Applications of large random matrix to high dimensional statistical signalprocessing

Gia-Thuy Pham

École Doctorale Mathématiques et Sciences et Technologie de l'Information et de la Communication

THÈSE

pour obtenir le grade de docteur délivré par

Université Paris-Est

Spécialité: Traitement du signal

présentée et soutenue publiquement par

Gia-Thuy PHAM

le 28 Février 2017

Applications des grandes matrices aléatoires au traitement du signal de grandes dimensions

Applications of large random matrix to high dimensional statistical signal processing

Directeur de thèse : **Philippe LOUBATON**

Composition de la commission d'examen

Xavier MESTRE,	Professeur	Rapporteur
Jianfeng YAO,	Professeur	Rapporteur
Frédéric PASCAL,	Professeur	Examinateur
Walid HACHEM,	Professeur	Examinateur
Inbar FIJALKOW,	Professeur	Examinateur

Contents

Remerciements

Mes premiers remerciements iront sans doute à Philippe Loubaton. Il m'a inspiré de sa patience, sa compréhension et sa persévérance. Son expertise et ses connaissances sont splendides, cela aide à résoudre des problèmes plus rapidement. Perturbé par les tâches administratives du LabEx, du CNRS, et les enseignements, Philippe a toujours su se montrer disponible, de bonne humeur, d'efficacité. En particulier, pour moi, Philippe est comme le deuxième père, qui m'écoute et qui me donne des conseils. Phillipe, tu es un exemple que je vais continuer à suivre pendant toute ma carrière !

I also want to thank Xavier Mestre and Jian-Feng Yao for accepting to be rapporteurs of this thesis, and I thank them for their positive comments. I have met Xavier in many occasions, especially during SSP 2016 in Palma de Mallorca with a very nice dinner. In addition, inspired by Xavier's idea in his paper, back in 2006, that we could have a publication. My gratitude also goes to Walid Hachem,Frédéric Pascal and Inbar Fijalkow for accepting to participate to the jury of my thesis.

Ce fut un plaisir de travailler avec Pascal Vallet. C'était le début de ma thèse, Pascal n'a pas hésité de me donner un cours, transmettant ses connaissances de manière très enthousiaste. Ses idées sont enfin matérialisées par notre publication sur TSP. J'espère sincèrement pouvoir continuer à travailler avec lui dans l'avenir.

Durant ces trois années, j'ai travaillé au sein de l'équipe signal du laboratoire IGM de Marne-la-Vallée. Je remercie tous les collègues pour la bonne ambiance qu'ils ont su apporter chaque jour: Sonja et Mohamed Ali, mes colocataires de bureau 4B048, qui ont dû me supporter durant ces années, les thésards Feriel, Yosra, Aurélie, Audrey, Mai-Quyen, Mireille, Safa, Zakaria..., à Jamal et Abdellatif pour les discussions enrichissantes, sans oublier Samuele pour les parties échecs, Carlos, Siqi, Daniel, Alain,... pour les matchs de foot, et aussi Emilie et Jean-Christophe pour leurs conseils. Corinne Palescandolo et Patrice Herault méritent amplement des remerciements, pour leur aide et leurs grandes compétences dans les domaines administratif et informatique, respectivement.

Mes derniers remerciements seront consacrés à mes proches. Je remercie mes parents pour leur soutien et toute leur aide, surtout pour leur confiance continue en leur fils qui se trouve à l'autre bout du monde, à 12000 km d'eux. Je remercie ma mère pour son esprit fort, qui m'a appris à ne jamais abandonner dans la vie. Je remercie mon père pour son éducation, ses cours de maths, d'échecs, de foot,etc, malgré qu'il soit hospitalisé depuis quelques temps, mais il garde toujours un bon esprit et de l'optimisme. Je remercie également ma petite soeur d'avoir partagé notre enfance, et qui me motive à devenir un bon exemplaire du grand frère. Je remercie également tous mes amis, pour leur encouragement et pour avoir pris régulièrement des nouvelles. Finalement, mes remerciements vont à Thai-Hoa, celle qui a partagé ses 7 ans de jeunesse avec moi, qui restait à côté de moi et qui m'a soutenu durant tous ces années. Surtout, je la remercie d'avoir découvert mon anomalie cardiaque, qui a sauvé ma vie. Malgré que l'on ait pris des chemins différents, j'espère que l'on se croisera un jour.

Résumé

Cette thèse porte sur des problèmes de statistiques mettant en jeu une série temporelle multivariable $\mathbf{y}_n$ de grande dimension M définie comme la somme d'un bruit gaussien blanc temporellement et spatialement et d'un signal utile généré comme la sortie d'un filtre 1 entrée / M sorties à réponse impulsionnelle finie excité par une séquence déterministe scalaire non observable. Si l'on suppose que $\mathbf{y}$ est observé entre les instants 1 et N, un bon nombre de techniques existantes sont basées sur des fonctionnelles de la matrice de covariance empirique $\hat{\mathbf{R}}_L$ des vecteurs de dimensions ML $(\mathbf{y}_n^{(L)})_{n=1,\dots,N}$ obtenus en empilant les vecteurs $\mathbf{y}_k$ entre les instants n et $n+L-1$, où L est un paramètre bien choisi. Lorsque l'on est en mesure de collecter un nombre d'observations très nettement plus grand que la dimension ML des vecteurs $(\mathbf{y}_n^{(L)})_{n=1,\dots,N}$, $\hat{\mathbf{R}}_L$ a le même comportement en norme spectrale que son espérance mathématique, et cela permet d'étudier les techniques d'inférences basés sur $\hat{\mathbf{R}}_L$ par le biais de techniques classiques de statistique asymptotique. Dans cette thèse, nous nous intéressons au cas où ML et N sont du même ordre de grandeur, ce que nous modélisons par des régimes asymptotiques dans lesquels M et N tendent tous les deux vers l'infini, et où le rapport ML/N converge vers une constante non nulle, L pouvant aussi croître avec M et N. Les problèmes que nous résolvons dans ce travail nécessitent d'étudier le comportement des éléments propres de la grande matrice aléatoire $\hat{\mathbf{R}}_L$. Compte tenu de la structure particulière des vecteurs $(\mathbf{y}_n^{(L)})_{n=1,\dots,N}$, $\hat{\mathbf{R}}_L$ coïncide avec la matrice de Gram d'une matrice Hankel par bloc $\mathbf{\Sigma}_L$, et cette spécificité nécessite le développement de techniques appropriées.

Dans le chapitre 2, nous nous intéressons au cas où le nombre de coefficients P de la réponse impulsionnelle générant le signal utile et le paramètre L restent fixes quand M et N grandissent. La matrice $\mathbf{\Sigma}_L$ est alors une perturbation de rang fini de la matrice Hankel par bloc $\mathbf{W}_L$ constituée à partir du bruit additif. Nous montrons que les éléments propres de $\hat{\mathbf{R}}_L$ se comportent comme si la matrice $\mathbf{W}_L$ était à éléments indépendants et identiquement distribués. Cela nous permet d'aborder l'étude de tests de détection du signal utile portant sur les plus grandes valeurs propres de $\hat{\mathbf{R}}_L$ ainsi que la mise en évidence de nouvelles stratégies de détermination du paramètre de régularisation de filtres de Wiener spatio-temporels estimés à partir d'une séquence d'apprentissage. Techniquement, ce dernier point est abordé en caractérisant le comportement asymptotique des éléments de la résolvente de la matrice $\hat{\mathbf{R}}_L$. Nous montrons enfin également que ces résultats permettent d'analyser le comportement d'algorithmes sous-espace de localisation de sources bande étroite utilisant la technique du lissage spatial.

Dans le chapitre 3, motivés par le cas où P et L peuvent tendre vers l'infini, nous nous écartons quelque peu du modèle initial, et supposons que la matrice $\mathbf{\Sigma}_L$ est la somme de la matrice aléatoire Hankel par bloc $\mathbf{W}_L$ avec une matrice déterministe sans structure particulière. En utilisant des approaches basées sur la transformée de Stieltjes et des outils adaptés au caractère gaussien du bruit, nous montrons que la distribution empirique des valeurs propres de $\hat{\mathbf{R}}_L$ a un comportement déterministe que nous caractérisons. Sous réserve L^2/MN tende vers 0, nous faisons de même pour les éléments de la résolvante de $\hat{\mathbf{R}}_L$.

Dans le chapitre 4, nous revenons au modèle initial, mais supposons que P et L tendent vers l'infini au même rythme. Dans ce contexte, la contribution du signal utile à la matrice $\mathbf{\Sigma}_L$ est une matrice dont le rang tend vers l'infini, et les techniques utilisées dans le chapitre 2 ne sont plus applicables. En utilisant

les résultats du chapitre 3, nous établissons que si L^2/MN tende vers 0, les éléments de la résolvante de $\hat{\mathbf{R}}_L$ se comportent comme les éléments d'une matrice déterministe qui coïncide avec l'équivalent déterministe de la resolvente d'un modèle information plus bruit dans lequel les éléments de la matrice de bruit sont indépendants et identiquement distribués. Dans le cas où L/M tend vers 0, ceci nous permet d'étendre les résultats du chapitre 2 relatifs à la détermination du paramètre de régularisation des filtres de Wiener spatio-temporels estimés à partir d'une séquence d'apprentissage.

Abstract

This thesis focuses on statistical problems involving a multivariate time series $\mathbf{y}_n$ of large dimension M defined as the sum of gaussian white noise temporally and spatially and a useful signal defined as the output of an unknown finite impulse response single input multiple outputs system driven by a deterministic scalar nonobservable sequence. Supposing $(\mathbf{y}_n)_{n=1,\dots,N}$ is available, a number of existing methods are based on the functionals of empirical covariance matrix $\hat{\mathbf{R}}_L$ of ML–dimensional vectors $(\mathbf{y}_n^{(L)})_{n=1,\dots,N}$ obtained by stacking the vectors $(\mathbf{y}_k)_{k=n,\dots,n+L-1}$, where L is a relevant parameter. In the case where the number of observations N is much larger than ML the dimension of vectors $(\mathbf{y}_n^{(L)})_{n=1,\dots,N}$, $\hat{\mathbf{R}}_L$ behaves as its mathematical expectation in the sense of spectral norm. This allows us to study the inference technique based on $\hat{\mathbf{R}}_L$ via classical techniques of asymptotic statistics. In this thesis, we interested in the case where ML and N have the same order of magnitude, we call this the asymptotic regimes in which M and N converge towards infinity, such that the ratio $\frac{ML}{N}$ converges towards a strictly positive constant, given that L may scale with M, N. To solve the problems in this work, it is necessary to investigate the behaviour of the eigenvalues and eigenvectors of the random matrix $\hat{\mathbf{R}}_L$. Taking account of the particular structure of vectors $(\mathbf{y}_n^{(L)})_{n=1,\dots,N}$, $\hat{\mathbf{R}}_L$ coincides with the Gram matrix of a block-Hankel matrix $\mathbf{\Sigma}_L$, and this specificity requires the development of appropriate techniques.

In chapter 2, we interested to the case where the number of coefficients P of the finite impulse response generated the useful signal and the parameter L remain fixed when M, N grow large. As a consequence, the matrix $\mathbf{\Sigma}_L$ is a finite rank perturbation of block-Hankel matrix $\mathbf{W}_L$ composed of additive noise. We prove that eigenvalues and eigenvectors of $\hat{R}_L$ behave as if the entries of matrix $\mathbf{W}_L$ are independent and identically distributed. This allows us to construct detection tests of useful signal based on largest eigenvalues of $\hat{\mathbf{R}}_L$ and to develop new estimation strategies of the regularization parameter of the spatio-temporal Wiener filter estimated from a training sequence. This approach is characterized by the asymptotic behaviour of the resolvent of matrix $\hat{\mathbf{R}}_L$. We also prove that these results provide consistent subspace estimation methods for source localization using spatial-smoothing scheme.

In chapter 3, motivated by the case where P and L may converge towards infinity, we move off somewhat the initial model. We suppose that the matrix $\mathbf{\Sigma}_L$ is the sum of the block-Hankel random matrix $\mathbf{W}_L$ with a deterministic matrix without particular structure. Using the approaches based on Stieltjes transform and tools adapted to gaussian noise, we prove that the empirical eigenvalue distribution of $\hat{\mathbf{R}}_L$ has deterministic behaviour which we shall describe. Provided $\frac{L^2}{MN}$ converges towards 0, we do likewise for the elements of the resolvent of $\hat{\mathbf{R}}_L$.

In chapter 4, we return to the initial model, but we suppose that P, L converge towards infinity with the same rate. In this context, matrix $\mathbf{\Sigma}_L$ is a matrix whose rank goes to infinity, and thus the techniques employed in chapter 2 are not applicable. Using the results obtained in chapter 3, we establish that when $\frac{L^2}{MN}$ goes to 0, the elements of the resolvent of $\hat{\mathbf{R}}_L$ behave as the elements of a deterministic matrix which coincides with the deterministic equivalent of the resolvent of the information plus noise model in which entries of the noise matrix are independent and identically distributed. In the case where $\frac{L}{M}$ goes to 0, this allows us to extend the results of chapter 2 related to the determination of the regularization parameter of the spatial-temporal Wiener filter estimated from a training sequence.

Chapter 1

Introduction

1.1 Introduction

Due to the spectacular evolution of data acquisition devices and sensor networks, it becomes common to be faced to multivariate signals of high dimension. Very often, the sample size that can be used in practice to perform statistical inference cannot be much larger than the dimension of the observation because the duration of the signals are limited, or because their statistics are not time-invariant over large enough temporal windows. In this context, it is well established that fundamental statistical signal processing techniques implemented in existing systems (*e.g.*, source detection, source localisation, estimation of various kinds of spatial and spatio-temporal filters, or of equalizers, blind source separation, blind deconvolution and equalization,...) show poor performance. It is therefore of crucial importance to revisit the corresponding problems, and to be able to propose new algorithms with enhanced performance. In the last decade, a number of mathematical tools were thus developed in the context of high-dimensional statistical signal processing. Among others, we mention the use of possible sparsity of certain parameters of interest, and large random matrices. This thesis concentrates on the development of large random matrix tools, and to their applications to important high-dimensional statistical signal processing problems.

Large random matrices have been proved to be of fundamental importance in mathematics (high dimensional probability and statistics, operator algebras, combinatorics, number theory,...) and in physics (nuclear physics, quantum fields theory, quantum chaos,..) for a long time. The introduction of large random matrix theory in electrical engineering is more recent. It was introduced at the end of the nineties in the context of digital communications in order to analyse the performance of large CDMA and MIMO systems. Except the pioneering work of Girko ([26], [27]), the first works using large random matrix theory in the context of multivariate statistical signal processing were published in the second part of the 2000s, see e.g. [53], [54], [55], [40], [57], [43], [58]. These works were followed by a number of subsequent contributions, e.g. [12], [16], [17], [18], [19], [72], [33], [74], [79], [80]. The common point of the above mentioned works is to address statistical inference problems for the narrow band array processing model, also called the linear static factor model in the statistical terminology. In this context, the observation is a M-dimensional signal $\mathbf{y}_n$ defined as a noisy version of a low rank K useful signal on which various informations have to be retrieved from samples $(\mathbf{y}_n)_{n=1,\dots,\mathrm{N}}$. In particular, one may want to detect the presence or the absence of the useful signal, to estimate its rank K, the corresponding signal subspace, or to estimate certain parameters (e.g. direction of arrival). A number of statistical inference methods are based on the observation that the relevant informations are contained in the "true" spatial covariance matrix $\mathbf{R} = \mathbb{E}(\mathbf{y}_n \mathbf{y}_n^*)$, and that, in the case where $\mathrm{M} << \mathrm{N}$, the empirical spatial covariance matrix $\hat{\mathbf{R}}$ behaves as $\mathbf{R}$. Therefore, many detection / estimation schemes use functionals of $\hat{\mathbf{R}}$, which, when

$M << N$, behave as the corresponding functionals of $\mathbf{R}$. When M and N are large and of the same order of magnitude, it is well established that $\hat{\mathbf{R}}$ is a poor estimate of $\mathbf{R}$ in the sense that functionals of $\hat{\mathbf{R}}$ do not behave as the functionals of $\mathbf{R}$. The above mentioned papers developed methodologies that allow to evaluate the behaviour of functionals of $\hat{\mathbf{R}}$ in the asymptotic regime M and N converge towards infinity at the same rate, and to take benefit of the corresponding results to propose improved performance statistical inference schemes.

In the present thesis, we address statistical inference problems for the wide band array processing model in which the useful signal is low rank in the frequency domain, or, in other words, coincides with the output of an unknown K inputs / M outputs filter (with $K < M$) driven by a K–dimensional time series. In this case, the empirical spatial covariance matrix does not convey enough information on the low rank component, and several inference schemes are rather based on empirical spatio-temporal covariance matrices $\hat{\mathbf{R}}^{(L)}$ defined as the empirical covariance matrices of augmented ML–dimensional vectors $\mathbf{y}_n^{(L)} = (\mathbf{y}_n^T, \ldots, \mathbf{y}_{n+L-1}^T)^T$ where L is a relevant parameter. Our goal is thus to study the asymptotic behaviour of functionals of random matrices $\hat{\mathbf{R}}^{(L)}$ when M and N converge towards infinity, to use the corresponding results to analyze the performance of traditional detection / estimation schemes, and to propose improved methods. The main originality of this research program follows from the particular structure of the $ML \times N$ augmented observation matrix $\mathbf{Y}_L = (\mathbf{y}_1^{(L)}, \ldots, \mathbf{y}_N^{(L)})$. This matrix has a block-Hankel structure, so that the analysis of the asymptotic behaviour of $\hat{\mathbf{R}}^{(L)} = \mathbf{Y}_L \mathbf{Y}_L^* / N$ needs the establishment of new results.

1.2 Review of useful known results.

In this section, we review some fundamental results concerning large empirical covariance matrices.

1.2.1 The Marcenko-Pastur distribution.

The Marcenko-Pastur distribution was introduced 40 years ago in [52], and plays a key role in a number of high-dimensional statistical signal processing problems. In this section, $(\mathbf{v}_n)_{n=1,\ldots,N}$ denotes a sequence of i.i.d. zero mean complex Gaussian random M–dimensional vectors for which $\mathbb{E}(\mathbf{v}_n \mathbf{v}_n^*) = \sigma^2 \mathbf{I}_M$. We consider the empirical covariance matrix

$$\frac{1}{N} \sum_{n=1}^{N} \mathbf{v}_n \mathbf{v}_n^*$$

which can also be written as

$$\frac{1}{N} \sum_{n=1}^{N} \mathbf{v}_n \mathbf{v}_n^* = \mathbf{W}_N \mathbf{W}_N^*$$

where matrix $\mathbf{W}_N$ is defined by $\mathbf{W}_N = \frac{1}{\sqrt{N}}(\mathbf{v}_1, \ldots, \mathbf{v}_N)$. $\mathbf{W}_N$ is thus a complex Gaussian matrix with independent identically distributed $\mathcal{N}_c(0, \frac{\sigma^2}{N})$ entries. When $N \to +\infty$ while M remains fixed, matrix $\mathbf{W}_N \mathbf{W}_N^*$ converges towards $\sigma^2 \mathbf{I}_M$ in the spectral norm sense. In the high-dimensional asymptotic regime defined by

$$M \to +\infty, N \to +\infty, d_N = \frac{M}{N} \to d > 0 \tag{1.2.1}$$

it is well understood that $\|\mathbf{W}_N \mathbf{W}_N^* - \sigma^2 \mathbf{I}_M\|$ does not converge towards 0. In particular, the empirical distribution $\hat{\mu}_N = \frac{1}{M} \sum_{m=1}^{M} \delta_{\hat{\lambda}_{m,N}}$ of the eigenvalues $\hat{\lambda}_{1,N} \geq \ldots \geq \hat{\lambda}_{M,N}$ of $\mathbf{W}_N \mathbf{W}_N^*$ does not converge towards the Dirac measure at point $\lambda = \sigma^2$. More precisely, we denote by μ_{d,σ^2} the Marcenko-Pastur distribution of parameters (d, σ^2) defined as probability measure

$$d\mu_{d,\sigma^2}(\lambda) = \delta_0[1-d^{-1}]_+ + \frac{\sqrt{(\lambda-\lambda^-)(\lambda^+-\lambda)}}{2\sigma^2 d\pi\lambda} \mathbb{1}_{[\lambda^-,\lambda^+]}(\lambda)\, d\lambda$$

with $\lambda^- = \sigma^2(1-\sqrt{d})^2$ and $\lambda^+ = \sigma^2(1+\sqrt{d})^2$. Then, the following result holds.

Theorem 1.2.1. *The empirical eigenvalue value distribution* $\hat{\mu}_N$ *converges weakly almost surely towards* μ_{d,σ^2} *when both* M *and* N *converge towards* $+\infty$ *in such a way that* $d_N = \frac{M}{N}$ *converges towards* $d > 0$. *Moreover, it holds that*

$$\hat{\lambda}_{1,N} \to \sigma^2(1+\sqrt{d})^2 \ a.s. \tag{1.2.2}$$

$$\hat{\lambda}_{\min(M,N)} \to \sigma^2(1-\sqrt{d})^2 \ a.s. \tag{1.2.3}$$

We also observe that Theorem 1.2.1 remains valid if $\mathbf{W}_N$ is a non necessarily Gaussian matrix whose i.i.d. elements have a finite fourth order moment (see e.g. [5]). Theorem 1.2.1 means that when ratio $\frac{M}{N}$ is not small enough, the eigenvalues of the empirical spatial covariance matrix of a temporally and spatially white noise tend to spread out around the variance of the noise, and that almost surely, for N large enough, all the eigenvalues are located in a neighbourhood of interval $[\lambda^-, \lambda^+]$. While the spreading of the eigenvalues is not an astonishing phenomenon, it is remarkable that the behaviour of the eigenvalues can be characterized very precisely.

In order to establish the convergence of $\hat{\mu}_N$ towards μ_{d,σ^2}, the simplest approach consists in studying the asymptotic behaviour of the Stieltjes transform $\hat{m}_N(z)$ of $\hat{\mu}_N$ defined as

$$\hat{m}_N(z) = \int_{\mathbf{R}^+} \frac{1}{\lambda - z}\, d\hat{\mu}_N(z) = \frac{1}{M}\sum_{m=1}^{M} \frac{1}{\hat{\lambda}_{m,N} - z} \tag{1.2.4}$$

and to establish that for each $z \in \mathbb{C} - \mathbb{R}^+$, it holds that

$$\lim_{N\to+\infty} \hat{m}_N(z) = m_{d,\sigma^2}(z)\ a.s., \tag{1.2.5}$$

where $m_{d,\sigma^2}(z) = \int_{\mathbf{R}^+} \frac{1}{\lambda - z}\, d\mu_{d,\sigma^2}(\lambda)$ represents the Stieltjes transform of μ_{d,σ^2}. Function m_{d,σ^2} is known to be the unique Stieltjes transform of a probability measure carried by $\mathbb{R}^+$ satisfying the equation

$$m_{d,\sigma^2}(z) = \frac{1}{-z + \frac{\sigma^2}{1+\sigma^2 d\, m_{d,\sigma^2}(z)}} \tag{1.2.6}$$

for each z. $m_{d,\sigma^2}(z)$ can also be defined as the first component of the solution of the coupled equation

$$\begin{aligned} m_{d,\sigma^2}(z) &= \frac{1}{-z\left(1+\sigma^2 \tilde{m}_{d,\sigma^2}(z)\right)} \\ \tilde{m}_{d,\sigma^2}(z) &= \frac{1}{-z\left(1+\sigma^2 d m_{d,\sigma^2}(z)\right)} \end{aligned} \tag{1.2.7}$$

The second component $\tilde{m}_{d,\sigma^2}(z)$ of the solution of (1.2.7) then coincides with $-(1-d)/z + d\, m_{d,\sigma^2}(z)$ i.e. with the Stieltjes transform of $(1-d)\delta_0 + d\,\mu_{d,\sigma^2}$. It is easily seen that

$$(1-d)\delta_0 + d\,\mu_{d,\sigma^2} = \mu_{d^{-1},\sigma^2 d} \tag{1.2.8}$$

Therefore, it also holds that

$$\tilde{m}_{d,\sigma^2}(z) = m_{d^{-1},\sigma^2 d}(z) \tag{1.2.9}$$

for each $z \in \mathbb{C} - \mathbb{R}^+$. In order to establish (1.2.5), it is sufficient to use a Gaussian concentration argument, and to prove that $\mathbb{E}(\hat{m}_N(z))$ satisfies a perturbed version of (1.2.6). In the Gaussian case, this last point can be addressed using the integration by parts formula and the Poincaré-Nash inequality, see below for more details. We also mention that function $w_{d,\sigma^2}(z)$ defined by

$$w_{d,\sigma^2}(z) = \frac{1}{z m_{d,\sigma^2}(z) \tilde{m}_{d,\sigma^2}(z)} \tag{1.2.10}$$

is analytic on $\mathbb{C} - [\lambda^-, \lambda^+]$, verifies $\frac{\mathrm{Im}(w(z))}{\mathrm{Im}(z)} > 0$ if $z \in \mathbb{C} - \mathbb{R}$, $w_{d,\sigma^2}(\lambda^-) = -\sigma^2\sqrt{d}$ and $w_{d,\sigma^2}(\lambda^+) = \sigma^2\sqrt{d}$. Moreover, $w_{d,\sigma^2}(\lambda)$ increases from $-\infty$ to $-\sigma^2\sqrt{d}$ when λ increases from $-\infty$ to λ^-, $w_{d,\sigma^2}(\lambda)$ increases from $\sigma^2\sqrt{d}$ to $+\infty$ when λ increases from λ^+ to $+\infty$, and the set $\{w_{d,\sigma^2}(\lambda), \lambda \in [\lambda^-, \lambda^+]\}$ coincides with the half circle $\{\sigma^2\sqrt{d}\, e^{i\theta}\}$ where θ increases from $-\pi$ to 0. Using (1.2.7), it is possible to express $m_{d,\sigma^2}(z)$ in terms of $w_{d,\sigma^2}(z)$. More precisely, $\frac{1}{z\tilde{m}_{d,\sigma^2}(z)} = -(1 + \sigma^2 d m_{d,\sigma^2}(z))$ so that

$$w_{d,\sigma^2}(z) = -\frac{1 + \sigma^2 d m_{d,\sigma^2}(z)}{m_{d,\sigma^2}(z)}$$

Solving w.r.t. $m_{d,\sigma^2}(z)$ leads to

$$m_{d,\sigma^2}(z) = -\frac{1}{w_{d,\sigma^2}(z) + \sigma^2 d} \tag{1.2.11}$$

It can be shown similarly that

$$\tilde{m}_{d,\sigma^2}(z) = -\frac{1}{w_{d,\sigma^2}(z) + \sigma^2} \tag{1.2.12}$$

Plugging (1.2.11) into (1.2.6), we obtain that $w_{d,\sigma^2}(z)$ is solution of the equation

$$\phi_{d,\sigma^2}\left(w_{d,\sigma^2}(z)\right) = z \tag{1.2.13}$$

for each $z \in \mathbb{C}$ where $\phi_{d,\sigma^2}(w)$ is the function defined by

$$\phi_{d,\sigma^2}(w) = \frac{(w + \sigma^2)(w + \sigma^2 d)}{w} \tag{1.2.14}$$

It is useful to notice that if $\mathbf{Q}_{W,N}(z)$ denotes the resolvent of matrix $\mathbf{W}_N\mathbf{W}_N^*$ defined by

$$\mathbf{Q}_{W,N}(z) = \left(\mathbf{W}_N\mathbf{W}_N^* - z\mathbf{I}_M\right)^{-1} \tag{1.2.15}$$

then $\hat{m}_N(z)$ coincides with $\frac{1}{M}\mathrm{Tr}\left(\mathbf{Q}_{W,N}(z)\right)$. We also note that it is possible to establish a stronger result, i.e. for each $z \in \mathbb{C} - \mathbb{R}^+$,

$$\mathbf{a}_N^* \left(\mathbf{Q}_{W,N}(z) - m_{d,\sigma^2}(z)\mathbf{I}_M\right)\mathbf{b}_N \to 0 \; a.s. \tag{1.2.16}$$

for each deterministic vectors $\mathbf{a}_N, \mathbf{b}_N$ for which $\sup_N(\|\mathbf{a}_N\|, \mathbf{b}_N\|) < +\infty$. A similar result holds for $\tilde{\mathbf{Q}}_{W,N}(z)$ defined as the resolvent of $\mathbf{W}_N^*\mathbf{W}_N$, i.e.

$$\tilde{\mathbf{Q}}_{W,N}(z) = \left(\mathbf{W}_N^*\mathbf{W}_N - z\mathbf{I}_N\right)^{-1} \tag{1.2.17}$$

More precisely, it holds that

$$\tilde{\mathbf{a}}_N^* \left(\tilde{\mathbf{Q}}_{W,N}(z) - \tilde{m}_{d,\sigma^2}(z)\mathbf{I}_N\right)\tilde{\mathbf{b}}_N \to 0 \; a.s. \tag{1.2.18}$$

for each deterministic vectors $\tilde{\mathbf{a}}_N, \tilde{\mathbf{b}}_N$ for which $\sup_N(\|\tilde{\mathbf{a}}_N\|, \tilde{\mathbf{b}}_N\|) < +\infty$. Moreover, for each $z \in \mathbb{C} - \mathbb{R}^+$, it holds that

$$\mathbf{a}_N^* \left(\mathbf{Q}_{W,N}(z)\mathbf{W}_N\right)\tilde{\mathbf{b}}_N \to 0 \; a.s. \tag{1.2.19}$$

Finally, convergence properties (1.2.16, 1.2.18, 1.2.19) hold uniformly w.r.t. z on each compact subset of $\mathbb{C}^* - [\lambda^-, \lambda^+]$.

1.2.2 The Information plus noise spiked models.

An information plus noise model is a $M \times N$ random matrix model $\mathbf{\Sigma}_N$ defined as

$$\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N \tag{1.2.20}$$

where complex Gaussian random matrix $\mathbf{W}_N$ is defined as above, and where $\mathbf{B}_N$ is a deterministic $M \times N$ matrix. In the asymptotic regime (1.2.1), this kind of models were initially studied by Girko (see e.g. [26]) , later by Dozier-Silverstein (see [20], [21]), and by [32] in the case where the entries of $\mathbf{W}_N$ are independent but not necessarily identically distributed. We denote by $(\hat{\lambda}_{m,N})_{m=1,...,M}$ the eigenvalues of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ arranged in the decreasing order. When the rank of matrix $\mathbf{B}_N$ is a fixed value K that does not scale with M and N, it is clear that the eigenvalue distribution $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ still converges towards the Marcenko-Pastur distribution μ_{d,σ^2}. More importantly, [9] and [10] proposed an elementary analysis that allows to characterize quite explicitly the largest eigenvalues and corresponding eigenvectors of matrix $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$.

More precisely, we denote by $\lambda_{1,N} > \lambda_{2,N} \ldots > \lambda_{K,N}$ the non zero eigenvalues of matrix $\mathbf{B}_N\mathbf{B}_N^*$ arranged in decreasing order, and by $(\mathbf{u}_{k,N})_{k=1,...,K}$ and $(\tilde{\mathbf{u}}_{k,N})_{k=1,...,K}$ the associated left and right singular vectors of $\mathbf{B}_N$. The singular value decomposition of $\mathbf{B}_N$ is thus given by

$$\mathbf{B}_N = \sum_{k=1}^{K} \lambda_{k,N}^{1/2} \mathbf{u}_{k,N}\tilde{\mathbf{u}}_{k,N}^* = \mathbf{U}_N \mathbf{\Lambda}_N^{1/2} \tilde{\mathbf{U}}_N^*$$

Moreover, we assume that:

Assumption 1.2.1. *The* K *non zero eigenvalues* $(\lambda_{k,N})_{k=1,...,K}$ *of matrix* $\mathbf{B}_N\mathbf{B}_N^*$ *converge towards* $\lambda_1 > \lambda_2 > \ldots > \lambda_K$ *when* $N \to +\infty$.

Here, for ease of exposition, we assume that the eigenvalues $(\lambda_{k,N})_{k=1,...,K}$ have multiplicity 1 and that $\lambda_k \neq \lambda_l$ for $k \neq l$. However, the forthcoming result can be easily adapted if some λ_k coincide.

Theorem 1.2.2. *We denote by* K_s, $0 \le K_s \le K$, *the largest integer for which*

$$\lambda_{K_s} > \sigma^2\sqrt{d} \tag{1.2.21}$$

Then, for $k = 1, \ldots, K_s$, *it holds that*

$$\hat{\lambda}_{k,N} \xrightarrow[N\to\infty]{a.s.} \rho_k = \phi_{d,\sigma^2}(\lambda_k) = \frac{(\lambda_k + \sigma^2)(\lambda_k + \sigma^2 d)}{\lambda_k} > \lambda^+. \tag{1.2.22}$$

Moreover, for $k = K_s + 1, \ldots, K$, *it holds that*

$$\hat{\lambda}_{k,N} \to \lambda^+ \ a.s. \tag{1.2.23}$$

Finally, for all deterministic sequences of M*–dimensional unit vectors* $(\mathbf{a}_N)$, $(\mathbf{b}_N)$, *we have for* $k = 1, \ldots, K_s$

$$\mathbf{a}_N^*\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*\mathbf{b}_N = \frac{\lambda_k^2 - \sigma^4 d}{\lambda_k(\lambda_k + \sigma^2 d)} \mathbf{a}_N^*\mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\mathbf{b}_N + o(1) \quad a.s. \tag{1.2.24}$$

This result implies that if the some of the non zero eigenvalues of $\mathbf{B}_N\mathbf{B}_N^*$ are large enough, then the corresponding greatest eigenvalues of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ escape from $[\lambda^-, \lambda^+]$ that can be interpreted as the noise eigenvalue interval. Moreover, each eigenvector $\hat{\mathbf{u}}_{k,N}$, $k = 1, \ldots, K_s$ associated to such an eigenvalue has a non trivial correlation with eigenvector $\mathbf{u}_{k,N}$ (take $\mathbf{a}_N = \mathbf{b}_N = \mathbf{u}_{k,N}$ in (1.2.24)) while $\mathbf{u}_{l,N}^*\hat{\mathbf{u}}_{k,N} \to 0$ for each

$l \neq k$. We note that the term $\frac{\lambda_k^2-\sigma^4 d}{\lambda_k(\lambda_k+\sigma^2 d)}$ is less than 1, and is close to 0 when λ_k is close to the threshold $\sigma^2\sqrt{d}$. We also remark that if $h_{d,\sigma^2}(z)$ represents the function defined by

$$h_{d,\sigma^2}(z) = \frac{\left[w_{d,\sigma^2}(z)\right]^2 - \sigma^4 d}{w_{d,\sigma^2}(z)(w_{d,\sigma^2}(z)+\sigma^2 d)} \tag{1.2.25}$$

then, (1.2.13) implies that

$$\frac{\lambda_k^2 - \sigma^4 d}{\lambda_k(\lambda_k+\sigma^2 d)} = h_{d,\sigma^2}(\rho_k) \tag{1.2.26}$$

As eigenvalue $\hat{\lambda}_{k,\mathrm{N}}$ converges towards ρ_k, (1.2.26) leads to

$$h_{d,\sigma^2}(\hat{\lambda}_{k,\mathrm{N}}) \to \frac{\lambda_k^2 - \sigma^4 d}{\lambda_k(\lambda_k+\sigma^2 d)}$$

Therefore, $\frac{\lambda_k^2-\sigma^4 d}{\lambda_k(\lambda_k+\sigma^2 d)}$ and each bilinear form of $\mathbf{u}_{k,\mathrm{N}}\mathbf{u}_{k,\mathrm{N}}^*$ can be estimated consistently from $\hat{\lambda}_{k,\mathrm{N}}$ and $\hat{\mathbf{u}}_{k,\mathrm{N}}$ as soon as σ^2 is known.

In order to understand which particular properties of matrix $\mathbf{W}_\mathrm{N}$ play a role in this result, we provide a sketch of proof of Theorem 1.2.2. We first justify (1.2.22). In the following, we prefer to follow the approach used in [17], which, while equivalent to [10], is more direct. For this, we denote by $\mathbf{Q}_\mathrm{N}(z)$ and $\tilde{\mathbf{Q}}_\mathrm{N}(z)$ the resolvents of matrices $\mathbf{\Sigma}_\mathrm{N}\mathbf{\Sigma}_\mathrm{N}^*$ and $\mathbf{\Sigma}_\mathrm{N}^*\mathbf{\Sigma}_\mathrm{N}$ respectively. The proof is based on the observation that $\det(\mathbf{\Sigma}_\mathrm{N}\mathbf{\Sigma}_\mathrm{N}^* - z\mathbf{I})$ can be expressed in terms of $\det(\mathbf{W}_\mathrm{N}\mathbf{W}_\mathrm{N}^* - z\mathbf{I})$, and that, in regime (1.2.1), the corresponding expression allows to check whether some of the K largest eigenvalues of $\mathbf{\Sigma}_\mathrm{N}\mathbf{\Sigma}_\mathrm{N}^*$ may escape from $[\lambda^- - \epsilon, \lambda^+ + \epsilon]$ where $\epsilon > 0$ may be arbitrarily small. We express $\mathbf{\Sigma}_\mathrm{N}\mathbf{\Sigma}_\mathrm{N}^* - z\mathbf{I}$ as

$$\mathbf{\Sigma}_\mathrm{N}\mathbf{\Sigma}_\mathrm{N}^* - z\mathbf{I} = \mathbf{W}_\mathrm{N}\mathbf{W}_\mathrm{N}^* - z\mathbf{I} + (\mathbf{U}_\mathrm{N}, \mathbf{W}_\mathrm{N}\tilde{\mathbf{U}}_\mathrm{N}\mathbf{\Lambda}_\mathrm{N}^{1/2}) \begin{pmatrix} \mathbf{\Lambda}_\mathrm{N} & \mathbf{I}_\mathrm{K} \\ \mathbf{I}_\mathrm{K} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{U}_\mathrm{N}^* \\ \mathbf{\Lambda}_\mathrm{N}^{1/2}\tilde{\mathbf{U}}_\mathrm{N}^*\mathbf{W}_\mathrm{N}^* \end{pmatrix} \tag{1.2.27}$$

Using (1.2.2) and (1.2.3), we obtain that if $z \neq 0$ is chosen real and outside $[\lambda^- - \epsilon, \lambda^+ + \epsilon]$, then matrix $\mathbf{W}_\mathrm{N}\mathbf{W}_\mathrm{N}^* - z\mathbf{I}$ is invertible. Therefore, for such z, $\mathbf{\Sigma}_\mathrm{N}\mathbf{\Sigma}_\mathrm{N}^* - z\mathbf{I}$ can be written as

$$\mathbf{\Sigma}_\mathrm{N}\mathbf{\Sigma}_\mathrm{N}^* - z\mathbf{I} = \left(\mathbf{W}_\mathrm{N}\mathbf{W}_\mathrm{N}^* - z\mathbf{I}\right)\left(\mathbf{I} + \mathbf{Q}_{\mathrm{W,N}}(z)(\mathbf{U}_\mathrm{N}, \mathbf{W}_\mathrm{N}\tilde{\mathbf{U}}_\mathrm{N}\mathbf{\Lambda}_\mathrm{N}^{1/2}) \begin{pmatrix} \mathbf{\Lambda}_\mathrm{N} & \mathbf{I}_\mathrm{K} \\ \mathbf{I}_\mathrm{K} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{U}_\mathrm{N}^* \\ \mathbf{\Lambda}_\mathrm{N}^{1/2}\tilde{\mathbf{U}}_\mathrm{N}^*\mathbf{W}_\mathrm{N}^* \end{pmatrix}\right) \tag{1.2.28}$$

Therefore, if $z \neq 0$ is real and outside $[\lambda^- - \epsilon, \lambda^+ + \epsilon]$, z is eigenvalue of $\mathbf{\Sigma}_\mathrm{N}\mathbf{\Sigma}_\mathrm{N}^*$ if and only if

$$\det\left(\mathbf{I} + \mathbf{Q}_{\mathrm{W,N}}(z)(\mathbf{U}_\mathrm{N}, \mathbf{W}_\mathrm{N}\tilde{\mathbf{U}}_\mathrm{N}\mathbf{\Lambda}_\mathrm{N}^{1/2}) \begin{pmatrix} \mathbf{\Lambda}_\mathrm{N} & \mathbf{I}_\mathrm{K} \\ \mathbf{I}_\mathrm{K} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{U}_\mathrm{N}^* \\ \mathbf{\Lambda}_\mathrm{N}^{1/2}\tilde{\mathbf{U}}_\mathrm{N}^*\mathbf{W}_\mathrm{N}^* \end{pmatrix}\right) = 0$$

or equivalently, if and only if $\det(\mathbf{F}_\mathrm{N}(z)) = 0$ where $\mathbf{F}_\mathrm{N}(z)$ is the $2\mathrm{K} \times 2\mathrm{K}$ matrix defined by

$$\mathbf{F}_\mathrm{N}(z) = \mathbf{I}_{2\mathrm{K}} + \begin{pmatrix} \mathbf{U}_\mathrm{N}^* \\ \mathbf{\Lambda}_\mathrm{N}^{1/2}\tilde{\mathbf{U}}_\mathrm{N}^*\mathbf{W}_\mathrm{N}^* \end{pmatrix} \mathbf{Q}_{\mathrm{W,N}}(z)(\mathbf{U}_\mathrm{N}, \mathbf{W}_\mathrm{N}\tilde{\mathbf{U}}_\mathrm{N}\mathbf{\Lambda}_\mathrm{N}^{1/2}) \begin{pmatrix} \mathbf{\Lambda}_\mathrm{N} & \mathbf{I}_\mathrm{K} \\ \mathbf{I}_\mathrm{K} & 0 \end{pmatrix}$$

It turns out that it is possible to evaluate the behaviour of the entries of matrix $\mathbf{F}_\mathrm{N}(z)$ when $\mathrm{N} \to +\infty$. More precisely, the entries of $\mathbf{F}_\mathrm{N}(z)$ depend on bilinear forms of matrices $\mathbf{Q}_{\mathrm{W,N}}(z)$, $\mathbf{Q}_{\mathrm{W,N}}(z)\mathbf{W}_\mathrm{N}$, and $\mathbf{W}_\mathrm{N}^*\mathbf{Q}_\mathrm{N}(z)\mathbf{W}_\mathrm{N} = \mathbf{I} + z\tilde{\mathbf{Q}}_{\mathrm{W,N}}(z)$. (1.2.16, 1.2.18, 1.2.19) imply immediately that $\mathbf{F}_\mathrm{N}(z)$ converge towards matrix $\mathbf{F}(z)$ given by

$$\mathbf{F}(z) = \begin{pmatrix} \mathbf{I}_\mathrm{K} + m(z)\,\mathbf{\Lambda} & m(z)\,\mathbf{I}_\mathrm{K} \\ (1 + z\,\tilde{m}(z))\,\mathbf{\Lambda} & \mathbf{I}_\mathrm{K} \end{pmatrix}$$

where $\mathbf{\Lambda}$ is the diagonal matrix $\mathbf{\Lambda} = \mathrm{Diag}(\lambda_1, \ldots, \lambda_K)$ and where we have denoted Stieltjes transforms $m_{d,\sigma^2}(z)$ and $\tilde{m}_{d,\sigma^2}(z)$ by $m(z)$ and $\tilde{m}(z)$ in order to simplify the notations. Therefore, if $z \neq 0$ is chosen real and outside $[\lambda^- - \epsilon, \lambda^+ + \epsilon]$, the limit form of equation $\det(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* - z\mathbf{I}) = 0$ is

$$\det(\mathbf{\Lambda} - w(z)\mathbf{I}_K) = 0 \tag{1.2.29}$$

where $w(z)$ is defined by (1.2.10). Using the properties of function $w(z)$, we obtain immediately that (1.2.29) has K_s solutions that coincide with the $\phi(\lambda_k) = \rho_k$ for $k = 1, \ldots, K_s$. This, and some extra technical details lead to (1.2.22).

We now justify (1.2.24). Again, we do not follow [10], and rather use (1.2.28) as well as the approach developed in [75]. For this, we consider $k \leq K_s$. As $\hat{\lambda}_{k,N}$ converges towards $\rho_k > \lambda^+$, projection matrix $\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*$ can be written as

$$\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^* = -\frac{1}{2i\pi}\int_{\mathscr{C}_k} \left(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^* - z\mathbf{I}\right)^{-1} \, dz \tag{1.2.30}$$

where $\mathscr{C}_k$ is a contour enclosing the eigenvalue $\hat{\lambda}_{k,N}$, ρ_k, and not the other eigenvalues of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$. (1.2.24) is based on the observation that it is possible to evaluate the almost sure asymptotic behaviour of the bilinear forms of $\left(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^* - z\mathbf{I}\right)^{-1} = \mathbf{Q}_N(z)$ for each $z \in \mathscr{C}_k$. For this, we express $\mathbf{Q}_N(z)$ in terms of $\mathbf{Q}_{W,N}(z)$ by taking the inverse of Eq. (1.2.27). After some algebra, we obtain that

$$\mathbf{Q}_N = \mathbf{Q}_{W,N} - \mathbf{Q}_{W,N}(\mathbf{U}_N, \mathbf{W}_N\tilde{\mathbf{U}}_N\mathbf{\Lambda}_N^{1/2}) \begin{pmatrix} \mathbf{\Lambda}_N & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix} \times \tag{1.2.31}$$

$$\left[\mathbf{I} + \begin{pmatrix} \mathbf{U}_N^* \\ \mathbf{\Lambda}_N^{1/2}\tilde{\mathbf{U}}_N^*\mathbf{W}_N^* \end{pmatrix} \mathbf{Q}_W(\mathbf{U}_N, \mathbf{W}_N\tilde{\mathbf{U}}_N\mathbf{\Lambda}_N^{1/2}) \begin{pmatrix} \mathbf{\Lambda} & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix}\right]^{-1} \begin{pmatrix} \mathbf{U}_N^* \\ \mathbf{\Lambda}_N^{1/2}\tilde{\mathbf{U}}_N^*\mathbf{W}_N^* \end{pmatrix} \mathbf{Q}_W$$

Using (1.2.16, 1.2.18, 1.2.19) as above, we obtain that

$$\begin{pmatrix} \mathbf{U}_N^* \\ \mathbf{\Lambda}_N^{1/2}\tilde{\mathbf{U}}_N^*\mathbf{W}_N^* \end{pmatrix} \mathbf{Q}_W(\mathbf{U}_N, \mathbf{W}_N\tilde{\mathbf{U}}_N\mathbf{\Lambda}_N^{1/2}) \xrightarrow{a.s} \begin{pmatrix} m(z)\mathbf{I} & 0 \\ 0 & (1 + z\tilde{m}(z))\,\mathbf{\Lambda} \end{pmatrix}$$

Using again (1.2.16, 1.2.18, 1.2.19), we obtain after some algebra that for each sequence $\mathbf{a}_N, \mathbf{b}_N$ of M–dimensional unit vectors, it holds that

$$\mathbf{a}_N^* \left(\mathbf{Q}_N(z) - \mathbf{S}_N(z)\right) \mathbf{b}_N \to 0 \ a.s. \tag{1.2.32}$$

for each $z \in \mathbb{C} - \mathbb{R}^+$, where $\mathbf{S}_N(z)$ is the $M \times M$ matrix-valued function defined by

$$\mathbf{S}_N(z) = \left(-z(1 + \sigma^2\tilde{m}(z)) + \frac{\mathbf{B}_N\mathbf{B}_N^*}{1 + \sigma^2 d m(z)}\right)^{-1} \tag{1.2.33}$$

Moreover, it is easily seen that the convergence in (1.2.32) is uniform on each compact subset of $\mathbb{C} - \mathbb{R}^+$. This property is however not sufficient to claim that

$$\mathbf{a}_N^* \left(\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^* - \frac{1}{2i\pi}\int_{\mathscr{C}_k} \mathbf{S}_N(z)\, dz\right) \mathbf{b}_N \to 0 \tag{1.2.34}$$

To justify this, we check that (1.2.32) holds uniformly on contour $\mathscr{C}_k$. For this, we first verify that function $\mathbf{S}_N(z)$ is analytic in a neighbourhood of $\mathscr{C}_k$. By (1.2.7), we obtain that $w(z) = z(1 + \sigma^2\tilde{m}(z))(1 + \sigma^2 d m(z))$. Using (1.2.11), $\mathbf{S}_N(z)$ can thus also be written as

$$\mathbf{S}_N(z) = \frac{w(z)}{w(z) + \sigma^2 d} \left(\mathbf{B}_N\mathbf{B}_N^* - w(z)\,\mathbf{I}\right)^{-1}$$

Using the properties of $w(z)$ recalled in paragraph 1.2.1, it is easily checked that when z describes contour $\mathscr{C}_k$, $w(z)$ describes a contour $\mathscr{D}_k$ enclosing λ_k, eigenvalue $\lambda_{k,\mathrm{N}}$, and not the other eigenvalues of $\mathbf{B}_\mathrm{N}\mathbf{B}_\mathrm{N}^*$. Therefore, $\mathbf{B}_\mathrm{N}\mathbf{B}_\mathrm{N}^* - w(z)\mathbf{I}$ is invertible in a neighbourhood of $\mathscr{C}_k$, leading to the conclusion that $\mathbf{S}_\mathrm{N}(z)$ is analytic in any such neighbourhood. As almost surely for N large enough, $\mathbf{Q}_\mathrm{N}(z)$ is also analytic there, (1.2.32) also holds uniformly on $\mathscr{C}_k$. This justifies (1.2.34). In order to evaluate the contour integral in (1.2.34), we write $\mathbf{S}_\mathrm{N}(z)$ as $\mathbf{S}_\mathrm{N}(z) = \frac{1}{w'(z)}\mathbf{S}_\mathrm{N}(z)\, w'(z)$, and obtain that

$$\frac{1}{2i\pi}\int_{\mathscr{C}_k} \mathbf{S}_\mathrm{N}(z)\,dz = \frac{1}{2i\pi}\int_{\mathscr{C}_k} \frac{1}{w'(z)}\,\frac{w(z)}{w(z)+\sigma^2 d}\left(\mathbf{B}_\mathrm{N}\mathbf{B}_\mathrm{N}^* - w(z)\mathbf{I}\right)^{-1} w'(z)\,dz$$

Differentiating (1.2.13) w.r.t. z, we get immediately that

$$w'(z) = \frac{w^2(z)}{w^2(z)-\sigma^4 d} \tag{1.2.35}$$

Therefore, it holds that

$$\frac{1}{2i\pi}\int_{\mathscr{C}_k} \mathbf{S}_\mathrm{N}(z)\,dz = \frac{1}{2i\pi}\int_{\mathscr{D}_k} \frac{w^2-\sigma^4 d}{w(w+\sigma^2 d)}\left(\mathbf{B}_\mathrm{N}\mathbf{B}_\mathrm{N}^* - w(z)\mathbf{I}\right)^{-1} dw$$

Using residue theorem, we obtain that this contour integral coincides with $\frac{\lambda_{k,\mathrm{N}}^2-\sigma^4 d}{\lambda_{k,\mathrm{N}}(\lambda_{k,\mathrm{N}}+\sigma^2 d)}\,\mathbf{u}_{k,\mathrm{N}}\mathbf{u}_{k,\mathrm{N}}^*$. Therefore, (1.2.34) implies (1.2.24).

Remark 1. *It is important to notice that the properties that have been used in the above discussion are the following:*

- *(i) The eigenvalue distribution of* $\mathbf{W}_\mathrm{N}\mathbf{W}_\mathrm{N}^*$ *converges almost surely towards the Marcenko-Pastur distribution*
- *(ii) For each* $\epsilon > 0$*, all the non zero eigenvalues of* $\mathbf{W}_\mathrm{N}\mathbf{W}_\mathrm{N}^*$ *are located inside* $[\lambda^- - \epsilon, \lambda^+ + \epsilon]$ *almost surely for* N *large enough*
- *(iii) (1.2.16, 1.2.18, 1.2.19) hold uniformly w.r.t.* z *on each compact subset of* $\mathbb{C}^* - [\lambda^-, \lambda^+]$.

Therefore, Theorem 1.2.2 holds if Gaussian random matrix with i.i.d. entries $\mathbf{W}_\mathrm{N}$ *is replaced by a random matrix verifying the conditions (i,ii, iii).*

Remark 2. *It is possible to evaluate the asymptotic behaviour of the bilinear forms of matrix* $\tilde{\mathbf{Q}}_\mathrm{N}(z)$ *using the same kind of calculations. After some algebra, we obtain that*

$$\tilde{\mathbf{a}}_\mathrm{N}^*\left(\tilde{\mathbf{Q}}_\mathrm{N}(z) - \tilde{\mathbf{S}}_\mathrm{N}(z)\right)\tilde{\mathbf{b}}_\mathrm{N} \to 0 \ a.s. \tag{1.2.36}$$

for each $z \in \mathbb{C} - \mathbb{R}^+$*, where* $\tilde{\mathbf{S}}_\mathrm{N}(z)$ *is the* $\mathrm{N}\times\mathrm{N}$ *matrix-valued function defined by*

$$\tilde{\mathbf{S}}_\mathrm{N}(z) = \left(-z(1+\sigma^2 d\, m(z)) + \frac{\mathbf{B}_\mathrm{N}^*\mathbf{B}_\mathrm{N}}{1+\sigma^2\tilde{m}(z)}\right)^{-1} \tag{1.2.37}$$

Moreover, (1.2.36) holds uniformly on each compact subset of $\mathbb{C} - \mathbb{R}^+$.

1.2.3 The Information plus Noise model.

We now briefly review some results related to the information plus noise model $\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N$ where $\mathbf{B}_N$ is a deterministic matrix whose rank may scale with N. We refer the reader to [20] and to [72] for more details. In order to simplify the presentation, we assume that $\mathbf{B}_N$ satisfies the condition

$$\sup_N \|\mathbf{B}_N\| < +\infty \tag{1.2.38}$$

It is well known that the empirical eigenvalue distribution sequence $(\hat{\mu}_N)_{N\geq 1}$ of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ has almost surely the same asymptotic behaviour than a sequence of probability measure $(\mu_N)_{N\geq 1}$ carried by $\mathbb{R}^+$. Measure μ_N is characterized by its Stieltjes transform $m_N(z)$ which appears as the unique solution of the equation

$$m_N(z) = \frac{1}{M}\mathrm{Tr}\left[-z(1+\sigma^2\tilde{m}_N(z)) + \frac{\mathbf{B}_N\mathbf{B}_N^*}{1+\sigma^2 d_N m_N(z)}\right]^{-1} \tag{1.2.39}$$

in the class $\mathcal{S}(\mathbb{R}^+)$ of Stieltjes transforms of probability measures carried by $\mathbb{R}^+$. In (1.2.39), $\tilde{m}_N(z)$ is defined by $\tilde{m}_N(z) = d_N m_N(z) - \frac{1-d_N}{z}$, and thus coincides with the Stieltjes transform of measure $\tilde{\mu}_N = d_N\mu_N + (1-d_N)\delta_0$. Therefore, it holds that

$$\frac{1}{M}\mathrm{Tr}(\mathbf{Q}_N(z)) - m_N(z) \to 0 \; a.s.$$

for each $z \in \mathbb{C} - \mathbb{R}^+$. Matrix $\mathbf{T}_N(z)$ defined by

$$\mathbf{T}_N(z) = \left[-z(1+\sigma^2\tilde{m}_N(z)) + \frac{\mathbf{B}_N\mathbf{B}_N^*}{1+\sigma^2 d_N m_N(z)}\right]^{-1} \tag{1.2.40}$$

coincides with the Stieltjes transform of a M × M positive matrix valued measure $\boldsymbol{\mu}_N$ carried by $\mathbb{R}^+$ satisfying $\boldsymbol{\mu}_N(\mathbb{R}^+) = \mathbf{I}_M$, i.e.

$$\mathbf{T}_N(z) = \int_{\mathbb{R}^+} \frac{d\boldsymbol{\mu}_N(\lambda)}{\lambda - z}\, d\lambda$$

$\mathbf{T}_N(z)$ is thus holomorphic on $\mathbb{C} - \mathbb{R}^+$. It can be shown that for each sequence of deterministic M × M matrices $\mathbf{D}_N$ such that $\sup_N \|\mathbf{D}_N\| < +\infty$, it holds that

$$\frac{1}{M}\mathrm{Tr}\left((\mathbf{Q}_N(z) - \mathbf{T}_N(z))\mathbf{D}_N\right) \to 0 \; a.s. \tag{1.2.41}$$

for each $z \in \mathbb{C} - \mathbb{R}^+$. Moreover, for each sequence of deterministic unit vectors $\mathbf{a}_N, \mathbf{b}_N$, we have

$$\mathbf{a}_N^*\left(\mathbf{Q}_N(z) - \mathbf{T}_N(z)\right)\mathbf{b}_N \to 0 \; a.s. \tag{1.2.42}$$

for each $z \in \mathbb{C} - \mathbb{R}^+$. It is interesting to point out the differences between $\mathbf{T}_N(z)$ and $\mathbf{S}_N(z)$ defined by (1.2.33): the expressions are the same, except that m_N and $\tilde{m}_N$ are replaced by their Marcenko-Pastur analogues m and $\tilde{m}$. As m_N and $\tilde{m}_N$ have the same asymptotic behaviour that m and $\tilde{m}$ in the case where $K/N \to 0$, the entries of $\mathbf{T}_N(z)$ and $\mathbf{S}_N(z)$ have the same behaviour.

We also note that matrix $\tilde{\mathbf{T}}_N(z)$ defined by

$$\tilde{\mathbf{T}}_N(z) = \left[-z(1+\sigma^2 d_N m_N(z)) + \frac{\mathbf{B}_N^*\mathbf{B}_N}{1+\sigma^2\tilde{m}_N(z)}\right]^{-1} \tag{1.2.43}$$

is the Stieltjes transform of a N × N positive matrix valued measure $\tilde{\boldsymbol{\mu}}_N$ carried by $\mathbb{R}^+$ satisfying $\tilde{\boldsymbol{\mu}}_N(\mathbb{R}^+) = \mathbf{I}_N$, and $\tilde{m}_N(z)$ coincides with $\tilde{m}_N(z) = \frac{1}{N}\mathrm{Tr}\left(\tilde{\mathbf{T}}_N(z)\right)$. Moreover, (1.2.41) and (1.2.42) still remain valid when $\mathbf{Q}_N(z)$ and $\mathbf{T}_N(z)$ are replaced by $\tilde{\mathbf{Q}}_N(z)$ and $\tilde{\mathbf{T}}_N(z)$ respectively. When $K/N \to 0$, the entries of $\tilde{\mathbf{T}}_N(z)$ and of $\tilde{\mathbf{S}}_N(z)$ defined by (1.2.36) have the same behaviour.

We finally mention that the support $\mathcal{S}_N$ of measure μ_N can be characterized (see [72]), and that it was proved that almost surely, for N large enough, there is no eigenvalue of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ in intervals $[a, b]$ for which $[a-\epsilon, b+\epsilon] \cap \mathcal{S}_N = \emptyset$ for each N large enough ([48]).

1.3 Behaviour of the eigenvalues of the empirical spatio-temporal covariance matrices of temporally and spatially complex Gaussian noise.

We consider a M-dimensional time series $(\mathbf{y}_n)_{n=1,\ldots,N}$ observed between time 1 and time $N+L-1$ where L is a certain integer. We consider the ML–dimensional random vector $\mathbf{y}_n^{(L)}$ defined by

$$\mathbf{y}_n^{(L)} = (\mathbf{y}_{1,n},\ldots,\mathbf{y}_{1,n+L-1},\ldots,\mathbf{y}_{M,n},\ldots,\mathbf{y}_{M,n+L-1})^T \tag{1.3.1}$$

The $ML \times ML$ empirical covariance matrix

$$\hat{\mathbf{R}}_{y,N}^{(L)} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{y}_n^{(L)} \mathbf{y}_n^{(L)*} \tag{1.3.2}$$

is usually called an empirical spatio-temporal covariance matrix. This kind of matrix plays of course an important role in a number of statistical inference problems related to multivariate time series. Therefore, it is potentially useful to study the properties of the eigenvalues of $\hat{\mathbf{R}}_{y,N}^{(L)}$. When $N \to +\infty$ and that ML remains fixed, the law of large number implies that $\|\hat{\mathbf{R}}_{y,N}^{(L)} - \mathbb{E}(\mathbf{y}_n^{(L)}\mathbf{y}_n^{(L)*})\| \to 0$. However, in a number of contexts, N is not much larger than ML, and the above mentioned traditional regime does not allow to predict the properties of $\hat{\mathbf{R}}_{y,N}^{(L)}$. It is therefore of great interest to study the matrix $\hat{\mathbf{R}}_{y,N}^{(L)}$ in the asymptotic regime

$$N \to +\infty, ML \to +\infty,\ c_N = \frac{ML}{N} \to c > 0 \tag{1.3.3}$$

This kind of problems was essentially studied in previous works when $(\mathbf{y}_n)_{n=1,\ldots,N}$ coincides with a sequence $(\mathbf{v}_n)_{n=1,\ldots,N}$ of i.i.d. zero mean complex Gaussian random M–dimensional vectors for which $\mathbb{E}(\mathbf{v}_n\mathbf{v}_n^*) = \sigma^2 \mathbf{I}_M$. In order to simplify the notations, we denote by $(w_{m,n})_{m=1,\ldots,M,n=1,\ldots,N}$ the i.i.d. sequence of $\mathcal{N}_c(0,\frac{\sigma^2}{N})$ random variables defined by $w_{m,n} = \frac{\mathbf{v}_{m,n}}{\sqrt{N}}$ where $\mathbf{v}_{m,n}$ represents element m of vector $\mathbf{v}_n$. For each $m = 1,\ldots,M$, $\mathbf{W}_N^{(m)}$ represents the $L \times N$ Hankel matrix whose entries are given by

$$\left(\mathbf{W}_N^{(m)}\right)_{i,j} = w_{m,i+j-1},\ 1 \le i \le L, 1 \le j \le N \tag{1.3.4}$$

If we define $\mathbf{W}_N$ as the $ML \times N$ matrix

$$\mathbf{W}_N = \begin{pmatrix} \mathbf{W}_N^{(1)} \\ \mathbf{W}_N^{(2)} \\ \vdots \\ \mathbf{W}_N^{(M)} \end{pmatrix} \tag{1.3.5}$$

then, it is easily seen that the empirical spatio-temporal covariance matrix $\hat{\mathbf{R}}_{v,N}^{(L)}$ coincides with matrix $\mathbf{W}_N\mathbf{W}_N^*$.

Matrix $\mathbf{W}_N$ can be interpreted as a block line matrix whose M line blocks $(\mathbf{W}_N^{(m)})_{m=1,\ldots,M}$ are independent and identically distributed. When L does not scale with N, i.e. if M and N are of the same order of magnitude, the works of Girko ([26], Chapter 16) and of [24] allow to conclude that the empirical eigenvalue distribution of $\mathbf{W}_N\mathbf{W}_N^*$ converges towards the Marcenko-Pastur distribution μ_{c,σ^2}. When M is reduced to 1, matrix $\mathbf{W}_N$ is reduced to a Hankel matrix. When the $w_{1,n}$ for $N < n < N+L$ are forced to 0, matrix $\mathbf{W}_N\mathbf{W}_N^*$ coincides with the traditional empirical estimate of the autocovariance matrix of vector $(w_{1,n},\ldots,w_{1,n+L-1})^T$. Using the moment method, it was shown in this context in [6] that the empirical eigenvalue distribution of $\mathbf{W}_N\mathbf{W}_N^*$ converges towards a non compactly distributed limit distribution. The case where $M \to +\infty$ while L may also converge towards $+\infty$ is studied in [49]. As in the case where L is finite, it is shown that the eigenvalue distribution converges towards μ_{c,σ^2}. More importantly, it is established in [49] that if $L = \mathcal{O}(N^{\alpha})$ with $\alpha < 2/3$, then, almost surely, all the eigenvalues of $\mathbf{W}_N \mathbf{W}_N^*$ are localized in a neighbourhood of the support of μ_{c,σ^2}. More precisely, the main result of [49] is the following Theorem.

Theorem 1.3.1. *When* M *and* N *converge towards* ∞ *in such a way that* $c_N = \frac{ML}{N}$ *converges towards* $c \in (0, +\infty)$*, the eigenvalue distribution of* $\mathbf{W}_N \mathbf{W}_N^*$ *converges weakly almost surely towards the Marcenko-Pastur distribution with parameters* (σ^2, c)*. If moreover*

$$L = \mathcal{O}(N^{\alpha}) \tag{1.3.6}$$

where $\alpha < 2/3$*, then, for each* $\epsilon > 0$*, almost surely, for* N *large enough, all the eigenvalues of* $\mathbf{W}_N \mathbf{W}_N^*$ *are located in the interval* $[\sigma^2(1-\sqrt{c})^2 - \epsilon, \sigma^2(1+\sqrt{c})^2 + \epsilon]$ *if* $c \le 1$*. If* $c > 1$*, almost surely, for* N *large enough,* 0 *is eigenvalue of* $\mathbf{W}_N \mathbf{W}_N^*$ *with multiplicity* $ML - N$*, and the* N *non zero eigenvalues of* $\mathbf{W}_N \mathbf{W}_N^*$ *are located in the interval* $[\sigma^2(1-\sqrt{c})^2 - \epsilon, \sigma^2(1+\sqrt{c})^2 + \epsilon]$

It is standard that the convergence towards the Marcenko-Pastur distribution μ_{c,σ^2} and the almost sure location of the non zero eigenvalues in a neighbourhood of $[\sigma^2(1-\sqrt{c})^2, \sigma^2(1+\sqrt{c})^2]$ imply the almost sure convergence of $\hat{\lambda}_{1,N}$ towards $\sigma^2(1+\sqrt{c})^2$ and of $\hat{\lambda}_{\min(ML,N),N}$ towards $\sigma^2(1-\sqrt{c})^2$. Therefore, Theorem 1.3.1 implies that if the rate of convergence of L towards $+\infty$ is not too fast (i.e. if (1.3.6) holds), or equivalently if $M = \mathcal{O}(N^{\beta})$ with $1/3 < \beta \le 1$), then the eigenvalues of $\mathbf{W}_N \mathbf{W}_N^*$ are localized as if the entries of $\mathbf{W}_N$ were independent identically distributed, a property which is of course not verified. As shown below, properties (1.2.16, 1.2.18, 1.2.19) are also verified for $d = c$. Consequently, matrix $\mathbf{W}_N$ satisfies the properties mentioned in Remark 1. The behaviour of the largest eigenvalues of matrices such as $(\mathbf{B}_N + \mathbf{W}_N)(\mathbf{B}_N + \mathbf{W}_N)^*$ where $\mathbf{B}_N$ represents a deterministic matrix whose rank does not scale with M, N, L will thus appear to be governed by Theorem 1.2.2.

1.4 Contributions of the thesis.

The general topic of the thesis is to study detection/estimation problems for M dimensional signals $(\mathbf{y}_n)_{n\in\mathbb{Z}}$ that can be written as

$$\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p} + \mathbf{v}_n = [\mathbf{h}(z)] s_n + \mathbf{v}_n \tag{1.4.1}$$

where $(s_n)_{n\in\mathbb{Z}}$ is a non observable scalar deterministic sequence and where $\mathbf{h}(z) = \sum_{p=0}^{P-1} \mathbf{h}_p z^{-p}$ is the transfer function of an unknown 1–input / M–outputs linear system. While $\mathbf{h}(z)$ is assumed unknown, P, which represents an upper bound on the number of non zero coefficients of $\mathbf{h}(z)$, is assumed to be known. $(\mathbf{v}_n)_{n\in\mathbb{Z}}$ represents a temporally and spatially complex Gaussian white noise of variance σ^2. In model (1.4.1), signal $[\mathbf{h}(z)] s_n$ represents a useful wideband signal on which various informations have to be inferred from the observation of N samples $(\mathbf{y}_n)_{n\in\mathbb{Z}}$. We have considered the case of a single wideband signal to simplify the exposition, but we feel that the various results we obtained could be generalized in the presence of K wideband signals where K does not scale with M and N.

In this thesis, we address the following problems when M and N converge towards $+\infty$:

- Problem 1. Detection of an unknown wideband signal from $(\mathbf{y}_n)_{n=1,\dots,N}$.
- Problem 2. Estimation of the parameter of a regularized spatio-temporal Wiener filter in the case where a length N training sequence $(s_n)_{n=1,\dots,N}$ is available.
- Problem 3. Subspace estimation of directions of arrivals of narrowband sources using spatial smoothing schemes when the number of snapshots N is much smaller than the number of antennas M.

The various approaches that are proposed are based on the study of the eigenvalue/eigenvector decomposition of the spatio-temporal covariance matrix $\hat{\mathbf{R}}_y^{(L)}$ defined by (1.3.2). In order to understand the structure of $\hat{\mathbf{R}}_y^{(L)}$, we remark that augmented vector $\mathbf{y}_n^{(L)}$ defined by (1.3.1) can be written as

$$\mathbf{y}_n^{(L)} = \mathbf{H}^{(L)} \mathbf{s}_n^{(L)} + \mathbf{v}_n^{(L)} \tag{1.4.2}$$

where $\mathbf{s}_n^{(L)}$ represents $P + L - 1$–dimensional vector $\mathbf{s}_n^{(L)} = (s_{n-(P-1)}, \dots, s_{n+L-1})^T$ and where $ML \times (P + L - 1)$ matrix $\mathbf{H}^{(L)}$ is a block row matrix

$$\mathbf{H}^{(L)} = \begin{pmatrix} \mathbf{H}_1^{(L)} \\ \vdots \\ \mathbf{H}_M^{(L)} \end{pmatrix}$$

where each block $\mathbf{H}_m^{(L)}$ is a $L \times (P + L - 1)$ Toeplitz matrix corresponding to the convolution between sequence s and the impulse response $(\mathbf{h}_{m,p})_{p=0,\dots,P-1}$ of component m $\mathbf{h}_m(z)$ of filter $\mathbf{h}(z)$. If we denote by $\mathbf{Y}^{(L)}$ the $ML \times N$ matrix defined by

$$\mathbf{Y}^{(L)} = (\mathbf{y}_1^{(L)}, \dots, \mathbf{y}_N^{(L)})$$

then (1.4.2) implies that

$$\mathbf{Y}^{(L)} = \mathbf{H}^{(L)} \mathbf{S}^{(L)} + \mathbf{V}^{(L)} \tag{1.4.3}$$

Therefore, $\mathbf{Y}^{(L)}$ is the sum of the low rank $P + L - 1$ deterministic matrix $\mathbf{H}^{(L)} \mathbf{S}^{(L)}$ with random matrix $\mathbf{V}^{(L)}$. In order to simplify the notations, we denote by $\mathbf{\Sigma}^{(L)}$, $\mathbf{B}^{(L)}$, $\mathbf{W}^{(L)}$ the normalized matrices

$$\mathbf{\Sigma}^{(L)} = \frac{\mathbf{Y}^{(L)}}{\sqrt{N}}, \ \mathbf{B}^{(L)} = \frac{\mathbf{H}^{(L)} \mathbf{S}^{(L)}}{\sqrt{N}}, \ \mathbf{W}^{(L)} = \frac{\mathbf{V}^{(L)}}{\sqrt{N}}$$

and remark that $\hat{\mathbf{R}}_y^{(L)} = \mathbf{\Sigma}^{(L)} \mathbf{\Sigma}^{(L)*}$ and that

$$\mathbf{\Sigma}^{(L)} = \mathbf{B}^{(L)} + \mathbf{W}^{(L)} \tag{1.4.4}$$

Therefore, $\hat{\mathbf{R}}_y^{(L)}$ coincides with the Gram matrix $\mathbf{\Sigma}^{(L)} \mathbf{\Sigma}^{(L)*}$ of a low rank deterministic perturbation of the structured random matrix $\mathbf{W}^{(L)}$ which has exactly the same properties than matrix $\mathbf{W}_N$ defined by (1.3.5). The study of the properties of $\mathbf{\Sigma}^{(L)} \mathbf{\Sigma}^{(L)*}$ crucially depend on the behaviour of the rank $P + L - 1$ [1] of $\mathbf{B}^{(L)}$.

In chapter 2, it is assumed that P and L do not scale with M, N, and we address Problem 1 and Problem 2 under this hypothesis. Problem 3 is also studied in chapter 2 because it appears equivalent to the study of the eigenvalue/eigenvector decomposition of the Gram matrix of a structured fixed rank information plus noise model similar to $\mathbf{\Sigma}^{(L)}$. Chapter 3 is devoted to the case where P and L may scale with M and N, and Problem 2 is revisited in this more difficult context.

Contributions of Chapter 2.

- In section 1, we first establish that the largest eigenvalues / eigenvectors of matrix $\mathbf{\Sigma}^{(L)} \mathbf{\Sigma}^{(L)*}$ are governed by Theorem 1.2.2 where parameter d should be replaced by $c = \left(\lim_{N \to +\infty} \frac{M}{N}\right) L$. As briefly mentioned at the end of section 1.3, this result holds because matrix $\mathbf{W}^{(L)}$ satisfies the 3 properties mentioned in Remark 1. As the 2 first properties follow directly from Theorem 1.3.1, it is sufficient to establish (1.2.16), (1.2.18), (1.2.19).

[1] The rank of $\mathbf{H}^{(L)}$ may be smaller than $P + L - 1$ if the components of $\mathbf{h}(z)$ share a comon zero. As this property is generically not satisfied, we assume in the following for the sake of simplication that $\text{Rank}(\mathbf{H}^{(L)}) = P + L - 1$

- In section 2, we consider detection Problem 1. It can be formulated as the following composite hypothesis testing problem in which the hypothesis H_0 corresponds to $\mathbf{y}_n = \mathbf{v}_n$ for $n = 1,\dots,N$ and in which the hypothesis H_1 corresponds to $\mathbf{y}_n = [\mathbf{h}(z)]s_n + \mathbf{v}_n$ for some scalar deterministic sequence s and some SIMO filter $\mathbf{h}(z)$ of maximal degree $P-1$ where P is supposed to be known. The generalized likelihood ratio test cannot be implemented in this context because the maximum likelihood ratio estimate of s and $\mathbf{h}(z)$ cannot be expressed in closed form. Motivated by the GLRT in the context of narrow band signals ([12]), we thus study the reasonable pragmatic test statistics consisting in comparing the sum of the $P+L-1$ largest eigenvalues of matrix $\hat{\mathbf{R}}_y^{(L)}$ to a threshold. Using the results of section 1, we study the conditions under which the above test is consistant, and discuss on the optimum value of the parameter L.

- In section 3, we consider the problem of estimating a regularized spatio-temporal Wiener filter from the observations $(\mathbf{y}_n)_{n=1,\dots,N}$ when a training sequence $(s_n)_{n=1,\dots}$ is available at the receiver side. The regularized spatio-temporal Wiener estimate of sequence s is defined by

$$\hat{s}_n = \hat{\mathbf{g}}_\lambda^{(L)*}\, \mathbf{y}_n^{(L)}$$

where the regularized spatio-temporal Wiener filter $\hat{\mathbf{g}}^{(L)}$ is the ML–dimensional vector given by

$$\hat{\mathbf{g}}^{(L)} = \left(\hat{\mathbf{R}}_y^{(L)} + \lambda \mathbf{I}_{ML}\right)^{-1} \hat{\mathbf{r}}_Y^{(L)}$$

where $\hat{\mathbf{r}}_Y^{(L)}$ is the empirical cross covariance between the augmented observations $(\mathbf{y}_n^{(L)})_{n=1,\dots,N}$ and the training sequence $(s_n)_{n=1,\dots,N}$. Our main contribution is the derivation of a new estimation scheme of the regularization parameter λ that consists in choosing λ so as to maximize the signal to noise plus interference ratio $\mathrm{SINR}(\lambda)$ provided by the regularized spatio-temporal Wiener estimate. The SINR appears to be a random variable depending on the noise samples corrupting the observations $(\mathbf{y}_n)_{n=1,\dots,N}$. However, we prove that in the high dimension regime, it converges almost surely towards a deterministic term $\phi(\lambda)$, depending on λ and on the coefficients of $\mathbf{h}(z)$, and that can be expressed in closed form. Although the coefficients of $\mathbf{h}(z)$ are unknown, we prove that $\phi(\lambda)$ can be estimated consistently from the observations $(\mathbf{y}_n)_{n=1,\dots,N}$ by a term $\hat{\phi}(\lambda)$ for each λ, and we propose to choose the regularization parameter that maximizes $\hat{\phi}(\lambda)$. In order to mention the technical results that we develop, we consider the normalized matrix $\mathbf{\Sigma}^{(L)}$ and the resolvent $\mathbf{Q}(z)$ of matrix $\mathbf{\Sigma}^{(L)}\mathbf{\Sigma}^{(L)*}$. Our results are based on the characterization of the asymptotic behaviour of bilinear forms of matrices $\mathbf{Q}(z)$, $\mathbf{Q}(z)\mathbf{\Sigma}^{(L)}$, $\mathbf{\Sigma}^{(L)*}\mathbf{Q}^2(z)\mathbf{\Sigma}^{(L)}$, $\mathbf{\Sigma}^{(L)*}\mathbf{Q}(z)\mathbf{H}^{(L)}\mathbf{H}^{(L)*}\mathbf{\Sigma}^{(L)}$ when z lies on the negative real axis. For this, we use the hypothesis that $P+L-1$ does not scale with M and N, and take (1.2.31) as a starting point to express the above mentioned bilinear forms in terms of bilinear forms of matrices depending on the resolvent $\mathbf{Q}_{W^{(L)}}(z)$ of the noise part $\mathbf{W}^{(L)}$ of $\mathbf{\Sigma}^{(L)}$ whose behaviour is known.

- In section 4, we address the estimation of the directions of arrival of K narrow band sources impinging on a large uniform linear array of sensors in the case where the number of snapshots N is large, but much smaller than the number of sensors M. In this context, it is standard to use spatial smoothing technics in order to generate L non overlapping subarrays of $M-L+1$ sensors, and to multiply by L the number of snapshots. More precisely, the observations $(\mathbf{y}_n)_{n=1,\dots,N}$ follow the classical narrow band array processing model

$$\mathbf{y}_n = \mathbf{A}_M\, \mathbf{s}_n + \mathbf{v}_n$$

where $\mathbf{A}_M = (\mathbf{a}_M(\theta_1),\dots,\mathbf{a}_M(\theta_K))$ is the $M\times K$ matrix of directional vectors associated to the K sources.

For each n, we denote by $\mathscr{Y}_n^{(L)}$ the $(M-L+1)\times L$ Hankel matrix defined by

$$\mathscr{Y}_n^{(L)} = \begin{pmatrix} \mathbf{y}_{1,n} & \mathbf{y}_{2,n} & \cdots & \cdots & \mathbf{y}_{L,n} \\ \mathbf{y}_{2,n} & \mathbf{y}_{3,n} & \cdots & \cdots & \mathbf{y}_{L+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_{M-L+1,n} & \mathbf{y}_{M-L+2,n} & \cdots & \cdots & \mathbf{y}_{M,n} \end{pmatrix} \tag{1.4.5}$$

Column l of matrix $\mathscr{Y}_n^{(L)}$ corresponds to the observation on subarray l at time n. Collecting all the observations on the various subarrays allows to obtain NL snapshots, thus increasing artificially the number of observations. We define $\mathbf{Y}_N^{(L)}$ as the $(M-L+1)\times NL$ block-Hankel matrix given by

$$\mathbf{Y}_N^{(L)} = \left(\mathscr{Y}_1^{(L)}, \ldots, \mathscr{Y}_N^{(L)}\right) \tag{1.4.6}$$

Matrix $\mathbf{Y}_N^{(L)}$ has a block Hankel structure similar to matrix defined by (1.4.3), the difference being that the Hankel matrices are block columns in (1.4.6) while they are block lines in (1.4.3). It is therefore expected that the mathematical results developed to address Problems 1,2 can be used in the context of Problem 3. It is easily seen that $\mathbf{Y}_N^{(L)}$ is the sum of a rank K matrix whose range coincides with the space generated by the $(M-L+1)$–dimensional vectors $\mathbf{a}_{M-L+1}(\theta_1),\ldots,\mathbf{a}_{M-L+1}(\theta_K))$ with the noise matrix $\mathbf{V}_N^{(L)}$ defined in the same way than $\mathbf{Y}_N^{(L)}$. As

$$\mathbb{E}\left(\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{NL}\right) = \sigma^2 \mathbf{I}_{M-L+1}$$

it possible to develop consistent subspace estimation methods of the angles $(\theta_k)_{k=1,\ldots,K}$ in the asymptotic regime where $NL \to +\infty$ and $M-L+1$ remains fixed. This regime corresponds to values of L for which the virtual subarrays have a small number of elements, thus leading to poor resolution methods. It is thus much more relevant to address the case where $M-L+1 \simeq M$, i.e. $\frac{L}{M} \to 0$. We thus consider an asymptotic regime in which $N \to +\infty$, $M \to +\infty$ and $c_N = \frac{M}{NL} \to c$ where $c > 0$. Adapting the above mentioned results to the present context, we are able to characterize the largest eigenvalues and corresponding eigenvectors of the empirical covariance matrix $\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL$ provided $N = \mathcal{O}(M^\beta)$ for $1/3 < \beta \leq 1$, and to deduce from this a consistent subspace estimation method of the directional parameters.

Contributions of Chapter 3. In Chapter 3, we address the case where the observation $\mathbf{y}$ is generated by a general model

$$\mathbf{y}_n = \mathbf{x}_n + \mathbf{v}_n$$

where $\mathbf{x}$ is a deterministic signal which is not necessarily defined as a filtered version of a scalar sequence. In this context, matrix $\mathbf{Y}^{(L)} = \mathbf{X}^{(L)} + \mathbf{V}^{(L)}$ where $\mathbf{X}^{(L)}$ is not necessarily rank deficient. In this more general context, we study the behaviour of the empirical eigenvalue distribution of normalized matrix $\mathbf{\Sigma}^{(L)}\mathbf{\Sigma}^{(L)*}$ in the case where $L = \mathcal{O}(N^\alpha)$ for $\alpha < 2/3$. Generalizing the Gaussian tools used in [49], we prove that the normalized traces and the bilinear forms of the resolvent $\mathbf{Q}(z)$ of matrix $\mathbf{\Sigma}^{(L)}\mathbf{\Sigma}^{(L)*}$ have the same behaviour than the normalized traces and the bilinear forms of a deterministic matrix-valued function $\mathbf{T}(z)$ defined as the solution of a certain (complicated) equation. This result implies that the empirical eigenvalue distribution $\hat{\mu}_N$ of $\mathbf{\Sigma}^{(L)}\mathbf{\Sigma}^{(L)*}$ has a deterministic behaviour. However, we have not been able to characterize the properties of the corresponding deterministic approximation of $\hat{\mu}_N$ (in particular its support), and to obtain results concerning the almost sure location of the eigenvalues of $\mathbf{\Sigma}^{(L)}\mathbf{\Sigma}^{(L)*}$. However, the results of Chapter 3 are useful in Chapter 4.

Contributions of Chapter 4. In Chapter 4, we consider again the wide band model (1.4.1), but assume that P and L may scale with N, and that $c_N = \frac{ML}{N}$ converges towards a non zero constant c. We assume that when P and L converge towards $+\infty$, $P = \mathcal{O}(L)$. This implies that the rank of normalized matrix $\mathbf{B}^{(L)} = \frac{\mathbf{H}^{(L)}\mathbf{S}^{(L)}}{\sqrt{N}}$ may scale with M and N, but that $\text{Rank}(\mathbf{B}^{(L)})/N \to 0$.

- In section 3 and 4, we establish that matrix $\mathbf{T}(z)$ can be replaced by matrix defined by (1.2.40) corresponding to a standard information plus noise model in which the noise matrix is i.i.d. Moreover, scalar Stieltjes transform $m_N(z)$ can also be replaced by the Stieltjes transform of the Marcenko-Pastur with parameters (c, σ^2). If more information could be obtained on the location of the eigenvalues of $\boldsymbol{\Sigma}^{(L)}\boldsymbol{\Sigma}^{(L)*}$, this result could be useful to understand the behaviour of the projection matrices on certain eigenspaces.

- In section 5, we take benefit of the results of Chapter 3 to revisit Problem 2 in the case where P and L may converge towards $+\infty$. In particular, using Gaussian tools, we are able to generalize the asymptotic behaviours of bilinear forms of $\mathbf{Q}(z)\boldsymbol{\Sigma}^{(L)}$, $\boldsymbol{\Sigma}^{(L)*}\mathbf{Q}^2(z)\boldsymbol{\Sigma}^{(L)}$ found in section 3 of Chapter 2 when $L = \mathcal{O}(N^\alpha)$ for $\alpha < 2/3$. The study of the SINR also needs to evaluate bilinear forms of $\boldsymbol{\Sigma}^{(L)*}\mathbf{Q}(z)\mathbf{H}^{(L)}\mathbf{H}^{(L)*}\boldsymbol{\Sigma}^{(L)}$, which, in the case where P and L are fixed, is an easy task. When P and L scale with M and N, the problem appears more difficult. We however found a satisfying result provided $\frac{L}{M} \to 0$ (a condition nearly equivalent to $L = \mathcal{O}(N^\alpha)$ for $\alpha < 1/2$) because a number of complicated terms vanish. Under this extra assumption, the SINR has the same asymptotic expression than in the case where P and L are fixed. However, we feel that the asymptotic behaviour of $\boldsymbol{\Sigma}^{(L)*}\mathbf{Q}(z)\mathbf{H}^{(L)}\mathbf{H}^{(L)*}\boldsymbol{\Sigma}^{(L)}$ could also be characterized when $L = \mathcal{O}(N^\alpha)$ for $1/2 \le \alpha < 2/3$, and that a correcting term could appear in the expression of the limit form of the SINR.

We finally mention the publications connected to this thesis.

- G.T. Pham, P. Loubaton, P. Vallet, "Performance analysis of spatial smoothing schemes in the context of large arrays", IEEE Trans. on Signal Processing, vol. 64, no. 1, pp. 160-172, January 2016.
- G.T. Pham, P. Loubaton, P. Vallet, "Performance analysis of spatial smoothing schemes in the context of large arrays", Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on Year: 2015, Pages: 2824 - 2828.
- G.T. Pham; P. Loubaton, "Applications of large empirical spatio-temporal covariance matrix in multipath channels detection", Signal Processing Conference (EUSIPCO), Nice, September 2015.
- G.T. Pham; P. Loubaton, "Performances des filtres de Wiener spatio-temporels entrainés: le cas des grandes dimensions", Proc. Colloque Gretsi, Lyon, September 2015.
- G.T. Pham; P. Loubaton, "Optimization of the loading factor of regularized estimated spatial-temporal Wiener filters in large system case", Proc. of Statistical Signal Processing Workshop (SSP), Palma de Majorque, June 2016.

Chapter 2

Spatial-temporal Gaussian information plus noise spiked model

2.1 Introduction

In this chapter, we assume that the M dimensional signals $(\mathbf{y}_n)_{n\in\mathbb{Z}}$ is generated as

$$\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p} + \mathbf{v}_n = [\mathbf{h}(z)] s_n + \mathbf{v}_n \tag{2.1.1}$$

where $(s_n)_{n\in\mathbb{Z}}$ is a non observable scalar deterministic sequence and where $\mathbf{h}(z) = \sum_{p=0}^{P-1} \mathbf{h}_p z^{-p}$ is the transfer function of an unknown 1–input / M–outputs linear system. While $\mathbf{h}(z)$ is assumed unknown, P, which represents an upper bound on the number of non zero coefficients of $\mathbf{h}(z)$, is assumed to be known. $(\mathbf{v}_n)_{n\in\mathbb{Z}}$ represents a temporally and spatially complex Gaussian white noise of variance σ^2.

In the following, we propose approaches based on the empirical spatio-temporal covariance matrices of $\mathbf{y}$ to address 3 well defined inference problems. We recall that if $\mathbf{y}_n^{(L)}$ is the ML–dimensional vector defined by

$$\mathbf{y}_n^{(L)} = (\mathbf{y}_{1,n}, \dots, \mathbf{y}_{1,n+L-1}, \dots, \mathbf{y}_{M,n}, \dots, \mathbf{y}_{M,n+L-1})^T$$

then the spatio-temporal covariance matrix $\hat{\mathbf{R}}_y^{(L)}$ is the ML × ML matrix defined by

$$\hat{\mathbf{R}}_{y,N}^{(L)} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{y}_n^{(L)} \mathbf{y}_n^{(L)*} = \frac{\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}}{N} \tag{2.1.2}$$

where $\mathbf{Y}_N^{(L)} = (\mathbf{y}_1^{(L)}, \dots, \mathbf{y}_N^{(L)})$. As mentioned in Chapter 1, matrix $\mathbf{Y}_N^{(L)}$ can be written as

$$\mathbf{Y}_N^{(L)} = \mathbf{H}^{(L)} \mathbf{S}_N^{(L)} + \mathbf{V}_N^{(L)} \tag{2.1.3}$$

where ML × (P + L − 1) matrix $\mathbf{H}^{(L)}$ is the block row matrix

$$\mathbf{H}^{(L)} = \begin{pmatrix} \mathbf{H}_1^{(L)} \\ \vdots \\ \mathbf{H}_M^{(L)} \end{pmatrix}$$

where each row block $\mathbf{H}_m^{(L)}$ is a L × (P + L − 1) Toeplitz matrix corresponding to the convolution between sequence s and the impulse response $(\mathbf{h}_{m,p})_{p=0,\dots,P-1}$ of component m $\mathbf{h}_m(z)$ of filter $\mathbf{h}(z)$. In (2.1.3), (P + L − 1) × N matrix $\mathbf{S}_N^{(L)}$ is the Hankel matrix whose entries are defined by $(\mathbf{S}_N^{(L)})_{i,j} = s_{i+j-P}$.

Therefore, $\mathbf{Y}_N^{(L)}$ is the sum of the low rank $P+L-1$ deterministic matrix $\mathbf{H}^{(L)}\mathbf{S}_N^{(L)}$ with random matrix $\mathbf{V}_N^{(L)}$. All along this chapter, we assume that P and L do not scale with M and N. Therefore, if we denote $\boldsymbol{\Sigma}_N^{(L)}$, $\mathbf{B}_N^{(L)}$, $\mathbf{W}_N^{(L)}$ the normalized matrices

$$\boldsymbol{\Sigma}_N^{(L)} = \frac{\mathbf{Y}_N^{(L)}}{\sqrt{N}},\ \mathbf{B}_N^{(L)} = \frac{\mathbf{H}^{(L)}\mathbf{S}_N^{(L)}}{\sqrt{N}},\ \mathbf{W}_N^{(L)} = \frac{\mathbf{V}_N^{(L)}}{\sqrt{N}}$$

then normalized matrix $\boldsymbol{\Sigma}_N^{(L)} = \mathbf{B}_N^{(L)} + \mathbf{W}_N^{(L)}$ can be interpreted as a finite rank deterministic perturbation of the structured random matrix $\mathbf{W}_N^{(L)}$. We first characterize in section 2.2 the largest eigenvalues and corresponding eigenvectors of $\boldsymbol{\Sigma}_N^{(L)}\boldsymbol{\Sigma}_N^{(L)*} = \hat{\mathbf{R}}_{y,N}^{(L)}$, and establish that they behave as if block line Hankel random matrix $\mathbf{W}_N^{(L)}$ was i.i.d. In section 2.3, we study the problem of detecting the wideband signal $[\mathbf{h}(z)]s_n$ and in section 2.4 we propose a new approach to select the regularization parameter of estimated spatio-temporal Wiener filters when a lenght N training sequence $(s_n)_{n=1,\ldots,N}$ is available at the receiver side. Finally, we study in section 2.5 the performance of spatial smoothing schemes for the subspace estimation of the directions of arrival of K narrow band sources impinging on a uniform linear array of sensors in the case where the number N of snapshots is large, but much smaller than the number of antennas M.

2.2 Behaviour of the largest eigenvalues and corresponding eigenvectors of $\boldsymbol{\Sigma}_N^{(L)}\boldsymbol{\Sigma}_N^{(L)*}$.

In this section, we establish that the largest eigenvalues and corresponding eigenvectors of $\boldsymbol{\Sigma}_N^{(L)}\boldsymbol{\Sigma}_N^{(L)*}$ behave as if the entries of noise matrix $\mathbf{W}_N^{(L)}$ were i.i.d. (see Theorem 1.2.2). We first define some notations. We denote by d_N and c_N the ratios $d_N = \frac{M}{N}$ and $c_N = \frac{ML}{N}$. It is assumed that $d_N \to d$ so that $c_N \to c$ where $c = dL$. $(\hat{\lambda}_{k,N}^{(L)})_{k=1,\ldots,ML}$ represent the eigenvalues of $\boldsymbol{\Sigma}^{(L)}\boldsymbol{\Sigma}^{(L)*}$ arranged in the decreasing order and $(\hat{\mathbf{u}}_k)_{k=1,\ldots,ML}$ are the corresponding eigenvectors. $\lambda_{1,N}^{(L)} \geq \lambda_{2,N}^{(L)} \ldots \geq \lambda_{P+L-1,N}^{(L)}$ are the non zero eigenvalues of $\mathbf{B}_N^{(L)}\mathbf{B}_N^{(L)*}$ and $(\mathbf{u}_k)_{k=1,\ldots,P+L-1}$ are the corresponding eigenvectors.

We first remark that matrix $\mathbf{W}_N^{(L)}$ verifies Theorem 1.3.1. Therefore, the empirical eigenvalue distribution of $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ converges almost surely towards the Marcenko-Pastur distribution μ_{dL,σ^2} while almost surely, for N large enough, all the non zero eigenvalues $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ are localized in a neighbourhood of $[\lambda_L^-, \lambda_L^+]$ where $\lambda_L^- = \sigma^2(1-\sqrt{dL})^2$ and $\lambda_L^+ = \sigma^2(1+\sqrt{dL})^2$.

In this section, we formulate the following assumption:

Assumption 2.2.1. *The* $P+L-1$ *non zero eigenvalues* $(\lambda_{k,N}^{(L)})_{k=1,\ldots,K}$ *of matrix* $\mathbf{B}_N^{(L)}\mathbf{B}_N^{(L)*}$ *converge towards* $\lambda_1^{(L)} > \lambda_2^{(L)} > \ldots > \lambda_{P+L-1}^{(L)}$ *when* $N \to +\infty$.

Then, the following Theorem holds.

Theorem 2.2.1. *We denote by* K_L, $0 \leq K_L \leq P+L-1$, *the largest integer for which* $\lambda_{K_L}^{(L)} > \sigma^2\sqrt{dL}$. *Then, for* $k = 1,\ldots,K_L$, *it holds that*

$$\hat{\lambda}_{k,N}^{(L)} \xrightarrow[N\to\infty]{a.s.} \rho_k^{(L)} = \phi_{dL,\sigma^2}(\lambda_k^{(L)}) = \frac{(\lambda_k^{(L)}+\sigma^2)(\lambda_k^{(L)}+\sigma^2 dL)}{\lambda_k^{(L)}} > \lambda_L^+ = \sigma^2(1+\sqrt{dL})^2$$

while for $k = K_L+1,\ldots,P+L-1$, $\hat{\lambda}_{k,N}^{(L)} \to \lambda_L^+$ *a.s.*

Finally, for all deterministic sequences of ML*–dimensional unit vectors* $(\mathbf{a}_N)$, $(\mathbf{b}_N)$, *for* $k = 1, \dots, K_L$, *it holds that*

$$\mathbf{a}_N^* \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \mathbf{b}_N = \frac{(\lambda_k^{(L)})^2 - \sigma^4 dL}{\lambda_k^{(L)}(\lambda_k^{(L)} + \sigma^2 dL)} \mathbf{a}_N^* \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \mathbf{b}_N + o(1) \quad a.s. \tag{2.2.1}$$

Proof. In order to simplify the notations, we will denote matrix $\mathbf{W}_N^{(L)}$ by $\mathbf{W}_N$ in the course of the proof. The resolvents of matrices $\mathbf{W}_N^{(L)} \mathbf{W}_N^{(L)*}$ and $\mathbf{W}_N^{(L)*} \mathbf{W}_N^{(L)}$ will also be denoted $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$ respectively. Finally, the Stieltjes transforms m_{c,σ^2} and $\tilde{m}_{c,\sigma^2}$ will be denoted by $m_c(z)$ and $\tilde{m}_c(z)$.

As matrix $\mathbf{W}_N$ verifies Theorem 1.3.1, Remark 1 implies that it is sufficient to check the following properties:

- (i) For each deterministic vectors $\mathbf{a}_N, \mathbf{b}_N$ for which $\sup_N(\|\mathbf{a}_N\|, \mathbf{b}_N\|) < +\infty$,

$$\mathbf{a}_N^* \left(\mathbf{Q}_N(z) - m_c(z)\mathbf{I}_M\right) \mathbf{b}_N \to 0 \; a.s. \tag{2.2.2}$$

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

- (ii) For each deterministic vectors $\tilde{\mathbf{a}}_N, \tilde{\mathbf{b}}_N$ for which $\sup_N(\|\tilde{\mathbf{a}}_N\|, \tilde{\mathbf{b}}_N\|) < +\infty$.

$$\tilde{\mathbf{a}}_N^* \left(\tilde{\mathbf{Q}}_N(z) - \tilde{m}_c(z)\mathbf{I}_N\right) \tilde{\mathbf{b}}_N \to 0 \; a.s. \tag{2.2.3}$$

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

- (iii) For each deterministic vectors $\mathbf{a}_N, \tilde{\mathbf{b}}_N$ such that $\sup_N(\|\mathbf{a}_N\|, \tilde{\mathbf{b}}_N\|) < +\infty$.

$$\mathbf{a}_N^* \left(\mathbf{Q}_N(z)\mathbf{W}_N\right) \tilde{\mathbf{b}}_N \to 0 \; a.s. \tag{2.2.4}$$

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

- (iv) Convergence properties (2.2.2, 2.2.3, 2.2.4) hold uniformly w.r.t. z on each compact subset of $\mathbb{C}^* - [\lambda_c^-, \lambda_c^+]$.

Properties (2.2.2, 2.2.3, 2.2.4) hold as soon as $\frac{ML}{N} \to c$, $c > 0$ and $M = \mathcal{O}(N^\beta)$ where $\beta > 0$, or equivalently $L = \mathcal{O}(N^\alpha)$, $\alpha < 1$. In particular, condition L finite or $L = \mathcal{O}(N^\alpha)$, $\alpha < 2/3$ are not needed. We thus use the minimal conditions in the proofs. Item (iv) however needs that $L = \mathcal{O}(N^\alpha)$, $\alpha < 2/3$ holds.

We first establish (2.2.2). This property is nearly proved in [49], and appears as a particular case of more general results in Chapter 3. However, to make the present proof reasonably self-contained, we provide in the following a sketch of proof synthesizing the various points developed in [49]. We first remark that

$$\mathbf{a}_N^* \left(\mathbf{Q}_N(z) - m_c(z)\mathbf{I}\right) \mathbf{b}_N = \mathbf{a}_N^* \left(\mathbf{Q}_N(z) - \mathbb{E}(\mathbf{Q}_N(z))\right) \mathbf{b}_N + \mathbf{a}_N^* \left(\mathbb{E}(\mathbf{Q}_N(z)) - m_c(z)\mathbf{I}\right) \mathbf{b}_N$$

and establish that the 2 terms at the right hand side of the above equation converge towards 0. In order to simplify the notations, we denote by ξ the first term. The almost sure convergence of ξ towards 0 follows from the Poincaré-Nash inequality (see e.g. Proposition 2 of [49] as well as Chapter 3 of the present Thesis). More precisely, ξ can be considered as a smooth function $\xi(\mathbf{W}_N, \overline{\mathbf{W}}_N)$ of the entries of $\mathbf{W}_N$ and of matrix $\overline{\mathbf{W}}_N$ whose entries are the complex conjugates of the entries of $\mathbf{W}_N$. For each $m = 1, \dots, M$, we denote by $\mathbf{W}_{i,j}^m$ the entry (i,j) of $L \times N$ matrix $\mathbf{W}_N^{(m)}$. Then, the Poincaré-Nash inequality is a concentration inequality which states that

$$\begin{aligned}\mathrm{Var}(\xi) \quad \leq \quad & \sum_{m,m'} \sum_{i,j,i',j'} \mathbb{E}\left[\frac{\partial \xi}{\partial \mathbf{W}_{i,j}^m} \mathbb{E}\left(\mathbf{W}_{i,j}^m (\mathbf{W}_{i',j'}^{m'})^*\right) \left(\frac{\partial \xi}{\partial \mathbf{W}_{i',j'}^{m'}}\right)^*\right] + \\ & \sum_{m,m'} \sum_{i,j,i',j'} \mathbb{E}\left[\left(\frac{\partial \xi}{\partial \overline{\mathbf{W}}_{i,j}^m}\right)^* \mathbb{E}\left(\mathbf{W}_{i,j}^m (\mathbf{W}_{i',j'}^{m'})^*\right) \left(\frac{\partial \xi}{\partial \overline{\mathbf{W}}_{i',j'}^{m'}}\right)\right]\end{aligned}$$

We notice that the structure of $\mathbf{W}_N$ implies that

$$\mathbb{E}\left(\mathbf{W}^m_{i,j}(\mathbf{W}^{m'}_{i',j'})^*\right) = \frac{\sigma^2}{N}\delta(m-m')\delta(i+j=i'+j') \tag{2.2.5}$$

so that the above sums reduce to simpler terms. The above upper bound of $\mathrm{Var}(\xi)$ was evaluated in Proposition 3-1 in [49] (see Eq. (3.2)): $\mathrm{Var}(\xi) = \mathbb{E}|\xi|^2 = \mathcal{O}(\frac{L}{N}) = \mathcal{O}(\frac{1}{M})$. As $M \to +\infty$, this implies that ξ converges in probability towards 0. In order to prove the almost sure convergence, we briefly justify that for each k, it holds that

$$\mathbb{E}|\xi|^{2k} = \mathcal{O}\left((1/M)^k\right) \tag{2.2.6}$$

(2.2.6) is established by induction on k. As mentioned above, (2.2.6) is verified for $k=1$. We now assume that it holds until integer $k-1$, and prove (2.2.6). For this, we use the obvious relation:

$$\mathbb{E}|\xi|^{2k} = \left(\mathbb{E}|\xi|^k\right)^2 + \mathrm{Var}(\xi^k)$$

In order to manage $\mathrm{Var}(\xi^k)$, we use again the Poincaré-Nash inequality. As

$$\frac{\partial \xi^k}{\partial \mathbf{W}^m_{i,j}} = k\xi^{k-1}\frac{\partial \xi}{\partial \mathbf{W}^m_{i,j}}$$

the Poincaré-Nash inequality leads to

$$\begin{aligned}\mathrm{Var}(\xi^k) \quad \le \quad & k^2 \sum_{m,m'}\sum_{i,j,i',j'} \mathbb{E}\left[|\xi|^{2k-2}\frac{\partial \xi}{\partial \mathbf{W}^m_{i,j}}\mathbb{E}\left(\mathbf{W}^m_{i,j}(\mathbf{W}^{m'}_{i',j'})^*\right)\left(\frac{\partial \xi}{\partial \mathbf{W}^{m'}_{i',j'}}\right)^*\right] + \\ & k^2 \sum_{m,m'}\sum_{i,j,i',j'} \mathbb{E}\left[|\xi|^{2k-2}\left(\frac{\partial \xi}{\partial \overline{\mathbf{W}}^m_{i,j}}\right)^*\mathbb{E}\left(\mathbf{W}^m_{i,j}(\mathbf{W}^{m'}_{i',j'})^*\right)\left(\frac{\partial \xi}{\partial \overline{\mathbf{W}}^{m'}_{i',j'}}\right)\right]\end{aligned}$$

Following the proof of Proposition 3-1 in [49], it is easy to check that the Poincaré-Nash inequality leads to

$$\mathrm{Var}(\xi^k) \le C\frac{L}{N}\mathbb{E}(|\xi|^{2k-2})$$

where C is a constant that depends on z but not on the dimensions L, M, N. As (2.2.6) is assumed to hold until integer $k-1$, this implies that $\mathrm{Var}(\xi^k) = \mathcal{O}\left((L/N)^k\right)$. The Schwartz inequality leads immediately to

$$\left(\mathbb{E}|\xi|^k\right)^2 \le \mathbb{E}(|\xi|^2)\,\mathbb{E}(|\xi|^{2k-2})$$

which is a $\mathcal{O}\left((L/N)^k\right)$ term. This establishes (2.2.6). As $M = \mathcal{O}(N^\beta)$ with $\beta > 0$, it is clear that it exists k_0 for which $\frac{1}{M^{k_0}} < \frac{1}{N^{1+\delta}}$ for some $\delta > 0$. Therefore, (2.2.6) for $k = k_0$ leads to

$$\mathbb{E}\left(|\xi|^{2k_0}\right) = \mathcal{O}\left(\frac{1}{N^{1+\delta}}\right)$$

The use of the Markov inequality and of the Borel-Cantelli lemma imply that ξ converges towards 0 almost surely as expected.

It remains to justify that

$$\mathbf{a}^*_N\left(\mathbb{E}(\mathbf{Q}_N(z)) - m_c(z)\mathbf{I}\right)\mathbf{b}_N \to 0 \tag{2.2.7}$$

For this, we first simplify the notations and denote by $\mathbf{W}, \mathbf{Q}, \tilde{\mathbf{Q}}$ the matrices $\mathbf{W}_N, \mathbf{Q}_N(z), \tilde{\mathbf{Q}}_N(z)$. Moreover, $\mathbf{Q}$ is a $ML \times ML$ block matrix, so that we denote by $\mathbf{Q}^{m_1,m_2}_{i_1,i_2}$ its entry $(i_1 + (m_1-1)L, i_2 + (m_2-1)L)$. We also denote $(\mathbf{w}_j)_{j=1,\dots,N}$ the columns of $\mathbf{W}$. Although it is not stated explicitly in [49], (2.2.7) can be deduced from various intermediate evaluations. In order to be more specific, we mention that it is proved in [49] that matrix $\mathbb{E}(\mathbf{Q})$ can be written as

$$\mathbb{E}(\mathbf{Q}(z)) = \mathbf{I}_M \otimes \mathbf{R}_N(z) + \mathbf{\Delta}_N(z) \tag{2.2.8}$$

(see Eq. (4.14) in [49]) where $\mathbf{R}_N(z)$ is a $L \times L$ matrix whose expression is omitted, and where $\mathbf{\Delta}_N(z)$ is shown to verify $\mathbf{a}_N^* \mathbf{\Delta}_N(z)\mathbf{b}_N \to 0$ using the Poincaré-Nash inequality (see Eq. (5.3) in [49]). As it will be useful to establish (2.2.3), we give some insights on the proof of (2.2.8). [49] uses the identity

$$\mathbb{E}\left[\mathbf{Q}_{i_1,i_2}^{m_1,m_2}\right] = -\frac{1}{z}\delta(i_1-i_2)\delta(m_1-m_2) + \frac{1}{z}\mathbb{E}\left[\left(\mathbf{QWW}^*\right)_{i_1,i_2}^{m_1,m_2}\right] \tag{2.2.9}$$

It turns out that the second term of the righthandside of (2.2.9) can be expressed in terms of the entries of $\mathbb{E}(\mathbf{Q})$. To obtain the corresponding expression, [49] evaluates $\mathbb{E}\left[\left(\mathbf{Qw}_k\mathbf{w}_j^*\right)_{i_1,i_2}^{m_1,m_2}\right] = \mathbb{E}\left[(\mathbf{Qw}_k)_{i_1}^{m_1}\left(\mathbf{w}_j^*\right)_{i_2}^{m_2}\right]$ for each k, j, i_1, i_2, m_1, m_2. For this, the identity

$$\mathbb{E}\left[(\mathbf{Qw}_k)_{i_1}^{m_1}\left(\mathbf{w}_j^*\right)_{i_2}^{m_2}\right] = \sum_{i_3,m_3}\mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3}\mathbf{W}_{i_3,k}^{m_3}\overline{\mathbf{W}}_{i_2,j}^{m_2}\right)$$

and the so-called the integration by parts formula

$$\mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3}\mathbf{W}_{i_3,k}^{m_3}\overline{\mathbf{W}}_{i_2,j}^{m_2}\right) = \sum_{i',j'}\mathbb{E}\left(\mathbf{W}_{i_3,k}^{m_3}\overline{\mathbf{W}}_{i',j'}^{m_3}\right)\mathbb{E}\left[\frac{\partial\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3}\overline{\mathbf{W}}_{i_2,j}^{m_2}\right)}{\partial\overline{\mathbf{W}}_{i',j'}^{m_3}}\right]$$

are used. After some calculations, this allows to express

$$\mathbb{E}\left[\left(\mathbf{QWW}^*\right)_{i_1,i_2}^{m_1,m_2}\right] = \sum_{j=1}^{N}\left(\mathbf{Qw}_j\mathbf{w}_j^*\right)_{i_1,i_2}^{m_1,m_2}$$

in terms of the entries of $\mathbb{E}(\mathbf{Q})$, and to plug the corresponding expression into (2.2.9). This, in turn, leads to (2.2.8).

In order to complete the proof of (2.2.7), it remains to justify that

$$\mathbf{a}_N^*\left(\mathbf{I}_M \otimes \mathbf{R}_N(z) - m_c(z)\mathbf{I}\right)\mathbf{b}_N \to 0$$

or equivalently that

$$\mathbf{a}_N^*\left(\mathbf{I}_M \otimes \mathbf{R}_N(z) - m_{c_N}(z)\mathbf{I}\right)\mathbf{b}_N \to 0 \tag{2.2.10}$$

where $m_{c_N}(z)$ is the Stieltjes transform of the Marcenko-Pastur distribution of parameters (c_N, σ^2), which, of course, verifies $m_{c_N}(z) - m_c(z) \to 0$ because $c_N \to c$. The reader may check that (2.2.10) follows from Corollary 5.1, Theorem 7.1 and Eq. (7.3) in [49].

Sketch of proof of (2.2.3). Using the Poincaré-Nash inequality, it can be proved as above that

$$\tilde{\mathbf{a}}_N^*\left(\tilde{\mathbf{Q}} - \mathbb{E}(\tilde{\mathbf{Q}})\right)\tilde{\mathbf{b}}_N \to 0 \; a.s.$$

and establish that

$$\tilde{\mathbf{a}}_N^*\left(\mathbb{E}(\tilde{\mathbf{Q}}(z)) - \tilde{m}_c(z)\mathbf{I}\right)\tilde{\mathbf{b}}_N \to 0 \tag{2.2.11}$$

for each $z \in \mathbb{C} - \mathbb{R}^+$. The behaviour of matrix $\mathbb{E}(\tilde{\mathbf{Q}})$ is not studied in [49]. However, it can be evaluated using the results of [49]. We first briefly justify that

$$\mathbb{E}(\tilde{\mathbf{Q}}) = \tilde{\mathbf{R}} + \tilde{\mathbf{\Delta}} \tag{2.2.12}$$

where $\tilde{\mathbf{R}}$ is a certain $N \times N$ matrix, and where $\tilde{\boldsymbol{\Delta}}$ verifies $\tilde{\mathbf{a}}_N^* \tilde{\boldsymbol{\Delta}} \tilde{\mathbf{b}}_N \to 0$. The proof of (2.2.12) uses the same ingredients than the proof of (2.2.8). We first remark that

$$\mathbf{W}^* \mathbf{Q} \mathbf{W} = \tilde{\mathbf{Q}} \mathbf{W}^* \mathbf{W} = \mathbf{I} + z\tilde{\mathbf{Q}} \tag{2.2.13}$$

The above mentioned evaluation of $\mathbb{E}\left[(\mathbf{Q}\mathbf{w}_k)_{i_1}^{m_1} \left(\mathbf{w}_j^*\right)_{i_2}^{m_2} \right]$ for each k, j, i_1, i_2, m_1, m_2 allows to calculate $\mathbb{E}\left((\mathbf{W}^*\mathbf{Q}\mathbf{W})_{j,k}\right)$ in terms of the entries of $\mathbb{E}(\mathbf{Q})$ and of $\mathbb{E}(\tilde{\mathbf{Q}})$. Plugging this relation as well as (2.2.8) into (2.2.13) leads to the expression (2.2.12). As previously, $\tilde{\mathbf{a}}_N^* \tilde{\boldsymbol{\Delta}} \tilde{\mathbf{b}}_N \to 0$ is obtained using the Poincaré-Nash inequality.

The proof of $\tilde{\mathbf{a}}_N^* \left(\tilde{\mathbf{R}}(z) - \tilde{m}_c(z)\mathbf{I}\right) \tilde{\mathbf{b}}_N \to 0$ is omitted as it needs the introduction of several notations of [49].

We establish (2.2.4). For this, we first remark that for each $\theta \in \mathbb{R}$, the distribution of matrix $\mathbf{W}_N e^{i\theta}$ coincides with the distribution of $\mathbf{W}_N$. Therefore, it holds that

$$\mathbb{E}\left(\mathbf{Q}_N(z)\mathbf{W}_N e^{i\theta}\right) = \mathbb{E}\left(\mathbf{Q}_N(z)\mathbf{W}_N\right)$$

which implies that $\mathbb{E}\left(\mathbf{Q}_N(z)\mathbf{W}_N\right) = 0$. In order to complete the proof of (1.2.19), it is sufficient to establish that if we denote by κ_N the random variable $\kappa_N = \mathbf{a}_N^* \left(\mathbf{Q}_N(z)\mathbf{W}_N\right) \tilde{\mathbf{b}}_N$, then, for each $p \geq 1$, it holds that

$$\mathbb{E}\left|\kappa_N - \mathbb{E}(\kappa_N)\right|^{2p} = \mathcal{O}\left(\left(\frac{L}{N}\right)^p\right) \tag{2.2.14}$$

Choosing p large enough leads to $\kappa_N - \mathbb{E}(\kappa_N) = \kappa_N \to 0$ *a.s.* as expected. (2.2.14) can be proved as above by using the Poincaré-Nash inequality.

We finally justify that, (2.2.2), (2.2.3), (2.2.4) hold uniformly w.r.t. z on each compact subset of $\mathbb{C}^* - [\lambda_c^-, \lambda_c^+]$ when $L = \mathcal{O}(N^\alpha)$ for $\alpha < 2/3$. We just prove that it is the case for (2.2.4). By Theorem 1.3.1, for each $\epsilon > 0$, almost surely, for N large enough, there is no eigenvalue of $\mathbf{W}_N\mathbf{W}_N^*$ outside $[\lambda_c^- - \epsilon, \lambda_c^+ + \epsilon] \cup \{0\}$. Therefore, almost surely, for N large enough, function $z \to \kappa_N(z)$ is analytic on $\mathbb{C}^* - [\lambda_c^- - \epsilon, \lambda_c^+ + \epsilon]$. We use a standard argument based on Montel's theorem ([64], p.282). We first justify that for each compact subset $\mathcal{K} \subset \mathbb{C}^* - [\lambda_c^- - \epsilon, \lambda_c^+ + \epsilon]$, then it exists a constant η such that

$$\sup_{z \in \mathcal{K}} |\kappa_N(z)| \leq \eta \tag{2.2.15}$$

for each N large enough. We consider the singular value decomposition of matrix $\mathbf{W}_N$:

$$\mathbf{W}_N = \boldsymbol{\Gamma}_N \boldsymbol{\Delta}_N \boldsymbol{\Theta}_N^*$$

where $\boldsymbol{\Delta}_N$ represents the diagonal matrix of non zero singular values of $\mathbf{W}_N$. $\kappa_N(z)$ can be written as

$$\kappa_N(z) = \mathbf{a}_N^* \boldsymbol{\Gamma}_N \left(\boldsymbol{\Delta}_N^2 - z\mathbf{I}\right)^{-1} \boldsymbol{\Delta}_N \boldsymbol{\Theta}_N^* \tilde{\mathbf{b}}_N$$

Therefore, it holds that

$$|\kappa_N(z)| \leq \left\| \left(\boldsymbol{\Delta}_N^2 - z\mathbf{I}\right)^{-1} \boldsymbol{\Delta}_N \right\| \|\mathbf{a}_N\| \|\tilde{\mathbf{b}}_N\|$$

The entries of $\boldsymbol{\Delta}_N^2$ are located almost surely into $[\lambda_c^- - \epsilon, \lambda_c^+ + \epsilon] \cup \{0\}$ for each N large enough. Therefore, for each $z \in \mathcal{K}$, it holds that

$$\left\| \left(\boldsymbol{\Delta}_N^2 - z\mathbf{I}\right)^{-1} \boldsymbol{\Delta}_N \right\| \leq \frac{1}{\operatorname{dist}([\lambda_c^- - \epsilon, \lambda_c^+ + \epsilon] \cup \{0\}, \mathcal{K})}$$

The conclusion follows from the hypothesis that vectors $\mathbf{a}_N$ and $\tilde{\mathbf{b}}_N$ satisfy $\sup_N(\|\mathbf{a}_N\|, \|\tilde{\mathbf{b}}_N\|) < +\infty$. (2.2.15) implies that the sequence of analytic functions $(\kappa_N)_{N\geq 1}$ is a normal family . Therefore, it exists a subsequence extracted from $(\kappa_N)_{N\geq 1}$ that converges uniformly on each compact subset of $\mathbb{C}^* - [\lambda_c^- - \epsilon, \lambda_c^+ + \epsilon]$ towards a certain analytic function κ_*. As (2.2.4) holds for each $z \in \mathbb{C} - \mathbb{R}^+$, function κ_* is identically zero. We have thus shown that each converging subsequence extracted from $(\kappa_N)_{N\geq 1}$ converges uniformly towards 0 on each compact subset of $\mathbb{C}^* - [\lambda_c^- - \epsilon, \lambda_c^+ + \epsilon]$. This, in turn, shows that the whole sequence converges uniformly on each compact subset of $\mathbb{C}^* - [\lambda_c^- - \epsilon, \lambda_c^+ + \epsilon]$. As this property holds for each $\epsilon > 0$, the sequence also converges uniformly on each compact subset of $\mathbb{C}^* - [\lambda_c^-, \lambda_c^+]$ as expected.

Remark 3. *Remark 1 leads to the conclusion that convergence properties (2.2.2, 2.2.3, 2.2.4) and their uniformity w.r.t. z on each compact subset of $\mathbb{C}^* - [\lambda_c^-, \lambda_c^+]$ implies that the bilinear forms of matrices $(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^* - z\mathbf{I})^{-1}$ and $(\boldsymbol{\Sigma}_N^*\boldsymbol{\Sigma}_N - z\mathbf{I})^{-1}$ have the same behaviour than the bilinear forms of matrices $\mathbf{S}_N(z)$ and $\tilde{\mathbf{S}}_N(z)$ defined by (1.2.33) and (1.2.37).*

2.3 Detection of a wideband signal.

2.3.1 Introduction

The multi-antenna detection of low rank non observable narrow band signals corrupted by an additive spatially and temporally white Gaussian noise is a fundamental problem that was studied extensively e.g. in the contexts of array processing (see e.g. [45], [11]) and more recently of spectrum sensing (see among others [65], [47], [25]). The most popular method to solve the above problem is the GLRT test, which, in the present case, can be expressed in closed form. In order to obtain some insights on the statistical performance of the generalized likelihood ratio test (GLRT), it is standard to assume that the number of observations N converge towards ∞, and to characterize the asymptotic distribution of the GLRT statistics under the null and the alternative hypothesis. In practice, this approach provides reasonable results when N is much larger than the number of antennas M at the receiver side. When the antenna array is large, the assumption that $N >> M$ is often not justified and the standard asymptotic analysis does not provide reliable results (see e.g. [36] in the context of supervised detection). In this context, it is now standard to consider the large system regime $M \to +\infty$, $N \to +\infty$ in such a way that $\frac{M}{N} \to d$ where $d > 0$. We refer the reader to the papers [57], [43], [12] in which this approach is developed.

In this section, we assume that M and N are large and of the same order of magnitude. We address the detection of a single signal in a multipath propagation channel, i.e. its contribution to the observation coincides with the output of an unknown finite impulse response SIMO filter driven by an unobservable deterministic scalar sequence $s = (s_n)_{n\in\mathbb{Z}}$. The signal to be detected is thus a rank 1 signal in the frequency domain. We assume moreover that the number of paths P, or equivalently the number of coefficients of the SIMO filter is much smaller than M. [65] studied the GLRT when s is an i.i.d. Gaussian sequence and the filter has an infinite impulse response, or equivalently when $P = +\infty$. Under certain assumptions, [65] proposed to evaluate the log likelihood ratio using the Witthle approximation, and obtained an expression based on integrals over the frequency domain. When P is finite, the GLRT cannot be expressed in closed form because the maximization of the likelihood over sequence s and the filter coefficients $(\mathbf{h}_p)_{p=0,\ldots,P-1}$ has no explicit solution.

As the GLRT cannot be used, a pragmatic approach is to observe that the signal to be detected can be interpreted as a superposition of P narrow band deterministic signals. Therefore, it is possible to use the corresponding GLRT which consists in comparing the sum of the P greatest eigenvalues of the empirical spatial covariance matrix of the observation to a threshold, at least if the noise variance is known. However, it is intuitively more appealing to consider the greatest eigenvalues of the empirical spatio-temporal covariance matrix in order to take benefit of the particular convolutive structure of the signal to be detected. We compare these 2 approaches in the asymptotic regime $M \to +\infty$, $N \to +\infty$ in such a way that $\frac{M}{N} \to d$ where $d > 0$. In this regime, the first order behaviour of the largest eigenvalues of the empirical spatial covariance matrix is well known, and this allows to evaluate the relevance of the "narrow band" test. In this section, we use Theorem 2.2.1 in order to evaluate the behaviour of the greatest eigenvalues of the empirical spatio-temporal covariance matrix. This allows to have a clear understanding of the advantages of the use of the spatio-temporal covariance matrix.

This section is organized as follows. In subsection 2.3.2, we precise the signal models and the underlying assumptions. In subsection 2.3.3, we deduce from Theorem 2.2.1 the first order behaviour of the detection test based on the greatest eigenvalues of the empirical spatio-temporal covariance matrix. Finally, subsection 2.3.4 present numerical experiments sustaining our theoretical results.

In the following, $\mathcal{N}_c(\mathbf{x}, \mathbf{\Gamma})$ represents the M-variate complex Gaussian (i.e. circular) distribution with mean $\mathbf{x}$ and covariance matrix $\mathbf{\Gamma}$.

2.3.2 Problem formulation.

In the following, we denote by $(\mathbf{y}_n)_{n=1,\dots,N}$ the M-dimensional signal received on the M-sensors array. Under hypothesis H_0, the observation is reduced to a spatially and temporally complex Gaussian noise, i.e.

$$\mathbf{y}_n = \mathbf{v}_n, \; n = 1, \dots, N \tag{2.3.1}$$

where $(\mathbf{v}_n)_{n=1,\dots,N}$ are i.i.d. $\mathcal{N}_c(0, \sigma^2\mathbf{I})$ distributed random vectors. We assume from now on that σ^2 is known in order to simplify the exposition, but our results can be easily generalized if σ^2 is unknown (see below). Under hypothesis H_1, the observation is given by

$$\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p} + \mathbf{v}_n, \; n = 1, \dots, N \tag{2.3.2}$$

where $(s_n)_{n\in\mathbb{Z}}$ is a non observable deterministic scalar sequence and where the $M \times 1$ transfer function $\mathbf{h}(z) = \sum_{p=0}^{P-1} \mathbf{h}_p z^{-p}$ is unknown. We however assume that P is known, which, in practice, means that an upper bound of the number of paths is available. s is assumed deterministic in order to avoid to formulate restrictive hypotheses,e.g. that signal $(s_n)_{n\in\mathbb{Z}}$ is an i.i.d. Gaussian sequence.

In order to test hypothesis H_0 versus H_1, the GLRT cannot be implemented because, under H_1, the maximum likelihood estimator of filter $\mathbf{h}(z)$ and sequence $(s_n)_{n=-(P-1),\dots,N}$ cannot be expressed in closed form (see e.g. [66], [38]). We note that when s is an i.i.d. Gaussian sequence, [65] derived an approximate GLRT based on the Whittle approximation, but without assuming that filter $\mathbf{h}(z)$ is FIR. Moreover, the approach of [65] needs the observation of at least M independent realizations of the observation $(\mathbf{y}_n)_{n=1,\dots,N}$, an hypothesis which is not formulated in the present paper. Finally, the approach of [65] cannot be adapted to the case of a deterministic signal $(s_n)_{n\in\mathbb{Z}}$.

As the GLRT cannot be implemented, we study pragmatic alternative approaches. The most obvious solution is based on the observation that signal $[\mathbf{h}(z)]s(n) = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p}$ can be interpreted as a superposition of P narrow band signals. It is thus possible to test the hypothesis H_0 against hypothesis $H_1^{'}$ defined by

$$\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_n^{(p)} + \mathbf{v}_n \tag{2.3.3}$$

where signals $(s^{(p)})_{p=0,\dots,P-1}$ are non observable deterministic signals. Hypothesis $H_1^{'}$ is of course not equivalent to H_1 because the particular structure of $s_n^{(p)} = s_{n-p}$ is ignored in the formulation of $H_1^{'}$. We denote by $\mathbf{Y}$ the $M \times N$ matrix defined by $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_N)$. Then, $H_1^{'}$ is of course equivalent to

$$\mathbf{Y} = \mathbf{A} + \mathbf{V} \tag{2.3.4}$$

where $\mathbf{A}$ is a rank P deterministic matrix, and where $\mathbf{V}$ is defined as $\mathbf{Y}$. The corresponding GLRT is easy to derive, and consists in comparing the statistics

$$\eta_N = \sum_{i=1}^{P} \lambda_i \left(\frac{\mathbf{Y}\mathbf{Y}^*}{N} \right) \tag{2.3.5}$$

to a threshold. Here, $\left(\lambda_i\left(\frac{\mathbf{YY}^*}{N}\right)\right)_{i=1,\dots,M}$ represent the eigenvalues of $\frac{\mathbf{YY}^*}{N}$ arranged in decreasing order.

The matrix $\frac{\mathbf{YY}^*}{N}$ coincides with the empirical spatial covariance matrix of the observations. In order to take benefit of the particular convolutive structure of signal $[\mathbf{h}(z)]s_n$, it seems however more appropriate to consider a statistics based on the largest eigenvalues of spatio-temporal covariance matrices. If L is an integer, we denote by $\mathbf{y}_n^{(L)}$ the ML–dimensional vector defined by

$$\mathbf{y}_n^{(L)} = (\mathbf{y}_{1,n}, \dots, \mathbf{y}_{1,n+L-1}, \dots, \mathbf{y}_{M,n}, \dots, \mathbf{y}_{M,n+L-1})^T$$

and by $\mathbf{Y}^{(L)}$ the $ML \times N$ block-Hankel matrix defined by $\mathbf{Y}^{(L)} = (\mathbf{y}_1^{(L)}, \dots, \mathbf{y}_N^{(L)})$. We note that $\mathbf{Y}^{(L)}$ depends on $(\mathbf{y}_n)_{n=1,\dots,N+L-1}$ while, in principle, the observation $\mathbf{y}_n$ is available until $n = N$. As we consider in the following asymptotic regimes in which $N \to +\infty$ while L remains fixed, the above mentioned end effect has no consequence on our results.

Under hypothesis H_0, matrix $\mathbf{Y}^{(L)}$ is reduced to $\mathbf{V}^{(L)}$, and under H_1, $\mathbf{Y}^{(L)}$ is given by

$$\mathbf{Y}^{(L)} = \mathbf{H}^{(L)}\mathbf{S}^{(L)} + \mathbf{V}^{(L)} \tag{2.3.6}$$

(see (2.1.3)). Instead of using η_N defined by (2.3.5), we propose to consider the statistics $\eta_N^{(L)}$ given by

$$\eta_N^{(L)} = \sum_{i=1}^{P+L-1} \lambda_i \left(\frac{\mathbf{Y}^{(L)}\mathbf{Y}^{(L)*}}{N} \right) \tag{2.3.7}$$

for a suitable value of integer L. We note that if $L = 1$, then $\mathbf{Y}^{(1)}$ and $\eta_N^{(1)}$ coincide with $\mathbf{Y}$ and η_N.

In order to obtain some insights on the merits of statistics $\eta_N^{(L)}$ in the case where M and N are large and of the same order of magnitude, we evaluate under both hypotheses the first order behaviour of $\eta_N^{(L)}$ in the asymptotic regime $M \to +\infty$, $N \to +\infty$ in such a way that $d_N = \frac{M}{N} \to d$ where $d > 0$. We also assume that P and L do not scale with M, N. In the following, $N \to +\infty$ should be understood as the above asymptotic regime. The study of $\eta_N^{(L)}$ when $N \to +\infty$ is equivalent to the study of the largest eigenvalues of matrix $\frac{\mathbf{Y}^{(L)}\mathbf{Y}^{(L)*}}{N}$.

Remark 4. The case σ^2 unknown. *When σ^2 is unknown, the GLRT corresponding to hypotheses H_0 and $H_1^{'}$ given by (2.3.4) consists in comparing statistics $\frac{\eta_N}{\frac{1}{M}\mathrm{Tr}(\mathbf{YY}^*/N)}$ to a threshold. Therefore, it is relevant to replace statistics $\eta_N^{(L)}$ by $\frac{\eta_N^{(L)}}{\frac{1}{ML}\mathrm{Tr}(\mathbf{Y}^{(L)}\mathbf{Y}^{(L)*}/N)}$. It is easily seen that $\frac{1}{ML}\mathrm{Tr}(\mathbf{Y}^{(L)}\mathbf{Y}^{(L)*}/N)$ converges almost surely towards σ^2 in the absence and in the presence of signal. Therefore, the characterization of the first order asymptotic behaviours of $\eta_N^{(L)}$ and of its normalized version are equivalent.*

2.3.3 Asymptotic behaviour of $\eta_N^{(L)}$.

$\mathbf{W}_N^{(L)}$, $\mathbf{B}_N^{(L)}$ and $\mathbf{\Sigma}_N^{(L)}$ are the normalized matrices defined by $\mathbf{W}_N^{(L)} = \mathbf{V}_N^{(L)}/\sqrt{N}$, $\mathbf{B}_N^{(L)} = \frac{1}{\sqrt{N}}\mathbf{H}^{(L)}\mathbf{S}^{(L)}$ and $\mathbf{\Sigma}_N^{(L)} = \mathbf{B}_N^{(L)} + \mathbf{W}_N^{(L)}$. We notice that $\mathbf{Y}_N^{(L)}/\sqrt{N}$ coincides with $\mathbf{W}_N^{(L)}$ under H_0 and with $\mathbf{\Sigma}_N^{(L)}$ under H_1. The content of this chapter is based on Theorem 2.2.1 which, allows to evaluate the asymptotic behaviour of the largest eigenvalues of $\mathbf{\Sigma}_N^{(L)}\mathbf{\Sigma}_N^{(L)*}$.

In the following, we denote by $(\hat{\lambda}_{k,N}^{(L)})_{k=1,\dots,ML}$ the eigenvalues of $\mathbf{\Sigma}_N^{(L)}\mathbf{\Sigma}_N^{(L)*}$, and by $\lambda_{1,N}^{(L)} \geq \lambda_{2,N}^{(L)} \dots \geq \lambda_{P+L-1,N}^{(L)}$ the non zero eigenvalues of $\mathbf{B}_N^{(L)}\mathbf{B}_N^{(L)*}$.

In order to simplify the following discussion, we formulate the following hypotheses on vectors $(\mathbf{h}_p)_{p=0,\dots,P-1}$ and on signal $(s_n)_{n\in\mathbb{Z}}$:

Assumption 2.3.1.

- *(i) When* $N \to +\infty$, *matrix* $\mathbf{H}^*\mathbf{H}$ *converges towards a* $P\times P$ *matrix* $\mathbf{\Delta}$
- *(ii) For each integers* $i, j \geq 1$, $\frac{1}{N}\sum_{n=1}^{N} s_{n+i-P}s_{n+j-P}^*$ *converges towards a limit. In this case, the limit only depends on* $i-j$, *and is denoted* R_{i-j}.

As the entries of matrix $\mathbf{H}^{(L)*}\mathbf{H}^{(L)}$ depend on the entries of $\mathbf{H}^*\mathbf{H}$, (i) implies that $\mathbf{H}^{(L)*}\mathbf{H}^{(L)}$ converges towards a matrix $\mathbf{\Delta}^{(L)}$ whose entries depend on the entries of $\mathbf{\Delta}$. In the following, we also denote by $\mathbf{R}^{(L)}$ the $(P+L-1)\times(P+L-1)$ Toeplitz matrix defined by $\mathbf{R}_{i,j}^{(L)} = R_{i-j}$.

As the non zero eigenvalues of $\mathbf{B}_N^{(L)}\mathbf{B}_N^{(L)*}$ coincide with the eigenvalues of matrix $\mathbf{H}^{(L)*}\mathbf{H}^{(L)}\frac{\mathbf{S}^{(L)}\mathbf{S}^{(L)*}}{N}$, it is clear that Assumption 2.3.1 implies that Assumption 2.2.1 holds, and that $\lambda_k^{(L)} = \lambda_k\left(\mathbf{\Delta}^{(L)}\mathbf{R}^{(L)}\right)$.

As Assumption 2.2.1 holds, Theorem 2.2.1 holds. It is clear that $\eta_N^{(L)}$ converges almost surely $\eta^{(L)}$ defined by

$$\eta^{(L)} = \sum_{k=1}^{K_L}\rho_k^{(L)} + (P+L-1-K_L)\sigma^2(1+\sqrt{dL})^2) \tag{2.3.8}$$

We note that if $K_L = 0$, or equivalently, if the largest limit eigenvalue $\lambda_1^{(L)}$ of matrix $\mathbf{B}_N^{(L)}\mathbf{B}_N^{(L)*}$ is below the detectability threshold $\sigma^2\sqrt{dL}$, then the first order asymptotic behaviour of $\eta_N^{(L)}$ under hypotheses H_0 and H_1 coincide. In this case, the test based on $\eta_N^{(L)}$ is not consistent, in the sense that it does not allow to distinguish between the 2 hypotheses when $N \to +\infty$. If however $\lambda_1^{(L)}$ is greater than $\sigma^2\sqrt{dL}$, the asymptotic behaviours of $\eta_N^{(L)}$ under H_0 and H_1 do not coincide and the test is consistent. In other words, the test based on $\eta_N^{(L)}$ is consistent if and only if

$$\frac{\lambda_1^{(L)}}{\sigma^2\sqrt{dL}} > 1 \tag{2.3.9}$$

This condition implies that the value of L for which $\frac{\lambda_1^{(L)}}{\sqrt{dL}}$ is maximum can be considered as optimal from the consistency of the GLRT test point of view. In order to obtain some insights on the optimal choice of L, we first assume that $(s_n)_{n\in\mathbb{Z}}$ coincides with a realization of a unit variance zero mean i.i.d. sequence and that the limit $\mathbf{\Delta}$ of matrix $\mathbf{H}^*\mathbf{H}$ is diagonal, a condition meaning that the P paths are decorrelated. In order to simplify the notations, we denote by $\delta_0,\dots,\delta_{P-1}$ the diagonal entries of $\mathbf{\Delta}$ which represent the powers of the various paths. We notice that $\sum_{p=0}^{P-1}\delta_p$ coincides with the power of the signal to be detected. It is easily seen that for each L, matrix $\mathbf{\Delta}^{(L)}$ is diagonal as well, and that its largest entry is equal to $\sum_{p=0}^{P-1}\delta_p$ if $L \geq P$, and to $\max_{k=0,P-L}\sum_{p=0}^{L-1}\delta_{p+k}$ if $L \leq P$. As matrix $\mathbf{R}^{(L)}$ is equal to $\mathbf{I}_{P-L+1}$, this implies that the largest limit eigenvalue $\lambda_1^{(L)}$ is equal to $\sum_{p=0}^{P-1}\delta_p$ if $L \geq P$, and to $\max_{k=0,P-L}\sum_{p=0}^{L-1}\delta_{p+k}$ if $L \leq P$. If $L \geq P$, the left handside of (2.3.9) is equal to

$$\frac{\sum_{p=0}^{P-1}\delta_p}{\sigma^2\sqrt{dL}}$$

while it is equal to

$$\frac{\max_{k=0,\dots,P-L}\sum_{p=0}^{L-1}\delta_{p+k}}{\sigma^2\sqrt{dL}}$$

if $L \leq P$. The optimal value of L of course depends on the particular values of $\delta_0, \ldots, \delta_{P-1}$. If the powers all coincide with a common term δ, the optimal value is $L = P$, and the test based on $\eta_N^{(P)}$ is consistent if and only if

$$\delta > \frac{\sigma^2 \sqrt{d}}{\sqrt{P}} \tag{2.3.10}$$

In contrast, we mention that if $L = 1$, the consistency condition is

$$\delta > \sigma^2 \sqrt{d} \tag{2.3.11}$$

Therefore, choosing $L = P$ allows to gain a factor $\sqrt{P}$ w.r.t. $L = 1$. If $L > P$, the consistency condition is equivalent to

$$\delta > \frac{\sigma^2 \sqrt{d}}{\sqrt{P}} \sqrt{L/P} \tag{2.3.12}$$

thus showing that an overdetermination of L may also induce a loss of performance.

We now consider a more realistic scenario in which matrix $\mathbf{\Delta}$ is not diagonal. We assume that the signal to be detected is a sampled version of a continuous time linearly modulated signal $\sum_n s_n g_a(t - nT)$ where $(s_n)_{n \in \mathbb{Z}}$ is an i.i.d. sequence of symbols and where $g_a(t)$ is a continuous time shaping filter. The propagation channel is a Rayleigh multipath channel with Q uncorrelated paths with time-delays $\tau_0, \ldots, \tau_{Q-1}$. In this context, vectors $(\mathbf{h}_p)_{p=0,\ldots,P-1}$ are given by

$$\mathbf{h}_p = \sum_{q=0}^{Q-1} \boldsymbol{\lambda}_q \, g_a(pT - \tau_q) \tag{2.3.13}$$

where vectors $(\boldsymbol{\lambda}_q)_{q=0,\ldots,Q-1}$ are the realizations of independent zero-mean random Gaussian vectors. We denote by $\mathbf{\Lambda}$ the $M \times Q$ matrix $\mathbf{\Lambda} = (\boldsymbol{\lambda}_0, \ldots, \boldsymbol{\lambda}_{Q-1})$, and assume that matrix $\mathbf{\Lambda}^* \Lambda$ converges towards $\mu \mathbf{I}_Q$. In practice, this hypothesis means that the Q paths share the same power. As $\mathbf{H} = (\mathbf{h}_{P-1}, \ldots, \mathbf{h}_0)$ is given by $\mathbf{H} = \mathbf{\Lambda} \mathbf{G}$ where $\mathbf{G} = (\mathbf{g}_{P-1}, \ldots, \mathbf{g}_0)$ and where each Q–dimensional vector $\mathbf{g}_p$ is given by $\mathbf{g}_p = (g_a((P-1)T - \tau_0), \ldots, g_a((P-1)T - \tau_{Q-1}))^T$, it is clear that matrix $\mathbf{H}^* \mathbf{H}$ converges towards $\mathbf{\Delta} = \mu \mathbf{G}^* \mathbf{G}$, and that $\mathbf{H}^{(L)*} \mathbf{H}^{(L)}$ converges towards $\mathbf{\Delta}^{(L)} = \mu \mathbf{G}^{(L)*} \mathbf{G}^{(L)}$ where matrix $\mathbf{G}^{(L)}$ is the $QL \times (P+L-1)$ block-Toeplitz matrix with first block line $(\mathbf{g}_{P-1}, \ldots, \mathbf{g}_0, 0, \ldots, 0)$. Therefore, the largest eigenvalue $\lambda_1^{(L)}$ of $\mathbf{\Delta}^{(L)}$ is equal to $\mu \lambda_1(\mathbf{G}^{(L)*} \mathbf{G}^{(L)})$. The optimal value of L thus depends on the way the largest eigenvalue of $\mathbf{G}^{(L)*} \mathbf{G}^{(L)}$ increases with L. As the optimal value of L cannot be found using analytical arguments, we give a numerical example. We assume that $g_a(t)$ is a square root Nyquist filter with excess bandwidth 0.5 which is truncated to interval $[-2.5T, 2.5T]$. Moreover, $Q = 2$, $\tau_0 = 0$, $\tau_1 = 2T$, $d = 1/2$ and the SNR $\frac{\mu}{\sigma^2}$ is equal to 2 dB. In figure 2.1, we plot the largest eigenvalue of $\mathbf{G}^{(L)*} \mathbf{G}^{(L)}$ and the lefthandside of (2.3.9) versus L. It is seen that the optimal value of L is equal to 3, it is thus different from P, which, is the present context is equal to $P = 7$.

2.3.4 Simulation results.

In this section, we provide numerical simulations illustrating the results given in the previous paragraphs. We first consider the case where matrix $\mathbf{H}$ coincides with a realization of a Gaussian random matrix with i.i.d. $\mathcal{N}_{\mathbb{C}}(0, \mathbf{I}/(MP))$ entries. In this context, matrix $\mathbf{H}^* \mathbf{H}$ converge towards $\mathbf{I}_P / P$. Sequence $(s_n)_{n=2-P,\ldots,N-P+1}$ is a realization of an i.i.d. sequence taking values ± 1 with probability $1/2$. In this context, we have shown before that the optimal value of L is equal to P. In order to illustrate this behaviour, we consider the case $M = 80, N = 160$ and $P = 5$, and represent in figure 2.2 the ROC curves, evaluated using Monte-Carlo simulations, corresponding to the statistics $\eta_N^{(L)}, \eta_N, \lambda_1(\frac{\mathbf{Y}^{(L)} \mathbf{Y}^{(L)*}}{N})$ and $\lambda_1(\frac{\mathbf{Y}\mathbf{Y}*}{N})$, referred to as spatio-temporal, spatial, lmax-st and lmax-s in figure 2.2. The numerical results confirm that the

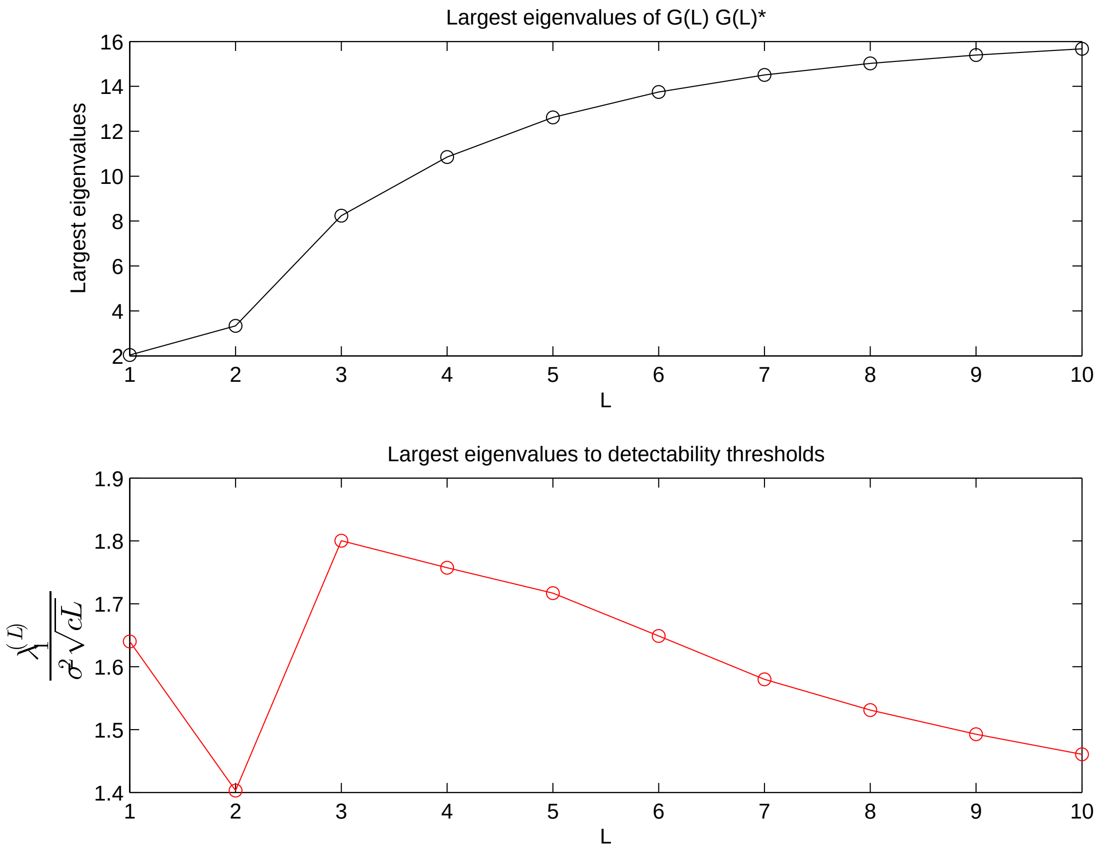

Figure 2.1 – Ratio Largest eigenvalue over detectability threshold versus L

use of $\eta_N^{(5)}$ leads to much better results than the use of η_N which corresponds to $L = 1$, and that it is indeed beneficial to take into account the $P + L - 1$ largest eigenvalues of the empirical spatio-temporal covariance matrix, and not only the largest one.

We now generate vectors $(\mathbf{h}_p)_{p=0,\dots,P-1}$ according to model (2.3.13) for $Q = 2, \tau_0 = 0, \tau_1 = 2T, P = 7$ and when $g_a(t)$ is a square root Nyquist filter with excess bandwidth 0.5 which is truncated to interval $[-2.5T, 2.5T]$. In figure 2.3, we assume that $M = 80, N = 160$ and again represent the ROC curves corresponding to the statistics $\eta_N^{(L)}, \eta_N, \lambda_1(\frac{\mathbf{Y}^{(L)}\mathbf{Y}^{(L))*}}{N})$ and $\lambda_1(\frac{\mathbf{YY}*}{N})$ for $L = 3$. This time, it is seen that it is not beneficial to take into account the $L + P - 1$ largest eigenvalues of $\mathbf{Y}^{(L)}\mathbf{Y}^{(L)*}/N$, and that the best strategy is to consider the largest eigenvalue, which, for $L = 3$, provides the best results.

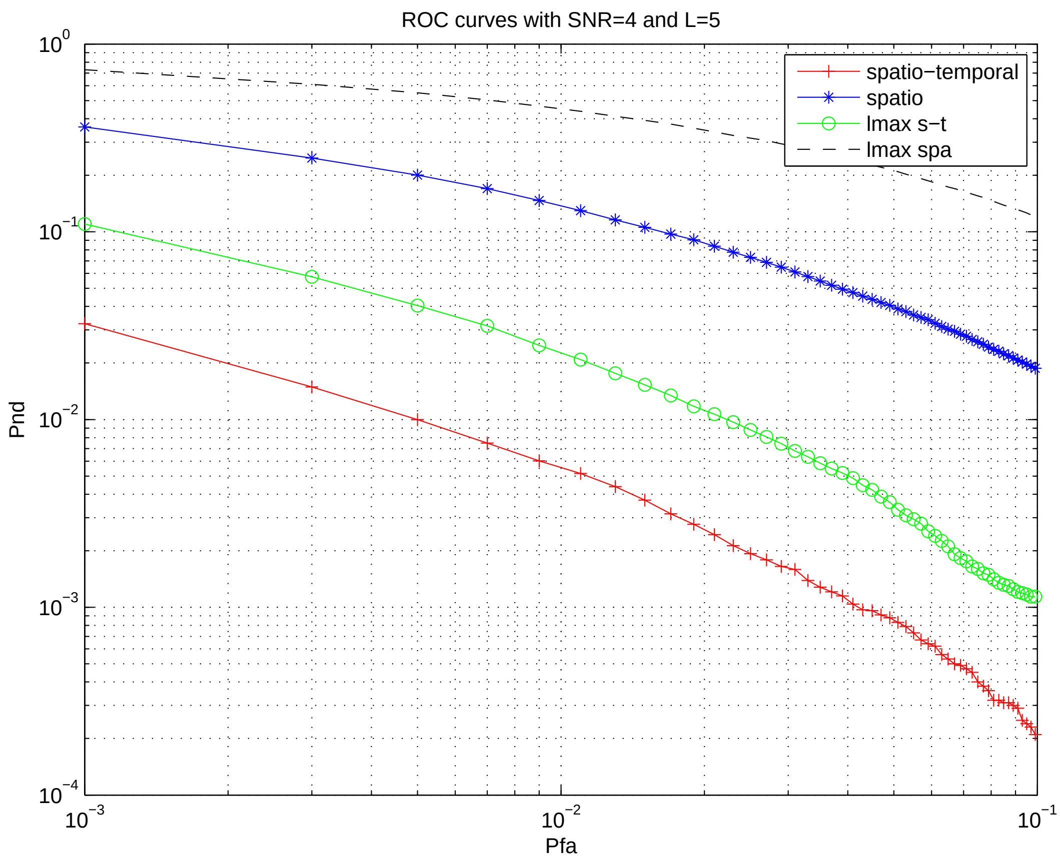

Figure 2.2 – ROC curves of different statistics, $\boldsymbol{\Delta}$ diagonal

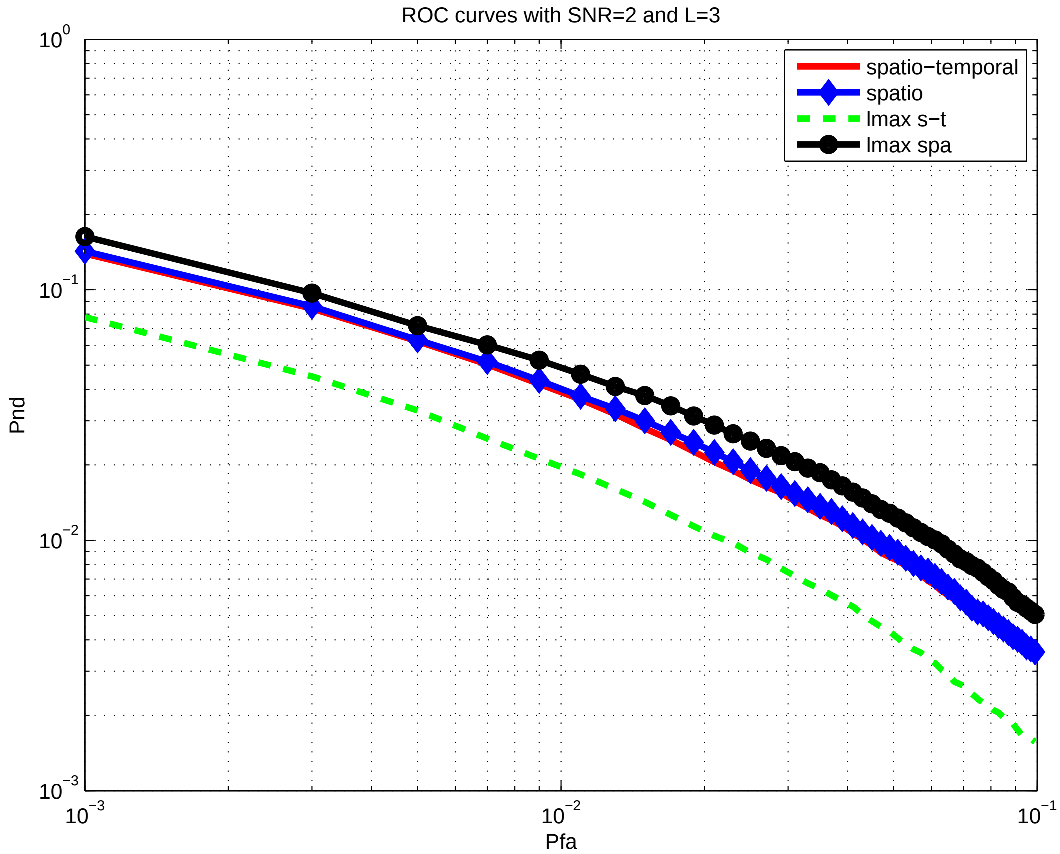

Figure 2.3 – ROC curves of different statistics, $\boldsymbol{\Delta}$ not diagonal

2.4 Estimation of regularized spatio-temporal Wiener filters

2.4.1 Introduction

Finite impulse response spatio-temporal Wiener filter estimation using a training sequence is a very classical problem. When the useful signal is corrupted by an additive temporally and spatially white Gaussian noise, the optimal estimator is known to be the standard least-squares estimate defined as the action of the inverse of the empirical spatio-temporal covariance matrix on the empirical cross correlation between the observation and the training sequence. However, it is known for a long time that regularizing the empirical spatio-temporal covariance matrix by a multiple of the identity matrix may enhance the performance of the estimate because this matrix can be ill-conditioned or even non invertible when the size of the training sequence is smaller than the dimension of the vector associated to the Wiener filter. The choice of the regularization parameter appears to be a crucial issue that was addressed in a heuristic manner in a number of references (see e.g. [39], [78, p. 748], [50] and [42]) because classical figures of merit such as the signal to interference plus noise ratio (SINR) produced by the estimated Wiener filter are difficult to estimate in the general case. In the context of large dimensional systems where the number of sensors and the length of training sequence are both large, the situation appears more favourable due to some subtle self-averaging effects. The existing related works addressed the purely spatial context. Ledoit and Wolf proposed in [44] to find the loading factor so as to minimize the mean-square error of the estimated empirical covariance matrix, and showed that the optimal value can be estimated consistently. This approach was generalized in [16] to the Tyler estimator in the context of robust estimation. Mestre and Lagunas [53] considered the case where the array response is a priori known (no training sequence) and where the interference plus noise covariance matrix is unknown. It is shown in [53] that the SINR produced by the regularized estimated Wiener filter can be consistently estimated from the available observations, and [53] proposed to estimate the loading factor as the argument of the SINR maximization. The optimization of the SINR was also considered in [83] in the context of robust estimation.

In the present section, we assume that the observation is a M–dimensional time series defined as a noisy output of an unknown SIMO finite impulse response system driven by the sequence of interest. We assume that a length N training sequence is available at the receiver side in order to estimate a regularized degree $L-1$ FIR spatio-temporal Wiener filter from the N M–dimensional observations collected during the transmission of the training sequence. In the large system context in which M and N both converge towards $+\infty$ at the same rate and where L remains fixed, we establish that the SINR produced by the regularized estimated Wiener filter, which, in principle, depends on the additive noise corrupting the N available observations, converges towards a deterministic term depending on the loading factor, the noise variance, assumed to be known, and the unknown filter. We show that, while the channel filter is unknown, the above limit SINR can be estimated consistently from the N available observations for each value of the regularization parameter, and propose to estimate the loading factor as the argument of its maximum.

This section is organized as follows. In subsection 2.4.2, we present the signal models and the underlying assumptions. In subsection 2.4.3, we present some useful technical results. In subsection 2.4.4, we establish that the SINR converges towards a deterministic term, and subsection 2.4.5 addresses the consistent estimation of the limit SINR. Finally, subsection 2.4.6 presents numerical experiments sustaining our theoretical results, and comparing our proposal to the Ledoit-Wolf ([44]) estimator of the regularization parameter and to other empirical schemes proposed in the past ([39], [78, p. 748], [50] and [42]).

2.4.2 Problem formulation.

We assume that the observation is a M–dimensional time series $(\mathbf{y}_n)_{n\in\mathbb{Z}}$ defined by

$$\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p} + \mathbf{v}_n, n = 1, ..., N \tag{2.4.1}$$

where $\mathbf{h}(z) = \sum_{p=0}^{P-1} \mathbf{h}_p z^{-p}$ represents the transfer function of the unknown FIR SIMO system and $(\mathbf{v}_n)_{n\in\mathbb{Z}}$ is an i.i.d sequence of complex Gaussian random vectors with spatial covariance matrix $\sigma^2\mathbf{I}$. Although $\mathbf{h}(z)$ is not known, we assume that P is known, i.e. in practice, that an upper bound of the support of the impulse response associated to $\mathbf{h}(z)$ is available. We assume that a length N training sequence $(s_n)_{n=1,...,N}$ is available at the receiver side, and estimate from $(\mathbf{y}_n)_{n=1,...,N}$ the Wiener spatio-temporal filter $\mathbf{g} = (\mathbf{g}_0^T, \ldots, \mathbf{g}_{L-1}^T)^T$ defined in such a way that $\sum_{l=0}^{L-1} \mathbf{g}_l^* \mathbf{y}_{n+l}$ represents the minimum mean-square estimate of s_n. If we denote by $\mathbf{y}_n^{(L)}$ the ML–dimensional vector defined by

$$\mathbf{y}_n^{(L)} = (\mathbf{y}_{1,n}, \ldots, \mathbf{y}_{1,n+L-1}, \ldots, \mathbf{y}_{M,n}, \ldots, \mathbf{y}_{M,n+L-1})^T$$

we study the performance of the estimated regularized Wiener filter $\hat{\mathbf{g}}_\lambda$ defined by

$$\hat{\mathbf{g}}_\lambda = \left(\frac{1}{N} \sum_{n=1}^{N} \mathbf{y}_n^{(L)} \mathbf{y}_n^{(L)*} + \lambda \mathbf{I}_{ML} \right)^{-1} \left(\frac{1}{N} \sum_{n=1}^{N} \mathbf{y}_n^{(L)} s_n^* \right) \tag{2.4.2}$$

$\hat{\mathbf{g}}_\lambda$ is destined to estimate the unknown transmitted data $(s_n)_{n>N}$. In the following, for each $m = 1, \ldots, M$, we denote by $\mathbf{H}_m$ the $L \times (P + L - 1)$ Toeplitz matrix corresponding to the convolution of signal $(s_n)_{n\in\mathbb{Z}}$ with sequence $(\mathbf{h}_{m,p})_{p=0,...,P-1}$, and define $ML \times (P + L - 1)$ block-Hankel matrix $\mathbf{H}$ by $\mathbf{H} = (\mathbf{H}_1^T, ..., \mathbf{H}_M^T)^T$. Assuming sequence $(s_n)_{n>N}$ i.i.d., the signal to interference plus noise ratio produced by $\hat{\mathbf{g}}_\lambda^{(L)}$ is easily seen to be equal to

$$\mathsf{SINR}(\hat{\mathbf{g}}_\lambda) = \frac{|\hat{\mathbf{g}}_\lambda^* \mathbf{h}_P|^2}{\hat{\mathbf{g}}_\lambda^* \mathbf{H}_{-P} \mathbf{H}_{-P}^* \hat{\mathbf{g}}_\lambda + \sigma^2 \|\hat{\mathbf{g}}_\lambda\|^2} \tag{2.4.3}$$

where $\mathbf{h}_P$ is column P of $\mathbf{H}$, and matrix $\mathbf{H}_{-P}$ is obtained by deleting column P from matrix $\mathbf{H}$. $\mathsf{SINR}(\hat{\mathbf{g}}_\lambda^{(L)})$ is random in the sense that it depends on the vectors $(\mathbf{y}_n)_{n=1,...,N}$, which are random themselves due to the presence of the additive noise. When $N \to +\infty$ and M, L remain fixed, it is easy to see that if $\lambda = 0$, the filter $\hat{\mathbf{g}}_0$ converges towards Wiener filter $(\mathbf{H}\mathbf{H}^* + \sigma^2\mathbf{I})^{-1}\mathbf{h}_P$ and that $\mathsf{SINR}(\hat{\mathbf{g}}_0)$ converges towards γ defined by

$$\gamma = \frac{\mathbf{h}_P^* \left(\mathbf{H}\mathbf{H}^* + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{h}_P}{1 - \mathbf{h}_P^* \left(\mathbf{H}\mathbf{H}^* + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{h}_P} \tag{2.4.4}$$

Similar results hold when $\lambda > 0$. On the contrary, when M, N are of the same order of magnitude, the analysis of the behaviour of $\mathsf{SINR}(\hat{\mathbf{g}}_\lambda)$ is different and requires much more work. From now on, we assume that

M, N $\to +\infty$, the ratio $d_N = \frac{M}{N} \to d > 0$, and P and L remain fixed.

To simplify the notations, $N \to +\infty$ should be understood as the above asymptotic regime. We also denote by c_N the ratio $c_N = \frac{ML}{N}$ which converges towards $c = dL$.

In the following, we again consider the normalized block-Hankel matrices $\mathbf{\Sigma}_N$ and $\mathbf{W}_N$ defined by

$$\mathbf{\Sigma}_N = \frac{1}{\sqrt{N}} (\mathbf{y}_1^{(L)}, ..., \mathbf{y}_N^{(L)}), \mathbf{W}_N = \frac{1}{\sqrt{N}} (\mathbf{v}_1^{(L)}, ..., \mathbf{v}_N^{(L)})$$

respectively. Then, the relations between the available observations and sequence $(s_n)_{n=1,...,N}$ can be expressed as

$$\mathbf{\Sigma}_N = \mathbf{H}\mathbf{U}_N + \mathbf{W}_N \tag{2.4.5}$$

where $\mathbf{U}_N$ is the $(P+L-1)\times N$ Hankel matrix defined by $(\mathbf{U}^{(L)})_{i,n} = s_{n+i-P}/\sqrt{N}$. As the dimensions of matrices $\mathbf{H}$ and $\mathbf{U}_N$ scale with M, N, we have to formulate the following hypothesis:

Assumption 2.4.1. $\sup_N \|\mathbf{H}\| < +\infty$, $\sup_N \|\mathbf{U}_N\| < +\infty$

Without loss of generality, we can assume that $\mathbf{U}_N\mathbf{U}_N^* = \mathbf{I}$, because it is possible to replace $\mathbf{H}$ by $\mathbf{H}(\mathbf{U}_N\mathbf{U}_N^*)^{1/2}$ and $\mathbf{U}_N$ by $\mathbf{U}_N\,(\mathbf{U}_N\mathbf{U}_N^*)^{-1/2}$ without modifying the model and Assumption 2.4.1.

In the following, we define $\mathbf{Q}_N(z)$ (resp. $\tilde{\mathbf{Q}}_N(z)$) as the resolvent of matrix $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ (resp. of matrix $\mathbf{\Sigma}_N^*\mathbf{\Sigma}_N$), and remark that the estimated Wiener filter $\hat{\mathbf{g}}_\lambda$ can be written as

$$\hat{\mathbf{g}}_\lambda^{(L)} = \mathbf{Q}_N(-\lambda)\mathbf{\Sigma}_N\mathbf{u}_N^*$$

where $\mathbf{u}_N = \frac{1}{\sqrt{N}}(s_1,\dots,s_N)$ is the P-th row of matrix $\mathbf{U}_N$. To evaluate the behaviour of the SINR given by formula (4.4.2) when $N \to +\infty$, it is sufficient to study the following 3 terms:

$$r_{1,N} = \mathbf{h}_P^*\hat{\mathbf{g}}_\lambda = \mathbf{h}_P^*\mathbf{Q}_N(-\lambda)\mathbf{\Sigma}_N\mathbf{u}_N^* \tag{2.4.6}$$

$$r_{2,N} = \mathbf{H}_{-P}^*\hat{\mathbf{g}}_\lambda = \mathbf{H}_{-P}^*\mathbf{Q}_N(-\lambda)\mathbf{\Sigma}_N\mathbf{u}_N^* \tag{2.4.7}$$

$$r_{3,N} = \|\hat{\mathbf{g}}_\lambda\|^2 = \mathbf{u}_N\mathbf{\Sigma}_N^*\mathbf{Q}_N(-\lambda)^2\mathbf{\Sigma}_N\mathbf{u}_N^* \tag{2.4.8}$$

$r_{1,N}$ and each component of $P+L-2$–dimensional vector $r_{2,N}$ are bilinear forms of matrix $\mathbf{Q}_N(z)\mathbf{\Sigma}_N$ for $z=-\lambda$. As P and L do not scale with M,N, the asymptotic behaviour of $\|r_{2,N}\|^2$ is equivalent to that of each component of $r_{2,N}$. Therefore, evaluating the behaviour of each bilinear form of $\mathbf{Q}_N(z)\mathbf{\Sigma}_N$ for $z=-\lambda$ allows to characterize $\left|\mathbf{h}_P^*\hat{\mathbf{g}}_\lambda\right|^2$ and $\|\mathbf{H}_{-P}^*\hat{\mathbf{g}}_\lambda\|^2$ when $N\to+\infty$.

In order to study $r_{3,N}$, we remark that the derivative w.r.t. z of $\mathbf{Q}_N(z)$ coincides with $(\mathbf{Q}_N(z))^2$. Therefore,

$$\mathbf{\Sigma}_N^*\mathbf{Q}_N(z)^2\mathbf{\Sigma}_N = \frac{\partial}{\partial z}\left[\mathbf{\Sigma}_N^*\mathbf{Q}_N(z)\mathbf{\Sigma}_N\right]$$

Using the identity $\mathbf{\Sigma}_N^*\mathbf{Q}_N(z)\mathbf{\Sigma}_N = \tilde{\mathbf{Q}}_N(z)\mathbf{\Sigma}_N^*\mathbf{\Sigma}_N = \mathbf{I} + z\tilde{\mathbf{Q}}_N(z)$, we finally obtain that

$$\mathbf{\Sigma}_N^*\mathbf{Q}_N(z)^2\mathbf{\Sigma}_N = \frac{\partial}{\partial z}\left[z\tilde{\mathbf{Q}}_N(z)\right] \tag{2.4.9}$$

Hence, in order to evaluate the asymptotic behaviour of $r_{3,N}$, it is sufficient to study the behaviour of the bilinear forms of $z\tilde{\mathbf{Q}}_N(z)$, and to differentiate w.r.t. z for $z \in \mathbb{R}^{-,*}$.

2.4.3 Asymptotic behaviour of the bilinear forms of $\mathbf{Q}_N(z)\mathbf{\Sigma}_N$, $z\tilde{\mathbf{Q}}_N(z)$, and $\left(z\tilde{\mathbf{Q}}_N(z)\right)'$.

In this paragraph, we evaluate the behaviour of the bilinear forms of $\mathbf{Q}_N(z)\mathbf{\Sigma}_N$, $z\tilde{\mathbf{Q}}_N(z)$ and $\left(z\tilde{\mathbf{Q}}_N(z)\right)'$. For this, we introduce the following notations: If $\mathbf{E}_N$ and $\mathbf{F}_N$ represent 2 sequence of $ML\times ML$ or $N\times N$ matrices, we will write that $\mathbf{E}_N \simeq \mathbf{F}_N$ if for each sequence of unit vectors $\mathbf{a}_N, \mathbf{b}_N$, it holds that $\mathbf{a}_N^*\mathbf{E}_N\mathbf{b}_N - \mathbf{a}_N^*\mathbf{F}_N\mathbf{b}_N \to 0$.

As $\mathbf{Q}_N(z)\mathbf{\Sigma}_N = \mathbf{Q}_N(z)\mathbf{H}\mathbf{U}_N + \mathbf{Q}_N(z)\mathbf{W}_N$, we have to evaluate separately the bilinear forms of $\mathbf{Q}_N(z)\mathbf{W}_N$ and of $\mathbf{Q}_N(z)$ (recall Assumption 2.4.1). For this, we denote by $\mathbf{Q}_{W,N}(z)$ and $\tilde{\mathbf{Q}}_{W,N}(z)$ the resolvents of matrices $\mathbf{W}_N\mathbf{W}_N^*$ and $\mathbf{W}_N^*\mathbf{W}_N$. (2.2.2, 2.2.3, 2.2.4) imply that

$$\mathbf{Q}_{W,N}(z) \simeq m_c(z)\mathbf{I}_{ML},\ \tilde{\mathbf{Q}}_{W,N}(z) \simeq \tilde{m}_c(z)\mathbf{I}_N,\ \mathbf{Q}_{W,N}(z)\mathbf{W}_N \simeq 0 \tag{2.4.10}$$

for each $z \in \mathbb{C}-\mathbb{R}^+$, where $m_c(z)$ represents the Stieltjes transform of the Marcenko-Pastur distribution of parameters (c,σ^2). Using this, it is easy to check that the calculations of paragraph 1.2.2 devoted to the case where matrix $\mathbf{W}_N$ has i.i.d. entries are still valid (see Remark 3)). In particular, (1.2.32) remains valid, i.e.

$$\mathbf{Q}_N(z) \simeq \mathbf{S}_N(z)$$

where $\mathbf{S}_N(z)$ is the $M \times M$ matrix-valued function defined by

$$\mathbf{S}_N(z) = \left(-z(1+\sigma^2 \tilde{m}_c(z)) + \frac{\mathbf{H}\mathbf{H}^*}{1+\sigma^2 c m_c(z)}\right)^{-1}$$

or equivalently by

$$\mathbf{S}_N(z) = m_c(z) \left(\mathbf{I} - z m_c(z) \tilde{m}_c(z) \mathbf{H}\mathbf{H}^*\right)^{-1} \tag{2.4.11}$$

Therefore, it holds that

$$\mathbf{Q}_N(z)\mathbf{H}\mathbf{U}_N \simeq m_c(z) \left(\mathbf{I} - z m_c(z) \tilde{m}_c(z) \mathbf{H}\mathbf{H}^*\right)^{-1} \mathbf{H}\mathbf{U}_N \tag{2.4.12}$$

In order to evaluate the bilinear forms of $\mathbf{Q}_N(z)\mathbf{W}_N$, we express $\mathbf{Q}_N(z)$ in terms of $\mathbf{Q}_{W,N}(z)$ by a formula similar to (1.2.31), except that the singular value decomposition of the deterministic part of $\mathbf{\Sigma}_N$ is replaced by factorization $\mathbf{H}\mathbf{U}_N$ with $\mathbf{U}_N\mathbf{U}_N^* = \mathbf{I}_{P+L-1}$. We have:

$$\mathbf{Q}_N = \mathbf{Q}_{W,N} - \mathbf{Q}_{W,N}(\mathbf{H}, \mathbf{W}_N\mathbf{U}_N^*) \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix} \times \tag{2.4.13}$$

$$\left[\mathbf{I} + \begin{pmatrix} \mathbf{H}^* \\ \mathbf{U}_N\mathbf{W}_N^* \end{pmatrix} \mathbf{Q}_W(\mathbf{H}, \mathbf{W}_N\mathbf{U}_N^*) \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix}\right]^{-1} \begin{pmatrix} \mathbf{U}_N^* \\ \mathbf{U}_N\mathbf{W}_N^* \end{pmatrix} \mathbf{Q}_W$$

Using (2.4.10), we obtain after some algebra that

$$\mathbf{Q}_N(z)\mathbf{W}_N \simeq -m_c(z)(1+z\tilde{m}_c(z)) \left(\mathbf{I} - z m_c(z) \tilde{m}_c(z) \mathbf{H}\mathbf{H}^*\right)^{-1} \mathbf{H}\mathbf{U}_N \tag{2.4.14}$$

Therefore, it holds that

$$\mathbf{Q}_N(z)\mathbf{\Sigma}_N \simeq -z m_c(z) \tilde{m}_c(z) \left(\mathbf{I} - z m_c(z) \tilde{m}_c(z) \mathbf{H}\mathbf{H}^*\right)^{-1} \mathbf{H}\mathbf{U}_N \tag{2.4.15}$$

or equivalently,

$$\mathbf{Q}_N(z)\mathbf{\Sigma}_N \simeq \left(\mathbf{H}\mathbf{H}^* - w_c(z)\mathbf{I}\right)^{-1} \mathbf{H}\mathbf{U}_N \tag{2.4.16}$$

As for $z\tilde{\mathbf{Q}}_N(z)$, (1.2.36) remains valid, i.e.

$$\tilde{\mathbf{Q}}_N(z) \simeq \tilde{\mathbf{S}}_N(z)$$

where $\tilde{\mathbf{S}}_N(z)$ is the $N \times N$ matrix-valued function defined by

$$\tilde{\mathbf{S}}_N(z) = \left(-z(1+\sigma^2 c m_c(z)) + \frac{\mathbf{U}_N^*\mathbf{H}^*\mathbf{H}\mathbf{U}_N}{1+\sigma^2 \tilde{m}_c(z)}\right)^{-1}$$

which can also be written as

$$\tilde{\mathbf{S}}_N(z) = \left(1+\sigma^2 \tilde{m}_c(z)\right) \left(\mathbf{U}_N^*\mathbf{H}^*\mathbf{H}\mathbf{U}_N - w_c(z)\mathbf{I}\right)^{-1} = -\frac{1}{z m_c(z)} \left(\mathbf{U}_N^*\mathbf{H}^*\mathbf{H}\mathbf{U}_N - w_c(z)\mathbf{I}\right)^{-1} \tag{2.4.17}$$

Therefore, it holds that

$$z\tilde{\mathbf{Q}}_N(z) \simeq -\frac{1}{m_c(z)} \left(\mathbf{U}_N^*\mathbf{H}^*\mathbf{H}\mathbf{U}_N - w_c(z)\right)^{-1} \tag{2.4.18}$$

or, using the expression (1.2.11) of $m_c(z)$ in terms of $w_c(z)$,

$$z\tilde{\mathbf{Q}}_N(z) \simeq (w_c(z) + \sigma^2 c) \left(\mathbf{U}_N^*\mathbf{H}^*\mathbf{H}\mathbf{U}_N - w_c(z)\right)^{-1} \tag{2.4.19}$$

Writing the deterministic equivalent of $z\tilde{\mathbf{Q}}_N(z)$ in terms of $w_c(z)$ is useful because the calculation of the derivative w.r.t. z of the righthandside of (2.4.20) appears easy due to the rather simple expression of $w_c^{'}(z) = \frac{w_c^2(z)}{w_c^2(z)-\sigma^4 c}$ (see 1.2.35).

In order to obtain an approximation of the bilinear forms of $\left(z\tilde{\mathbf{Q}}_N(z)\right)^{'}$, we remark that it is possible to differentiate (2.4.20) because the convergence of the corresponding bilinear forms is uniform on each compact subset of $\mathbb{C}-\mathbb{R}^+$. After some algebra, we obtain that

$$\left(z\tilde{\mathbf{Q}}_N(z)\right)^{'} \simeq \frac{w_c^2(z)}{w_c^2(z)-\sigma^4 c}\left(\mathbf{U}_N^*\mathbf{H}^*\mathbf{H}\mathbf{U}_N - w_c(z)\right)^{-2}\left(\mathbf{U}_N^*\mathbf{H}^*\mathbf{H}\mathbf{U}_N + \sigma^2 c\right) \tag{2.4.20}$$

2.4.4 Asymptotic behaviour of the SINR

Using (2.4.16) and (2.4.20), we obtain after some algebra the following result.

Proposition 2.4.1. *For each $\lambda > 0$, when $N \to +\infty$, the three terms $\|\mathbf{h}_P^*\hat{\mathbf{g}}_\lambda\|^2$, $\|\mathbf{H}_{-P}^*\hat{\mathbf{g}}_\lambda\|^2$ et $\|\hat{\mathbf{g}}_\lambda\|^2$ can be approximated (i.e. have the same almost sure behaviour) by the following deterministic quantities:*

- $$\|\mathbf{h}_P^*\hat{\mathbf{g}}_\lambda\|^2 \simeq (\mathbf{h}_P^*(\mathbf{H}\mathbf{H}^* - w_c(-\lambda)\mathbf{I})^{-1}\mathbf{h}_P)^2 \tag{2.4.21}$$

- $$\|\mathbf{H}_{-P}^*\hat{\mathbf{g}}_\lambda\|^2 \simeq w_c(-\lambda)\mathbf{h}_P^*(\mathbf{H}\mathbf{H}^* - w_c(-\lambda)\mathbf{I})^{-2}\mathbf{h}_P + \mathbf{h}_P^*(\mathbf{H}\mathbf{H}^* - w_c(-\lambda)\mathbf{I})^{-1}\mathbf{h}_P\left(1-\mathbf{h}_P^*(\mathbf{H}\mathbf{H}^* - w_c(-\lambda)\mathbf{I})^{-1}\mathbf{h}_P\right) \tag{2.4.22}$$

- $$\|\hat{\mathbf{g}}_\lambda\|^2 \simeq \frac{\sigma^2 c}{w_c(-\lambda)-\sigma^4 c}\left(1-\mathbf{h}_P^*(\mathbf{H}\mathbf{H}^* - w_c(-\lambda)\mathbf{I})^{-1}\mathbf{h}_P\right) + \frac{w_c(-\lambda)(\sigma^2 c + w_c(-\lambda))}{w_c(-\lambda)-\sigma^4 c}\mathbf{h}_P^*(\mathbf{H}\mathbf{H}^* - w_c(-\lambda)\mathbf{I})^{-2}\mathbf{h}_P \tag{2.4.23}$$

Moreover, if we introduce $\alpha(\lambda) = \mathbf{h}_P^*(\mathbf{H}\mathbf{H}^* - w_c(-\lambda)\mathbf{I})^{-1}\mathbf{h}_P$ and $\beta(\lambda) = \mathbf{h}_P^*(\mathbf{H}\mathbf{H}^* - w_c(-\lambda)\mathbf{I})^{-2}\mathbf{h}_P$, it holds that

$$\mathit{SINR}(\hat{\mathbf{g}}_\lambda^{(L)}) - \phi(\lambda) \to 0 \tag{2.4.24}$$

almost surely, where $\phi(\lambda)$ is defined by

$$\phi(\lambda) = \frac{\alpha(\lambda)^2}{[1-\alpha(\lambda)][(\alpha(\lambda) + \frac{\sigma^4 c}{w_c^2(-\lambda)-\sigma^4 c}] + \frac{w_c^2(-\lambda)(w_c(-\lambda)+\sigma^2)}{w_c^2(-\lambda)-\sigma^4 c}\beta(\lambda)} \tag{2.4.25}$$

When $c = d\mathrm{L} < 1$, almost surely, for N large enough, $\mathbf{Q}_N(z)$ is analytic in a neighbourhood of 0. Therefore, the results of Proposition (2.4.1) also hold for $\lambda = 0$. Using the observation that $w_c(0) = -\sigma^2$, we obtain immediately that

$$\phi(0) = \gamma\,\frac{(1-d\mathrm{L})\gamma}{\gamma + d\mathrm{L}} \tag{2.4.26}$$

where γ is the SINR corresponding to the true Wiener filter (see Eq. (2.4.4)). Consequently, the estimation of the Wiener filter by $\hat{\mathbf{g}}_{(0)}^{\mathrm{L}}$ produces a SINR loss equal to $(1-d\mathrm{L})\frac{\gamma}{\gamma+d\mathrm{L}}$, which, of course, is considerable when $d\mathrm{L}$ is close from 1. As shown below, the use of a convenient regularization coefficient allows to improve considerably the SINR in this context.

2.4.5 Consistent estimators of the SINR

It is clear that function $\lambda \to \phi(\lambda)$ depends on matrix $\mathbf{H}$ which is unknown. We establish in this subsection that it is possible to estimate $\phi(\lambda)$ consistently for each $\lambda > 0$. For this, it is sufficient to estimate $\alpha(\lambda)$ and $\beta(\lambda)$ (see (2.4.25)). In the following, we show that $\alpha(\lambda)$ and $\beta(\lambda)$ can be expressed in terms of $\tilde{\mathbf{S}}_N(z)$ given by (2.4.17) and of the derivative of $z\tilde{\mathbf{S}}_N(z)$, and build the estimators from the observation that $\tilde{\mathbf{Q}}_N(z) \simeq \tilde{\mathbf{S}}_N(z)$ and from $\left(z\tilde{\mathbf{Q}}_N(z)\right)' \simeq \left(z\tilde{\mathbf{S}}_N(z)\right)'$ for each $z \in \mathbb{C} - \mathbb{R}^+$.

We first observe that

$$\mathbf{H}^* \left(\mathbf{H}\mathbf{H}^* - w_c(z)\mathbf{I}\right)^{-1} \mathbf{H} = \left(\mathbf{H}^*\mathbf{H} - w_c(z)\mathbf{I}\right)^{-1} \mathbf{H}^*\mathbf{H} = \mathbf{I} + w_c(z) \left(\mathbf{H}^*\mathbf{H} - w_c(z)\mathbf{I}\right)^{-1}$$

Taking the (P, P) entry of this matrix equation, we obtain that $\alpha(\lambda)$ can be written as

$$\alpha(\lambda) = 1 + w_c(-\lambda)((\mathbf{H}^*\mathbf{H} - w_c(-\lambda)\mathbf{I})^{-1})_{P,P}$$

We also remark that

$$((\mathbf{H}^*\mathbf{H} - w_c(-\lambda)\mathbf{I})^{-1})_{P,P} = \mathbf{u}_N\, (\mathbf{U}_N^*\mathbf{H}^*\mathbf{H}\mathbf{U}_N - w_c(-\lambda)\mathbf{I})^{-1})\, \mathbf{u}_N^*$$

(2.4.17) implies that

$$(\mathbf{U}_N^*\mathbf{H}^*\mathbf{H}\mathbf{U}_N - w_c(z)\mathbf{I})^{-1}) = \frac{1}{1+\sigma^2 \tilde{m}_c(z)}\, \tilde{\mathbf{S}}_N(z)$$

and that

$$w_c(z)\, (\mathbf{U}_N^*\mathbf{H}^*\mathbf{H}\mathbf{U}_N - w_c(z)\mathbf{I})^{-1}) = -\frac{1}{\tilde{m}_c(z)}\, \tilde{\mathbf{S}}_N(z)$$

As it holds that $\tilde{\mathbf{Q}}_N(z) \simeq \tilde{\mathbf{S}}_N(z)$, we conclude that $((\mathbf{H}^*\mathbf{H} - w_c(-\lambda)\mathbf{I})^{-1})_{P,P}$ can be consistently estimated $-\frac{1}{1+\sigma^2 \tilde{m}_c(-\lambda)}\, \mathbf{u}_N\, \tilde{\mathbf{Q}}_N(-\lambda)\, \mathbf{u}_N^*$ and that

$$\hat{\alpha}(\lambda) = 1 - \frac{\mathbf{u}_N \tilde{\mathbf{Q}}_N(-\lambda) \mathbf{u}_N^*}{\tilde{m}_d(-\lambda)} \tag{2.4.27}$$

is a consistent estimate of $\alpha(\lambda)$. In order to obtain an estimator $\hat{\beta}(\lambda)$ of $\beta(\lambda)$, we observe that

$$\mathbf{h}_P^*(\mathbf{H}\mathbf{H}^* - w_c(-\lambda)\mathbf{I})^{-2}\mathbf{h}_P = ((\mathbf{H}^*\mathbf{H} - w_c(-\lambda)\mathbf{I})^{-1})_{P,P} + w_c(-\lambda)((\mathbf{H}^*\mathbf{H} - w_c(-\lambda)\mathbf{I})^{-2})_{P,P} \tag{2.4.28}$$

As the first term of the righthandside of (2.4.28) can be consistently estimated, it remains to be able to estimate $((\mathbf{H}^*\mathbf{H} - w_c(-\lambda)\mathbf{I})^{-2})_{P,P}$. For this, we remark that (2.4.20) leads to

$$\mathbf{u}_N\, z\tilde{\mathbf{Q}}_N(z)\, \mathbf{u}_N^* \simeq (w_c(z) + \sigma^2 c)((\mathbf{H}^*\mathbf{H} - w_c(z)\mathbf{I})^{-1})_{P,P}$$

Differentiating this w.r.t. z, we obtain immediately that $(\mathbf{H}^*\mathbf{H} - w_c(-\lambda)\mathbf{I})^{-2})_{P,P}$ can be consistently estimated and that $\beta(\lambda)$ can be estimated by $\hat{\beta}(\lambda)$ defined by

$$\begin{aligned}\hat{\beta}(\lambda) = &-\frac{\sigma^2 c}{(w_c(-\lambda)+\sigma^2 c)^2}\lambda \mathbf{u}_N \tilde{\mathbf{Q}}_N(-\lambda)\mathbf{u}_N^* \\ &+ \frac{w_c^2(-\lambda) - \sigma^4 c}{w_c(-\lambda)(w_c(-\lambda)+\sigma^2 c)}\left[\mathbf{u}_N(\tilde{\mathbf{Q}}_N(-\lambda) - \lambda\tilde{\mathbf{Q}}_N^2(-\lambda))\mathbf{u}_N^*\right]\end{aligned} \tag{2.4.29}$$

Replacing $\alpha(\lambda)$ and $\beta(\lambda)$ by $\hat{\alpha}(\lambda)$ and $\hat{\beta}(\lambda)$ in formula (2.4.25), we obtain immediately a consistent estimator $\hat{\phi}(\lambda)$ of the asymptotic SINR $\phi(\lambda)$. Moreover, it is possible to establish that function $\phi(\lambda) - \hat{\phi}(\lambda)$ converges uniformly towards 0 on each compact subset of $\mathbb{R}_*^+$. Therefore, if we denote by λ_{opt} and $\hat{\lambda}_{opt}$ the argument of the maximum of ϕ and $\hat{\phi}$ on a fixed compact of $\mathbb{R}_*^+$, it holds that $\lambda_{opt} - \hat{\lambda}_{opt} \to 0$. Therefore, maximizing function $\lambda \to \hat{\phi}(\lambda)$ allows to estimate a regularization parameter for which the true asymptotic SINR $\phi(\lambda)$ is maximum. We also notice that this approach allows to choose the smoothing factor L: it is sufficient to evaluate $\hat{\phi}(\hat{\lambda}_{opt})$ for each choice of L, and to select the smoothing factor for which the latter term is maximum. This is of course not a computationally efficient procedure because it needs to evaluate matrix $\mathbf{Q}_N(-\lambda)$ et $\tilde{\mathbf{Q}}_N(-\lambda)$ for each λ and each integer L.

2.4.6 Numerical experiments.

In this section, we provide numerical simulations illustrating the results given in the previous sections. We first illustrate the accuracy of the approximation $\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)}) \simeq \phi(\lambda)$ where we recall that $\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)})$ is the true SINR defined by (4.4.2). Matrix $(\mathbf{h}_0, ..., \mathbf{h}_{P-1})$ is a realization of a normalized version (so as to obtain a Frobenius norm equal to 1) of random matrix $(\mathbf{a}(\theta_0), ..., \mathbf{a}(\theta_{P-1}))$, with $\mathbf{a}(\theta) = \frac{1}{\sqrt{M}}(1, ..., e^{i(M-1)\theta})^T$, and where the angles are drawn uniformly on $[0, 2\pi]$. The sequence $(s_n)_{n=1,...,N}$ is a realization of an i.i.d ± 1 sequence with probability $1/2$. The signal to noise ratio SNR is thus equal to $1/\sigma^2$. In the following experiments, $N = 200, M = 40$ and $P = 5$. In figure 2.4, SNR is equal to 8dB, $L = 5$, and we evaluate by Monte-Carlo simulations (10.000 realizations are generated) function $\lambda \to \text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)})$. We represent the graph of the function $\phi(\lambda)$ along with 2 plots representing the lower and upper bounds of the 95% confidence interval of $\lambda \to \text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)})$. We can notice that the 3 graphs are close one from each other.

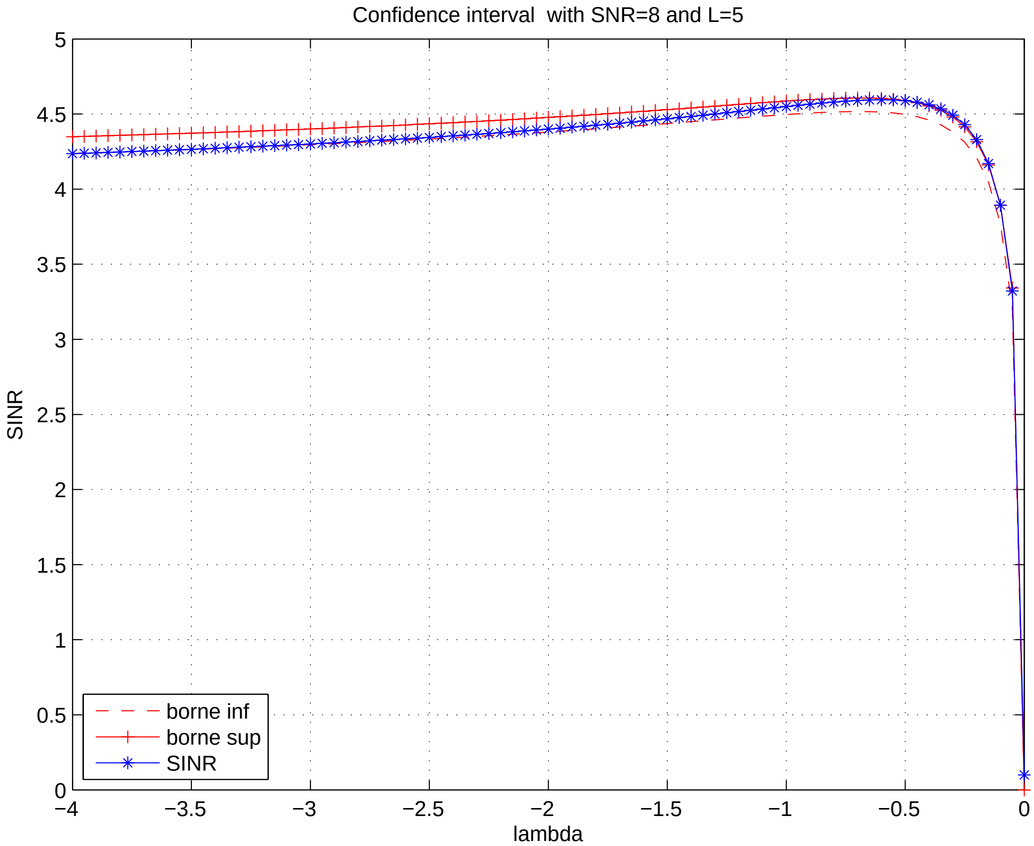

Figure 2.4 – Confidence region and asymptotic curve of SINR versus λ

We now evaluate the performance of the estimator $\hat{\lambda}_{opt}$ of λ_{opt} and evaluate by Monte-Carlo simulation the root relative least mean squares error of $\phi(\hat{\lambda}_{opt}) - \phi(\lambda_{opt})$. We also evaluate the same quantity, but when λ_{opt} is estimated by other existing schemes: the Ledoit-Wolf estimator ([44]), 3 empirical methods mentioned in [53] to be referred to as M1 ([39], [78, p. 748]), M2 [50], M3 [42] in the figure 2.5, and the naive estimate obtained by maximizing w.r.t. λ the expression (4.4.2) in which matrix $\mathbf{H}$ is replaced by $\mathbf{\Sigma U}^*$, which, of course, is not supposed to be a good estimator when M and N are of the same order of magnitude. The various root relative mean squares errors are given in figure 2.5 for various values of the smoothing parameter L.

We finally justify that our approach may be used in order to estimate a relevant value of the smoothing parameter L. Keeping the same parameters as above, we first represent in Figure 2.6 function $\lambda \to$

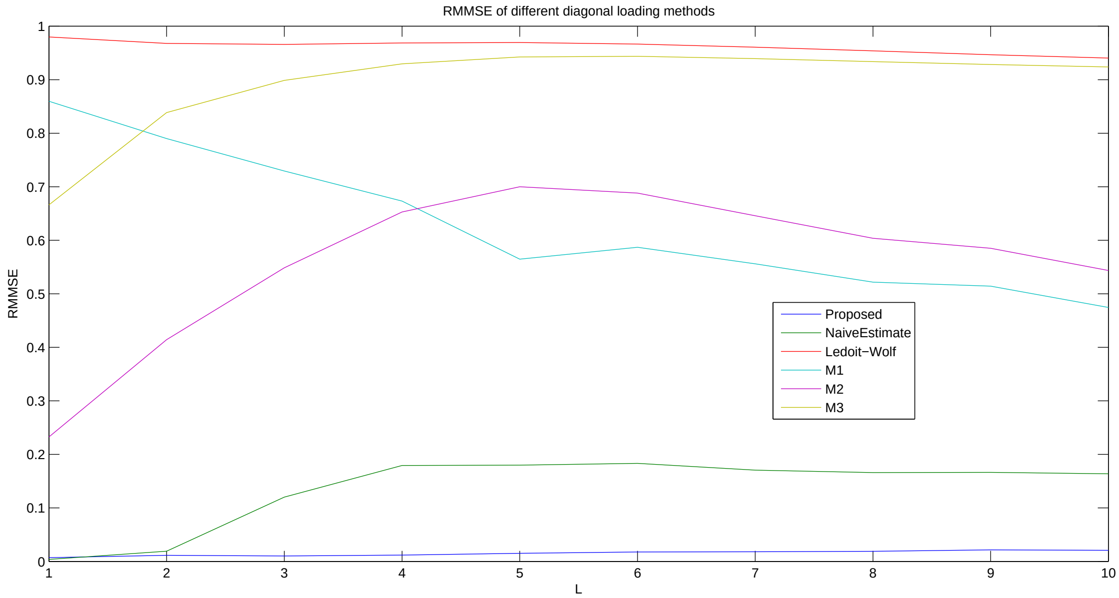

Figure 2.5 – RMMSE of different diagonal loading methods versus L

$\phi(\lambda)$ for $L = 1,2,3,4,5,6,7,8$, and conclude that $L = P = 5$ maximizes $\phi(\lambda_{opt})$, but that $L = 6,7,8$ also provide reasonable optimum asymptotic SINR.

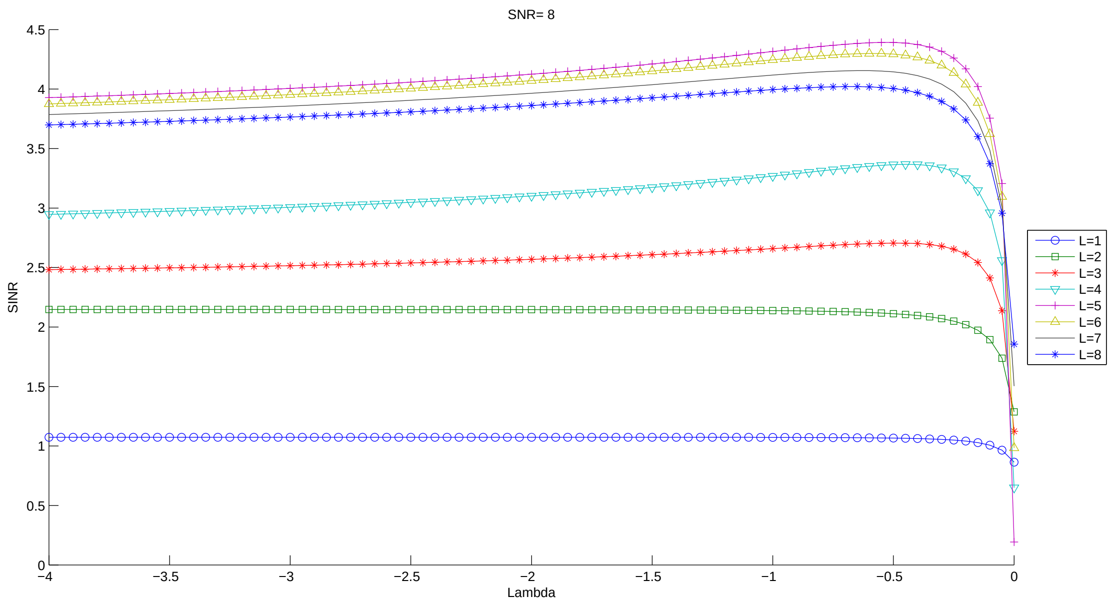

Figure 2.6 – Asymptotic SINR versus L and λ

In order to estimate the optimum value of L, we evaluate $\hat{\phi}(\hat{\lambda}_{opt})$ for each possible value of L, and propose to select the value L_{opt} of L for which the latter term is maximum. We represent in Figure 2.7 the histogram of 10.000 realizations of L_{opt} when L_{opt} is forced to belong to $\{1,2,3,4,5,6,7,8,9,10\}$. Figure

2.7 shows that the selected values of L provide reasonably optimal SINR.

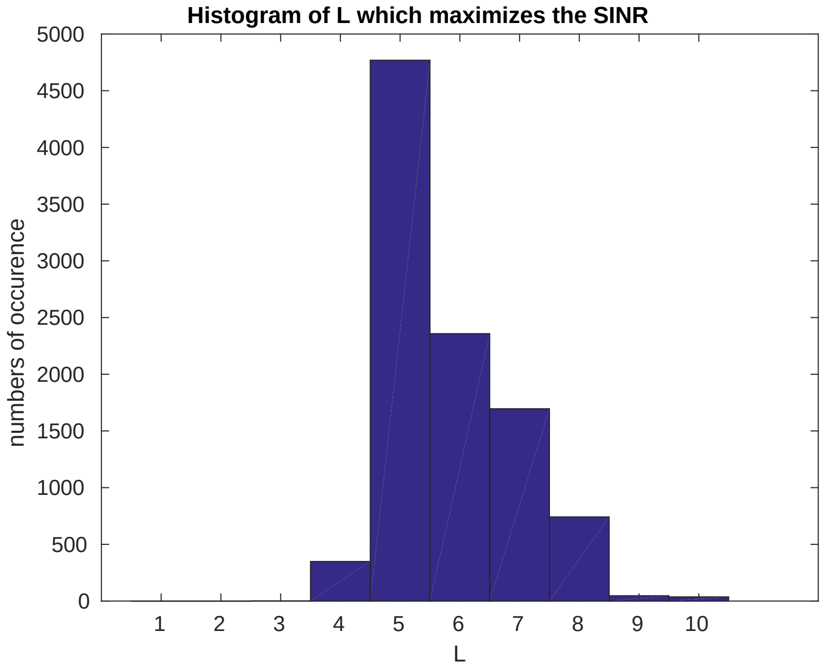

Figure 2.7 – Histogram of L_{opt}

2.5 Performance analysis of spatial smoothing schemes in the context of large arrays

2.5.1 Introduction

The statistical analysis of subspace DoA estimation methods using an array of sensors is a topic that has received a lot of attention since the seventies. Most of the works were devoted to the case where the number of available samples N of the observed signal is much larger than the number of sensors M of the array (see e.g. [69] and the references therein). More recently, the case where M and N are large and of the same order of magnitude was addressed for the first time in [53] using large random matrix theory. [53] was followed by various works such as [40], [72], [34], [33]. The number of observations may also be much smaller than the number of sensors. In this context, it is well established that spatial smoothing schemes, originally developed to address coherent sources ([22], [68], [62]), can be used to artificially increase the number of snapshots (see e.g. [69] and the references therein, see also the recent related contributions [70], [71] devoted to the case where $N = 1$). Spatial smoothing consists in considering $L < M$ overlapping arrays with $M-L+1$ sensors, and allows to generate artificially NL snapshots observed on a virtual array of $M-L+1$ sensors. The corresponding $(M-L+1) \times NL$ matrix, denoted $\mathbf{Y}_N^{(L)}$, collecting the observations is the sum of a low rank component generated by $(M-L+1)$-dimensional steering vectors with a noise matrix having a block-Hankel structure. Subspace methods can still be developed, but the statistical analysis of the corresponding DoA estimators was addressed in the standard regime where $M-L+1$ remains fixed while NL converges towards ∞. This context is not the most relevant when M is large because L must be chosen in such a way that the number of virtual sensors $M-L+1$ be small enough w.r.t. NL, thus limiting the statistical performance of the estimates. In this paper, we study the statistical performance of spatial smoothing subspace DoA estimators in asymptotic regimes where $M - L + 1$ and NL both converge towards ∞ at the same rate, where $\frac{L}{M} \to 0$ in order to not affect the aperture of the virtual array, and where the number of sources K does not scale with M, N, L. For this, it is necessary to evaluate the behaviour of the K largest eigenvalues and corresponding eigenvectors of the empirical covariance matrix $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{NL}$. To address this issue, we prove that the above eigenvalues and eigenvectors have the same asymptotic behaviour as if the noise contribution $\mathbf{V}_N^{(L)}$ to matrix $\mathbf{Y}_N^{(L)}$, a block-Hankel random matrix, was a Gaussian random matrix with independent identically distributed entries. To establish this result, we use the same approach than in the proof of Theorem 2.2.1. This allows to obtain a characterization of the behaviour of the largest eigenvalues and eigenvectors of $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{NL}$. We deduce from this improved subspace estimators, called DoA G-MUSIC SS (spatial smoothing) estimators, which are similar to those of [72] and [33]. We deduce from the results of [74] that when the DoAs do not scale with M, N, L, i.e. if the DoAs are widely spaced compared to aperture array, then both G-MUSIC SS and traditional MUSIC SS estimators are consistent and converge at a rate faster than $\frac{1}{M}$. Moreover, when the DoAs are spaced of the order of $\frac{1}{M}$, the behaviour of G-MUSIC SS estimates remains unchanged, but the convergence rate of traditional subspace estimates is lower.

This section is organized as follows. In subsection 2.5.2, we precise the signal model, the underlying assumptions, and formulate our main results. In subsection 2.5.3, we establish that the largest eigenvalues and eigenvectors of matrix $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{NL}$ behave as if the entries of $\mathbf{V}_N^{(L)}$ were i.i.d., and obtain a result similar to Theorem 2.2.1. In subsection 2.5.4, we use the main result of subsection 2.5.3, and follow [33] and [74] in order to propose a G-MUSIC algorithm adapted to the spatial smoothing context of this paper. The consistency and the convergence speed of the G-MUSIC SS estimates and of the traditional MUSIC SS estimates are then deduced from the results of [74]. Finally, subsection 2.5.5 present numerical experiments sustaining our theoretical results.

2.5.2 Problem formulation and main results.

Problem formulation.

We assume that K narrow-band and far-field source signals are impinging on a uniform linear array of M sensors, with $K < M$. In this context, the M–dimensional received signal $(\mathbf{y}_n)_{n\geq 1}$ can be written as

$$\mathbf{y}_n = \mathbf{A}_M \mathbf{s}_n + \mathbf{v}_n,$$

where

- $\mathbf{A}_M = [\mathbf{a}_M(\theta_1), \ldots, \mathbf{a}_M(\theta_K)]$ is the $M \times K$ matrix of M–dimensionals steering vectors $\mathbf{a}_M(\theta_1), \ldots, \mathbf{a}_M(\theta_K)$, with $\theta_1, \ldots, \theta_K$ the source signals DoA, and $\mathbf{a}_M(\theta) = \frac{1}{\sqrt{M}}[1, \ldots, e^{i(M-1)\theta}]^T$;
- $\mathbf{s}_n \in \mathbb{C}^K$ contains the source signals received at time n, considered as unknown deterministic ;
- $(\mathbf{v}_n)_{n\geq 1}$ is a temporally and spatially white complex Gaussian noise with spatial covariance $\mathbb{E}[\mathbf{v}_n\mathbf{v}_n^*] = \sigma^2\mathbf{I}$.

The received signal is observed between time 1 and time N, and we collect the available observations in the $M \times N$ matrix $\mathbf{Y}_N$ defined

$$\mathbf{Y}_N = [\mathbf{y}_1, \ldots, \mathbf{y}_N] = \mathbf{A}_M \mathbf{S}_N + \mathbf{V}_N, \tag{2.5.1}$$

with $\mathbf{S}_N = [\mathbf{s}_1, \ldots, \mathbf{s}_N]$ and $\mathbf{V}_N = [\mathbf{v}_1, \ldots, \mathbf{v}_N]$. We assume that $\text{Rank}(\mathbf{S}_N) = K$ for each M, N greater than K. The DoA estimation problem consists in estimating the K DoA $\theta_1, \ldots, \theta_K$ from the matrix of samples $\mathbf{Y}_N$.

When the number of observations N is much less than the number of sensors M, the standard subspace method fails. In this case, it is standard to use spatial smoothing schemes in order to artificially increase the number of observations. In particular, it is well established that spatial smoothing schemes allow to use subspace methods even in the single snapshot case, i.e. when $N = 1$ (see e.g. [69] and the references therein). If $L < M$, spatial smoothing consists in considering L overlapping subarrays of dimension $M-L+1$. At each time n, L snapshots of dimension $M-L+1$ are thus available, and the scheme provides NL observations of dimension $M-L+1$. In order to be more specific, we introduce the following notations. If L is an integer less than M, we denote by $\mathscr{Y}_n^{(L)}$ the $(M-L+1)\times L$ Hankel matrix defined by

$$\mathscr{Y}_n^{(L)} = \begin{pmatrix} \mathbf{y}_{1,n} & \mathbf{y}_{2,n} & \cdots & \cdots & \mathbf{y}_{L,n} \\ \mathbf{y}_{2,n} & \mathbf{y}_{3,n} & \cdots & \cdots & \mathbf{y}_{L+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_{M-L+1,n} & \mathbf{y}_{M-L+2,n} & \cdots & \cdots & \mathbf{y}_{M,n} \end{pmatrix} \tag{2.5.2}$$

Column l of matrix $\mathscr{Y}_n^{(L)}$ corresponds to the observation on subarray l at time n. Collecting all the observations on the various subarrays allows to obtain NL snapshots, thus increasing artificially the number of observations. We define $\mathbf{Y}_N^{(L)}$ as the $(M-L+1)\times NL$ block-Hankel matrix given by

$$\mathbf{Y}_N^{(L)} = \left(\mathscr{Y}_1^{(L)}, \ldots, \mathscr{Y}_N^{(L)}\right) \tag{2.5.3}$$

In order to express $\mathbf{Y}_N^{(L)}$, we consider the $(M-L+1)\times L$ Hankel matrix $\mathscr{A}^{(L)}(\theta)$ defined from vector $\mathbf{a}_M(\theta)$ in the same way than $\mathscr{Y}_n^{(L)}$. We remark that $\mathscr{A}^{(L)}(\theta)$ is rank 1, and can be written as

$$\mathscr{A}^{(L)}(\theta) = \sqrt{L(M-L+1)/M}\, \mathbf{a}_{M-L+1}(\theta)\, (\mathbf{a}_L(\theta))^T \tag{2.5.4}$$

We consider the $(M-L+1)\times KL$ matrix $\mathbf{A}^{(L)}$

$$\mathbf{A}^{(L)} = \left(\mathscr{A}^{(L)}(\theta_1), \mathscr{A}^{(L)}(\theta_2), \ldots, \mathscr{A}^{(L)}(\theta_K)\right) \tag{2.5.5}$$

which, of course, is a rank K matrix whose range coincides with the subspace generated by the $(M-L+1)$-dimensional vectors $\mathbf{a}_{M-L+1}(\theta_1), \ldots, \mathbf{a}_{M-L+1}(\theta_K)$. $\mathbf{Y}_N^{(L)}$ can be written as

$$\mathbf{Y}_N^{(L)} = \mathbf{A}^{(L)}\, (\mathbf{S}_N \otimes \mathbf{I}_L) + \mathbf{V}_N^{(L)} \tag{2.5.6}$$

where matrix $\mathbf{V}_N^{(L)}$ is the block Hankel matrix corresponding to the additive noise. As matrix $\mathbf{S}_N \otimes \mathbf{I}_L$ is full rank, the extended observation matrix $\mathbf{Y}_N^{(L)}$ appears as a noisy version of a low rank component whose range is the K–dimensional subspace generated by vectors $\mathbf{a}_{M-L+1}(\theta_1), \ldots, \mathbf{a}_{M-L+1}(\theta_K)$. Moreover, it is easy to check that

$$\mathbb{E}\left(\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{NL}\right) = \sigma^2 \mathbf{I}_{M-L+1}$$

Therefore, it is potentially possible to estimate the DoAs $(\theta_k)_{k=1,\ldots,K}$ using a subspace approach based on the eigenvalues / eigenvectors decomposition of matrix $\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL$. The asymptotic behaviour of spatial smoothing subspace methods is standard in the regimes where $M-L+1$ remains fixed while NL converges towards ∞. This is due to the law of large numbers which implies that the empirical covariance matrix $\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL$ has the same asymptotic behaviour than $\mathbf{A}^{(L)}\left(\mathbf{S}_N\mathbf{S}_N^* \otimes \mathbf{I}_L/NL\right)\mathbf{A}^{(L)*} + \sigma^2\mathbf{I}_{M-L+1}$, In this context, the orthogonal projection matrix $\hat{\mathbf{\Pi}}_N^{(L)}$ onto the eigenspace associated to the $M-L+1-K$ smallest eigenvalues of $\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL$ is a consistent estimate of the orthogonal projection matrix $\mathbf{\Pi}^{(L)}$ on the noise subspace, i.e. the orthogonal complement of $\text{sp}\{\mathbf{a}_{M-L+1}(\theta_1),\ldots,\mathbf{a}_{M-L+1}(\theta_K)\}$. In other words, it holds that

$$\left\|\hat{\mathbf{\Pi}}_N^{(L)} - \mathbf{\Pi}^{(L)}\right\| \to 0 \; a.s. \tag{2.5.7}$$

The traditional pseudo-spectrum estimate $\hat{\eta}_N^{(t)}(\theta)$ defined by

$$\hat{\eta}_N^{(t)}(\theta) = \mathbf{a}_{M-L+1}(\theta)^* \hat{\mathbf{\Pi}}_N^{(L)} \mathbf{a}_{M-L+1}(\theta)$$

thus verifies

$$\sup_{\theta\in[-\pi,\pi]} \left|\hat{\eta}_N^{(t)}(\theta) - \eta(\theta)\right| \xrightarrow[N\to\infty]{a.s.} 0. \tag{2.5.8}$$

where $\eta(\theta) = \mathbf{a}_{M-L+1}(\theta)^*\mathbf{\Pi}^{(L)}\mathbf{a}_{M-L+1}(\theta)$ is the MUSIC pseudo-spectrum. Moreover, the K MUSIC traditional DoA estimates, defined formally, for $k = 1,\ldots,K$, by

$$\hat{\theta}_{k,N}^{(t)} = \underset{\theta\in\mathcal{I}_k}{\operatorname{argmin}}\, \hat{\eta}_N^{(t)}(\theta), \tag{2.5.9}$$

where $\mathcal{I}_k$ is a compact interval containing θ_k and such that $\mathcal{I}_k \cap \mathcal{I}_l = \emptyset$ for $k \neq l$, are consistent, i.e.

$$\hat{\theta}_{k,N}^{(t)} \xrightarrow[N\to\infty]{a.s.} \theta_k. \tag{2.5.10}$$

However, the regime where $M-L+1$ remains fixed while NL converges towards ∞ is not very interesting in practice because the size $M-L+1$ of the subarrays may be much smaller that the number of antennas M, thus reducing the resolution of the method. We therefore study spatial smoothing schemes in regimes where the dimensions $M-L+1$ and NL of matrix $\mathbf{Y}_N^{(L)}$ are of the same order of magnitude and where $\frac{L}{M} \to 0$ in order to keep unchanged the aperture of the array. More precisely, we assume that integers N and L depend on M and that

$$M \to +\infty, N = \mathcal{O}(M^\beta),\ \frac{1}{3} < \beta \le 1,\ c_N = \frac{M-L+1}{NL} \to c \tag{2.5.11}$$

In regime (2.5.11), N thus converges towards ∞ but at a rate that may be much lower than M thus modelling contexts in which N is much smaller than M. As $N \to +\infty$, it also holds that $\frac{M}{NL} \to c$. Therefore, it is clear that $L = \mathcal{O}(M^\alpha)$ where $\alpha = 1-\beta$ verifies with $0 \le \alpha < 2/3$. L may thus converge towards ∞ (even faster than N if $\beta < 1/2$) but condition $\alpha < 2/3$ (or equivalently $\beta > 1/3$) implies that the convergence speed of L to $+\infty$ is not arbitrarily fast. As explained in paragraph 2.5.2, condition $L = \mathcal{O}(M^\alpha)$ with $\alpha < 2/3$ implies that matrix $\mathbf{V}_N^{(L)}$, behaves, in some sense, as a random matrix with i.i.d. entries, and that the results of [33] and [74] obtained in the case $L = 1$ can be extended to asymptotic regime (2.5.11).

As in regime (2.5.11) N depends on M, it could be appropriate to index the various matrices and DoA estimators by integer M rather than by integer N as in definitions (2.5.5) and (2.5.9). However, we prefer to use the index N in the following in order to keep the notations unchanged. We also denote projection matrix $\mathbf{\Pi}^{(L)}$ and pseudo-spectrum $\eta(\theta)$ by $\mathbf{\Pi}_N^{(L)}$ and $\eta_N(\theta)$ because they depend on M. Moreover, in the following, the notation $N \to +\infty$ should be understood as regime (2.5.11) for some $\beta \in (1/3, 1]$.

Main results.

In regime (2.5.11), the empirical covariance matrix $\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL$ is not a good estimate of the true covariance matrix $\mathbf{A}^{(L)}\left(\mathbf{S}_N\mathbf{S}_N^* \otimes \mathbf{I}_L/NL\right)\mathbf{A}^{(L)*} + \sigma^2\mathbf{I}_{M-L+1}$ in the sense that

$$\left\|\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL - \left(\mathbf{A}^{(L)}\left(\mathbf{S}_N\mathbf{S}_N^* \otimes \mathbf{I}_L/NL\right)\mathbf{A}^{(L)*} + \sigma^2\mathbf{I}_{M-L+1}\right)\right\|$$

does not converge towards 0 almost surely. Roughly speaking, this is because the true covariance matrix depends on $\frac{(M-L+1)(M-L)}{2} = O(M^2)$ parameters, and that the number of independent random variables that are available for estimation is equal to MN, which, in regime (2.5.11), is of course not sufficient. Therefore, (2.5.7) is no more valid, and hence, (2.5.10) is questionable. In this chapter, we show that it is possible to generalize the G-MUSIC estimators introduced in [33] and [74] in the case where $L = 1$ to the context of spatial smoothing schemes in regime (2.5.11). In order to explain this informally, we denote by $\mathbf{X}_N$, $\mathbf{Z}_N$, and $\mathbf{B}_N$ the matrices defined by $\mathbf{X}_N = \frac{\mathbf{Y}_N^{(L)}}{\sqrt{NL}}$, $\mathbf{Z}_N = \frac{\mathbf{V}_N^{(L)}}{\sqrt{NL}}$, and $\mathbf{B}_N = \frac{1}{\sqrt{NL}}\mathbf{A}^{(L)}\left(\mathbf{S}_N \otimes \mathbf{I}_L\right)$ (we do not mention that these matrices depend on L in order to simplify the notations), and observe that

$$\mathbf{X}_N = \mathbf{B}_N + \mathbf{Z}_N$$

We denote by $(\mathbf{u}_{k,N})_{k=1,\ldots,K}$ and $(\lambda_{k,N})_{k=1,\ldots,K}$ the non zero eigenvalues of matrix $\mathbf{B}_N\mathbf{B}_N^*$, and by $(\hat{\mathbf{u}}_{k,N})_{k=1,\ldots,M}$ and $(\hat{\lambda}_{k,N})_{k=1,\ldots,M}$ the eigenvalues of matrix $\mathbf{X}_N\mathbf{X}_N^*$. Matrix $\mathbf{X}_N$ coincides with the sum of rank K deterministic matrix $\mathbf{B}_N$ and bloc-Hankel random matrix $\mathbf{Z}_N$ due to the noise, and it is of course of fundamental interest to precise the behaviour of the K largest eigenvalues $(\hat{\lambda}_{k,N})_{k=1,\ldots,K}$ and related eigenvectors $(\hat{\mathbf{u}}_{k,N})_{k=1,\ldots,K}$ in the asymptotic regime (2.5.11). If matrix $\mathbf{Z}_N$ was i.i.d., Theorem 1.2.2 would imply that, under the so-called separation condition

$$\lambda_{K,N} > \sigma^2\sqrt{c}, \text{ for each N large enough} \tag{2.5.12}$$

then, for each $k = 1,\ldots,K$, it would hold that

$$\mathbf{a}_{M-L+1}(\theta)^*\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*\mathbf{a}_{M-L+1}(\theta) = h(\hat{\lambda}_{k,N})\mathbf{a}_{M-L+1}(\theta)^*\mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\mathbf{a}_{M-L+1}(\theta) + o(1) \quad a.s., \tag{2.5.13}$$

for each θ, where function h coincides with function h_{c,σ^2} defined by (1.2.25). This would immediately imply that the traditional pseudo-spectrum estimate $\hat{\eta}_N^{(t)}(\theta)$ would verify

$$\hat{\eta}_N^{(t)}(\theta) = 1 - \sum_{k=1}^{K} h(\hat{\lambda}_{k,N})\mathbf{a}_{M-L+1}(\theta)^*\mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\mathbf{a}_{M-L+1}(\theta) + o(1) \quad a.s., \tag{2.5.14}$$

and that the true MUSIC pseudo-spectrum

$$\begin{aligned}\eta_N(\theta) &= \mathbf{a}_{M-L+1}(\theta)^*\mathbf{\Pi}_N^{(L)}\mathbf{a}_{M-L+1}(\theta)\\ &= 1 - \mathbf{a}_{M-L+1}(\theta)^*\sum_{k=1}^{K}\mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\mathbf{a}_{M-L+1}(\theta)\end{aligned}$$

could be estimated consistently by

$$\hat{\eta}_N(\theta) = (\mathbf{a}_{M-L+1}(\theta))^*\left(\mathbf{I} - \sum_{k=1}^{K}\frac{1}{h\left(\hat{\lambda}_{k,N}\right)}\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*\right)\mathbf{a}_{M-L+1}(\theta) \tag{2.5.15}$$

While matrix $\mathbf{Z}_N$ is of course not i.i.d. as soon as $L > 1$, we prove that the fundamental identity (2.5.13), in principle valid when $\mathbf{Z}_N$ is i.i.d., still holds in the asymptotic regime (2.5.11). For this, we again use

Remark 1, and check the corresponding items. This not only implies (2.5.14) and the consistency of $\hat{\eta}_N(\theta)$ for each θ, but also that

$$\sup_{\theta\in[-\pi,\pi]} \left| \hat{\eta}_N^{(t)}(\theta) - \left(1 - \sum_{k=1}^{K} h(\hat{\lambda}_{k,N}) \mathbf{a}_{M-L+1}(\theta)^* \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \mathbf{a}_{M-L+1}(\theta)\right) \right| \to 0 \quad a.s. \tag{2.5.16}$$

and

$$\sup_{\theta\in[-\pi,\pi]} \left| \hat{\eta}_N(\theta) - \eta_N(\theta) \right| \to 0 \quad a.s. \tag{2.5.17}$$

These uniform consistency properties allow to study the asymptotic behaviour of the traditional MUSIC SS estimates $(\hat{\theta}_{k,N}^{(t)})_{k=1,\dots,K}$ and of the G MUSIC SS estimates $(\hat{\theta}_{k,N})_{k=1,\dots,K}$ defined as the K most significant local minima of $|\hat{\eta}_N(\theta)|$. More precisely, (2.5.16) and (2.5.17) allow to generalize immediately in the asymptotic regime (2.5.11) the proof of Theorem 3 of [33] and the proof of Theorem 5 of [74] (these theorems address the case $L = 1$), and to conclude that, under the separation condition (2.5.12), it holds that:

- $(\hat{\theta}_{k,N}^{(t)})_{k=1,\dots,K}$ and $(\hat{\theta}_{k,N})_{k=1,\dots,K}$ are consistent and verify

$$M(\hat{\theta}_{k,N}^{(t)} - \theta_k) \quad \to \quad 0 \; a.s., \tag{2.5.18}$$

$$M(\hat{\theta}_{k,N} - \theta_k) \quad \to \quad 0 \; a.s. \tag{2.5.19}$$

(2.5.18) and (2.5.19) hold when the DoA $(\theta_k)_{k=1,\dots,K}$ are fixed parameters that do not depend on M and N. In practice, this assumption corresponds to practical situations where the DoA are widely spaced because when the DoA $(\theta_k)_{k=1,\dots,K}$ are fixed, the ratio

$$\frac{\min_{k\neq l} |\theta_k - \theta_l|}{(2\pi)/M}$$

converges towards ∞. Adapting the proof of Theorem 6 of [74], we obtain that:

- If $K = 2$, $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N} \to \mathbf{I}_2$, and if the 2 DoAs scale with M, N is such a way that $\theta_{2,N} - \theta_{1,N} = \mathcal{O}(\frac{1}{M})$, then the G-MUSIC SS estimates still verify (2.5.19) while the traditional MUSIC SS estimates no longer verify (2.5.18)

As in the case $L = 1$, the separation condition (2.5.12) ensures that the K largest eigenvalues of the empirical covariance matrix $(\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*})/NL$ correspond to the K sources, and the signal and noise subspaces can be separated. In order to obtain some insights on this condition, and on the potential benefit of the spatial smoothing, we study the separation condition when M and N converge towards ∞ at the same rate, i.e. when $\frac{M}{N} \to d$, or equivalently that $\beta = 1$ and that L does not scale with N. In this case, it is clear that c coincides with $c = d/L$. Under the assumption that $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}$ converges towards a diagonal matrix $\mathbf{D}$ when N increases, then we establish that the separation condition holds if

$$\lambda_K\left(\mathbf{A}_{M-L+1}^*\mathbf{A}_{M-L+1}\mathbf{D}\right) > \frac{\sigma^2\sqrt{d}}{\sqrt{L}} \tag{2.5.20}$$

for each (M, N) large enough. If $L = 1$, the separation condition introduced in the context of (unsmoothed) G-MUSIC algorithms ([33]) is of course recovered, i.e.

$$\lambda_K\left(\mathbf{A}_M^*\mathbf{A}_M\mathbf{D}\right) > \sigma^2\sqrt{d}$$

If M is large and that L << M, matrix $\mathbf{A}_{M-L+1}^* \mathbf{A}_{M-L+1}$ is close from $\mathbf{A}_M^* \mathbf{A}_M$ and the separation condition is nearly equivalent to

$$\lambda_K \left(\mathbf{A}_M^* \mathbf{A}_M \mathbf{D} \right) > \frac{\sigma^2 \sqrt{d}}{\sqrt{L}}$$

Therefore, it is seen that the use of the spatial smoothing scheme allows to reduce the threshold $\sigma^2 \sqrt{d}$ corresponding to G-MUSIC method without spatial smoothing by the factor $\sqrt{L}$. Therefore, if M and N are the same order of magnitude, our asymptotic analysis allows to predict an improvement of the performance of the G-MUSIC SS methods when L increases provided L << M. If L becomes too large, the above rough analysis is no more justified and the impact of the diminution of the number of antennas becomes dominant, and the performance tends to decrease.

2.5.3 Asymptotic behaviour of the largest eigenvalues and eigenvectors of $\frac{\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}}{NL}$.

In this subsection, N, M, L still satisfy (2.5.11) while K is a fixed integer that does not scale with N. We consider the $(M+L-1) \times NL$ block-Hankel random matrix $\mathbf{V}_N^{(L)}$ defined previously, as well as matrices $\mathbf{Z}_N$ defined by $\mathbf{Z}_N = \frac{1}{\sqrt{NL}} \mathbf{V}_N^{(L)}$, $\mathbf{B}_N = \frac{1}{\sqrt{NL}} \mathbf{A}^{(L)} (\mathbf{S}_N \otimes \mathbf{I}_L)$, and $\mathbf{X}_N = \mathbf{B}_N + \mathbf{Z}_N = \frac{1}{\sqrt{NL}} \mathbf{Y}_N^{(L)}$. The entries of $\mathbf{Z}_N$ have of course variance σ^2/NL. We assume from now on that $K \times N$ matrix $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N}$ verifies

$$\sup_N \| \frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \| < +\infty \tag{2.5.21}$$

a condition which implies that $\mathbf{B}_N$ satisfies

$$\sup_N \| \mathbf{B}_N \| < +\infty \tag{2.5.22}$$

We denote by $\lambda_{1,N} > \lambda_{2,N} \ldots > \lambda_{K,N}$ the non zero eigenvalues of matrix $\mathbf{B}_N \mathbf{B}_N^*$ arranged in decreasing order, and by $(\mathbf{u}_{k,N})_{k=1,\ldots,K}$ the associated eigenvectors. Moreover, we assume that:

Assumption 2.5.1. *The* K *non zero eigenvalues* $(\lambda_{k,N})_{k=1,\ldots,K}$ *of matrix* $\mathbf{B}_N \mathbf{B}_N^*$ *converge towards* $\lambda_1 > \lambda_2 > \ldots > \lambda_K$ *when* $N \to +\infty$.

Here, for ease of exposition, we assume that the eigenvalues $(\lambda_{k,N})_{k=1,\ldots,K}$ have multiplicity 1 and that $\lambda_k \neq \lambda_l$ for $k \neq l$. However, the forthcoming results can be easily adapted if some λ_k coincide.

The purpose of this subsection is to formalize claim (2.5.13), and to establish that the K largest eigenvalues $(\hat{\lambda}_{k,N})_{k=1,\ldots,K}$ of matrix $\mathbf{X}_N \mathbf{X}_N^*$ as well as of their corresponding eigenvectors $(\hat{\mathbf{u}}_{k,N})_{k=1,\ldots,K}$ behave as if the entries of $\mathbf{Z}_N$ were i.i.d.

We denote by $m_c(z)$ the Stieltjes transform of the Marcenko-Pastur distribution μ_{c,σ^2} of parameters (c, σ^2) and by $\tilde{m}_c(z)$ the Stieltjes transform of $c\mu_{c,\sigma^2} + (1-c)\delta_0$. $w_c(z)$ represents function $w_c(z) = \frac{1}{z m_c(z) \tilde{m}_c(z)}$ and $h(z)$ is defined by $h(z) = \frac{w_c(z)^2 - \sigma^4 c}{w_c(z)(w_c(z) + \sigma^2 c)}$. Finally, we denote by $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$ the resolvents of matrices $\mathbf{Z}_N \mathbf{Z}_N^*$ and $\mathbf{Z}_N^* \mathbf{Z}_N$ respectively. Then, the following Proposition holds.

Proposition 2.5.1.

- *(i) The eigenvalue distribution of matrix* $\mathbf{Z}_N \mathbf{Z}_N^*$ *converges almost surely towards the Marcenko-Pastur distribution* μ_{c,σ^2}

- *(ii) For each* $\epsilon > 0$*, almost surely, for* N *large enough, all the eigenvalues of* $\mathbf{Z}_N \mathbf{Z}_N^*$ *belong to* $[\sigma^2(1-\sqrt{c})^2 - \epsilon, \sigma^2(1+\sqrt{c})^2 + \epsilon]$ *if* $c \leq 1$*, and to* $[\sigma^2(1-\sqrt{c})^2 - \epsilon, \sigma^2(1+\sqrt{c})^2 + \epsilon] \cup \{0\}$ *if* $c > 1$.

- *(iii) Moreover, if* $\mathbf{a}_N, \mathbf{b}_N$ *are* $(M-L+1)$*–dimensional deterministic vectors satisfying* $\sup_N(\|\mathbf{a}_N\|, \|\mathbf{b}_N\|) < +\infty$ *, then it holds that for each* $z \in \mathbb{C} - \mathbb{R}^+$

$$\mathbf{a}_N^* \left(\mathbf{Q}_N(z) - m_c(z)\mathbf{I}\right)\mathbf{b}_N \to 0 \ a.s. \tag{2.5.23}$$

Similarly, if $\tilde{\mathbf{a}}_N$ *and* $\tilde{\mathbf{b}}_N$ *are* NL*–dimensional deterministic vectors verifying* $\sup_N(\|\tilde{\mathbf{a}}_N\|, \|\tilde{\mathbf{b}}_N\|) < +\infty$, *then for each* $z \in \mathbb{C} - \mathbb{R}^+$*, it holds that*

$$\tilde{\mathbf{a}}_N^* \left(\tilde{\mathbf{Q}}_N(z) - \tilde{m}_c(z)\mathbf{I}\right)\tilde{\mathbf{b}}_N \to 0 \ a.s. \tag{2.5.24}$$

Moreover, for each $z \in \mathbb{C} - \mathbb{R}^+$*, it holds that*

$$\mathbf{a}_N^* \left(\mathbf{Q}_N(z)\mathbf{Z}_N\right)\tilde{\mathbf{b}}_N \to 0 \ a.s. \tag{2.5.25}$$

Finally, for each $\epsilon > 0$*, convergence properties (2.5.23, 2.5.24, 2.5.25) hold uniformly w.r.t. z on each compact subset of* $\mathbb{C}^* - [\sigma^2(1-\sqrt{c})^2, \sigma^2(1+\sqrt{c})^2]$.

Proof. Proposition 2.5.1 follows directly from Theorem 1.3.1 and from the results established in the course of the proof of Theorem 2.2.1. In order to explain this, we denote by $\mathbf{W}_N$ the $NL \times (M-L+1)$ matrix defined by

$$\mathbf{W}_N = \frac{1}{\sqrt{c_N}}\mathbf{Z}_N^*$$

$\mathbf{W}_N$ can be written as $\mathbf{W}_N = (\mathbf{W}_N^{(1)T}, \ldots, \mathbf{W}_N^{(N)T})^T$ where matrices $(\mathbf{W}_N^{(n)})_{n=1,\ldots,N}$ are independent identically distributed $L \times (M-L+1)$ Hankel matrices built from i.i.d. standard complex Gaussian sequences with variance $\frac{\sigma^2}{M-L+1}$. As $\mathbf{Z}_N^*\mathbf{Z}_N = c_N\mathbf{W}_N\mathbf{W}_N^*$ and that $\mathbf{Z}_N\mathbf{Z}_N^* = c_N\mathbf{W}_N^*\mathbf{W}_N$, the use of Theorem 1.3.1 when (M,N) is exchanged by $(N, M-L+1)$ immediately implies that item (i) holds. More precisely, the empirical eigenvalue distributions of $\mathbf{W}_N\mathbf{W}_N^*$ and of $\mathbf{W}_N^*\mathbf{W}_N$ converge almost surely towards μ_{c^{-1},σ^2} and $\mu_{c,\sigma^2 c^{-1}}$ respectively. Therefore, the empirical eigenvalue distributions of $\mathbf{Z}_N^*\mathbf{Z}_N$ and of $\mathbf{Z}_N\mathbf{Z}_N^*$ converge towards $\mu_{c^{-1},\sigma^2 c}$ and μ_{c,σ^2} respectively as expected. (ii) holds for similar reasons. Item (iii) eventually follows from the application of (2.2.2, 2.2.3, 2.2.4) to matrix $\mathbf{W}_N$ as well as from the uniformity of the convergence of (2.2.2, 2.2.3, 2.2.4) on the compact subsets $\mathbb{C}^* - [\sigma^2(1-\sqrt{c^{-1}})^2, \sigma^2(1+\sqrt{c^{-1}})^2]$.

Using Remark 1, we obtain immediately the following Theorem.

Theorem 2.5.1. *We denote by* s*,* $0 \le s \le K$*, the largest integer for which*

$$\lambda_s > \sigma^2\sqrt{c} \tag{2.5.26}$$

Then, for $k = 1, \ldots, s$*, it holds that*

$$\hat{\lambda}_{k,N} \xrightarrow[N\to\infty]{a.s.} \rho_k = \phi(\lambda_k) = \frac{(\lambda_k + \sigma^2)(\lambda_k + \sigma^2 c)}{\lambda_k} > \sigma^2(1+\sqrt{c})^2. \tag{2.5.27}$$

Moreover, for $k = s+1, \ldots, K$*, it holds that*

$$\hat{\lambda}_{k,N} \to \sigma^2(1+\sqrt{c})^2 \ a.s. \tag{2.5.28}$$

Finally, for all deterministic sequences of unit vectors $(\mathbf{d}_{1,N})$*,* $(\mathbf{d}_{2,N})$*, we have for* $k = 1, \ldots, s$

$$\mathbf{d}_{1,N}^*\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*\mathbf{d}_{2,N} = h(\rho_k)\mathbf{d}_{1,N}^*\mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\mathbf{d}_{2,N} + o(1) \quad a.s., \tag{2.5.29}$$

2.5.4 Derivation of a consistent G-MUSIC method.

We now use the results of section 2.5.3. We recall that $(\hat{\lambda}_{k,N})_{k=1,\dots,M-L+1}$ and $(\hat{\mathbf{u}}_{k,N})_{k=1,\dots,M-L+1}$ represent the eigenvalues and eigenvectors of the empirical covariance matrix $\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL$, and that $(\lambda_{k,N})_{k=1,\dots,K}$ and $(\mathbf{u}_{k,N})_{k=1,\dots,K}$ are the non zero eigenvalues and corresponding eigenvectors of $\frac{1}{L}\mathbf{A}^{(L)}\left(\mathbf{S}_N\mathbf{S}_N^*/N\otimes\mathbf{I}_L\right)\mathbf{A}^{(L)*}$. We recall that $\mathbf{\Pi}_N^{(L)}$ represents the orthogonal projection matrix onto the noise subspace, i.e. the orthogonal complement of the space generated by vectors $(\mathbf{a}_{M-L+1}(\theta_k))_{k=1,\dots,K}$ and that $\eta_N(\theta)$ is the corresponding MUSIC pseudo-spectrum

$$\eta_N(\theta) = \mathbf{a}_{M-L+1}(\theta)^*\,\mathbf{\Pi}_N^{(L)}\,\mathbf{a}_{M-L+1}(\theta)$$

Theorem 2.5.1 allows to generalize immediately the results of [33] and [74] concerning the consistency of G-MUSIC and MUSIC DoA estimators in the case $L = 1$. More precisely:

Theorem 2.5.2. *Assume that the* K *non zero eigenvalues* $(\lambda_{k,N})_{k=1,\dots,K}$ *converge towards deterministic terms* $\lambda_1 > \lambda_2 > \ldots > \lambda_K$ *and that*

$$\lambda_K > \sigma^2\sqrt{c} \tag{2.5.30}$$

Then, the estimator $\hat{\eta}_N(\theta)$ *of the pseudo-spectrum* $\eta_N(\theta)$ *defined by*

$$\hat{\eta}_N(\theta) = (\mathbf{a}_{M-L+1}(\theta))^*\left(\mathbf{I}-\sum_{k=1}^{K}\frac{1}{h\left(\hat{\lambda}_{k,N}\right)}\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*\right)\mathbf{a}_{M-L+1}(\theta) \tag{2.5.31}$$

verifies

$$\sup_{\theta\in[-\pi,\pi]}\left|\hat{\eta}_N(\theta)-\eta_N(\theta)\right| \xrightarrow[N\to\infty]{a.s.} 0, \tag{2.5.32}$$

(2.5.31) is a direct consequence of (2.5.29) and (2.5.27). (2.5.32) can be proved as Proposition 1 in [33]. We notice that the proof of this Proposition uses extensively Lemma 5 in [33], which, in the context of the present paper has to be replaced by item (iii) of Proposition 2.5.1.

In order to obtain some insights on condition (2.5.30) and on the potential benefits of the spatial smoothing, we explicit the separation condition (2.5.30) when M and N converge towards ∞ at the same rate, i.e. when $\frac{M}{N}\to d$, or equivalently that $\beta = 1$ and that L does not scale with N. In this case, it is clear that c coincides with $c = d/L$. It is easily seen that

$$\frac{1}{L}\mathbf{A}^{(L)}\left(\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}\otimes\mathbf{I}_L\right)\mathbf{A}^{(L)*} = (M-L+1/M)\ \mathbf{A}_{M-L+1}\left(\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}\bullet\mathbf{A}_L^T\overline{\mathbf{A}}_L\right)\mathbf{A}_{M-L+1}^* \tag{2.5.33}$$

where $\bullet$ represents the Hadamard (i.e. element wise) product of matrices, and where $\overline{\mathbf{B}}$ stands for the complex conjugation operator of the elements of matrix $\mathbf{B}$. If we assume that $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}$ converges towards a diagonal matrix $\mathbf{D}$ when N increases, then $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}\bullet(\mathbf{A}_L^T\overline{\mathbf{A}}_L)$ converges towards the diagonal matrix $\mathbf{D}\bullet\mathrm{Diag}\left(\mathbf{A}_L^T\overline{\mathbf{A}}_L\right)=\mathbf{D}$. Therefore, $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}\bullet(\mathbf{A}_L^T\overline{\mathbf{A}}_L)\simeq\mathbf{D}$ when N is large enough. Using that $\frac{L}{M}\to 0$, we obtain that the separation condition is nearly equivalent to

$$\lambda_K\left(\mathbf{A}_{M-L+1}\mathbf{D}\,\mathbf{A}_{M-L+1}^*\right) > \frac{\sigma^2\sqrt{d}}{\sqrt{L}}$$

or to

$$\lambda_K\left(\mathbf{A}_{M-L+1}^*\mathbf{A}_{M-L+1}\mathbf{D}\right) > \frac{\sigma^2\sqrt{d}}{\sqrt{L}} \tag{2.5.34}$$

for each (M,N) large enough. If $L = 1$, the separation condition introduced in the context of (unsmoothed) G-MUSIC algorithms ([33], [74]) is of course recovered, i.e.

$$\lambda_K\left(\mathbf{A}_M^*\mathbf{A}_M\mathbf{D}\right) > \sigma^2\sqrt{d}$$

for each (M,N) large enough. If M is large and that $L << M$, matrix $\mathbf{A}^*_{M-L+1}\mathbf{A}_{M-L+1}$ is close from $\mathbf{A}^*_M\mathbf{A}_M$ and the separation condition is nearly equivalent to

$$\lambda_K\left(\mathbf{A}^*_M\mathbf{A}_M\mathbf{D}\right) > \frac{\sigma^2\sqrt{d}}{\sqrt{L}}$$

Therefore, it is seen that the use of the spatial smoothing scheme allows to reduce the threshold $\sigma^2\sqrt{d}$ corresponding to G-MUSIC method without spatial smoothing by the factor $\sqrt{L}$. Hence, if M and N are the same order of magnitude, our asymptotic analysis allows to predict an improvement of the performance of the G-MUSIC methods based on spatial smoothing when L increases provided $L << M$. If L becomes too large, the above rough analysis is no more justified and the impact of the diminution of the number of antennas becomes dominant, and the performance tends to decrease. This analysis is sustained by the numerical simulations presented in subsection 2.5.5.

We define the DoA G-MUSIC SS estimates $(\hat{\theta}_{k,N})_{k=1,\ldots,K}$ by

$$\hat{\theta}_{k,N} = \underset{\theta\in\mathscr{I}_k}{\operatorname{argmin}}\left|\hat{\eta}_N(\theta)\right|, \tag{2.5.35}$$

where $\mathscr{I}_k$ is a compact interval containing θ_k and such that $\mathscr{I}_k \cap \mathscr{I}_l = \emptyset$ for $k \neq l$. As in [33], the uniform consistency (2.5.32) as well as the particular structure of directional vectors $\mathbf{a}_{M-L+1}(\theta)$ imply the following result which can be proved as Theorem 3 of [33]

Theorem 2.5.3. *Under condition (2.5.30), the DoA G-MUSIC SS estimates* $(\hat{\theta}_{k,N})_{k=1,\ldots,K}$ *verify*

$$M\left(\hat{\theta}_{k,N} - \theta_k\right) \to 0 \ a.s. \tag{2.5.36}$$

for each $k = 1,\ldots,K$.

Remark 5. *We remark that under the extra assumption that* $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}$ *converges towards a diagonal matrix, [33] (see also [74] for more general matrices* $\mathbf{S}$*) proved when* $L = 1$ *that* $M^{3/2}\left(\hat{\theta}_{k,N} - \theta_k\right)$ *converges in distribution towards a Gaussian distribution. It would be interesting to generalize the results of [33] and [74] to the G-MUSIC estimators with spatial smoothing in the asymptotic regime (2.5.11).*

Theorem 2.5.1 also allows to generalize immediately the results of [74] concerning the consistency of the traditional estimates $(\hat{\theta}^{(t)}_{k,N})_{k=1,\ldots,K}$ in the case $L = 1$. In particular, while the traditional estimate $\hat{\eta}^{(t)}_N(\theta)$ of the pseudo-spectrum is not consistent, it is shown in [74] (see Theorem 5) that if $L = 1$, then the arguments of its local minima $(\hat{\theta}^{(t)}_{k,N})_{k=1,\ldots,K}$ are consistent and verify

$$M\left(\hat{\theta}^{(t)}_{k,N} - \theta_k\right) \to 0 \ a.s. \tag{2.5.37}$$

for each $k = 1,\ldots,K$ if the separation condition is verified. The proof of Theorem 5 in [74] can be immediately adapted to the context of the present paper. For this, it is sufficient to follow the proof of [74], and to use Theorem 2.5.1, as well as the uniform consistency property

$$\sup_{\theta\in[-\pi,\pi]}\left|\hat{\eta}^{(t)}_N(\theta) - \left(1 - \sum_{k=1}^{K} h(\rho_k)\mathbf{a}_{M-L+1}(\theta)^*\mathbf{u}_{k,N}\mathbf{u}^*_{k,N}\mathbf{a}_{M-L+1}(\theta)\right)\right| \to 0 \quad a.s. \tag{2.5.38}$$

which can be proved in the same way that (2.5.32). We note that, as $\hat{\lambda}_{k,N} \to \rho_k$, then (2.5.38) and (2.5.16) are equivalent. Therefore, the following result holds.

Theorem 2.5.4. *Under condition (2.5.30), the DoA traditional MUSIC SS estimates* $(\hat{\theta}^{(t)}_{k,N})_{k=1,\dots,K}$ *verify*

$$M\left(\hat{\theta}^{(t)}_{k,N} - \theta_k\right) \to 0 \ a.s. \tag{2.5.39}$$

for each $k = 1, \dots, K$.

Remark 6. *It is established in [74] in the case* $L = 1$ *that if* $\frac{S_N S_N^*}{N}$ *converges towards a diagonal matrix, then* $M^{3/2}\left(\hat{\theta}^{(t)}_{k,N} - \theta_k\right)$ *has a Gaussian behaviour, and that the corresponding variance coincides with the asymptotic variance of* $M^{3/2}\left(\hat{\theta}_{k,N} - \theta_k\right)$. *In particular, if* $L = 1$, *the asymptotic performance of MUSIC and G-MUSIC estimators coincide. It would be interesting to check whether this result still holds true for the MUSIC and G-MUSIC estimators with spatial smoothing.*

Theorems 2.5.2 and 2.5.3 as well as (2.5.37) are valid when the DoAs $(\theta_k)_{k=1,\dots,K}$ are fixed parameters, i.e. do not depend on M and N. Therefore, the ratio

$$\frac{\min_{k \neq l} |\theta_k - \theta_l|}{(2\pi)/M}$$

converges towards $+\infty$. In practice, this context is able to model practical situations in which $\sup_{k \neq l} |\theta_k - \theta_l|$ is significantly larger than the aperture of the array. In the case $L = 1$, [74] also addressed the case where the DoA's $(\theta_{k,N})_{k=1,\dots,K}$ depend on N, M and verify $\theta_{k,N} - \theta_{l,N} = \mathcal{O}(\frac{1}{M})$. This context allows to capture practical situations in which the DoA's are spaced of the order of a beamwidth. In order to simplify the calculations, [74] considered the case $K = 2$, $\theta_{2,N} = \theta_{1,N} + \frac{\alpha}{N}$ and where matrix $\frac{S_N S_N^*}{N} \to \mathbf{I}_2$. However, the results can be generalized easily to more general situations. It is shown in [74] that the G-MUSIC estimates still verify (2.5.39), but that, in general, $M\left(\hat{\theta}^{(t)}_{k,N} - \theta_k\right)$ does not converge towards 0. The results of [74] can be generalized immediately to the context of G-MUSIC estimators with spatial smoothing in regime (2.5.11). For this, we have to assume that $\theta_{2,N} = \theta_{1,N} + \frac{\kappa}{M}$ (in [74], M and N are of the same order of magnitude so that the assumptions $\theta_{2,N} = \theta_{1,N} + \frac{\alpha}{N}$ and $\theta_{2,N} = \theta_{1,N} + \frac{\kappa}{M}$ are equivalent), and to follow the arguments of section 3 in [74]. The conclusion of this discussion is the following Theorem.

Theorem 2.5.5. *Assume* $K = 2$, $\theta_{2,N} = \theta_{1,N} + \frac{\kappa}{M}$, *and that* $\frac{S_N S_N^*}{N} \to \mathbf{I}_2$. *If the separation condition*

$$1 - |\mathrm{sinc}\,\kappa/2| > \sigma^2 c \tag{2.5.40}$$

holds, then the G-MUSIC SS estimates $(\hat{\theta}_{k,N})_{k=1,2}$ *defined by*

$$\hat{\theta}_{k,N} = \operatorname*{argmin}_{\theta \in \mathcal{I}_{k,N}} \left|\hat{\eta}_N(\theta)\right|, \tag{2.5.41}$$

where $\mathcal{I}_{k,N} = [\theta_{k,N} - \frac{\kappa - \epsilon}{2N}, \theta_{k,N} + \frac{\kappa - \epsilon}{2N}]$ *for* ϵ *small enough, verify*

$$M\left(\hat{\theta}_{k,N} - \theta_{k,N}\right) \to 0 \ a.s. \tag{2.5.42}$$

In general, the traditional MUSIC SS estimates defined by (2.5.41) when the G-MUSIC estimate $\hat{\eta}_N(\theta)$ *is replaced by the traditional spectrum estimate* $\hat{\eta}^{(t)}_N(\theta)$ *are such that* $M\left(\hat{\theta}^{(t)}_{k,N} - \theta_{k,N}\right)$ *does not converge towards* 0.

2.5.5 Numerical examples

In this section, we provide numerical simulations illustrating the results given in the previous sections. We first consider 2 closely spaced sources with DoAs $\theta_1 = 0$ and $\theta_2 = \frac{\pi}{2M}$, and we assume that $M = 160$ and $N = 20$. The $2 \times N$ signal matrix is obtained by normalizing a realization of a random matrix with $\mathcal{N}_{\mathbb{C}}(0,1)$ i.i.d. entries in such a way that the 2 source signals have power 1. The signal to noise ratio is thus equal to $\mathrm{SNR} = 1/\sigma^2$. Table 2.1 provides the minimum value of SNR for which the separation condition, in its finite length version (i.e. when the limits $(\lambda_k)_{k=1,\ldots,K}$ and c are replaced by $(\lambda_{k,N})_{k=1,\ldots,K}$ and c_N respectively) holds, i.e.

$$(\sigma^2)^{-1} = \frac{1}{\lambda_{K,N}} \sqrt{(M-L+1)/NL}$$

It is seen that the minimal SNR first decreases but that it increases if L is large enough. This confirms the discussion of the previous paragraph on the effect of L on the separation condition.

L	2	4	8	16	32	64	96	128
SNR	33.46	30.30	27.46	25.31	24.70	28.25	36.11	51.52

Table 2.1 – Minimum value of SNR for separation condition

In figure 2.8, we represent the mean-square errors of the G-MUSIC SS estimator $\hat{\theta}_1$ for $L = 2, 4, 8, 16$ versus SNR. The corresponding Cramer-Rao bounds are also represented. As expected, it is seen that the performance tends to increase with L until $L = 16$. In figure 2.9, L is equal to 16, 32, 64, 96, 128.

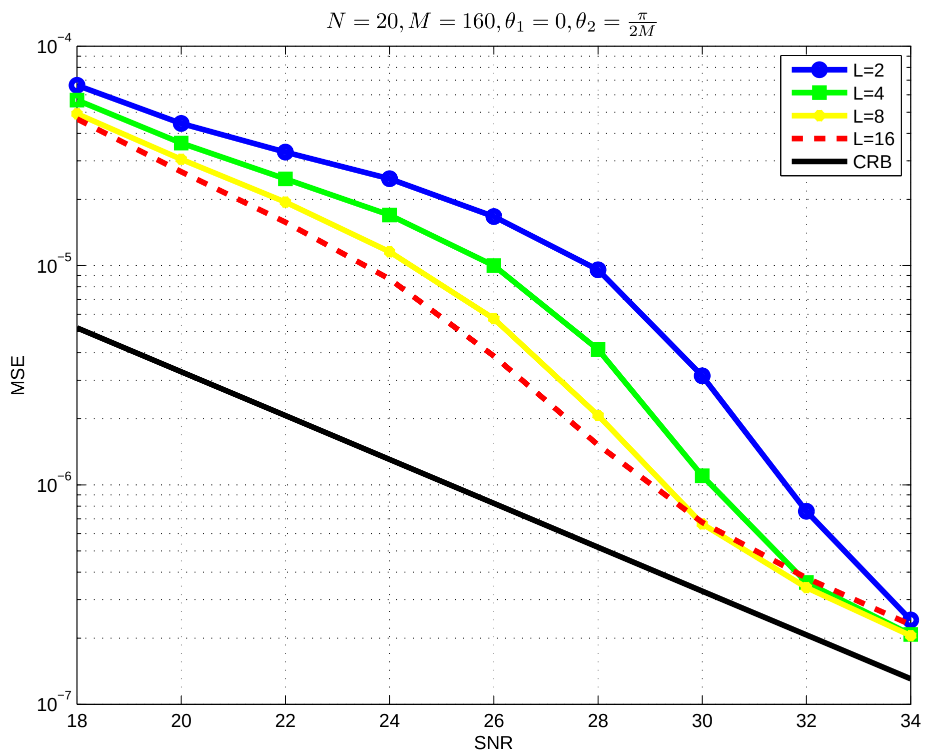

Figure 2.8 – Empirical MSE of G-MUSIC SS estimator $\hat{\theta}_1$ versus SNR

For $L = 32$, it is seen that the MSE tends to degrade at high SNR w.r.t. $L = 16$, while the performance severely degrades for larger values of L.

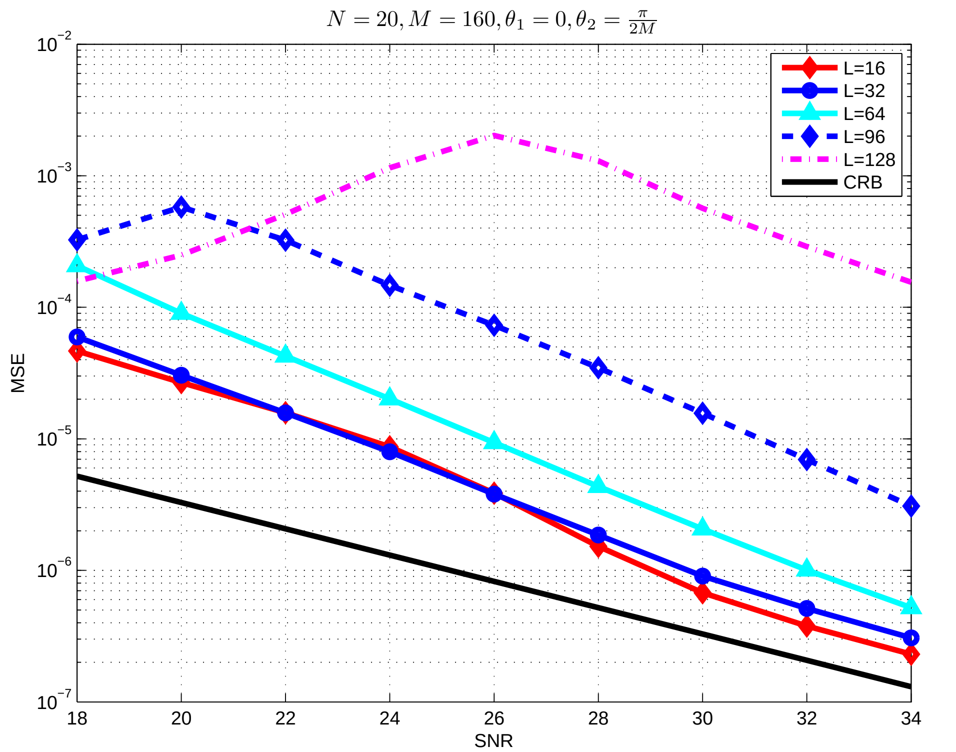

Figure 2.9 – Empirical MSE of G-MUSIC SS estimator $\hat{\theta}_1$ versus SNR

In Figure 2.10, parameter L is equal to 16. We compare the performance of G-MUSIC SS with the standard MUSIC method with spatial smoothing. We also represent the MSE provided by G-MUSIC and MUSIC for $L = 1$. The standard unsmoothed MUSIC method of course completely fails, while the use of the G-MUSIC SS provides a clear improvement of the performance w.r.t. MUSIC SS and unsmoothed G-MUSIC.

We finally consider the case $L = 128$, and compare as above G-MUSIC SS, MUSIC SS, unsmoothed G-MUSIC and unsmoothed MUSIC. G-MUSIC SS completely fails because L and M are of the same order of magnitude. Theorem 2.5.2 is thus no more valid, and the pseudo-spectrum estimate is not consistent.

We now consider 2 widely spaced sources with DoAs $\theta_1 = 0$ and $\theta_2 = 5\frac{2\pi}{M}$, and keep the same parameters as above. We consider the case $L = 16$, and represent in Fig. 2.12 the performance of MUSIC, G-MUSIC, MUSIC-SS, and G-MUSIC-SS. It is first observed that, in contrast with the case of closely spaced DoAs, MUSIC-SS and G-MUSIC-SS have the same performance when the SNR is above the threshold 6 dB. This is in accordance with Theorem 2.5.4, and tends to indicate that, as in the case $L = 1$, if $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N}$ converges towards a diagonal matrix, then the asymptotic performance of G-MUSIC-SS and MUSIC-SS coincide (see Remark 6). The comparison between the methods with and without spatial smoothing also confirm that the use of spatial smoothing schemes allow to improve the performance.

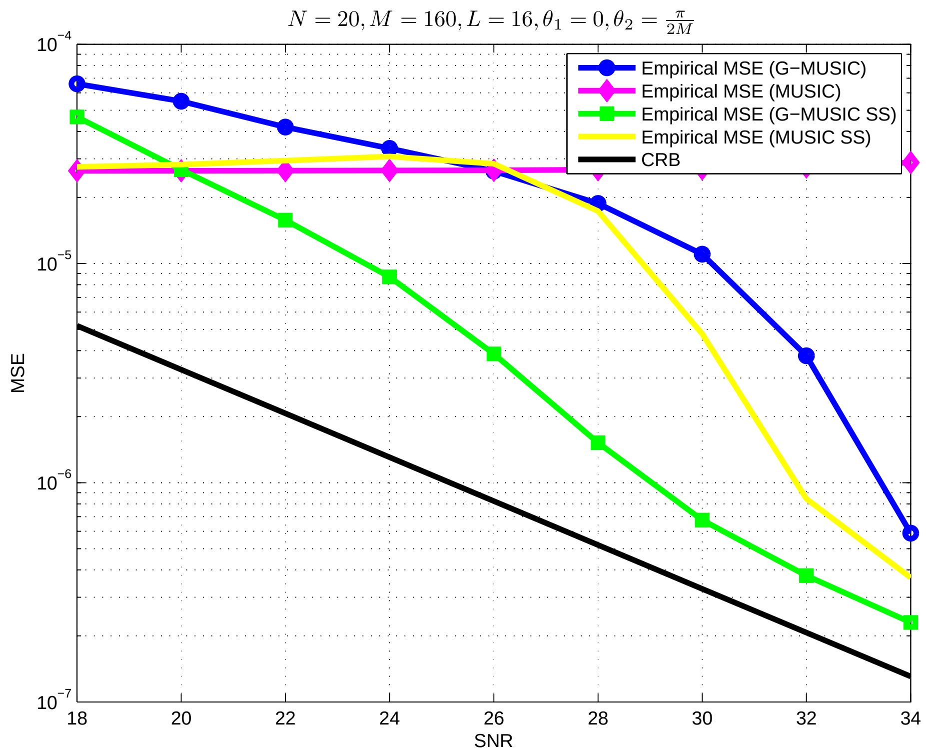

Figure 2.10 – Empirical MSE of different estimators of θ_1 when L=16

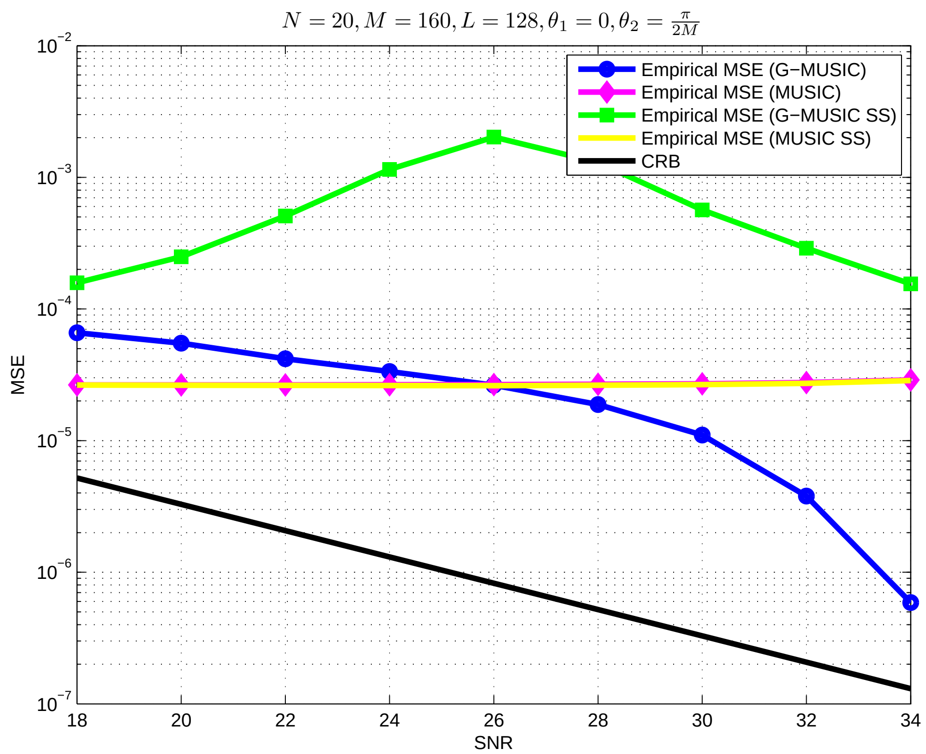

Figure 2.11 – Empirical MSE of different estimators of θ_1 when L=128

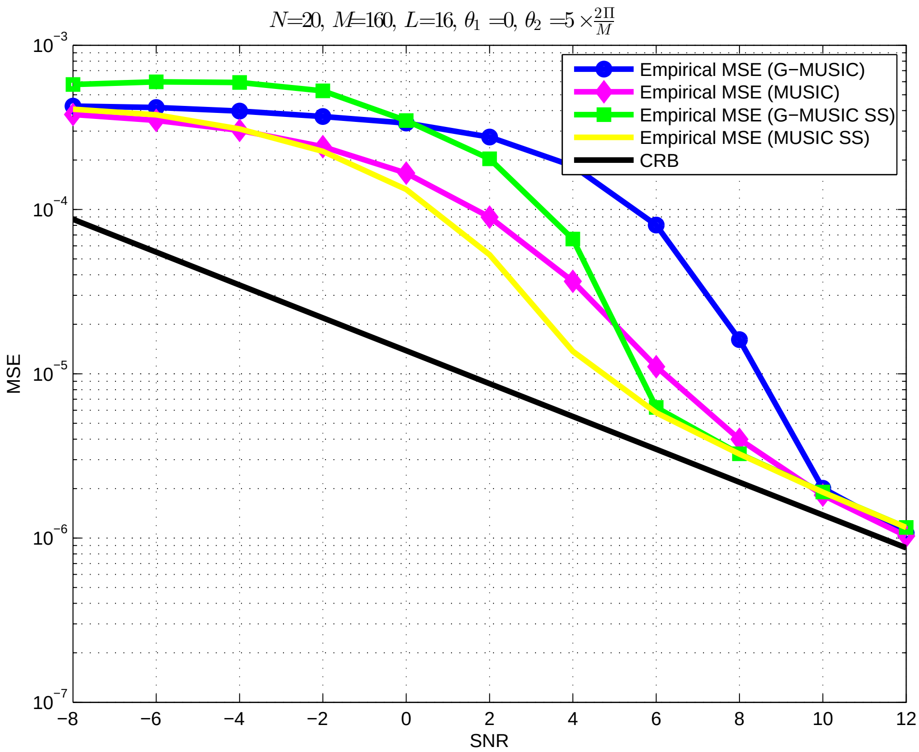

Figure 2.12 – Empirical MSE of different estimators of θ_1 when L=16 and widely spaced DoAs

Chapter 3

Complex Gaussian information plus noise models: the deterministic equivalents

Before introducing the content of this chapter devoted to the study of information plus noise block-Hankel large random matrix models, we have to define some useful notations and tools.

3.1 Notations and useful tools.

3.1.1 Complex numbers and nice polynomials.

$\mathbb{C}^+$ denotes the set of complex numbers with strictly positive imaginary parts. The conjugate of a complex number z is denoted by z^* or $\overline{z}$ depending on the context.

A nice polynomial is a polynomial with positive coefficients whose degree and coefficients do not depend on the dimensions L, M, N of the matrices under consideration.

3.1.2 Matrix notations and Toeplitzification operators

If $\mathbf{A}$ is a matrix, $\|\mathbf{A}\|$ represents the spectral norm of matrix $\mathbf{A}$, and $\mathbf{A}^*$ (resp. $\mathbf{A}^T$) denotes the conjugate transpose (resp. transpose) of $\mathbf{A}$. If $\mathbf{A}$ and $\mathbf{B}$ are 2 matrices, $\mathbf{A} \otimes \mathbf{B}$ represents the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$, i.e the block matrix whose block (i, j) is $\mathbf{A}_{i,j}\mathbf{B}$.

If $\mathbf{A}$ is a ML × ML matrix, we denote $\mathbf{A}_{i_1,i_2}^{m_1,m_2}$ the entry $(i_1 + (m_1 - 1)L, i_2 + (m_2 - 1)L)$ of matrix $\mathbf{A}$, while $\mathbf{A}^{m_1,m_2}$ represents the L × L matrix $\mathbf{A}_{1\le(i_1,i_2)\le L}^{m_1,m_2}$. We also denote by $\hat{\mathbf{A}}$ the L × L matrix defined by

$$\hat{\mathbf{A}} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{A}^{m,m} \tag{3.1.1}$$

It is clear that $\|\hat{\mathbf{A}}\| \le \|\mathbf{A}\|$, and that for each pair $(\mathbf{b}, \mathbf{c})$ of L–dimensional vectors, it holds that

$$\mathbf{b}^*\hat{\mathbf{A}}\,\mathbf{c} = \frac{1}{M}\mathrm{Tr}\left(\mathbf{A}(\mathbf{I}_M \otimes \mathbf{c}\mathbf{b}^*)\right) \tag{3.1.2}$$

For each $1 \le i \le L$ and $1 \le m \le M$, $\mathbf{f}_i^m$ represents the vector of the canonical basis of $\mathbb{C}^{ML}$ whose non zero component is located at index $i + (m - 1)L$, $\mathbf{f}^m$ and $\mathbf{f}_i$ represent the vectors of the canonical basis of $\mathbb{C}^M$ and $\mathbb{C}^L$ respectively. If $1 \le j \le N$, $\mathbf{e}_j$ is the j^{th}-vector of the canonical basis on $\mathbb{C}^N$.

If $\mathbf{A}$ is a square matrix, Re($\mathbf{A}$) and Im($\mathbf{A}$) represent the Hermitian matrices $\mathrm{Re}(A) = \frac{\mathbf{A}+\mathbf{A}^*}{2}$ and $\mathrm{Im}(\mathbf{A}) = \frac{\mathbf{A}-\mathbf{A}^*}{2i}$ respectively.

If $(\mathbf{A}_N)_{N\geq 1}$ (resp. $((\mathbf{b}_N)_{N\geq 1})$) is a sequence of matrices (resp. vectors) whose dimensions increase with N, $(\mathbf{A}_N)_{N\geq 1}$ (resp. $((\mathbf{b}_N)_{N\geq 1})$) is said to be uniformly bounded if $\sup_{N\geq 1}\|\mathbf{A}_N\| < +\infty$ (resp. $\sup_{N\geq 1}\|\mathbf{b}_N\| < +\infty$).

We now introduce certain Toeplitzification operators that will play an important role in the following:

Definition 1.

- *If* $\mathbf{A}$ *is a* $K\times K$ *Toeplitz matrix, we denote by* $(\mathbf{A}(k))_{k=-(K-1),\ldots,K-1}$ *the sequence such that* $\mathbf{A}_{k,l} = \mathbf{A}(k-l)$.
- *For any integer* K, J_K *is the* $K\times K$ *"shift" matrix defined by* $(J_K)_{i,j} = \delta(j-i=1)$. *In order to short the notations, matrix* J_K^* *is denoted by* J_K^{-1}, *although* J_K *is of course not invertible.*
- *For any* $PK\times PK$ *block matrix* $\mathbf{A}$ *with* $K\times K$ *blocks* $(\mathbf{A}^{p_1,p_2})_{1\leq(p_1,p_2)\leq P}$, *we define* $(\tau^{(P)}(\mathbf{A})(k))_{k=-(K-1),\ldots,K-1}$ *as the sequence*

$$\tau^{(P)}(\mathbf{A})(k) = \frac{1}{PK}\mathrm{Tr}\left[\mathbf{A}(\mathbf{I}_P\otimes\mathbf{J}_K^k)\right] = \frac{1}{PK}\sum_{i-j=k}\sum_{p=1}^{P}\mathbf{A}_{i,j}^{(p,p)} = \frac{1}{PK}\sum_{p=1}^{P}\sum_{u=1}^{K}\mathbf{A}_{k+u,u}^{p,p}\,\mathbb{1}_{1\leq k+u\leq K} \tag{3.1.3}$$

- *For any* $PK\times PK$ *block matrix* $\mathbf{A}$ *and for 2 integers* R *and* Q *such that* $Q\leq K$, *matrix* $\mathcal{T}_{R,Q}^{(P)}(\mathbf{A})$ *represents the* $R\times R$ *Toeplitz matrix given by*

$$\mathcal{T}_{R,Q}^{(P)}(\mathbf{A}) = \sum_{q=-(Q-1)}^{Q-1}\tau^{(P)}(\mathbf{A})(q)\,\mathbf{J}_R^{*q} \tag{3.1.4}$$

In other words, for $(i,j)\in\{1,2,\ldots,R\}$, *it holds that*

$$\left(\mathcal{T}_{R,Q}^{(P)}(\mathbf{A})\right)_{i,j} = \tau^{(P)}(\mathbf{A})(i-j)\,\mathbb{1}_{|i-j|\leq Q-1} \tag{3.1.5}$$

When $P=1$, *sequence* $(\tau^{(1)}(\mathbf{A})(k))_{k=-(K-1),\ldots,K-1}$ *and matrix* $\mathcal{T}_{R,Q}^{(1)}(\mathbf{A})$ *are denoted by* $(\tau(\mathbf{A})(k))_{k=-(K-1),\ldots,K-1}$ *and matrix* $\mathcal{T}_{R,Q}(\mathbf{A})$ *in order to simplify the notations. We note that if* $\mathbf{A}$ *is a* $PK\times PK$ *block matrix matrix, then, it holds that*

$$\tau^{(P)}(\mathbf{A})(k) = \tau(\hat{\mathbf{A}})(k) \tag{3.1.6}$$

for each k, *where we recall that matrix* $\hat{\mathbf{A}}$ *is defined by* $\hat{\mathbf{A}} = \frac{1}{P}\sum_{p=1}^{P}\mathbf{A}^{(p,p)}$.

The reader may check that the following straightforward identities hold:

- If $\mathbf{A}$ is a $R\times R$ Toeplitz matrix, for any $R\times R$ matrix $\mathbf{B}$, it holds that

$$\frac{1}{R}\mathrm{Tr}(\mathbf{AB}) = \sum_{k=-(R-1)}^{R-1}\mathbf{A}(-k)\tau(\mathbf{B})(k) = \frac{1}{R}\mathrm{Tr}\left(\mathbf{A}\mathcal{T}_{R,R}(\mathbf{B})\right) \tag{3.1.7}$$

- If $\mathbf{A}$ and $\mathbf{B}$ are both $R\times R$ matrices, and if $Q\leq R$, then,

$$\frac{1}{R}\mathrm{Tr}\left(\mathcal{T}_{R,Q}(\mathbf{A})\mathbf{B}\right) = \sum_{q=-(Q-1)}^{Q-1}\tau(\mathbf{A})(-q)\,\tau(\mathbf{B})(q) = \frac{1}{R}\mathrm{Tr}\left(\mathbf{A}\mathcal{T}_{R,Q}(\mathbf{B})\right) \tag{3.1.8}$$

- If $\mathbf{A}$ is a $PK\times PK$ matrix, if $\mathbf{B}$ is a $R\times R$ matrix, and if $R\geq Q$ and $Q\leq K$, then it holds that

$$\frac{1}{R}\mathrm{Tr}\left(\mathbf{B}\mathcal{T}_{R,Q}^{(P)}(\mathbf{A})\right) = \sum_{k=-(Q-1)}^{Q-1}\tau(\mathbf{B})(k)\,\tau^{(P)}(\mathbf{A})(-k) \tag{3.1.9}$$

  Moreover when $Q=K$, we have

$$\frac{1}{R}\mathrm{Tr}\left(\mathbf{B}\mathcal{T}_{R,K}^{(P)}(\mathbf{A})\right) = \frac{1}{PK}\mathrm{Tr}\left[\left(\mathbf{I}_P\otimes\mathcal{T}_{K,K}(\mathbf{B})\right)\mathbf{A}\right] \tag{3.1.10}$$

- If $\mathbf{C}$ is a PK × PK matrix, $\mathbf{B}$ is a K × K matrix and $\mathbf{D}, \mathbf{E}$ R × R matrices with K ≤ R, then, it holds that

$$\frac{1}{\mathrm{K}}\mathrm{Tr}\left[\mathbf{B}\mathscr{T}_{\mathrm{K,K}}\left(\mathbf{D}\mathscr{T}^{(\mathrm{P})}_{\mathrm{R,K}}(\mathbf{C})\mathbf{E}\right)\right] = \frac{1}{\mathrm{PK}}\mathrm{Tr}\left[\mathbf{C}\left(\mathbf{I}_{\mathrm{P}} \otimes \mathscr{T}_{\mathrm{K,K}}[\mathbf{E}\mathscr{T}_{\mathrm{R,K}}(\mathbf{B})\mathbf{D}]\right)\right] \tag{3.1.11}$$

The above various Toeplitzified matrices have integral representations that allow to simplify a number of issues. In order to introduce these representations, for each integer K, we define $\mathbf{a}_{\mathrm{K}}(\nu)$ and $\mathbf{d}_{\mathrm{K}}(\nu)$ as the K–dimensional vectors defined by

$$\mathbf{a}_{\mathrm{K}}(\nu) = \frac{1}{\sqrt{\mathrm{K}}}\left(1, e^{2i\pi\nu}, \ldots, e^{2i\pi(\mathrm{K}-1)\nu}\right)^{\mathrm{T}} \tag{3.1.12}$$

and

$$\mathbf{d}_{\mathrm{K}}(\nu) = \left(1, e^{2i\pi\nu}, \ldots, e^{2i\pi(\mathrm{K}-1)\nu}\right)^{\mathrm{T}} \tag{3.1.13}$$

Then, the following representation holds.

Proposition 3.1.1. *If* $\mathbf{A}$ *is a* PK × PK *block matrix, then, it holds that*

$$\sum_{k=-(\mathrm{K}-1)}^{\mathrm{K}-1} \tau^{(\mathrm{P})}(\mathbf{A})(k)e^{-2i\pi\nu} = \mathbf{a}_{\mathrm{K}}(\nu)^*\,\hat{\mathbf{A}}\,\mathbf{a}_{\mathrm{K}}(\nu) \tag{3.1.14}$$

Moreover, matrix $\mathscr{T}^{(\mathrm{P})}_{\mathrm{R,K}}(\mathbf{A})$ *can be written as*

$$\mathscr{T}^{(\mathrm{P})}_{\mathrm{R,K}}(\mathbf{A}) = \int_0^1 \mathbf{d}_{\mathrm{R}}(\nu)\mathbf{d}_{\mathrm{R}}(\nu)^*\,\mathbf{a}_{\mathrm{K}}(\nu)^*\,\hat{\mathbf{A}}\,\mathbf{a}_{\mathrm{K}}(\nu)\,d\nu \tag{3.1.15}$$

Proof. Using the expression (3.1.6) of $\tau^{(\mathrm{P})}(\mathbf{A})(k)$, the term $\sum_{k=-(\mathrm{K}-1)}^{\mathrm{K}-1} \tau^{(\mathrm{P})}(\mathbf{A})(k)e^{-2i\pi k\nu}$ can be written as

$$\mathrm{Tr}\left(\hat{\mathbf{A}}\left(\frac{1}{\mathrm{K}}\sum_{k=-(\mathrm{K}-1)}^{\mathrm{K}-1} e^{-2i\pi k\nu}\mathbf{J}_{\mathrm{K}}^k\right)\right)$$

It is easily seen that

$$\frac{1}{\mathrm{K}}\sum_{k=-(\mathrm{K}-1)}^{\mathrm{K}-1} e^{-2i\pi k\nu}\mathbf{J}_{\mathrm{K}}^k = \mathbf{a}_{\mathrm{K}}(\nu)\mathbf{a}_{\mathrm{K}}(\nu)^*$$

from which (3.1.14) and (3.1.15) follow immediately.

We now establish some useful corollaries of Proposition 3.1.1.

Proposition 3.1.2. *If* $\mathbf{A}$ *is a* PK × PK *positive definite matrix, then, for each integer* R *it holds that*

$$\mathscr{T}^{(\mathrm{P})}_{\mathrm{R,K}}(\mathbf{A}) > 0 \tag{3.1.16}$$

Proof. As $\mathbf{A}$ is positive definite, matrix $\hat{\mathbf{A}}$ is also positive definite which implies that $\mathbf{a}_{\mathrm{K}}(\nu)^*\,\hat{\mathbf{A}}\,\mathbf{a}_{\mathrm{K}}(\nu) > 0$ for each ν. Therefore, matrix $\mathscr{T}^{(\mathrm{P})}_{\mathrm{R,K}}(\mathbf{A})$ is hermitian. For each R–dimensional vector $\mathbf{b}$, the integral representation (3.1.15) leads to

$$\mathbf{b}^*\mathscr{T}^{(\mathrm{P})}_{\mathrm{R,K}}(\mathbf{A})\mathbf{b} = \int_0^1 \left|\mathbf{b}^*\mathbf{d}_{\mathrm{R}}(\nu)\right|^2 \mathbf{a}_{\mathrm{K}}(\nu)^*\,\hat{\mathbf{A}}\,\mathbf{a}_{\mathrm{K}}(\nu)\,d\nu$$

which is of course strictly positive as expected.

Proposition 3.1.3. *If* $\mathbf{A}$ *is a* $\mathrm{PK} \times \mathrm{PK}$ *matrix, then, for each integer* R*, it holds that*

$$\left\| \mathscr{T}^{(\mathrm{P})}_{\mathrm{R,K}}(\mathbf{A}) \right\| \leq \sup_{\nu \in [0,1]} \left| \mathbf{a}_{\mathrm{K}}(\nu)^* \hat{\mathbf{A}}\, \mathbf{a}_{\mathrm{K}}(\nu) \right| \leq \|\mathbf{A}\| \tag{3.1.17}$$

Proof. The first inequality of (3.1.17) is a well known property of Toeplitz matrices (see e.g. [13]). We however provide the proof for the reader's convenience. For each R–dimensional vectors $\mathbf{b}$ and $\mathbf{c}$,

$$\left| \mathbf{b}^* \mathscr{T}^{(\mathrm{P})}_{\mathrm{R,K}}(\mathbf{A}) \mathbf{c} \right| \leq \sup_{\nu \in [0,1]} \left| \mathbf{a}_{\mathrm{K}}(\nu)^* \hat{\mathbf{A}}\, \mathbf{a}_{\mathrm{K}}(\nu) \right| \int_0^1 |\mathbf{b}^* \mathbf{d}_{\mathrm{R}}(\nu)| |\mathbf{d}_{\mathrm{R}}(\nu)^* \mathbf{c}| \, d\nu$$

The Schwartz inequality leads to

$$\int_0^1 |\mathbf{b}^* \mathbf{d}_{\mathrm{R}}(\nu)| |\mathbf{d}_{\mathrm{R}}(\nu)^* \mathbf{c}| \, d\nu \leq \left(\int_0^1 \mathbf{b}^* \mathbf{d}_{\mathrm{R}}(\nu) \mathbf{d}_{\mathrm{R}}(\nu)^* \mathbf{b}\, d\nu \right)^{1/2} \left(\int_0^1 \mathbf{c}^* \mathbf{d}_{\mathrm{R}}(\nu) \mathbf{d}_{\mathrm{R}}(\nu)^* \mathbf{c}\, d\nu \right)^{1/2}$$

and the conclusion follows from the observation that

$$\int_0^1 \mathbf{b}^* \mathbf{d}_{\mathrm{R}}(\nu) \mathbf{d}_{\mathrm{R}}(\nu)^* \mathbf{b}\, d\nu = \|\mathbf{b}\|^2$$

Therefore, we have checked that

$$\left| \mathbf{b}^* \mathscr{T}^{(\mathrm{P})}_{\mathrm{R,K}}(\mathbf{A}) \mathbf{c} \right| \leq \sup_{\nu \in [0,1]} \left| \mathbf{a}_{\mathrm{K}}(\nu)^* \hat{\mathbf{A}}\, \mathbf{a}_{\mathrm{K}}(\nu) \right| \|\mathbf{b}\|\, \|\mathbf{c}\|$$

which implies the first inequality of (3.1.17). The second inequality of (3.1.17) follows directly from $\|\hat{\mathbf{A}}\| \leq \|\mathbf{A}\|$.

Remark 7. *We can notice that*

$$\begin{aligned}
\mathbf{a}^*_{\mathrm{K}}(\nu) \left(\frac{1}{\mathrm{P}} \sum_{p=1}^{\mathrm{P}} \mathbf{A}^{p,p} \right) \mathbf{a}_{\mathrm{K}}(\nu) &= \frac{1}{\mathrm{P}} \sum_{p=1}^{\mathrm{M}} \mathrm{Tr}\left(\mathbf{A}^{p,p} \mathbf{a}_{\mathrm{K}}(\nu) \mathbf{a}^*_{\mathrm{K}}(\nu)\right) \\
&= \frac{1}{\mathrm{P}} \mathrm{Tr}\left(\mathbf{A}\left(\mathbf{I}_{\mathrm{P}} \otimes \mathbf{a}_{\mathrm{K}}(\nu) \mathbf{a}^*_{\mathrm{K}}(\nu)\right)\right) \\
&= \frac{1}{\mathrm{P}} \mathrm{Tr}\left(\mathbf{A}\left(\left(\sum_{p=1}^{\mathrm{P}} \mathbf{f}^p (\mathbf{f}^p)^{\mathrm{T}}\right) \otimes \mathbf{a}_{\mathrm{K}}(\nu) \mathbf{a}^*_{\mathrm{K}}(\nu)\right)\right) \\
&= \frac{1}{\mathrm{P}} \mathrm{Tr}\left(\sum_{p=1}^{\mathrm{P}} \mathbf{A}(\mathbf{f}^p \otimes \mathbf{a}_{\mathrm{K}}(\nu))(\mathbf{f}^p \otimes \mathbf{a}_{\mathrm{K}}(\nu))^*\right) \\
&= \frac{1}{\mathrm{P}} \sum_{p=1}^{\mathrm{P}} (\mathbf{f}^p \otimes \mathbf{a}_{\mathrm{K}}(\nu))^* \mathbf{A}\, (\mathbf{f}^p \otimes \mathbf{a}_{\mathrm{K}}(\nu))
\end{aligned} \tag{3.1.18}$$

with $(\mathbf{f}^1, \ldots, \mathbf{f}^{\mathrm{P}})$ *the canonical base of* $\mathbb{C}^{\mathrm{P}}$*. So that from* (3.1.17)*, it holds immediately*

$$\left\| \mathscr{T}^{(\mathrm{P})}_{\mathrm{R,K}}(\mathbf{A}) \right\| \leq \sup_{\nu \in [0,1]} \left| \frac{1}{\mathrm{P}} \sum_{p=1}^{\mathrm{P}} (\mathbf{f}^p \otimes \mathbf{a}_{\mathrm{K}}(\nu))^* \mathbf{A}\, (\mathbf{f}^p \otimes \mathbf{a}_{\mathrm{K}}(\nu)) \right|. \tag{3.1.19}$$

We finally establish the following result.

Proposition 3.1.4. *If* $\mathbf{A}$ *is a* $\mathrm{K} \times \mathrm{K}$ *matrix and if* R *is an integer, then, it holds that*

$$\mathscr{T}_{\mathrm{R,K}}(\mathbf{A}) \left(\mathscr{T}_{\mathrm{R,K}}(\mathbf{A})\right)^* \leq \mathscr{T}_{\mathrm{R,K}}(\mathbf{A}\mathbf{A}^*) \tag{3.1.20}$$

Proof. The proof takes benefit of the integral representation (3.1.15) and follows directly from a version of the matrix-valued Schwartz inequality: If $\mathbf{u}(\nu)$ and $\mathbf{v}(\nu)$ are R–dimensional vector-valued functions defined on $[0,1]$, then it holds that

$$[\mathbf{u},\mathbf{v}]\left([\mathbf{v},\mathbf{v}]\right)^{-1}[\mathbf{u},\mathbf{v}]^* \leq [\mathbf{u},\mathbf{u}] \tag{3.1.21}$$

where $[\mathbf{u},\mathbf{v}]$ represents the $\mathrm{R}\times\mathrm{R}$ matrix

$$[\mathbf{u},\mathbf{v}] = \int_0^1 \mathbf{u}(\nu)\mathbf{v}(\nu)^* \, d\nu$$

We use (3.1.21) for $\mathbf{u}(\nu) = \mathbf{a}_{\mathrm{K}}(\nu)^*\hat{\mathbf{A}}\,\mathbf{a}_{\mathrm{K}}(\nu)\,\mathbf{d}_{\mathrm{R}}(\nu)$ and $\mathbf{v}(\nu) = \mathbf{d}_{\mathrm{R}}(\nu)$. It is clear that $[\mathbf{u},\mathbf{v}]$ coincides with $\mathscr{T}_{\mathrm{R,K}}(\mathbf{A})$, that $[\mathbf{v},\mathbf{v}] = \mathbf{I}$, and that

$$[\mathbf{u},\mathbf{u}] = \int_0^1 \left|\mathbf{a}_{\mathrm{K}}(\nu)^*\hat{\mathbf{A}}\,\mathbf{a}_{\mathrm{K}}(\nu)\right|^2 \mathbf{d}_{\mathrm{R}}(\nu)\,\mathbf{d}_{\mathrm{R}}(\nu)^* \, d\nu$$

The conclusion follows from the inequalities

$$\left|\mathbf{a}_{\mathrm{K}}(\nu)^*\hat{\mathbf{A}}\,\mathbf{a}_{\mathrm{K}}(\nu)\right|^2 \leq \mathbf{a}_{\mathrm{K}}(\nu)^*\hat{\mathbf{A}}\hat{\mathbf{A}}^*\mathbf{a}_{\mathrm{K}}(\nu)$$

and

$$\left(\frac{1}{\mathrm{P}}\sum_{p=1}^{\mathrm{P}}\mathbf{A}^{(p,p}\right)\left(\frac{1}{\mathrm{P}}\sum_{p=1}^{\mathrm{P}}\mathbf{A}^{(p,p}\right)^* \leq \frac{1}{\mathrm{P}}\sum_{p=1}^{\mathrm{P}}\mathbf{A}^{(p,p}\mathbf{A}^{(p,p)*} \leq \frac{1}{\mathrm{P}}\sum_{p=1}^{\mathrm{P}}(\mathbf{A}\mathbf{A}^*)^{(p,p)}$$

which can be seen as a matrix Jensen inequality.

3.1.3 Resolvents and Stieltjes transforms.

Let μ be a finite positive measure with support $supp(\mu) \subset \mathbb{R}$. Its Stieltjes transform $m_\mu(z)$ is the function defined by :

$$m_\mu(z) = \int_{\mathbb{R}} \frac{d\mu(\lambda)}{\lambda - z} \quad \forall z \in \mathbb{C} \setminus supp(\mu).$$

m_μ satisfies the following properties:

Property 3.1.1.

- *m_μ is holomorphic on $\mathbb{C} \setminus supp(\mu)$;*
- *m_μ verifies $(m_\mu(z))^* = m_\mu(z^*)$;*
- *$z \in \mathbb{C}^+$ implies $m_\mu(z) \in \mathbb{C}^+$;*
- *$m_\mu(z)$ satisfies*

$$|m_\mu(z)| \leq \frac{\mu(\mathbb{R})}{dist(z, supp(\mu))} \tag{3.1.22}$$

for each $z \in \mathbb{C} \setminus supp(\mu)$, and

$$|m_\mu(z)| \leq \frac{\mu(\mathbb{R})}{|\mathrm{Im}(z)|} \tag{3.1.23}$$

for each $z \in \mathbb{C} \setminus \mathbb{R}$.

If moreover $supp(\mu) \subset \mathbb{R}^+$, then it holds that

- *$z \in \mathbb{C}^+$ implies $z\, m_\mu(z) \in \mathbb{C}^+$;*

- *$m_\mu(z)$ satisfies*

$$|m_\mu(z)| \leq \frac{\mu(\mathbb{R})}{dist(z, supp(\mu))} \leq \frac{\mu(\mathbb{R})}{dist(z, \mathbb{R}^+)} \tag{3.1.24}$$

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

The following property shows how to recover the measure μ from its Stieltjes transform $m_\mu(z)$.

Property 3.1.2. *The mass $\mu(\mathbb{R})$ can be recovered through the formula*

$$\mu(\mathbb{R}) = \lim_{y\to\infty} -iy m_\mu(iy). \tag{3.1.25}$$

Moreover, for all $\phi \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$, the set of all compactly supported real valued smooth functions, it holds that

$$\int_{\mathbb{R}} \phi(\lambda) d\mu(\lambda) = \frac{1}{\pi} \lim_{y\downarrow 0} \mathrm{Im}\left(\int_{\mathbb{R}} \phi(x) m_\mu(x+iy) dx\right). \tag{3.1.26}$$

Finally, if a and b are continuity points of μ,

$$\mu([a,b]) = \lim_{y\downarrow 0} \int_a^b \mathrm{Im}(m_\mu(x+iy)) dx. \tag{3.1.27}$$

The Stieltjes transforms are characterized by the following properties.

Property 3.1.3. *If a function $m(z)$ verifies the following conditions*

1. *$m(z)$ is holomorphic on $\mathbb{C}^+$*
2. *$z \in \mathbb{C}^+$ implies $m(z) \in \mathbb{C}^+$*
3. *$\limsup_{y\to\infty} |iy\, m(iy)| < \infty$*

then $m(z)$ is the Stieltjes transform of a uniquely defined positive measure μ. If moreover $zm(z) \in \mathbb{C}^+$ when $z \in \mathbb{C}^+$, then μ is carried by $\mathbb{R}^+$.

The concept of Stieltjes transform can be extended to finite positive matrix valued measures defined on $\mathbb{R}$. A finite positive matrix valued measure $\boldsymbol{\mu}$ is a σ–additive positive matrix valued function defined on the set of all Borel sets of $\mathbb{R}$ for which $\|\boldsymbol{\mu}(\mathbb{R})\| < +\infty$. For such a measure, the support of $\boldsymbol{\mu}$ is defined as the support of the scalar positive measure Tr($\boldsymbol{\mu}$). The Stieltjes transform $m_{\boldsymbol{\mu}}(z)$ of $\boldsymbol{\mu}$ is the function defined for $z \in \mathbb{C} \setminus supp(\mu)$ by

$$m_{\boldsymbol{\mu}}(z) = \int_{\mathbb{R}} \frac{1}{\lambda - z} d\boldsymbol{\mu}(\lambda) \tag{3.1.28}$$

$m_{\boldsymbol{\mu}}(z)$ satisfies properties that can be interpreted as matrix generalizations of Properties 3.1.1, 3.1.2 and 3.1.3. In particular, we have:

Property 3.1.4.

- *$m_{\boldsymbol{\mu}}$ is holomorphic on $\mathbb{C} \setminus supp(\boldsymbol{\mu})$;*
- *$m_{\boldsymbol{\mu}}$ verifies $(m_{\boldsymbol{\mu}}(z))^* = m_{\boldsymbol{\mu}}(z^*)$;*
- *$z \in \mathbb{C}^+$ implies $\mathrm{Im}\left(m_{\boldsymbol{\mu}}(z)\right) > 0$;*

- *If* $\boldsymbol{\mu}(\mathbb{R}) = \mathbf{I}$, $m_{\boldsymbol{\mu}}(z)$ *satisfies*

$$m_{\boldsymbol{\mu}}(z)\left(m_{\boldsymbol{\mu}}(z)\right)^* \leq \frac{\mathbf{I}}{\left[dist(z, supp(\boldsymbol{\mu}))\right]^2} \tag{3.1.29}$$

for each $z \in \mathbb{C} \setminus supp(\boldsymbol{\mu})$, *and*

$$m_{\boldsymbol{\mu}}(z)\left(m_{\boldsymbol{\mu}}(z)\right)^* \leq \frac{\mathbf{I}}{(\mathrm{Im}(z))^2} \tag{3.1.30}$$

for each $z \in \mathbb{C} \setminus \mathbb{R}$.

If moreover $supp(\boldsymbol{\mu}) \subset \mathbb{R}^+$, *then it holds that*

- $z \in \mathbb{C}^+$ *implies* $\mathrm{Im}\left(z\, m_{\boldsymbol{\mu}}(z)\right) > 0$;
- *If* $\boldsymbol{\mu}(\mathbb{R}) = \mathbf{I}$, $m_{\boldsymbol{\mu}}(z)$ *satisfies*

$$m_{\boldsymbol{\mu}}(z)\left(m_{\boldsymbol{\mu}}(z)\right)^* \leq \frac{\mathbf{I}}{\left[dist(z, supp(\boldsymbol{\mu}))\right]^2} \leq \frac{\mathbf{I}}{[dist(z, \mathbb{R}^+)]^2} \tag{3.1.31}$$

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

Proof. Bounds (3.1.29, 3.1.30, 3.1.31) are not classical so that we provide the proof of e.g. (3.1.23). The proof is based on a version of the matrix Schwartz inequality. We denote by $\mathbb{L}^2(\boldsymbol{\mu})$ the space of all row vector-valued functions $\mathbf{u}(\lambda)$ defined on $\mathbb{R}$ satisfying $\int_{\mathbb{R}} \mathbf{u}(\lambda)\, d\boldsymbol{\mu}(\lambda)\, \mathbf{u}(\lambda)^* < +\infty$ (see [63] for more details on the definition of the above integral). $\mathbb{L}^2(\boldsymbol{\mu})$ endowed with the scalar product $< \mathbf{u}, \mathbf{v} >$ defined by

$$< \mathbf{u}, \mathbf{v} >= \int_{\mathbb{R}} \mathbf{u}(\lambda)\, d\boldsymbol{\mu}(\lambda)\, \mathbf{v}(\lambda)^*$$

is a Hilbert space. Then, if $\mathbf{U}(\lambda) = (\mathbf{u}_1(\lambda)^T, \ldots, \mathbf{u}_p(\lambda))^T)^T$ and $\mathbf{V}(\lambda) = (\mathbf{v}_1(\lambda)^T, \ldots, \mathbf{v}_p(\lambda))^T)^T$ are matrices whose rows are elements of $\mathbb{L}^2(\boldsymbol{\mu})$, it holds that

$$[\mathbf{U}, \mathbf{V}]\ ([\mathbf{V}, \mathbf{V}])^{-1}\ [\mathbf{U}, \mathbf{V}]^* \leq [\mathbf{U}, \mathbf{U}] \tag{3.1.32}$$

where $[\mathbf{U}, \mathbf{V}]$ is matrix defined by $([\mathbf{U}, \mathbf{V}])_{i,j} =< \mathbf{u}_i, \mathbf{v}_j >$. Using (3.1.32) for $\mathbf{U}(\lambda) = \frac{\mathbf{I}}{\lambda - z}$ and $\mathbf{V} = \mathbf{I}$, and remarking that $|\lambda - z|^2 \geq |\mathrm{Im}(z)|^2$ for each λ, we obtain immediately (3.1.30).

Proposition (3.1.3) can also be extended as follows:

Property 3.1.5. *If a square matrix valued function* $m(z)$ *verifies the following conditions*

1. $m(z)$ *is holomorphic on* $\mathbb{C}^+$
2. $z \in \mathbb{C}^+$ *implies* $\mathrm{Im}(m(z)) > 0$
3. $\limsup_{y \to +\infty} |iy\, \mathrm{Tr}(m(iy))| < \infty$

then $m(z)$ *is the Stieltjes transform of a uniquely defined positive matrix valued measure* $\boldsymbol{\mu}$. *If moreover* $\mathrm{Im}(zm(z)) > 0$ *when* $z \in \mathbb{C}^+$, *then* $\boldsymbol{\mu}$ *is carried by* $\mathbb{R}^+$.

Definition 2. *In the following, we will denote by* $\mathcal{S}_P(\mathbb{R}^+)$ *the set of all Stieltjes transforms of positive* $\mathbb{C}^{P\times P}$-*valued measures* $\boldsymbol{\mu}$ *carried by* $\mathbb{R}^+$ *and satisfying* $\boldsymbol{\mu}(\mathbb{R}^+) = \mathbf{I}_P$.

We will often use the following consequence of Montel's theorem ([64]).

Proposition 3.1.5. *If* $(f_n(z))_{n\in\mathbb{N}}$ *and* $(g_n(z))_{n\in\mathbb{N}}$ *are Stieltjes transforms of positive measures* $(\mu_n)_{n\in\mathbb{N}}$ *and* $(\nu_n)_{n\in\mathbb{N}}$ *carried by* $\mathbb{R}^+$ *and satisfying* $\sup_n \mu_n(\mathbb{R}^+) \leq \kappa$ *and* $\sup_n \nu_n(\mathbb{R}^+) \leq \kappa$, *and if* $f_n(z) - g_n(z) \to 0$ *in a certain domain* $\mathcal{D}$ *of* $\mathbb{C}^+$, *then,* $f_n - g_n$ *converges towards* 0 *uniformly on each compact subset of* $\mathbb{C} \setminus \mathbb{R}^+$.

Proof. For each compact subset $\mathcal{K} \in \mathcal{S}_1(\mathbb{R}^+)$, (3.1.24) implies that $|f_n(z) - g_n(z)| \leq \frac{2\kappa}{\text{dist}(\mathcal{K},\mathbb{R}^+)}$ for each $z \in \mathcal{K}$. Therefore, $(f_n - g_n)_{n\in\mathbb{N}}$ is a normal family on $\mathbb{C} \setminus \mathbb{R}^+$. Each convergent subsequence extracted from $(f_n - g_n)_{n\in\mathbb{N}}$ thus converges uniformly on each compact subset of $\mathbb{C} \setminus \mathbb{R}^+$ towards a function $h(z)$ which is analytic on $\mathbb{C} \setminus \mathbb{R}^+$. As $f_n(z) - g_n(z)$ converges towards 0 for each $z \in \mathcal{D}$, then $h(z) = 0$ for each $z \in \mathcal{D}$. As h is analytic on $\mathbb{C} \setminus \mathbb{R}^+$, h must be zero on $\mathbb{C} \setminus \mathbb{R}^+$. This shows that the limits of each convergent subsequence extracted from $(f_n - g_n)_{n\in\mathbb{N}}$ are zero, thus showing that the whole sequence $(f_n - g_n)_{n\in\mathbb{N}}$ converges uniformly towards 0 on each compact subset of $\mathbb{C} \setminus \mathbb{R}^+$. □

If $\mathbf{A}$ is a $\mathrm{P} \times \mathrm{P}$ matrix, we recall that the resolvent of $\mathbf{AA}^*$ is defined as the matrix valued function

$$\mathbf{Q}(z) = \left(\mathbf{AA}^* - z\mathbf{I}\right)^{-1} \tag{3.1.33}$$

If $\mathbf{AA}^* = \sum_{k=1}^{\mathrm{P}} \lambda_k \mathbf{u}_k \mathbf{u}_k^*$ is the eigenvalue / eigenvector decomposition of $\mathbf{AA}^*$, then, $\mathbf{Q}$ coincides with the Stieltjes transform of the positive matrix-valued measure carried by $\mathbb{R}^+$ $\boldsymbol{\mu}$ defined by

$$\boldsymbol{\mu} = \sum_{k=1}^{\mathrm{P}} \delta_{\lambda_k} \mathbf{u}_k \mathbf{u}_k^*$$

where δ_{λ_k} represents the Dirac distribution at point λ_k. As $\boldsymbol{\mu}(\mathbb{R}) = \mathbf{I}$, $\mathbf{Q}$ verifies properties (3.1.29, 3.1.30, 3.1.31). We also recall that $\frac{1}{\mathrm{P}}\mathrm{Tr}(\mathbf{Q}(z))$ is the Stieltjes transform of the empirical eigenvalue distribution of $\mathbf{AA}^*$. We finally mention that $\mathbf{Q}(z)$ satisfies the equation

$$\mathbf{Q}(z) = -\frac{\mathbf{I}_{\mathrm{P}}}{z} + \frac{1}{z}\mathbf{Q}(z)\mathbf{AA}^* \tag{3.1.34}$$

3.2 Random variables notations and tools.

If x is a complex-valued random variable, the variance of x, denoted by $\mathrm{Var}(x)$, is defined by

$$\mathrm{Var}(x) = \mathbb{E}(|x|^2) - |\mathbb{E}(x)|^2$$

The zero-mean random variable $x - \mathbb{E}(x)$ is denoted by x°. x is said to be $\mathcal{N}_c(\alpha, \sigma^2)$ (complex circular) distributed if $\mathbb{E}(x) = \alpha$ and if $\mathrm{Re}(x)$ and $\mathrm{Im}(x)$ are independent real Gaussian variables of variance $\sigma^2/2$. Multivariable complex circular distribution is denoted $\mathcal{N}_c(\boldsymbol{\alpha}, \boldsymbol{\Gamma})$ in the following.

We now present the 2 fundamental tools that are used in the following.

Proposition 3.2.1. Integration by parts formula. *Let* $\boldsymbol{\xi} = [\xi_1, \ldots, \xi_{\mathrm{K}}]^{\mathrm{T}}$ *be a* $\mathcal{N}_c(0, \boldsymbol{\Omega})$ *distributed complex Gaussian random vector. If* $\boldsymbol{\Gamma} : (\boldsymbol{\xi}) \mapsto \boldsymbol{\Gamma}(\boldsymbol{\xi}, \overline{\boldsymbol{\xi}})$ *is a* $\mathcal{C}^1$ *complex valued function polynomially bounded together with its derivatives, then*

$$\mathbb{E}[\xi_p \boldsymbol{\Gamma}(\boldsymbol{\xi})] = \sum_{m=1}^{\mathrm{K}} \boldsymbol{\Omega}_{pm} \mathbb{E}\left[\frac{\partial \boldsymbol{\Gamma}(\boldsymbol{\xi})}{\partial \overline{\xi}_m}\right]. \tag{3.2.1}$$

Proposition 3.2.2. Poincaré-Nash inequality. *Let* $\boldsymbol{\xi} = [\xi_1, \ldots, \xi_{\mathrm{K}}]^{\mathrm{T}}$ *be a* $\mathcal{N}_c(0, \boldsymbol{\Omega})$ *distributed complex Gaussian random vector. If* $\boldsymbol{\Gamma} : (\boldsymbol{\xi}) \mapsto \boldsymbol{\Gamma}(\boldsymbol{\xi}, \overline{\boldsymbol{\xi}})$ *is a* $\mathcal{C}^1$ *complex valued function polynomially bounded together with its derivatives, then, noting* $\nabla_{\boldsymbol{\xi}} \boldsymbol{\Gamma} = [\frac{\partial \boldsymbol{\Gamma}}{\partial \xi_1}, \ldots, \frac{\partial \boldsymbol{\Gamma}}{\partial \xi_{\mathrm{K}}}]^{\mathrm{T}}$ *and* $\nabla_{\overline{\boldsymbol{\xi}}} \boldsymbol{\Gamma} = [\frac{\partial \boldsymbol{\Gamma}}{\partial \overline{\xi}_1}, \ldots, \frac{\partial \boldsymbol{\Gamma}}{\partial \overline{\xi}_{\mathrm{K}}}]^{\mathrm{T}}$*,*

$$\mathrm{Var}(\boldsymbol{\Gamma}(\boldsymbol{\xi})) \leq \mathbb{E}\left[\nabla_{\boldsymbol{\xi}} \boldsymbol{\Gamma}(\boldsymbol{\xi})^{\mathrm{T}}\, \boldsymbol{\Omega}\, \overline{\nabla_{\boldsymbol{\xi}} \boldsymbol{\Gamma}(\boldsymbol{\xi})}\right] + \mathbb{E}\left[\nabla_{\overline{\boldsymbol{\xi}}} \boldsymbol{\Gamma}(\boldsymbol{\xi})^*\, \boldsymbol{\Omega}\, \nabla_{\overline{\boldsymbol{\xi}}} \boldsymbol{\Gamma}(\boldsymbol{\xi})\right] \tag{3.2.2}$$

3.3 Overview of the results of Chapter.

We are now in position to present the results of this chapter, and first introduce the random matrix model under study.

We consider independent identically distributed $\mathcal{N}_c(0,\sigma^2/N)$ random variables $(w_{m,n})_{m=1,\dots,M,n=1,\dots,N+L-1}$ where M,N,L are integers. We define the $L\times N$ matrices $(\mathbf{W}_N^{(m)})_{m=1,\dots,M}$ as the Hankel matrices whose entries are given by

$$\left(\mathbf{W}_N^{(m)}\right)_{i,j} = w_{m,i+j-1},\ 1\le i\le L, 1\le j\le N \tag{3.3.1}$$

and $\mathbf{W}_N$ represents the $ML\times N$ matrix

$$\mathbf{W}_N = \begin{pmatrix} \mathbf{W}_N^{(1)} \\ \mathbf{W}_N^{(2)} \\ \vdots \\ \mathbf{W}_N^{(M)} \end{pmatrix} \tag{3.3.2}$$

In this Chapter, we study the complex Gaussian block-Hankel information plus noise $ML\times N$ random matrix model defined by :

$$\mathbf{Y}_N = \mathbf{A}_N + \mathbf{W}_N \tag{3.3.3}$$

where

- matrix $\mathbf{A}_N$ is deterministic and satisfies $\sup_N \|\mathbf{A}_N\| < \infty$,
- matrix $\mathbf{W}_N$ is defined as in (3.3.2).

in the asymptotic regime:

Assumption A-1: $M\to+\infty, N\to+\infty$ *in such a way that* $c_N = \frac{ML}{N}\to c$, *where* $0<c<+\infty$

In order to avoid technicalities, we however consider the following regime which is nearly equivalent to Assumption A-1

Assumption A-2: $N\to+\infty$, $c_N = \frac{ML}{N}\to c$, *where* $0<c<+\infty$, $L=\mathcal{O}(N^\alpha)$ *for* $\alpha<1$.

To shorten the notations, $N\to+\infty$ should be understood as the asymptotic regime defined by Assumption A-2.

Model (3.3.3) is a generalization of the normalized random matrix models $\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N$ considered in Chapter 2 in which the rank of deterministic matrix $\mathbf{B}_N$ does not scale with N. In this Chapter, we study the behaviour of the resolvents $\mathbf{Q}_N(z) = \left(\mathbf{Y}_N\mathbf{Y}_N^* - z\mathbf{I}\right)^{-1}$ and $\tilde{\mathbf{Q}}_N(z) = \left(\mathbf{Y}_N^*\mathbf{Y}_N - z\mathbf{I}\right)^{-1}$, and show that they behave as deterministic matrices $\mathbf{T}_N(z)$ and $\tilde{\mathbf{T}}_N(z)$, called the deterministic equivalents of $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$, which are defined as the solutions of the following system of equations:

$$\begin{cases} \mathbf{T}_N(z) = \left[-z\left(\mathbf{I}_{ML} + \sigma^2\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_N^T(z))\right) + \mathbf{A}_N\left(\mathbf{I}_N + \sigma^2 c_N\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_N^T(z))\right)^{-1}\mathbf{A}_N^*\right]^{-1} \\ \tilde{\mathbf{T}}_N(z) = \left[-z\left(\mathbf{I}_N + \sigma^2 c_N\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_N^T(z))\right) + \mathbf{A}_N^*\left(\mathbf{I}_{ML} + \sigma^2\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_N^T(z))\right)^{-1}\mathbf{A}_N\right]^{-1} \end{cases} \tag{3.3.4}$$

$\mathbf{T}_N(z)$ and $\tilde{\mathbf{T}}_N(z)$ are the Stieltjes transforms of matrix valued measures whose masses are equal to $\mathbf{I}_{ML}$ and $\mathbf{I}_N$ respectively, and $\frac{1}{ML}\mathrm{Tr}(\mathbf{T}_N(z))$ is the Stieltjes transform of a probability measure denoted μ_N. We establish that if $\mathbf{B}_N$ and $\tilde{\mathbf{B}}_N$ are uniformly bounded $ML\times ML$ and $N\times N$ matrices, then, under Assumption A-2, it holds that

$$\frac{1}{ML}\mathrm{Tr}\left[(\mathbf{Q}_N(z)-\mathbf{T}_N(z))\mathbf{B}_N\right]\to 0, a.s.\ \frac{1}{N}\mathrm{Tr}\left[\left(\tilde{\mathbf{Q}}_N(z)-\tilde{\mathbf{T}}_N(z)\right)\tilde{\mathbf{B}}_N\right]\to 0, a.s.$$

If moreover $\alpha < 2/3$, a condition nearly equivalent to $\frac{L}{M^2} \to 0$, then, we prove that bounded bilinear forms of $(\mathbf{Q}_N(z) - \mathbf{T}_N(z))$ and $\left(\tilde{\mathbf{Q}}_N(z) - \tilde{\mathbf{T}}_N(z)\right)$ also converge towards 0 almost surely. These results imply that the empirical eigenvalue distribution of $\mathbf{Y}_N\mathbf{Y}_N^*$ has the same behaviour that μ_N. The deterministic behaviour of the bilinear forms of $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$ can be of interest to study eigenvectors of $\mathbf{Y}_N\mathbf{Y}_N^*$ provided informations on the location of the eigenvalues can be collected: evaluation of the support of μ_N and derivation of results showing that no eigenvalue of $\mathbf{Y}_N\mathbf{Y}_N^*$ can belong to intervals that are outside the support of μ_N for each N large enough. However, we have not been able to obtain the corresponding results. In particular, the equation (3.3.4) looks complicated, and this makes the analysis of the support of μ_N very difficult. As for the derivation of almost sure location of the eigenvalues of $\mathbf{Y}_N\mathbf{Y}_N^*$, the most direct approach would be to generalize the Haagerup-Thornbjornsen-Schultz approach ([35], [67]) used in [49] in the case $\mathbf{A}_N = 0$. The method consists in expanding of $\mathbb{E}\left(\frac{1}{ML}\mathrm{Tr}(\mathbf{Q}_N(z))\right)$ up to the required order, and to check that the various coefficients are Stieltjes transforms of distributions whose support are included into the support of μ_N. Even if $\mathbf{A}_N = 0$ (μ_N is the Marcenko-Pastur distribution of parameters (c_N, σ^2)), the calculations are very complicated, and we have chosen to address different topics. However, the study of the behaviour of the bilinear forms of $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$ are used in Chapter 4 in order to generalize the results of section 2.4 to the case where P and L scale with N.

In order to study the asymptotic behaviour of $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$, we use the integration by parts formula and the Poincaré-Nash inequality. We first establish that the variance of the various functionals converge towards 0, and evaluate $\mathbb{E}(\mathbf{Q}_N(z))$ and $\mathbb{E}(\tilde{\mathbf{Q}}_N(z))$ using the integration by parts formula. This allows to establish that $\mathbb{E}(\mathbf{Q}_N(z))$ and $\mathbb{E}(\tilde{\mathbf{Q}}_N(z))$ have the same behaviour than deterministic matrix valued Stieltjes transforms $\mathbf{R}_N(z)$ and $\tilde{\mathbf{R}}_N(z)$ satisfying a perturbed version of equation (3.3.4). We then establish the convergence towards 0 of the normalized traces and of the bilinear forms of $\mathbb{E}(\mathbf{Q}_N(z)) - \mathbf{T}_N(z)$ and of $\mathbb{E}(\tilde{\mathbf{Q}}_N(z)) - \tilde{\mathbf{T}}_N(z)$. Generally speaking, we evaluate the above terms for $z \in \mathbb{C}^+$ because this allows to simplify the notations. A number of upper bounds are products of polynomials of $|z|$ and of $\frac{1}{\mathrm{Im}(z)}$ that do not depend on the dimensions L, M, N (such polynomials are called nice polynomials in the following). However, the above evaluations can be extended for $z \in \mathbb{C} - \mathbb{R}^+$ by replacing $\mathrm{Im}(z)$ by the distance between z and $\mathbb{R}^+$.

This chapter is organized as follows. In section 3.4, we evaluate the variances of various useful functionals of $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$. In sections 3.5 and 3.6, we use the integration by parts formula and the Poincaré-Nash inequality to establish that $\mathbb{E}(\mathbf{Q}_N(z))$ and $\mathbb{E}(\tilde{\mathbf{Q}}_N(z))$ are close from the above mentioned matrix-valued functions $\mathbf{R}_N(z)$ and $\tilde{\mathbf{R}}_N(z)$. The rate of convergence of various functionals is in particular evaluated. In section 3.7, we show that equation (3.3.4) has a unique solution in the relevant class of matrix-valued functions, and in sections 3.8, 3.9, we prove that the normalized traces and bilinear forms of $\mathbb{E}(\mathbf{Q}_N(z)) - \mathbf{T}_N(z)$ and of $\mathbb{E}(\tilde{\mathbf{Q}}_N(z)) - \tilde{\mathbf{T}}_N(z)$ converge towards 0. Note that in contrast with the case $\mathbf{A}_N = 0$ studied in [49], due to the lack of time, we have not been able to evaluate the corresponding rate of convergence. This issue will be addressed in the next future.

3.4 Poincaré-Nash variance evaluations

In this section, we take benefit of the Poincaré-Nash inequality to evaluate the variance of certain important terms. In particular, we prove the following useful result.

Proposition 3.4.1. *Let* $\mathbf{B}$ *be a deterministic* $ML \times ML$ *matrix for which* $\sup_N \|\mathbf{B}\| \leq \kappa$*, and consider 2* ML*–dimensional deterministic vectors* $\mathbf{b}_1, \mathbf{b}_2$ *such that* $\sup_N \|\mathbf{b}_i\| \leq \kappa$ *for* $i = 1, 2$ *as well as 2* L*–dimensional* *deterministic vectors* $\mathbf{u}_1, \mathbf{u}_2$ *such that* $\sup_N \|\mathbf{u}_i\| \leq \kappa$ *for* $i = 1, 2$. *Then, for each* $z \in \mathbb{C}^+$, *it holds that*

$$\operatorname{Var}\left(\frac{1}{ML}\operatorname{Tr}(\mathbf{BQ}(z))\right) \leq C(z)\,\kappa^2\,\frac{1}{MN} \tag{3.4.1}$$

$$\operatorname{Var}\left(\mathbf{b}_1^*\mathbf{Q}(z)\mathbf{b}_2\right) \leq C(z)\,\kappa^4\,\frac{L}{N} \tag{3.4.2}$$

$$\operatorname{Var}\left(\mathbf{u}_1^*\left[\frac{1}{M}\sum_{m=1}^{M}(\mathbf{Q}(z))^{m,m}\right]\mathbf{u}_2\right) \leq C(z)\,\kappa^4\,\frac{L}{MN} \tag{3.4.3}$$

where $C(z)$ *can be written as* $C(z) = P_1(|z|)P_2\left(\frac{1}{\operatorname{Im}(z)}\right)$ *for some nice polynomials* P_1 *and* P_2. *Moreover, if* $\mathbf{G}$ *is a* $N \times N$ *deterministic matrix verifying* $\sup_N \|\mathbf{G}\| \leq \kappa$, *the following evaluations hold:*

$$\operatorname{Var}\left(\frac{1}{ML}\operatorname{Tr}\left(\mathbf{BQ}(z)\mathbf{YGY}^*\right)\right) \leq C(z)\,\kappa^4\,\frac{1}{MN} \tag{3.4.4}$$

$$\operatorname{Var}\left(\mathbf{b}_1^*\mathbf{Q}(z)\mathbf{YGY}^*\mathbf{b}_2\right) \leq C(z)\,\kappa^6\,\frac{L}{N} \tag{3.4.5}$$

$$\operatorname{Var}\left(\mathbf{u}_1^*\left[\frac{1}{M}\sum_{m=1}^{M}\left(\mathbf{Q}(z)\mathbf{YGY}^*\right)^{m,m}\right]\mathbf{u}_2\right) \leq C(z)\,\kappa^6\,\frac{L}{MN} \tag{3.4.6}$$

where $C(z)$ *can be written as above. Moreover, if* $\tilde{\mathbf{B}}$ *is a* $ML \times N$ *deterministic matrix verifying* $\sup_N \|\tilde{\mathbf{B}}\| \leq \kappa$, *the following evaluations hold :*

$$\operatorname{Var}\left(\frac{1}{ML}\operatorname{Tr}\left(\mathbf{Q}(z)\tilde{\mathbf{B}}\mathbf{Y}^*\right)\right) \leq C(z)\,\kappa^2\,\frac{1}{MN} \tag{3.4.7}$$

$$\operatorname{Var}\left(\mathbf{b}_1^*\mathbf{Q}(z)\tilde{\mathbf{B}}\mathbf{Y}^*\mathbf{b}_2\right) \leq C(z)\,\kappa^6\,\frac{L}{N} \tag{3.4.8}$$

$$\operatorname{Var}\left(\mathbf{u}_1^*\left[\frac{1}{M}\sum_{m=1}^{M}\left(\mathbf{Q}(z)\tilde{\mathbf{B}}\mathbf{Y}^*\right)^{m,m}\right]\mathbf{u}_2\right) \leq C(z)\,\kappa^6\,\frac{L}{MN} \tag{3.4.9}$$

Proof. We first establish (3.4.1) and denote by ξ the random variable $\xi = \frac{1}{ML}\operatorname{Tr}(\mathbf{BQ}(z))$. As the various entries of 2 different blocks $\mathbf{W}^{m_1}, \mathbf{W}^{m_2}$ are independent, the Poincaré-Nash inequality can be written as

$$\operatorname{Var}\xi \leq \sum_{m,i_1,i_2,j_1,j_2} \mathrm{E}\left[\left(\frac{\partial\xi}{\partial\overline{\mathbf{W}}^m_{i_1,j_1}}\right)^* \mathbb{E}\left(\mathbf{W}^m_{i_1,j_1}\overline{\mathbf{W}}^m_{i_2,j_2}\right)\frac{\partial\xi}{\partial\overline{\mathbf{W}}^m_{i_2,j_2}}\right] + \tag{3.4.10}$$

$$\sum_{m,i_1,i_2,j_1,j_2} \mathrm{E}\left[\frac{\partial\xi}{\partial\mathbf{W}^m_{i_1,j_1}}\mathbb{E}\left(\mathbf{W}^m_{i_1,j_1}\overline{\mathbf{W}}^m_{i_2,j_2}\right)\left(\frac{\partial\xi}{\partial\mathbf{W}^m_{i_2,j_2}}\right)^*\right] \tag{3.4.11}$$

In the following, we just evaluate the right hand side of (3.4.10), denoted by α, because the behaviour of the term defined by (3.4.11) can be established similarly. It is easy to check that

$$\frac{\partial\mathbf{Q}}{\partial\overline{\mathbf{W}}^m_{i,j}} = -\mathbf{QYe}_j(\mathbf{f}^m_i)^T\mathbf{Q}$$

so that

$$\frac{\partial\xi}{\partial\overline{\mathbf{W}}^m_{i,j}} = -\frac{1}{ML}\operatorname{Tr}\left(\mathbf{BQYe}_j(\mathbf{f}^m_i)^T\mathbf{Q}\right)$$

which can also be written $-\frac{1}{ML}(\mathbf{f}^m_i)^T\mathbf{QBQYe}_j$. We recall that $\mathbb{E}\left(\mathbf{W}^m_{i_1,j_1}\overline{\mathbf{W}}^m_{i_2,j_2}\right) = \frac{\sigma^2}{N}\delta(i_1 - i_2 = j_2 - j_1)$ (see (2.2.5)). Therefore, α is equal to the mathematical expectation of the term

$$\frac{1}{(ML)^2}\frac{\sigma^2}{N}\sum_{m,i_1,i_2,j_1,j_2}\delta(j_2 - j_1 = i_1 - i_2)\mathbf{e}_{j_1}^T\mathbf{Y}^*\mathbf{Q}^*\mathbf{B}^*\mathbf{Q}^*\mathbf{f}^m_{i_1}(\mathbf{f}^m_{i_2})^T\mathbf{QBQYe}_{j_2}$$

We put $u = i_1 - i_2$ and remark that $\sum_{m,i_1-i_2=u} \mathbf{f}_{i_1}^m (\mathbf{f}_{i_2}^m)^{\mathrm{T}} = \mathbf{I}_{\mathrm{M}} \otimes \mathbf{J}_{\mathrm{L}}^{*u}$. We thus obtain that

$$\alpha = \frac{1}{(\mathrm{ML})^2} \frac{\sigma^2}{\mathrm{N}} \mathbb{E}\left[\sum_{u=-(\mathrm{L}-1)}^{\mathrm{L}-1} \sum_{j_2-j_1=u} \mathbf{e}_{j_1}^{\mathrm{T}} \mathbf{Y}^* \mathbf{Q}^* \mathbf{B}^* \mathbf{Q}^* (\mathbf{I}_{\mathrm{M}} \otimes \mathbf{J}_{\mathrm{L}}^{*u}) \mathbf{QBQYe}_{j_2} \right]$$

Using that $\sum_{j_2-j_1=u} \mathbf{e}_{j_2} \mathbf{e}_{j_1}^{\mathrm{T}} = \mathbf{J}_{\mathrm{N}}^{*u}$, we get that

$$\alpha = \frac{1}{\mathrm{ML}} \frac{\sigma^2}{\mathrm{N}} \mathbb{E}\left[\sum_{u=-(\mathrm{L}-1)}^{\mathrm{L}-1} \frac{1}{\mathrm{ML}} \mathrm{Tr}\left(\mathbf{QBQYJ}_{\mathrm{N}}^{*u} \mathbf{Y}^* \mathbf{Q}^* \mathbf{B}^* \mathbf{Q}^* (\mathbf{I}_{\mathrm{M}} \otimes \mathbf{J}_{\mathrm{L}}^{*u}) \right) \right]$$

If $\mathbf{C}$ is a $\mathrm{ML} \times \mathrm{N}$ matrix, the Schwartz inequality as well as the inequality $(xy)^{1/2} \leq 1/2(x+y)$ lead to

$$\left| \frac{1}{\mathrm{ML}} \mathrm{Tr}\left(\mathbf{CJ}_{\mathrm{N}}^{*u} \mathbf{C}^* (\mathbf{I}_{\mathrm{M}} \otimes \mathbf{J}_{\mathrm{L}}^{*u}) \right) \right| \leq \frac{1}{2\mathrm{ML}} \mathrm{Tr}\left(\mathbf{CJ}_{\mathrm{N}}^{*u} \mathbf{J}_{\mathrm{N}}^{u} \mathbf{C}^* \right) + \frac{1}{2\mathrm{ML}} \mathrm{Tr}\left(\mathbf{C}^* (\mathbf{I}_{\mathrm{M}} \otimes \mathbf{J}_{\mathrm{L}}^{*u} \mathbf{J}_{\mathrm{L}}^{u}) \mathbf{C} \right)$$

It is clear that matrices $\mathbf{J}_{\mathrm{N}}^{*u} \mathbf{J}_{\mathrm{N}}^{u}$ and $\mathbf{J}_{\mathrm{L}}^{*u} \mathbf{J}_{\mathrm{L}}^{u}$ are less than $\mathbf{I}_{\mathrm{N}}$ and $\mathbf{I}_{\mathrm{L}}$ respectively. Therefore,

$$\left| \frac{1}{\mathrm{ML}} \mathrm{Tr}\left(\mathbf{CJ}_{\mathrm{N}}^{*u} \mathbf{C}^* (\mathbf{I}_{\mathrm{M}} \otimes \mathbf{J}_{\mathrm{L}}^{*u}) \right) \right| \leq \frac{1}{\mathrm{ML}} \mathrm{Tr}\left(\mathbf{CC}^* \right) \tag{3.4.12}$$

Using (3.4.12) for $\mathbf{C} = \mathbf{QBQY}$ for each u leads to

$$\alpha \leq \frac{\sigma^2}{\mathrm{MN}} \mathbb{E}\left[\frac{1}{\mathrm{ML}} \mathrm{Tr}\left(\mathbf{QBQYY}^* \mathbf{Q}^* \mathbf{B}^* \mathbf{Q}^* \right) \right]$$

The resolvent identity (3.1.34) can also be written as $\mathbf{QYY}^* = \mathbf{I} + z\mathbf{Q}$. This implies that the greatest eigenvalue of $\mathbf{QYY}^*\mathbf{Q}^*$ coincides with the greatest eigenvalue of $(\mathbf{I} + z\mathbf{Q})\mathbf{Q}^*$ which is itself less than $\|\mathbf{Q}\| + |z|\|\mathbf{Q}\|^2$. As $\|\mathbf{Q}\| \leq \frac{1}{\mathrm{Im}z}$, we obtain that

$$\mathbf{QYY}^* \mathbf{Q}^* \leq \frac{1}{\mathrm{Im}z} \left(1 + \frac{|z|}{\mathrm{Im}z} \right) \mathbf{I}. \tag{3.4.13}$$

Therefore, it holds that

$$\alpha \leq \frac{1}{\mathrm{Im}z} \left(1 + \frac{|z|}{\mathrm{Im}z} \right) \frac{1}{\mathrm{MN}} \mathbb{E}\left[\frac{1}{\mathrm{ML}} \mathrm{Tr}\left(\mathbf{QBB}^* \mathbf{Q}^* \right) \right] \tag{3.4.14}$$

We eventually obtain that $\alpha \leq \mathrm{C}(z) \kappa^2 \frac{1}{\mathrm{MN}}$ where $\mathrm{C}(z)$ is defined by

$$\mathrm{C}(z) = \frac{1}{(\mathrm{Im}z)^3} \left(1 + \frac{|z|}{\mathrm{Im}z} \right)$$

The conclusion follows from the observation that $\mathrm{C}(z)$ verifies

$$\mathrm{C}(z) \leq \left[\frac{1}{(\mathrm{Im}z)^3} + \frac{1}{(\mathrm{Im}z)^4} \right] (|z| + 1)$$

In order to prove (3.4.2) and (3.4.3), we remark that

$$\begin{aligned} \mathbf{b}_1^* \mathbf{Q} \mathbf{b}_2 &= \mathrm{ML} \frac{1}{\mathrm{ML}} \mathrm{Tr}\left(\mathbf{Qb}_2 \mathbf{b}_1^* \right) \\ \mathbf{u}_1^* \left[\frac{1}{\mathrm{M}} \sum_{m=1}^{\mathrm{M}} (\mathbf{Q}(z))^{m,m} \right] \mathbf{u}_2 &= \mathrm{L} \frac{1}{\mathrm{ML}} \mathrm{Tr}\left(\mathbf{Q}(\mathbf{I}_{\mathrm{M}} \otimes \mathbf{u}_2 \mathbf{u}_1^*) \right) \end{aligned}$$

(3.4.2) and (3.4.3) follow immediately from this and inequality (3.4.14) used in the case $\mathbf{B} = \mathbf{b}_2\mathbf{b}_1^*$ and $\mathbf{B} = \mathbf{I}_M \otimes \mathbf{u}_2\mathbf{u}_1^*$ respectively.

We provide a sketch of proof of (3.4.4), and omit the proof of (3.4.6) and (3.4.5) which can be obtained as above. We still denote by ξ the random variable $\xi = \frac{1}{ML}\mathrm{Tr}\,(\mathbf{Q}(z)\mathbf{YGY}^*)$, and only evaluate the behaviour of the right hand side α of (3.4.10) denoted α. After easy calculations, we obtain that

$$\alpha \leq \frac{2\sigma^2}{MN}\mathbb{E}\left[\frac{1}{ML}\mathrm{Tr}\left(\mathbf{QYGY}^*\mathbf{BQYY}^*\mathbf{Q}^*\mathbf{B}^*\mathbf{YG}^*\mathbf{Y}^*\mathbf{Q}^*\right)\right] + \tag{3.4.15}$$

$$\frac{2\sigma^2}{MN}\mathbb{E}\left[\frac{1}{ML}\mathrm{Tr}\left(\mathbf{G}^*\mathbf{Y}^*\mathbf{Q}^*\mathbf{B}^*\mathbf{BQYG}\right)\right] \tag{3.4.16}$$

The term defined by (3.4.16) is easy to handle because $\mathbf{Q}^*\mathbf{B}^*\mathbf{BQ} \leq \frac{\kappa^2}{(\mathrm{Im}(z))^2}\mathbf{I}$. Therefore, (3.4.16) is less than $\frac{2\sigma^2\kappa^2}{(\mathrm{Im}(z))^2}\frac{1}{MN}\mathbb{E}\left[\frac{1}{MN}\mathrm{Tr}\,(\mathbf{YGG}^*\mathbf{Y}^*)\right]$ which is itself lower bounded by $\frac{1}{MN}\frac{2\sigma^2\kappa^4}{(\mathrm{Im}(z))^2}$. To evaluate the righthandside of (3.4.15), we use (3.4.13) twice, and obtain immediately that is less than $\frac{C(z)\kappa^4}{MN}$.

We finally provide a sketch of proof of (3.4.7). We still denote by ξ the random variable $\xi = \frac{1}{ML}\mathrm{Tr}\left(\mathbf{Q}(z)\tilde{\mathbf{B}}\mathbf{Y}^*\right)$, and only evaluate the behaviour of the right hand side α of (3.4.10) still denoted α. We obtain that

$$\alpha \leq \frac{2\sigma^2}{MN}\mathbb{E}\left[\frac{1}{ML}\mathrm{Tr}\left(\mathbf{QYY}^*\mathbf{Q}^*\mathbf{Y}\tilde{\mathbf{B}}^*\tilde{\mathbf{B}}\mathbf{Y}^*\right)\right] + \tag{3.4.17}$$

$$\frac{2\sigma^2}{MN}\mathbb{E}\left[\frac{1}{ML}\mathrm{Tr}\left(\mathbf{Q}\tilde{\mathbf{B}}\tilde{\mathbf{B}}^*\mathbf{Q}^*)\right)\right] \tag{3.4.18}$$

For the same reasons that above, we obtain immediately that the term α is $\mathcal{O}(\frac{1}{MN})$ term. We omit the proof of (3.4.8) and (3.4.9), which can be obtained similarly.

Remark 8. *The above Poincaré-Nash evaluations imply that under Assumption A-2, all the functionals considered in Proposition 3.4.1 have almost surely the same asymptotic behaviour than their mathematical expectation. We just briefly justify that*

$$\frac{1}{ML}\mathrm{Tr}\,(\mathbf{BQ}(z)) - \mathbb{E}\left[\frac{1}{ML}\mathrm{Tr}\,(\mathbf{BQ}(z))\right] \to 0, \ a.s. \tag{3.4.19}$$

and

$$\mathbf{b}_1^*\mathbf{Q}(z)\mathbf{b}_2 - \mathbb{E}\left(\mathbf{b}_1^*\mathbf{Q}(z)\mathbf{b}_2\right) \to 0, \ a.s. \tag{3.4.20}$$

We denote by χ_1 the term defined by (3.4.19). Then, (3.4.1) leads to $\mathbb{E}(|\chi_1|^2) = \mathcal{O}(\frac{1}{N^{2-\alpha}})$, which, using the Markov inequality and the Borel-Cantelli lemma implies that χ_1 converges towards 0 almost surely. We denote by χ_2 the term defined by (3.4.20). This time, $\mathbb{E}(|\chi_2|^2) = \mathcal{O}(\frac{1}{M}) = \mathcal{O}(\frac{1}{N^{1-\alpha}})$, which does not allow to conclude that $\chi_2 \to 0$. However, using the same approach as in the proof of (2.2.6), it is possible to establish that $\mathbb{E}(|\chi_2|^{2k}) = \mathcal{O}\left(\frac{1}{M^k})\right) = \mathcal{O}\left(\frac{1}{N^{k(1-\alpha)}})\right)$ for each integer k. Choosing k in such a way that $k(1-\alpha) > 1$ leads to $\chi_2 \to 0$ almost surely.

This discussion shows that in order to evaluate the asymptotic behaviour of terms such as $\frac{1}{ML}\mathrm{Tr}\,(\mathbf{BQ}(z))$, $\mathbf{b}_1^*\mathbf{Q}(z)\mathbf{b}_2$, or $\mathbf{b}_1^*\mathbf{Q}(z)\tilde{\mathbf{B}}\mathbf{Y}^*\mathbf{b}_2$, it is sufficient to evaluate the behaviour of their mathematical expectation. This is the purpose of the next sections.

□

3.5 Expression of matrices $\mathbb{E}(\mathbf{Q})$ and $\mathbb{E}(\tilde{\mathbf{Q}})$ obtained using the integration by parts formula

In this section, we use the integration by parts formula in order to express $\mathbb{E}(\mathbf{Q}(z)), \mathbb{E}(\tilde{\mathbf{Q}}(z))$ as terms which will appear to be close from $\mathbf{T}(z), \tilde{\mathbf{T}}(z)$ where we recall that $\mathbf{T}(z), \tilde{\mathbf{T}}(z)$ represent the solutions of the equation (3.3.4). For this, we have first to introduce useful matrix valued functions of the complex variable z and to study their properties.

Lemma 3.5.1. *For each $z \in \mathbb{C}^+$, matrices $\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbb{E}(\mathbf{Q}(z)))$ and $\mathbf{I}_{ML} + \sigma^2 \mathbf{I}_M \otimes \mathcal{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}(z)))$ are invertible. We denote by $\mathbf{H}(z)$ and $\tilde{\mathbf{H}}(z)$ their inverses, i.e.*

$$\begin{cases} \mathbf{H}(z) = \left[\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbb{E}(\mathbf{Q}(z)))\right]^{-1} \\ \tilde{\mathbf{H}}(z) = \left[\mathbf{I}_{ML} + \sigma^2 \mathbf{I}_M \otimes \mathcal{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}))\right]^{-1} \end{cases} \tag{3.5.1}$$

Then, functions $z \to \mathbf{H}(z)$ and $z \to \tilde{\mathbf{H}}(z)$ are holomorphic in $\mathbb{C}^+$ and verify

$$\begin{cases} \mathbf{H}(z)\mathbf{H}(z)^* \leq \left(\frac{|z|}{\operatorname{Im} z}\right)^2 \mathbf{I}_N \\ \tilde{\mathbf{H}}(z)\tilde{\mathbf{H}}(z)^* \leq \left(\frac{|z|}{\operatorname{Im} z}\right)^2 \mathbf{I}_{ML} \end{cases} \tag{3.5.2}$$

Similarly, for each $z \in \mathbb{C}^+$, these following matrices

$$\begin{cases} \mathbf{R}(z) = \left[-z\left(\mathbf{I}_{ML} + \sigma^2 \mathbf{I}_M \otimes \mathcal{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}^T))\right) + \mathbf{A}\mathbf{H}^T\mathbf{A}^*\right]^{-1} \\ \tilde{\mathbf{R}}(z) = \left[-z\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbb{E}(\mathbf{Q}^T))\right) + \mathbf{A}^*\tilde{\mathbf{H}}^T\mathbf{A}\right]^{-1} \end{cases} \tag{3.5.3}$$

are well-defined and functions $z \to \mathbf{R}(z)$ and $z \to \tilde{\mathbf{R}}(z)$ are holomorphic in $\mathbb{C}^+$. Moreover, it exists positive matrix-valued measures $\boldsymbol{\mu}$ and $\tilde{\boldsymbol{\mu}}$ carried by $\mathbf{R}^+$ satisfying $\boldsymbol{\mu}(\mathbb{R}^+) = \mathbf{I}_{ML}$ and $\tilde{\boldsymbol{\mu}}(\mathbb{R}^+) = \mathbf{I}_N$, and for which

$$\mathbf{R}(z) = \int_{\mathbb{R}^+} \frac{d\boldsymbol{\mu}(\lambda)}{\lambda - z} \quad ; \tilde{\mathbf{R}}(z) = \int_{\mathbb{R}^+} \frac{d\tilde{\boldsymbol{\mu}}(\lambda)}{\lambda - z}$$

Therefore, it holds that

$$\begin{cases} \mathbf{R}(z)\mathbf{R}(z)^* \leq \left(\frac{1}{\operatorname{Im} z}\right)^2 \mathbf{I}_{ML} \\ \tilde{\mathbf{R}}(z)\tilde{\mathbf{R}}(z)^* \leq \left(\frac{1}{\operatorname{Im} z}\right)^2 \mathbf{I}_N \end{cases} \tag{3.5.4}$$

Finally, $\mathbf{H}, \tilde{\mathbf{H}}, \mathbf{R}, \tilde{\mathbf{R}}$ are analytic on $\mathbb{C} \setminus \mathbb{R}^+$.

Proof. The proof is sketched in the appendix.

We introduce the main result of this section :

Proposition 3.5.1. *The expectation of matrices $\mathbf{Q}$ and $\tilde{\mathbf{Q}}$ can be expressed as*

$$\begin{cases} \mathbb{E}(\mathbf{Q}) = \mathbf{R} + \boldsymbol{\Upsilon} \\ \mathbb{E}(\tilde{\mathbf{Q}}) = \tilde{\mathbf{R}} + \tilde{\boldsymbol{\Upsilon}} \end{cases} \tag{3.5.5}$$

where $\boldsymbol{\Upsilon}$ and $\tilde{\boldsymbol{\Upsilon}}$ are terms which converge towards zero in a certain sense.

Proof.
In order to be able the integration by parts formula, we use the identity (3.1.34) which implies that

$$\mathbb{E}\left[\mathbf{Q}_{i_1,i_2}^{m_1,m_2}\right] = -\frac{1}{z}\delta(i_1-i_2)\delta(m_1-m_2) + \frac{1}{z}\mathbb{E}\left[\left(\mathbf{Q}\mathbf{Y}\mathbf{Y}^*\right)_{i_1,i_2}^{m_1,m_2}\right] \tag{3.5.6}$$

We express $(\mathbf{Q}\mathbf{Y}\mathbf{Y}^*)_{i_1,i_2}^{m_1,m_2}$ as

$$\left(\mathbf{Q}\mathbf{Y}\mathbf{Y}^*\right)_{i_1,i_2}^{m_1,m_2} = \sum_{j=1}^{N}\left(\mathbf{Q}\mathbf{y}_j\mathbf{y}_j^*\right)_{i_1,i_2}^{m_1,m_2} = \sum_{j=1}^{N}\left(\mathbf{Q}\mathbf{y}_j\right)_{i_1}^{m_1}\overline{\mathbf{Y}}_{i_2,j}^{m_2}$$

where we recall that $(\mathbf{y}_j)_{j=1,\dots,N}$ represent the columns of $\mathbf{Y}$. In order to be able to evaluate $\mathbb{E}\left[\left(\mathbf{Q}\mathbf{y}_j\mathbf{y}_j^*\right)_{i_1,i_2}^{m_1,m_2}\right]$, it is necessary to express $\mathbb{E}\left[\left(\mathbf{Q}\mathbf{y}_k\mathbf{y}_j^*\right)_{i_1,i_2}^{m_1,m_2}\right] = \mathbb{E}\left[\left(\mathbf{Q}\mathbf{y}_k\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right]$ for each pair (k,j). For this, we use the identity

$$\mathbb{E}\left[\left(\mathbf{Q}\mathbf{y}_k\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right] = \mathbb{E}\left[\left(\mathbf{Q}\mathbf{w}_k\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right] + \mathbb{E}\left[\left(\mathbf{Q}\mathbf{a}_k\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right] \tag{3.5.7}$$

where $(\mathbf{w}_j)_{j=1,\dots,N}$ and $(\mathbf{a}_j)_{j=1,\dots,N}$ represent the columns of $\mathbf{W}$ and $\mathbf{A}$, and use the integration by parts formula to evaluate the 2 terms of the right-hand-side of (3.5.7). As for the first term :

$$\mathbb{E}\left[\left(\mathbf{Q}\mathbf{w}_k\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right] = \sum_{i_3,m_3}\mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3}\mathbf{W}_{i_3,k}^{m_3}\overline{\mathbf{Y}}_{i_2,j}^{m_2}\right) \tag{3.5.8}$$

We apply the integration by parts

$$\mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3}\mathbf{W}_{i_3,k}^{m_3}\overline{\mathbf{Y}}_{i_2,j}^{m_2}\right) = \sum_{i',j'}\mathbb{E}\left(\mathbf{W}_{i_3,k}^{m_3}\overline{\mathbf{W}}_{i',j'}^{m_3}\right)\mathbb{E}\left[\frac{\partial\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3}\overline{\mathbf{Y}}_{i_2,j}^{m_2}\right)}{\partial\overline{\mathbf{W}}_{i',j'}^{m_3}}\right]$$

It is easy to check that

$$\frac{\partial\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3}\overline{\mathbf{Y}}_{i_2,j}^{m_2}\right)}{\partial\overline{\mathbf{W}}_{i',j'}^{m_3}} = \mathbf{Q}_{i_1,i_3}^{m_1,m_3}\delta(m_2=m_3)\delta(i'=i_2)\delta(j=j') - \left(\mathbf{Q}\mathbf{y}_{j'}\right)_{i_1}^{m_1}\mathbf{Q}_{i',i_3}^{m_3,m_3}\overline{\mathbf{Y}}_{i_2,j}^{m_2}$$

(3.3.1) implies that $\mathbb{E}\left(\mathbf{W}_{i_3,k}^{m_3}\overline{\mathbf{W}}_{i',j'}^{m_3}\right) = \frac{\sigma^2}{N}\delta(i_3-i'=j'-k)$. Therefore, we obtain that

$$\begin{aligned}\mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3}\mathbf{W}_{i_3,k}^{m_3}\overline{\mathbf{Y}}_{i_2,j}^{m_2}\right) &= \frac{\sigma^2}{N}\delta(i_3-i_2=j-k)\delta(m_2=m_3)\mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3}\right)\\ &- \frac{\sigma^2}{N}\sum_{i',j'}\delta(i_3-i'=j'-k)\mathbb{E}\left[\left(\mathbf{Q}\mathbf{y}_{j'}\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\mathbf{Q}_{i',i_3}^{m_3,m_3}\right]\end{aligned}$$

and that

$$\sum_{i_3,m_3}\mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3}\mathbf{W}_{i_3,k}^{m_3}\overline{\mathbf{Y}}_{i_2,j}^{m_2}\right) = \frac{\sigma^2}{N}\sum_{i_3,m_3}\delta(i_3-i_2=j-k)\delta(m_2=m_3)\mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3}\right) \tag{3.5.9}$$

$$- \frac{\sigma^2}{N}\sum_{i_3,m_3}\sum_{i',j'}\delta(i_3-i'=j'-k)\mathbb{E}\left[\left(\mathbf{Q}\mathbf{y}_{j'}\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\mathbf{Q}_{i',i_3}^{m_3,m_3}\right] \tag{3.5.10}$$

The term on the right-hand-side of (3.5.9) is reduced to :

$$\frac{\sigma^2}{N}\mathbb{E}\left(\mathbf{Q}_{i_1,i_2-(k-j)}^{m_1,m_2}\right)\mathbb{1}_{1\le i_2-(k-j)\le L} \tag{3.5.11}$$

As for the term of (3.5.10), we put $i = i' - i_3$, and get that

$$-\sigma^2 c_N \sum_{i=-(L-1)}^{L-1} \mathbb{1}_{1\le k-i\le N}\,\mathbb{E}\left[\left(\mathbf{Q}\mathbf{y}_{k-i}\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\frac{1}{ML}\sum_{i'-i_3=i}\sum_{m_3}\mathbf{Q}_{i',i_3}^{m_3,m_3}\right] \tag{3.5.12}$$

Using the definition (3.1.3), (3.5.12) is equal to

$$-\sigma^2 c_N \sum_{i=-(L-1)}^{L-1} \mathbb{1}_{1\le k-i\le N}\,\mathbb{E}\left[\tau^{(M)}(\mathbf{Q})(i)\left(\mathbf{Q}\mathbf{y}_{k-i}\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right] \tag{3.5.13}$$

Plugging ((3.5.11),(3.5.13)) into ((3.5.9),(3.5.10)) and using (3.5.8), we obtain that

$$\begin{aligned}\mathbb{E}\left[\left(\mathbf{Q}\mathbf{w}_k\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right] &= \frac{\sigma^2}{N}\mathbb{E}\left(\mathbf{Q}_{i_1,i_2-(k-j)}^{m_1,m_2}\right)\mathbb{1}_{1\le i_2-(k-j)\le L}\\ &-\sigma^2 c_N \sum_{i=-(L-1)}^{L-1} \mathbb{1}_{1\le k-i\le N}\,\mathbb{E}\left[\tau^{(M)}(\mathbf{Q})(i)\left(\mathbf{Q}\mathbf{y}_{k-i}\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right]\end{aligned} \tag{3.5.14}$$

Setting $u = k - i$, the second term of the right-hand-side of (3.5.14) can be written as

$$-\sigma^2 c_N \mathbb{E}\left[\sum_{u=1}^{N}\tau^{(M)}(\mathbf{Q})(k-u)\mathbb{1}_{-(L-1)\le k-u\le L-1}\left(\mathbf{Q}\mathbf{y}_u\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right]$$

or, using the observation that $\tau^{(M)}(\mathbf{Q})(k-u)\mathbb{1}_{-(L-1)\le k-u\le L-1} = \left(\mathscr{T}_{N,L}^{(M)}(\mathbf{Q})\right)_{k,u}$ (see Eq. (3.1.5)), as

$$-\sigma^2 c_N\,\mathbb{E}\left[\mathbf{e}_k^T\mathscr{T}_{N,L}^{(M)}(\mathbf{Q})\begin{pmatrix}\left(\mathbf{Q}\mathbf{y}_1\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\\ \left(\mathbf{Q}\mathbf{y}_2\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\\ \vdots\\ \left(\mathbf{Q}\mathbf{y}_N\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\end{pmatrix}\right]$$

We express matrix $\mathbf{Q}$ as $\mathbf{Q} = \mathbb{E}(\mathbf{Q})+\mathbf{Q}^\circ$ and define the following $N\times N$ matrices $\mathbf{X}_{i_1,i_2}^{m_1,m_2}, \mathbf{B}_{i_1,i_2}^{m_1,m_2}, \mathbf{B}^{\circ\,m_1,m_2}_{i_1,i_2}, \mathbf{U}_{i_1,i_2}^{m_1,m_2}$

$$\left(\mathbf{X}_{i_1,i_2}^{m_1,m_2}\right)_{k,j} = \mathbb{E}\left[\left(\mathbf{Q}\mathbf{y}_k\right)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right] \tag{3.5.15}$$

$$\left(\mathbf{B}_{i_1,i_2}^{m_1,m_2}\right)_{k,j} = \mathbb{E}\left[\mathbf{Q}_{i_1,i_2-(k-j)}^{m_1,m_2}\mathbb{1}_{1\le i_2-(k-j)\le L}\right] \tag{3.5.16}$$

$$\left(\mathbf{B}^{\circ\,m_1,m_2}_{i_1,i_2}\right)_{k,j} = \mathbf{Q}^{\circ\,m_1,m_2}_{i_1,i_2-(k-j)}\mathbb{1}_{1\le i_2-(k-j)\le L} \tag{3.5.17}$$

$$\mathbf{U}_{i_1,i_2}^{m_1,m_2} = -\sigma^2 c_N\,\mathbb{E}\left[\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^\circ)\begin{pmatrix}\left(\mathbf{Q}\mathbf{y}_1\right)_{i_1}^{m_1}\\ \left(\mathbf{Q}\mathbf{y}_2\right)_{i_1}^{m_1}\\ \vdots\\ \left(\mathbf{Q}\mathbf{y}_N\right)_{i_1}^{m_1}\end{pmatrix}\begin{pmatrix}\left(\mathbf{y}_1^*\right)_{i_2}^{m_2} & \left(\mathbf{y}_2^*\right)_{i_2}^{m_2} & \dots & \left(\mathbf{y}_N^*\right)_{i_2}^{m_2}\end{pmatrix}\right]$$

We notice that matrix

$$\left[\begin{pmatrix} (\mathbf{Q}\mathbf{y}_1)_{i_1}^{m_1} \\ (\mathbf{Q}\mathbf{y}_2)_{i_1}^{m_1} \\ \vdots \\ (\mathbf{Q}\mathbf{y}_N)_{i_1}^{m_1} \end{pmatrix} \begin{pmatrix} (\mathbf{y}_1^*)_{i_2}^{m_2} & (\mathbf{y}_2^*)_{i_2}^{m_2} & \dots & (\mathbf{y}_N^*)_{i_2}^{m_2} \end{pmatrix}\right]$$

can also be written as

$$\begin{pmatrix} \mathbf{y}_1^T \mathbf{Q}^T \\ \vdots \\ \mathbf{y}_N^T \mathbf{Q}^T \end{pmatrix} \left(\mathbf{f}_{i_1}^{m_1}\right) \left(\mathbf{f}_{i_2}^{m_2}\right)^T (\overline{\mathbf{y}}_1, \dots, \overline{\mathbf{y}}_N)$$

or as

$$\mathbf{Y}^T \mathbf{Q}^T \left(\mathbf{f}_{i_1}^{m_1}\right) \left(\mathbf{f}_{i_2}^{m_2}\right)^T \overline{\mathbf{Y}}$$

Therefore,

$$\mathbf{U}_{i_1,i_2}^{m_1,m_2} = -\sigma^2 c_N \, \mathbb{E}\left[\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^\circ)\, \mathbf{Y}^T \mathbf{Q}^T \left(\mathbf{f}_{i_1}^{m_1}\right) \left(\mathbf{f}_{i_2}^{m_2}\right)^T \overline{\mathbf{Y}}\right] \tag{3.5.18}$$

It is useful to notice that matrix $\mathbf{B}_{i_1,i_2}^{m_1,m_2}$ is a band Toeplitz matrix whose (k,l) element is zero if $|k-l| \geq L$. It is clear that Eq. (3.5.14) is equivalent to

$$\mathbb{E}\left[(\mathbf{Q}\mathbf{w}_k)_{i_1}^{m_1} \left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right] = \frac{\sigma^2}{N}\left(\mathbf{B}_{i_1,i_2}^{m_1,m_2}\right)_{k,j} - \sigma^2 c_N \left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}))\mathbf{X}_{i_1,i_2}^{m_1,m_2}\right)_{k,j} + \left(\mathbf{U}_{i_1,i_2}^{m_1,m_2}\right)_{k,j} \tag{3.5.19}$$

Now we evaluate the second term of the right-hand-side of (3.5.7). It is clear that

$$\mathbb{E}\left[(\mathbf{Q}\mathbf{a}_k)_{i_1}^{m_1} \left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right] = \mathbb{E}\left[(\mathbf{Q}\mathbf{a}_k)_{i_1}^{m_1} \left(\mathbf{a}_j^*\right)_{i_2}^{m_2}\right] + \mathbb{E}\left[(\mathbf{Q}\mathbf{a}_k)_{i_1}^{m_1} \left(\mathbf{w}_j^*\right)_{i_2}^{m_2}\right] \tag{3.5.20}$$

We define $N \times N$ matrix $\mathbf{D}_{i_1,i_2}^{m_1,m_2}$ as

$$\left(\mathbf{D}_{i_1,i_2}^{m_1,m_2}\right)_{k,j} = \mathbb{E}\left[(\mathbf{Q}\mathbf{a}_k)_{i_1}^{m_1} \left(\mathbf{a}_j^*\right)_{i_2}^{m_2}\right] \tag{3.5.21}$$

We can see easily that

$$\mathbb{E}\left[(\mathbf{Q}\mathbf{a}_k)_{i_1}^{m_1} \left(\mathbf{w}_j^*\right)_{i_2}^{m_2}\right] = \sum_{i_3,m_3} \mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3} \mathbf{A}_{i_3,k}^{m_3} \overline{\mathbf{W}}_{i_2,j}^{m_2}\right) \tag{3.5.22}$$

Applying the integration by parts formula, we get that

$$\mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3} \mathbf{A}_{i_3,k}^{m_3} \overline{\mathbf{W}}_{i_2,j}^{m_2}\right) = \sum_{i',j'} \mathbb{E}\left(\overline{\mathbf{W}}_{i_2,j}^{m_2} \mathbf{W}_{i',j'}^{m_2}\right) \mathbb{E}\left[\frac{\partial\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3} \mathbf{A}_{i_3,k}^{m_3}\right)}{\partial \mathbf{W}_{i',j'}^{m_2}}\right]$$

It is easy to check that

$$\frac{\partial\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3} \mathbf{A}_{i_3,k}^{m_3}\right)}{\partial \mathbf{W}_{i',j'}^{m_2}} = -\mathbf{Q}_{i_1,i'}^{m_1,m_2} \mathbf{e}_{j'}^T \mathbf{Y}^* \mathbf{Q} \mathbf{f}_{i_3}^{m_3} \mathbf{A}_{i_3,k}^{m_3}$$

(3.3.1) implies that $\mathbb{E}\left(\overline{\mathbf{W}}_{i_2,j}^{m_2} \mathbf{W}_{i',j'}^{m_2}\right) = \frac{\sigma^2}{N}\delta(i_2 - i' = j' - j)$. Therefore

$$\sum_{m_3,i_3} \mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{m_1,m_3} \mathbf{A}_{i_3,k}^{m_3} \overline{\mathbf{W}}_{i_2,j}^{m_2}\right) = -\frac{\sigma^2}{N} \sum_{i',j'} \delta(i_2 - i' = j' - j)\mathbb{E}\left[\mathbf{Q}_{i_1,i'}^{m_1,m_2} \mathbf{e}_{j'}^T \mathbf{Y}^* \mathbf{Q} \sum_{m_3,i_3} \mathbf{f}_{i_3}^{m_3} \mathbf{A}_{i_3,k}^{m_3}\right]$$

It is easy to check that $\sum_{m_3,i_3} \mathbf{A}^{m_3}_{i_3,k}\mathbf{f}^{m_3}_{i_3} = \mathbf{a}_k = \mathbf{A}\mathbf{e}_k$, by (3.5.22), the last equation leads to

$$\mathbb{E}\left[(\mathbf{Q}\mathbf{a}_k)^{m_1}_{i_1}\left(\mathbf{w}^*_j\right)^{m_2}_{i_2}\right] = -\frac{\sigma^2}{N}\sum_{i',j'}\delta(i_2 - i' = j' - j)\mathbb{E}\left[\mathbf{Q}^{m_1,m_2}_{i_1,i'}\mathbf{e}^T_{j'}\mathbf{Y}^*\mathbf{Q}\mathbf{A}\mathbf{e}_k\right]$$

We put $i = i_2 - i'$ in the above sum and get that

$$\begin{aligned}\mathbb{E}\left[(\mathbf{Q}\mathbf{a}_k)^{m_1}_{i_1}\left(\mathbf{w}^*_j\right)^{m_2}_{i_2}\right] &= -\frac{\sigma^2}{N}\sum_{i=-(L-1)}^{L-1}\mathbb{1}_{1\le i_2-i\le L}\mathbb{E}\left[\mathbf{Q}^{m_1,m_2}_{i_1,i_2-i}\mathbb{1}_{1\le i+j\le N}\mathbf{e}^T_{i+j}\mathbf{Y}^*\mathbf{Q}\mathbf{A}\mathbf{e}_k\right] \\ &= -\frac{\sigma^2}{N}\sum_{i=-(L-1)}^{L-1}\mathbb{1}_{1\le i_2-i\le L}\mathbb{E}\left[\mathbf{Q}^{m_1,m_2}_{i_1,i_2-i}\mathbb{1}_{1\le i+j\le N}(\mathbf{Y}^*\mathbf{Q}\mathbf{A})_{i+j,k}\right] \end{aligned} \tag{3.5.23}$$

By decorrelating the term

$$\mathbb{E}\left[\mathbf{Q}^{m_1,m_2}_{i_1,i_2-i}(\mathbf{Y}^*\mathbf{Q}\mathbf{A})_{i+j,k}\right] = \mathbb{E}(\mathbf{Q}^{m_1,m_2}_{i_1,i_2-i})\mathbb{E}\left[(\mathbf{Y}^*\mathbf{Q}\mathbf{A})_{i+j,k}\right] + \mathbb{E}\left[\mathbf{Q}^{\circ\, m_1,m_2}_{\ i_1,i_2-i}(\mathbf{Y}^*\mathbf{Q}\mathbf{A})_{i+j,k}\right]$$

We can see that it is necessary to calculate the expectation of matrix $\mathbf{Y}^*\mathbf{Q}\mathbf{A} = \mathbf{W}^*\mathbf{Q}\mathbf{A} + \mathbf{A}^*\mathbf{Q}\mathbf{A}$. For this, for all $1 \le p, q \le N$, we can easily see that

$$\mathbb{E}\left[(\mathbf{W}^*\mathbf{Q}\mathbf{A})_{p,q}\right] = \sum_{m_3,i_3}\mathbb{E}\left[(\mathbf{Q}\mathbf{a}_q)^{m_3}_{i_3}(\mathbf{w}^*_p)^{m_3}_{i_3}\right]$$

Using (3.5.23) when $m_1 = m_2 = m_3$, $i_1 = i_2 = i_3$, $k = q$ and $j = p$, we get that

$$\mathbb{E}\left[(\mathbf{Q}\mathbf{a}_q)^{m_3}_{i_3}(\mathbf{w}^*_p)^{m_3}_{i_3}\right] = -\frac{\sigma^2}{N}\sum_{i=-(L-1)}^{L-1}\mathbb{E}\left[\mathbf{Q}^{m_3,m_3}_{i_3,i_3-i}\mathbb{1}_{1\le i_3-i\le L}(\mathbf{Y}^*\mathbf{Q}\mathbf{A})_{i+p,q}\mathbb{1}_{1\le p+i\le N}\right]$$

Therefore

$$\mathbb{E}\left[(\mathbf{W}^*\mathbf{Q}\mathbf{A})_{p,q}\right] = -\sigma^2 c_N\sum_{i=-(L-1)}^{L-1}\mathbb{E}\left[\frac{1}{ML}\sum_{m_3,i_3}\mathbf{Q}^{m_3,m_3}_{i_3,i_3-i}\mathbb{1}_{1\le i_3-i\le L}(\mathbf{Y}^*\mathbf{Q}\mathbf{A})_{i+p,q}\mathbb{1}_{1\le p+i\le N}\right]$$

As $\frac{1}{ML}\sum_{m_3,i_3}\mathbf{Q}^{m_3,m_3}_{i_3,i_3-i}\mathbb{1}_{1\le i_3-i\le L} = \tau^{(M)}(\mathbf{Q})(i)$ (see Eq. (3.1.3)), we obtain that

$$\mathbb{E}\left[(\mathbf{W}^*\mathbf{Q}\mathbf{A})_{p,q}\right] = -\sigma^2 c_N\sum_{i=-(L-1)}^{L-1}\mathbb{E}\left[\tau^{(M)}(\mathbf{Q})(i)(\mathbf{Y}^*\mathbf{Q}\mathbf{A})_{i+p,q}\mathbb{1}_{1\le p+i\le N}\right]$$

Setting $r = p + i$, the above equation can be written as

$$\mathbb{E}\left[(\mathbf{W}^*\mathbf{Q}\mathbf{A})_{p,q}\right] = -\sigma^2 c_N\sum_{r=1}^{N}\mathbb{E}\left(\tau^{(M)}(\mathbf{Q})(r-p)\mathbb{1}_{-(L-1)\le r-p\le L-1}(\mathbf{Y}^*\mathbf{Q}\mathbf{A})_{r,q}\right)$$

Moreover, it holds that $\tau^{(M)}(\mathbf{Q})(r-p)\mathbb{1}_{-(L-1)\le r-p\le L-1} = \left(\mathcal{T}^{(M)}_{N,L}\mathbb{E}(\mathbf{Q})\right)_{r,p}$ (see Eq. (3.1.5)), and using $\mathbf{Q} = \mathbb{E}(\mathbf{Q}) + \mathbf{Q}^\circ$, we get

$$\mathbb{E}\left[(\mathbf{W}^*\mathbf{Q}\mathbf{A})_{p,q}\right] = -\sigma^2 c_N\sum_{r=1}^{N}\left(\mathcal{T}^{(M)}_{N,L}(\mathbb{E}(\mathbf{Q}))\right)_{r,p}\mathbb{E}\left((\mathbf{Y}^*\mathbf{Q}\mathbf{A})_{r,q}\right) - \sigma^2 c_N\sum_{r=1}^{N}\mathbb{E}\left((\mathcal{T}^{(M)}_{N,L}(\mathbf{Q}^\circ))_{r,p}(\mathbf{Y}^*\mathbf{Q}\mathbf{A})_{r,q}\right)$$

As a consequence, we obtain that

$$\begin{aligned}\mathbb{E}\left[(\mathbf{W}^*\mathbf{QA})_{p,q}\right] &= -\sigma^2 c_N \left[\mathcal{T}_{N,L}^{(M)}\left(\mathbb{E}(\mathbf{Q}^T)\right)\mathbb{E}(\mathbf{Y}^*\mathbf{QA})\right]_{p,q} - \sigma^2 c_N \left[\mathbb{E}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QA}\right)\right]_{p,q} \\ &= -\sigma^2 c_N \left[\mathcal{T}_{N,L}^{(M)}\left(\mathbb{E}(\mathbf{Q}^T)\right)\left(\mathbb{E}(\mathbf{W}^*\mathbf{QA}) + \mathbb{E}(\mathbf{A}^*\mathbf{QA})\right)\right]_{p,q} - \sigma^2 c_N \left[\mathbb{E}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QA}\right)\right]_{p,q}\end{aligned} \tag{3.5.24}$$

In order to evaluate matrix $\mathbb{E}(\mathbf{W}^*\mathbf{QA})$, we express (3.5.24) as

$$\left[\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}\left(\mathbb{E}(\mathbf{Q}^T)\right)\right]\mathbb{E}(\mathbf{W}^*\mathbf{QA}) = -\sigma^2 c_N \mathcal{T}_{N,L}^{(M)}\mathbb{E}(\mathbf{Q}^T)\mathbb{E}(\mathbf{A}^*\mathbf{QA}) - \sigma^2 c_N \mathbb{E}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QA}\right) \tag{3.5.25}$$

Lemma 3.5.1 implies that matrix $\left[\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}(z)))\right]$ is invertible for each $z \in \mathbb{C}^+$, and we recall that its inverse is denoted by $\mathbf{H}(z)$. (3.5.25) leads to

$$\mathbb{E}(\mathbf{W}^*\mathbf{QA}) = -\sigma^2 c_N \mathbf{H}^T \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\mathbb{E}(\mathbf{A}^*\mathbf{QA}) - \sigma^2 c_N \mathbf{H}^T \mathbb{E}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QA}\right)$$

We recall that $\mathbf{Y}^*\mathbf{QA} = \mathbf{W}^*\mathbf{QA} + \mathbf{A}^*\mathbf{QA}$. Therefore, the above equation can be expressed as

$$\mathbb{E}(\mathbf{Y}^*\mathbf{QA}) = \left[\mathbf{I}_N - \sigma^2 c_N \mathbf{H}^T \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\right]\mathbb{E}(\mathbf{A}^*\mathbf{QA}) - \sigma^2 c_N \mathbf{H}^T \mathbb{E}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QA}\right)$$

Using that $\mathbf{H}^T = \left[\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}\left(\mathbb{E}(\mathbf{Q}^T)\right)\right]^{-1}$, we obtain immediately that $\mathbf{I}_N - \sigma^2 c_N \mathbf{H}^T \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T)) = \mathbf{H}^T$. Therefore

$$\mathbb{E}(\mathbf{Y}^*\mathbf{QA}) = \mathbf{H}^T\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A} - \sigma^2 c_N \mathbf{H}^T \mathbb{E}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QA}\right) \tag{3.5.26}$$

Recalling (3.5.23) and setting $r = i + j$, we get

$$\begin{aligned}\mathbb{E}\left[(\mathbf{Qa}_k)_{i_1}^{m_1}\left(\mathbf{w}_j^*\right)_{i_2}^{m_2}\right] &= -\frac{\sigma^2}{N}\sum_{i=-(L-1)}^{L-1} \mathbb{1}_{1\le i_2-i\le L}\mathbb{E}\left[\mathbf{Q}_{i_1,i_2-i}^{m_1,m_2}\mathbb{1}_{1\le i+j\le N}(\mathbf{Y}^*\mathbf{QA})_{i+j,k}\right] \\ &= -\frac{\sigma^2}{N}\sum_{r=1}^{N}\mathbb{E}\left[\mathbf{Q}_{i_1,i_2-(r-j)}^{m_1,m_2}\mathbb{1}_{1\le i_2-(r-j)\le L}(\mathbf{Y}^*\mathbf{QA})_{r,k}\right] \\ &= -\frac{\sigma^2}{N}\sum_{r=1}^{N}\mathbb{E}(\mathbf{Q}_{i_1,i_2-(r-j)}^{m_1,m_2})\mathbb{1}_{1\le i_2-(r-j)\le L}\mathbb{E}(\mathbf{Y}^*\mathbf{QA})_{r,k} - \frac{\sigma^2}{N}\sum_{r=1}^{N}\mathbb{E}\left[\mathbf{Q}^{\circ}{}_{i_1,i_2-(r-j)}^{m_1,m_2}\mathbb{1}_{1\le i_2-(r-j)\le L}(\mathbf{Y}^*\mathbf{QA})_{r,k}\right]\end{aligned}$$

We plug (3.5.26) into the expression of $\mathbb{E}(\mathbf{Y}^*\mathbf{QA})$ in the above equation, we get

$$\begin{aligned}\mathbb{E}\left[(\mathbf{Qa}_k)_{i_1}^{m_1}\left(\mathbf{w}_j^*\right)_{i_2}^{m_2}\right] = &-\frac{\sigma^2}{N}\sum_{r=1}^{N}\mathbb{E}\left(\mathbf{Q}_{i_1,i_2-(r-j)}^{m_1,m_2})\mathbb{1}_{1\le i_2-(r-j)\le L}\right)\left(\mathbf{H}^T\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A} - \sigma^2 c_N \mathbf{H}^T\mathbb{E}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QA}\right)\right)_{r,k} \\ &-\frac{\sigma^2}{N}\sum_{r=1}^{N}\mathbb{E}\left[\mathbf{Q}^{\circ}{}_{i_1,i_2-(r-j)}^{m_1,m_2}\mathbb{1}_{1\le i_2-(r-j)\le L}(\mathbf{Y}^*\mathbf{QA})_{r,k}\right]\end{aligned}$$

Using the definitions of $\mathbf{B}_{i_1,i_2}^{m_1,m_2}$ and $\mathbf{B}^{\circ}{}_{i_1,i_2}^{m_1,m_2}$ in ((3.5.16), (3.5.17)), we obtain that

$$
\begin{aligned}
\mathbb{E}\left[(\mathbf{Qa}_k)_{i_1}^{m_1}\left(\mathbf{w}_j^*\right)_{i_2}^{m_2}\right] &= -\frac{\sigma^2}{\mathrm{N}}\sum_{r=1}^{\mathrm{N}}\left(\mathbf{B}_{i_1,i_2}^{m_1,m_2}\right)_{r,j}\left(\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}-\sigma^2 c_{\mathrm{N}}\mathbf{H}^{\mathrm{T}}\mathbb{E}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\circ\mathrm{T}})\mathbf{Y}^*\mathbf{QA}\right)\right)_{r,k} \\
&\quad -\frac{\sigma^2}{\mathrm{N}}\sum_{r=1}^{\mathrm{N}}\mathbb{E}\left[(\mathbf{B}^{\circ}{}_{i_1,i_2}^{m_1,m_2})_{r,j}(\mathbf{Y}^*\mathbf{QA})_{r,k}\right] \\
&= -\frac{\sigma^2}{\mathrm{N}}\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\right)_{j,k}+\frac{\sigma^4}{\mathrm{N}}c_{\mathrm{N}}\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbb{E}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\circ\mathrm{T}})\mathbf{Y}^*\mathbf{QA}\right)\right)_{j,k} \\
&\quad -\frac{\sigma^2}{\mathrm{N}}\mathbb{E}\left((\mathbf{B}^{\circ}{}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{Y}^*\mathbf{QA}\right)_{j,k} \\
&= -\frac{\sigma^2}{\mathrm{N}}\left(\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\right)^{\mathrm{T}}\right)_{k,j}+\frac{\sigma^4}{\mathrm{N}}c_{\mathrm{N}}\left(\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbb{E}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\circ\mathrm{T}})\mathbf{Y}^*\mathbf{QA}\right)\right)^{\mathrm{T}}\right)_{k,j} \\
&\quad -\frac{\sigma^2}{\mathrm{N}}\mathbb{E}\left(\left((\mathbf{B}^{\circ}{}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{Y}^*\mathbf{QA}\right)^{\mathrm{T}}\right)_{k,j}
\end{aligned}
\tag{3.5.27}
$$

Finally, recall that

$$
\mathbb{E}\left[(\mathbf{Qy}_k)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right]=\mathbb{E}\left[(\mathbf{Qw}_k)_{i_1}^{m_1}\left(\mathbf{y}_j^*\right)_{i_2}^{m_2}\right]+\mathbb{E}\left[(\mathbf{Qa}_k)_{i_1}^{m_1}\left(\mathbf{w}_j^*\right)_{i_2}^{m_2}\right]+\mathbb{E}\left[(\mathbf{Qa}_k)_{i_1}^{m_1}\left(\mathbf{a}_j^*\right)_{i_2}^{m_2}\right]
$$

By (3.5.15), (3.5.19), (3.5.21) and (3.5.27), we can rewrite the above equation as

$$
\begin{aligned}
\left(\mathbf{X}_{i_1,i_2}^{m_1,m_2}\right)_{k,j} &= \frac{\sigma^2}{\mathrm{N}}\left(\mathbf{B}_{i_1,i_2}^{m_1,m_2}\right)_{k,j}-\sigma^2 c_{\mathrm{N}}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}))\mathbf{X}_{i_1,i_2}^{m_1,m_2}\right)_{k,j}+\left(\mathbf{U}_{i_1,i_2}^{m_1,m_2}\right)_{k,j} \\
&\quad -\frac{\sigma^2}{\mathrm{N}}\left(\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\right)^{\mathrm{T}}\right)_{k,j}+\frac{\sigma^4}{\mathrm{N}}c_{\mathrm{N}}\left(\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbb{E}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\circ\mathrm{T}})\mathbf{Y}^*\mathbf{QA}\right)\right)^{\mathrm{T}}\right)_{k,j} \\
&\quad -\frac{\sigma^2}{\mathrm{N}}\mathbb{E}\left(\left((\mathbf{B}^{\circ}{}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{Y}^*\mathbf{QA}\right)^{\mathrm{T}}\right)_{k,j}+\left(\mathbf{D}_{i_1,i_2}^{m_1,m_2}\right)_{k,j}
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
\left[\mathbf{I}_{\mathrm{N}}+\sigma^2 c_{\mathrm{N}}\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}))\right]\mathbf{X}_{i_1,i_2}^{m_1,m_2} &= \frac{\sigma^2}{\mathrm{N}}\mathbf{B}_{i_1,i_2}^{m_1,m_2}+\left(\mathbf{U}_{i_1,i_2}^{m_1,m_2}\right)-\frac{\sigma^2}{\mathrm{N}}\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\right)^{\mathrm{T}} \\
&\quad +\frac{\sigma^4}{\mathrm{N}}c_{\mathrm{N}}\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbb{E}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\circ\mathrm{T}})\mathbf{Y}^*\mathbf{QA}\right)\right)^{\mathrm{T}}-\frac{\sigma^2}{\mathrm{N}}\mathbb{E}\left((\mathbf{B}^{\circ}{}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{Y}^*\mathbf{QA}\right)^{\mathrm{T}}+\mathbf{D}_{i_1,i_2}^{m_1,m_2}
\end{aligned}
$$

Using that $\mathbf{H}=\left[\mathbf{I}_{\mathrm{N}}+\sigma^2 c_{\mathrm{N}}\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}))\right]^{-1}$, matrix $\mathbf{X}_{i_1,i_2}^{m_1,m_2}$ is equal to

$$
\begin{aligned}
\mathbf{X}_{i_1,i_2}^{m_1,m_2} &= \frac{\sigma^2}{\mathrm{N}}\mathbf{HB}_{i_1,i_2}^{m_1,m_2}+\mathbf{HU}_{i_1,i_2}^{m_1,m_2}-\frac{\sigma^2}{\mathrm{N}}\mathbf{H}\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\right)^{\mathrm{T}} \\
&\quad +\frac{\sigma^4}{\mathrm{N}}c_{\mathrm{N}}\mathbf{H}\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbb{E}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\circ\mathrm{T}})\mathbf{Y}^*\mathbf{QA}\right)\right)^{\mathrm{T}}-\frac{\sigma^2}{\mathrm{N}}\mathbf{H}\left(\mathbb{E}\left((\mathbf{B}^{\circ}{}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{Y}^*\mathbf{QA}\right)\right)^{\mathrm{T}}+\mathbf{HD}_{i_1,i_2}^{m_1,m_2}
\end{aligned}
\tag{3.5.28}
$$

The term $\mathbb{E}(\mathbf{QYY}^*)_{i_1,i_2}^{m_1,m_2}$ coincides with $\mathrm{Tr}\left(\mathbf{X}_{i_1,i_2}^{m_1,m_2}\right)$. Therefore, taking the trace of both side of (3.5.28), we obtain

$$
\begin{aligned}
\mathbb{E}\left(\mathbf{QYY}^*\right)_{i_1,i_2}^{m_1,m_2} &= \frac{\sigma^2}{\mathrm{N}}\mathrm{Tr}\left(\mathbf{H}\,\mathbf{B}_{i_1,i_2}^{m_1,m_2}\right)+\mathrm{Tr}\left(\mathbf{H}\,\mathbf{U}_{i_1,i_2}^{m_1,m_2}\right)-\frac{\sigma^2}{\mathrm{N}}\mathrm{Tr}\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{AH}^{\mathrm{T}}\right) \\
&\quad +\frac{\sigma^4}{\mathrm{N}}c_{\mathrm{N}}\mathrm{Tr}\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbb{E}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\circ\mathrm{T}})\mathbf{Y}^*\mathbf{QA}\right)\mathbf{H}^{\mathrm{T}}\right)-\frac{\sigma^2}{\mathrm{N}}\mathrm{Tr}\left(\mathbb{E}\left((\mathbf{B}^{\circ}{}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{Y}^*\mathbf{QA}\right)\mathbf{H}^{\mathrm{T}}\right)+\mathrm{Tr}\left(\mathbf{HD}_{i_1,i_2}^{m_1,m_2}\right)
\end{aligned}
\tag{3.5.29}
$$

We can easily compute $\mathrm{Tr}\left(\mathbf{H}\,\mathbf{D}_{i_1,i_2}^{m_1,m_2}\right)$, where we recall $\mathbf{D}_{i_1,i_2}^{m_1,m_2}$ is defined by (3.5.21)

$$\begin{aligned}
\mathrm{Tr}\left(\mathbf{H}\mathbf{D}_{i_1,i_2}^{m_1,m_2}\right) &= \sum_{(k,j)=1}^{\mathrm{N}} \mathbb{E}\left[(\mathbf{Q}\mathbf{a}_k)_{i_1}^{m_1}(\mathbf{a}_j^*)_{i_2}^{m_2}\mathbf{H}_{j,k}\right] \\
&= \sum_{(k,j)=1}^{\mathrm{N}} \mathbb{E}\left[(\mathbf{Q}\mathbf{A}\mathbf{e}_k)_{i_1}^{m_1}\mathbf{H}_{k,j}^{\mathrm{T}}(\mathbf{e}_j^{\mathrm{T}}\mathbf{A}^*)_{i_2}^{m_2}\right] \\
&= \mathbb{E}\left(\mathbf{Q}\mathbf{A}\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\right)_{i_1,i_2}^{m_1,m_2}
\end{aligned}$$

As matrix $\mathbf{B}_{i_1,i_2}^{m_1,m_2}$ is Toeplitz, it holds that (see Eq. (3.1.7))

$$\frac{1}{\mathrm{N}}\mathrm{Tr}\left(\mathbf{H}\,\mathbf{B}_{i_1,i_2}^{m_1,m_2}\right) = \sum_{u=-(\mathrm{N}-1)}^{\mathrm{N}-1} \tau(\mathbf{H})(u)\mathbb{E}\left(\mathbf{Q}_{i_1,i_2+u}^{m_1,m_2}\right)\mathbb{1}_{1\le i_2+u\le \mathrm{L}}$$

which also coincides with

$$\frac{1}{\mathrm{N}}\mathrm{Tr}\left(\mathbf{H}\,\mathbf{B}_{i_1,i_2}^{m_1,m_2}\right) = \sum_{u=-(\mathrm{L}-1)}^{\mathrm{L}-1} \tau(\mathbf{H})(u)\mathbb{E}\left(\mathbf{Q}_{i_1,i_2+u}^{m_1,m_2}\right)\mathbb{1}_{1\le i_2+u\le \mathrm{L}}$$

because $\mathbb{1}_{1\le i_2+u\le \mathrm{L}} = 0$ if $|u| \ge \mathrm{L}$. Setting $v = i_2 + u$, this term can be written as

$$\frac{1}{\mathrm{N}}\mathrm{Tr}\left(\mathbf{H}\,\mathbf{B}_{i_1,i_2}^{m_1,m_2}\right) = \sum_{v=1}^{\mathrm{L}} \mathbb{E}\left(\mathbf{Q}_{i_1,v}^{m_1,m_2}\right)\tau(\mathbf{H})(v-i_2)$$

or, using definition (3.1.5), as

$$\begin{aligned}
\frac{1}{\mathrm{N}}\mathrm{Tr}\left(\mathbf{H}\,\mathbf{B}_{i_1,i_2}^{m_1,m_2}\right) &= \sum_{v=1}^{\mathrm{L}} \mathbb{E}\left(\mathbf{Q}_{i_1,v}^{m_1,m_2}\right)\left(\mathscr{T}_{\mathrm{L,L}}(\mathbf{H})\right)_{v,i_2} \\
&= \left(\mathbb{E}(\mathbf{Q}^{m_1,m_2})\mathscr{T}_{\mathrm{L,L}}(\mathbf{H})\right)_{i_1,i_2}
\end{aligned}$$

As the entry (i_1,i_2) of matrix $\mathbb{E}(\mathbf{Q}^{m_1,m_2})\,\mathscr{T}_{\mathrm{L,L}}(\mathbf{H})$ coincides with $\left(\mathbb{E}(\mathbf{Q})\,\mathbf{I}_{\mathrm{M}}\otimes\mathscr{T}_{\mathrm{L,L}}(\mathbf{H})\right)_{i_1,i_2}^{m_1,m_2}$, it holds that

$$\frac{1}{\mathrm{N}}\mathrm{Tr}\left(\mathbf{H}\,\mathbf{B}_{i_1,i_2}^{m_1,m_2}\right) = \left(\mathbb{E}(\mathbf{Q})\,\mathbf{I}_{\mathrm{M}}\otimes\mathscr{T}_{\mathrm{L,L}}(\mathbf{H})\right)_{i_1,i_2}^{m_1,m_2}$$

Computing in the same way we get the following relations

- $$\begin{aligned}\frac{1}{\mathrm{N}}\mathrm{Tr}\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^{\mathrm{T}}\right) &= \left(\mathbb{E}(\mathbf{Q}^{m_1,m_2})\mathscr{T}_{\mathrm{L,L}}\left(\mathbf{H}\mathbf{A}^{\mathrm{T}}\mathbb{E}(\mathbf{Q}^{\mathrm{T}})\overline{\mathbf{A}}\mathbf{H}\right)\right)_{i_1,i_2} \\ &= \left(\mathbb{E}(\mathbf{Q})\;\left(\mathbf{I}_{\mathrm{M}}\otimes\mathscr{T}_{\mathrm{L,L}}(\mathbf{H}\mathbf{A}^{\mathrm{T}}\mathbb{E}(\mathbf{Q}^{\mathrm{T}})\overline{\mathbf{A}}\mathbf{H})\right)\right)_{i_1,i_2}^{m_1,m_2}\end{aligned}$$

- $$\begin{aligned}\frac{1}{\mathrm{N}}\mathrm{Tr}\left((\mathbf{B}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\,\mathbb{E}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\circ\mathrm{T}})\mathbf{Y}^*\mathbf{Q}\mathbf{A}\right)\mathbf{H}^{\mathrm{T}}\right) &= \left(\mathbb{E}(\mathbf{Q}^{m_1,m_2})\mathscr{T}_{\mathrm{L,L}}\left(\mathbf{H}\mathbb{E}(\mathbf{A}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\overline{\mathbf{Y}}\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\circ}))\mathbf{H}\right)\right)_{i_1,i_2} \\ &= \left(\mathbb{E}(\mathbf{Q})\;\left(\mathbf{I}_{\mathrm{M}}\otimes\mathscr{T}_{\mathrm{L,L}}\left(\mathbf{H}\mathbb{E}\left(\mathbf{A}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\overline{\mathbf{Y}}\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\circ})\right)\mathbf{H}\right)\right)\right)_{i_1,i_2}^{m_1,m_2}\end{aligned}$$

- $$\begin{aligned}\frac{1}{\mathrm{N}}\mathrm{Tr}\left(\mathbb{E}\left((\mathbf{B}^{\circ}{}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{Y}^*\mathbf{Q}\mathbf{A}\right)\mathbf{H}^{\mathrm{T}}\right) &= \mathbb{E}\left(\mathbf{Q}^{\circ\,m_1,m_2}\mathscr{T}_{\mathrm{L,L}}(\mathbf{H}\mathbf{A}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\overline{\mathbf{Y}})\right)_{i_1,i_2} \\ &= \mathbb{E}\left(\mathbf{Q}^{\circ}\;\left(\mathbf{I}_{\mathrm{M}}\otimes\mathscr{T}_{\mathrm{L,L}}(\mathbf{H}\mathbf{A}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\overline{\mathbf{Y}})\right)\right)_{i_1,i_2}^{m_1,m_2}\end{aligned}$$

(3.5.18) implies that

$$\mathrm{Tr}\left(\mathbf{H}\,\mathbf{U}_{i_1,i_2}^{m_1,m_2}\right) = -\sigma^2 c_N \mathbb{E}\left[\mathbf{QY}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^\circ)\right)^T \mathbf{H}^T\mathbf{Y}^*\right]_{i_1,i_2}^{m_1,m_2}$$

To simplify the notations, we define

$$\boldsymbol{\Delta} = \sigma^4 c_N \mathbb{E}(\mathbf{Q})\left(\mathbf{I}_M \otimes \mathcal{T}_{L,L}\left(\mathbf{H}\mathbb{E}\left(\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}}\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^\circ)\right)\mathbf{H}\right)\right) - \sigma^2\mathbb{E}\left(\mathbf{Q}^\circ\ \left(\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\mathbf{HA}^T\mathbf{Q}^T\overline{\mathbf{Y}})\right)\right)$$
$$-\sigma^2 c_N \mathbb{E}\left(\mathbf{QY}\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{H}^T\mathbf{Y}^*\right) \tag{3.5.30}$$

Eq. (3.5.29) eventually leads to

$$\mathbb{E}\left(\mathbf{QYY}^*\right) = \sigma^2\mathbb{E}(\mathbf{Q})\ \left(\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\mathbf{H})\right) - \sigma^2\mathbb{E}(\mathbf{Q})\ \left(\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\mathbf{HA}^T\mathbb{E}(\mathbf{Q}^T)\overline{\mathbf{A}}\mathbf{H})\right) + \mathbb{E}(\mathbf{Q})\mathbf{AH}^T\mathbf{A}^* + \boldsymbol{\Delta} \tag{3.5.31}$$

By (3.1.34), it holds that $\mathbf{QYY}^* = \mathbf{I}_{ML} + z\mathbf{Q}$. Therefore, we deduce from (3.5.31) that

$$\mathbb{E}(\mathbf{Q})\ \left[-z\mathbf{I}_{ML} + \mathbf{AH}^T\mathbf{A}^*\right] = \mathbf{I}_{ML} + \sigma^2\mathbb{E}(\mathbf{Q})\left[\mathbf{I}_M\otimes\mathcal{T}_{L,L}\left(-\mathbf{H} + \mathbf{HA}^T\mathbb{E}(\mathbf{Q}^T)\overline{\mathbf{A}}\mathbf{H}\right)\right] - \boldsymbol{\Delta} \tag{3.5.32}$$

We claim that

Lemma 3.5.2. *The term* $-\mathbf{H} + \mathbf{HA}^T\mathbb{E}(\mathbf{Q}^T)\overline{\mathbf{A}}\mathbf{H}$ *is an approximation of* $z\mathbb{E}(\tilde{\mathbf{Q}}^T)$ *in a sense to be defined.*

Proof. To prove lemma 3.5.2, we calculate the expectation of $\tilde{\mathbf{Q}} = (\mathbf{Y}^*\mathbf{Y} - z\mathbf{I}_N)^{-1}$. For this we remark that $\mathbf{Y}^*\mathbf{QY} = \tilde{\mathbf{Q}}\mathbf{Y}^*\mathbf{Y} = z\tilde{\mathbf{Q}} + \mathbf{I}$. The above calculate allow to compute $\mathbb{E}(\mathbf{Y}^*\mathbf{QY})$. Indeed, for $1 \le j,k \le N$, we express

$$\mathbb{E}(\mathbf{Y}^*\mathbf{QY})_{j,k} = \sum_{m,i}\mathbb{E}[(\mathbf{y}_j^*)_i^m(\mathbf{Qy}_k)_i^m] = \sum_{m,i}(\mathbf{X}_{i,i}^{m,m})_{k,j} \tag{3.5.33}$$

Using the equation (3.5.28) when $m_1 = m_2 = m$ and $i_1 = i_2 = i$, we can express $\mathbf{X}_{i,i}^{m,m}$

$$\mathbf{X}_{i,i}^{m,m} = \frac{\sigma^2}{N}\mathbf{HB}_{i,i}^{m,m} + \mathbf{HU}_{i,i}^{m,m} - \frac{\sigma^2}{N}\mathbf{H}\left((\mathbf{B}_{i,i}^{m,m})^T\mathbf{H}^T\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\right)^T$$
$$+\frac{\sigma^4}{N}c_N\mathbf{H}\left((\mathbf{B}_{i,i}^{m,m})^T\mathbf{H}^T\mathbb{E}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QA}\right)\right)^T - \frac{\sigma^2}{N}\mathbf{H}\left(\mathbb{E}\left((\mathbf{B}^{\circ}{}_{i,i}^{m,m})^T\mathbf{Y}^*\mathbf{QA}\right)\right)^T + \mathbf{HD}_{i,i}^{m,m}$$

Summing over m and i, we can notice that

- $\frac{1}{ML}\sum_{m,i}\left(\mathbf{B}_{i,i}^{m,m}\right)_{k,j} = \frac{1}{ML}\sum_{m,i}\mathbb{E}\left[\mathbf{Q}_{i,i-(k-j)}^{m,m}\mathbb{1}_{1\le i-(k-j)\le L}\right] = \tau^{(M)}(\mathbb{E}(\mathbf{Q}))(k-j) = \left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}))\right)_{k,j}$
- $\frac{1}{ML}\sum_{m,i}\left(\mathbf{B}^{\circ}{}_{i,i}^{m,m}\right)_{k,j} = \frac{1}{ML}\sum_{m,i}\left[\mathbf{Q}^{\circ}{}_{i,i-(k-j)}^{m,m}\mathbb{1}_{1\le i-(k-j)\le L}\right] = \tau^{(M)}(\mathbf{Q}^\circ)(k-j) = \left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^\circ)\right)_{k,j}$
- $\sum_{m,i}\left(\mathbf{U}_{i,i}^{m,m}\right)_{k,j} = -\sigma^2 c_N\sum_{m,i}\mathbb{E}\left[\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^\circ)\,\mathbf{Y}^T\mathbf{Q}^T\left(\mathbf{f}_i^m\right)\left(\mathbf{f}_i^m\right)^T\overline{\mathbf{Y}}\right]_{k,j} = -\sigma^2 c_N\mathbb{E}\left[\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^\circ)\,\mathbf{Y}^T\mathbf{Q}^T\overline{\mathbf{Y}}\right]_{k,j}$
- $\sum_{m,i}\left(\mathbf{D}_{i,i}^{m,m}\right)_{k,j} = \sum_{m,i}\mathbb{E}\left[(\mathbf{Qa}_k)_i^m(\mathbf{a}_j^*)_i^m\right] = \sum_{m,i}\mathbb{E}\left[(\mathbf{a}_j^*)_i^m(\mathbf{Qa}_k)_i^m\right] = \mathbb{E}\left[\mathbf{A}^*\mathbf{QA}\right]_{j,k} = \mathbb{E}\left[\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{A}}\right]_{k,j}$

So that we eventually get

$$\sum_{m,i}\mathbf{X}_{i,i}^{m,m} = \sigma^2 c_N\mathbf{H}\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q})) - \sigma^2 c_N\mathbb{E}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^\circ)\,\mathbf{Y}^T\mathbf{Q}^T\overline{\mathbf{Y}}\right) - \sigma^2 c_N\mathbf{H}\left(\mathcal{T}_{N,L}^{(M)}\left(\mathbb{E}(\mathbf{Q}^T)\right)\mathbf{H}^T\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\right)^T$$
$$+\sigma^4(c_N)^2\mathbf{H}\left(\mathcal{T}_{N,L}^{(M)}\left(\mathbb{E}(\mathbf{Q}^T)\right)\mathbf{H}^T\mathbb{E}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QA}\right)\right)^T - \sigma^2 c_N\mathbf{H}\left(\mathbb{E}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QA}\right)\right)^T + \mathbf{HA}^T\mathbb{E}(\mathbf{Q}^T)\overline{\mathbf{A}}$$

By (3.5.33), we get

$$\mathbb{E}(\mathbf{Y}^*\mathbf{Q}\mathbf{Y})_{j,k} = \sigma^2 c_N \left(\mathbf{H}\mathscr{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}))\right)_{k,j} - \sigma^2 c_N \left[\mathbf{H}\mathbb{E}\left(\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^\circ)\mathbf{Y}^T\mathbf{Q}^T\overline{\mathbf{Y}}\right)\right]_{k,j}$$
$$-\sigma^2 c_N \left[\mathbf{H}\left(\mathscr{T}_{N,L}^{(M)}\left(\mathbb{E}(\mathbf{Q}^T)\right)\mathbf{H}^T\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\right)^T\right]_{k,j} + \sigma^4(c_N)^2\left[\mathbf{H}\left(\mathscr{T}_{N,L}^{(M)}\left(\mathbb{E}(\mathbf{Q}^T)\right)\mathbf{H}^T\mathbb{E}\left(\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{Q}\mathbf{A}\right)\right)^T\right]_{k,j}$$
$$-\sigma^2 c_N \left[\mathbf{H}\left(\mathbb{E}\left(\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{Q}\mathbf{A}\right)\right)^T\right]_{k,j} + \left(\mathbf{H}\mathbf{A}^T\mathbb{E}(\mathbf{Q}^T)\overline{\mathbf{A}}\right)_{k,j}$$

So that we can obtain

$$\begin{aligned}\mathbb{E}(\mathbf{Y}^*\mathbf{Q}\mathbf{Y}) = {}& \sigma^2 c_N \mathscr{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\mathbf{H}^T + \mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^T - \sigma^2 c_N\left(\mathscr{T}_{N,L}^{(M)}\left(\mathbb{E}(\mathbf{Q}^T)\right)\right)\mathbf{H}^T\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^T \\ & -\sigma^2 c_N\mathbb{E}\left(\mathbf{Y}^*\mathbf{Q}\mathbf{Y}\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\right)\mathbf{H}^T + \sigma^4(c_N)^2\left(\mathscr{T}_{N,L}^{(M)}\left((\mathbb{E}(\mathbf{Q}^T))\right)\right)\mathbf{H}^T\mathbb{E}\left(\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{Q}\mathbf{A}\right)\mathbf{H}^T \\ & -\sigma^2 c_N\mathbb{E}\left(\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{Q}\mathbf{A}\right)\mathbf{H}^T\end{aligned} \tag{3.5.34}$$

To short the notations, we define

$$\begin{aligned}\tilde{\boldsymbol{\Delta}} = {}& -\sigma^2 c_N\mathbb{E}\left(\mathbf{Y}^*\mathbf{Q}\mathbf{Y}\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\right)\mathbf{H}^T + \sigma^4(c_N)^2\left(\mathscr{T}_{N,L}^{(M)}\left(\mathbb{E}(\mathbf{Q}^T)\right)\right)\mathbf{H}^T\mathbb{E}\left(\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{Q}\mathbf{A}\right)\mathbf{H}^T \\ & -\sigma^2 c_N\mathbb{E}\left(\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{Q}\mathbf{A}\right)\mathbf{H}^T\end{aligned} \tag{3.5.35}$$

Recalling that $\mathbf{Y}^*\mathbf{Q}\mathbf{Y} = \tilde{\mathbf{Q}}\mathbf{Y}^*\mathbf{Y} = z\tilde{\mathbf{Q}} + \mathbf{I}$, (3.5.34) can be rewritten as

$$z\mathbb{E}(\tilde{\mathbf{Q}}) = -\mathbf{I}_N + \sigma^2 c_N\left(\mathscr{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\right)\mathbf{H}^T + \mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^T - \sigma^2 c_N\mathscr{T}_{N,L}^{(M)}\left(\mathbb{E}(\mathbf{Q}^T)\right)\mathbf{H}^T\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^T + \tilde{\boldsymbol{\Delta}}$$

which can be factorized into

$$\begin{aligned}z\mathbb{E}(\tilde{\mathbf{Q}}) &= \left[-\mathbf{I}_N + \sigma^2 c_N\mathscr{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\mathbf{H}^T\right]\left[\mathbf{I}_N - \mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^T\right] + \tilde{\boldsymbol{\Delta}} \\ &= \left[-(\mathbf{H}^T)^{-1} + \sigma^2 c_N\mathscr{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\right]\mathbf{H}^T\left[\mathbf{I}_N - \mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^T\right] + \tilde{\boldsymbol{\Delta}}\end{aligned}$$

Using the definition of $\mathbf{H} = \left(\mathbf{I}_N + \sigma^2 c_N\mathscr{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}))\right)^{-1}$, $z\mathbb{E}(\tilde{\mathbf{Q}})$ is reduced to

$$z\mathbb{E}(\tilde{\mathbf{Q}}) = -\mathbf{H}^T\left[\mathbf{I}_N - \mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^T\right] + \tilde{\boldsymbol{\Delta}}$$

We obtain thus the relation

$$-\mathbf{H} + \mathbf{H}\mathbf{A}^T\mathbb{E}(\mathbf{Q}^T)\overline{\mathbf{A}}\mathbf{H} = z\mathbb{E}(\tilde{\mathbf{Q}}^T) - \tilde{\boldsymbol{\Delta}}^T \tag{3.5.36}$$

Equation (3.5.36) shows us that $z\mathbb{E}(\tilde{\mathbf{Q}}^T)$ can be approximated by $-\mathbf{H} + \mathbf{H}\mathbf{A}^T\mathbb{E}(\mathbf{Q}^T)\overline{\mathbf{A}}\mathbf{H}$ since $\tilde{\boldsymbol{\Delta}}$ is an error term with its normalized trace and bilinear form converge towards zero. This point will be proved in next section 3.6. □

We resume our previous calculations. Plugging (3.5.36) into (3.5.32), we get

$$\mathbb{E}(\mathbf{Q})\left[-z\mathbf{I}_{ML} + \mathbf{A}\mathbf{H}^T\mathbf{A}^*\right] = \mathbf{I}_{ML} + \sigma^2\mathbb{E}(\mathbf{Q})\left[\mathbf{I}_M \otimes \mathscr{T}_{L,L}(z\mathbb{E}(\tilde{\mathbf{Q}}^T) - \tilde{\boldsymbol{\Delta}}^T)\right] - \boldsymbol{\Delta}$$

which leads to the equation

$$\mathbb{E}(\mathbf{Q})\left[-z\left(\mathbf{I}_{ML} + \sigma^2\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}^T))\right) + \mathbf{A}\mathbf{H}^T\mathbf{A}^*\right] = \mathbf{I}_{ML} - \sigma^2\mathbb{E}(\mathbf{Q})(\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\boldsymbol{\Delta}}^T)) - \boldsymbol{\Delta}$$

Lemma 3.5.1 implies that matrix $\left[-z\left(\mathbf{I}_{ML}+\sigma^2\mathbf{I}_M\otimes\mathscr{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}^T))\right)+\mathbf{A}\mathbf{H}^T\mathbf{A}^*\right]$ is invertible for each $z\in\mathbb{C}^+$, and we recall that its inverse is denoted by $\mathbf{R}(z)$. We obtain that:

$$\begin{aligned}\mathbb{E}(\mathbf{Q}) &= \mathbf{R}-\sigma^2\mathbb{E}(\mathbf{Q})\mathbf{I}_M\otimes\mathscr{T}_{L,L}(\tilde{\mathbf{\Delta}}^T)\mathbf{R}-\mathbf{\Delta}\mathbf{R}\\ &= \mathbf{R}+\mathbf{\Upsilon}(\mathbf{\Delta},\tilde{\mathbf{\Delta}}) \qquad (3.5.37)\end{aligned}$$

with

$$\mathbf{\Upsilon}(\mathbf{\Delta},\tilde{\mathbf{\Delta}})=-\sigma^2\mathbb{E}(\mathbf{Q})\mathbf{I}_M\otimes\mathscr{T}_{L,L}(\tilde{\mathbf{\Delta}}^T)\mathbf{R}-\mathbf{\Delta}\mathbf{R} \qquad (3.5.38)$$

Now, plugging Eq. (3.5.37) into Eq. (3.5.36), we eventually get

$$\mathbb{E}(\tilde{\mathbf{Q}})=\frac{1}{z}\left[-\mathbf{H}^T+\mathbf{H}^T\mathbf{A}^*\mathbf{R}\mathbf{A}\mathbf{H}^T+\mathbf{H}^T\mathbf{A}^*\mathbf{\Upsilon}(\mathbf{\Delta},\tilde{\mathbf{\Delta}})\mathbf{A}\mathbf{H}^T+\tilde{\mathbf{\Delta}}\right] \qquad (3.5.39)$$

By lemma 3.5.1, matrices

$$\tilde{\mathbf{H}}=\left[\mathbf{I}_{ML}+\sigma^2\mathbf{I}_M\otimes\mathscr{T}_{L,L}\left(\mathbb{E}(\tilde{\mathbf{Q}}(z))\right)\right]^{-1}$$

and

$$\tilde{\mathbf{R}}=\left[-z\left(\mathbf{I}_N+\sigma^2c_N\mathscr{T}^{(M)}_{N,L}(\mathbb{E}(\mathbf{Q}^T))\right)+\mathbf{A}^*\tilde{\mathbf{H}}^T\mathbf{A}\right]^{-1}$$

are well-defined for each $z\in\mathbb{C}^+$. Moreover, we claim that

Lemma 3.5.3. *The following equality holds*

$$-\mathbf{H}^T+\mathbf{H}^T\mathbf{A}^*\mathbf{R}\mathbf{A}\mathbf{H}^T=z\tilde{\mathbf{R}} \qquad (3.5.40)$$

Proof. We have that $\mathbf{R}\mathbf{A}\mathbf{H}^T=\tilde{\mathbf{H}}^T\mathbf{A}\tilde{\mathbf{R}}$ because it holds that

$$\begin{cases}\mathbf{R}(z)=\left[-z\left(\mathbf{I}_{ML}+\sigma^2\mathbf{I}_M\otimes\mathscr{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}^T))\right)+\mathbf{A}\mathbf{H}^T\mathbf{A}^*\right]^{-1}\\ \tilde{\mathbf{R}}(z)=\left[-z\left(\mathbf{I}_N+\sigma^2c_N\mathscr{T}^{(M)}_{N,L}(\mathbb{E}(\mathbf{Q}^T))\right)+\mathbf{A}^*\tilde{\mathbf{H}}^T\mathbf{A}\right]^{-1}\end{cases}\iff\begin{cases}\mathbf{R}^{-1}=-z(\tilde{\mathbf{H}}^T)^{-1}+\mathbf{A}\mathbf{H}^T\mathbf{A}^*\\ \tilde{\mathbf{R}}^{-1}=-z(\mathbf{H}^T)^{-1}+\mathbf{A}^*\tilde{\mathbf{H}}^T\mathbf{A}\end{cases}$$

which give us

$$\begin{cases}\mathbf{I}_{ML}=-z\mathbf{R}(\tilde{\mathbf{H}}^T)^{-1}+\mathbf{R}\mathbf{A}\mathbf{H}^T\mathbf{A}^* & (3.5.41)\\ \mathbf{I}_N=-z(\tilde{\mathbf{H}}^T)^{-1}\tilde{\mathbf{R}}+\mathbf{A}^*\tilde{\mathbf{H}}^T\mathbf{A}\tilde{\mathbf{R}} & (3.5.42)\end{cases}$$

Multiplying (3.5.42) by $\mathbf{R}\mathbf{A}\mathbf{H}^T$, we get

$$\mathbf{R}\mathbf{A}\mathbf{H}^T=-z\mathbf{R}\mathbf{A}\tilde{\mathbf{R}}+\mathbf{R}\mathbf{A}\mathbf{H}^T\mathbf{A}^*\tilde{\mathbf{H}}^T\mathbf{A}\tilde{\mathbf{R}}$$

By (3.5.41), it is clear that $\mathbf{R}\mathbf{A}\mathbf{H}^T\mathbf{A}^*=\mathbf{I}_{ML}+z\mathbf{R}(\tilde{\mathbf{H}}^T)^{-1}$. Plugging into the above equation we get

$$\mathbf{R}\mathbf{A}\mathbf{H}^T=-z\mathbf{R}\mathbf{A}\tilde{\mathbf{R}}+(\mathbf{I}_{ML}+z\mathbf{R}(\tilde{\mathbf{H}}^T)^{-1})\tilde{\mathbf{H}}^T\mathbf{A}\tilde{\mathbf{R}}=\mathbf{H}^T\mathbf{A}\tilde{\mathbf{R}}$$

Thus we have

$$-\mathbf{H}^T+\mathbf{H}^T\mathbf{A}^*\mathbf{R}\mathbf{A}\mathbf{H}^T=\mathbf{H}^T(-\mathbf{I}_N+\mathbf{A}^*\tilde{\mathbf{H}}^T\mathbf{A}\tilde{\mathbf{R}})=\mathbf{H}^T(-\tilde{\mathbf{R}}^{-1}+\mathbf{A}^*\tilde{\mathbf{H}}^T\mathbf{A})\tilde{\mathbf{R}}=\mathbf{H}^Tz(\mathbf{H}^T)^{-1}\tilde{\mathbf{R}}=z\tilde{\mathbf{R}}$$

It finishes the proof of the lemma. □

As a consequence, we can rewrite (3.5.39) as

$$\begin{aligned}\mathbb{E}(\tilde{\mathbf{Q}}) &= \tilde{\mathbf{R}} + \frac{1}{z}\left[\mathbf{H}^{\mathrm{T}}\mathbf{A}^{*}\boldsymbol{\Upsilon}(\boldsymbol{\Delta},\tilde{\boldsymbol{\Delta}})\mathbf{A}\mathbf{H}^{\mathrm{T}} + \tilde{\boldsymbol{\Delta}}\right] \\ &= \tilde{\mathbf{R}} + \tilde{\boldsymbol{\Upsilon}}(\boldsymbol{\Delta},\tilde{\boldsymbol{\Delta}}) \end{aligned} \tag{3.5.43}$$

with

$$\tilde{\boldsymbol{\Upsilon}}(\boldsymbol{\Delta},\tilde{\boldsymbol{\Delta}}) = \frac{1}{z}\left[\mathbf{H}^{\mathrm{T}}\mathbf{A}^{*}\boldsymbol{\Upsilon}(\boldsymbol{\Delta},\tilde{\boldsymbol{\Delta}})\mathbf{A}\mathbf{H}^{\mathrm{T}} + \tilde{\boldsymbol{\Delta}}\right] \tag{3.5.44}$$

By (3.5.37) and (3.5.43), we finally get that

$$\begin{cases}\mathbb{E}(\mathbf{Q}) = \mathbf{R} + \boldsymbol{\Upsilon}(\boldsymbol{\Delta},\tilde{\boldsymbol{\Delta}}) \\ \mathbb{E}(\tilde{\mathbf{Q}}) = \tilde{\mathbf{R}} + \tilde{\boldsymbol{\Upsilon}}(\boldsymbol{\Delta},\tilde{\boldsymbol{\Delta}})\end{cases}$$

It completes the proposition 3.5.1. □

3.6 Preliminary controls of the error terms $\boldsymbol{\Upsilon}, \tilde{\boldsymbol{\Upsilon}}$

In this section, to short the notation, we will drop $\boldsymbol{\Delta}$ and $\tilde{\boldsymbol{\Delta}}$ in the error terms $\boldsymbol{\Upsilon}(\boldsymbol{\Delta},\tilde{\boldsymbol{\Delta}})$ and $\tilde{\boldsymbol{\Upsilon}}(\boldsymbol{\Delta},\tilde{\boldsymbol{\Delta}})$, and just denote by $\boldsymbol{\Upsilon}$ and $\tilde{\boldsymbol{\Upsilon}}$. We evaluate the behaviour of various terms depending on $\boldsymbol{\Upsilon}$, i.e. normalized traces $\frac{1}{\mathrm{ML}}\mathrm{Tr}\boldsymbol{\Upsilon}\mathbf{B}$, quadratic forms $\mathbf{b}_1^*\boldsymbol{\Upsilon}\mathbf{b}_2$, quadratic forms of matrix $\hat{\boldsymbol{\Upsilon}} = \frac{1}{\mathrm{M}}\sum_{m=1}^{\mathrm{M}}\boldsymbol{\Upsilon}^{m,m}$. Using rough estimates based on the results of section 3.4 and the Schwartz inequality, we establish that the normalized traces are $\mathcal{O}(\frac{\mathrm{L}}{\mathrm{MN}})$, and that two other terms are $\mathcal{O}(\delta_{\mathrm{N}})$ where $\delta_{\mathrm{N}} = \sup\left(\frac{\mathrm{L}}{\mathrm{M}^2}, \sqrt{\frac{\mathrm{L}}{\mathrm{M}^3}}\right)$. We first establish the following proposition.

Proposition 3.6.1. *Let* $\mathbf{B}$ *and* $\tilde{\mathbf{B}}$ *be* $\mathrm{ML}\times\mathrm{ML}$ *and* $\mathrm{N}\times\mathrm{N}$ *matrices such that* $\sup_{\mathrm{N}}\|\mathbf{B}\| \le \kappa$ *and* $\sup_{\mathrm{N}}\|\tilde{\mathbf{B}}\| \le \kappa$. *Then, it holds that*

$$\left|\frac{1}{\mathrm{ML}}\mathrm{Tr}\boldsymbol{\Upsilon}\mathbf{B}\right| \le C(z)\frac{\mathrm{L}}{\mathrm{MN}}\kappa \tag{3.6.1}$$

$$\left|\frac{1}{\mathrm{N}}\mathrm{Tr}\tilde{\boldsymbol{\Upsilon}}\tilde{\mathbf{B}}\right| \le C(z)\frac{\mathrm{L}}{\mathrm{MN}}\kappa \tag{3.6.2}$$

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

Proof. We only prove (3.6.1), (3.6.2) can be obtained similarly. We first express $\frac{1}{\mathrm{ML}}\mathrm{Tr}\boldsymbol{\Upsilon}\mathbf{B}$, for this we recall that

$$\boldsymbol{\Upsilon} = -\sigma^2\mathbb{E}(\mathbf{Q})\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}\left(\tilde{\boldsymbol{\Delta}}^{\mathrm{T}}\right)\mathbf{R} - \boldsymbol{\Delta}\mathbf{R}$$

We get immediately that

$$\frac{1}{\mathrm{ML}}\mathrm{Tr}\boldsymbol{\Upsilon}\mathbf{B} = -\frac{\sigma^2}{\mathrm{ML}}\mathrm{Tr}\left[\mathbb{E}(\mathbf{Q})\left(\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\tilde{\boldsymbol{\Delta}}^{\mathrm{T}})\right)\mathbf{R}\mathbf{B}\right] - \frac{1}{\mathrm{ML}}\mathrm{Tr}\left(\boldsymbol{\Delta}\mathbf{R}\mathbf{B}\right)$$

We denote by γ_1 the term $-\frac{\sigma^2}{\mathrm{ML}}\mathrm{Tr}\left[\mathbb{E}(\mathbf{Q})\left(\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\tilde{\boldsymbol{\Delta}}^{\mathrm{T}})\right)\mathbf{R}\mathbf{B}\right]$ and by γ_2 the term $-\frac{1}{\mathrm{ML}}\mathrm{Tr}\left(\boldsymbol{\Delta}\mathbf{R}\mathbf{B}\right)$.

We first evaluate the term γ_1. Using (3.1.10), we can see that

$$\begin{aligned}\frac{1}{\mathrm{ML}}\mathrm{Tr}\left[\left(\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\tilde{\boldsymbol{\Delta}}^{\mathrm{T}})\right)\mathbf{R}\mathbf{B}\mathbb{E}(\mathbf{Q})\right] &= \frac{1}{\mathrm{ML}}\mathrm{Tr}\left[\left(\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\tilde{\boldsymbol{\Delta}}^{\mathrm{T}})\right)(\mathbf{R}\mathbf{B}\mathbb{E}(\mathbf{Q}))^{\mathrm{T}}\right] \\ &= \frac{1}{\mathrm{N}}\mathrm{Tr}\left[\tilde{\boldsymbol{\Delta}}\left(\mathcal{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{R}\mathbf{B}\mathbb{E}(\mathbf{Q}))\right)^{\mathrm{T}}\right]\end{aligned}$$

We recall the expression of $\tilde{\mathbf{\Delta}}$:

$$\begin{aligned}\tilde{\mathbf{\Delta}} &= -\sigma^2 c_N \mathbb{E}\left(\mathbf{Y}^*\mathbf{QY}\left(\mathscr{T}^{(M)}_{N,L}(\mathbf{Q}^{\circ T})\right)\right)\mathbf{H}^T + \sigma^4 (c_N)^2 \left(\mathscr{T}^{(M)}_{N,L}(\mathbb{E}(\mathbf{Q}^T)\right)\mathbf{H}^T \mathbb{E}\left(\mathscr{T}^{(M)}_{N,L}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QA}\right)\mathbf{H}^T \\ &\quad - \sigma^2 c_N \mathbb{E}\left(\mathscr{T}^{(M)}_{N,L}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QA}\right)\mathbf{H}^T \\ &= \tilde{\mathbf{\Delta}}_1 + \tilde{\mathbf{\Delta}}_2 + \tilde{\mathbf{\Delta}}_3\end{aligned}$$

So that we have

$$\begin{aligned}\gamma_1 &= -\frac{\sigma^2}{N}\mathrm{Tr}\left[\tilde{\mathbf{\Delta}}_1\left(\mathscr{T}^{(M)}_{N,L}(\mathbf{RB}\mathbb{E}(\mathbf{Q}))\right)^T\right] - \frac{\sigma^2}{N}\mathrm{Tr}\left[\tilde{\mathbf{\Delta}}_2\left(\mathscr{T}^{(M)}_{N,L}(\mathbf{RB}\mathbb{E}(\mathbf{Q}))\right)^T\right] - \frac{\sigma^2}{N}\mathrm{Tr}\left[\tilde{\mathbf{\Delta}}_3\left(\mathscr{T}^{(M)}_{N,L}(\mathbf{RB}\mathbb{E}(\mathbf{Q}))\right)^T\right] \\ &= \gamma_{1,1} + \gamma_{1,2} + \gamma_{1,3}\end{aligned}$$

We will prove that the term $\gamma_{1,3}$ is a $\mathscr{O}(\frac{L}{MN})$ quantity and omit to check $\gamma_{1,1}$ and $\gamma_{1,2}$. As we can see that

$$\gamma_{1,3} = \sigma^2 c_N \mathbb{E}\left[\frac{1}{ML}\mathrm{Tr}\left(\mathscr{T}^{(M)}_{N,L}(\mathbf{Q}^{\circ T})\mathbf{Y}^*\mathbf{QAH}^T\left(\mathscr{T}^{(M)}_{N,L}(\mathbf{RB}\mathbb{E}(\mathbf{Q}))\right)^T\right)\right]$$

Using Eq. (3.1.9) and the identity $\tau^{(M)}\left((\mathbf{Q}^\circ)^T\right)(-u) = \tau^{(M)}(\mathbf{Q}^\circ)(u)$, we get that

$$\gamma_{1,3} = \sigma^2 c_N \mathbb{E}\left[\sum_{u=-(L-1)}^{L-1}\tau^{(M)}(\mathbf{Q}^\circ)(u)\,\frac{1}{ML}\mathrm{Tr}\left(\mathbf{QAH}^T\left(\mathscr{T}^{(M)}_{N,L}(\mathbf{RB}\mathbb{E}(\mathbf{Q}))\right)^T\mathbf{J}_N^u\mathbf{Y}^*\right)\right]$$

We can easily notice that $\left\|\mathscr{T}^{(M)}_{N,L}(\mathbf{RB}\mathbb{E}(\mathbf{Q}))\right\| \leq \|\mathbf{R}\|\,\|\mathbf{B}\|\,\|\mathbb{E}(\mathbf{Q})\| \leq \frac{\kappa}{\mathrm{Im}(z)}$, since $\|\mathbf{R}\|$ and $\|\mathbb{E}(\mathbf{Q})\|$ are upper-bounded by $\frac{1}{\mathrm{Im}(z)}$ and $\|\mathbf{B}\| \leq \kappa$.

Equations (3.4.1), (3.4.7) imply that $\mathbb{E}\left|\tau^{(M)}(\mathbf{Q}^\circ)(-u)\right|^2$ and $\mathrm{Var}\left(\frac{1}{ML}\mathrm{Tr}\left(\mathbf{QAH}^T\left(\mathscr{T}^{(M)}_{N,L}(\mathbf{RB}\mathbb{E}(\mathbf{Q}))\right)^T\mathbf{J}_N^u\mathbf{Y}^*\right)\right)$ are upperbounded by $\frac{C(z)}{MN}\kappa^2$. The Cauchy-Schwartz inequality thus implies immediately that $\gamma_{1,3} \leq C(z)\kappa\frac{L}{MN}$.
We now evaluate the term γ_2, recall the expression of $\mathbf{\Delta}$

$$\begin{aligned}\mathbf{\Delta} &= \sigma^4 c_N \mathbb{E}(\mathbf{Q})\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}\left(\mathbf{H}\mathbb{E}\left(\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}}\mathscr{T}^{(M)}_{N,L}(\mathbf{Q}^\circ)\right)\mathbf{H}\right)\right) - \sigma^2\mathbb{E}\left(\mathbf{Q}^\circ\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\mathbf{HA}^T\mathbf{Q}^T\overline{\mathbf{Y}})\right)\right) \\ &\quad -\sigma^2 c_N \mathbb{E}\left(\mathbf{QY}\mathscr{T}^{(M)}_{N,L}(\mathbf{Q}^{\circ T})\mathbf{H}^T\mathbf{Y}^*\right) \\ &= \mathbf{\Delta}_1 + \mathbf{\Delta}_2 + \mathbf{\Delta}_3\end{aligned}$$

So that we have,

$$\gamma_2 = -\frac{1}{ML}\mathrm{Tr}\left[(\mathbf{\Delta}_1 + \mathbf{\Delta}_2 + \mathbf{\Delta}_3)\mathbf{RB}\right]$$

$\frac{1}{ML}\mathrm{Tr}\left[(\mathbf{\Delta}_2 + \mathbf{\Delta}_3)\mathbf{RB}\right]$ can be evaluated in the same way than $\gamma_{1,3}$. As for the term $\frac{1}{ML}\mathrm{Tr}(\mathbf{\Delta}_1\mathbf{RB})$, by (3.1.10) we notice that

$$\frac{1}{ML}\mathrm{Tr}\left[\mathbb{E}(\mathbf{Q})\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}\left(\mathbf{H}\mathbb{E}\left(\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}}\mathscr{T}^{(M)}_{N,L}(\mathbf{Q}^\circ)\right)\mathbf{H}\right)\right)\mathbf{RB}\right] = \frac{1}{N}\mathrm{Tr}\left[\mathbf{H}\mathbb{E}\left(\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}}\mathscr{T}^{(M)}_{N,L}(\mathbf{Q}^\circ)\right)\mathbf{H}\left(\mathscr{T}^{(M)}_{N,L}(\mathbf{RB}\mathbb{E}(\mathbf{Q}))\right)\right]$$

Expanding $\mathscr{T}^{(M)}_{N,L}(\mathbf{Q}^\circ)$ as $\sum_{u=-(L-1)}^{L-1}\tau(\mathbf{Q}^\circ)(u)\mathbf{J}^{u*}$, and using the same arguments than previously, we obtain that $\left|\frac{1}{ML}\mathrm{Tr}(\mathbf{\Delta}_1\mathbf{RB})\right| \leq C(z)\frac{L}{MN}\kappa$. Therefore (3.6.1) holds.
Recall that $\tilde{\mathbf{\Upsilon}} = \frac{1}{z}\left[\sigma^2\mathbf{H}^T\mathbf{A}^*\mathbf{\Upsilon}\mathbf{AH}^T + \tilde{\mathbf{\Delta}}\right]$. Since $\|\mathbf{H}\| \leq \frac{|z|}{\mathrm{Im}(z)}$ and $\sup_N \|\mathbf{A}\| < +\infty$, using the same arguments than previously, we can easily get that

$$\left|\frac{1}{N}\mathrm{Tr}\tilde{\mathbf{\Upsilon}}\tilde{\mathbf{B}}\right| \leq C(z)\frac{L}{MN}\kappa$$

We now evaluate the behaviour of quadratic forms of matrices $\mathbf{\Upsilon}$, $\hat{\mathbf{\Upsilon}}$ and $\tilde{\mathbf{\Upsilon}}$

Proposition 3.6.2. *Let* $\delta_N = \sup\left(\frac{L^2}{MN}, \sqrt{\frac{L}{M^3}}\right)$. *Let* $\mathbf{b}_1$ *and* $\mathbf{b}_2$ *2* ML*–dimensional vectors such that* $\sup_N \|\mathbf{b}_i\| \le \kappa$ *for* $i = 1,2$ *. Then, it holds that*

$$|\mathbf{b}_1^* \boldsymbol{\Upsilon} \mathbf{b}_2| \le \kappa^2 C(z)\,\delta_N \tag{3.6.3}$$

for each $z \in \mathbb{C}^+$.
Let $\mathbf{c}_i$*,* $i = 1,2$ *be 2 deterministic* L*–dimensional vectors such that* $\sup_N \|\mathbf{c}_i\| \le \kappa$*. Then, it holds that*

$$\left|\mathbf{c}_1^* \left(\frac{1}{M}\sum_{m=1}^{M} \boldsymbol{\Upsilon}^{m,m}\right)\mathbf{c}_2\right| \le \kappa^2 C(z)\delta_N \tag{3.6.4}$$

for each $z \in \mathbb{C}^+$.
Moreover, for $\tilde{\mathbf{b}}_1$ *and* $\tilde{\mathbf{b}}_2$ *2* N*–dimensional vectors such that* $\sup_N \|\tilde{\mathbf{b}}_i\| \le \kappa$ *for* $i = 1,2$*. Then, it holds that*

$$|\tilde{\mathbf{b}}_1^* \tilde{\boldsymbol{\Upsilon}} \tilde{\mathbf{b}}_2| \le \kappa^2 C(z)\,\delta_N \tag{3.6.5}$$

for each $z \in \mathbb{C}^+$*. Where* $C(z)$ *can be written as* $C(z) = P_1(|z|)\,P_2\left((\mathrm{Im} z)^{-1}\right)$ *for some nice polynomials* P_1 *and* P_2.

Proof. We recall that

$$\boldsymbol{\Upsilon} = -\sigma^2 \mathbb{E}(\mathbf{Q})\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\boldsymbol{\Delta}}^T)\right)\mathbf{R} - \boldsymbol{\Delta}\mathbf{R}$$

So that we can express

$$\mathbf{b}_1^* \boldsymbol{\Upsilon} \mathbf{b}_2 = -\sigma^2 \mathbf{b}_1^* \mathbb{E}(\mathbf{Q})\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\boldsymbol{\Delta}}^T)\right)\mathbf{R}\mathbf{b}_2 - \mathbf{b}_1^* \boldsymbol{\Delta}\mathbf{R}\mathbf{b}_2$$

By equation (3.1.10), we can see that

$$\begin{aligned}
\mathbf{b}_1^* \mathbb{E}(\mathbf{Q})\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\boldsymbol{\Delta}}^T)\right)\mathbf{R}\mathbf{b}_2 &= \mathrm{Tr}\left(\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\boldsymbol{\Delta}}^T)\right)\mathbf{R}\mathbf{b}_2\mathbf{b}_1^*\mathbb{E}(\mathbf{Q})\right) \\
&= \frac{ML}{N}\mathrm{Tr}\left(\tilde{\boldsymbol{\Delta}}^T \mathscr{T}_{N,L}^{(M)}(\mathbf{R}\mathbf{b}_2\mathbf{b}_1^*\mathbb{E}(\mathbf{Q}))\right) \\
&= c_N \sum_{u=-(L-1)}^{L-1} \frac{1}{ML}\mathrm{Tr}\left(\tilde{\boldsymbol{\Delta}}^T \mathbf{J}_N^{*u}\right)\left(\mathbf{b}_1^*\mathbb{E}(\mathbf{Q})\left(\mathbf{I}_M \otimes \mathbf{J}_L^u\right)\mathbf{R}\mathbf{b}_2\right)
\end{aligned}$$

We notice that $\left|\mathbf{b}_1^*\mathbb{E}(\mathbf{Q})\left(\mathbf{I}_M \otimes \mathbf{J}_L^u\right)\mathbf{R}\mathbf{b}_2\right| \le \frac{\kappa^2}{(\mathrm{Im}(z))^2}$. By (3.6.1), we have that

$$\left|\mathbf{b}_1^*\mathbb{E}(\mathbf{Q})\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\boldsymbol{\Delta}}^T)\right)\mathbf{R}\mathbf{b}_2\right| \le \kappa^2 C(z) L\frac{L}{MN} = \kappa^2 C(z)\frac{L^2}{MN}$$

As for the term $\mathbf{b}_1^*\boldsymbol{\Delta}\mathbf{R}\mathbf{b}_2$, we recall that $\boldsymbol{\Delta} = \boldsymbol{\Delta}_1 + \boldsymbol{\Delta}_2 + \boldsymbol{\Delta}_3$ (see Eq. (3.5.30)). We will control these terms separately, we have that

$$\begin{aligned}
\mathbf{b}_1^*\boldsymbol{\Delta}_2\mathbf{R}\mathbf{b}_2 &= -\sigma^2 \mathbf{b}_1^*\mathbb{E}\left(\mathbf{Q}^\circ\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\mathbf{H}\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}})\right)\right)\mathbf{R}\mathbf{b}_2 \\
&= -\sigma^2\mathbb{E}\left[\mathrm{Tr}\left(\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\mathbf{H}\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}})\right)\mathbf{R}\mathbf{b}_2\mathbf{b}_1^*\mathbf{Q}^\circ\right)\right] \\
&= -\sigma^2\mathbb{E}\left[\sum_{u=-(L-1)}^{L-1}\tau(\mathbf{H}\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}})(u)\mathrm{Tr}\left((\mathbf{Q}^\circ\left(\mathbf{I}_M \otimes \mathbf{J}_L^{*u}\right)\mathbf{R}\mathbf{b}_2\mathbf{b}_1^*\right)\right] \\
&= -\sigma^2\mathbb{E}\left[\sum_{u=-(L-1)}^{L-1}\tau(\mathbf{H}\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}})(u)(\mathbf{b}_1^*\mathbf{Q}^\circ\left(\mathbf{I}_M \otimes \mathbf{J}_L^{*u}\right)\mathbf{R}\mathbf{b}_2)\right]
\end{aligned}$$

(3.4.2, 3.4.7) and the Schwartz inequality lead immediately to

$$|\mathbf{b}_1^*\boldsymbol{\Delta}_2\mathbf{R}\mathbf{b}_2| \le \kappa^2 C(z) L\frac{1}{\sqrt{MN}}\sqrt{\frac{L}{N}} = \kappa^2 C(z)\sqrt{\frac{L}{M}}\frac{L}{N}$$

Using the same tricks we can prove that the term $\mathbf{b}_1^* \boldsymbol{\Delta}_3 \mathbf{R} \mathbf{b}_2$ is upperbounded by $\kappa^2 C(z)\sqrt{\frac{L}{M}\frac{L}{N}}$.
To evaluate the term $\mathbf{b}_1^* \boldsymbol{\Delta}_1 \mathbf{R} \mathbf{b}_2$, we can see that

$$\begin{aligned}
\mathbf{b}_1^* \boldsymbol{\Delta}_1 \mathbf{R} \mathbf{b}_2 &= \sigma^4 c_N \mathbf{b}_1^* \mathbb{E}(\mathbf{Q}) \left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}\left(\mathbf{H}\mathbb{E}(\mathbf{A}^T \mathbf{Q}^T \overline{\mathbf{Y}} \mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^\circ))\mathbf{H}\right)\right) \mathbf{R} \mathbf{b}_2 \\
&= \sigma^4 c_N \mathbb{E}\left[\mathrm{Tr}\left(\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}\left((\mathbf{H}\mathbf{A}^T \mathbf{Q}^T \overline{\mathbf{Y}} \mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^\circ)\mathbf{H}\right)\right) \mathbf{R}\mathbf{b}_2 \mathbf{b}_1^* \mathbb{E}(\mathbf{Q})\right)\right] \\
&= \sigma^4 c_N \mathbb{E}\left[\sum_{u=-(L-1)}^{L-1} \tau^{(M)}(\mathbf{Q}^\circ)(u) \mathrm{Tr}\left(\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}\left(\mathbf{H}\mathbf{A}^T \mathbf{Q}^T \overline{\mathbf{Y}} \mathbf{J}_N^{*u} \mathbf{H}\right)\right) \mathbf{R}\mathbf{b}_2 \mathbf{b}_1^* \mathbb{E}(\mathbf{Q})\right)\right] \\
&= \sigma^4 c_N \mathbb{E}\left[\sum_{u=-(L-1)}^{L-1} \sum_{l=-(L-1)}^{L-1} \tau^{(M)}(\mathbf{Q}^\circ)(u) \tau(\mathbf{H}\mathbf{A}^T \mathbf{Q}^T \overline{\mathbf{Y}} \mathbf{J}_N^{*u} \mathbf{H})(l) \mathrm{Tr}\left((\mathbf{I}_M \otimes \mathbf{J}_L^{*l}) \mathbf{R}\mathbf{b}_2 \mathbf{b}_1^* \mathbb{E}(\mathbf{Q})\right)\right] \\
&= \sigma^4 c_N \mathbb{E}\left[\sum_{u=-(L-1)}^{L-1} \sum_{l=-(L-1)}^{L-1} \tau^{(M)}(\mathbf{Q}^\circ)(u) \tau(\mathbf{H}\mathbf{A}^T \mathbf{Q}^T \overline{\mathbf{Y}} \mathbf{J}_N^{*u} \mathbf{H})(l)\, \mathbf{b}_1^* \left(\mathbb{E}(\mathbf{Q})(\mathbf{I}_M \otimes \mathbf{J}_L^{*l})\mathbf{R}\right) \mathbf{b}_2\right]
\end{aligned}$$

As $\left|\mathbf{b}_1^* (\mathbb{E}(\mathbf{Q})(\mathbf{I}_M \otimes \mathbf{J}_L^{*l})\mathbf{R})\mathbf{b}_2\right| \le \frac{\kappa^2}{(\mathrm{Im}(z))^2}$, (3.4.2, 3.4.7) and the Schwartz inequality lead immediately to

$$|\mathbf{b}_1^* \boldsymbol{\Delta}_1 \mathbf{R} \mathbf{b}_2| \le L^2 \kappa^2 \frac{1}{\sqrt{MN}} \frac{1}{\sqrt{MN}} C(z) = \kappa^2 C(z) \frac{L^2}{MN}$$

As N has the same order of magnitude than ML, it holds that $\mathscr{O}\left(\frac{L^2}{MN}\right) = \mathscr{O}\left(\frac{L}{M^2}\right)$ and $\mathscr{O}\left(\sqrt{\frac{L}{M}\frac{L}{N}}\right) = \mathscr{O}\left(\sqrt{\frac{L}{M^3}}\right)$, thus the convergence rate is $\sup\left(\frac{L^2}{MN}, \sqrt{\frac{L}{M^3}}\right)$. This establishes (3.6.3).
We now establish (3.6.4). Firstly, we remark that

$$\begin{aligned}
\mathbf{c}_1^* \left(\frac{1}{M}\sum_{m=1}^{M} \boldsymbol{\Upsilon}^{m,m}\right) \mathbf{c}_2 &= \frac{1}{M} \mathrm{Tr}\left((\mathbf{I}_M \otimes \mathbf{c}_2 \mathbf{c}_1^*)\boldsymbol{\Upsilon}\right) \\
&= -\sigma^2 \frac{1}{M} \mathrm{Tr}\left((\mathbf{I}_M \otimes \mathbf{c}_2 \mathbf{c}_1^*)\mathbb{E}(\mathbf{Q}) \mathbf{I}_M \otimes \mathscr{T}_{L,L}\left(\tilde{\boldsymbol{\Delta}}^T\right)\mathbf{R}\right) - \frac{1}{M}\mathrm{Tr}\left((\mathbf{I}_M \otimes \mathbf{c}_2 \mathbf{c}_1^*)\boldsymbol{\Delta}\mathbf{R}\right)
\end{aligned}$$

Since $\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\boldsymbol{\Delta}}^T) = \left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\boldsymbol{\Delta}})\right)^T = \left(\sum_{l=-(L-1)}^{L-1} \tau(\tilde{\boldsymbol{\Delta}})(l) \mathbf{I}_M \otimes \mathbf{J}_L^{*l}\right)^T = \sum_{l=-(L-1)}^{L-1} \tau(\tilde{\boldsymbol{\Delta}})(l) \mathbf{I}_M \otimes \mathbf{J}_L^{l}$, we can rewrite the term

$$\begin{aligned}
\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}_M \otimes \mathbf{c}_2 \mathbf{c}_1^*)\mathbb{E}(\mathbf{Q})\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\boldsymbol{\Delta}}^T)\right)\mathbf{R}\right) &= \frac{1}{M}\mathrm{Tr}\left((\mathbf{I}_M \otimes \mathbf{c}_2 \mathbf{c}_1^*)\mathbb{E}(\mathbf{Q}) \sum_{l=-(L-1)}^{L-1} \tau(\tilde{\boldsymbol{\Delta}})(l) \mathbf{I}_M \otimes \mathbf{J}_L^{l} \mathbf{R}\right) \\
&= \sum_{l=-(L-1)}^{L-1} \tau(\tilde{\boldsymbol{\Delta}})(l) \frac{1}{M}\mathrm{Tr}\left((\mathbf{I}_M \otimes \mathbf{c}_2 \mathbf{c}_1^*)\mathbb{E}(\mathbf{Q})\left(\mathbf{I}_M \otimes \mathbf{J}_L^{l}\right)\mathbf{R}\right) \\
&= \sum_{l=-(L-1)}^{L-1} \tau(\tilde{\boldsymbol{\Delta}})(l) \frac{1}{M}\sum_{m=1}^{M} \mathbf{c}_1^* (\mathbb{E}(\mathbf{Q})(\mathbf{I}_M \otimes \mathbf{J}_L^{*k})\mathbf{R})^{m,m} \mathbf{c}_2
\end{aligned}$$

We can notice that $\|\mathbb{E}(\mathbf{Q})(\mathbf{I}_M \otimes \mathbf{J}_L^{*k})\mathbf{R}\| \le \frac{1}{(\mathrm{Im}(z))^2}$. Since $(\mathbb{E}(\mathbf{Q})(\mathbf{I}_M \otimes \mathbf{J}_L^{l})\mathbf{R})^{m,m}$ is a sub-matrix of $\mathbb{E}(\mathbf{Q})(\mathbf{I}_M \otimes \mathbf{J}_L^{l})\mathbf{R}$, we can easily deduce that $\left|\mathbf{c}_1^* (\mathbb{E}(\mathbf{Q})(\mathbf{I}_M \otimes \mathbf{J}_L^{l})\mathbf{R})^{m,m} \mathbf{c}_2\right| \le \frac{\kappa^2}{(\mathrm{Im}(z))^2}$. It yields immediately that

$$\left|\frac{1}{M}\sum_{m=1}^{M} \mathbf{c}_1^* (\mathbb{E}(\mathbf{Q})(\mathbf{I}_M \otimes \mathbf{J}_L^{l})\mathbf{R})^{m,m} \mathbf{c}_2\right| \le \frac{\kappa^2}{(\mathrm{Im}(z))^2} \tag{3.6.6}$$

(3.6.1) implies that $\tau(\tilde{\boldsymbol{\Delta}}^T)(l)$ is an $\mathscr{O}\left(\frac{L}{MN}\right)$. The term $\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}_M \otimes \mathbf{c}_2 \mathbf{c}_1^*)\mathbb{E}(\mathbf{Q})\left(\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\boldsymbol{\Delta}}^T)\right)\mathbf{R}\right)$ is thus upperbounded by $\kappa^2 \frac{L^2}{MN} C(z)$.

As for the term $\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}_M \otimes \mathbf{c}_2\mathbf{c}_1^*)\boldsymbol{\Delta}\mathbf{R}\right)$, we recall that $\boldsymbol{\Delta} = \boldsymbol{\Delta}_1 + \boldsymbol{\Delta}_2 + \boldsymbol{\Delta}_3$. We can express each term separately like as the following

- $$\begin{aligned}\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}\otimes\mathbf{c}_2\mathbf{c}_1^*)\boldsymbol{\Delta}_1\mathbf{R}\right) &= \sigma^4 c_N \frac{1}{M}\mathrm{Tr}\left((\mathbf{I}_M\otimes\mathbf{c}_2\mathbf{c}_1^*)\mathbb{E}(\mathbf{Q})\mathbf{I}_M\otimes\mathscr{T}_{L,L}\left(\mathbf{H}\mathbb{E}\left(\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}}\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^\circ)\right)\mathbf{H}\right)\mathbf{R}\right)\\ &= \sigma^4 c_N \sum_{k=-(L-1)}^{L-1}\tau\left(\mathbf{H}\mathbb{E}\left(\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}}\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^\circ)\right)\mathbf{H}\right)(k)\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}_M\otimes\mathbf{c}_2\mathbf{c}_1^*)\mathbb{E}(\mathbf{Q})(\mathbf{I}_M\otimes\mathbf{J}_L^{*k})\mathbf{R}\right)\\ &= \sigma^4 c_N \sum_{k=-(L-1)}^{L-1}\mathbb{E}\left[\tau\left(\mathbf{H}\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}}\sum_{l=-(L-1)}^{L-1}\tau(\mathbf{Q}^\circ)(l)\mathbf{J}_N^{*l}\mathbf{H}\right)(k)\right]\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}_M\otimes\mathbf{c}_2\mathbf{c}_1^*)\mathbb{E}(\mathbf{Q})(\mathbf{I}_M\otimes\mathbf{J}_L^{*k})\mathbf{R}\right)\\ &= \sigma^4 c_N \sum_{k,l=-(L-1)}^{L-1}\mathbb{E}\left[\tau(\mathbf{Q}^\circ)(l)\tau\left(\mathbf{H}\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}}\mathbf{J}_N^{*l}\mathbf{H}\right)(k)\right]\frac{1}{M}\sum_{m=1}^{M}\mathbf{c}_1^*(\mathbb{E}(\mathbf{Q})(\mathbf{I}_M\otimes\mathbf{J}_L^{*k})\mathbf{R})^{m,m}\mathbf{c}_2\end{aligned}$$

(3.4.1) and (3.4.7) imply that $\mathbb{E}|\tau(\mathbf{Q}^\circ)|^2$ and $\mathbb{E}\left|\tau\left(\mathbf{H}\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}}\mathbf{J}_N^{*l}\mathbf{H}\right)(k)\right|^2$ are upperbounded by terms of the form $C(z)\,\frac{1}{MN}$. The Cauchy-Schwartz inequality thus implies immediately that

$$\left|\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}\otimes\mathbf{c}_2\mathbf{c}_1^*)\boldsymbol{\Delta}_1\mathbf{R}\right)\right| \le \kappa^2\frac{L^2}{MN}C(z)$$

As for the second term

- $$\begin{aligned}\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}\otimes\mathbf{c}_2\mathbf{c}_1^*)\boldsymbol{\Delta}_2\mathbf{R}\right) &= -\sigma^2\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}_M\otimes\mathbf{c}_2\mathbf{c}_1^*)\mathbb{E}(\mathbf{Q}^\circ\mathbf{I}_M\otimes\mathscr{T}_{L,L}(\mathbf{H}\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}}))\mathbf{R}\right)\\ &= \frac{\sigma^2}{M}\mathbb{E}\left[\mathrm{Tr}\left((\mathbf{I}_M\otimes\mathbf{c}_2\mathbf{c}_1^*)\mathbf{Q}^\circ\sum_{l=-(L-1)}^{L-1}\tau(\mathbf{H}\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}})(l)(\mathbf{I}_M\otimes\mathbf{J}_L^{*l})\mathbf{R}\right)\right]\\ &= \frac{\sigma^2}{M}\sum_{l=-(L-1)}^{L-1}\mathbb{E}\left[\mathrm{Tr}\left((\mathbf{I}_M\otimes\mathbf{c}_2\mathbf{c}_1^*)(\mathbf{Q}^\circ(\mathbf{I}_M\otimes\mathbf{J}_L^{*l})\mathbf{R})\right)\tau(\mathbf{H}\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}})(l)\right]\\ &= \sigma^2\sum_{l=-(L-1)}^{L-1}\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\left[\left((\mathbf{f}^m\otimes\mathbf{c}_1)^*(\mathbf{Q}^\circ(\mathbf{I}_M\otimes\mathbf{J}_L^{*l})\mathbf{R})(\mathbf{f}^m\otimes\mathbf{c}_2)\right)\tau(\mathbf{H}\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}})(l)\right]\end{aligned}$$

((3.4.2),(3.4.1)) imply that $\mathbb{E}|(\mathbf{f}^m\otimes\mathbf{c}_1)^*(\mathbf{Q}^\circ(\mathbf{I}_M\otimes\mathbf{J}_L^{*l})\mathbf{R})(\mathbf{f}^m\otimes\mathbf{c}_2)|^2$ and $\mathbb{E}|\tau(\mathbf{H}\mathbf{A}^T\mathbf{Q}^T\overline{\mathbf{Y}})(l)|^2$ are upperbounded by $\kappa^4 C(z)\frac{L}{N}$ and $C(z)\frac{1}{MN}$ terms respectively. Hence, by Cauchy-Schwartz inequality, it holds that

$$\left|\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}\otimes\mathbf{c}_2\mathbf{c}_1^*)\boldsymbol{\Delta}_1\mathbf{R}\right)\right| \le \kappa^2 L\sqrt{\frac{L}{N}}\sqrt{\frac{1}{MN}}C(z) = \kappa^2\sqrt{\frac{L}{M}}\frac{L}{N}C(z)$$

As for the third term

- $$\begin{aligned}\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}\otimes\mathbf{c}_2\mathbf{c}_1^*)\boldsymbol{\Delta}_3\mathbf{R}\right) &= -\sigma^2 c_N\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}_M\otimes\mathbf{c}_2\mathbf{c}_1^*)\mathbb{E}(\mathbf{Q}\mathbf{Y}\mathscr{T}_{N,L}^{(M)}(\mathbf{Q}^{\circ T})\mathbf{H}^T\mathbf{Y}^*)\mathbf{R}\right)\\ &= -\sigma^2 c_N\frac{1}{M}\mathbb{E}\left[\mathrm{Tr}\left((\mathbf{I}_M\otimes\mathbf{c}_2\mathbf{c}_1^*)\mathbf{Q}\mathbf{Y}\sum_{l=-(L-1)}^{L-1}\tau(\mathbf{Q}^\circ)(l)\mathbf{J}_N^l\mathbf{H}^T\mathbf{Y}^*\mathbf{R}\right)\right]\\ &= -\sigma^2 c_N\sum_{l=-(L-1)}^{L-1}\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\left[\left((\mathbf{f}^m\otimes\mathbf{c}_1)^*\mathbf{Q}\mathbf{Y}\mathbf{J}_N^l\mathbf{H}^T\mathbf{Y}^*\mathbf{R}(\mathbf{f}^m\otimes\mathbf{c}_2)\right)\tau(\mathbf{Q}^\circ)(l)\right]\end{aligned}$$

Using ((3.4.1),(3.4.5) and the Schwartz inequality, we obtain that $|\frac{1}{M}\mathrm{Tr}\left((\mathbf{I}\otimes\mathbf{c}_2\mathbf{c}_1^*)\boldsymbol{\Delta}_3\mathbf{R}\right)|$ is upperbounded by $\kappa^2\sqrt{\frac{L}{M}}\frac{L}{N}C(z)$. Thus we can conclude (3.6.4).
Finally, we recall that $\tilde{\boldsymbol{\Upsilon}} = \frac{1}{z}\left[\sigma^2\mathbf{H}^T\mathbf{A}^*\boldsymbol{\Upsilon}\mathbf{A}\mathbf{H}^T + \tilde{\boldsymbol{\Delta}}\right]$. Using the same tricks as previously, it is easy to prove

(3.6.5) □

Proposition 3.6.2 implies that the bilinear forms of $\boldsymbol{\Upsilon}$ and $\tilde{\boldsymbol{\Upsilon}}$ converge towards 0 if $\delta_N \to 0$. As $\frac{L^2}{MN} = \mathcal{O}(\frac{L}{M^2})$, it is clear that $\delta_N \to 0$ if and only if $\frac{L}{M^2} \to 0$. As it is assumed that $L = \mathcal{O}(N^\alpha)$ with $\alpha < 1$, $\frac{L}{M^2} \to 0$ holds if and only if the following assumption is verified:

Assumption A-3: $L = \mathcal{O}(N^\alpha)$ *with* $\alpha < 2/3$.

We finally remark that Proposition 3.6.2 implies the following stronger properties.

Corollary 3.6.1. *Under Assumption A-3, it holds that*

$$\|\boldsymbol{\Upsilon}\| = \|\mathbb{E}(\mathbf{Q}) - \mathbf{R}\| \leq C(z) \sup\left(\frac{L^2}{MN}, \sqrt{\frac{L}{M^3}}\right) \to 0, \tag{3.6.7}$$

$$\|\tilde{\boldsymbol{\Upsilon}}\| = \|\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{R}}\| \leq C(z) \sup\left(\frac{L^2}{MN}, \sqrt{\frac{L}{M^3}}\right) \to 0. \tag{3.6.8}$$

for each $z \in \mathbb{C}^+$ *where* $C(z)$ *can be written as* $C(z) = P_1(|z|) P_2\left((\text{Im} z)^{-1}\right)$ *for some nice polynomials* P_1 *and* P_2*. Moreover,*

$$\|\mathcal{T}^{(M)}_{N,L}\left(\mathbb{E}(\mathbf{Q}) - \mathbf{R}\right)\| \leq C(z) \sup\left(\frac{L^2}{MN}, \sqrt{\frac{L}{M^3}}\right) \tag{3.6.9}$$

$$\|\mathcal{T}_{L,L}\left(\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{R}}\right)\| \leq C(z) \sup\left(\frac{L^2}{MN}, \sqrt{\frac{L}{M^3}}\right) \tag{3.6.10}$$

Proof. As $C(z)$ in (3.6.3) does not depend on L, M, N and on $\mathbf{b}_1, \mathbf{b}_2$, (3.6.3) implies that

$$\sup_{\|\mathbf{b}_1\|=1, \|\mathbf{b}_2\|=1} \left|\mathbf{b}_1^* \boldsymbol{\Upsilon} \mathbf{b}_2\right| = \|\boldsymbol{\Upsilon}\| \leq C(z)\delta_N \to 0$$

(3.6.8) holds for the same reason. (3.6.9) and (3.6.9) follow from Proposition 3.1.3.

□

3.7 The deterministic equivalents: existence and uniqueness

In this section, we prove the existence and uniqueness of the solution of the equation governing the matrix-valued deterministic equivalents of $\mathbf{Q}(z)$ and $\tilde{\mathbf{Q}}(z)$.

Theorem 3.7.1. *The deterministic system of equations :*

$$\begin{cases} \mathbf{T}(z) = \left[-z\left(\mathbf{I}_{ML} + \sigma^2 \mathbf{I}_M \otimes \mathcal{T}_{L,L}(\tilde{\mathbf{T}}^T(z))\right) + \mathbf{A}\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T(z))\right)^{-1} \mathbf{A}^*\right]^{-1} \\ \tilde{\mathbf{T}}(z) = \left[-z\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T(z))\right) + \mathbf{A}^*\left(\mathbf{I}_{ML} + \sigma^2 \mathbf{I}_M \otimes \mathcal{T}_{L,L}(\tilde{\mathbf{T}}^T(z))\right)^{-1} \mathbf{A}\right]^{-1} \end{cases} \tag{3.7.1}$$

admits a unique solution $(\mathbf{T}(z), \tilde{\mathbf{T}}(z))$ *for which* $\mathbf{T} \in \mathcal{S}_{ML}(\mathbb{R}^+)$ *and* $\tilde{\mathbf{T}} \in \mathcal{S}_N(\mathbb{R}^+)$*, i.e.*

$$\begin{cases} \boldsymbol{\mu}(\mathbb{R}^+) = \mathbf{I}_{ML} \quad, \tilde{\boldsymbol{\mu}}(\mathbb{R}^+) = \mathbf{I}_N \\ \mathbf{T}(z) = \int_{\mathbb{R}^+} \frac{\boldsymbol{\mu}(d\lambda)}{\lambda - z} \quad, \tilde{\mathbf{T}}(z) = \int_{\mathbb{R}^+} \frac{\tilde{\boldsymbol{\mu}}(d\lambda)}{\lambda - z} \end{cases} \tag{3.7.2}$$

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

Remark 9. *We recall that as* $\mathbf{T}(z), \tilde{\mathbf{T}}(z)$ *belong to* $\mathscr{S}_{\mathrm{ML}}(\mathbb{R}^+)$ *and* $\mathscr{S}_{\mathrm{N}}(\mathbb{R}^+)$ *respectively, they verify*

$$\mathbf{T}(z)\mathbf{T}(z)^* \leq \frac{\mathbf{I}_{\mathrm{ML}}}{(\mathrm{Im}(z))^2} \quad ; \tilde{\mathbf{T}}(z)\tilde{\mathbf{T}}(z)^* \leq \frac{\mathbf{I}_{\mathrm{N}}}{(\mathrm{Im}(z))^2} \tag{3.7.3}$$

for every $z \in \mathbb{C}^+$ *(see (3.1.30)).*

Proof of Theorem 3.7.1. We will prove that the solution of (3.7.1) exists and is unique. For this let us consider the following propositions :

Proposition 3.7.1 (Uniqueness of solution)**.** *Assume that there exist two solutions* $(\mathbf{T}_1(z), \tilde{\mathbf{T}}_1(z))$ *and* $(\mathbf{T}_2(z), \tilde{\mathbf{T}}_2(z))$ *of the system* (3.7.1). *Then it holds that* $(\mathbf{T}_1, \tilde{\mathbf{T}}_1) = (\mathbf{T}_2, \tilde{\mathbf{T}}_2)$.

Proof. Let us calculate $\mathbf{T}_1(z) - \mathbf{T}_2(z)$ for $z \in \mathbb{C}^+$:

$$\mathbf{T}_1(z) - \mathbf{T}_2(z) = \mathbf{T}_1(z)(\mathbf{T}_2^{-1}(z) - \mathbf{T}_1^{-1}(z))\mathbf{T}_2(z)$$

Standard calculation gives us that

$$\begin{aligned}\mathbf{T}_1(z) - \mathbf{T}_2(z) = \ & \sigma^2 z \mathbf{T}_1(z)\mathbf{I}_{\mathrm{M}} \otimes \mathscr{T}_{\mathrm{L,L}}(\tilde{\mathbf{T}}_1^{\mathrm{T}}(z) - \tilde{\mathbf{T}}_2^{\mathrm{T}}(z))\mathbf{T}_2(z) \\ & + \sigma^2 c_{\mathrm{N}} \mathbf{T}_1(z)\mathbf{A}\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathscr{T}^{(\mathrm{M})}_{\mathrm{N,L}}(\mathbf{T}_2^{\mathrm{T}}(z))\right)^{-1} \mathscr{T}^{(\mathrm{M})}_{\mathrm{N,L}}(\mathbf{T}_1^{\mathrm{T}}(z) - \mathbf{T}_2^{\mathrm{T}}(z))\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathscr{T}^{(\mathrm{M})}_{\mathrm{N,L}}(\mathbf{T}_1^{\mathrm{T}}(z))\right)^{-1}\mathbf{A}^*\mathbf{T}_2(z)\end{aligned}$$

By (3.1.17), $\|\mathbf{I}_{\mathrm{M}} \otimes \mathscr{T}_{\mathrm{L,L}}(\tilde{\mathbf{T}}_1^{\mathrm{T}}(z) - \tilde{\mathbf{T}}_2^{\mathrm{T}}(z))\| \leq \|\tilde{\mathbf{T}}_1(z) - \tilde{\mathbf{T}}_2(z)\|$ and $\|\mathscr{T}^{(\mathrm{M})}_{\mathrm{N,L}}(\mathbf{T}_1^{\mathrm{T}}(z) - \mathbf{T}_2^{\mathrm{T}}(z))\| \leq \|\mathbf{T}_1(z) - \mathbf{T}_2(z)\|$, so that we get

$$\begin{aligned}\|\mathbf{T}_1(z) - \mathbf{T}_2(z)\| \leq \ & \sigma^2 |z| \|\mathbf{T}_1(z)\| \, \|\mathbf{T}_2(z)\| \, \|\tilde{\mathbf{T}}_1(z) - \tilde{\mathbf{T}}_2(z)\| \\ & + \sigma^2 c_{\mathrm{N}} \|\mathbf{A}\|^2 \, \|\mathbf{T}_1(z)\| \, \|\mathbf{T}_2(z)\| \left\|\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathscr{T}^{(\mathrm{M})}_{\mathrm{N,L}}(\mathbf{T}_1(z))\right)^{-1}\right\| \left\|\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathscr{T}^{(\mathrm{M})}_{\mathrm{N,L}}(\mathbf{T}_2(z))\right)^{-1}\right\| \|\mathbf{T}_1(z) - \mathbf{T}_2(z)\|\end{aligned}$$

As $(\mathbf{T}_i(z), \tilde{\mathbf{T}}_i(z))_{i=1,2}$ satisfy (3.7.2), it holds that

$$\|\mathbf{T}_i(z)\| \leq \frac{1}{\mathrm{Im}(z)} \tag{3.7.4}$$

Moreover, we can prove for $i = 1, 2$ that $\left\|\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathscr{T}^{(\mathrm{M})}_{\mathrm{N,L}}(\mathbf{T}_i(z))\right)^{-1}\right\| \leq \frac{|z|}{\mathrm{Im}(z)}$ by noticing that $-\frac{\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathscr{T}^{(\mathrm{M})}_{\mathrm{N,L}}(\mathbf{T}_i(z))\right)^{-1}}{z}$ is an element of $\mathscr{S}_{\mathrm{N}}(\mathbb{R}^+)$. We omit the proof which it is similar to the proof of lemma 3.5.1.
We introduce the following maxima :

$$\mathbf{M}(z) = \max\{\|\mathbf{T}_1(z) - \mathbf{T}_2(z)\|, \|\tilde{\mathbf{T}}_1(z) - \tilde{\mathbf{T}}_2(z)\|\} \text{ and } \overline{c}_{\mathrm{N}} = \max\left(\frac{\mathrm{ML}}{\mathrm{N}}, \frac{\mathrm{N}}{\mathrm{ML}}\right)$$

Note that $\overline{c}_{\mathrm{N}} \geq c_{\mathrm{N}} = \frac{\mathrm{ML}}{\mathrm{N}}$ and $\overline{c}_{\mathrm{N}} \geq 1$ and recall that $\sup_{\mathrm{N}} \|\mathbf{A}\| \leq \mathrm{C} < +\infty$ .We obtain thus the following upperbound of $\|\mathbf{T}_1(z) - \mathbf{T}_2(z)\|$

$$\|\mathbf{T}_1(z) - \mathbf{T}_2(z)\| \leq \overline{c}_{\mathrm{N}} \left(\sigma^2 \frac{|z|}{\mathrm{Im}(z)^2} + \sigma^2 \mathrm{C}^2 \frac{|z|^2}{(\mathrm{Im}(z))^4}\right)\mathbf{M}(z) = \xi(z)\mathbf{M}(z) \tag{3.7.5}$$

where $\xi(z) = \frac{\overline{c}_{\mathrm{N}} \sigma^2 |z|}{\mathrm{Im}(z)^2}\left(1 + \mathrm{C}^2 \frac{|z|}{(\mathrm{Im}(z))^2}\right)$
The same kind of calculation yields

$$\begin{aligned}\tilde{\mathbf{T}}_1(z) - \tilde{\mathbf{T}}_2(z) = \ & \sigma^2 c_{\mathrm{N}} z \tilde{\mathbf{T}}_1(z) \mathscr{T}^{(\mathrm{M})}_{\mathrm{N,L}}(\mathbf{T}_1(z) - \mathbf{T}_2(z))\tilde{\mathbf{T}}_2(z) \\ & + \sigma^2 \tilde{\mathbf{T}}_1(z)\mathbf{A}^*\left(\mathbf{I}_{\mathrm{ML}} + \sigma^2 \mathbf{I}_{\mathrm{M}} \otimes \mathscr{T}_{\mathrm{L,L}}(\tilde{\mathbf{T}}_2^{\mathrm{T}}(z))\right)^{-1} \mathbf{I}_{\mathrm{ML}} \otimes \mathscr{T}_{\mathrm{L,L}}(\tilde{\mathbf{T}}_1^{\mathrm{T}}(z) - \tilde{\mathbf{T}}_2^{\mathrm{T}}(z))\left(\mathbf{I}_{\mathrm{ML}} + \sigma^2 \mathbf{I}_{\mathrm{ML}} \otimes \mathscr{T}_{\mathrm{L,L}}(\tilde{\mathbf{T}}_1^{\mathrm{T}}(z))\right)^{-1}\mathbf{A}\tilde{\mathbf{T}}_2(z)\end{aligned}$$

which gives us the inequality

$$\|\tilde{\mathbf{T}}_1(z) - \tilde{\mathbf{T}}_2(z)\| \leq \xi(z)\mathbf{M}(z) \tag{3.7.6}$$

Finally, gathering (3.7.5) and (3.7.6) we obtain

$$\mathbf{M}(z) \leq \xi(z)\mathbf{M}(z) \tag{3.7.7}$$

If z verifies $|\frac{z}{\mathrm{Im}(z)}| \leq 2$ and $\mathrm{Im}(z)$ large enough, one has $\xi(z) \leq \frac{1}{2}$, which, by (3.7.7), implies that $\mathbf{M}(z) = 0$, i.e. that $\mathbf{T}_1(z) = \mathbf{T}_2(z)$ and $\tilde{\mathbf{T}}_1(z) = \tilde{\mathbf{T}}_2(z)$ for these values of z. The analyticity of $(\mathbf{T}_i, \tilde{\mathbf{T}}_i)_{i=1,2}$ implies that $\mathbf{T}_1 = \mathbf{T}_2$ and $\tilde{\mathbf{T}}_1 = \tilde{\mathbf{T}}_2$ on $\mathbb{C} - \mathbb{R}^+$. Proposition 3.7.1 is proved.

□

Proposition 3.7.2 (Existence of solutions). *There exists* $(\mathbf{T}, \tilde{\mathbf{T}}) \in \mathcal{S}_{\mathrm{ML}}(\mathbb{R}^+) \times \mathcal{S}_{\mathrm{N}}(\mathbb{R}^+)$ *, solution of equation* (3.7.1).

Proof. We construct the desired solution by induction. Let

$$\mathbf{T}^{(0)}(z) = -\frac{\mathbf{I}_{\mathrm{ML}}}{z}$$
$$\tilde{\mathbf{T}}^{(0)}(z) = -\frac{\mathbf{I}_{\mathrm{N}}}{z}$$

It is clear that $\mathbf{T}^{(0)}(z)$ and $\tilde{\mathbf{T}}^{(0)}(z)$ satisfy (3.7.2). For $p \geq 0$, let

$$\mathbf{T}^{(p+1)}(z) = \left[-z\left(\mathbf{I}_{\mathrm{ML}} + \sigma^2 \mathbf{I}_{\mathrm{M}} \otimes \mathcal{T}_{\mathrm{L,L}}((\tilde{\mathbf{T}}^{(p)}(z))^{\mathrm{T}})\right) + \mathbf{A}\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathcal{T}^{(\mathrm{M})}_{\mathrm{N,L}}((\mathbf{T}^{(p)}(z))^{\mathrm{T}})\right)^{-1} \mathbf{A}^*\right]^{-1}$$
$$\tilde{\mathbf{T}}^{(p+1)}(z) = \left[-z\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathcal{T}^{(\mathrm{M})}_{\mathrm{N,L}}((\mathbf{T}^{(p)}(z))^{\mathrm{T}})\right) + \mathbf{A}^*\left(\mathbf{I}_{\mathrm{ML}} + \sigma^2 \mathbf{I}_{\mathrm{M}} \otimes \mathcal{T}_{\mathrm{L,L}}((\tilde{\mathbf{T}}^{(p)}(z))^{\mathrm{T}})\right)^{-1} \mathbf{A}\right]^{-1}$$

Using the same arguments as in the proof of lemma 3.5.1, we can check that for each $p \geq 0$, $\mathbf{T}^{(p)}$ and $\tilde{\mathbf{T}}^{(p)}$ belong to $\mathcal{S}_{\mathrm{ML}}(\mathbb{R}^+)$ and $\mathcal{S}_{\mathrm{N}}(\mathbb{R}^+)$ respectively.
We denote by

$$\mathbf{M}^{(p)}(z) = \max\{\|\mathbf{T}^{(p+1)}(z) - \mathbf{T}^{(p)}(z)\|, \|\tilde{\mathbf{T}}^{(p+1)}(z) - \tilde{\mathbf{T}}^{(p)}(z)\|\}$$

The same calculations as in the proof of Proposition 3.7.2 yield

$$\mathbf{M}^{(p+1)}(z) \leq \xi(z)\mathbf{M}^{(p)}(z),$$

where $\xi(z) = \frac{\overline{c}_{\mathrm{N}}\sigma^2|z|}{\mathrm{Im}(z)^2}\left(1 + \kappa^2 \frac{|z|}{(\mathrm{Im}(z))^2}\right)$.
Let $z \in \mathbb{C}^+$ be such that $\frac{|z|}{\mathrm{Im}(z)} \leq 2$. For $\mathrm{Im}(z)$ large enough, $\mathbf{T}^{(p)}(z)$ and $\tilde{\mathbf{T}}^{(p)}(z)$ are Cauchy sequences of matrices w.r.t to spectral norms on $\mathbb{C}^{\mathrm{ML}\times\mathrm{ML}}$ and $\mathbb{C}^{\mathrm{N}\times\mathrm{N}}$ respectively. Denote by $\mathbf{T}(z)$ and $\tilde{\mathbf{T}}(z)$ their corresponding limits. Moreover, $(\mathbf{T}^{(p)}(z))_p$ is a normal family over $\mathbb{C}^+$, since each $\mathbf{T}^{(p)}(z)$ satisfies (3.7.2), its norm is upperbounded on every compact set included in $\mathbb{C}^+$ uniformly in p (see (3.7.3)). Therefore one can extract, by Montel's theorem, converging subsequence whose limit is analytic over $\mathbb{C}^+$. Since the limit of any converging subsequence is equal to $\mathbf{T}$ in the domain $\{z \in \mathbb{C}^+, \frac{|z|}{\mathrm{Im}(z)} \leq 2, \mathrm{Im}(z) \text{ large enough}\}$, $\mathbf{T}^{(p)}(z)$ converges towards an analytic function on $\mathbb{C}^+$, so that $\mathbf{T}(z)$ is analytic. Similar arguments can be applied to prove the analyticity of $\tilde{\mathbf{T}}(z)$.

We now prove that $\mathbf{T}(z)$ and $\tilde{\mathbf{T}}(z)$ satisfy (3.7.2). The convergence of $\mathbf{T}^{(p)}(z)$ immediately yields that

$$\begin{cases} \mathrm{Im}(\mathbf{T}^{(p)}(z)) \geq 0 \implies \mathrm{Im}(\mathbf{T}(z)) \geq 0 \\ \mathrm{Im}(z\mathbf{T}^{(p)}(z)) \geq 0 \implies \mathrm{Im}(z\mathbf{T}(z)) \geq 0 \\ \mathbf{T}^{(p)}(z)\mathbf{T}^{(p)}(z)^* \leq \frac{\mathbf{I}_{\mathrm{ML}}}{(\mathrm{Im}(z))^2} \implies \mathbf{T}(z)\mathbf{T}(z)^* \leq \frac{\mathbf{I}_{\mathrm{ML}}}{(\mathrm{Im}(z))^2} \end{cases} \tag{3.7.8}$$

on $\mathbb{C}^+$.

Finally, as $\mathbf{T}^{(p)}(z) \xrightarrow[p\to+\infty]{} \mathbf{T}(z)$, $\mathbf{T}(z)$ is therefore written as

$$\mathbf{T}(z) = \left[-z\left(\mathbf{I}_{\mathrm{ML}} + \sigma^2 \mathbf{I}_{\mathrm{M}} \otimes \mathscr{T}_{\mathrm{L,L}}(\tilde{\mathbf{T}}^{\mathrm{T}}(z))\right) + \mathbf{A}\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{T}^{\mathrm{T}}(z))\right)^{-1} \mathbf{A}^*\right]^{-1}$$

It is easy to see that $\lim_{y\to+\infty} -iy\mathbf{T}(iy) = \mathbf{I}_{\mathrm{ML}}$, this and (3.7.8) prove that $\mathbf{T}(z) \in \mathscr{S}_{\mathrm{ML}}(\mathbb{R}^+)$. We can prove similarly that $\tilde{\mathbf{T}}(z)$ verifies (3.7.2). Proposition 3.7.2 is proved.

□

3.8 Convergence towards the deterministic equivalents: the normalized traces

In the following, we establish that

$$\begin{cases} \frac{1}{\mathrm{ML}} \mathrm{Tr}\,(\mathbb{E}(\mathbf{Q}(z)) - \mathbf{T}(z)) \to 0 \\ \frac{1}{\mathrm{N}} \mathrm{Tr}\,(\mathbb{E}(\tilde{\mathbf{Q}}(z) - \tilde{\mathbf{T}}(z)) \to 0 \end{cases} \tag{3.8.1}$$

for each $z \in \mathbb{C} \setminus \mathbb{R}^+$, with $\mathbf{T}$ and $\tilde{\mathbf{T}}$ are defined as in Theorem 3.7.1.

In the following, we thus prove (3.8.1). Proposition 3.6.1 implies that for uniformly bounded $\mathrm{ML} \times \mathrm{ML}$ and $\mathrm{N} \times \mathrm{N}$ matrices $\mathbf{B}, \tilde{\mathbf{B}}$, it holds that

$$\begin{aligned} \frac{1}{\mathrm{ML}} \mathrm{Tr}\,[(\mathbb{E}(\mathbf{Q}(z)) - \mathbf{R}(z))\mathbf{B}] &= \mathscr{O}(\frac{\mathrm{L}}{\mathrm{MN}}) \\ \frac{1}{\mathrm{N}} \mathrm{Tr}\,[(\mathbb{E}(\tilde{\mathbf{Q}}(z)) - \tilde{\mathbf{R}}(z))\tilde{\mathbf{B}}] &= \mathscr{O}(\frac{\mathrm{L}}{\mathrm{MN}}) \end{aligned} \tag{3.8.2}$$

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

$$\begin{cases} \mathbf{R}(z) = \left[-z\left(\mathbf{I}_{\mathrm{ML}} + \sigma^2 \mathbf{I}_{\mathrm{M}} \otimes \mathscr{T}_{\mathrm{L,L}}(\mathbb{E}(\tilde{\mathbf{Q}}^{\mathrm{T}}))\right) + \mathbf{A}\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\right]^{-1} \\ \tilde{\mathbf{R}}(z) = \left[-z\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}^{\mathrm{T}}))\right) + \mathbf{A}^*\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{A}\right]^{-1} \end{cases}$$

with

$$\begin{cases} \mathbf{H}(z) = \left[\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}(z)))\right]^{-1} \\ \tilde{\mathbf{H}}(z) = \left[\mathbf{I}_{\mathrm{ML}} + \sigma^2 \mathbf{I}_{\mathrm{M}} \otimes \mathscr{T}_{\mathrm{L,L}}(\mathbb{E}(\tilde{\mathbf{Q}}))\right]^{-1} \end{cases}$$

and

$$\begin{cases} \mathbf{T}(z) = \left[-z\left(\mathbf{I}_{\mathrm{ML}} + \sigma^2 \mathbf{I}_{\mathrm{M}} \otimes \mathscr{T}_{\mathrm{L,L}}(\tilde{\mathbf{T}}^{\mathrm{T}}(z))\right) + \mathbf{A}\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{T}^{\mathrm{T}}(z))\right)^{-1} \mathbf{A}^*\right]^{-1} \\ \tilde{\mathbf{T}}(z) = \left[-z\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{T}^{\mathrm{T}}(z))\right) + \mathbf{A}^*\left(\mathbf{I}_{\mathrm{ML}} + \sigma^2 \mathbf{I}_{\mathrm{M}} \otimes \mathscr{T}_{\mathrm{L,L}}(\tilde{\mathbf{T}}^{\mathrm{T}}(z))\right)^{-1} \mathbf{A}\right]^{-1} \end{cases}$$

We will establish that

$$\begin{aligned} \frac{1}{\mathrm{ML}} \mathrm{Tr}\,[(\mathbf{R}(z) - \mathbf{T}(z))\mathbf{B}] &\to 0 \\ \frac{1}{\mathrm{N}} \mathrm{Tr}\,[(\tilde{\mathbf{R}}(z) - \tilde{\mathbf{T}}(z))\tilde{\mathbf{B}}] &\to 0 \end{aligned} \tag{3.8.3}$$

For this, we first mention that straightforward computations lead to

$$\mathbf{R} - \mathbf{T} = \sigma^2 z \mathbf{R}\, \mathbf{I}_{\mathrm{M}} \otimes \mathscr{T}_{\mathrm{L,L}}(\mathbb{E}(\tilde{\mathbf{Q}}^{\mathrm{T}}) - \tilde{\mathbf{T}}^{\mathrm{T}})\, \mathbf{T} + \sigma^2 c_{\mathrm{N}} \mathbf{R}\mathbf{A}\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}} \mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{T}^{\mathrm{T}})\right)^{-1} \mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}^{\mathrm{T}}) - \mathbf{T}^{\mathrm{T}})\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\mathbf{T}$$

Writing that $\mathbb{E}(\mathbf{Q}) - \mathbf{T} = \mathbb{E}(\mathbf{Q}) - \mathbf{R} + \mathbf{R} - \mathbf{T}$ and $\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{T}} = \mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{R}} + \tilde{\mathbf{R}} - \tilde{\mathbf{T}}$, we obtain that

$$\begin{aligned}\mathbf{R} - \mathbf{T} &= \sigma^2 z \mathbf{R}\, \mathbf{I}_M \otimes \mathcal{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}^T) - \tilde{\mathbf{R}}^T)\, \mathbf{T} + \sigma^2 c_N \mathbf{R}\mathbf{A}\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T)\right)^{-1} \mathcal{T}^{(M)}_{N,L}(\mathbb{E}(\mathbf{Q}^T) - \mathbf{R}^T)\mathbf{H}^T\mathbf{A}^*\mathbf{T} \\ &+ \sigma^2 z \mathbf{R}\, \mathbf{I}_M \otimes \mathcal{T}_{L,L}(\tilde{\mathbf{R}}^T - \tilde{\mathbf{T}}^T)\, \mathbf{T} + \sigma^2 c_N \mathbf{R}\mathbf{A}\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T)\right)^{-1} \mathcal{T}^{(M)}_{N,L}(\mathbf{R}^T - \mathbf{T}^T)\mathbf{H}^T\mathbf{A}^*\mathbf{T} \qquad (3.8.4)\end{aligned}$$

so that we have

$$\begin{aligned}\frac{1}{ML}\mathrm{Tr}\,[(\mathbf{R}-\mathbf{T})\mathbf{B}] &= \sigma^2 z \frac{1}{ML}\mathrm{Tr}\left[(\mathbf{T}\mathbf{B}\mathbf{R})^T\, \mathbf{I}_M \otimes \mathcal{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{R}})\right] \\ &+ \sigma^2 c_N \frac{1}{ML}\mathrm{Tr}\left[\left(\mathbf{H}^T\mathbf{A}^*\mathbf{T}\mathbf{B}\mathbf{R}\mathbf{A}\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T)\right)^{-1}\right)^T \mathcal{T}^{(M)}_{N,L}(\mathbb{E}(\mathbf{Q}) - \mathbf{R})\right] \\ &+ \sigma^2 z \frac{1}{ML}\mathrm{Tr}\left[(\mathbf{T}\mathbf{B}\mathbf{R})^T\, \mathbf{I}_M \otimes \mathcal{T}_{L,L}(\tilde{\mathbf{R}} - \tilde{\mathbf{T}})\right] \\ &+ \sigma^2 c_N \frac{1}{ML}\mathrm{Tr}\left[\left(\mathbf{H}^T\mathbf{A}^*\mathbf{T}\mathbf{B}\mathbf{R}\mathbf{A}\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T)\right)^{-1}\right)^T \mathcal{T}^{(M)}_{N,L}(\mathbf{R} - \mathbf{T})\right]\end{aligned}$$

Direct application of (3.1.10) to the case P = M, K = L, R = N implies that

$$\begin{aligned}\frac{1}{ML}\mathrm{Tr}\,[(\mathbf{R}-\mathbf{T})\mathbf{B}] &= \sigma^2 z \frac{1}{N}\mathrm{Tr}\left[(\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{R}})\, \mathcal{T}^{(M)}_{N,L}((\mathbf{T}\mathbf{B}\mathbf{R})^T)\right] \\ &+ \sigma^2 \frac{1}{ML}\mathrm{Tr}\left[(\mathbb{E}(\mathbf{Q}) - \mathbf{R})\, \mathbf{I}_M \otimes \mathcal{T}_{L,L}\left(\mathbf{H}^T\mathbf{A}^*\mathbf{T}\mathbf{B}\mathbf{R}\mathbf{A}\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T)\right)^{-1}\right)^T\right] \\ &+ \sigma^2 z \frac{1}{N}\mathrm{Tr}\left[(\tilde{\mathbf{R}} - \tilde{\mathbf{T}})\, \mathcal{T}^{(M)}_{N,L}((\mathbf{T}\mathbf{B}\mathbf{R})^T)\right] \\ &+ \sigma^2 \frac{1}{ML}\mathrm{Tr}\left[(\mathbf{R} - \mathbf{T})\, \mathbf{I}_M \otimes \mathcal{T}_{L,L}\left(\mathbf{H}^T\mathbf{A}^*\mathbf{T}\mathbf{B}\mathbf{R}\mathbf{A}\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T)\right)^{-1}\right)^T\right]\end{aligned} \qquad (3.8.5)$$

We define matrix $\mathbf{G}(\mathbf{B})$ by

$$\mathbf{G}(\mathbf{B}) = \mathcal{T}_{L,L}\left(\mathbf{H}^T\mathbf{A}^*\mathbf{T}\mathbf{B}\mathbf{R}\mathbf{A}\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T)\right)^{-1}\right)^T \qquad (3.8.6)$$

We now prove that

$$\sup_{\|\mathbf{C}\|\le 1}\left|\frac{1}{ML}\mathrm{Tr}\,((\mathbf{R}-\mathbf{T})\mathbf{C})\right| \to 0$$
$$\sup_{\|\tilde{\mathbf{C}}\|\le 1}\left|\frac{1}{N}\mathrm{Tr}\,((\tilde{\mathbf{R}}-\tilde{\mathbf{T}})\tilde{\mathbf{C}})\right| \to 0$$

when z belongs to a certain domain. For this, we first remark that Proposition 3.1.3 implies that

$$\|\mathcal{T}^{(M)}_{N,L}((\mathbf{T}\mathbf{B}\mathbf{R})^T)\| \le \|\mathbf{T}\|\,\|\mathbf{R}\|\,\|\mathbf{B}\|$$
$$\|\mathbf{G}(\mathbf{A})\| \le \|\mathbf{H}\|\,\|\mathbf{T}\|\,\|\mathbf{R}\|\,\|\mathbf{A}\|^2\left\|\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T)\right)^{-1}\right\|\,\|\mathbf{B}\|$$

We assume that $z \in \mathbb{C}^+$. By lemma 3.5.1, it holds that $\|\mathbf{H}\| \le \frac{|z|}{(\mathrm{Im}(z))}$ and $\|\mathbf{R}\| \le \frac{1}{\mathrm{Im}(z)}$. By (3.7.4), we have that $\|\mathbf{T}\| \le \frac{1}{\mathrm{Im}(z)}$ and $\left\|\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T)\right)^{-1}\right\| \le \frac{|z|}{\mathrm{Im}(z)}$. Consequently, we get that

$$\|\mathcal{T}^{(M)}_{N,L}((\mathbf{T}\mathbf{B}\mathbf{R})^T)\| \le \frac{1}{(\mathrm{Im}(z))^2}\|\mathbf{B}\|$$
$$\|\mathbf{G}(\mathbf{B})\| \le \frac{|z|^2}{(\mathrm{Im}(z))^4}\,\|\mathbf{A}\|^2\,\|\mathbf{B}\|$$

We recall that $\sup_N \|\mathbf{A}\| \leq C < +\infty$. For each $ML \times ML$ matrix $\mathbf{B}$ such that $\|\mathbf{B}\| \leq 1$, it holds that $\|\mathcal{T}_{N,L}^{(M)}(\mathbf{B})\| \leq \|\mathbf{B}\| \leq 1$ and that $\|\mathbf{I}_M \otimes \mathcal{T}_{L,L}(\mathbf{B})\| \leq \|B\| \leq 1$. Thus, it follows

- $\left|\frac{1}{N}\mathrm{Tr}\left[(\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{R}})\,\mathcal{T}_{N,L}^{(M)}((\mathbf{TBR})^T)\right]\right| \leq \frac{1}{(\mathrm{Im}(z))^2} \sup_{\|\tilde{\mathbf{C}}\|\leq 1} \left|\frac{1}{N}\mathrm{Tr}\,((\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{R}})\tilde{\mathbf{C}})\right|$
- $\left|\frac{1}{N}\mathrm{Tr}\left[(\tilde{\mathbf{R}} - \tilde{\mathbf{T}})\,\mathcal{T}_{N,L}^{(M)}((\mathbf{TBR})^T)\right]\right| \leq \frac{1}{(\mathrm{Im}(z))^2} \sup_{\|\tilde{\mathbf{C}}\|\leq 1} \left|\frac{1}{N}\mathrm{Tr}\,((\tilde{\mathbf{R}} - \tilde{\mathbf{T}})\tilde{\mathbf{C}})\right|$
- $\left|\frac{1}{ML}\mathrm{Tr}\,[(\mathbb{E}(\mathbf{Q}) - \mathbf{R})\,\mathbf{I}_M \otimes \mathbf{G}(\mathbf{B})]\right| \leq \frac{\mathbf{C}^2|z|^2}{(\mathrm{Im}(z))^4} \sup_{\|\mathbf{C}\|\leq 1} \left|\frac{1}{ML}\mathrm{Tr}\,((\mathbb{E}(\mathbf{Q}) - \mathbf{R})\mathbf{C})\right|$
- $\left|\frac{1}{ML}\mathrm{Tr}\,[(\mathbf{R} - \mathbf{T})\,\mathbf{I}_M \otimes \mathbf{G}(\mathbf{B})]\right| \leq \frac{\mathbf{C}^2|z|^2}{(\mathrm{Im}(z))^4} \sup_{\|\mathbf{C}\|\leq 1} \left|\frac{1}{ML}\mathrm{Tr}\,((\mathbf{R} - \mathbf{T})\mathbf{C})\right|$

Proposition 3.6.1 implies that

$$\sup_{\|\mathbf{C}\|\leq 1} \left|\frac{1}{ML}\mathrm{Tr}\,((\mathbb{E}(\mathbf{Q}) - \mathbf{R})\mathbf{C})\right| = \mathcal{O}(\frac{L}{MN})$$
$$\sup_{\|\tilde{\mathbf{C}}\|\leq 1} \left|\frac{1}{N}\mathrm{Tr}\,((\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{R}})\tilde{\mathbf{C}})\right| = \mathcal{O}(\frac{L}{MN})$$

This and Eq. (3.8.5) eventually imply that

$$\sup_{\|\mathbf{C}\|\leq 1} \left|\frac{1}{ML}\mathrm{Tr}\,((\mathbf{R}-\mathbf{T})\mathbf{C})\right| \leq \mathcal{O}(\frac{L}{MN}) + \frac{\sigma^2|z|}{(\mathrm{Im}(z))^2} \sup_{\|\tilde{\mathbf{C}}\|\leq 1} \left|\frac{1}{N}\mathrm{Tr}\,((\tilde{\mathbf{R}}-\tilde{\mathbf{T}})\tilde{\mathbf{C}})\right| + \frac{\sigma^2\kappa^2|z|^2}{(\mathrm{Im}(z))^4} \sup_{\|\mathbf{C}\|\leq 1} \left|\frac{1}{ML}\mathrm{Tr}\,((\mathbf{R}-\mathbf{T})\mathbf{C})\right| \tag{3.8.7}$$

Using similar tricks, we obtain

$$\sup_{\|\tilde{\mathbf{C}}\|\leq 1} \left|\frac{1}{ML}\mathrm{Tr}\,((\tilde{\mathbf{R}}-\tilde{\mathbf{T}})\tilde{\mathbf{C}})\right| \leq \mathcal{O}(\frac{L}{MN}) + \frac{\sigma^2 c_N|z|}{(\mathrm{Im}(z))^2} \sup_{\|\mathbf{C}\|\leq 1} \left|\frac{1}{ML}\mathrm{Tr}\,((\mathbf{R}-\mathbf{T})\mathbf{C})\right| + \frac{\sigma^2 c_N\kappa^2|z|^2}{(\mathrm{Im}(z))^4} \sup_{\|\tilde{\mathbf{C}}\|\leq 1} \left|\frac{1}{N}\mathrm{Tr}\,((\tilde{\mathbf{R}}-\tilde{\mathbf{T}})\tilde{\mathbf{C}})\right| \tag{3.8.8}$$

Denote by $\delta = \sup_{\|\mathbf{C}\|\leq 1} \left|\frac{1}{ML}\mathrm{Tr}\,((\mathbf{R}-\mathbf{T})\mathbf{C})\right|$ and $\tilde{\delta} = \sup_{\|\tilde{\mathbf{C}}\|\leq 1} \left|\frac{1}{ML}\mathrm{Tr}\,((\tilde{\mathbf{R}}-\tilde{\mathbf{T}})\tilde{\mathbf{C}})\right|$, we have that

$$\begin{cases} \delta \leq \frac{\sigma^2|z|}{(\mathrm{Im}(z))^2}\tilde{\delta} + \frac{\sigma^2\kappa^2|z|^2}{(\mathrm{Im}(z))^4}\delta + \mathcal{O}(\frac{L}{MN}) \\ \tilde{\delta} \leq \frac{\sigma^2 c_N|z|}{(\mathrm{Im}(z))^2}\delta + \frac{\sigma^2 c_N\kappa^2|z|^2}{(\mathrm{Im}(z))^4}\tilde{\delta} + \mathcal{O}(\frac{L}{MN}) \end{cases} \tag{3.8.9}$$

which is equivalent to

$$\begin{cases} (1 - \frac{\sigma^2\kappa^2|z|^2}{(\mathrm{Im}(z))^4})\delta \leq \frac{\sigma^2|z|}{(\mathrm{Im}(z))^2}\tilde{\delta} + \mathcal{O}(\frac{L}{MN}) \\ (1 - \frac{\sigma^2 c_N\kappa^2|z|^2}{(\mathrm{Im}(z))^4})\tilde{\delta} \leq \frac{\sigma^2 c_N|z|}{(\mathrm{Im}(z))^2}\delta + \mathcal{O}(\frac{L}{MN}) \end{cases} \tag{3.8.10}$$

Let $z \in \mathbb{C}^+$ such that $\frac{|z|}{\mathrm{Im}(z)} \leq 2$. For $\mathrm{Im}(z)$ large enough, $1 - \frac{\sigma^2\kappa^2|z|^2}{(\mathrm{Im}(z))^4} \leq \frac{1}{2}$ and $1 - \frac{\sigma^2 c_N\kappa^2|z|^2}{(\mathrm{Im}(z))^4} \leq \frac{1}{2}$. (3.8.10) thus verify that

$$\begin{cases} \delta \leq \frac{\sigma^2|z|}{(\mathrm{Im}(z))^2}\frac{1}{1-\frac{\sigma^2\kappa^2|z|^2}{(\mathrm{Im}(z))^4}}\tilde{\delta} + \mathcal{O}(\frac{L}{MN}) \\ \tilde{\delta} \leq \frac{\sigma^2 c_N|z|}{(\mathrm{Im}(z))^2}\frac{1}{1-\frac{\sigma^2 c_N\kappa^2|z|^2}{(\mathrm{Im}(z))^4}}\delta + \mathcal{O}(\frac{L}{MN}) \end{cases} \tag{3.8.11}$$

which leads to

$$\begin{cases} \delta \leq \frac{\sigma^4 c_N|z|^2}{(\mathrm{Im}(z))^4}\frac{1}{(1-\frac{\sigma^2\kappa^2|z|^2}{(\mathrm{Im}(z))^4})(1-\frac{\sigma^2 c_N\kappa^2|z|^2}{(\mathrm{Im}(z))^4})}\delta + \mathcal{O}(\frac{L}{MN}) \\ \tilde{\delta} \leq \frac{\sigma^4 c_N|z|^2}{(\mathrm{Im}(z))^4}\frac{1}{(1-\frac{\sigma^2\kappa^2|z|^2}{(\mathrm{Im}(z))^4})(1-\frac{\sigma^2 c_N\kappa^2|z|^2}{(\mathrm{Im}(z))^4})}\tilde{\delta} + \mathcal{O}(\frac{L}{MN}) \end{cases}$$

Therefore, if $z \in \mathbb{C}^+$ such that $\frac{|z|}{\mathrm{Im}(z)} \leq 2$, for $\mathrm{Im}(z)$ large enough, we obtain that

$$\delta = \sup_{\|\mathbf{C}\|\leq 1} \left| \frac{1}{ML} \mathrm{Tr}\left((\mathbf{R}-\mathbf{T})\mathbf{C}\right) \right| = \mathscr{O}(\frac{L}{MN})$$
$$\tilde{\delta} = \sup_{\|\tilde{\mathbf{C}}\|\leq 1} \left| \frac{1}{N} \mathrm{Tr}\left((\tilde{\mathbf{R}}-\tilde{\mathbf{T}})\tilde{\mathbf{C}}\right) \right| = \mathscr{O}(\frac{L}{MN})$$

This establish (3.8.3) for uniformly bounded $ML \times ML$ and $N \times N$ matrices $\mathbf{B}, \tilde{\mathbf{B}}$ whenever z is well chosen. Moreover, for these values of z, $\frac{1}{ML}\mathrm{Tr}\left((\mathbf{R}-\mathbf{T})\mathbf{B}\right)$ and $\frac{1}{N}\mathrm{Tr}\left((\tilde{\mathbf{R}}-\tilde{\mathbf{T}})\tilde{\mathbf{B}}\right)$ are $\mathscr{O}\left(\frac{L}{MN}\right)$ terms and thus $\frac{1}{ML}\mathrm{Tr}\left((\mathbb{E}(\mathbf{Q})-\mathbf{T})\mathbf{B}\right)$ and $\frac{1}{N}\mathrm{Tr}\left((\mathbb{E}(\tilde{\mathbf{Q}})-\tilde{\mathbf{T}})\tilde{\mathbf{B}}\right)$ are $\mathscr{O}\left(\frac{L}{MN}\right)$ terms. A standard application of Montel's theorem (see Proposition 3.1.5) implies that (3.8.3) holds on $\mathbb{C} \setminus \mathbb{R}^+$. This, in turn, establish (3.8.1)

□

3.9 Convergence towards the deterministic equivalents: the bilinear forms

In this section, we prove the following Proposition. We recall that $\delta_N = \sup\left(\frac{L^2}{MN}, \sqrt{\frac{L}{M^3}}\right)$.

Proposition 3.9.1. *Under Assumption A-3, it holds that*

$$\|\mathbf{R}(z)-\mathbf{T}(z)\| = \sup_{\|\mathbf{b}_1\|=\|\mathbf{b}_2\|=1} \left|\mathbf{b}_1^*(\mathbf{R}(z)-\mathbf{T}(z))\mathbf{b}_2\right| \to 0 \quad (3.9.1)$$
$$\|\tilde{\mathbf{R}}(z)-\tilde{\mathbf{T}}(z)\| = \sup_{\|\tilde{\mathbf{b}}_1\|=\|\tilde{\mathbf{b}}_2\|=1} \left|\tilde{\mathbf{b}}_1^*(\tilde{\mathbf{R}}(z)-\tilde{\mathbf{T}}(z))\tilde{\mathbf{b}}_2\right| \to 0 \quad (3.9.2)$$

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

Proof. We consider 2 unit vectors $\mathbf{b}_1$ and $\mathbf{b}_2$. It is easy to check that

$$\begin{aligned}
\mathbf{b}_1^*(\mathbf{R}-\mathbf{T})\mathbf{b}_2 &= \sigma^2 z \mathbf{b}_1^* \mathbf{R}\, \mathbf{I}_M \otimes \mathscr{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}^T) - \tilde{\mathbf{R}}^T)\, \mathbf{T}\mathbf{b}_2 \\
&+ \sigma^2 c_N \mathbf{b}_1^* \mathbf{R}\mathbf{A}\left(\mathbf{I}_N + \sigma^2 c_N \mathscr{T}_{N,L}^{(M)}(\mathbf{T}^T)\right)^{-1} \mathscr{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T) - \mathbf{R}^T)\mathbf{H}^T\mathbf{A}^*\mathbf{T}\mathbf{b}_2 \\
&+ \sigma^2 z \mathbf{b}_1^* \mathbf{R}\, \mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\mathbf{R}}^T - \tilde{\mathbf{T}}^T)\, \mathbf{T}\mathbf{b}_2 \\
&+ \sigma^2 c_N \mathbf{b}_1^* \mathbf{R}\mathbf{A}\left(\mathbf{I}_N + \sigma^2 c_N \mathscr{T}_{N,L}^{(M)}(\mathbf{T}^T)\right)^{-1} \mathscr{T}_{N,L}^{(M)}(\mathbf{R}^T - \mathbf{T}^T)\mathbf{H}^T\mathbf{A}^*\mathbf{T}\mathbf{b}_2
\end{aligned}$$

So that we have

$$\begin{aligned}
|\mathbf{b}_1^*(\mathbf{R}-\mathbf{T})\mathbf{b}_2| &\leq \sigma^2 |z|\, \|\mathbf{b}_1\|\, \|\mathbf{b}_2\|\, \|\mathbf{R}\|\, \|\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}^T) - \tilde{\mathbf{R}}^T)\|\, \|\mathbf{T}\| \\
&+ \sigma^2 c_N \|\mathbf{b}_1\|\, \|\mathbf{b}_2\|\, \|\mathbf{R}\|\, \|\mathbf{A}\|\, \|\left(\mathbf{I}_N + \sigma^2 c_N \mathscr{T}_{N,L}^{(M)}(\mathbf{T}^T)\right)^{-1}\|\, \|\mathscr{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T) - \mathbf{R}^T)\|\, \|\mathbf{H}^T\|\, \|\mathbf{A}^*\|\, \|\mathbf{T}\| \\
&+ \sigma^2 |z| \|\mathbf{b}_1\|\, \|\mathbf{b}_2\|\, \|\mathbf{R}\|\, \|\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\mathbf{R}}^T - \tilde{\mathbf{T}}^T)\|\, \|\mathbf{T}\| \\
&+ \sigma^2 c_N \|\mathbf{b}_1\|\, \|\mathbf{b}_2\|\, \|\mathbf{R}\|\, \|\mathbf{A}\|\, \|\left(\mathbf{I}_N + \sigma^2 c_N \mathscr{T}_{N,L}^{(M)}(\mathbf{T}^T)\right)^{-1}\|\, \|\mathscr{T}_{N,L}^{(M)}(\mathbf{R}^T - \mathbf{T}^T)\|\, \|\mathbf{H}^T\|\, \|\mathbf{A}^*\|\, \|\mathbf{T}\|
\end{aligned}$$

We assume that $z \in \mathbb{C}^+$. We recall that $\sup_N \|\mathbf{A}\| \leq C < +\infty$, $\|\mathbf{H}\| \leq \frac{|z|}{\mathrm{Im}(z)}$, $\|\mathbf{R}\| \leq \frac{1}{(\mathrm{Im}(z))}$, $\|\mathbf{T}\| \leq \frac{1}{\mathrm{Im}(z)}$ and $\left\|\left(\mathbf{I}_N + \sigma^2 c_N \mathscr{T}_{N,L}^{(M)}(\mathbf{T}^T)\right)^{-1}\right\| \leq \frac{|z|}{\mathrm{Im}(z)}$. Therefore, it holds that

$$\begin{aligned}
\sup_{\|\mathbf{b}_1\|=\|\mathbf{b}_2\|=1} |\mathbf{b}_1^*(\mathbf{R}-\mathbf{T})\mathbf{b}_2| = \|\mathbf{R}-\mathbf{T}\| &\leq \frac{\sigma^2|z|}{(\mathrm{Im}(z))^2} \|\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{R}})\| + \frac{\sigma^2 c_N \kappa^2 |z|^2}{(\mathrm{Im}(z))^4} \|\mathscr{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}) - \mathbf{R})\| \\
&+ \frac{\sigma^2|z|}{(\mathrm{Im}(z))^2} \|\mathbf{I}_M \otimes \mathscr{T}_{L,L}(\tilde{\mathbf{R}} - \tilde{\mathbf{T}})\| + \frac{\sigma^2 c_N \kappa^2 |z|^2}{(\mathrm{Im}(z))^4} \|\mathscr{T}_{N,L}^{(M)}(\mathbf{R} - \mathbf{T})\|
\end{aligned}$$

We recall that $\|\mathbb{E}(\mathbf{Q}) - \mathbf{R}\| = \mathcal{O}(\delta_N)$ and that $\|\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{R}}\| = \mathcal{O}(\delta_N)$. Proposition 3.1.3 thus implies that,

$$\|\mathbf{R}-\mathbf{T}\| \leq \mathcal{O}(\delta_N) + \frac{\sigma^2|z|}{(\mathrm{Im}(z))^2}\|\tilde{\mathbf{R}}-\tilde{\mathbf{T}})\| + \frac{\sigma^2 c_N \kappa^2 |z|^2}{(\mathrm{Im}(z))^4}\|\mathbf{R}-\mathbf{T}\| \tag{3.9.3}$$

We denote by $\gamma = \|\mathbf{R}-\mathbf{T}\|$ and $\tilde{\gamma} = \|\tilde{\mathbf{R}}-\tilde{\mathbf{T}}\|$. (3.9.3) leads to

$$\gamma \leq \mathcal{O}(\delta_N) + \frac{\sigma^2|z|}{(\mathrm{Im}(z))^2}\tilde{\gamma} + \frac{\sigma^2 c_N \kappa^2 |z|^2}{(\mathrm{Im}(z))^4}\gamma$$

Using the same calculations, we get that

$$\tilde{\gamma} \leq \mathcal{O}(\delta_N) + \frac{\sigma^2 c_N|z|}{(\mathrm{Im}(z))^2}\gamma + \frac{\sigma^2 \kappa^2 |z|^2}{(\mathrm{Im}(z))^4}\tilde{\gamma}$$

From this, it is straightforward to check that if $z \in \mathbb{C}^+$ verifies $\frac{|z|}{\mathrm{Im}(z)} \leq 2$ and $\mathrm{Im}(z)$ large enough, we have

$$\gamma = \|\mathbf{R}(z) - \mathbf{T}(z)\| = \mathcal{O}(\delta_N)$$
$$\tilde{\gamma} = \|\tilde{\mathbf{R}}(z) - \tilde{\mathbf{T}}(z)\| = \mathcal{O}(\delta_N)$$

We now establish that $\|\mathbf{R}(z) - \mathbf{T}(z)\|$ and $\|\tilde{\mathbf{R}}(z) - \tilde{\mathbf{T}}(z)\|$ converge towards 0 for each $z \in \mathbb{C} \setminus \mathbb{R}^+$ using an argument "à la Montel". For this, we assume that it exists $z_0 \in \mathbb{C} \setminus \mathbb{R}^+$ for which $\|\mathbf{R}_N(z_0) - \mathbf{T}_N(z_0)\|$ does not converge towards 0. We consider 2 unit vectors $\mathbf{b}_{1,N}(z_0), \mathbf{b}_{2,N}(z_0)$ such that $\|\mathbf{R}_N(z_0) - \mathbf{T}_N(z_0)\| = |\mathbf{b}_{1,N}(z_0)^*(\mathbf{R}_N(z_0) - \mathbf{T}_N(z_0))\mathbf{b}_{2,N}(z_0)|$. Then, sequence $\mathbf{b}_{1,N}(z_0)^*(\mathbf{R}_N(z_0) - \mathbf{T}_N(z_0))\mathbf{b}_{2,N}(z_0)$ does not converge towards 0. However, if we consider the sequence of analytic functions $f_N(z)$ defined by $\mathbf{b}_{1,N}(z_0)^*(\mathbf{R}_N(z) - \mathbf{T}_N(z))\mathbf{b}_{2,N}(z_0)$, $f_N(z)$ converges towards 0 when $z \in \mathbb{C}^+$ verifies $\frac{|z|}{\mathrm{Im}(z)} \leq 2$ and $\mathrm{Im}(z)$ large enough. Sequence f_N is a normal family on $\mathbb{C} \setminus \mathbb{R}^+$. Therefore, by Montel's theorem, $f_N(z)$ converges towards 0 for each $z \in \mathbb{C} \setminus \mathbb{R}^+$, and also for $z = z_0$, a contradiction. Therefore, (3.9.1) holds for each $z \in \mathbb{C} \setminus \mathbb{R}^+$. (3.9.2) is proved in the same way.

Remark 10. *We remark that Proposition3.9.1 does not provide the rate of convergence of* $\|\mathbf{R}(z) - \mathbf{T}(z)\|$ *and* $\|\tilde{\mathbf{R}}(z) - \tilde{\mathbf{T}}(z)\|$ *because analytic continuation arguments does not allow to preserve the rates of convergence. We suspect that it is possible to establish that* $\|\mathbf{R}(z) - \mathbf{T}(z)\|$ *and* $\|\tilde{\mathbf{R}}(z) - \tilde{\mathbf{T}}(z)\|$ *are* $\mathcal{O}(\delta_N)$ *for each* $z \in \mathbb{C} \setminus \mathbb{R}^+$. *However, the proof of this result does not seem to be an easy task.*

As we have shown the $\|\mathbb{E}(\mathbf{Q}(z)) - \mathbf{R}(z)\|$ and $\|\mathbb{E}(\tilde{\mathbf{Q}}(z)) - \tilde{\mathbf{R}}(z)\|$ converge towards 0 (see Corollary 3.6.1), the following Corollary holds.

Corollary 3.9.1. *Under Assumption A-3, it holds that*

$$\|\mathbb{E}(\mathbf{Q}(z)) - \mathbf{T}(z)\| = \sup_{\|\mathbf{b}_1\|=\|\mathbf{b}_2\|} \left|\mathbf{b}_1^*(\mathbb{E}(\mathbf{Q}(z)) - \mathbf{T}(z))\mathbf{b}_2\right| \to 0 \tag{3.9.4}$$

$$\|\mathbb{E}(\tilde{\mathbf{Q}}(z)) - \tilde{\mathbf{T}}(z)\| = \sup_{\|\tilde{\mathbf{b}}_1\|=\|\tilde{\mathbf{b}}_2\|} \left|\tilde{\mathbf{b}}_1^*(\mathbb{E}(\tilde{\mathbf{Q}}(z)) - \tilde{\mathbf{T}}(z))\tilde{\mathbf{b}}_2\right| \to 0 \tag{3.9.5}$$

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

Proof. It just remains to justify that $\|\mathbb{E}(\mathbf{Q}(z)) - \mathbf{R}(z)\|$ and $\|\mathbb{E}(\tilde{\mathbf{Q}}(z)) - \tilde{\mathbf{R}}(z)\|$ converge towards 0 for $z \in \mathbb{C} \setminus \mathbb{R}^+$. For this, it is sufficient to use again the argument "à la Montel" used in the proof of Proposition 3.9.1.

□

3.A Proof of lemma 3.5.1

We use the same ingredients than in the proof of Lemma 5-1 of [28]. The invertibility of $\mathbf{I}_N+\sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbb{E}(\mathbf{Q}(z)))$ for $z \in \mathbb{C}^+$ is a direct consequence of $\text{Im}(\mathbf{Q}(z)) > 0$ on $\mathbb{C}^+$ as well as of Proposition 3.1.2 . In order to prove (3.5.2), we first establish that function $\mathbf{G}_1(z)$ defined by

$$\mathbf{G}_1(z) = -\frac{\mathbf{H}(z)}{z}$$

coincides with the Stieltjes transform of a positive $\mathbb{C}^{N\times N}$ matrix valued measure $\mathbf{v}_1$ carried by $\mathbb{R}^+$ such that $\mathbf{v}_1(\mathbb{R}^+) = \mathbf{I}_N$, i.e

$$\mathbf{G}_1(z) = \int_{\mathbb{R}^+} \frac{d\mathbf{v}_1}{\lambda - z}$$

For this it is sufficient to check that $\text{Im}(\mathbf{G}_1(z))$ and $\text{Im}(z\mathbf{G}_1(z))$ are both positive on $\mathbb{C}^+$, and that $\lim_{y\to+\infty} -iy\mathbf{G}_1(iy) = \mathbf{I}_N$ (see proof of Lemma 5-1 of [28]).
$\text{Im}(\mathbf{G}_1(z))$ can be written as

$$\text{Im}(\mathbf{G}_1(z)) = \frac{\mathbf{H}(z)}{z}\frac{1}{2i}\left[z\mathbf{H}^{-1}(z) - z^*(\mathbf{H}^{-1}(z))^*\right]\frac{\mathbf{H}(z)^*}{z^*}$$

which is equivalent to

$$\text{Im}(\mathbf{G}_1(z)) = \frac{\mathbf{H}(z)}{z}\mathbb{E}\left[\text{Im}(z) + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\text{Im}(z\mathbf{Q}(z)))\right]\frac{\mathbf{H}(z)^*}{z^*}$$

As $\text{Im}(z\mathbf{Q}(z)) > 0$ on $\mathbb{C}^+$, this implies that

$$\text{Im}(\mathbf{G}_1(z)) > \frac{\text{Im}(z)}{|z|^2}\mathbf{H}(z)\mathbf{H}(z)^* \geq 0$$

The term $z\mathbf{G}_1(z) = -\mathbf{H}(z)$, so that we have

$$\text{Im}(z\mathbf{G}_1(z)) = -\text{Im}(\mathbf{H}(z)) = -\mathbf{H}(z)\left[(\mathbf{H}^{-1}(z))^* - \mathbf{H}^{-1}(z)\right]\mathbf{H}(z)^*$$

which is equivalent to

$$\text{Im}(z\mathbf{G}_1(z)) = \sigma^2 c_N \mathbf{H}(z)\mathcal{T}^{(M)}_{N,L}(\mathbb{E}(\text{Im}(\mathbf{Q}(z)))\mathbf{H}(z)^*$$

As $\text{Im}(\mathbf{Q}(z)) > 0$ on $\mathbb{C}^+$, thus $\text{Im}(z\mathbf{G}_1(z))$ is positive.
Moreover, since $\|\mathbf{Q}(z)\| \leq \frac{1}{\text{Im}(z)}$, we can easily get

$$\lim_{y\to+\infty} -iy\mathbf{G}_1(iy) = \lim_{y\to+\infty}\left[\mathbf{I}_N + \sigma^2 c_N \mathcal{T}^{(M)}_{N,L}(\mathbb{E}(\mathbf{Q}(iy)))\right]^{-1} = \mathbf{I}_N$$

Therefore, since $\mathbf{G}_1(z)$ is the Stieltjes transform of $\mathbf{v}_1$, it is clear that

$$\text{Im}(\mathbf{G}_1(z)) = \text{Im}(z)\int_{\mathbb{R}^+}\frac{d\mathbf{v}_1}{|\lambda - z|^2} \leq \frac{1}{\text{Im}(z)}\mathbf{I}_N$$

So that we have

$$\frac{1}{\text{Im}(z)} \geq \text{Im}(\mathbf{G}_1(z)) > \frac{\text{Im}(z)}{|z|^2}\mathbf{H}(z)\mathbf{H}(z)^*$$

which implies

$$\mathbf{H}(z)\mathbf{H}(z)^* \leq \left(\frac{|z|}{\text{Im}(z)}\right)^2 \mathbf{I}_N$$

The other statements of lemma 3.5.1 are proved similarly.

Chapter 4

Convergence towards spiked model : the case of wideband array processing models

4.1 Introduction

In this chapter, we take benefit of the results of chapter 3 to revisit the information plus noise block-Hankel random matrix model built from samples collected from the M–dimensional time series $(\mathbf{y}_n)_{n\in\mathbb{Z}}$ given by

$$\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p} + \mathbf{v}_n, \; n = 1, ..., N$$

between time 1 and time N. Here, $\mathbf{h}(z) = \sum_{p=0}^{P-1} \mathbf{h}_p z^{-p}$ represents the transfer function of an unknown FIR(Finite Impulse Response) SIMO complex channel, s is a scalar non observable sequence and $(\mathbf{v}_n)_{n\in\mathbb{Z}}$ is an i.i.d sequence of $\mathcal{N}_c(0, \sigma^2 \mathbf{I})$ random vectors.

If L is an integer, for each $m = 1, \ldots, M$, we define for $n = 1, \ldots, N$ the L–dimensional vector $\mathbf{y}_{m,n}^{(L)}$ given by

$$\mathbf{y}_{m,n}^{(L)} = (y_{m,n}, ..., y_{m,n+L-1})^T$$

and the ML–dimensional vector $\mathbf{y}_n^{(L)}$ given by

$$\mathbf{y}_n^{(L)} = ((\mathbf{y}_{1,n}^{(L)})^T, (\mathbf{y}_{2,n}^{(L)})^T, ..., (\mathbf{y}_{M,n}^{(L)})^T)^T$$

We denote by $\mathbf{Y}_N^{(L)}$ the ML × N matrix defined by

$$\mathbf{Y}_N^{(L)} = (\mathbf{y}_1^{(L)}, ..., \mathbf{y}_N^{(L)})$$

and notice that $\mathbf{Y}_N^{(L)}$ is given by

$$\mathbf{Y}_N^{(L)} = \begin{bmatrix} \mathbf{Y}_{1,N}^{(L)} \\ \vdots \\ \mathbf{Y}_{M,N}^{(L)} \end{bmatrix}$$

in which, for each m, matrix $\mathbf{Y}_{m,N}^{(L)}$ is defined as

$$\mathbf{Y}_{m,N}^{(L)} = \begin{pmatrix} \mathbf{y}_{m,1} & \mathbf{y}_{m,2} & \cdots & \mathbf{y}_{m,N} \\ \mathbf{y}_{m,2} & \mathbf{y}_{m,3} & \cdots & \mathbf{y}_{m,N+1} \\ \mathbf{y}_{m,3} & \cdots & \cdots & \mathbf{y}_{m,N+2} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_{m,L} & \mathbf{y}_{m,L+1} & \cdots & \mathbf{y}_{m,N+L-1} \end{pmatrix}$$

Moreover we notice that for each m, matrix $\mathbf{Y}^{(L)}_{m,N}$ is equal to $\mathcal{H}^{(L)}_{m,N}\mathbf{S}^{(L)}_N + \mathbf{V}^{(L)}_{m,N}$, where $\mathbf{V}^{(L)}_{m,N}$ has the same structure of $\mathbf{Y}^{(L)}_{m,N}$, and $\mathcal{H}^{(L)}_{m,N}$ is the $L \times (P+L-1)$ Toeplitz matrix defined by

$$\mathcal{H}^{(L)}_{m,N} = \begin{pmatrix} h_{m,P-1} & \cdots & h_{m,0} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h_{m,P-1} & \cdots & h_{m,0} \end{pmatrix}$$

and $\mathbf{S}^{(L)}_N$ is the $(P+L-1) \times N$ Hankel matrix defined by

$$\mathbf{S}^{(L)}_N = \begin{pmatrix} s_{2-P} & \cdots & s_{N-P+1} \\ \vdots & & \vdots \\ s_L & \cdots & s_{N+L-1} \end{pmatrix}$$

Therefore, it holds that

$$\mathbf{Y}^{(L)}_N = \begin{pmatrix} \mathcal{H}^{(L)}_{1,N} \\ \vdots \\ \mathcal{H}^{(L)}_{M,N} \end{pmatrix} \mathbf{S}^{(L)}_N + \begin{pmatrix} \mathbf{V}^{(L)}_{1,N} \\ \vdots \\ \mathbf{V}^{(L)}_{M,N} \end{pmatrix} = \mathcal{H}^{(L)}_N \mathbf{S}^{(L)}_N + \mathbf{V}^{(L)}_N$$

Finally, we rather consider normalized versions of the above matrices, i.e. define $\mathbf{\Sigma}_N = \frac{1}{\sqrt{N}}\mathbf{Y}^{(L)}_N$, $\mathbf{A}_N = \mathcal{H}^{(L)}_N \frac{1}{\sqrt{N}}\mathbf{S}^{(L)}_N$ and $\mathbf{W}_N = \frac{1}{\sqrt{N}}\mathbf{V}^{(L)}_N$. The model thus becomes

$$\mathbf{\Sigma}_N = \mathbf{A}_N + \mathbf{W}_N \tag{4.1.1}$$

As in the previous chapter, we study this model in the asymptotic regime

Assumption A-4: *We assume that* ML *and* N *converge towards* $+\infty$ *in such a way that* $c_N = \frac{ML}{N}$ *satisfies* $c_N \to c$ *where* $0 < c < +\infty$*, and that moreover,* P, L *are* $\mathcal{O}(N^{\alpha})$ *with* $0 < \alpha < \frac{2}{3}$.

Matrix $\mathbf{\Sigma}_N$ is a particular case of random matrix models defined in the previous chapter by Eqs. (3.3.3), except that matrix $\mathbf{A}_N$ has the particular structure $\mathbf{A}_N = \mathcal{H}^{(L)}_N \frac{1}{\sqrt{N}}\mathbf{S}^{(L)}_N$. $\mathbf{A}_N$ is in particular a rank $P+L-1$ matrix, which, under the Assumption A-4, is much smaller than N. Matrix $\mathbf{\Sigma}_N$ is also similar to model (2.4.5) of chapter 2 except that P and L may scale with M and N in the context of the present chapter. Therefore, the methods used in chapter 2 are no more valid. The purpose of the present chapter is to extend the results of section 2.4 of chapter 2 to the case where P and L may converge towards $+\infty$ according to Assumption A-(4). For this, we take benefit of the general results derived in chapter 3.

4.2 Overview of the results.

As in the previous chapters, we denote by $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$ the resolvents of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ and $\mathbf{\Sigma}_N^*\mathbf{\Sigma}_N$ respectively. As $P+L-1 = o(N)$, it is clear that the empirical eigenvalue distribution of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ converges towards the Marcenko-Pastur distribution μ_{c,σ^2} of parameters (c,σ^2), i.e. the normalized trace of $\mathbf{Q}_N(z)$ has the same asymptotic behaviour than the Stieltjes transform of μ_{c,σ^2}. However, this property is no longer true for the bilinear forms of $\mathbf{Q}_N(z)$ whose asymptotic behaviour is useful to extend the results of section 2.4 of chapter 2 to the case where P and L may converge towards $+\infty$.

In section 4.3, we study the asymptotic behaviour of the bilinear forms of $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$ in the context of model (4.1.1). The results of chapter 3 show that these bilinear forms have the same behaviour than the forms of $\mathbf{T}_N(z)$ and $\tilde{\mathbf{T}}_N(z)$. As matrices $\mathbf{T}_N(z)$ and $\tilde{\mathbf{T}}_N(z)$ have complicated expressions, we use the particular structure of matrix $\mathbf{A}_N = \mathcal{H}_N^{(L)} \frac{1}{\sqrt{N}} \mathbf{S}_N^{(L)}$ to establish that the bilinear forms of $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$ behave as if P and L were fixed integers. More precisely, we define $\mathbf{T}_{1,N}(z)$ and $\tilde{\mathbf{T}}_{1,N}(z)$ as the solutions of the equation

$$\begin{cases} \mathbf{T}_{1,N}(z) = \left(-z(1+\sigma^2 \tilde{t}_{1,N}(z))\mathbf{I}_{ML} + \frac{\mathbf{A}_N\mathbf{A}_N^*}{1+\sigma^2 c_N t_{1,N}(z)}\right)^{-1} \\ \tilde{\mathbf{T}}_{1,N}(z) = \left(-z(1+\sigma^2 c_N t_{1,N}(z))\mathbf{I}_N + \frac{\mathbf{A}_N^*\mathbf{A}_N}{1+\sigma^2 \tilde{t}_{1,N}(z)}\right)^{-1} \end{cases} \tag{4.2.1}$$

where $t_{1,N}(z) = \frac{1}{ML}\mathrm{Tr}(\mathbf{T}_{1,N}(z))$ and where $\tilde{t}_{1,N}(z) = \frac{1}{N}\mathrm{Tr}(\tilde{\mathbf{T}}_{1,N}(z))$ also coincides with $c_N t_{1,N}(z) - (1-c_N)/z$. It is clear that $\mathbf{T}_{1,N}(z)$ and $\tilde{\mathbf{T}}_{1,N}(z)$ correspond to the deterministic equivalents of the resolvents of matrices $\mathbf{\Sigma}_{1,N}\mathbf{\Sigma}_{1,N}^*$ and $\mathbf{\Sigma}_{1,N}^*\mathbf{\Sigma}_{1,N}$ where $\mathbf{\Sigma}_{1,N}$ represents a virtual information plus noise model

$$\mathbf{\Sigma}_{1,N} = \mathbf{A}_N + \mathbf{W}_{1,N} \tag{4.2.2}$$

where $\mathbf{W}_{1,N}$ represents a $ML \times N$ random matrix with $\mathcal{N}_c(0, \sigma^2/N)$ **independent, identically distributed entries**, see . Eq (1.2.40, 1.2.43) in chapter 1. Then, we establish that $\mathbf{T}_N(z) \simeq \mathbf{T}_{1,N}(z)$ and $\tilde{\mathbf{T}}_N(z) \simeq \tilde{\mathbf{T}}_{1,N}(z)$ where we recall that the symbol $\simeq$ means that the bilinear forms of the 2 matrices have the same asymptotic behaviour. Moreover, as $P+L-1 = o(N)$, Stieltjes transforms $t_{1,N}(z)$ and $\tilde{t}_{1,N}(z)$ can be replaced by the Stieltjes transforms $m_c(z)$ and $\tilde{m}_c(z)$ associated to the Marcenko-Pastur distribution μ_{c,σ^2}. If we denote by $\mathbf{S}_N(z)$ and $\tilde{\mathbf{S}}_N(z)$ the functions defined by

$$\begin{cases} \mathbf{S}_N(z) = \left(-z(1+\sigma^2 \tilde{m}_c(z))\mathbf{I}_{ML} + \frac{\mathbf{A}_N\mathbf{A}_N^*}{1+\sigma^2 c m_c(z)(z)}\right)^{-1} \\ \tilde{\mathbf{S}}_N(z) = \left(-z(1+\sigma^2 c m_c(z))\mathbf{I}_N + \frac{\mathbf{A}_N^*\mathbf{A}_N}{1+\sigma^2 \tilde{m}_c(z)}\right)^{-1} \end{cases} \tag{4.2.3}$$

then, it also holds that $\mathbf{T}_{1,N}(z) \simeq \mathbf{S}_N(z)$ and $\tilde{\mathbf{T}}_{1,N}(z) \simeq \tilde{\mathbf{S}}_N(z)$. It is shown in chapter 2 that $\mathbf{Q}_N(z) \simeq \mathbf{S}_N(z)$ and $\tilde{\mathbf{Q}}_N(z) \simeq \tilde{\mathbf{S}}_N(z)$ when P and L do not scale with N. Therefore, we generalize this result to the case where P and L converge towards $+\infty$ at rate strictly less than $\mathcal{O}(N^{2/3})$.

4.3 Simplified behaviour of the bilinear forms of $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$.

In this section, we prove that $\mathbf{T}_N(z) \simeq \mathbf{T}_{1,N}(z)$ and $\tilde{\mathbf{T}}_N(z) \simeq \tilde{\mathbf{T}}_{1,N}(z)$ for each $z \in \mathbb{C}^+$. Montel's theorem will thus imply that this equivalence also holds for $z \in \mathbb{C} \setminus \mathbb{R}^+$. The proof of $\mathbf{T}_{1,N}(z) \simeq \mathbf{S}_N(z)$ and $\tilde{\mathbf{T}}_{1,N}(z) \simeq \tilde{\mathbf{S}}_N(z)$ is straightforward, and thus omitted.

The above equivalence holds true however under certain hypotheses on filter $\mathbf{h}(z)$ and on sequence s. More precisely, we formulate the following assumptions. $(h^{(m)}(z))_{m=1,\ldots,M}$ represent the components of $\mathbf{h}(z)$, and if $\nu \in [0,1]$, we denote by $h^{(m)}(\nu)$ and $\mathbf{h}(\nu)$ the terms $h^{(m)}(e^{2i\pi\nu})$ and $\mathbf{h}(e^{2i\pi\nu})$.

Assumption A-5: *We assume that*

$$\sup_M \sup_{\nu\in[0,1]} \sum_{m=1}^{M} |h^{(m)}(\nu)|^2 < +\infty.$$

Remark 11. *Assumption A-5 implies that* $\sup_N \|\mathcal{H}_N^{(L)}\| < +\infty$. *To see this, we can notice that*

$$\mathcal{H}_m^{(L)} = \int_{[0,1]} e^{2i\pi(P-1)\nu}\, h^{(m)}(\nu)\, \mathbf{d}_L(\nu)\mathbf{d}_{P+L-1}^*(\nu)\, d\nu$$

and that

$$\mathcal{H}_N^{(L)} = \begin{pmatrix} \mathcal{H}_{1,N}^{(L)} \\ \vdots \\ \mathcal{H}_{M,N}^{(L)} \end{pmatrix} = \int_{[0,1]} (e^{2i\pi(P-1)\nu}\mathbf{h}(\nu) \otimes \mathbf{d}_L(\nu))\mathbf{d}_{P+L-1}^*(\nu)\, d\nu$$

Using the definition of spectral norm, we can write

$$\|\mathcal{H}_N^{(L)}\| = \sup_{\substack{\mathbf{b}_1\in\mathbb{C}^{ML};\mathbf{b}_2\in\mathbb{C}^{P+L-1}\\ \|\mathbf{b}_1\|=\|\mathbf{b}_2\|=1}} \left|\mathbf{b}_1^*\mathcal{H}_N^{(L)}\mathbf{b}_2\right| = \sup_{\substack{\mathbf{b}_1\in\mathbb{C}^{ML};\mathbf{b}_2\in\mathbb{C}^{P+L-1}\\ \|\mathbf{b}_1\|=\|\mathbf{b}_2\|=1}} \left|\int_{[0,1]} \mathbf{b}_1^*(e^{2i\pi(P-1)\nu}\mathbf{h}(\nu)\otimes\mathbf{d}_L(\nu))\mathbf{d}_{P+L-1}^*(\nu)\mathbf{b}_2 d\nu\right|$$

We express $\mathbf{b}_1 = (\mathbf{b}_{1,1}^T, ..., \mathbf{b}_{M,1}^T)^T$*, with* $(\mathbf{b}_{m,1})_{m=1,...,M}$ L*–dimensional vectors. Using Cauchy-Schwartz inequality, we obtain that*

$$\begin{aligned}
\left|\mathbf{b}_1^*\mathcal{H}_N^{(L)}\mathbf{b}_2\right| &= \left|\int_{[0,1]}\left(\sum_{m=1}^{M} e^{2i\pi(P-1)\nu}\, h^{(m)}(\nu)\mathbf{b}_{m,1}^*\mathbf{d}_L(\nu)\right)\mathbf{d}_{P+L-1}^*(\nu)\mathbf{b}_2 d\nu\right| \\
&\le \sqrt{\int_{[0,1]}\left|\sum_{m=1}^{M} h^{(m)}(\nu)\mathbf{b}_{m,1}^*\mathbf{d}_L(\nu)\right|^2 d\nu}\sqrt{\int_{[0,1]}\left|\mathbf{d}_{P+L-1}^*(\nu)\mathbf{b}_2\right|^2 d\nu} \\
&\le \sqrt{\int_{[0,1]}\sum_{m=1}^{M}\left|h^{(m)}(\nu)\right|^2\sum_{m=1}^{M}\left|\mathbf{b}_{m,1}^*\mathbf{d}_L(\nu)\right|^2 d\nu}\,\|\mathbf{b}_2\| \\
&\le \sqrt{\sup_{\nu\in[0,1]}\sum_{m=1}^{M}\left|h^{(m)}(\nu)\right|^2}\sqrt{\int_{[0,1]}\sum_{m=1}^{M}\left|\mathbf{b}_{m,1}^*\mathbf{d}_L(\nu)\right|^2 d\nu}\,\|\mathbf{b}_2\| \\
&\le \sqrt{\sup_{\nu\in[0,1]}\sum_{m=1}^{M}\left|h^{(m)}(\nu)\right|^2}\sqrt{\sum_{m=1}^{M}\left\|\mathbf{b}_{m,1}\right\|^2 d\nu}\,\|\mathbf{b}_2\| \\
&\le \sqrt{\sup_{\nu\in[0,1]}\sum_{m=1}^{M}\left|h^{(m)}(\nu)\right|^2}\,\|\mathbf{b}_1\|\,\|\mathbf{b}_2\|
\end{aligned}$$

Therefore, it holds that

$$\sup_N \|\mathcal{H}_N^{(L)}\| \le \sup_M \sqrt{\sup_{\nu\in[0,1]}\sum_{m=1}^{M}\left|h^{(m)}(\nu)\right|^2} < +\infty$$

■

Finally, we formulate the following assumption on signal s_n:

Assumption A-6: *We assume that* $(s_n)_{n\in\mathbb{Z}}$ *is a* ***real*** *stationary random process defined as*

$$s_n = \sum_{j=0}^{\infty}\alpha_j\epsilon_{n-j}, \;\; \sum_{j=0}^{\infty}|\alpha_j| < +\infty \;\; , \alpha_0 = 1 \tag{4.3.1}$$

where ϵ_n *is an i.i.d. sequence of zero mean unit variance random variables having finite moments of all order, and whose characteristic function* $\mathcal{Q}(\theta)$*, satisfies*

$$\sup_{|\theta|>\theta_0>0} |\mathcal{Q}(\theta)| = \beta(\theta_0) < 1$$

and $\mathbb{E}(\epsilon_n) = 0$ *and* $\mathbb{E}(\epsilon_n^2) = 1$.
In addition, we assume that, there exists a certain $\eta > \frac{1}{\alpha} - 1$ *for which*

$$\sum_{j=m}^{\infty}|\alpha_j| = \mathcal{O}\left(m^{-\eta}\right) \tag{4.3.2}$$

In the following, we denote by $\phi(\nu)$ the continuous function defined by

$$\phi(\nu) = \left|\sum_{j=0}^{+\infty} \alpha_j e^{-2i\pi\nu j}\right|^2 = \sum_{u=-\infty}^{+\infty} \gamma_u e^{-2i\pi\nu u} \tag{4.3.3}$$

where $\gamma_u = \sum_{l=0}^{+\infty} \alpha_{u+l}\overline{\alpha}_l \mathbb{1}_{u+l\geq 0} = \mathbb{E}(s_{m+u}s_m^*)$ and remark that $\phi(\nu)$ coincides with the spectral density of s.

The hypothesis that sequence s is real is certainly not necessary, and could probably be replaced by s complex circular, i.e. in representation (4.3.1), the real part and the imaginary parts of ϵ_n are mutually independent sequences sharing the same probability distribution. However, the technical results we use in the following (a moderate deviation result related to the periodogram of sequence $(s_n)_{n=1,\dots,\mathrm{N}}$) is apparently only available in the real case. Due to the lack of time, we have not tried to extend it to the complex circular case.

Lemma 4.3.1. *Under Assumption A-6, when* $\mathrm{N}\to\infty$, *the empirical covariance matrix* $\frac{1}{\mathrm{N}}\mathbf{S}_{\mathrm{N}}^{(\mathrm{L})}\mathbf{S}_{\mathrm{N}}^{(\mathrm{L})*}$ *converges almost surely towards the* $(\mathrm{P}+\mathrm{L}-1)\times(\mathrm{P}+\mathrm{L}-1)$ *Toeplitz matrix* $\mathbf{R}_{\mathrm{S}}$ *defined by*

$$\mathbf{R}_{\mathrm{S}} = \int_{[0,1]} \phi(\nu)\mathbf{d}_{\mathrm{P}+\mathrm{L}-1}(\nu)\mathbf{d}^*_{\mathrm{P}+\mathrm{L}-1}(\nu)d\nu \tag{4.3.4}$$

in the spectral norm sense. Moreover, it holds that

$$\left\|\frac{1}{\mathrm{N}}\mathbf{S}_{\mathrm{N}}^{(\mathrm{L})}\mathbf{S}_{\mathrm{N}}^{(\mathrm{L})*} - \mathbf{R}_{\mathrm{S}}\right\| = \mathcal{O}\left(\sqrt{\frac{\log(\mathrm{L})}{\mathrm{N}^{1-\alpha}}}\right), \; a.s$$

Proof. The proof is given in the appendix.

We recall that the deterministic equivalents $\mathbf{T},\tilde{\mathbf{T}}$ of the resolvents $\mathbf{Q} = (\mathbf{\Sigma}\mathbf{\Sigma}^* - z\mathbf{I}_{\mathrm{ML}})^{-1}$ and $\tilde{\mathbf{Q}} = (\mathbf{\Sigma}^*\mathbf{\Sigma} - z\mathbf{I}_{\mathrm{N}})^{-1}$ are defined by

$$\begin{cases} \mathbf{T}(z) = \left[-z\left(\mathbf{I}_{\mathrm{ML}} + \sigma^2\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\tilde{\mathbf{T}}^{\mathrm{T}}(z))\right) + \mathbf{A}\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}}\mathcal{T}^{(\mathrm{M})}_{\mathrm{N,L}}(\mathbf{T}^{\mathrm{T}}(z))\right)^{-1}\mathbf{A}^*\right]^{-1} \\ \tilde{\mathbf{T}}(z) = \left[-z\left(\mathbf{I}_{\mathrm{N}} + \sigma^2 c_{\mathrm{N}}\mathcal{T}^{(\mathrm{M})}_{\mathrm{N,L}}(\mathbf{T}^{\mathrm{T}}(z))\right) + \mathbf{A}^*\left(\mathbf{I}_{\mathrm{ML}} + \sigma^2\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\tilde{\mathbf{T}}^{\mathrm{T}}(z))\right)^{-1}\mathbf{A}\right]^{-1} \end{cases}$$

We will show that $\mathbf{T},\tilde{\mathbf{T}}$ converge in the spectral norm sense towards the deterministic equivalents of classical complex Gaussian information plus noise model (see e.g [9],[48],[32]) which we denote by $\mathbf{T}_1,\tilde{\mathbf{T}}_1$. These quantity $\mathbf{T}_1,\tilde{\mathbf{T}}_1$ are defined by

$$\begin{cases} \mathbf{T}_1 = \left(-z(1+\sigma^2\tilde{t}_1(z))\mathbf{I}_{\mathrm{ML}} + \frac{\mathbf{A}\mathbf{A}^*}{1+\sigma^2 c_{\mathrm{N}} t_1(z)}\right)^{-1} \\ \tilde{\mathbf{T}}_1 = \left(-z(1+\sigma^2 c_{\mathrm{N}} t_1(z))\mathbf{I}_{\mathrm{N}} + \frac{\mathbf{A}^*\mathbf{A}}{1+\sigma^2\tilde{t}_1(z)}\right)^{-1} \end{cases} \tag{4.3.5}$$

where $t_1(z) = \frac{1}{\mathrm{ML}}\mathrm{Tr}(\mathbf{T}_1)$ and $\tilde{t}_1(z) = \frac{1}{\mathrm{N}}\mathrm{Tr}(\tilde{\mathbf{T}}_1)$.

Proposition 4.3.1. *Under assumptions A-6 and A-5, it holds that*

$$\begin{cases} \|\mathbf{T}(z) - \mathbf{T}_1(z)\| = \sup_{\|\mathbf{b}_1\|=\|\mathbf{b}_2\|=1} |\mathbf{b}_1^*(\mathbf{T}(z) - \mathbf{T}_1(z))\mathbf{b}_2| \to 0 \\ \|\tilde{\mathbf{T}}(z) - \tilde{\mathbf{T}}_1(z)\| = \sup_{\|\tilde{\mathbf{b}}_1\|=\|\tilde{\mathbf{b}}_2\|=1} \tilde{\mathbf{b}}_1^*(\tilde{\mathbf{T}}(z) - \tilde{\mathbf{T}}_1(z))\tilde{\mathbf{b}}_2| \to 0 \end{cases} \tag{4.3.6}$$

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

Proof: By standard algebra, we have that

- $\mathbf{T}-\mathbf{T}_1 = z\sigma^2\mathbf{T}_1\,\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}^T-\tilde{\mathbf{T}}_1^T)\,\mathbf{T}+\frac{\sigma^2 c_N}{1+\sigma^2 c_N t_1(z)}\mathbf{T}_1\,\mathbf{A}\left(\mathbf{I}_N+\sigma^2 c_N\mathcal{T}_{N,L}^{(M)}(\mathbf{T}^T)\right)^{-1}\mathcal{T}_{N,L}^{(M)}(\mathbf{T}^T-\mathbf{T}_1^T)\,\mathbf{A}^*\mathbf{T}$

$$+z\sigma^2\mathbf{T}_1\,\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1^T-\tilde{t}_1(z)\mathbf{I}_N)\,\mathbf{T}+\frac{\sigma^2 c_N}{1+\sigma^2 c_N t_1(z)}\mathbf{T}_1\,\mathbf{A}\left(\mathbf{I}_N+\sigma^2 c_N\mathcal{T}_{N,L}^{(M)}(\mathbf{T}^T)\right)^{-1}\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1^T-t_1(z)\mathbf{I}_{ML})\,\mathbf{A}^*\mathbf{T} \tag{4.3.7}$$

- $\tilde{\mathbf{T}}-\tilde{\mathbf{T}}_1 = z\sigma^2 c_N\tilde{\mathbf{T}}_1\,\mathcal{T}_{N,L}^{(M)}(\mathbf{T}^T-\mathbf{T}_1^T)\,\tilde{\mathbf{T}}+\frac{\sigma^2}{1+\sigma^2\tilde{t}_1(z)}\tilde{\mathbf{T}}_1\mathbf{A}^*\left(\mathbf{I}_{ML}+\sigma^2\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}^T)\right)^{-1}\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}^T-\tilde{\mathbf{T}}_1^T)\mathbf{A}\tilde{\mathbf{T}}$

$$z\sigma^2 c_N\tilde{\mathbf{T}}_1\,\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1^T-t_1\mathbf{I}_{ML})\,\tilde{\mathbf{T}}+\frac{\sigma^2}{1+\sigma^2\tilde{t}_1(z)}\tilde{\mathbf{T}}_1\mathbf{A}^*\left(\mathbf{I}_{ML}+\sigma^2\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}^T)\right)^{-1}\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1^T-\tilde{t}_1(z)\mathbf{I}_N)\mathbf{A}\tilde{\mathbf{T}} \tag{4.3.8}$$

To prove Proposition 4.3.1, we first prove that

$$\|\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1-t_1\mathbf{I}_{ML})\|\le C(z)\sqrt{\frac{\log(L)}{N^{1-\alpha}}} \tag{4.3.9}$$

$$\|\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1-\tilde{t}_1\mathbf{I}_N)\|\le C(z)\sqrt{\frac{\log(N)}{N^{1-\alpha}}} \tag{4.3.10}$$

for each $z\in\mathbf{C}^+$.
First step: Proof of (4.3.9)
We first express $\mathbf{T}_1-t_1(z)\mathbf{I}_{ML}$. Recall that $\mathbf{T}_1=\left(-z(1+\sigma^2\tilde{t}_1(z))\mathbf{I}_{ML}+\frac{\mathbf{AA}^*}{1+\sigma^2 c_N t_1(z)}\right)^{-1}$. We notice that

$$\begin{aligned}\mathbf{T}_1-\frac{-1}{z(1+\sigma^2\tilde{t}_1(z))}\mathbf{I}_{ML} &= \mathbf{T}_1\left(-z(1+\sigma^2\tilde{t}_1(z)-\mathbf{T}_1^{-1}\right)\frac{-1}{z(1+\sigma^2\tilde{t}_1(z))}\\ &= \mathbf{T}_1\mathbf{AA}^*\frac{1}{z(1+\sigma^2\tilde{t}_1(z))(1+\sigma^2 c_N t_1(z))}\end{aligned} \tag{4.3.11}$$

Taking the normalized trace of both side, and noticing that $t_1(z)=\frac{1}{ML}\mathrm{Tr}(\mathbf{T}_1)$ we get that

$$t_1(z)=\frac{-1}{z(1+\sigma^2\tilde{t}_1(z))}+\frac{1}{ML}\mathrm{Tr}(\mathbf{T}_1\mathbf{AA}^*)\frac{1}{z(1+\sigma^2\tilde{t}_1(z))(1+\sigma^2 c_N t_1(z))}$$

Thus, we can express $\mathbf{T}_1-t_1(z)\mathbf{I}_{ML}$ as

$$\mathbf{T}_1-t_1(z)\mathbf{I}_{ML}=\mathbf{T}_1-\frac{-1}{z(1+\sigma^2\tilde{t}(z))}-\frac{1}{ML}\mathrm{Tr}(\mathbf{T}_1\mathbf{AA}^*)\frac{1}{z(1+\sigma^2\tilde{t}_1(z))(1+\sigma^2 c_N t_1(z))} \tag{4.3.12}$$

Plugging (4.3.11) into (4.3.12), we obtain

$$\mathbf{T}_1-t_1(z)\mathbf{I}_{ML}=\frac{1}{z(1+\sigma^2\tilde{t}_1(z))(1+\sigma^2 c_N t_1(z))}(\mathbf{T}_1\mathbf{AA}^*-\frac{1}{ML}\mathrm{Tr}(\mathbf{T}_1\mathbf{AA}^*)\mathbf{I}_{ML}) \tag{4.3.13}$$

Since the Toeplitzification operator $\mathcal{T}_{N,L}^{(M)}(.)$ is linear and $\mathcal{T}_{N,L}^{(M)}(\mathbf{I}_{ML})=\mathbf{I}_N$, we obtain that

$$\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1-t_1\mathbf{I}_{ML})=\frac{1}{z(1+\sigma^2\tilde{t}_1(z))(1+\sigma^2 c_N t_1(z))}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1\mathbf{AA}^*)-\frac{1}{ML}\mathrm{Tr}(\mathbf{T}_1\mathbf{AA}^*)\mathbf{I}_N\right) \tag{4.3.14}$$

We recall (see paragraph 1.2.3)) that $\mathbf{T}_1$ belongs to $\mathcal{S}_{ML}(\mathbb{R}^+)$ coincides with the Stieltjes transform of a positive matrix-valued measure $\boldsymbol{\mu}_N$ carried by $\mathbb{R}^+$ such that $\boldsymbol{\mu}_N(\mathbb{R}^+)=\mathbf{I}_{ML}$ i.e

$$\mathbf{T}_1=\int_{\mathbb{R}^+}\frac{d\boldsymbol{\mu}(\lambda)}{\lambda-z}$$

Therefore, it is clear that $t_1(z)$ and $\tilde{t}_1(z)$ are Stieltjes transforms of the probability measures $\mu_N = \frac{1}{ML}\boldsymbol{\mu}_N$ and $c_N\mu_N + (1-c_N)\delta_0$, so that we have $|t_1(z)|$ and $|\tilde{t}_1(z)|$ are bounded by $\frac{1}{\mathrm{Im}(z)}$.
From section 3.1.3, in chapter 3, it is easy to check that $\frac{1}{z(1+\sigma^2 c_N t(z))}$ and $\frac{1}{z(1+\sigma^2\tilde{t}_1(z))}$ are also Stieltjes transforms of probability measure carried by $\mathbb{R}^+$, which particularly implies that $\left|\frac{1}{z(1+\sigma^2 c_N t(z))}\right|$ and $\left|\frac{1}{z(1+\sigma^2\tilde{t}_1(z))}\right|$ are upper-bounded by $\frac{1}{\mathrm{Im}(z)}$.
Therefore, taking the spectral norm of both side of (4.3.14), we get that

$$\begin{aligned}\left\|\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1 - t_1\mathbf{I}_{ML})\right\| &\le \left|\frac{1}{z(1+\sigma^2\tilde{t}_1(z))(1+\sigma^2 c_N t_1(z))}\right|\left(\left\|\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1\mathbf{A}\mathbf{A}^*)\right\| + \left|\frac{1}{ML}\mathrm{Tr}(\mathbf{T}_1\mathbf{A}\mathbf{A}^*)\right|\right)\\ &\le \frac{|z|}{(\mathrm{Im}(z))^2}\left(\left\|\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1\mathbf{A}\mathbf{A}^*)\right\| + \left|\frac{1}{ML}\mathrm{Tr}(\mathbf{T}_1\mathbf{A}\mathbf{A}^*)\right|\right)\end{aligned} \tag{4.3.15}$$

Since $\|\mathbf{T}_1\| \le \frac{1}{\mathrm{Im}(z)}$ and the rank of $\mathbf{A}$ is the same as that of $\boldsymbol{\mathcal{H}}^{(L)}$, i.e $P+L-1$, it holds that

$$\frac{1}{ML}\mathrm{Tr}\left[\mathbf{T}_1\mathbf{A}\mathbf{A}^*\right] \le \frac{C(z)}{M}$$

Thus, (4.3.15) becomes

$$\left\|\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1 - t_1\mathbf{I}_{ML})\right\| \le \frac{|z|}{(\mathrm{Im}(z))^2}\left\|\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1\mathbf{A}\mathbf{A}^*)\right\| + \frac{C(z)}{M}$$

It remains to prove that $\left\|\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1\mathbf{A}\mathbf{A}^*)\right\| = \mathcal{O}\left(\sqrt{\frac{\log(L)}{N^{1-\alpha}}}\right)$. For this, we first recall from (3.1.19) that

$$\|\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1\mathbf{A}\mathbf{A}^*)\| \le \sup_{\nu\in[0,1]}\left|\frac{1}{M}\sum_{m=1}^{M}(\mathbf{f}^m\otimes\mathbf{a}_L(\nu))^*\,\mathbf{T}_1\mathbf{A}\mathbf{A}^*(\mathbf{f}^m\otimes\mathbf{a}_L(\nu))\right|$$

with $(\mathbf{f}^1,\dots,\mathbf{f}^M)$ the canonical base of $\mathbb{C}^M$. By Cauchy-Schwarz inequality, we have that

$$|(\mathbf{f}^m\otimes\mathbf{a}_L(\nu))^*\,\mathbf{T}_1\mathbf{A}\mathbf{A}^*(\mathbf{f}^m\otimes\mathbf{a}_L(\nu))| \le \left((\mathbf{f}^m\otimes\mathbf{a}_L(\nu))^*\,\mathbf{T}_1\mathbf{T}_1^*(\mathbf{f}^m\otimes\mathbf{a}_L(\nu))\right)^{1/2}\left((\mathbf{f}^m\otimes\mathbf{a}_L(\nu))^*\,(\mathbf{A}\mathbf{A}^*)^2(\mathbf{f}^m\otimes\mathbf{a}_L(\nu))\right)^{1/2}$$

As $\|\mathbf{T}_1\| \le \frac{1}{\mathrm{Im}(z)}$, the following inequality holds

$$\sup_{\nu\in[0,1]}\left((\mathbf{f}^m\otimes\mathbf{a}_L(\nu))^*\,\mathbf{T}_1\mathbf{T}_1^*(\mathbf{f}^m\otimes\mathbf{a}_L(\nu))\right)^{1/2} \le \|\mathbf{T}_1\| \le \frac{1}{\mathrm{Im}(z)}$$

Moreover, as we assumed that $\sup_N\|\mathbf{A}\| \le C < +\infty$, we obtain that

$$\begin{aligned}\left((\mathbf{f}^m\otimes\mathbf{a}_L(\nu))^*\,(\mathbf{A}\mathbf{A}^*)^2(\mathbf{f}^m\otimes\mathbf{a}_L(\nu))\right)^{1/2} &\le \|\mathbf{A}\|\left((\mathbf{f}^m\otimes\mathbf{a}_L(\nu))^*\,\mathbf{A}\mathbf{A}^*(\mathbf{f}^m\otimes\mathbf{a}_L(\nu))\right)^{1/2}\\ &\le C\left((\mathbf{f}^m\otimes\mathbf{a}_L(\nu))^*\,\mathbf{A}\mathbf{A}^*(\mathbf{f}^m\otimes\mathbf{a}_L(\nu))\right)^{1/2}\end{aligned}$$

Thus, the following inequality is obtained

$$\|\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1\mathbf{A}\mathbf{A}^*)\| \le \frac{C}{\mathrm{Im}(z)}\sup_{\nu\in[0,1]}\frac{1}{M}\sum_{m=1}^{M}\left((\mathbf{f}^m\otimes\mathbf{a}_L(\nu))^*\,\mathbf{A}\mathbf{A}^*(\mathbf{f}^m\otimes\mathbf{a}_L(\nu))\right)^{1/2}$$

On the other hand, we recall that $\mathbf{A}\mathbf{A}^* = \boldsymbol{\mathcal{H}}^{(L)}\frac{\mathbf{S}^{(L)}\mathbf{S}^{(L)*}}{N}\boldsymbol{\mathcal{H}}^{(L)*}$. Therefore,

$$\mathbf{A}\mathbf{A}^* = \boldsymbol{\mathcal{H}}^{(L)}\left(\frac{\mathbf{S}^{(L)}\mathbf{S}^{(L)*}}{N} - \mathbf{R}_S\right)\boldsymbol{\mathcal{H}}^{(L)*} + \boldsymbol{\mathcal{H}}^{(L)}\mathbf{R}_S\boldsymbol{\mathcal{H}}^{(L)*}$$

In this way we obtain

$$
\begin{aligned}
(\mathbf{f}^m \otimes \mathbf{a}_L(\nu))^* \mathbf{A}\mathbf{A}^* (\mathbf{f}^m \otimes \mathbf{a}_L(\nu)) &= (\mathbf{f}^m \otimes \mathbf{a}_L(\nu))^* \, \mathcal{H}^{(L)} \left(\frac{\mathbf{S}^{(L)}\mathbf{S}^{(L)*}}{N} - \mathbf{R}_S \right) \mathcal{H}^{(L)*} (\mathbf{f}^m \otimes \mathbf{a}_L(\nu)) \\
&\quad + (\mathbf{f}^m \otimes \mathbf{a}_L(\nu))^* \, \mathcal{H}^{(L)} \mathbf{R}_S \mathcal{H}^{(L)*} (\mathbf{f}^m \otimes \mathbf{a}_L(\nu)) \\
&\le \|\mathcal{H}^{(L)}\|^2 \left\| \frac{\mathbf{S}^{(L)}\mathbf{S}^{(L)*}}{N} - \mathbf{R}_S \right\| + (\mathbf{f}^m \otimes \mathbf{a}_L(\nu))^* \, \mathcal{H}^{(L)} \mathbf{R}_S \mathcal{H}^{(L)*} (\mathbf{f}^m \otimes \mathbf{a}_L(\nu))
\end{aligned}
$$

By lemma 4.3.1, we have that $\left\| \frac{\mathbf{S}^{(L)}\mathbf{S}^{(L)*}}{N} - \mathbf{R}_S \right\| = \mathcal{O}\left(\sqrt{\frac{\log(L)}{N^{1-\alpha}}} \right)$. Moreover, $\|\mathcal{H}^{(L)}\|$ is uniformly bounded by Remark 11. Hence, $\|\mathcal{H}^{(L)}\|^2 \left\| \frac{\mathbf{S}^{(L)}\mathbf{S}^{(L)*}}{N} - \mathbf{R}_S \right\| = \mathcal{O}\left(\sqrt{\frac{\log(L)}{N^{1-\alpha}}} \right)$. Therefore, in order to prove $\|\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1\mathbf{A}\mathbf{A}^*)\| = \mathcal{O}\left(\sqrt{\frac{\log(L)}{N^{1-\alpha}}} \right)$, we will prove that $\sup_{\nu\in[0,1]} \frac{1}{M} \sum_{m=1}^{M} (\mathbf{f}^m \otimes \mathbf{a}_L(\nu))^* \, \mathcal{H}^{(L)} \mathbf{R}_S \mathcal{H}^{(L)*} (\mathbf{f}^m \otimes \mathbf{a}_L(\nu)) = \mathcal{O}\left(\frac{1}{M}\right) = \mathcal{O}\left(\frac{1}{N^{1-\alpha}}\right)$.
We notice that

$$
\mathcal{H}^{(L)} \mathbf{R}_S \mathcal{H}^{(L)*} = \begin{pmatrix} \mathcal{H}_1^{(L)} \mathbf{R}_S \mathcal{H}_1^{(L)*} & \cdots & \mathcal{H}_1^{(L)} \mathbf{R}_S \mathcal{H}_M^{(L)} \\ \vdots & & \vdots \\ \mathcal{H}_M^{(L)} \mathbf{R}_S \mathcal{H}_1^{(L)*} & \cdots & \mathcal{H}_M^{(L)} \mathbf{R}_S \mathcal{H}_M^{(L)} \end{pmatrix}
$$

It is easy to see that

$$
\frac{1}{M} \sum_{m=1}^{M} (\mathbf{f}^m \otimes \mathbf{a}_L(\nu))^* \, \mathcal{H}^{(L)} \mathbf{R}_S \mathcal{H}^{(L)*} (\mathbf{f}^m \otimes \mathbf{a}_L(\nu)) = \mathbf{a}_L(\nu)^* \left(\frac{1}{M} \sum_{m=1}^{M} \mathcal{H}_m^{(L)} \mathbf{R}_S \mathcal{H}_m^{(L)*} \right) \mathbf{a}_L(\nu)
$$

In order to calculate $\tau(\mathcal{H}_m^{(L)} \mathbf{R}_S \mathcal{H}_m^{(L)*})(l)$, we express $(\mathcal{H}_m^{(L)} \mathbf{R}_S \mathcal{H}_m^{(L)*})_{j,k}$, for $1 \le j, k \le L$ and $1 \le m \le M$.
We can notice that

$$
(\mathcal{H}_m^{(L)} \mathbf{R}_S \mathcal{H}_m^{(L)*})_{j,k} = \mathbf{e}_j^T \mathcal{H}_m^{(L)} \mathbf{R}_S \mathcal{H}_m^{(L)*} \mathbf{e}_k
$$

with $(\mathbf{e}_1, ..., \mathbf{e}_L)$ the canonical base of $\mathbb{C}^L$.
Recall that

$$
\mathbf{R}_S = \int_{[0,1]} \phi(\theta) \mathbf{d}_{P+L-1}(\theta) \mathbf{d}_{P+L-1}^*(\theta) \, d\theta
$$

It is clear that

- $\mathbf{e}_j^T \mathcal{H}_m^{(L)} = (\underbrace{0, ..., 0}_{j-1}, \underbrace{h_{m,P-1}, ..., h_{m,0}}_{P}, \underbrace{0, ..., 0}_{L-j})$
- $\mathcal{H}_m^{(L)*} \mathbf{e}_k = (\underbrace{0, ..., 0}_{k-1}, \underbrace{\overline{h_{m,P-1}}, ..., \overline{h_{m,0}}}_{P}, \underbrace{0, ..., 0}_{L-k})^T$

We eventually get

$$
\begin{aligned}
\mathbf{e}_j^T \mathcal{H}_m^{(L)} \mathbf{R}_S \mathcal{H}_m^{(L)*} \mathbf{e}_k &= \int_{[0,1]} \phi(\theta) \left(\sum_{t=0}^{P-1} e^{2i\pi(t+j)\theta} h_{m,P-1-t} \right) \left(\sum_{u=0}^{P-1} e^{-2i\pi(u+k)\theta} \overline{h_{m,P-1-u}} \right) d\theta \\
&= \int_{[0,1]} \phi(\theta) e^{2i\pi(j-k)\theta} \left(\sum_{t=0}^{P-1} e^{2i\pi t\theta} h_{m,P-1-t} \right) \left(\sum_{u=0}^{P-1} e^{-2i\pi u\theta} \overline{h_{m,P-1-u}} \right) d\theta \\
&= \int_{[0,1]} \phi(\theta) |h^{(m)}(\theta)|^2 e^{2i\pi(j-k)\theta} d\theta
\end{aligned}
$$

where we recall that $h^{(m)}(\theta) = \sum_{u=0}^{P-1} h_{m,u} e^{-2i\pi\theta l}$.
It yields immediately that

$$\mathcal{H}_m^{(L)} \mathbf{R}_S \mathcal{H}_m^{(L)*} = \int_{[0,1]} \phi(\theta) |h^{(m)}(\theta)|^2 \mathbf{d}_L(\theta) \mathbf{d}_L^*(\theta) d\theta$$

As $\mathbf{a}_L^*(\nu)\mathbf{d}_L(\theta) = \frac{1}{\sqrt{L}} \sum_{l=0}^{L} e^{2i\pi(\theta-\nu)l}$, we get that

$$\mathbf{a}_L(\nu)^* \mathcal{H}_m^{(L)} \mathbf{R}_S \mathcal{H}_m^{(L)*} \mathbf{a}_L(\nu) = \int_{[0,1]} |h^{(m)}(\theta)|^2 \phi(\theta) \frac{1}{L} \left| \sum_{l=0}^{L-1} e^{2i\pi(\theta-\nu)l} \right|^2 d\theta$$

Finally, we have that

$$\frac{1}{M} \sum_{m=1}^{M} \mathbf{a}_L(\nu)^* \mathcal{H}_m^{(L)} \mathbf{R}_S \mathcal{H}_m^{(L)*} \mathbf{a}_L(\nu) = \frac{1}{M} \int_{[0,1]} \left(\sum_{m=1}^{M} |h^{(m)}(\theta)|^2 \right) \phi(\theta) \frac{1}{L} \left| \sum_{l=0}^{L-1} e^{-i2\pi(\theta-\nu)l} \right|^2 d\theta \qquad (4.3.16)$$

Assumption A-5 give us that $\sup_M \sup_\theta \sum_{m=1}^{M} |h^{(m)}(\theta)|^2 < +\infty$ and $\phi(\theta)$ is the spectral density which is finite positive and bounded, thus there exists a constant C such that

$$\sup_M \sup_\theta \phi(\theta) \left| \sum_{m=1}^{M} h^{(m)}(\theta) \right|^2 \leq C$$

As a consequence, it yields

$$\frac{1}{M} \sum_{m=1}^{M} \mathbf{a}_L(\nu)^* \mathcal{H}_m^{(L)} \mathbf{R}_S \mathcal{H}_m^{(L)*} \mathbf{a}_L(\nu) \leq \frac{C}{M} \int \frac{1}{L} \left| \sum_{l=0}^{L-1} e^{-i2\pi(\theta-\nu)l} \right|^2 d\theta$$

Moreover, Parseval identity leads to

$$\frac{1}{L} \int_{[0,1]} \left| \sum_{l=0}^{L-1} e^{-i2\pi(\theta-\nu)l} \right|^2 d\theta = \frac{1}{L} L = 1$$

By (4.3.16), it follows

$$\left| \frac{1}{M} \sum_{m=1}^{M} \mathbf{a}_L(\nu)^* \mathcal{H}_m^{(L)} \mathbf{R}_S \mathcal{H}_m^{(L)*} \mathbf{a}_L(\nu) \right| \leq \frac{C}{M}$$

for all $\nu \in [0,1]$, and C is a constant independent of N, M and L. We conclude that

$$\left\| \mathcal{T}_{N,L}^{(M)} \left(\mathcal{H}^{(L)} \mathbf{R}_S \mathcal{H}^{(L)*} \right) \right\| \leq \frac{C}{M},$$

from which (4.3.9) follows immediately. □

Second step: Proof of (4.3.10)

We recall that $\tilde{\mathbf{T}}_1 = \left(-z(1+\sigma^2 c_N t_1(z)) \mathbf{I}_N + \frac{\mathbf{A}^*\mathbf{A}}{1+\sigma^2 \tilde{t}_1(z)} \right)^{-1}$ and $\tilde{t}_1(z) = \frac{1}{N} \mathrm{Tr}(\tilde{\mathbf{T}}_1)$.
Using the same approach as in (4.3.11), we obtain that

$$\tilde{t}(z) = \frac{-1}{z(1+\sigma^2 c_N t_1(z))} + \frac{1}{N} \mathrm{Tr}(\tilde{\mathbf{T}}_1 \mathbf{A}^* \mathbf{A}) \frac{1}{z(1+\sigma^2 \tilde{t}_1(z))(1+\sigma^2 c_N t_1(z))}$$

It follows

$$\tilde{\mathbf{T}}_1 - \tilde{t}(z) \mathbf{I}_N = \frac{1}{z(1+\sigma^2 \tilde{t}_1(z))(1+\sigma^2 c_N t_1(z))} \left(\tilde{\mathbf{T}}_1 \mathbf{A}^* \mathbf{A} - \frac{1}{N} \mathrm{Tr}(\tilde{\mathbf{T}}_1 \mathbf{A}^* \mathbf{A}) \mathbf{I}_N \right)$$

since $\mathcal{T}_{L,L}(.)$ is linear, we get that

$$\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1 - \tilde{t}_1\mathbf{I}_N) = \frac{1}{z(1+\sigma^2\tilde{t}_1(z))(1+\sigma^2 c_N t_1(z))}\left(\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1\mathbf{A}^*\mathbf{A}) - \frac{1}{N}\mathrm{Tr}\,(\tilde{\mathbf{T}}_1\mathbf{A}^*\mathbf{A})\mathbf{I}_L\right)$$

taking the spectral norm, we get

$$\|\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1 - \tilde{t}_1\mathbf{I}_N)\| \leq \frac{1}{|z(1+\sigma^2\tilde{t}_1(z))(1+\sigma^2 c_N t_1(z))|}\left(\left\|\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1\mathbf{A}^*\mathbf{A})\right\| + \left|\frac{1}{N}\mathrm{Tr}\,(\tilde{\mathbf{T}}_1\mathbf{A}^*\mathbf{A})\right|\right)$$

We recall that $\|\tilde{\mathbf{T}}_1\| \leq \frac{1}{\mathrm{Im}(z)}$ and we remark that the rank of matrix $\mathbf{A}$ is $P+L-1$ and P and L have the same order of magnitude. Hence

$$\left|\frac{1}{N}\mathrm{Tr}\,(\tilde{\mathbf{T}}_1\mathbf{A}^*\mathbf{A})\right| \leq C(z)\frac{L}{N} = \mathcal{O}\left(\frac{1}{M}\right) = \mathcal{O}\left(\frac{1}{N^{1-\alpha}}\right)$$

It remains to prove that

$$\left\|\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1\mathbf{A}^*\mathbf{A})\right\| = \mathcal{O}\left(\sqrt{\frac{\log(N)}{N^{1-\alpha}}}\right)$$

By (3.1.17), we have that

$$\|\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1\mathbf{A}^*\mathbf{A})\| \leq \sup_{\nu\in[0,1]} |\mathbf{a}_N^*(\nu)\tilde{\mathbf{T}}_1\mathbf{A}^*\mathbf{A}\mathbf{a}_N(\nu)| \tag{4.3.17}$$

By Cauchy-Schwarz inequality, for all $\nu \in [0,1]$, it yields

$$|\mathbf{a}_N^*(\nu)\tilde{\mathbf{T}}_1\mathbf{A}^*\mathbf{A}\mathbf{a}_N(\nu)| \leq (\mathbf{a}_N^*(\nu)\tilde{\mathbf{T}}_1\tilde{\mathbf{T}}^*\mathbf{a}_N(\nu))^{1/2}(\mathbf{a}_N^*(\nu)(\mathbf{A}^*\mathbf{A})^2\mathbf{a}_N(\nu))^{1/2}$$

As $\|\mathbf{T}_1\| \leq \frac{1}{\mathrm{Im}(z)}$, $\sup_N \|\mathbf{A}\| \leq C$, and recall that $\mathbf{A}^*\mathbf{A} = \frac{\mathbf{S}^{(L)*}}{\sqrt{N}}\boldsymbol{\mathcal{H}}^{(L)*}\boldsymbol{\mathcal{H}}^{(L)}\frac{\mathbf{S}^{(L)}}{\sqrt{N}}$, we obtain

$$\begin{aligned}
|\mathbf{a}_N^*(\nu)\tilde{\mathbf{T}}_1\mathbf{A}^*\mathbf{A}\mathbf{a}_N(\nu)| &\leq \frac{C}{\mathrm{Im}(z)}\left(\mathbf{a}_N^*(\nu)\mathbf{A}^*\mathbf{A}\mathbf{a}_N(\nu)\right)^{1/2} \\
&\leq \frac{C}{\mathrm{Im}(z)}\left(\mathbf{a}_N^*(\nu)\frac{\mathbf{S}^{(L)*}}{\sqrt{N}}\boldsymbol{\mathcal{H}}^{(L)*}\boldsymbol{\mathcal{H}}^{(L)}\frac{\mathbf{S}^{(L)}}{\sqrt{N}}\mathbf{a}_N(\nu)\right)^{1/2} \\
&\leq \frac{C\|\boldsymbol{\mathcal{H}}^{(L)}\|}{\mathrm{Im}(z)}\left(\mathbf{a}_N^*(\nu)\frac{\mathbf{S}^{(L)*}\mathbf{S}^{(L)}}{N}\mathbf{a}_N(\nu)\right)^{1/2} \\
&\leq C(z)\left(\mathbf{a}_N^*(\nu)\frac{\mathbf{S}^{(L)*}\mathbf{S}^{(L)}}{N}\mathbf{a}_N(\nu)\right)^{1/2}
\end{aligned} \tag{4.3.18}$$

(4.3.18) is obtained by recalling that $\sup_N \|\boldsymbol{\mathcal{H}}^{(L)}\| < +\infty$, thus it exists a constant $C_0 > 0$, such that $\sup_N \|\boldsymbol{\mathcal{H}}^{(L)}\| \leq C_0$.
Combining (4.3.17) and (4.3.18), we get

$$\|\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1\mathbf{A}^*\mathbf{A})\| \leq C(z)\sup_{\nu\in[0,1]}\left(\mathbf{a}_N^*(\nu)\frac{\mathbf{S}^{(L)*}\mathbf{S}^{(L)}}{N}\mathbf{a}_N(\nu)\right)^{1/2} \tag{4.3.19}$$

It remains to prove that $\sup_{\nu\in[0,1]}\mathbf{a}_N^*(\nu)\frac{\mathbf{S}^{(L)*}\mathbf{S}^{(L)}}{N}\mathbf{a}_N(\nu) = \mathcal{O}\left(\frac{\log N}{M}\right)$. For this we can notice that

$$\frac{\mathbf{S}}{\sqrt{N}}\mathbf{a}_N(\nu) = \frac{1}{N}\begin{pmatrix} s_{2-P} & \cdots & s_{N-P+1} \\ \vdots & & \vdots \\ s_L & \cdots & s_{N+L-1}\end{pmatrix}\begin{pmatrix} 1 \\ \vdots \\ e^{2i\pi(N-1)\nu}\end{pmatrix} = \frac{1}{N}\begin{pmatrix} \sum_{k=1}^N s_{1-P+k}e^{2i\pi k\nu} \\ \vdots \\ \sum_{k=1}^N s_{L+k}e^{2i\pi k\nu}\end{pmatrix}$$

and that

$$\left\| \frac{\mathbf{S}}{\sqrt{N}} \mathbf{a}_N(\nu) \right\|^2 = \frac{1}{N} \sum_{l=1}^{P+L-1} \left| \frac{1}{\sqrt{N}} \sum_{k=1}^{N} s_{-P+l+k} e^{2i\pi k\nu} \right|^2 \tag{4.3.20}$$

We notice that for $1 \leq l \leq P{+}L{-}1$, $\left| \frac{1}{\sqrt{N}} \sum_{k=1}^{N} s_{-P+l+k} \right|^2$ is the periodogram of the sequence $(s_{-P+l+1}, s_{-P+l+2}, \ldots, s_{-P+l+N})$ that we denote by $I_N^{(l)}(\nu)$. And we use the following result due to [3]

Remark 12. *Under Assumption A-6, it holds*

$$\lim_{N\to\infty} \max_{\nu\in[0,1]} \frac{I_N(\nu)}{\phi(\nu)\log N} = 1, \ \ a.s \tag{4.3.21}$$

with $I_N(\nu) = \frac{1}{N} \left| \sum_{n=1}^{N} s_n e^{-2i\pi\nu n} \right|^2$ *the periodogram of sequence* $(s_1, \ldots, s_N)$ *and* $\phi(\nu) = |\sum_{j=0}^{+\infty} \alpha_j e^{-2i\pi\nu j}|^2$ *the spectral density of* s.

By (4.3.21),for $1 \leq l \leq P + L - 1$, $\forall \epsilon > 0$, almost surely, for N large enough, we have that

$$\max_{\nu\in[0,1]} \frac{I_N^{(l)}(\nu)}{\phi(\nu)\log(N)} \leq 1 + \epsilon$$

Almost surely, for N large enough, we thus have

$$\forall \nu \in [0,1], \ \ I_N^{(l)}(\nu) \leq (1+\epsilon)\phi(\nu)\log(N)$$

Applying this inequality to (4.3.20), almost surely, for N large enough,

$$\begin{aligned} \forall \nu \in [0,1] \quad \left\| \frac{\mathbf{S}}{\sqrt{N}} \mathbf{a}_N(\nu) \right\|^2 &\leq \frac{1}{N} \sum_{l=1}^{P+L-1} (1+\epsilon)\phi(\nu)\log(N) \\ &\leq (1+\epsilon)\phi(\nu) \frac{P+L-1}{N} \log(N) \end{aligned}$$

Since P, L are $\mathcal{O}(N^{\alpha})$, with $0 < \alpha < 1$, $\sup_{\nu\in[0,1]} \left\| \frac{\mathbf{S}}{\sqrt{N}} \mathbf{a}_N(\nu) \right\|^2 = \mathcal{O}\left(\frac{\log(N)}{N^{1-\alpha}} \right)$. This and (4.3.19) give us that

$$\|\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1 \mathbf{A}^* \mathbf{A})\| = \mathcal{O}\left(\sqrt{\frac{\log(N)}{N^{1-\alpha}}} \right)$$

This completes the proof of (4.3.10). □

Now, let $(\mathbf{b}_1, \mathbf{b}_2)$ two unitary ML–dimensional vectors, and $(\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2)$ two unitary N–dimensional vectors. Recall the expressions of (4.3.7) and (4.3.8), we take the bilinear forms

$$\begin{aligned} \mathbf{b}_1(\mathbf{T} - \mathbf{T}_1)\mathbf{b}_2 &= z\sigma^2 \mathbf{b}_1 \mathbf{T}_1 \mathbf{I}_M \otimes \mathcal{T}_{L,L}(\tilde{\mathbf{T}}^T - \tilde{\mathbf{T}}_1^T) \mathbf{T} \mathbf{b}_2 \\ &+ \frac{\sigma^2 c_N}{1+\sigma^2 c_N t_1(z)} \mathbf{b}_1 \mathbf{T}_1 \mathbf{A} \left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbf{T}^T) \right)^{-1} \mathcal{T}_{N,L}^{(M)}(\mathbf{T}^T - \mathbf{T}_1^T) \mathbf{A}^* \mathbf{T} \mathbf{b}_2 \\ &+ z\sigma^2 \mathbf{b}_1 \mathbf{T}_1 \mathbf{I}_M \otimes \mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1^T - \tilde{t}_1(z) \mathbf{I}_N) \mathbf{T} \mathbf{b}_2 \\ &+ \frac{\sigma^2 c_N}{1+\sigma^2 c_N t_1(z)} \mathbf{b}_1 \mathbf{T}_1 \mathbf{A} \left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbf{T}^T) \right)^{-1} \mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1^T - t_1(z) \mathbf{I}_{ML}) \mathbf{A}^* \mathbf{T} \mathbf{b}_2 \end{aligned}$$

$$\begin{aligned}\tilde{\mathbf{b}}_1(\tilde{\mathbf{T}}-\tilde{\mathbf{T}}_1)\tilde{\mathbf{b}}_2 &= z\sigma^2 c_N \tilde{\mathbf{b}}_1 \tilde{\mathbf{T}}_1 \mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T - \mathbf{T}_1^T)\tilde{\mathbf{T}}\tilde{\mathbf{b}}_2 \\ &+ \frac{\sigma^2}{1+\sigma^2\tilde{t}_1(z)}\tilde{\mathbf{b}}_1\tilde{\mathbf{T}}_1\mathbf{A}^*\left(\mathbf{I}_{ML}+\sigma^2\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}^T)\right)^{-1}\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}^T-\tilde{\mathbf{T}}_1^T)\mathbf{A}\tilde{\mathbf{T}}\tilde{\mathbf{b}}_2 \\ &+ z\sigma^2 c_N\tilde{\mathbf{b}}_1\tilde{\mathbf{T}}_1\mathcal{T}^{(M)}_{N,L}(\mathbf{T}_1^T - t_1\mathbf{I}_{ML})\tilde{\mathbf{T}}\tilde{\mathbf{b}}_2 \\ &+ \frac{\sigma^2}{1+\sigma^2\tilde{t}_1(z)}\tilde{\mathbf{b}}_1\tilde{\mathbf{T}}_1\mathbf{A}^*\left(\mathbf{I}_{ML}+\sigma^2\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}^T)\right)^{-1}\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1^T-\tilde{t}_1(z)\mathbf{I}_N)\mathbf{A}\tilde{\mathbf{T}}\tilde{\mathbf{b}}_2\end{aligned}$$

It is clear that

$$\begin{aligned}|\mathbf{b}_1(\mathbf{T}-\mathbf{T}_1)\mathbf{b}_2| &\le |z|\sigma^2\|\mathbf{b}_1\|\,\|\mathbf{T}_1\|\,\|\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}^T-\tilde{\mathbf{T}}_1^T)\|\,\|\mathbf{T}\|\,\|\mathbf{b}_2\| \\ &+ \frac{\sigma^2 c_N}{|1+\sigma^2 c_N t_1(z)|}\|\mathbf{b}_1\|\,\|\mathbf{T}_1\|\,\|\mathbf{A}\|\left\|\left(\mathbf{I}_N+\sigma^2 c_N\mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T)\right)^{-1}\right\|\,\|\mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T-\mathbf{T}_1^T)\|\,\|\mathbf{A}^*\|\,\|\mathbf{T}\|\,\|\mathbf{b}_2\| \\ &+ |z|\sigma^2\|\mathbf{b}_1\|\,\|\mathbf{T}_1\|\,\|\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1^T-\tilde{t}_1(z)\mathbf{I}_N)\|\,\|\mathbf{T}\|\,\|\mathbf{b}_2\| \\ &+ \frac{\sigma^2 c_N}{1+\sigma^2 c_N t_1(z)}\|\mathbf{b}_1\|\,\|\mathbf{T}_1\|\,\|\mathbf{A}\|\left\|\left(\mathbf{I}_N+\sigma^2 c_N\mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T)\right)^{-1}\right\|\,\|\mathcal{T}^{(M)}_{N,L}(\mathbf{T}_1^T-t_1(z)\mathbf{I}_{ML})\|\,\|\mathbf{A}^*\|\,\|\mathbf{T}\|\,\|\mathbf{b}_2\|\end{aligned}$$

Since $\mathbf{b}_1, \mathbf{b}_2$ are unitary vectors, $\|\mathbf{A}\| \le \kappa$, $\|\mathbf{T}\| \le \frac{1}{\mathrm{Im}(z)}$, $\|\mathbf{T}_1\| \le \frac{1}{\mathrm{Im}(z)}$, $\left\|\left(\mathbf{I}_N+\sigma^2 c_N\mathcal{T}^{(M)}_{N,L}(\mathbf{T}^T)\right)^{-1}\right\| \le \frac{|z|}{\mathrm{Im}(z)}$, $\frac{1}{|1+\sigma^2 c_N t_1(z)|} \le \frac{1}{\mathrm{Im}(z)}$, $\|\mathcal{T}^{(M)}_{N,L}(\mathbf{T}_1 - t_1\mathbf{I}_{ML})\| = \mathcal{O}\left(\sqrt{\frac{\log(L)}{N^{1-\alpha}}}\right)$, and $\|\mathcal{T}_{L,L}(\tilde{\mathbf{T}}_1-\tilde{t}_1\mathbf{I}_N)\| = \mathcal{O}\left(\frac{\log(N)}{N^{1-\alpha}}\right)$, it yields

$$\begin{aligned}|\mathbf{b}_1(\mathbf{T}-\mathbf{T}_1)\mathbf{b}_2| &\le \frac{|z|\sigma^2}{(\mathrm{Im}(z)^2}\|\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\tilde{\mathbf{T}}-\tilde{\mathbf{T}}_1)\| \\ &+ \frac{\sigma^2 c_N\kappa^2|z|}{(\mathrm{Im}(z))^4}\|\mathcal{T}^{(M)}_{N,L}(\mathbf{T}-\mathbf{T}_1)\| + \mathcal{O}\left(\sqrt{\frac{\log(L)}{N^{1-\alpha}}}\right)\end{aligned}$$

Recall that $\|\mathcal{T}_{L,L}(\tilde{\mathbf{T}}^T-\tilde{\mathbf{T}}_1^T)\| \le \sup_{\|\tilde{\mathbf{b}}_1\|,\|\tilde{\mathbf{b}}_2\|=1}|\tilde{\mathbf{b}}_1(\tilde{\mathbf{T}}-\tilde{\mathbf{T}}_1)\tilde{\mathbf{b}}_2|$ and $\|\mathcal{T}^{(M)}_{N,L}(\mathbf{T}-\mathbf{T}_1)\| \le \sup_{\|\mathbf{b}_1\|,\|\mathbf{b}_2\|=1}|\mathbf{b}_1(\mathbf{T}-\mathbf{T}_1)\mathbf{b}_2|$, we can simply get that

$$\begin{aligned}\sup_{\|\mathbf{b}_1\|,\|\mathbf{b}_2\|=1}|\mathbf{b}_1(\mathbf{T}-\mathbf{T}_1)\mathbf{b}_2| &\le \frac{|z|\sigma^2}{(\mathrm{Im}(z))^2}\sup_{\|\tilde{\mathbf{b}}_1\|,\|\tilde{\mathbf{b}}_2\|=1}|\tilde{\mathbf{b}}_1(\tilde{\mathbf{T}}-\tilde{\mathbf{T}}_1)\tilde{\mathbf{b}}_2| \\ &+ \frac{\sigma^2 c_N\kappa^2|z|}{(\mathrm{Im}(z))^4}\sup_{\|\mathbf{b}_1\|,\|\mathbf{b}_2\|=1}|\mathbf{b}_1(\mathbf{T}-\mathbf{T}_1)\mathbf{b}_2| + \mathcal{O}\left(\sqrt{\frac{\log(L)}{N^{1-\alpha}}}\right)\end{aligned} \tag{4.3.22}$$

Similar calculations give us that

$$\begin{aligned}\sup_{\|\tilde{\mathbf{b}}_1\|,\|\tilde{\mathbf{b}}_2\|=1}|\tilde{\mathbf{b}}_1(\tilde{\mathbf{T}}-\tilde{\mathbf{T}}_1)\tilde{\mathbf{b}}_2| &\le \frac{|z|\sigma^2 c_N}{(\mathrm{Im}(z))^2}\sup_{\|\mathbf{b}_1\|,\|\mathbf{b}_2\|=1}|\mathbf{b}_1(\mathbf{T}-\mathbf{T}_1)\mathbf{b}_2| \\ &+ \frac{\sigma^2\kappa^2|z|}{(\mathrm{Im}(z))^4}\sup_{\|\tilde{\mathbf{b}}_1\|,\|\tilde{\mathbf{b}}_2\|=1}|\tilde{\mathbf{b}}_1(\tilde{\mathbf{T}}-\tilde{\mathbf{T}}_1)\tilde{\mathbf{b}}_2| + \mathcal{O}\left(\sqrt{\frac{\log(L)}{N^{1-\alpha}}}\right)\end{aligned} \tag{4.3.23}$$

Solving the system of linear inequalities (4.3.22), (4.3.23), we obtain that for $z \in \mathbb{C}^+$ well chosen $\|\mathbf{T}(z) - \mathbf{T}_1(z)\| = \sup_{\|\mathbf{b}_1\|,\|\mathbf{b}_2\|=1}|\mathbf{b}_1^*(\mathbf{T}-\mathbf{T}_1)\mathbf{b}_2|$ and $\|\tilde{\mathbf{T}}(z)-\tilde{\mathbf{T}}_1(z)\| = \sup_{\|\tilde{\mathbf{b}}_1\|,\|\tilde{\mathbf{b}}_2\|=1}|\tilde{\mathbf{b}}_1^*(\tilde{\mathbf{T}}-\tilde{\mathbf{T}}_1)\tilde{\mathbf{b}}_2|$ are both $\sqrt{\frac{\log(N)}{N^{1-\alpha}}}$ terms. Using the argument "à la Montel" used in the course of the proof of Proposition 3.9.1, we obtain the convergence towards 0 of $\|\mathbf{T}(z)-\mathbf{T}_1(z)\|$ and of $\|\tilde{\mathbf{T}}(z)-\tilde{\mathbf{T}}_1(z)\|$ for each $z \in \mathbb{C}\setminus\mathbb{R}^+$.

We finally mention that $\mathbf{T}_1$ and $\tilde{\mathbf{T}}_1$ behave as $\mathbf{S}$ and $\tilde{\mathbf{S}}$ defined by (4.2.3), and that this implies that $\|\mathbf{T}-\mathbf{S}\|$ and $\|\tilde{\mathbf{T}}-\tilde{\mathbf{S}}\|$ converge towards 0. More precisely, the following Corollary holds.

Corollary 4.3.1. *For each $z \in \mathbb{C} \setminus \mathbb{R}^+$, it holds that*

$$\|\mathbf{T}_1(z) - \mathbf{S}(z)\| \rightarrow 0 \quad (4.3.24)$$
$$\|\tilde{\mathbf{T}}_1(z) - \tilde{\mathbf{S}}(z)\| \rightarrow 0 \quad (4.3.25)$$

and that

$$\|\mathbf{T}(z) - \mathbf{S}(z)\| \rightarrow 0 \quad (4.3.26)$$
$$\|\tilde{\mathbf{T}}(z) - \tilde{\mathbf{S}}(z)\| \rightarrow 0 \quad (4.3.27)$$

The proof of (4.3.24, 4.3.25) is based on the observation that $t_1(z) - m_c(z) \rightarrow 0$ and that $t_1(z) - m_c(z) \rightarrow 0$, while (4.3.26, 4.3.27) follows directly from (4.3.6).

□

4.4 Application to regularized estimated spatial-temporal Wiener filters in large system case

In this section, we revisit Section 2.4 in the case where P and L may converge towards $+\infty$. We recall that $(s_n)_{n\in\mathbb{Z}}$ represents an i.i.d. sequence available from $n = 1$ to $n = \mathrm{N}$. We study the performance of the estimated regularized Wiener filter $\hat{\mathbf{g}}_\lambda$ defined by

$$\hat{\mathbf{g}}_\lambda = \left(\frac{1}{\mathrm{N}}\sum_{n=1}^{\mathrm{N}} \mathbf{y}_n^{(\mathrm{L})}\mathbf{y}_n^{(\mathrm{L})*} + \lambda\mathbf{I}_{\mathrm{ML}}\right)^{-1}\left(\frac{1}{\mathrm{N}}\sum_{n=1}^{\mathrm{N}} \mathbf{y}_n^{(\mathrm{L})} s_n^*\right) \quad (4.4.1)$$

$\hat{\mathbf{g}}_\lambda$ is destined to estimate the unknown transmitted data $(s_n)_{n>\mathrm{N}}$. We recall that the signal to interference plus noise ratio produced by $\hat{\mathbf{g}}_\lambda^{(\mathrm{L})}$ is equal to

$$\mathsf{SINR}(\hat{\mathbf{g}}_\lambda) = \frac{|\hat{\mathbf{g}}_\lambda^* \mathbf{h}_{\mathrm{P}}^{(\mathrm{L})}|^2}{\hat{\mathbf{g}}_\lambda^* \boldsymbol{\mathcal{H}}_{-\mathrm{P}}^{(\mathrm{L})}\boldsymbol{\mathcal{H}}_{-\mathrm{P}}^{(\mathrm{L})*}\hat{\mathbf{g}}_\lambda + \sigma^2\|\hat{\mathbf{g}}_\lambda\|^2} \quad (4.4.2)$$

where $\mathbf{h}_{\mathrm{P}}^{(\mathrm{L})}$ is column P of $\boldsymbol{\mathcal{H}}^{(\mathrm{L})}$, and matrix $\boldsymbol{\mathcal{H}}_{-\mathrm{P}}^{(\mathrm{L})}$ is obtained by deleting column P from matrix $\boldsymbol{\mathcal{H}}^{(\mathrm{L})}$. $\mathsf{SINR}(\hat{\mathbf{g}}_\lambda^{(\mathrm{L})})$ is random in the sense that it depends on the vectors $(\mathbf{y}_n)_{n=1,\dots,\mathrm{N}}$, which are random themselves due to the presence of the additive noise. In the following, we generalize the results of Section 2.4 and establish that, provided $\frac{\mathrm{L}}{\mathrm{M}} \rightarrow 0$, $\mathsf{SINR}(\hat{\mathbf{g}}_\lambda)$ has exactly the same asymptotic behaviour than in the case where P and L do not scale with N, i.e. $\mathsf{SINR}(\hat{\mathbf{g}}_\lambda)$ converges almost surely towards term $\phi(\lambda)$ defined by (2.4.25). The reader may check that in the present context, the estimate $\hat{\phi}(\lambda)$ of $\phi(\lambda)$ introduced in Section 2.4 is still consistent in the context of the present chapter. From now on, we assume that

Assumption A-7: ML *and* N *converge towards* $+\infty$ *in such a way that* $c_{\mathrm{N}} = \frac{\mathrm{ML}}{\mathrm{N}}$ *satisfies* $c_{\mathrm{N}} \rightarrow c_*$ *where* $0 < c_* < +\infty$. P, L *are* $\mathcal{O}(\mathrm{N}^\alpha)$ *with* $0 < \alpha < \frac{1}{2}$, *and that* $\mathrm{M} = \mathcal{O}\left(\mathrm{N}^{1-\alpha}\right)$.

and that

Assumption A-8: *Sequence s is an independent identically distributed sequence verifying Assumption A-6*

To simplify the notations, $\mathrm{N} \rightarrow +\infty$ should be understood as the above asymptotic regime.
As in Section 2.4, we denote by $\mathbf{U}_\mathrm{N}$ the matrix $(\mathrm{P}+\mathrm{L}-1) \times \mathrm{N}$ defined by $\mathbf{U}_\mathrm{N} = \frac{1}{\sqrt{\mathrm{N}}}\mathbf{S}_\mathrm{N}^{(\mathrm{L})}$. We also denote by $\Sigma_\mathrm{N}^{(\mathrm{L})}$ and $\mathbf{W}_\mathrm{N}^{(\mathrm{L})}$ the normalized observed and noise matrices $\mathbf{Y}_\mathrm{N}^{(\mathrm{L})}/\sqrt{\mathrm{N}}$ and $\mathbf{V}_\mathrm{N}^{(\mathrm{L})}/\sqrt{\mathrm{N}}$. In Section 2.4, it was

assumed without restriction that $\mathbf{U}_N$ verified $\mathbf{U}_N\mathbf{U}_N^* = \mathbf{I}$ because replacing $\mathcal{H}$ by $\mathcal{H}\left(\frac{1}{\sqrt{N}}\mathbf{S}_N^{(L)}\right)^{1/2}$ and $\mathbf{U}_N$ by $\left(\frac{1}{\sqrt{N}}\mathbf{S}_N^{(L)}\right)^{-1/2}\mathbf{U}_N$ had no impact. In the present context, this is however apparently not true because replacing $\mathcal{H}$ and $\mathbf{U}_N$ by $\mathcal{H}\left(\frac{1}{\sqrt{N}}\mathbf{S}_N^{(L)}\right)^{1/2}$ and $\left(\frac{1}{\sqrt{N}}\mathbf{S}_N^{(L)}\right)^{-1/2}\mathbf{U}_N$ breaks the Hankel structures of the 2 matrices. The Hankel structures were important to establish that $\|\mathbf{T}(z)-\mathbf{T}_1(z)\|$ and $\|\tilde{\mathbf{T}}(z)-\tilde{\mathbf{T}}_1(z)\|$ converge towards 0. However, as sequence s verifies Assumption A-8, matrix $\mathbf{U}_N\mathbf{U}_N^*$ verifies

$$\|\mathbf{U}_N\mathbf{U}_N^* - \mathbf{I}_{P+L-1}\| \to 0$$

and it is easy to check that the replacement of $\mathcal{H}$ by $\mathcal{H}\left(\frac{1}{\sqrt{N}}\mathbf{S}_N^{(L)}\right)^{1/2}$ and of $\mathbf{U}_N$ by $\left(\frac{1}{\sqrt{N}}\mathbf{S}_N^{(L)}\right)^{-1/2}\mathbf{U}_N$ has no impact on the proof of Proposition 4.3.1. We therefore assume without restriction that $\mathbf{U}_N\mathbf{U}_N^* = \mathbf{I}$ as in Section 2.4.

In the following, we define $\mathbf{Q}(z)$ as the resolvent of matrix $\mathbf{\Sigma\Sigma}^*$ defined by $\mathbf{Q}(z) = (\mathbf{\Sigma\Sigma}^* - z\mathbf{I}_{ML})^{-1}$, and recall that the estimated Wiener filter $\hat{\mathbf{g}}_\lambda$ can be written as

$$\hat{\mathbf{g}}_\lambda^{(L)} = \mathbf{Q}(-\lambda)\mathbf{\Sigma}\mathbf{u}^*$$

where $\mathbf{u} = \frac{1}{\sqrt{N}}(s_1,\ldots,s_N)$ is the P-th row of matrix $\mathbf{U}$. To evaluate the behaviour of the SINR given by formula (4.4.2) when $N \to +\infty$, it is necessary to study $|\mathbf{h}_P^*\hat{\mathbf{g}}_\lambda|^2$,$\|\mathbf{H}_1^*\hat{\mathbf{g}}_\lambda\|^2$, and $\|\hat{\mathbf{g}}_\lambda\|^2$. These terms depend on bilinear forms of matrices $\mathbf{Q}(-\lambda)$ and $\mathbf{Q}(-\lambda)^2$ whose asymptotic behaviour have thus to be evaluated. In the following, we prove that the above 3 terms behave as in the case where P and L do not scale with N, thus showing that the SINR has also the same behaviour.

4.5 Asymptotic behaviour of the SINR

In this section, we study the asymptotic behaviour of the SINR and evaluate $|\hat{\mathbf{g}}_\lambda^*\mathbf{h}_P^{(L)}|^2$, $\hat{\mathbf{g}}_\lambda^*\mathcal{H}_{-P}^{(L)}\mathcal{H}_{-P}^{(L)*}\hat{\mathbf{g}}_\lambda$ and $\|\hat{\mathbf{g}}_\lambda\|^2$.

4.5.1 Evaluation of $\hat{\mathbf{g}}_\lambda^*\mathbf{h}_P^{(L)}$

This subsection is devoted to the evaluation of $\hat{\mathbf{g}}_\lambda^*\mathbf{h}_P^{(L)}$, or more generally, of $\mathbf{b}^*\hat{\mathbf{g}}_\lambda = \mathbf{b}^*\mathbf{Q}(-\lambda)\mathbf{\Sigma}\mathbf{u}^*$, where $\mathbf{b}$ is a deterministic ML–dimensional vector such that $\sup_N \|\mathbf{b}\| \le \kappa$.
Now let $z \in \mathbb{C}\setminus\mathbb{R}^+$, and consider $\mathbf{b}^*\mathbf{Q}(z)\mathbf{\Sigma}\mathbf{u}^*$. To short the notation, we omit the variable z in the calculations of $\mathbf{Q}$. Remark 8 implies that

$$\mathbf{b}^*[\mathbf{Q\Sigma} - \mathbb{E}(\mathbf{Q\Sigma})]\mathbf{u}^* \to 0$$

almost surely. It is thus sufficient to evaluate $\mathbb{E}(\mathbf{Q\Sigma})$

Expression of $\mathbb{E}(\mathbf{Q\Sigma})$

For this, we express $\mathbb{E}(\mathbf{Q\Sigma})$ as $\mathbb{E}(\mathbf{Q\Sigma}) = \mathbb{E}(\mathbf{QW}) + \mathbb{E}(\mathbf{QA})$, and we calculate $\mathbb{E}(\mathbf{QW})$ using the integration by parts formula. We recall the notation $(\mathbf{QW})_{i,j}^m$, the entry $(i+(m-1)L, n)$ of $ML\times N$ matrix $\mathbf{QW}$ for $1 \le m \le M$, $1 \le i \le L$ and $1 \le j \le N$ and use the identity

$$\mathbb{E}\left[(\mathbf{QW})_{i,j}^m\right] = \sum_{m_1,i_1}\mathbb{E}\left(\mathbf{Q}_{i,i_1}^{m,m_1}\mathbf{W}_{i_1,j}^{m_1}\right)$$

By the integration by parts formula, it yields

$$\mathbb{E}\left(\mathbf{Q}_{i,i_1}^{m,m_1}\mathbf{W}_{i_1,j}^{m_1}\right) = \sum_{i',j'} \mathbb{E}\left(\mathbf{W}_{i_1,j}^{m_1}\overline{\mathbf{W}}_{i',j'}^{m_1}\right)\mathbb{E}\left(\frac{\partial \mathbf{Q}_{i,i_1}^{m,m_1}}{\partial \overline{\mathbf{W}}_{i',j'}^{m_1}}\right)$$

It is easy to check that

$$\frac{\partial \mathbf{Q}_{i,i_1}^{m,m_1}}{\partial \overline{\mathbf{W}}_{i',j'}^{m_1}} = -(\mathbf{Q}\mathbf{\Sigma})_{i,j'}^{m}\mathbf{Q}_{i',i_1}^{m_1,m_1}$$

(3.3.1) implies that $\mathbb{E}\left(\mathbf{W}_{i_1,j}^{m_1}\overline{\mathbf{W}}_{i',j'}^{m_1}\right) = \frac{\sigma^2}{N}\delta(i_1 - i' = j' - j)$. Therefore, we obtain that

$$\mathbb{E}\left(\mathbf{Q}_{i,i_1}^{m,m_1}\mathbf{W}_{i_1,j}^{m_1}\right) = -\frac{\sigma^2}{N}\sum_{i',j'}\delta(i_1 - i' = j' - j)\mathbb{E}\left[(\mathbf{Q}\mathbf{\Sigma})_{i,j'}^{m}\mathbf{Q}_{i',i_1}^{m_1,m_1}\right]$$

We put $u = i_1 - i'$ in the above sum, and get that

$$\mathbb{E}\left(\mathbf{Q}_{i,i_1}^{m,m_1}\mathbf{W}_{i_1,j}^{m_1}\right) = -\frac{\sigma^2}{N}\sum_{u=-(L-1)}^{L-1}\mathbb{E}\left((\mathbf{Q}\mathbf{W})_{i,u+j}^{m}\mathbf{Q}_{i_1-u,i_1}^{m_1,m_1}\right)\mathbb{1}_{1\le u+j\le N}\mathbb{1}_{1\le i_1-u\le L}$$

and that

$$\begin{aligned}\mathbb{E}\left[(\mathbf{Q}\mathbf{W})_{i,j}^{m}\right] &= \sum_{m_1,i_1}\mathbb{E}\left[\mathbf{Q}_{i,i_1}^{m,m_1}\mathbf{W}_{i_1,j}^{m_1}\right]\\ &= -\frac{\sigma^2}{N}\sum_{u=-(L-1)}^{L-1}\mathbb{E}\left((\mathbf{Q}\mathbf{W})_{i,u+j}^{m}\mathbb{1}_{1\le u+j\le N}\sum_{m_1,i_1}\mathbf{Q}_{i_1-u,i_1}^{m_1,m_1}\mathbb{1}_{1\le i_1-u\le L}\right)\end{aligned}$$

Using the definition $\tau^{(M)}(\mathbf{Q})(-u) = \frac{1}{ML}\sum_{m_1,i_1}\mathbf{Q}_{i_1-u,i_1}^{m_1,m_1}\mathbb{1}_{1\le i_1-u\le L}$, we get that

$$\mathbb{E}\left[(\mathbf{Q}\mathbf{W})_{i,j}^{m}\right] = -\sigma^2\frac{ML}{N}\sum_{u=-(L-1)}^{L-1}\mathbb{E}\left((\mathbf{Q}\mathbf{W})_{i,u+j}^{m}\mathbb{1}_{1\le u+j\le N}\tau^{(M)}(\mathbf{Q})(-u)\right)$$

Setting $k = u + j$, the righthandside of the above equation can also be written as

$$-\sigma^2 c_N\sum_{k=1}^{N}\mathbb{E}\left((\mathbf{Q}\mathbf{\Sigma})_{i,k}^{m}\tau^{(M)}(\mathbf{Q})(-k+j)\mathbb{1}_{-(L-1)\le k-j\le L-1}\right)$$

or, using the observation that $\tau^{(M)}(\mathbf{Q})(j-k)\mathbb{1}_{-(L-1)\le j-k\le L-1} = \left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q})\right)_{j,k}$ (see Eq. (3.1.5)), as

$$-\sigma^2 c_N\sum_{k=1}^{N}\mathbb{E}\left((\mathbf{Q}\mathbf{\Sigma})_{i,k}^{m}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^T)\right)_{k,j}\right)$$

Therefore,

$$\begin{aligned}\mathbb{E}\left[(\mathbf{Q}\mathbf{W})_{i,j}^{m}\right] &= -\sigma^2 c_N\sum_{k=1}^{N}\mathbb{E}\left((\mathbf{Q}\mathbf{\Sigma})_{i,k}^{m}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^T)\right)_{k,j}\right)\\ &= -\sigma^2 c_N\mathbb{E}\left[\left(\mathbf{Q}\mathbf{\Sigma}\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^T)\right)_{i,j}^{m}\right]\end{aligned}$$

It is clear that

$$\mathbb{E}[\mathbf{Q}\mathbf{W}] = -\sigma^2 c_N\mathbb{E}\left[\mathbf{Q}\mathbf{\Sigma}\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^T)\right]$$

Thus, we obtain that

$$\begin{aligned}\mathbb{E}[\mathbf{Q}\boldsymbol{\Sigma}] &= \mathbb{E}[\mathbf{Q}\mathbf{W}] + \mathbb{E}[\mathbf{Q}\mathbf{A}] \\ &= -\sigma^2 c_N \mathbb{E}\left[\mathbf{Q}\boldsymbol{\Sigma}\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^T)\right] + \mathbb{E}[\mathbf{Q}]\mathbf{A}\end{aligned}$$

We express matrix $\mathbf{Q}$ as $\mathbf{Q} = \mathbb{E}(\mathbf{Q}) + \mathbf{Q}^\circ$, and obtain that

$$\mathbb{E}[\mathbf{Q}\boldsymbol{\Sigma}] = \mathbb{E}(\mathbf{Q})\mathbf{A} - \sigma^2 c_N \mathbb{E}[\mathbf{Q}\boldsymbol{\Sigma}]\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T)) - \sigma^2 c_N \mathbb{E}\left[\mathbf{Q}\boldsymbol{\Sigma}\mathcal{T}_{N,L}^{(M)}\left((\mathbf{Q}^\circ)^T\right)\right]$$

Hence, we get that

$$\mathbb{E}[\mathbf{Q}\boldsymbol{\Sigma}]\left[\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\right] = \mathbb{E}(\mathbf{Q})\mathbf{A} - \mathbb{E}\left[\mathbf{Q}\boldsymbol{\Sigma}\mathcal{T}_{N,L}^{(M)}\left((\mathbf{Q}^\circ)^T\right)\right]$$

The invertibility of $\left[\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\right]$ is proven in section 3.A, and we recall that its inverse is denoted by $\mathbf{H}$. We obtain that

$$\mathbb{E}[\mathbf{Q}\boldsymbol{\Sigma}] = \mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H} + \mathbb{E}\left[\mathbf{Q}\boldsymbol{\Sigma}\mathcal{T}_{N,L}^{(M)}\left((\mathbf{Q}^\circ)^T\right)\right]\mathbf{H} \tag{4.5.1}$$

Convergence of bilinear form $\mathbf{b}^*\mathbb{E}(\mathbf{Q}\boldsymbol{\Sigma})\mathbf{u}^*$

In this subsection, we prove that the bilinear form $\mathbf{b}^*\mathbb{E}(\mathbf{Q}\boldsymbol{\Sigma})\mathbf{u}^*$ converges towards a deterministic quantity. First of all, we prove the lemma

Lemma 4.5.1. *Under assumption A-7, it holds that*

$$\mathbf{b}^*\left[\mathbb{E}[\mathbf{Q}\boldsymbol{\Sigma}] - \mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}\right]\mathbf{u}^* \to 0$$

when $N \to \infty$.

Proof of lemma (4.5.1): Using the equation (4.5.1) and the definition $\mathcal{T}_{N,L}^{(M)}\left((\mathbf{Q}^\circ)^T\right) = \sum_{l=-(L-1)}^{L-1} \tau^{(M)}(\mathbf{Q}^\circ)(l)\mathbf{J}^l$, we can obtain that:

$$\begin{aligned}\mathbf{b}^*\left[\mathbb{E}[\mathbf{Q}\boldsymbol{\Sigma}] - \mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}\right]\mathbf{u}^* &= \mathbf{b}^*\left(\mathbb{E}\left[\mathbf{Q}\boldsymbol{\Sigma}\mathcal{T}_{N,L}^{(M)}\left((\mathbf{Q}^\circ)^T\right)\right]\mathbf{H}\right)\mathbf{u}^* \\ &= \mathrm{Tr}\left(\mathbb{E}\left[\mathbf{Q}\boldsymbol{\Sigma}\sum_{l=-(L-1)}^{L-1}\tau^{(M)}(\mathbf{Q}^\circ)(l)\mathbf{J}^l\right]\mathbf{H}\mathbf{u}^*\mathbf{b}^*\right) \\ &= \mathbb{E}\left[\sum_{l=-(L-1)}^{L-1}\tau^{(M)}(\mathbf{Q}^\circ)(l)\mathbf{b}^*\mathbf{Q}\boldsymbol{\Sigma}\mathbf{J}^l\mathbf{H}\mathbf{u}^*\right]\end{aligned}$$

(3.4.1,3.4.7) imply that $\mathbb{E}|\tau^{(M)}(\mathbf{Q}^\circ)(l)|^2$ and $\mathrm{Var}\left(\mathbf{b}^*\mathbf{Q}\boldsymbol{\Sigma}\mathbf{J}^l\mathbf{H}\mathbf{u}^*\right)$ are upperbounded by terms of the form $\frac{C(z)}{MN}$ and $\kappa^2 C(z)\frac{L}{N}$ respectively. The Cauchy-Schwartz inequality thus implies that

$$|\mathbf{b}^*\left[\mathbb{E}[\mathbf{Q}\boldsymbol{\Sigma}] - \mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}\right]\mathbf{u}^*| \le \kappa C(z) L\sqrt{\frac{L}{MN^2}} = \mathcal{O}\left(\sqrt{\frac{L}{M}\frac{L}{N}}\right) = \mathcal{O}\left(\sqrt{\frac{L}{M^3}}\right)$$

As $\frac{L}{M^2} \to 0$, the lemma is thus proved. □

We now prove that

Proposition 4.5.1.

$$\mathbf{b}^*\left[\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H} - \frac{1}{1+\sigma^2 c_N t_1(z)}\mathbf{T}_1\mathbf{A}\right]\mathbf{u}^* \to 0 \tag{4.5.2}$$

with $\mathbf{T}_1 = \left(-z(1+\sigma^2\tilde{t}_1(z))\mathbf{I}_{ML} + \frac{\mathbf{A}\mathbf{A}^*}{1+\sigma^2 c_N t_1(z)}\right)^{-1}$, $\tilde{\mathbf{T}}_1 = \left(-z(1+\sigma^2 c_N t_1(z))\mathbf{I}_N + \frac{\mathbf{A}^*\mathbf{A}}{1+\sigma^2\tilde{t}_1(z)}\right)^{-1}$, $t_1(z) = \frac{1}{ML}\mathrm{Tr}(\mathbf{T}_1)$ *and* $\tilde{t}_1(z) = \frac{1}{N}\mathrm{Tr}(\tilde{\mathbf{T}}_1)$.

Proof. We recall that $\mathbf{H} = \left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}))\right)^{-1}$. It is clear that

$$\begin{aligned}
&\mathbf{b}^* \left[\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H} - \frac{1}{1+\sigma^2 c_N t_1(z)}\mathbf{T}_1\mathbf{A}\right]\mathbf{u}^* \\
&= \mathbf{b}^*(\mathbb{E}(\mathbf{Q}) - \mathbf{T}_1)\mathbf{A}\,\mathbf{H}\,\mathbf{u}^* + \mathbf{b}^*\mathbf{T}_1\mathbf{A}\left[\mathbf{H} - \frac{1}{1+\sigma^2 c_N t_1(z)}\mathbf{I}_N\right]\mathbf{u}^* \\
&= \mathbf{b}^*(\mathbb{E}(\mathbf{Q}) - \mathbf{T}_1)\mathbf{A}\,\mathbf{H}\,\mathbf{u}^* + \mathbf{b}^*\mathbf{T}_1\mathbf{A}\,\mathbf{H}\left[(1+\sigma^2 c_N t_1(z))\mathbf{I}_N - \mathbf{H}^{-1}\right]\frac{1}{1+\sigma^2 c_N t_1(z)}\mathbf{u}^* \\
&= \mathbf{b}^*(\mathbb{E}(\mathbf{Q}) - \mathbf{T}_1)\mathbf{A}\,\mathbf{H}\,\mathbf{u}^* + \mathbf{b}^*\mathbf{T}_1\mathbf{A}\,\mathbf{H}\left[(1+\sigma^2 c_N t_1(z))\mathbf{I}_N - \left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}))\right)\right]\frac{1}{1+\sigma^2 c_N t_1(z)}\mathbf{u}^* \\
&= \mathbf{b}^*(\mathbb{E}(\mathbf{Q}) - \mathbf{T}_1)\mathbf{A}\,\mathbf{H}\,\mathbf{u}^* \\
&+ \frac{\sigma^2 c_N}{1+\sigma^2 c_N t_1(z)}\mathbf{b}^*\mathbf{T}_1\mathbf{A}\,\mathbf{H}\left[t_1(z)\mathbf{I}_N - \mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1^T) + \mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1^T) - \mathcal{T}_{N,L}^{(M)}\left(\mathbb{E}(\mathbf{Q}^T)\right)\right]\mathbf{u}^* \\
&= \mathbf{b}^*(\mathbb{E}(\mathbf{Q}) - \mathbf{T}_1)\mathbf{A}\mathbf{H}\mathbf{u}^* + \frac{\sigma^2 c_N}{1+\sigma^2 c_N t_1(z)}\mathbf{b}^*\mathbf{T}_1\mathbf{A}\,\mathbf{H}\left(\mathcal{T}_{N,L}^{(M)}\left(t_1(z)\mathbf{I}_{ML} - \mathbf{T}_1^T\right)\right)\mathbf{u}^* \\
&+ \frac{\sigma^2 c_N}{1+\sigma^2 c_N t_1(z)}\mathbf{b}^*\mathbf{T}_1\mathbf{A}\,\mathbf{H}\left(\mathcal{T}_{N,L}^{(M)}\left(\mathbf{T}_1^T - \mathbb{E}(\mathbf{Q}^T)\right)\right)\mathbf{u}^*
\end{aligned} \tag{4.5.3}$$

Since $\sup_N \|\mathbf{A}\| < +\infty$, $\|\mathbf{H}\| \le \frac{|z|}{\text{Im}(z)}$, Corollary 3.9.1, Proposition 4.3.1 and ((4.3.9)) imply (4.5.2). □

Therefore, we have the following result.

Proposition 4.5.2. *Under Assumption A-7, it holds that*

$$\mathbf{b}^* \left[\mathbf{Q}\mathbf{\Sigma} - \frac{1}{1+\sigma^2 c_N t_1(z)}\mathbf{T}_1\mathbf{A}\right]\mathbf{u}^* \to 0,\ a.s \tag{4.5.4}$$

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

In order to connect this with the corresponding result in Section 2.4 (see Eq. (2.4.16)), we remark that $t_1(z)$ and $\mathbf{T}_1(z)$ can be replaced in (4.5.4) by $m_c(z)$ and $\mathbf{S}(z)$ where $m_c(z)$ is the Stieltjes of the Marcenko-Pastur distribution of parameters (c, σ^2) and where $\mathbf{S}(z)$ is given by

$$\mathbf{S}(z) = \left(-z(1+\sigma^2 \tilde{m}_c(z)) + \frac{\mathcal{H}\mathcal{H}^*}{1+\sigma^2 c m_c(z)}\right)^{-1}$$

(see Corollary 4.3.1). Moreover, $\mathbf{A} = \mathcal{H}\mathbf{U}$ and $\mathbf{A}\mathbf{u}^* = \mathcal{H}\mathbf{U}\mathbf{u}^*$. As $\mathbf{U}$ is assumed orthogonal, $\mathbf{U}\mathbf{u}^* = \mathbf{e}_P$ where $\mathbf{e}_P$ is the P-th vector of the canonical basis of $\mathbb{C}^{P+L-1}$. Therefore, (4.5.4) is equivalent to

$$\mathbf{b}^*\mathbf{Q}\mathbf{\Sigma}\mathbf{u}^* - \frac{1}{1+\sigma^2 c m_c(z)}\mathbf{b}^*\mathbf{S}\,\mathbf{h}_P \to 0,\ a.s \tag{4.5.5}$$

for each $z \in \mathbb{C} \setminus \mathbb{R}^+$, which is precisely the evaluation that follows from (2.4.16). in section 2.4.

4.5.2 Evaluation of $\|\hat{\mathbf{g}}_\lambda\|^2$

In this subsection, we study the behaviour of $\|\hat{\mathbf{g}}_\lambda\|^2$, when N converges towards ∞. In order to do this, we express $\|\hat{\mathbf{g}}_\lambda\|^2 = \mathbf{u}\mathbf{\Sigma}^*\mathbf{Q}^2(-\lambda)\mathbf{\Sigma}\mathbf{u}^*$. As we can notice that $\frac{d}{dz}\mathbf{Q}(z) = \mathbf{Q}^2(z)$, we can obtain

$$\begin{aligned}
\mathbf{u}\mathbf{\Sigma}^*\mathbf{Q}^2(z)\mathbf{\Sigma}\mathbf{u}^* &= \frac{d}{dz}\left(\mathbf{u}\mathbf{\Sigma}^*\mathbf{Q}(z)\mathbf{\Sigma}\mathbf{u}^*\right) \\
&= \mathbf{u}\frac{d}{dz}\left(\mathbf{\Sigma}^*\mathbf{Q}(z)\mathbf{\Sigma}\right)\mathbf{u}^*
\end{aligned}$$

We recall that $\mathbf{\Sigma}^*\mathbf{Q}(z)\mathbf{\Sigma} = \tilde{\mathbf{Q}}(z)\mathbf{\Sigma}^*\mathbf{\Sigma} = \mathbf{I}_N + z\tilde{\mathbf{Q}}(z)$. It follows that

$$\begin{aligned}\mathbf{u}\mathbf{\Sigma}^*\mathbf{Q}^2(z)\mathbf{\Sigma}\mathbf{u}^* &= \mathbf{u}\frac{d}{dz}\left(\mathbf{\Sigma}^*\mathbf{Q}(z)\mathbf{\Sigma}\right)\mathbf{u}^* \\ &= \mathbf{u}\frac{d}{dz}\left(\mathbf{I}_N + z\tilde{\mathbf{Q}}(z)\right)\mathbf{u}^* \\ &= \mathbf{u}\frac{d}{dz}\left(z\tilde{\mathbf{Q}}(z)\right)\mathbf{u}^* \\ &= \frac{d}{dz}\left(z\mathbf{u}\tilde{\mathbf{Q}}(z)\mathbf{u}^*\right)\end{aligned}$$

We have thus to evaluate the behaviour of $z\mathbf{u}\tilde{\mathbf{Q}}(z)\mathbf{u}^*$ for $z \in \mathbb{C}\setminus\mathbb{R}^+$ and to differentiate w.r.t z to obtain the behaviour of $\|\hat{\mathbf{g}}_\lambda\|^2$. As $\mathbf{u}\tilde{\mathbf{Q}}(z)\mathbf{u}^* - \mathbf{u}\mathbb{E}(\tilde{\mathbf{Q}}(z))\mathbf{u}^* \to 0$ almost surely for each $z \in \mathbb{C}\setminus\mathbb{R}^+$, we have just to study $\mathbb{E}(\mathbf{u}\tilde{\mathbf{Q}}(z)\mathbf{u}^*)$. Again, the result we obtain is the same as in Section 2.4. In particular, we have the following result.

Proposition 4.5.3. *For all $z \in \mathbb{C}\setminus\mathbb{R}^+$, it holds that*

$$\mathbf{u}\frac{d}{dz}\left[z(\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{T}}_1)\right]\mathbf{u}^* \to 0 \tag{4.5.6}$$

when $N \to \infty$.

Proof. We express $\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{T}}_1 = \mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{T}} + \tilde{\mathbf{T}} - \tilde{\mathbf{T}}_1$. (3.9.4) and (4.3.6) imply that $\mathbf{u}(\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{T}})\mathbf{u}^*$ and $\mathbf{u}(\tilde{\mathbf{T}} - \tilde{\mathbf{T}}_1)\mathbf{u}^*$ both converge towards 0, when $N \to \infty$, for all $z \in \mathbb{C}\setminus\mathbb{R}^+$. Hence, it holds that

$$\mathbf{u}(\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{T}}_1)\mathbf{u}^* \to 0$$

for all $\mathbf{z} \in \mathbb{C}^+$. (4.5.6) is obtained by noticing that the convergence of $z\mathbf{u}(\mathbb{E}(\tilde{\mathbf{Q}}) - \tilde{\mathbf{T}}_1)\mathbf{u}^*$ towards 0 is uniform on each compact subset of $\mathbb{C}\setminus\mathbb{R}^+$. □

In order to retrieve the result of Section 2.4, we remark that, by Corollary 4.3.1, $\tilde{\mathbf{T}}_1(z)$ can be replaced in (4.5.6) by $\tilde{\mathbf{S}}(z)$ given by

$$\tilde{\mathbf{S}}(z) = \left(-z(1+\sigma^2 c m_c(z)) + \frac{\mathbf{U}^*\mathcal{H}^*\mathcal{H}\mathbf{U}}{1+\sigma^2\tilde{m}_c(z)}\right)^{-1}$$

4.5.3 Evaluation of $\hat{\mathbf{g}}_\lambda^*\mathcal{H}_{-P}^{(L)}\mathcal{H}_{-P}^{(L)*}\hat{\mathbf{g}}_\lambda$

This subsection is devoted to study the behaviour of the quantity $\hat{\mathbf{g}}_\lambda^*\mathcal{H}_{-P}^{(L)}\mathcal{H}_{-P}^{(L)*}\hat{\mathbf{g}}_\lambda$ when N grows large. First of all, we express

$$\hat{\mathbf{g}}_\lambda^*\mathcal{H}_{-P}^{(L)}\mathcal{H}_{-P}^{(L)*}\hat{\mathbf{g}}_\lambda = \mathbf{u}\mathbf{\Sigma}^*\mathbf{Q}\mathcal{H}_{-P}^{(L)}\mathcal{H}_{-P}^{(L)*}\mathbf{Q}\mathbf{\Sigma}\mathbf{u}^*$$

We denote $\mathcal{D} = \mathcal{H}_{-P}^{(L)}\mathcal{H}_{-P}^{(L)*}$. By Poincaré-Nash inequality, we can prove easily that $\mathbf{u}(\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{\Sigma} - \mathbb{E}(\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{\Sigma}))\mathbf{u}^* \to 0$, almost surely. Therefore, it is necessary to calculate matrix $\mathbb{E}(\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{\Sigma})$ and to study the behaviour of its bilinear forms. For this, we still use the integration by parts formula. However, the calculations are this time very tedious because it is necessary to evaluate $\mathbb{E}(\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{\Sigma})$ which is more complicated than $\mathbb{E}(\mathbf{Q}\mathbf{\Sigma})$. We therefore do not present all the calculations we did, and just provide a summary of the main steps.
First of all, for $1 \le j, k \le N$, it is clear that

$$\left(\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{\Sigma}\right)_{j,k} = \sum_{m_1,m_2,i_1,i_2} (\mathbf{\Sigma}^*)_{j,i_1}^{m_1} (\mathbf{Q}\mathcal{D}\mathbf{Q})_{i_1,i_2}^{m_1,m_2} \mathbf{\Sigma}_{i_2,k}^{m_2}$$

Recalling that $\mathbf{\Sigma} = \mathbf{A} + \mathbf{W}$, we evaluate the following term.

$$\mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q})_{i_1,i_3}^{m_1,m_3}\mathbf{\Sigma}_{i_3,k}^{m_3}(\mathbf{\Sigma}^*)_{j,i_2}^{m_2}\right] = \mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q})_{i_1,i_3}^{m_1,m_3}\mathbf{W}_{i_3,k}^{m_3}(\mathbf{\Sigma}^*)_{j,i_2}^{m_2}\right] + \mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q})_{i_1,i_3}^{m_1,m_3}\mathbf{A}_{i_3,k}^{m_3}(\mathbf{W}^*)_{j,i_2}^{m_2}\right] + \mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q})_{i_1,i_3}^{m_1,m_3}\mathbf{A}_{i_3,k}^{m_3}(\mathbf{A}^*)_{j,i_2}^{m_2}\right] \tag{4.5.7}$$

Remark 13. *We recall that ≃ between matrices means that their spectral norms have the same asymptotic behaviour.*

We use the integration by parts to evaluate the 2 first terms on the righthandside of (4.5.7). We can easily get that

$$\mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q})_{i_1,i_3}^{m_1,m_3}\mathbf{W}_{i_3,k}^{m_3}(\mathbf{\Sigma}^*)_{j,i_2}^{m_2}\right] = \sum_{i',j'}\frac{\sigma^2}{\mathrm{N}}\delta(i_3 - i' = j' - k)\times$$
$$\mathbb{E}\left[-(\mathbf{Q\Sigma})_{i_1,j'}^{m_1}(\mathbf{Q}\mathcal{D}\mathbf{Q})_{i',i_3}^{m_3,m_3}(\mathbf{\Sigma}^*)_{j,i_2}^{m_2} - (\mathbf{QDQ\Sigma})_{i_1,j'}^{m_1}\mathbf{Q}_{i',i_3}^{m_3,m_3}(\mathbf{\Sigma}^*)_{j,i_2}^{m_2} + (\mathbf{Q}\mathcal{D}\mathbf{Q})_{i_1,i_3}^{m_1,m_3}\delta(m_2 = m_3)\delta(i_2 = i')\delta(j = j')\right] \tag{4.5.8}$$

and that

$$\mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q})_{i_1,i_3}^{m_1,m_3}\mathbf{A}_{i_3,k}^{m_3}(\mathbf{W}^*)_{j,i_2}^{m_2}\right] = \sum_{i',j'}\frac{\sigma^2}{\mathrm{N}}\delta(i_2 - i' = j' - j)\mathbb{E}\left[-\mathbf{Q}_{i_1,i'}^{m_1,m_2}(\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q})_{j',i_3}^{m_3}\mathbf{A}_{i_3,k}^{m_3} - (\mathbf{Q}\mathcal{D}\mathbf{Q})_{i_1,i'}^{m_1,m_2}(\mathbf{\Sigma}^*\mathbf{Q})_{j',i_3}^{m_3}\mathbf{A}_{i_3,k}^{m_3}\right] \tag{4.5.9}$$

Summing both side of (4.5.8) and (4.5.9) over m_3 and i_3, we obtain that

$$\mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{W})_{i_1,k}^{m_1}(\mathbf{\Sigma}^*)_{j,i_2}^{m_2}\right] = -\sigma^2 c_{\mathrm{N}}\mathbb{E}\left[\left(\mathbf{Q\Sigma}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}\mathcal{D}\mathbf{Q})\right)^{\mathrm{T}}\right)_{i_1,k}^{m_1}(\mathbf{\Sigma}^*)_{j,i_2}^{m_2}\right]$$
$$-\sigma^2 c_{\mathrm{N}}\mathbb{E}\left[\left(\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{\Sigma}\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\mathrm{T}})\right)_{i_1,k}^{m_1}(\mathbf{\Sigma}^*)_{j,i_2}^{m_2}\right] + \frac{\sigma^2}{\mathrm{N}}\mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q})_{i_1,i_2-(k-j)}^{m_1,m_2}\right]\mathbb{1}_{1\leq i_2-(k-j)\leq \mathrm{L}} \tag{4.5.10}$$

and that

$$\mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A})_{i_1,k}^{m_1}(\mathbf{W}^*)_{j,i_2}^{m_2}\right] = -\frac{\sigma^2}{\mathrm{N}}\sum_{i=-(\mathrm{L}-1)}^{\mathrm{L}-1}\mathbb{E}\left[\mathbf{Q}_{i_1,i+i_2}^{m_1,m_2}(\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A})_{j-i,k} + (\mathbf{Q}\mathcal{D}\mathbf{Q})_{i_1,i+i_2}^{m_1,m_2}(\mathbf{\Sigma}^*\mathbf{QA})_{j-i,k}\right]\mathbb{1}_{1\leq i+i_2\leq \mathrm{L}}\mathbb{1}_{1\leq j-i\leq \mathrm{N}} \tag{4.5.11}$$

We recall that $z\tilde{\mathbf{Q}} = \mathbf{\Sigma}^*\mathbf{Q\Sigma} - \mathbf{I}_{\mathrm{N}}$. From (4.5.10), setting $m_1 = m_2$, $i_1 = i_2$ and summing over m_1, i_1, we get that

$$\begin{aligned}
\mathbb{E}\left[\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{W}\right]_{j,k} &= -\sigma^2 c_{\mathrm{N}}\mathbb{E}\left[\mathbf{\Sigma}^*\mathbf{Q\Sigma}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{QDQ})\right)^{\mathrm{T}}\right]_{j,k} - \sigma^2 c_{\mathrm{N}}\mathbb{E}\left[\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{\Sigma}\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\mathrm{T}})\right]_{j,k} + \sigma^2 c_{\mathrm{N}}\mathbb{E}\left[\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}\mathcal{D}\mathbf{Q})\right]_{k,j} \\
&= -\sigma^2 c_{\mathrm{N}}\mathbb{E}\left[(\mathbf{\Sigma}^*\mathbf{Q\Sigma} - \mathbf{I}_{\mathrm{N}})\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}\mathcal{D}\mathbf{Q})\right)^{\mathrm{T}}\right]_{j,k} - \sigma^2 c_{\mathrm{N}}\mathbb{E}\left[\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{\Sigma}\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\mathrm{T}})\right]_{j,k} \\
&= -\sigma^2 c_{\mathrm{N}} z\mathbb{E}\left[\tilde{\mathbf{Q}}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}\mathcal{D}\mathbf{Q})\right)^{\mathrm{T}}\right]_{j,k} - \sigma^2 c_{\mathrm{N}}\mathbb{E}\left[\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{\Sigma}\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\mathrm{T}})\right]_{j,k}
\end{aligned} \tag{4.5.12}$$

We obtain immediately that

$$\begin{aligned}
\mathbb{E}\left[\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{W}\right] &= -\sigma^2 c_{\mathrm{N}} z\mathbb{E}\left[\tilde{\mathbf{Q}}\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}\mathcal{D}\mathbf{Q})\right)^{\mathrm{T}}\right] - \sigma^2 c_{\mathrm{N}}\mathbb{E}\left[\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{\Sigma}\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbf{Q}^{\mathrm{T}})\right] \\
&\simeq -\sigma^2 c_{\mathrm{N}} z\mathbb{E}\left(\tilde{\mathbf{Q}}\right)\left(\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^{\mathrm{T}} - \sigma^2 c_{\mathrm{N}}\mathbb{E}\left(\mathbf{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{\Sigma}\right)\mathscr{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}^{\mathrm{T}}))
\end{aligned} \tag{4.5.13}$$

Similarly, from (4.5.11), setting $m_1 = m_2$, $i_1 = i_2$ and summing over m_1, i_1, we get that

$$\mathbb{E}\left[\mathbf{W}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right]_{j,k} = -\sigma^2 c_N \mathbb{E}\left[\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^T)\,\boldsymbol{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right]_{j,k} - \sigma^2 c_N \mathbb{E}\left[\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}\mathcal{D}\mathbf{Q})\right)^T\,\boldsymbol{\Sigma}^*\mathbf{Q}\mathbf{A}\right]_{j,k} \tag{4.5.14}$$

It follows that

$$\begin{aligned}
\mathbb{E}\left[\mathbf{W}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right] &= -\sigma^2 c_N \mathbb{E}\left[\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}^T)\,\boldsymbol{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right] - \sigma^2 c_N \mathbb{E}\left[\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{Q}\mathcal{D}\mathbf{Q})\right)^T\,\boldsymbol{\Sigma}^*\mathbf{Q}\mathbf{A}\right] \\
&\simeq -\sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\mathbb{E}\left(\boldsymbol{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right) - \sigma^2 c_N \left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^T \mathbb{E}\left(\boldsymbol{\Sigma}^*\mathbf{Q}\mathbf{A}\right) \\
&\simeq -\sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\mathbb{E}\left(\mathbf{W}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right) - \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\mathbb{E}\left(\mathbf{A}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right) \\
&- \sigma^2 c_N \left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^T \mathbb{E}\left(\boldsymbol{\Sigma}^*\mathbf{Q}\mathbf{A}\right)
\end{aligned} \tag{4.5.15}$$

(3.5.26) in chapter 3 implies that

$$\mathbb{E}\left[\boldsymbol{\Sigma}^*\mathbf{Q}\mathbf{A}\right] \simeq \mathbf{H}^T\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}$$

while we recall that $\mathbf{H} = \left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}))\right)^{-1}$. Therefore, (4.5.15) becomes

$$\left[\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\right]\mathbb{E}\left[\mathbf{W}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right] \simeq -\sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\mathbf{A}^*\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q})\,\mathbf{A} - \sigma^2 c_N \left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^T \mathbf{H}^T\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}$$

It follows that

$$\mathbb{E}\left[\mathbf{W}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right] \simeq -\sigma^2 c_N \mathbf{H}^T\,\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\,\mathbb{E}\left(\mathbf{A}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right) - \sigma^2 c_N \mathbf{H}^T\left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^T \mathbf{H}^T\,\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A} \tag{4.5.16}$$

By (4.5.13),(4.5.16), it is clear that

$$\begin{aligned}
\mathbb{E}\left[\boldsymbol{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\boldsymbol{\Sigma}\right] &= \mathbb{E}\left[\boldsymbol{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{W}\right] + \mathbb{E}\left[\mathbf{W}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right] + \mathbb{E}\left[\mathbf{A}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right] \\
&\simeq -\sigma^2 c_N z\mathbb{E}\left(\tilde{\mathbf{Q}}\right)\left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^T - \sigma^2 c_N \mathbb{E}\left(\boldsymbol{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\boldsymbol{\Sigma}\right)\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T)) \\
&- \sigma^2 c_N \mathbf{H}^T\,\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\,\mathbb{E}\left(\mathbf{A}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right) - \sigma^2 c_N \mathbf{H}^T\left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^T \mathbf{H}^T\,\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A} + \mathbb{E}\left[\mathbf{A}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right]
\end{aligned}$$

Hence, we eventually obtain

$$\begin{aligned}
\mathbb{E}\left[\boldsymbol{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\boldsymbol{\Sigma}\right]\left[\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\right] &\simeq -\sigma^2 c_N z\mathbb{E}\left(\tilde{\mathbf{Q}}\right)\left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^T \\
- \sigma^2 c_N \mathbf{H}^T\left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^T \mathbf{H}^T\,\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A} &+ \left[\mathbf{I}_N - \sigma^2 c_N \mathbf{H}^T\,\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T))\right]\mathbb{E}\left(\mathbf{A}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right)
\end{aligned}$$

Noticing that $\mathbf{I}_N - \sigma^2 c_N \mathbf{H}^T\,\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}^T)) = \mathbf{H}^T$, it holds that

$$\begin{aligned}
\mathbb{E}\left[\boldsymbol{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\boldsymbol{\Sigma}\right] &\simeq -\sigma^2 c_N z\mathbb{E}\left(\tilde{\mathbf{Q}}\right)\left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^T \mathbf{H}^T \\
&- \sigma^2 c_N \mathbf{H}^T\left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^T \mathbf{H}^T\,\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^T + \mathbf{H}^T\,\mathbb{E}\left(\mathbf{A}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A}\right)\mathbf{H}^T
\end{aligned} \tag{4.5.17}$$

We notice immediately that it is necessary to calculate the expectation of $\mathbf{Q}\mathcal{D}\mathbf{Q}$. To do this, we can observe that $\mathbf{Q}\mathcal{D}\mathbf{Q}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^* = z\mathbf{Q}\mathcal{D}\mathbf{Q} + \mathbf{Q}\mathcal{D}$. Therefore, we will evaluate $\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^*)$. However, simply plugging (4.5.10) and (4.5.11) to the expression

$$\mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q}\boldsymbol{\Sigma})_{i_1,k}^{m_1}(\boldsymbol{\Sigma}^*)_{j,i_2}^{m_2}\right] = \mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{W})_{i_1,k}^{m_1}(\boldsymbol{\Sigma}^*)_{j,i_2}^{m_2}\right] + \mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A})_{i_1,k}^{m_1}(\mathbf{W}^*)_{j,i_2}^{m_2}\right] + \mathbb{E}\left[(\mathbf{Q}\mathcal{D}\mathbf{Q}\mathbf{A})_{i_1,k}^{m_1}\mathbf{A}_{j,i_2}^{m_2}\right]$$

and summing over $k=j$, will not allow us to have a closed form equation of $\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q})$.
To circumvent this problem, we will use the same technique as in chapter 3. More precisely, we define the following $\mathrm{N}\times\mathrm{N}$ matrices

$$\left[\boldsymbol{\mathcal{X}}_{i_1,i_2}^{m_1,m_2}\right]_{k,j} = \mathbb{E}\left[(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}\boldsymbol{\Sigma})_{i_1,k}^{m_1}(\boldsymbol{\Sigma}^*)_{j,i_2}^{m_1}\right] \tag{4.5.18}$$

$$\left[\boldsymbol{\mathcal{Y}}_{i_1,i_2}^{m_1,m_2}\right]_{k,j} = \mathbb{E}\left[(\mathbf{Q}\boldsymbol{\Sigma})_{i_1,k}^{m_1}(\boldsymbol{\Sigma}^*)_{j,i_2}^{m_1}\right] \tag{4.5.19}$$

$$\left[\boldsymbol{\mathcal{B}}_{i_1,i_2}^{m_1,m_2}\right]_{k,j} = \mathbb{E}\left[(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q})_{i_1,i_2-(k-j)}^{m_1,m_2}\right]\mathbb{1}_{1\le i_2-(k-j)\le \mathrm{L}} \tag{4.5.20}$$

$$\left[\boldsymbol{\mathcal{C}}_{i_1,i_2}^{m_1,m_2}\right]_{k,j} = \mathbb{E}\left[\mathbf{Q}_{i_1,i_2-(k-j)}^{m_1,m_2}\right]\mathbb{1}_{1\le i_2-(k-j)\le \mathrm{L}} \tag{4.5.21}$$

and notice that $\mathbb{E}\left[(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^*)_{i_1,i_2}^{m_1,m_2}\right] = \mathrm{Tr}\,(\boldsymbol{\mathcal{X}}_{i_1,i_2}^{m_1,m_2})$. Therefore, we will calculate matrix $\boldsymbol{\mathcal{X}}_{i_1,i_2}^{m_1,m_2}$.
Using (4.5.12), (4.5.14), after a few calculations, we obtain that

$$\begin{aligned}\boldsymbol{\mathcal{X}}_{i_1,i_2}^{m_1,m_2} \simeq{}& \frac{\sigma^2}{\mathrm{N}}\mathbf{H}\boldsymbol{\mathcal{B}}_{i_1,i_2}^{m_1,m_2} - \sigma^2 c_{\mathrm{N}}\mathcal{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))\boldsymbol{\mathcal{Y}}_{i_1,i_2}^{m_1,m_2} - \frac{\sigma^2}{\mathrm{N}}\mathbf{H}\mathbf{A}^{\mathrm{T}}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))^{\mathrm{T}}\overline{\mathbf{A}}\mathbf{H}\boldsymbol{\mathcal{C}}_{i_1,i_2}^{m_1,m_2}\\
&+\frac{\sigma^2}{\mathrm{N}}\mathbf{H}\left[\sigma^2 c_{\mathrm{N}}(\boldsymbol{\mathcal{C}}_{i_1,i_2}^{m_1,m_2})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\left(\mathcal{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))\right)^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A} - \left(\boldsymbol{\mathcal{B}}_{i_1,i_2}^{m_1,m_2}\right)\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\mathbb{E}(\mathbf{Q})\mathbf{A}\right]^{\mathrm{T}}\\
&+\mathbf{H}\mathbf{A}^{\mathrm{T}}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))^{\mathrm{T}}\mathbf{f}_{i_1}^{m_1}\left(\mathbf{f}_{i_2}^{m_2}\right)^{\mathrm{T}}\overline{\mathbf{A}}\end{aligned} \tag{4.5.22}$$

Taking the trace of both side of (4.5.22), we obtain after some calculations that

$$\begin{aligned}\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^*) \simeq{}& \mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q})\left[\mathbf{A}\mathbf{H}^{\mathrm{T}}\mathbf{A}^* - \sigma^2 z\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\mathbb{E}(\tilde{\mathbf{Q}}^{\mathrm{T}}))\right] - \sigma^2\mathbb{E}(\mathbf{Q})\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\mathbf{H}\mathbf{A}^{\mathrm{T}}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))^{\mathrm{T}}\overline{\mathbf{A}}\mathbf{H})\\
&+\sigma^4 c_{\mathrm{N}}\mathbb{E}(\mathbf{Q})\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}\left(\mathbf{H}\mathbf{A}^{\mathrm{T}}\mathbb{E}(\mathbf{Q}^{\mathrm{T}})\overline{\mathbf{A}}\mathbf{H}\mathcal{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))\mathbf{H}\right) - \sigma^4 c_{\mathrm{N}}\mathbb{E}(\mathbf{Q})\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}\left(\mathbf{H}\mathcal{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))\mathbf{H}\right)\\
&-\sigma^2 c_{\mathrm{N}}\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^{\mathrm{T}}\left(\mathcal{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))\right)^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{A}^* + \sigma^4\mathbb{E}(\mathbf{Q})\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\mathbf{H}\mathcal{T}_{\mathrm{N,L}}^{(\mathrm{M})}\left(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))\mathbf{H}\mathbf{A}^{\mathrm{T}}\mathbb{E}(\mathbf{Q}^{\mathrm{T}})\overline{\mathbf{A}}\mathbf{H}\right)\end{aligned} \tag{4.5.23}$$

Now, we remark that

$$\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\mathbf{H}\mathbf{A}^{\mathrm{T}}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))^{\mathrm{T}}\overline{\mathbf{A}}\mathbf{H}) = \sum_{l=-(\mathrm{L}-1)}^{\mathrm{L}-1}\frac{1}{\mathrm{N}}\mathrm{Tr}\left(\mathbf{H}\mathbf{A}^{\mathrm{T}}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))^{\mathrm{T}}\overline{\mathbf{A}}\mathbf{H}\mathbf{J}^l\right)\mathbf{I}_{\mathrm{M}}\otimes\mathbf{J}_{\mathrm{L}}^{*l}$$

and we recall that the rank of $\mathbf{A}$ is $\mathrm{P}+\mathrm{L}-1$. As a consequence

$$\left|\frac{1}{\mathrm{N}}\mathrm{Tr}\left(\mathbf{H}\mathbf{A}^{\mathrm{T}}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))^{\mathrm{T}}\overline{\mathbf{A}}\mathbf{H}\mathbf{J}^l\right)\right| \le \mathrm{C}(z)\frac{\mathrm{L}}{\mathrm{N}}$$

It follows that

$$\left\|\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\mathbf{H}\mathbf{A}^{\mathrm{T}}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))^{\mathrm{T}}\overline{\mathbf{A}}\mathbf{H})\right\| \le \mathrm{C}(z)\frac{\mathrm{L}^2}{\mathrm{N}} = \mathcal{O}\left(\frac{\mathrm{L}}{\mathrm{M}}\right)$$

which converge towards 0 under assumption 7. Similarly, we have that

$$\left\|\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}\left(\mathbf{H}\mathbf{A}^{\mathrm{T}}\mathbb{E}(\mathbf{Q}^{\mathrm{T}})\overline{\mathbf{A}}\mathbf{H}\mathcal{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))\mathbf{H}\right)\right\| \to 0$$

$$\left\|\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\mathbf{H}\mathcal{T}_{\mathrm{N,L}}^{(\mathrm{M})}\left(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))\mathbf{H}\mathbf{A}^{\mathrm{T}}\mathbb{E}(\mathbf{Q}^{\mathrm{T}})\overline{\mathbf{A}}\mathbf{H}\right)\right\| \to 0$$

Therefore, (4.5.23) can be simplified as follows

$$\begin{aligned}z\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}) + \mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}) \simeq{}& \mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q})\left[\mathbf{A}\mathbf{H}^{\mathrm{T}}\mathbf{A}^* - \sigma^2 z\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}(\mathbb{E}(\tilde{\mathbf{Q}}^{\mathrm{T}}))\right]\\
&-\sigma^4 c_{\mathrm{N}}\mathbb{E}(\mathbf{Q})\mathbf{I}_{\mathrm{M}}\otimes\mathcal{T}_{\mathrm{L,L}}\left(\mathbf{H}\mathcal{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))\mathbf{H}\right) - \sigma^2 c_{\mathrm{N}}\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^{\mathrm{T}}\left(\mathcal{T}_{\mathrm{N,L}}^{(\mathrm{M})}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))\right)^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{A}^*\end{aligned}$$

or equivalently

$$\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q})\left[-z\left(\mathbf{I}_{ML}+\sigma^2\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}^T))\right)+\mathbf{A}\mathbf{H}^T\mathbf{A}^*\right]\simeq$$
$$\mathbb{E}(\mathbf{Q})\left[\mathcal{D}+\sigma^4 c_N\mathbf{I}_M\otimes\mathcal{T}_{L,L}\left(\mathbf{H}\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\mathbf{H}\right)+\sigma^2 c_N\mathbf{A}\mathbf{H}^T\left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^T\mathbf{H}^T\mathbf{A}^*\right]$$

Recalling that $\mathbf{R}=\left[-z\left(\mathbf{I}_{ML}+\sigma^2\mathbf{I}_M\otimes\mathcal{T}_{L,L}(\mathbb{E}(\tilde{\mathbf{Q}}^T))\right)+\mathbf{A}\mathbf{H}^T\mathbf{A}^*\right]^{-1}$, it holds that

$$\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q})\simeq\mathbb{E}(\mathbf{Q})\mathcal{D}\mathbf{R}+\sigma^4 c_N\mathbb{E}(\mathbf{Q})\mathbf{I}_M\otimes\mathcal{T}_{L,L}\left(\mathbf{H}\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\mathbf{H}\right)\mathbf{R}+\sigma^2 c_N\mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^T\left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q}))\right)^T\mathbf{H}^T\mathbf{A}^*\mathbf{R} \tag{4.5.24}$$

Lemma 4.5.2. *Under assumption A-7, for each $z\in\mathbb{C}\setminus\mathbb{R}^+$, it holds that*

$$\|\mathbb{E}(\mathbf{Q}\mathcal{D}\mathbf{Q})-\mathbf{T}\mathcal{D}\mathbf{T}\|\to 0 \tag{4.5.25}$$

Proof. This lemma will be proved in the appendix. □

Moreover, we have that $\|\mathbb{E}(\mathbf{Q})-\mathbf{T}\|\to 0$ and $\left\|\mathbb{E}(\tilde{\mathbf{Q}})-\tilde{\mathbf{T}}\right\|\to 0$. Denoting by $\mathbf{H}_\mathbf{T}=\left[\mathbf{I}_N+\sigma^2 c_N\mathcal{T}_{N,L}^{(M)}(\mathbf{T})\right]^{-1}$. (4.5.17) becomes

$$\mathbb{E}\left[\boldsymbol{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\boldsymbol{\Sigma}\right]\simeq-\sigma^2 c_N z\tilde{\mathbf{T}}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{T}\mathcal{D}\mathbf{T})\right)^T\mathbf{H}_\mathbf{T}^T$$
$$-\sigma^2 c_N\mathbf{H}_\mathbf{T}^T\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{T}\mathcal{D}\mathbf{T})\right)^T\mathbf{H}_\mathbf{T}^T\mathbf{A}^*\mathbf{T}\mathbf{A}\mathbf{H}_\mathbf{T}^T+\mathbf{H}_\mathbf{T}^T\mathbf{A}^*\mathbf{T}\mathcal{D}\mathbf{T}\mathbf{A}\mathbf{H}_\mathbf{T}^T \tag{4.5.26}$$

Since $\mathcal{D}=\mathcal{H}_{-P}^{(L)}\mathcal{H}_{-P}^{(L)*}$, the rank of $\mathcal{D}$ is $P+L-2$. As a consequence, the rank of $\mathbf{T}\mathcal{D}\mathbf{T}$ is $P+L-2$. Therefore, by remarking that

$$\mathcal{T}_{N,L}^{(M)}(\mathbf{T}\mathcal{D}\mathbf{T})=\sum_{l=-(L-1)}^{L-1}\frac{1}{ML}\mathrm{Tr}\left(\mathbf{T}\mathcal{D}\mathbf{T}\left(\mathbf{I}_M\otimes\mathbf{J}_L^l\right)\right)\mathbf{J}_N^{*l}$$

we can have immediately that $\left\|\mathcal{T}_{N,L}^{(M)}(\mathbf{T}\mathcal{D}\mathbf{T})\right\|=\mathcal{O}\left(\frac{L}{M}\right)$ which converges towards 0. Therefore,

$$\|\sigma^2 c_N z\tilde{\mathbf{T}}\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{T}\mathcal{D}\mathbf{T})\right)^T\mathbf{H}_\mathbf{T}^T-\sigma^2 c_N\mathbf{H}_\mathbf{T}^T\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{T}\mathcal{D}\mathbf{T})\right)^T\mathbf{H}_\mathbf{T}^T\mathbf{A}^*\mathbf{T}\mathbf{A}\mathbf{H}_\mathbf{T}^T\|\to 0 \tag{4.5.27}$$

Hence, equation (4.5.26) can be simplified into

$$\mathbb{E}\left[\boldsymbol{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\boldsymbol{\Sigma}\right]\simeq\mathbf{H}_\mathbf{T}^T\mathbf{A}^*\mathbf{T}\mathcal{D}\mathbf{T}\mathbf{A}\mathbf{H}_\mathbf{T}^T \tag{4.5.28}$$

Recalling that $(\mathbf{T}_1,\tilde{\mathbf{T}}_1)$ are defined by

$$\begin{cases}\mathbf{T}_1=\left(-z(1+\sigma^2\tilde{t}_1(z))\mathbf{I}_{ML}+\frac{\mathbf{A}\mathbf{A}^*}{1+\sigma^2 c_N t_1(z)}\right)^{-1}\\ \tilde{\mathbf{T}}_1=\left(-z(1+\sigma^2 c_N t_1(z))\mathbf{I}_N+\frac{\mathbf{A}^*\mathbf{A}}{1+\sigma^2\tilde{t}_1(z)}\right)^{-1}\end{cases}$$

where $t_1(z)=\frac{1}{ML}\mathrm{Tr}(\mathbf{T}_1)$ and $\tilde{t}_1(z)=\frac{1}{N}\mathrm{Tr}(\tilde{\mathbf{T}}_1)$, under assumption A-7, we have that $\|\mathbf{T}-\mathbf{T}_1\|\to 0$ and $\|\mathcal{T}_{N,L}^{(M)}(\mathbf{T}_1)-t_1(z)\mathbf{I}_{ML}\|\to 0$, with $t_1(z)=\frac{1}{ML}\mathrm{Tr}(\mathbf{T}_1)$. It is clear that

$$\left\|\mathbf{H}_\mathbf{T}-\frac{1}{1+c_N t_1(z)}\mathbf{I}_N\right\|\to 0$$

Remarking that $t_1(z)$ and $\mathbf{T}_1(z)$ can be replaced by $m_c(z)$ and $\mathbf{S}(z)$, we obtain that

$$\mathbb{E}\left[\boldsymbol{\Sigma}^*\mathbf{Q}\mathcal{D}\mathbf{Q}\boldsymbol{\Sigma}\right]\simeq\frac{1}{(1+\sigma^2 c_N m_c(z))^2}\mathbf{A}^*\mathbf{S}\mathcal{D}\mathbf{S}\mathbf{A}=\left(\mathbf{U}^*\mathcal{H}^*\frac{\mathbf{S}}{1+\sigma^2 c_N m_c(z)}\mathcal{H}_{-P}\mathbf{U}\right)^2 \tag{4.5.29}$$

which is exactly what can be obtained from (2.4.16) when P and L remain fixed.

Appendix

4.A Proof of lemma 4.3.1

To prove lemma 4.3.1, we first consider the $(P+L-1) \times (N+2(P+L-1))$ matrix

$$\tilde{\mathbf{S}}^{(L)} = \begin{pmatrix} 0 & \cdots & 0 & s_{2-P} & \cdots & s_{N-P+1} & s_{N-P+2} & \cdots & s_{N+L-1} \\ \vdots & \iddots & \iddots & \vdots & & \vdots & \iddots & \iddots & 0 \\ 0 & \iddots & \iddots & \vdots & & \vdots & \iddots & \iddots & \vdots \\ s_{2-P} & \cdots & s_{L-1} & s_L & \cdots & s_{N+L-1} & 0 & \cdots & 0 \end{pmatrix} \tag{4.A.1}$$

and we define the following $(P+L-1) \times (P+L-2)$ matrices

$$\mathbf{S}_-^{(L)} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \iddots & s_{2-P} \\ 0 & \iddots & \vdots \\ s_{2-P} & \cdots & s_{L-1} \end{pmatrix} \tag{4.A.2}$$

and

$$\mathbf{S}_+^{(L)} = \begin{pmatrix} s_{N-P+2} & \cdots & s_{N+L-1} \\ \vdots & \iddots & 0 \\ s_{N+L-1} & \iddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \tag{4.A.3}$$

We can notice that $\tilde{\mathbf{S}}^{(L)} = (\mathbf{S}_-^{(L)}, \mathbf{S}^{(L)}, \mathbf{S}_+^{(L)})$. As a consequence we have

$$\frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L} = \frac{\mathbf{S}_-^{(L)}\mathbf{S}_-^{(L)*}}{N+P+L} + \frac{\mathbf{S}_+^{(L)}\mathbf{S}_+^{(L)*}}{N+P+L} + \frac{N}{N+P+L}\frac{\mathbf{S}^{(L)}\mathbf{S}^{(L)*}}{N}$$

which gives us

$$\frac{\mathbf{S}^{(L)}\mathbf{S}^{(L)*}}{N} = \frac{N+P+L}{N}\frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L} - \frac{\mathbf{S}_-^{(L)}\mathbf{S}_-^{(L)*}}{N} - \frac{\mathbf{S}_+^{(L)}\mathbf{S}_+^{(L)*}}{N}$$

It follows that

$$\begin{aligned} \frac{\mathbf{S}^{(L)}\mathbf{S}^{(L)*}}{N} - \int_{[0,1]} \phi(\nu)\mathbf{d}_{P+L-1}(\nu)\mathbf{d}_{P+L-1}^*(\nu)d\nu = \frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L} - \int_{[0,1]} \phi(\nu)\mathbf{d}_{P+L-1}(\nu)\mathbf{d}_{P+L-1}^*(\nu)d\nu \\ + \frac{P+L}{N}\frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L} - \frac{\mathbf{S}_-^{(L)}\mathbf{S}_-^{(L)*}}{N} - \frac{\mathbf{S}_+^{(L)}\mathbf{S}_+^{(L)*}}{N} \end{aligned}$$

The introduction of matrix $\tilde{\mathbf{S}}^{(L)}$ is motivated by the observation that matrix $\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}/N+L-1$ is a Toeplitz matrix whose symbol $\hat{\phi}(\nu)$ coincides with a windowed periodogramm of sequence $(s_n)_{n=2-P,\ldots,s_{N+L-1}}$. Existing results will be used to establish that $\sup_\nu |\hat{\phi}(\nu) - \phi(\nu)| \to 0$, and this will imply that

$$\left\| \frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L} - \int_{[0,1]} \phi(\nu)\mathbf{d}_{P+L-1}(\nu)\mathbf{d}_{P+L-1}^*(\nu)d\nu \right\| \to 0$$

The strategy of the proof consists in proving the three following properties.

- $$\left\| \frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L} - \int_{[0,1]} \phi(\nu)\mathbf{d}_{P+L-1}(\nu)\mathbf{d}^*_{P+L-1}(\nu)d\nu \right\| = \mathcal{O}\left(\sqrt{\frac{\log(L)}{N^{1-\alpha}}}\right) \ , a.s \tag{4.A.4}$$
- $$\left\| \frac{\mathbf{S}^{(L)}_{-}\mathbf{S}^{(L)*}_{-}}{N} \right\| = \mathcal{O}\left(\frac{L\log(L)}{N}\right) \ , a.s \tag{4.A.5}$$
- $$\left\| \frac{\mathbf{S}^{(L)}_{+}\mathbf{S}^{(L)*}_{+}}{N} \right\| = \mathcal{O}\left(\frac{L\log(L)}{N}\right) \ , a.s \tag{4.A.6}$$

It is clear that

$$\frac{P+L}{N}\left\| \frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L} \right\| \leq \frac{P+L}{N}\left\| \frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L} - \int_{[0,1]} \phi(\nu)\mathbf{d}_{P+L-1}(\nu)\mathbf{d}^*_{P+L-1}(\nu)d\nu \right\| + \frac{P+L}{N}\left\| \int_{[0,1]} \phi(\nu)\mathbf{d}_{P+L-1}(\nu)\mathbf{d}^*_{P+L-1}(\nu)d\nu \right\|$$

and that

$$\left\| \int_{[0,1]} \phi(\nu)\mathbf{d}_{P+L-1}(\nu)\mathbf{d}^*_{P+L-1}(\nu)d\nu \right\| \leq \sup_{\nu\in[0,1]} |\phi(\nu)| < +\infty$$

Moreover, since $\frac{P+L}{N} = \mathcal{O}\left(\frac{1}{N^{1-\alpha}}\right)$, we can easily conclude that $\left\| \frac{P+L}{N}\frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L} \right\| = \mathcal{O}\left(\frac{1}{N^{1-\alpha}}\right)$, almost surely.

4.A.1 Proof of (4.A.4)

We prove the first point (4.A.4). We denote $(\mathbf{e}_1, ..., \mathbf{e}_{P+L-1})$ the canonical basis of $\mathbb{C}^{P+L-1}$. For $1 \leq k, l \leq P+L-1$, it holds that $\left(\frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L}\right)_{k,l} = \frac{1}{N+P+L}\mathbf{e}_k^T\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}\mathbf{e}_l$, and we can easily see that

- $\mathbf{e}_k^T\tilde{\mathbf{S}}^{(L)} = (\underbrace{0,\cdots,0}_{P+L-1-k}, \underbrace{s_{2-P},\cdots,s_{N+L-1}}_{N+P+L-2}, \underbrace{0,\cdots,0}_{k-1})$
- $\tilde{\mathbf{S}}^{(L)*}\mathbf{e}_l = (\underbrace{0,\cdots,0}_{P+L-1-l}, \underbrace{\overline{s_{2-P}},\cdots,\overline{s_{N+L-1}}}_{N+P+L-2}, \underbrace{0,\cdots,0}_{l-1})^T$

Thus, it yields

$$\left(\frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L}\right)_{k,l} = \begin{cases} \frac{1}{N+P+L}\sum_{m=2-P}^{N+L-1-(k-l)} s_{m+(k-l)}s_m^* \ , k \geq l \\ \frac{1}{N+P+L}\sum_{m=2-P}^{N+L-1-|k-l|} s_m s^*_{m+|k-l|} \ , k < l \end{cases} \triangleq \hat{\gamma}_{k-l}$$

Therefore, $\frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L}$ is Toeplitz, and its entries are the $\hat{\alpha}_{k-l}$. We denote by $\hat{\phi}(\nu)$ the well-known windowed periodogram estimate of $\phi(\nu)$:

$$\hat{\phi}(\nu) = \sum_{u=-(P+L-2)}^{P+L-2} \hat{\gamma}_u e^{-2i\pi\nu u} = \mathbf{a}^*_{P+L-1}(\nu)\frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L}\mathbf{a}_{P+L-1}(\nu) \tag{4.A.7}$$

and notice that

$$\frac{\tilde{\mathbf{S}}^{(L)}\tilde{\mathbf{S}}^{(L)*}}{N+P+L} = \int_{[0,1]} \hat{\phi}(\nu)\mathbf{d}_{P+L-1}(\nu)\mathbf{d}^*_{P+L-1}(\nu)d\nu$$

From this, we obtain immediately the following result:

Lemma 4.A.1. *Almost surely,for* N *large enough, if*

$$\sup_{\nu\in[0,1]} |\hat{\phi}(\nu)-\phi(\nu)| = \mathcal{O}\left(\sqrt{\frac{\log(\mathrm{L})}{\mathrm{N}^{1-\alpha}}}\right)$$

then we have

$$\left\| \frac{\tilde{\mathbf{S}}^{(\mathrm{L})}\tilde{\mathbf{S}}^{(\mathrm{L})*}}{\mathrm{N}+\mathrm{P}+\mathrm{L}} - \int_{[0,1]} \phi(\nu)\mathbf{d}_{\mathrm{P}+\mathrm{L}-1}(\nu)\mathbf{d}^*_{\mathrm{P}+\mathrm{L}-1}(\nu)d\nu \right\| = \mathcal{O}\left(\sqrt{\frac{\log(\mathrm{L})}{\mathrm{N}^{1-\alpha}}}\right)$$

almost surely.

Proof: By the definition of spectral norm, we have that

$$\begin{aligned}
\left\| \frac{\tilde{\mathbf{S}}^{(\mathrm{L})}\tilde{\mathbf{S}}^{(\mathrm{L})*}}{\mathrm{N}+\mathrm{P}+\mathrm{L}} - \int_{[0,1]} \phi(\nu)\mathbf{d}_{\mathrm{P}+\mathrm{L}-1}(\nu)\mathbf{d}^*_{\mathrm{P}+\mathrm{L}-1}(\nu)d\nu \right\| &= \sup_{\substack{\mathbf{b}_1,\mathbf{b}_2\in\mathbb{C}^{\mathrm{P}+\mathrm{L}-1}\\ \|\mathbf{b}_1\|=\|\mathbf{b}_2\|=1}} \left|\mathbf{b}_1^*\left(\int_{[0,1]}(\hat{\phi}(\nu)-\phi(\nu))\mathbf{d}_{\mathrm{P}+\mathrm{L}-1}(\nu)\mathbf{d}^*_{\mathrm{P}+\mathrm{L}-1}(\nu)d\nu\right)\mathbf{b}_2\right| \\
&\leq \sup_{\substack{\mathbf{b}_1,\mathbf{b}_2\in\mathbb{C}^{\mathrm{P}+\mathrm{L}-1}\\ \|\mathbf{b}_1\|=\|\mathbf{b}_2\|=1}} \int_{[0,1]} |\hat{\phi}(\nu)-\phi(\nu)|\,|\mathbf{b}_1^*\mathbf{d}_{\mathrm{P}+\mathrm{L}-1}(\nu)\mathbf{d}^*_{\mathrm{P}+\mathrm{L}-1}(\nu)\mathbf{b}_2|d\nu \\
&\leq \sup_{\nu\in[0,1]}|\hat{\phi}(\nu)-\phi(\nu)| \sup_{\substack{\mathbf{b}_1,\mathbf{b}_2\in\mathbb{C}^{\mathrm{P}+\mathrm{L}-1}\\ \|\mathbf{b}_1\|=\|\mathbf{b}_2\|=1}} \int_{[0,1]} |\mathbf{b}_1^*\mathbf{d}_{\mathrm{P}+\mathrm{L}-1}(\nu)\mathbf{d}^*_{\mathrm{P}+\mathrm{L}-1}(\nu)\mathbf{b}_2|d\nu \\
&\leq \sup_{\nu\in[0,1]}|\hat{\phi}(\nu)-\phi(\nu)| \sup_{\substack{\mathbf{b}_1,\mathbf{b}_2\in\mathbb{C}^{\mathrm{P}+\mathrm{L}-1}\\ \|\mathbf{b}_1\|=\|\mathbf{b}_2\|=1}} \left(\int_{[0,1]}|\mathbf{b}_1^*\mathbf{d}_{\mathrm{P}+\mathrm{L}-1}(\nu)|^2d\nu\right)^{1/2}\left(\int_{[0,1]}|\mathbf{d}^*_{\mathrm{P}+\mathrm{L}-1}(\nu)\mathbf{b}_2|^2d\nu\right)^{1/2} \\
&\leq \sup_{\nu\in[0,1]}|\hat{\phi}(\nu)-\phi(\nu)|
\end{aligned}$$

This concludes the lemma. ■

According to the result obtained in the lemma 4.A.1, we will prove (4.A.4) in two steps, $\sup_{\nu\in[0,1]}|\mathbb{E}(\hat{\phi}(\nu))-\phi(\nu)| = \mathcal{O}\left(\frac{1}{\mathrm{N}^{1-\alpha}}\right)$ and $\sup_{\nu\in[0,1]}|\hat{\phi}(\nu)-\mathbb{E}(\hat{\phi}(\nu))| = \mathcal{O}\left(\sqrt{\frac{\mathrm{L}\log(\mathrm{L})}{\mathrm{N}}}\right)$

First step.

The first step consists in showing that

$$\sup_{\nu\in[0,1]}|\mathbb{E}(\hat{\phi}(\nu))-\phi(\nu)| = \mathcal{O}\left(\frac{1}{\mathrm{N}^{1-\alpha}}\right)$$

for large N.

For this, we can rewrite the spectral density as

$$\begin{aligned}
\phi(\nu) &= \left(\sum_{k=0}^{+\infty}\alpha_k e^{-2i\pi\nu k}\right)\left(\sum_{l=0}^{+\infty}\overline{\alpha}_l e^{2i\pi\nu l}\right) \\
&= \sum_{u=-\infty}^{+\infty}\sum_{l=0}^{+\infty}\alpha_{u+l}\overline{\alpha}_l \mathbb{1}_{u+l\geq 0}e^{-2i\pi\nu u} \\
&= \sum_{u=-\infty}^{+\infty}\gamma_u e^{-2i\pi\nu u}
\end{aligned}$$

where

$$\gamma_u = \sum_{l=0}^{+\infty}\alpha_{u+l}\overline{\alpha}_l\mathbb{1}_{u+l\geq 0} = \mathbb{E}(s_{m+u}s_m^*)$$

represent the autocovariance coefficients.
It follows that

$$\mathbb{E}(\hat{\phi}(\nu)) - \phi(\nu) = \sum_{u=P+L-2}^{P+L-2} (\mathbb{E}(\hat{\gamma}_u) - \gamma_u) e^{-2i\pi\nu u} + \sum_{|u|\geq P+L-1} \gamma_u e^{-2i\pi\nu u}$$

Hence, for all $\nu \in [0,1]$

$$|\mathbb{E}(\hat{\phi}(\nu)) - \phi(\nu)| \leq \sum_{u=P+L-2}^{P+L-2} |\mathbb{E}(\hat{\gamma}_u) - \gamma_u| + \sum_{|u|\geq P+L-1} |\gamma_u| \tag{4.A.8}$$

Since, $\hat{\gamma}_u = \frac{1}{N+P+L} \sum_{m=2-P}^{N+L-1-|u|} s_{m+u} s_m^*$, it is easily seen that

$$\mathbb{E}(\hat{\gamma}_u) = \frac{1}{N+P+L} \sum_{m=2-P}^{N+L-1-|u|} \gamma_u = \left(1 - \frac{2+|u|}{N+P+L}\right) \gamma_u$$

This leads to

$$\begin{aligned}
\sum_{u=P+L-2}^{P+L-2} |\mathbb{E}(\hat{\gamma}_u) - \gamma_u| &\leq \frac{1}{N+P+L} \sum_{u=P+L-2}^{P+L-2} (2+|u|)|\gamma_u| \\
&\leq \frac{2(P+L-1)-1}{N+P+L} \left(\sum_{u=P+L-2}^{P+L-2} |\gamma_u| \right)
\end{aligned}$$

Now since $\sum_{u=P+L-2}^{P+L-2} |\gamma_u| < +\infty$, it is clear that

$$|\mathbb{E}(\hat{\gamma}_u) - \gamma_u| = \mathcal{O}\left(\frac{L}{N}\right) = \mathcal{O}\left(\frac{1}{N^{1-\alpha}}\right) \tag{4.A.9}$$

As for the term $\sum_{|u|\geq P+L-1} |\gamma_u|$, we can see that

$$\begin{aligned}
\sum_{|u|\geq P+L-1} |\gamma_u| &\leq \sum_{|u|\geq P+L-1} \sum_{l=0}^{+\infty} \alpha_{u+l} \overline{\alpha}_l \mathbb{1}_{u+l\geq 0} \\
&\leq \sum_{|u|\geq P+L-1} \sum_{|l|>\lfloor \frac{P+L-1}{2} \rfloor} |\alpha_{u+l}||\alpha_l| + \sum_{|u|\geq P+L-1} \sum_{|l|\leq\lfloor \frac{P+L-1}{2} \rfloor} |\alpha_{u+l}||\alpha_l|
\end{aligned}$$

We can easily determine an upperbound for the term

$$\begin{aligned}
\sum_{|u|\geq P+L-1} \sum_{|l|>\lfloor \frac{P+L-1}{2} \rfloor} |\alpha_{u+l}||\alpha_l| &= \sum_{|l|>\lfloor \frac{P+L-1}{2} \rfloor} \sum_{|u|\geq P+L-1} |\alpha_{u+l}||\alpha_l| \\
&\leq \left(\sum_{u\in\mathbb{Z}} |\alpha_u| \right) \left(\sum_{|l|>\lfloor \frac{P+L-1}{2} \rfloor} |\alpha_l| \right)
\end{aligned}$$

By (4.3.2), $\sum_{|l|>\lfloor \frac{P+L-1}{2} \rfloor} |\alpha_l| = \mathcal{O}\left(\frac{1}{L^\eta}\right)$, which implies that

$$\sum_{|u|\geq P+L-1} \sum_{|l|>\lfloor \frac{P+L-1}{2} \rfloor} |\alpha_{u+l}||\alpha_l| = \mathcal{O}\left(\frac{1}{L^\eta}\right) \tag{4.A.10}$$

Using the same trick, we obtain similarly that

$$\begin{aligned}
\sum_{|u|\geq P+L-1} \sum_{|l|\leq\lfloor \frac{P+L-1}{2} \rfloor} |\alpha_{u+l}||\alpha_l| &\leq \sum_{|l|\leq\lfloor \frac{P+L-1}{2} \rfloor} |\alpha_l| \sum_{|u|\geq P+L-1} |\alpha_{u+l}| \\
&\leq \sum_{|l|\leq\lfloor \frac{P+L-1}{2} \rfloor} |\alpha_l| \left(\sum_{|k|>\lfloor \frac{P+L-1}{2} \rfloor} |\alpha_k| \right) = \mathcal{O}\left(\frac{1}{L^\eta}\right)
\end{aligned} \tag{4.A.11}$$

as $\sum_{|l|\leq\lfloor\frac{P+L-1}{2}\rfloor}|\alpha_l| \leq \sum_{l=0}^{+\infty}|\alpha_l| < +\infty$.
(4.A.10) and (4.A.11) imply that

$$\sum_{|u|\geq P+L-1}|\gamma_u| = \mathcal{O}\left(\frac{1}{L^{\eta}}\right) = \mathcal{O}\left(\frac{1}{N^{\eta\alpha}}\right)$$

Since $\eta > \frac{1}{\alpha}-1$, the convergence rate $\frac{1}{N^{\eta\alpha}}$ is faster than $\frac{1}{N^{1-\alpha}}$. This, (4.A.9) and (4.A.8) imply that

$$\sup_{\nu\in[0,1]}|\mathbb{E}(\hat{\phi}(\nu))-\phi(\nu)| = \mathcal{O}\left(\frac{1}{N^{1-\alpha}}\right)$$

□

Second step.
The second step consists in showing that, almost surely, for N sufficiently large

$$\sup_{\nu\in[0,1]}|\hat{\phi}(\nu)-\mathbb{E}(\hat{\phi}(\nu))| = \mathcal{O}\left(\sqrt{\frac{L\log(L)}{N}}\right) = \mathcal{O}\left(\sqrt{\frac{\log(L)}{N^{1-\alpha}}}\right) \tag{4.A.12}$$

We recall the expression of the estimator of the spectral density $\hat{\phi}(\nu) = \sum_{u=-(P+L-2)}^{P+L-2}\hat{\gamma}_u e^{-2i\pi\nu u}$. Since $\hat{\phi}(\nu)$ is a trigonometric polynomial of order 2L, we can bound its maximum by the maximum over a discrete grid. More precisely, we use lemma 3 in [82]:

Lemma 4.A.2. *Let* $S(\nu) = \frac{1}{2}a_0 + \sum_{k=1}^{n}[a_k cos(2\pi\nu k) + b_k + sin(2\pi\nu k)]$ *be a trigonometric polynomial of order n. For any* $\delta > 0$ *and* $l > 2(1+\delta)n$*, we denote by* $\nu_j = \frac{j}{l}$ *for* $0 \leq j \leq l$*. Then, it holds that*

$$\max_{\nu\in[0,1]}|S(\nu)| \leq (1+\frac{1}{\delta})\max_{0\leq j\leq l}|S(\nu_j)|.$$

For $\delta > 0$, let $\nu_j = \frac{j}{\lceil 4(1+\delta)L\rceil}$, for $0 \leq j \leq \lceil 4(1+\delta)L\rceil$, then by lemma 4.A.2,

$$\sup_{\nu\in[0,1]}|\hat{\phi}(\nu)-\mathbb{E}(\hat{\phi}(\nu))| \leq (1+\frac{1}{\delta})\max_j|\hat{\phi}(\nu_j)-\mathbb{E}(\hat{\phi}(\nu_j))| \tag{4.A.13}$$

Moreover, we have for each $\epsilon > 0$ that

$$\mathbb{P}(\max_j|\hat{\phi}(\nu_j)-\mathbb{E}(\hat{\phi}(\nu_j))| > \epsilon) \leq \sum_j \mathbb{P}[|\hat{\phi}(\nu_j)-\mathbb{E}(\hat{\phi}(\nu_j))| > \epsilon] \tag{4.A.14}$$

We evaluate then the quantity $\mathbb{P}[|\hat{\phi}(\nu)-\mathbb{E}(\hat{\phi}(\nu))| > \epsilon]$ for each $\nu \in [0,1]$. For this we express the estimator of the spectral density as

$$\begin{aligned}\hat{\phi}(\nu) &= \sum_{u=-(P+L-2)}^{P+L-2}\left(\frac{1}{N+P+L}\sum_{\substack{m,n=2-P\\ m-n=u}}^{N+L-1} s_m s_n^*\right)e^{-2i\pi\nu u}\\ &= \frac{1}{N+P+L}\sum_{u=-(P+L-2)}^{P+L-2}\sum_{\substack{m,n=2-P\\ m-n=u}}^{N+L-1} s_m s_n^* e^{-2i\pi\nu(m-n)}\end{aligned}$$

We can notice that $\hat{\phi}(\nu)$ has the form $\frac{Q_N}{N+P+L}$, with Q_N is the quadratic form of s_n defined as

$$Q_N = \sum_{2-P\leq m,n\leq N+L-1} a_{m,n} s_m s_n^*$$

with $a_{m,n} = e^{-2i\pi\nu(m-n)}\mathbb{1}_{|m-n|\leq P+L-2}$. In addition, $\sup_{m,n}|a_{m,n}| \leq 1$ and $a_{m,n} = 0$ when $|m-n| > P+L-2$. This remark leads us to adapt Theorem 10 from [82] to our case. More precisely, the theorem 10 from [82] gives us a bound on the quantity $\mathbb{P}[|Q_N - \mathbb{E}(Q_N)| > \epsilon]$. To apply this, we have to justify some technical conditions.
Firstly, we propose the following lemma which assures that s_n is in $\mathscr{L}^p$, $\forall p \geq 1$:

Lemma 4.A.3. *For every $p \geq 1$, $s_n \in \mathscr{L}^p$*

Proof. We recall the expression of $s_n = \sum_{k=0}^{+\infty} \alpha_k \epsilon_{n-k}$.

- For $p = 1$,

$$\begin{aligned}\mathbb{E}|s_n| &= \mathbb{E}\left|\sum_{k=0}^{+\infty} \alpha_k \epsilon_{n-k}\right| \\ &\leq \sum_{k=0}^{+\infty} |\alpha_k| \mathbb{E}|\epsilon_{n-k}| < +\infty\end{aligned}$$

as $\sum_{k=0}^{+\infty} |\alpha_k|$ and $\mathbb{E}|\epsilon_{n-k}|$ are both bounded for all n, k

- For $p > 1$, there exists an integer q which satisfy $\frac{1}{p} + \frac{1}{q} = 1$, using Hölder inequality, we get

$$\begin{aligned}\mathbb{E}|s_n|^p &= \mathbb{E}\left|\sum_{k=0}^{+\infty} \alpha_k \epsilon_{n-k}\right|^p \\ &\leq \mathbb{E}\left(\sum_{k=0}^{+\infty} |\alpha_k||\epsilon_{n-k}|\right)^p \\ &\leq \mathbb{E}\left(\sum_{k=0}^{+\infty} |\alpha_k|^{1/q}|\alpha_k|^{1/p}|\epsilon_{n-k}|\right)^p \\ &\leq \mathbb{E}\left(\left(\sum_{k=0}^{+\infty} |\alpha_k|\right)^{1/q}\left(\sum_{k=0}^{+\infty} |\alpha_k||\epsilon_{n-k}|^p\right)^{1/p}\right)^p \\ &\leq \left(\sum_{k=0}^{+\infty} |\alpha_k|\right)^{p/q}\left(\sum_{k=0}^{+\infty} |\alpha_k|\mathbb{E}|\epsilon_{n-k}|^p\right)\end{aligned}$$

Since $\sum_{k=0}^{+\infty} |\alpha_k|$ and $\mathbb{E}|\epsilon_{n-k}|^p$ are both bounded for all n, k, we can conclude that $\mathbf{s}_n \in \mathscr{L}^p$. ■

For $p \geq 1$, we define

$$\Theta_p(t) = \sum_{n=t}^{\infty} \delta_p(n), \quad t \geq 2-P, \quad \text{where } \delta_p(n) = \|s_n - s'_n\|_p$$

where $s'_n = \sum_{k=0; k\neq 2-P}^{\infty} \alpha_k \epsilon_{n-k} + \alpha_{n+P-2}\epsilon'_{2-P}$ and ϵ'_{2-P} is an i.i.d copy of ϵ_{2-P}.
In Wu (2005) [81], the quantity $\delta_p(n)$ is called *physical dependence measure*. We make the convention that $\delta_p(n) = 0$ for $n < 2-P$.
We will prove that the process s_n satisfies the *short-range dependence condition* $\Theta_p := \Theta_p(2-P) < \infty$, $\forall p \geq 1$. For this we calculate

$$\begin{aligned}\delta_p(n) = \|s_n - s'_n\|_p = \|\alpha_{n+P-2}(\epsilon_{2-P} - \epsilon'_{2-P})\|_p &= |\alpha_{n+P-2}|\left(\mathbb{E}|\epsilon_{2-P} - \epsilon'_{2-P}|^p\right)^{1/p} \\ &\leq |\alpha_{n+P-2}|\left(\mathbb{E}(|\epsilon_{2-P}| + |\epsilon'_{2-P}|)^p\right)^{1/p} \\ &\leq C_p^{1/p}|\alpha_{n+P-2}|\end{aligned}$$

with

$$C_p = \sum_{i=0}^{p} \begin{pmatrix} i \\ p \end{pmatrix} M_{p-1} M_i$$

where $M_i = \mathbb{E}(|\epsilon_{2-P}|^i)$, $i = 0, ..., p$ the $p+1$ moments of ϵ_{2-P}, so that we have

$$\Theta_p(2-P) = \sum_{n=2-P}^{\infty} \delta_p(n) \leq C_p^{1/p} \sum_{n=2-P}^{\infty} |\alpha_{n+P-2}| < \infty$$

Moreover we can easily see that $\Theta_p(m) = \mathcal{O}(m^{-\eta})$, $\eta > 0$ since $\sum_{n=m}^{\infty} |\alpha_n| = \mathcal{O}(m^{-\eta})$. We can now adapt the theorem 10 from [82]

Theorem 4.A.1. *Assume that* $s_n \in \mathcal{L}^p$, $p > 4$, $\mathbb{E}(s_n) = 0$, *and* $\Theta_p(m) = \mathcal{O}(m^{-\eta})$. *Set* $c_p = (p+4)e^{p/4}\Theta_4^2$. *For any* $a > 1$, *let* $x_N = 2c_p\sqrt{NL\log(L)a}$. *Assume that* $L \to \infty$ *and* $L = \mathcal{O}(N^\alpha)$ *for some* $0 < \alpha < 1$. *Then for any* $\alpha < \beta < 1$, *there exists a constant* $C_{p,a,\beta} > 0$ *such that*

$$\mathbb{P}\left(\frac{|Q_N - \mathbb{E}(Q_N)|}{N+P+L} \geq \frac{x_N}{N+P+L}\right) \leq C_{p,a,\beta} x_N^{p/2}(\log(N))\left[(NL)^{p/4} N^{-\eta\beta p/2} + NL^{p/2-1-\eta\beta p/2} + N\right] + C_{p,a,\beta}L^{-a} \tag{4.A.15}$$

We recall that $\hat{\phi}(\nu) = \frac{Q_N}{P+P+L}$. Therefore, it is clear that $|\hat{\phi}(\nu) - \mathbb{E}(\hat{\phi}(\nu))| = \frac{|Q_N - \mathbb{E}(Q_N)|}{N+P+L}$. Applying theorem 4.A.1, we get that

$$\mathbb{P}\left(|\hat{\phi}(\nu) - \mathbb{E}(\hat{\phi}(\nu))| \geq \frac{2c_p\sqrt{NL\log(L)a}}{N+P+L}\right) \leq C_{p,a,\beta} x_N^{p/2}(\log(N))\left[(NL)^{p/4} N^{-\eta\beta p/2} + NL^{p/2-1-\eta\beta p/2} + N\right] + C_{p,a,\beta}L^{-a} \tag{4.A.16}$$

for every $\nu \in [0,1]$.
(4.A.16) and (4.A.14) imply that

$$\begin{aligned}\mathbb{P}\left(\max_j |\hat{\phi}(\nu_j) - \mathbb{E}(\hat{\phi}(\nu_j))| \geq \frac{2c_p\sqrt{NL\log(L)a}}{N+P+L}\right) &\leq \sum_j \mathbb{P}\left(|\hat{\phi}(\nu_j) - \mathbb{E}(\hat{\phi}(\nu_j))| > \frac{2c_p\sqrt{NL\log(L)a}}{N+P+L}\right) \\ &\leq L\left(C_{p,a,\beta} x_N^{p/2}(\log(N))\left[(NL)^{p/4} N^{-\eta\beta p/2} + NL^{p/2-1-\eta\beta p/2} + N\right] + C_{p,a,\beta}L^{-a}\right)\end{aligned}$$

Recall that $L = \mathcal{O}(N^\alpha)$, where $0 < \alpha < 1$. We choose $\alpha < \beta < 1$ and consider the values of a and p verifying

- $a > 1 + \frac{1}{\alpha}$ and
- $p > \max\left(\frac{2(\alpha+1)}{\eta\beta}, \frac{8}{1-\alpha+2\eta\beta\alpha}, 4\left(1+\frac{1}{\alpha}\right)\right)$

Then, it is easy to check that

$$L\left(C_{p,a,\beta} x_N^{p/2}(\log(N))\left[(NL)^{p/4} N^{-\eta\beta p/2} + NL^{p/2-1-\eta\beta p/2} + N\right] + C_{p,a,\beta}L^{-a}\right)$$

is $\mathcal{O}\left(\frac{1}{N^{1+\delta}}\right)$, with a certain $\delta > 0$. Consequently, almost surely for N large enough, it holds that

$$\max_j |\hat{\phi}(\nu_j) - \mathbb{E}(\hat{\phi}(\nu_j))| \leq C\sqrt{\frac{L\log(L)}{N}}$$

This and (4.A.13) imply that almost surely, for N large enough

$$\sup_{\nu\in[0,1]} |\hat{\phi}(\nu) - \mathbb{E}(\hat{\phi}(\nu))| = \mathcal{O}\left(\sqrt{\frac{L\log(L)}{N}}\right) = \mathcal{O}\left(\sqrt{\frac{\log(L)}{N^{1-\alpha}}}\right)$$

This finish the proof of (4.A.4). □

4.A.2 Proof of (4.A.5) and (4.A.6)

We only prove (4.A.5), (4.A.6) is obtained similarly. For this, we recall the expression of $\mathbf{S}_{-}^{(L)}$

$$\mathbf{S}_{-}^{(L)} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \iddots & s_{2-P} \\ 0 & \iddots & \vdots \\ s_{2-P} & \cdots & s_{L-1} \end{pmatrix}$$

We can notice that spectral norm of $\frac{\mathbf{S}_{-}^{(L)}}{\sqrt{N}}$ is equal to spectral norm of the $(P+L-2)\times(P+L-2)$ matrix

$$\mathbf{S}_0 = \frac{1}{\sqrt{N}} \begin{pmatrix} s_{2-P} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ s_{L-1} & \cdots & \cdots & s_{2-P} \end{pmatrix}$$

Otherwise, $\mathbf{S}_0$ can be expressed as

$$\mathbf{S}_0 = \frac{1}{\sqrt{N}} \int_{[0,1]} \sum_{l=2-P}^{L-1} s_l \, e^{-2i\pi(l+P-2)\nu} \mathbf{d}_{P+L-2}(\nu) \mathbf{d}_{P+L-2}^*(\nu)$$

As a result, we have the following inequality

$$\left\| \frac{\mathbf{S}_{-}^{(L)} \mathbf{S}_{-}^{(L)*}}{N} \right\| \leq \sup_{\nu\in[0,1]} \frac{1}{N} \left| \sum_{l=2-P}^{L-1} s_l \, e^{-2i\pi(l+P-2)\nu} \right|^2$$

Since it is easy to remark that $\frac{1}{P+L-2} \left| \sum_{l=2-P}^{L-1} s_l \, e^{-2i\pi(l+P-2)\nu} \right|^2$ is the periodogram of sequence $(s_{2-P}, ..., s_{L-1})$, (4.3.21) implies that

$$\lim_{L\to\infty} \sup_{\nu\in[0,1]} \frac{\frac{1}{P+L-2} \left| \sum_{l=2-P}^{L-1} s_l \, e^{-2i\pi(l+P-2)\nu} \right|^2}{\phi(\nu) \log(P+L-2)} = 1, \ a.s$$

from which we obtain that, $\forall \epsilon > 0, \exists N_0$ such that $\forall N > N_0$, $\forall \nu \in [0,1]$

$$1-\epsilon \leq \frac{\frac{1}{N} \left| \sum_{l=2-P}^{L-1} s_l \, e^{-2i\pi(l+P-2)\nu} \right|}{\frac{(P+L-2)\log(P+L-2)}{N} \phi(\nu)} \leq 1+\epsilon, \ \ a.s$$

It follows

$$(1-\epsilon) \frac{(P+L-2)\log(P+L-2)}{N} \phi(\nu) \leq \frac{1}{N} \left| \sum_{l=2-P}^{L-1} s_l \, e^{-2i\pi(l+P-2)\nu} \right|^2 \leq (1+\epsilon) \frac{(P+L-2)\log(P+L-2)}{N} \phi(\nu), \ \ a.s$$

This implies directly that

$$\left\| \frac{\mathbf{S}_{-}^{(L)} \mathbf{S}_{-}^{(L)*}}{N} \right\| = \mathcal{O}\left(\frac{L \log(L)}{N} \right)$$

Therefore, we get (4.A.5). (4.A.6) is obtained in the same way.
Combining (4.A.4), (4.A.5),(4.A.6), we conclude that

$$\left\| \frac{1}{N} \mathbf{S}_N^{(L)} \mathbf{S}_N^{(L)*} - \mathbf{R}_S \right\| = \mathcal{O}\left(\sqrt{\frac{\log(L)}{N^{1-\alpha}}} \right)$$

□

4.B Proof of lemma 4.5.2

Recalling that $\mathbf{H_T} = \left[\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbf{T})\right]^{-1}$, we introduce the linear functional

$$\Phi(\mathbf{X}) = \sigma^4 c_N \mathbf{T}\mathbf{I}_M \otimes \mathcal{T}_{L,L}\left(\mathbf{H_T}\mathcal{T}_{N,L}^{(M)}(\mathbf{X})\mathbf{H_T}\right)\mathbf{T} + \sigma^2 c_N \mathbf{T}\mathbf{A}\mathbf{H}_\mathbf{T}^T\left(\mathcal{T}_{N,L}^{(M)}(\mathbf{X})\right)^T \mathbf{H}_\mathbf{T}^T\mathbf{A}^*\mathbf{T}$$

It is proved in chapter 3 that $\|\mathbb{E}(\mathbf{Q}) - \mathbf{T}\|$ and $\|\mathbf{R} - \mathbf{T}\|$ converge towards 0. Therefore, the approximation (4.5.24)

$$\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}) \simeq \mathbb{E}(\mathbf{Q})\boldsymbol{\mathcal{D}}\mathbf{R} + \sigma^4 c_N \mathbb{E}(\mathbf{Q})\mathbf{I}_M \otimes \mathcal{T}_{L,L}\left(\mathbf{H}\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))\mathbf{H}\right)\mathbf{R} + \sigma^2 c_N \mathbb{E}(\mathbf{Q})\mathbf{A}\mathbf{H}^T\left(\mathcal{T}_{N,L}^{(M)}(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}))\right)^T \mathbf{H}^T\mathbf{A}^*\mathbf{R}$$

can be written as

$$\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}) \simeq \mathbf{T}\boldsymbol{\mathcal{D}}\mathbf{T} + \Phi(\mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q})) \tag{4.B.1}$$

Moreover, by equation (4.5.27), we have that

$$\|\Phi(\mathbf{T}\boldsymbol{\mathcal{D}}\mathbf{T})\| \to 0$$

Therefore, denoting by $\boldsymbol{\Delta} = \mathbb{E}(\mathbf{Q}\boldsymbol{\mathcal{D}}\mathbf{Q}) - \mathbf{T}\boldsymbol{\mathcal{D}}\mathbf{T}$, we can express (4.B.1) as

$$\boldsymbol{\Delta} = \Phi(\boldsymbol{\Delta}) + \boldsymbol{\delta}$$

with $\boldsymbol{\delta} = \Phi(\mathbf{T}\boldsymbol{\mathcal{D}}\mathbf{T}) + \boldsymbol{\epsilon}$. We mention that the matrix $\boldsymbol{\epsilon}$ is the negligible matrix which appears implicitly in the approximation (4.B.1). Its spectral norm converges towards 0. Thus, it is clear that $\|\boldsymbol{\delta}\| \to 0$.
We can notice that the functional Φ verifies

$$\|\Phi(\mathbf{X})\| \le C(z)\|\mathbf{X}\|$$

where $C(z)$ is a nice polynomial which does not depend on M, N, L.
Thus, for $z \in \mathbb{C}\backslash\mathbb{R}^+$ well chosen such that $|C(z)| < \frac{1}{2}$, for all fixed dimensional matrices $\mathbf{X}$, the series $\sum_{n=0}^{\infty}\Phi^{(n)}(\mathbf{X})$ converges.
We notice that

$$\boldsymbol{\Delta} = \sum_{k=0}^{K}\Phi^{(k)}(\boldsymbol{\delta}) + \Phi^{(K+1)}(\boldsymbol{\Delta})$$

which implies immediately that

$$\boldsymbol{\Delta} = \sum_{k=0}^{\infty}\Phi^{(k)}(\boldsymbol{\delta})$$

It is clear that

$$\|\boldsymbol{\Delta}\| \le C(z)\|\boldsymbol{\delta}\|$$

Therefore, for z well chosen $\|\boldsymbol{\Delta}\| \to 0$. We can extend this convergence on $\mathbb{C}\backslash\mathbb{R}^+$, for this, it is sufficient to use again the argument "à la Montel" used in the proof of Proposition 3.9.1.

□

Bibliography

[1] K. Abed-Meraim, E. Moulines, Ph. Loubaton, "Prediction error method for second-order blind identification", *IEEE Trans. on Signal Processing*, vol. 45, no. 3, pp. 694-705, March 1997.

[2] N.I. Akhiezer, "The classical moment problem", Oliver and Boyd, Edinburgh and London, 1965.

[3] Hong-Zhi An, Zhao-Guo Chen, E.J Hannan, "The maximum of the periodogram", *Journal of Multivariate Analysis* , Volume 13, Issue 3, September 1983, Pages 383-400.

[4] G.W. Anderson, A. Guionnet, O. Zeitouni, "An Introduction to Random Matrices", *Cambridge Studies in Advanced Mathematics*, vol. 118, Cambridge University Press, 2010.

[5] Z.D. Bai and J.W. Silverstein. *Spectral analysis of large dimensional random matrices.* Springer Series in Statistics, 2nd edition, 2010.

[6] A. Basak, A. Bose, S. Sen, "Limiting spectral distribution of sample autocovariance matrices", Bernoulli, vol. 20, no. 3 (2014), 1234-1259.

[7] R. Basu, A. Bose, S. Ganguly, R.S. Hazra, "Limiting spectral distribution of block matrices with Toeplitz block structure", *Statist. and Probab. Lett.*, 82 (2012), no. 7, 1430-1438.

[8] W. Bryc, A. Dembo, T. Jiang, "Spectral measure of large random Hankel, Markov and Toeplitz matrices", Annals of Probability, vol. 34, no. 1 (2006), 1-38.

[9] Florent Benaych-Georges and Raj Rao Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. in Math.*, 227(1):494–521, 2011.

[10] Florent Benaych-Georges and Raj Rao Nadakuditi. The singular values and vectors of low rank perturbations of large rectangular random matrices. *Journal of Multivariate Analysis*, 111:120–135, 2012.

[11] O. Besson, S. Kraut, L.L. Sharf, "Detection of an unknown rank-one component in white noise", *IEEE Trans. on Signal Process.*, vol. 54, no. 7, pp. 2835-2839, Jul. 2006.

[12] P. Bianchi, M. Debbah, M. Maëda and J. Najim, "Performance of Statistical Tests for Single Source Detection using Random Matrix Theory", *IEEE Inf. Theory*, vol. 57, no.4, pp. 2400–2419, april 2011

[13] A. Böttcher, B. Silbermann, "Introduction to large truncated Toeplitz matrices", Springer Verlag, New York, 1999.

[14] M. Capitaine, C. Donati-Martin, "Strong asymptotic freeness of Wigner and Wishart matrices", *Indiana Univ. Math. Journal*, vol. 25, pp. 295-309, 2007

[15] M. Capitaine, C. Donati-Martin, D. Féral, "The largest eigenvalue of finite rank deformation of large Wigner matrices: convergence and non-universality of the fluctuations", *Annals of Probability*, vol. 37, no. 1, pp. 1-47, 2009.

[16] R. Couillet, M. McKay, “Large Dimensional Analysis and Optimization of Robust Shrinkage Covariance Matrix Estimators”, Elsevier Journal of Multivariate Analysis, vol. 131, pp. 99-120, 2014.

[17] R. Couillet, "Robust spiked random matrices and a robust G-MUSIC estimator", *Journal of Multivariate Analysis*, 140(2015), 139-161.

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

[19] Couillet, R., A. Kammoun, and F. Pascal, "Second order statistics of robust estimators of scatter. Application to GLRT detection for elliptical signals", Elsevier Journal of Multivariate Analysis, vol. 143, pp. 249-274, 2016.

[20] Dozier, R.B. and Silverstein, J.W., "On the empirical distribution of eigenvalues of large dimensional information-plus-noise-type matrices", Journal of Multivariate Analysis, vol. 98, no. 4, pp. 678-694, 2007.

[21] Dozier, R.B. and Silverstein, J.W., "Analysis of the limiting spectral distribution of large dimensional information-plus-noise type matrices", Elsevier Journal of Multivariate Analysis, vol. 98, no. 6, pp. 1099-1122, 2007

[22] J.E Evans, D.F Sun, and J.R Johnson. Application of advanced signal processing techniques to angle estimation in ATC navigation and surveillance systems. 1982. M.I.T. Lincoln LAB, Lexington, M.A., Rep. 582.

[23] J. Evans, D.C.N. Tse, "Large system performance of linear multiuser receivers in multipath fading channels", *IEEE Trans. on Information Theory*, pp. 2059-2078, September 2000.

[24] R.R. Far, T. Oraby, W. Bryc, R. Speicher, "Spectra of large block matrices", Preprint available on Arxiv, arXiv:cs/0610045.

[25] A. Taherpour, M. Naisri-Kenari, S. Gazor, "Multiple antenna based spectrum sensing in cognitive radios, *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 814-823, Feb. 2010.

[26] V.L. Girko, "Theory of stochastic canonical equations", Mathematics and its Applications, Kluwer Academic Publishers, Dordrecht, 2001

[27] V.L. Girko, "An introduction to statistical analysis of random arrays", VSP, The Netherlands, 1998.

[28] W. Hachem, P. Loubaton, J. Najim, "Deterministic equivalents for certain functional of large random matrices", *Ann. Appl. Prob.*, vol. 17, no. 3, pp. 875-930, 2007.

[29] Walid Hachem, Philippe Loubaton, Jamal Najim, Xavier Mestre, Pascal Vallet, "Large information plus noise random matrix models and consistent subspace estimation in large sensor networks ", Random Matrices, Theory and Applications (RMTA), vol. 1, no. 2 (2012).

[30] Walid Hachem, Philippe Loubaton, Jamal Najim, Xavier Mestre, Pascal Vallet, "A subspace estimator for fixed rank perturbations of large random matrices", Journal of Multivariate Analysis, 114 (2013), pp. 427-447.

[31] Walid Hachem, Philippe Loubaton, Jamal Najim, Pascal Vallet, "On bilinear forms based on the resolvent of large random matrices", *Ann. Inst. H. Poincaré Probab. Statis.*, vol. 49, no. 1, 2013, pp. 36-63.

[32] W. Hachem, Ph. Loubaton, J. Najim, "Deterministic equivalents for certain functionals of large random matrices", Annals of Applied Probability, vol. 17, pp. 875-930, 2007.

[33] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet. A Subspace Estimator for Fixed Rank Perturbations of Large Random Matrices. *Journal of Multivariate Analysis*, 114:427–447, 2012. arXiv:1106.1497.

[34] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet. Large information plus noise random matrix models and consistent subspace estimation in large sensor networks. *Random Matrices: Theory and Applications*, 1(2), 2012.

[35] U. Haagerup, S. Thorbjornsen, "A new application of random matrices: Ext($C^*_{red}(F_2)$) is not a group", *Annals of Mathematics*, vol. 162, no. 2, 2005.

[36] S. Hiltunen, P. Loubaton, P. Chevalier, "Large system analysis of a GLRT for detection with large sensor arrays in temporally white noise", IEEE Trans. on Signal Processing, vol. 63, no. 20, pp. 5409 - 5423, October 15 2015.

[37] R. Horn, C.R. Johnson, "Matrix Analysis", 2nd. Ed., Cambridge Univ. Press, 2013.

[38] Y. Hua, "Fast blind maximum likelihood for blind identification of multiple FIR channels", *IEEE Trans. on Signal Process.*, vol. 44, no.3, pp. 661-672, Mar. 1996.

[39] E. Hung and R. Turner, "A fast beamforming algorithm for large arrays," IEEE Transactions on Aerospace and Electronic Systems, vol. 19, pp. 598–607, Jul. 1983.

[40] B.A. Johnson, Y.I. Abramovich, and X. Mestre. MUSIC, G-MUSIC, and Maximum-Likelihood performance breakdown . *IEEE Transactions on Signal Processing*, 56(8):3944 – 3958, 2008.

[41] Kammoun, A., R. Couillet, J. Najim, and M. Debbah, "Performance of mutual information inference methods under unknown interference", IEEE Transactions on Information Theory, vol. 59, issue 2, pp. 1129-1148, 02/2013.

[42] Y. Kim, S. Pillai, and J. Guerci, "Optimal loading factor for minimal sample support space-time adaptive radar," in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, (Seattle), pp. 2505–2508, 1998.

[43] S. Kritchman, B. Nadler, "Non-parametric Detection of the Number of Signals, Hypothesis Testing and Random Matrix Theory, *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3930–3941, 2009.

[44] O.Ledoit, M.Wolf, "A Well-Conditioned Estimator For Large-Dimensional Covariance Matrices", Journal of Multivariate Analysis 88 (2004) 365–411.

[45] A. Leshem, A.J. Van der Veen, "Multichannel detection of Gaussian signals with uncalibrated receivers", *IEEE Signal Process. Letters*, vol. 8, pp. 120-122, 2001

[46] T.Y. Li, D.Z. Liu, Z.D. Wang, "Limit distributions of eigenvalues for random block Toeplitz and Hankel matrices", *J. Theor. Probab.*, 24, no. 4, pp. 1063-1086 (2011).

[47] P. Wang, J. Fang, N. Han, H. Li, "Multiantenna-assisted spectrum sensing for cognitive radio", *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1791-1800, May 2010.

[48] P. Loubaton, P. Vallet, "Almost sure localization of the eigenvalues in a gaussian information plus noise model. Applications to the spiked models", Electronic J. on Probability, October 2011, pp. 1934-1959.

[49] P. Loubaton, "On the almost sure location of the singular values of certain Gaussian block-Hankel large random matrices", J. of Theoretical Probability, vol. 29, no. 4, pp. 1339-1443, December 2016.

[50] N.Ma and J. Goh, "Efficient method to determine diagonal loading value," in Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (Vol. V), pp. 341–344, 2003.

[51] C. Male, "The norm of polynomials in large random and deterministic matrices", *Probab. Theory Related Fields*, 154 (2012), no. 3-4, 477-532.

[52] V.A. Marcenko and L.A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1:457, 1967.

[53] X. Mestre, M. A. Lagunas. "Finite sample size effect on Minimum Variance beamformers: optimum diagonal loading factor for large arrays". IEEE Transactions on Signal Processing, vol. 54, no. 1, pp. 69-82, January 2006.

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

[55] X. Mestre, "Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates", IEEE Trans. Inform. Theory, 54 (11), pp. 5113-5129, 2008.

[56] E. Moulines, P. Duhamel, J.F. Cardoso, S. Mayrargue, "Subspace methods for blind identification of multichannel FIR filters", *IEEE Trans. on Signal Processing*, vol. 43, pp. 516-525, February 1995.

[57] R.R. Nadakuditi, A. Edelman, "Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples", *IEEE Transactions of Signal Processing*, vol. 56, no. 7, pp. 2625-2637, July 2008.

[58] Nadakuditi, R.R. and Silverstein, J.W., "Fundamental Limit of Sample Generalized Eigenvalue Based Detection of Signals in Noise Using Relatively Few Signal-Bearing and Noise-Only Samples", *IEEE Journal of Selected Topics in Signal Processing*, Vol. 4, No. 3, pp. 468 -480, June 2010.

[59] L.A. Pastur, "A simple approach for the study of the global regime of large random matrices", Ukrainian Math. J., vol. 57, no. 6, pp. 936-966, June 2005.

[60] L.A Pastur, M.Shcherbina, "Eigenvalue Distribution of Large Random Matrices", American Mathematical Society, 2011.

[61] G.T.Pham, P.Loubaton, P.Vallet, "Performance analysis of spatial smoothing schemes in the context of large arrays", IEEE Trans. on Signal Processing, vol. 64, no. 1, pp. 160-172, January 2016.

[62] S.U. Pillai and B.H. Kwon. Performance analysis of MUSIC type high resolution estimators for direction finding in correlated and coherent scenes . *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(8):1176–1189, 1989.

[63] Rozanov, Yu. A., Stationary random processes, Holden Day, San Francisco, CA, 1967. MR0214134

[64] W. Rudin. *Real and Complex Analysis.* McGraw-Hill, 3nd edition, 1987.

[65] D. Ramirez, G. Vazquez-Vilar, R. Lopez-Valcarce, J. Via, I. Santamaria, "Detection of rank P signals in cognitive radio netwroks with uncalibrated multiple antennas", *IEEE Trans. on Signal Process.*, vol. 59, no. 8, August 2011.

[66] D.T.M. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction", *in Proc. ICASSP 1994*, 19th IEEE International Conference on Acoustics, Speech, and Signal Processing, 19-22 April 1994, Adelaide, Australia.

[67] H.Schultz, "Non-commutative polynomials of independent Gaussian random matrices. The real and symplectic cases." Theory Relat. Fields (2005) 131: 261. doi:10.1007/s00440-004-0366-7.

[68] T.J. Shan, M. Wax, and T. Kailath. On spatial smoothing for estimation of coherent signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(8):806–811, 1985.

[69] P. Stoica and A. Nehorai. MUSIC, maximum likelihood, and Cramer-Rao bound. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 37(5):720–741, 1989.

[70] A. Thakre, M. Haardt, and K. Giridhar. Single snapshot spatial smoothing with improved effective array aperture . *IEEE Signal Processing Letters*, 16(6):505–508, 2008.

[71] A. Thakre, M. Haardt, and K. Giridhar. Tensor-based spatial smoothing (TB-SS) using multiple snapshots . *IEEE Transactions on Signal Processing*, 58(5):2715–2728, 2010.

[72] P. Vallet, P. Loubaton, and X. Mestre. Improved Subspace Estimation for Multivariate Observations of High Dimension: The Deterministic Signal Case. *IEEE Transactions on Information Theory*, 58(2), Feb. 2012. arXiv: 1002.3234.

[73] P. Vallet, X. Mestre, and P. Loubaton. A CLT for G-MUSIC DOA estimator. In *European Signal Processing Conf (EUSIPCO)*, 2012.

[74] P. Vallet, X. Mestre, P. Loubaton, "Performance analysis of an improved MUSIC DoA estimator", IEEE Trans. on Signal Processing, vol. 63, no. 23, pp. 6407-6422, December 1 2015.

[75] P. Vallet, X. Mestre, P. Loubaton, "A CLT for an improved subspace estimator with observations of increasing dimensions", Preprint arXiv:1502.02501, June 2015.

[76] A.J. Van der Veen, S. Talwar, A. Paulraj, "A subspace approach to blind space-time signal processing for wireless communication systems", *IEEE Trans. on Signal Processing*, vol. 45, no. 1, January 1997.

[77] A.J. Van der Veen, M. Vanderveen, A. Paulraj, "Joint angle and delay estimation using shift-invariant tehniques", *IEEE Trans. on Signal Processing*, vol. 46, no. 2, pp. 405-418, February 1998.

[78] H. Van Trees, Optimum Array Processing. New York: John Wiley and Sons, 2002. Part IV of "Detection, Estimation and Modulation Theory".

[79] J. Vinogradova, R. Couillet and W. Hachem, “Statistical Inference in Large Antenna Arrays under Unknown Noise Pattern”, IEEE Transactions on Signal Processing, Vol. 61, no. 22, pp. 5633-5645, 2013.

[80] J. Vinogradova, R. Couillet, and W. Hachem, “Estimation of Toeplitz covariance matrices in large dimensional regime with application to source detection”, *IEEE Transactions on Signal Processing*, 63 (18), pages 4903-4913, September 2015.

[81] Wu, W.B (2005) "Nonlinear system theory : Another look at dependence". Proc. Natl. Acad. Sci. USA 102 14150-14154 (electronic). MR2172215

[82] Han Xiao, Wei Biao Wu, "Covariance matrix estimation for stationary time series", *Annals of Statistics* . Volume 40, Number 1 (2012), 466-493.

[83] L. Yang, R. Couillet, M. McKay, "A Robust Statistics Approach to Minimum Variance Portfolio Optimization", IEEE Transactions on Signal Processing, vol. 63, no. 24, pp. 6684–6697, 2015

[84] Zygmund, A. (2002). "Trigonometric Series". Vols I, II, 3rd ed. Cambridge Univ.Press, Cambridge. MR1963498.