \\
&\xrightarrow[N \to \infty]{\text{a.s.}} 0.
\end{aligned}
\tag{4.1}
$$

From Lemma 4.1, $(\mathbf{A}^*\mathbf{A})^{-1}$ converges to $\mathbf{I}_K$ as $N \to \infty$. Since $(\tilde{\theta}_{k,N})$ is bounded, we can extract a converging subsequence $(\tilde{\theta}_{k,\varphi(N)})$. Let $\alpha_N = \tilde{\theta}_{k,\varphi(N)} - \theta_k$. From Lemma 4.1, if $\alpha_N \to \alpha \neq 0$ as $N \to \infty$, then

$$\mathbf{a}(\tilde{\theta}_{k,\varphi(N)})^* \mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*\mathbf{a}(\tilde{\theta}_{k,\varphi(N)}) \xrightarrow[N \to \infty]{\text{a.s.}} 0, \tag{4.2}$$

and thus $\eta_N(\tilde{\theta}_{k,\varphi(N)}) \to 1$, a contradiction with (4.1). This implies that the whole sequence $(\tilde{\theta}_{k,N})$ converges towards θ_k. If $N|\tilde{\theta}_{k,N} - \theta_k|$ is not bounded, we can extract a subsequence such that $\phi(N)|\tilde{\theta}_{k,\phi(N)} - \theta_k| \to +\infty$ and Lemma 4.1 again implies that (4.2) holds, a contradiction. $N|\tilde{\theta}_{k,N} - \theta_k|$ is thus bounded, and we consider a subsequence such that $N(\tilde{\theta}_{k,\varphi(N)} - \theta_k) \to \beta/2$ where $\beta \in (-\pi, \pi]$. From Lemma 4.1,

if $\beta \neq 0$, we get

$$\eta_{\varphi(N)}(\tilde{\theta}_{k,\varphi(N)}) \xrightarrow[N\to\infty]{\text{a.s.}} 1 - \left(\frac{\sin(\beta/2)}{\beta/2}\right)^2 > 0,$$

which is again in contradiction with (4.1). Therefore, $\beta = 0$ and all converging subsequences of $(N|\hat{\theta}_{k,\varphi(N)} - \theta_k|)$ converge to 0, which of course implies that the whole sequence $(N|\hat{\theta}_{k,N} - \theta_k|)$ converges to 0. We finally end up with $N(\tilde{\theta}_{k,N} - \theta_k) \to 0$ w.p.1., as $N \to \infty$.

Appendix A

A.1. *Proof of Lemma 3.2: Estimate of $\mathbb{E}[\hat{m}'_N(z)]$*

We first give the following useful technical result. Its proof, based on Poincaré's inequality, is elementary and therefore omitted.

Lemma A.1. *Let $(\mathbf{M}_N(z))$ a sequence of deterministic complex $M \times M$ matrix-valued functions defined on $\mathbb{C}\backslash\mathbb{R}$ such that*

$$\|\mathbf{M}_N(z)\| \leq \mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1}).$$

Then, under Assumptions A-1, A-2 and A-4,

$$\mathrm{Var}\left[\frac{1}{N}\mathrm{Tr}\,\mathbf{Q}_N(z)\mathbf{M}_N(z)\right] \leq \frac{1}{N^2}\mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1}),$$

$$\mathrm{Var}\left[\frac{1}{N}\mathrm{Tr}\,\mathbf{\Sigma}_N^*\mathbf{Q}_N(z)\mathbf{M}_N(z)\right] \leq \frac{1}{N^2}\mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1}).$$

Moreover, the same results still hold when $\mathbf{Q}_N(z)$ is replaced by $\mathbf{Q}_N(z)^2$.

We are now in position to establish Lemma 3.2. We have to establish that

$$|\mathbb{E}[\hat{m}'_N(z)] - m'_N(z)| \leq \frac{1}{N^2}\mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1}). \tag{A.1}$$

For clarity, we recall results from [10, 23, 14], on which the proof heavily relies. We have first to introduce some new notations extensively used in [10, 23, 14]. We define $\delta_N(z) = \sigma c_N m_N(z) = \sigma\frac{1}{N}\mathrm{Tr}\,(\mathbf{T}_N(z))$, as well as $\tilde{\delta}_N(z) = \delta_N(z) - \frac{\sigma(1-c_N)}{z}$ which coincides with the Stieltjes transform of finite measures $c_N\mu_N + (1 - c_N)\delta_0$. In the following, matrix $\tilde{\mathbf{T}}_N(z)$ is defined by

$$\tilde{\mathbf{T}}_N(z) = \left(-z(1 + \sigma\delta_N(z))\mathbf{I}_M + \frac{\mathbf{B}_N^*\mathbf{B}_N}{1 + \sigma\tilde{\delta}_N(z)}\right)^{-1},$$

and is related to $\tilde{\delta}_N(z)$ through the equation $\tilde{\delta}_N(z) = \sigma\frac{1}{N}\mathrm{Tr}\,(\tilde{\mathbf{T}}_N(z))$ (cf. [13, 23]). We also remark that matrix $\mathbf{T}_N(z)$ can be written as

$$\mathbf{T}_N(z) = \left(-z(1 + \sigma\tilde{\delta}_N(z))\mathbf{I}_M + \frac{\mathbf{B}_N\mathbf{B}_N^*}{1 + \sigma\delta_N(z)}\right)^{-1},$$

and that $w_N(z)$ coincides with $z(1 + \sigma\delta_N(z))(1 + \sigma\tilde{\delta}_N(z))$. We also denote $\tilde{\mathbf{Q}}_N(z)$ the resolvent of matrix $\boldsymbol{\Sigma}_N^*\boldsymbol{\Sigma}_N$, i.e.

$$\tilde{\mathbf{Q}}_N(z) = (\boldsymbol{\Sigma}_N^*\boldsymbol{\Sigma}_N - z\mathbf{I}_N)^{-1}$$

and define $\alpha_N(z) = \mathbb{E}[\frac{\sigma}{N}\operatorname{Tr}\mathbf{Q}_N(z)]$, $\tilde{\alpha}_N(z) = \mathbb{E}[\frac{\sigma}{N}\operatorname{Tr}\tilde{\mathbf{Q}}_N(z)]$, and the matrices

$$\mathbf{R}_N(z) = \left(-z(1 + \sigma\tilde{\alpha}_N(z))\mathbf{I}_M + \frac{\mathbf{B}_N\mathbf{B}_N^*}{1 + \sigma\alpha_N(z)}\right)^{-1},$$

$$\tilde{\mathbf{R}}_N(z) = \left(-z(1 + \sigma\alpha_N(z))\mathbf{I}_N + \frac{\mathbf{B}_N^*\mathbf{B}_N}{1 + \sigma\tilde{\alpha}_N(z)}\right)^{-1}.$$

It is shown in [10, 23] that the entries of $\mathbf{Q}_N(z)$ (respectively, $\tilde{\mathbf{Q}}_N(z)$) have the same behavior as the entries of $\mathbf{R}_N(z)$ and $\mathbf{T}_N(z)$ (respectively, of $\tilde{\mathbf{R}}_N(z)$ and $\tilde{\mathbf{T}}_N(z)$). It is also useful to recall that $|\alpha_N(z)|$, $|\tilde{\alpha}_N(z)|$, $|-z(1 + \sigma\alpha_N(z))|^{-1}$, $|-z(1 + \sigma\tilde{\alpha}_N(z))|^{-1}$, $\|\mathbf{T}_N(z)\|$, $\|\tilde{\mathbf{T}}_N(z)\|$, $\|\mathbf{R}_N(z)\|$ and $\|\tilde{\mathbf{R}}_N(z)\|$ are bounded on $\mathbb{C}\backslash\mathbb{R}$ by $\mathrm{P}_1(|z|)\mathrm{P}_2(|\operatorname{Im}(z)|^{-1})$. We remark that our new notations are symmetrical with respect to the substitution $\boldsymbol{\Sigma}_N \rightarrow \boldsymbol{\Sigma}_N^*$, and are easier to use in the forthcoming calculations.

We first notice that (A.1) is equivalent to

$$|\alpha_N'(z) - \delta_N'(z)| \leq \frac{1}{N^2}\mathrm{P}_1(|z|)\mathrm{P}_2(|\operatorname{Im}(z)|^{-1}). \tag{A.2}$$

In order to prove (A.2), we first show that

$$\left|\alpha_N'(z) - \frac{\sigma}{N}\operatorname{Tr}\mathbf{R}_N'(z)\right| \leq \frac{1}{N^2}\mathrm{P}_1(|z|)\mathrm{P}_2(|\operatorname{Im}(z)|^{-1}) \tag{A.3}$$

and deduce from this that (A.2) holds. Using results on the behavior of $\alpha_N(z) - \frac{\sigma}{N}\operatorname{Tr}\mathbf{R}_N(z)$ established in [10, 23, 14], we first establish that (A.3) holds. For this, we recall the following lemma.

Lemma A.2 ([10], [23, Proof of Proposition 6]). *Under Assumptions* A-1, A-2 *and* A-4, *for* $z \in \mathbb{C}\backslash\mathbb{R}$, *it holds that*

$$\mathbb{E}[\mathbf{Q}_N(z)] = \mathbf{R}_N(z) + \boldsymbol{\Delta}_N(z)\mathbf{R}_N(z) + \left(\frac{\sigma^2}{N}\operatorname{Tr}\boldsymbol{\Delta}_N(z)\right)\mathbb{E}[\mathbf{Q}_N(z)]\mathbf{R}_N(z), \tag{A.4}$$

where $\boldsymbol{\Delta}_N(z)$ *is given by* $\boldsymbol{\Delta}_N(z) = \boldsymbol{\Delta}_{1,N}(z) + \boldsymbol{\Delta}_{2,N}(z) + \boldsymbol{\Delta}_{3,N}(z)$ *with*

$$\boldsymbol{\Delta}_{1,N}(z) = \frac{\sigma}{1 + \sigma\alpha_N(z)}\mathbb{E}\left[\mathbf{Q}_N(z)\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\frac{\sigma}{N}\operatorname{Tr}\left(\mathbf{Q}_N(z) - \mathbb{E}[\mathbf{Q}_N(z)]\right)\right],$$

$$\boldsymbol{\Delta}_{2,N}(z) = \frac{\sigma^2}{1 + \sigma\alpha_N(z)}\mathbb{E}\left[(\mathbf{Q}_N(z) - \mathbb{E}[\mathbf{Q}_N(z)])\frac{\sigma}{N}\operatorname{Tr}\boldsymbol{\Sigma}_N^*\mathbf{Q}_N(z)\mathbf{B}_N\right],$$

$$\boldsymbol{\Delta}_{3,N}(z) = -\frac{\sigma^2}{(1+\sigma\alpha_N(z))^2}\mathbb{E}[\mathbf{Q}_N(z)]\mathbb{E}\left[\frac{\sigma}{N}\mathrm{Tr}\left(\mathbf{Q}_N(z)\right.\right.$$

$$\left.\left. - \mathbb{E}[\mathbf{Q}_N(z)]\right)\frac{\sigma}{N}\mathrm{Tr}\,\boldsymbol{\Sigma}_N^*\mathbf{Q}_N(z)\mathbf{B}_N\right].$$

If $\mathbf{M}_N(z)$ is a sequence of deterministic complex matrix-valued functions defined on $\mathbb{C}\backslash\mathbb{R}$ such that

$$\|\mathbf{M}_N(z)\| \le \mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1}),$$

then $\frac{1}{N}\mathrm{Tr}(\boldsymbol{\Delta}_{i,N}(z)\mathbf{M}_N(z))$, for $i=1,2,3$, is bounded by $\frac{1}{N^2}\mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$. Therefore,

$$\alpha_N(z) = \frac{\sigma}{N}\mathrm{Tr}\,\mathbf{R}_N(z) + \frac{\epsilon_{1,N}(z)}{N^2}, \tag{A.5}$$

where $\epsilon_{1,N}(z)$ is the holomorphic function on $\mathbb{C}\backslash\mathbb{R}$ defined by

$$\frac{\epsilon_{1,N}(z)}{N^2} = \left(\frac{\sigma}{N}\mathrm{Tr}\,\boldsymbol{\Delta}_N(z)\mathbf{R}_N(z) + \frac{\sigma}{N}\mathrm{Tr}\,\mathbb{E}[\mathbf{Q}_N(z)]\mathbf{R}_N(z)\frac{\sigma^2}{N}\mathrm{Tr}\,\boldsymbol{\Delta}_N(z)\right),$$

and satisfies $|\epsilon_{1,N}(z)| \le \mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$. Finally, $\tilde{\alpha}_N(z)$ can also be written as

$$\tilde{\alpha}_N(z) = \frac{\sigma}{N}\mathrm{Tr}\,\tilde{\mathbf{R}}_N(z) + \frac{\tilde{\epsilon}_{1,N}(z)}{N^2}, \tag{A.6}$$

where $\tilde{\epsilon}_{1,N}(z)$ is equal to

$$\tilde{\epsilon}_{1,N}(z) = \frac{1+\sigma\tilde{\alpha}_N(z)}{1+\sigma\alpha_N(z)}\epsilon_{1,N}(z),$$

and satisfies $|\tilde{\epsilon}_{1,N}(z)| \le \mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$.

In order to evaluate the behavior of $\alpha_N'(z) - \frac{\sigma}{N}\mathrm{Tr}\,\mathbf{R}_N'(z)$, we differentiate (A.5) with respect to z and get the following result.

Proposition A.1. *Under Assumptions A-1, A-2 and A-4, for $z \in \mathbb{C}\backslash\mathbb{R}$, the derivatives $\epsilon_{1,N}'(z)$ and $\tilde{\epsilon}_{1,N}'(z)$ of $\epsilon_{1,N}(z)$ and $\tilde{\epsilon}_{1,N}(z)$ with respect to z satisfy $|\epsilon_{1,N}'(z)| \le \mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$ and $|\tilde{\epsilon}_{1,N}'(z)| \le \mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$.*

Proof. The proof uses Lemma A.1 and the observation that the spectral norms $\|\mathbf{R}_N'(z)\|$ and $\|\tilde{\mathbf{R}}_N'(z)\|$ are bounded by $\mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$. The details are omitted. □

In order to complete the proof of the lemma, we establish that

Proposition A.2. *Under Assumptions A-1, A-2 and A-4, for $z \in \mathbb{C}\backslash\mathbb{R}$,*

$$\alpha_N'(z) = \delta_N'(z) + \frac{\epsilon_{2,N}(z)}{N^2},$$

$$\tilde{\alpha}_N'(z) = \tilde{\delta}_N'(z) + \frac{\tilde{\epsilon}_{2,N}(z)}{N^2},$$

where $|\epsilon_{2,N}(z)|$ and $|\tilde{\epsilon}_{2,N}(z)|$ are both bounded by $\mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$.

Proof. We first observe that (A.5) and (A.6) imply that

$$\alpha_N'(z) - \delta_N'(z) = \sigma\frac{1}{N}\mathrm{Tr}\left(\mathbf{R}_N'(z)\right) - \delta_N'(z) + \frac{\epsilon_{1,N}'(z)}{N^2},\tag{A.7}$$

$$\tilde{\alpha}_N'(z) - \tilde{\delta}_N'(z) = \sigma\frac{1}{N}\mathrm{Tr}\left(\tilde{\mathbf{R}}_N'(z)\right) - \tilde{\delta}_N'(z) + \frac{\tilde{\epsilon}_{1,N}'(z)}{N^2}.\tag{A.8}$$

We start with the classical identities

$$\mathbf{R}_N(z) - \mathbf{T}_N(z) = \mathbf{R}_N(z)(\mathbf{T}_N(z)^{-1} - \mathbf{R}_N(z)^{-1})\mathbf{T}_N(z),$$

$$\tilde{\mathbf{R}}_N(z) - \tilde{\mathbf{T}}_N(z) = \tilde{\mathbf{R}}_N(z)(\tilde{\mathbf{T}}_N(z)^{-1} - \tilde{\mathbf{R}}_N(z)^{-1})\tilde{\mathbf{T}}_N(z),$$

and get that

$$\frac{\sigma}{N}\mathrm{Tr}\left(\mathbf{R}_N(z) - \mathbf{T}_N(z)\right) = (\tilde{\alpha}_N(z) - \tilde{\delta}_N(z))zv_N(z)$$
$$+ (\alpha_N(z) - \delta_N(z))u_N(z),\tag{A.9}$$

$$\frac{\sigma}{N}\mathrm{Tr}\left(\tilde{\mathbf{R}}_N(z) - \tilde{\mathbf{T}}_N(z)\right) = (\tilde{\alpha}_N(z) - \tilde{\delta}_N(z))\tilde{u}_N(z)$$
$$+ (\alpha_N(z) - \delta_N(z))z\tilde{v}_N(z),\tag{A.10}$$

with

$$u_N(z) = \frac{\sigma^2}{N}\mathrm{Tr}\frac{\mathbf{R}_N(z)\mathbf{B}_N\mathbf{B}_N^*\mathbf{T}_N(z)}{(1 + \sigma\alpha_N(z))(1 + \sigma\delta_N(z))},$$

$$\tilde{u}_N(z) = \frac{\sigma^2}{N}\mathrm{Tr}\frac{\tilde{\mathbf{R}}_N(z)\mathbf{B}_N^*\mathbf{B}_N\tilde{\mathbf{T}}_N(z)}{(1 + \sigma\tilde{\alpha}_N(z))(1 + \sigma\tilde{\delta}_N(z))},$$

and

$$v_N(z) = \frac{\sigma^2}{N}\mathrm{Tr}\,\mathbf{R}_N(z)\mathbf{T}_N(z) \quad \tilde{v}_N(z) = \frac{\sigma^2}{N}\mathrm{Tr}\,\tilde{\mathbf{R}}_N(z)\tilde{\mathbf{T}}_N(z).$$

Note that it is easy to check that $u_N(z) = \tilde{u}_N(z)$. We differentiate (A.9), (A.10) with respect to z, we use (A.7), (A.8) and Proposition A.1, and recall that both $|\alpha_N(z) - \delta_N(z)|$ and $|\tilde{\alpha}_N(z) - \tilde{\delta}_N(z)|$ are bounded that $\frac{1}{N^2}\mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$ (see [23]). We check that $u_N(z)$, $zv_N(z)$, $z\tilde{v}_N(z)$ and their derivatives are bounded by $\mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$, and obtain eventually that

$$\begin{bmatrix} \alpha_N'(z) - \delta_N'(z) \\ \tilde{\alpha}_N'(z) - \tilde{\delta}_N'(z) \end{bmatrix} = \begin{bmatrix} u_N(z) & zv_N(z) \\ z\tilde{v}_N(z) & u_N(z) \end{bmatrix}\begin{bmatrix} \alpha_N'(z) - \delta_N'(z) \\ \tilde{\alpha}_N'(z) - \tilde{\delta}_N'(z) \end{bmatrix} + \frac{1}{N^2}\begin{bmatrix} \epsilon_{3,N}(z) \\ \tilde{\epsilon}_{3,N}(z) \end{bmatrix},$$

with $|\epsilon_{3,N}(z)|$, $|\tilde{\epsilon}_{3,N}(z)|$ bounded by $\mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$. We denote by $\Delta_N(z)$ the determinant of the above system, i.e.

$$\Delta_N(z) = (1 - u_N(z))^2 - zv_N(z)\tilde{v}_N(z).\tag{A.11}$$

The determinant $\Delta_N(z)$ was studied in [14] and in [23] where it was proved that $|\Delta_N(z)^{-1}| \leq \mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$ on a subset $\mathcal{D}_N$ of $\mathbb{C}$ defined as

$$\mathcal{D}_N = \left\{ z \in \mathbb{C} - \mathbb{R}, \frac{1}{N^2}\mathrm{Q}_1(|z|)\mathrm{Q}_2(|\mathrm{Im}(z)|^{-1}) < 1 \right\}$$

where Q_1 and Q_2 are two polynomials independent of N. Thus, we can invert the previous system on $\mathcal{D}_N$ to get

$$\begin{bmatrix} \alpha_N'(z) - \delta_N'(z) \\ \tilde{\alpha}_N'(z) - \tilde{\delta}_N'(z) \end{bmatrix} = \frac{1}{\Delta_N(z)} \begin{bmatrix} 1 - u_N(z) & zv_N(z) \\ z\tilde{v}_N(z) & 1 - u_N(z) \end{bmatrix} \frac{1}{N^2} \begin{bmatrix} \epsilon_{3,N}(z) \\ \tilde{\epsilon}_{3,N}(z) \end{bmatrix}.$$

This implies that $|\alpha_N'(z) - \delta_N'(z)|$ is bounded by $\frac{1}{N^2}\mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$ on $\mathcal{D}_N$. If $z \in \mathbb{C}\backslash\{\mathbb{R} \cup \mathcal{D}_N\}$, we use the trick in [12]. We remark that

$$|\alpha_N'(z) - \delta_N'(z)| \leq |\alpha_N'(z)| + |\delta_N'(z)| \leq \frac{C}{|\mathrm{Im}\,z|},$$

for each z, and that $1 \leq \frac{1}{N^2}\mathrm{Q}_1(|z|)\mathrm{Q}_2(|\mathrm{Im}(z)|^{-1})$ on $\mathbb{C}\backslash\{\mathbb{R} \cup \mathcal{D}_N\}$. Therefore,

$$|\alpha_N'(z) - \delta_N'(z)| \leq \frac{C}{|\mathrm{Im}\,z|}\frac{1}{N^2}\mathrm{Q}_1(|z|)\mathrm{Q}_2(|\mathrm{Im}(z)|^{-1}) \leq \frac{1}{N^2}\mathrm{P}_1(|z|)\mathrm{P}_2(|\mathrm{Im}(z)|^{-1})$$

on $\mathbb{C}\backslash\{\mathbb{R} \cup \mathcal{D}_N\}$. This in turn shows that (A.2) holds on $\mathbb{C}\backslash\mathbb{R}$. □

A.2. Proof of Lemma 3.4: Differentiability of $\frac{1}{M}\mathrm{Tr}\,\tilde{\psi}(\hat{\Omega}_N)$

We first need to establish the following useful lemma.

Lemma A.3. *Given an integer $D > 0$, let f be a continuous real function on $\mathbb{R}^D$. Let $\mathcal{O}$ be an open set of $\mathbb{R}^D$ such that $\mathbb{R}^D\backslash\mathcal{O}$ has a zero Lebesgue measure. Assume that f is a $\mathcal{C}^1$ function on $\mathcal{O}$ and that its gradient f' on $\mathcal{O}$ can be continuously extended to $\mathbb{R}^D$. Then f is $\mathcal{C}^1$ on the whole $\mathbb{R}^D$ with gradient f'.*

Proof. We only need to prove that for any $x \in \mathbb{R}^D - \mathcal{O}$ and any sequence $x_n \to x$,

$$f(x_n) - f(x) = \langle f'(x), x_n - x \rangle + o(d_n),$$

where $d_n = \|x_n - x\|$. Since f is uniformly continuous on any small neighborhood of x, there exists a sequence δ_n such that for every y and y' in this neighborhood for which $\|y - y'\| < \delta_n$, $|f(y) - f(y')| \leq d_n^2$. Since $\mathbb{R}^D - \mathcal{O}$ has a zero Lebesgue measure, there exists y_n and z_n in $\mathcal{O}$ such that

$$\|x_n - y_n\| < \min(\delta_n, d_n^2) \quad \text{and} \quad \|x - z_n\| < \min(\delta_n, d_n^2).$$

Therefore, it holds that $\max(|f(x_n) - f(y_n)|, |f(z_n) - f(x)|) < d_n^2$. Writing $f(x_n) - f(x) = f(x_n) - f(y_n) + f(y_n) - f(z_n) + f(z_n) - f(x)$, we obtain that $f(x_n) - f(x) = f(y_n) - f(z_n) + o(d_n)$. By differentiability of f on $\mathcal{O}$ and continuity of f' at x,

$$f(y_n) - f(z_n) = \langle f'(z_n), y_n - z_n \rangle + o(\|y_n - z_n\|) = \langle f'(x), x_n - x \rangle + o(d_n)$$

which proves the lemma. □

We now complete the proof of the lemma. We consider $\tilde{\psi} \in \mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R})$, and establish that, considered as a function of the real and imaginary parts of $\mathbf{W}_N$, function $\frac{1}{M}\mathrm{Tr}\,\tilde{\psi}(\hat{\boldsymbol{\Omega}}_N)$ is continuously differentiable on $\mathbb{R}^{2MN}$, i.e. that for each pair (i,j), the partial derivatives

$$\frac{\partial}{\partial W_{i,j,N}}\left\{\frac{1}{M}\mathrm{Tr}\,\tilde{\psi}(\hat{\boldsymbol{\Omega}}_N)\right\}$$

exist, and are continuous.[b] We denote by $\mathcal{O}$ the open subset of $\mathbb{R}^{2MN}$ for which the eigenvalues $(\hat{\lambda}_{l,N})_{l=1,\ldots,M}$ of $\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*$ have multiplicity 1. It is clear that $\mathbb{R}^{2MN}\backslash\mathcal{O}$ has a zero Lebesgue measure. On $\mathcal{O}$, it is standard that the eigenvalues $(\hat{\lambda}_{l,N})_{l=1,\ldots,M}$ are $\mathcal{C}^1$ functions and that

$$\frac{\partial\hat{\lambda}_{l,N}}{\partial W_{i,j,N}} = [\boldsymbol{\Sigma}_N^*\hat{\boldsymbol{\Pi}}_{l,N}]_{j,i}. \tag{A.12}$$

Using Lemma 4.6 in Haagerup–Thorbjornsen [12], we obtain

$$\frac{\partial}{\partial W_{i,j,N}}\{\mathrm{Tr}\,\tilde{\psi}(\hat{\boldsymbol{\Omega}}_N)\} = \mathrm{Tr}\left(\tilde{\psi}'(\hat{\boldsymbol{\Omega}}_N)\frac{\partial}{\partial W_{i,j,N}}\{\hat{\boldsymbol{\Omega}}_N\}\right)$$

$$= \left[\boldsymbol{\Sigma}_N^*\sum_{l=1}^M[\tilde{\psi}'(\hat{\boldsymbol{\Omega}}_N)]_{ll}\hat{\boldsymbol{\Pi}}_{l,N}\right]_{j,i} \tag{A.13}$$

and get that $\frac{1}{M}\mathrm{Tr}\,\tilde{\psi}(\hat{\boldsymbol{\Omega}}_N)$ is a $\mathcal{C}^1$ on $\mathcal{O}$. By Lemma A.3, it remains to establish that the right-hand side of (A.13) can be continuously extended to any point $\mathbf{W}_N^0$ of $\mathbb{R}^{2MN}\backslash\mathcal{O}$. For this, we first prove the following useful result.

Lemma A.4. *If* $\hat{\lambda}_{k,N} = \hat{\lambda}_{l,N}$, *then* $[\tilde{\psi}(\hat{\boldsymbol{\Omega}}_N)]_{kk} = [\tilde{\psi}(\hat{\boldsymbol{\Omega}}_N)]_{ll}$.

Proof. We start by observing that for any integers $m_1, m_2, \ldots, m_t$, matrix $\mathbf{A} = \hat{\boldsymbol{\Lambda}}_N^{m_1}\mathbf{1}\mathbf{1}^T\hat{\boldsymbol{\Lambda}}_N^{m_2}\cdots\mathbf{1}\mathbf{1}^T\hat{\boldsymbol{\Lambda}}_N^{m_t}$ writes

$$\mathbf{A} = \begin{bmatrix} \hat{\lambda}_{1,N}^{m_1} & \cdots & \hat{\lambda}_{1,N}^{m_1} \\ \vdots & \vdots & \vdots \\ \hat{\lambda}_{M,N}^{m_1} & \cdots & \hat{\lambda}_{M,N}^{m_1} \end{bmatrix}\cdots\begin{bmatrix} \hat{\lambda}_{1,N}^{m_{t-1}} & \cdots & \hat{\lambda}_{1,N}^{m_{t-1}} \\ \vdots & \vdots & \vdots \\ \hat{\lambda}_{M,N}^{m_{t-1}} & \cdots & \hat{\lambda}_{M,N}^{m_{t-1}} \end{bmatrix}\hat{\boldsymbol{\Lambda}}_N^{m_t},$$

hence $[\mathbf{A}]_{kk} = [\mathbf{A}]_{ll}$ if $\hat{\lambda}_{k,N} = \hat{\lambda}_{l,N}$. The same can be said about $\mathbf{1}\mathbf{1}^T\mathbf{A}$ and $\mathbf{A}\mathbf{1}\mathbf{1}^T$. Consequently, the result of the lemma is true when $\tilde{\psi}$ is a polynomial. Since any continuous function $\tilde{\psi}$ is the uniform limit of a sequence of polynomials on compact subsets of $\mathbb{R}$, the result is true for such $\tilde{\psi}$. $\square$

---

[b]$\tilde{\psi}$ is real-valued, the partial derivatives with respect to $\overline{W}_{i,j,N}$ thus coincide with the complex conjugate of the partial derivative with respect to $W_{i,j,N}$. It is therefore sufficient to consider these derivatives.

We consider an element $\mathbf{W}_N^0$ of $\mathbb{R}^{2MN}\backslash\mathcal{O}$, and denote by $m_1,\ldots,m_L$, with $M = \sum_{l=1}^L m_l$, the respective multiplicities of the eigenvalues of $\boldsymbol{\Sigma}_N^0\boldsymbol{\Sigma}_N^{0*}$ where $\boldsymbol{\Sigma}_N^0 = \mathbf{B}_N + \mathbf{W}_N^0$. We also denote by $(\overline{\boldsymbol{\Pi}}_{l,N})_{l=1,\ldots,L}$ the orthogonal projection matrices over the corresponding eigenspaces. Lemma A.4 implies that for each $i = 1,\ldots,L$,

$$[\tilde{\psi}'(\hat{\boldsymbol{\Omega}})]_{m_1+\cdots+m_i,m_1+\cdots+m_i}$$

$$= \cdots = [\tilde{\psi}'(\hat{\boldsymbol{\Omega}})]_{m_1+\cdots+m_i+m_{i+1}-1,m_1+\cdots+m_i+m_{i+1}-1} = \kappa_i.$$

Therefore, for any sequence $(\mathbf{W}_{N,n})_{n\in\mathbb{N}}$ converging toward $\mathbf{W}_N^0$, it holds that

$$\lim_{n\to\infty}\frac{\partial}{\partial W_{i,j,N}}\left\{\frac{1}{M}\mathrm{Tr}\,\tilde{\psi}(\hat{\boldsymbol{\Omega}}_N)\right\}\bigg|_{\mathbf{W}_N=\mathbf{W}_{N,n}} = \left[\boldsymbol{\Sigma}_N^*\sum_{l=1}^L\kappa_l\overline{\boldsymbol{\Pi}}_{l,N}\right]_{j,i}.$$

This completes the proof of Lemma 3.4.

A.3. *Proof of Lemma 3.5: Uniform boundedness of* $\mathbb{E}[\|\mathbf{W}_N\|^p]$

It is clear that it is sufficient to prove the boundedness of $\mathbb{E}[\|\mathbf{W}_N\|^p]$ if the entries of $\mathbf{W}_N$ are real. We thus consider the case of real matrices and denote by X_N the largest singular value of $\frac{\mathbf{W}_N}{\sigma}$. The following concentration result is well-known.

Theorem A.1 ([7, Theorem II.13]). *Under Assumption* A-4, *it holds that* $\mathbb{E}[X_N] \leq 1 + \sqrt{c_N}$ *and for all* $t > 0$, $\mathbb{P}(X_N > 1 + \sqrt{c_N} + t) \leq \exp(-Nt^2/2)$.

Using Theorem A.1 and for $p \geq 2$ the inequality,

$$\mathbb{E}[X_N^p] = \int_0^{+\infty}\mathbb{P}(X_N \geq t)pt^{p-1}\mathrm{d}t$$

$$\leq p(1 + \sqrt{c_N})^p + \int_0^{+\infty}\mathbb{P}(X_N \geq t + 1 + \sqrt{c_N})p(t + 1 + \sqrt{c_N})^{p-1}\mathrm{d}t,$$

we easily obtain $\mathbb{E}[X_N^p] \leq K < \infty$, with K a constant independent of N, for all $p \in \mathbb{N}$.

A.4. *Proof of Lemma 3.9: Differentiability of the regularization factor*

We first establish that $\det\phi(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*)$ is a $\mathcal{C}^1$ function, and that (3.28) holds. We use the same approach as in Haagerup–Thorbjornsen [12, Lemma 4.6]. We begin by showing that the differential of $\det\phi(\mathbf{X})$ is given by

$$\det\phi(\mathbf{X})' \cdot \mathbf{H} = \mathrm{Tr}(\mathrm{adj}(\phi(\mathbf{X}))\phi'(\mathbf{X})\mathbf{H}). \qquad (\mathrm{A}.14)$$

As $\det(\mathbf{X})' \cdot \mathbf{H} = \mathrm{Tr}(\mathrm{adj}(\mathbf{X})\mathbf{H})$ and $(\mathbf{X}^n)' \cdot \mathbf{H} = \sum_{i=0}^{n-1}\mathbf{X}^i\mathbf{H}\mathbf{X}^{n-1-i}$ for any $n \in \mathbb{N}$, we have

$$\det(\mathbf{X}^n)' \cdot \mathbf{H} = \mathrm{Tr}(\mathrm{adj}(\mathbf{X}^n)(n\mathbf{X}^{n-1})\mathbf{H})$$

since $\mathrm{adj}(\mathbf{X}^n)$ and $\mathbf{X}$ commute. So (A.14) is true when ϕ is a polynomial. By choosing a sequence of polynomials P_n such that $P_n \to \phi$ and $P_n' \to \phi'$ uniformly

on compact subsets of $\mathbb{R}$, we generalize (A.14) to any $\phi \in \mathcal{C}_1$. Now one can check that

$$\frac{\partial(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*)}{\partial W_{i,j,N}} = \mathbf{e}_i \mathbf{e}_j^* \mathbf{\Sigma}_N^*, \tag{A.15}$$

and it remains to apply the composition formula for differentials to obtain (3.28).

We also remark that at a point $\mathbf{W}_N$ for which there exists a $\hat{\lambda}_{l,N} \notin \mathrm{supp}(\phi)$, we have

$$\mathrm{adj}(\phi(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*))\phi'(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*) = \sum_{l=1}^{M} \left(\prod_{k \neq l} \phi(\hat{\lambda}_{k,N}) \right) \phi'(\hat{\lambda}_{l,N}) \hat{\mathbf{u}}_{l,N} \hat{\mathbf{u}}_{l,N}^* = \mathbf{0},$$

hence the derivative (3.28) is zero on $\mathcal{A}_{1,N}^c$.

It is easy to check that $\det \phi(\hat{\mathbf{\Omega}}_N)$ is a $\mathcal{C}^1$ function on the open set $\mathcal{O}$ of all matrices $\mathbf{W}_N$ for which the eigenvalues of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ are simple, and that (3.29) holds if $\mathbf{W}_N \in \mathcal{O}$, i.e. on a set of probability one. In order to show that $\det \phi(\hat{\mathbf{\Omega}}_N)$ is a $\mathcal{C}^1$ function on $\mathbb{R}^{2MN} \backslash \mathcal{O}$, we use again Lemma A.3, and verify that (3.29) can be continuously extended to $\mathbb{R}^{2MN} \backslash \mathcal{O}$. For this, we claim that

$$[\mathrm{adj}(\phi(\hat{\mathbf{\Omega}}_N))\phi'(\hat{\mathbf{\Omega}}_N)]_{k,k} = [\mathrm{adj}(\phi(\hat{\mathbf{\Omega}}_N))\phi'(\hat{\mathbf{\Omega}}_N)]_{l,l} \tag{A.16}$$

if $\hat{\lambda}_{k,N} = \hat{\lambda}_{l,N}$. Indeed, given $\varepsilon > 0$, let $\phi_\varepsilon(x) = \phi(x) + \varepsilon$. Since $\phi_\varepsilon(\hat{\mathbf{\Omega}}_N) > 0$,

$$\mathrm{adj}(\phi_\varepsilon(\hat{\mathbf{\Omega}}_N))\phi_\varepsilon'(\hat{\mathbf{\Omega}}_N) = \det(\phi_\varepsilon(\hat{\mathbf{\Omega}}_N))\phi_\varepsilon^{-1}(\hat{\mathbf{\Omega}}_N)\phi_\varepsilon'(\hat{\mathbf{\Omega}}_N).$$

Applying Lemma A.4 to $\tilde{\psi} = \phi_\varepsilon^{-1} \times \phi_\varepsilon'$, we obtain that

$$[\mathrm{adj}(\phi_\varepsilon(\hat{\mathbf{\Omega}}_N))\phi_\varepsilon'(\hat{\mathbf{\Omega}}_N)]_{kk} = [\mathrm{adj}(\phi_\varepsilon(\hat{\mathbf{\Omega}}_N))\phi_\varepsilon'(\hat{\mathbf{\Omega}}_N)]_{ll} \quad \text{if } \hat{\lambda}_{k,N} = \hat{\lambda}_{l,N}$$

and letting $\varepsilon \to 0$, we obtain the same result for $\mathrm{adj}(\phi(\hat{\mathbf{\Omega}}_N))\phi'(\hat{\mathbf{\Omega}}_N)$. Similarly to the proof of Lemma 3.4, this proves that (3.29) can be continuously extended to $\mathbb{R}^{2MN} \backslash \mathcal{O}$.

A.5. *Proof of Lemma* 3.10: *Various estimates*

In this section, we denote by $\alpha_{r,N}(z), \tilde{\alpha}_{r,N}(z), \mathbf{R}_{r,N}(z)$ and $\tilde{\mathbf{R}}_{r,N}(z)$ the regularized versions of the respective functions $\alpha_N(z), \tilde{\alpha}_N(z), \mathbf{R}_N(z)$ and $\tilde{\mathbf{R}}_N(z)$ defined in Sec. A.1, i.e.

$$\alpha_{r,N}(z) = \sigma \mathbb{E}\left(\frac{1}{N} \mathrm{Tr}\,(\mathbf{Q}_N(z))\chi_N \right) \quad \text{and} \quad \tilde{\alpha}_{r,N}(z) = \sigma \mathbb{E}\left(\frac{1}{N} \mathrm{Tr}\,(\tilde{\mathbf{Q}}_N(z))\chi_N \right),$$

and

$$\mathbf{R}_{r,N}(z) = \left(\frac{\mathbf{B}_N \mathbf{B}_N^*}{1 + \sigma \alpha_{r,N}(z)} - z(1 + \sigma \tilde{\alpha}_{r,N}(z)) \right)^{-1},$$

$$\tilde{\mathbf{R}}_{r,N}(z) = \left(\frac{\mathbf{B}_N^* \mathbf{B}_N}{1 + \sigma \tilde{\alpha}_{r,N}(z)} - z(1 + \sigma \alpha_{r,N}(z)) \right)^{-1}.$$

It is clear that $\alpha_{r,N}$ and $\tilde{\alpha}_{r,N}$ are the Stieltjes transforms of positive measures carried by $\mathbb{C}\backslash\mathrm{supp}(\phi)$ and $\mathbb{C}^*\backslash\mathrm{supp}(\phi)$ respectively and with mass $\sigma c_N \mathbb{E}[\chi_N]$ and $\sigma\mathbb{E}[\chi_N]$. This implies that the following uniform bounds hold: Let $\mathcal{K}$ and $\tilde{\mathcal{K}}$ be compact subsets of $\mathbb{C}\backslash\mathrm{supp}(\phi)$ and $\mathbb{C}^*\backslash\mathrm{supp}(\phi)$ respectively, then we have

$$\sup_{z\in\mathcal{K}}|\alpha_{r,N}(z)| < C \quad \text{and} \quad \sup_{z\in\tilde{\mathcal{K}}}|\tilde{\alpha}_{r,N}(z)| < C. \tag{A.17}$$

In order to establish Lemma 3.10, it is necessary to show that similar bounds hold for functions $\frac{1}{1+\sigma\alpha_{r,N}(z)}$, $\|\mathbf{R}_{r,N}(z)\|$ and $\|\tilde{\mathbf{R}}_{r,N}(z)\|$. For this, we introduce function $w_{r,N}(z) = z(1+\sigma\alpha_{r,N}(z))(1+\sigma\tilde{\alpha}_{r,N}(z))$ and prove the following lemma.

Lemma A.5. *Under Assumptions* A-1–A-6, *for any compact subset* $\mathcal{K}$ *of* $\mathbb{C}\backslash\mathrm{supp}(\phi)$, *it holds that*

$$\sup_{z\in\mathcal{K}}|\alpha_{r,N}(z) - \delta_N(z)| \xrightarrow[N\to\infty]{} 0, \tag{A.18}$$

$$\inf_{z\in\mathcal{K}}\min_{k=1,\ldots,M}|\lambda_{k,N} - w_{r,N}(z)| > C > 0. \tag{A.19}$$

Proof. Define $\kappa_N(z) := \alpha_{r,N}(z) - \delta_N(z)$ where we recall that $\delta_N(z) = \sigma c_N m_N(z) = \frac{\sigma}{N}\mathrm{Tr}\,(\mathbf{T}_N(z))$. Since $\delta_N(z)$ and $\alpha_{r,N}(z)$ are Stieltjes transforms of positive measures carried by $\mathbb{C}\backslash\mathrm{supp}(\phi)$, κ_N is holomorphic on $\mathbb{C}\backslash\mathrm{supp}(\phi)$ and satisfies

$$|\kappa_N(z)| \leq \frac{C}{\mathrm{d}(z,\mathrm{supp}(\phi))}.$$

This implies that the sequence (κ_N) is uniformly bounded on each compact subset of $\mathbb{C}\backslash\mathrm{supp}(\phi)$. By Montel's theorem, (κ_N) is a normal family. Let $(\kappa_{\psi(N)})$ a subsequence of (κ_N) which converges uniformly to κ on each compact subset of $\mathbb{C}\backslash\mathrm{supp}(\phi)$. Then κ is holomorphic on $\mathbb{C}\backslash\mathrm{supp}(\phi)$. From [23, Proposition 6], $\mathbb{E}[\frac{1}{N}\mathrm{Tr}\,\mathbf{Q}_N(z)] - \frac{1}{N}\mathrm{Tr}\,\mathbf{T}_N(z) \xrightarrow[N]{} 0$ for $z\in\mathbb{C}\backslash\mathbb{R}^+$ and since $\chi_N \xrightarrow[N]{} 1$ a.s., dominated convergence theorem implies

$$\kappa_N(z) = \mathbb{E}\left[\frac{\sigma}{N}\mathrm{Tr}\,\mathbf{Q}_N(z)\chi_N\right] - \frac{\sigma}{N}\mathrm{Tr}\,\mathbf{T}_N(z) \xrightarrow[N]{} 0$$

for $z\in\mathbb{C}\backslash\mathbb{R}^+$. Thus, $\kappa(z) = 0$ for $z\in\mathbb{C}\backslash\mathbb{R}^+$, and by analytic continuation, $\kappa(z) = 0$ for all $z\in\mathbb{C}\backslash\mathrm{supp}(\phi)$. Therefore, all converging subsequences extracted from the normal family $(\kappa_N(z))$ converge to 0 uniformly on each compact subset of $\mathbb{C}\backslash\mathrm{supp}(\phi)$. Consequently, the whole sequence (κ_N) converges uniformly to 0 on each compact subset of $\mathbb{C}\backslash\mathrm{supp}(\phi)$. This completes the proof of (A.18). We also notice that

$$\tilde{\alpha}_{r,N}(z) = \alpha_{r,N}(z) - \frac{\sigma(1-c_N)}{z} + \frac{\sigma(1-c_N)}{z}(1-\mathbb{E}(\chi_N)) \tag{A.20}$$

and recall that $\tilde{\delta}_N(z) = \delta_N(z) - \frac{\sigma(1-c_N)}{z}$. As $1-\mathbb{E}(\chi_N) = \mathcal{O}(\frac{1}{N^p})$ for each integer p, (A.18) implies

$$\sup_{z\in\mathcal{K}}|z(\tilde{\alpha}_{r,N}(z) - \tilde{\delta}_N(z))| \to 0.$$

Hence, it holds that

$$\sup_{z \in \mathcal{K}} |w_{r,N}(z) - w_N(z)| \to 0.$$

Thus, (A.19) follows immediately from (2.12). □

Lemma A.5 immediately implies that the following uniform bounds hold.

Lemma A.6. *Let $\mathcal{K}$ and $\tilde{\mathcal{K}}$ be compact subsets of $\mathbb{C} - \operatorname{supp}(\phi)$ and $\mathbb{C}^* - \operatorname{supp}(\phi)$ respectively. Under Assumptions A-1–A-6, for N large enough, we have*

$$\sup_{z \in \mathcal{K}} \left| \frac{1}{1 + \sigma \alpha_{r,N}(z)} \right| < C, \tag{A.21}$$

$$\sup_{z \in \mathcal{K}} \|\mathbf{R}_{r,N}(z)\| < C, \tag{A.22}$$

$$\sup_{z \in \tilde{\mathcal{K}}} \|\tilde{\mathbf{R}}_{r,N}(z)\| < C, \tag{A.23}$$

$$\sup_{z \in \mathcal{K}} \|\mathbf{R}_{r,N}(z) - \mathbf{T}_N(z)\| \to 0, \tag{A.24}$$

$$\sup_{z \in \tilde{\mathcal{K}}} \|\tilde{\mathbf{R}}_{r,N}(z) - \tilde{\mathbf{T}}_N(z)\| \to 0. \tag{A.25}$$

Proof. We first recall that inequality (2.9) holds. Therefore, the uniform convergence result (A.18) implies that

$$\inf_{z \in \mathcal{K}} |1 + \sigma \alpha_{r,N}(z)| > \frac{1}{4}$$

for N large enough. This establishes (A.21) that holds for N large enough. In order to prove (A.22), we express $\mathbf{R}_{r,N}(z)$ as

$$\mathbf{R}_{r,N}(z) = (1 + \sigma \alpha_{r,N}(z))(\mathbf{B}_N \mathbf{B}_N^* - w_{r,N}(z))^{-1}$$

and use (A.17) and (A.19). The proof of (A.23) is similar, and is based on the identity

$$\tilde{\mathbf{R}}_{r,N}(z) = (1 + \sigma \tilde{\alpha}_{r,N}(z))(\mathbf{B}_N^* \mathbf{B}_N - w_{r,N}(z))^{-1}.$$

We remark that function $\tilde{\alpha}_{r,N}(z)$ has a pole at $z = 0$. Hence, any compact $\tilde{\mathcal{K}}$ over which $\|\tilde{\mathbf{R}}_{r,N}(z)\|$ is supposed to be uniformly bounded should not contain 0. The proof of (A.24) follows immediately from (A.18) and from (A.21), (A.22), (A.23). Finally, to establish (A.25), we remark that

$$\tilde{\mathbf{R}}_{r,N}(z) = \frac{\mathbf{B}_N^* \mathbf{R}_{r,N}(z) \mathbf{B}_N}{w_{r,N}(z)} - \frac{\mathbf{I}_N}{1 + \sigma \alpha_{r,N}(z)},$$

$$\tilde{\mathbf{T}}_{r,N}(z) = \frac{\mathbf{B}_N^* \mathbf{T}_N(z) \mathbf{B}_N}{w_N(z)} - \frac{\mathbf{I}_N}{1 + \sigma \delta_N(z)},$$

and that $|w_{r,N}(z)|$ and $|w_N(z)|$ are uniformly bounded from below by (2.12) and (A.19) (recall that 0 is one of the eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$). □

We now establish (3.34) and (3.35). In order to prove that $\alpha_N(z) - \delta_N(z) = \mathcal{O}(\frac{1}{N^2})$ on $\mathbb{C}\backslash\mathbb{R}^+$, [10, 23] used the integration by parts formula (see e.g. [19]) and the Poincaré inequality to show that the entries of $\mathbb{E}[\mathbf{Q}_N(z)]$ are close from the entries of $\mathbf{R}_N(z)$ (see the fundamental equation (A.4)). Then, $\alpha_N(z) - \delta_N(z)$ was evaluated by solving a linear system whose determinant $\Delta_N(z)$ given by (A.11) was shown to be bounded from below. Lemma A.6 allows to follow exactly the same approach to establish (3.34) and (3.35). However, functions $\alpha_N, \tilde{\alpha}_N, \mathbf{R}_N, \tilde{\mathbf{R}}_N$ have to be replaced by their regularized versions. The following results show that the presence of the regularization term χ_N does not modify essentially the calculations of [10, 23]. We first indicate how the integration by parts formula is modified. $\mathrm{Vec}(\cdot)$ denotes the column by column vectorization operator of a matrix.

Lemma A.7. *Let $(f_N)_{N\geq 1}$ be a sequence of continuously differentiable functions defined on $\mathbb{C}^{M(M+N)}$ with polynomially bounded partial derivatives satisfying the condition*

$$\sup_{z\in\partial\mathcal{R}_y} |f_N(\mathrm{Vec}(\mathbf{Q}_N(z)), \mathrm{Vec}(\mathbf{\Sigma}_N))\chi_N| < C.$$

Then, under Assumptions A-1–A-6, for all $p \in \mathbb{N}$, we have

$$\mathbb{E}[f(\mathrm{Vec}(\mathbf{Q}_N(z)), \mathrm{Vec}(\mathbf{\Sigma}_N))\chi_N] = \mathbb{E}[f(\mathrm{Vec}(\mathbf{Q}_N(z)), \mathrm{Vec}(\mathbf{\Sigma}_N))\chi_N^k] + \frac{\epsilon_{1,N}(z)}{N^p},$$
(A.26)

for all $k \in \mathbb{N}^*$, and

$$\mathbb{E}[W_{ij,N} f(\mathrm{Vec}(\mathbf{Q}_N(z)), \mathrm{Vec}(\mathbf{\Sigma}_N))\chi_N]$$
$$= \frac{\sigma^2}{N}\mathbb{E}\left[\frac{\partial f(\mathrm{Vec}(\mathbf{Q}_N(z)), \mathrm{Vec}(\mathbf{\Sigma}_N))}{\partial \overline{W}_{ij,N}}\chi_N\right] + \frac{\epsilon_{2,N}(z)}{N^p}, \qquad \text{(A.27)}$$

$$\mathbb{E}[\overline{W}_{ij,N} f(\mathrm{Vec}(\mathbf{Q}_N(z)), \mathrm{Vec}(\mathbf{\Sigma}_N))\chi_N]$$
$$= \frac{\sigma^2}{N}\mathbb{E}\left[\frac{\partial f(\mathrm{Vec}(\mathbf{Q}_N(z)), \mathrm{Vec}(\mathbf{\Sigma}_N))}{\partial W_{ij,N}}\chi_N\right] + \frac{\epsilon_{3,N}(z)}{N^p}, \qquad \text{(A.28)}$$

with $\sup_{z\in\partial\mathcal{R}_y} |\epsilon_{i,N}(z)| \leq C < \infty$.

As for the use of the Poincaré inequality, we have:

Lemma A.8. *Let $(\mathbf{M}_N(z))$ a sequence of deterministic complex $M \times M$ matrix-valued functions defined on $\mathbb{C}\backslash\mathbb{R}$ such that*

$$\sup_{z\in\partial\mathcal{R}_y} \|\mathbf{M}_N(z)\| \leq C.$$

Then, under Assumptions A-1–A-6,

$$\sup_{z\in\partial\mathcal{R}_y} \mathrm{Var}\left[\frac{1}{N}\mathrm{Tr}\,\mathbf{Q}_N(z)\mathbf{M}_N(z)\chi_N\right] \leq \frac{C}{N^2},$$

and for $\mathbf{a}_N \in \mathbb{C}^M$ *such that* $\sup_N \|\mathbf{a}_N\| < \infty,$

$$\sup_{z \in \partial \mathcal{R}_y} \mathrm{Var}[\mathbf{a}_N^* \mathbf{Q}_N(z) \mathbf{M}_N(z) \mathbf{a}_N \chi_N] \leq \frac{C}{N}.$$

Moreover, the same kind of uniform bounds still hold when $\mathbf{Q}_N(z)$ *is replaced by* $\mathbf{Q}_N(z)^2.$

The proofs of these results are based on elementary arguments, and are thus omitted. Following the calculations of [10, 23], we obtain that

$$\mathbb{E}[\mathbf{Q}_N(z)\chi_N] = \mathbf{R}_{r,N}(z) + \boldsymbol{\Delta}_{r,N}(z)\mathbf{R}_{r,N}(z)$$

$$+ \mathbb{E}[\mathbf{Q}_N(z)\chi_N]\mathbf{R}_{r,N}(z)\frac{\sigma^2}{N}\mathrm{Tr}\,\boldsymbol{\Delta}_{r,N}(z) + \boldsymbol{\Theta}_N(z)\mathbf{R}_{r,N}(z) \quad \text{(A.29)}$$

for each $z \in \mathbb{C}\backslash\mathrm{supp}(\phi)$ where $\boldsymbol{\Theta}_N(z)$ is a matrix whose elements are uniformly bounded on $\partial\mathcal{R}_y$ by $\frac{C}{N^p}$ for each p, and where $\boldsymbol{\Delta}_{r,N}(z)$ is the regularized version of matrix $\boldsymbol{\Delta}_N(z)$ introduced in Lemma A.2 defined by

$$\boldsymbol{\Delta}_{r,N}(z) = -\frac{1}{(1+\sigma\alpha_{r,N}(z))^2}\mathbb{E}[\mathbf{Q}_N(z)\chi_N]$$

$$\times \mathbb{E}\left[\left(\frac{\sigma^2}{N}\mathrm{Tr}\,\mathbf{Q}_N(z)\chi_N - \mathbb{E}\left[\frac{\sigma^2}{N}\mathrm{Tr}\,\mathbf{Q}_N(z)\chi_N\right]\right)\frac{\sigma^2}{N}\mathrm{Tr}\,\boldsymbol{\Sigma}_N^*\mathbf{Q}_N(z)\mathbf{B}_N\chi_N\right]$$

$$+ \frac{1}{1+\sigma\alpha_{r,N}(z)}\mathbb{E}\left[\left(\frac{\sigma^2}{N}\mathrm{Tr}\,\boldsymbol{\Sigma}_N^*\mathbf{Q}_N(z)\mathbf{B}_N\chi_N\right.\right.$$

$$\left.\left. - \mathbb{E}\left[\frac{\sigma^2}{N}\mathrm{Tr}\,\boldsymbol{\Sigma}_N^*\mathbf{Q}_N(z)\mathbf{B}_N\chi_N\right]\right)\mathbf{Q}_N(z)\chi_N\right]$$

$$+ \frac{1}{1+\sigma\alpha_{r,N}(z)}\mathbb{E}\left[\left(\frac{\sigma^2}{N}\mathrm{Tr}\,\mathbf{Q}_N(z)\chi_N - \mathbb{E}\left[\frac{\sigma^2}{N}\mathrm{Tr}\,\mathbf{Q}_N(z)\chi_N\right]\right)\right.$$

$$\left. \times \mathbf{Q}_N(z)\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\chi_N\right]. \quad \text{(A.30)}$$

After some calculations using Lemmas A.6–A.8, we eventually obtain that

$$\sup_{z \in \partial\mathcal{R}_y} |\mathbf{a}_N^*(\mathbb{E}[\mathbf{Q}_N(z)\chi_N] - \mathbf{R}_{r,N}(z))\mathbf{a}_N| \leq \frac{C}{N^{\frac{3}{2}}},$$

$$\sup_{z \in \partial\mathcal{R}_y} \left|\alpha_{r,N}(z) - \frac{\sigma}{N}\mathrm{Tr}\,(\mathbf{R}_{r,N}(z))\right| \leq \frac{C}{N^2}, \quad \text{(A.31)}$$

$$\sup_{z \in \partial\mathcal{R}_y} \left|\tilde{\alpha}_{r,N}(z) - \frac{\sigma}{N}\mathrm{Tr}\,(\tilde{\mathbf{R}}_{r,N}(z))\right| \leq \frac{C}{N^2}, \quad \text{(A.32)}$$

for all large N. In order to prove (3.34) and (3.35), it remains to handle the terms involving the difference $\mathbf{R}_{r,N}(z) - \mathbf{T}_N(z)$. We show in the following that

$$\sup_{z \in \partial\mathcal{R}_y} |\mathbf{a}_N^*(\mathbf{R}_{r,N}(z) - \mathbf{T}_N(z))\mathbf{a}_N| \leq \frac{C}{N^2} \quad \text{(A.33)}$$

for all large N. We start as usual with the identity

$$\mathbf{R}_{r,N}(z) - \mathbf{T}_N(z) = \mathbf{R}_{r,N}(z)(\mathbf{T}_N(z)^{-1} - \mathbf{R}_{r,N}(z)^{-1})\mathbf{T}_N(z),$$

to get

$$\mathbf{a}_N^*(\mathbf{R}_{r,N}(z) - \mathbf{T}_N(z))\mathbf{a}_N$$

$$= \sigma \frac{\alpha_{r,N}(z) - \delta_N(z)}{(1 + \sigma\alpha_{r,N}(z))(1 + \sigma\delta_N(z))}\mathbf{a}_N^*\mathbf{R}_{r,N}(z)\mathbf{B}_N\mathbf{B}_N^*\mathbf{T}_N(z)\mathbf{a}_N$$

$$+ z\sigma(\tilde{\alpha}_{r,N}(z) - \tilde{\delta}_N(z))\mathbf{a}_N^*\mathbf{R}_{r,N}(z)\mathbf{T}_N(z)\mathbf{a}_N.$$

The expression (A.20) of $\tilde{\alpha}_{r,N}$ implies that $z(\tilde{\alpha}_{r,N}(z) - \tilde{\delta}_N(z)) = z(\alpha_{r,N}(z) - \delta_N(z)) + \mathcal{O}(\frac{1}{N^p})$ for each integer p. Thus, to prove (3.34) and (3.35), it is sufficient to check that

$$\sup_{z\in\partial\mathcal{R}_y} |\alpha_{r,N}(z) - \delta_N(z)| \leq \frac{C}{N^2}.$$

We will use the same ideas as in Sec. A.1 and remark that $(\alpha_{r,N}(z) - \delta_N(z), \tilde{\alpha}_{r,N}(z) - \tilde{\delta}_N(z))$ can be interpreted as the solution of a 2×2 linear system whose determinant is a regularized version of (A.11), and appears uniformly bounded away from zero on $\partial\mathcal{R}_y$.

Using again the previous expression of $\mathbf{R}_{r,N}(z) - \mathbf{T}_N(z)$ together with (A.31), (A.32) and repeating the procedure for $\tilde{\mathbf{R}}_{r,N}(z) - \tilde{\mathbf{T}}_N(z)$, we obtain

$$\begin{bmatrix} \alpha_{r,N}(z) - \delta_N(z) \\ \tilde{\alpha}_{r,N}(z) - \tilde{\delta}_N(z) \end{bmatrix} = \begin{bmatrix} u_{r,N}(z) & zv_{r,N}(z) \\ z\tilde{v}_{r,N}(z) & u_{r,N}(z) \end{bmatrix} \begin{bmatrix} \alpha_{r,N}(z) - \delta_N(z) \\ \tilde{\alpha}_{r,N}(z) - \tilde{\delta}_N(z) \end{bmatrix}$$

$$+ \frac{1}{N^2}\begin{bmatrix} \epsilon_N(z) \\ \tilde{\epsilon}_N(z) \end{bmatrix}, \tag{A.34}$$

with

$$u_{r,N}(z) = \frac{\sigma^2}{N}\mathrm{Tr}\,\frac{\mathbf{R}_{r,N}(z)\mathbf{B}_N^*\mathbf{B}_N\mathbf{T}_N(z)}{(1 + \sigma\alpha_{r,N}(z))(1 + \sigma\delta_N(z))}, \quad v_{r,N}(z) = \frac{\sigma^2}{N}\mathrm{Tr}\,\mathbf{R}_{r,N}(z)\mathbf{T}_N(z)$$

and

$$\tilde{v}_{r,N}(z) = \frac{\sigma^2}{N}\mathrm{Tr}\,\tilde{\mathbf{R}}_{r,N}(z)\tilde{\mathbf{T}}_N(z).$$

The quantities $\epsilon_N(z)$, $\tilde{\epsilon}_N(z)$ satisfy $\sup_{z\in\partial\mathcal{R}_y}|\epsilon_N(z)| < C$, $\sup_{z\in\partial\mathcal{R}_y}|\tilde{\epsilon}_N(z)| < C$. The determinant of the system is given by

$$\Delta_{r,N}(z) = (1 - u_{r,N}(z))^2 - z^2 v_{r,N}(z)\tilde{v}_{r,N}(z).$$

Lemma A.6 implies that for all large N, $u_{r,N}(z)$, $v_{r,N}(z)$ and $\tilde{v}_{r,N}(z)$ are uniformly bounded on $\partial\mathcal{R}_y$. Therefore, to conclude the proof of (A.33), it remains to check that for all large N,

$$\inf_{z\in\partial\mathcal{R}_y} |\Delta_{r,N}(z)| \geq C > 0.$$

Consider the function $\check{\Delta}_N(z)$ where we have replaced the matrix $\mathbf{R}_{r,N}(z)$ and $\tilde{\mathbf{R}}_{r,N}(z)$ by $\mathbf{T}_N(z)$ and $\tilde{\mathbf{T}}_N(z)$, i.e.

$$\check{\Delta}_N(z) = (1 - \check{u}_N(z))^2 - z^2 \check{v}_N(z)\check{\tilde{v}}_N(z),$$

with

$$\check{u}_N(z) = \frac{\sigma^2}{N}\mathrm{Tr}\,\frac{\mathbf{T}_N(z)\mathbf{B}_N^*\mathbf{B}_N\mathbf{T}_N(z)}{(1 + \sigma\delta_N(z))(1 + \sigma\tilde{\delta}_N(z))}, \quad \check{v}_N(z) = \frac{\sigma^2}{N}\mathrm{Tr}\,\mathbf{T}_N(z)^2,$$

and

$$\check{\tilde{v}}_N(z) = \frac{\sigma^2}{N}\mathrm{Tr}\,\tilde{\mathbf{T}}_N(z)^2.$$

Lemmas A.5 and A.6 imply that $|\check{u}_N(z) - u_{r,N}(z)|$, $|v_{r,N}(z) - \check{v}_N(z)|$ and $|\tilde{v}_{r,N}(z) - \check{\tilde{v}}_N(z)|$ converge to 0 uniformly on $\partial\mathcal{R}_y$ which of course implies

$$\sup_{z\in\partial\mathcal{R}_y} |\Delta_N(z) - \check{\Delta}_N(z)| \xrightarrow[N\to\infty]{} 0. \tag{A.35}$$

Using Cauchy–Schwarz inequality, we get

$$|\check{\Delta}_N(z)| \geq \underline{\Delta}_N(z) := (1 - \underline{u}_N(z))^2 - |z|^2 \underline{v}_N(z)\underline{\tilde{v}}_N(z),$$

with

$$\underline{u}_N(z) = \frac{\sigma^2}{N}\mathrm{Tr}\,\frac{\mathbf{T}_N(z)\mathbf{B}_N\mathbf{B}_N^*\mathbf{T}_N(z)^*}{|1 + \sigma\delta_N(z)|^2},$$

$$\underline{v}_N(z) = \frac{\sigma^2}{N}\mathrm{Tr}\,\mathbf{T}_N(z)\mathbf{T}_N(z)^*$$

and

$$\underline{\tilde{v}}_N(z) = \frac{\sigma^2}{N}\mathrm{Tr}\,\tilde{\mathbf{T}}_N(z)\tilde{\mathbf{T}}_N(z)^*.$$

Now, we use the following lemma.

Lemma A.9. *Under Assumptions* A-1–A-6, *there exists a constant* $C > 0$ *independent of* N *such that*

$$\inf_{z\in\partial\mathcal{R}_y} \underline{\Delta}_N(z) \geq C.$$

Proof. It is shown in [23, 14] that $\underline{\Delta}_N(z)$ is the determinant of the following 2×2 linear system

$$\begin{bmatrix} \mathrm{Im}(\delta_N(z)) \\ \mathrm{Im}(z\tilde{\delta}_N(z)) \end{bmatrix} = \begin{bmatrix} \underline{u}_N(z) & \underline{v}_N(z) \\ |z|^2\underline{\tilde{v}}_N(z) & \underline{u}_N(z) \end{bmatrix} \begin{bmatrix} \mathrm{Im}(\delta_N(z)) \\ \mathrm{Im}(z\tilde{\delta}_N(z)) \end{bmatrix} + \frac{\mathrm{Im}(z)}{\sigma} \begin{bmatrix} \underline{v}_N(z) \\ \underline{u}_N(z) \end{bmatrix}, \tag{A.36}$$

and that for $z \in \mathbb{C}\backslash\mathbb{R}$, $\underline{\Delta}_N(z) > 0$. Solving the system, and looking at the corresponding expression of $\mathrm{Im}(\delta_N(z))$, we easily get that

$$\underline{\Delta}_N(z) = \frac{\mathrm{Im}(z)}{\mathrm{Im}(\delta_N(z))}\frac{\sigma}{N}\mathrm{Tr}\,\mathbf{T}_N(z)\mathbf{T}_N(z)^*$$

for $z \in \mathbb{C}\backslash\mathbb{R}$. Expressing $\mathbf{T}_N(z) - \mathbf{T}_N(z)^*$ as $\mathbf{T}_N(z)(\mathbf{T}_N(z)^{-*} - \mathbf{T}_N(z)^{-1})\mathbf{T}_N(z)^*$, and using the equation $\delta_N(z) = \frac{\sigma}{N}\mathrm{Tr}\,(\mathbf{T}_N(z))$, we obtain that

$$\mathrm{Im}(\delta_N(z)) = \mathrm{Im}(w_N(z))\frac{\sigma}{N}\mathrm{Tr}\,\mathbf{T}_N(z)\mathbf{T}_N(z)^*$$

and deduce the useful formula

$$\underline{\Delta}_N(z) = \frac{\mathrm{Im}(z)}{\mathrm{Im}(w_N(z))}. \tag{A.37}$$

Using the integral representation $\delta_N(z) = \int_{\mathcal{S}_N} \frac{\mathrm{d}\mu_N(\lambda)}{\lambda - z}$, we obtain after straightforward computations that the expression $\frac{\mathrm{Im}(z)}{\mathrm{Im}(w_N(z))}$ extends to $\mathbb{C}\backslash\mathcal{S}_N$, and therefore to $\mathbb{C}\backslash\mathrm{supp}(\phi)$, and satisfies

$$\sup_{z \in \partial \mathcal{R}_y} \frac{|\mathrm{Im}(w_N(z))|}{|\mathrm{Im}(z)|} < C.$$

Equation (A.37) thus implies that

$$\sup_{z \in \partial \mathcal{R}_y} |\underline{\Delta}_N(z)|^{-1} \leq C,$$

which concludes the proof. $\qquad\square$

We deduce from this that $\inf_{z \in \partial \mathcal{R}_y} |\Delta_N(z)| \geq C > 0$ for all large N. Therefore, we can invert the system (A.34) and obtain

$$\sup_{z \in \partial \mathcal{R}_y} |\alpha_{r,N}(z) - \delta_N(z)| \leq \frac{C}{N^2},$$

for all large N. This establishes (3.35) and completes the proof of (3.34).

The proof of (3.36) is similar to the proof of Lemma 3.2, but as above, $\alpha_N(z), \tilde{\alpha}_N(z), \mathbf{R}_N(z)$ and $\tilde{\mathbf{R}}_N(z)$ have to be replaced by their regularized versions $\alpha_{r,N}(z), \tilde{\alpha}_{r,N}(z), \mathbf{R}_{r,N}(z)$ and $\tilde{\mathbf{R}}_{r,N}(z)$. The reader can check that the properties of these regularized functions allow to follow the various steps of the proof of Lemma 3.2.

Acknowledgment

References

[1] Z. D. Bai and J. W. Silverstein, No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices, *Ann. Probab.* **26**(1) (1998) 316–345.

[2] Z. D. Bai and J. W. Silverstein, Exact separation of eigenvalues of large dimensional sample covariance matrices, *Ann. Probab.* **27**(3) (1999) 1536–1555.

[3] F. Benaych-Georges and R. R. Nadakuditi, The singular values and vectors of low rank perturbations of large rectangular random matrices, submitted, arXiv:1103.2221.

[4] M. Capitaine and C. Donati-Martin, Strong asymptotic freeness of Wigner and Wishart matrices, *Indiana Univ. Math. J.* **56** (2007) 295–309.

[5] 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. Probab.* **37**(1) (2009) 1–47.

[6] L. H. Y. Chen, An inequality for the multivariate normal distribution, *J. Multivariate Anal.* **12**(2) (1982) 306–315.

[7] K. R. Davidson and S. J. Szarek, Local operator theory, random matrices and Banach spaces, in *Handbook of the Geometry of Banach Spaces* (2001), pp. 317–366.

[8] R. B. Dozier and J. W. Silverstein, Analysis of the limiting spectral distribution of large dimensional information plus noise type matrices, *J. Multivariate Anal.* **98**(6) (2007) 1099–1122.

[9] R. B. Dozier and J. W. Silverstein, On the empirical distribution of eigenvalues of large dimensional information plus noise type matrices, *J. Multivariate Anal.* **98**(4) (2007) 678–694.

[10] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton and J. Najim, On the capacity achieving covariance matrix for Rician MIMO channels: An asymptotic approach, *IEEE Trans. Information Theory* **56**(3) (2010) 1048–1069.

[11] V. L. Girko, *Theory of Stochastic Canonical Equations, Vol. I*, Mathematics and its applications, Vol. 535 (Kluwer Academic Publishers, Dordrecht, 2001).

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

[13] W. Hachem, P. Loubaton and J. Najim, Deterministic equivalents for certain functionals of large random matrices, *Ann. Appl. Probab.* **17**(3) (2007) 875–930.

[14] W. Hachem, P. Loubaton, J. Najim and P. Vallet, On bilinear forms based on the resolvent of large random matrices, to appear in *Ann. Inst. Henri Poincaré*.

[15] E. J. Hannan, The estimation of frequency, *J. Appl. Probab.* **10**(3) (1973) 510–519.

[16] A. T. James, Distributions of matrix variates and latent roots derived from normal samples, *Ann. Math. Stat.* **35**(2) (1964) 475–501.

[17] P. Loubaton and P. Vallet, Almost sure localization of the eigenvalues in a gaussian information plus noise model. Application to the spiked models, to appear in *Electronic J. Probab.*

[18] X. Mestre and M. A. Lagunas, Modified subspace algorithms for DoA estimation with large arrays, *IEEE Trans. Signal Process.* **56**(2) (2008) 598.

[19] L. Pastur, A simple approach to the global regime of Gaussian ensembles of random matrices, *Ukrainian Math. J.* **57**(6) (2005) 936–966.

[20] J. W. Silverstein and S. Choi, Analysis of the limiting spectral distribution of large dimensional random matrices, *J. Multivariate Anal.* **52**(2) (1995) 175–192.

[21] P. Stoica and N. Arye, Music, maximum likelihood, and Cramer–Rao bound, *IEEE Trans. Acoustics, Speech Signal Process.* **37**(5) (1989) 720–741.

[22] P. Vallet, W. Hachem, P. Loubaton, X. Mestre and J. Najim, An improved MUSIC algorithm based on low rank perturbation of large random matrices, in *Proc. of the 2011 IEEE Workshop on Statistical Signal Processing* (IEEE Signal Processing Society, 2011).

[23] P. Vallet, P. Loubaton and X. Mestre, Improved subspace estimation for multivariate observations of high dimension: The deterministic signal case, to appear in *IEEE Trans. Information Theory*, arXiv:1002.3234.