Estimation and fluctuations of functionals of large random matrices

Jianfeng Yao

2013-ENST-0080

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Signal et Images »

présentée et soutenue publiquement par

Jianfeng YAO

le 09.12.2013

Estimation et Fluctuations de Fonctionnelles de Grandes Matrices Aléatoires

Directeur de thèse : **Jamal NAJIM**
Co-encadrement de la thèse : **Eric MOULINES**

Jury

M. Charles BORDENAVE, Université de Toulouse	Rapporteur
Mme Florence MERLEVÈDE, Université Paris Est	Rapporteur
Mme Catherine DONATI-MARTIN, Université de Versailles-Saint-Quentin	Examinateur
Mme Sandrine PÉCHÉ, Université Paris Diderot	Examinateur
M. Walid HACHEM, Télécom Paristech	Examinateur
M. Jamal NAJIM, Université Paris Est	Directeur
M. Éric MOULINES, Télécom Paristech	Co-directeur

T H È S E

TELECOM ParisTech
École de l'Institut Mines-Télécom - membre de ParisTech
46 rue Barrault 75013 Paris - (+33) 1 45 81 77 77 - www.telecom-paristech.fr

Table des matières

Remerciements

Je tiens d'abord à remercier chaleureusement mon directeur de thèse Jamal Najim. Sa générosité constante, la richesse de ses intuitions et sa grande disponibilité ont rendu ces années très enrichissantes. J'ai appris non seulement les mathématiques, mais aussi la rigueur et la qualité de la rédaction. Il est bien évident que sans lui, cette thèse n'aurait pas vu le jour et je serais loin de profiter des mathématiques qu'il m'a transmises. Sa façon de travailler restera pour moi un modèle.

Je remercie sincèrement mon co-directeur de thèse Éric Moulines pour m'avoir proposé un sujet très riche et d'intégrer un laboratoire où l'ambiance est conviviale. Son aide a été chaleureuse et précieuse. Je voudrais aussi mentionner Romain Couillet, Mérouane Debbah, Walid Hachem, Abla Kammoun et Philippe Loubaton pour des discussions productives. En 2011, j'ai eu la chance d'assister à l'école d'été organisée par Alice Guionnet et Zhidong Bai à Changchun et je les remercie pour ce séjour confortable et profitable scientifiquement.

Je remercie Charles Bordenave et Florence Merlevède d'avoir accepté de relire cette thèse. Leurs commentaires et leurs remarques m'ont été précieux. C'est en outre un grand honneur pour moi d'avoir Catherine Donati-Martin, Sandrine Péché, Walid Hachem dans mon jury, et c'est avec plaisir que je les remercie. Catherine Donati-Martin est aussi le premier professeur qui m'a présenté les matrices aléatoires pendant mon M2 à l'Université Paris-Sud.

Je remercie mes collègues pour les repas partagés dans la bonne humeur : Alexandre, Amandine, Andrés, Charanpal, Charles, Cristina, Yu Ding, Émilie C., Émilie K., Éric, Joffrey, Marc, Mony, Nicolas, Olaf, Olivier, Onur, Sylvain L.C., Sylivain R., Virgile, Yafei Xing. Je remercie Fabrice Planche pour ses aides informatiques et ses blagues ainsi que Laurence Zelmar pour l'aide administrative.

Je voudrais aussi remercier mes amis chinois pour leur amitié : mes amis de l'ENS (Yu Chen, Lie Fu, Fangzhou Jin, Shinan Liu, Zhenkai Lv, Zhe Tong, Bin Xu, Fangjia Yan), mes amis de classe préparatoire (Yuan Bian, Jiaming Chen, Fanding Gu, Yue Huang, Wenfeng Li, Wentian Liu), mes amis de bridge (Yizuo Chen, Zhoujie Ding, Hui Jiang, Shuoyan Liu, Yuntao Ni, Songyan Xie) et bien d'autres (Xin He, Bin Li, Meilin Li, Yang Wang, Yang Yang, Yixian Zhou, etc). Parmi eux, j'aimerais exprimer un remerciement particulier au couple Shize Dai et Yang Lu pour les nombreux repas et voyages partagés et leur souhaite plein de bonheur.

J'exprime toute ma gratitude à mes parents pour leur soutien constant.

Résumé

L'objectif principal de la thèse est : l'étude des fluctuations de fonctionnelles du spectre de grandes matrices aléatoires, la construction d'estimateurs consistants et l'étude de leurs performances, dans la situation où la dimension des observations est du même ordre que le nombre des observations disponibles. Un tel contexte se rencontre très fréquemment en communications numériques, principal domaine applicatif de la thèse. On retrouve les mêmes problématiques dès lors que l'on traite des observations de très grande dimension (théorie de l'apprentissage, biostatistiques, etc.), ou que des contraintes fortes de temps réel (réduisant le nombre d'observations que l'on peut acquérir) entrent en jeu (radio cognitive, radio intelligente, réseaux de capteurs).

Il y aura deux grandes parties dans cette thèse. La première concerne la contribution méthodologique. Nous ferons l'étude des fluctuations pour les statistiques linéaires des valeurs propres du modèle 'information-plus-bruit' pour des fonctionnelles analytiques, et étendrons ces résultats au cas des fonctionnelles non analytiques. Le procédé d'extension est fondé sur des méthodes d'interpolation avec des quantités gaussiennes. Ce procédé est appliqué aux grandes matrices de covariance empirique. L'autre grande partie sera consacrée à l'estimation des valeurs propres de la vraie covariance à partir d'une matrice de covariance empirique en grande dimension et l'étude de son comportement. Nous proposons un nouvel estimateur consistant et étudions ces fluctuations. En communications sans fil, cette procédure permet à un réseau secondaire d'établir la présence de ressources spectrales disponibles.

Les objectifs scientifiques de la thèse présentée seront :

– Fluctuations de statistiques linéaires pour grandes matrices aléatoires,
– Construction d'estimateurs consistants dans un contexte de traitement d'antenne et de canaux de propagation MIMO et l'étude de la performance de ces estimateurs.

Abstract

The principal objective of this thesis is : the study of the fluctuations of functionals of spectrum for large random matrices, the construction of consistent estimators and the study of their performances, in the situation where the dimension of observations is with the same order as the number of the available observations. Such a situation shows up very frequently in numerical communications, which is the principal application domain of the thesis. It is also an extremely large situation which appears as soon as the observations have very large dimensions (machine learning, biostatistics, etc.), or when some strong constraints of real time, (reducing the number of observations we have acquired) play a decisive role (cognitive radio, intelligent radio, wireless sensor networks).

There will be two parts in the report : the methodological contribution and the estimation in large-dimensional data. As to the methodological contribution, we will study the fluctuations for spectral linear statistics of the model 'information-plus-noise' for analytic functionals, and the extension for non-analytic functionals. The extension is based on the interpolation between random variables and Gaussian terms. This method can be applied to empirical covariance matrices. Another part consists in the estimation of the eigenvalues of the real covariance from the empirical covariance for high dimensional data and the study of its performance. We propose a new consistent estimator and the fluctuation of the estimator will be studied . In wireless communications, this procedure permits a secondary network to ensure the presence of the available spectral resources.

The scientific objectives of the presented thesis are :

– Fluctuations of functionals for large dimensional random matrices,
– Construction of consistent estimators in the context of signal processing and radio channel MIMO and the study of the performance of these estimators.

Acronyms

a.s.	almost surely
CDMA	code-division multiple access
CLT	central limit theorem
DF	distributed function
ESD	empirical spectral distribution
i.i.d.	independent and identically distributed
LHS	left-hand side
LSD	limiting spectral distribution
MIMO	multiple-input multiple-output
MP	Marčenko-Pastur
RHS	right-hand side
RMT	random matrix theory

Notations

Linear algebra

$\mathbf{X}$	matrix
$\mathbf{I}_N$	identity matrix of size $N \times N$
$\text{diag}(x_1, \cdots, x_N)$	diagonal matrix with entries $(x_1, \cdots, x_N)$
$\text{eig}(\mathbf{X})$	the set of eigenvalues of $\mathbf{X}$
X_{ij},$[X]_{ij}$,x_{ij}	(i, j) entry of matrix $\mathbf{X}$
$\mathbf{X}^T$	transpose of matrix $\mathbf{X}$
$\mathbf{X}^H$ or $\mathbf{X}^*$	complex conjugate transpose of $\mathbf{X}$
$\bar{\mathbf{X}}$	complex conjugate of $\mathbf{X}$
$\text{vdiag}(\mathbf{X})$	vector with entries $(\mathbf{X}_{11}, \cdots, \mathbf{X}_{NN})$
$\text{Tr}(\mathbf{X})$, $\det(\mathbf{X})$, $\text{rank}(\mathbf{X})$	trace, determinant and rank of $\mathbf{X}$
$\|\mathbf{X}\|$	spectral norm of $\mathbf{X}$
$\mathbf{x}$	column vector
x_i	i-th entry of $\mathbf{x}$
$\leq, \geq, >, <$	matrix inequalities, $\mathbf{A} \geq \mathbf{B}$ means $\mathbf{A} - \mathbf{B}$ is nonnegative while $\mathbf{A} > \mathbf{B}$ means $\mathbf{A} - \mathbf{B}$ is nonnegative definite

Probability theory

$(\Omega, \mathcal{F}, \mathcal{P})$	probability space with σ-algebra $\mathcal{F}$ and measure $\mathcal{P}$
$\mathbb{E}(X)$	expectation of X
$\mathbb{P}$	probability
F_X	distribution function of X, i.e., $F_X(x) = \mathbb{P}(X \leq x)$
supp(F)	support of the distribution function F
$\xrightarrow{a.s.}$	almost surely convergence
$\xrightarrow{\mathcal{P}}$, $\xrightarrow{\mathcal{D}}$	convergence in probability and distribution
$\mathcal{CN}(\mathbf{m}, \mathbf{Q})$	complex Gaussian distribution with mean $\mathbf{m}$ and covariance $\mathbf{Q}$
$\mathcal{N}(\mathbf{m}, \mathbf{Q})$	Gaussian distribution with mean $\mathbf{m}$ and covariance $\mathbf{Q}$
$\mathbf{X}(\omega)$	the realization of the variable $\mathbf{X}$ at point ω
δ_x	the Dirac probability measure at x

Analysis

$\mathbb{C}, \mathbb{R}, \mathbb{N}$	the complex, real and natural number
$\mathrm{Re}(z)$, $\mathrm{Im}(z)$	real and imaginary part of z
$\bar{z}$	conjugate of z
$\mathbb{C}^+$	$\{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$
$\mathbb{R}^+, \mathbb{R}^-$	nonnegative and nonpositive real numbers
$\mathbb{C}^*$	$\mathbb{C}\backslash\{0\}$
$\mathbf{i}$	$\mathbf{i} = \sqrt{\mathbf{-1}}$ with $\mathrm{Im}(\mathbf{i}) = 1$
$\lvert z\rvert$	modulus of z
$\mathbb{I}_A$	indicator function
$(x_n)_{n\geq 1}$	sequence of $x_1, x_2, \cdots$
$\mathcal{O}(y_n)$	$x_n = \mathcal{O}(y_n)$ means there exists C and n_0 such that $x_n \leq Cy_n$ for $n \geq n_0$
$o(y_n)$	$x_n = o(y_n)$ means for all $\epsilon > 0$ and there exists n_0 such that $x_n \leq \epsilon y_n$ for $n \geq n_0$
$f'(x)$,$f''(x)$, $f^{(n)}(x)$	first, second and n-th derivative of f
$\limsup_n x_n$, $\liminf_n x_n$	limit superior and inferior of (x_n)
$\delta_{i,j}$	Kronecker's symbole, i.e., $\delta_{i,j} = 1$ if $i = j$, 0 otherwise

Chapitre 1

Introduction

L'origine de l'étude de matrices aléatoires remonte aux travaux du statisticien Wishart en 1928 [98] qui s'intéresse au comportement asymptotique de la matrice de covariance empirique définie par

$$\mathbf{R}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^*,$$

où $\mathbf{x}_i \in \mathbb{C}^N$ est i.i.d. centrée réduite.

En 1955 [96], et pour des motivations de physique théorique [97], le physicien Wigner s'intéresse au comportement asymptotique du spectre d'une matrice aléatoire symétrique, à entrées indépendantes et correctement normalisées quand la dimension de la matrice tend vers l'infini. Wigner établit la convergence de la mesure empirique des valeurs propres vers la loi semi-circulaire.

Des résultats similaires voient le jour pour de grandes matrices de covariance empirique pour lesquelles les dimensions tendent vers l'infini au même rythme (rapport entre les dimensions constant). Le résultat emblématique est le théorème de Marčenko-Pastur [62] qui décrit le comportement asymptotique de la mesure spectrale d'une grande matrice de covariance empirique ; s'en suivent l'étude des fluctuations de la plus grande valeur propre [88], l'étude des fluctuations du spectre [11], certains théorèmes de grandes déviations [13], etc.

D'un point de vue des applications, la théorie des grandes matrices aléatoires a suscité un grand intérêt en physique [63], en biologie [3], en finance [77], en théorie des communications numériques [89, 94, 87, 33], etc.

Nos travaux ont porté sur l'estimation et l'étude des fluctuations de fonctionnelles de grandes matrices aléatoires de covariance empirique pour des modèles de matrices structurées : matrice de *Gram* et matrices du type *information-plus-bruit*. Dans cette introduction, nous présentons brièvement nos travaux ainsi que certains outils utilisés dans la suite de ce manuscrit.

1.1 CLT pour le modèle 'information-plus-bruit'

Considérons une matrice aléatoire $\mathbf{Y}_n = (y_{ij}^n)$ donnée par :

$$\mathbf{Y}_n = \frac{1}{\sqrt{n}} \mathbf{X}_n + \mathbf{A}_n \ ,$$

où $\mathbf{X}_n$ est une matrice $N \times n$ dont les entrées $(x_{ij}^n\ ; i, j, n)$ sont réelles ou complexes, i.i.d. de moyenne 0 et variance 1 ; la matrice $\mathbf{A}_n$ a la même dimension et est déterministe. La matrice $\mathbf{Y}_n$ est parfois appelée matrice du type "*information-plus-bruit*" dans la littérature.

Dans le Chapitre 2, nous étudierons les fluctuations de statistiques linéaires de la forme :

$$\mathrm{Tr} f(\mathbf{Y}_n \mathbf{Y}_n^*) = \sum_{i=1}^{N} f(\lambda_i) \ ,$$

où les λ_i sont les valeurs propres de $\mathbf{Y}_n \mathbf{Y}_n^*$, et f est une fonction analytique, dans le régime où les dimensions de n et N tendent vers l'infini à la même vitesse :

$$N, n \to \infty \quad \text{et} \quad 0 \ < \ \liminf \frac{N}{n} \ \leq \ \limsup \frac{N}{n} \ < \ \infty \ .$$

Les grandes matrices du type information-plus-bruit, et plus généralement, les grandes matrices aléatoires non centrées, ont récemment attiré l'attention de chercheurs comme dans Loubaton et al. [93], Hachem et al. [46], Capitaine et al. [22]. Elles constituent un modèle important pour les applications. Sous des conditions peu restrictives sur les moments des entrées de la matrice $\mathbf{X}_n$ et sur la norme spectrale de la matrice $\mathbf{A}_n$, le comportement asymptotique de la distribution empirique des valeurs propres de $\mathbf{Y}_n \mathbf{Y}_n^*$:

$$F^{\mathbf{Y}_n \mathbf{Y}_n^*} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i} \ , \tag{1.1.1}$$

a été étudié par Girko [35, Chapitre 7], Dozier et Silverstein [30] et Hachem et al. [44]. En suivant ces résultats, les propriétés variées du spectre asymptotique ont été étudiées, voir par exemple [29, 58, 7, 21]. Dans le contexte plus applicatif, les résultats ont été développés en communications numériques et traitement du signal [33, 87, 68, 32, 42, 43].

Concernant les fluctuations de fonctionnelles du spectre de grandes matrices aléatoires, les références se retrouvent dans les monographies [1, 6, 74, 86]. Dans le cas spécifique de fluctuations des grandes matrices de covariance, nous pouvons citer Jonsson [49], Johansson [47], Cabanal-Duvillard [19], Guionnet [37], Bai et Silverstein [11], Hachem et al. [45, 39], Pan et al. [73, 12], Chatterjee [23], Lytova et Pastur [60], Shcherbina [74], Najim [70], Khorunzhiy et al. [40], Benaych-Georges et al. [14], mais beaucoup d'autres articles méritent d'être cités.

Dans le cas de matrices non centrées du type information-plus-bruit, peu de résultats ont été développés concernant leurs fluctuations. Dans le cas spécifique où la fonctionnelle d'intérêt est

$$\log\det\left(\mathbf{I}_n + \mathbf{Y}_n\mathbf{Y}_n^*\right) = \sum_{i=1}^{N} \log(1+\lambda_i)\ ,$$

les fluctuations ont été décrites à l'aide de techniques physiciennes [69] dans le cas où les entrées sont gaussiennes complexes, et le CLT a été établi dans le cas général par Hachem et al. dans [39].

Notre étude portera sur les fluctuations de $\mathrm{Tr} f(\mathbf{Y}_n\mathbf{Y}_n^*)$ dans le cas où f est analytique. Dans ce cas, la formule de Cauchy entraîne que

$$\mathrm{Tr} f(\mathbf{Y}_n\mathbf{Y}_n^*) = -\frac{1}{2i\pi}\oint_C f(z)\,\mathrm{Tr}\left(\mathbf{Y}_n\mathbf{Y}_n^* - zI_N\right)^{-1} dz\ . \tag{1.1.2}$$

Soit $\mathbf{Q}_n(z) = (\mathbf{Y}_n\mathbf{Y}_n^* - z\mathbf{I}_N)^{-1}$ (resp. $\tilde{\mathbf{Q}}_n(z) = (\mathbf{Y}_n^*\mathbf{Y}_n - z\mathbf{I}_n)^{-1}$) la résolvante de $\mathbf{Y}_n\mathbf{Y}_n^*$ (resp. $\mathbf{Y}_n^*\mathbf{Y}_n$). Considérons $\delta_n, \tilde{\delta}_n : \mathbb{C}^+ \mapsto \mathbb{C}^+$, qui satisfont

$$\begin{cases} \delta_n(z) = \frac{1}{n}\mathrm{Tr}\left(-z(1+\tilde{\delta}_n)\mathbf{I}_N + \frac{\mathbf{A}_n\mathbf{A}_n^*}{1+\delta_n}\right)^{-1} \\ \tilde{\delta}_n(z) = \frac{1}{n}\mathrm{Tr}\left(-z(1+\delta_n)\mathbf{I}_n + \frac{\mathbf{A}_n^*\mathbf{A}_n}{1+\tilde{\delta}_n}\right)^{-1} \end{cases}, \tag{1.1.3}$$

et $\mathbf{T}_n$ (resp. $\tilde{\mathbf{T}}_n$) l'équivalent déterministe de $\mathbf{Q}_n$ (resp. $\tilde{\mathbf{Q}}_n$) défini par

$$\begin{cases} \mathbf{T}_n(z) = \left(-z(1+\tilde{\delta}_n)\mathbf{I}_N + \frac{\mathbf{A}_n\mathbf{A}_n^*}{1+\delta_n}\right)^{-1} \\ \tilde{\mathbf{T}}_n(z) = \left(-z(1+\delta_n)\mathbf{I}_n + \frac{\mathbf{A}_n^*\mathbf{A}_n}{1+\tilde{\delta}_n}\right)^{-1} \end{cases}. \tag{1.1.4}$$

Dans [46], ils démontrent que $\frac{n}{N}\delta_n$ (resp. $\tilde{\delta}_n(z)$) est la transformée de Stieljes d'une mesure de probabilité F_n (resp. $\tilde{F}_n$) et $\mathbf{T}_n$ (resp. $\tilde{\mathbf{T}}_n$) est une bonne approximation de $\mathbf{Q}_n$ (resp. $\tilde{\mathbf{Q}}_n$) dans le sens où

$$\frac{1}{n}\mathrm{Tr}\mathbf{T}_n(z) - \frac{1}{n}\mathrm{Tr}\mathbf{Q}_n(z) \xrightarrow[N,n\to\infty]{p.s.} 0, \quad \mathbf{u}^*\mathbf{Q}_n\mathbf{v} - \mathbf{u}^*\mathbf{T}_n\mathbf{v} \xrightarrow[N,n\to\infty]{p.s.} 0,$$

où $\mathbf{u}, \mathbf{v}$ sont deux vecteurs $N \times 1$ avec la norme bornée.

La statistique linéaire qui nous intéresse peut s'écrire sous la forme

$$M_n(z) \stackrel{\triangle}{=} \mathrm{Tr}\mathbf{Q}_n(z) - \mathrm{Tr}\mathbf{T}_n(z). \tag{1.1.5}$$

On peut décomposer (1.1.5) en deux parties :

$$\begin{aligned} \mathrm{Tr}(\mathbf{Q}_n(z) - \mathbf{T}_n(z)) &= \underbrace{\mathrm{Tr}(\mathbf{Q}_n(z) - \mathbb{E}\,\mathbf{Q}_n(z))}_{\text{Fluctuations}} + \underbrace{\mathrm{Tr}(\mathbb{E}\,\mathbf{Q}_n(z) - \mathbf{T}_n(z))}_{\text{Biais}} \\ &\stackrel{\triangle}{=} M_n^1(z) + M_n^2(z)\ . \end{aligned}$$

Le premier terme, qui est centré, donnera lieu aux fluctuations de la statistique linéaire et sera géré par la méthode des martingales [1]. Cette stratégie a été appliquée avec succès dans [11, 73, 45, 53, 42, 70, 12]. Le second terme est complètement déterministe et donnera le biais. Dans le cas où $\mathbf{A}_n = 0$, nous retrouverons l'expression de la variance dans [11] et du biais dans [70] (cf. Section 2.2.3).

Trois difficultés majeures apparaissent dans la preuve :

1. Comme nous travaillons avec des variables générales ayant un 4^e moment fini, deux termes supplémentaires apparaîtrons dans l'étude de la variance asymptotique : un terme proportionnel au cumulant d'ordre 4, et un terme proportionnel à $\mathbb{E}(x_{11}^n)^2$. Naturellement, ces termes dépendent de la matrice $\mathbf{A}_n\mathbf{A}_n^*$. Mais malheureusement, ils ne dépendent pas uniquement du spectre de $\mathbf{A}_n\mathbf{A}_n^*$, mais aussi des vecteurs propres de $\mathbf{A}_n\mathbf{A}_n^*$. Cela entraîne qu'il est difficile de formuler des hypothèses naturelles sur $\mathbf{A}_n\mathbf{A}_n^*$ permettant la convergence de ces deux termes. Au lieu d'exprimer le résultat sous la forme habituelle
$$\mathrm{Tr} f(\mathbf{Y}_n\mathbf{Y}_n^*) - N\int f(x)F_n(dx) \xrightarrow[N,n\to\infty]{\mathcal{D}} \mathcal{N}(\mathcal{B}_\infty, \Theta_\infty),$$
on l'exprimera sous la forme
$$d_{\mathcal{LP}}\left(\mathrm{Tr} f(\mathbf{Y}_n\mathbf{Y}_n^*) - N\int f(x)F_n(dx), \mathcal{N}(\mathcal{B}_n, \Theta_n)\right) \xrightarrow[N,n\to\infty]{} 0$$
à l'aide de la distance de Lévy-Prohorov entre la statistique linéaire d'intérêt et une famille de variables gaussiennes, qui s'ajustera à la famille des matrices. Cela permettra d'éviter de formuler des hypothèses peu naturelles sur la convergence jointe des vecteurs propres et du spectre de $\mathbf{A}_n\mathbf{A}_n^*$. (cf. Section 2.2)

2. L'introduction d'une famille de gaussiennes $\mathcal{N}(\mathcal{B}_n, \Theta_n)$ rend nécessaire différents contrôles asymptotiques, par exemple, la tension de la famille gaussienne, une borne uniforme de la variance et du biais. Malheureusement, la définition de leurs paramètres $\mathcal{B}_n$ et Θ_n repose sur les équations implicites définissant δ_n et $\tilde{\delta}_n$ (cf. Eq (1.1.3)). Pour contourner ce problème, nous introduisons un méta-modèle qui permet de transférer les propriétés de notre statistique linéaire à cette famille de gaussiennes. Le méta-modèle consiste à introduire un nouveau paramètre $M \in \mathbb{N}$ pour que δ_n soit la limite de la résolvante normalisée associée à la matrice définie par
$$\mathbf{A}_n(M) = \begin{bmatrix} \mathbf{A}_n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{A}_n \end{bmatrix}$$
et
$$\mathbf{Y}_n(M) = \frac{1}{\sqrt{nM}}\mathbf{X}_n(M) + \mathbf{A}_n(M),$$
où $\mathbf{X}_n(M)$ est une matrice $NM \times nM$ centrée réduite et $\mathbf{A}_n(M)$ est une matrice par bloc dont les éléments sur la diagonale par bloc est $\mathbf{A}_n$. L'intérêt d'introduire la matrice $\mathbf{A}_n(M)$

1. Dans [35, Introduction] (voir aussi les références dans [35]), Girko appelle cette méthode la méthode REFORM (REsolvent, FORmula, Martingales).

repose sur le fait que $\mathbf{Y}_n(M)\mathbf{Y}_n(M)^*$ a le même équivalent déterministe que $\mathbf{Y}_n\mathbf{Y}_n^*$, *i.e.*,

$$\frac{1}{NM}\mathrm{Tr}\mathbf{Q}_{nM}(z) - \frac{1}{N}\mathrm{Tr}\mathbf{T}_n(z) \xrightarrow[M\to\infty]{} 0$$

où $\mathbf{Q}_{nM} = (\mathbf{Y}_n(M)\mathbf{Y}_n{}^*(M) - z\mathbf{I}_{NM})^{-1}$. $\delta_{nM} = \delta_n$ est indépendant de M du fait de la nature bloc-diagonale de $\mathbf{A}_n(M)$. N, n étant fixés, notons $\mathcal{N}(\mathcal{B}_{n,M}, \Theta_{n,M})$ la famille gaussienne de la loi limite associée à $\mathbf{Q}_{nM}$, nous pouvons montrer sans difficulté que

$$\forall M \geq 1, \quad \mathcal{B}_{n,M} = \mathcal{B}_n, \quad \Theta_{n,M} = \Theta_n,$$

et transférer les propriétés de $\mathrm{Tr}\mathbf{Q}_n$ aux $\mathcal{B}_n$ et Θ_n. Cette méthode sera utilisée de manière systématique dans l'étude de CLT (cf. Section 2.3.6, 3.3.1).

3. Le calcul de la covariance s'inspire fortement des calculs faits dans Hachem et al. [39] dans le cadre particulier de la fonctionnelle $\log\det(\mathbf{I}_N + \mathbf{Y}_n\mathbf{Y}_n^*)$. Du fait de la formule de dérivation
$$\frac{\partial}{\partial\rho}\log\det\left(\rho\mathbf{I}_n + \mathbf{Y}_n\mathbf{Y}_n^*\right) = \sum_{i=1}^{N}\frac{1}{\rho+\lambda_i}\,, \quad \text{pour } \rho > 0,$$
il existe des liens profonds avec les calculs que nous devons mener. Une différence importante réside sur le fait que nous travaillons dans le plan complexe alors que Hachem et al. travaillent sur l'axe réel. Cette différence induit des difficultés techniques substantielles (cf. Section 2.4).

1.2 Fluctuations pour des fonctionnelles non analytiques

Dans le Chapitre 3, nous nous proposons d'étudier les fluctuations des fonctionnelles non-analytiques pour le modèle $\mathbf{Y}_n = \frac{1}{\sqrt{n}}\mathbf{R}_n^{1/2}\mathbf{X}_n$ où les entrées sont supposées réelles, centrées réduites et $\mathbf{R}_n$ est une matrice symmétrique avec la norme spectrale bornée. La représentation (1.1.2) est la clé pour la méthode des martingales et impose une contrainte très forte sur la nature de la fonction f, qui doit être analytique sur un contour spécifique.

Afin d'affaiblir l'hypothèse d'analycité de f, nous allons procéder différemment et adopter le point de vue de Lytova et Pastur [60]. Pour des entrées réelles, une autre manière de décomposer une fonction est la transformée de Fourier : $\hat{f}(t) = (2\pi)^{-1}\int e^{-it\lambda}f(\lambda)\,d\lambda$ et sa formule d'inversion $f(\lambda) = \int e^{\mathbf{i}\lambda t}\hat{f}(t)\,dt$, qui amène la représentation :

$$\mathrm{Tr}f(\mathbf{Y}_n\mathbf{Y}_n^T) = \int \hat{f}(t)\,\mathrm{Tr}e^{\mathbf{i}t\mathbf{Y}_n\mathbf{Y}_n^T}\,dt\,. \tag{1.2.1}$$

Nous pouvons remarquer que la représentation précédente ne nécessite pas que f soit analytique, et c'est là l'un des grands avantages pour décomposer une fonction dans le domaine de Fourier plutôt que par la formule de Cauchy (dans le domaine de Stieljes).

La stratégie est maintenant claire : La fonction $z \mapsto e^{\mathbf{i}tz}$ étant analytique sur $\mathbb{C}$, nous pouvons décrire les fluctuations de

$$\mathrm{Tr}[e^{\mathbf{i}t\mathbf{Y}_n\mathbf{Y}_n^T}] = -\frac{1}{2i\pi}\oint e^{itz}\mathrm{Tr}\left(\mathbf{Y}_n\mathbf{Y}_n^T - z\mathbf{I}_N\right)^{-1} dz ,$$

à l'aide de l'extension du théorème de Bai et Silverstein comme développé dans [11], et un argument de tension va fournir un CLT pour

$$\int_{-K}^{K} \hat{f}(t)\,\mathrm{Tr}[e^{\mathbf{i}t\mathbf{Y}_n\mathbf{Y}_n^T}]\,dt. \tag{1.2.2}$$

Pour retrouver la représentation (1.2.1), nous avons besoin d'étendre le CLT précédent sur l'axe réel ($K \to \infty$). L'ingrédient clé sera un contrôle, à priori, de la variance de la variable aléatoire $u_n(t) = \mathrm{Tr}[e^{\mathbf{i}t\mathbf{Y}_n\mathbf{Y}_n^T}]$ qui sera un polynôme de t. Ce contrôle utilise les techniques d'interpolation développées par Lytova et Pastur [60]. L'idée principale consiste à interpoler entre les statistiques linéaires $u(t) = \mathrm{Tr}[e^{\mathbf{i}t\mathbf{Y}_n\mathbf{Y}_n^T}]$ et les statistiques linéaires gaussiennes réelles $\hat{u}(t) = \mathrm{Tr}[e^{\mathbf{i}t\hat{\mathbf{Y}}_n\hat{\mathbf{Y}}_n^T}]$ en considérant la matrice d'interpolation

$$\mathbf{X}_s = \sqrt{s}\mathbf{X}_n + \sqrt{1-s}\hat{\mathbf{X}}_n,$$

et les quantités correspondantes.

Dans le cas gaussien, une application directe de l'inégalité de Nash-Poincaré [41, 75, 22] aboutit à :

$$\mathrm{Var}\{\hat{u}(t)\} \le Kt^2.$$

Les techniques d'interpolation nous permettent d'obtenir le contrôle suivant

$$\mathrm{Var}[\mathrm{Tr}e^{\mathbf{i}t\mathbf{Y}_n\mathbf{Y}_n^T}] \le K(1+|t|^4)^2. \tag{1.2.3}$$

Nos travaux sont constitués de 3 étapes :

1. Adapter les techniques d'interpolation de Lytova et Pastur au modèle

$$\mathbf{Y}_n = \frac{1}{\sqrt{n}}\mathbf{R}_n^{1/2}\mathbf{X}_n + \mathbf{A}_n$$

   où les normes spectrales de deux matrices déterministes $\mathbf{A}_n$ et $\mathbf{R}_n$ sont supposées bornées. L'ajout de la matrice $\mathbf{A}_n$ ne présente pas de difficulté supplémentaires, aussi nous établissons la borne (1.2.3) pour ce modèle général (cf. Section 3.2.4).

   Dans le reste du chapitre, nous considérerons $\mathbf{A}_n = \mathbf{0}$. En prenant $\mathbf{R}_n = \mathbf{I}_n$, cette borne permettrait d'appliquer cette technique au modèle information-plus-bruit étudié au Chapitre 2.

2. Établir le CLT au modèle $\mathbf{Y}_n = \mathbf{R}_n^{1/2}\mathbf{X}_n$ avec des entrées réelles pour $\mathrm{Tr}f(\mathbf{Y}_n\mathbf{Y}_n^T)$. Par la représentation (1.2.1), l'estimation de $\mathrm{Tr}[e^{\mathbf{i}t\mathbf{Y}_n\mathbf{Y}_n^T}]$ aboutit à l'inégalité

$$\mathrm{Var}^{1/2}[\hat{f}(t)\mathrm{Tr}e^{\mathbf{i}t\mathbf{Y}_n\mathbf{Y}_n^T}] \le K|\hat{f}(t)|(1+|t|^4)$$

qui est par hypothèse, intégrable sur $\mathbb{R}$. Un théorème de transfert [50, Theorem 4.28] nous assure la convergence de l'intégrale (1.2.2) vers une gaussienne quand $K \to \infty$ (cf. Section 3.3).

3. Étudier le biais $\mathbb{E} \int f(\lambda) F^{\mathbf{Y}_n \mathbf{Y}_n^T}(d\lambda) - N \int f(\lambda) F_n(d\lambda)$ où F_n est la measure associée à l'équivalent déterministe $\mathbf{T}_n(z)$. Une approche directe du biais semble compliquée, aussi nous introduisons un intermédiare de calcul gaussien

$$\begin{aligned} &\mathbb{E} \int f(\lambda) F^{\mathbf{Y}_n \mathbf{Y}_n^T}(d\lambda) - N \int f(\lambda) F_n(d\lambda) \\ &= \left(\mathbb{E} \int f(\lambda) F^{\mathbf{Y}_n \mathbf{Y}_n^T}(d\lambda) - \mathbb{E} \int f(\lambda) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(d\lambda) \right) \\ &\quad + \left(\mathbb{E} \int f(\lambda) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(d\lambda) - N \int f(\lambda) F_n(d\lambda) \right), \end{aligned} \tag{1.2.4}$$

où $\hat{\mathbf{Y}}_c = \frac{1}{\sqrt{n}} \mathbf{R}_n^{1/2} \hat{\mathbf{X}}_c$ et $\hat{\mathbf{X}}_c$ est une matrice gaussienne complexe standard. Le premier terme se gère sans difficulté à l'aide de techniques d'interpolation. Pour le second terme, par un calcul gaussien développé dans [46, 93], il existe deux polynômes P_{12} et P_{17} tels que

$$|\mathbb{E}\mathrm{Tr}\mathbf{Q}_c(z) - \mathrm{Tr}\mathbf{T}(z)| \leq \frac{1}{n} P_{17}(|\mathrm{Im} z|^{-1}) P_{12}(|z|) \tag{1.2.5}$$

où $\mathbf{Q}_c$ est la résolvante associée à $\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*$ et P_k désigne un polynôme de coefficients strictement positifs de degré k. Pour une fonction régulière (par exemple, de classe C^{18}), en utilisant la représentation

$$N\mathbb{E} \int f(\lambda) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(d\lambda) - N \int f(\lambda) F_n(d\lambda) = \frac{1}{\pi} \lim_{y \downarrow 0} \int_{\mathbb{R}} f(x) \mathrm{Im}\{\mathrm{Tr}\mathbf{Q}_c(x+\mathbf{i}y) - \mathrm{Tr}\mathbf{T}(x+\mathbf{i}y)\} dx,$$

la méthode de Haagerup et Thorbjørnsen [38] assure que,

$$\left| N\mathbb{E} \int f(\lambda) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(d\lambda) - N \int f(\lambda) F_n(d\lambda) \right| \leq \frac{1}{n} \int |(1+D)^p f(x)| \frac{P_{12}(|x|)}{|\mathrm{Im} z|^{18-p}} dx,$$

où l'opérateur $D = \frac{d}{dx}$. Autrement dit, la régularité de f compense l'ordre de la singularité de $|\mathrm{Im} z|^{-1}$ près de l'axe réel (cf. Section 3.4).

En ce qui concerne l'étude des fluctuations de $\mathrm{Tr} f(\mathbf{Y}_n \mathbf{Y}_n^T) - \mathbb{E}\mathrm{Tr} f(\mathbf{Y}_n \mathbf{Y}_n^T)$, nos hypothèses sur f sont comparables à celles de Lytova et Pastur [60] et Shcherbina [82] sur plusieurs aspects. En effet, nous imposons

$$\int_{\mathbb{R}} |\hat{f}(t)|(1+|t|^4) dt < \infty,$$

alors que la condition dans Pastur et Lytova [60] est

$$\int_{\mathbb{R}} (1+|t|^5) |\hat{f}(t)| dt < \infty,$$

et celle dans Shcherbina [82] est

$$\int_{\mathbb{R}}(1+|t|)^{3+\varepsilon}|\hat{f}(t)|^2dt<\infty,$$

pour les entrées qui ont le moment d'ordre $4+\varepsilon$ fini.

En ce qui concerne l'étude du biais, l'hypothèse sur f est beaucoup plus forte (f de classe C^{18}). Cela est compatible avec les travaux de Haagerup et Thorbjørnsen [38] (cas de Wigner gaussien complexe, où f est de classe C^8) et ceux de Schultz [81] (cas de Wigner réel, où f est de classe C^{14}). Il semble que la structure de notre modèle nécessite une hypothèse plus forte. La différence est due au contrôle de la variable gaussienne (cf. (1.2.5)).

1.3 Estimation en grande dimension

Dans le cas où on traite des données de grande dimension, c'est-à-dire, lorsque la dimension des données est du même ordre que la taille de l'échantillon, les méthodes classiques comme par exemple l'approximation de la matrice de covariance des observations par la matrice de covariance empirique échouent au motif que la structure de bruit dans un tel objet de grande dimension est trop importante pour qu'une moyennisation la fasse disparaître. Dans ce cadre-là, la construction d'estimateurs consistants est délicate mais des tendances se dégagent :

- il est illusoire de vouloir estimer de manière consistante des objets de grande taille $M\times M$ tels que la matrice de covariance de notre modèle,
- il est possible d'estimer de manière consistante des fonctions de tels objets au motif que, moralement, le fait d'estimer une fonction d'une grande matrice et non la grande matrice elle-même opère une réduction drastique de la dimension de l'objet à estimer.

Deux ingrédients majeurs interviennent dans l'estimation consistante de fonctions de grandes matrices :

- la compréhension du comportement asymptotique de la fonction à estimer. Dans le cadre de la théorie des matrices aléatoires, cette fonction dépend en général d'une matrice aléatoire dont le comportement asymptotique est traduit par des équations déterministes associées au modèle étudié (cf. [36], et [44] pour les équations régissant le modèle de Rice corrélé par exemple). Cela a plusieurs conséquences pratiques importantes : la construction de l'estimateur d'une fonction donnée, par exemple, transformée de Stieltjes de la mesure spectrale d'une matrice de covariance donnée, dépendra de la fonction en question mais aussi des équations déterministes relatives au modèle matriciel considéré (ex : canal de Rayleigh, de Rice, etc.) ; autrement dit, la construction d'estimateurs se fait au cas par cas.
- le contrôle local du spectre de matrices aléatoires (propriétés de *séparation spectrale*) permettant d'assurer la consistance des estimateurs construits.

Pour certains types de canaux et certains modèle de grandes matrices aléatoires, les travaux théoriques [9, 10] ont précisé les constructions de séparation spectrale nécessaires pour la construction d'estimateurs consistants [65, 91, 90].

1.3.1 Présentation de la problématique dans un cas particulier

Nous allons dans ce paragraphe présenter un exemple de situation permettant de bien comprendre la nature des problèmes qu'il convient de résoudre. Considérons un signal aléatoire $(\mathbf{y}_j)_{j\in\mathbb{Z}}$ de dimension N supposé observé entre les instants 1 et n. On suppose que les vecteurs $(\mathbf{y}_j)_{j=1,\cdots,n}$ sont de moyenne nulle, indépendants et de même loi. Désignons par $\mathbf{R}_N$ la matrice de covariance des vecteurs $(y_j)_{j=1,\cdots,n}$.

Supposons que l'on souhaite estimer un paramètre scalaire dépendant de $\mathbf{R}_N$, par exemple $\frac{1}{N}\mathrm{Tr}(\mathbf{R}_N + x\mathbf{I}_N)^{-1}$ où $x \geq 0$. Ce problème est important car de nombreux autres estimateurs sont basés sur celui de $\frac{1}{N}\mathrm{Tr}(\mathbf{R}_N + x\mathbf{I}_N)^{-1}$. En absence d'hypothèse sur la loi de probabilité des vecteurs observés, l'usage consiste à estimer $\mathbf{R}_N$ par la matrice de covariance empirique $\hat{\mathbf{R}}_N$ définie par

$$\hat{\mathbf{R}}_N = \frac{1}{n}\sum_{j=1}^{n}\mathbf{y}_j\mathbf{y}_j^H$$

et d'estimer $\frac{1}{M}\mathrm{Tr}(\mathbf{R}_N+x\mathbf{I}_N)^{-1}$ par $\frac{1}{N}\mathrm{Tr}(\hat{\mathbf{R}}_N+x\mathbf{I}_N)^{-1}$. Cette pratique est essentiellement justifiée par le fait que si $n \to \infty$ et que N reste fixe, cet estimateur est consistant, c'est-à-dire que

$$\lim_{n\to\infty}\frac{1}{N}\mathrm{Tr}(\hat{\mathbf{R}}_N + x\mathbf{I}_N)^{-1} = \frac{1}{N}\mathrm{Tr}(\mathbf{R}_N + x\mathbf{I}_N)^{-1}.$$

Ceci provient évidemment du fait que $\|\hat{\mathbf{R}}_N - \mathbf{R}_N\|$ tend presque surement vers 0 dans ce régime (N fixé, $n \to \infty$). En pratique, ceci signifie que si n est beaucoup plus grand que N, alors il est raisonnable d'estimer $\frac{1}{N}\mathrm{Tr}(\mathbf{R}_N + x\mathbf{I}_N)^{-1}$ par $\frac{1}{N}\mathrm{Tr}(\hat{\mathbf{R}}_N + x\mathbf{I}_N)^{-1}$.

Il existe cependant des situations d'intérêt pratique dans lesquelles n, tout en restant supérieur à N, est du même ordre de grandeur que N. Afin d'évaluer la pertinence de l'estimateur dans ce cas, il est naturel d'étudier son comportement lorsque n et N tendent vers $+\infty$ de telle sorte que le rapport $\frac{N}{n} \to c$, où $0 < c < 1$. Dans ce contexte, la matrice $\hat{\mathbf{R}}_N$ est très loin de se comporter comme $\mathbf{R}_N$, et il n'y a donc aucune raison pour que $\frac{1}{N}\mathrm{Tr}(\hat{\mathbf{R}}_N + x\mathbf{I}_N)^{-1}$ soit un estimateur consistant de $\frac{1}{N}\mathrm{Tr}(\mathbf{R}_N + x\mathbf{I}_N)^{-1}$.

Les matrices aléatoires telles que $\hat{\mathbf{R}}_N$ ont fait l'objet de nombreux travaux. En particulier, il a été établi par que la distribution des valeurs propres de $\hat{\mathbf{R}}_N$ converge vers une distribution déterministe dépendant de $\mathbf{R}_N$ [83]. Afin d'aboutir à ce résultat, il convient d'étudier la résolvante de $\hat{\mathbf{R}}_N$ définie pour tout $x \geq 0$ comme la matrice $(\hat{\mathbf{R}}_N + x\mathbf{I}_N)^{-1}$.

En traitement statistique du signal, l'estimateur proposé par X. Mestre [65, 64] repose sur une intégrale de contours. Considérons le modèle $\mathbf{Y}_n = \frac{1}{\sqrt{n}}\mathbf{R}_n^{1/2}\mathbf{X}_n$ où la mesure spectrale de $\mathbf{R}_n$ est supposée

$$F^{\mathbf{R}_n} = \sum_{k=1}^{L}\frac{N_k}{N}\delta_{\lambda_k}, \quad L \text{ fixé et connu.}$$

La *condition de séparabilité* suppose que les valeurs propres λ_k sont suffisamment éloignées pour qu'il y ait L compacts disjoints dans la loi limite (voir Section 4.1.1 pour plus de détails). Sous cette condition, λ_k peut s'écrire comme

$$\lambda_k = -\frac{1}{2\mathbf{i}\pi N_k}\oint_{\mathcal{C}_k}\sum_{i=1}^{N}\frac{1}{\lambda_i - z}dz$$

où le contour $\mathcal{C}_k$ ne contient que λ_k. Par le changement de variable $z = -\frac{1}{\tilde{\delta}_n(z)}$, λ_k peut s'exprimer comme

$$\lambda_k = \frac{N}{2\mathbf{i}\pi N_k}\int_{\mathcal{R}_k} z\frac{\tilde{\delta}_n'(z)}{\tilde{\delta}_n(z)}dz$$

où le contour $\mathcal{C}_k$ ne contient que le k^e compact. L'estimateur consistant est construit par le remplacement de l'equivalent déterministe $\tilde{\delta}_n(z)$ par la trace normalisée de la résolvante associée $\tilde{m}_n(z) = \frac{1}{n}\mathrm{Tr}(\mathbf{Y}_n^*\mathbf{Y}_n - z\mathbf{I}_n)^{-1}$.

Suivant les travaux de Mestre, la première partie dans le Chapitre 4 établit un CLT pour son estimateur dans le cas gaussien. La méthode consiste à établir le CLT pour $\mathrm{Tr}(\mathbf{Y}_n^*\mathbf{Y}_n - z\mathbf{I}_n)^{-1}$ et sa dérivée $\mathrm{Tr}(\mathbf{Y}_n^*\mathbf{Y}_n - z\mathbf{I}_n)^{-2}$ par le théorème de Bai et Silverstein [11]. Un argument de tension permet de conclure le CLT. Nous proposons également un estimateur consistant de la variance. La seconde partie est consacrée à l'estimation sans condition de séparabilité par la méthode des moments et l'étude de son comportement.

Les points délicats sont :

1. Les propriétés sur le spectre de grandes matrices aléatoires sont en général faites pour la plus grande ou plus petite valeur propre. En revanche, il y a peu de résultats sur les positions de valeurs propres en dehors du contour $\mathcal{R}_k$. Dans [11], pour tout $k \in \mathbb{N}$, et μ_r (resp. μ_ℓ), plus grand (resp. petit) que la plus grande (resp. petite) valeur du support $\mathcal{S}_n$ de l'équivalent déterministe $\tilde{F}_n$, alors
$$\mathbb{P}(\lambda_{\max}^{\mathbf{Y}_n\mathbf{Y}_n^*} > \mu_r) = o(n^{-k}), \quad \mathbb{P}(\lambda_{\min}^{\mathbf{Y}_n\mathbf{Y}_n^*} < \mu_\ell) = o(n^{-k}).$$
Par l'inégalité de Nash-Poincaré, nous allons montrer que
$$\mathbb{P}(\max_i d(\lambda_i^{\mathbf{Y}_n\mathbf{Y}_n^*}, \mathcal{S}_n) > \varepsilon) = o(n^{-k}), \quad \text{pour } k \in \mathbb{N},$$
où $d(a, \mathcal{S}_n)$ est la distance euclidienne entre a et $\mathcal{S}_n$ (cf. Proposition 4.1.1).
2. L'estimation de la variance demande un calcul de résidu soigneux puisqu'il s'agit de calculer
$$\oint_{\mathcal{R}_k}\oint_{\mathcal{R}_\ell'}\left(\frac{\tilde{m}_n'(z_1)\tilde{m}_n'(z_2)}{(\tilde{m}_n(z_1)-\tilde{m}_n(z_2))^2} - \frac{1}{(z_1-z_2)^2}\right)\frac{1}{\tilde{m}_n(z_1)\tilde{m}_n(z_2)}d\,z_1 d\,z_2$$
où $\tilde{m}_n(z) = \frac{1}{n}\mathrm{Tr}(\mathbf{Y}_n^*\mathbf{Y}_n - z\mathbf{I}_n)^{-1}$ (cf. Section 4.1.5).
3. L'estimation par la méthode des moments pose en général un problème de simulation. Trouver un bon algorithme pour assurer la bonne performance de simulation reste un problème ouvert. (cf. Section 4.2.4)

1.4 Organisation du rapport

Le rapport est organisé de la façon suivante.

– le Chapitre 2 établit un CLT de statistiques linéaires pour les fonctionnelles analytiques par la méthode classique sur les martingales. Nous exprimerons la covariance sous une seule formule dans le cas réel ou complexe. Cela fait l'objet du papier

  J. Yao et J. Najim. "*Fluctuations for linear spectral statistics of large information-plus-noise type random matrices*", en préparation.

– Le Chapitre 3 étudit le CLT pour les fonctions non-analytiques en utilisant la transformée de Fourier. Nous préparons le papier

  J. Yao et J. Najim. "*Gaussian fluctuations for linear spectral statistics of large random covariance matrices*", en préparation.

– Dans le Chapitre 4, nous allons étudier les fluctuations de l'estimateur de Mestre dans le cas gaussien. Une application sur radio cognitive sera aussi proposée. La seconde partie du Chapitre 4 construit un estimateur consistant sans " condition de séparabilité". Les fluctuations seront également étudiées. La preuve mélange la méthode par l'intégrale de contours et la méthode des moments. C'est l'objet des articles

  J. Yao, R. Couillet, J. Najim, and M. Debbah. "*Fluctuations of an improved population eigenvalue estimator in sample covariance matrix models*", IEEE Trans. Inform. Theory, 59(2) :1149-1163, 2013.

  J. Yao, R. Couillet, J. Najim, E. Moulines, and M. Debbah. *CLT for eigeninference methods in cognitive radios.* In *Acoustics, Speech and Signal Processing* (ICASSP), 2011 IEEE International Conference on, pages 2980-2983, may 2011.

  J. Yao, A. Kammoun, et J. Najim. "*Eigenvalue estimation of parameterized covariance matrices of large dimensional data.*" IEEE Trans. Signal Process., 60(11) :5893-5905, 2012.

Chapitre 2

Fluctuations for linear spectral statistics of information-plus-noise type random matrices

This chapter is inspired from [71]. Consider an $N \times n$ random matrix $\mathbf{Y}_n = (y_{ij}^n)$ given by :

$$\mathbf{Y}_n = \frac{1}{\sqrt{n}} \mathbf{X}_n + \mathbf{A}_n \ , \tag{2.0.1}$$

where $\mathbf{X}_n$ is an $N \times n$ matrix whose entries $(x_{ij}^n\ ; i, j, n)$ are real or complex, i.i.d. with mean 0 and variance 1 ; matrix $\mathbf{A}_n$ has the same dimensions and is deterministic. Matrix $\mathbf{Y}_n$ is sometimes coined as "Information-plus-noise" type matrix in the literature.

The purpose of this chapter is to study the fluctuations of linear spectral statistics of the form :

$$\mathrm{Tr} f(\mathbf{Y}_n \mathbf{Y}_n^*) = \sum_{i=1}^{N} f(\lambda_i) \ , \tag{2.0.2}$$

where the λ_i's are the eigenvalues of $\mathbf{Y}_n \mathbf{Y}_n^*$, and f is an analytic function, under the regime where the dimensions n and $N = N(n)$ go to infinity at the same pace :

$$N, n \to \infty \quad \text{and} \quad 0 \ < \ \liminf \frac{N}{n} \ \leq \ \limsup \frac{N}{n} \ < \ \infty \ . \tag{2.0.3}$$

This condition will simply be referred to as $N, n \to \infty$ in the sequel.

2.1 Introduction

Resolvent, canonical equations and deterministic equivalents

Denote by $\mathbf{Q}_n(z)$ and $\tilde{\mathbf{Q}}_n(z)$ the resolvents of matrices $\mathbf{Y}_n \mathbf{Y}_n^*$ and $\mathbf{Y}_n^* \mathbf{Y}_n$:

$$\mathbf{Q}_n(z) = (\mathbf{Y}_n\mathbf{Y}_n^* - z\mathbf{I}_N)^{-1} \ , \qquad \tilde{\mathbf{Q}}_n(z) = (\mathbf{Y}_n^*\mathbf{Y}_n - z\mathbf{I}_n)^{-1} \ , \tag{2.1.1}$$

and by $m_{\mathbf{Y}_n\mathbf{Y}_n^*}(z)$ and $m_{\mathbf{Y}_n^*\mathbf{Y}_n}(z)$ their normalized trace, which are respectively the Stieltjes transforms of the empirical distribution of $\mathbf{Y}_n\mathbf{Y}_n^*$'s eigenvalues and of $\mathbf{Y}_n^*\mathbf{Y}_n$'s eigenvalues :

$$m_{\mathbf{Y}_n\mathbf{Y}_n^*}(z) = \frac{1}{N}\mathrm{Tr}\,\mathbf{Q}_n(z) \ , \quad m_{\mathbf{Y}_n^*\mathbf{Y}_n}(z) = \frac{1}{n}\mathrm{Tr}\,\tilde{\mathbf{Q}}_n(z) \ .$$

The following canonical equations admit a unique solution $(\delta, \tilde{\delta})$ in the class of Stieltjes transforms of nonnegative measures with support in $\mathbb{R}^+$ (see for instance [44, 39], see also [35, Section 7.11]) :

$$\left\{ \begin{array}{rcl} \delta_n(z) & = & \frac{1}{n}\mathrm{Tr}\left(-z(1+\tilde{\delta}_n)\mathbf{I}_N + \frac{\mathbf{A}_n\mathbf{A}_n^*}{1+\delta_n(z)}\right)^{-1} \\ \tilde{\delta}_n(z) & = & \frac{1}{n}\mathrm{Tr}\left(-z(1+\delta_n)\mathbf{I}_n + \frac{\mathbf{A}_n^*\mathbf{A}_n}{1+\tilde{\delta}_n(z)}\right)^{-1} \end{array} \right. . \tag{2.1.2}$$

Moreover, $\frac{n}{N}\delta_n$ and $\tilde{\delta}_n$ are Stieltjes transforms of probability measures, i.e. there exist probability measures $\mathcal{F}_n$ and $\tilde{\mathcal{F}}_n$ over $\mathbf{R}^+$ such that :

$$\frac{n}{N}\delta_n(z) = \int_{\mathbf{R}^+} \frac{\mathcal{F}_n(d\lambda)}{\lambda - z} \qquad \text{and} \qquad \tilde{\delta}_n(z) = \int_{\mathbf{R}^+} \frac{\tilde{\mathcal{F}}_n(d\lambda)}{\lambda - z} \ . \tag{2.1.3}$$

Functions δ and $\tilde{\delta}$ being introduced, we can now define the following $N \times N$ and $n \times n$ matrices :

$$\left\{ \begin{array}{l} \mathbf{T}_n(z) = \left(-z(1+\tilde{\delta}_n(z))\mathbf{I}_N + \frac{\mathbf{A}_n\mathbf{A}_n^*}{1+\delta_n(z)}\right)^{-1} \\ \tilde{\mathbf{T}}_n(z) = \left(-z(1+\delta_n(z))\mathbf{I}_n + \frac{\mathbf{A}_n^*\mathbf{A}_n}{1+\tilde{\delta}_n(z)}\right)^{-1} \end{array} \right. . \tag{2.1.4}$$

Matrices $\mathbf{T}_n$ and $\tilde{\mathbf{T}}_n$ can be thought of as deterministic equivalents of the resolvents $\mathbf{Q}_n$ and $\tilde{\mathbf{Q}}_n$ in the sense that they approximate the resolvents in various ways. For instance, for $z \in \mathbb{C}\backslash\mathbb{R}^+$:

$$\frac{1}{n}\mathrm{Tr}(\mathbf{Q}_n(z) - \mathbf{T}_n(z)) \xrightarrow[N,n\to\infty]{a.s.} 0 \quad \text{and} \quad \frac{1}{n}\mathrm{Tr}(\tilde{\mathbf{Q}}_n(z) - \tilde{\mathbf{T}}_n(z)) \xrightarrow[N,n\to\infty]{a.s.} 0 \ .$$

Interestingly, not only $\mathbf{T}_n$ and $\tilde{\mathbf{T}}_n$ convey information on the limiting spectrum of the resolvents but also on their eigenvectors. It has been proved indeed in [46] that

$$u_n^*\mathbf{Q}_n v_n - u_n^*\mathbf{T}_n v_n \xrightarrow[N,n\to\infty]{} 0$$

(and a symmetric result for $\tilde{\mathbf{Q}}_n$ and $\tilde{\mathbf{T}}_n$) where u_n and v_n are deterministic $N \times 1$ vectors with uniformly bounded euclidian norms (in N).

Representation of linear spectral statistics

We can now properly center the linear spectral statistics and introduce the main object of which we shall study the fluctuations :

$$L_n(f) = \sum_{i=1}^{N} f(\lambda_i) - N \int f(x)\ \mathcal{F}_n(dx)$$

as $N, n \to \infty$, where $\mathcal{F}_n$ is defined by (2.1.3). In the case where function f is analytic in a neighborhood of the limiting support of the spectrum of $\mathbf{Y}_n \mathbf{Y}_n^*$ (to be properly defined), Cauchy's integral formula yields (denote by Γ a contour surrounding this limiting spectrum on which f is analytic) :

$$L_n(f) = -\frac{1}{2i\pi} \oint_\Gamma f(z) \mathrm{Tr}(\mathbf{Q}_n(z) - \mathbf{T}_n(z))\, dz\ , \tag{2.1.5}$$

which follows from the identities :

$$\begin{aligned} \sum_{i=1}^{N} f(\lambda_i) &= -\frac{1}{2i\pi} \oint_\Gamma f(z) \mathrm{Tr}\, \mathbf{Q}_n(z) dz\ , \\ N \int f(x)\, \mathcal{F}_n(dx) &= -\frac{1}{2i\pi} \oint_\Gamma f(z) \mathrm{Tr}\, \mathbf{T}_n(z) dz\ , \end{aligned}$$

the last equality being an immediate consequence of (2.1.3), Cauchy's integral formula and Fubini's theorem.

Fluctuations and bias

It is clear from (2.1.5) that in order to describe the fluctuations of $L_n(f)$, a good approach is to study the fluctuations of the process $(\mathrm{Tr}(\mathbf{Q}_n(z) - \mathbf{T}_n(z))\ ; z \in \Gamma))$. In order to proceed, we split

$$M_n(z) \stackrel{\triangle}{=} \mathrm{Tr}\mathbf{Q}_n(z) - \mathrm{Tr}\mathbf{T}_n(z) \tag{2.1.6}$$

into two terms :

$$\begin{aligned} \mathrm{Tr}(\mathbf{Q}_n(z) - \mathbf{T}_n(z)) &= \underbrace{\mathrm{Tr}(\mathbf{Q}_n(z) - \mathbb{E}\, \mathbf{Q}_n(z))}_{\text{Fluctuations}} + \underbrace{\mathrm{Tr}(\mathbb{E}\, \mathbf{Q}_n(z) - \mathbf{T}_n(z))}_{\text{Bias}} \\ &\stackrel{\triangle}{=} M_n^1(z) + M_n^2(z)\ . \end{aligned}$$

The first term, which is centered, will give rise to the fluctuations of the linear spectral statistics and will be handled by martingale techniques. This strategy has been successfully applied in [11, 73, 45, 53, 39, 70, 12] and will be followed. The second term is completely deterministic and will give rise to a bias.

Entries with non-null fourth cumulant and a family of gaussian random variables

It is well known since the paper by Khorunzhiy et al. [55] that if the fourth moment of the entries differs from the fourth moment of a Gaussian random variable, then a term appears in the variance of the trace of the resolvent, which is proportional to the fourth cumulant κ of the entries :

$$\kappa = \mathbb{E}\,|x_{11}^n|^4 - |\vartheta|^2 - 2 \quad \text{where} \quad |\vartheta|^2 = |\mathbb{E}(x_{11}^n)^2|^2 \ .$$

The same phenomenon will occur here but the convergence of this additional term may fail to happen under usual assumptions (such as the convergence of $F^{\mathbf{A}_n\mathbf{A}_n^*}$ toward [1] a probability measure when $N, n \to \infty$ and the convergence of the ratio Nn^{-1} toward some positive constant c).

As will appear later, the reason of this lack of convergence lies in the fact that this additional term not only depends upon the spectrum of $\mathbf{A}_n\mathbf{A}_n^*$ but also on its eigenvectors. In order to avoid cumbersome assumptions enforcing the joint convergence of $\mathbf{A}_n\mathbf{A}_n^*$'s spectrum and eigenvalues, we shall express our fluctuation results in the same way as in [70] and prove that the distribution of the linear statistics $L_n(f)$ becomes close to a family of Gaussian distributions, whose parameters (mean and variance) may not converge. Namely, we shall establish that there exists a Gaussian random variable $\mathcal{N}(\mathcal{B}_n, \Theta_n)$ such that :

$$d_{\mathcal{L}P}\left(L_n(f), \mathcal{N}(\mathcal{B}_n, \Theta_n)\right) \xrightarrow[N,n\to\infty]{} 0 \ , \tag{2.1.7}$$

where $d_{\mathcal{L}P}$ denotes the L¨|vy-Prohorov distance (and in particular metrizes the convergence of laws).

Outline of the chapter

In Section 2.2, we state the main results of the chapter ; in Section 2.3, we establish the proof of the CLT (Theorem 2.2.1) except the identification of the covariance of the normalized trace, which is postponed to Section 2.4. In Section 2.5, we compute the limiting bias and prove Theorem 2.2.2.

2.2 Statement of the Central Limit Theorem

We state below the main assumptions of the chapter. Recall the fact that $N = N(n)$ and the asymptotic regime (2.0.3) where $N, n \to \infty$ and denote by

$$c_n \triangleq \frac{N}{n} \ , \quad c_- \triangleq \liminf \frac{N}{n} \quad \text{and} \quad c_+ \triangleq \limsup \frac{N}{n} \ .$$

1. Here $F^{\mathbf{M}}$ denotes the empirical distribution of matrix $\mathbf{M}$'s eigenvalues.

Assumption 2.2.1 *The random variables $(x_{ij}^n\ ;\ 1 \le i \le N(n), 1 \le j \le n, n \ge 1)$ are real or complex, independent and identically distributed (i.i.d.). They satisfy*

$$\mathbb{E}x_{ij}^n = 0\ , \qquad \mathbb{E}|x_{ij}^n|^2 = 1 \quad \text{and} \quad \mathbb{E}|x_{ij}^n|^4 < \infty\ .$$

Associated to these moments are the quantities :

$$\vartheta = \mathbb{E}(x_{ij}^2) \quad \text{and} \quad \kappa = \mathbb{E}|x_{ij}|^4 - |\vartheta|^2 - 2\ .$$

Notice that in the case where the x_{ij}'s are real, $\vartheta = 1$; in the case where the random variables are complex with decorrelated real and imaginary part with equal variance, $\vartheta = 0$.

Assumption 2.2.2 *The family of deterministic $N \times n$ complex matrices $(\mathbf{A}_n)$ is bounded for the spectral norm :*

$$a_{\max} \stackrel{\triangle}{=} \sup_{n \ge 1} \|\mathbf{A}_n\| < \infty.$$

Assumption 2.2.3 *Function $f : \mathbf{R} \to \mathbf{R}$ is analytic on an open region containing $[0, \ell_+]$ where*

$$\ell_+ = 2(1 + \sqrt{c_+})^2 + 2a_{\max}\ .$$

Remark 2.2.1 *Since $\mathbf{Y}_n\mathbf{Y}_n^* \le 2(n^{-1}\mathbf{X}_n\mathbf{X}_n^* + \mathbf{A}_n\mathbf{A}_n^*)$, the largest eigenvalue of $\mathbf{Y}_n\mathbf{Y}_n^*$ will eventually lie below ℓ_+ and representation (2.1.5) will eventually hold true for a contour Γ containing $[0, \ell_+]$. Assumption 2.2.3 may be slightly strengthened, but in this case estimates over the extreme eigenvalues of $\mathbf{Y}_n\mathbf{Y}_n^*$ (see for instance Lemma 2.3.4 below) would be more challenging to obtain.*

Denote by $d_{\mathcal{LP}}$ the Lévy-Prohorov distance between two probability measures P, Q defined as :

$$d_{\mathcal{LP}}(P,Q) = \inf\{\varepsilon > 0 : P(A) \le Q(A^\varepsilon) + \varepsilon \quad \text{for all Borel sets } A \subset \mathbb{R}^d\},$$

where A^ε is an ε-blow up of A (see [18, Chapter 1, Section 6] for more details). If X and Y are random variables with distributions $\mathcal{L}(X)$ and $\mathcal{L}(Y)$, we simply write (with a slight abuse of notations) $d_{\mathcal{LP}}(X,Y)$ instead of $d_{\mathcal{LP}}(\mathcal{L}(X), \mathcal{L}(Y))$. It is well-known that the Lévy-Prohorov distance metrizes the convergence in distribution (see for instance [31, Chapter 11]).

In the sequel, we shall often drop subscripts and superscripts n and for example write $\mathbf{Q}$ instead of $\mathbf{Q}_n$ and x_{ij} instead of x_{ij}^n, also we shall further simplify and denote :

$$\delta_{1,2} = \delta(z_{1,2})\ , \quad \text{and} \quad \mathbf{T}_{1,2} = \mathbf{T}(z_{1,2})\ ,$$

where δ and $\mathbf{T}$ are respectively defined in (2.1.2) and (2.1.4).

We now introduce a number of quantities that will help in expressing the covariance and bias in the CLT.

$$\begin{aligned} \gamma(z_1,z_2) &= \frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{T}_2\ , \quad \tilde{\gamma}(z_1,z_2) = \frac{1}{n}\mathrm{Tr}\tilde{\mathbf{T}}_1\tilde{\mathbf{T}}_2\ , \\ \gamma^\dagger(z_1,z_2) &= \frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{T}_2^T\ , \quad \tilde{\gamma}^\dagger(z_1,z_2) = \frac{1}{n}\mathrm{Tr}\tilde{\mathbf{T}}_1\tilde{\mathbf{T}}_2^T\ , \end{aligned} \tag{2.2.1}$$

In order to express the variance and bias, we need the following notations : For $\mathbf{M} = \mathbf{I}$ or $\mathbf{T}$,

$$\begin{aligned}
\nu^{\dagger}_{\mathbf{M}}(z_1,z_2) &= \frac{1}{n}\frac{\mathrm{Tr}\mathbf{A}^*\mathbf{T}_1\mathbf{M}\mathbf{T}_2^T\bar{\mathbf{A}}}{(1+\delta_1)(1+\delta_2)}, &(2.2.2)\\
\tilde{\nu}^{\dagger}_{\mathbf{M}}(z_1,z_2) &= \frac{1}{n}\frac{\mathrm{Tr}\mathbf{A}^T\mathbf{T}_1\mathbf{M}\mathbf{T}_2^T\mathbf{A}}{(1+\delta_1)(1+\delta_2)}, &(2.2.3)\\
\gamma^{\dagger}_{\mathbf{M}}(z_1,z_2) &= \frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{M}\mathbf{T}_2^T &(2.2.4)\\
\eta(z_1,z_2) &= \frac{1}{n}\sum_{j=1}^{n} z_1z_2\tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2) &(2.2.5)\\
\omega^{\dagger}(z_1,z_2) &= \frac{1}{n}\sum_{k=1}^{n}\sum_{\ell=1,\ell\neq k}^{n}\frac{\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^T\mathbf{T}_2^T\bar{\mathbf{a}}_k}{(1+\delta_1)^2(1+\delta_2)^2}. &(2.2.6)
\end{aligned}$$

As we shall see, $\mathbf{M}$ will be equal to the identity matrix $\mathbf{I}$ in the computation of the covariance, and we will simply denote

$$\begin{aligned}
\nu^{\dagger}_{\mathbf{I}} &= \nu^{\dagger}\ ,\\
\tilde{\nu}^{\dagger}_{\mathbf{I}} &= \tilde{\nu}^{\dagger}\ ,\\
\gamma^{\dagger}_{\mathbf{I}} &= \gamma^{\dagger}\ .
\end{aligned}$$

In the computation of the bias where $z_1 = z_2 = z$, $\mathbf{M} = \mathbf{T}(z)$.

Remark 2.2.2 *Notice that notation $\dagger$ refers to matrix products involving $\mathbf{T}$ and $\mathbf{T}^T$. In the case where matrix $\mathbf{A}$ is real, $\nu^{\dagger}_{\mathbf{M}} = \tilde{\nu}^{\dagger}_{\mathbf{M}}$; however, these terms differ if $\mathbf{A}$ has complex entries.*

Define :

$$\begin{aligned}
\Delta_n(z_1,z_2) &= \left(1 - \frac{\mathrm{Tr}\mathbf{T}_1\mathbf{A}\mathbf{A}^*\mathbf{T}_2}{n(1+\delta_1)(1+\delta_2)}\right)^2 - z_1z_2\gamma\tilde{\gamma}, &(2.2.7)\\
\Delta_n^{\vartheta}(z_1,z_2) &= \left(1-\vartheta\nu^{\dagger}\right)\left(1-\bar{\vartheta}\tilde{\nu}^{\dagger}\right) - |\vartheta|^2 z_1z_2\gamma^{\dagger}\tilde{\gamma}^{\dagger}\ , &(2.2.8)\\
\Upsilon_n(z) &= \left(1 - \frac{z}{n}\mathrm{Tr}\mathbf{T}^2 - \frac{1}{n}\frac{\mathrm{Tr}\mathbf{T}\mathbf{A}\mathbf{A}^*\mathbf{T}}{(1+\delta)^2}\right)^{-1}\ . &(2.2.9)
\end{aligned}$$

Remark 2.2.3 *From (2.1.4),*

$$\begin{cases}
\mathbf{T}^T(z) = \left(-z\left(1+\tilde{\delta}(z)\right)\mathbf{I}_N + \frac{\bar{\mathbf{A}}\bar{\mathbf{A}}^*}{1+\delta(z)}\right)^{-1}\\
\tilde{\mathbf{T}}^T(z) = \left(-z\left(1+\delta(z)\right)\mathbf{I}_n + \frac{\bar{\mathbf{A}}^*\bar{\mathbf{A}}}{1+\tilde{\delta}(z)}\right)^{-1}.
\end{cases}$$

Notice that $\mathbf{T}^T(z)$ and $\tilde{\mathbf{T}}^T(z)$ differ from $\bar{\mathbf{T}}(z)$ and $\bar{\tilde{\mathbf{T}}}(z)$ due to the presence of z. However, the following holds true :

$$\mathbf{T}^*(z) = \mathbf{T}(\bar{z}) = \left(-\bar{z}\left(1+\tilde{\delta}(\bar{z})\right)\mathbf{I}_N + \frac{\mathbf{A}\mathbf{A}^*}{1+\delta(\bar{z})}\right)^{-1}.$$

2.2.1 Statement of the Central Limit Theorem

Recall that $L_n(f)$ writes :

$$\begin{aligned} L_n(f) &= -\frac{1}{2i\pi}\oint_\Gamma f(z)M_n(z)dz \\ &= -\frac{1}{2i\pi}\oint_\Gamma f(z)M_n^1(z)dz - \frac{1}{2i\pi}\oint_\Gamma f(z)M_n^2(z)dz \\ &\triangleq L_n^1(f) + L_n^2(f)\ . \end{aligned}$$

The first term above accounts for the fluctuations and is handled in Theorem 2.2.1 while the second one yields a bias and is handled in Theorem 2.2.2.

Theorem 2.2.1 *Let Assumptions 2.2.1 and 2.2.2 hold true ; let k be a given integer and $\mathbf{f} = (f_i\ ;\ 1 \le i \le k)$ where the f_i's are functions satisfying Assumption 2.2.3. Consider the random vector*

$$L_n^1(\mathbf{f}) \triangleq \left(L_n^1(f_1), \cdots, L_n^1(f_k)\right)\ .$$

Then

$$d_{\mathcal{LP}}(L_n^1(\mathbf{f}), \mathcal{N}_n^1(\mathbf{f})) \xrightarrow[N,n\to\infty]{} 0,$$

where $\mathcal{N}_n^1(\mathbf{f})$ is a centered Gaussian vector $\mathcal{N}_k(\mathbf{0}, \mathbf{V})$ with covariance matrix $\mathbf{V} = (V_{ij}\ ;\ 1 \le i, j \le k)$ defined as

$$V_{ij} = -\frac{1}{4\pi^2}\oint_{\Gamma_1}\oint_{\Gamma_2} f_i(z_1)f_j(z_2)\Theta_n(z_1,z_2)dz_1dz_2, \tag{2.2.10}$$

where

$$\Theta_n(z_1,z_2) \triangleq \Theta_{0,n}(z_1,z_2) + \Theta_{1,n}(\vartheta, z_1, z_2) + \Theta_{2,n}(\kappa, z_1, z_2)\ , \tag{2.2.11}$$

with

$$\Theta_{0,n}(z_1,z_2) = \frac{s_n'(z_1)s_n'(z_2)}{(s_n(z_1)-s_n(z_2))^2} - \frac{1}{(z_1-z_2)^2}, \tag{2.2.12}$$

$$\Theta_{1,n}(\vartheta, z_1, z_2) = -\frac{\partial}{\partial z_2}\left(\frac{\partial \Delta_n^\vartheta}{\partial z_1}\frac{1}{\Delta_n^\vartheta}\right), \tag{2.2.13}$$

$$\Theta_{2,n}(\kappa, z_1, z_2) = \kappa\frac{\partial^2}{\partial z_1 \partial z_2}\left\{\frac{z_1z_2}{n^2}\sum_{i=1}^N t_{ii}(z_1)t_{ii}(z_2)\sum_{j=1}^n \tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2)\right\}, \tag{2.2.14}$$

and where Δ_n^ϑ is defined in (2.2.8) and where s_n is defined as :

$$s_n(z) = z(1+\tilde{\delta}_n)(1+\delta_n)\ , \tag{2.2.15}$$

moreover, the contours Γ_1, Γ_2 are non-overlapping, taken in the positive direction, and each enclosing $[0, \ell_+]$, where ℓ_+ is defined in Assumption 2.2.3.

In the case where the random variables x_{ij}^n's and matrix $\mathbf{A}$ are real (in particular $\vartheta = 1$), then $\Theta_{0,n} = \Theta_{1,n}$.

Proof of Theorem 2.2.1 closely follows the ideas in Bai and Silverstein [11] and in Najim [70], and is postponed to Section 2.3, except for the computation of the covariance. This part of the proof is postponed to Section 2.4, and the computations therein are strongly inspired by those in Hachem et al. [39].

Theorem 2.2.2 *Let Assumptions 2.2.1 and 2.2.2 hold true ; assume that f is a function satisfying Assumption 2.2.3. Then*

$$L_n^2(f) - B(f) \xrightarrow[N,n\to\infty]{} 0\ ,$$

where

$$B(f) \triangleq -\frac{1}{2i\pi}\oint_\Gamma f(z)B(z)\,dz$$

with Γ a contour taken in the positive direction and enclosing $[0,\ell_+]$ (ℓ_+ being defined in Assumption 2.2.3) and

$$\begin{aligned} B(z) \quad &= \quad \frac{\Upsilon_n}{\Delta_n^\vartheta}\Big[|\vartheta|^2\gamma_{\mathbf{T}}^\dagger\tilde{\gamma}^\dagger + \vartheta\nu_{\mathbf{T}}^\dagger + \bar{\vartheta}\tilde{\nu}_{\mathbf{T}}^\dagger - |\vartheta|^2(\nu_{\mathbf{T}}^\dagger\tilde{\nu}^\dagger + \tilde{\nu}_{\mathbf{T}}^\dagger\nu^\dagger) \\ &\qquad + |\vartheta|^2\gamma_{\mathbf{T}}^\dagger\omega^\dagger + \bar{\vartheta}|\vartheta|^2\eta(\tilde{\nu}^\dagger\gamma_{\mathbf{T}}^\dagger - \tilde{\nu}_{\mathbf{T}}^\dagger\gamma^\dagger)\Big] \\ &\qquad + \kappa\frac{z^2\Upsilon_n}{n^2}\sum_{j=1}^n \tilde{t}_{jj}^2\sum_{i=1}^N t_{ii}[T^2]_{ii}\ , \end{aligned} \tag{2.2.16}$$

where all the relevant quantities are evaluated at $(z_1, z_2) = (z, z)$.

2.2.2 Specialization to the diagonal and Gaussian cases

Three cases deserve a specific statement : The (pseudo-)diagonal case, where $a_{ij} = 0$ outside the diagonal, the real gaussian case and the complex gaussian case. Due to the orthogonal (resp. unitary) invariance of matrices with real (resp. complex) i.i.d. gaussian entries, these three cases are strongly connected, as we shall see.

The case where matrix A is pseudo-diagonal

When $\mathbf{A}$ is pseudo-diagonal, *i.e.* $a_{ij} = 0$ for $i \neq j$, $\mathbf{T}$ is a diagonal matrix and $\mathbf{T}^T = \mathbf{T}$. We notice that $\mathbf{a}_k^*\mathbf{T}\mathbf{a}_j = 0$ when $k \neq j$, which implies that $\omega^\dagger = 0$. Moreover, as all quantities depending on $\mathbf{T}$ and $\mathbf{A}$ are diagonal, we have :

$$\nu_{\mathbf{T}}^\dagger \;=\; \frac{1}{n}\frac{\mathrm{Tr}\mathbf{A}^*\mathbf{T}^3\bar{\mathbf{A}}}{(1+\delta)^2}\ ,\quad \tilde{\nu}_{\mathbf{T}}^\dagger \;=\; \frac{1}{n}\frac{\mathrm{Tr}\mathbf{A}^T\mathbf{T}^3\mathbf{A}}{(1+\delta)^2}\ ,\quad \gamma_{\mathbf{T}}^\dagger \;=\; \frac{1}{n}\mathrm{Tr}\mathbf{T}^3\ .$$

In this case, the bias writes :

$$
\begin{aligned}
B(z) =& \frac{\Upsilon_n}{\Delta_n^{\vartheta}} \Bigg[\frac{z^2|\vartheta|^2}{n^2} \mathrm{Tr}\mathbf{T}^3 \mathrm{Tr}\tilde{\mathbf{T}}^2 + \frac{\vartheta}{n} \frac{\mathrm{Tr}\mathbf{A}^*\mathbf{T}^3\bar{\mathbf{A}}}{(1+\delta)^2} + \frac{\bar{\vartheta}}{n} \frac{\mathrm{Tr}\mathbf{A}^T\mathbf{T}^3\mathbf{A}}{(1+\delta)^2} \\
&- \frac{|\vartheta|^2}{n^2(1+\delta)^4} \left(\mathrm{Tr}(\mathbf{A}^*\mathbf{T}^3\bar{\mathbf{A}})\mathrm{Tr}(\mathbf{A}^T\mathbf{T}^2\mathbf{A}) + \mathrm{Tr}(\mathbf{A}^T\mathbf{T}^3\mathbf{A})\mathrm{Tr}(\mathbf{A}^T\bar{\mathbf{T}}^2\mathbf{A}) \right) \\
&+ \frac{\bar{\vartheta}|\vartheta|^2\eta}{n^2(1+\delta)^2} \left(\mathrm{Tr}(\mathbf{A}^T\mathbf{T}^3\mathbf{A})\mathrm{Tr}(\mathbf{T}^2) - \mathrm{Tr}(\mathbf{A}^T\mathbf{T}^2\mathbf{A})\mathrm{Tr}(\mathbf{T}^3) \right) \Bigg] \\
&+ \frac{\kappa \Upsilon_n z^2}{n^2} \mathrm{Tr}\tilde{\mathbf{T}}^2 \mathrm{Tr}\mathbf{T}^3 .
\end{aligned}
\tag{2.2.17}
$$

The case where matrix X is real gaussian and matrix A is real

In this case

$$\vartheta = 1 \, , \quad \kappa = 0 \, , \quad \bar{\mathbf{A}} = \mathbf{A} \, , \quad \mathbf{A}^* = \mathbf{A}^T \quad \text{and} \quad \mathbf{T}^T = \mathbf{T} \, .$$

This yields minor simplifications in the bias formula (2.2.17). More important simplifications occur for the covariance since $\Theta_{0,n} = \Theta_{1,n}$ and $\Theta_{2,n} = 0$. Equation (2.2.11) writes

$$V_{ij} = -\frac{1}{2\pi^2} \oint \oint f_i(z_1) f_j(z_2) \frac{s_n'(z_1)s_n'(z_2)}{(s_n(z_1) - s_n(z_2))^2} dz_1 dz_2 .$$

The case where matrix X is complex gaussian

As $\vartheta = 0$ and $\kappa = 0$, the bias is zero and the variance is

$$V_{ij} = -\frac{1}{4\pi^2} \oint \oint f_i(z_1) f_j(z_2) \frac{s_n'(z_1)s_n'(z_2)}{(s_n(z_1) - s_n(z_2))^2} dz_1 dz_2 .$$

This is coherent with several papers. In complex gaussian case, it is well known that the bias is zero, (cf. [93, 32]) while the variance has the same form as [11].

2.2.3 Further computations

We further specialize the previous results and consider the case where $\mathbf{A}_n = 0$. A natural question is then to compare the results provided in Theorems 2.2.1 and 2.2.2 with those in Bai and Silverstein [11] and in Najim [70], in the case where the considered population matrix is equal to the identity.

Comparison of the covariance

In the case where x_{ij} are i.i.d. complex random variables with $\mathbf{A} = \mathbf{0}$, $\kappa = 0$ and $\vartheta = 0$, i.e. Marčenko-Pastur case (cf. [62]), we have $\tilde{\delta}_n(1 + \delta_n) = -\frac{1}{z}$. Then

$$s_n(z) = -\frac{1 + \tilde{\delta}_n(z)}{\tilde{\delta}_n(z)} ,$$

and

$$\frac{s_n'(z_1)s_n'(z_2)}{(s_n(z_1)-s_n(z_2))^2} = \frac{\tilde\delta_n'(z_1)\tilde\delta_n'(z_2)}{\tilde\delta_n^2(z_1)\tilde\delta_n^2(z_2)\left(\frac{1+\tilde\delta_n(z_1)}{\tilde\delta_n(z_1)}-\frac{1+\tilde\delta_n(z_2)}{\tilde\delta_n(z_2)}\right)^2} = \frac{\tilde\delta_n'(z_1)\tilde\delta_n'(z_2)}{(\tilde\delta_n(z_1)-\tilde\delta_n(z_2))^2}$$

which coincides with the formula in [11].

Comparison of the bias

When $\mathbf{A}=0$, we will identify the bias (2.2.16) with the expression in Najim [70, Eq. (2.20), (2.21)] . Notice that in this case

$$\mathbf{T} = -(z(1+\tilde\delta))^{-1}\mathbf{I}_N, \quad \tilde{\mathbf{T}} = -(z(1+\delta))^{-1}\mathbf{I}_n.$$

Taking trace and dividing by n, we have

$$\delta = -c_n(z(1+\tilde\delta))^{-1}, \quad \tilde\delta = -(z(1+\delta))^{-1}. \tag{2.2.18}$$

Injecting these terms into [70, Eq. (2.20), (2.21)], the bias is

$$\begin{aligned}\mathcal{B}^1(\vartheta,z) &= \frac{|\vartheta|^2c_n\tilde\delta^3}{(1+\tilde\delta)^3\left(1-c_n\frac{\tilde\delta^2}{(1+\tilde\delta)^2}\right)\left(1-c_n|\vartheta|^2\frac{\tilde\delta^2}{(1+\tilde\delta)^2}\right)}\\ \mathcal{B}^2(\kappa,z) &= \frac{\kappa c_n\tilde\delta^3}{(1+\tilde\delta)^3\left(1-c_n\frac{\tilde\delta^2}{(1+\tilde\delta)^2}\right)}.\end{aligned}$$

With Theorem 2.2.2 , since for $\mathbf{M}=\mathbf{I}$ or $\mathbf{T}$,

$$\nu_{\mathbf{M}}^\dagger = \tilde\nu_{\mathbf{M}}^\dagger = 0, \quad \Upsilon_n = \left(1-c_n\frac{1}{z(1+\tilde\delta)^2}\right)^{-1}, \quad \Delta_n^\vartheta = 1-c_n|\vartheta|^2\frac{\tilde\delta^2}{(1+\tilde\delta)^2},$$

we have

$$\begin{aligned}B^1(\vartheta,z) &= -\frac{c_n|\vartheta|^2\tilde\delta^2}{z(1+\tilde\delta)^3\left(1-c_n\frac{1}{z(1+\tilde\delta)^2}\right)\left(1-c_n|\vartheta|^2\frac{\tilde\delta^2}{(1+\tilde\delta)^2}\right)}\\ B^2(\kappa,z) &= -\frac{\kappa c_n\tilde\delta^2}{z(1+\tilde\delta)^3\left(1-c_n\frac{1}{z(1+\tilde\delta)^2}\right)}.\end{aligned}$$

To show $\mathcal{B}^1(\vartheta,z)=B^1(\vartheta,z)$, with comparison, it suffices to show that

$$-z\tilde\delta\left(1-c_n\frac{\tilde\delta^2}{(1+\tilde\delta)^2}\right)^{-1} = \left(1-c_n\frac{1}{z(1+\tilde\delta)^2}\right)^{-1},$$

which is equivalent to show

$$-z\tilde{\delta}\left(1-c_n\frac{1}{z(1+\tilde{\delta})^2}\right)=1-c_n\frac{\tilde{\delta}^2}{(1+\tilde{\delta})^2}.$$

From (2.2.18), we have

$$1-c_n\frac{\tilde{\delta}^2}{(1+\tilde{\delta})^2}+z\tilde{\delta}\left(1-c_n\frac{1}{z(1+\tilde{\delta})^2}\right)=-z\delta\tilde{\delta}-\frac{c_n\tilde{\delta}}{1+\tilde{\delta}}\quad=0.$$

This proves $\mathcal{B}^1(\vartheta,z)=B^1(\vartheta,z)$. The same method yields also that $\mathcal{B}^2(\kappa,z)=B^2(\kappa,z)$.

2.3 Proof of Theorem 2.2.1 (I)

2.3.1 Outline of the proof

Recall that

$$\begin{aligned}L_n^1(f) &= -\frac{1}{2i\pi}\oint_\Gamma f(z)\left(\mathrm{Tr}\mathbf{Q}_n(z)-\mathrm{Tr}\mathbb{E}\mathbf{Q}_n(z)\right)dz\\ &= -\frac{1}{2i\pi}\oint_\Gamma f(z)M_n^1(z)\,dz\end{aligned}$$

where Γ is a positively oriented contour enclosing the limiting spectrum on which f is analytic. In order to establish the fluctuations of the term $L_n^1(f)$, we closely follow and adapt the strategy in Bai and Silverstein [11]. We first establish the gaussian fluctuations of the process (M_n^1) on the contour Γ and then prove that (M_n^1) is tight over Γ. The gaussian fluctuations of $M_n^1(z)$ are established via the following central limit theorem for martingales, which is a variation on [18, Theorem 35.12] (see also Lemmas 4.7 and 4.8 in [70]) :

Theorem 2.3.1 *Suppose that for each n $(Y_{nj};1\le j\le r_n)$ is a $\mathbb{C}^d$-valued martingale difference sequence with respect to the increasing σ-field $\{\mathcal{F}_{n,j};1\le j\le r_n\}$ having second moments. Write :*

$$Y_{nj}^T=(Y_{nj}^1,\cdots,Y_{nj}^d)\ .$$

Assume moreover that $(\Theta_n(k,\ell))_n$ and $(\tilde{\Theta}_n(k,\ell))_n$ are uniformly bounded sequences of complex numbers, for $1\le k,\ell\le d$. If

$$\sum_{j=1}^{r_n}\mathbb{E}\left(Y_{nj}^k\bar{Y}_{nj}^d\mid\mathcal{F}_{n,j-1}\right)-\Theta_n(k,\ell)\quad\xrightarrow[n\to\infty]{\mathcal{P}}\quad 0\ ,\tag{2.3.1}$$

$$\sum_{j=1}^{r_n}\mathbb{E}\left(Y_{nj}^kY_{nj}^\ell\mid\mathcal{F}_{n,j-1}\right)-\tilde{\Theta}_n(k,\ell)\quad\xrightarrow[n\to\infty]{\mathcal{P}}\quad 0\ ,\tag{2.3.2}$$

and for each $\varepsilon > 0$, the following Lyapunov condition holds true :

$$\sum_{j=1}^{r_n} \mathbb{E}\left(\|Y_{nj}\|^2 \mathbb{I}_{\|Y_{nj}\|>\varepsilon} \right) \xrightarrow[n\to\infty]{} 0 \ . \tag{2.3.3}$$

Then

$$d_{\mathcal{L}P}\left(\sum_{j=1}^{r_n} Y_{nj} \ , Z_n \right) \xrightarrow[n\to\infty]{} 0 \ ,$$

where Z_n is a $\mathbb{C}^d$-valued centered Gaussian random vector with parameters

$$\mathbb{E} Z_n Z_n^* = (\Theta_n(k,\ell))_{k,\ell} \qquad \textit{and} \qquad \mathbb{E} Z_n Z_n^T = (\tilde{\Theta}_n(k,\ell))_{k,\ell} \ .$$

In view of Theorem 2.3.1, we shall study the fluctuations of

$$\left(M_n^1(z_1), \cdots, M_n^1(z_d) \right)$$

with $(z_1, \cdots, z_d) \in \mathbb{C}^+$. Here are the different steps of the proof :

1. We first perform in Section 2.3.2 a double truncation step : we truncate the random variables (x_{ij})'s and modify the process (M_n) near the real axis where it may blow up.
2. In Section 2.3.4, we verify Lyapunov's condition for $(M_n^1(z_i))_{1\le i\le d}$.
3. Convergences (2.3.1) and (2.3.2) for $(M_n^1(z_i))$ are demanding, and are postponed to Section 2.4.
4. The tightness of the process (M_n^1) over a given contour is established in Section 2.3.5.
5. In order to establish the tightness of the gaussian process $(\mathcal{N}_n(z), ; z \in \Gamma)$, we introduce a meta matrix model in Section 2.3.6.

2.3.2 Truncation

We closely follow Bai and Silverstein [11] and proceed with two truncation steps. In the first step the entries are truncated at an appropriate level, then centered and normalized, with no effect on the fluctuations of the linear spectral statistics. Truncation of the entries essentially enables to deal with entries having a fourth moment finite, which is the optimal assumption for our results to hold. The second truncation step is concerned with the truncation of the process associated to the normalized resolvent near the real axis. Since

$$\mathrm{Tr}\, \mathbf{Q}_n(z) = \sum_{i=1}^{N} \frac{1}{\lambda_i - z}$$

may blow up on the real axis, it is replaced by a process that coincides with it except on a thin strip covering the real axis. If the strip is thin enough, there are no losses at a fluctuation level.

This is packaged in the following results whose proofs are deferred to Appendix 2.6.1.

Proposition 2.3.1 (Truncation of the entries) *Let Assumptions 2.2.1 and 2.2.2 hold true and let f satisfy Assumption 2.2.3, then there exists a sequence $\eta_n \downarrow 0$ with $\eta_n n^{1/5} \to \infty$ such that if $\tilde{\mathbf{X}}_n$ is the $N \times n$ matrix with centered and normalized entries*

$$\tilde{x}_{ij} = \frac{\hat{x}_{ij} - \mathbb{E}\hat{x}_{ij}}{\sqrt{\mathbb{E}\,|\hat{x}_{ij} - \mathbb{E}\hat{x}_{ij}|^2}} \ , \qquad \hat{x}_{ij} = x_{ij}\mathbb{I}_{\{|x_{ij}|\le \eta_n\sqrt{n}\}}$$

and $\tilde{\mathbf{Y}}_n = n^{-1/2}\tilde{\mathbf{X}}_n + \mathbf{A}_n$ then

$$\mathbb{E}\tilde{x}_{ij}^2 \xrightarrow[n\to\infty]{} \vartheta = \mathbb{E}x_{ij}^2, \quad \mathbb{E}\tilde{x}_{ij}^4 \xrightarrow[n\to\infty]{} \mathbb{E}|x_{ij}|^4.$$

Moreover

$$\operatorname{Tr} f(\mathbf{Y}_n\mathbf{Y}_n^*) - \operatorname{Tr} f(\tilde{\mathbf{Y}}_n\tilde{\mathbf{Y}}_n^*) = o_{\mathbb{P}}(1)$$

as $N, n \to \infty$.

We shall now introduce a truncated version of M_n. Let $\mu_\ell < 0$ and $\mu_r > \ell^+$ where ℓ^+ is defined in Assumption 2.2.3; let $d > 0$ and let ε_n be a real sequence decreasing to zero and satisfying, for some $\alpha \in (0,1)$:

$$\varepsilon_n \ge N^{-\alpha} \ .$$

Consider the contour

$$\Gamma = \mathcal{C} \cup \overline{\mathcal{C}} \ , \tag{2.3.4}$$

where

$$\mathcal{C} = \{\mu_\ell + \mathbf{i}v;\ v \in [0,d]\} \cup \{x + \mathbf{i}d;\ x \in [\mu_\ell, \mu_r]\} \cup \{\mu_r + \mathbf{i}v;\ v \in [0,d]\}. \tag{2.3.5}$$

Denote by

$$\mathcal{C}_n = \{\mu_\ell + \mathbf{i}v;\ v \in [\varepsilon_n/n, d]\} \cup \{x + \mathbf{i}d;\ x \in [\mu_\ell, \mu_r]\} \cup \{\mu_r + \mathbf{i}v;\ v \in [\varepsilon_n/n, d]\}. \tag{2.3.6}$$

We can now introduce the truncated process $\hat{M}_n(z)$ defined for $z \in \mathcal{C}$ by :

$$\hat{M}_n(z) = \begin{cases} M_n(z) & \text{for } z \in \mathcal{C}_n \\ M_n(\mu_r + \mathbf{i}\varepsilon_n/n) & \text{for } x = \mu_r, v \in [0, \varepsilon_n/n] \\ M_n(\mu_\ell + \mathbf{i}\varepsilon_n/n) & \text{for } x = \mu_\ell, v \in [0, \varepsilon_n/n] \end{cases} ,$$

and for $z \in \overline{\mathcal{C}}$ by $\hat{M}_n(z) = \overline{\hat{M}_n(\bar{z})}$.

Proposition 2.3.2 (Truncation of the process) *Let Assumptions 2.2.1 and 2.2.2 hold true and let f satisfy Assumption 2.2.3. Then*

$$\oint_\Gamma f(z)\left(M_n(z) - \hat{M}_n(z)\right)\, dz = o_{\mathbb{P}}(1)$$

as $N, n \to \infty$.

From now on, and without further mention below, we assume that the entries are truncated, centered and normalized (and write x_{ij} instead of $\hat{x}_{ij}$), and that the process M_n is truncated near the real axis (and write M_n instead of $\hat{M}_n$).

2.3.3 More notations and useful estimates

Denote by $\mathbf{y}_j$, $\mathbf{a}_j$ and $\mathbf{r}_j$ the j-th columns of matrices $\mathbf{Y}_n$, $\mathbf{A}_n$ and $n^{-1/2}\mathbf{X}_n$. Recall that $\mathbf{Q}(z) = (\mathbf{Y}_n\mathbf{Y}_n^* - z\mathbf{I}_N)^{-1}$. If $z, z_1, z_2 \in \mathbb{C}^+$, let

$$v = |\mathrm{Im} z| \ , \quad v_i = |\mathrm{Im} z_i| \ , \ i = 1, 2 \ .$$

Denote by

$$\begin{aligned}
\mathbf{Y}_j\mathbf{Y}_j^* &= \mathbf{Y}\mathbf{Y}^* - \mathbf{y}_j\mathbf{y}_j^* && (2.3.7)\\
\mathbf{Q}_j &= (\mathbf{Y}\mathbf{Y}^* - \mathbf{y}_j\mathbf{y}_j^* - z\mathbf{I}_N)^{-1}, && (2.3.8)\\
b_j &= \frac{1}{1 + n^{-1}\mathbb{E}\mathrm{Tr}\mathbf{Q}_j + \mathbf{a}_j^*(\mathbb{E}\mathbf{Q}_j)\mathbf{a}_j}, && (2.3.9)\\
b &= \frac{1}{1 + n^{-1}\mathbb{E}\mathrm{Tr}\mathbf{Q}}, && (2.3.10)\\
\gamma_j &= \mathbf{y}_j^*\mathbf{Q}_j\mathbf{y}_j - n^{-1}\mathbb{E}(\mathrm{Tr}\mathbf{Q}_j) - \mathbf{a}_j^*(\mathbb{E}\mathbf{Q}_j)\mathbf{a}_j, && (2.3.11)\\
\hat{\gamma}_j &= \mathbf{y}_j^*\mathbf{Q}_j\mathbf{y}_j - n^{-1}\mathrm{Tr}\mathbf{Q}_j - \mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j, && (2.3.12)\\
\beta_j &= \frac{1}{1 + \mathbf{y}_j^*\mathbf{Q}_j\mathbf{y}_j}, && (2.3.13)\\
\bar{\beta}_j &= \frac{1}{1 + n^{-1}\mathrm{Tr}\mathbf{Q}_j + \mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j}, && (2.3.14)\\
\alpha_j &= \mathbf{y}_j^*\mathbf{Q}_j^2\mathbf{y}_j - n^{-1}\mathrm{Tr}\mathbf{Q}_j^2 - \mathbf{a}_j^*\mathbf{Q}_j^2\mathbf{a}_j. && (2.3.15)
\end{aligned}$$

Lemma 2.3.1 *[9, Lemma 2.7] Let* $\mathbf{x} = (x_1, \cdots, x_n)$ *be an* $n \times 1$ *vector where* x_i *are centered i.i.d. complex random variables with unit variance and bounded fourth moment. Let* $\mathbf{M}$ *be a Hermitian complex matrix. Then for any* $p \geq 2$*, there exists a constant* K_p *depending only on* p *for which*

$$\mathbb{E}|\mathbf{x}^*\mathbf{M}\mathbf{x} - \mathrm{Tr}\mathbf{M}|^p \leq K_p((\mathbb{E}|x_1|^4\mathrm{Tr}\mathbf{M}\mathbf{M}^*)^{p/2} + \mathbb{E}|x_1|^{2p}\mathrm{Tr}(\mathbf{M}\mathbf{M}^*)^{p/2}).$$

A direct consequence of this lemma is the following proposition.

Lemma 2.3.2 *[11, Eq. (3.2)] Let* $\mathbf{x}_n$ *be a sequence of random vectors as in the statement of the precedent lemma and* $\mathbf{M}_n$ *be a sequence of* $n \times n$ *matrices independent of* $\mathbf{x}_n$ *with uniformly bounded spectral norm. Denote by* $\mathbf{y}_n = \frac{1}{\sqrt{n}}\mathbf{x}_n + \mathbf{a}_n$ *where* $\mathbf{a}_n$ *is a deterministic vector with uniformly bounded Euclidian norm. Then*

$$\max\left(\mathbb{E}\left|\frac{\mathbf{x}_n^*\mathbf{M}_n\mathbf{x}_n}{n} - \frac{\mathrm{Tr}\mathbf{M}_n}{n}\right|^2, \mathbb{E}\left|\mathbf{y}_n^*\mathbf{M}_n\mathbf{y}_n - \frac{\mathrm{Tr}\mathbf{M}_n}{n} - \mathbf{a}_n^*\mathbf{M}_n\mathbf{a}_n\right|^2\right) \leq \frac{K}{n}.$$

Another consequence of Lemma 2.3.1 is the following lemma.

Lemma 2.3.3 *Under the same setting as Lemma 2.3.2, we have*

$$\mathbb{E}\left|\mathbf{y}_n^*\mathbf{M}_n\mathbf{y}_n - \frac{\mathrm{Tr}\mathbf{M}_n}{n} - \mathbf{a}_n^*\mathbf{M}_n\mathbf{a}_n\right|^4 \le \frac{K\eta_n^4}{n} \ ,$$

where η_n is defined in Proposition 2.3.1.

Proofs of Lemma 2.3.2 and 2.3.3 are postponed to Appendices 2.6.3 and 2.6.4.

2.3.4 The Lyapunov condition for the process (M_n^1)

We first decompose M_n^1 into martingale increments. Let $\mathbb{E}_0$ denote the expectation and $\mathbb{E}_j$ the conditional expectation with respect to the σ-field $\mathcal{F}_{n,j}$ generated by $\{\mathbf{x}_\ell\,,\ 1\le \ell\le j\}$. Noticing that

$$\mathbf{Q}-\mathbf{Q}_j = -\mathbf{Q}_j\mathbf{y}_j\mathbf{y}_j^*\mathbf{Q}_j\beta_j, \tag{2.3.16}$$

and

$$\beta_j = \bar\beta_j - \beta_j\bar\beta_j\hat\gamma_j = \bar\beta_j - \bar\beta_j^2\hat\gamma_j + \bar\beta_j^2\beta_j\hat\gamma_j^2 \ , \tag{2.3.17}$$

we have

$$\begin{aligned} M_n^1 &= \mathrm{Tr}(\mathbf{Q}-\mathbb{E}\mathbf{Q}) \\ &= \mathrm{Tr}\sum_{j=1}^n(\mathbb{E}_j-\mathbb{E}_{j-1})(\mathbf{Q}-\mathbf{Q}_j) \ = \ \sum_{j=1}^n(\mathbb{E}_j-\mathbb{E}_{j-1})\mathrm{Tr}\,(\mathbf{Q}-\mathbf{Q}_j)\end{aligned}$$

Otherwise stated, $(M_n^1(z_1),\cdots M_n^1(z_d))$ is a martingale difference sequence with respect to the increasing σ-field generated by $\mathbf{X}$'s columns and we are in position to apply Theorem 2.3.1. Write :

$$\begin{aligned} M_n^1 &= \sum_{j=1}^n(\mathbb{E}_j-\mathbb{E}_{j-1})\mathrm{Tr}\,(\mathbf{Q}-\mathbf{Q}_j) \\ &= -\sum_{j=1}^n(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_j\mathbf{y}_j^*\mathbf{Q}_j^2\mathbf{y}_j \\ &= -\sum_{j=1}^n(\mathbb{E}_j-\mathbb{E}_{j-1})(\bar\beta_j-\bar\beta_j^2\hat\gamma_j+\bar\beta_j^2\beta_j\hat\gamma_j^2)(\alpha_j+n^{-1}\mathrm{Tr}\mathbf{Q}_j^2+\mathbf{a}_j^*\mathbf{Q}_j^2\mathbf{a}_j) \\ &\overset{(a)}{=} -\sum_{j=1}^n\mathbb{E}_j(\bar\beta_j\alpha_j-\bar\beta_j^2\hat\gamma_j(n^{-1}\mathrm{Tr}\mathbf{Q}_j^2+\mathbf{a}_j^*\mathbf{Q}_j^2\mathbf{a}_j)) \\ &\qquad +\sum_{j=1}^n(\mathbb{E}_j-\mathbb{E}_{j-1})\bar\beta_j^2(\hat\gamma_j\alpha_j-\beta_j\alpha_j\hat\gamma_j^2) \\ &\qquad\qquad -\sum_{j=1}^n(\mathbb{E}_j-\mathbb{E}_{j-1})\bar\beta_j^2\beta_j\hat\gamma_j^2(n^{-1}\mathrm{Tr}\mathbf{Q}_j^2+\mathbf{a}_j^*\mathbf{Q}_j^2\mathbf{a}_j)\ , \\ &\overset{\triangle}{=} \sum_{j=1}^n P_j + \sum_{j=1}^n P_j' + \sum_{j=1}^n P_j'' \end{aligned} \tag{2.3.18}$$

where we have used in (a) the fact that conditionally on $\mathbf{y}_j$, the expectation of α_j and $\hat{\gamma}_j$ is zero. Consider $(z_\ell\,, 1 \le \ell \le d)$ in $\mathbb{C}^+$. We first check Lyapunov's condition (2.3.3) and shall prove :

$$\lim_{N,n} \sum_{j=1}^{n} \mathbb{E}\left(\sum_{\ell=1}^{d} |(P_j + P_j' + P_j'')(z_\ell)|^2 \right) \mathbb{I}_{\{\sum_{\ell=1}^{d} |(P_j+P_j'+P_j'')(z_\ell)|^2 \ge \varepsilon^2\}} = 0 \ . \tag{2.3.19}$$

Let $z \in \mathbb{C}^+$; one can easily check that (2.3.20)-(2.3.23) below (to be proven hereafter) imply (2.3.19).

$$\lim_{N,n} \quad \sum_{j=1}^{n} \quad \mathbb{E}|P_j'(z)|^2 = 0 \ , \tag{2.3.20}$$

$$\lim_{N,n} \quad \sum_{j=1}^{n} \quad \mathbb{E}|P_j''(z)|^2 = 0 \ , \tag{2.3.21}$$

$$\sup_{N,n} \quad \sum_{j=1}^{n} \quad \mathbb{E}\left(\sum_{\ell=1}^{d} |P_j(z_\ell)|^2 \right) \le K < \infty \ , \tag{2.3.22}$$

$$\lim_{N,n} \quad \sum_{j=1}^{n} \quad \mathbb{E}\left(\sum_{\ell=1}^{d} |P_j(z_\ell)|^2 \right) \mathbb{I}_{\{\sum_{\ell=1}^{d} |P_j(z_\ell)|^2 \ge \varepsilon^2\}} = 0 \ . \tag{2.3.23}$$

Let $z \in \mathbb{C}^+$ and recall that $v = |\mathrm{Im} z|$. Taking into account the fact that $-z^{-1}\bar{\beta}_j(z)$, $-z^{-1}\beta_j(z)$ and $-z^{-1}b_j(z)$ are Stieltjes transforms of probability measures, a fact that can be proved using [44, Proposition 2.2], we easily obtain the estimates :

$$\max\left(\left|\frac{\bar{\beta}_j(z)}{z}\right| \ , \ \left|\frac{\beta_j(z)}{z}\right| \ , \ \left|\frac{b_j(z)}{z}\right| \right) \le \frac{1}{v} \ . \tag{2.3.24}$$

Write :

$$\begin{aligned}
\sum_{j=1}^{n} \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\bar{\beta}_j^2(z)\hat{\gamma}_j(z)\alpha_j(z)|^2 \quad &\le \quad 4\sum_{j=1}^{n} \mathbb{E}|\bar{\beta}_j^2(z)\hat{\gamma}_j(z)\alpha_j(z)|^2 \\
&\overset{(a)}{\le} \quad \frac{|z|^4}{v^4} \sum_{j=1}^{n} \left\{\mathbb{E}|\hat{\gamma}_j(z)|^4\right\}^{1/2} \left\{\mathbb{E}|\alpha_j(z)|^4\right\}^{1/2} \\
&\overset{(b)}{\le} \quad K\,\eta_n^4 \ ,
\end{aligned}$$

where (a) follows from (2.3.24) and (b) from Lemma 2.3.3. A similar argument yields :

$$\mathbb{E}\left| \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j-1})\beta_j(z)\alpha_j(z)\hat{\gamma}_j^2(z) \right|^2 \quad \xrightarrow[N,n\to\infty]{} \quad 0.$$

Gathering the two previous results yields (2.3.20) ; (2.3.21) can be proved similarly. Recall that

$$P_j(z) = -\mathbb{E}_j \left[\bar{\beta}_j(z)\alpha_j(z) - \bar{\beta}_j^2(z)\hat{\gamma}_j(z)(\frac{1}{n}\mathrm{Tr}\mathbf{Q}_j^2(z) + \mathbf{a}_j^*\mathbf{Q}_j^2(z)\mathbf{a}_j) \right] \ . \tag{2.3.25}$$

By Lemma 2.3.2 and (2.3.24), we have :

$$\mathbb{E}|P_j(z)|^2 \le K\left(\frac{|z|^2}{v^2}\mathbb{E}|\alpha_j(z)|^2 + \frac{|z|^4}{v^8}\mathbb{E}|\hat{\gamma}(z)|^2\right) \le \frac{K}{n} \ , \tag{2.3.26}$$

hence (2.3.22). By Lemma 2.3.3 and (2.3.24), we have :

$$\mathbb{E}|P_j(z)|^4 \le K\left(\frac{|z|^4}{v^4}\mathbb{E}|\alpha_j(z)|^4 + \frac{|z|^8}{v^{16}}\mathbb{E}|\hat{\gamma}_j(z)|^4\right) = o(n^{-1}) \ .$$

Now,

$$\begin{aligned}
&\sum_{j=1}^n \mathbb{E}\left(\sum_{\ell=1}^d |P_j(z_\ell)|^2\right)\mathbb{I}_{\{\sum_{\ell=1}^d |P_j(z_\ell)|^2 \ge \varepsilon^2\}} \\
&\quad \le \ \frac{1}{\varepsilon^2}\sum_{j=1}^n \mathbb{E}\left(\sum_{\ell=1}^d |P_j(z_\ell)|^2\right)^2 \ \le \ \frac{K}{\varepsilon^2}\sum_{j=1}^n\sum_{\ell=1}^d \mathbb{E}|P_j(z_\ell)|^4 \xrightarrow[N,n\to\infty]{} 0 \ ,
\end{aligned}$$

hence (2.3.23). Lyapunov's condition (2.3.19) is therefore established.

For future use, we show

$$\sup_n \mathbb{E}|M_n^1(z)|^2 \le K \ . \tag{2.3.27}$$

Using decomposition (2.3.18), we have

$$\begin{aligned}
\mathbb{E}|M_n^1(z)|^2 &= \mathbb{E}\left|\sum_{j=1}^n \left(P_j(z) + P_j'(z) + P_j''(z)\right)\right|^2 \ , \\
&= \sum_{j=1}^n \mathbb{E}\left|P_j(z) + P_j'(z) + P_j''(z)\right|^2 \ , \\
&\le 3\sum_{j=1}^n \mathbb{E}\left(|P_j(z)|^2 + |P_j'(z)|^2 + |P_j''(z)|^2\right) \ .
\end{aligned}$$

Estimate (2.3.27) then follows from (2.3.20)-(2.3.22).

It now remains to prove the counterparts of (2.3.1)-(2.3.2) for (M_n^1). Since $\overline{M}_n^1(z) = M_n^1(\bar{z})$, it suffices to prove

$$\sum_{i=1}^n \mathbb{E}\left(M_n^1(z_k)M_n^1(z_\ell) \mid \mathcal{F}_{n,i-1}\right) - \Theta_n(z_k, z_\ell) \xrightarrow[N,n\to\infty]{\mathcal{P}} 0$$

for $z_k, z_\ell \in \Gamma$, with Θ_n given by (2.2.11). Section 2.4 is devoted to the computation of this key-estimate.

Therefore, we achieve the finite-dimensional convergence of the process $M_n^1(z)$.

2.3.5 Tightness of the process (M_n^1)

We closely follow the computations of the tightness in Bai and Silverstein [11, Section 3]. Based on [18, Theorem 13.1] after the truncation step (cf. Section 2.3.2), it suffices to prove that there exists $z_0 \in \mathcal{C}_n$ and a constant K such that $M_n^1(z)$ satisfies :

$$\sup_n \mathbb{E}|M_n^1(z_0)|^2 < \infty \quad \text{and} \quad \sup_{n,z_1,z_2 \in \mathcal{C}_n} \frac{\mathbb{E}|(M_n^1(z_1) - M_n^1(z_2))|^2}{|z_1 - z_2|^2} \le K. \tag{2.3.28}$$

The estimate $\mathbb{E}|M_n^1(z_0)|^2 < \infty$ immediatly follows from (2.3.22). To prove the second estimate (2.3.28), we start with the estimation of the eigenvalues.

Lemma 2.3.4 *Assume that Assumptions 2.2.1 and 2.2.2 hold true and that the random variables (x_{ij}^n) are truncated as in Section 2.3.2. Recall the definition of μ_ℓ and μ_r in Section 2.3.2. Then*

$$\forall k \in \mathbb{N} \ , \quad \mathbb{P}(\lambda_{\max}^{\mathbf{Y}_n\mathbf{Y}_n^*} > \mu_r) = o(n^{-k}) \quad \textit{and} \quad \mathbb{P}(\lambda_{\min}^{\mathbf{Y}_n\mathbf{Y}_n^*} < \mu_\ell) = 0 \ .$$

Proof: Since $\mu_\ell < 0$, we clearly have $\mathbb{P}(\lambda_{\min}^{\mathbf{Y}_n\mathbf{Y}_n^*} < \mu_\ell) = 0$. Now

$$\begin{aligned}
\mathbb{P}(\lambda_{\max}^{\mathbf{Y}_n\mathbf{Y}_n^*} > \mu_r) &\le \mathbb{P}(2\lambda_{\max}^{\frac{1}{n}\mathbf{X}_n\mathbf{X}_n^*} + 2\lambda_{\max}^{\mathbf{A}_n\mathbf{A}_n^*} > \mu_r) \\
&= \mathbb{P}(\lambda_{\max}^{\frac{1}{n}\mathbf{X}_n\mathbf{X}_n^*} > \frac{\mu_r}{2} - \lambda_{\max}^{\mathbf{A}_n\mathbf{A}_n^*}) \\
&\le \mathbb{P}(\lambda_{\max}^{\frac{1}{n}\mathbf{X}_n\mathbf{X}_n^*} > (1+\sqrt{c})^2) \quad = \quad o(n^{-k}) \ ,
\end{aligned}$$

where the last result follows from [11, Eq. (1.9a)]. □

Recall notations introduced in Section 2.3.3, the definition of η_n in Section 2.3.2 and of $\mathcal{C}_n$ in (2.3.6).

Lemma 2.3.5 *Let Assumptions 2.2.1 and 2.2.2 hold true and that the random variables (x_{ij}^n) are truncated as in Section 2.3.2. Denote by*

$$\mathbf{M}_j = \mathbf{Q}_j^{k_1}(z_1)\mathbf{Q}_j^{k_2}(z_2) \ .$$

Then for $z_1, z_2 \in \mathcal{C}_n$, $q \in \mathbb{N}$, there exists a constant K_p which depends only on p, such that

$$\max\left\{\sup_{z\in\mathcal{C}_n} \mathbb{E}\|\mathbf{Q}_j(z)\|^p, \sup_{z\in\mathcal{C}_n} \mathbb{E}|\mathbf{y}_j^*\mathbf{Q}_j^q(z)\mathbf{y}_j|^p\right\} \le K_{qp}, \tag{2.3.29}$$

$$\mathbb{E}\left|\mathbf{y}_j^*\mathbf{M}_j\mathbf{y}_j - n^{-1}\mathrm{Tr}\mathbf{M}_j - \mathbf{a}_j^*\mathbf{M}_j\mathbf{a}_j\right|^2 \le \frac{K_{2(k_1+k_2)}}{n}, \tag{2.3.30}$$

$$\mathbb{E}\left|\mathbf{y}_j^*\mathbf{M}_j\mathbf{y}_j - n^{-1}\mathrm{Tr}\mathbf{M}_j - \mathbf{a}_j^*\mathbf{M}_j\mathbf{a}_j\right|^4 \le \frac{K_{4(k_1+k_2)}\eta_n^4}{n} \ , \tag{2.3.31}$$

$$\mathbb{E}|\gamma_j(z)|^2 \le K_2/n, \quad \mathbb{E}|\gamma_j(z)|^4 \le K_4\eta_n^4/n \ . \tag{2.3.32}$$

$$\max\left\{\sup_{z\in\mathcal{C}_n} \mathbb{E}|\beta_j(z)|^p, \sup_{z\in\mathcal{C}_n} |b_j(z)|^p\right\} \le K_p. \tag{2.3.33}$$

Proof of Lemma 2.3.5 is postponed to Appendix 2.6.5. By the resolvent identity, we have :

$$\mathbf{Q}(z_1) - \mathbf{Q}(z_2) = (z_1 - z_2)\mathbf{Q}(z_1)\mathbf{Q}(z_2) \ .$$

Since $(\mathbb{E}_j - \mathbb{E}_{j-1})\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2) = 0$, we obtain :

$$\begin{aligned} \frac{M_n^1(z_1) - M_n^1(z_2)}{z_1 - z_2} &= \sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1})\mathrm{Tr}\mathbf{Q}(z_1)\mathbf{Q}(z_2) \\ &= \sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1})\mathrm{Tr}\left(\mathbf{Q}(z_1)\mathbf{Q}(z_2) - \mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)\right) \ . \end{aligned}$$

Using (2.3.16), we have :

$$\begin{aligned} &\mathbf{Q}(z_1)\mathbf{Q}(z_2) - \mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2) \\ &\quad = (\mathbf{Q}(z_1) - \mathbf{Q}_j(z_1))(\mathbf{Q}(z_2) - \mathbf{Q}_j(z_2)) \\ &\qquad +(\mathbf{Q}(z_1) - \mathbf{Q}_j(z_1))\mathbf{Q}_j(z_2) + \mathbf{Q}_j(z_1)(\mathbf{Q}(z_2) - \mathbf{Q}_j(z_2)) \\ &\quad = \beta_j(z_1)\beta_j(z_2)\mathbf{Q}_j(z_1)\mathbf{y}_j\mathbf{y}_j^*\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j\mathbf{y}_j^*\mathbf{Q}_j(z_2) \\ &\qquad -\beta_j(z_1)\mathbf{Q}_j(z_1)\mathbf{y}_j\mathbf{y}_j^*\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2) - \beta_j(z_2)\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j\mathbf{y}_j^*\mathbf{Q}_j(z_2), \end{aligned}$$

which yields

$$\begin{aligned} &\mathrm{Tr}(\mathbf{Q}(z_1)\mathbf{Q}(z_2) - \mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)) \\ &\quad = \beta_j(z_1)\beta_j(z_2)(\mathbf{y}_j^*\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j)^2 \\ &\qquad -\beta_j(z_1)\mathbf{y}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j - \beta_j(z_2)\mathbf{y}_j^*\mathbf{Q}_j^2(z_2)\mathbf{Q}_j(z_1)\mathbf{y}_j. \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{M_n^1(z_1) - M_n^1(z_2)}{z_1 - z_2} &= \sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1})\beta_j(z_1)\beta_j(z_2)(\mathbf{y}_j^*\mathbf{Q}_j(z_1)\mathbf{Q}_j\mathbf{y}_j)^2 \\ &\quad - \sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1})\beta_j(z_1)\mathbf{y}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j \\ &\qquad - \sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1})\beta_j(z_2)\mathbf{y}_j^*\mathbf{Q}_j^2(z_2)\mathbf{Q}_j(z_1)\mathbf{y}_j \\ &\triangleq T_1 + T_2 + T_3 \ . \end{aligned} \tag{2.3.34}$$

Our goal is to show the absolute second moment of (2.3.34) is uniformly bounded over the contour $\mathcal{C}_n$. From the relation $\beta_j(z) = b_j(z) - b_j(z)\beta_j(z)\gamma_j(z)$, the term T_2 writes :

$$T_2 = \sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1})\beta_j(z_1)\mathbf{y}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j$$

$$\begin{aligned}
&= \sum_{j=1}^{n}(\mathbb{E}_j-\mathbb{E}_{j-1})\left\{b_j(z_1)\mathbf{y}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j-\beta_j(z_1)b_j(z_1)\mathbf{y}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j\gamma_j(z_1)\right\}\\
&= \sum_{j=1}^{n} b_j(z_1)\mathbb{E}_j\left\{\mathbf{y}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j-n^{-1}\mathrm{Tr}\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)-\mathbf{a}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{a}_j\right\}\\
&\quad -\sum_{j=1}^{n} b_j(z_1)(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_j(z_1)\mathbf{y}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j\gamma_j(z_1)\\
&\triangleq W_1-W_2.
\end{aligned}$$

By Lemma 2.3.5 with $k_1=2, k_2=1$ and (2.3.33), we have

$$\begin{aligned}
\mathbb{E}|W_1|^2 = \sum_{j=1}^{n}|b_j(z)|^2\mathbb{E}\Big|\mathbb{E}_j[\mathbf{y}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j\\
-n^{-1}\mathrm{Tr}\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)-\mathbf{a}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{a}_j]\Big|^2 \le K.
\end{aligned}\tag{2.3.35}$$

For W_2, taking $\eta_\ell\in(\mu_\ell,0)$ and $\eta_r\in(\ell_+,\mu_r)$ we have :

$$\begin{aligned}
\mathbb{E}|W_2|^2 &= \sum_{j=1}^{n}|b_j(z_1)|^2\mathbb{E}|(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_j(z_1)\mathbf{y}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j\gamma_j(z_1)|^2\\
&\overset{(a)}{\le} K\sum_{j=1}^{n}\mathbb{E}\left|\beta_j(z_1)\mathbf{y}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j\gamma_j(z_1)\right|^2\\
&\overset{(b)}{\le} K\sum_{j=1}^{n}\left(\mathbb{E}|\gamma_j(z_1)|^2+v^{-8}\mathbb{P}(\lambda_{\max}^{\mathbf{Y}_j\mathbf{Y}_j^*}\ge\eta_r \quad\text{or}\quad \lambda_{\min}^{\mathbf{Y}_j\mathbf{Y}_j^*}\le\eta_\ell)\right)\\
&\overset{(c)}{\le} K,
\end{aligned}$$

where (a) uses (2.3.33), (b) is from the fact that $\beta_j(z_1)\mathbf{y}_j^*\mathbf{Q}_j^2(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j$ is bounded when $\eta_\ell\le\lambda_{\max}^{\mathbf{Y}_j\mathbf{Y}_j^*}\le\eta_r$ and (c) is from (2.3.30). The term T_3 in (2.3.34) can be handled similarly.

We now handle T_1 :

$$\begin{aligned}
&\sum_{j=1}^{n}(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_j(z_1)\beta_j(z_2)(\mathbf{y}_j^*\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j)^2\\
&=\sum_{j=1}^{n} b_j(z_1)b_j(z_2)(\mathbb{E}_j-\mathbb{E}_{j-1})[(\mathbf{y}_j^*\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j)^2\\
&\quad-(n^{-1}\mathrm{Tr}\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)+\mathbf{a}_j^*\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)\mathbf{a}_j)^2]\\
&\quad-\sum_{j=1}^{n} b_j(z_2)(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_j(z_1)\beta_j(z_2)(\mathbf{y}_j^*\mathbf{Q}_j(z_1)\mathbf{Q}_j\mathbf{y}_j)^2\gamma_j(z_2)\\
&\quad-\sum_{j=1}^{n} b_j(z_1)b_j(z_2)(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_j(z_1)(\mathbf{y}_j^*\mathbf{Q}_j(z_1)\mathbf{Q}_j\mathbf{y}_j)^2\gamma_j(z_2)\\
&\triangleq Y_1-Y_2-Y_3.
\end{aligned}$$

Both Y_2 and Y_3 can be handled in the same way as W_2 and one can prove that :

$$\mathbb{E}|Y_2|^2 \leq K, \quad \mathbb{E}|Y_3|^2 \leq K.$$

Moreover,

$$\begin{aligned}
\mathbb{E}|Y_1|^2 &\leq K \sum_{j=1}^{n} \mathbb{E}|(\mathbf{y}_j^* \mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j)^2 - (n^{-1}\mathrm{Tr}\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2) + \mathbf{a}_j^*\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)\mathbf{a}_j)^2|^2 \\
&\overset{(a)}{\leq} K \sum_{j=1}^{n} \mathbb{E}|\mathbf{y}_j^* \mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)\mathbf{y}_j - n^{-1}\mathrm{Tr}\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2) - \mathbf{a}_j^*\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)\mathbf{a}_j|^2 \\
&\quad + K v_1^{-4} v_2^{-4} \mathbb{P}(\lambda_{\max}^{\mathbf{Y}_j\mathbf{Y}_j^*} \geq \eta_r \quad \text{or} \quad \lambda_{\min}^{\mathbf{Y}_j\mathbf{Y}_j^*} \leq \eta_\ell) \\
&\overset{(b)}{\leq} K
\end{aligned}$$

where (a) comes from the fact that $\|\mathbf{Q}_j(z_1)\mathbf{Q}_j(z_2)\|$ is bounded when $\eta_\ell \leq \lambda_{\max}^{\mathbf{Y}_j\mathbf{Y}_j^*} \leq \eta_r$ and (b) is again a consequence of (2.3.30) and Lemma 2.3.4.

We conclude that all terms in (2.3.34) are indeed bounded uniformly over $\mathcal{C}_n$. This completes the proof of the tightness of the process $M_n^1(z)$.

2.3.6 Tightness of the Gaussian process

The last task consists in establishing the tightness for the Gaussian process $N_n(z) = \mathcal{N}(0, \Theta_n(z))$ where Θ_n is defined in (2.2.11). We will transfer the tightness of the process $M_n^1(z)$ to the tightness of the Gaussian process $N_n(z)$ with the help of meta-model.

With the extra parameter M, consider the $NM \times nM$ matrix

$$\mathbf{A}_n^M = \begin{bmatrix} \mathbf{A}_n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{A}_n \end{bmatrix}$$

where $\mathbf{A}_n^M$ is the block matrix whose diagonal is composed of M matrices $\mathbf{A}_n$ and 0 elsewhere. Consider now the random matrix

$$\mathbf{Y}_n^M = \frac{1}{\sqrt{nM}} \mathbf{X}_n^M + \mathbf{A}_n^M.$$

Denote $M_{n,M}^1(z) = \frac{1}{NM}\sum_{j=1}^{NM} \frac{1}{\lambda_j - z} - \frac{1}{NM}\sum_{j=1}^{NM} \mathbb{E}\frac{1}{\lambda_j - z}$ where $(\lambda_i)_{1\leq i\leq NM}$ are eigenvalues of $\mathbf{Y}_n^M \mathbf{Y}_n^{M*}$. By a simple calculation of the bloc matrix, $\mathbf{Y}_n^M\mathbf{Y}_n^{M*}$ has the same deterministic equivalent δ_n as $\mathbf{Y}_n\mathbf{Y}_n{}^*$. With the precedent martingale argument, we can show that for n, N fixed, $M_{n,M}^1(z)$ converges to a Gaussian process with

$$\mathrm{Cov}[M_{n,M}^1(z_1), M_{n,M}^1(z_2)] \xrightarrow[M\to\infty]{} \Theta_n(z_1, z_2)$$

where $\Theta(z_1, z2) = \Theta_{0,n}(z_1, z_2) + \Theta_{1,n}(\vartheta, z_1, z_2) + \Theta_{2,n}(\kappa, z_1, z_2)$.

With these notations at hand, we shall show the tightness of the Gaussian process. According to [18, Chapter 2, Section 7], a sequence of Gaussian process is tight if and only if it is relatively compact in distribution. Consider the set of matrices :

$$\{\mathbf{A}_n^M : M \geq 1, \mathbf{A}_n \text{ is an } N \times n \text{ matrix which satisfies Assumption 2.2.2}\}.$$

Since $\|\mathbf{A}_n^M\| = \|\mathbf{A}_n\|$ for all $M \geq 1$, we have

$$\sup_{M\geq 1, N,n\to\infty} \|\mathbf{A}_n^M\| = \sup_{N,n\to\infty} \|\mathbf{A}_n\| < \infty$$

by Assumption 2.2.2. The arguments with Billingsley's condition (cf. (2.3.28)) show that the family $\{M_{n,M} : M \geq 1\}_{n,N\to\infty}$ forms a tight sequence over $\mathcal{C}_N$, hence relatively compact in distribution. As the distribution $\mathcal{L}(N_n)$ of the Gaussian process N_n is the limit in M of the distribution of $M_{n,M}$, belongs to the closure of $\{M_{n,M} : M \geq 1\}_{n,N\to\infty}$, which is relatively compact. Finally, $\mathcal{L}(N_n)$ is included in a compact set, hence is relatively compact. In particular, the family of gaussian processes (N_n) is tight. This proves that the variance $\Theta_n(z)$ is uniformly bounded.

2.4 Proof of Theorem 2.2.1 (II) : Computation of the covariance

In Hachem et al. [39], a CLT for mutual information has been established

$$\frac{1}{N}\log\det(\mathbf{Y}_n\mathbf{Y}_n^* + \rho\mathbf{I}), \quad \text{for } \rho > 0.$$

Be aware that

$$\mathrm{Tr}(\mathbf{Y}_n\mathbf{Y}_n^* + \rho\mathbf{I})^{-1} = \frac{\partial}{\partial\rho}\log\det(\mathbf{Y}_n\mathbf{Y}_n^* + \rho\mathbf{I}),$$

and based on this idea, in this section we shall calculate the covariance $\mathrm{Cov}(M_n^1(z_1), M_n^1(z_2))$. We denote :

$$\Gamma_j(z) = \frac{\mathbf{y}_j^*\mathbf{Q}_j(z)\mathbf{y}_j - (\frac{1}{n}\mathrm{Tr}\mathbf{Q}_j(z) + \mathbf{a}_j^*\mathbf{Q}_j(z)\mathbf{a}_j)}{1 + \frac{1}{n}\mathrm{Tr}\mathbf{Q}_j(z) + \mathbf{a}_j^*\mathbf{Q}_j(z)\mathbf{a}_j}, \tag{2.4.1}$$

$$\tilde{b}_j(z) = \frac{-1}{z(1 + \mathbf{a}_j^*\mathbf{Q}_j(z)\mathbf{a}_j + \frac{1}{n}\mathrm{Tr}\mathbf{Q}_j(z))}, \tag{2.4.2}$$

$$\tilde{q}_{jj}(z) = \frac{-1}{z(1 + \mathbf{y}_j^*\mathbf{Q}_j(z)\mathbf{y}_j)}, \tag{2.4.3}$$

$$\mathbf{T}_j(z) = \left(-z(1+\tilde{\delta}_n)\mathbf{I}_N + \frac{\mathbf{A}_j\mathbf{A}_j^*}{1+\delta_n(z)}\right)^{-1}, \tag{2.4.4}$$

$$\tilde{t}_{jj} = \frac{-1}{z(1 + \mathbf{a}_j^*\mathbf{T}_j\mathbf{a}_j + \delta_n)}, \tag{2.4.5}$$

where $\mathbf{A}_j$ is the matrix $\mathbf{A}$ where the j-th column has been removed and $\delta_n, \tilde{\delta}_n$ are defined in (3.1.2). From (2.4.1)-(2.4.5), we have then

$$\begin{aligned}
\mathbf{Q}(z) &= \mathbf{Q}_j(z) + z\tilde{q}_{jj}(z)\mathbf{Q}_j(z)\mathbf{y}_j\mathbf{y}_j^*\mathbf{Q}_j(z), && (2.4.6)\\
\mathbf{Q}_j(z) &= \mathbf{Q}(z) + \frac{\mathbf{Q}(z)\mathbf{y}_j\mathbf{y}_j^*\mathbf{Q}(z)}{1-\mathbf{y}_j^*\mathbf{Q}(z)\mathbf{y}_j}, && (2.4.7)\\
1+\mathbf{y}_j^*\mathbf{Q}_j(z)\mathbf{y}_j &= \frac{1}{1-\mathbf{y}_j^*\mathbf{Q}(z)\mathbf{y}_j}, && (2.4.8)\\
\mathbf{y}_j^*\mathbf{Q}(z) &= -z\tilde{q}_{jj}(z)\mathbf{y}_j^*\mathbf{Q}_j(z), && (2.4.9)\\
1+z\tilde{t}_{\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_\ell\mathbf{a}_\ell &= -z\tilde{t}_{\ell\ell}(1+\delta), && (2.4.10)\\
-z\tilde{t}_{jj}\mathbf{a}_j^*\mathbf{T}_j\mathbf{b} &= \frac{\mathbf{a}_j^*\mathbf{T}\mathbf{b}}{1+\delta_n}, \quad \text{for any vector } \mathbf{b}, && (2.4.11)
\end{aligned}$$

which will be extremely useful in the sequel.

The following theorem gives the estimates of the terms.

Theorem 2.4.1 *[39, Theorem 3.1] Let $(\mathbf{u}_n)$ and $(\mathbf{v}_n)$ be two sequences of deterministic complex $N \times 1$ vectors bounded by*

$$\sup_{n\geq 1} \max(\|\mathbf{u}_n\|, \|\mathbf{v}_n\|) < \infty,$$

and let $(\mathbf{U}_n)$ be a sequence of deterministic $N \times N$ matrix with bounded spectral norm

$$\sup_{n\geq 1} \|\mathbf{U}_n\| < \infty.$$

Then, we have the following inequalities,

$$\begin{aligned}
\sum_{j=1}^{n} \mathbb{E}|\mathbf{u}_n^*\mathbf{Q}_j\mathbf{a}_j|^2 &\leq K, && (2.4.12)\\
\left|\frac{1}{n}\mathrm{Tr}\mathbf{U}(\mathbf{T}-\mathbb{E}\mathbf{Q})\right| &\leq \frac{K}{n}, && (2.4.13)\\
\mathbb{E}|\mathbf{u}_n^*(\mathbf{Q}-\mathbf{T})\mathbf{v}_n|^2 &\leq \frac{K_p}{n}, && (2.4.14)\\
\mathbb{E}|\mathbf{u}_n^*(\mathbf{Q}_j-\mathbf{T}_j)\mathbf{v}_n|^2 &\leq \frac{K_p}{n}, && (2.4.15)\\
\mathbb{E}|\mathrm{Tr}\mathbf{U}(\mathbf{Q}-\mathbb{E}\mathbf{Q})|^2 &\leq K, && (2.4.16)\\
\mathbb{E}|\tilde{q}_{jj}-\tilde{b}_{jj}|^p &\leq \frac{K_p}{n^{p/2}}, && (2.4.17)\\
\mathbb{E}|\tilde{q}_{jj}-\tilde{t}_{jj}|^p &\leq \frac{K_p}{n^{p/2}}, && (2.4.18)
\end{aligned}$$

where $p \in [1,2]$, K is a constant and K_p is a constant which depends only on p.

Recall the decomposition (cf. (2.3.18))

$$M_n^1(z) = \sum_{j=1}^{n} P_j + \sum_{j=1}^{n} P_j' + \sum_{j=1}^{n} P_j''$$

where

$$\lim_{N,n}\sum_{j=1}^{n}\mathbb{E}|P_j'(z)|^2=0,\quad \lim_{N,n}\sum_{j=1}^{n}\mathbb{E}|P_j''(z)|^2=0,$$

and the definition of P_j in (2.3.25). To calculate the covariance between $M_n^1(z_1)$ and $M_n^1(z_2)$, by Condition (2.3.1) and (2.3.2), we consider the sum

$$\mathbb{E}_{j-1}P_j(z_1)P_j(z_2) \tag{2.4.19}$$

and we have the following proposition.

Proposition 2.4.1 *Recall the definition of Γ_j in (2.4.1) and consider the sum*

$$A_n(z_1,z_2)=\sum_{j=1}^{n}\mathbb{E}_{j-1}\left\{\mathbb{E}_j\Gamma_j(z_1)\mathbb{E}_j\Gamma_j(z_2)\right\}. \tag{2.4.20}$$

Then

$$\frac{\partial^2}{\partial z_1\partial z_2}A_n(z_1,z_2)=(2.4.19), \tag{2.4.21}$$

and if there exists a term $L_n(z_1,z_2)$ such that

$$d_{\mathcal{LP}}(A_n(z_1,z_2),L_n(z_1,z_2))\xrightarrow[N,n\to\infty]{}0,$$

then

$$d_{\mathcal{LP}}\left(\mathrm{Cov}(M_n^1(z_1),M_n^1(z_2)),\frac{\partial^2}{\partial z_1\partial z_2}L_n(z_1,z_2)\right)\xrightarrow[N,n\to\infty]{}0. \tag{2.4.22}$$

Moreover, recall the definition of $\Theta_{0,n}$ (resp. $\Theta_{1,n}$, $\Theta_{2,n}$) in (2.2.12) (resp. (2.2.13), (2.2.14)). For all $z_1,z_2\in\mathbb{C}\backslash\mathbb{R}$, we have

$$\mathrm{Cov}(M_n^1(z_1),M_n^1(z_2))=\Theta_{0,n}(z_1,z_2)+\Theta_{1,n}(\vartheta,z_1,z_2)+\Theta_{2,n}(\kappa,z_1,z_2). \tag{2.4.23}$$

Remark 2.4.1 *In [39], they have shown that for $z=-\rho$ with $\rho>0$,*

$$\mathrm{Var}(A_n(z))=-\log\Delta_n(z,z)-\log\Delta_n^{\vartheta}(z,z)+\frac{\kappa z^2}{n^2}\sum_{i=1}^{N}t_{ii}^2(z)\sum_{j=1}^{n}\tilde{t}_{jj}^2(z)+o(1).$$

The main task is to generate the precedent formula of $\mathrm{Cov}(A_n(z_1),A_n(z_2))$ for $z_1,z_2\in\mathbb{C}^+$. However, $\log z_1z_2=\log z_1+\log z_2$ is no more true in complex case. We follow closely the method in [39] and will conquer this difficulty. In addition, [39] supposes that $\mathbb{E}|x_{ij}|^{16}<\infty$ which will be relaxed later. In particular, the moment condition will be replaced by Lemma 2.3.3 in the proof of Lemma 2.4.6.

Proof: We begin with the proof of (2.4.21) in Proposition 2.4.1. By Cauchy-Schwarz,

$$|\mathbb{E}(M_n^1(z_1)M_n^1(z_2))|\leq(\mathbb{E}|M_n^1(z_1)|^2)^{1/2}(\mathbb{E}|M_n^1(z_2)|^2)^{1/2}.$$

With the dominated convergence theorem with expectation, it suffices to show that $\sup_n \mathbb{E}|M_n(z_1)|^2 < \infty$ which is done in (2.3.27). This proves (2.4.21).

Now we will prove (2.4.22) based on the following lemma.

Lemma 2.4.1 *[11, Lemma 2.3] Let $(f_n)_{n\geq 1}$ be analytic in D, a connected open set of $\mathbb{E}$, satisfying $|f_n(z)| \leq M$ for every n and z in D, and $f_n(z)$ converges, as $n \to \infty$ for each z in a subset of D having a limit point in D. Then there exists a function f, analytic in D for which $f_n(z) \to f(z)$ and $f_n'(z) \to f'(z)$ for all $z \in D$. Moreover, on any set bounded by a contour interior to D the convergence is uniform and $f_n(z)$ is uniformly bounded by $2M/\varepsilon$, where ε is the distance between the contour and the boundary of D.*

We first prove that $|A_n(z_1, z_2)|$ and $|L_n(z_1, z_2)|$ are bounded. Recall $v_i = |\mathrm{Im} z_i|$ for $i = 1, 2$. Using (2.3.24) and Lemma 2.3.2, we have

$$\left| \sum_{j=1}^{n} \mathbb{E}_{j-1}\{\mathbb{E}_j \Gamma_j(z_1)\mathbb{E}_j\Gamma_j(z_2)\} \right| \leq K|z_1 z_2|(v_1 v_2)^{-2}.$$

And by the argument of meta-model (cf. Section 2.3.6), the above inequality implies that $\sum_{j=1}^{n} \mathbb{E}_{j-1}\{\mathbb{E}_j\Gamma_j(z_1)\mathbb{E}_j\Gamma_j(z_2)\}$ is bounded. Then the limit L_n is in the closure of A_n, hence also bounded. This implies that L_n is bounded.

Now we will apply Lemma 2.4.1. Denote v_0 a lower bound of $|\mathrm{Im} z_i|$ and $\{z_i\} \subset D = \{z : v_0 \leq |\mathrm{Im} z| \leq K\}$. Then by a diagonalization argument, for any subsequence of the natural numbers, there is a further subsequence such that, with probability one, $A_n - L_n$ converges for each pair $z_k, z_\ell \in \{z_i\}$. Write $A_n - L_n$ as $f_n(z_1, z_2)$. We concentrate on this subsequence and on one realization for which convergence holds. For each $z_\ell \in \{z_i\}$, we apply Lemma 2.4.1 on the set of $\{z : v_0/2 \leq \mathrm{Im} z \leq K\} \cup \{z : -K \leq \mathrm{Im} z \leq -v_0/2\}$ to get the convergence of $f_n(z, z_\ell)$ to 0, satisfying $\frac{\partial}{\partial z} f_n(z, z_l) \to 0$. From Lemma 2.4.1 we see that $\frac{\partial}{\partial z} f_n(z, w)$ is bounded in w and n for all $w \in D$. Applying again Lemma 2.4.1 on the remaining variable we see that $\frac{\partial^2}{\partial z \partial w} f_n(z, z_l) \to 0$. This proves (2.4.22).

Thus, it suffices to concentrate on the convergence of A_n which is the central object in the next section. □

2.4.1 Computation of the covariance of $A_n(z)$

In this section , we will calculate the limit of $A_n(z)$ defined in (2.4.20). Noticing $\Gamma_j = -z\tilde{b}_j\hat{\gamma}_j$, we first prove that

$$\begin{aligned} &\sum_{j=1}^{n} \mathbb{E}_{j-1}(\mathbb{E}_j z_1 \tilde{b}_j(z_1)\hat{\gamma}_j(z_1))(\mathbb{E}_j z_2 \tilde{b}_j(z_2)\hat{\gamma}_j(z_2)) \\ &\qquad - \sum_{j=1}^{n} z_1 z_2 \tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2)\mathbb{E}_{j-1}(\mathbb{E}_j\hat{\gamma}_j(z_1))(\mathbb{E}_j\hat{\gamma}_j(z_2)) \xrightarrow[n\to\infty]{\mathcal{P}} 0. \end{aligned} \tag{2.4.24}$$

Remark 2.4.2 *$z_1, z_2 \in \mathcal{C}_n$ are fixed in this section. Then $|z_i|$ and $|\mathrm{Im} z_i|$ are bounded for $i = 1, 2$. We shall not mention the dependence on $|z_i|$ or $|\mathrm{Im} z_i|$ in the sequel.*

From (2.4.17) and (2.4.18) in Theorem 2.4.1, $\mathbb{E}|\tilde{b}_j-\tilde{t}_{jj}|^2 \leq 2\mathbb{E}|\tilde{b}_j-\tilde{q}_{jj}|^2+2\mathbb{E}|\tilde{q}_{jj}-\tilde{t}_{jj}|^2 \leq K/n$. Therefore, by Cauchy-Schwarz and Lemma 2.3.3,

$$\begin{aligned}
&\mathbb{E}\Big|\mathbb{E}_{j-1}\{\mathbb{E}_j[z_1\tilde{b}_j(z_1)\hat{\gamma}_j(z_1)]\mathbb{E}_j[z_2\tilde{b}_j(z_2)\hat{\gamma}_j(z_2)]\}\\
&\qquad -z_1z_2\tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2)\mathbb{E}_{j-1}\{\mathbb{E}_j\hat{\gamma}_j(z_1)\mathbb{E}_j\hat{\gamma}_j(z_2)\}\Big|\\
&\quad\leq \mathbb{E}\Big|z_1z_2\mathbb{E}_{j-1}\{\mathbb{E}_j[(\tilde{b}_j(z_1)-\tilde{t}_{jj}(z_1))\hat{\gamma}_j(z_1)]\mathbb{E}_j[\tilde{b}_j(z_2)\hat{\gamma}_j(z_2)]\}\\
&\qquad -z_1z_2\tilde{t}_{jj}(z_1)\mathbb{E}_{j-1}\{\mathbb{E}_j[\hat{\gamma}_j(z_1)\mathbb{E}_j(\tilde{b}_{jj}(z_2)-\tilde{t}_{jj}(z_2))\hat{\gamma}_j(z_2)]\}\Big|\\
&\quad\leq K\mathbb{E}^{1/2}|\tilde{b}_j(z_1)-\tilde{t}_{jj}(z_1)|^2\mathbb{E}^{1/2}\left|\hat{\gamma}_j(z_1)\mathbb{E}_j[\hat{\gamma}_j(z_2)]\right|^2\\
&\qquad +K\mathbb{E}^{1/2}|\tilde{b}_j(z_2)-\tilde{t}_{jj}(z_2)|^2\mathbb{E}^{1/2}\left|\hat{\gamma}_j(z_2)\mathbb{E}_j[\hat{\gamma}_j(z_1)]\right|^2\\
&\quad\leq \frac{K\eta_n^2}{n}.
\end{aligned}$$

By summing the precedent inequalities over j, we achieve Equation (2.4.24).

Now we will calculate the covariance. The formula below plays a key role in the calculation.

Proposition 2.4.2 *[39, Equation (3.20)] Let $\mathbf{x}=(x_1,\cdots,x_N)^T$ be an $N\times 1$ vector where $\mathbf{x}_i$ are centered i.i.d. complex random variables with unit variance. Let $\mathbf{y}=N^{-1/2}\mathbf{x}$, $\mathbf{M}=(m_{ij})$ and $\mathbf{P}=(p_{ij})$ be $N\times N$ deterministic complex matrices and let $\mathbf{u}$ be an $N\times 1$ deterministic vector.* $\mathrm{vdiag}(\mathbf{M})$ *stands for the $N\times 1$ vector $[\mathbf{M}_{11},\cdots,\mathbf{M}_{NN}]^T$. Denote $\Upsilon(\mathbf{M})$, the random variable :*

$$\Upsilon(\mathbf{M})=(\mathbf{y}+\mathbf{u})^*\mathbf{M}(\mathbf{y}+\mathbf{u}).$$

Then $\mathbb{E}\Upsilon(\mathbf{M})=\frac{1}{N}\mathrm{Tr}\mathbf{M}+\mathbf{u}^*\mathbf{M}\mathbf{u}$ and the covariance between $\Upsilon(\mathbf{M})$ and $\Upsilon(\mathbf{P})$ is :

$$\begin{aligned}
&\mathbb{E}[(\Upsilon(\mathbf{M})-\mathbb{E}\Upsilon(\mathbf{M}))(\Upsilon(\mathbf{P})-\mathbb{E}\Upsilon(\mathbf{P}))]\\
&\quad=\frac{1}{N^2}\mathrm{Tr}(\mathbf{MP})+\frac{1}{N}(\mathbf{u}^*\mathbf{MPu}+\mathbf{u}^*\mathbf{PMu})+\frac{|\mathbb{E}[x_1]^2|^2}{N^2}\mathrm{Tr}(\mathbf{MP}^T)\\
&\qquad+\frac{\mathbb{E}[x_1^2]}{N}\mathbf{u}^*\mathbf{PM}^T\bar{\mathbf{u}}+\frac{\mathbb{E}[\bar{x}_1^2]}{N}\mathbf{u}^T\mathbf{M}^T\mathbf{Pu}\\
&\qquad\quad+\frac{\mathbb{E}[|x_1|^2x_1]}{N^{3/2}}(\mathbf{u}^*\mathbf{P}\mathrm{vdiag}(\mathbf{M})+\mathbf{u}^*\mathbf{M}\mathrm{vdiag}(\mathbf{P}))\\
&\qquad\qquad+\frac{\mathbb{E}[|x_1|^2\bar{x}_1]}{N^{3/2}}(\mathrm{vdiag}(\mathbf{P})^T\mathbf{Mu}+\mathrm{vdiag}(\mathbf{M})^T\mathbf{Pu})+\frac{\kappa}{N^2}\sum_{i=1}^N m_{ii}p_{ii},
\end{aligned}$$

where $\kappa=\mathbb{E}|x_1|^4-2-|\mathbb{E}x_1^2|^2$.

Noticing that $\mathbb{E}_j(\mathbf{Q}_j(z_1))$ is $\mathcal{F}_{n,j-1}$ measurable, by the precedent formula with $\mathbf{M}=\mathbb{E}_j(\mathbf{Q}_j(z_1))$, $\mathbf{P}=\mathbb{E}_j(\mathbf{Q}_j(z_2))$, we have

$$\begin{aligned}
&\sum_{j=1}^n z_1z_2\tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2)\mathbb{E}_{j-1}[\mathbb{E}_j\hat{\gamma}_j(z_1)\mathbb{E}_j\hat{\gamma}_j(z_2)]\\
&\triangleq\sum_{j=1}^n\xi_{1j}+\sum_{j=1}^n\xi_{2j}+\sum_{j=1}^n\xi_{2j}'+\sum_{j=1}^n\xi_{3j}+\sum_{j=1}^n\xi_{4j},
\end{aligned}$$

where

$$\sum_{j=1}^{n} \xi_{1j} \triangleq \frac{\kappa}{n^2} \sum_{j=1}^{n} z_1 z_2 \tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2) \sum_{i=1}^{N} \mathbb{E}_j[Q_j(z_1)]_{ii}\mathbb{E}_j[Q_j(z_2)]_{ii} \tag{2.4.25}$$

$$\sum_{j=1}^{n} \xi_{2j} \triangleq \frac{1}{n} \sum_{j=1}^{n} z_1 z_2 \tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2) \mathbb{E}(|x_{11}|^2 x_{11}) \times \tag{2.4.26}$$

$$\left(\frac{\mathbf{a}_j^*(\mathbb{E}_j \mathbf{Q}_j(z_1))\mathrm{vdiag}(\mathbb{E}_j \mathbf{Q}_j(z_2))}{\sqrt{n}} + \frac{\mathbf{a}_j^*(\mathbb{E}_j \mathbf{Q}_j(z_2))\mathrm{vdiag}(\mathbb{E}_j \mathbf{Q}_j(z_1))}{\sqrt{n}}\right)$$

$$\sum_{j=1}^{n} \xi'_{2j} \triangleq \frac{1}{n} \sum_{j=1}^{n} z_1 z_2 \tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2) \mathbb{E}(|x_{11}|^2 \bar{x}_{11}) \times \tag{2.4.27}$$

$$\left(\frac{\mathrm{vdiag}(\mathbb{E}_j \mathbf{Q}_j(z_2))^T(\mathbb{E}_j \mathbf{Q}_j(z_1))\mathbf{a}_j}{\sqrt{n}} + \frac{\mathrm{vdiag}(\mathbb{E}_j \mathbf{Q}_j(z_1))^T(\mathbb{E}_j \mathbf{Q}_j(z_2))\mathbf{a}_j}{\sqrt{n}}\right)$$

$$\sum_{j=1}^{n} \xi_{3j} \triangleq \frac{1}{n} \sum_{j=1}^{n} z_1 z_2 \tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2) \left(\frac{1}{n}\mathrm{Tr}(\mathbb{E}_j \mathbf{Q}_j(z_1))(\mathbb{E}_j \mathbf{Q}_j(z_2))\right. \tag{2.4.28}$$

$$\left. + \mathbf{a}_j^*(\mathbb{E}_j \mathbf{Q}_j(z_1))(\mathbb{E}_j \mathbf{Q}_j(z_2))\mathbf{a}_j + \mathbf{a}_j^*(\mathbb{E}_j \mathbf{Q}_j(z_2))(\mathbb{E}_j \mathbf{Q}_j(z_1))\mathbf{a}_j\right)$$

$$\sum_{j=1}^{n} \xi_{4j} \triangleq \frac{1}{n} \sum_{j=1}^{n} z_1 z_2 \tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2) \left(\frac{|\vartheta|^2}{n}\mathrm{Tr}(\mathbb{E}_j \mathbf{Q}_j(z_1))(\mathbb{E}_j \mathbf{Q}_j^T(z_2))\right. \tag{2.4.29}$$

$$\left. + \vartheta \mathbf{a}_j^*(\mathbb{E}_j \mathbf{Q}_j(z_2))(\mathbb{E}_j \mathbf{Q}_j^T(z_1))\bar{\mathbf{a}}_j + \bar{\vartheta}\mathbf{a}_j^T(\mathbb{E}_j \mathbf{Q}_j^T(z_1))(\mathbb{E}_j \mathbf{Q}_j(z_2))\mathbf{a}_j\right).$$

$\sum_{j=1}^n \xi_{1j}$ contributes the term with κ in (2.4.23). $\sum_{j=1}^n \xi_{2j}$ and $\sum_{j=1}^n \xi'_{2j}$ will not contribute in the final expression. $\sum_{j=1}^n \xi_{3j}$ corresponds to the term $\Theta_{0,n}$ while $\sum_{j=1}^n \xi_{4j}$ corresponds to Δ_n^{ϑ}. The first two terms are easy to compute. The computations for $\sum_{j=1}^n \xi_{3j}$ and $\sum_{j=1}^n \xi_{4j}$ are similar which demand more carefulness and a good understanding how the terms interact. We start the computations with $\sum_{j=1}^n \xi_{1j}$.

Lemma 2.4.2 *Recall the expression of $\sum_{j=1}^n \xi_{1j}$ in (2.4.25). We have*

$$\sum_{j=1}^{n} \xi_{1j} - \frac{\kappa z_1 z_2}{n^2} \sum_{i=1}^{N} t_{ii}(z_1)t_{ii}(z_2) \sum_{j=1}^{n} \tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2) \xrightarrow[N,n\to\infty]{\mathcal{P}} 0.$$

Proof: The idea is trying to replace $\mathbb{E}_j[Q_j]_{ii}$ by t_{ii}. Write

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{N} \mathbb{E}_j[Q_j(z_1)]_{ii}\mathbb{E}_j[Q_j(z_2)]_{ii} - \frac{1}{n}\sum_{i=1}^{N} \mathbb{E}_j[Q_j(z_1)]_{ii} t_{ii}(z_2) \\
=&\frac{1}{n}\sum_{i=1}^{N} \mathbb{E}_j[Q_j(z_1)]_{ii}\mathbb{E}_j([Q_j(z_2)]_{ii} - [Q(z_2)]_{ii}) + \frac{1}{n}\sum_{i=1}^{N} \mathbb{E}_j[Q_j(z_1)]_{ii}(\mathbb{E}_j[Q(z_2)]_{ii} - t_{ii}(z_2)) \\
\triangleq&\varepsilon_{1,j} + \varepsilon_{2,j}.
\end{aligned}$$

To estimate $\varepsilon_{1,j}$, we use rank-one perturbation lemma (see [45, Lemma 6.3] and [83, Lemma 2.6]).

Lemma 2.4.3 *The resolvent* $\mathbf{Q}$ *and the perturbed resolvent* $\mathbf{Q}_j$ *satisfy for* $z \neq 0$ *and any* $N \times N$ *matrix* $\mathbf{A}$ *:*

$$|\mathrm{Tr}\mathbf{A}(\mathbf{Q} - \mathbf{Q}_j)| \leq \frac{\|\mathbf{A}\|}{|\mathrm{Im}(z)|}.$$

Thanks to the rank-one perturbation lemma, the first term

$$\varepsilon_{1,j} = \frac{1}{n}|\mathbb{E}_j \mathrm{Tr}[\mathrm{diag}(\mathbb{E}_j \mathbf{Q}_j(z_1))(\mathbf{Q}_j(z_2) - \mathbf{Q}(z_2))]|$$

is with the order $\mathcal{O}(n^{-1})$. The second item

$$\mathbb{E}|\varepsilon_{2,j}| = \mathcal{O}(n^{-1/2})$$

is a direct consequence of Cauchy-Schwarz and (2.4.18) in Lemma 2.4.1. Hence

$$\sum_{j=1}^{n} \xi_{1j} - \frac{\kappa z_1 z_2}{n^2} \sum_{i=1}^{N} \sum_{j=1}^{n} \tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2)\mathbb{E}_j[Q_j(z_1)]_{ii} t_{ii}(z_2) \xrightarrow[n\to\infty]{\mathcal{P}} 0.$$

With the same method, we can replace $\mathbb{E}_j[Q_j]_{ii}(z_1)$ by $t_{ii}(z_1)$ successively without losing the order. This leads to the result. □

Lemma 2.4.4 *Recall the definition of* $\sum_{j=1}^n \xi_{2j}$ *and* $\sum_{j=1}^n \xi'_{2j}$ *in (2.4.26) and (2.4.27) respectively, we have*

$$\sum_{j=1}^{n} \xi_{2j} \xrightarrow[N,n\to\infty]{\mathcal{P}} 0,$$
$$\sum_{j=1}^{n} \xi'_{2j} \xrightarrow[N,n\to\infty]{\mathcal{P}} 0.$$

Proof: $\sum_{j=1}^n \xi_{2j}$ and $\sum_{j=1}^n \xi'_{2j}$ can be treated similarly, we focus on the proof of $\sum_{j=1}^n \xi_{2j}$.

From

$$\mathbb{E}\left|\sum_{j=1}^{n} \xi_{2j}\right| \leq \frac{K}{n}\sum_{j=1}^{n} \mathbb{E}\left|\frac{\mathbf{a}_j^*(\mathbb{E}_j \mathbf{Q}_j(z_1))\mathrm{vdiag}(\mathbf{Q}_j(z_2))}{\sqrt{n}}\right| + \frac{K}{n}\sum_{j=1}^{n} \mathbb{E}\left|\frac{\mathbf{a}_j^*(\mathbb{E}_j \mathbf{Q}_j(z_2))\mathrm{vdiag}(\mathbf{Q}_j(z_1))}{\sqrt{n}}\right|,$$

it suffices to consider the first term. The other term can be treated in the same way. Consider the term

$$\frac{1}{n}\sum_{j=1}^{n} \mathbb{E}\left|\frac{\mathbf{a}_j^*(\mathbb{E}_j \mathbf{Q}_j(z_1))\mathrm{vdiag}(\mathbf{Q}_j(z_2))}{\sqrt{n}}\right| \leq \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}\left|\frac{\mathbf{a}_j^*\mathbf{Q}_j(z_1)\mathrm{vdiag}(\mathbf{T}(z_2))}{\sqrt{n}}\right|$$

$$+\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left|\frac{\mathbf{a}_j^*(\mathbb{E}_j\mathbf{Q}_j(z_1))\mathrm{vdiag}(\mathbf{Q}(z_2)-\mathbf{T}(z_2))}{\sqrt{n}}\right|$$
$$+\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left|\frac{\mathbf{a}_j^*(\mathbb{E}_j\mathbf{Q}_j(z_1))\mathrm{vdiag}(\mathbf{Q}_j(z_2)-\mathbf{Q}(z_2))}{\sqrt{n}}\right|. \tag{2.4.30}$$

The first term in (2.4.30) satisfies

$$\sum_{j=1}^{n}\mathbb{E}\left|\frac{\mathbf{a}_j^*\mathbf{Q}_j(z_1)\mathrm{vdiag}(\mathbf{T}(z_2))}{\sqrt{n}}\right| \le \sqrt{n}\left(\sum_{j=1}^{n}\mathbb{E}\left|\frac{\mathbf{a}_j^*\mathbf{Q}_j(z_1)\mathrm{vdiag}(\mathbf{T}(z_2))}{\sqrt{n}}\right|^2\right)^{1/2}.$$

As $\|n^{-1/2}\mathrm{vdiag}(\mathbf{T})\| = (n^{-1}\sum_{i=1}^{n} t_{ii}^2)^{1/2} \le K$, (2.4.12) in Theorem 2.4.1 can be applied and

$$\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left|\frac{\mathbf{a}_j^*\mathbf{Q}_j(z_1)\mathrm{vdiag}(\mathbf{T}(z_2))}{\sqrt{n}}\right| = \mathcal{O}(n^{-1/2}).$$

We now deal with the second term in (2.4.30) :

$$\begin{aligned}
&\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left|\frac{\mathbf{a}_j^*(\mathbb{E}_j\mathbf{Q}_j(z_1))\mathrm{vdiag}(\mathbf{Q}(z_2)-\mathbf{T}(z_2))}{\sqrt{n}}\right| \\
&\quad\le\quad K\mathbb{E}\left\|\frac{\mathrm{vdiag}(\mathbf{Q}(z_2)-\mathbf{T}(z_2))}{\sqrt{n}}\right\| \\
&\quad\le\quad K\left(\frac{1}{n}\sum_{i=1}^{N}\mathbb{E}|q_{ii}(z_2)-t_{ii}(z_2)|^2\right)^{1/2} \\
&\quad=\quad \mathcal{O}(n^{-1/2}),
\end{aligned}$$

by Equation (2.4.18) in Theorem 2.4.1. Since $\|\mathbf{a}_j^*(\mathbb{E}_j\mathbf{Q}_j)\|$ is uniformly bounded by K, the third term in (2.4.30) can be estimated by using Lemma 2.4.3 :

$$\begin{aligned}
&\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left|\frac{\mathbf{a}_j^*(\mathbb{E}_j\mathbf{Q}_j(z_1))\mathrm{vdiag}(\mathbf{Q}_j(z_2)-\mathbf{Q}(z_2))}{\sqrt{n}}\right| \\
&\quad=\quad \frac{1}{n^{3/2}}\sum_{j=1}^{n}\mathbb{E}|\mathrm{Tr}[\mathrm{diag}(\mathbf{a}_j^*(\mathbb{E}_j\mathbf{Q}_j(z_1)))(\mathbf{Q}_j(z_2)-\mathbf{Q}(z_2))]| \\
&\quad\le\quad \frac{K}{\sqrt{n}}.
\end{aligned}$$

□

Now we concentrate on the proof of the following lemma whose idea is from [39, Lemma 4.3].

Lemma 2.4.5 *Recall the expression of $\sum_{j=1}^{n}\xi_{3j}$ in (2.4.28). Under the same assumptions, we* *have*

$$\frac{\partial^2}{\partial z_1 \partial z_2} \sum_{j=1}^n \xi_{3j} - \frac{s_n'(z_1)s_n'(z_2)}{(s_n(z_1)-s_n(z_2))^2} + \frac{1}{(z_1-z_2)^2} \xrightarrow[n\to\infty]{\mathcal{P}} 0,$$

where $s_n(z) = z(1+\delta)(1+\tilde{\delta})$.

Proof: For future use, for a Hermitian matrix $\mathbf{M}$ with bounded spectral norm, consider the notations :

$$\begin{aligned}
\psi_j^{\mathbf{M}}(z_1,z_2) &= \frac{1}{n}\mathrm{Tr}\mathbb{E}\left\{[\mathbb{E}_j\mathbf{Q}(z_1)]\mathbf{M}[\mathbb{E}_j\mathbf{Q}(z_2)]\right\} = \frac{1}{n}\mathrm{Tr}\mathbb{E}\{[\mathbb{E}_j\mathbf{Q}(z_1)]\mathbf{M}\mathbf{Q}(z_2)\},\\
\zeta_{kj}^{\mathbf{M}}(z_1,z_2) &= \mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}(z_1)]\mathbf{M}[\mathbb{E}_j\mathbf{Q}(z_2)]\mathbf{a}_k\} = \mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}(z_1)]\mathbf{M}\mathbf{Q}(z_2)\mathbf{a}_k\},\\
\theta_{kj}^{\mathbf{M}}(z_1,z_2) &= \mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}_k(z_1)]\mathbf{M}[\mathbb{E}_j\mathbf{Q}_k(z_2)\mathbf{a}_k]\} = \mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}_k(z_1)]\mathbf{M}\mathbf{Q}_k(z_2)\mathbf{a}_k\},\\
\phi_j^{\mathbf{M}}(z_1,z_2) &= \frac{1}{n}\sum_{k=1}^j z_1 z_2 \tilde{t}_{kk}(z_1)\tilde{t}_{kk}(z_2)\theta_{kj}^{\mathbf{M}}.
\end{aligned}$$

Remark 2.4.3 *In particular, we will take* $\mathbf{M} = \mathbf{I}_N$ *to calculate the variance and* $\mathbf{M} = \mathbf{T}$ *in the calculation of the bias.*

We will decompose $\sum_{j=1}^n \xi_{3j}$. The following lemma allows to replace $\sum_{j=1}^n \xi_{3j}$ by its expectation.

Lemma 2.4.6 *For any* $N \times 1$ *vector* $\mathbf{a}$ *with bounded Euclidean norm, we have*

$$\max_j \mathrm{Var}[\mathbf{a}^*\mathbb{E}_j\mathbf{Q}(z_1)\mathbf{M}\mathbb{E}_j\mathbf{Q}(z_2)\mathbf{a}] = o(n^{-1/2})$$

and

$$\max_j \mathrm{Var}[\mathrm{Tr}\mathbb{E}_j\mathbf{Q}(z_1)\mathbf{M}\mathbb{E}_j\mathbf{Q}(z_2)] = o(n^{1/2}).$$

The proof is postponed to Appendix 2.6.6.

Remark 2.4.4 *In [39, Lemma 5.1], with the hypothesis* $\mathbb{E}|x_{ij}|^{16} < \infty$*, it has been shown that* $\max_j \mathrm{Var}[\mathbf{a}^*\mathbb{E}_j\mathbf{Q}(z_1)\mathbf{M}\mathbb{E}_j\mathbf{Q}(z_2)\mathbf{a}] = o(n^{-1/2}) = \mathcal{O}(n^{-1})$ *which is stronger than the precedent bound. However, Lemma 2.4.6 is sufficient to ensure the replacement of* $\sum_{j=1}^n \xi_{3j}$ *by its expectation.*

Remark 2.4.5 *Although the computation is a little lengthy, there is a key rule during the computation. Let* $\mathbf{y}_j = \mathbf{r}_j + \mathbf{a}_j$*. For matrices* $\mathbf{M}_1$*,* $\mathbf{M}_2$ *with bounded spectral norms and independent of* $\mathbf{y}_j$*. Consider the term*

$$\mathbb{E}\mathbf{r}_j^*\mathbf{M}_1\mathbf{y}_j\mathbf{y}_j^*\mathbf{M}_2\mathbf{a}_j.$$

Developing $\mathbf{y}_j\mathbf{y}_j = \mathbf{r}_j\mathbf{r}_j^* + \mathbf{a}_j\mathbf{r}_j^* + \mathbf{r}_j\mathbf{a}_j^* + \mathbf{a}_j\mathbf{a}_j^*$*, we will show that the only term which is not negligible is* $\mathbf{r}_j\mathbf{a}_j^*$*, then we have :*

$$\mathbb{E}\mathbf{r}_j^*\mathbf{M}_1\mathbf{y}_j\mathbf{y}_j^*\mathbf{M}_2\mathbf{a}_j = \mathbb{E}\mathbf{r}_j^*\mathbf{M}_1\mathbf{r}_j\mathbf{a}_j^*\mathbf{M}_2\mathbf{a}_j + o(1).$$

By Lemma 2.4.6, $\mathbf{r}_j^*\mathbf{M}_1\mathbf{r}_j$ and $\mathbf{a}_j^*\mathbf{M}_2\mathbf{a}_j$ are uncorrelated, and we get

$$\mathbb{E}\mathbf{r}_j^*\mathbf{M}_1\mathbf{y}_j\mathbf{y}_j^*\mathbf{M}_2\mathbf{a}_j = \frac{1}{n}\mathbb{E}\mathrm{Tr}\mathbf{M}_1\mathbb{E}\mathbf{a}_j^*\mathbf{M}_2\mathbf{a}_j + o(1).$$

The rest of the computation is the use of identities. We will be more precise during the proof.

By Lemma 2.4.6 and 2.4.3, observe that

$$\begin{aligned}\sum_{i=1}^n \xi_{3j} =& \frac{1}{n}\sum_{j=1}^n z_1 z_2 \tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2)\left(\frac{1}{n}\mathrm{Tr}\mathbb{E}[\mathbb{E}_j\mathbf{Q}(z_1)\mathbb{E}_j\mathbf{Q}(z_2)]\right.\\ &\left.+\theta_{jj}^{\mathbf{I}_N}(z_1,z_2)+\theta_{jj}^{\mathbf{I}_N}(z_2,z_1)\right)+o(n^{-1/2}),\end{aligned}$$

Taking $\mathbf{M} = \mathbf{I}_N$, it suffices to study

$$\frac{1}{n}\sum_{j=1}^n z_1 z_2 \tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2)\left(\psi_j^{\mathbf{I}_N}+\theta_{jj}^{\mathbf{I}_N}(z_1,z_2)+\theta_{jj}^{\mathbf{I}_N}(z_2,z_1)\right).$$

To simplify the notations, all terms which depend on z_1(resp. z_2) will be denoted by $\mathbf{Q}_1 = \mathbf{Q}(z_1)$, $\mathbf{T}_1 = \mathbf{T}(z_1)$, $\mathbf{Q}_{1,\ell} = \mathbf{Q}_\ell(z_1)$, (resp. $\mathbf{Q}_2 = \mathbf{Q}(z_2)$, $\mathbf{T}_2 = \mathbf{T}(z_2)$,$\mathbf{Q}_{2,\ell} = \mathbf{Q}_\ell(z_2)$) etc. We will look for the equations verified by $\zeta_{kj}^{\mathbf{M}}$, $\psi_j^{\mathbf{M}}$ and $\phi_j^{\mathbf{M}}$.

The idea is to have the terms which are displayed in Remark 2.4.5. There are four steps. In the first step, starting with $\zeta_{kj}^{\mathbf{M}}$, with $\mathbf{A}^{-1}-\mathbf{B}^{-1} = -\mathbf{A}^{-1}(\mathbf{A}-\mathbf{B})\mathbf{B}^{-1}$, we obtain the first relation satisfied by $\zeta_{kj}^{\mathbf{M}}$, $\psi_j^{\mathbf{M}}$ and $\phi_j^{\mathbf{M}}$. In the second step, beginning with ψ_j, we search for the relation between $\psi_j^{\mathbf{M}}$ and $\phi_j^{\mathbf{M}}$. In Step 3, we start again with $\zeta_{kj}^{\mathbf{M}}$, with (2.4.6), we achieve the third equation with $\zeta_{kj}^{\mathbf{M}}$, $\psi_j^{\mathbf{M}}$ and $\phi_j^{\mathbf{M}}$. Hence, we obtain three equations with three parameters. This allows us to conclude the computation which is the goal in Step 4.

Step 1 : Expression of $\zeta_{kj}^{\mathbf{M}}$

In this section, we will show that : $\zeta_{kj}^{\mathbf{M}}, \psi_j^{\mathbf{M}}, \phi_j^{\mathbf{M}}$ satisfy

$$\begin{aligned}\zeta_{kj}^{\mathbf{M}} =& \mathbf{a}_k^*\mathbf{T}_1\mathbf{M}\mathbf{T}_2\mathbf{a}_k + \psi_j^{\mathbf{M}}\left(\sum_{\ell=1}^j \frac{\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_2\mathbf{a}_k}{(1+\delta_1)(1+\delta_2)} + \mathbf{a}_k^*\mathbf{T}_1\mathbf{T}_2\mathbf{a}_k\frac{1}{n}\sum_{\ell=1}^j z_1 z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\right)\\ &+\mathbf{a}_k^*\mathbf{T}_1\mathbf{T}_2\mathbf{a}_k\phi_j^{\mathbf{M}} + \mathcal{O}(n^{-1/2}).\end{aligned} \tag{2.4.31}$$

Writing

$$\mathbf{Q} = \mathbf{T} + \mathbf{T}(\mathbf{T}^{-1}-\mathbf{Q}^{-1})\mathbf{Q} = \mathbf{T} + \mathbf{T}(-z\tilde{\delta}\mathbf{I}_N + (1+\delta)^{-1}\mathbf{A}\mathbf{A}^* - \mathbf{Y}\mathbf{Y}^*)^{-1}\mathbf{Q}, \tag{2.4.32}$$

we have

$$
\begin{aligned}
\zeta_{kj}^{\mathbf{M}} &= \mathbb{E}[\mathbf{a}_k^*(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}(\mathbb{E}_j\mathbf{Q}_2)\mathbf{a}_k] \\
&= \mathbb{E}[\mathbf{a}_k^*(\mathbb{E}_j[\mathbf{T}_1 + \mathbf{T}_1(-z_1\tilde{\delta}_1\mathbf{I}_N + (1+\delta_1)^{-1}\mathbf{A}\mathbf{A}^* - \mathbf{Y}\mathbf{Y}^*)\mathbf{Q}_1]\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] \\
&= \mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] - z_1\tilde{\delta}_1\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] \\
&\quad + (1+\delta_1)^{-1}\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{A}\mathbf{A}^*(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] - \mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1(\mathbb{E}_j\mathbf{Y}\mathbf{Y}^*\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] \\
&= \mathbf{a}_k^*\mathbf{T}_1\mathbf{M}\mathbf{T}_2\mathbf{a}_k - z_1\tilde{\delta}_1\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] + X + Z + \varepsilon,
\end{aligned}
\tag{2.4.33}
$$

where

$$
X \triangleq (1+\delta_1)^{-1}\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{A}\mathbf{A}^*(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{a}_k], \tag{2.4.34}
$$

$$
Z \triangleq -\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1(\mathbb{E}_j\mathbf{Y}\mathbf{Y}^*\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{a}_k], \tag{2.4.35}
$$

and $|\varepsilon| = \mathcal{O}(n^{-1/2})$ by (2.4.14) in Theorem 2.4.1. Recall $\mathbf{y}_j = \mathbf{r}_j + \mathbf{a}_j$, the term X is

$$
\begin{aligned}
X &= (1+\delta_1)^{-1}\sum_{\ell=1}^{n}\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] \\
&= (1+\delta_1)^{-1}\sum_{\ell=1}^{n}\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] \\
&\quad +(1+\delta_1)^{-1}\sum_{\ell=1}^{n}\mathbb{E}[z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell}\mathbf{y}_\ell\mathbf{y}_\ell^*\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] + \epsilon_1 \\
&= (1+\delta_1)^{-1}\sum_{\ell=1}^{n}\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] \\
&\quad +(1+\delta_1)^{-1}\sum_{\ell=1}^{n}\mathbb{E}[z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,l}\mathbf{a}_\ell\mathbf{y}_\ell^*\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] + \epsilon_1 + \epsilon_2 \\
&= (1+\delta_1)^{-1}\sum_{\ell=1}^{n}\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,l})\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] \\
&\quad +(1+\delta_1)^{-1}\sum_{\ell=1}^{n}\mathbb{E}[z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell(\mathbb{E}_j\mathbf{y}_\ell^*\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] + \epsilon_1 + \epsilon_2 + \epsilon_3 \\
&\triangleq X_1 + X_2 + \epsilon_1 + \epsilon_2 + \epsilon_3,
\end{aligned}
\tag{2.4.36}
$$

where

$$
\begin{aligned}
\epsilon_1 &\triangleq (1+\delta_1)^{-1}\sum_{\ell=1}^{n}\mathbb{E}[\tilde{t}_{1,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell(\mathbb{E}_j(z_1\tilde{q}_{1,\ell\ell} - z_1\tilde{t}_{1,\ell\ell})\mathbf{a}_\ell^*\mathbf{Q}_{1,\ell}\mathbf{y}_\ell\mathbf{y}_\ell^*\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{a}_k], \\
\epsilon_2 &\triangleq (1+\delta_1)^{-1}\sum_{\ell=1}^{n}\mathbb{E}[z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell}\mathbf{r}_\ell\mathbf{y}_\ell^*\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{a}_k], \\
\epsilon_3 &\triangleq (1+\delta_1)^{-1}\sum_{\ell=1}^{n}\mathbb{E}[z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell(\mathbb{E}_j\mathbf{a}_\ell^*(\mathbf{Q}_{1,\ell} - \mathbf{T}_{1,\ell})\mathbf{a}_\ell\mathbf{y}_\ell^*\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{a}_k].
\end{aligned}
$$

Using Formula (2.4.3) and (2.4.9),

$$\epsilon_1 = (1+\delta_1)^{-1}\mathbb{E}[\mathbb{E}_j(\mathbf{a}_k^*\mathbf{T}_1\mathbf{A}\mathrm{diag}(\chi_{1,\ell})\mathbf{Y}^*\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{a}_k], \tag{2.4.37}$$

where $\chi_{1,\ell} = -z_1(\tilde{q}_{1,\ell\ell} - \tilde{t}_{1,\ell\ell})(1+\mathbf{y}_\ell^*\mathbf{Q}_{1,\ell}\mathbf{y}_\ell)\mathbf{a}_\ell^*\mathbf{Q}_{1,\ell}\mathbf{y}_\ell$.

Notice that

$$\|\mathbf{Y}^*\mathbf{Q}\| \le \|\mathbf{Y}\|\|\mathbf{Q}\| < \infty,$$

where by truncation, $\|\mathbf{Y}\| < \infty$ and $\|\mathbf{Q}\| \le |\mathrm{Im} z|^{-1}$. The same method yields that $\|\mathbf{Y}^*\mathbf{Q}_j\|$ is also bounded and we obtain

$$\begin{aligned}
|\epsilon_1| &\le K\mathbb{E}\|\mathbf{a}_k^*\mathbf{T}_1\mathbf{A}\mathrm{diag}(\chi_{1,\ell})\| \\
&\le K\mathbb{E}^{1/2}\left[\sum_{\ell=1}^n |[a_k^*T_1A]_\ell|^2|\chi_{1,\ell}|^2\right] \\
&\le K\mathbb{E}^{1/2}\left[\sum_{\ell=1}^n |[a_k^*T_1A]_\ell|^2|\tilde{q}_{1,\ell\ell} - \tilde{t}_{1,\ell\ell}|^2|\mathbf{a}_\ell^*\mathbf{Q}_{1,\ell}\mathbf{y}_\ell|^2\right] \\
&\le K\mathbb{E}^{1/2}\left[\sum_{\ell=1}^n |[a_k^*T_1A]_\ell|^2|\tilde{q}_{1,\ell\ell} - \tilde{t}_{1,\ell\ell}|^2\right] \\
&\le \frac{K}{\sqrt{n}},
\end{aligned}$$

where the last inequality follows from (2.4.18) in Theorem 2.4.1. A similar result for ϵ_2 and ϵ_3 can be shown that $\epsilon_2 = \mathcal{O}(n^{-1/2})$ and $\epsilon_3 = \mathcal{O}(n^{-1/2})$. We now develop X_2,

$$\begin{aligned}
X_2 &= (1+\delta_1)^{-1}\sum_{\ell=1}^n \mathbb{E}[z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] \\
&\quad + (1+\delta_1)^{-1}\sum_{\ell=1}^j \mathbb{E}[z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] \\
&\triangleq U_1 + U_2.
\end{aligned} \tag{2.4.38}$$

By formula (2.4.6), the term U_2 can be expressed as

$$\begin{aligned}
U_2 &= (1+\delta_1)^{-1}\sum_{\ell=1}^j \mathbb{E}[z_1z_2\tilde{t}_{1,\ell\ell}\tilde{q}_{2,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{y}_\ell\mathbf{y}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k] \\
&= \sum_{\ell=1}^j z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell\frac{\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{y}_\ell\mathbf{y}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k]}{(1+\delta_1)} + \mathcal{O}(n^{-1/2}).
\end{aligned}$$

We are in the framework of Remark 2.4.5. Write $\mathbf{y}_\ell\mathbf{y}_\ell^* = \mathbf{a}_\ell\mathbf{a}_\ell^* + \mathbf{a}_\ell\mathbf{r}_\ell^* + \mathbf{r}_\ell\mathbf{r}_\ell^* + \mathbf{r}_\ell\mathbf{a}_\ell^*$, we will show that the only term which is not negligible is $\mathbf{r}_\ell\mathbf{a}_\ell^*$. The term in $\mathbf{a}_\ell\mathbf{a}_\ell^*$ is :

$$\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k] = 0.$$

Turning to the term $\mathbf{a}_\ell \mathbf{r}_\ell^*$, we have $\mathbb{E}|\mathbf{r}_\ell^*(\mathbb{E}_j \mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{a}_\ell \mathbf{r}_\ell^* \mathbf{Q}_{2,\ell}\mathbf{a}_k| = \mathcal{O}(n^{-1})$, hence

$$\sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell \frac{\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j \mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{a}_\ell \mathbf{r}_\ell^* \mathbf{Q}_{2,\ell}\mathbf{a}_k]}{(1+\delta_1)}$$

$$\leq \frac{K}{n}\sum_{\ell=1}^{j}|\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell| \leq \frac{K}{\sqrt{n}}\left(\sum_{\ell=1}^{j}|\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell|^2\right)^{1/2} \leq \frac{K}{\sqrt{n}}.$$

Moreover, for the term in $\mathbf{r}_j\mathbf{r}_j^*$, we have

$$\sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell \frac{\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j \mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,l}\mathbf{r}_\ell \mathbf{r}_\ell^* \mathbf{Q}_{2,\ell}\mathbf{a}_k]}{(1+\delta_1)}$$

$$= \sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell \frac{\mathbb{E}\{[\mathbf{r}_\ell^*(\mathbb{E}_j \mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell - n^{-1}\mathrm{Tr}(\mathbb{E}_j \mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}]\mathbf{r}_\ell^* \mathbf{Q}_{2,\ell}\mathbf{a}_k\}}{(1+\delta_1)}$$

$$\leq \frac{K}{n}\sum_{\ell=1}^{j}|\mathbf{a}_k^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell| = \mathcal{O}(n^{-1/2}).$$

The term in $\mathbf{r}_\ell \mathbf{a}_\ell^*$ is written as

$$\sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell \frac{\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j \mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell \mathbf{a}_\ell^* \mathbf{Q}_{2,\ell}\mathbf{a}_k]}{(1+\delta_1)}$$

$$\stackrel{(a)}{=} \sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell \frac{\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j \mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell]\mathbb{E}[\mathbf{a}_\ell^* \mathbf{Q}_{2,\ell}\mathbf{a}_k]}{(1+\delta_1)} + \mathcal{O}(n^{-1/2})$$

$$\stackrel{(b)}{=} \psi_j^{\mathbf{M}} \sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell \frac{\mathbb{E}[\mathbf{a}_\ell^* \mathbf{Q}_{2,\ell}\mathbf{a}_k]}{(1+\delta_1)} + \mathcal{O}(n^{-1/2}),$$

where (a) is deduced from the fact that $\mathbf{r}_\ell^*(\mathbb{E}_j \mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell$ and $\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k$ are uncorrelated (cf. Lemma 2.4.6) and (b) is from Lemma 2.4.3. The remaining term in the right hand side can be handled by the following lemma.

Lemma 2.4.7 *Let* $(\mathbf{u}) = (u_n)_{n\in\mathbb{N}}$ *be a sequence of vectors with bounded Euclidean norms. Let* $(\alpha_\ell)_{1\leq\ell\leq n} = (\alpha_{\ell,n})_{1\leq\ell\leq n}$ *be an array of bounded real numbers. Then*

$$\sum_{\ell=1}^{j}\alpha_\ell \mathbf{u}^*\mathbf{T}_1\mathbf{a}_\ell\mathbb{E}[\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{u}] = -\sum_{\ell=1}^{j}\frac{\alpha_\ell\mathbf{u}^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_2\mathbf{u}}{z_2\tilde{t}_{2,\ell\ell}(1+\delta_2)} + \mathcal{O}(n^{-1/2}).$$

The proof is postponed in Appendix 2.6.7.

Applying the precedent lemma with $\mathbf{u} = \mathbf{a}_k$ and $\alpha_\ell = z_1 z_2 \tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell$, we obtain

$$U_2 = -\psi_j^{\mathbf{M}}\sum_{\ell=1}^{j}\frac{z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_2\mathbf{a}_k}{(1+\delta_1)(1+\delta_2)} + \mathcal{O}(n^{-1/2}).$$

Gathering (2.4.36) and (2.4.38), and using the identity (2.4.10), we obtain

$$\begin{aligned} X = &- \sum_{\ell=1}^{n} z_1 \tilde{t}_{1,\ell\ell} \mathbf{a}_k^* \mathbf{T}_1 \mathbf{a}_\ell \mathbb{E}[\mathbf{a}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k] \\ &- \psi_j^{\mathbf{M}} \sum_{\ell=1}^{j} \frac{z_1 \tilde{t}_{1,\ell\ell} \mathbf{a}_k^* \mathbf{T}_1 \mathbf{a}_\ell \mathbf{a}_\ell^* \mathbf{T}_{1,\ell} \mathbf{a}_\ell \mathbf{a}_\ell^* \mathbf{T}_2 \mathbf{a}_k}{(1+\delta_1)(1+\delta_2)} + \mathcal{O}(n^{-1/2}). \end{aligned} \tag{2.4.39}$$

We now turn to the term Z in (2.4.35). By (2.4.9),

$$\begin{aligned} Z &= -\mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 (\mathbb{E}_j \mathbf{Y}\mathbf{Y}^* \mathbf{Q}_1) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k] \\ &= \sum_{\ell=1}^{n} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 (\mathbb{E}_j z_1 \tilde{q}_{1,\ell\ell} \mathbf{y}_\ell \mathbf{y}_\ell^* \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k] \\ &= z_1 \sum_{\ell=1}^{n} \tilde{t}_{1,\ell\ell} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 (\mathbb{E}_j \mathbf{y}_\ell \mathbf{y}_\ell^* \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k] + \epsilon, \end{aligned}$$

where

$$\epsilon = -\sum_{\ell=1}^{n} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 (\mathbb{E}_j z (\tilde{q}_{1,\ell\ell} - \tilde{t}_{1,\ell\ell}) \mathbf{y}_\ell \mathbf{y}_\ell^* \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k]$$

satisfies $\epsilon = \mathcal{O}(n^{-1/2})$ (cf. the argument for ϵ_1 in (2.4.37)). Writing $\mathbf{y}_\ell \mathbf{y}_\ell^* = \mathbf{a}_\ell \mathbf{a}_\ell^* + \mathbf{r}_\ell \mathbf{r}_\ell^* + \mathbf{a}_\ell \mathbf{r}_\ell^* + \mathbf{r}_\ell \mathbf{a}_\ell^*$, we obtain :

$$\begin{aligned} Z = &\, z_1 \sum_{\ell=1}^{n} \tilde{t}_{1,\ell\ell} \mathbf{a}_k^* \mathbf{T}_1 \mathbf{a}_\ell \mathbb{E}[\mathbf{a}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k] \\ &+ z_1 \left(\sum_{\ell=1}^{j} \tilde{t}_{1,\ell\ell} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 \mathbf{r}_\ell \mathbf{r}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k] + \frac{1}{n} \sum_{\ell=j+1}^{n} \tilde{t}_{1,\ell\ell} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k] \right) \\ &+ z_1 \sum_{\ell=1}^{j} \tilde{t}_{1,\ell\ell} \mathbf{a}_k^* \mathbf{T}_1 \mathbf{a}_\ell \mathbb{E}[\mathbf{r}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k] \\ &+ z_1 \sum_{\ell=1}^{n} \tilde{t}_{1,\ell\ell} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 \mathbf{r}_\ell \mathbf{a}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k] + \mathcal{O}(n^{-1/2}) \\ \triangleq &\, Z_1 + Z_2 + Z_3 + Z_4 + \mathcal{O}(n^{-1/2}). \end{aligned} \tag{2.4.40}$$

The term Z_1 cancels with the first term in the decomposition of X (cf. (2.4.39)). By developing $\mathbf{Q}_2$ with Formula (2.4.6), the term Z_2 can be written as :

$$
\begin{aligned}
Z_2 &= \left(z_1 \sum_{\ell=1}^{j} \tilde{t}_{1,\ell\ell} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 \mathbf{r}_\ell \mathbf{r}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_{2,\ell} \mathbf{a}_k] + \frac{1}{n} z_1 \sum_{\ell=j+1}^{n} \tilde{t}_{1,\ell\ell} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k] \right) \\
&\quad + \sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell} \tilde{t}_{2,\ell\ell} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 \mathbf{r}_\ell \mathbf{r}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_{2,\ell} \mathbf{y}_\ell \mathbf{y}_\ell^* \mathbf{Q}_{2,\ell} \mathbf{a}_k] + \epsilon \\
&\triangleq W_1 + W_2 + \epsilon,
\end{aligned}
\tag{2.4.41}
$$

where $\epsilon = \mathcal{O}(n^{-1/2})$ can be shown as (2.4.37). Consider first W_1 :

$$
W_1 = \left(\frac{z_1}{n} \sum_{\ell=1}^{j} \tilde{t}_{1,\ell\ell} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_{2,\ell} \mathbf{a}_k] + \frac{z_1}{n} \sum_{\ell=j+1}^{n} \tilde{t}_{1,\ell\ell} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k] \right).
$$

Denote $\mathbf{Y}_{1:j} = (\mathbf{y}_1, \cdots, \mathbf{y}_j)$. Writing :

$$
(\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_{2,\ell} - (\mathbb{E}_j \mathbf{Q}_1) \mathbf{M} \mathbf{Q}_2 = (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} (\mathbf{Q}_{2,\ell} - \mathbf{Q}_2) + (\mathbb{E}_j \mathbf{Q}_{1,\ell} - \mathbb{E}_j \mathbf{Q}_1) \mathbf{M} \mathbf{Q}_2
$$

and using (2.4.7) and (2.4.8), we have

$$
\begin{aligned}
&\left| \frac{z_1}{n} \sum_{\ell=1}^{j} \tilde{t}_{1,\ell\ell} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} (\mathbf{Q}_{2,\ell} - \mathbf{Q}_2) \mathbf{a}_k] \right| \\
&\quad = \left| \frac{z_1}{n} \sum_{\ell=1}^{j} \tilde{t}_{1,\ell\ell} \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} (1 + \mathbf{y}_j^* \mathbf{Q}_{2,j} \mathbf{y}_j) \mathbf{Q}_2 \mathbf{y}_j \mathbf{y}_j^* \mathbf{Q}_2 \mathbf{a}_k] \right| \\
&\quad \leq \frac{K}{n} \sum_{\ell=1}^{j} (\mathbb{E}|1 + \mathbf{y}_\ell^* \mathbf{Q}_2 \mathbf{y}_\ell|^2)^{1/2} (\mathbb{E}|\mathbf{y}_\ell^* \mathbf{Q}_2 \mathbf{a}_k|^2)^{1/2} \\
&\quad \leq \frac{K}{n} \left(\sum_{\ell=1}^{j} \mathbb{E}|1 + \mathbf{y}_\ell^* \mathbf{Q}_2 \mathbf{y}_\ell|^2 \right)^{1/2} \left(\sum_{j=1}^{n} \mathbb{E}|\mathbf{y}_\ell^* \mathbf{Q}_2 \mathbf{a}_k|^2 \right)^{1/2} \\
&\quad \leq \frac{K}{\sqrt{n}} (\mathbb{E} \mathbf{a}_k^* \mathbf{Q}_2 \mathbf{Y}_{1:j} \mathbf{Y}_{1:j}^* \mathbf{Q}_2^* \mathbf{a}_k)^{1/2} \\
&\quad = \mathcal{O}(n^{-1/2}),
\end{aligned}
$$

and the same argument applies to the term $(\mathbb{E}_j \mathbf{Q}_{1,\ell} - \mathbb{E}_j \mathbf{Q}_1) \mathbf{M} \mathbf{Q}_2$. Hence,

$$
W_1 = z_1 \tilde{\delta}_1 \mathbb{E}[\mathbf{a}_k^* \mathbf{T}_1 (\mathbb{E}_j \mathbf{Q}_1) \mathbf{M} \mathbf{Q}_2 \mathbf{a}_k] + \mathcal{O}(n^{-1/2}). \tag{2.4.42}
$$

Turning to W_2 and developing $\mathbf{y}_\ell\mathbf{y}_\ell^* = \mathbf{a}_\ell\mathbf{a}_\ell^* + \mathbf{a}_\ell\mathbf{r}_\ell^* + \mathbf{r}_\ell\mathbf{a}_\ell^* + \mathbf{r}_\ell\mathbf{r}_\ell^*$, for the term $\mathbf{r}_\ell\mathbf{a}_\ell^*$, we have

$$\begin{aligned}&\sum_{\ell=1}^{j} z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{r}_\ell\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k]\\ &=\sum_{\ell=1}^{j} z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbb{E}\left[\mathbf{a}_k^*\mathbf{T}_1\mathbf{r}_\ell\left(\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell - \frac{1}{n}\mathrm{Tr}(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\right)\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k\right],\end{aligned}$$

whose modulus is of order $\mathcal{O}(n^{-1/2})$. The estimate

$$\sum_{\ell=1}^{j} z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{r}_\ell\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{a}_\ell\mathbf{r}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k] = \mathcal{O}(n^{-1/2})$$

can be handled similarly and the term

$$\begin{aligned}&\sum_{\ell=1}^{j} z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{r}_\ell\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k]\\ &\quad = \frac{1}{n}\sum_{\ell=1}^{j} z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k]\end{aligned}$$

is bounded by $Kn^{-1/2}$. Finally,

$$\begin{aligned}W_2 &= \sum_{\ell=1}^{j} z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{r}_\ell\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell\mathbf{r}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k] + \mathcal{O}(n^{-1/2})\\ &\overset{(a)}{=} \sum_{\ell=1}^{j} z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell]\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{r}_\ell\mathbf{r}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k] + \mathcal{O}(n^{-1/2}) \qquad (2.4.43)\\ &\overset{(b)}{=} \psi_j^{\mathbf{M}}\mathbf{a}_k^*\mathbf{T}_1\mathbf{T}_2\mathbf{a}_k\frac{1}{n}\sum_{\ell=1}^{j} z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell} + \mathcal{O}(n^{-1/2}),\end{aligned}$$

where (a) follows from the fact that $\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell$ and $\mathbf{a}_k^*\mathbf{T}_1\mathbf{r}_\ell\mathbf{r}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k$ are uncorrelated and (b) is from Lemma 2.4.3 and Equation (2.4.15) (cf. Remark 2.4.5).

The term Z_3 satisfies

$$Z_3 = z_1z_2\sum_{\ell=1}^{j}\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{y}_\ell\mathbf{y}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k].$$

We are again in the framework of Remark 2.4.5. By developing $\mathbf{y}_\ell\mathbf{y}_\ell^*$ and relying on arguments as those already used, only the term with $\mathbf{r}_j\mathbf{a}_j^*$ is not negligible and we have :

$$
\begin{aligned}
Z_3 &= z_1z_2\sum_{\ell=1}^{j}\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k] \\
&\overset{(a)}{=} z_1z_2\sum_{\ell=1}^{j}\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell]\mathbb{E}[\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k] \\
&\overset{(b)}{=} \psi_j^{\mathbf{M}}\sum_{\ell=1}^{j}z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbb{E}[\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k] + \mathcal{O}(n^{-1/2}) \\
&\overset{(c)}{=} -\psi_j^{\mathbf{M}}\sum_{\ell=1}^{j}\frac{z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_2\mathbf{a}_k}{(1+\delta_2)} + \mathcal{O}(n^{-1/2}),
\end{aligned}
\tag{2.4.44}
$$

where (a) follows from the fact that $\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell$ and $\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k$ are uncorrelated, (b) is a consequence of Lemma 2.4.3 and (c) is from Lemma 2.4.7. Similarly,

$$
\begin{aligned}
Z_4 &= z_1z_2\sum_{\ell=1}^{j}\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1\mathbf{r}_\ell\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{y}_\ell\mathbf{y}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{a}_k] + \mathcal{O}(n^{-1/2}) \\
&= \mathbf{a}_k^*\mathbf{T}_1\mathbf{T}_2\mathbf{a}_k\frac{1}{n}\sum_{\ell=1}^{j}z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbb{E}[\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{a}_\ell] + \mathcal{O}(n^{-1/2}) \\
&= \mathbf{a}_k^*\mathbf{T}_1\mathbf{T}_2\mathbf{a}_k\phi_j^{\mathbf{M}} + \mathcal{O}(n^{-1/2}).
\end{aligned}
\tag{2.4.45}
$$

Plugging (2.4.42), (2.4.43), (2.4.44), (2.4.45) into (2.4.40) , we obtain :

$$
\begin{aligned}
Z = \sum_{\ell=1}^{n} & z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbb{E}[\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] + z_1\tilde{\delta}_1\mathbb{E}[\mathbf{a}_k^*\mathbf{T}_1(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] \\
& + \psi_j^{\mathbf{M}}\mathbf{a}_k^*\mathbf{T}_1\mathbf{T}_2\mathbf{a}_k\frac{1}{n}\sum_{\ell=1}^{j}z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell} - \psi_j^{\mathbf{M}}\sum_{\ell=1}^{j}\frac{z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_2\mathbf{a}_k}{(1+\delta_2)} \\
& + \mathbf{a}_k^*\mathbf{T}_1\mathbf{T}_2\mathbf{a}_k\phi_j^{\mathbf{M}} + \mathcal{O}(n^{-1/2}).
\end{aligned}
$$

Plugging this and (2.4.39) into (2.4.33), and noticing that $-z\tilde{t}_{\ell\ell}(\mathbf{a}_\ell^*\mathbf{T}_\ell\mathbf{a}_\ell(1+\delta)^{-1}+1) = (1+\delta)^{-1}$ (cf. (2.4.10)), we obtain (2.4.31).

Step 2 : Expression of $\psi_j^{\mathbf{M}}$

In this section, we will show the following equation satisfied by $\psi_j^{\mathbf{M}}$ and $\zeta_{kj}^{\mathbf{M}}$:

$$
\begin{aligned}
\psi_j^{\mathbf{M}} = & \frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{M}\mathbf{T}_2 + \psi_j^{\mathbf{M}}\left(\frac{1}{n}\sum_{\ell=1}^{j}\frac{\mathbf{a}_\ell^*\mathbf{T}_2\mathbf{T}_1\mathbf{a}_\ell}{(1+\delta_1)(1+\delta_2)} + \frac{\gamma}{n}\sum_{\ell=1}^{j}z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\right) \\
& + \gamma\phi_j^{\mathbf{M}} + \mathcal{O}(n^{-1/2}).
\end{aligned}
\tag{2.4.46}
$$

Using identity (2.4.32), we obtain :

$$\begin{aligned}\psi_j^{\mathbf{M}} &= \frac{1}{n}\mathrm{Tr}\mathbb{E}[(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2] \\ &= \frac{1}{n}\mathrm{Tr}\mathbb{E}[\mathbf{T}_1\mathbf{M}\mathbf{Q}_2] - \frac{z_1\tilde{\delta}_1}{n}\mathbb{E}[\mathbf{T}_1(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2] \\ &\quad + \frac{1}{n(1+\delta_1)}\mathrm{Tr}\mathbb{E}[\mathbf{T}_1\mathbf{A}\mathbf{A}^*(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2] - \frac{1}{n}\mathrm{Tr}\mathbb{E}[\mathbf{T}_1(\mathbb{E}_j\mathbf{Y}\mathbf{Y}^*\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2] \\ &\triangleq \frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{M}\mathbf{T}_2 - \frac{z_1\tilde{\delta}_1}{n}\mathbb{E}[\mathbf{T}_1(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2] + X + Z + \epsilon,\end{aligned} \tag{2.4.47}$$

where X and Z are the last two terms and $\epsilon = \mathcal{O}(n^{-1/2})$ by Theorem 2.4.1. X and Z can be treated in a similar way in the precedent section. More detailed, X satisfies

$$\begin{aligned} X &= \frac{1}{n(1+\delta_1)}\sum_{\ell=1}^n \mathbb{E}[\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{T}_1\mathbf{a}_\ell] \\ &= \frac{1}{n(1+\delta_1)}\sum_{\ell=1}^n \mathbb{E}[\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{T}_1\mathbf{a}_\ell] \\ &\quad + \frac{1}{n(1+\delta_1)}\sum_{\ell=1}^n \mathbb{E}[z_1\tilde{t}_{1,\ell\ell}(\mathbb{E}_j\mathbf{a}_\ell^*\mathbf{Q}_{1,\ell}\mathbf{y}_\ell\mathbf{y}_\ell^*\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{T}_1\mathbf{a}_\ell] + \mathcal{O}(n^{-1/2}) \\ &= \frac{1}{n(1+\delta_1)}\sum_{\ell=1}^n \mathbb{E}[\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{T}_1\mathbf{a}_\ell] \\ &\quad + \frac{1}{n(1+\delta_1)}\sum_{\ell=1}^n \mathbb{E}[z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{T}_1\mathbf{a}_\ell] \\ &\quad + \frac{1}{n(1+\delta_1)}\sum_{\ell=1}^j \mathbb{E}[z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{T}_1\mathbf{a}_\ell] + \mathcal{O}(n^{-1/2}).\end{aligned}$$

By (2.4.10), we have

$$\begin{aligned} X &= -\frac{1}{n}\sum_{\ell=1}^n \mathbb{E}[z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{T}_1\mathbf{a}_\ell] \\ &\quad + \frac{1}{n(1+\delta_1)}\sum_{\ell=1}^j \mathbb{E}[z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{T}_1\mathbf{a}_\ell] + \mathcal{O}(n^{-1/2})\end{aligned} \tag{2.4.48}$$

and the second term satisfies

$$\begin{aligned} &\frac{1}{n(1+\delta_1)}\sum_{\ell=1}^j z_1\tilde{t}_{1,\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_2\mathbf{T}_1\mathbf{a}_\ell] \\ &= \frac{1}{n(1+\delta_1)}\sum_{\ell=1}^j z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{y}_\ell\mathbf{y}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{T}_1\mathbf{a}_\ell] + \mathcal{O}(n^{-1/2}) \\ &\overset{(a)}{=} \frac{1}{n(1+\delta_1)}\sum_{\ell=1}^j z_1z_2\tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_{1,\ell}\mathbf{a}_\ell\mathbb{E}[\mathbf{r}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{r}_\ell\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{T}_1\mathbf{a}_\ell] + \mathcal{O}(n^{-1/2})\end{aligned}$$

$$\overset{(b)}{=} \frac{\psi_j^{\mathbf{M}}}{n} \sum_{\ell=1}^{j} \frac{\mathbf{a}_\ell^* \mathbf{T}_{1,\ell} \mathbf{a}_\ell \mathbf{a}_\ell^* \mathbf{T}_2 \mathbf{T}_1 \mathbf{a}_\ell}{(1+\delta_1)^2(1+\delta_2)} + \mathcal{O}(n^{-1/2}),$$

where (a) follows from Remark 2.4.5 and (b) follows from (2.4.11). The term Z can be expressed as :

$$\begin{aligned}
Z &= \frac{1}{n}\sum_{\ell=1}^{n} \mathrm{Tr}\mathbb{E}[z_1 \tilde{t}_{1,\ell\ell} \mathbf{T}_1 (\mathbb{E}_j \mathbf{y}_\ell \mathbf{y}_\ell^* \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2] + \mathcal{O}(n^{-1/2}) \\
&= \frac{1}{n}\sum_{\ell=1}^{n} \mathbb{E}[z_1 \tilde{t}_{1,\ell\ell} \mathbf{a}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{T}_1 \mathbf{a}_\ell] \\
&\quad + \left(\frac{1}{n}\sum_{\ell=1}^{j} \mathbb{E}[z_1 \tilde{t}_{1,\ell\ell} \mathbf{r}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{T}_1 \mathbf{r}_\ell] + \frac{1}{n}\sum_{\ell=j+1}^{n} \frac{1}{n} \mathrm{Tr}\mathbb{E}[z_1 \tilde{t}_{1,\ell\ell} (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{T}_1] \right) \\
&\qquad + \frac{1}{n}\sum_{\ell=1}^{j} \mathbb{E}[z_1 \tilde{t}_{1,\ell\ell} \mathbf{y}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{T}_1 \mathbf{a}_\ell] \\
&\qquad\quad + \frac{1}{n}\sum_{\ell=1}^{j} \mathbb{E}[z_1 \tilde{t}_{1,\ell\ell} \mathbf{a}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_2 \mathbf{T}_1 \mathbf{r}_\ell] + \mathcal{O}(n^{-1/2}) \\
&\triangleq Z_1 + Z_2 + Z_3 + Z_4 + \mathcal{O}(n^{-1/2}).
\end{aligned}$$

The term Z_1 cancels with the first term of X's decomposition in (2.4.48). The term Z_2, Z_3 and Z_4 satisfy :

$$\begin{aligned}
Z_2 &= \frac{1}{n}\sum_{\ell=1}^{j} \mathbb{E}[z_1 z_2 \tilde{t}_{1,\ell\ell} \tilde{t}_{2,\ell\ell} \mathbf{y}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_{2,\ell} \mathbf{y}_\ell \mathbf{y}_\ell^* \mathbf{Q}_{2,\ell} \mathbf{T}_1 \mathbf{r}_\ell] \\
&\quad + \frac{z_1 \delta_1}{n} \mathrm{Tr}\mathbb{E}[\mathbf{T}_1 (\mathbb{E}_j \mathbf{Q}_1) \mathbf{M} \mathbf{Q}_2] + \mathcal{O}(n^{-1/2}) \\
&= \frac{1}{n} \psi_j^{\mathbf{M}} \mathrm{Tr} \mathbf{T}_1 \mathbf{T}_2 \frac{1}{n} \sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell} \tilde{t}_{2,\ell\ell} + \frac{z_1 \delta_1}{n} \mathrm{Tr}\mathbb{E}[\mathbf{T}_1 (\mathbb{E}_j \mathbf{Q}_1) \mathbf{M} \mathbf{Q}_2] + \mathcal{O}(n^{-1/2}),
\end{aligned}$$

and by Remark 2.4.5,

$$\begin{aligned}
Z_3 &= \frac{1}{n}\sum_{\ell=1}^{j} \mathbb{E}[z_1 z_2 \tilde{t}_{1,\ell\ell} \tilde{t}_{2,\ell\ell} \mathbf{r}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_{2,\ell} \mathbf{y}_\ell \mathbf{y}_\ell^* \mathbf{Q}_{2,\ell} \mathbf{T}_1 \mathbf{a}_\ell] + \mathcal{O}(n^{-1/2}) \\
&= \frac{1}{n}\sum_{\ell=1}^{j} \mathbb{E}[z_1 z_2 \tilde{t}_{1,\ell\ell} \tilde{t}_{2,\ell\ell} \mathbf{r}_\ell^* (\mathbb{E}_j \mathbf{Q}_{1,\ell}) \mathbf{M} \mathbf{Q}_{2,\ell} \mathbf{r}_\ell \mathbf{a}_\ell^* \mathbf{Q}_{2,\ell} \mathbf{T}_1 \mathbf{a}_\ell] + \mathcal{O}(n^{-1/2}) \\
&= -\psi_j^{\mathbf{M}} \frac{1}{n} \sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell} \tilde{t}_{2,\ell\ell} \mathbf{a}_\ell^* \mathbf{T}_{2,\ell} \mathbf{T}_1 \mathbf{a}_\ell + \mathcal{O}(n^{-1/2}) \\
&= -\psi_j^{\mathbf{M}} \frac{1}{n} \sum_{\ell=1}^{j} z_1 \tilde{t}_{1,\ell\ell} \frac{\mathbf{a}_\ell^* \mathbf{T}_2 \mathbf{T}_1 \mathbf{a}_\ell}{1+\delta_2} + \mathcal{O}(n^{-1/2}),
\end{aligned}$$

where the last equality is from (2.4.11). For Z_4, recall the definition of γ in (2.2.1),

$$\begin{aligned}
Z_4 &= \frac{1}{n}\sum_{\ell=1}^{j}\mathbb{E}[z_1 z_2 \tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{y}_\ell\mathbf{y}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{T}_1\mathbf{r}_\ell] + \mathcal{O}(n^{-1/2}) \\
&= \frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{T}_2\frac{1}{n}\sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\mathbb{E}[\mathbf{a}_\ell^*(\mathbb{E}_j\mathbf{Q}_{1,\ell})\mathbf{M}\mathbf{Q}_{2,\ell}\mathbf{a}_\ell] + \mathcal{O}(n^{-1/2}) \\
&= \gamma\phi_j^{\mathbf{M}} + \mathcal{O}(n^{-1/2}).
\end{aligned}$$

Plugging these terms in (2.4.47), by using (2.4.11) and (2.4.5), we obtain :

$$\begin{aligned}
\psi_j^{\mathbf{M}} &= \frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{M}\mathbf{T}_2 + \psi_j^{\mathbf{M}}\left(\frac{1}{n}\sum_{\ell=1}^{j}\mathbf{a}_\ell^*\mathbf{T}_2\mathbf{T}_1\mathbf{a}_\ell\left(\frac{\mathbf{a}_\ell^*\mathbf{T}_1\mathbf{a}_\ell}{(1+\delta_1)^2(1+\delta_2)} - \frac{z_1\tilde{t}_{1,\ell\ell}}{1+\delta_2}\right)\right. \\
&\quad \left. + \frac{\gamma}{n}\sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\right) + \gamma\phi_j^{\mathbf{M}} + \mathcal{O}(n^{-1/2}) \\
&= \frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{M}\mathbf{T}_2 + \psi_j^{\mathbf{M}}\left(\frac{1}{n}\sum_{\ell=1}^{j}\frac{\mathbf{a}_\ell^*\mathbf{T}_2\mathbf{T}_1\mathbf{a}_\ell}{(1+\delta_1)(1+\delta_2)} + \frac{\gamma}{n}\sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell}\tilde{t}_{2,\ell\ell}\right) \\
&\quad + \gamma\phi_j^{\mathbf{M}} + \mathcal{O}(n^{-1/2}).
\end{aligned}$$

Step 3 : Relation between $\zeta_{kj}^{\mathbf{M}}$ and $\theta_{kj}^{\mathbf{M}}$

With the similar computations, in this section, we will show that :

$$\zeta_{kj}^{\mathbf{M}} = z_1 z_2 \tilde{t}_{1,kk}\tilde{t}_{2,kk}(1+\delta_1)(1+\delta_2)\theta_{kj}^{\mathbf{M}} + \frac{\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_k\mathbf{a}_k^*\mathbf{T}_2\mathbf{a}_k}{(1+\delta_1)(1+\delta_2)}\psi_j^{\mathbf{M}} + \mathcal{O}(n^{-1/2}). \tag{2.4.49}$$

By (2.4.6), the term ζ_{kj} can be written as

$$\begin{aligned}
\zeta_{kj}^{\mathbf{M}} &= \mathbb{E}[a_k^*(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{a}_k] \\
&= \mathbb{E}[\mathbf{a}_k^*\mathbb{E}_j(\mathbf{Q}_{1,k} + z_1\tilde{q}_{1,kk}\mathbf{Q}_{1,k}\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_{1,k})\mathbf{M}(\mathbf{Q}_{2,k} + z_2\tilde{q}_{2,kk}\mathbf{Q}_{2,k}\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_{2,k})\mathbf{a}_k] \\
&= \theta_{kj}^{\mathbf{M}} + z_1\tilde{t}_{1,kk}\mathbb{E}[\mathbf{a}_k^*\mathbb{E}_j[\mathbf{Q}_{1,k}\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_{1,k}]\mathbf{M}\mathbf{Q}_{2,k}\mathbf{a}_k] \\
&\quad + z_2\tilde{t}_{2,kk}\mathbb{E}[\mathbf{a}_k^*(\mathbb{E}_j\mathbf{Q}_{1,k})\mathbf{M}\mathbf{Q}_{2,k}\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_{2,k}\mathbf{a}_k] \\
&\quad\quad + z_1 z_2\tilde{t}_{1,kk}\tilde{t}_{2,kk}\mathbb{E}[\mathbf{a}_k^*\mathbb{E}_j[\mathbf{Q}_{1,k}\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_{1,k}]\mathbf{M}\mathbf{Q}_{2,k}\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_{2,k}\mathbf{a}_k] + \mathcal{O}(n^{-1/2}) \\
&\triangleq \theta_{kj}^{\mathbf{M}} + X_1 + X_2 + X_3 + \mathcal{O}(n^{-1/2}).
\end{aligned}$$

Using similar arguments as before, we get

$$\begin{aligned}
X_1 &= z_1\tilde{t}_{1,kk}\mathbf{a}_k^*\mathbf{T}_{1,k}\mathbf{a}_k\mathbb{E}[\mathbf{a}_k^*\mathbb{E}_j[\mathbf{Q}_{1,k}]\mathbf{M}\mathbf{Q}_{2,k}\mathbf{a}_k] + \mathcal{O}(n^{-1/2}) \\
&= z_1\tilde{t}_{1,kk}\mathbf{a}_k^*\mathbf{T}_{1,k}\mathbf{a}_k\theta_{kj}^{\mathbf{M}} + \mathcal{O}(n^{-1/2}), \\
X_2 &= z_2\tilde{t}_{2,kk}\mathbf{a}_k^*\mathbf{T}_{2,k}\mathbf{a}_k\theta_{kj}^{\mathbf{M}} + \mathcal{O}(n^{-1/2}).
\end{aligned}$$

For $k \leq j$,

$$\begin{aligned} X_3 &= z_1 z_2 \tilde{t}_{1,kk} \tilde{t}_{2,kk} (\mathbf{a}_k^* \mathbf{T}_{1,k} \mathbf{a}_k)(\mathbf{a}_k^* \mathbf{T}_{2,k} \mathbf{a}_k) \mathbb{E}[\mathbf{y}_k^* (\mathbb{E}_j \mathbf{Q}_{1,k}) \mathbf{M} \mathbf{Q}_{2,k} \mathbf{y}_k] + \mathcal{O}(n^{-1/2}) \\ &= z_1 z_2 \tilde{t}_{1,kk} \tilde{t}_{2,kk} (\mathbf{a}_k^* \mathbf{T}_{1,k} \mathbf{a}_k)(\mathbf{a}_k^* \mathbf{T}_{2,k} \mathbf{a}_k)(\theta_{kj}^{\mathbf{M}} + \psi_j^{\mathbf{M}}) + \mathcal{O}(n^{-1/2}). \end{aligned}$$

Using the identity (2.4.10), we finally obtain (2.4.49).

Step 4 : A system of perturbed linear equations.

Combining the expression of $\zeta_{kj}^{\mathbf{M}}$ (cf. (2.4.49)) with $\zeta_{kj}^{\mathbf{M}}$ (cf. (2.4.31)), we get

$$\begin{aligned} z_1 z_2 \tilde{t}_{1,kk} \tilde{t}_{2,kk} \theta_{kj}^{\mathbf{M}} = \frac{\mathbf{a}_k^* \mathbf{T}_1 \mathbf{M} \mathbf{T}_2 \mathbf{a}_k}{(1+\delta_1)(1+\delta_2)} &+ \psi_j^{\mathbf{M}} \left(\sum_{\ell=1}^{j} \frac{\mathbf{a}_k^* \mathbf{T}_1 \mathbf{a}_\ell \mathbf{a}_\ell^* \mathbf{T}_2 \mathbf{a}_k}{(1+\delta_1)^2 (1+\delta_2)^2} \right. \\ &\left. + \frac{\mathbf{a}_k^* \mathbf{T}_1 \mathbf{T}_2 \mathbf{a}_k}{(1+\delta_1)(1+\delta_2)} \frac{1}{n} \sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell} \tilde{t}_{2,\ell\ell} - \frac{\mathbf{a}_k^* \mathbf{T}_1 \mathbf{a}_k \mathbf{a}_k^* \mathbf{T}_2 \mathbf{a}_k}{(1+\delta_1)^2 (1+\delta_2)^2} \right) \\ &+ \frac{\mathbf{a}_k^* \mathbf{T}_1 \mathbf{T}_2 \mathbf{a}_k}{(1+\delta_1)(1+\delta_2)} \phi_j^{\mathbf{M}} + \mathcal{O}(n^{-1/2}) \end{aligned} \tag{2.4.50}$$

which implies that $\phi_j^{\mathbf{M}}$ satisfies

$$(1-\nu_j)\phi_j^{\mathbf{M}} - (\omega_j + \nu_j \eta_j)\psi_j^{\mathbf{M}} = \nu_j^{\mathbf{M}} + \mathcal{O}(n^{-1/2})$$

where

$$\nu_j = \frac{1}{n} \sum_{k=1}^{j} \frac{a_k^* \mathbf{T}_1 \mathbf{T}_2 a_k}{(1+\delta_1)(1+\delta_2)} \tag{2.4.51}$$

$$\nu_j^{\mathbf{M}} = \frac{1}{n} \sum_{k=1}^{j} \frac{a_k^* \mathbf{T}_1 \mathbf{M} \mathbf{T}_2 a_k}{(1+\delta_1)(1+\delta_2)} \tag{2.4.52}$$

$$\eta_j = \frac{1}{n} \sum_{\ell=1}^{j} z_1 z_2 \tilde{t}_{1,\ell\ell} \tilde{t}_{2,\ell\ell} \tag{2.4.53}$$

$$\omega_j = \frac{1}{n} \sum_{k=1}^{j} \sum_{\ell=1,\ell\neq k}^{j} \frac{\mathbf{a}_k^* \mathbf{T}_1 \mathbf{a}_\ell \mathbf{a}_\ell^* \mathbf{T}_2 \mathbf{a}_k}{(1+\delta_1)^2 (1+\delta_2)^2}. \tag{2.4.54}$$

With these notations, Equation (2.4.46) is rewritten as

$$-\gamma \phi_j^{\mathbf{M}} + (1 - \nu_j - \gamma \eta_j)\psi_j^{\mathbf{M}} = \gamma^{\mathbf{M}} + \mathcal{O}(n^{-1/2}),$$

where

$$\gamma^{\mathbf{M}} = \frac{1}{n} \mathrm{Tr} \mathbf{T}_1 \mathbf{M} \mathbf{T}_2.$$

We end up with a system of two perturbed linear equations in $\phi_j^{\mathbf{M}}$ and $\psi_j^{\mathbf{M}}$:

$$\begin{cases}(1-\nu_j)\phi_j^{\mathbf{M}} - (\omega_j + \nu_j\eta_j)\psi_j^{\mathbf{M}} = \nu_j^{\mathbf{M}} + \mathcal{O}(n^{-1/2}) \\ -\gamma\phi_j^{\mathbf{M}} + (1-\nu_j-\gamma\eta_j)\psi_j^{\mathbf{M}} = \gamma^{\mathbf{M}} + \mathcal{O}(n^{-1/2}).\end{cases} \tag{2.4.55}$$

The determinant of the system is

$$\boldsymbol{\Delta}_j = (1-\nu_j)^2 - \gamma\eta_j - \gamma\omega_j,$$

which has several interesting characters.

Lemma 2.4.8 *We have :*

1. *The determinant of the system $\boldsymbol{\Delta}_j$ coincides with Δ_n in (2.2.7) when $j = n$.*
2. *The sequence (Δ_n) satisfies for all $z_1, z_2 \in \mathbb{C}^+$ and $z_1 \neq z_2$,*

$$\frac{1}{\Delta_n} = \frac{s_n(z_1) - s_n(z_2)}{z_1 - z_2},$$

where $s_n(z) = z(1+\delta_n)(1+\tilde{\delta}_n)$.

3. *For all $z_1, z_2 \in \mathbb{C}^+$,*

$$\liminf_n \inf_{j\in\{1,\cdots,n\}} |\boldsymbol{\Delta}_j(z_1,z_2)| > 0.$$

The proof is in Appendix 2.6.8. With the precedent lemma, the system has the solution

$$\begin{bmatrix}\phi_j^{\mathbf{M}} \\ \psi_j^{\mathbf{M}}\end{bmatrix} = \frac{1}{\boldsymbol{\Delta}_j}\begin{bmatrix}\nu_j^{\mathbf{M}}(1-\nu_j-\gamma\eta_j) + \gamma^{\mathbf{M}}(\omega_j+\nu_j\eta_j) \\ \gamma\nu_j^{\mathbf{M}} + \gamma^{\mathbf{M}}(1-\nu_j)\end{bmatrix} + \mathcal{O}(n^{-1/2}). \tag{2.4.56}$$

Now we take $\mathbf{M} = \mathbf{I}_N$ and the convention that $\nu_0 = \omega_0 = \eta_0 = 0$. Noticing that $\nu_j^{\mathbf{I}_N} = \nu_j$ and $\gamma^{\mathbf{I}_N} = \gamma$, and replacing the solution into (2.4.50), we obtain :

$$\begin{aligned}&\frac{z_1z_2\tilde{t}_{1,jj}\tilde{t}_{2,jj}\theta_{jj}^{\mathbf{I}_N}(z_1,z_2)}{n} \\ =&(\nu_j-\nu_{j-1}) + \left(\frac{1}{2}(\omega_j-\omega_{j-1}) + \eta_j(\nu_j-\nu_{j-1})\right)\psi_j^{\mathbf{I}_N}(z_1,z_2) \\ &+ (\nu_j-\nu_{j-1})\phi_j^{\mathbf{I}_N}(z_1,z_2) + \mathcal{O}(n^{-3/2}) \\ =&(\nu_j-\nu_{j-1}) + \frac{\frac{1}{2}\gamma(\omega_j-\omega_{j-1}) + \gamma\eta_j(\nu_j-\nu_{j-1})}{\boldsymbol{\Delta}_j} \\ &+ \frac{(\nu_j-\nu_{j-1})(\nu_j(1-\nu_j)+\gamma\nu_j)}{\boldsymbol{\Delta}_j} + \mathcal{O}(n^{-3/2}).\end{aligned} \tag{2.4.57}$$

Since z_1, z_2 are symmetric, we can achieve the same equation by interchanging z_1 and z_2. Moreover, by diagonalization,

$$\mathbf{T}_1\mathbf{T}_2 = \mathbf{T}_2\mathbf{T}_1,$$

which implies $\nu_j(z_1,z_2)=\nu_j(z_2,z_1)$. Thus, we get

$$\frac{z_1z_2\tilde{t}_{1,jj}\tilde{t}_{2,jj}\theta_{jj}^{\mathbf{I}_N}(z_2,z_1)}{n}=\frac{z_1z_2\tilde{t}_{1,jj}\tilde{t}_{2,jj}\theta_{jj}^{\mathbf{I}_N}(z_1,z_2)}{n}+\mathcal{O}(n^{-3/2})$$

which leads to

$$\begin{aligned}&\frac{1}{n}z_1z_2\tilde{t}_{1,jj}\tilde{t}_{2,jj}(\psi_j^{\mathbf{I}_N}(z_1,z_2)+\theta_{jj}^{\mathbf{I}_N}(z_1,z_2)+\theta_{jj}^{\mathbf{I}_N}(z_2,z_1))\\=&\frac{1}{n}\frac{\gamma(\eta_j-\eta_{j-1})}{\mathbf{\Delta}_j}+\frac{2}{n}z_1z_2\tilde{t}_{1,jj}\tilde{t}_{2,jj}\theta_{jj}^{\mathbf{I}_N}+\mathcal{O}(n^{-3/2})\\=&\frac{2(\nu_j-\nu_{j-1})(1-\nu_j)+\gamma(\eta_j-\eta_{j-1})+\gamma(\omega_j-\omega_{j-1})}{\mathbf{\Delta}_j}+\mathcal{O}(n^{-3/2}).\end{aligned}$$

On the other hand, taking $\mathbf{\Delta}_0=1$ and noticing that $\mathbf{\Delta}_{j-1}-\mathbf{\Delta}_j=2(\nu_j-\nu_{j-1})(1-\nu_j)+\gamma(\eta_j-\eta_{j-1})+\gamma(\omega_j-\omega_{j-1})+\mathcal{O}(n^{-2})$, then we have

$$\frac{1}{n}z_1z_2\tilde{t}_{1,jj}\tilde{t}_{2,jj}(\psi_j^{\mathbf{I}_N}(z_1,z_2)+\theta_{jj}^{\mathbf{I}_N}(z_1,z_2)+\theta_{jj}^{\mathbf{I}_N}(z_2,z_1))=\frac{\mathbf{\Delta}_{j-1}-\mathbf{\Delta}_j}{\mathbf{\Delta}_j}+\mathcal{O}(n^{-3/2}).\tag{2.4.58}$$

Since for all $j\in\{1,\cdots,n\}$,

$$\left|\frac{1}{n}z_1z_2\tilde{t}_{1,jj}\tilde{t}_{2,jj}(\psi_j^{\mathbf{I}_N}(z_1,z_2)+\theta_{jj}^{\mathbf{I}_N}(z_1,z_2)+\theta_{jj}^{\mathbf{I}_N}(z_2,z_1))\right|\le\frac{K}{n},$$

then

$$\left|\frac{\mathbf{\Delta}_{j-1}-\mathbf{\Delta}_j}{\mathbf{\Delta}_j}\right|\le\frac{K}{n}.\tag{2.4.59}$$

Now we introduce the logarithm function for complex number. For $z\in\mathbb{C}^*$, denote

$$\log(z)=\log|z|+\mathbf{i}\arg(z),\quad \arg(z)\in[-\pi,\pi).$$

We can verify easily that $\log(1+z)=z+o(|z|)$ when $|z|\to 0$ and $\frac{d}{dz}\log(1+z)=\frac{1}{1+z}$.

Combining (2.4.58) and (2.4.59), we have

$$\begin{aligned}&\frac{1}{n}\sum_{j=1}^{n}z_1z_2\tilde{t}_{1,jj}\tilde{t}_{2,jj}(\psi_j^{\mathbf{I}_N}(z_1,z_2)+\theta_{jj}^{\mathbf{I}_N}(z_1,z_2)+\theta_{jj}^{\mathbf{I}_N}(z_2,z_1))\\&=\sum_{j=1}^{n}\frac{\mathbf{\Delta}_{j-1}-\mathbf{\Delta}_j}{\mathbf{\Delta}_j}+\mathcal{O}(n^{-1/2})\\&=\sum_{j=1}^{n}\log\left(1+\frac{\mathbf{\Delta}_{j-1}-\mathbf{\Delta}_j}{\mathbf{\Delta}_j}\right)+\mathcal{O}(n^{-1/2})\\&=\sum_{j=1}^{n}\log\frac{\mathbf{\Delta}_{j-1}}{\mathbf{\Delta}_j}+\mathcal{O}(n^{-1/2}).\end{aligned}\tag{2.4.60}$$

Now we will apply Lemma 2.4.1 to derive the function. Since

$$\begin{aligned}\left|\sum_{j=1}^{n}\xi_{3j}-\sum_{j=1}^{n}\log\frac{\boldsymbol{\Delta}_{j-1}}{\boldsymbol{\Delta}_j}\right| &\xrightarrow[n\to\infty]{} 0,\\ \max\left\{\left|\sum_{j=1}^{n}\xi_{3j}\right|,\left|\sum_{j=1}^{n}\log\frac{\boldsymbol{\Delta}_{j-1}}{\boldsymbol{\Delta}_j}\right|\right\} &\leq K,\end{aligned}$$

then by Lemma 2.4.1, with the same argument as Proposition 2.4.1,we have

$$\left|\frac{\partial}{\partial z_1}\sum_{j=1}^{n}\xi_{3j}-\frac{\partial}{\partial z_1}\sum_{j=1}^{n}\log\frac{\boldsymbol{\Delta}_{j-1}}{\boldsymbol{\Delta}_j}\right|\xrightarrow[n\to\infty]{} 0,$$

and

$$\left|\frac{\partial}{\partial z_1}\sum_{j=1}^{n}\xi_{3j}-\frac{\partial}{\partial z_1}\sum_{j=1}^{n}\log\frac{\boldsymbol{\Delta}_{j-1}}{\boldsymbol{\Delta}_j}\right|\leq K.$$

By another application of Lemma 2.4.1, we have

$$\left|\frac{\partial^2}{\partial z_1\partial z_2}\sum_{j=1}^{n}\xi_{3j}-\frac{\partial^2}{\partial z_1\partial z_2}\sum_{j=1}^{n}\log\frac{\boldsymbol{\Delta}_{j-1}}{\boldsymbol{\Delta}_j}\right|\xrightarrow[n\to\infty]{} 0. \tag{2.4.61}$$

By a simple calculation,

$$\frac{\partial}{\partial z_1}\sum_{j=1}^{n}\log\frac{\boldsymbol{\Delta}_{j-1}}{\boldsymbol{\Delta}_j}=\sum_{j=1}^{n}\left(\frac{1}{\boldsymbol{\Delta}_{j-1}}\frac{\partial\boldsymbol{\Delta}_{j-1}}{\partial z_1}-\frac{1}{\boldsymbol{\Delta}_j}\frac{\partial\boldsymbol{\Delta}_j}{\partial z_1}\right)=-\frac{1}{\boldsymbol{\Delta}_n}\frac{\partial\boldsymbol{\Delta}_n}{\partial z_1}.$$

This along with (1) and (2) in Lemma 2.4.8 conclude the proof of Lemma 2.4.5.

□

A similar result holds true for $\sum_{j=1}^{n}\xi_{4j}$.

Lemma 2.4.9 *Recall the expression of $\sum_{j=1}^{n}\xi_{4j}$ in (2.4.29). We have*

$$\frac{\partial^2}{\partial z_1\partial z_2}\sum_{j=1}^{n}\xi_{4j}-\Theta_{1,n}\xrightarrow[n\to\infty]{\mathcal{P}} 0$$

where $\Theta_{1,n}$ is defined in (2.2.13).

The same method as the proof of Lemma 2.4.5 applies for this lemma. We just recall some elements of the proof. We first calculate $\mathbf{a}_j^*(\mathbb{E}_j\mathbf{Q}_j(z_1))(\mathbb{E}_j\mathbf{Q}_j^T(z_2))\bar{\mathbf{a}}_j$ and denote :

$$\begin{aligned}\underline{\psi}_j(z_1,z_2) &= \frac{1}{n}\mathrm{Tr}\mathbb{E}\{[\mathbb{E}_j\mathbf{Q}(z_1)]\mathbb{E}_j\mathbf{Q}^T(z_2)\}=\frac{1}{n}\mathrm{Tr}\mathbb{E}\{[\mathbb{E}_j\mathbf{Q}(z_1)]\mathbf{Q}^T(z_2)\},\\ \underline{\zeta}_{kj}(z_1,z_2) &= \mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}(z_1)]\mathbb{E}_j\mathbf{Q}^T(z_2)\bar{\mathbf{a}}_k\}=\mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}(z_1)]\mathbf{Q}^T(z_2)\bar{\mathbf{a}}_k\},\\ \underline{\theta}_{kj}(z_1,z_2) &= \mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}_k(z_1)]\mathbb{E}_j\mathbf{Q}_k^T(z_2)\bar{\mathbf{a}}_k\}=\mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}_k(z_1)]\mathbf{Q}_k^T(z_2)\bar{\mathbf{a}}_k\},\end{aligned}$$

$$\underline{\phi}_j = \frac{1}{n}\sum_{k=1}^{j} z_1 z_2 \tilde{t}_{kk}(z_1)\tilde{t}_{kk}(z_2)\underline{\theta}_{kj}.$$

Similar derivations as the proof of Lemma 2.4.5 yield the perturbed system :

$$\begin{cases}(1-\vartheta\nu_j^\dagger)\underline{\phi}_j - (\bar{\vartheta}\omega_j^\dagger + |\vartheta|^2\nu_j^\dagger\eta_j)\underline{\psi}_j = \nu_j^\dagger + \mathcal{O}(n^{-1/2})\\ -\vartheta\gamma^\dagger\underline{\phi}_j + (1-\bar{\vartheta}\tilde{\nu}_j^\dagger - |\vartheta|^2\gamma^\dagger\eta_j)\underline{\psi}_j = \gamma^\dagger + \mathcal{O}(n^{-1/2}).\end{cases} \tag{2.4.62}$$

where

$$\begin{aligned}
\nu_j^\dagger &= \frac{1}{n}\sum_{k=1}^{j}\frac{\mathbf{a}_k^*\mathbf{T}_1\mathbf{T}_2^T\bar{\mathbf{a}}_k}{(1+\delta_1)(1+\delta_2)}\\
\tilde{\nu}_j^\dagger &= \frac{1}{n}\sum_{k=1}^{j}\frac{\mathbf{a}_k^T\mathbf{T}_1^T\mathbf{T}_2\mathbf{a}_k}{(1+\delta_1)(1+\delta_2)}\\
\gamma^\dagger &= \frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{T}_2^T\\
\omega_j^\dagger &= \frac{1}{n}\sum_{k=1}^{j}\sum_{\ell=1,\ell\neq k}^{j}\frac{\mathbf{a}_k^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^T\mathbf{T}_2^T\bar{\mathbf{a}}_k}{(1+\delta_1)^2(1+\delta_2)^2}.
\end{aligned}$$

We have similar proprieties for the determinant of the system

$$\mathbf{\Delta}_j^\vartheta(z_1,z_2) = (1-\vartheta\nu_j^\dagger)(1-\bar{\vartheta}\tilde{\nu}_j^\dagger) - |\vartheta|^2\gamma^\dagger(\eta_j+\omega_j^\dagger).$$

Lemma 2.4.10 *For all $z_1, z_2 \in \mathbb{C}^+$, We have*

1. *The determinant of the system $\mathbf{\Delta}_j^\vartheta$ coincides with Δ_n^ϑ in (2.2.8) when $j = n$.*
2. *For all $z_1, z_2 \in \mathbb{C}^+$,*
$$\liminf_n \inf_{j\in\{1,\cdots,n\}} |\mathbf{\Delta}_j(z_1,z_2)| > 0.$$

The proof is in Appendix 2.6.9.

Then the resolution of the system yields that

$$\begin{bmatrix}\underline{\phi}_j\\ \underline{\psi}_j\end{bmatrix} = \frac{1}{\mathbf{\Delta}_j^\vartheta}\begin{bmatrix}\nu_j^\dagger(1-\bar{\vartheta}\tilde{\nu}_j^\dagger)+\bar{\vartheta}\gamma^\dagger\omega_j^\dagger\\ \gamma^\dagger\end{bmatrix} + \mathcal{O}(n^{-1/2}), \tag{2.4.63}$$

The other term in $\sum_{j=1}^n \xi_{4j}$ is $\mathbf{a}_j^T(\mathbb{E}_j\mathbf{Q}_j^T(z_1))(\mathbb{E}_j\mathbf{Q}_j(z_2))\mathbf{a}_j$. With Remark 2.2.3, we also consider

$$\begin{aligned}
\underline{\tilde{\psi}}_j(z_1,z_2) &= \frac{1}{n}\mathrm{Tr}\mathbb{E}\{[\mathbb{E}_j\mathbf{Q}^T(z_1)]\mathbb{E}_j\mathbf{Q}(z_2)\} = \frac{1}{n}\mathrm{Tr}\mathbb{E}\{[\mathbb{E}_j\mathbf{Q}^T(z_1)]\mathbf{Q}(z_2)\},\\
\underline{\tilde{\zeta}}_{kj}(z_1,z_2) &= \mathbb{E}\{\mathbf{a}_k^T[\mathbb{E}_j\mathbf{Q}^T(z_1)]\mathbb{E}_j\mathbf{Q}(z_2)\mathbf{a}_k\} = \mathbb{E}\{\mathbf{a}_k^T[\mathbb{E}_j\mathbf{Q}^T(z_1)]\mathbf{Q}(z_2)\mathbf{a}_k\},\\
\underline{\tilde{\theta}}_{kj}(z_1,z_2) &= \mathbb{E}\{\mathbf{a}_k^T[\mathbb{E}_j\mathbf{Q}_k^T(z_1)]\mathbb{E}_j\mathbf{Q}_k(z_2)\mathbf{a}_k\} = \mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}_k^T(z_1)]\mathbf{Q}_k(z_2)\bar{\mathbf{a}}_k\},\\
\underline{\tilde{\phi}}_j &= \frac{1}{n}\sum_{k=1}^{j} z_1z_2\tilde{t}_{kk}(z_1)\tilde{t}_{kk}(z_2)\underline{\tilde{\theta}}_{kj}.
\end{aligned}$$

Then we have the perturbed system

$$\begin{cases}(1-\bar{\vartheta}\tilde{\nu}_j^\dagger)\underline{\tilde{\phi}}_j-(\vartheta\omega_j^\dagger+|\vartheta|^2\tilde{\nu}_j^\dagger\eta_j)\underline{\tilde{\psi}}_j=\tilde{\nu}_j^\dagger+\mathcal{O}(n^{-1/2})\\ -\bar{\vartheta}\gamma^\dagger\underline{\tilde{\phi}}_j+(1-\vartheta\nu_j^\dagger-|\vartheta|^2\gamma^\dagger\eta_j)\underline{\tilde{\psi}}_j=\gamma^\dagger+\mathcal{O}(n^{-1/2}).\end{cases} \tag{2.4.64}$$

The determinant of this system is again $\boldsymbol{\Delta}_j^\vartheta$ and the solution of the system is

$$\begin{bmatrix}\underline{\tilde{\phi}}_j\\ \underline{\tilde{\psi}}_j\end{bmatrix}=\frac{1}{\boldsymbol{\Delta}_j^\vartheta}\begin{bmatrix}\tilde{\nu}_j^\dagger(1-\vartheta\nu_j^\dagger)+\vartheta\gamma^\dagger\omega_j^\dagger\\ \gamma^\dagger\end{bmatrix}+\mathcal{O}(n^{-1/2}), \tag{2.4.65}$$

Remark 2.4.6 *Notice that in [39], when $z_1,z_2\in\mathbb{R}$,*

$$\tilde{\nu}_j^\dagger=\bar{\nu}_j^\dagger$$

which is coherent with their result.

Then it suffices to show that

$$\frac{\partial^2}{\partial z_1\partial z_2}\frac{1}{n}\sum_{j=1}^n z_1z_2\tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2)(|\vartheta|^2\underline{\psi}_j+\vartheta\underline{\theta}_{jj}+\bar{\vartheta}\underline{\tilde{\theta}}_{jj})-\Theta_{1,n}\xrightarrow[n\to\infty]{}0.$$

Since

$$\begin{aligned}&\frac{1}{n}\sum_{j=1}^n(z_1z_2\tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2)|\vartheta|^2\underline{\psi}_j+z_1z_2\tilde{t}_{jj}(z_1)\tilde{t}_{jj}(z_2)(\vartheta\underline{\theta}_{jj}+\bar{\vartheta}\underline{\bar{\theta}}_{jj})\\ =&\sum_{j=1}^n\frac{\boldsymbol{\Delta}_{j-1}^\vartheta-\boldsymbol{\Delta}_j^\vartheta}{\boldsymbol{\Delta}_j^\vartheta}+\mathcal{O}(n^{-1/2})\\ =&\sum_{j=1}^n\log\frac{\boldsymbol{\Delta}_{j-1}^\vartheta}{\boldsymbol{\Delta}_j^\vartheta}+\mathcal{O}(n^{-1/2}),\end{aligned}$$

with a similar application of Lemma 2.4.1, Lemma 2.4.9 is proved.

Gathering Proposition 2.4.1 , Lemma 2.4.2, 2.4.4, 2.4.5 and 2.4.9, we achieve the expression (2.2.11) of the covariance between $M_n(z_1)$ and $M_n(z_2)$. This ends the computation of the covariance.

2.5 Proof of Theorem 2.2.2 : Computation of the bias

In this section, we will calculate the bias

$$L_n^2(f)=\mathbb{E}(\sum_{i=1}^N f(\lambda_i))-N\int_{\mathbb{R}}f(x)\delta_n(x)dx.$$

This issue has been studied in [11, 39, 93, 38].

Theorem 2.5.1 *Under Assumptions 2.2.1 and 2.2.2, we have*

$$L_n^2(f) - B_n(f) \xrightarrow[n\to\infty]{} 0,$$

where $B_n(f)$ is defined in (2.2.16).

Proof: We will show that

1. For all $z \in \mathbb{C}^+$,
$$|M_n^2(z) - B(z)| \xrightarrow[N,n\to\infty]{} 0$$
2. $M_n^2(z) - B(z)$ forms an equicontinuous family for $z \in \mathcal{C}_n$.

From (3.1.2) and the relation that $\delta_n - \tilde{\delta}_n = \frac{1-c_n}{z}$,

$$\delta_n = \frac{1}{n}\mathrm{Tr}\left(-z(1+\delta_n - \frac{1-c_n}{z})\mathbf{I}_N + \frac{\mathbf{A}\mathbf{A}^*}{1+\delta_n}\right)^{-1}. \tag{2.5.1}$$

Then by replacing δ_n by $\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}$, we consider

$$A_n = \frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q} - \frac{1}{n}\mathrm{Tr}\left[-z\left(1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q} - \frac{1-c_n}{z}\right)\mathbf{I}_N + \frac{\mathbf{A}\mathbf{A}^*}{1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}}\right]^{-1}. \tag{2.5.2}$$

Subtracting (2.5.1) by (2.5.2), we have

$$\mathrm{Tr}\mathbb{E}\mathbf{Q} - n\delta_n = nA_n\left(1 - \frac{z}{n}\mathrm{Tr}\mathbf{T}^2 - \frac{1}{n}\mathrm{Tr}\frac{\mathbf{T}\mathbf{A}\mathbf{A}^*\mathbf{T}}{(1+\delta_n)^2}\right)^{-1} + o(1). \tag{2.5.3}$$

Now we will calculate nA_n. To simplify the notations, we denote

$$\mathbf{S} = -z(1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q} - \frac{1-c_n}{z})\mathbf{I}_N + \frac{\mathbf{A}\mathbf{A}^*}{1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}}. \tag{2.5.4}$$

By $\mathbf{A}^{-1} - \mathbf{B}^{-1} = -\mathbf{B}^{-1}(\mathbf{A}-\mathbf{B})\mathbf{A}^{-1}$, we have

$$nA_n = -\mathbb{E}\mathrm{Tr}\mathbf{S}^{-1}[\sum_{j=1}^{n}\mathbf{y}_j\mathbf{y}_j^* - z\mathbf{I}_N - \mathbf{S}](\mathbf{Y}\mathbf{Y}^* - z\mathbf{I}_N)^{-1}. \tag{2.5.5}$$

From $\mathbf{Y}\mathbf{Y}^* - z\mathbf{I}_N + z\mathbf{I}_N = \sum_{j=1}^n \mathbf{y}_j\mathbf{y}_j^*$, taking the inverse of $\mathbf{Y}\mathbf{Y}^* - z\mathbf{I}_N$ from both sides along with (2.4.9), we have

$$\mathbf{I}_N + z(\mathbf{Y}\mathbf{Y}^* - z\mathbf{I}_N)^{-1} = \sum_{j=1}^{n}\beta_j\mathbf{y}_j\mathbf{y}_j^*\mathbf{Q}_j.$$

Taking trace on both sides and dividing by n, we have

$$\begin{aligned} c_n + \frac{z}{n}\mathrm{Tr}(\mathbf{Y}\mathbf{Y}^* - z\mathbf{I}_N)^{-1} &= \frac{1}{n}\sum_{j=1}^n \beta_j \mathbf{y}_j^*\mathbf{Q}_j\mathbf{y}_j \\ &= 1 - \frac{1}{n}\sum_{j=1}^n \beta_j. \end{aligned}$$

Thus

$$\frac{1}{n}\mathrm{Tr}(\mathbf{Y}^*\mathbf{Y} - z\mathbf{I}_n)^{-1} = -\frac{1}{nz}\sum_{j=1}^n \beta_j.$$

Plugging the precedent equality and (2.4.9) into (2.5.5), we have

$$nA_n = -\mathbb{E}\sum_{j=1}^n\{\beta_j\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{y}_j - \frac{1}{n}\mathbb{E}\beta_j\mathbb{E}\mathrm{Tr}\mathbf{S}^{-1}\mathbf{Q}\} + \sum_{j=1}^n \frac{1}{1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}\mathbf{S}^{-1}\mathbf{a}_j. \tag{2.5.6}$$

Then by (2.4.6), we have

$$\begin{aligned} &\mathbb{E}\mathrm{Tr}\mathbf{S}^{-1}\mathbf{Q} - \mathbb{E}\mathrm{Tr}\mathbf{S}^{-1}\mathbf{Q}_j \\ &= -\mathbb{E}\mathrm{Tr}\beta_j\mathbf{S}^{-1}\mathbf{Q}_j\mathbf{y}_j\mathbf{y}_j^*\mathbf{Q}_j \\ &= -b_j\mathbb{E}(1-\beta_j\gamma_j)\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{Q}_j\mathbf{y}_j. \end{aligned} \tag{2.5.7}$$

By Proposition 2.3.2,

$$|\mathbb{E}\beta_j\gamma_j\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{Q}_j\mathbf{y}_j| \le \frac{K}{\sqrt{n}}.$$

Then

$$\mathbb{E}\mathrm{Tr}\mathbf{S}^{-1}\mathbf{Q} - \mathbb{E}\mathrm{Tr}\mathbf{S}^{-1}\mathbf{Q}_j = -b_j\mathbb{E}\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{Q}_j\mathbf{y}_j + o(1).$$

Since $\beta_j = b_j - b_j^2\gamma_j + \beta_j b_j^2\gamma_j^2$, we have

$$\begin{aligned} &\sum_{j=1}^n\{\mathbb{E}\beta_j\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{y}_j - \frac{1}{n}\mathbb{E}\beta_j\mathbb{E}\mathrm{Tr}\mathbf{S}^{-1}\mathbf{Q}_j - \mathbb{E}\beta_j\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\} \\ &= \sum_{j=1}^n\left(-b_j^2\mathbb{E}\gamma_j\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{y}_j + b_j^2\mathbb{E}\gamma_j\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\right) \\ &\quad + \sum_{j=1}^n\left(b_j^2[\mathbb{E}\beta_j\gamma_j^2\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{y}_j - \mathbb{E}\beta_j\gamma_j^2\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j - \frac{1}{n}\mathbb{E}\beta_j\gamma_j^2\mathbb{E}\mathrm{Tr}\mathbf{S}^{-1}\mathbf{Q}_j]\right) \\ &= \sum_{j=1}^n\left(-b_j^2\mathbb{E}\gamma_j\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{y}_j + b_j^2\mathbb{E}\gamma_j\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\right) \\ &\quad + \sum_{j=1}^n\left(b_j^2[\mathbb{E}\beta_j\gamma_j^2\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{y}_j - \mathbb{E}\beta_j\gamma_j^2\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j - \frac{1}{n}\mathbb{E}\beta_j\gamma_j^2\mathrm{Tr}\mathbf{S}^{-1}\mathbf{Q}_j]\right) \\ &\quad + \sum_{j=1}^n \mathrm{Cov}(\beta_j\gamma_j^2, \frac{1}{n}\mathrm{Tr}\mathbf{S}^{-1}\mathbf{Q}_j). \end{aligned}$$

By Lemma 2.3.3, we have

$$\left|\sum_{j=1}^{n}\left(b_j^2[\mathbb{E}\beta_j\gamma_j^2\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{y}_j-\mathbb{E}\beta_j\gamma_j^2\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j-\frac{1}{n}\mathbb{E}\beta_j\gamma_j^2\mathrm{Tr}\mathbf{S}^{-1}\mathbf{Q}_j]\right)\right|\leq K\eta_n^2$$

and

$$\left|\sum_{j=1}^{n}\mathrm{Cov}(\beta_j\gamma_j^2,\frac{1}{n}\mathrm{Tr}\mathbf{S}^{-1}\mathbf{Q}_j)\right|\leq K\eta_n^2.$$

Then

$$\begin{aligned}&\sum_{j=1}^{n}[\mathbb{E}\gamma_j\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{y}_j-\mathbb{E}\gamma_j\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j]\\&=\sum_{j=1}^{n}\left[\mathbb{E}\left(\mathbf{y}_j^*\mathbf{Q}_j\mathbf{y}_j-n^{-1}\mathrm{Tr}\mathbf{Q}_j-\mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j\right)\left(\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{y}_j-\frac{1}{n}\mathrm{Tr}\mathbf{Q}_j\mathbf{S}^{-1}-\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\right)\right]\\&\quad+\sum_{j=1}^{n}\mathrm{Cov}\left(\frac{1}{n}\mathrm{Tr}\mathbf{Q}_j+\mathbf{a}_j\mathbf{Q}_j\mathbf{a}_j,\frac{1}{n}\mathrm{Tr}\mathbf{Q}_j\mathbf{S}^{-1}\right).\end{aligned}\tag{2.5.8}$$

From Lemma 2.3.2, the second term is of the order $\frac{1}{N}$. Now we will show that

$$\sum_{j=1}^{n}\frac{1}{1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}\mathbf{S}^{-1}\mathbf{a}_j-\sum_{j=1}^{n}\mathbb{E}\beta_j\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j=-\sum_{j=1}^{n}\frac{b_j^2}{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j+o(1).\tag{2.5.9}$$

From the relation that $\mathbf{Q}=\mathbf{Q}_j-\beta_j\mathbf{Q}_j\mathbf{y}_j\mathbf{y}_j^*\mathbf{Q}_j$, we have

$$\begin{aligned}&\sum_{j=1}^{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}\mathbf{S}^{-1}\mathbf{a}_j\\&=\sum_{j=1}^{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j-\sum_{j=1}^{n}\mathbb{E}\beta_j\mathbf{a}_j^*\mathbf{Q}_j\mathbf{y}_j\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\\&\overset{(a)}{=}\sum_{j=1}^{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j-\sum_{j=1}^{n}\frac{b_j}{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j-\sum_{j=1}^{n}b_j\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j+o(1)\\&\overset{(b)}{=}\sum_{j=1}^{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j-\sum_{j=1}^{n}\frac{b_j}{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j-\sum_{j=1}^{n}b_j\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j+o(1)\end{aligned}$$

where (a) follows from $\mathbb{E}\beta_j=b_j+o(1)$ and (b) is using the fact that $\mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j$ and $\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j$ are uncorrelated. From the definition of b_j, (cf. (2.3.9)),

$$\frac{1}{1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}}-b_j=\frac{b_j\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j}{1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}}+o(1),$$

Then we have

$$
\begin{aligned}
&\sum_{j=1}^{n}\frac{1}{1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}_j}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j-\sum_{j=1}^{n}\mathbb{E}\beta_j\mathbf{a}_j^*\mathbf{Q}\mathbf{S}^{-1}\mathbf{a}_j\\
&\quad-\sum_{j=1}^{n}\frac{b_j}{1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}_j}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\\
&= \sum_{j=1}^{n}\left(\frac{1}{1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}_j}-b_j\right)\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\\
&\quad-\sum_{j=1}^{n}\frac{b_j}{1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}_j}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j+o(1)\\
&= o(1).
\end{aligned}
\tag{2.5.10}
$$

Moreover

$$
\begin{aligned}
&\sum_{j=1}^{n}\frac{b_j}{n(1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}_j)}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j-\sum_{j=1}^{n}\frac{b_j^2}{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\\
&= \sum_{j=1}^{n}\frac{b_j}{n(1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}_j)}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j-\sum_{j=1}^{n}\frac{b_j^2}{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\\
&= \sum_{j=1}^{n}\frac{b_j^2\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j}{n(1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}_j)}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j+o(1)\\
&= \sum_{j=1}^{n}\frac{b_j^2}{n(1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}_j)}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j\mathbf{a}_j^*\mathbf{Q}_j\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\\
&\quad-\sum_{j=1}^{n}\frac{b_j^2}{n(1+\frac{1}{n}\mathrm{Tr}\mathbb{E}\mathbf{Q}_j)}\mathrm{Cov}\left(\mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j,\mathbf{a}_j^*\mathbf{Q}_j\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\right)+o(1).
\end{aligned}
\tag{2.5.11}
$$

From Lemma 2.3.1, the second term is of the order $\mathcal{O}\left(\frac{1}{n}\right)$ and the first term is also of the order $\mathcal{O}\left(\frac{1}{n}\right)$ as

$$\left\|\mathbb{E}\mathbf{Q}_j\mathbf{a}_j\mathbf{a}_j^*\mathbf{Q}_j\mathbf{Q}_j\mathbf{S}^{-1}\right\|=\mathcal{O}(n^{-1}).$$

Thus, (2.5.10) and (2.5.11) imply (2.5.9).

Since for all matrices $\mathbf{M}_1,\mathbf{M}_2$ with bounded spectral measures,

$$\frac{1}{n}\mathrm{Tr}\mathbf{M}_1\mathbf{S}^{-1}\mathbf{M}_2-\frac{1}{n}\mathrm{Tr}\mathbf{M}_1\mathbf{T}\mathbf{M}_2\xrightarrow[N,n\to\infty]{}0,$$

gathering (2.5.6), (2.5.7), (2.5.8) and (2.5.9), we have

$$nA_n = -\sum_{j=1}^{n}\frac{b_j^2}{n^2}\mathbb{E}\mathrm{Tr}\mathbf{Q}_j\mathbf{T}\mathbf{Q}_j-\sum_{j=1}^{n}\frac{b_j^2}{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{T}\mathbf{Q}_j\mathbf{a}_j-\sum_{j=1}^{n}\frac{b_j^2}{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{Q}_j\mathbf{T}\mathbf{a}_j$$

$$+\sum_{j=1}^{n} b_j^2 \left[\mathbb{E}\left(\mathbf{y}_j^*\mathbf{Q}_j\mathbf{y}_j - n^{-1}\mathrm{Tr}\mathbf{Q}_j - \mathbf{a}_j^*\mathbf{Q}_j\mathbf{a}_j\right) \times \left(\mathbf{y}_j^*\mathbf{Q}_j\mathbf{T}\mathbf{y}_j - \frac{1}{n}\mathrm{Tr}\mathbf{Q}_j\mathbf{T} - \mathbf{a}_j^*\mathbf{Q}_j\mathbf{T}\mathbf{a}_j\right)\right] + o(1).$$

By Proposition 2.4.2, and with the same argument as Lemma 2.4.4, the terms with $\mathbb{E}[|x_1|^2 x_1]$ and $\mathbb{E}[|x_1|^2\bar{x}_1]$ will not contribute in the sum. Then we have

$$\begin{aligned} nA_n = &\sum_{j=1}^{n} \frac{|\vartheta|^2 b_j^2}{n^2}\mathbb{E}\mathrm{Tr}\mathbf{Q}_j\mathbf{T}\mathbf{Q}_j^T + \sum_{j=1}^{n}\frac{\vartheta b_j^2}{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{T}\mathbf{Q}_j^T\bar{\mathbf{a}}_j \\ &+ \sum_{j=1}^{n}\frac{\bar{\vartheta} b_j^2}{n}\mathbb{E}\mathbf{a}_j^T\mathbf{Q}_j^T\mathbf{Q}_j\mathbf{T}\mathbf{a}_j + \kappa\sum_{j=1}^{n}\frac{b_j^2}{n^2}\mathbb{E}\sum_{i=1}^{N}[Q_j]_{ii}[Q_jT]_{ii} + o(1). \end{aligned}$$

By Lemma 2.4.1,

$$\frac{1}{n}\mathbb{E}\sum_{i=1}^{N}[Q_j]_{ii}[Q_jT]_{ii} - \frac{1}{n}\sum_{i=1}^{N} t_{ii}[T^2]_{ii} \xrightarrow[n\to\infty]{} 0.$$

Hence, as $b_j = -z\tilde{t}_{jj} + o(1)$, the term on κ is

$$\kappa\sum_{j=1}^{n}\frac{b_j^2}{n^2}\mathbb{E}\sum_{i=1}^{N}[Q_j]_{ii}[Q_jT]_{ii} = \frac{\kappa z^2}{n^2}\sum_{j=1}^{n}\tilde{t}_{jj}^2\sum_{i=1}^{N}t_{ii}[T^2]_{ii} + o(1).$$

Now we will calculate

$$\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}b_j^2\mathrm{Tr}\mathbf{Q}_j\mathbf{T}\mathbf{Q}_j^T.$$

From the argument in the proof of Lemma 2.4.4 and 2.4.9, consider

$$\begin{aligned} \underline{\psi}_j^{\mathbf{T}}(z_1,z_2) &= \frac{1}{n}\mathrm{Tr}\mathbb{E}\{[\mathbb{E}_j\mathbf{Q}(z_1)]\mathbf{T}\mathbb{E}_j\mathbf{Q}^T(z_2)\} = \frac{1}{n}\mathrm{Tr}\mathbb{E}\{[\mathbb{E}_j\mathbf{Q}(z_1)]\mathbf{T}\mathbf{Q}^T(z_2)\}, \\ \underline{\zeta}_{kj}^{\mathbf{T}}(z_1,z_2) &= \mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}(z_1)]\mathbf{T}\mathbb{E}_j\mathbf{Q}^T(z_2)\bar{\mathbf{a}}_k\} = \mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}(z_1)]\mathbf{T}\mathbf{Q}^T(z_2)\bar{\mathbf{a}}_k\}, \\ \underline{\theta}_{kj}^{\mathbf{T}}(z_1,z_2) &= \mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}_k(z_1)]\mathbf{T}\mathbb{E}_j\mathbf{Q}_k^T(z_2)\bar{\mathbf{a}}_k\} = \mathbb{E}\{\mathbf{a}_k^*[\mathbb{E}_j\mathbf{Q}_k(z_1)]\mathbf{T}\mathbf{Q}_k^T(z_2)\bar{\mathbf{a}}_k\}, \\ \underline{\phi}_j^{\mathbf{T}} &= \frac{1}{n}\sum_{k=1}^{j} z_1z_2\tilde{t}_{kk}(z_1)\tilde{t}_{kk}(z_2)\underline{\theta}_{kj}^{\mathbf{T}}. \end{aligned}$$

The perturbed system is :

$$\begin{cases} (1-\vartheta\nu_j^{\dagger})\underline{\phi}_j^{\mathbf{T}} - (\bar{\vartheta}\omega_j^{\dagger} + |\vartheta|^2\nu_j^{\dagger}\eta_j)\underline{\psi}_j^{\mathbf{T}} = \nu_j^{\dagger}(\mathbf{T}) + \mathcal{O}(n^{-1/2}) \\ -\vartheta\gamma^{\dagger}\underline{\phi}_j^{\mathbf{T}} + (1-\bar{\vartheta}\tilde{\nu}_j^{\dagger} - |\vartheta|^2\gamma^{\dagger}\eta_j)\underline{\psi}_j^{\mathbf{T}} = \gamma_{\mathbf{T}}^{\dagger} + \mathcal{O}(n^{-1/2}). \end{cases} \tag{2.5.12}$$

where

$$\nu_j^{\dagger}(\mathbf{T}) = \frac{1}{n}\sum_{k=1}^{j}\frac{\mathbf{a}_k^*\mathbf{T}_1\mathbf{T}\mathbf{T}_2^T\bar{\mathbf{a}}_k}{(1+\delta_1)(1+\delta_2)}$$

$$\gamma_{\mathbf{T}}^{\dagger} = \frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{T}\mathbf{T}_2^T.$$

The determinant of this system is $\mathbf{\Delta}_j^{\vartheta}$ and the resolution of the system yields that

$$\begin{bmatrix} \underline{\phi}_j^{\mathbf{T}} \\ \underline{\psi}_j^{\mathbf{T}} \end{bmatrix} = \frac{1}{\mathbf{\Delta}_j^{\vartheta}} \begin{bmatrix} \underline{F}_j^{\mathbf{T}}(1-\bar{\vartheta}\tilde{\nu}_j^{\dagger} - |\vartheta|^2\gamma^{\dagger}\eta_j) + \gamma_j^{\mathbf{T}}(\bar{\vartheta}\omega_j^{\dagger} + |\vartheta|^2\nu_j^{\dagger}\eta_j) \\ \vartheta\gamma^{\dagger}\nu_j^{\dagger}(\mathbf{T}) + \gamma_{\mathbf{T}}^{\dagger}(1-\vartheta\nu_j^{\dagger}) \end{bmatrix} + \mathcal{O}(n^{-1/2}). \tag{2.5.13}$$

Then we have

$$\begin{aligned} \frac{1}{n^2}\sum_{j=1}^n \mathbb{E}b_j^2\mathrm{Tr}\mathbf{Q}_j\mathbf{T}\mathbf{Q}_j^T &= \frac{1}{n^2}\sum_{j=1}^n \mathbb{E}b_j^2\mathrm{Tr}\mathbf{Q}\mathbf{T}\mathbf{Q}^T + o(1) \\ &= \frac{z^2}{n^2}\underline{\psi}_n^{\mathbf{T}}(z,z)\sum_{j=1}^n \tilde{t}_{jj}^2 + o(1). \end{aligned}$$

Now we will calculate $\sum_{j=1}^n \frac{\bar{\vartheta}b_j^2}{n}\mathbb{E}\mathbf{a}_j^T\mathbf{Q}_j^T\mathbf{Q}_j\mathbf{T}\mathbf{a}_j$. We first show that

$$\left|\sum_{j=1}^n \frac{\bar{\vartheta}b_j^2}{n}\left(\mathbb{E}\mathbf{a}_j^T\mathbf{Q}_j^T\mathbf{Q}_j\mathbf{T}\mathbf{a}_j - \mathbb{E}\mathbf{a}_j^T\mathbf{Q}_j^T\mathbf{T}\mathbf{Q}_j\mathbf{a}_j\right)\right| \xrightarrow[n\to\infty]{} 0.$$

By Lemma 2.4.3, it suffices to prove that

$$\left|\sum_{j=1}^n \frac{\bar{\vartheta}b_j^2}{n}\left(\mathbb{E}\mathbf{a}_j^T\mathbf{Q}^T\mathbf{Q}\mathbf{T}\mathbf{a}_j - \mathbb{E}\mathbf{a}_j^T\mathbf{Q}^T\mathbf{T}\mathbf{Q}\mathbf{a}_j\right)\right| \xrightarrow[n\to\infty]{} 0. \tag{2.5.14}$$

We have

$$\begin{aligned} &\left|\sum_{j=1}^n \frac{\bar{\vartheta}b_j^2}{n}\mathbb{E}\mathbf{a}_j^T\mathbf{Q}^T\mathbf{Q}\mathbf{T}\mathbf{a}_j - \sum_{j=1}^n \frac{\bar{\vartheta}b_j^2}{n}\mathbb{E}\mathbf{a}_j^T\mathbf{Q}^T\mathbf{T}\mathbf{T}\mathbf{a}_j\right| \\ &\leq \frac{K}{n}\sum_{j=1}^n \left|\mathbb{E}\mathbf{a}_j^T\mathbf{Q}^T(\mathbf{Q}\mathbf{T}-\mathbf{T}\mathbf{T})\mathbf{a}_j\right| \\ &\overset{(a)}{\leq} \frac{K}{n}\sum_{j=1}^n \mathbb{E}\left\|(\mathbf{Q}\mathbf{T}-\mathbf{T}\mathbf{T})\mathbf{a}_j\right\| \\ &\overset{(b)}{=} o(1), \end{aligned}$$

where (a) follows from the fact that $\|\mathbf{Q}\| \leq (\mathrm{Im}z)^{-1}$ and (b) is from (2.4.14) in Theorem 2.4.1. Same result holds true for $\frac{\bar{\vartheta}b_j^2}{n}\mathbb{E}\mathbf{a}_j^T\mathbf{Q}^T\mathbf{T}\mathbf{Q}\mathbf{a}_j$:

$$\left|\sum_{j=1}^n \frac{\bar{\vartheta}b_j^2}{n}\mathbb{E}\mathbf{a}_j^T\mathbf{Q}^T\mathbf{T}\mathbf{T}\mathbf{a}_j - \sum_{j=1}^n \frac{\bar{\vartheta}b_j^2}{n}\mathbb{E}\mathbf{a}_j^T\mathbf{Q}^T\mathbf{T}\mathbf{Q}\mathbf{a}_j\right| = o(1).$$

This proves (2.5.14).

Then it suffices to consider

$$\begin{aligned}
\underline{\tilde{\psi}}_j^{\mathbf{T}}(z_1,z_2) &= \frac{1}{n}\mathrm{Tr}\mathbb{E}[\mathbb{E}_j\mathbf{Q}^T(z_1)\mathbf{T}\mathbb{E}_j\mathbf{Q}(z_2)] = \frac{1}{n}\mathrm{Tr}\mathbb{E}[\mathbb{E}_j\mathbf{Q}^T(z_1)\mathbf{T}\mathbf{Q}(z_2)],\\
\underline{\tilde{\zeta}}_{kj}^{\mathbf{T}}(z_1,z_2) &= \mathbb{E}[\mathbf{a}_k^T\mathbb{E}_j\mathbf{Q}^T(z_1)\mathbf{T}\mathbb{E}_j\mathbf{Q}(z_2)\mathbf{a}_k] = \mathbb{E}[\mathbf{a}_k^T\mathbb{E}_j\mathbf{Q}^T(z_1)\mathbf{T}\mathbf{Q}(z_2)\mathbf{a}_k],\\
\underline{\tilde{\theta}}_{kj}^{\mathbf{T}}(z_1,z_2) &= \mathbb{E}[\mathbf{a}_k^T\mathbb{E}_j\mathbf{Q}_k^T(z_1)\mathbf{T}\mathbb{E}_j\mathbf{Q}_k(z_2)\mathbf{a}_k] = \mathbb{E}[\mathbf{a}_k^T\mathbb{E}_j\mathbf{Q}_k^T(z_1)\mathbf{T}\mathbf{Q}_k(z_2)\mathbf{a}_k],\\
\underline{\tilde{\phi}}_j^{\mathbf{T}} &= \frac{1}{n}\sum_{k=1}^{j} z_1z_2\tilde{t}_{kk}(z_1)\tilde{t}_{kk}(z_2)\underline{\tilde{\theta}}_{kj}^{\mathbf{T}}.
\end{aligned}$$

We have the solution

$$\begin{bmatrix}\underline{\tilde{\phi}}_j^{\mathbf{T}}\\ \underline{\tilde{\psi}}_j^{\mathbf{T}}\end{bmatrix} = \frac{1}{\mathbf{\Delta}_j^{\vartheta}}\begin{bmatrix}\tilde{\nu}_j^{\dagger}(\mathbf{T})(1-\vartheta\nu_j^{\dagger}-|\vartheta|^2\gamma^{\dagger}\eta_j)+\underline{\tilde{\gamma}}^{\mathbf{T}}(\vartheta\omega_j^{\dagger}+|\vartheta|^2\tilde{\nu}_j^{\dagger}\eta_j)\\ \bar{\vartheta}\gamma^{\dagger}\tilde{\nu}_j^{\dagger}(\mathbf{T})+\tilde{\gamma}_{\mathbf{T}}^{\dagger}(1-\bar{\vartheta}\tilde{\nu}_j^{\dagger})\end{bmatrix}+\mathcal{O}(n^{-1/2}), \tag{2.5.15}$$

where

$$\begin{aligned}
\tilde{\nu}_j^{\dagger}(\mathbf{T}) &= \frac{1}{n}\sum_{k=1}^{j}\frac{\mathbf{a}_k^T\mathbf{T}_1^T\mathbf{T}\mathbf{T}_2\mathbf{a}_k}{(1+\delta_1)(1+\delta_2)}\\
\tilde{\gamma}_{\mathbf{T}}^{\dagger} &= \frac{1}{n}\mathrm{Tr}\mathbf{T}_1^T\mathbf{T}\mathbf{T}_2.
\end{aligned}$$

Then with (2.5.13) and (2.5.15), we have

$$\begin{aligned}
&\frac{|\vartheta|^2}{n}\sum_{j=1}^{n}\mathbb{E}b_j^2\mathrm{Tr}\mathbf{Q}_j\mathbf{T}\mathbf{Q}_j^T+\sum_{j=1}^{n}\frac{\vartheta b_j^2}{n}\mathbb{E}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{T}\mathbf{Q}_j^T\bar{\mathbf{a}}_j+\sum_{j=1}^{n}\frac{\bar{\vartheta}b_j^2}{n}\mathbb{E}\mathbf{a}_j^T\mathbf{Q}_j^T\mathbf{Q}_j\mathbf{T}\mathbf{a}_j\\
=&|\vartheta|^2\omega(z,z)\underline{\psi}_n^{\mathbf{T}}(z,z)+\vartheta\underline{\phi}_n^{\mathbf{T}}(z,z)+\bar{\vartheta}\underline{\tilde{\phi}}_n^{\mathbf{T}}(z,z)\\
=&\frac{1}{\Delta_n^{\vartheta}}\Big[|\vartheta|^2\eta(\vartheta\gamma^{\dagger}\nu_{\mathbf{T}}^{\dagger}+\gamma_{\mathbf{T}}^{\dagger}(1-\vartheta\nu^{\dagger}))+\vartheta\nu_{\mathbf{T}}^{\dagger}(1-\bar{\vartheta}\tilde{\nu}^{\dagger}-|\vartheta|^2\gamma^{\dagger}\eta)\\
&+\vartheta\gamma_{\mathbf{T}}^{\dagger}(\bar{\vartheta}\omega^{\dagger}+|\vartheta|^2\nu^{\dagger}\eta)+\bar{\vartheta}\tilde{\nu}_{\mathbf{T}}^{\dagger}(1-\vartheta\nu^{\dagger}-|\vartheta|^2\gamma^{\dagger}\eta)+\bar{\vartheta}\underline{\tilde{\gamma}}^{\mathbf{T}}(\vartheta\omega^{\dagger}+|\vartheta|^2\tilde{\nu}^{\dagger}\eta)\Big]+o(1),
\end{aligned}$$

where all terms are evaluated on $(z_1,z_2)=(z,z)$.

Combining this with (2.5.3) and recall the definition of Υ_n in (2.2.9), and $\eta_n+\omega_n^{\dagger}=\tilde{\gamma}^{\dagger}$, we achieve the final expression of the bias in (2.2.16).

The proof is completed by showing that $\mathbb{E}(\mathrm{Tr}\mathbf{Q})-n\delta_n$ forms an equicontinuous family for $z\in\mathcal{C}_n$. With the precedent work, it suffices to show that $L_n'(z)$ is bounded for $z\in\mathcal{C}_n$ where

$$L_n(z)=\sum_{j=1}^{n}b_j^2\left[\mathbb{E}\hat{\gamma}_j\left(\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{y}_j-\frac{1}{n}\mathrm{Tr}\mathbf{Q}_j\mathbf{S}^{-1}-\mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\right)\right].$$

The terms in $L'_n(z)$ are linear combinations of

$$\begin{aligned}
L_n^1(z) &= 2b_jb'_j\left[\mathbb{E}\hat{\gamma}_j\left(\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{y}_j - \frac{1}{n}\mathrm{Tr}\mathbf{Q}_j\mathbf{S}^{-1} - \mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\right)\right]\\
L_n^2(z) &= b_j^2\left[\mathbb{E}\hat{\gamma}'_j\left(\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{y}_j - \frac{1}{n}\mathrm{Tr}\mathbf{Q}_j\mathbf{S}^{-1} - \mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-1}\mathbf{a}_j\right)\right]\\
L_n^3(z) &= b_j^2\left[\mathbb{E}\hat{\gamma}_j\left(\mathbf{y}_j^*\mathbf{Q}_j^2\mathbf{S}^{-1}\mathbf{y}_j - \frac{1}{n}\mathrm{Tr}\mathbf{Q}_j^2\mathbf{S}^{-1} - \mathbf{a}_j^*\mathbf{Q}_j^2\mathbf{S}^{-1}\mathbf{a}_j\right)\right]\\
L_n^4(z) &= -b_j^2\left[\mathbb{E}\hat{\gamma}_j\left(\mathbf{y}_j^*\mathbf{Q}_j\mathbf{S}^{-2}\mathbf{S}'\mathbf{y}_j - \frac{1}{n}\mathrm{Tr}\mathbf{Q}_j\mathbf{S}^{-2}\mathbf{S}' - \mathbf{a}_j^*\mathbf{Q}_j\mathbf{S}^{-2}\mathbf{S}'\mathbf{a}_j\right)\right].
\end{aligned}$$

According to Lemma 2.3.5, all terms are indeed bounded. This concludes the computations of the bias.

□

2.6 Appendices

2.6.1 Proof of Proposition 2.3.1

We begin the proof with the replacement of the entries x_{ij} satisfying that there exists a sequence $\eta_n \to 0$ such that $\eta_n n^{1/5} \to \infty$, $|x_{ij}| < \eta_n\sqrt{n}$. By Assumption 2.2.1, we may select $\eta_n \downarrow 0$ such that :

$$\frac{1}{\eta_n^4}\mathbb{E}|x_{11}|^4\mathbb{I}(|x_{11}| \geq \eta_n\sqrt{n}) \xrightarrow[N,n\to\infty]{} 0.$$

Let $\widehat{\mathbf{Y}}_n = (\frac{1}{\sqrt{n}}\widehat{\mathbf{X}}_n + \mathbf{A}_n)$ with $\widehat{\mathbf{X}}_n$, an $N\times n$ matrix with the entries $\hat{x}_{ij} = x_{ij}\mathbb{I}_{|x_{ij}|\leq\eta_n\sqrt{n}}$. We have then

$$\begin{aligned}
\mathbb{P}(\mathbf{Y}_n\mathbf{Y}_n^* \neq \widehat{\mathbf{Y}}_n\widehat{\mathbf{Y}}_n^*) &\leq \sum_{ij}\mathbb{P}(|x_{ij}| \geq \eta_n\sqrt{n})\\
&\leq \frac{1}{\eta^4}\mathbb{E}|x_{11}|^4\mathbb{I}(|x_{11}| \geq \eta_n\sqrt{n}) = o(1).
\end{aligned} \tag{2.6.1}$$

Denote $\widetilde{\mathbf{Y}}_n = (\frac{1}{\sqrt{n}}\widetilde{\mathbf{X}}_n + \mathbf{A}_n)$ with the entries $\tilde{x}_{ij} = (\hat{x}_{ij} - \mathbb{E}\hat{x}_{ij})/\sigma_{ij}$, where

$$\sigma_{ij}^2 = \mathbb{E}|\hat{x}_{ij} - \mathbb{E}\hat{x}_{ij}|^2.$$

We will show that

$$\mathbb{E}(\tilde{x}_{11}^2) \xrightarrow[n\to\infty]{} \vartheta, \quad \mathbb{E}|\tilde{x}_{11}|^4 \xrightarrow[n\to\infty]{} \mathbb{E}|x_{11}|^4,$$

and

$$\mathrm{Tr}f(\widetilde{\mathbf{Y}}_n\widetilde{\mathbf{Y}}_n^*) = \mathrm{Tr}f(\mathbf{Y}_n\mathbf{Y}_n^*) + o_{\mathbb{P}}(1). \tag{2.6.2}$$

Since $\mathbb{E}|x_{11}|^2 < \infty$, in particular, $\mathbb{E}|x_{11}| < \infty$. By dominated convergence theorem,

$$|\mathbb{E}\hat{x}_{11}| \xrightarrow[n\to\infty]{} 0, \quad |\sigma_{11} - 1| \xrightarrow[n\to\infty]{} 0, \quad |\mathbb{E}\hat{x}_{11}^2 - x_{11}^2| \xrightarrow[n\to\infty]{} 0, \tag{2.6.3}$$

which implies that

$$\left|\mathbb{E}(\tilde{x}_{11}^2) - \vartheta\right| = \left|\mathbb{E}\left(\frac{\hat{x}_{11} - \mathbb{E}\hat{x}_{11}}{\sigma_{11}}\right)^2 - \mathbb{E}x_{11}^2\right| \xrightarrow[n\to\infty]{} 0.$$

Same method yields that

$$\mathbb{E}|\tilde{x}_{11}|^4 \xrightarrow[n\to\infty]{} \mathbb{E}|x_{11}|^4.$$

We shall use the following proposition (cf. [6, Corollary A.42]) to show (2.6.2).

Proposition 2.6.1 *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two* $N \times n$ *matrices and the empirical estimate distribution of* $\mathbf{S}^{\mathbf{A}} = \mathbf{A}\mathbf{A}^*$ *and* $\mathbf{S}^{\mathbf{B}} = \mathbf{B}\mathbf{B}^*$ *be denoted by* $F^{\mathbf{S}^{\mathbf{A}}}$ *and* $F^{\mathbf{S}^{\mathbf{B}}}$*. Then,*

$$d^4_{\mathcal{LP}}(F^{\mathbf{S}^{\mathbf{A}}}, F^{\mathbf{S}^{\mathbf{B}}}) \le \frac{2}{N^2}(\mathrm{Tr}(\mathbf{A}\mathbf{A}^* + \mathbf{B}\mathbf{B}^*))(\mathrm{Tr}[\mathbf{A}-\mathbf{B}][\mathbf{A}-\mathbf{B}]^*).$$

As

$$\left(\frac{1}{\sqrt{n}}\widehat{\mathbf{X}}_n + \mathbf{A}_n\right)\left(\frac{1}{\sqrt{n}}\widehat{\mathbf{X}}_n + \mathbf{A}_n\right)^* \le 2\left(\frac{1}{n}\widehat{\mathbf{X}}_n\widehat{\mathbf{X}}_n^* + \mathbf{A}_n\mathbf{A}_n^*\right), \text{[2]}$$

denote $\lambda^{\mathbf{M}}_{\max}$, the largest eigenvalue of the matrix $\mathbf{M}$, then we get

$$\limsup \lambda_{\max}^{\widehat{\mathbf{Y}}_n\widehat{\mathbf{Y}}_n^*} \le 2\limsup \lambda_{\max}^{\frac{1}{n}\widehat{\mathbf{X}}_n\widehat{\mathbf{X}}_n^*} + 2\lambda_{\max}^{\mathbf{A}_n\mathbf{A}_n^*}.$$

From [5, Theorem 5.11], $\lambda_{\max}^{\frac{1}{n}\widehat{\mathbf{X}}_n\widehat{\mathbf{X}}_n^*}$ is bounded almost surely, then $\limsup \lambda_{\max}^{\widehat{\mathbf{Y}}_n\widehat{\mathbf{Y}}_n^*}$ is bounded almost surely as well by Assumption 2.2.2. Same method yields that $\limsup \lambda_{\max}^{\widetilde{\mathbf{Y}}_n\widetilde{\mathbf{Y}}_n^*}$ is bounded almost surely. As the deterministic equivalent δ depends only on $\mathbf{A}$ (see Eq. (2.1.4)), we can denote $\widehat{G}_n(x) = N \times (F^{\widehat{\mathbf{Y}}\widehat{\mathbf{Y}}^*} - \mathcal{F}_n)$ and $\widetilde{G}_n(x) = N(F^{\widetilde{\mathbf{Y}}\widetilde{\mathbf{Y}}^*} - \mathcal{F}_n)$ and let $\lambda_i^{\mathbf{A}}$ denote the i-th smallest eigenvalue of $\mathbf{A}$. Using Proposition 2.6.1, for an analytic function f which satisfies Assumption 2.2.3, we have,

$$\begin{aligned}
&\mathbb{E}\left|\int f(x)d\widehat{G}_n(x) - \int f(x)d\widetilde{G}_n(x)\right| \qquad (2.6.4)\\
\le\; &\mathbb{E}\left|\sum_{j=1}^N f(\hat{\lambda}_j) - f(\tilde{\lambda}_j)\right| \\
\le\; &K\mathbb{E}\sum_{j=1}^N \left|\hat{\lambda}_j - \tilde{\lambda}_j\right| \\
\le\; &\frac{2K}{n}\mathbb{E}^{1/2}\left\{\mathrm{Tr}(\widehat{\mathbf{X}}_n - \widetilde{\mathbf{X}}_n)(\widehat{\mathbf{X}}_n - \widetilde{\mathbf{X}}_n)^*\right\}\mathbb{E}^{1/2}\left\{\mathrm{Tr}[\widehat{\mathbf{Y}}\widehat{\mathbf{Y}}^* + \widetilde{\mathbf{Y}}\widetilde{\mathbf{Y}}^*]\right\}.
\end{aligned}$$

where K is the bound of $|f'(x)|$ in the compact which contains both supports of $\widehat{\mathbf{Y}}\widehat{\mathbf{Y}}^*$ and $\widetilde{\mathbf{Y}}\widetilde{\mathbf{Y}}^*$.

2. For two complex Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$, we denote $\mathbf{A} \le \mathbf{B}$ if $\mathbf{B} - \mathbf{A}$ is a Hermitian nonnegative matrix.

From the definition of $\hat{x}_{ij}$,

$$
\begin{aligned}
|\mathbb{E}|\hat{x}_{11}|^2 - 1|^2 &= (\mathbb{E}|x_{11}|^2 \mathbb{I}_{|x_{11}|\geq \eta_n\sqrt{n}})^2 \\
&\leq \frac{1}{\eta_n^4 n^2}(\mathbb{E}|x_{11}|^4 \mathbb{I}_{|x_{11}|\geq \eta_n\sqrt{n}})^2 = o(n^{-2}).
\end{aligned}
$$

Moreover,

$$
|\mathbb{E}\hat{x}_{11}|^2 \leq \frac{1}{n^3\eta_n^6}\left(\mathbb{E}|x_{11}|^4 \mathbb{I}(|x_{11}| \geq \eta_n\sqrt{n})\right)^2 = o(n^{-2}).
$$

These give us

$$
\begin{aligned}
& \mathbb{E}\mathrm{Tr}(\widehat{\mathbf{X}}_n - \widetilde{\mathbf{X}}_n)(\widehat{\mathbf{X}}_n - \widetilde{\mathbf{X}}_n) \\
\leq \;& 2\sum_{ij}[(1-\sigma_{ij}^{-1})^2\mathbb{E}|\hat{x}_{ij}|^2 + \sigma_{ij}^{-2}|\mathbb{E}\hat{x}_{ij}|^2] \\
\leq \;& 2\sum_{ij}[(1-\sigma_{ij}^{-1})^2 + |\mathbb{E}\hat{x}_{ij}|^2] = o(1).
\end{aligned} \tag{2.6.5}
$$

Similarly,

$$
\begin{aligned}
& \frac{1}{n}\mathbb{E}\left\{\mathrm{Tr}\left[\widehat{\mathbf{Y}}\widehat{\mathbf{Y}}^* + \widetilde{\mathbf{Y}}\widetilde{\mathbf{Y}}^*\right]\right\} \\
\leq \;& \frac{1}{n}\sum_{ij}\mathbb{E}\left[\left|\frac{1}{\sqrt{n}}\hat{x}_{ij} + a_{ij}\right|^2 + \left|\frac{1}{\sqrt{n}}\tilde{x}_{ij} + a_{ij}\right|^2\right] \\
\leq \;& \frac{2}{n}\sum_{ij}\mathbb{E}\left[\frac{1}{n}|\hat{x}_{ij}|^2 + \frac{1}{n}|\tilde{x}_{ij}|^2 + 2\,|a_{ij}|^2\right] \leq K.
\end{aligned} \tag{2.6.6}
$$

Plugging the estimations (2.6.5) and (2.6.6) into (2.6.4), along with (2.6.1), we conclude that

$$
\mathrm{Tr}f(\widetilde{\mathbf{Y}}\widetilde{\mathbf{Y}}^*) = \mathrm{Tr}f(\mathbf{Y}\mathbf{Y}^*) + o_{\mathbb{P}}(1).
$$

2.6.2 Proof of Proposition 2.3.2

It has been shown that the largest value of the limiting support $\mathcal{S}$ of the eigenvalues of $\frac{1}{n}\mathbf{X}_n\mathbf{X}_n^*$ is $(1+\sqrt{c})^2$ (cf. [9, 102]). Let the contour $\mathcal{C}_n$, ε_n be defined in Section 2.3.2 and $\mu_\ell < 0$ and $\mu_r > \ell_+$.

The partition of $\mathcal{C}_n$ is identical to that used in [11, Section 1]. With probability one (see [9] and [10]), for all $\epsilon > 0$,

$$
\lim \sup_{\lambda \in \mathrm{eig}(\mathbf{Y}_n\mathbf{Y}_n^*)} d(\lambda, \mathcal{S}_n) < \epsilon
$$

where $d(x,S)$ is the Euclidean distance of x to the set S, $\mathrm{eig}(\mathbf{Y}_n\mathbf{Y}_n^*)$ denotes the set of eigenvalues of $\mathbf{Y}_n\mathbf{Y}_n^*$ and $\mathcal{S}_n$ is the support associated with $\mathcal{F}_n$. Then with probability one, for all N large, as f is continuous over $\mathcal{C}$,

$$
\left|\oint f(z)\Big(M_n(z) - \hat{M}_n(z)\Big)dz\right| \leq 4\sup_{z\in\mathcal{C}}|f(z)|\varepsilon_n(|\max(\ell_+, \lambda_{\max}^{\mathbf{Y}_n\mathbf{Y}_n^*}) - \mu_r)|^{-1} + |\mu_\ell|^{-1}).
$$

This term converges to zero as $n \to \infty$.

2.6.3 Proof of Proposition 2.3.2

$$\mathbb{E}\left|\frac{\mathbf{x}^*\mathbf{M}_n\mathbf{x}_n}{n} - \frac{\mathrm{Tr}\mathbf{M}_n}{n}\right|^2 \leq \frac{K}{n} \tag{2.6.7}$$

has been proved in [11, Eq. (3.2)]. Write

$$\begin{aligned}
&\mathbb{E}\left|\mathbf{y}_n\mathbf{M}_n\mathbf{y}_n - \frac{1}{n}\mathrm{Tr}\mathbf{M}_n - \mathbf{a}_n\mathbf{M}_n\mathbf{a}_n\right|^2 \\
\leq & K\left[\mathbb{E}\left|\frac{1}{n}\mathbf{x}_n\mathbf{M}_n\mathbf{x}_n - \frac{1}{n}\mathrm{Tr}\mathbf{M}_n\right|^2 + \frac{K}{n}\mathbb{E}|\mathbf{a}_n^*\mathbf{M}_n\mathbf{x}_n|^2 + \frac{K}{n}\mathbb{E}|\mathbf{x}_n^*\mathbf{M}_n\mathbf{a}_n|^2\right].
\end{aligned}$$

The first term is with the order $\mathcal{O}(n^{-1})$ by (2.6.7). The second and the third terms are similar and we treat $\frac{1}{n}\mathbb{E}|\mathbf{a}_n^*\mathbf{M}_n\mathbf{x}_n|^2$:

$$\begin{aligned}
\frac{1}{n}\mathbb{E}|\mathbf{a}_n^*\mathbf{M}_n\mathbf{x}_n|^2 &= \frac{1}{n}\mathbb{E}(\mathbf{x}_n^*\mathbf{M}_n^*\mathbf{a}_n\mathbf{a}_n^*\mathbf{M}_n\mathbf{x}_n) \\
&\leq \frac{K}{n}\mathbb{E}\left|\frac{1}{n}\mathbf{x}_n^*\mathbf{M}_n^*\mathbf{a}_n\mathbf{a}_n^*\mathbf{M}_n\mathbf{x}_n - \frac{1}{n}\mathrm{Tr}(\mathbf{M}_n^*)\mathbf{a}_j\mathbf{a}_j^*\mathbf{M}_n\right| \\
&\quad + \frac{K}{n}\mathbb{E}\left|\frac{1}{n}\mathrm{Tr}\mathbf{M}_n\mathbf{a}_n\mathbf{a}_n^*\mathbf{M}_n\right|.
\end{aligned}$$

The first term in the above inequality is with the order $\mathcal{O}(n^{-1})$ by (2.6.7) and the second term is of the right order.

2.6.4 Proof of Lemma 2.3.3

From

$$\begin{aligned}
&\mathbb{E}\left|\frac{1}{n}\mathbf{x}_n^*\mathbf{M}_n\mathbf{x}_n - \frac{1}{n}\mathrm{Tr}\mathbf{M}_n + \frac{1}{\sqrt{n}}\mathbf{a}_n^*\mathbf{M}_n\mathbf{x}_n + \frac{1}{\sqrt{n}}\mathbf{x}_n^*\mathbf{M}_n\mathbf{a}_n\right|^4 \\
\leq & K\mathbb{E}\left|\frac{1}{n}\mathbf{x}_n^*\mathbf{M}_n\mathbf{x}_n - \frac{1}{n}\mathrm{Tr}\mathbf{M}_n\right|^4 + \frac{1}{n^2}\mathbb{E}|\mathbf{a}_n^*\mathbf{M}_n\mathbf{x}_n|^4 + \frac{1}{n^2}\mathbb{E}|\mathbf{x}_n^*\mathbf{M}_n\mathbf{a}_n|^4,
\end{aligned}$$

by [11, Eq. (3.2)], we have

$$\mathbb{E}\left|\frac{1}{n}\mathbf{x}_n^*\mathbf{M}_n\mathbf{x}_n - \frac{1}{n}\mathrm{Tr}\mathbf{M}_n\right|^4 \leq \frac{K\eta_n^4}{n}.$$

The rest can be treated as following :

$$\begin{aligned}
&\frac{1}{n^2}\mathbb{E}|\mathbf{x}_n^*\mathbf{M}_n\mathbf{a}_n|^4 \\
&= \frac{1}{n^2}\mathbb{E}|\mathbf{x}_n^*\mathbf{M}_n\mathbf{a}_n\mathbf{a}_n^*\mathbf{M}_n^*\mathbf{x}_n|^2 \\
&\leq 2\left(\mathbb{E}\left|\frac{1}{n}\mathbf{x}_n^*\mathbf{M}_n\mathbf{a}_n\mathbf{a}_n^*\mathbf{M}_n\mathbf{x}_n - \frac{1}{n}\mathrm{Tr}\mathbf{M}_n\mathbf{a}_n\mathbf{a}_n^*\mathbf{M}_n\right|^2 + \mathbb{E}\left|\frac{1}{n}\mathrm{Tr}\mathbf{M}_n\mathbf{a}_n\mathbf{a}_n^*\mathbf{M}_n\right|^2\right).
\end{aligned}$$

Thanks to Lemma 2.3.2, the first term is with the order $\mathcal{O}(n^{-2})$ and the second term is directly of the right order while $\frac{1}{n^2}\mathbb{E}|\mathbf{a}_n^*\mathbf{M}_n\mathbf{x}_n|^4$ can be treated similarly. Then we have

$$\mathbb{E}\left|\frac{1}{n}\mathbf{x}_n^*\mathbf{M}_n\mathbf{x}_n - \frac{1}{n}\mathrm{Tr}\mathbf{M}_n + \frac{1}{\sqrt{n}}\mathbf{a}_n^*\mathbf{M}_n\mathbf{x}_n + \frac{1}{\sqrt{n}}\mathbf{x}_n^*\mathbf{M}_n\mathbf{a}_n\right|^4 \leq \frac{K\eta_n^4}{n}.$$

2.6.5 Proof of Lemma 2.3.5

We first prove (2.3.29). After the truncation step (cf. Section 2.3.2), for $z \in \{x + \mathbf{i}d; x \in [\mu_\ell, \mu_r]\}$

$$\mathbb{E}\|\mathbf{Q}(z)\|^p \leq K\mathbb{E}\|\mathbf{Q}(z)\| < \infty.$$

Take $\eta_\ell \in (\mu_\ell, 0)$ and $\eta_r \in (\ell_+, \mu_r)$. For $z \in \{\mu_\ell + \mathbf{i}v; v \in [\varepsilon_n/n, d]\} \cup \{\mu_\ell + \mathbf{i}v; v \in [\varepsilon_n/n, d]\}$, we get

$$\begin{aligned}\mathbb{E}\|\mathbf{Q}_j(z)\|^p &\leq K + Kv^{-p}\mathbb{P}(\lambda_{\max}^{\mathbf{Y}_j\mathbf{Y}_j^*} \geq \eta_r \quad \text{or} \quad \lambda_{\min}^{\mathbf{Y}_j\mathbf{Y}_j^*} \leq \eta_\ell)\\ &\leq K + Kn^p\varepsilon_n^{-p}n^{-k} \leq K_p,\end{aligned}$$

for sufficiently large k. This proves

$$\mathbb{E}\|\mathbf{Q}_j(z)\|^p \leq K_p.$$

Now we will show

$$\mathbb{E}|\mathbf{y}_j^*\mathbf{Q}_j^q(z)\mathbf{y}_j|^p \leq K_{pq}. \tag{2.6.8}$$

With the same method, (2.6.8) holds true for $p = 1$. For $p \geq 2$,

$$\begin{aligned}&\mathbb{E}|\mathbf{y}_j^*\mathbf{Q}_j^q(z)\mathbf{y}_j|^p\\ &\leq \quad K + Kv^{-qp}\mathbb{E}|\mathbf{y}_j^*\mathbf{y}_j|^p\mathbb{P}(\lambda_{\max}^{\mathbf{Y}_j\mathbf{Y}_j^*} \geq \eta_r \quad \text{or} \quad \lambda_{\min}^{\mathbf{Y}_j\mathbf{Y}_j^*} \leq \eta_\ell)\\ &\leq \quad K + Kn^{qp}\varepsilon_n^{-qp}\eta_n^{2p-4}\mathbb{E}|\mathbf{y}_j^*\mathbf{y}_j|^2n^{-k} \leq K_{qp},\end{aligned}$$

for large enough k. This terminates the proof of (2.3.29).

Following the same idea, we will prove (2.3.30). From Lemma 2.3.2, (2.3.30) holds true for $z \in \{x + \mathbf{i}d; x \in [\mu_\ell, \mu_r]\}$. For $z \in \{\mu_\ell + \mathbf{i}v; v \in [\varepsilon_n/n, d]\} \cup \{\mu_\ell + \mathbf{i}v; v \in [\varepsilon_n/n, d]\}$, noticing that

$$\|\mathbf{M}_j\| = \|\mathbf{Q}_j^{k_1}(z_1)\mathbf{Q}_j^{k_2}(z_2)\| \leq v_1^{-k_1}v_2^{-k_2},$$

we have then

$$\begin{aligned}&\mathbb{E}\left|\mathbf{y}_j^*\mathbf{M}_j\mathbf{y}_j - n^{-1}\mathrm{Tr}\mathbf{M}_j - \mathbf{a}_j^*\mathbf{M}_j\mathbf{a}_j\right|^2\\ &\leq \quad K/n + Kv_1^{-k_1}v_2^{-k_2}\mathbb{P}(\lambda_{\max}^{\mathbf{YY}^*} \geq \eta_r \quad \text{or} \quad \lambda_{\min}^{\mathbf{YY}^*} \leq \eta_\ell)\\ &\leq \quad K + Kn^{k_1+k_2}\varepsilon_n^{-k_1-k_2}n^{-k} \leq K_{2k_1+2k_2}/n,\end{aligned}$$

for k sufficiently high.

(2.3.31) can be shown in a similar way.

We now show that for $z \in \mathcal{C}_n$,

$$\mathbb{E}|\gamma_j(z)|^2 \leq K/n. \tag{2.6.9}$$

From [11, Eq (3.5)],

$$\mathbb{E}|n^{-1}\mathbf{x}_j^*\mathbf{Q}_j\mathbf{x}_j - n^{-1}\mathbb{E}\mathrm{Tr}\mathbf{Q}_j|^2 \leq K/n.$$

Then

$$\begin{aligned}\mathbb{E}|\gamma_j(z)|^2 &= \mathbb{E}|n^{-1}\mathbf{x}_j^*\mathbf{Q}_j\mathbf{x}_j - n^{-1}\mathbb{E}\mathrm{Tr}\mathbf{Q}_j + n^{-1/2}\mathbf{x}_j^*\mathbf{Q}_j\mathbf{a}_j + n^{-1/2}\mathbf{a}_j^*\mathbf{Q}_j\mathbf{x}_j|^2 \\ &\leq 3\mathbb{E}|n^{-1}\mathbf{x}_j^*\mathbf{Q}_j\mathbf{x}_j - n^{-1}\mathbb{E}\mathrm{Tr}\mathbf{Q}_j|^2 + \frac{3}{n}\mathbb{E}|\mathbf{x}_j^*\mathbf{Q}_j\mathbf{a}_j|^2 + \frac{3}{n}\mathbb{E}|\mathbf{a}_j^*\mathbf{Q}_j\mathbf{x}_j|^2 \\ &\leq K/n.\end{aligned}$$

where the last inequality follows from the fact that by (2.3.29),

$$\max\left\{\frac{1}{n}\mathbb{E}|\mathbf{a}_j^*\mathbf{Q}_j\mathbf{x}_j|^2, \frac{1}{n}\mathbb{E}|\mathbf{x}_j^*\mathbf{Q}_j\mathbf{a}_j|^2\right\} \leq K/n.$$

The same way also handles

$$\mathbb{E}|\gamma_j(z)|^4 \leq K\eta_n^4/n.$$

Since $\beta_j = 1 - \mathbf{y}_j^*\mathbf{Q}\mathbf{y}_j$, same method yields that, for $p \geq 2$,

$$\begin{aligned}\mathbb{E}|\beta_j|^p &= \mathbb{E}|1 - \mathbf{y}_j\mathbf{Q}\mathbf{y}_j|^p \\ &\leq K + Kv^{-p}\mathbb{E}|\mathbf{y}_j\mathbf{y}_j^*|^p\mathbb{P}(\lambda_{\max}^{\mathbf{YY}^*} \geq \eta_r \quad \text{or} \quad \lambda_{\min}^{\mathbf{YY}^*} \leq \eta_\ell) \\ &\leq K + Kn^p\varepsilon_n^{-p}\eta_n^{2p}n^{-k} \leq K_p.\end{aligned}$$

Since $b_j = \beta_j + b_j\beta_j\gamma_j$ along with (2.6.9), we have

$$|b_j| \leq |\mathbb{E}\beta_j + \mathbb{E}b_j\beta_j\gamma_j| \leq K_1 + K_2|b_j|n^{-1/2},$$

which implies

$$|b_j(z)| \leq \frac{K_1}{1 - K_2 n^{-1/2}} < \infty.$$

This terminates the proof of (2.3.33).

2.6.6 Proof of Lemma 2.4.6

From Equation (2.4.6), we have

$$\begin{aligned}&\mathbf{a}^*\mathbb{E}_j\mathbf{Q}_1\mathbf{M}\mathbb{E}_j\mathbf{Q}_2\mathbf{a} - \mathbb{E}\mathbf{a}^*\mathbb{E}_j\mathbf{Q}_1\mathbf{M}\mathbb{E}_j\mathbf{Q}_2\mathbf{a} \\ &= \sum_{i=1}^{j}(\mathbb{E}_i - \mathbb{E}_{i-1})(\mathbf{a}^*\mathbb{E}_j\mathbf{Q}_1\mathbf{M}\mathbb{E}_j\mathbf{Q}_2\mathbf{a}) \\ &= \sum_{i=1}^{j}(\mathbb{E}_i - \mathbb{E}_{i-1})(\mathbf{a}^*\mathbb{E}_j[\mathbf{Q}_{1,i} + z_1\tilde{q}_{1,ii}\mathbf{Q}_{1,i}\mathbf{y}_i\mathbf{y}_i^*\mathbf{Q}_{1,i}]\mathbf{M} \times \\ &\qquad \mathbb{E}_j[\mathbf{Q}_{2,i} + z_2\tilde{q}_{2,ii}\mathbf{Q}_{2,j}\mathbf{y}_i\mathbf{y}_i^*\mathbf{Q}_{2,i}]\mathbf{a})\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^{j} z_2(\mathbb{E}_i - \mathbb{E}_{i-1})[\mathbf{a}^*(\mathbb{E}_j\mathbf{Q}_{1,i})\mathbf{M}(\mathbb{E}_j\tilde{q}_{2,ii}\mathbf{Q}_{2,j}\mathbf{y}_i\mathbf{y}_i^*\mathbf{Q}_{2,i})\mathbf{a} \\
&\quad + \sum_{i=1}^{j} \mathbf{a}^* z_1\tilde{q}_{1,ii}\mathbf{Q}_{1,i}\mathbf{y}_i\mathbf{y}_i^*\mathbf{Q}_{1,i}\mathbf{M}(\mathbb{E}_j\mathbf{Q}_{2,i})\mathbf{a} \\
&\quad\quad + \sum_{i=1}^{j} z_1z_2\mathbb{E}_j(\tilde{q}_{1,ii}\tilde{q}_{2,ii}\mathbf{a}^*\mathbf{Q}_{1,i}\mathbf{y}_i\mathbf{y}_i^*\mathbf{Q}_{1,i})\mathbf{M}(\mathbb{E}_j\mathbf{Q}_{2,i}\mathbf{y}_i\mathbf{y}_i^*\mathbf{Q}_{2,i}\mathbf{a})] \\
&\triangleq X_1 + X_2 + Z. \qquad (2.6.10)
\end{aligned}$$

The variance of $\mathbf{a}^*\mathbb{E}_j\mathbf{Q}_1\mathbf{M}\mathbb{E}_j\mathbf{Q}_2\mathbf{a}$ is the sum the variance of these martingale increments. Recall that $\tilde{q}_{ii} = \tilde{b}_i + z\tilde{q}_{ii}\tilde{b}_i\hat{\gamma}_i$,

$$\begin{aligned}
X_1 &= z_2\sum_{i=1}^{j}(\mathbb{E}_i - \mathbb{E}_{i-1})(\tilde{b}_{2,i}\mathbf{y}_i^*\mathbf{Q}_{2,i}\mathbf{a}\mathbf{a}^*(\mathbb{E}_j\mathbf{Q}_{1,i})\mathbf{M}\mathbf{Q}_{2,i}\mathbf{y}_i) \\
&\quad + z_2^2\sum_{i=1}^{j}(\mathbb{E}_i - \mathbb{E}_{i-1})(\tilde{b}_{2,i}\tilde{q}_{2,ii}\hat{\gamma}_{2,i}\mathbf{y}_i^*\mathbf{Q}_{2,i}\mathbf{a}\mathbf{a}^*(\mathbb{E}_j\mathbf{Q}_{1,i})\mathbf{M}\mathbf{Q}_{2,i}\mathbf{y}_i) \\
&\triangleq X_1' + X_1''.
\end{aligned}$$

Let $\mathbf{M}_i(z_1, z_2) = \mathbf{Q}_{2,i}\mathbf{a}\mathbf{a}^*(\mathbb{E}_j\mathbf{Q}_{1,i})\mathbf{M}\mathbf{Q}_{2,i}$. X_1' satisfies :

$$\mathbb{E}|X_1'|^2 = |z_2|^2\sum_{i=1}^{j}\mathbb{E}\left|\mathbb{E}_i\tilde{b}_i\left(\mathbf{r}_i^*\mathbf{M}_i\mathbf{r}_i - \frac{1}{n}\mathrm{Tr}\mathbf{M}_i + \mathbf{r}_i^*\mathbf{M}_i\mathbf{a}_i + \mathbf{a}_i\mathbf{M}_i\mathbf{r}_i\right)\right|^2.$$

Since $\mathbf{M}_i$ is a rank one matrix, by Lemma 2.4.1, $\sum_{i=1}^{j}\mathbb{E}|\mathbf{r}_i^*\mathbf{M}_i\mathbf{r}_i^* - \mathrm{Tr}\mathbf{M}_i/n|^2 \leq K/n$. Moreover,

$$\begin{aligned}
\sum_{i=1}^{j}\mathbb{E}|\mathbf{r}_i^*\mathbf{M}_i\mathbf{a}_i|^2 &= \frac{1}{n}\sum_{i=1}^{j}\mathbb{E}(\mathbf{a}^*\mathbf{Q}_{2,i}^*\mathbf{Q}_{2,i}\mathbf{a}|\mathbf{a}^*(\mathbb{E}_j\mathbf{Q}_{1,i})\mathbf{M}\mathbf{Q}_{2,i}\mathbf{a}_i|^2) \\
&\leq \frac{K}{n}\sum_{i=1}^{j}\mathbb{E}|\mathbf{a}^*(\mathbb{E}_j\mathbf{Q}_{1,i})\mathbf{M}\mathbf{Q}_{2,i}\mathbf{a}_i|^2.
\end{aligned}$$

Then

$$\begin{aligned}
&|\mathbf{a}^*(\mathbb{E}_j\mathbf{Q}_{1,i})\mathbf{M}\mathbf{Q}_{2,i}\mathbf{a}_i|^2 \\
\leq&4(|\mathbf{a}^*(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{a}_i|^2 + |\mathbf{a}^*(\mathbb{E}_j\mathbf{Q}_{1,i} - \mathbf{Q}_1)\mathbf{M}(\mathbf{Q}_{2,i} - \mathbf{Q}_2)\mathbf{a}_i|^2 \\
&+ |\mathbf{a}^*(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}(\mathbf{Q}_{2,i} - \mathbf{Q}_2)\mathbf{a}_i|^2 + |\mathbf{a}^*(\mathbb{E}_j(\mathbf{Q}_{1,i} - \mathbf{Q}_1))\mathbf{M}\mathbf{Q}_2\mathbf{a}_i|^2) \\
\triangleq&4(W_{i,1} + W_{i,2} + W_{i,3} + W_{i,4}).
\end{aligned}$$

Denote by $\mathbf{A}_{1:j} = [\mathbf{a}_1, \cdots, \mathbf{a}_j]$, we have :

$$\left|\sum_{i=1}^{j} W_{i,1}\right| = \mathbf{a}^*(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}\mathbf{Q}_2\mathbf{A}_{1:j}\mathbf{A}_{1:j}^*(\mathbb{E}_j\mathbf{Q}_1^*)\mathbf{M}\mathbf{Q}_2^*\mathbf{a} \leq K.$$

Recalling Equation (2.4.7) and (2.4.8), recalling $1/\beta_i = 1 + \mathbf{y}_i^*\mathbf{Q}_i\mathbf{y}_i$, we get

$$W_{i,2} \leq \frac{\mathbf{a}^*\mathbf{Q}_2\mathbf{y}_i\mathbf{y}_i^*\mathbf{Q}_2^*\mathbf{a}_i}{1-\mathbf{y}_i^*\mathbf{Q}_2\mathbf{y}_i} \times \mathbf{a}^*\mathbb{E}_j\left(\frac{\mathbf{Q}_1\mathbf{y}_i\mathbf{y}_i^*\mathbf{Q}_1}{\beta_{1,i}}\right)\mathbf{M}\frac{\mathbf{Q}_2\mathbf{y}_i\mathbf{y}_i^*\mathbf{Q}_2}{1-\mathbf{y}_i^*\mathbf{Q}_2\mathbf{y}_i}\mathbf{M}\mathbb{E}_j\left(\frac{\mathbf{Q}_1\mathbf{y}_i\mathbf{y}_i^*\mathbf{Q}_1}{\beta_{1,i}}\right)\mathbf{a}.$$

As $\|(1-\mathbf{y}_i^*\mathbf{Q}\mathbf{y}_i)^{-1}\mathbf{Q}\mathbf{y}_i\mathbf{y}_i^*\mathbf{Q}\| = \|\mathbf{Q}-\mathbf{Q}_i\| \leq K$ and $\|\mathbf{Q}\mathbf{Y}\| \leq K$, we have

$$\sum_{i=1}^{j}\mathbb{E}W_{i,2} \leq \sum_{i=1}^{j}\mathbb{E}\left[\frac{|\mathbf{a}^*\mathbf{Q}_1\mathbf{y}_i|^2}{|\beta_{1,i}|^2}\right].$$

Noticing that $\beta_i^{-1} = -(z\tilde{b}_i)^{-1} + \hat{\gamma}_i$, by Lemma 2.3.2 we obtain :

$$\begin{aligned}\sum_{i=1}^{j}\mathbb{E}W_{i,2} &\leq 2\mathbb{E}\mathbf{a}^*\mathbf{Q}_1\mathbf{Y}_{1:j}\mathrm{diag}(-(z_1\tilde{b}_{1,1})^{-1},\cdots,-(z_1\tilde{b}_{1,j})^{-1})\mathbf{Y}_{1:j}\mathbf{Q}_1\mathbf{a} + K\sum_{i=1}^{j}\mathbb{E}|\hat{\gamma}_1|^2 \\ &\leq K.\end{aligned}$$

The terms $W_{i,3}$ and $W_{i,4}$ can be handled similarly. Therefore we get $\sum_{i=1}^{j}\mathbb{E}|\mathbf{y}_i^*\mathbf{M}_i\mathbf{a}_i|^2 \leq K/n$ and $\sum_{i=1}^{j}\mathbb{E}|\mathbf{a}_i^*\mathbf{M}_i\mathbf{y}_i|^2 \leq K/n$ which implies that $\mathbb{E}|X_1|^2 \leq K/n$.

We now consider X_2. Since $\mathbb{E}|X_2|^2 \leq 2|z_2|^4\sum_{i=1}^{j}\mathbb{E}|\tilde{b}_{2,i}\tilde{q}_{2,ii}\hat{\gamma}_{2,i}\mathbf{y}_i^*\mathbf{M}_i\mathbf{y}_i|^2$, we have

$$\begin{aligned}\sum_{i=1}^{j}\mathbb{E}|\tilde{b}_{2,i}\tilde{q}_{2,ii}\hat{\gamma}_{2,i}\mathbf{a}_i^*\mathbf{M}_i\mathbf{a}_i|^2 &\leq K\sum_{i=1}^{j}\mathbb{E}\mathbb{E}^{(i)}|\hat{\gamma}_{2,i}\mathbf{a}_i^*\mathbf{M}_i\mathbf{a}_i|^2 \\ &\leq \frac{K}{n}\sum_{i=1}^{j}\mathbb{E}|\mathbf{a}_i^*\mathbf{M}_i\mathbf{a}_i|^2 \\ &\leq \frac{K}{n}\sum_{i=1}^{j}\mathbb{E}|\mathbf{a}_i^*\mathbf{Q}_{2,i}\mathbf{a}_i|^2 \leq \frac{K}{n},\end{aligned}$$

where $\mathbb{E}^{(i)} = \mathbb{E}[.|\mathbf{r}_1,\cdots,\mathbf{r}_{i-1},\mathbf{r}_{i+1},\cdots,\mathbf{r}_n]$. Moreover, by Lemma 2.3.3,

$$\sum_{i=1}^{j}\mathbb{E}|\tilde{b}_{2,i}\tilde{q}_{2,ii}\hat{\gamma}_{2,i}\mathbf{r}_i^*\mathbf{M}_i\mathbf{a}_i|^2 \leq K\sum_{i=1}^{j}\mathbb{E}^{1/2}|\hat{\gamma}_{2,i}|^4\mathbb{E}^{1/2}|\mathbf{r}_i^*\mathbf{M}_i\mathbf{a}_i|^4 \leq \frac{K\eta_n^2}{\sqrt{n}},$$

and similarly for the terms in $\mathbf{a}_i^*\mathbf{M}_i\mathbf{r}_i$ and $\mathbf{r}_i^*\mathbf{M}_i\mathbf{r}_i$. We get that $\mathbb{E}|X_2|^2 = o(n^{-1/2})$.

We now turn to the term Z of Equation (2.6.10). By developing $\mathbf{y}_i = \mathbf{r}_i + \mathbf{a}_i$, we have three terms to control,

$$\begin{aligned}Z_1 &= \sum_{i=1}^{j}(\mathbb{E}_i - \mathbb{E}_{i-1})(\mathbb{E}_j\tilde{q}_{1,ii}\mathbf{a}^*\mathbf{Q}_{1,i}\mathbf{r}_i\mathbf{y}_i^*\mathbf{Q}_{1,i})\mathbf{M}(\mathbb{E}_j\tilde{q}_{2,ii}\mathbf{Q}_{2,i}\mathbf{r}_i\mathbf{a}_i^*\mathbf{Q}_{2,i}\mathbf{a}) \\ Z_2 &= \sum_{i=1}^{j}(\mathbb{E}_i - \mathbb{E}_{i-1})(\mathbb{E}_j\tilde{q}_{1,ii}\mathbf{a}^*\mathbf{Q}_{1,i}\mathbf{r}_i\mathbf{y}_i^*\mathbf{Q}_{1,i})\mathbf{M}(\mathbb{E}_j\tilde{q}_{2,ii}\mathbf{Q}_{2,i}\mathbf{r}_i\mathbf{y}_i^*\mathbf{Q}_{2,i}\mathbf{a}) \\ Z_3 &= \sum_{i=1}^{j}(\mathbb{E}_i - \mathbb{E}_{i-1})(\mathbb{E}_j\tilde{q}_{1,ii}\mathbf{a}^*\mathbf{Q}_{1,i}\mathbf{r}_i\mathbf{a}_i^*\mathbf{Q}_{1,i})\mathbf{M}(\mathbb{E}_j\tilde{q}_{2,ii}\mathbf{Q}_{2,i}\mathbf{r}_i\mathbf{a}_i^*\mathbf{Q}_{2,i}\mathbf{a}).\end{aligned}$$

The first term satisfies

$$\begin{aligned}
\mathbb{E}|Z_1|^2 &\leq 2\sum_{i=1}^{j}\mathbb{E}|(\mathbb{E}_j\tilde{q}_{1,ii}\mathbf{a}^*\mathbf{Q}_{1,i}\mathbf{r}_i\mathbf{a}_i^*\mathbf{Q}_{1,i})\mathbf{M}(\mathbb{E}_j\tilde{q}_{2,ii}\mathbf{Q}_{2,i}\mathbf{r}_i\mathbf{a}_i^*\mathbf{Q}_{2,i}\mathbf{a})|^2 \\
&\leq K\mathbb{E}|(\mathbb{E}_j\mathbf{a}^*\mathbf{Q}_{1,i}\mathbf{r}_i\mathbf{a}_i^*\mathbf{Q}_{1,i})\mathbf{M}(\mathbf{Q}_{2,i}\mathbf{r}_i\mathbf{a}_i^*\mathbf{Q}_{2,i}\mathbf{a})|^2 \\
&= K\sum_{i=1}^{j}\mathbb{E}|\mathbf{a}_i^*\mathbf{Q}_{2,i}\mathbf{a}|^2\mathbb{E}^{(i)}|(\mathbb{E}_j\mathbf{a}^*\mathbf{Q}_{1,i}\mathbf{r}_i\mathbf{a}_i^*\mathbf{Q}_{1,i})\mathbf{M}\mathbf{Q}_{2,i}\mathbf{r}_i|^2 \\
&\leq \frac{K}{n}\sum_{i=1}^{j}\mathbb{E}|\mathbf{a}_i^*\mathbf{Q}_{2,i}\mathbf{a}|^2 = \mathcal{O}(n^{-1}),
\end{aligned}$$

where the second inequality comes from $\mathbb{E}|\mathbb{E}_j(X)\mathbb{E}_j(Y)|^2 = \mathbb{E}|\mathbb{E}_j(X\mathbb{E}_i(Y))|^2 \leq \mathbb{E}|X\mathbb{E}_j(Y)|^2$. The terms Z_2 and Z_3 can be handled similarly. Hence $\mathbb{E}|Z|^2 \leq \frac{K}{n}$.

Therefore, $\mathrm{Var}(\mathbf{a}^*(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}(\mathbb{E}_j\mathbf{Q}_2)a) = o(1/\sqrt{n})$. $\mathrm{Var}(\mathrm{Tr}[(\mathbb{E}_j\mathbf{Q}_1)\mathbf{M}(\mathbb{E}_j\mathbf{Q}_2)]) = o(\sqrt{n})$ can be handled in a similar way.

2.6.7 Proof of Lemma 2.4.7

From the expression (2.4.10) and (2.4.14) in Theorem 2.4.1, we have

$$\sum_{\ell=1}^{j}\frac{\alpha_\ell\mathbf{u}^*\mathbf{T}_1\mathbf{a}_\ell\mathbb{E}[\mathbf{a}_\ell^*\mathbf{Q}_2\mathbf{u}]}{1+z_2\tilde{t}_{2,\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_{2,\ell}\mathbf{a}_\ell} = \sum_{\ell=1}^{j}\frac{-\alpha_\ell\mathbf{u}^*\mathbf{T}_1\mathbf{a}_\ell\mathbf{a}_\ell^*\mathbf{T}_2\mathbf{u}}{z_2\tilde{t}_{2,\ell\ell}(1+\delta_2)} + \mathcal{O}(n^{-1/2}).$$

Moreover, by Equation (2.4.6),

$$\begin{aligned}
&\sum_{\ell=1}^{j}\alpha_\ell\mathbf{u}^*\mathbf{T}_1\mathbf{a}_\ell\left(\frac{\mathbb{E}[\mathbf{a}_\ell^*\mathbf{Q}_2\mathbf{u}]}{1+z_2\tilde{t}_{2,\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_{2,\ell}\mathbf{a}_\ell} - \mathbb{E}[\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{u}]\right) \\
&= \sum_{\ell=1}^{j}\alpha_\ell\mathbf{u}^*\mathbf{T}_1\mathbf{a}_\ell\mathbb{E}\left[\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{u}\left(\frac{1+z_2\tilde{q}_{2,\ell\ell}\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{y}_\ell}{1+z_2\tilde{t}_{2,\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_{2,\ell}\mathbf{a}_\ell} - 1\right)\right] \\
&\quad + \sum_{\ell=1}^{j}\alpha_\ell\mathbf{u}^*\mathbf{T}_1\mathbf{a}_\ell\frac{\mathbb{E}[z_2\tilde{q}_{2,\ell\ell}\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{y}_\ell\mathbf{r}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{u}]}{1+z_2\tilde{t}_{2,\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_{2,\ell}\mathbf{a}_\ell} \\
&\triangleq \epsilon_1 + \epsilon_2.
\end{aligned}$$

We have then $\epsilon_1 = \sum_{\ell=1}^{j}\alpha_\ell\mathbf{u}^*\mathbf{T}_1\mathbf{a}_\ell\mathbb{E}[\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{u}\mu_{2,\ell}]$ where $\mathbb{E}|\mu_{2,\ell}|^2 \leq Kn^{-1}$ by (2.4.18) again. It follows that

$$|\epsilon_1| \leq (\sum_{\ell=1}^{j}\alpha_\ell^2|\mathbf{u}^*\mathbf{T}_1\mathbf{a}_\ell|^2\mathbb{E}|\mu_{2,\ell}|^2)^{1/2}(\sum_{\ell=1}^{j}\mathbb{E}|\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{u}|^2)^{1/2} \leq \frac{K}{\sqrt{n}}$$

by (2.4.12) in Theorem 2.4.1. By writing

$$\epsilon_2 = \sum_{\ell=1}^{j}\alpha_\ell\mathbf{u}^*\mathbf{T}_1\mathbf{a}_\ell\frac{\mathbb{E}[z_2(\tilde{q}_{2,\ell\ell}\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{y}_\ell - \mathbb{E}[\tilde{q}_{2,\ell\ell}\mathbf{a}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{y}_\ell])\mathbf{r}_\ell^*\mathbf{Q}_{2,\ell}\mathbf{u}]}{1+z_2\tilde{t}_{2,\ell\ell}\mathbf{a}_\ell^*\mathbf{T}_{2,\ell}\mathbf{a}_\ell}$$

and proceeding similarly to ϵ_1, we obtain $|\epsilon_2| \leq K/\sqrt{n}$ which completes the proof.

2.6.8 Proof of Lemma 2.4.8

For a matrix $\mathbf{A}$, define $\tilde{\mathbf{A}}_{1:j} = [\tilde{\mathbf{a}}_1, \cdots, \tilde{\mathbf{a}}_j, 0, \cdots]$, where $\mathbf{a}_j$ it the j-th column of $\mathbf{A}$. We will show that

$$\eta_j + \omega_j = \frac{z_1 z_2}{n} \mathrm{Tr} \tilde{\mathbf{T}}_{1:j} \tilde{\mathbf{T}}_{1:j}^*. \tag{2.6.11}$$

In particular, $j = n$ implies (1).

From the definition of ω_j, we have

$$\omega_j = \frac{1}{n}(1+\delta_1)^2(1+\delta_2)^2 \mathrm{Tr} \mathbf{A}_{1:j}^* \mathbf{T}_1 \mathbf{A}_{1:j} \mathbf{A}_{1:j}^* \mathbf{T}_2 \mathbf{A}_{1:j} - \frac{1}{n} \sum_{k=1}^{j} \frac{\mathbf{a}_k^* \mathbf{T}_1 \mathbf{a}_k \mathbf{a}_k^* \mathbf{T}_2 \mathbf{a}_k}{(1+\delta_1)^2 (1+\delta_2)^2}.$$

From (2.4.9) and (2.4.5), $(1+\delta)^{-2} \mathbf{a}_k^* \mathbf{T}_k \mathbf{a}_k = (1+\delta)^{-1} + z\tilde{t}_{kk}$. Hence

$$\frac{1}{n} \sum_{k=1}^{j} \frac{\mathbf{a}_k^* \mathbf{T}_1 \mathbf{a}_k \mathbf{a}_k^* \mathbf{T}_2 \mathbf{a}_k}{(1+\delta_1)^2 (1+\delta_2)^2} = \frac{1}{n} \sum_{k=1}^{j} z_1 z_2 \tilde{t}_{1,kk} \tilde{t}_{2,kk} + \frac{1}{n} \sum_{k=1}^{j} \frac{z_1 \tilde{t}_{1,kk}}{1+\delta_2} + \frac{1}{n} \sum_{k=1}^{j} \frac{\mathbf{a}_k^* \mathbf{T}_2 \mathbf{a}_k}{(1+\delta_1)(1+\delta_2)^2}.$$

By $-z\tilde{\mathbf{T}}_{1:j} = (1+\delta)^{-1} \mathbf{I}_{1:j} - (1+\delta)^{-2} \mathbf{A}_{1:j}^* \mathbf{T} \mathbf{A}_{1:j}$, we have

$$\begin{aligned}
\eta_j + \omega_j &= -\frac{1}{n(1+\delta_2)} \mathrm{Tr} \tilde{\mathbf{T}}_{1:j}(z_1) - \frac{1}{n(1+\delta_1)(1+\delta_2)^2} \mathrm{Tr} \mathbf{A}_{1:j}^* \mathbf{T}_2 \mathbf{A}_{1:j} \\
&\quad + \frac{1}{n(1+\delta_1)^2 (1+\delta_2)^2} \mathrm{Tr} \mathbf{A}_{1:j}^* \mathbf{T}_1 \mathbf{A}_{1:j} \mathbf{A}_{1:j}^* \mathbf{T}_2 \mathbf{A}_{1:j} \\
&= \frac{z_1 z_2}{n} \mathrm{Tr}[\tilde{\mathbf{T}}_{1:j}(z_1) \tilde{\mathbf{T}}_{1:j}(z_2)].
\end{aligned}$$

Now we will prove (2). From (3.1.1),

$$\frac{\mathbf{T}(z)}{1+\delta} = (-z(1+\tilde{\delta}(z))(1+\delta(z))\mathbf{I}_N + \mathbf{A}\mathbf{A}^*)^{-1} \tag{2.6.12}$$

$$\frac{\tilde{\mathbf{T}}(z)}{1+\tilde{\delta}} = (-z(1+\delta(z))(1+\tilde{\delta}(z))\mathbf{I}_n + \mathbf{A}^*\mathbf{A})^{-1}. \tag{2.6.13}$$

By (2.6.12)- (2.6.13) and taking trace, we have

$$\delta - \tilde{\delta} = \frac{n-N}{n} \frac{1}{z}. \tag{2.6.14}$$

Using the resolvent formula, we get

$$\frac{\mathbf{T}_1}{1+\delta_1} - \frac{\mathbf{T}_2}{1+\delta_2} = \frac{z_1(1+\tilde{\delta}_1)(1+\delta_1) - z_2(1+\tilde{\delta}_2)(1+\delta_2)}{(1+\delta_1)(1+\delta_2)} \mathbf{T}_1 \mathbf{T}_2. \tag{2.6.15}$$

Taking trace on both sides of the precedent equality, we have

$$\begin{aligned}
\gamma &= \frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{T}_2 \\
&= \frac{\delta_1(1+\delta_2)-\delta_2(1+\delta_1)}{z_1(1+\tilde{\delta}_1)(1+\delta_1)-z_2(1+\tilde{\delta}_2)(1+\delta_2)} \\
&= \frac{\delta_1-\delta_2}{z_1(1+\tilde{\delta}_1)(1+\delta_1)-z_2(1+\tilde{\delta}_2)(1+\delta_2)}.
\end{aligned}$$

In the same way, we get that

$$\tilde{\gamma} = \frac{\tilde{\delta}_1-\tilde{\delta}_2}{z_1(1+\tilde{\delta}_1)(1+\delta_1)-z_2(1+\tilde{\delta}_2)(1+\delta_2)}.$$

We will express $\frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{A}\mathbf{A}^*\mathbf{T}_2$ with the help of δ_1, δ_2. We obtain,

$$\begin{aligned}
\frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{A}\mathbf{A}^*\mathbf{T}_2 &= \frac{1}{n}\mathrm{Tr}\mathbf{T}_1(\mathbf{A}\mathbf{A}^* - z_1(1+\tilde{\delta}_1)(1+\delta_1)+z_1(1+\tilde{\delta}_1)(1+\delta_1))\mathbf{T}_2 \\
&= (1+\delta_1)\delta_2 + \frac{z_1(1+\tilde{\delta}_1)(1+\delta_1)}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{T}_2 \\
&= (1+\delta_1)(1+\delta_2)\frac{z_1(1+\tilde{\delta}_1)\delta_1 - z_2(1+\tilde{\delta}_2)\delta_2}{z_1(1+\tilde{\delta}_1)(1+\delta_1)-z_2(1+\tilde{\delta}_2)(1+\delta_2)}.
\end{aligned}$$

Then

$$1-\frac{1}{n}\frac{1}{1+\delta_1}\frac{1}{1+\delta_2}\mathrm{Tr}\mathbf{T}_1\mathbf{A}\mathbf{A}^*\mathbf{T}_2 = \frac{z_1(1+\tilde{\delta}_1)-z_2(1+\tilde{\delta}_2)}{z_1(1+\tilde{\delta}_1)(1+\delta_1)-z_2(1+\tilde{\delta}_2)(1+\delta_2)}. \tag{2.6.16}$$

Replace the expressions of γ, $\tilde{\gamma}$ and $\frac{1}{n}\mathrm{Tr}\mathbf{T}_1\mathbf{A}\mathbf{A}^*\mathbf{T}_2$ in Δ_n, we get :

$$\Delta_n = \frac{(z_1(1+\tilde{\delta}_1)-z_2(1+\tilde{\delta}_2))^2 - z_1z_2(\delta_1-\delta_2)(\tilde{\delta}_1-\tilde{\delta}_2))}{(z_1(1+\tilde{\delta}_1)(1+\delta_1)-z_2(1+\tilde{\delta}_2)(1+\delta_2))^2},$$

and

$$\begin{aligned}
&(z_1(1+\tilde{\delta}_1)-z_2(1+\tilde{\delta}_2))^2 - z_1z_2(\delta_1-\delta_2)(\tilde{\delta}_1-\tilde{\delta}_2) \\
&= \quad z_1^2(1+\tilde{\delta}_1)^2 + z_2^2(1+\tilde{\delta}_2)^2 - 2z_1z_2(1+\tilde{\delta}_1)(1+\tilde{\delta}_2) \\
&\qquad -z_1z_2(\tilde{\delta}_1 + \frac{n-N}{nz_1} - \tilde{\delta}_2 - \frac{n-N}{nz_2})(\tilde{\delta}_1-\tilde{\delta}_2) \\
&= \quad z_1^2(1+\tilde{\delta}_1)^2 + z_2^2(1+\tilde{\delta}_2)^2 - z_1z_2 - z_1z_2((\tilde{\delta}_1+1)^2+(\tilde{\delta}_2+1)^2) \\
&\qquad +\frac{n-N}{n}(z_1-z_2)(\tilde{\delta}_1-\tilde{\delta}_2) \\
&= \quad (z_1-z_2)\left(z_1(1+\tilde{\delta}_1)^2 - z_2(1+\tilde{\delta}_2)^2 + \frac{n-N}{n}(\tilde{\delta}_1-\tilde{\delta}_2)\right) \\
&= \quad (s_n(z_1)-s_n(z_2))(z_1-z_2)
\end{aligned}$$

where the last equality follows from Equation (2.6.14). On the other hand,

$$\begin{aligned}&(z_1(1+\tilde{\delta}_1)(1+\delta_1) - z_2(1+\tilde{\delta}_2)(1+\delta_2))^2 \\ =\ & (z_1(1+\tilde{\delta}_1)(1+\tilde{\delta}_1+\frac{n-N}{nz_1}) - z_2(1+\tilde{\delta}_2)(1+\delta_2))^2 \\ =\ & (s_n(z_1) - s_n(z_2))^2.\end{aligned}$$

Thus,

$$\frac{1}{\Delta_n(z_1,z_2)} = \frac{s_n(z_1)-s_n(z_2)}{z_1-z_2}.$$

The last task is to show that for $z_1, z_2 \in \mathbb{C}^+$,

$$\liminf_n \inf_{j\in\{1,\cdots,n\}} |\mathbf{\Delta}_j(z_1,z_2)| > 0.$$

We first prove that for all $j \in \{1,\cdots,n\}$ and $z \in \mathbb{C}^+$,

$$\begin{aligned}\operatorname{Im}[z(1+\delta)(1+\tilde{\delta})] &> 0, &(2.6.17)\\ \nu_j(z,\bar{z}) &< 1, &(2.6.18)\\ \liminf_n \inf_{j\in\{1,\cdots,n\}} \mathbf{\Delta}_j(z,\bar{z}) &> 0. &(2.6.19)\end{aligned}$$

From (2.1.4),

$$\frac{\delta}{1+\delta} = \frac{1}{n}\sum_{j=1}^{n} \frac{1}{a_i - z(1+\delta)(1+\tilde{\delta})}$$

where $a_i \in \mathbb{R}^+$ are eigenvalues of $\mathbf{AA}^*$.

As $\operatorname{Im}(\delta) > 0$, we get

$$\operatorname{Im}\frac{\delta}{1+\delta} = \operatorname{Im}\left(1 - \frac{1}{1+\delta}\right) > 0,$$

which leads to

$$\operatorname{Im}[z(1+\delta)(1+\tilde{\delta})] > 0. \tag{2.6.20}$$

Recall the definition of ν_j in (2.4.51) and by (2.6.16),

$$1-\nu_n(z,\bar{z}) = \frac{z(1+\tilde{\delta}(z)) - \bar{z}(1+\tilde{\delta}(\bar{z}))}{z(1+\tilde{\delta}(z))(1+\delta(z)) - \bar{z}(1+\tilde{\delta}(\bar{z}))(1+\delta(\bar{z}))} > 0,$$

hence (2.6.18).

From (2) in Lemma 2.4.8,

$$0 < \frac{1}{\mathbf{\Delta}_n(z,\bar{z})} = \frac{s_n(z)-s_n(\bar{z})}{z-\bar{z}} \le K.$$

This proves

$$\liminf_n \mathbf{\Delta}_n(z,\bar{z}) > 0. \tag{2.6.21}$$

As $\nu_j(z,\bar{z})$ is increasing to $\nu_n(z,\bar{z})$, we get

$$\nu_j(z,\bar{z}) < 1.$$

From the definition of $\mathbf{\Delta}_n$ and (1) in Lemma 2.4.8,

$$\mathbf{\Delta}_n(z,\bar{z}) = (1-\nu_n(z,\bar{z}))^2 - \gamma(z,\bar{z})(\eta_n(z,\bar{z})+\omega_n(z,\bar{z})).$$

$\nu_j(z,\bar{z})$ (resp. $\eta_j(z,\bar{z})$, $\omega_j(z,\bar{z})$) is increasing to ν_n (resp. $\eta_n(z,\bar{z})$, $\omega_n(z,\bar{z})$), we have

$$\liminf_n \inf_{j\in\{1,\cdots,n\}} \mathbf{\Delta}_j(z,\bar{z}) \geq \liminf_n \mathbf{\Delta}_n(z,\bar{z}) > 0.$$

Hence (2.6.19).

With these tools, we have

$$|\nu_j(z_1,z_2)| \leq \sqrt{\nu_j(z_1,\bar{z}_1)}\sqrt{\nu_j(z_2,\bar{z}_2)} < 1. \tag{2.6.22}$$

We end up the proof by showing that for all $z_1, z_2 \in \mathbb{C}^+$,

$$|\mathbf{\Delta}_j(z_1,z_2)| \geq (\mathbf{\Delta}_j(z_1,\bar{z}_1)\mathbf{\Delta}_j(z_2,\bar{z}_2))^{1/2}. \tag{2.6.23}$$

In particular, this directly implies

$$\liminf_n \inf_j |\mathbf{\Delta}_j(z_1,z_2)| \geq \left(\liminf_n \inf_j \mathbf{\Delta}_j(z_1,\bar{z}_1) \liminf_n \inf_j \mathbf{\Delta}_j(z_2,\bar{z}_2)\right)^{1/2} > 0.$$

From (2.6.11), we get

$$\begin{aligned}|\mathbf{\Delta}_j(z_1,z_2)| &\geq (1-|\nu_j(z_1,z_2)|)^2 - \frac{|z_1z_2|}{n}|\gamma||\mathrm{Tr}\tilde{\mathbf{T}}_{1:j}(z_1)\tilde{\mathbf{T}}^*_{1:j}(z_2)| \\ &\geq (1-\nu_j^{1/2}(z_1,\bar{z}_1)\nu_j^{1/2}(z_2,\bar{z}_2))^2 - \frac{|z_1z_2|}{n}\gamma^{1/2}(z_1,\bar{z}_1)\gamma^{1/2}(z_2,\bar{z}_2)\times \\ &\quad (\mathrm{Tr}\tilde{\mathbf{T}}_{1:j}(z_1)\tilde{\mathbf{T}}^*_{1:j}(z_1))^{1/2}(\mathrm{Tr}\tilde{\mathbf{T}}_{1:j}(z_2)\tilde{\mathbf{T}}^*_{1:j}(z_2))^{1/2}.\end{aligned} \tag{2.6.24}$$

By a direct calculation, we have : for $0 \leq a, b \leq 1$,

$$(1-\sqrt{ab})^2 \geq (1-a)(1-b). \tag{2.6.25}$$

We first show that the right side of (2.6.24) is positive. As $(1-\nu_j(z,\bar{z}))^2 \geq \frac{|z|^2}{n}\gamma_j(z,\bar{z})\mathrm{Tr}\mathbf{T}_{1:j}(z)\mathbf{T}^*_{1:j}(z)$, we have

$$\begin{aligned}&\frac{|z_1z_2|}{n}\gamma^{1/2}(z_1,\bar{z}_1)\gamma^{1/2}(z_2,\bar{z}_2)(\mathrm{Tr}\tilde{\mathbf{T}}_{1:j}(z_1)\tilde{\mathbf{T}}^*_{1:j}(z_1))^{1/2}(\mathrm{Tr}\tilde{\mathbf{T}}_{1:j}(z_2)\tilde{\mathbf{T}}^*_{1:j}(z_2))^{1/2} \\ &\leq (1-\nu_j(z_1,\bar{z}_1))(1-\nu_j(z_2,\bar{z}_2)) \leq (1-\nu_j^{1/2}(z_1,\bar{z}_1)\nu_j^{1/2}(z_2,\bar{z}_2))^2.\end{aligned}$$

Then we will show that the right hand side of (2.6.24) is greater than $(\mathbf{\Delta}_j(z_1,\bar{z}_1)\mathbf{\Delta}_j(z_1,\bar{z}_1))^{1/2}$.

Denote for $i = 1, 2$,

$$a_i = \nu_j(z_i, \bar{z}_i), \quad b_i = \frac{|z_i|^2}{n}\gamma(z_i, \bar{z}_i)(\mathrm{Tr}\tilde{\mathbf{T}}_{1:j}(z_i)\tilde{\mathbf{T}}_{1:j}^*(z_i)).$$

We will show that

$$[(1 - \sqrt{a_1}\sqrt{a_2})^2 - b_1^{1/2}b_2^{1/2}]^2 \geq ((1 - a_1)^2 - b_1)((1 - a_2)^2 - b_2).$$

It suffices to show that

$$(1 - \sqrt{a_1}\sqrt{a_2})^2((1 - \sqrt{a_1}\sqrt{a_2})^2 - 2b_1^{1/2}b_2^{1/2}) \geq (1 - a_1)^2(1 - a_2)^2 - b_1(1 - a_2)^2 - b_2(1 - a_1)^2.$$

By $(1 - \sqrt{a_1}\sqrt{a_2})^2 \geq (1 - a_1)(1 - a_2)$, it suffices to show that

$$(1 - \sqrt{a_1}\sqrt{a_2})^2 - 2b_1^{1/2}b_2^{1/2} \geq (1 - a_1)(1 - a_2) - b_1\frac{1 - a_2}{1 - a_1} - b_2\frac{1 - a_1}{1 - a_2},$$

which is immediate.

2.6.9 Proof of Lemma 2.4.10

Following the same proof as Lemma 2.4.8, noticing that $|\vartheta| \leq 1$, we have

$$\begin{aligned}
|\boldsymbol{\Delta}_j^{\vartheta}(z_1, z_2)| &\geq (1 - |\vartheta||\nu_j^{\dagger}|)(1 - |\bar{\vartheta}||\tilde{\nu}_j^{\dagger}|) - \frac{|\vartheta^2 z_1 z_2|}{n}\gamma^{\dagger}|\mathrm{Tr}\tilde{\mathbf{T}}_{1:j}(z_1)\tilde{\mathbf{T}}_{1:j}^T(z_2)| \\
&\geq (1 - \nu_j^{1/2}(z_1, \bar{z}_1)\nu_j^{1/2}(z_2, \bar{z}_2))^2 - \frac{|z_1 z_2|}{n}\gamma^{1/2}(z_1, \bar{z}_1)\gamma^{1/2}(z_2, \bar{z}_2)\times \\
&\quad (\mathrm{Tr}\tilde{\mathbf{T}}_{1:j}(z_1)\tilde{\mathbf{T}}_{1:j}^*(z_1))^{1/2}(\mathrm{Tr}\tilde{\mathbf{T}}_{1:j}(z_2)\tilde{\mathbf{T}}_{1:j}^*(z_2))^{1/2}.
\end{aligned}$$

This is the right hand side of (2.6.24). Thus,

$$\liminf_n \inf_j |\Delta_j^{\vartheta}(z_1, z_2)| > 0.$$

Chapitre 3

Fluctuations for non-analytic functionals of linear spectral statistics

The content below is inspired from [72]. Consider an $N \times n$ matrix $\mathbf{Y}_n$ with real random entries. Of particular interest is the study of the fluctuations of linear spectral statistics of large random covariance matrices $\mathbf{Y}_n \mathbf{Y}_n^T$:

$$\operatorname{Tr} f(\mathbf{Y}_n \mathbf{Y}_n^T) = \sum_{i=1}^{N} f(\lambda_i), \quad (\lambda_i) \text{ eigenvalues of } \mathbf{Y}_n \mathbf{Y}_n^T,$$

as the dimensions of matrix $\mathbf{Y}_n$ go to infinity at the same pace :

$$N, n \to \infty \quad \text{and} \quad 0 < \liminf \frac{N}{n} \leq \limsup \frac{N}{n} < \infty , \tag{3.0.1}$$

(a condition that will be simply referred as $N, n \to \infty$ in the sequel). This subject has a rich history (see for instance [2, 11, 60, 70] and the references therein) among which we single out the contributions [11, 73, 70, 71].

3.1 General principles and main results

The main idea is in Introduction 1.2. If φ is a random variable, denote by $\overset{\circ}{\varphi} = \varphi - \mathbb{E}\varphi$.

3.1.1 Two lemmas

Lemma 3.1.1 *Let $(\varphi_n(t), t \in \mathbb{R})_{n \in \mathbf{N}}$ and $(\psi(t), t \in \mathbf{R})$ be complex-valued continuous random processes. Assume that ψ is centered and that :*

(i) The following convergence in distribution holds true :

$$\forall t_i \in \mathbb{R},\ 1 \leq i \leq d, \quad \left(\overset{\circ}{\varphi}_n (t_1), \cdots, \overset{\circ}{\varphi}_n (t_d)\right) \xrightarrow[n \to \infty]{\mathcal{D}} (\psi(t_1), \cdots, \psi(t_d)).$$

(ii) For all $T > 0$, $\varphi_n(t)$ is tight on $[-T, T]$.

(iii) The following estimates hold true

$$(\mathrm{Var}\varphi_n(t))^{1/2} \le K(t) \quad \textit{and} \quad (\mathrm{Var}\psi(t))^{1/2} \le K(t) \ ,$$

where $t \mapsto K(t)$ does not depend on n and is such that $\int_{\mathbb{R}} |K(t)|\, dt < \infty$.

Then,

$$\int_{\mathbb{R}} \overset{\circ}{\varphi}_n (t)\, dt \xrightarrow[n\to\infty]{\mathcal{D}} \int_{\mathbb{R}} \psi(t)\, dt$$

Lemma 3.1.1 should be compared with [50, Th. 4.28] and the proof is in Appendix 3.5.1. In the context of large random matrices, an interesting phenomenon occurs [70, 71], where sometimes there is not a limiting process $\psi(t)$ but where there exists instead a sequence of processes (ψ_n) whose distribution gets closer to the distribution of the random process of interest $(\overset{\circ}{\varphi}_n)$. The following lemma formalizes this.

Lemma 3.1.2 *Recall the definition of Lévy-Prohorov distance in Section 2.2 and let $(\varphi_n(t), t \in \mathbb{R})_{n\in\mathbf{N}}$ and $(\psi_n(t), t \in \mathbb{R})_{n\in\mathbf{N}}$ be complex-valued continuous random processes. Assume that all the ψ_n's are centered and that :*

(i) The following convergence in distribution holds true :

$$\forall t_i \in \mathbb{R},\ 1 \le i \le d, \quad d_{\mathcal{L}P}\left((\overset{\circ}{\varphi}_n (t_1), \cdots, \overset{\circ}{\varphi}_n (t_d)), (\psi_n(t_1), \cdots, \psi_n(t_d))\right) \xrightarrow[n\to\infty]{} 0$$

(ii) For all $T > 0$, $\varphi_n(t)$ and $\psi_n(t)$ are tight on $[-T, T]$.

(iii) The following estimates hold true

$$\forall n \in \mathbf{N}\ , \quad (\mathrm{Var}\varphi_n(t))^{1/2} \le K(t) \quad \textit{and} \quad (\mathrm{Var}\psi_n(t))^{1/2} \le K(t)\ ,$$

where $t \mapsto K(t)$ does not depend on n and is such that $\int_{\mathbb{R}} |K(t)|\, dt < \infty$.

Then,

$$d_{\mathcal{L}P}\left(\int_{\mathbb{R}} \overset{\circ}{\varphi}_n (t)\, dt\ , \int_{\mathbb{R}} \psi_n(t)\, dt\right) \xrightarrow[n\to\infty]{} 0.$$

The proof is the same as Lemma 3.1.1 and is omitted here.

3.1.2 Fluctuations of the linear statistics

We consider the model

$$\mathbf{Y}_n = \frac{1}{\sqrt{n}} \mathbf{R}_n^{1/2} \mathbf{X}_n.$$

For the reason of readability, we often drop subscripts n if there is no confusion.

Assumption 3.1.1 *The random variables $(x_{ij})_{1\le i\le N, 1\le j\le n}$ are real, independent and identically distributed (i.i.d.). They satisfy*

$$\mathbb{E}x_{11} = 0, \qquad \mathbb{E}|x_{11}|^2 = 1 \quad \text{and} \quad \mu_4 = \mathbb{E}|x_{11}|^4 < \infty.$$

Assumption 3.1.2 *When $N, n \to \infty$,*

$$0 < \liminf \frac{N}{n} \leq \limsup \frac{N}{n} < \infty,$$

which will be simply denoted by $N, n \to \infty$ in the sequel.

Assumption 3.1.3 *The family of $N \times N$ symmetric, non-negative definite matrix $(\mathbf{R}_n)$ is bounded for the spectral norm :*

$$r_{\max} = \sup_{n \geq 1} \|\mathbf{R}_n\| < \infty.$$

Denote

$$\mathbf{U}_n(t) = e^{\mathbf{i}t\mathbf{Y}_n\mathbf{Y}_n^T}, \quad u(t) = \mathrm{Tr}\mathbf{U}_n(t), \quad \mathring{u}_n(t) = u_n(t) - \mathbb{E}u_n(t), \quad V_n(t) = \mathrm{Var}[u(t)].$$

The purpose of the chapter is, under the precedent conditions to study the fluctuations of the linear statistics of the eigenvalues of $\frac{1}{n}\mathbf{R}^{1/2}\mathbf{X}\mathbf{X}^T\mathbf{R}^{1/2}$ for non-analytic functionals.

For a sufficiently regular function f (for example : of class C^4), We will study the fluctuations of linear spectral analysis

$$\sum_{j=1}^{n} f(\lambda_j) = \mathrm{Tr} f(\mathbf{Y}_n\mathbf{Y}_n^T),$$

where λ_j are eigenvalues of $\mathbf{Y}_n\mathbf{Y}_n^T$. This issue was initially studied by Jonsson [49] for the moments of the eigenvalues of sample covariance matrix. The linear spectral analysis is first discovered by Bai and Silverstein in [11]. In [39], Hachem *et al.* establishes a CLT for mutual information $\log\det(\mathbf{Y}_n\mathbf{Y}_n^T + \rho\mathbf{I}_N)$. Recently, Najim [70] made a fully study of the model for analytic functionals and Najim and Yao [71] generalized the result for information-plus-noise model. However, all studies are based on the assumption that f is an analytic function on a certain open region.

Some more notations

The *Stieljes transform* associated to a random matrix $\mathbf{Y}_n\mathbf{Y}_n^T$ (resp. $\mathbf{Y}_n^T\mathbf{Y}_n$) is defined by $m_n(z) = \frac{1}{N}\mathrm{Tr}\mathbf{Q}_n(z)$ (resp. $\tilde{m}_n(z) = \frac{1}{n}\mathrm{Tr}\tilde{\mathbf{Q}}_n(z)$) where $\mathbf{Q}_n(z) = (\mathbf{Y}_n\mathbf{Y}_n^T - z\mathbf{I}_n)^{-1}$ (resp. $\tilde{\mathbf{Q}}_n(z) = (\mathbf{Y}_n^T\mathbf{Y}_n - z\mathbf{I}_n)^{-1}$) is the resolvent associated to $\mathbf{Y}_n\mathbf{Y}_n^T$ (resp.$\mathbf{Y}_n^T\mathbf{Y}_n$). We define the matrix $\mathbf{T}_n(z)$ by

$$\mathbf{T}_n(z) = (-z(\mathbf{I}_N + \tilde{t}(z)\mathbf{R}_n))^{-1}. \tag{3.1.1}$$

For $z \in \mathbb{C}\backslash\mathbb{R}^+$, the system

$$\begin{cases} t_n(z) = \frac{1}{N}\mathrm{Tr}(-z(\mathbf{I}_N + \tilde{t}(z)\mathbf{R}_n))^{-1}, \\ \tilde{t}_n(z) = -\frac{1}{z(1+\frac{1}{n}\mathrm{Tr}\mathbf{R}_n\mathbf{T}_n)} \end{cases} \tag{3.1.2}$$

admits a unique solution $(t_n, \tilde{t}_n)$ from $\mathbb{C}^+$ to $\mathbb{C}^+$. $\mathbf{T}_n$ approximate the resolvents $\mathbf{Q}_n(z)$ in the sense that for $z \in \mathbb{C}\backslash\mathbb{R}^+$,

$$\frac{1}{n}\mathrm{Tr}(\mathbf{Q}_n(z) - \mathbf{T}_n(z)) \xrightarrow[n\to\infty]{a.s.} 0.$$

It has also been proved in [11] that t_n (resp. $\tilde{t}_n$) is the Stieljes transform of a probability distribution which is denoted by F_n (resp. $\tilde{F}_n$).

By Silverstein and Choi [84], the limit of $\tilde{t}_n$ exists as $z \in \mathbb{C}^+$ approaches the real axis :

$$\lim_{z\in\mathbb{C}^+\to x} \tilde{t}_n(z) = \tilde{t}_n(x), \quad \text{for } x \in \mathbb{R}^*.$$

In the sequel, we denote $\mathcal{S}_n$ the support associated with the measure $\tilde{F}_n$ of deterministic equivalent $\tilde{t}_n$.

Denote $G_n = N(F^{\mathbf{Y}_n\mathbf{Y}_n^T} - F_n)$. The central object is the study of the random variable

$$\mathcal{N}(f) \triangleq \int_{\mathbb{R}} f(x)G(dx). \tag{3.1.3}$$

In particular, we will cut the random variable into random part $\mathcal{N}^1(f)$ and deterministic part $\mathcal{N}^2(f)$:

$$\mathcal{N}(f) \triangleq \mathcal{N}^1(f) + \mathcal{N}^2(f)$$

where

$$\begin{aligned}
\mathcal{N}^1(f) &= N\left(\int_{\mathbb{R}} f(x)F^{\mathbf{Y}_n\mathbf{Y}_n^T}(dx) - \mathbb{E}\int_{\mathbb{R}} f(x)F^{\mathbf{Y}_n\mathbf{Y}_n^T}(dx)\right) &\quad (3.1.4)\\
\mathcal{N}^2(f) &= N\left(\mathbb{E}\int_{\mathbb{R}} f(x)F^{\mathbf{Y}_n\mathbf{Y}_n^T}(dx) - \int_{\mathbb{R}} f(x)F_n(dx)\right). &\quad (3.1.5)
\end{aligned}$$

For $s_1, s_2 \in \mathbb{R}$, we denote

$$\begin{aligned}
\mathcal{K}_n(s_1,s_2) &= \frac{s_1 s_2 \tilde{t}_n(s_1)\tilde{t}_n(s_2)}{n}\sum_{i=1}^N [R_n^{1/2}T_n(s_1)R_n^{1/2}]_{ii}[R_n^{1/2}T_n(s_2)R_n^{1/2}]_{ii}, &\quad (3.1.6)\\
\mathcal{K}_n^\dagger(s_1,s_2) &= \frac{s_1 s_2 \overline{\tilde{t}_n(s_1)}\tilde{t}_n(s_2)}{n}\sum_{i=1}^N [R_n^{1/2}\overline{T_n(s_1)}R_n^{1/2}]_{ii}[R_n^{1/2}T_n(s_2)R_n^{1/2}]_{ii}. &\quad (3.1.7)
\end{aligned}$$

Notice that if $s_1 \notin \mathcal{S}_n$ or $s_2 \notin \mathcal{S}_n$, $\mathcal{K}_n(s_1,s_2) = \mathcal{K}_n^\dagger(s_1,s_2) = 0$. We will show the following theorem.

Theorem 3.1.1 *Under Assumptions 3.1.1, 3.1.2 3.1.3, let $k \in \mathbb{N}$ and f be an integrable function over $\mathbb{R}$ satisfying*

$$\int_{\mathbb{R}} |\hat{f}(t)|(1+|t|^4)dt < \infty. \tag{3.1.8}$$

Recall the definition of $\tilde{t}_n$ in (3.1.2) with the support $\mathcal{S}_n$ and the random vector $\mathcal{N}_n^1(\mathbf{f})$ in (3.1.4) *and define the Gaussian vector*

$$Z_n^1(\mathbf{f}) = (Z_n^1(f_1), \cdots, Z_n^1(f_k))$$

with zero mean and the covariance

$$\begin{aligned}\mathrm{Cov}[Z_n^1(f_i), Z_n^1(f_j)] = &\frac{1}{2\pi^2}\int_{\mathcal{S}_n}\int_{\mathcal{S}_n} f_i'(s_1)f_j'(s_2)\log\left(1+\frac{4\mathrm{Im}[\tilde{t}_n(s_1)]\mathrm{Im}[\tilde{t}_n(s_2)]}{|\tilde{t}_n(s_1)-\tilde{t}_n(s_2)|^2}\right)ds_1ds_2 \\ &-\frac{\kappa_4}{2\pi^2}\int_{\mathcal{S}_n}\int_{\mathcal{S}_n} f_i'(s_1)f_j'(s_2)\mathrm{Re}(\mathcal{K}_n(s_1,s_2)-\mathcal{K}_n^{\dagger}(s_1,s_2))ds_1ds_2,\end{aligned} \tag{3.1.9}$$

where $\kappa_4 = \mu_4 - 3$ is the fourth cumulant of X_{ij}. Then

$$d_{\mathcal{LP}}(\mathcal{N}_n^1(\mathbf{f}), Z_n^1(\mathbf{f})) \xrightarrow[N,n\to\infty]{} 0.$$

The expression of the covariance coincides with [11, Eq. (1.20)] and [8, Eq. (1.5)].

Remark 3.1.1 *Loosely speaking, condition (3.1.8) implies that de derivatives $f^{(k)}$ is bounded over $\mathbb{R}$ for $k = 0,1,2,3,4$.*

3.1.3 Expression of the bias

In this section, we will express the bias $\mathcal{N}^2(f)$ defined in (3.1.5). Define

$$\mathcal{B}_{1,n}(z) \triangleq -\frac{z^2\tilde{t}_n^3\frac{1}{n}\mathrm{Tr}\mathbf{R}_n^2\mathbf{T}_n^3}{(1-z^2\tilde{t}_n^2\frac{1}{n}\mathrm{Tr}\mathbf{R}_n^2\mathbf{T}_n^2)^2} \tag{3.1.10}$$

$$\mathcal{B}_{2,n}(z) \triangleq -\frac{z^3\tilde{t}_n^3\frac{1}{n}\sum_{i=1}^N[R_n^{1/2}T_nR_n^{1/2}]_{ii}[R_n^{1/2}T_n^2R_n^{1/2}]_{ii}}{1-z^2\tilde{t}_n^2\frac{1}{n}\mathrm{Tr}\mathbf{R}_n^2\mathbf{T}_n^2}, \tag{3.1.11}$$

Theorem 3.1.2 *Let f be a function of class C^{18}, integrable on $\mathbb{R}$ satisfying*

$$\int (1+|t|^4)|\hat{f}(t)|dt < \infty.$$

Recall the definition of $\mathcal{N}_n^2(\mathbf{f}) = (\mathcal{N}_n^2(f_1), \cdots, \mathcal{N}_n^2(f_k))$ where $\mathcal{N}_n^2(f_i)$ is defined in (3.1.5) and define the vector $Z_n^2(\mathbf{f}) = (Z_n^2(f_1), \cdots, Z_n^2(f_k))$ where

$$Z_n^2(f_k) = \frac{1}{\pi}\int_{\mathcal{S}_n} f_k(t)\mathrm{Im}[\mathcal{B}_n(t)]dt,$$

and

$$\mathcal{B}_n(t) = \mathcal{B}_{1,n}(t) + \kappa_4\mathcal{B}_{2,n}(t).$$

Then

$$\|\mathcal{N}_n^2(\mathbf{f}) - Z_n^2(\mathbf{f})\| \xrightarrow[N,n\to\infty]{} 0.$$

3.1.4 Estimation of the variance

The key point in the study of the fluctuations is an estimate of $\mathrm{Var}[u_n(t)]$. We shall estimate the variance in more general cases. Consider the model

$$\mathbf{Y}_n = \frac{1}{\sqrt{n}}\mathbf{R}_n^{1/2}\mathbf{X}_n + \mathbf{A}_n,$$

where $\mathbf{A_n}$ is a sequence of deterministic matrices with bounded spectral measure and $\mathbf{R}_n$, $\mathbf{X}_n$ satisfy Assumptions 3.1.1, 3.1.2, 3.1.3. We first need a truncated version of the entries.

Truncation of the random variables

As suggested in [11], Najim [70], Najim and Yao [71], the first step consists in the truncation of the random variables, *i.e.*, the replacement of the entries X_{ij} satisfying that there exists a sequence $\eta_n \to 0$ such that $\eta_n n^{1/5} \to \infty$, $|X_{ij}| < \eta_n\sqrt{n}$. More precisely, by Assumption 3.1.1, we may select $\eta_n \downarrow 0$ such that :

$$\frac{1}{\eta_n^4}\mathbb{E}|X_{11}|^4\mathbb{I}(|X_{11}| \geq \eta_n\sqrt{n}) \to 0, \quad n \to \infty.$$

Let $\tilde{\mathbf{Y}}_n = \frac{1}{\sqrt{n}}\mathbf{R}_n^{1/2}\tilde{\mathbf{X}}_n + \mathbf{A}_n$ with $\tilde{\mathbf{X}}_n$, an $N \times n$ matrix with the entries $\tilde{X}_{ij} = (\check{X}_{ij} - \mathbb{E}\check{X}_{ij})/\sigma_{ij}$ where $\check{X}_{ij} = X_{ij}\mathbb{I}_{|X_{ij}| \leq \eta_n\sqrt{n}}$ and $\sigma_{ij} = \mathrm{Var}\check{X}_{ij}$. We have then

$$\mathbb{P}(\mathbf{Y}_n\mathbf{Y}_n^T \neq \tilde{\mathbf{Y}}_n\tilde{\mathbf{Y}}_n^T) \xrightarrow[n\to\infty]{} 0,$$

and

$$\max\left(|\mathbb{E}\tilde{X}_{ij} - \mathbb{E}X_{ij}|, \mathbb{E}|\tilde{X}_{ij}|^2 - 1, \mathbb{E}|\tilde{X}_{ij}|^4 - \mathbb{E}|X_{ij}|^4\right) \xrightarrow[n\to\infty]{} 0.$$

Moreover, for future use, for any function f of class C^4,

$$\mathbb{P}\left[\left|\int f(\lambda)dF^{\mathbf{Y}_n\mathbf{Y}_n^T}(d\lambda) - \int f(\lambda)dF^{\tilde{\mathbf{Y}}_n\tilde{\mathbf{Y}}_n^T}(d\lambda)\right| > \epsilon\right] \xrightarrow[N,n\to\infty]{} 0.$$

Hence it is sufficient to consider the truncated, centralized and renormalized variables. This will be assumed throughout the chapter.

Then we have the estimate of the variance.

Theorem 3.1.3 *Under Assumptions 3.1.1, 3.1.2, 3.1.3, let $V_n(t) = \mathrm{Var}[u_n(t)]$, we have*

$$V_n(t) \leq C_0(\mu_4)(1 + |t|^4)^2,$$

where C_0 is a constant which only depends on μ_4.

The idea behind this theorem is Fourier transform. For an integrable function f, we have : For $x \in \mathbb{R}$,

$$\hat{f}(t) = \frac{1}{2\pi}\int_{\mathbb{R}} f(x)e^{-\mathrm{i}xt}dx$$

and by the inverse formula, we have

$$f(x) = \int_{\mathbb{R}} \hat{f}(t) e^{\mathbf{i}xt} dt.$$

Given a function f, the linear statistics of the eigenvalue of $\mathbf{Y}\mathbf{Y}^T$ is

$$\mathcal{N}(f) = \mathrm{Tr} f\left(\mathbf{Y}\mathbf{Y}^T\right) = \sum_{i=1}^{N} f(\lambda_i),$$

where the λ_i's are the eigenvalues of matrix $\mathbf{Y}\mathbf{Y}^T$. Applying the precedent theorem, one has immediately that

$$\mathrm{Var}(\mathcal{N}(f)) \leq \left(C_0(\mu_4) \int (1+|t|^4)|\hat{f}(t)|dt\right)^2. \tag{3.1.12}$$

Such an inequality has proved to be useful to establish CLTs for the linear statistics of large random covariance matrices, (cf. [60, 74]) and should be compared to logarithmic sobolev inequalities(LSI) and Talagrand concentration inequalities.

In the case of LSI, the constraint over f is lighter (f only needs to be Lipschitz with bounded Lipschitz norm - a fact that is induced by the condition that the right hand side of (3.1.12) should be bounded,) but the condition over the entries is much stronger as their distribution needs to satisfy a LSI, which essentially induces subgaussiannity.

In the case of Talagrand concentration inequalities, the regularity constraint over f is the same as LSI but there is a strong structure constraint as f needs to be convex as well (or to have a limited number of inflection points).

The chapter is organized as below. In Section 3.2, the proof of Theorem 3.1.3 is given for $\mathbf{Y}_n = \frac{1}{\sqrt{n}}\mathbf{R}_n^{1/2}\mathbf{X}_n + \mathbf{A}_n$ in real case. When $\mathbf{A}_n = \mathbf{0}$, Section 3.3 gives the proof of Theorem 3.1.1 by means of Lemma 3.1.2. The expression of real integrals of the variance is also proposed. The computation of the bias in this case is in Section 3.4. Some technical proofs are in Appendices.

3.2 Proof of Theorem 3.1.3 : Estimation of the variance

3.2.1 Some preparations

More notations

We denote $\hat{X}_{ij}$, the $N \times n$ random matrix with i.i.d. real standard gaussian entries. The terms related to $\hat{X}_{ij}$ will be denoted as

$$\hat{\mathbf{Y}}_n = \frac{1}{n}\mathbf{R}_n^{1/2}\hat{\mathbf{X}}_n + \mathbf{A}_n, \quad \hat{\mathbf{U}}(t) = e^{\mathbf{i}t\hat{\mathbf{Y}}_n\hat{\mathbf{Y}}_n^T}, \quad \hat{u}(t) = \mathrm{Tr}\hat{\mathbf{U}}(t).$$

Denote by μ_ℓ the moments $\mathbb{E}X_{ij}^\ell$ and κ_ℓ the associated cumulants.

The technique proposed by Lytova and Pastur [60] consists in considering the interpolating matrix

$$\mathbf{X}_s = \sqrt{s}\mathbf{X} + \sqrt{1-s}\hat{\mathbf{X}}$$

and the related terms will be denoted by $\mathbf{Y}(s) = \frac{1}{\sqrt{n}}\mathbf{R}_n^{1/2}\mathbf{X}_s + \mathbf{A}_n$ or $\mathbf{Y}_s$ when there is no confusion. Other related terms will be denoted by $\mathbf{U}(t,s)$ and $u(t,s)$. Associated to these three matrices are the following quantities :

$\mathbf{X}$	$\hat{\mathbf{X}}$	$\mathbf{X}_s$
$\mathbf{Y} = \frac{1}{\sqrt{n}}\mathbf{R}^{1/2}\mathbf{X} + \mathbf{A}$	$\hat{\mathbf{Y}} = \frac{1}{\sqrt{n}}\mathbf{R}^{1/2}\hat{\mathbf{X}} + \mathbf{A}$	$\mathbf{Y}_s = \frac{1}{\sqrt{n}}\mathbf{R}^{1/2}\mathbf{X}_s + \mathbf{A}$
$\mathbf{U}(t) = e^{\mathbf{i}t\mathbf{Y}\mathbf{Y}^T}$	$\hat{\mathbf{U}}(t) = e^{\mathbf{i}t\hat{\mathbf{Y}}\hat{\mathbf{Y}}^T}$	$\mathbf{U}(t,s) = e^{\mathbf{i}t\mathbf{Y}_s\mathbf{Y}_s^T}$
$u(t) = \mathrm{Tr}\mathbf{U}(t)$	$\hat{u} = \mathrm{Tr}\hat{\mathbf{U}}(t)$	$u(t,s) = \mathrm{Tr}\mathbf{U}(t,s)$

The idea is to show Theorem 3.1.3 under Gaussian case. Then matrix $\mathbf{X}_s$ will naturally appear when considering the difference

$$\alpha(t) = u(t) - \hat{u}(t) = \int_0^1 \frac{d}{ds}u(t,s)ds$$

which is at the heart of the interpolation procedure.

Differentiation formulas

Differentiating matrix functionals with respect to matrices' entries will be important in the sequel. This can be done with the help of Duhamel's formula which can be shown by a simple integration by parts :

Proposition 3.2.1 (Duhamel Formula) *Let $\mathbf{M}_1$, $\mathbf{M}_2$ be two $n\times n$ real symmetric or Hermitian matrices and $t \in \mathbb{R}$. Then we have*

$$e^{\mathbf{i}\mathbf{M}_2 t} = e^{\mathbf{i}\mathbf{M}_1 t} + \mathbf{i}\int_0^t e^{\mathbf{i}\mathbf{M}_1(t-s)}(\mathbf{M}_2 - \mathbf{M}_1)e^{\mathbf{i}\mathbf{M}_2 s}ds.$$

To describe the differential, we introduce the convolution notations. If f and g are two functions of t,

$$f * g(t) = \int_0^t f(t-t_1)g(t_1)dt_1.$$

Lemma 3.2.1 *First order derivatives*

$$\begin{aligned}
\frac{\partial}{\partial Y_{jk}}[U(t)]_{mn} &= \mathbf{i}([Y^TU(t)]_{km} * [U(t)]_{jn} + [U(t)]_{mj} * [Y^TU(t)]_{kn}), && (3.2.1)\\
\frac{\partial}{\partial Y_{jk}}[Y^TU(t)]_{kj} &= [U(t)]_{jj} + \mathbf{i}[Y^TU(t)Y]_{kk} * [U(t)]_{jj} \\
&\quad +\mathbf{i}[Y^TU(t)]_{kj} * [Y^TU(t)]_{kj}, && (3.2.2)\\
\frac{\partial}{\partial Y_{jk}}u(t) &= 2\mathbf{i}t[Y^TU(t)]_{kj}, && (3.2.3)\\
\frac{\partial}{\partial Y_{jk}}[Y^TU(t)Y]_{kk} &= 2[Y^TU(t)]_{kj} + 2\mathbf{i}[Y^TU(t)Y]_{kk} * [Y^TU(t)]_{kj}. && (3.2.4)
\end{aligned}$$

Second order derivatives

$$\begin{aligned}\frac{\partial^2}{\partial(Y_{jk})^2}[Y^TU(t)]_{kj} &= 6\mathbf{i}[U(t)]_{jj} * [Y^TU(t)]_{kj} - 6[Y^TU(t)Y]_{kk} * [Y^TU(t)]_{kj} * [U(t)]_{jj} \\ &\quad -2[Y^TU(t)]_{kj} * [Y^TU(t)]_{kj} * [Y^TU(t)]_{kj}, \qquad (3.2.5) \\ \frac{\partial^2}{\partial(Y_{jk})^2}u(t) &= 2\mathbf{i}t([U(t)]_{jj} + \mathbf{i}[Y^TU(t)Y]_{kk} * [U(t)]_{jj} + [U(t)Y]_{jk} * [U(t)Y]_{jk}). \qquad (3.2.6)\end{aligned}$$

The proof is in Appendix 3.5.2.

In general case, we introduce also the following differentiation operator :

$$D_{\alpha\beta} = \frac{\partial}{\partial X_{\alpha\beta}}, \quad \text{and} \quad D_{\alpha\beta}(s) = \frac{\partial}{\partial[X_s]_{\alpha\beta}}$$

with the obvious iterations $D^\ell_{\alpha\beta} = \frac{\partial^\ell}{\partial(X_{\alpha\beta})^\ell}$.

Since

$$\mathbf{Y} = \frac{1}{\sqrt{n}}\mathbf{R}^{1/2}\mathbf{X} + \mathbf{A},$$

we obtain that

$$\begin{aligned}D_{jk}[R^{1/2}U(t)R^{1/2}]_{mn} &= \frac{\mathbf{i}}{\sqrt{n}}([R^{1/2}U(t)Y]_{mk} * [R^{1/2}U(t)R^{1/2}]_{jn} \\ &\quad +[R^{1/2}U(t)R^{1/2}]_{mj} * [Y^TU(t)R^{1/2}]_{kn}), \qquad (3.2.7) \\ D_{jk}[Y^TU(t)R^{1/2}]_{kj} &= \frac{1}{\sqrt{n}}[R^{1/2}U(t)R^{1/2}]_{jj} + \frac{\mathbf{i}}{\sqrt{n}}[Y^TU(t)Y]_{kk} * [R^{1/2}U(t)R^{1/2}]_{jj} \\ &\quad +\frac{\mathbf{i}}{\sqrt{n}}[Y^TU(t)R^{1/2}]_{kj} * [Y^TU(t)R^{1/2}]_{kj}, \qquad (3.2.8) \\ D_{jk}u(t) &= \frac{2\mathbf{i}t}{\sqrt{n}}[Y^TU(t)R^{1/2}]_{kj}, \qquad (3.2.9) \\ D_{jk}[Y^TU(t)Y]_{kk} &= \frac{2}{\sqrt{n}}[Y^TU(t)R^{1/2}]_{jk} + \frac{2\mathbf{i}}{\sqrt{n}}[Y^TU(t)Y]_{kk} * [Y^TU(t)R^{1/2}]_{kj}. \qquad (3.2.10)\end{aligned}$$

and second order derivatives

$$\begin{aligned}D^2_{jk}[Y^TU(t)R^{1/2}]_{kj} &= \frac{6\mathbf{i}}{n}[R^{1/2}U(t)R^{1/2}]_{jj} * [Y^TU(t)R^{1/2}]_{kj} \qquad (3.2.11) \\ &\quad -\frac{6}{n}[Y^TU(t)Y]_{kk} * [Y^TU(t)R^{1/2}]_{kj} * [R^{1/2}U(t)R^{1/2}]_{jj} \\ &\quad -\frac{2}{n}[Y^TU(t)R^{1/2}]_{kj} * [Y^TU(t)R^{1/2}]_{kj} * [Y^TU(t)R^{1/2}]_{kj}, \\ D^2_{jk}u(t) &= \frac{2\mathbf{i}t}{n}([R^{1/2}U(t)R^{1/2}]_{jj} + \mathbf{i}[Y^TU(t)Y]_{kk} * [R^{1/2}U(t)R^{1/2}]_{jj} \\ &\quad +[Y^TU(t)R^{1/2}]_{jk} * [Y^TU(t)R^{1/2}]_{jk}). \qquad (3.2.12)\end{aligned}$$

Of course we may extend the previous formulas for

$$\hat{D}_{jk} = \frac{\partial}{\partial \hat{X}_{jk}} \quad \text{and} \quad D_{jk}(s) = \frac{\partial}{\partial [X_s]_{jk}}.$$

There are some bridges between these notations. Recall that $\mathbf{X}_s = \sqrt{s}\mathbf{X} + \sqrt{1-s}\hat{\mathbf{X}}$. If $\Phi = \Phi(\mathbf{X}_s)$, then Φ can be considered as a function of $\mathbf{X}$ or $\hat{\mathbf{X}}$ as well and we may use the notational shortcut $\Phi = \Phi(\mathbf{X})$ (resp. $\Phi = \Phi(\hat{\mathbf{X}})$). In this regard, we may use the convenient (but somehow hazardous) notation :

$$\begin{aligned} D_{jk}\Phi &= \sqrt{s} D_{jk}(s)\Phi \\ \hat{D}_{jk}\Phi &= \sqrt{1-s} D_{jk}(s)\Phi. \end{aligned} \tag{3.2.13}$$

In the different terms, the function Φ is respectively assumed to be a function of $\mathbf{X}_s$, then of $\mathbf{Y}$, then of $\hat{\mathbf{Y}}$; the parameter of Φ being indicated by the differentiation operator.

Remark 3.2.1 *Beware that if Φ is not a function of $\mathbf{Y}_s$, then (3.2.13) may not hold any more. Take for instance the simple example :*

$$\Phi = u(t) - \hat{u}(t).$$

According to (3.2.9) : $D_{jk}\Phi = \frac{2\mathrm{i}t}{\sqrt{n}}[Y^T U(t) R^{1/2}]_{kj}$ *while* $\hat{D}_{jk}\Phi = -\frac{2\mathrm{i}t}{\sqrt{n}}[\hat{Y}^T \hat{U}(t) R^{1/2}]_{kj}$.

By iterating the derivatives, although the expression of $D_{jk}^{\ell}[Y^T U(t) R^{1/2}]_{kj}$ might be lengthy, only three terms appear frequently : $[Y^T U(t) R^{1/2}]_{kj}$, $[R^{1/2} U(t) R^{1/2}]_{jj}$ and $[Y^T U(t) Y]_{kk}$. This is important to understand the technique.

Generalized integration by part formula

We recall the generalized integration by part formula, as exposed in Pastur and Shcherbina in [74, Proposition 18.1.4].

Proposition 3.2.2 *Let ξ be a random variable such that $\mathbb{E}\{|\xi|^{p+2}\} < \infty$ for a certain nonnegative integer p. Then for any function $\Phi : \mathbb{R} \to \mathbb{C}$ of the class C^{p+1} with bounded derivatives $\Phi^{(\ell)}$, $\ell = 1, ..., p+1$, we have*

$$\mathbb{E}\{\xi\Phi(\xi)\} = \sum_{\ell=0}^{p} \frac{\kappa_{\ell+1}}{\ell!} \mathbb{E}\{\Phi^{(\ell)}(\xi)\} + \varepsilon_p,$$

where the remainder term ϵ_p admits the bound

$$|\varepsilon_p| \le C_p \mathbb{E}\{|\xi|^{p+2}\} \sup_{t\in\mathbb{R}} |\Phi^{(p+1)}(t)|, \qquad C_p \le \frac{1 + (3+2p)^{p+2}}{(p+1)!}.$$

Remark 3.2.2 *Notice that if ξ has compact support $[a, b]$, then the bound over ε_p takes the sharper form :*

$$|\varepsilon_p| \le C_p \mathbb{E}\{|\xi|^{p+2}\} \sup_{t\in[a,b]} |\Phi^{(p+1)}(t)|.$$

In order to see this, consider the function

$$\tilde{\Phi} = \begin{cases} \Phi(t) & \text{if } t \in [a,b], \\ \sum_{k=0}^{p+1} \Phi^{(k)}(a) \frac{(t-a)^k}{k!} & \text{if } t < a, \\ \sum_{k=0}^{p+1} \Phi^{(k)}(b) \frac{(t-b)^k}{k!} & \text{if } t > b. \end{cases}$$

Then $\tilde{\Phi}$ *is of the class* $C^{(p+1)}$*, the following equalities hold true :* $\mathbb{E}\{\xi\tilde{\Phi}(\xi)\} = \mathbb{E}\{\xi\Phi(\xi)\}$, $\mathbb{E}\tilde{\Phi}^{(k)}(\xi) = \mathbb{E}\Phi^{(k)}(\xi)$ *and*

$$\sup_{t\in\mathbb{R}} |\tilde{\Phi}^{(p+1)}(t)| = \sup_{t\in[a,b]} |\tilde{\Phi}^{(p+1)}(t)|.$$

Hence the sharper estimate.

Strategy of the proof

The general strategy of proof has been exposed in Lytova and Pastur [60]. The main idea consists in interpolating between the linear statistics $u(t) = \mathrm{Tr}e^{\mathbf{i}t\mathbf{Y}\mathbf{Y}^T}$ (based on truncated random variables) and the Gaussian linear statistics $\hat{u}(t) = \mathrm{Tr}e^{\mathbf{i}t\hat{\mathbf{Y}}\hat{\mathbf{Y}}^T}$. In the Gaussian case, a straightforward application of Poincaré-Nash inequality [41, 75, 22] yields :

$$\mathrm{Var}\{\hat{u}(t)\} \leq Kt^2.$$

Since the entries of $\mathbf{Y}$ are assumed to have a fourth finite moment (which is the optimal condition for a CLT to hold for the same normalization), one can only use the generalized integration by part formula (cf. Proposition 3.2.2) up to the fourth cumulant, and this is not sufficient. In order to circumvent this issue, Lytova and Pastur perform a two-stage interpolation.

Since

$$\begin{aligned} V_n &= \mathrm{Var}(u(t)) \\ &= \mathbb{E}\mathring{u}(t)\mathring{u}(-t) \\ &= \mathbb{E}(u(t) - \hat{u}(t))(\mathring{u}(-t) - \mathring{\hat{u}}(-t)) - \mathbb{E}\hat{u}(t)\mathring{\hat{u}}(-t) \\ &= \mathbb{E}\alpha(t)\mathring{\alpha}(-t) - \mathbb{E}\hat{u}(t)\mathring{\hat{u}}(-t), \end{aligned} \tag{3.2.14}$$

in order to perform the first interpolation, we shall introduce three terms $T_1 - B$, T_2 and ε_2 that arise from an application of the generalized integration by parts formula and write :

$$\mathbb{E}(u(t) - \hat{u}(t))\mathring{\alpha}(-t) = \mathbf{i}t \int_0^1 ((T_1 - B) + T_2 + \varepsilon_2)ds.$$

A deeper study on ε_2 will yield the second interpolation :

$$\varepsilon_2 = \varepsilon_2' + \mathbf{i}t \int_0^1 ((T_1' - B') + T_2' + \varepsilon_2'')ds_1.$$

Controlling successively the terms $T_1 - B$ and T_2, then the terms ε_2', $T_1' - B'$, T_2' and ε_2'' will yield the master inequality

$$V_n \leq C_3(t)V_n^{1/2} + C_8(t),$$

where $C_\ell(t)$ is a polynomial in $|t|$ with nonnegative coefficients and degree $\ell \geq 0$. This inequality yields the conclusion.

3.2.2 First interpolation

Useful estimates

Denote $C_\ell(t)$ a polynomial in |t| with nonnegative coefficients and degree ℓ. The coefficients of $C_\ell(t)$ may change from line to line, however the degree is ℓ. As introduced in the precedent paragraph,

$$V_n = \mathrm{Var}(u(t)) = \mathbb{E}[(u(t) - \hat{u}(t))\mathring{\alpha}(-t)] - \mathbb{E}\hat{u}(t)\mathring{\hat{u}}(-t).$$

Notice that $\mathrm{Var}\{u(t)\} = \mathrm{Var}\{u(-t)\}$. We first give an estimate of $\mathrm{Var}[\hat{u}(t)]$.

Proposition 3.2.3 *The following estimates hold true*

$$\begin{aligned} \mathrm{Var}(\hat{u}(\pm t)) &\leq C_2(t) && (3.2.15)\\ \mathbb{E}^{1/2}|\mathring{\alpha}(\pm t)|^2 &\leq C_0 V_n^{1/2} + C_1(t). && (3.2.16) \end{aligned}$$

Proof: By Nash-Poincaré inequality (see for instance [74, Chapter 7]) and 3.2.9, we have

$$\begin{aligned} \mathrm{Var}[\hat{u}(t)] &\leq \sum_{i,j} \mathbb{E}\left|\frac{\partial \hat{u}(t)}{\partial \hat{X}_{ij}}\right|^2 \\ &\leq \frac{C_2(t)}{n}\mathbb{E}[\mathrm{Tr}\hat{\mathbf{U}}(t)\mathbf{R}\hat{\mathbf{U}}^*(t)\hat{\mathbf{Y}}\hat{\mathbf{Y}}^T] \\ &\leq \frac{C_2(t)}{n}\mathbb{E}[\mathrm{Tr}\hat{\mathbf{Y}}\hat{\mathbf{Y}}^T] \\ &\leq C_2(t). \end{aligned}$$

This proves (3.2.15).

(3.2.16) can be shown by Cauchy-Schwarz.

$$\begin{aligned} \mathbb{E}^{1/2}|\mathring{\alpha}(t)|^2 &\leq \sqrt{2}\mathbb{E}^{1/2}(|\mathring{u}_n(t)|^2 + |\mathring{\hat{u}}_n(t)|^2) \\ &\leq \sqrt{2}\sqrt{V_n + C_2(t)} \\ &\leq C_0 V_n^{1/2} + C_1(t). \end{aligned}$$

□

The following lemma is the key to estimate $[Y^T U(t) R^{1/2}]_{kj}$ and $[Y^T U(t) Y]_{kk}$.

Lemma 3.2.2 *We have :*

$$\begin{aligned} |[Y^T U(t) Y]_{kk}(t)| &\leq \sum_\alpha (Y_{\alpha k})^2, && (3.2.17)\\ 0 \leq \ell \leq 3, \quad |D^\ell_{jk}[Y^T U(t) R^{1/2}]_{kj}| &\leq \frac{C_\ell(t)}{n^{\ell/2}}\left(\sum_\alpha Y^2_{\alpha k}\right)^{(\ell+1)/2}, && (3.2.18) \end{aligned}$$

$$a, b \geq 0, \quad |D_{pq}^a D_{jk}^b [Y^T U(t) R^{1/2}]_{kj}| \quad \leq \quad \frac{C_{a+b}(t)}{n^{(a+b)/2}} \left(\sum_\alpha Y_{\alpha k}^2 \right)^{(b+1)/2} \left(\sum_\alpha Y_{\alpha q}^2 \right)^{a/2}, \tag{3.2.19}$$

$$1 \leq \ell \leq 3, \quad |D_{jk}^\ell u(t)| \quad \leq \quad \frac{C_\ell(t)}{n^\ell} \left(\sum_\alpha Y_{\alpha k}^2 \right)^{\ell/2}. \tag{3.2.20}$$

The proof is postponed to Appendix 3.5.3.

The following proposition gives an estimate of the right hand side of the terms in the precedent lemma.

Proposition 3.2.4 *For $\ell = 1, 2, 3, 4$,*

$$\mathbb{E}|\sum_\alpha Y_{\alpha k}^2|^\ell \leq C_0.$$

The proof is in Appendix 3.5.5.

Introduction of the terms $T_1 - B$, T_2 and ε_2

We now treat the term $\mathbb{E}(u(t) - \hat{u}(t))\mathring{\alpha}(-t)$. By considering the interpolating matrix

$$\mathbf{Y}(s) = s^{1/2} \frac{\mathbf{R}^{1/2}\mathbf{X}}{\sqrt{n}} + (1-s)^{1/2} \frac{\mathbf{R}^{1/2}\hat{\mathbf{X}}}{\sqrt{n}} + \mathbf{A},$$

we have

$$\begin{aligned} u(t) - \hat{u}(t) &= \int_0^1 \frac{d}{ds} u(t,s) ds \\ &= \frac{\mathbf{i}t}{\sqrt{n}} \int_0^1 \operatorname{Tr} \left[\exp(\mathbf{i}t \mathbf{Y}_s {\mathbf{Y}_s}^T) \left(\frac{\mathbf{R}^{1/2}\mathbf{X}}{2\sqrt{s}} - \frac{\mathbf{R}^{1/2}\hat{\mathbf{X}}}{2\sqrt{1-s}} \right) \mathbf{Y}_s^T \right] ds \\ &\quad + \frac{\mathbf{i}t}{\sqrt{n}} \int_0^1 \operatorname{Tr} \left[\exp(\mathbf{i}t \mathbf{Y}_s \mathbf{Y}_s^T) \mathbf{Y}_s \left(\frac{\mathbf{R}^{1/2}\mathbf{X}^T}{2\sqrt{s}} - \frac{\mathbf{R}^{1/2}\hat{\mathbf{X}}^T}{2\sqrt{1-s}} \right) \right] ds \\ &= \frac{\mathbf{i}t}{\sqrt{n}} \int_0^1 \operatorname{Tr} \left[\left(\frac{\mathbf{R}^{1/2}\mathbf{X}}{\sqrt{s}} - \frac{\mathbf{R}^{1/2}\hat{\mathbf{X}}}{\sqrt{1-s}} \right) \mathbf{Y}_s^T \mathbf{U}(t,s) \right] ds. \end{aligned} \tag{3.2.21}$$

Hence the relation,

$$\mathbb{E}[(u(t) - \hat{u}(t))\mathring{\alpha}(-t)] \triangleq \mathbf{i}t \int_0^1 [A_n - B_n] ds, \tag{3.2.22}$$

where

$$A_n = \frac{1}{\sqrt{ns}} \sum_{j,k=1}^n \mathbb{E}\{X_{jk}\Phi_n\}, \quad B_n = \frac{1}{\sqrt{n(1-s)}} \sum_{j,k=1}^n \mathbb{E}\{\hat{X}_{jk}\Phi_n\}$$

with

$$\Phi_n = [{Y_s}^T U(t,s) R^{1/2}]_{kj} \mathring{\alpha}(-t).$$

By the generalized integration by part formula (cf. Proposition 3.2.2) with $p = 2$, and beware that $\kappa_1 = 0$, (cf. Section 3.1.4) we have

$$A_n = T_1 + T_2 + \varepsilon_2$$

where

$$\begin{aligned}
T_1 &= \frac{\kappa_2}{\sqrt{sn}} \sum_{j,k}^{n} \mathbb{E}\left\{ \frac{\partial}{\partial X_{jk}} [Y_s^T U(t,s) R^{1/2}]_{kj} \mathring{\alpha}(-t) \right\} \\
&\quad + \frac{\kappa_2}{\sqrt{sn}} \sum_{j,k}^{n} \mathbb{E}\left\{ [Y_s^T U(t,s) R^{1/2}]_{kj} \frac{\partial}{\partial X_{jk}} \mathring{\alpha}(-t) \right\}, \\
T_2 &= \frac{\kappa_3}{2\sqrt{sn}} \sum_{j,k}^{n} \mathbb{E}\left\{ \frac{\partial^2}{\partial (X_{jk})^2} [Y_s^T U(t,s) R^{1/2}]_{kj} \mathring{\alpha}(-t) \right\} \\
&\quad + \frac{\kappa_3}{\sqrt{sn}} \sum_{j,k}^{n} \mathbb{E}\left\{ \frac{\partial}{\partial (X_{jk})} [Y_s^T U(t,s) R^{1/2}]_{kj} \frac{\partial}{\partial (X_{jk})} \mathring{\alpha}(-t) \right\} \\
&\quad + \frac{\kappa_3}{2\sqrt{sn}} \sum_{j,k}^{n} \mathbb{E}\left\{ [Y_s^T U(t,s) R^{1/2}]_{kj} \frac{\partial^2}{\partial (X_{jk})^2} \mathring{\alpha}(-t) \right\}
\end{aligned}$$

and

$$|\varepsilon_2| \leq \frac{C_0 \mu_4}{\sqrt{ns}} \sum_{j,k=1}^{n} \sup_{|W| \leq \eta_n \sqrt{n}} \left| \frac{\partial^3}{\partial (X_{jk})^3} \mathbb{E}\left\{ \Phi_n(X) | X_{jk} = W \right\} \right|.$$

The crux of this interpolation procedure is the partial cancelations that occur between the terms T_1 and B. The aim of the this section is to prove that

$$\begin{aligned}
|\mathbb{E}\alpha(t)\mathring{\alpha}(-t)| &= |t| \left| \int_0^1 (A_n - B_n) ds \right| \\
&\leq C_3(t) V_n^{1/2} + C_4(t) + |t| \int_0^1 |\varepsilon_2| ds.
\end{aligned} \tag{3.2.23}$$

Estimation for $T_1 - B_n$

Recall

$$\begin{aligned}
T_1 &= \frac{\kappa_2}{\sqrt{sn}} \sum_{j,k}^{n} \mathbb{E}\left\{ \frac{\partial}{\partial X_{jk}} [Y_s^T U(t,s) R^{1/2}]_{kj} \mathring{\alpha}(-t) \right\} \\
&\quad + \frac{\kappa_2}{\sqrt{sn}} \sum_{j,k}^{n} \mathbb{E}\left\{ [Y_s^T U(t,s) R^{1/2}]_{kj} \frac{\partial}{\partial X_{jk}} \mathring{\alpha}(-t) \right\}, \\
B_n &= \frac{\kappa_2}{\sqrt{n(1-s)}} \sum_{j,k=1}^{n} \mathbb{E}\{\hat{X}_{jk} \Phi_n\}.
\end{aligned}$$

In this section, we will show that

$$|T_1 - B_n| \leq \frac{C_2(t)}{\sqrt{s}} + \frac{C_2(t)}{\sqrt{1-s}}. \tag{3.2.24}$$

By the generalized integration by part,

$$B_n = \frac{\kappa_2}{\sqrt{(1-s)n}} \sum_{j,k}^{n} \mathbb{E}\left\{ \frac{\partial}{\partial \hat{X}_{jk}} [Y_s^T U(t,s) R^{1/2}]_{kj} \mathring{\alpha}(-t) \right\}$$
$$+ \frac{\kappa_2}{\sqrt{(1-s)n}} \sum_{j,k}^{n} \mathbb{E}\left\{ [Y_s^T U(t,s) R^{1/2}]_{kj} \frac{\partial}{\partial \hat{X}_{jk}} \mathring{\alpha}(-t) \right\},$$

Notice that it is possible to use $D_{jk}(s)$ in the first term of the r.h.s. (see Section 3.2.1) but not in the second where we use instead D_{jk} (see Remark 3.2.1). With this remark, we have

$$\begin{aligned} T_1 &= \frac{\kappa_2}{\sqrt{n}} \sum_{j,k}^{n} \mathbb{E}\{D_{jk}(s)[Y_s^T U(t,s) R^{1/2}]_{kj} \mathring{\alpha}(-t)\} \\ &\quad + \frac{\kappa_2}{\sqrt{sn}} \sum_{j,k}^{n} \mathbb{E}\{[Y_s^T U(t,s) R^{1/2}]_{kj} D_{jk} \mathring{\alpha}(-t)\}, \\ B_n &= \frac{\kappa_2}{\sqrt{n}} \sum_{j,k}^{n} \mathbb{E}\{D_{jk}(s)[Y_s^T U(t,s) R^{1/2}]_{kj} \mathring{\alpha}(-t)\} \\ &\quad + \frac{\kappa_2}{\sqrt{(1-s)n}} \sum_{j,k}^{n} \mathbb{E}\{[Y_s^T U(t,s) R^{1/2}]_{kj} \hat{D}_{jk} \mathring{\alpha}(-t)\}. \end{aligned}$$

and using (3.2.9),

$$\begin{aligned} T_1 - B_n &= \frac{\kappa_2}{\sqrt{sn}} \sum_{j,k}^{n} \mathbb{E}\{[Y_s^T U(t,s) R^{1/2}]_{kj} 2\mathbf{i}(-t)[Y^T U(-t) R^{1/2}]_{kj}\} \\ &\quad - \frac{\kappa_2}{\sqrt{(1-s)n}} \sum_{j,k}^{n} \mathbb{E}\{[Y_s^T U(t,s) R^{1/2}]_{kj} 2\mathbf{i}(-t)[\hat{Y}^T \hat{U}(-t) R^{1/2}]_{kj}\} \\ &\triangleq T_{1a} + T_{1b}. \end{aligned}$$

We deal with T_{1a} :

$$\begin{aligned} |T_{1a}| &\leq \frac{C_1(t)}{\sqrt{sn}} \left| \mathbb{E}[\mathrm{Tr}(\mathbf{R}^{1/2}\mathbf{U}(-t)\mathbf{Y}\mathbf{Y}_s^T \mathbf{U}(t,s)\mathbf{R}^{1/2})] \right| \\ &\leq \frac{C_1(t)}{\sqrt{sn}} \mathbb{E}^{1/2}[\mathrm{Tr}(\mathbf{R}\mathbf{U}(t,s)^* \mathbf{Y}_s \mathbf{Y}_s^T \mathbf{U}(t,s))] \mathbb{E}^{1/2}[\mathrm{Tr}(\mathbf{R}\mathbf{U}(-t)^* \mathbf{Y}\mathbf{Y}^T \mathbf{U}(-t))] \\ &\leq \frac{C_1(t)}{\sqrt{s}}, \end{aligned}$$

and $|T_{1b}|$ can be treated in a similar way. This ends the estimation of (3.2.24).

Estimation for T_2

Recall the expression of T_2 :

$$\begin{aligned}
T_2 &= \frac{\kappa_3}{2\sqrt{sn}}\sum_{j,k}^{n}\mathbb{E}\left\{\frac{\partial^2}{\partial(X_{jk})^2}[Y_s^TU(t,s)R^{1/2}]_{kj}\mathring{\alpha}(-t)\right\}\\
&\quad+\frac{\kappa_3}{\sqrt{sn}}\sum_{j,k}^{n}\mathbb{E}\left\{\frac{\partial}{\partial(X_{jk})}[Y_s^TU(t,s)R^{1/2}]_{kj}\frac{\partial}{\partial(X_{jk})}\mathring{\alpha}(-t)\right\}\\
&\quad+\frac{\kappa_3}{2\sqrt{sn}}\sum_{j,k}^{n}\mathbb{E}\left\{[Y_s^TU(t,s)R^{1/2}]_{kj}\frac{\partial^2}{\partial(X_{jk})^2}\mathring{\alpha}(-t)\right\}\\
&\triangleq T_{2a}+T_{2b}+T_{2c}.
\end{aligned}$$

In this section, we will show that

$$|T_2| \le \frac{C_2(t)V_n^{1/2}+C_3(t)}{\sqrt{s}}. \tag{3.2.25}$$

We start with T_{2a}. Using the differentiation formula (3.2.11) and remark 3.2.1, we obtain

$$\begin{aligned}
T_{2a} &= \frac{\sqrt{s}\kappa_3}{2n^{1/2}}\sum_{j,k}^{n}\mathbb{E}\left\{D_{jk}^2(s)[Y_s^TU(t,s)R^{1/2}]_{kj}\mathring{\alpha}(-t)\right\}\\
&= \frac{3\mathbf{i}\kappa_3\sqrt{s}}{n^{3/2}}\sum_{j,k}\mathbb{E}\left\{([R^{1/2}U(t,s)R^{1/2}]_{jj}*[Y_s^TU(t,s)R^{1/2}]_{kj})\mathring{\alpha}(-t)\right\}\\
&\quad-\frac{3\kappa_3\sqrt{s}}{n^{3/2}}\sum_{j,k}\mathbb{E}\left\{[Y_s^TU(t,s)Y_s]_{kk}*[Y_s^TU(t,s)R^{1/2}]_{kj}*[R^{1/2}U(t,s)R^{1/2}]_{jj}\mathring{\alpha}(-t)\right\}\\
&\quad-\frac{\kappa_3\sqrt{s}}{n^{3/2}}\sum_{j,k}\mathbb{E}\left\{[Y_s^TU(t,s)R^{1/2}]_{kj}*[Y_s^TU(t,s)R^{1/2}]_{kj}*[Y_s^TU(t,s)R^{1/2}]_{kj}\mathring{\alpha}(-t)\right\}.
\end{aligned}$$

In order to handle T_{2a}, we rely on the following proposition whose proof is in Appendix 3.5.4.

Proposition 3.2.5 *Let Z be a random variable with $(\mathbb{E}|Z|^2)^{1/2} \le K_Z$, then the following estimates hold true, for $t_1, t_2, t_3 \in \mathbb{R}$*

$$\left|\mathbb{E}\{\textstyle\sum_{j,k}[R^{1/2}U(t_1,s)R^{1/2}]_{jj}[Y_s^TU(t_2,s)R^{1/2}]_{kj}Z\}\right| \le C_0n^{3/2}K_Z, \tag{3.2.26}$$

$$\left|\textstyle\sum_{j,k}\mathbb{E}\{[Y_s^TU(t_1,s)Y_s]_{kk}[Y_s^TU(t_2,s)R^{1/2}]_{kj}[R^{1/2}U(t_3,s)R^{1/2}]_{jj}Z\}\right| \le C_0n^{3/2}K_Z, \tag{3.2.27}$$

$$\left|\textstyle\sum_{j,k}\mathbb{E}\{[Y_s^TU(t_1,s)R^{1/2}]_{kj}[Y_s^TU(t_2,s)R^{1/2}]_{kj}[Y_s^TU(t_3,s)R^{1/2}]_{kj}\}\right| \le C_0n^{3/2}K_Z. \tag{3.2.28}$$

Remark 3.2.3 *The same result holds true for the terms $[R^{1/2}U(t_1)R^{1/2}]_{jj}$, $[Y_s^TU(t_2,s)R^{1/2}]_{kj}$,* i.e.

$$\left|\mathbb{E}\{\sum_{j,k}[R^{1/2}U(t_1)R^{1/2}]_{jj}[Y^TU(t_2)R^{1/2}]_{kj}Z\}\right| \leq C_0 n^{3/2}K_Z.$$

We obtain similar results for the other two terms. In the sequel, we shall freely use the proposition with $U(t)$ and $U(t,s)$.

Replacing Z by $\mathring{\alpha}(-t)$ with $K_Z = C_0 V_n^{1/2} + C_1(t)$ (cf. (3.2.16)), and taking into account the fact that every convolution product adds a factor t, we end up with :

$$\begin{aligned}|T_{2a}| &\leq C_1(t)V_n^{1/2} + C_2(t) + C_2(t)V_n^{1/2} + C_3(t) + C_2(t)V_n^{1/2} + C_3(t)\\ &\leq C_2(t)V_n^{1/2} + C_3(t).\end{aligned} \tag{3.2.29}$$

Recall that

$$\begin{aligned}
T_{2b} &= \frac{\kappa_3}{2\sqrt{sn}}\sum_{j,k}\mathbb{E}\{D_{jk}[Y_s^TU(t,s)R^{1/2}]_{kj}D_{jk}\mathring{\alpha}(-t)\}\\
&= \frac{\kappa_3}{2n^{1/2}}\sum_{j,k}\mathbb{E}\{D_{jk}(s)[Y_s^TU(t,s)R^{1/2}]_{kj}D_{jk}u(-t)\}\\
&= \frac{\mathbf{i}t\kappa_3}{n^{3/2}}\sum_{j,k}\mathbb{E}\{[R^{1/2}U(t,s)R^{1/2}]_{jj}\times[Y^TU(t)R^{1/2}]_{kj}\}\\
&\quad -\frac{t\kappa_3}{n^{3/2}}\sum_{j,k}\mathbb{E}\{[Y^TU(t)R^{1/2}]_{kj}\times[Y_s^TU(t,s)Y_s]_{kk}*[R^{1/2}U(t,s)R^{1/2}]_{jj}\}\\
&\quad -\frac{t\kappa_3}{n^{3/2}}\sum_{j,k}\mathbb{E}\{[Y^TU(t)R^{1/2}]_{kj}\times[Y_s^TU(t,s)R^{1/2}]_{kj}*[Y_s^TU(t,s)R^{1/2}]_{kj}\}.
\end{aligned}$$

We use again Proposition 3.2.5 with $Z=1$ and obtain

$$|T_{2b}| \leq C_2(t). \tag{3.2.30}$$

Taking advantage from (3.2.12), we handle the term T_{2c} as follows.

$$\begin{aligned}
T_{2c} &= \frac{\kappa_3}{2\sqrt{ns}}\sum_{j,k}\mathbb{E}\{[Y_s^TU(t,s)R^{1/2}]_{kj}D_{jk}^2\mathring{\alpha}(-t)\}\\
&= \frac{\mathbf{i}t\kappa_3}{n^{3/2}\sqrt{s}}\sum_{j,k}\mathbb{E}\{[Y_s^TU(t,s)R^{1/2}]_{kj}\times[R^{1/2}U(-t)R^{1/2}]_{jj}\}\\
&\quad -\frac{t\kappa_3}{n^{3/2}\sqrt{s}}\sum_{j,k}\mathbb{E}\{[Y_s^TU(t,s)R^{1/2}]_{kj}\times[Y^TU(-t)Y]_{kk}*[R^{1/2}U(-t)R^{1/2}]_{jj}\}\\
&\quad -\frac{t\kappa_3}{n^{3/2}\sqrt{s}}\sum_{j,k}\mathbb{E}\{[Y_s^TU(t,s)R^{1/2}]_{kj}\times[Y^TU(-t)R^{1/2}]_{kj}*[Y^TU(-t)R^{1/2}]_{kj}\}.
\end{aligned}$$

By Proposition 3.2.5 with $Z=1$, as there is one convolution product, a factor t appears :

$$|T_{2c}| \leq \frac{C_2(t)}{\sqrt{s}}. \tag{3.2.31}$$

Gathering the estimates (3.2.29) , (3.2.48) and (3.2.31) finally yield (3.2.25).

Gathering (3.2.24) (3.2.25), along with the first interpolation (Formula (3.2.22)), we have

$$|V_n| \leq C_3(t)V_n^{1/2} + C_4(t) + |t| \int_0^1 |\varepsilon_2| ds. \tag{3.2.32}$$

It now remains to handle ε_2.

3.2.3 Decomposition of ε_2

Now we will evaluate ε_2. Recall that

$$|\varepsilon_2| \leq \frac{C_0}{\sqrt{ns}} \sum_{j,k} \sup_{|W| \leq \eta_n \sqrt{n}} \left| \frac{\partial^3}{\partial W^3} \mathbb{E}\left\{\Phi(W) | X_{jk} = W\right\} \right|$$

with

$$\Phi(W) = \mathbb{E}\{[Y_s^T U(t_1, s) R^{1/2}]_{kj} \mathring{\alpha}(-t) | X_{jk} = W\}.$$

In particular, we will show that

$$|\varepsilon_2(s)| \leq \frac{C_7(t)}{\sqrt{s}} \tag{3.2.33}$$

As Y_{jk} is bounded (cf. Section 3.1.4) and $\hat{\mathbf{Y}}$ has all moments, we can change the order between integration and the derivation and expand the third derivative :

$$\begin{aligned} D^3_{jk}\left([Y_s^T U(t_1, s) R^{1/2}]_{kj} \mathring{\alpha}(-t)\right) =& D^3_{jk}[Y_s^T U(t, s) R^{1/2}]_{kj} \mathring{\alpha}(-t) \\ &+ \sum_{\ell=1}^{3} C_3^\ell \left(D^{3-\ell}_{jk} [Y_s^T U(t, s) R^{1/2}]_{kj} D^\ell_{jk} \mathring{\alpha}(-t) \right). \end{aligned}$$

Applying (3.2.13), we have

$$|\varepsilon_2| \leq \frac{C_0}{\sqrt{n}} \sum_{j,k} S_{jk} + |\varepsilon_2'|, \tag{3.2.34}$$

where

$$S_{jk} = \sup_{|W| \leq \eta_n \sqrt{n}} \left| \mathbb{E}\{D^3_{jk}(s)[Y_s^T U(t, s) R^{1/2}]_{kj} \mathring{\alpha}(-t) | X_{jk} = W\} \right|, \tag{3.2.35}$$

$$|\varepsilon_2'| = \frac{C_0}{\sqrt{sn}} \sum_{j,k} \sup_{|W| \leq \eta_n \sqrt{n}} \left| \mathbb{E} \sum_{\ell=1}^{3} C_3^\ell \left(D^{3-\ell}_{jk} [Y_s^T U(t, s) R^{1/2}]_{kj} D^\ell_{jk} \mathring{\alpha}(-t) | X_{jk} = W \right) \right|. \tag{3.2.36}$$

Estimation of ε_2'

Recall the expression of ε_2' in (3.2.36). The aim of this section is to show that

$$|\epsilon_2'| \leq \frac{C_3(t)}{\sqrt{s}}.$$

Using (3.2.18) and (3.2.20), these immediately yield

$$
\begin{aligned}
&\left|\sum_{\ell=1}^{3} C_3^\ell D_{jk}^{3-\ell}[Y_s^T U(t,s)R^{1/2}]_{kj} D_{jk}^\ell u(-t)\right| \\
=&\left|\sum_{\ell=1}^{3} C_3^\ell s^{(3-\ell)/2} D_{jk}^{3-\ell}(s)[Y_s^T U(t,s)R^{1/2}]_{jk} D_{jk}^\ell u(-t)\right| \\
\leq&\frac{C_3(t)}{n^{3/2}}\sum_{\ell=1}^{3}\mathbb{E}\left[\left|\sum_\alpha [Y_s]_{\alpha k}^2\right|^{(4-\ell)/2}\left|\sum_\alpha [Y_{\alpha k}]^2\right|^{\ell/2}\right].
\end{aligned}
\tag{3.2.37}
$$

Now we will show that for all $\ell = 1, 2, 3$,

$$
\mathbb{E}\left[\left|\sum_\alpha [Y_s]_{\alpha k}^2\right|^{(4-\ell)/2}\left|\sum_\alpha (Y_{\alpha k})^2\right|^{\ell/2}\right] \leq C_0. \tag{3.2.38}
$$

When $\ell = 2$,

$$
\begin{aligned}
&\mathbb{E}\left\{\left(\sum_\alpha [Y_s]_{\alpha k}^2\right)\left(\sum_\alpha (Y_{\alpha k})^2\right)\right\} \\
&\leq \quad 2\mathbb{E}\left\{\left(\sum_\alpha [Y_{\alpha k}]^2 + [\hat{Y}_{\alpha k}]^2\right)\left(\sum_\alpha (Y_{\alpha k})^2\right)\right\} \\
&\leq \quad C_0,
\end{aligned}
$$

where the last inequality follows from the hypothesis that $\mathbb{E}|x_{ij}|^4 < \infty$.

When $\ell = 1$,

$$
\begin{aligned}
&\mathbb{E}\left\{\left(\sum_\alpha [Y_s]_{\alpha k}^2\right)^{3/2}\left(\sum_\alpha (Y_{\alpha k})^2\right)^{1/2}\right\} \\
&\leq \quad \mathbb{E}^{1/2}\left\{\left(\sum_\alpha [Y_s]_{\alpha k}^2\right)^2\right\}\mathbb{E}^{1/2}\left\{\left(\sum_\alpha [Y_s]_{\alpha k}^2\right)\left(\sum_\alpha (Y_{\alpha k})^2\right)\right\} \\
&\leq \quad C_0.
\end{aligned}
$$

And $\ell = 3$ can be treated similarly. (3.2.38) is proved. Then by (3.2.37) and (3.2.38),

$$
\begin{aligned}
\frac{1}{\sqrt{n}}\left|\mathbb{E}\sum_{\ell=1}^{3} C_3^\ell\left(D_{jk}^{3-\ell}[Y_s^T U(t,s)R^{1/2}]_{kj} D_{jk}^\ell \mathring{\alpha}(-t)|X_{jk} = W\right)\right| &\leq \frac{C_3(t)}{n^2}(1 + |Y_{jk}|^4) \\
&\leq \frac{C_3(t)}{n^2}.
\end{aligned}
\tag{3.2.39}
$$

Plugging this in (3.2.36), we get

$$
|\varepsilon_2'| \leq \frac{C_3(t)}{\sqrt{s}}. \tag{3.2.40}
$$

3.2.4 Second interpolation

Recall that

$$S_{jk} = \sup_{|W|\le \eta_n\sqrt{n}} \left| \mathbb{E}\{D^3_{jk}(s)[Y_s^T U(t,s)R^{1/2}]_{kj}\mathring{\alpha}(-t)|X_{jk}=W\}\right|.$$

In order to estimate S_{jk}, we perform an interpolation procedure on the quantity $\mathring{\alpha}(-t)$. With the same method as (3.2.22),

$$\begin{aligned}
&\mathbb{E}\left\{D^3_{jk}(s)[Y_s^T U(t,s)R^{1/2}]_{kj}\mathring{\alpha}(-t)|X_{jk}=W\right\}\\
=&\mathbb{E}\left\{D^3_{jk}(s)[Y_s^T U(t,s)R^{1/2}]_{kj}\int_0^1 ds_1\frac{d}{ds_1}u(-t,s_1)|X_{jk}=W\right\}\\
=&-\frac{\mathbf{i}t}{\sqrt{n}}\int_0^1 \mathbb{E}\left\{D^3_{jk}(s)[Y_s^T U(t,s)R^{1/2}]_{kj}\mathrm{Tr}\left[\left(\frac{\mathbf{R}^{1/2}\mathbf{X}}{\sqrt{s_1}}-\frac{\mathbf{R}^{1/2}\hat{\mathbf{X}}}{\sqrt{1-s_1}}\right)\mathbf{Y}^T_{s_1}\mathbf{U}(-t,s_1)\right]ds_1|X_{jk}=W\right\}\\
\triangleq&-\mathbf{i}t\int_0^1(A_0'-B')ds_1,
\end{aligned}$$

where

$$\begin{aligned}
A_0' &= \frac{1}{\sqrt{n}}\mathbb{E}\left\{D^3_{jk}(s)[Y_s^T U(t,s)R^{1/2}]_{kj}\mathrm{Tr}\left[\frac{\mathbf{X}}{\sqrt{s_1}}\mathbf{Y}^T_{s_1}\mathbf{U}(-t,s_1)\mathbf{R}^{1/2}\right]|X_{jk}=W\right\},\\
B' &= \frac{1}{\sqrt{n}}\mathbb{E}\left\{D^3_{jk}(s)[Y_s^T U(t,s)R^{1/2}]_{kj}\mathrm{Tr}\left[\frac{\hat{\mathbf{X}}}{\sqrt{1-s_1}}\mathbf{Y}^T_{s_1}\mathbf{U}(-t,s_1)\mathbf{R}^{1/2}\right]|X_{jk}=W\right\}.
\end{aligned}$$

By distinguishing $(p,q)\neq(j,k)$, we have

$$\begin{aligned}
A_0' &= \frac{1}{\sqrt{n}}\mathbb{E}\left\{D^3_{jk}(s)[Y_s^T U(t,s)R^{1/2}]_{kj}\mathrm{Tr}\left[\frac{\mathbf{X}}{\sqrt{s_1}}\mathbf{Y}^T_{s_1}\mathbf{U}(-t,s_1)\mathbf{R}^{1/2}\right]|X_{jk}=W\right\}\\
&= \frac{1}{\sqrt{ns_1}}\sum_{p,q}{}' \mathbb{E}\left\{X_{pq}[Y^T_{s_1}U(-t,s_1)R^{1/2}]_{qp}D^3_{jk}(s)[Y_s^T U(t,s)R^{1/2}]_{kj}|X_{jk}=W\right\}\\
&\quad+\frac{X_{jk}}{\sqrt{ns_1}}\mathbb{E}\left\{[Y^T_{s_1}U(-t,s_1)R^{1/2}]_{jk}D^3_{jk}(s)[Y_s^T U(t,s)R^{1/2}]_{kj}|X_{jk}=W\right\}\\
&\triangleq A'+A_{0a}',
\end{aligned}$$

where $\sum_{pq}'$ stands for the sum over the indices $\{p,q\}\neq\{j,k\}$. We first prove that

$$|A_{0a}'|\le\frac{C_3(t)W}{\sqrt{s_1}n^{3/2}}.$$

By Lemma 3.2.2,

$$|A_{0a}| \le C_3(t)\frac{W}{\sqrt{s_1}n^2}\mathbb{E}\left\{\left(\sum_\alpha[Y_{s_1}]^2_{\alpha k}\right)^{1/2}\left(\sum_\alpha[Y_s]^2_{\alpha k}\right)^2|X_{jk}=W\right\}$$

$$
\begin{aligned}
&\leq \frac{C_3(t)W}{\sqrt{s_1}n^2}\mathbb{E}\left\{\left(\sum_\alpha (Y_{\alpha k})^2+\hat{Y}_{\alpha k}^2\right)^{5/2}|X_{jk}=W\right\}\\
&\leq \frac{C_3(t)W}{\sqrt{s_1}n^2}\mathbb{E}\left\{\left(\sum_\alpha (Y_{\alpha k})^2\right)^{5/2}|X_{jk}=W\right\}+C_3(t)\frac{W}{\sqrt{s_1}n^2}\mathbb{E}\left\{\left(\sum_\alpha (\hat{Y}_{\alpha k})^2\right)^{5/2}|X_{jk}=W\right\}\\
&\leq \frac{C_3(t)W}{\sqrt{s_1}n^2}\mathbb{E}^{1/2}\left\{\left(\sum_\alpha (Y_{\alpha k})^2\right)^{3}|X_{jk}=W\right\}\mathbb{E}^{1/2}\left\{\left(\sum_\alpha (Y_{\alpha k})^2\right)^{2}|X_{jk}=W\right\}+\frac{C_3(t)W}{n^2\sqrt{s_1}}\\
&\overset{(a)}{\leq} \frac{C_3(t)W}{\sqrt{s_1}n^2}
\end{aligned}
$$

where (a) is from Proposition 3.2.4. Hence

$$|A'_{0a}|\leq \frac{C_3(t)W}{\sqrt{s_1}n^2}\leq \frac{C_3(t)\eta_n}{\sqrt{s_1}n^{3/2}}.$$

Plugging this and (3.2.40) into (3.2.34), we have

$$|\varepsilon_2|\leq C_4(t)+\frac{C_0|t|}{\sqrt{n}}\sup_{|W|\leq\eta_n\sqrt{n}}\int_0^1\left|A'-B'\right|ds_1$$

where

$$
\begin{aligned}
A' &= \frac{1}{\sqrt{ns_1}}\sum_{p,q}{}'\,\mathbb{E}\left\{X_{pq}[Y_{s_1}^TU(-t,s_1)R^{1/2}]_{qp}D_{jk}^3(s)[Y_s^TU(t,s)R^{1/2}]_{kj}|X_{jk}=W\right\}\\
B' &= \frac{1}{\sqrt{ns_1}}\sum_{p,q}{}'\,\mathbb{E}\left\{\hat{X}_{pq}[Y_{s_1}^TU(-t,s_1)R^{1/2}]_{qp}D_{jk}^3(s)[Y_s^TU(t,s)R^{1/2}]_{kj}|X_{jk}=W\right\}.
\end{aligned}
$$

By integration by part formula, (cf. Proposition 3.2.2), we have

$$A'=T'_1+T'_2+\epsilon''_2 \tag{3.2.41}$$

where

$$
\begin{aligned}
T'_1 &= \frac{\kappa_2}{\sqrt{ns_1}}\sum_{p,q}{}'\,\mathbb{E}D_{pq}\left\{[Y_{s_1}^TU(-t,s_1)R^{1/2}]_{qp}D_{jk}^3(s)[Y_s^TU(t,s)R^{1/2}]_{kj}|X_{jk}=W\right\}\\
T'_2 &= \frac{\kappa_3}{\sqrt{s_1n}}\sum_{p,q}{}'\,\mathbb{E}D_{pq}^2\left\{[Y_{s_1}^TU(-t,s_1)R^{1/2}]_{qp}D_{jk}^3(s)[Y_s^TU(t,s)R^{1/2}]_{kj}|X_{jk}=W\right\}\\
B' &= \frac{\kappa_2}{\sqrt{n(1-s_1)}}\sum_{p,q}{}'\,\mathbb{E}\hat{D}_{pq}\left\{[Y_{s_1}^TU(-t,s_1)R^{1/2}]_{qp}D_{jk}^3(s)[Y_s^TU(t,s)R^{1/2}]_{kj}|X_{jk}=W\right\},
\end{aligned}
$$

and

$$|\varepsilon''_2|\leq\frac{C_0}{\sqrt{ns_1}}\sum_{p,q}{}'\sup_{|Z|\leq\eta_n\sqrt{n}}\left|\frac{\partial}{\partial Z^3}\mathbb{E}\left\{[Y_{s_1}^TU_{t,s_1}R^{1/2}]_{qp}D_{jk}^3(s)[Y_s^TU(t,s)R^{1/2}]_{kj}|X_{pq}=Z,X_{jk}=W\right\}\right|.$$

All concentrations should be paid to estimate these terms. As in the first interpolation, B' cancels with T_1'. However, as more terms appear, we need a more delicate method to estimate these terms. One must understand all terms in the development.

Estimation of $T_1' - B'$

We now handle the term $T_1' - B'$. From (3.2.41) where some cancelation occurs. Recall the expression of T_1' and B' :

$$\begin{aligned} T_1' &= \frac{\kappa_2}{\sqrt{ns_1}} \sum_{p,q}' \mathbb{E} D_{pq} \left\{ [Y_{s_1}^T U(-t, s_1) R^{1/2}]_{qp} D_{jk}^3(s) [Y_s^T U(t,s) R^{1/2}]_{kj} | X_{jk} = W \right\} \\ B' &= \frac{\kappa_2}{\sqrt{n(1-s_1)}} \sum_{p,q}' \mathbb{E} \hat{D}_{pq} \left\{ [Y_{s_1}^T U(-t, s_1) R^{1/2}]_{qp} D_{jk}^3(s) [Y_s^T U(t,s) R^{1/2}]_{kj} | X_{jk} = W \right\}. \end{aligned}$$

We shall prove that

$$\sum_{|W| \le \eta_n \sqrt{n}} \frac{|t|}{\sqrt{n}} \int_0^1 |T_1' - B'| ds_1 \le C_5(t). \tag{3.2.42}$$

As in the first interpolation along with Remark 3.2.1,

$$\begin{aligned} D_{pq} [Y_{s_1}^T U(-t, s_1) R^{1/2}]_{qp} &= \sqrt{s_1} D_{pq}(s_1) [Y_{s_1}^T U(-t, s_1) R^{1/2}]_{qp} \\ \hat{D}_{pq} [Y_{s_1}^T U(-t, s_1) R^{1/2}]_{qp} &= \sqrt{1-s_1} D_{pq}(s_1) [Y_{s_1}^T U(-t, s_1) R^{1/2}]_{qp}. \end{aligned}$$

Then we have

$$\begin{aligned} \frac{1}{\sqrt{n}} (T_1' - B') &= \frac{\kappa_2}{\sqrt{s_1} n} \sum_{p,q}' \mathbb{E} \left\{ [Y_{s_1}^T U(-t, s_1) R^{1/2}]_{qp} D_{pq} D_{jk}^3(s) [Y_s^T U(t,s) R^{1/2}]_{kj} | X_{jk} = W \right\} \\ &\quad - \frac{\kappa_2}{\sqrt{1-s_1} n} \sum_{p,q}' \mathbb{E} \left\{ [Y_{s_1}^T U(-t, s_1) R^{1/2}]_{qp} \hat{D}_{pq} D_{jk}^3(s) [Y_s^T U(t,s) R^{1/2}]_{kj} | X_{jk} = W \right\} \\ &\triangleq T_{1a}' + T_{1b}'. \end{aligned}$$

The goal in the following is to show that

$$\begin{aligned} |T_{1a}'| &\le \frac{C_4(t)}{\sqrt{s_1}} \\ |T_{1b}'| &\le \frac{C_4(t)}{\sqrt{1-s_1}}. \end{aligned} \tag{3.2.43}$$

This implies immediately (3.2.42). T_{1a}' and T_{1b}' can be handled similarly. We shall focus on T_{1a}'.

Adding the term corresponding to $(p, q) = (j, k)$ does not affect the order of magnitude of

T'_{1a} since by (3.2.19) and Proposition 3.2.4,

$$\begin{aligned}
&\frac{\kappa_2}{\sqrt{s_1}n}\sum_{j,k}\mathbb{E}\left\{[Y_{s_1}^T U(-t,s_1)R^{1/2}]_{kj} D_{jk} D_{jk}^3(s)[Y_s^T U(t,s)R^{1/2}]_{kj} | X_{jk}=W\right\} \\
&\leq \frac{C_4(t)}{\sqrt{s_1}n^3}\sum_{j,k}\mathbb{E}\left\{\left(\sum_\alpha [Y_{s_1}]_{\alpha k}^2\right)^{1/2}\left(\sum_\alpha [Y_s]_{\alpha k}^2\right)^{5/2} | X_{jk}=W\right\} \\
&\leq \frac{C_4(t)}{\sqrt{s_1}n}\left(1+\frac{W^6}{n^3}\right).
\end{aligned}$$

Now we will disclose all terms in the development of

$$[Y_{s_1}^T U(-t,s_1)R^{1/2}]_{qp} D_{pq}(s) D_{jk}^3 [Y_s^T U(t,s)R^{1/2}]_{kj}.$$

Since the factor before the sum is of order n^{-1}, a rough estimate of each term of the sum as done before will not be sufficient here. One needs to look carefully to the terms depending on p and q in the derivative $D_{pq}D_{jk}^3(s)[Y_s^T U(t,s)R^{1/2}]_{kj}$ and to understand how they combine with $[Y_{s_1}^T U(t,s_1)R^{1/2}]_{pq}$ when summing up. We first need to understand all terms in $D_{jk}^3(s)[Y_s^T U(t,s)R^{1/2}]_{jk}$. Based upon the terms appearing in the second derivative $D_{jk}^2(s)[Y_s^T U(t,s)R^{1/2}]_{kj}$ (cf. (3.2.11)) and the first order derivatives, all terms in $D_{jk}^3(s)[Y_s^T U(t,s)R^{1/2}]_{kj}$ are :

$$\begin{aligned}
L_1 &= [R^{1/2}U(t,s)R^{1/2}]_{jj} * [R^{1/2}U(t,s)R^{1/2}]_{jj} \\
L_2 &= [R^{1/2}U(t,s)R^{1/2}]_{jj} * [Y_s^T U(t,s)R^{1/2}]_{kj} * [Y_s^T U(t,s)R^{1/2}]_{kj} \\
L_3 &= [R^{1/2}U(t,s)R^{1/2}]_{jj} * [R^{1/2}U(t,s)R^{1/2}]_{jj} * [Y_s^T U(t,s)Y_s]_{kk} \\
L_4 &= [R^{1/2}U(t,s)R^{1/2}]_{jj} * [Y_s^T U(t,s)R^{1/2}]_{kj} * [Y_s^T U(t,s)R^{1/2}]_{kj} * [Y_s^T U(t,s)Y_s]_{kk} \\
L_5 &= [R^{1/2}U(t,s)R^{1/2}]_{jj} * [R^{1/2}U(t,s)R^{1/2}]_{jj} * [Y_s^T U(t,s)Y_s]_{kk} * [Y_s^T U(t,s)Y_s]_{jj} \\
L_6 &= [Y_s^T U(t,s)R^{1/2}]_{kj} * [Y_s^T U(t,s)R^{1/2}]_{kj} * [Y_s^T U(t,s)R^{1/2}]_{kj} * [Y_s^T U(t,s)R^{1/2}]_{kj}.
\end{aligned}$$

In order to understand the action of D_{pq} upon all these terms, it is sufficient to understand the action of $D_{pq}(s)$ over the three recurring terms :

$$[R^{1/2}U(t,s)R^{1/2}]_{jj}, \quad [Y_s^T U(t,s)R^{1/2}]_{kj}, \quad [Y_s^T U(t,s)Y_s]_{kk}.$$

Computation yields :

$$\begin{aligned}
D_{pq}[R^{1/2}U(t,s)R^{1/2}]_{jj} &= 2\mathbf{i}\sqrt{s}[Y_s U(t,s)R^{1/2}]_{qj} * [R^{1/2}U(t,s)R^{1/2}]_{jp}, \\
D_{pq}[Y_s^T U(t,s)R^{1/2}]_{kj} &= \sqrt{s}\delta_{q,k}[R^{1/2}U(t,s)R^{1/2}]_{pj} + \mathbf{i}\sqrt{s}[Y_s^T U(t,s)Y_s]_{qk} * [R^{1/2}U(t,s)R^{1/2}]_{pj} \\
&\quad + \sqrt{s}[Y_s^T U(t,s)R^{1/2}]_{qj} * [Y_s^T U(t,s)R^{1/2}]_{kp}, \\
D_{pq}[Y_s^T U(t,s)Y_s]_{kk} &= 2\sqrt{s}\delta_{q,k}[R^{1/2}U(t,s)R^{1/2}]_{kp} + 2\mathbf{i}\sqrt{s}[Y_s^T U(t,s)R^{1/2}]_{kp} * [Y_s^T U(t,s)Y_s]_{qk}.
\end{aligned}$$

Now consider the sets

$\mathcal{B}$	$\mathcal{C}$
$[Y_s^T U(t,s)R^{1/2}]_{qj} * [R^{1/2}U(t,s)R^{1/2}]_{jp}$	$[R^{1/2}U(t,s)R^{1/2}]_{jj}$
$\delta_{q,k}[R^{1/2}U(t,s)R^{1/2}]_{pj}$	$[Y_s^T U(t,s)R^{1/2}]_{kj}$
$[Y_s^T U(t,s)Y_s]_{qk} * [R^{1/2}U(t,s)R^{1/2}]_{pj}$	$[Y_s^T U(t,s)Y_s]_{kk}$
$[Y_s^T U(t,s)R^{1/2}]_{qj} * [Y_s^T U(t,s)R^{1/2}]_{kp}$	
$\delta_{q,k}[R^{1/2}U(t,s)R^{1/2}]_{kp}$	
$[Y_s^T U(t,s)R^{1/2}]_{kp} * [Y_s^T U(t,s)Y_s]_{qk}$	

and introduce the set

$$\mathcal{G} = \{B(p,q) * C_1(j,k), \quad B(p,q) * C_1(j,k) * C_2(j,k), \quad B(p,q) * C_1(j,k) * C_2(j,k) * C_3(j,k)\},$$

where $B(p,q) \in \mathcal{B}, C_i \in \mathcal{C}$ for $i = 1,2,3$.

Up to a constant, the generic terms G that appear in T'_{1a} are

$$G = \sum_{p,q} [Y_{s_1}^T U(t,s_1)R^{1/2}]_{pq} \times g,$$

where $g \in \mathcal{G}$ and

$$T'_{1a} = \frac{\kappa_2}{n\sqrt{s_1}} \sum_{g \in \mathcal{G}} \mathbb{E}\{G|X_{jk} = W\}.$$

We now assert that :

$$\forall g \in \mathcal{G}, \sup_{|W| \leq \eta_n \sqrt{n}} \mathbb{E}\{|G||X_{jk} = W\} \leq nC_4(t).$$

The factor $C_4(t)$ comes from the fact that the maximum power of t in the estimate above is give by the maximum number of convolution products which is 4.

Computations are similar and we only handle one term as an example. Taking

$$\begin{aligned} G = \sum_{p,q} &[Y_{s_1}^T U(t,s_1)R^{1/2}]_{qp} \times [Y_s^T U(t,s)R^{1/2}]_{kp} * [Y_s^T U(t,s)Y_s]_{qk} \\ &* [R^{1/2}U(t,s)R^{1/2}]_{jj} * [R^{1/2}U(t,s)R^{1/2}]_{jj} * [Y_s^T U(t,s)Y_s]_{kk}, \end{aligned} \tag{3.2.44}$$

we have

$$\begin{aligned} &\Big| \sum_{p,q} [Y_{s_1}^T U(t,s_1)R^{1/2}]_{qp} \times [Y_s^T U(t_1,s)R^{1/2}]_{kp} [Y_s^T U(t_2,s)Y_s]_{qk} \\ &\qquad \times [R^{1/2}U(t_3,s)R^{1/2}]_{jj} [R^{1/2}U(t_4,s)R^{1/2}]_{jj} [Y_s^T U(t_5,s)Y_s]_{kk} \Big| \\ &= \big| [Y_s^T U(t_1,s)RU(t,s)Y_{s_1} Y_s^T U(t_2,s)Y_s]_{kk} \\ &\qquad \times [R^{1/2}U(t_3,s)R^{1/2}]_{jj} [R^{1/2}U(t_4,s)R^{1/2}]_{jj} [Y_s^T U(t_5,s)Y_s]_{kk} \big| \\ &\leq C_0 \left| \sum_\alpha [Y_s^T U(t_1,s)RU(t,s)Y_{s_1}]_{k\alpha} [Y_s^T U(t_2,s)Y_s]_{\alpha k} \right| \times \left(\sum_\beta [Y_{s,\beta k}]^2 \right) \end{aligned}$$

$$\leq \quad C_0 \sum_{\alpha} \left(\sum_{\beta} [Y_{s_1}]^2_{\beta\alpha} \right)^{1/2} \left(\sum_{\beta} [Y_s]^2_{\beta\alpha} \right)^{1/2} \left(\sum_{\beta} [Y_s]^2_{\beta k} \right)^{2} . \tag{3.2.45}$$

This kind of term can be handled as before and its expectation is of order n. Hence the estimate

$$\sup_{|W| \leq \eta_n \sqrt{n}} |t| \int_0^1 |T'_{1a}| ds_1 \leq C_5(t). \tag{3.2.46}$$

This ends the proof of (3.2.42).

Estimation of T'_2

In this section, we shall prove that

$$\sup_{|W| \leq \eta_n \sqrt{n}} \frac{|t|}{\sqrt{n}} \int_0^1 |T'_2| ds_1 \leq C_6(t). \tag{3.2.47}$$

From (3.2.41), we have

$$\begin{aligned}
\frac{1}{\sqrt{n}} T'_2 \quad &= \quad \frac{\kappa_3}{\sqrt{s_1} n} \sum_{p,q}{}' \mathbb{E} D^2_{pq} \left\{ [Y^T_{s_1} U(-t, s_1) R^{1/2}]_{qp} D^3_{jk}(s) [Y^T_s U(t,s) R^{1/2}]_{kj} | X_{jk} = W \right\} \\
&= \quad \frac{\kappa_3 \sqrt{s_1}}{n} \sum_{p,q}{}' \mathbb{E} \left\{ D^2_{pq}(s_1) [Y^T_{s_1} U(-t, s_1) R^{1/2}]_{qp} D^3_{jk}(s) [Y^T_s U(t,s) R^{1/2}]_{kj} | X_{jk} = W \right\} \\
&\quad + \frac{\kappa_3}{n} \sum_{p,q}{}' \mathbb{E} \left\{ D_{pq}(s_1) [Y^T_{s_1} U(-t, s_1) R^{1/2}]_{qp} D_{pq} D^3_{jk}(s) [Y^T_s U(t,s) R^{1/2}]_{kj} | X_{jk} = W \right\} \\
&\quad + \frac{\kappa_3}{\sqrt{s_1} n} \sum_{p,q}{}' \mathbb{E} \left\{ [Y^T_{s_1} U(-t, s_1) R^{1/2}]_{qp} D^2_{pq} D^3_{jk}(s) [Y^T_s U(t,s) R^{1/2}]_{kj} | X_{jk} = W \right\} \\
&\triangleq \quad T'_{2a} + T'_{2b} + T'_{2c}.
\end{aligned}$$

Beware that we use the relation $D_{pq} = \sqrt{s_1} D_{pq}(s_1)$ and apply the differentiation operator $D_{pq}(s_1)$ to the quantity $[Y^T_{s_1} U(-t, s_1) R^{1/2}]_{qp}$ but not the quantity $[Y^T_s U(t,s) R^{1/2}]_{kj}$ which depends on Y_s but not on Y_{s_1}.

Notice that by Lemma 3.2.2,

$$\sup_{|W| \leq \eta_n \sqrt{n}} \mathbb{E} \left(\left| D^3_{jk} [Y^T_s U(t,s) R^{1/2}]_{kj} \right|^2 | X_{jk} = W \right)^{1/2} \leq \frac{C_3(t)}{n^{3/2}}.$$

Hence, T'_{2a} can be handled in the same way as the term T_{2a} with the replacement of $\mathring{\alpha}_n$ by $D^3_{jk} [Y^T_s U(t,s) R^{1/2}]_{kj}$. This yields

$$|T'_{2a}| \leq C_5(t).$$

Now we will handle the term T'_{2b} which is by far the most delicate.

$$T'_{2b} = \frac{\kappa_3}{n} \sum_{p,q}' \mathbb{E}\left\{ D_{pq}(s_1)[Y_{s_1}^T U(-t,s_1)R^{1/2}]_{qp} D_{pq} D_{jk}^3(s)[Y_s^T U(t,s)R^{1/2}]_{kj} | X_{jk} = W \right\}$$

In order to proceed, we shall perform the same analysis as for the term in T'_{1a}. The analysis is based on a precise description of the generic term that appears when expanding the derivative $D_{pq}(s_1)[Y_{s_1}^T U(t,s_1)R^{1/2}]_{pq}]$, the derivative $D_{pq}D_{jk}^3[Y_s^T U(t,s)R^{1/2}]_{jk}$ and their product.

Up to a constant, the generic terms are

$$G' = \sum_{p,q} A \times g,$$

where

$$\begin{aligned} A \in \mathcal{A} = \Big\{ &[R^{1/2}U(-t,s_1)R^{1/2}]_{pp}, \quad [Y_{s_1}^T U(-t,s_1)Y_{s_1}]_{qq} * [U(-t,s_1)]_{pp}, \\ &[Y_{s_1}^T U(-t,s_1)R^{1/2}]_{pq} * [Y_{s_1}^T U(-t,s_1)R^{1/2}]_{pq} \Big\} \end{aligned}$$

and $g \in \mathcal{G}$ with $\mathcal{G}$ defined in (3.2.49). Then we can show without difficulties that,

$$|T'_{2b}| = \left| \frac{\kappa_3}{\sqrt{s_1} n^{3/2}} \sum_{A \in \mathcal{A}, g \in \mathcal{G}} \mathbb{E}\{G' | X_{jk} = W\} \right| \leq C_5(t). \tag{3.2.48}$$

We make an example as usual :

$$\begin{aligned} G' = \sum_{p,q} &[Y_{s_1}^T U(t,s_1)R^{1/2}]_{qp} * [Y_{s_1}^T U(t,s_1)R^{1/2}]_{qp} \times [Y_s^T U(t,s)R^{1/2}]_{kp} * [Y_s^T U(t,s)Y_s]_{qk} \\ &* [R^{1/2}U(t,s)R^{1/2}]_{jj} * [R^{1/2}U(t,s)R^{1/2}]_{jj} * [Y_s^T U(t,s)Y_s]_{kk}, \end{aligned} \tag{3.2.49}$$

we have

$$\begin{aligned} &\mathbb{E}\Bigg| \sum_{p,q} [Y_{s_1}^T U(t,s_1)R^{1/2}]_{qp}[Y_{s_1}^T U(t_1,s_1)R^{1/2}]_{qp} \times [Y_s^T U(t_2,s)R^{1/2}]_{kp}[Y_s^T U(t_3,s)Y_s]_{qk} \\ &\times [R^{1/2}U(t_4,s)R^{1/2}]_{jj}[R^{1/2}U(t_5,s)R^{1/2}]_{jj}[Y_s^T U(t_6,s)Y_s]_{kk} \Bigg| \\ &\leq C_0 \mathbb{E}\left| \sum_{p,q} \left(\sum_\alpha [Y_{s_1}]_{\alpha q}^2 \right)^{1/2} [Y_{s_1}^T U(t_1,s_1)R^{1/2}]_{pq}[Y_s^T U(t_2,s)R^{1/2}]_{kp}[Y_s^T U(t_3,s)Y_s]_{qk}[Y_s^T U(t_6,s)Y_s]_{kk} \right| \\ &\leq \sqrt{n} C_0 \mathbb{E}\left| [Y_s^T U(t_2,s)RU(t_1,s_1)Y_{s_1}Y_s^T U(t_3,s)Y_s]_{kk} \sum_\alpha [Y_s]_{\alpha k}^2 \right| \\ &\leq C_0 n^{3/2}, \end{aligned}$$

where the last inequality follows from the same argument as T'_{1a}. The degree 5 of the polynomial comes from the fact that we have 5 convolutions. Thus we get (3.2.48).

Now we will evaluate T'_{2c}.

$$\begin{aligned}
|T'_{2c}| &= \left| \frac{\kappa_3}{\sqrt{s_1}n} \sum_{p,q}{}' \mathbb{E}\left\{ [Y_{s_1}^T U(-t,s_1)R^{1/2}]_{qp} D_{pq}^2 D_{jk}^3(s)[Y_s^T U(t,s)R^{1/2}]_{kj} | X_{jk} = W \right\} \right| \\
&\le \frac{\kappa_3}{\sqrt{s_1}n} \left(\sum_{p,q}{}' \mathbb{E}\left\{ \left| [Y_{s_1}^T U(-t,s_1)R^{1/2}]_{qp} \right|^2 | X_{jk} = W \right\} \right)^{1/2} \\
&\times \left(\sum_{p,q}{}' \mathbb{E}\left\{ \left| D_{pq}^2 D_{jk}^3(s)[Y_s^T U(t,s)R^{1/2}]_{kj} \right|^2 | X_{jk} = W \right\} \right)^{1/2} \\
&\le \frac{C_5(t)}{\sqrt{s_1}},
\end{aligned}$$

where the last inequality follows from Lemma 3.2.2 and Proposition 3.2.4. This ends the proof of (3.2.47).

Estimation of ϵ_2''

Recall that

$$|\epsilon_2''| \le \frac{C_0}{\sqrt{ns_1}} \sum_{p,q}{}' \sup_{|Z|\le \eta_n\sqrt{n}} \left| \frac{\partial}{\partial Z^3} \mathbb{E}\left\{ [Y_{s_1}^T U_{t,s_1} R^{1/2}]_{qp} D_{jk}^3(s)[Y_s^T U(t,s)R^{1/2}]_{kj} | X_{pq} = Z, X_{jk} = W \right\} \right|. \tag{3.2.50}$$

We shall prove that

$$\sup_{\eta_n\sqrt{n}} \frac{|t|}{\sqrt{n}} \int_0^1 |\epsilon_2''| ds_1 \le C_7(t). \tag{3.2.51}$$

Changing the order between the derivation and the integration yields :

$$\begin{aligned}
&\frac{1}{n} \left| \frac{\partial}{\partial Z^3} \mathbb{E}\left\{ [Y_{s_1}^T U(t,s_1)R^{1/2}]_{qp} D_{jk}^3(s)[Y_s^T U(t,s)R^{1/2}]_{kj} | X_{pq} = Z, X_{jk} = W \right\} \right| \\
&= \frac{1}{n} \left| \mathbb{E} \frac{\partial}{\partial Z^3} \left\{ [Y_{s_1}^T U(t,s_1)R^{1/2}]_{qp} D_{jk}^3(s)[Y_s^T U(t,s)R^{1/2}]_{kj} | X_{pq} = Z, X_{jk} = W \right\} \right| \\
&= \frac{1}{n} \left| \mathbb{E} \left\{ \sum_{\ell=0}^3 C_3^\ell \frac{\partial^\ell}{\partial Z^\ell} [Y_{s_1}^T U(t,s_1)R^{1/2}]_{qp} \frac{\partial^\ell}{\partial Z^{3-\ell}} D_{jk}^3(s)[Y_s^T U(t,s)R^{1/2}]_{kj} | X_{pq} = Z, X_{jk} = W \right\} \right| \\
&\overset{(a)}{\le} \frac{C_0}{n^4} \left| \mathbb{E} \left(\sum_{\ell=0}^3 C_\ell(t) \left(\sum_\alpha [Y_{s_1}]_{\alpha q}^2 \right)^{(\ell+1)/2} C_{6-\ell}(t) \left(\sum_\alpha [Y_s]_{\alpha k}^2 \right)^{(7-\ell)/2} | X_{pq} = Z, X_{jk} = W \right) \right| \\
&\overset{(b)}{\le} \frac{C_6(t)}{n^2} \sum_{\ell=0}^3 \left(1 + \left(\frac{|Z|}{\sqrt{n}} \right)^\ell \right) \left(1 + \left(\frac{|W|}{\sqrt{n}} \right)^{7-\ell} \right).
\end{aligned}$$

where (a) follows from Lemma 2 and (b) follows from Proposition 3.2.4.

Plugging the precedent inequality into (3.2.50), we get (3.2.51). Gathering (3.2.42), (3.2.47) and (3.2.51), with (3.2.32), finally we get

$$V_n \le C_3(t)V_n^{1/2} + C_8(t).$$

This implies that

$$V_n \le C_0(1+|t|^4)^2.$$

3.2.5 More interpolations

For future use, we prove the following proposition which will serve in the next section.

Proposition 3.2.6 *Recall $\alpha(t) = u(t) - \hat{u}(t)$. Under Assumptions 4.2.3, 3.1.2 and 3.1.3, we have*

$$|\mathbb{E}\alpha(t)| \le C_4(t)$$

where $C_4(t)$ is a polynomial in $|t|$ of degree 4.

Proof: By (3.2.21), we have

$$\mathbb{E}\alpha(t) \triangleq \mathbf{i}t\int_0^1 (A''_n - B''_n)ds \tag{3.2.52}$$

where

$$\begin{aligned} A''_n &= \frac{1}{\sqrt{ns}}\sum_{j,k}\mathbb{E}\{X_{jk}[Y_s^T U R^{1/2}]_{kj}\}, \\ B''_n &= \frac{1}{\sqrt{n(1-s)}}\sum_{j,k}\mathbb{E}\{\hat{X}_{jk}[Y_s^T U R^{1/2}]_{kj}\}. \end{aligned}$$

By Proposition 3.2.2, we have

$$A''_n = T''_1 + T''_2 + \varepsilon'''_2 \tag{3.2.53}$$

where

$$\begin{aligned} T''_1 &= \frac{\kappa_2}{\sqrt{sn}}\sum_{j,k}^n \mathbb{E}\left\{\frac{\partial}{\partial X_{jk}}[Y_s^T U(t,s)R^{1/2}]_{kj}\right\}, \\ T''_2 &= \frac{\kappa_3}{2\sqrt{sn}}\sum_{j,k}^n \mathbb{E}\left\{\frac{\partial^2}{\partial (X_{jk})^2}[Y_s^T U(t,s)R^{1/2}]_{kj}\right\}, \\ B''_n &= \frac{\kappa_2}{\sqrt{sn}}\sum_{j,k}^n \mathbb{E}\left\{\frac{\partial}{\partial \hat{X}_{jk}}[Y_s^T U(t,s)R^{1/2}]_{kj}\right\}, \\ |\varepsilon'''_2| &\le \frac{C_0\mu_4}{\sqrt{ns}}\sum_{j,k=1}^n \sup_{|W|\le\eta_n\sqrt{n}}\left|\frac{\partial^3}{\partial (X_{jk})^3}\mathbb{E}\left\{[Y_s^T U(t,s)R^{1/2}]_{kj}|X_{jk}=W\right\}\right|. \end{aligned}$$

Then

$$A''_n - B''_n = T''_1 - B''_n + T''_2 + \varepsilon'''_2.$$

As $[Y_s^T U(t,s)R^{1/2}]_{kj}$ is a function of $\mathbf{X}_s$, we have

$$\frac{1}{\sqrt{s}}D_{jk}[Y_s^T U(t,s)R^{1/2}]_{kj} = D_{jk}(s)[Y_s^T U(t,s)R^{1/2}]_{kj} = \frac{1}{\sqrt{1-s}}\hat{D}_{jk}[Y_s^T U(t,s)R^{1/2}]_{kj},$$

which implies

$$T_1'' - B_n'' = 0. \tag{3.2.54}$$

For T_2'', by (3.2.11),

$$\begin{aligned}
T_2'' &= \frac{\sqrt{s}\kappa_3}{2n^{1/2}} \sum_{j,k}^{n} \mathbb{E}\left\{D_{jk}^2(s)[Y_s^T U(t,s)R^{1/2}]_{kj}\right\} \\
&= \frac{3\mathbf{i}\kappa_3\sqrt{s}}{n^{3/2}} \sum_{j,k} \mathbb{E}\left\{[R^{1/2}U(t,s)R^{1/2}]_{jj} * [Y_s^T U(t,s)R^{1/2}]_{kj}\right\} \\
&\quad - \frac{3\kappa_3\sqrt{s}}{n^{3/2}} \sum_{j,k} \mathbb{E}\left\{[Y_s^T U(t,s)Y_s]_{kk} * [Y_s^T U(t,s)R^{1/2}]_{kj} * [R^{1/2}U(t,s)R^{1/2}]_{jj}\right\} \\
&\quad - \frac{\kappa_3\sqrt{s}}{n^{3/2}} \sum_{j,k} \mathbb{E}\left\{[Y_s^T U(t,s)R^{1/2}]_{kj} * [Y_s^T U(t,s)R^{1/2}]_{kj} * [Y_s^T U(t,s)R^{1/2}]_{kj}\right\}.
\end{aligned}$$

By Proposition 3.2.5 with $Z = 1$ and taking two convolutions into accnout, we get

$$|T_2''| \leq C_2(t). \tag{3.2.55}$$

For ε_2''', by (3.2.18),

$$\begin{aligned}
|\varepsilon_2'''| &\leq \frac{C_3(t)}{n^2} \sum_{j,k} \sup_{|W| \leq \eta_n\sqrt{n}} \left| \mathbb{E}\left\{ \left(\sum_{\alpha} [Y_{\alpha k}]^2 \right)^2 | X_{jk} = W \right\} \right| \\
&\leq C_3(t).
\end{aligned} \tag{3.2.56}$$

Gathering (3.2.54), (3.2.55) and (3.2.56) into (3.2.52), we obtain

$$|\mathbb{E}\alpha(t)| \leq C_4(t).$$

□

Another useful interpolation is between the real Gaussian matrices and the complex Gaussian matrices which will serve in the calculation of the bias.

Proposition 3.2.7 *Denote $\hat{\mathbf{X}}_r$ (resp. $\hat{\mathbf{X}}_c$), the standard real (resp. complex) Gaussian random matrix. Consider*

$$\begin{aligned}
\hat{\mathbf{Y}}_r &= \frac{1}{\sqrt{n}} \mathbf{R}_n^{1/2} \hat{\mathbf{X}}_r + \mathbf{A}_n \\
\hat{\mathbf{Y}}_c &= \frac{1}{\sqrt{n}} \mathbf{R}_n^{1/2} \hat{\mathbf{X}}_c + \mathbf{A}_n
\end{aligned}$$

where $\mathbf{R}_n$ and $\mathbf{A}_n$ satisfy Assumption 3.1.3. The relative terms are denoted by $\hat{u}_r(t)$, $\hat{\mathbf{U}}_r(t)$ (resp. $\hat{u}_c(t)$, $\hat{\mathbf{U}}_c(t)$). Then we have

$$|\mathbb{E}\hat{u}_r(t) - \mathbb{E}\hat{u}_c(t)| \leq C_2(|t|).$$

Proof: As before, consider the interpolation matrix

$$\mathbf{X}_s = \sqrt{s}\hat{\mathbf{X}}_r + \sqrt{1-s}\hat{\mathbf{X}}_c, \quad \mathbf{Y}_s = \frac{1}{\sqrt{n}}\mathbf{R}_n^{1/2}\mathbf{X}_s + \mathbf{A}_n,$$

and relative term $\mathbf{U}(t,s)$.

Write

$$\begin{aligned}
\mathbb{E}\hat{u}_r(t) - \mathbb{E}\hat{u}_c(t) &= \int_0^1 \frac{d}{ds}\mathbb{E}\mathrm{Tr}e^{\mathbf{i}t\mathbf{Y}_s\mathbf{Y}_s^*}ds \\
&= \mathbf{i}t\int_0^1 \mathbb{E}\mathrm{Tr}\left[e^{\mathbf{i}t\mathbf{Y}_s\mathbf{Y}_s^*}\left(\frac{1}{2\sqrt{s}}\hat{\mathbf{Y}}_r - \frac{1}{2\sqrt{1-s}}\hat{\mathbf{Y}}_c\right)\mathbf{Y}_s^*\right]ds \\
&\quad + \mathbf{i}t\int_0^1 \mathbb{E}\mathrm{Tr}\left[e^{\mathbf{i}t\mathbf{Y}_s\mathbf{Y}_s^*}\mathbf{Y}_s\left(\frac{1}{2\sqrt{s}}\hat{\mathbf{Y}}_r - \frac{1}{2\sqrt{1-s}}\hat{\mathbf{Y}}_c\right)^*\right]ds.
\end{aligned}$$

The two terms are similar and we shall focus on the first term.

We have

$$\int_0^1 \mathbb{E}\mathrm{Tr}\left[e^{\mathbf{i}t\mathbf{Y}_s\mathbf{Y}_s^*}\left(\frac{1}{\sqrt{s}}\hat{\mathbf{Y}}_r - \frac{1}{\sqrt{1-s}}\hat{\mathbf{Y}}_c\right)\mathbf{Y}_s^*\right]ds \triangleq \int_0^1(\hat{A}(s) - \hat{B}(s))ds$$

where

$$\begin{aligned}
\hat{A}(s) &\triangleq \frac{1}{\sqrt{ns}}\sum_{j,k}\mathbb{E}\{\hat{X}_{r,jk}[Y_s^*U(t,s)R^{1/2}]_{kj}\}, \\
\hat{B}(s) &\triangleq \frac{1}{\sqrt{n(1-s)}}\sum_{j,k}\mathbb{E}\{\hat{X}_{c,jk}[Y_s^*U(t,s)R^{1/2}]_{kj}\}.
\end{aligned}$$

For a complex Gaussian variable X and a function $\Phi(X)$ of class C^1, by integration by parts, we have

$$\mathbb{E}[X\Phi(X)] = \mathbb{E}|X|^2\mathbb{E}\frac{\partial\Phi(X)}{\partial\overline{X}}, \tag{3.2.57}$$

$$\mathbb{E}[\overline{X}\Phi(X)] = \mathbb{E}|X|^2\mathbb{E}\frac{\partial\Phi(X)}{\partial X}. \tag{3.2.58}$$

With the precedent formula, we have

$$\begin{aligned}
\hat{A}(t,s) &\triangleq \frac{1}{\sqrt{ns}}\sum_{j,k}\mathbb{E}\frac{\partial}{\partial\hat{X}_{r,jk}}[Y_s^*U(t,s)R^{1/2}]_{kj}, \\
\hat{B}(t,s) &\triangleq \frac{1}{\sqrt{n(1-s)}}\sum_{j,k}\mathbb{E}\frac{\partial}{\partial\hat{X}_{c,jk}}[Y_s^*U(t,s)R^{1/2}]_{kj}.
\end{aligned}$$

From (3.2.8),

$$\frac{\partial}{\partial\hat{X}_{r,jk}}[Y_s^*U(t,s)R^{1/2}]_{kj}$$

$$= \frac{\sqrt{s}}{\sqrt{n}}[R^{1/2}U(t,s)R^{1/2}]_{jj} + \frac{\mathbf{i}\sqrt{s}}{\sqrt{n}}[Y_s^*U(t,s)Y_s]_{kk} * [R^{1/2}U(t,s)R^{1/2}]_{jj}$$
$$+\frac{\mathbf{i}\sqrt{s}}{\sqrt{n}}[Y_s^*U(t,s)R^{1/2}]_{kj} * [Y_s^*U(t,s)R^{1/2}]_{kj}.$$

The same method yields that

$$\frac{\partial}{\partial \overline{\hat{X}_{c,jk}}}[Y_s^*U(t,s)R^{1/2}]_{kj} = \frac{\sqrt{1-s}}{\sqrt{n}}[R^{1/2}U(t,s)R^{1/2}]_{jj} + \frac{\mathbf{i}\sqrt{1-s}}{\sqrt{n}}[Y_s^*U(t,s)Y_s]_{kk} * [R^{1/2}U(t,s)R^{1/2}]_{jj}.$$

Remark 3.2.4 *As $\mathbf{U}(t,s) = e^{\mathbf{i}t\mathbf{Y}_s\mathbf{Y}_s^*}$, when we derive $\mathbf{U}(t,s)$ according to real Gaussian variable, both $\mathbf{Y}_s$ and $\mathbf{Y}_s^*$ depend on $\hat{X}_{r,jk}$. However, $\mathbf{Y}_s$ does not depend on $\overline{\hat{X}_{c,jk}}$.*

Then

$$\hat{A}(t,s) - \hat{B}(t,s) = \frac{\mathbf{i}}{n}\sum_{j,k}[Y_s^*U(t,s)R^{1/2}]_{kj} * [Y_s^*U(t,s)R^{1/2}]_{kj}.$$

With the same argument as Lemma 3.2.2 and take the convolution into account, (see also the argument for T_{1a} in (3.2.24))

$$|\hat{A}(t,s) - \hat{B}(t,s)| \le C_1(|t|),$$

and we have

$$|\mathbb{E}\hat{u}_r(t) - \mathbb{E}\hat{u}_c(t)| \le C_2(|t|).$$

□

3.3 Proof of Theorem 3.1.1 : CLT for non-analytic functionals

We return back to the model

$$\mathbf{Y}_n = \frac{1}{\sqrt{n}}\mathbf{R}_n^{1/2}\mathbf{X}_n,$$

and shall establish the CLT for the linear statistics $\mathcal{N}(f)$ defined in (3.1.3). Recently, the study of $\mathcal{N}(f)$ when f is analytic is investigated in Najim [70]. Recall the definition of $\mathcal{K}_n(z_1,z_2)$ in (3.1.6) and define

$$\Theta_{0,n} \triangleq \frac{\tilde{t}_n'(z_1)\tilde{t}_n'(z_2)}{(\tilde{t}_n(z_1)-\tilde{t}_n(z_2))^2} - \frac{1}{(z_1-z_2)^2} \quad (3.3.1)$$
$$\Theta_{1,n} \triangleq \frac{\partial^2}{\partial z_1 \partial z_2}\mathcal{K}_n(z_1,z_2). \quad (3.3.2)$$

Recall the following Central Limit Theorem.

Proposition 3.3.1 *[70] Let $k \in \mathbb{N}$ and f be a function which is analytic on an open region containing*

$$[\liminf_n \lambda_{\min}^{\mathbf{R}_n}\mathbb{I}_{0\le c\le 1}(1-\sqrt{c})^2, \limsup_n \lambda_{\max}^{\mathbf{R}_n}(1+\sqrt{c})^2]. \quad (3.3.3)$$

Consider the random vector

$$\mathcal{N}_n^1(\mathbf{f}) = (\mathcal{N}_n^1(f_1), \cdots, \mathcal{N}_n^1(f_k))$$

where $\mathcal{N}_n^1(f)$ is defined in (3.1.4). Define the Gaussian random vector

$$Z_n^1(\mathbf{f}) = (Z_n^1(f_1), \cdots, Z_n^1(f_k))$$

with zero mean and the covariance

$$\mathrm{Cov}(Z_n^1(f_i), Z_n^1(f_j)) = -\frac{1}{4\pi^2} \oint \oint f_i(z_1) f_j(z_2) \Theta_n(z_1, z_2) dz_1 dz_2, \quad \text{for } 1 \le i, j \le k$$

where $\Theta_n = 2\Theta_{0,n} + \kappa_4 \Theta_{1,n}$ and $\Theta_{0,n}$, $\Theta_{1,n}$ are defined in (3.3.1) and (3.3.2) respectively. The contours here are assumed to be non-overlapping and closed, taken in the positive direction in the complex plane and enclosing the interval (3.3.3).

Then, the sequence of $\mathbb{R}^k$-valued random variable $Z_n^1(\mathbf{f})$ is tight and the random vector $\mathcal{N}_n^1(\mathbf{f})$ satisfies :

$$d_{\mathcal{LP}}(\mathcal{N}_n^1(\mathbf{f}), Z_n^1(\mathbf{f})) \xrightarrow[N,n\to\infty]{} 0.$$

Outline of the proof :

By Fourier transform, as f is integrable in $\mathbb{R}$, for $x \in \mathbb{R}$,

$$\hat{f}(t) = \frac{1}{2\pi} \int_{\mathbb{R}} f(x) e^{-\mathbf{i}xt} dx$$

and by the inverse formula, we have

$$f(x) = \int_{\mathbb{R}} \hat{f}(t) e^{\mathbf{i}xt} dt.$$

Then

$$\mathcal{N}_n^1(f) = \int_{\mathbb{R}} \hat{f}(t) \left(\sum_{k=1}^N e^{\mathbf{i}\lambda_k t} - \sum_{k=1}^N \mathbb{E} e^{\mathbf{i}\lambda_k t} \right) dt.$$

It is natural to study the fluctuation of the process $\mathring{u}(t) = \sum_{k=1}^N e^{\mathbf{i}\lambda_k t} - \sum_{k=1}^N \mathbb{E} e^{\mathbf{i}\lambda_k t}$. We will verify all conditions in Lemma 3.1.2. In particular, the condition

$$\mathrm{Var}[\hat{f}(t)\mathring{u}_n(t)] \le \left(C_0(1+|t|^4)|\hat{f}(t)| \right)^2$$

is ensure by Theorem 3.1.3. Then the expression of of the variance will be investigated.

3.3.1 Convergence of the process $\mathcal{N}_n^1(f)$

We start the proof with the convergence of $u_n(x)$. In particular, we will show that all conditions in Lemma 3.1.2 are fulfilled.

Condition (i). Recall $m_n(z) = \frac{1}{N} \sum_{j=1}^N \frac{1}{\lambda_j - z}$ where λ_i are eigenvalues of $\mathbf{Y}\mathbf{Y}^T$ and $M_n^1(z) =$

$N(m_n(z) - \mathbb{E}m_n(z))$. By Cauchy formula,

$$\mathring{u}_n(x) = \sum_{k=1}^{N} e^{\mathbf{i}\lambda_k x} - \mathbb{E}e^{\mathbf{i}\lambda_k x} = \frac{1}{2\mathbf{i}\pi}\oint_{\mathcal{C}_n} e^{\mathbf{i}zx} M_n^1(z)dz$$

where the contour $\mathcal{C}_n$ is a rectangle in trigonometric direction which contains (3.3.3). As $g(z) = e^{\mathbf{i}zx}$ is clearly analytic over $\mathbb{C}$, by Proposition 3.3.1,

$$\forall t_i \in \mathbb{R},\ 1 \le i \le d, \quad d_{\mathcal{LP}}\Big((\mathring{u}_n\,(t_1), \cdots, \mathring{u}_n\,(t_d)), (\psi_n(t_1), \cdots, \psi_n(t_d))\Big) \xrightarrow[n\to\infty]{} 0, \tag{3.3.4}$$

where ψ_n denotes the sequence of centered Gaussian random variables with the covariance

$$\mathrm{Cov}(\psi_n(t_i), \psi_n(t_j)) = -\frac{1}{4\pi^2}\oint\oint e^{\mathbf{i}t_i z_1} e^{\mathbf{i}t_j z_2}\Theta_n(z_1, z_2)dz_1 dz_2, \quad 1 \le i, j \le d.$$

Thus the finite dimensional convergence of $\hat{f}(t)\mathring{u}_n(t)$ is achieved.

Condition (ii). Now we will prove the tightness of $\mathring{u}(x)$ over a certain interval $[-T, T]$. From [17, Theorem 12.3], it suffices to verify there exists $t_0 \in [-T, T]$ such that :

$$\mathbb{E}|\mathring{u}_n(t_0)|^2 < \infty, \quad \text{and} \quad \sup_{n, t_1, t_2 \in [-T,T]} \mathbb{E}\left|\frac{\mathring{u}_n(t_1) - \mathring{u}_n(t_2)}{t_1 - t_2}\right|^2 < \infty.$$

By Proposition 3.2.3 and 3.2.6,

$$\mathbb{E}|\mathring{u}_n(t_0)|^2 \le C_4(|t_0|) + \mathrm{Var}[\hat{u}_n(t_0)] < \infty.$$

Since
$$\frac{\mathring{u}_n(t_1) - \mathring{u}_n(t_2)}{t_1 - t_2} = \frac{1}{2\mathbf{i}\pi}\oint_{\mathcal{C}_n}\left(\frac{e^{\mathbf{i}zt_1} - e^{\mathbf{i}zt_2}}{t_1 - t_2}\right) M_n^1(z)dz,$$

we have

$$\mathbb{E}\left|\frac{\mathring{u}_n(t_1) - \mathring{u}_n(t_2)}{t_1 - t_2}\right|^2 = \frac{1}{4\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2}\left(\frac{e^{\mathbf{i}z_1 t_1} - e^{\mathbf{i}z_1 t_2}}{t_1 - t_2}\right)\left(\overline{\frac{e^{\mathbf{i}z_2 t_1} - e^{\mathbf{i}z_2 t_2}}{t_1 - t_2}}\right)\mathbb{E}\left\{M_n^1(z_1)M_n^1(\bar{z}_2)\right\}dz_1 dz_2,$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are two contours non-overlapping both of which contain the interval (3.3.3).

By a simple calculus, we have

$$\sup_{t_1, t_2 \in [-T,T]}\left|\frac{e^{\mathbf{i}t_1 z_1} - e^{\mathbf{i}t_2 z_1}}{t_1 - t_2}\right| = \sup_{t_1, t_2 \in [-T,T]}\left|z_1\int_0^1 e^{\mathbf{i}t_2 z_1(1-s)} e^{\mathbf{i}t_1 z_1 s} ds\right| < \infty.$$

Moreover, from Proposition 3.3.1,

$$\sup_{z_1 \in \mathcal{C}_1, z_2 \in \mathcal{C}_2}\left|\mathbb{E}M_n^1(z_1)M_n^1(\bar{z}_2)\right| \le \sup_{z_1 \in \mathcal{C}_1, z_2 \in \mathcal{C}_2}\sqrt{\mathrm{Var}M_n^1(z_1)}\sqrt{\mathrm{Var}M_n^1(z_2)} < \infty.$$

Therefore,

$$\sup_{n,t_1,t_2\in[-T,T]} \mathbb{E}\left|\frac{\mathring{u}_n(t_1)-\mathring{u}_n(t_2)}{t_1-t_2}\right|^2 < \infty.$$

The tightness of the process $\mathring{u}(x)$ is achieved which implies Condition (ii).

Condition (iii). It remains to estimate the variance. By the inverse formula of Fourier transform,

$$\sum_{i=1}^{N} f(\lambda_i) = \int_{\mathbb{R}} \hat{f}(x)u_n(x)dx.$$

By Theorem 3.1.3, we have

$$\mathrm{Var}^{1/2}[\hat{f}(t)u_n(t)] \leq C_0|\hat{f}(t)|(1+|t|^4) \tag{3.3.5}$$

which is integrable on $\mathbb{R}$ by hypothesis.

Now we will use the argument of meta-model to show that

$$\mathrm{Var}^{1/2}[\hat{f}(t)\psi_n(t)] \leq C_0|\hat{f}(t)|(1+|t|^4)$$

where ψ_n ia a Gaussian variable (cf. (3.3.4)).

Introduction of the meta-model

The initial idea of meta-model is from [70]. With the extra parameter $M \in \mathbb{N}$, consider the $NM \times nM$ matrix

$$\mathbf{R}_n(M) = \begin{bmatrix} \mathbf{R}_n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{R}_n \end{bmatrix}$$

where $\mathbf{R}_n(M)$ is the block matrix whose diagonal is composed of M matrices $\mathbf{R}_n$ and 0 elsewhere. Consider now the random matrix

$$\mathbf{Y}_n(M) = \frac{1}{\sqrt{nM}}\mathbf{R}_n^{1/2}(M)\mathbf{X}_n(M),$$

where $\mathbf{X}_n(M)$ is an $MN \times Mn$ matrix with i.i.d. random entries with the same distribution as the X_{ij}'s and satisfying (3.1.1).

Denote $M^1_{n,M}(z) = \sum_{j=1}^{NM} \frac{1}{\lambda_j - z} - \sum_{j=1}^{NM} \mathbb{E}\frac{1}{\lambda_j - z}$ where $(\lambda_i)_{1\leq i\leq NM}$ are eigenvalues of $\mathbf{Y}_n(M)\mathbf{Y}_n{}^*(M)$. By a simple calculation of the bloc matrix, $\mathbf{Y}_n(M)\mathbf{Y}_n{}^*(M)$ has the same deterministic equivalent t_n as $\mathbf{Y}_n\mathbf{Y}_n{}^*$. With Proposition 3.3.1, we can show that for n, N fixed, $M^1_{n,M}(z)$ converges to a Gaussian process with

$$\mathrm{Cov}[M^1_{n,M}(z_1), M^1_{n,M}(z_2)] \xrightarrow[M\to\infty]{} \Theta_n(z_1,z_2)$$

where $\Theta(z_1, z2) = 2\Theta_{0,n}(z_1,z_2) + \kappa_4\Theta_{1,n}(z_1,z_2)$.

With these notations at hand, consider the set of matrices :

$$\{\mathbf{R}_n(M) : M \geq 1, \mathbf{R}_n \text{ is an } N \times n \text{ matrix which satisfies Assumption 3.1.3}\}.$$

According to [18, Chapter 2, Section 7], a sequence of Gaussian process is tight if and only if it is relatively compact in distribution. Since $\|\mathbf{R}_n(M)\| = \|\mathbf{R}_n\|$ for all $M \geq 1$, we have

$$\sup_{M\geq 1, N,n\to\infty} \|\mathbf{R}_n(M)\| = \sup_{N,n\to\infty} \|\mathbf{R}_n\| < \infty$$

by Assumption 3.1.3. Then (3.3.5) holds true for $u_n^M(t) = \sum_{j=1}^{NM} e^{\mathbf{i}t\lambda_j}$ where λ_j are eigenvalues of $\mathbf{Y}_n(M)\mathbf{Y}_n^T(M)$:

$$\mathrm{Var}^{1/2}[\hat{f}(t)u_n^M(t)] \leq C_0|\hat{f}(t)|(1+|t|^4).$$

Since $|\hat{f}(t)|(1+|t|^4)$ is supposed integrable in $\mathbb{R}$, in particular $\hat{f}(t)u_n^M(t)$ is tight, hence relatively compact in distribution. As the distribution $\mathcal{L}(\psi_n)$ of the Gaussian process ψ_n is the limit in M of the distribution of u_n^M, belongs to the closure of $\{u_n^M : M \geq 1\}_{n,N\to\infty}$, which is relatively compact. Finally, $\mathcal{L}(\psi_n)$ is included in a compact set, hence is relatively compact. In particular, the family of gaussian processes (ψ_n) is tight and consider $h_K(x) = |x|^2 \wedge K$,

$$\mathbb{E}h_K(\psi_n(t)) = \lim_{M\to\infty} \mathbb{E}h_K(\mathring{u}_n^M(t)) \leq \limsup_{M\to\infty} \mathbb{E}|\mathring{u}_n^M(t)|^2 \leq C_0^2(1+|t|^4)^2.$$

Let $K \to \infty$ and by monotone convergence theorem, we get

$$\mathrm{Var}^{1/2}[\hat{f}(t)\psi_n(t)] \leq C_0|\hat{f}(t)|(1+|t|^4).$$

All conditions in Lemma 3.1.2 being verified, we conclude that

$$d_{\mathcal{LP}}(\mathcal{N}_n^1(\mathbf{f}), Z_n^1(\mathbf{f})) \xrightarrow[N,n\to\infty]{} 0$$

where the variance of $Z_n^1(\mathbf{f})$ will be determined in the next section.

3.3.2 Computation of the covariance

In this section, we will interpret the covariance with the real integral. As f is no more analytic, the existence of the contour integral will not be ensured in this case. By Proposition 3.3.1, and the inverse formula of Fourier transform, the covariance for two functions f_1, f_2 satisfying (3.1.8) is

$$\begin{aligned}
&\mathrm{Cov}\left[\sum_{j=1}^N f_1(\lambda_j), \sum_{j=1}^N f_2(\lambda_j)\right] \\
&= \mathrm{Cov}\left[\int_{\mathbb{R}} \sum_{j=1}^N e^{\mathbf{i}\lambda_j x_1}\hat{f}_1(x_1)dx_1, \int_{\mathbb{R}} \sum_{j=1}^N e^{\mathbf{i}\lambda_j x_2}\hat{f}_2(x_2)dx_2\right] \\
&= -\frac{1}{4\pi^2}\int_{\mathbb{R}}\int_{\mathbb{R}} \hat{f}_1(x_1)\hat{f}_2(x_2)dx_1dx_2 \oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\Theta_n(z_1,z_2)dz_1dz_2 + o(1).
\end{aligned}$$

Then by Proposition 3.3.1, recall the definitions (3.3.1), (3.3.2), we have

$$\begin{aligned}
&\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\Theta_n(z_1,z_2)dz_1dz_2 \\
&= -\frac{1}{2\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\Theta_{0,n}(z_1,z_2)dz_1dz_2 \\
&\quad -\frac{\kappa_4}{4\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\Theta_{1,n}(z_1,z_2)dz_1dz_2. \qquad (3.3.6)
\end{aligned}$$

The first double integral is in [11, Eq. (1.20)] and we recall some elements of the proof. By integration by parts twice, it suffices to study

$$\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\frac{\tilde{t}'(z_1)\tilde{t}'(z_2)}{(\tilde{t}(z_1)-\tilde{t}(z_2))^2}dz_1dz_2 = -\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} x_1x_2e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\log(\tilde{t}(z_1)-\tilde{t}(z_2))dz_1dz_2.$$

Fixing $z_2 \in \mathcal{C}_2$, we deform $\mathcal{C}_1$ and get :

$$\begin{aligned}
&\oint_{\mathcal{C}_1} e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\log(\tilde{t}(z_1)-\tilde{t}(z_2))dz_1 \\
&= \int_{\mathcal{S}_n} e^{\mathbf{i}x_1s_1}e^{\mathbf{i}z_2x_2}\left(\log(\overline{\tilde{t}(s_1)}-\tilde{t}(z_2)) - \log(\tilde{t}(s_1)-\tilde{t}(z_2))\right)ds_1,
\end{aligned}$$

where $\mathcal{S}_n$ is the support of the measure $\tilde{F}_n$ associated with $\tilde{t}_n$. With the same method, we develop the double integral according to z_2, we get

$$\begin{aligned}
&\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\log(\tilde{t}(z_1)-\tilde{t}(z_2))dz_1dz_2 \\
&= \int_{\mathcal{S}_n}\int_{\mathcal{S}_n} e^{\mathbf{i}x_1s_1}e^{\mathbf{i}s_2x_2}\left(\log|\tilde{t}(s_1)-\tilde{t}(s_2)|^2 - \log|\overline{\tilde{t}(s_1)}-\tilde{t}(s_2)|^2\right)ds_1ds_2 \qquad (3.3.7) \\
&= 2\int_{\mathcal{S}_n}\int_{\mathcal{S}_n} e^{\mathbf{i}x_1s_1}e^{\mathbf{i}x_2s_2}[\log|\tilde{t}(s_1)-\tilde{t}(s_2)| - \log|\overline{\tilde{t}(s_1)}-\tilde{t}(s_2)|]ds_1ds_2.
\end{aligned}$$

Then

$$\begin{aligned}
&-\frac{1}{\pi^2}\int_{\mathbb{R}}\int_{\mathbb{R}}\hat{f}_1(x_1)\hat{f}_2(x_2)dx_1dx_2\int_{\mathcal{S}_n}\int_{\mathcal{S}_n}\mathbf{i}x_1\mathbf{i}x_2e^{\mathbf{i}x_1s_1}e^{\mathbf{i}x_2s_2}\log\left|\frac{\tilde{t}(s_1)-\tilde{t}(s_2)}{\overline{\tilde{t}(s_1)}-\tilde{t}(s_2)}\right|ds_1ds_2 \\
&= -\frac{1}{\pi^2}\int_{\mathcal{S}_n}\int_{\mathcal{S}_n}\log\left|\frac{\tilde{t}(s_1)-\tilde{t}(s_2)}{\overline{\tilde{t}(s_1)}-\tilde{t}(s_2)}\right|ds_1ds_2\int_{\mathbb{R}}\int_{\mathbb{R}}\hat{f}_1(x_1)\hat{f}_2(x_2)\mathbf{i}x_1\mathbf{i}x_2e^{\mathbf{i}x_1s_1}e^{\mathbf{i}x_2s_2}dx_1dx_2 \qquad (3.3.8) \\
&= -\frac{1}{\pi^2}\int_{\mathcal{S}_n}\int_{\mathcal{S}_n}f_1'(s_1)f_2'(s_2)\log\left|\frac{\tilde{t}(s_1)-\tilde{t}(s_2)}{\overline{\tilde{t}(s_1)}-\tilde{t}(s_2)}\right|ds_1ds_2 \\
&= \frac{1}{2\pi^2}\int_{\mathcal{S}_n}\int_{\mathcal{S}_n}f_1'(s_1)f_2'(s_2)\log\left(1+\frac{4\mathrm{Im}[\tilde{t}(s_1)]\mathrm{Im}[\tilde{t}(s_2)]}{|\tilde{t}(s_1)-\tilde{t}(s_2)|^2}\right)ds_1ds_2.
\end{aligned}$$

The same method applies to deal with the second integral in (3.3.6) :

$$\begin{aligned}
&\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\Theta_{1,n}(z_1,z_2)dz_1dz_2 \\
&= -x_1x_2\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\mathcal{K}_n(z_1,z_2)dz_1dz_2.
\end{aligned}$$

where $\mathcal{K}_n(z_1,z_2)$ is defined in (3.1.6). It is easy to check that $\tilde{t}_n(\bar{s}) = \overline{\tilde{t}_n(s)}$ and $\mathbf{T}_n(\bar{s}) = \overline{\mathbf{T}_n(s)}$. Fixing z_2 and we get

$$\oint_{\mathcal{C}_1} e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\mathcal{K}_n(z_1,z_2)dz_1 = \int_{\mathcal{S}_n} e^{\mathbf{i}s_1x_1}e^{\mathbf{i}z_2x_2}(\mathcal{K}_n(s_1,z_2)-\mathcal{K}_n^{\dagger}(s_1,z_2))ds_1,$$

where $\mathcal{K}_n^{\dagger}(s_1,z_2)$ is defined in (3.1.7). The development for z_2 yields that

$$\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\mathcal{K}_n(z_1,z_2)dz_1dz_2 = 2\int_{\mathcal{S}_n}\int_{\mathcal{S}_n} e^{\mathbf{i}s_1x_1}e^{\mathbf{i}s_2x_2}\mathrm{Re}(\mathcal{K}_n(s_1,s_2)-\mathcal{K}_n^{\dagger}(s_1,s_2))ds_1ds_2,$$

which leads to

$$\begin{aligned}
&-\frac{\kappa_4}{4\pi^2}\int_{\mathbb{R}}\int_{\mathbb{R}}\hat{f}_1(x_1)\hat{f}_2(x_2)dx_1dx_2\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} e^{\mathbf{i}z_1x_1}e^{\mathbf{i}z_2x_2}\Theta_{1,n}(z_1,z_2)dz_1dz_2 \\
&= -\frac{\kappa_4}{2\pi^2}\int_{\mathcal{S}_n}\int_{\mathcal{S}_n} f_1'(s_1)f_2'(s_2)\mathrm{Re}(\mathcal{K}_n(s_1,s_2)-\mathcal{K}_n^{\dagger}(s_1,s_2))ds_1ds_2.
\end{aligned}$$

These terminate the computation of the expression (3.1.9).

3.4 Proof of Theorem 3.1.2 : Estimation of the bias

In this paragraph, we will calculate the bias. This issue has been studied in several papers as [11, 93, 38]. We first recall the bias calculated in [70] for analytic functions.

Proposition 3.4.1 *Under the same settings as Theorem 3.1.2, consider the deterministic vector*

$$\mathcal{N}_n^2(\mathbf{f}) = (\mathcal{N}_n^2(f_1),\cdots,\mathcal{N}_n^2(f_k))$$

where $\mathcal{N}_n^2(f)$ is defined in (3.1.5) and the vector

$$Z_n^2(\mathbf{f}) \triangleq (Z_n^2(f_1),\cdots,Z_n^2(f_k))$$

where

$$\mathbb{E}Z_n^2(f_j) = \frac{1}{2\mathbf{i}\pi}\oint f_j(z)\mathcal{B}_n(z)dz, \quad \text{for } 1\le j\le k,$$

with $\mathcal{B}_n(z) = \mathcal{B}_{1,n}+\kappa_4\mathcal{B}_{2,n}$, and $\mathcal{B}_{1,n}$, $\mathcal{B}_{2,n}$ being defined in (3.1.10) and (3.1.11) respectively. Then we have

$$\limsup_n |Z_n^2(f)| < \infty$$

and

$$\left\|\mathcal{N}_n^2(\mathbf{f}) - Z_n^2(\mathbf{f})\right\| \xrightarrow[N,n\to\infty]{} 0.$$

Denote the bias

$$\mathcal{N}_n^2(f) = N\mathbb{E}\int f(x)F^{\mathbf{Y}\mathbf{Y}^T}(dx) - N\int_{\mathbb{R}} f(x)F_n(dx),$$

where F_n is the measure associated with deterministic equivalent t_n (cf. (3.1.1)). Then consider

$$\mathcal{N}_n^2(f) \triangleq \mathcal{N}_n^{2\prime}(f) + \mathcal{N}_n^{2\prime\prime}(f) + \mathcal{N}_n^{2\prime\prime\prime}(f)$$

where

$$\mathcal{N}_n^{2\prime}(f) = N\mathbb{E}\int f(x)F^{\mathbf{Y}\mathbf{Y}^T}(dx) - N\mathbb{E}\int f(x)F^{\hat{\mathbf{Y}}_r\hat{\mathbf{Y}}_r^T}(dx) \tag{3.4.1}$$

$$\mathcal{N}_n^{2\prime\prime}(f) = N\mathbb{E}\int f(x)F^{\hat{\mathbf{Y}}_r\hat{\mathbf{Y}}_r^T}(dx) - N\mathbb{E}\int f(x)F^{\hat{\mathbf{Y}}_c\hat{\mathbf{Y}}_c^*}(dx) \tag{3.4.2}$$

$$\mathcal{N}_n^{2\prime\prime\prime}(f) = N\mathbb{E}\int f(x)F^{\hat{\mathbf{Y}}_c\hat{\mathbf{Y}}_c^*}(dx) - N\int f(x)F_n(dx), \tag{3.4.3}$$

and $\hat{\mathbf{Y}}_r$ (resp. $\hat{\mathbf{Y}}_c$) is the matrix associated with standard real (resp. complex) Gaussian entries.
Outline of the proof :

1. For complex Gaussian bias $\mathcal{N}_n^{2\prime\prime\prime}(f)$, following [93], we first establish the inequality

$$|N(\mathbb{E}\hat{m}_c(z) - t(z))| \le \frac{1}{n}C_{12}(|z|)C_{17}(|\mathrm{Im} z|^{-1}),$$

   where $C_\ell(t)$ is the polynomial of degree ℓ on |t| with positive coefficients and $\hat{m}_c$ is the Stieljes transform associated with complex Gaussian entries (cf. Proposition 3.4.2). Then from the idea in [38], for $\varphi \in C_c^{18}(\mathbb{R},\mathbb{R})$ where $C_c^{18}(\mathbb{R},\mathbb{R})$ denotes the set of function of class C^{18} with compact support, we show that

$$|\mathcal{N}_n^{2\prime\prime\prime}(\varphi)| \le \frac{K}{n}\int |(1+D)^{18}\varphi(x)|(1+|x|)^{12}dx,$$

   where the operator $D = \frac{d}{dx}$. Finally we approximate f by a function $\check{f}(x) \in C_c^{18}(\mathbb{R},\mathbb{R})$ to get $\left|\mathcal{N}_n^{2\prime\prime\prime}(f)\right| \xrightarrow[n\to\infty]{} 0$ (cf. Theorem 3.4.1).

2. Recall $u_n(t) = \mathrm{Tr}e^{\mathbf{i}t\mathbf{Y}\mathbf{Y}^T}$ and $\hat{u}_c(t) = \mathrm{Tr}e^{\mathbf{i}t\hat{\mathbf{Y}}_c\hat{\mathbf{Y}}_c^*}$. Using Fourier transform, the expression of the bias $\mathbb{E}u_n(t) - \mathbb{E}\hat{u}_c(t)$ is deduced from Proposition 3.4.1 and 3.4.2. Then Proposition 3.2.6 and 3.2.7 will provide the domination

$$|\mathbb{E}u_n(t) - \mathbb{E}\hat{u}_c(t)| \le |\mathbb{E}u_n(t) - \mathbb{E}\hat{u}_r(t)| + |\mathbb{E}\hat{u}_r(t) - \mathbb{E}\hat{u}_c(t)| \le C_4(|t|),$$

   in order to obtain

$$|\mathcal{N}_n^{2\prime}(f) + \mathcal{N}_n^{2\prime\prime}(f) - B| \xrightarrow[N,n\to\infty]{} 0,$$

where B is the bias to be calculated (cf. (3.4.44)).

3.4.1 Complex Gaussian bias

We start the proof with the estimation $\mathcal{N}_n^{2'''}(f)$. Denote $\hat{m}_c(z) = \frac{1}{N}\mathrm{Tr}(\hat{\mathbf{Y}}_c\hat{\mathbf{Y}}_c^* - z\mathbf{I}_N)^{-1}$ and $\tilde{m}_c(z) = \frac{1}{N}\mathrm{Tr}(\hat{\mathbf{Y}}_c^*\hat{\mathbf{Y}}_c - z\mathbf{I}_n)^{-1}$. In [93, 32, 11], it is known that under complex Gaussian case, the bias

$$N\mathbb{E}\hat{m}_c(z) = Nt(z) + \frac{1}{n}\xi_N(z)$$

where ξ_N is analytic in $\mathbb{C}\backslash\mathbb{R}^+$ and satisfies

$$|\xi_N(z)| \le (|z| + C)^k P(|\mathrm{Im}z|^{-1}), \quad \text{for each } z \in \mathbb{C}^+$$

where C is a constant, $k \in \mathbb{N}$ and P is a polynomial with positive coefficients independent of N. We shall specify the degree of polynomial P and the integer k which are useful in the estimation of the bias. Following the idea in [93], we have

Proposition 3.4.2 *Denote the bias in Gaussian case by $\hat{M}_c^2(z) = N\mathbb{E}\hat{m}_c(z) - Nt(z)$. Then*

$$|\hat{M}_c^2(z)| \le \frac{1}{n}C_{12}(|z|)C_{17}(|\mathrm{Im}z|^{-1}),$$

where $C_\ell(|z|)$ (resp. $C_\ell(|\mathrm{Im}z|^{-1})$) is a polynomial in $|z|$ (resp. $|\mathrm{Im}z|^{-1}$) with positive coefficient of degree ℓ.

Proof: For reason of readability, in the proof, we write $\mathbf{Y}$ instead of $\hat{\mathbf{Y}}_c$ to denote the standard complex Gaussian vector and denote the relative terms in complex Gaussian case by $\mathbf{Q} = (\hat{\mathbf{Y}}_c\hat{\mathbf{Y}}_c^* - z\mathbf{I})^{-1}$ and $\tilde{m}_c = \frac{1}{n}\mathrm{Tr}(\hat{\mathbf{Y}}_c^*\hat{\mathbf{Y}}_c - z\mathbf{I})^{-1}$. We have the derivatives

$$\frac{\partial[Q]_{pq}}{\partial X_{ij}} = -\frac{1}{\sqrt{n}}[QR^{1/2}]_{pi}[Y^*Q]_{jq}, \quad \frac{\partial[Q]_{pq}}{\partial \overline{X}_{ij}} = -\frac{1}{\sqrt{n}}[R^{1/2}Q]_{iq}[QY^*]_{pj}. \tag{3.4.4}$$

and define also intermediate matrices

$$\mathbf{S} = -\left[z(\mathbf{I} + \mathbb{E}\tilde{m}_c(z)\mathbf{R})\right]^{-1}, \tag{3.4.5}$$

$$\tilde{\mathbf{S}} = \tilde{\tau}\mathbf{I}, \tag{3.4.6}$$

and

$$\tilde{\tau} = -\frac{1}{z(1 + \frac{1}{n}\mathbb{E}\mathrm{Tr}\mathbf{RQ})}.$$

From $\mathbf{QYY}^* = \mathbf{I} + z\mathbf{Q}$, we have

$$\begin{aligned}\delta_{p,q} + z\mathbb{E}Q_{pq} &= \mathbb{E}\sum_{i,j}[Q]_{pi}Y_{ij}Y_{jq}^* \\ &= \frac{1}{\sqrt{n}}\mathbb{E}\sum_{i,j}[QR^{1/2}]_{pi}X_{ij}\overline{Y}_{qj}.\end{aligned} \tag{3.4.7}$$

By (3.2.58) with $\Phi(X) = [QR^{1/2}]_{pi}\overline{Y}_{qj}$ and (3.4.4), we have

$$\begin{aligned}\mathbb{E}[QR^{1/2}]_{pi}X_{ij}\overline{Y}_{qj} &= \mathbb{E}\frac{\partial[QR^{1/2}]_{pi}}{\partial \overline{X}_{ij}}\overline{Y}_{qj} + \mathbb{E}\frac{\partial \overline{Y}_{qj}}{\partial \overline{X}_{ij}}[QR^{1/2}]_{pi} \\ &= -\frac{1}{\sqrt{n}}\mathbb{E}[RQ]_{ii}[QY]_{pj}\overline{Y}_{qj} + \frac{\delta_{i,q}}{\sqrt{n}}\mathbb{E}[QR]_{pq}.\end{aligned}$$

Then (3.4.7) can be rewritten as :

$$\mathbb{E}[QYY^*]_{pq} = -\frac{1}{n}\mathbb{E}\{[QYY^*]_{pq}\mathrm{Tr}\mathbf{RQ}\} + \mathbb{E}[QR]_{pq}. \tag{3.4.8}$$

Let

$$\omega = \frac{1}{n}\mathrm{Tr}\mathbf{RQ}.$$

Noticing that $\omega = \mathring{\omega} + \mathbb{E}\omega$, and resolving (3.4.8) according to $\mathbb{E}[QYY^*]_{pq}$, we have

$$\delta_{p,q} + z\mathbb{E}Q_{pq} = \frac{1}{1+\mathbb{E}\omega}\mathbb{E}[QR]_{pq} - \frac{1}{1+\mathbb{E}\omega}\mathbb{E}\{[QYY^*]_{pq}\mathring{\omega}\}.$$

Therefore we get

$$\mathbf{I} + \mathbf{\Delta}(z) = \mathbb{E}\mathbf{Q}(-z\mathbf{I} - z\tilde{\tau}(z)\mathbf{R}) \tag{3.4.9}$$

where

$$\mathbf{\Delta}(z) \quad = \quad \frac{1}{1+\mathbb{E}\omega}\mathbb{E}\left[\mathbf{Q}(z)\mathbf{Y}\mathbf{Y}^*(\omega - \mathbb{E}\omega)\right]. \tag{3.4.10}$$

Lemma 3.4.1 *It holds that*

$$z(\mathbb{E}\tilde{m}_c(z) - \tilde{\tau}(z)) = -\frac{1}{n}\mathrm{Tr}\mathbf{\Delta}(z).$$

Proof: Taking trace in (3.4.9) and using the fact that $\tilde{m}_c(z) = c_n\hat{m}_c - \frac{1-c_n}{z}$, we have

$$\begin{aligned}-\frac{1}{n}\mathrm{Tr}\mathbf{\Delta}(z) &= \frac{z}{n}\mathbb{E}\mathrm{Tr}\mathbf{Q} + \frac{z}{n}\tilde{\tau}(z)\mathbb{E}\mathrm{Tr}\mathbf{QR} + c_n \\ &= z\mathbb{E}\tilde{m}_c(z) - z\tilde{\tau}(z).\end{aligned}$$

□

Then the right hand side in (3.4.9) is

$$\mathbb{E}\mathbf{Q}(-z\mathbf{I} - z\tilde{\tau}(z)\mathbf{R}) = \mathbb{E}\mathbf{Q}\mathbf{S}^{-1} + z(\mathbb{E}\tilde{m}_c(z) - \tilde{\tau}(z))\mathbb{E}\mathbf{Q}(z)\mathbf{R}.$$

Using Lemma 3.4.1 and Eq. (3.4.9), we have

$$\mathbb{E}\mathbf{Q}(z) - \mathbf{S}(z) = \mathbf{\Delta}(z)\mathbf{S}(z) + \frac{1}{n}(\mathrm{Tr}\mathbf{\Delta}(z))\mathbb{E}\mathbf{Q}(z)\mathbf{R}\mathbf{S}(z),$$

which implies that

$$N\mathbb{E}\hat{m}_c(z) - \mathrm{Tr}\mathbf{S}(z) = \mathrm{Tr}\boldsymbol{\Delta}(z)\mathbf{S}(z) + \frac{1}{n}\mathrm{Tr}\boldsymbol{\Delta}(z)\mathrm{Tr}[\mathbb{E}\mathbf{Q}(z)\mathbf{R}\mathbf{S}(z)]. \tag{3.4.11}$$

Now we will show that

Proposition 3.4.3 *For $z \in \mathbb{C}^+$, we have*

$$\begin{aligned} |\mathbb{E}\mathrm{Tr}\mathbf{Q}(z) - \mathrm{Tr}\mathbf{S}(z)| &\leq \frac{1}{n}C_3(|z|)C_7(|\mathrm{Im}z|^{-1}), & (3.4.12)\\ |\mathbb{E}\mathrm{Tr}\mathbf{R}\mathbf{Q}(z) - \mathrm{Tr}\mathbf{R}\mathbf{S}(z)| &\leq \frac{1}{n}C_3(|z|)C_7(|\mathrm{Im}z|^{-1}), & (3.4.13)\\ |n(\tilde{m}_c(z) - \tilde{\tau}(z))| &\leq \frac{1}{n}C_2(|z|)C_5(|\mathrm{Im}z|^{-1}) & (3.4.14) \end{aligned}$$

Proof: We first apply the following preliminary result which is based on Nash-Poincaré inequality (cf. [74, Proposition 2.1.6]).

Lemma 3.4.2 *Consider an $N \times N$ matrix $\mathbf{A}$ satisfying $\|\mathbf{A}\| < \infty$. For all $z \in \mathbb{C}^+$, we have*

$$\mathrm{Var}\left[\mathrm{Tr}\mathbf{Q}(z)\mathbf{A}\right] \leq \|\mathbf{A}\|^2 C_1(|z|)C_4(|\mathrm{Im}z|^{-1}).$$

The proof is in Appendix 3.5.6. We now complete the proof of Proposition 3.4.3. We shall estimate the two terms in (3.4.11). Remark that $\mathrm{Im}[\mathbb{E}\omega] = \frac{1}{2in}[\mathbb{E}\mathrm{Tr}\mathbf{R}\mathbf{Q}(z) - \mathbb{E}\mathrm{Tr}\mathbf{R}\mathbf{Q}(\bar{z})] > 0$, we have

$$\frac{1}{|z(1+\mathbb{E}\omega)|} \leq \frac{1}{|\mathrm{Im}z|}.$$

This along with the inequality $\|\mathbf{Q}(z)\| \leq \frac{1}{|\mathrm{Im}z|}$ lead to

$$\left|\frac{1}{n}\mathrm{Tr}\boldsymbol{\Delta}(z)\mathrm{Tr}[\mathbb{E}\mathbf{Q}(z)\mathbf{R}\mathbf{S}(z)]\right| \leq \frac{K}{|\mathrm{Im}z|^2}\left|\mathrm{Tr}\boldsymbol{\Delta}(z)\right|. \tag{3.4.15}$$

From the expression of $\boldsymbol{\Delta}$ in (3.4.10) and by Lemma 3.4.2 with $\mathbf{A} = \mathbf{I}$ and $\mathbf{A} = \mathbf{R}$, we have

$$\begin{aligned} |\mathrm{Tr}\boldsymbol{\Delta}| &\leq \left|\frac{z}{(1+\mathbb{E}\omega)}\mathbb{E}\left[(\mathrm{Tr}\mathbf{Q} - \mathbb{E}\mathrm{Tr}\mathbf{Q})(\omega - \mathbb{E}\omega)\right]\right| \\ &\leq \frac{|z|^2}{|\mathrm{Im}z|}\left|\mathbb{E}\left[(\mathrm{Tr}\mathbf{Q} - \mathbb{E}\mathrm{Tr}\mathbf{Q})(\omega - \mathbb{E}\omega)\right]\right| \\ &\leq \frac{1}{n}C_3(|z|)C_5(|\mathrm{Im}z|^{-1}) \end{aligned} .$$

Plugging the precedent estimate into (3.4.15), we have

$$\left|\frac{1}{n}\mathrm{Tr}\boldsymbol{\Delta}(z)\mathrm{Tr}[\mathbb{E}\mathbf{Q}(z)\mathbf{R}\mathbf{S}(z)]\right| \leq \frac{1}{n}C_3(|z|)C_7(|\mathrm{Im}z|^{-1}). \tag{3.4.16}$$

With the same approach, since $\|\mathbf{S}(z)\| \leq |\mathrm{Im} z|^{-1}$, we have

$$|\mathrm{Tr}[\boldsymbol{\Delta}(z)\mathbf{S}(z)]| \leq \frac{1}{n} C_3(|z|) C_6(|\mathrm{Im} z|^{-1}). \tag{3.4.17}$$

Gathering (3.4.11), (3.4.16) and (3.4.17), (3.4.12) is proved. The same method applies for (3.4.13). By Lemma 3.4.1 and the expression of $\boldsymbol{\Delta}$ in (3.4.10),

$$\begin{aligned} n|\tilde{m}_c(z) - \tilde{\tau}(z)| &= \left| \frac{1}{z(1+\mathbb{E}\omega)} \mathbb{E}\left[\mathbf{Q}(z)\mathbf{Y}\mathbf{Y}^*(\omega - \mathbb{E}\omega)\right] \right| \\ &\leq \frac{|z|}{|\mathrm{Im} z|} \left|\mathbb{E}\left[(\mathrm{Tr}\mathbf{Q} - \mathbb{E}\mathrm{Tr}\mathbf{Q})(\omega - \mathbb{E}\omega)\right]\right| \\ &\leq \frac{1}{n} C_2(|z|) C_5(|\mathrm{Im} z|^{-1}) \end{aligned}$$

Thus we achieve Proposition 3.4.3.

$\square$

Denote

$$\begin{aligned} \varepsilon_n(z) &= n(\mathbb{E}\tilde{m}_c - \tilde{\tau}) \\ \tilde{\varepsilon}_n(z) &= \mathbb{E}\mathrm{Tr}\mathbf{R}\mathbf{Q} - \mathrm{Tr}\mathbf{R}\mathbf{S}. \end{aligned}$$

We remark that

$$\begin{aligned} n(\mathbb{E}\tilde{m}_c(z) - \tilde{t}(z)) &= n(\tilde{\tau} - \tilde{t}) + \varepsilon_n(z), \\ \mathbb{E}\mathrm{Tr}\mathbf{R}\mathbf{Q} - \mathrm{Tr}\mathbf{R}\mathbf{T} &= \mathrm{Tr}\mathbf{R}\mathbf{S} - \mathrm{Tr}\mathbf{R}\mathbf{T} + \tilde{\varepsilon}_n(z). \end{aligned}$$

Using the expression of $\mathbf{T}$ (resp. $\tilde{\tau}$) in (3.1.2) (resp. (3.4.5)) and the fact that for two matrix $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} - \mathbf{B} = -\mathbf{A}(\mathbf{A}^{-1} - \mathbf{B}^{-1})\mathbf{B}$, we obtain that

$$\begin{bmatrix} n(\mathbb{E}\tilde{m}_c(z) - \tilde{t}(z)) \\ \mathbb{E}\mathrm{Tr}\mathbf{R}\mathbf{Q} - \mathrm{Tr}\mathbf{R}\mathbf{T} \end{bmatrix} = \mathbf{D}_0(z) \begin{bmatrix} n(\mathbb{E}\tilde{m}_c(z) - \tilde{t}(z)) \\ \mathbb{E}\mathrm{Tr}\mathbf{R}\mathbf{Q} - \mathrm{Tr}\mathbf{R}\mathbf{T} \end{bmatrix} + \begin{bmatrix} \varepsilon_n(z) \\ \tilde{\varepsilon}_n(z) \end{bmatrix}, \tag{3.4.18}$$

where

$$\mathbf{D}_0(z) = \begin{bmatrix} 0 & z v_0(z) \\ z\tilde{v}_0(z) & 0 \end{bmatrix}$$

with $v_0(z)$ and $\tilde{v}_0(z)$ defined by

$$\begin{aligned} v_0(z) &= \tilde{\tau}(z)\tilde{t}(z) \\ \tilde{v}_0(z) &= \frac{1}{n}\mathrm{Tr}[\mathbf{R}\mathbf{S}\mathbf{R}\mathbf{T}] = \frac{1}{n}\mathrm{Tr}[\mathbf{R}\mathbf{S}\mathbf{T}\mathbf{R}]. \end{aligned}$$

Then we have

$$(\mathbf{I} - \mathbf{D}_0(z)) \begin{bmatrix} n(\mathbb{E}\tilde{m}_c(z) - \tilde{t}(z)) \\ \mathbb{E}\mathrm{Tr}\mathbf{R}\mathbf{Q} - \mathrm{Tr}\mathbf{R}\mathbf{T} \end{bmatrix} = \begin{bmatrix} \varepsilon_n(z) \\ \tilde{\varepsilon}_n(z) \end{bmatrix}. \tag{3.4.19}$$

Now we will resolve this system. We first evaluate the determinant of the system. We introduce the matrices

$$\mathbf{D}(z) = \begin{bmatrix} 0 & v(z) \\ |z|^2\tilde{v}(z) & 0 \end{bmatrix} \tag{3.4.20}$$

$$\mathbf{D}'(z) = \begin{bmatrix} 0 & v'(z) \\ |z|^2\tilde{v}'(z) & 0 \end{bmatrix} \tag{3.4.21}$$

with $v(z)$, $\tilde{v}(z)$ $v'(z)$ and $\tilde{v}'(z)$ defined by

$$v(z) = \tilde{t}(z)\tilde{t}(\bar{z}) \quad , \quad \tilde{v}(z) = \frac{1}{n}\mathrm{Tr}[\mathbf{R}\mathbf{T}(z)\mathbf{T}(z)^*\mathbf{R}], \tag{3.4.22}$$

$$v'(z) = \tilde{\tau}(z)\tilde{\tau}(\bar{z}) \quad , \quad \tilde{v}'(z) = \frac{1}{n}\mathrm{Tr}[\mathbf{R}\mathbf{S}(z)\mathbf{S}(z)^*\mathbf{R}]. \tag{3.4.23}$$

We have the following proposition which will be proved in Appendix 3.5.7.

Proposition 3.4.4 *There exists a strictly positive constant η such that*

$$\det(\mathbf{I} - \mathbf{D}(z)) \geq \frac{K|\mathrm{Im}z|^4}{(\eta^2 + |z|^2)^2}, \tag{3.4.24}$$

for each $z \in \mathbb{C}^+$. Moreover, there exists an integer N_0 and polynomials $C_{12}(|z|)$ and $C_{16}(|\mathrm{Im}z|^{-1})$ such that for each $N > N_0$,

$$\det(\mathbf{I} - \mathbf{D}'(z)) \geq \frac{K|\mathrm{Im}z|^4}{(\eta^2 + |z|^2)^2}, \tag{3.4.25}$$

for each element of z in the set $\mathcal{E}$ defined by

$$\mathcal{E} = \{z \in \mathbb{C}^+, 1 - \frac{1}{n^2}C_{12}(|z|)C_{16}(|\mathrm{Im}z|^{-1}) > 0\}.$$

Finally, for each $N > N_0$ and $z \in \mathcal{E}$,

$$|\det(\mathbf{I} - \mathbf{D}_0(z))| \geq \sqrt{\det(\mathbf{I} - \mathbf{D}(z))}\sqrt{\det(\mathbf{I} - \mathbf{D}'(z))} \geq \frac{K|\mathrm{Im}z|^4}{(\eta^2 + |z|^2)^2}. \tag{3.4.26}$$

We complete the estimation of the bias. From (3.4.19),

$$n(\mathbb{E}\tilde{m}_c(z) - \tilde{t}(z)) = \frac{1}{\det(\mathbf{I} - \mathbf{D}_0)}[\varepsilon(z) + zv_0\tilde{\varepsilon}(z)].$$

For $z \in \mathcal{E}$, by Proposition 3.4.3 and $|v_0| \leq \frac{1}{|\mathrm{Im}z|^2}$,

$$|\mathbb{E}\tilde{m}_c(z) - \tilde{t}(z)| \leq \frac{1}{n^2}C_8(|z|)C_{13}(|\mathrm{Im}z|^{-1}). \tag{3.4.27}$$

For $z \in \mathbb{C}^+ \backslash \mathcal{E}$, as $1 - \frac{1}{n^2}C_{12}(|z|)C_{16}(|\mathrm{Im}z|^{-1}) > 0$,

$$2 < \frac{2}{n^2}C_{12}(|z|)C_{16}(|\mathrm{Im}z|^{-1}).$$

Then

$$|n(\mathbb{E}\tilde{m}_c(z) - \tilde{t}(z))| \leq \frac{2n}{|\mathrm{Im} z|} \leq \frac{1}{n} C_{12}(|z|) C_{17}(|\mathrm{Im} z|^{-1}). \quad (3.4.28)$$

Combining (3.4.27) and (3.4.28), we have

$$|n(\mathbb{E}\tilde{m}_c(z) - \tilde{t}(z))| \leq \frac{1}{n} C_{12}(|z|) C_{17}(|\mathrm{Im} z|^{-1}).$$

□

Next step consists in applying Proposition 3.4.2 to estimate $\mathcal{N}_n^{2'''}(f)$.

Theorem 3.4.1 *Recall the complex Gaussian bias*

$$\mathcal{N}^{2'''}(f) = N\left(\mathbb{E}\int_{\mathbb{R}} f(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) - \int_{\mathbb{R}} f(x) F_n(dx)\right).$$

For a function f of class C^{18}, we have

$$|\mathcal{N}^{2'''}(f)| \xrightarrow[n\to\infty]{} 0.$$

Proof: The proof is split into two parts :

1. For a function $\varphi \in C_c^{18}(\mathbb{R}, \mathbb{R})$ where $C_c^{18}(\mathbb{R}, \mathbb{R})$ is the set of compactly supported functions of class C^{18}.
2. Approximate f by a function $\check{f} \in C_c^{18}(\mathbb{R}, \mathbb{R})$

Denote $\hat{M}_c(z) = (\mathbb{E}\hat{m}_c(z) - t_n(z))$ and recall the deterministic equivalent t_n with the associated measure F_n. By the inverse Stieljes transform,

$$F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) = \lim_{y\to 0^+}\left[-\frac{1}{\pi}\mathrm{Im}[\hat{m}_c(x+\mathbf{i}y)]dx\right].$$

In particular, we have : For $\varphi \in C_c^{18}(\mathbb{R}, \mathbb{R})$,

$$\mathbb{E}\int_{\mathbb{R}} \varphi(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) = \lim_{y\to 0^+}\left[-\frac{1}{\pi}\mathbb{E}\int_{\mathbb{R}} \mathrm{Im}[\varphi(x)\hat{m}_c(x+\mathbf{i}y)]dx\right], \quad (3.4.29)$$

$$\int_{\mathbb{R}} \varphi(x) F_n(dx) = \lim_{y\to 0^+}\left[-\frac{1}{\pi}\int_{\mathbb{R}} \mathrm{Im}[\varphi(x) t_n(x+\mathbf{i}y)]dx\right]. \quad (3.4.30)$$

Subtracting (3.4.29) and (3.4.30), we have

$$\left|\mathbb{E}\int_{\mathbb{R}} \varphi(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) - \int_{\mathbb{R}} \varphi(x) F_n(dx)\right| \leq \frac{1}{\pi}\limsup_{y\to 0^+}\left|\int_{\mathbb{R}} [\varphi(x)\hat{M}_c(x+\mathbf{i}y)]dx\right|. \quad (3.4.31)$$

For $z \in \mathbb{C}^+$ and $p \in \mathbb{N}$, define

$$I_p(z) = \frac{1}{(p-1)!}\int_0^\infty \hat{M}_c(z+t) t^{p-1} e^{-t} dt. \quad (3.4.32)$$

Note that $I_p(z)$ is well defined because $\hat{M}_c(z)$ is uniformly bounded when $z \in \mathbb{C}^+$. Also, it is easy to check that $I_p(z)$ is an analytic function of z and the first derivative is given by

$$I_p'(z) = \frac{1}{(p-1)!}\int_0^\infty \hat{M}_c'(z+t)t^{p-1}e^{-t}dt.$$

Then by partial integration we get

$$\begin{aligned} I_1'(z) &= [\hat{M}_c(z+t)e^{-t}]_0^\infty + \int_0^\infty \hat{M}_c(z+t)e^{-t}dt \\ &= -\hat{M}_c(z) + I_1(z). \end{aligned}$$

In the same way for $p \geq 2$,

$$\begin{aligned} I_p'(z) &= \frac{1}{(p-1)!}\int_0^\infty \hat{M}_c'(z+t)t^{p-1}e^{-t}dt \\ &= -\frac{1}{(p-1)!}\int_0^\infty \hat{M}_c(z+t)((p-1)t^{p-2} - t^{p-1})e^{-t}dt \\ &= -I_{p-1}(z) + I_p(z). \end{aligned} \tag{3.4.33}$$

We have then :

$$\begin{aligned} \int_{\mathbb{R}} \varphi(x)\hat{M}_c(x+\mathbf{i}y)dx &= \int_{\mathbb{R}} \varphi(x)I_1(x+\mathbf{i}y) - \int_{\mathbb{R}} \varphi(x)I_1'(x+\mathbf{i}y)dx \\ &= \int_{\mathbb{R}} \varphi(x)I_1(x+\mathbf{i}y) - \int_{\mathbb{R}} \varphi'(x)I_1(x+\mathbf{i}y)dx \\ &= \int_{\mathbb{R}} ((1+D)\varphi)(x)I_1(x+\mathbf{i}y)dx, \end{aligned}$$

where $D = \frac{d}{dx}$. Using (3.4.33), we iterate the precedent procedure and we obtain : for all $p \geq 1$,

$$\int_{\mathbb{R}} \varphi(x)\hat{M}_c(x+\mathbf{i}y) = \int_{\mathbb{R}} ((1+D)^p\varphi)(x)I_p(x+\mathbf{i}y)dx.$$

Hence by (3.4.31), we have for all $p \in \mathbb{N}$;

$$\left|\mathbb{E}\int_{\mathbb{R}} \varphi(x)F^{\hat{\mathbf{Y}}_c\hat{\mathbf{Y}}_c^*}(dx) - \int_{\mathbb{R}} \varphi(x)F_n(dx)\right| \leq \frac{1}{\pi}\limsup_{y\to 0^+}\left|\int_{\mathbb{R}} ((1+D)^p\varphi)(x)I_p(x+\mathbf{i}y)dx\right|. \tag{3.4.34}$$

Next, we will use Proposition 3.4.2 to show that for $p = 18$, and $z \in \mathbb{C}^+$, one has

$$|I_{18}(z)| \leq \frac{C_{12}(|z|)}{n^2} \tag{3.4.35}$$

where $C_{12}(|z|)$ is a polynomial on $|z|$ of degree 12.

We fix $w \in \mathbb{C}^+$. To prove (3.4.35), we apply Cauchy integral theorem to the function

$$F(z) = \frac{1}{17!}\hat{M}_c(w+z)z^{17}e^{-z},$$

which is analytic in the half-plane $\mathrm{Im}z > -\mathrm{Im}w$. Hence for $r > 0$,

$$\int_{[0,r]} F(z)dz + \int_{[r,r+\mathbf{i}r]} F(z)dz + \int_{[r+\mathbf{i}r,0]} F(z)dz = 0,$$

where $[\alpha, \beta]$ denotes the line segment connecting α, β in $\mathbb{C}$ oriented from α to β.

Then as $|\hat{M}_c(z+w)| \leq \frac{1}{|\mathrm{Im}w|}$ for $z \in \mathbb{C}^+$, we get

$$\begin{aligned}\left|\int_{[r,r+\mathbf{i}r]} F(z)dz\right| &\leq \frac{2}{17!|\mathrm{Im}w|}\int_0^r |r+\mathbf{i}t|^{17}e^{-r}dt \\ &\leq \frac{2}{17!|\mathrm{Im}w|}(2r)^{17}re^{-r} \xrightarrow[r\to 0]{} 0.\end{aligned}$$

Therefore,

$$\begin{aligned}I_{18}(z) &= \frac{1}{17!}\int_0^\infty \hat{M}_c(z+t)t^{17}e^{-t}dt \\ &= \lim_{r\to\infty}\int_{[0,r]} F(t)dt \\ &= \lim_{r\to\infty}\int_{[0,r+\mathbf{i}r]} F(t)dt \\ &= \frac{1}{17!}\int_0^\infty \hat{M}_c(z+(1+\mathbf{i})t)[(1+\mathbf{i})t]^{17}e^{-(1+\mathbf{i})t}(1+\mathbf{i})dt.\end{aligned}$$

By Proposition 3.4.3,

$$|\hat{M}_c(z)| \leq \frac{1}{n^2}C_{12}(|z|)C_{17}(|\mathrm{Im}z|^{-1}).$$

Then

$$\begin{aligned}|I_{18}(z)| &\leq \frac{1}{n^2}\int_0^\infty C_{12}(|z+(1+\mathbf{i})t|)C_{17}(|\mathrm{Im}z+t|^{-1})(\sqrt{2}t)^{17}e^{-t}\sqrt{2}dt \\ &\leq \frac{1}{n^2}\int_0^\infty C_{12}(|z|)e^{-t}\sqrt{2}dt \\ &\leq \frac{C_{12}(|z|)}{n^2}.\end{aligned}$$

This proves (3.4.35). Combining (3.4.34) and (3.4.35), we have

$$\begin{aligned}\left|\mathbb{E}\int_{\mathbb{R}} \varphi(x)F^{\hat{\mathbf{Y}}_c\hat{\mathbf{Y}}_c^*}(dx) - \int_{\mathbb{R}} \varphi(x)F_n(dx)\right| &\leq \frac{1}{\pi}\limsup_{y\to 0^+}\left|\int_{\mathbb{R}}((1+D)^{18}\varphi)(x)I_{18}(x+\mathbf{i}y)dx\right| \\ &\leq \frac{K}{n^2}\int_{\mathbb{R}} |(1+D)^{18}\varphi(x)|(1+|x|)^{12}dx.\end{aligned}$$

In particular, for any function $\varphi(x) \in C_c^{18}(\mathbb{R},\mathbb{R})$,

$$\left|N\mathbb{E}\int_{\mathbb{R}} \varphi(x)F^{\hat{\mathbf{Y}}_c\hat{\mathbf{Y}}_c^*}(dx) - N\int_{\mathbb{R}} \varphi(x)F_n(dx)\right| \xrightarrow[n,N\to\infty]{} 0. \tag{3.4.36}$$

Next task is to extend (3.4.36) for a function f of class C^{18}. Let $\varepsilon > 0$. By [11, Appendix], there exists a compact K_1 such that for all $k \in \mathbb{N}$,

$$\mathbb{P}(\min_i d(\hat{\lambda}_i, K_1) \geq \varepsilon) = o(n^{-k}), \qquad \hat{\lambda}_i \text{ eigenvalues of } \hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*. \tag{3.4.37}$$

From [9], the support of the limiting measure t_n is a compact which is denoted by K_2. Now we take a compact interval $\mathcal{K}$ such that

$$K_1 \cup K_2 \subset \mathcal{K}.$$

Consider a function $\check{f} \in C_c^{18}(\mathbb{R}, \mathbb{R})$ defined by

$$\check{f}(x) = f(x), \quad \text{when } x \in \mathcal{K}.$$

As $\check{f}(x) \in C_c^{18}(\mathbb{R}, \mathbb{R})$, by (3.4.36), we have

$$\left| N \int_{\mathbb{R}} \mathbb{E}\check{f}(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) - N \int_{\mathbb{R}} \check{f}(x) F_n(dx) \right| \xrightarrow[N,n\to\infty]{} 0.$$

The proof of Theorem 3.4.1 is completed if one can show

$$\begin{aligned} &\left| N \int_{\mathbb{R}} \mathbb{E}f(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) - N \int_{\mathbb{R}} \mathbb{E}\check{f}(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) \right. \\ &\qquad \left. - N \int_{\mathbb{R}} f(x) F_n(dx) + N \int_{\mathbb{R}} \check{f}(x) F_n(dx) \right| \xrightarrow[N,n\to\infty]{} 0. \end{aligned} \tag{3.4.38}$$

In particular, we will show that

$$\left| N\mathbb{E} \int_{\mathbb{R}} f(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) - N\mathbb{E} \int_{\mathbb{R}} \check{f}(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) \right| \xrightarrow[N,n\to\infty]{} 0, \tag{3.4.39}$$

$$N \int_{\mathbb{R}} f(x) F_n(dx) - N \int_{\mathbb{R}} \check{f}(x) F_n(dx) = 0. \tag{3.4.40}$$

To prove (3.4.39), we have

$$\begin{aligned} &\left| N\mathbb{E} \int_{\mathbb{R}} f(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) - N\mathbb{E} \int_{\mathbb{R}} \check{f}(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) \right| \\ &\leq \left| N\mathbb{E} \int_{\mathcal{K}} f(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) - N\mathbb{E} \int_{\mathcal{K}} \check{f}(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) \right| \\ &\quad + \left| N\mathbb{E} \int_{\mathbb{R}\backslash\mathcal{K}} \check{f}(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) \right| + \left| N\mathbb{E} \int_{\mathbb{R}\backslash\mathcal{K}} f(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) \right| \\ &\leq \left| N\mathbb{E} \int_{\mathbb{R}\backslash\mathcal{K}} \check{f}(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) \right| + \left| N\mathbb{E} \int_{\mathbb{R}\backslash\mathcal{K}} f(x) F^{\hat{\mathbf{Y}}_c \hat{\mathbf{Y}}_c^*}(dx) \right|, \end{aligned}$$

and

$$\left|N\mathbb{E}\int_{\mathbb{R}\backslash\mathcal{K}} f(x)F^{\hat{\mathbf{Y}}_c\hat{\mathbf{Y}}_c^*}(dx)\right| = \left|N\mathbb{E}\int_{\mathbb{R}} f(x)\mathbb{I}_{x\notin\mathcal{K}}F^{\hat{\mathbf{Y}}_c\hat{\mathbf{Y}}_c^*}(dx)\right|$$
$$\leq N^2\|f\|_\infty\mathbb{P}(\hat{\lambda}_{\max}\notin\mathcal{K})$$
$$\overset{(a)}{\leq} Kn^{-1}$$

where $\hat{\lambda}_{\max}$ is the largest eigenvalue of $\hat{\mathbf{Y}}_c\hat{\mathbf{Y}}_c^*$ and (a) follows from (3.4.37) and the fact that f is bounded in $\mathbb{R}$ (cf. Remark 3.1.1). Same argument holds true for $\check{f}$ and this proves (3.4.39).

From the definition of K_2,

$$\int_{\mathbb{R}} f(x)F_n(dx) - \int_{\mathbb{R}} \check{f}(x)F_n(dx) = \int_{\mathcal{K}} f(x)F_n(dx) - \int_{\mathcal{K}} \check{f}(x)F_n(dx) = 0.$$

(3.4.40) is proved.

This ends the proof of Theorem 3.4.1. □

3.4.2 Bias for non-analytic functionals

Now we can terminate the computations of the bias. Recall that

$$\mathcal{N}_n^2(f) = N\mathbb{E}\int f(x)F^{\mathbf{YY}^T}(dx) - N\int_{\mathbb{R}} f(x)F_n(dx) = \mathcal{N}_n^{2\prime}(f) + \mathcal{N}_n^{2\prime\prime}(f) + \mathcal{N}_n^{2\prime\prime\prime}(f),$$

where $\mathcal{N}_n^{2\prime}(f)$ (resp. $\mathcal{N}_n^{2\prime\prime}(f)$, $\mathcal{N}_n^{2\prime\prime\prime}(f)$) is defined in (3.4.1) (resp. (3.4.2) ,(3.4.3)). From Theorem 3.4.1,

$$|\mathcal{N}_n^{2\prime\prime\prime}(f)| \xrightarrow[n\to\infty]{} 0. \tag{3.4.41}$$

By Fourier transform,

$$\mathcal{N}_n^{2\prime}(f) + \mathcal{N}_n^{2\prime\prime}(f) = \mathbb{E}\int f(x)F^{\mathbf{YY}^T}(dx) - \mathbb{E}\int f(x)F^{\hat{\mathbf{Y}}_c\hat{\mathbf{Y}}_c^*}(dx)$$
$$= \mathbb{E}\int \hat{f}(x)u(x)dx - \mathbb{E}\int \hat{f}(x)\hat{u}_c(x)dx.$$

From Proposition 3.4.1, for $z\in\mathbb{C}^+$,

$$|N(\mathbb{E}m_n(z) - t_n(z)) - \mathcal{B}_n| \xrightarrow[n\to\infty]{} 0, \tag{3.4.42}$$

where $\mathcal{B}_n$ is defined in Proposition 3.4.1. By Proposition 3.4.3, the complex Gaussian bias satisfies

$$|N(\mathbb{E}\hat{m}_c(z) - t_n(z))| \xrightarrow[n\to\infty]{} 0. \tag{3.4.43}$$

(3.4.42) and (3.4.43) imply that

$$|N(\mathbb{E}m_n(z) - \mathbb{E}\hat{m}_c(z)) - \mathcal{B}_n| \xrightarrow[n\to\infty]{} 0.$$

As $z \mapsto e^{\mathbf{i}tz}$ is analytic and $\limsup_n |\mathcal{B}_n| < \infty$, by Propostion 3.4.1 we have then

$$\left| \mathbb{E}u_n(t) - \mathbb{E}\hat{u}_c(t) - \frac{1}{2\mathbf{i}\pi} \oint e^{\mathbf{i}zt}\mathcal{B}_n(z)dz \right| \xrightarrow[n\to\infty]{} 0.$$

Moreover, by Proposition 3.2.6 and 3.2.7,

$$\begin{aligned} |\hat{f}(t)(\mathbb{E}u_n(t) - \mathbb{E}\hat{u}_c(t))| &\leq |\hat{f}(t)(\mathbb{E}u_n(t) - \mathbb{E}\hat{u}_r(t))| + |\hat{f}(t)(\mathbb{E}\hat{u}_r(t) - \mathbb{E}\hat{u}_c(t))| \\ &\leq |\hat{f}(t)|C_4(|t|) \end{aligned}$$

which is integrable. By the argument of meta-model (see Section 3.3.1), we can show that

$$\left| \hat{f}(t) \oint e^{\mathbf{i}zt}\mathcal{B}_n(z)dz \right| \leq |\hat{f}(t)|C_4(|t|),$$

which is integrable in $\mathbb{R}$.

The dominated convergence theorem concludes that

$$\left| \mathcal{N}_n^{2\prime}(f) + \mathcal{N}_n^{2\prime\prime}(f) - \frac{1}{2\mathbf{i}\pi} \int \oint \hat{f}(t)e^{\mathbf{i}zt}\mathcal{B}_n(z)dzdt \right| \xrightarrow[n\to\infty]{} 0. \tag{3.4.44}$$

(3.4.41) and (3.4.44) yield that

$$\left| \mathcal{N}_n^{2}(f) - \frac{1}{2\mathbf{i}\pi} \int \oint \hat{f}(t)e^{\mathbf{i}zt}\mathcal{B}_n(z)dzdt \right| \xrightarrow[n\to\infty]{} 0.$$

The last task is to calculate $\int \oint_{\mathcal{C}} \hat{f}(t)e^{\mathbf{i}zt}\mathcal{B}_n(z)dzdt$. From [11, Section 5] and Proposition 3.4.1, $\mathcal{B}_n(z)$ is uniformly bounded over the contour $\mathcal{C}$. With a similar method in Section 3.3.2,

$$\frac{1}{2\mathbf{i}\pi} \oint e^{\mathbf{i}zt}\mathcal{B}_n(z)dz = \frac{1}{\pi} \int_{\mathcal{S}_n} e^{\mathbf{i}st}\mathrm{Im}(\mathcal{B}_n(s))ds,$$

where $\mathcal{B}_n(s)$ is defined in (3.1.10).

We conclude that

$$\frac{1}{2\mathbf{i}\pi} \int \oint \hat{f}(t)e^{\mathbf{i}zt}\mathcal{B}_n(z)dzdt = \frac{1}{\pi} \int_{\mathcal{S}_n} f(t)\mathrm{Im}(\mathcal{B}_n(t))dt.$$

This terminates the computations of the bias.

3.5 Appendices

3.5.1 Proof of Lemma 3.1.1

We first recall the following proposition in [50].

Proposition 3.5.1 *[50, Theorem 4.28] Let ξ, ξ_n, η^T and η_n^T be random elements in a metric space (S, ρ) such that $\eta_n^T \xrightarrow{\mathcal{D}} \eta^T$ as $n \to \infty$ for fixed T, and moreover $\eta^T \xrightarrow{\mathcal{D}} \xi$ when $T \to \infty$.*

Then $\xi_n \to \xi$ holds under the further condition

$$\lim_{T\to\infty} \limsup_{n\to\infty} \mathbb{E}[\rho(\eta_n^T, \xi_n) \wedge 1] = 0. \tag{3.5.1}$$

Taking

$$\eta_n^T = \int_{-T}^{T} \mathring{\varphi}_n(t)dt, \quad \xi_n = \int_{\mathbb{R}} \mathring{\varphi}_n(t)dt.$$

As $\mathring{\varphi}_n(t) \to \psi(t)$ and $\mathring{\varphi}_n(t)$ is tight for all $t \in [-T, T]$,

$$\eta_n^T \xrightarrow[n\to\infty]{\mathcal{D}} \eta^T \triangleq \int_{-T}^{T} \psi(t)dt.$$

Since $\mathrm{Var}^{1/2}[\psi(t)] \le K(t)$ where $K(t)$ is integrable over $\mathbb{R}$,

$$\int_{-T}^{T} \psi(t)dt \xrightarrow[T\to\infty]{\mathcal{D}} \int_{\mathbb{R}} \psi(t)dt.$$

Now we verify Condition (3.5.1) :

$$\eta_n^T - \xi_n - \mathbb{E}(\eta_n^T - \xi_n) = \int_T^{\infty} \mathring{\varphi}_n(x)dx + \int_{-\infty}^{-T} \mathring{\varphi}_n(x)dx.$$

The two terms are similar and we treat the first term. By Cauchy-Schwarz,

$$\begin{aligned}\mathbb{E}\left[\left|\int_T^{\infty} \mathring{\varphi}_n(x)dx\right|^2\right] &= \mathbb{E}\left[\int_T^{\infty}\int_T^{\infty} \mathring{\varphi}_{1n}(x)\overline{\mathring{\varphi}}_n(x_2)dx_1dx_2\right] \\ &\le \left(\int_T^{\infty} \mathrm{Var}^{1/2}(\mathring{\varphi}_n(x))dx\right)^2.\end{aligned}$$

Since $\mathrm{Var}^{1/2}(\mathring{\varphi}_n(x)) \le K(t)$ which is integrable in $\mathbb{R}$, we have

$$\lim_{T\to\infty} \limsup_{n\to\infty} \mathbb{E}\left[\left|\int_T^{\infty} \mathring{\varphi}_n(x)dx\right|^2\right] = 0.$$

All conditions in Proposition 3.5.1 are satisfied and we conclude that

$$\int \mathring{\varphi}(x)dx \xrightarrow[n\to\infty]{\mathcal{D}} \int \psi(t)dt.$$

3.5.2 Proof of Lemma 3.2.1

For an $N\times n$ matrix $\mathbf{Y}$, By applying Duhamel formula (cf. Proposition 3.2.1) with $U(\mathbf{Y}+\mathbf{H}) = \exp(\mathbf{i}t(\mathbf{Y}+\mathbf{H})(\mathbf{Y}+\mathbf{H})^T)$ and $U(\mathbf{Y}) = \exp(\mathbf{i}t\mathbf{Y}\mathbf{Y}^T)$, we have :

$$U(\mathbf{Y}+\mathbf{H}) - U(\mathbf{Y}) = i\int_0^t \exp(\mathbf{i}(t-s)\mathbf{Y}\mathbf{Y}^T)(\mathbf{Y}\mathbf{H}^T + \mathbf{H}\mathbf{Y}^T + \mathbf{Y}\mathbf{Y}^T)\exp(\mathbf{i}s(\mathbf{Y}+\mathbf{H})(\mathbf{Y}+\mathbf{H})^T))ds.$$

Taking $\mathbf{H} = h\mathbf{E}_{jk}$ where $\mathbf{E}_{jk}$ is the elementary matrix, we get :

$$D_{jk}[U(t)]_{mn} = \mathbf{i}([Y^T U(t)]_{km} * [U(t)]_{jn} + [U(t)]_{mj} * [Y^T U(t)]_{kn}).$$

The rest can be calculated similarly.

3.5.3 Proof of Lemma 3.2.2

It is worth noticing that for a unitary matrix $\mathbf{U}$, we have for $\alpha_1 \neq \alpha_2$,

$$\sum_k U_{k\alpha_1} \bar{U}_{k\alpha_2} = 0,$$

and

$$\sum_k |U_{k\alpha_1}|^2 = 1.$$

We first prove (3.2.17) :

$$\begin{aligned}
|[Y^T U(t) Y]_{kk}(t)| &\leq |\sum_{\alpha_1} Y_{\alpha_1 k} [U(t)Y]_{\alpha_1 k}| \\
&\leq \left(\sum_{\alpha_1} (Y_{\alpha_1 k})^2\right)^{1/2} \left(\sum_{\alpha_1} |[U(t)Y]_{\alpha_1 k}|^2\right)^{1/2} \\
&\leq \left(\sum_{\alpha_1} (Y_{\alpha_1 k})^2\right)^{1/2} \left(\sum_{\alpha_1,\alpha_2,\alpha_3} [U(t)]_{\alpha_1\alpha_2} [\bar{U}(t)]_{\alpha_1\alpha_3} Y_{\alpha_2 k} Y_{\alpha_3 k}\right)^{1/2} \\
&\leq \left(\sum_{\alpha_1} (Y_{\alpha_1 k})^2\right)^{1/2} \left(\sum_{\alpha_1,\alpha_2=\alpha_3} |U_{\alpha_1\alpha_2}|^2 (Y_{\alpha_2 k})^2\right)^{1/2} \\
&\leq \left(\sum_{\alpha} (Y_{\alpha k})^2\right).
\end{aligned}$$

For (3.2.18), when $\ell = 0$, it is a simple application of Cauchy-Schwarz. By using (3.2.7)-(3.2.10) and iterating the derivatives, along with the precedent estimate, we have (3.2.18).

(3.2.20) is a direct consequence of (3.2.9) and (3.2.18).

Now we will prove (3.2.19). For $q \neq k$, by (3.2.7)-(3.2.10),

$$D_{pq} D^b_{jk} [Y^T U(t) R^{1/2}]_{kj} = \mathbf{i} D^b_{jk} ([R^{1/2} U(t) R^{1/2}]_{jp} * [Y^T U(t) Y]_{qk} + [Y^T U(t) R^{1/2}]_{qj} * [Y^T U(t) R^{1/2}]_{kp}).$$

Then noticing that all terms are composed of $[Y^T U(t) R^{1/2}]_{a_1 b_1}$, $[Y^T U(t) Y^{1/2}]_{a_2 b_2}$ and $[R^{1/2} U(t) R^{1/2}]_{a_3 b_3}$, by (3.2.18), we have

$$|D_{pq} D^b_{jk} [Y^T U(t) R^{1/2}]_{kj}| \leq \frac{C_{b+1}(t)}{n^{(1+b)/2}} \left(\sum_\alpha (Y_{\alpha k})^2\right)^{(b+1)/2} \left(\sum_\alpha (Y_{\alpha q})^2\right)^{1/2}.$$

If $q = k$,

$$\begin{aligned}D_{pk}D^b_{jk}[U(t)Y]_{jk} =& \mathbf{i}D^b_{jk}([R^{1/2}U(t)R^{1/2}]_{jp} * [Y^TU(t)Y]_{kk} + [Y^TU(t)R^{1/2}]_{qj} * [Y^TU(t)R^{1/2}]_{kp} \\ &+ [R^{1/2}U(t)R^{1/2}]_{jp}).\end{aligned}$$

Then we get

$$|D_{pq}D^b_{jk}[R^{1/2}U(t)R^{1/2}]_{kj}| \leq \frac{C_{b+1}(t)}{n^{(1+b)/2}} \left(\sum_\alpha (Y_{\alpha k})^2\right)^{(b+2)/2}.$$

We obtain that

$$|D_{pq}D^b_{jk}[Y^TU(t)R^{1/2}]_{kj}| \leq \frac{C_{b+1}(t)}{n^{(1+b)/2}} \left(\sum_\alpha (Y_{\alpha k})^2\right)^{(b+1)/2} \left(\sum_\alpha (Y_{\alpha q})^2\right)^{1/2}.$$

With the same method, we can easily extend the result for all $a \in \mathbb{N}$,

$$|D^a_{pq}D^b_{jk}[Y^TU(t)R^{1/2}]_{kj}| \leq \frac{C_{a+b}(t)}{n^{(a+b)/2}} \left(\sum_\alpha (Y_{\alpha k})^2\right)^{(b+1)/2} \left(\sum_\alpha (Y_{\alpha q})^2\right)^{a/2}. \tag{3.5.2}$$

3.5.4 Proof of Proposition 3.2.5

We begin with (3.2.26),

$$\begin{aligned}
&\left|\mathbb{E}\sum_{jk}[R^{1/2}U(t_1,s)R^{1/2}]_{jj}[Y^TU(t_2,s)R^{1/2}]_{kj}Z\right| \\
\leq\ & \mathbb{E}^{1/2}\left|\sum_{jk}[R^{1/2}U(t_1,s)R^{1/2}]_{jj}[Y_s^TU(t_2,s)R^{1/2}]_{kj}\right|^2 K_Z \\
\leq\ & \mathbb{E}^{1/2}\left|\sum_{j}[R^{1/2}U(t_1,s)R^{1/2}]_{jj}\left(\sum_k[Y_s^TU(t_2,s)R^{1/2}]_{kj}\right)\right|^2 K_Z \\
\leq\ & \mathbb{E}^{1/2}\sum_j\left|[R^{1/2}U(t_1,s)R^{1/2}]_{jj}\right|^2 \mathbb{E}^{1/2}\sum_j\left|\left(\sum_k[Y_s^TU(t_2,s)R^{1/2}]_{kj}\right)\right|^2 K_Z \\
\leq\ & C_0\sqrt{n}\mathbb{E}^{1/2}\left[\sum_j\left|\sum_k[Y^TU(t_2,s)R^{1/2}]_{kj}\right|^2\right] K_Z \\
\leq\ & C_0\sqrt{n}\mathbb{E}^{1/2}\left[\sum_j\left|\sum_{k,\alpha}[R^{1/2}U(t_2,s)]_{j\alpha}Y_{\alpha k}\right|^2\right] K_Z \\
\leq\ & C_0\sqrt{n}\mathbb{E}^{1/2}\left[\sum_{k_1,\alpha_1,k_2,\alpha_2}\left(\sum_j[R^{1/2}U(t_2,s)]_{j\alpha_1}[R^{1/2}\bar{U}(t_2,s)]_{j\alpha_2}\right)|[Y_s]_{\alpha_1k_1}[Y_s]_{\alpha_2k_2}|\right] K_Z \\
\leq\ & C_0\sqrt{n}\mathbb{E}^{1/2}\left[\sum_{k_1,\alpha_1,k_2,\alpha_2} R_{\alpha_1\alpha_2}|[Y_s]_{\alpha_1k_1}[Y_s]_{\alpha_2k_2}|\right] K_Z
\end{aligned}$$

$$\begin{aligned}
&\le C_0\sqrt{n}\mathbb{E}^{1/2}\left(\sum_{k_1,\alpha_1,k_2}|[Y_s]_{\alpha_1k_1}[Y_s]_{\alpha_1k_2}|\right)K_Z\\
&\le C_0n^{3/2}K_Z.
\end{aligned}$$

Now we estimate (3.2.28).

$$\begin{aligned}
&\left|\mathbb{E}\sum_{j,k}[Y_s^TU(t_1,s)R^{1/2}]_{kj}[Y_s^TU(t_2,s)R^{1/2}]_{kj}[Y_s^TU(t_3,s)R^{1/2}]_{kj}Z\right|\\
&\le \mathbb{E}^{1/2}\left|\sum_{j,k}[Y_s^TU(t_1,s)R^{1/2}]_{kj}[Y_s^TU(t_2,s)R^{1/2}]_{kj}[Y_s^TU(t_3,s)R^{1/2}]_{kj}\right|^2K_Z\\
&\le \mathbb{E}^{1/2}\left|\sum_{\alpha_1,\alpha_2\alpha_3}\left(\sum_j[R^{1/2}U(t_1,s)]_{j\alpha_1}[R^{1/2}U(t_2,s)]_{j\alpha_2}[R^{1/2}U(t_3,s)]_{j\alpha_3}\right)\right.\\
&\qquad\times\left.\left(\sum_k[Y_s]_{\alpha_1k}[Y_s]_{\alpha_2k}[Y_s]_{\alpha_3k}\right)\right|^2K_Z\\
&\le \mathbb{E}^{1/2}\left[\sum_{\alpha_1\alpha_2\alpha_3}\left|\sum_j[R^{1/2}U(t_1,s)]_{j\alpha_1}[R^{1/2}U(t_2,s)]_{j\alpha_2}[R^{1/2}U(t_3,s)]_{j\alpha_3}\right|^2\right]\\
&\qquad\times\mathbb{E}^{1/2}\left[\sum_{\alpha_1\alpha_2\alpha_3}\left|\sum_k[Y_s]_{\alpha_1k}[Y_s]_{\alpha_2k}[Y_s]_{\alpha_3k}\right|^2\right]K_Z,
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}^{1/2}\left[\sum_{\alpha_1\alpha_2\alpha_3}\left|\sum_j[R^{1/2}U(t_1,s)]_{j\alpha_1}[R^{1/2}U(t_2,s)]_{j\alpha_2}[R^{1/2}U(t_3,s)]_{j\alpha_3}\right|^2\right]\\
&= \mathbb{E}^{1/2}\left[\sum_{\alpha_1\alpha_2\alpha_3}\sum_{j_1,j_2}[R^{1/2}U(t_1,s)]_{j_1\alpha_1}[R^{1/2}U(t_2,s)]_{j_1\alpha_2}[R^{1/2}U(t_3,s)]_{j_1\alpha_3}\right.\\
&\qquad\left.\times[R^{1/2}\bar U(t_1,s)]_{j_2\alpha_1}[R^{1/2}\bar U(t_2,s)]_{j_2\alpha_2}[R^{1/2}\bar U(t_3,s)]_{j_2\alpha_3}\right]\\
&= \mathbb{E}^{1/2}\left[\sum_{j_1,j_2}\left(\sum_{\alpha_1}[R^{1/2}U(t_1,s)]_{j_1\alpha_1}[R^{1/2}\bar U(t_1,s)]_{j_2\alpha_1}\right)\left(\sum_{\alpha_2}[R^{1/2}U(t_2,s)]_{j_1\alpha_2}[R^{1/2}\bar U(t_2,s)]_{j_2\alpha_2}\right)\right.\\
&\qquad\left.\times\left(\sum_{\alpha_3}[R^{1/2}U(t_3,s)]_{j_1\alpha_3}[R^{1/2}\bar U(t_3,s)]_{j_2\alpha_3}\right)\right]\\
&= \mathbb{E}^{1/2}\left[\sum_{j_1,j_2}|R_{j_1j_2}|^3\right]\le C_0\mathbb{E}^{1/2}\left[\sum_{j_1,j_2}|R_{j_1j_2}|^2\right]\le C_0\sqrt{n}, \qquad (3.5.3)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}^{1/2}\left[\sum_{\alpha_1\alpha_2\alpha_3}\left|\sum_k [Y_s]_{\alpha_1 k}[Y_s]_{\alpha_2 k}[Y_s]_{\alpha_3 k}\right|^2\right] \\
&= \mathbb{E}^{1/2}\left[\sum_{\alpha_1\alpha_2\alpha_3}\sum_{k_1,k_2}[Y_s]_{\alpha_1 k_1}[Y_s]_{\alpha_2 k_1}[Y_s]_{\alpha_3 k_1}[Y_s]_{\alpha_1 k_2}[Y_s]_{\alpha_2 k_2}[Y_s]_{\alpha_3 k_2}\right] \\
&= \mathbb{E}^{1/2}\left[\sum_{k_1,k_2}[Y_s^TY_s]_{k_2k_1}[Y_s^TY_s]_{k_2k_1}[Y_s^TY_s]_{k_2k_1}\right] \\
&\le C_0\sqrt{n}\mathbb{E}^{1/2}\left[\sum_{k_1,k_2}[Y_s^TY_s]^2_{k_2k_1}\right] \\
&\le C_0\sqrt{n}\mathbb{E}^{1/2}\mathrm{Tr}[Y_s^TY_sY_s^TY_s] \le C_0 n. \qquad (3.5.4)
\end{aligned}$$

(3.5.3) and (3.5.4) imply (3.2.28).

Same method applies for (3.2.27) :

$$\begin{aligned}
&\left|\mathbb{E}\sum_{jk}[Y_s^TU(t_1,s)Y_s]_{kk}[Y_s^TU(t_2,s)R^{1/2}]_{kj}[R^{1/2}U(t_3,s)R^{1/2}]_{jj}Z\right| \\
&\le \mathbb{E}^{1/2}\left|\sum_k [Y_s^TU(t_1,s)Y_s]_{kk}\left(\sum_j [Y_s^TU(t_2,s)R^{1/2}]_{kj}[R^{1/2}U(t_3,s)R^{1/2}]_{jj}\right)\right|^2 K_Z \\
&\le C_0\mathbb{E}^{1/2}\left(\sum_k \left|[Y_s^TU(t_1,s)Y_s]_{kk}\right|^2\right)\mathbb{E}^{1/2}\left(\sum_k\left|\sum_j [Y_s^TU(t_2,s)R^{1/2}]_{kj}[R^{1/2}U(t_3,s)R^{1/2}]_{jj}\right|^2\right)K_Z \\
&\le C_0\sqrt{n}\mathbb{E}^{1/2}\left(\sum_k \left|[Y_s^TU(t_1,s)Y_s]_{kk}\right|^2\right)\mathbb{E}^{1/2}\left(\sum_{j,k}\left|[Y_s^TU(t_2,s)R^{1/2}]_{kj}\right|^2\right)K_Z \\
&\le C_0\sqrt{n}\mathbb{E}^{1/2}\left(\sum_k\left|\sum_\alpha [Y_s]^2_{\alpha k}\right|^2\right)\mathbb{E}^{1/2}\left(\mathrm{Tr}(Y_s^TU(t_2,s)R\bar{U}(t_2,s)Y_s)\right)K_Z \\
&\le C_0 n\mathbb{E}^{1/2}\left(\sum_k\sum_{\alpha_1,\alpha_2}[Y_s]^2_{\alpha_1 k}[Y_s]^2_{\alpha_2 k}\right)K_Z \\
&\le C_0 n^{3/2}K_Z.
\end{aligned}$$

3.5.5 Proof of Proposition 3.2.4

$\ell = 1, 2$ is trivial. We shall prove the proposition for $\ell = 3$ or 4. When $\ell = 3$,

$$\mathbb{E}\left|\sum_\alpha Y^2_{\alpha k}\right|^3 = \mathbb{E}\sum_{\alpha_1,\alpha_2,\alpha_3} Y^2_{\alpha_1 k}Y^2_{\alpha_2 k}Y^2_{\alpha_3 k}$$

$$\begin{aligned}
&= \mathbb{E}\sum_{\alpha_1=\alpha_2=\alpha_3} Y_{\alpha_1 k}^6 + \mathbb{E}\sum_{\alpha_1=\alpha_2\neq\alpha_3} 3Y_{\alpha_1 k}^4 Y_{\alpha_3 k}^2 \\
&\quad +\mathbb{E}\sum_{\alpha_1\neq\alpha_2\neq\alpha_3} 3Y_{\alpha_1 k}^2 Y_{\alpha_2 k}^2 Y_{\alpha_3 k}^2 \\
&\leq C_0\mathbb{E}\sum_{\alpha_1} Y_{\alpha_1 k}^4 + C_0 \\
&\leq C_0.
\end{aligned}$$

When $\ell = 4$,

$$\begin{aligned}
\mathbb{E}\left|\sum_{\alpha} Y_{\alpha k}^2\right|^4 &= \mathbb{E}\sum_{\alpha_1,\alpha_2,\alpha_3,\alpha_4} Y_{\alpha_1 k}^2 Y_{\alpha_2 k}^2 Y_{\alpha_3 k}^2 Y_{\alpha_4 k}^2 \\
&= \mathbb{E}\sum_{\alpha_1} Y_{\alpha_1 k}^8 + \mathbb{E}\sum_{\alpha_1=\alpha_2=\alpha_3\neq\alpha_4} 4Y_{\alpha_1 k}^6 Y_{\alpha_4 k}^2 + \mathbb{E}\sum_{\alpha_1=\alpha_2\neq\alpha_3\neq\alpha_4} 6Y_{\alpha_1 k}^4 Y_{\alpha_3 k}^2 Y_{\alpha_4 k}^2 \\
&\quad +\mathbb{E}\sum_{\alpha_1=\alpha_2\neq\alpha_3=\alpha_4} 3Y_{\alpha_1 k}^4 Y_{\alpha_3 k}^4 + \mathbb{E}\sum_{\alpha_1\neq\alpha_2\neq\alpha_3\neq\alpha_4} Y_{\alpha_1 k}^2 Y_{\alpha_2 k}^2 Y_{\alpha_3 k}^2 Y_{\alpha_4 k}^2 \\
&\leq C_0\mathbb{E}\sum_{\alpha_1} Y_{\alpha_1 k}^4 + C_0\mathbb{E}\sum_{\alpha_1\neq\alpha_4} Y_{\alpha_1 k}^4 Y_{\alpha_4 k}^2 + C_0 \\
&\leq C_0.
\end{aligned}$$

3.5.6 Proof of Lemma 3.4.2

With the derivatives in (3.4.4), we have

$$\frac{\partial \mathrm{Tr}\mathbf{Q}(z)\mathbf{A}}{\partial X_{ij}} = -\frac{1}{\sqrt{n}}[Y^*Q(z)AQ(z)R^{1/2}]_{ji}, \quad \frac{\partial \mathrm{Tr}\mathbf{Q}(z)\mathbf{A}}{\partial \bar{X}_{ij}} = -\frac{1}{\sqrt{n}}[R^{1/2}Q(z)AQ(z)Y^*]_{ij}.$$

Then the Nash-Poicaré inequality yields that

$$\begin{aligned}
\mathrm{Var}\left[\frac{1}{n}\mathrm{Tr}\mathbf{Q}(z)\mathbf{A}\right] &\leq \frac{1}{n}\sum_{i,j}\left[\mathbb{E}\left|\frac{1}{n}\frac{\partial \mathrm{Tr}\mathbf{Q}(z)\mathbf{A}}{\partial X_{ij}}\right|^2 + \mathbb{E}\left|\frac{1}{n}\frac{\partial \mathrm{Tr}\mathbf{Q}(z)\mathbf{A}}{\partial \bar{X}_{ij}}\right|^2\right] \\
&\leq \frac{1}{n^3}\sum_{i,j}\left[\mathbb{E}\left|[Y^*Q(z)AQ(z)R^{1/2}]_{ji}\right|^2 + \mathbb{E}\left|[R^{1/2}Q(z)AQ(z)Y^*]_{ij}\right|^2\right] \\
&\leq \frac{2}{n^3}\mathbb{E}[\mathrm{Tr}\mathbf{Y}\mathbf{Y}^*\mathbf{Q}(z)\mathbf{A}\mathbf{Q}(z)\mathbf{R}\mathbf{Q}^*(z)\mathbf{A}^*\mathbf{Q}^*(z)] \\
&\leq \frac{2\|\mathbf{A}\|^2}{n^3}\mathbb{E}[\mathrm{Tr}(\mathbf{I}+z\mathbf{Q}(z))\mathbf{Q}(z)\mathbf{R}\mathbf{Q}^*(z)\mathbf{Q}^*(z)] \\
&\leq \frac{\|\mathbf{A}\|^2}{n^2}C_1(|z|)C_4(|\mathrm{Im}z|^{-1}).
\end{aligned}$$

3.5.7 Proof of Proposition 3.4.4

We first establish (3.4.24). From the definition of $\tilde{t}$ in (3.1.2), we have

$$\mathrm{Im}\tilde{t} = \frac{1}{2\mathbf{i}}(\tilde{t}(z) - \tilde{t}(\bar{z})) = v\mathrm{Im}(\frac{z}{n}\mathrm{Tr}\mathbf{R}\mathbf{T}) + v\mathrm{Im}z,$$

where $v = \tilde{t}(z)\tilde{t}(\bar{z})$.

For each matrix $\mathbf{M}$, we define $\mathrm{Im}\mathbf{M} = \frac{1}{2\mathbf{i}}(\mathbf{M} - \mathbf{M}^*)$. Noticing that $\mathbf{T}^*(z) = \mathbf{T}(\bar{z})$ and

$$\begin{cases} \mathbf{T}^T = (-z(\mathbf{I} + \tilde{t}(z)\bar{\mathbf{R}}))^{-1} \\ \mathbf{T}^* = (-\bar{z}(\mathbf{I} + \tilde{t}(\bar{z})\mathbf{R}))^{-1} \end{cases}$$

we have :

$$\begin{aligned} \mathrm{Im}(z\mathbf{RT}) &= \frac{1}{2\mathbf{i}}(z\mathbf{RT}(z) - \bar{z}\mathbf{RT}(\bar{z})) \\ &= \frac{1}{2\mathbf{i}} z\mathbf{RT}(z)[(\tilde{t}(z) - \tilde{t}(\bar{z}))\mathbf{R}]\bar{z}\mathbf{T}(\bar{z}). \end{aligned}$$

Taking trace and divided by n, we have

$$\mathrm{Im}(\frac{z}{n}\mathrm{Tr}\mathbf{RT}) = |z|^2\tilde{v}\mathrm{Im}(\tilde{t}(z))$$

where $\tilde{v} = \frac{1}{n}\mathrm{Tr}[\mathbf{RT}(z)\mathbf{T}^*(z)\mathbf{R}]$.

Thus we get :

$$\begin{bmatrix} \mathrm{Im}(\tilde{t}) \\ \mathrm{Im}(\frac{z}{n}\mathrm{Tr}\mathbf{RT}) \end{bmatrix} = \mathbf{D}(z) \begin{bmatrix} \mathrm{Im}(\tilde{t}) \\ \mathrm{Im}(\frac{z}{n}\mathrm{Tr}\mathbf{RT}) \end{bmatrix} + \begin{bmatrix} v(z) \\ 0 \end{bmatrix} \mathrm{Im}(z) \tag{3.5.5}$$

where $\mathbf{D}(z)$ is defined in (3.4.20). This is equivalent to

$$\begin{cases} \mathrm{Im}(\tilde{t}) = v\mathrm{Im}(\frac{z}{n}\mathrm{Tr}\mathbf{RT}) + v\mathrm{Im}z \\ \mathrm{Im}(\frac{z}{n}\mathrm{Tr}\mathbf{RT}) = |z|^2\tilde{v}\mathrm{Im}(\tilde{t}) \end{cases}. \tag{3.5.6}$$

Then we have

$$\det(\mathbf{I} - \mathbf{D}) = 1 - |z|^2 v\tilde{v} = v\frac{\mathrm{Im}z}{\mathrm{Im}(\tilde{t})}. \tag{3.5.7}$$

As $\tilde{t}$ is a Stieljes transform of the measure $\tilde{F}_n$, we have

$$|\mathrm{Im}(\tilde{t})| \leq \frac{1}{|\mathrm{Im}z|}. \tag{3.5.8}$$

On the other hand, $\mathrm{Im}(\tilde{t}(z))$ can be written as

$$\mathrm{Im}(\tilde{t}(z)) = \mathrm{Im}z \int_{\mathbb{R}^+} \frac{d\tilde{F}_n(\lambda)}{|\lambda - z|^2}.$$

We recall that in [70], it is shown that the sequence $\tilde{F}_n$ is tight. This implies that it exists a constant $C > 0$ for which $F_n(]0, C[) \geq 1/2$. Then

$$\int_{\mathbb{R}^+} \frac{d\tilde{F}_n(\lambda)}{|\lambda - z|^2} > \int_0^C \frac{d\tilde{F}_n(\lambda)}{|\lambda - z|^2} > \frac{1}{4(C^2 + |z|^2)}$$

for all $n \in \mathbb{N}$. Therefore,

$$v \geq |\mathrm{Im}(\tilde{t})|^2 \geq \frac{|\mathrm{Im}z|^2}{16(C^2+|z|^2)^2}. \tag{3.5.9}$$

Plugging (3.5.8) and (3.5.9) into (3.5.7), we get

$$\det(\mathbf{I}-\mathbf{D}) = 1-|z|^2 v\tilde{v} \geq \frac{|\mathrm{Im}z|^4}{16(C^2+|z|^2)^2}.$$

We now establish (3.4.25). Recall the definition of $\mathbf{S}$ and $\tilde{\tau}$ in (3.4.5) and (3.4.6) respectively. We express $\mathrm{Im}(\mathbb{E}\tilde{m}_c(z))$ and $\frac{1}{n}\mathrm{Im}(z\mathbb{E}\mathrm{Tr}\mathbf{RQ})$ as

$$\begin{bmatrix} \mathrm{Im}(\mathbb{E}\tilde{m}_c(z)) \\ \mathrm{Im}(\frac{z}{n}\mathbb{E}\mathrm{Tr}\mathbf{RQ}) \end{bmatrix} = \begin{bmatrix} \mathrm{Im}(\tilde{\tau}(z)) \\ \mathrm{Im}(\frac{z}{n}\mathrm{Tr}\mathbf{RS}) \end{bmatrix} + \begin{bmatrix} \frac{1}{n}\mathrm{Im}\varepsilon(z) \\ \frac{1}{n}\mathrm{Im}z\tilde{\varepsilon}(z) \end{bmatrix},$$

where $\varepsilon(z) = n(\mathbb{E}\tilde{m}_c(z)-\tilde{\tau}(z))$ and $\tilde{\varepsilon}(z) = \mathbb{E}\mathrm{Tr}\mathbf{RQ}-\mathrm{Tr}\mathbf{RS}$.

Then by calculating the imaginary parts of $\tilde{\tau}(z)$ and $\frac{z}{n}\mathrm{Tr}\mathbf{RS}$, we have

$$\begin{aligned} \mathrm{Im}(\tilde{\tau}(z)) &= v'\mathrm{Im}(\frac{z}{n}\mathbb{E}\mathrm{Tr}\mathbf{RQ}) + v'\mathrm{Im}z \\ \mathrm{Im}(\frac{z}{n}\mathrm{Tr}\mathbf{RS}) &= |z|^2\tilde{v}'\mathrm{Im}\mathbb{E}\tilde{m}_c(z), \end{aligned}$$

where v' and $\tilde{v}'$ are defined in (3.4.23). Then

$$\begin{bmatrix} \mathrm{Im}(\mathbb{E}\tilde{m}_c(z)) \\ \mathrm{Im}(\frac{z}{n}\mathbb{E}\mathrm{Tr}\mathbf{RQ}) \end{bmatrix} = \mathbf{D}'(z)\begin{bmatrix} \mathrm{Im}(\mathbb{E}\tilde{m}_c(z)) \\ \mathrm{Im}(\frac{z}{n}\mathbb{E}\mathrm{Tr}\mathbf{RQ}) \end{bmatrix} + \begin{bmatrix} v'(z) \\ 0 \end{bmatrix}\mathrm{Im}z + \begin{bmatrix} \frac{1}{n}\mathrm{Im}\varepsilon(z) \\ \frac{1}{n}\mathrm{Im}z\tilde{\varepsilon}(z) \end{bmatrix}, \tag{3.5.10}$$

where $\mathbf{D}'$ is defined in (3.4.21). We get the system

$$\begin{cases} \mathrm{Im}(\mathbb{E}\tilde{m}_c(z)) = v'\mathrm{Im}(\frac{z}{n}\mathbb{E}\mathrm{Tr}\mathbf{RQ}) + v'\mathrm{Im}z + \frac{1}{n}\mathrm{Im}\varepsilon(z) \\ \mathrm{Im}(\frac{z}{n}\mathbb{E}\mathrm{Tr}\mathbf{RQ}) = |z|^2\tilde{v}'\mathrm{Im}(\mathbb{E}\tilde{m}_c(z)) + \frac{1}{n}\mathrm{Im}(z\tilde{\varepsilon}(z)) \end{cases}. \tag{3.5.11}$$

This equation is similar to (3.5.6) with two extra error terms $\frac{1}{n}\mathrm{Im}[\varepsilon(z)]$ and $\frac{1}{n}\mathrm{Im}(z\tilde{\varepsilon}(z))$. However, $\frac{1}{n}\mathrm{Im}[\varepsilon(z)]$ and $\frac{1}{n}\mathrm{Im}(z\tilde{\varepsilon}(z))$ are controlled in Proposition 3.4.3.

We first evaluate $v' = \tilde{\tau}(z)\tilde{\tau}(\bar{z})$. As $\tilde{\tau}(z)$ is the Stieljes transform of a probability measure ξ_n. By Proposition 3.4.3, for $z \in \mathbb{C}^+$, $\mathbb{E}\tilde{m}_c(z)-\tilde{\tau}(z) \to 0$ and $\mathbb{E}\tilde{m}_c(z)-\tilde{t}(z) \to 0$. These imply that

$$\tilde{\tau}(z)-\tilde{t}(z) \to 0.$$

Therefore, ξ_n converges weakly to $\tilde{F}_n$ and there exists an integer N_1 and C>0 such that for $n \geq N_1$,

$$\xi_n([0,C]) > \frac{1}{4}.$$

With the same argument as (3.5.9), we obtain

$$v' \geq \frac{|\mathrm{Im}z|^2}{64(C^2+|z|^2)^2}. \tag{3.5.12}$$

By Proposition 3.4.3,

$$\begin{aligned}|\varepsilon(z)| &\le \frac{1}{n}C_2(|z|)C_5(|\mathrm{Im}z|^{-1}),\\ |\tilde{\varepsilon}(z)| &\le \frac{1}{n}C_3(|z|)C_7(|\mathrm{Im}z|^{-1}),\end{aligned}$$

this along with (3.5.11) yield that

$$\mathrm{Im}(\mathbb{E}\tilde{m}_c(z)) > v'\mathrm{Im}z - \frac{1}{n}|\varepsilon| > \frac{|\mathrm{Im}z|^3}{64(C^2+|z|^2)^2} - \frac{1}{n^2}C_2(|z|)C_5(|\mathrm{Im}z|^{-1}). \tag{3.5.13}$$

Then we consider the set $\mathcal{E}_1$ defined by

$$\mathcal{E}_1 = \left\{z \in \mathbb{C}^+ : 1 - \frac{1}{n^2}C_6(|z|)C_8(|\mathrm{Im}z|^{-1}) > \frac{1}{2}\right\}.$$

For $z \in \mathcal{E}_1$,

$$\mathrm{Im}(\mathbb{E}\tilde{m}_c(z)) > \frac{|\mathrm{Im}z|^3}{128(C^2+|z|^2)^2}. \tag{3.5.14}$$

From (3.5.10),

$$\begin{aligned}\det(\mathbf{I}-\mathbf{D}') &= v'\frac{\mathrm{Im}z}{\mathrm{Im}(\mathbb{E}\tilde{m}_c(z))} + v'\frac{\mathrm{Im}(\frac{z}{n}\tilde{\varepsilon})}{\mathrm{Im}(\mathbb{E}\tilde{m}_c(z))} + \frac{\frac{1}{n}\mathrm{Im}\varepsilon}{\mathrm{Im}(\mathbb{E}\tilde{m}_c(z))}\\ &\ge v'|\mathrm{Im}z|^2 - v'\frac{|z\tilde{\varepsilon}|}{n\mathrm{Im}\mathbb{E}\tilde{m}_c} - \frac{|\varepsilon|}{n\mathrm{Im}\mathbb{E}\tilde{m}_c}\\ &\overset{(a)}{\ge} \frac{|\mathrm{Im}z|^4}{64(C^2+|z|^2)^2} - \frac{128|z|(C^2+|z|^2)^2C_3(|z|)C_7(|\mathrm{Im}z|^{-1})}{n^2|\mathrm{Im}z|^5}\\ &\quad - \frac{128(C^2+|z|^2)^2C_2(|z|)C_5(|\mathrm{Im}z|^{-1})}{n^2|\mathrm{Im}z|^3}\\ &\ge \frac{|\mathrm{Im}z|^4}{64(C^2+|z|^2)^2}(1 - \frac{1}{n^2}C_{12}(|z|)C_{16}(|\mathrm{Im}z|^{-1})).\end{aligned}$$

where (a) uses (3.5.14) (3.5.12), (3.4.12) and $v' \le \frac{1}{|\mathrm{Im}z|^2}$. Then we consider the second set

$$\mathcal{E}_2 = \left\{z \in \mathbb{C}^+ : 1 - \frac{1}{n^2}C_{12}(|z|)C_{16}(|\mathrm{Im}z|^{-1}) > \frac{1}{2}\right\},$$

and the set

$$\mathcal{E} = \left\{z \in \mathbb{C}^+ : 1 - \frac{2}{n^2}C_6(|z|)C_{10}(|\mathrm{Im}z|^{-1}) - \frac{2}{n^2}C_{12}(|z|)C_{16}(|\mathrm{Im}z|^{-1}) > 0\right\} \subset \mathcal{E}_1 \cap \mathcal{E}_2,$$

or equivalently :

$$\mathcal{E} = \left\{z \in \mathbb{C}^+ : 1 - \frac{1}{n^2}C_{12}(|z|)C_{16}(|\mathrm{Im}z|^{-1}) > 0\right\}.$$

Remark 3.5.1 *Readers should pay attention that the polynomials C_ℓ here are not the same.*

It is clear that for $z \in \mathcal{E}$, (3.4.25) holds.

To show (3.4.26), we have

$$\begin{aligned}|\det(\mathbf{I} - \mathbf{D}_0)| &= |1 - z^2 v_0 \tilde{v}_0| \\ &= 1 - |z|^2 |v_0||\tilde{v}_0|.\end{aligned}$$

By Cauchy-Schwarz, $|v_0| \leq |v|^{1/2}|v'|^{1/2}$ and $|\tilde{v}_0| \leq |\tilde{v}|^{1/2}|\tilde{v}'|^{1/2}$, which yield that

$$|\det(\mathbf{I} - \mathbf{D}_0)| \geq 1 - |z|^2 |v|^{1/2}|v'|^{1/2}|\tilde{v}|^{1/2}|\tilde{v}'|^{1/2}.$$

As $\det(\mathbf{I} - \mathbf{D}) = 1 - |z|^2 v\tilde{v}$ and $\det(\mathbf{I} - \mathbf{D}') = 1 - |z|^2 v'\tilde{v}'$ are positive for $z \in \mathcal{E}$,

$$|z|^2 v\tilde{v} \leq 1, \quad |z|^2 v'\tilde{v}' \leq 1,$$

which imply that

$$|z|^2 |v|^{1/2}|v'|^{1/2}|\tilde{v}|^{1/2}|\tilde{v}'|^{1/2} \leq 1.$$

By a direct calculus, (3.4.26) holds true.

Chapitre 4

Estimation for large dimensional random matrices

4.1 Fluctuations of an improved estimator in sample covariance matrix models

The content below is evolved from the papers [100, 99].

Problems of statistical inference based on M independent observations of an N-variate random variable $\mathbf{y}$, with $\mathbb{E}[\mathbf{y}] = 0$ and $\mathbb{E}[\mathbf{y}\mathbf{y}^H] = \mathbf{R}_N$ have drawn the attention of researchers from many fields for years : Portfolio optimization in finance [77], gene coexistence in biostatistics [59], channel capacity in wireless communications [52], power estimation in sensor networks [26], distance of targets in array processing [66], etc.

In particular, retrieving spectral properties of the *population covariance matrix* $\mathbf{R}_N$, based on the observation of M independent and identically distributed (i.i.d.) samples $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(M)}$, is paramount to many questions of general science. If M is large compared to N, then it is known that almost surely $\|\hat{\mathbf{R}}_N - \mathbf{R}_N\| \to 0$, as $M \to \infty$, for any standard matrix norm, where $\hat{\mathbf{R}}_N$ is the *sample covariance matrix* $\hat{\mathbf{R}}_N \triangleq \frac{1}{M}\sum_{m=1}^{M} \mathbf{y}^{(m)}\mathbf{y}^{(m)H}$. However, one cannot always afford a large number of samples. In order to cope with this issue, random matrix theory [5, 25] has proposed new estimators, mainly spurred by the *G-estimators* of Girko [34]. Other works include convex optimization methods [85, 54], free probability tools [79, 24], and regularized estimation (banding, tapering, thresholding, etc.) [15, 16, 20], when the structure of $\mathbf{R}_N$ is known. Many of those estimators are consistent in the sense that they are asymptotically unbiased as M, N grow large at the same rate. Nonetheless, only recently have techniques been unveiled which allow to estimate individual eigenvalues and functionals of eigenvectors of $\mathbf{R}_N$. The first contributor is Mestre [65]-[64] who studied the case where $\hat{\mathbf{R}}_N = \mathbf{R}_N^{1/2}\mathbf{U}_N\mathbf{U}_N^H\mathbf{R}_N^{1/2}$ with $\mathbf{R}_N$ having eigenvalues with large multiplicities and unknown eigenvectors, and $\mathbf{U}_N$ with i.i.d. entries. For this model, he provides an estimator for every eigenvalue of $\mathbf{R}_N$ with large multiplicity under some separability condition, see also Vallet *et al.* [92], Couillet *et al.* [26] for more elaborate models.

These estimators, although proven asymptotically unbiased, have nonetheless not been fully characterized in terms of their asymptotic performances. It is in particular fundamental to evaluate the variance of these estimators for not-too-large M, N. The purpose of this article is to study the asymptotic fluctuations of the population eigenvalue estimator of [64] in the case of structured population covariance matrices. A central limit theorem (CLT) is provided to describe the asymptotic fluctuations of the estimators with exact expression for the variance as M, N tend to infinity. An empirical, asymptotically accurate, approximation is also derived. For an application of these results in a cognitive radio context, see for instance [100].

The remainder of the section is structured as follows : In Section 4.1.1, the system model is introduced and the main results from [65, 64] are recalled. In Section 4.1.4, the CLT for the estimator in [64] is stated and the asymptotic variance derived. In Section 4.1.5, an empirical approximation for the variance is derived. Some simulation results are shown in this section.Technical proofs are postponed to the appendix.

4.1.1 Matrix Model

Consider an $N \times M$ matrix $\mathbf{X}_N = (X_{ij})$ whose entries are i.i.d. random variables, with distribution $\mathcal{CN}(0,1)$, i.e. $X_{ij} = U + \mathbf{i}V$, where U, V are both i.i.d. real Gaussian random variables $\mathcal{N}(0, \frac{1}{2})$. Let $\mathbf{R}_N$ be an $N \times N$ Hermitian matrix with L (L being fixed) distinct eigenvalues $\rho_1 < \cdots < \rho_L$ with respective multiplicities $N_1, \cdots, N_L$ (so that $\sum_{i=1}^{L} N_i = N$). Consider now

$$\mathbf{Y}_N = \mathbf{R}_N^{1/2} \mathbf{X}_N \ .$$

The matrix $\mathbf{Y}_N = [\mathbf{y}_1, \cdots, \mathbf{y}_M]$ is the concatenation of M independent observations $[\mathbf{y}_1, \cdots, \mathbf{y}_M]$, where each observation writes $\mathbf{y}_i = \mathbf{R}_N^{1/2} \mathbf{x}_i$ with $\mathbf{X}_N = [\mathbf{x}_1, \cdots, \mathbf{x}_M]$. In particular, the (population) covariance matrix of each observation $\mathbf{y}_i$ is $\mathbf{R}_N = \mathbb{E}\left(\mathbf{y}_i \mathbf{y}_i^H\right)$. In this article, we are interested in recovering information on $\mathbf{R}_N$ based on the observation

$$\hat{\mathbf{R}}_N = \frac{1}{M} \mathbf{R}_N^{1/2} \mathbf{X}_N \mathbf{X}_N^H \mathbf{R}_N^{1/2} \ ,$$

commonly referred to as the *sample covariance matrix* of the $\mathbf{y}_i$'s.

It is in general a complicated task to infer the spectral properties of $\mathbf{R}_N$ based on $\hat{\mathbf{R}}_N$ for all finite N, M. Instead, in the following, we assume that N and M are large, and consider the following asymptotic regime :

Assumption A1. The dimensions N, M and $(N_i)_{1 \le i \le L}$ satisfy the following conditions : when $N, M, N_i \to \infty$,

$$\frac{N}{M} \to c \in (0, \infty) \quad \text{and} \quad \frac{N_i}{M} \to c_i \in (0, \infty) \ , \ 1 \le i \le L. \tag{4.1.1}$$

This assumption will be shortly referred to as $N, M \to \infty$.

Assumption A2. The limiting support $\mathcal{S}$ of the eigenvalue distribution of $\hat{\mathbf{R}}_N$ is formed of L compact disjoint subsets $(\mathcal{S}_k; 1 \le k \le L)$, often referred to as *clusters* in the sequel.

From [64], one can also reformulate this condition in mathematical terms : The limiting support of $\hat{\mathbf{R}}_N$ is formed of L disjoint clusters if and only if for $1 \le i \le L$, $\inf_N \{ \frac{M}{N} - \Psi_N(i) \} > 0$, where $\Psi_N(i)$ is defined in (4.1.2) and $\alpha_1 \le \cdots \le \alpha_{L-1}$ are the $L-1$ distinct real ordered solutions

$$\Psi_N(i) = \begin{cases} \frac{1}{N}\sum_{r=1}^{L} N_r\left(\frac{\rho_r}{\rho_r-\alpha_1}\right)^2 & m=1, \\ \max\left\{\frac{1}{N}\sum_{r=1}^{L} N_r\left(\frac{\rho_r}{\rho_r-\alpha_{m-1}}\right)^2, \frac{1}{N}\sum_{r=1}^{L} N_r\left(\frac{\rho_r}{\rho_r-\alpha_m}\right)^2\right\} & 1<m<L, \\ \frac{1}{N}\sum_{r=1}^{L} N_r\left(\frac{\rho_r}{\rho_r-\alpha_{L-1}}\right)^2 & m=L \end{cases} \tag{4.1.2}$$

---

of the equation :

$$\frac{1}{N}\sum_{r=1}^{L} N_r \frac{\rho_r^2}{(\rho_r - x)^3} = 0.$$

This condition is also called the *separability* condition.

Figure 4.1 depicts the eigenvalues of a realization of the random matrix $\hat{\mathbf{R}}_N$ and the associated limiting distribution as N, M grow large, for $\rho_1 = 1$, $\rho_2 = 3$, $\rho_3 = 10$ and $N = 60, M = 600$ with $N_1 = N_2 = N_3 = 20$. The separability condition is illustrated there. Figure 4.2 shows another situation where the separability condition is not satisfied for $\rho_1 = 1$, $\rho_2 = 3$, $\rho_3 = 5$ and $N = 30, M = 80$ with $N_1 = N_2 = N_3 = 10$.

4.1.2 Mestre's Estimator of the population eigenvalues

In [64], an estimator of the population eigenvalues (ρ_k; $1 \le k \le L$) based on the observations $\hat{\mathbf{R}}_N$ is proposed.

Theorem 4.1.1 ([64, Th. 3]) *Let Assumptions* **A1** *and* **A2** *hold true and denote by $\hat\lambda_1 \le \cdots \le \hat\lambda_N$ the ordered eigenvalues of $\hat{\mathbf{R}}_N$. Then the following convergence holds true :*

$$\hat\rho_k - \rho_k \xrightarrow[M,N\to\infty]{a.s.} 0 \ , \tag{4.1.3}$$

where

$$\hat\rho_k = \frac{M}{N_k}\sum_{m\in\mathcal{N}_k}\left(\hat\lambda_m - \hat\mu_m\right) \ , \tag{4.1.4}$$

with $\mathcal{N}_k = \{\sum_{j=1}^{k-1} N_j + 1, \dots, \sum_{j=1}^{k} N_j\}$ and the $\hat\mu_i$'s defined[1] *as follows :*

– *If $N \le M$, then $\hat\mu_1 \le \cdots \le \hat\mu_N$ are the real ordered solutions of*

$$\frac{1}{N}\sum_{m=1}^{N}\frac{\hat\lambda_m}{\hat\lambda_m - \mu} = \frac{M}{N} \ . \tag{4.1.5}$$

– *If $N > M$, $\hat\mu_i = 0$ for $1 \le i \le N - M$ and $\hat\mu_{N-M+1}, \cdots, \hat\mu_N$ are the real solutions of the above equation.*

---

1. Another characterization of interest of the $\hat\mu_i$'s is the fact that they are the eigenvalues of $\mathrm{diag}(\hat{\boldsymbol\lambda}) - \frac{1}{M}\sqrt{\hat{\boldsymbol\lambda}}\sqrt{\hat{\boldsymbol\lambda}}^T$, where $\hat{\boldsymbol\lambda} = (\hat\lambda_1, \dots, \hat\lambda_N)^T$, see for instance [25, Chapter 8].

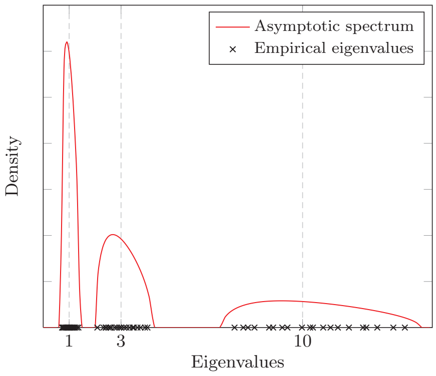

FIGURE 4.1 – Empirical and asymptotic eigenvalue distribution of $\hat{\mathbf{R}}_N$ for $L = 3$, $\rho_1 = 1$, $\rho_2 = 3$, $\rho_3 = 10$, $N/M = c = 0.1$, $N = 60$, $N_1 = N_2 = N_3 = 20$.

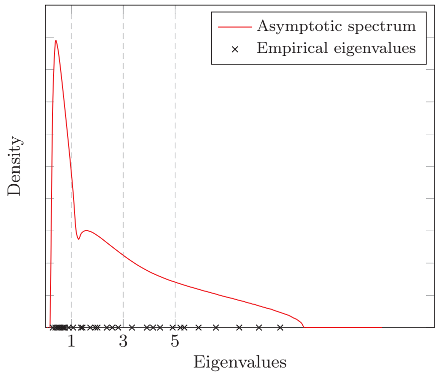

FIGURE 4.2 – Empirical and asymptotic eigenvalue distribution of $\hat{\mathbf{R}}_N$ for $L = 3$, $\rho_1 = 1$, $\rho_2 = 3$, $\rho_3 = 5$, $N/M = c = 3/8$, $N = 30$, $N_1 = N_2 = N_3 = 10$.

4.1.3 Integral representation of estimator $\hat{\rho}_k$ - Stieltjes transforms

The proof of Theorem 4.1.1 relies on large random matrix theory, and in particular on the *Stieltjes transform* of a probability distribution. The Stieltjes transform $m_{\mathbb{P}}$ of a probability distribution $\mathbb{P}$ over $\mathbb{R}^+$ is a $\mathbb{C}$-valued function defined by :

$$m_{\mathbb{P}}(z) = \int_{\mathbb{R}^+} \frac{\mathbb{P}(d\lambda)}{\lambda - z} \ , \quad z \in \mathbb{C}\backslash\mathbb{R}^+ \ .$$

There also exists an inverse formula to recover the probability distribution associated to a Stieljes transform : Let $a < b$ be two continuity points of the cumulative distribution function associated to $\mathbb{P}$, then

$$\mathbb{P}([a,b]) = \frac{1}{\pi} \lim_{y \downarrow 0} \Im \left[\int_a^b m_{\mathbb{P}}(x+iy)dx \right] .$$

In the case where $F^{\mathbf{Z}}$ is the spectral distribution associated to a Hermitian matrix $\mathbf{Z} \in \mathbb{C}^{N\times N}$ with eigenvalues $(\xi_i;\ 1 \le i \le N)$, the Stieltjes transform $m_{\mathbf{Z}}$ of $F^{\mathbf{Z}}$ takes the particular form :

$$\begin{aligned} m_{\mathbf{Z}}(z) &= \int \frac{F^{\mathbf{Z}}(d\,\lambda)}{\lambda - z} \\ &= \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\xi_i - z} = \frac{1}{N}\mathrm{Tr}\left(\mathbf{Z} - z\mathbf{I}_N\right)^{-1} \ , \end{aligned}$$

which is the normalized trace of the resolvent $(\mathbf{Z} - z\mathbf{I}_N)^{-1}$. Since the seminal paper of Marčenko and Pastur [62], the Stieltjes transform has proved to be extremely efficient to describe the limiting spectrum of large dimensional random matrices.

In the following, we recall some elements of the proof of Theorem 4.1.1, necessary for the remainder of the chapter. The following important result is due to Bai and Silverstein [83] (see also [62]).

Theorem 4.1.2 ([83]) *Let Assumption* **A1** *hold true and denote by F^R the limiting spectral distribution of $\mathbf{R}_N$,* i.e.

$$F^R(d\lambda) = \sum_{k=1}^{L} \frac{c_k}{c} \delta_{\rho_k}(d\lambda) \ .$$

Then, the spectral distribution $F^{\hat{\mathbf{R}}_N}$ of the sample covariance matrix $\hat{\mathbf{R}}_N$ converges (weakly and almost surely) to a probability distribution F as $M, N \to \infty$, whose Stieltjes transform $m(z)$ satisfies :

$$m(z) = \frac{1}{c}\underline{m}(z) - \left(1 - \frac{1}{c}\right)\frac{1}{z} \ , \tag{4.1.6}$$

for $z \in \mathbb{C}^+ = \{z \in \mathbb{C},\ \Im(z) > 0\}$ and where $\underline{m}(z)$ is defined as the unique solution in $\mathbb{C}^+$ of :

$$\underline{m}(z) = -\left(z - c\int \frac{t}{1 + t\underline{m}(z)} F^R(dt)\right)^{-1} .$$

Note that $\underline{m}(z)$ is also the Stieltjes transform of a probability distribution $\underline{F}$, which turns out to be the limiting spectral distribution of $F^{\underline{\hat{\mathbf{R}}}_N}$ where $\underline{\hat{\mathbf{R}}}_N$ is defined as :

$$\underline{\hat{\mathbf{R}}}_N \triangleq \frac{1}{M}\mathbf{X}_N^H \mathbf{R}_N \mathbf{X}_N \ .$$

Denote by $m_{\hat{\mathbf{R}}_N}(z)$ and $m_{\underline{\hat{\mathbf{R}}}_N}(z)$ the Stieltjes transforms of $F^{\hat{\mathbf{R}}_N}$ and $F^{\underline{\hat{\mathbf{R}}}_N}$. Note in particular that

$$m_{\hat{\mathbf{R}}_N}(z) = \frac{M}{N} m_{\underline{\hat{\mathbf{R}}}_N}(z) - \left(1 - \frac{M}{N}\right)\frac{1}{z} \ .$$

Remark 4.1.1 *This relation associated to (4.1.5) readily implies that for $\hat{\mu}_i \neq 0$, $m_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i) = 0$. Otherwise stated, the (non null) $\hat{\mu}_i$'s are the zeros of $m_{\underline{\hat{\mathbf{R}}}_N}$. This fact will be of importance in the sequel.*

Denote by $m_N(z)$ and $\underline{m}_N(z)$ the finite-dimensional counterparts of $m(z)$ and $\underline{m}(z)$, respectively, defined by the relations :

$$\begin{aligned}
\underline{m}_N(z) &= -\left(z - \frac{N}{M}\int \frac{t}{1 + t\underline{m}_N(z)} F^{\mathbf{R}_N}(dt)\right)^{-1} , \\
m_N(z) &= \frac{M}{N}\underline{m}_N(z) - \left(1 - \frac{M}{N}\right)\frac{1}{z} \ ,
\end{aligned} \tag{4.1.7}$$

where $\underline{m}_N(z)$ is the unique solution of (4.1.7) satisfying $\underline{m}_N(z) \in \mathbb{C}^+$ if $z \in \mathbb{C}^+$. It can be shown that m_N and $\underline{m}_N$ are Stieltjes transforms of probability measures F_N and $\underline{F}_N$, respectively (cf. [25, Theorem 3.2]).

With these notations at hand, we can now provide some elements of the proof of Theorem 4.1.1.

Proof :[Elements of proof for Theorem 4.1.1] By Cauchy's formula, write :

$$\rho_k = \frac{N}{N_k}\frac{1}{2i\pi}\oint_{\Gamma_k}\left(\frac{1}{N}\sum_{r=1}^{L} N_r \frac{w}{\rho_r - w}dw\right) ,$$

where Γ_k is a negatively oriented contour taking values on $\mathbb{C}\setminus\{\rho_1,\cdots,\rho_L\}$ and only enclosing ρ_k. With the change of variable $w = -\frac{1}{\underline{m}_M(z)}$, the condition that the limiting support $\mathcal{S}$ of the eigenvalue distribution of $\mathbf{R}_N$ is formed of L distinct clusters $(\mathcal{S}_k, 1 \le k \le L)$ (cf. Assumption **A2**), and standard properties of contour integrals, we can write :

$$\rho_k = \frac{M}{2i\pi N_k}\oint_{\mathcal{R}_k} z\frac{\underline{m}_N'(z)}{\underline{m}_N(z)}dz \ , \quad 1 \le k \le L \tag{4.1.8}$$

where $\mathcal{R}_k$ denotes a negatively oriented, rectangular and symmetric with respect to the abscissa axis, contour which only encloses the corresponding cluster $\mathcal{S}_k$. Defining

$$\hat{\rho}_k \triangleq \frac{M}{2\pi i N_k}\oint_{\mathcal{R}_k} z\frac{m'_{\underline{\hat{\mathbf{R}}}_N}(z)}{m_{\underline{\hat{\mathbf{R}}}_N}(z)}dz \ , \quad 1 \le k \le L \ , \tag{4.1.9}$$

dominated convergence arguments ensure that $\rho_k - \hat{\rho}_k \to 0$, almost surely. The integral form of $\hat{\rho}_k$ can then be explicitly computed thanks to residue calculus, and this finally yields (4.2.1).

□

Remark 4.1.2 (About the contour integrals) *If $\mathcal{R}_k'$ is another (rectangular and symmetric with respect to the abscissa axis) contour which only encloses the k-th cluster, then the value of the contour integrals in* (4.1.8) *and* (4.1.9) *remains unchanged. In particular, we can arbitrarily choose two non-overlapping contours $\mathcal{R}_k$ and $\mathcal{R}_k'$ of the same cluster $\mathcal{S}_k$ in the sequel.*

The main objective of this chapter is to study the performance of the estimators $(\hat{\rho}_k,\ 1 \le k \le L)$. More precisely, we will establish a CLT for $(M(\hat{\rho}_k - \rho_k),\ 1 \le k \le L)$ as $M, N \to \infty$, explicitly characterize the limiting covariance matrix $\mathbf{\Theta} = (\Theta_{k\ell})_{1\le k,\ell\le L}$, and finally provide an estimator for $\mathbf{\Theta}$.

4.1.4 Fluctuations of the population eigenvalue estimators

The Central Limit Theorem

The main result of this chapter is the following CLT which expresses the fluctuations of $(\hat{\rho}_k,\ 1 \le k \le L)$.

Theorem 4.1.3 *Let Assumptions* **A1** *and* **A2** *hold true and recall the definitions of the $\underline{m}(z)$ and $\underline{m}_N(z)$ in Section 4.1.6 and 4.1.7 respectively. Then :*

$$(n(\hat{\rho}_k - \rho_k),\ 1 \le k \le L) \xrightarrow[M,N\to\infty]{\mathcal{D}} \mathbf{x} \sim \mathcal{N}_L(0, \mathbf{\Theta}) \ ,$$

where $\mathbf{\Theta}$ *is an* $L \times L$ *matrix whose entries* $\Theta_{k\ell}$ *are given by*

$$\Theta_{k\ell} \quad = \quad -\frac{1}{4\pi^2 c_k c_\ell}\oint_{\mathcal{R}_k}\oint_{\mathcal{R}'_\ell}\left[\frac{\underline{m}'(z_1)\underline{m}'(z_2)}{(\underline{m}(z_1)-\underline{m}(z_2))^2}-\frac{1}{(z_1-z_2)^2}\right]\frac{1}{\underline{m}(z_1)\underline{m}(z_2)}dz_1 dz_2 \;. \quad (4.1.10)$$

The contours in (4.1.10) *are defined as follows. The contours* $(\mathcal{R}_k; 1 \leq k \leq L)$ *and* $(\mathcal{R}'_k; 1 \leq k \leq L)$ *are negatively oriented rectangles, symmetric with respect to the abscissa axis, and only enclosing the cluster* $\mathcal{S}_k$*. They also verify :*

$$\begin{aligned}\mathcal{R}_k \cap \mathcal{R}_\ell \quad &= \quad \mathcal{R}'_k \cap \mathcal{R}'_\ell \;=\; \emptyset \quad \text{for } k \neq \ell \;, \\ \mathcal{R}_k \cap \mathcal{R}'_\ell \quad &= \quad \emptyset \qquad \text{for all } k, \ell .\end{aligned}$$

In particular, the families $(\mathcal{R}_k)$ *and* $(\mathcal{R}'_k)$ *are non-overlapping.*

Remark 4.1.3 *In Theorem 4.1.3, the separability assumption* **A2** *can be relaxed to some extent. For example, if only the cluster associated to* ρ_k *satisfies the separability condition, one can study the fluctuations of* $\hat{\rho}_k$ *by relying on the same techniques.*

Proof of Theorem 4.1.3

We first outline the main steps of the proof and then provide the details.

Using the integral representation of $\hat{\rho}_k$ and ρ_k, we get :

$$M(\hat{\rho}_k - \rho_k) \quad = \quad \frac{M^2}{2\pi i N_k}\oint_{\mathcal{R}_k} z\left(\frac{m'_{\underline{\hat{\mathbf{R}}}_N}(z)}{m_{\underline{\hat{\mathbf{R}}}_N}(z)} - \frac{\underline{m}'_N(z)}{\underline{m}_N(z)}\right) dz \;.$$

Let $\mathcal{K}$ be the union of the $\mathcal{R}_k$'s and the $\mathcal{R}'_k$'s ; denote by $C(\mathcal{K}, \mathsf{X})$ the set of continuous functions from $\mathcal{K}$ to a Banach space X endowed with the supremum norm $\|u\|_\infty = \sup_{\mathcal{K}} |u|$. Consider the process :

$$\begin{aligned}&(X_N, X'_N, u_N, u'_N) : \mathcal{K} \to \mathbb{C}^4 \\ &\qquad\qquad z \mapsto (X_N(z), X'_N(z), u_N(z), u'_N(z))\end{aligned}$$

where

$$\begin{aligned}X_N(z) \quad &= \quad M\left(m_{\underline{\hat{\mathbf{R}}}_N}(z) - \underline{m}_N(z)\right) \;, \\ X'_N(z) \quad &= \quad M\left(m'_{\underline{\hat{\mathbf{R}}}_N}(z) - \underline{m}'_N(z)\right) \;, \\ u_N(z) \quad &= \quad m_{\underline{\hat{\mathbf{R}}}_N}(z) \;, \quad u'_N(z) \;=\; m'_{\underline{\hat{\mathbf{R}}}_N}(z) \;.\end{aligned}$$

Then from [9] (see also Proposition 4.1.1), (X_N, X'_N, u_N, u'_N) almost surely belongs to $C(\mathcal{K}, \mathbb{C}^4)$ for N, M large enough and $M(\hat{\rho}_k - \rho_k)$ writes :

$$\begin{aligned}M(\hat{\rho}_k - \rho_k) &= \frac{M}{2\pi i N_k}\oint_{\mathcal{R}_k} z\left(\frac{\underline{m}_N(z)X'_N(z) - u'_N(z)X_N(z)}{\underline{m}_N(z)u_N(z)}\right) dz \\ &\triangleq \Upsilon_N(X_N, X'_N, u_N, u'_N) \;,\end{aligned}$$

where
$$\Upsilon_N(x,x',u,u') = \frac{M}{2\pi i N_k}\oint_{\mathcal{R}_k} z\left(\frac{\underline{m}_N(z)x'(z)-u'(z)x(z)}{\underline{m}_N(z)u(z)}\right)dz. \tag{4.1.11}$$

Remark 4.1.4 *Note that Υ_N is a real random variable ; if needed, we shall explicitly indicate the dependence on the contour $\mathcal{R}_k$ and write $\Upsilon_N(x,x',u,u',\mathcal{R}_k)$.*

Remark 4.1.5 *Note that, due to formulas (4.1.8) and (4.1.9) and to Remark 4.1.2, the following equality holds true :*
$$\Upsilon_N(x,x',u,u',\mathcal{R}_k) = \Upsilon_N(x,x',u,u',\mathcal{R}_k'),$$
if $\mathcal{R}_k$ and $\mathcal{R}_k'$ are two contours which only contain the k-th cluster. This fact will be of importance later.

The main idea of the proof of the theorem lies in three steps :

(i) To prove the convergence in distribution of the process (X_N, X_N', u_N, u_N') to a Gaussian process.

(ii) To transfer this convergence to the quantity $\Upsilon_N(X_N, X_N', u_N, u_N', \mathcal{R}_k)$ with the help of the continuous mapping theorem [50].

(iii) To check that the limit (in distribution) of $\Upsilon_N(X_N, X_N', u_N, u_N')$ is Gaussian and to compute the limiting covariance between $\Upsilon_N(X_N, X_N', u_N, u_N', \mathcal{R}_k)$ and $\Upsilon_N(X_N, X_N', u_N, u_N', \mathcal{R}_\ell)$.

Remark 4.1.6 *Note that the convergence in step (i) is a distribution convergence at a process level, hence one has to first establish the finite dimensional convergence of the process and then to prove that the process is tight over $\mathcal{C}_k$ (see for instance [18, Theorem 13.1]). Tightness turns out to be difficult to establish due to the lack of control over the eigenvalues of $\underline{\hat{\mathbf{R}}}_N$ whenever the contour crosses the real line. In order to circumvent this issue, we shall introduce, following Bai and Silverstein [11], a process that approximates X_N and X_N'.*

Let us now start the proof of Theorem 4.1.3.

We begin by simple considerations on complex Gaussian random vectors. Consider a $\mathbb{C}^2$-valued, centered, random vector (U,V). If (U,V) is, as an $\mathbb{R}^4$-valued vector, Gaussian, then its distribution is fully characterized with the quantities :
$$\mathbb{E}U^2,\ \mathbb{E}V^2,\ \mathbb{E}U\bar{U},\ \mathbb{E}V\bar{V},\ \mathbb{E}UV \text{ and } \mathbb{E}U\bar{V}\ .$$

Lemma 4.1.1 *Let $\underline{\mathcal{S}}$ be the support of the distribution $\underline{F}$.*

1. *The function $\kappa : (\mathbb{C}\setminus\underline{\mathcal{S}})^2 \to \mathbb{C}$ defined by*
$$\kappa(z,\tilde{z}) = \begin{cases} \frac{\underline{m}'(z)\underline{m}'(\tilde{z})}{(\underline{m}(z)-\underline{m}(\tilde{z}))^2} - \frac{1}{(z-\tilde{z})^2} & \text{if } z\neq\tilde{z}\ , \\ \frac{\underline{m}'''(z)}{6\underline{m}'(z)} - \frac{\underline{m}''(z)^2}{4\underline{m}'(z)^2} & \text{if } z=\tilde{z}\ , \end{cases}$$
is continuous and admits partial derivatives up to order 2 over $(\mathbb{C}\setminus\underline{\mathcal{S}})^2$.

2. *Let Assumptions* **A1** *and* **A2** *hold true and consider a compact set $\mathcal{K} \subset \mathbb{C}$, symmetric with respect to the real axis (i.e. $z \in \mathcal{K} \Rightarrow \bar{z} \in \mathcal{K}$) which does not intersect $\underline{\mathcal{S}}$. Then, the process*

$$\begin{aligned} (X_N, X'_N) : \mathcal{K} &\to \mathbb{C}^2 \\ z &\mapsto (X_N(z), X'_N(z)) \end{aligned}$$

converges in distribution to a stochastic process (X, Y) satisfying $\overline{X(z)} = X(\bar{z})$ and $\overline{Y(z)} = Y(\bar{z})$. As an $\mathbb{R}^4$-valued real process, the process (X, Y) is a centered Gaussian process with mean function zero and covariance function defined as follows, for $z, \tilde{z} \in \mathcal{K}$:

$$\begin{aligned} \mathbb{E}\, X(z)X(\tilde{z}) &= \kappa(z, \tilde{z}) \ , \\ \mathbb{E}\, Y(z)X(\tilde{z}) &= \frac{\partial \kappa}{\partial z}(z, \tilde{z}) \ , \\ \mathbb{E}\, X(z)Y(\tilde{z}) &= \frac{\partial \kappa}{\partial \tilde{z}}(z, \tilde{z}) \ , \\ \mathbb{E}\, Y(z)Y(\tilde{z}) &= \frac{\partial^2}{\partial z \partial \tilde{z}} \kappa(z, \tilde{z}) \ . \end{aligned} \tag{4.1.12}$$

Remark 4.1.7 *Due to the properties of the process (X, Y), $\mathbb{E}\, X(z)\overline{X(w)} = \kappa(z, \bar{w})$ (and similarly for the other cross-conjugate quantities $\mathbb{E}\, X(z)\overline{Y(w)}$, etc.) ; moreover, the quantities $\mathbb{E}\, X(z)Y(z)$ and $\mathbb{E}\, Y(z)Y(z)$ can be computed by considering the limits $\lim_{\tilde{z}\to z} \frac{\partial \kappa}{\partial z}(z, \tilde{z})$ and $\lim_{\tilde{z}\to z} \frac{\partial^2 \kappa}{\partial z \partial \tilde{z}}(z, \tilde{z})$. The covariance structure of the process (X, Y) is hence fully described.*

Lemma 4.1.1 is the cornerstone to the proof of Theorem 4.2.1 ; its proof is postponed to Appendix 4.3.2 and relies on the following proposition, of independent interest :

Proposition 4.1.1 *Assume that* **A1** *and* **A2** *hold true and denote by $\mathcal{S}$ the support of the probability distribution associated to the Stieltjes transform m. Then, for every $\varepsilon > 0$, $\ell \in \mathbb{N}^*$:*

$$\mathbb{P}\left(\sup_{\lambda \in \mathrm{eig}(\hat{\mathbf{R}}_\mathbf{N})} d(\lambda, \mathcal{S}) > \varepsilon \right) = \mathcal{O}\left(\frac{1}{N^\ell} \right) \ ,$$

where $d(\lambda, \mathcal{S}) = \inf_{x \in \mathcal{S}} |\lambda - x|$.

The proof of Proposition 4.1.1 is postponed to Appendix 4.3.1.

As $(u_N, u'_N) \xrightarrow[N,M\to\infty]{a.s.} (\underline{m}, \underline{m}')$, a straightforward corollary of Lemma 4.1.1 yields the convergence in distribution of (X_N, X'_N, u_N, u'_N) to $(X, Y, \underline{m}, \underline{m}')$. This concludes the proof of step (i).

Consider two families of contours $(\mathcal{R}_k)$ and $(\mathcal{R}'_k)$ as described in Theorem 4.2.1. Denote by

$$\mathcal{K} = \bigcup_{k=1:L} \mathcal{R}_k \cup \bigcup_{k=1:L} \mathcal{R}'_k \ . \tag{4.1.13}$$

A direct consequence of Lemma 4.1.1 yields that $(X_N, X'_N, u_N, u'_N) : \mathcal{K} \to \mathbb{C}^4$ converges in distribution to the Gaussian process $(X, Y, \underline{m}, \underline{m}')$ with mean $(0, 0, \underline{m}, \underline{m}')$ and covariance structure inherited from the Gaussian process (X, Y). We are now in position to transfer the convergence of (X_N, X'_N, u_N, u'_N) to $\Upsilon_N(X_N, X'_N, u_N, u'_N)$ via the continuous mapping theorem, whose statement is reminded below.

Theorem 4.1.4 ([50, Th. 4.27]) *For any metric spaces S_1 and S_2, let ξ, $(\xi_n)_{n\geq 1}$ be random elements in S_1 with $\xi_n \xrightarrow[n\to\infty]{\mathcal{D}} \xi$ and consider some measurable mappings f, $(f_n)_{n\geq 1} : S_1 \mapsto S_2$ and a measurable set $\Gamma \subset S_1$ with $\xi \in \Gamma$ a.s. such that $f_n(s_n) \to f(s)$ as $s_n \to s \in \Gamma$. Then $f_n(\xi_n) \xrightarrow[n\to\infty]{\mathcal{D}} f(\xi)$.*

It remains to apply Theorem 4.2.5 to the process (X_N, X'_N, u_N, u'_N) and to the function Υ_N as defined in (4.1.11). Denote by [2]

$$\Upsilon(x, y, v, w) \;=\; \frac{1}{2\pi i c_k} \oint_{\mathcal{R}_k} z \left(\frac{\underline{m}(z) y(z) - w(z) x(z)}{\underline{m}(z) v(z)} \right) dz \ ,$$

and consider the set

$$\Gamma = \left\{ (x, y, v, w) \in C^4(\mathcal{K}, \mathbb{C}), \inf_{\mathcal{K}} |v| > 0, \quad \text{and} \quad x, y, w \text{ are continuous over } \mathcal{K} \right\}.$$

It is obvious that X, Y and $\underline{m}'$ are continuous over $\mathcal{K}$. Then, it is shown in [5, Section 9.12.1] that $\inf_{\mathcal{K}} |\underline{m}| > 0$, and, by a dominated convergence theorem argument, that $(x_N, y_N, v_N, w_N) \to (x, y, v, w) \in \Gamma$ implies that $\Upsilon_N(x_N, y_N, v_N, w_N) \to \Upsilon(x, y, v, w)$. Therefore, Theorem 4.2.5 applies to $\Upsilon_N(x_N, y_N, v_N, w_N)$ and the following convergence holds true :

$$\Upsilon_N(X_N, X'_N, u_N, u'_N) \xrightarrow[M,N\to\infty]{\mathcal{D}} \Upsilon(X, Y, \underline{m}, \underline{m}') \ ,$$

and step (ii) is established.

It now remains to prove step (iii), *i.e.* to check the Gaussianity of the random variable $\Upsilon(X, Y, \underline{m}, \underline{m}')$ and to compute the covariance between $\Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{C}_k)$ and $\Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{C}_\ell)$.

In order to propagate the Gaussianity of the deviations in the integrands of (4.1.9) to the fluctuations of the integral which defines $\hat{\rho}_k$, it suffices to notice that the integral can be written as the limit of a finite Riemann sum and that a finite Riemann sum of Gaussian random variables is still Gaussian. Therefore $M(\hat{\rho}_k - \rho_k)$ converges to a real Gaussian distribution (notice that $\Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{R}_k)$, being the limiting distribution of the real random variable Υ_N, is real as well). The same argument applies to the whole vector $(n(\hat{\rho}_k - \rho_k); 1 \leq k \leq L)$, which hence converges toward a Gaussian vector $\boldsymbol{\Upsilon} \sim \mathcal{N}_L(\mathbf{m}, \boldsymbol{\Theta})$:

$$\begin{pmatrix} M(\hat{\rho}_1 - \rho_1) \\ \vdots \\ M(\hat{\rho}_L - \rho_L) \end{pmatrix} \xrightarrow[M,N\to\infty]{\mathcal{D}} \boldsymbol{\Upsilon} \triangleq \begin{pmatrix} \Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{R}_1) \\ \vdots \\ \Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{R}_L) \end{pmatrix},$$

where $\mathbf{m}$ is a $L \times 1$ vector and $\boldsymbol{\Theta} = (\Theta_{k\ell})$ is a $L \times L$ covariance matrix.

2. As previously, we shall explicitly indicate the dependence on the contour $\mathcal{R}_k$ if needed and write $\Upsilon(x, x', u, u', \mathcal{R}_k)$.

As $\inf_{z\in\mathcal{K}} |\underline{m}(z)| > 0$, a straightforward application of Fubini's theorem together with the fact that $\mathbb{E}(X) = \mathbb{E}(Y) = 0$ yields :

$$\mathbb{E} \oint \left(z\frac{\underline{m}^{'}(z)X(z)}{\underline{m}^2(z)} - z\frac{Y(z)}{\underline{m}(z)} \right) dz = 0 \ ,$$

hence $\mathbf{m} = 0$.

It remains to compute the covariance between $\Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{R}_k)$ and $\Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{R}_\ell)$. As $\mathbf{\Theta} = \mathbb{E}(\mathbf{\Upsilon\Upsilon}^T)$, write :

$$\begin{aligned}
\Theta_{k\ell} &= \mathbb{E}\left(\Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{R}_k)\Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{R}_\ell)\right) \ , \\
&\overset{(a)}{=} \mathbb{E}\left(\Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{R}_k)\Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{R}'_\ell)\right) \ , \\
&= -\frac{1}{4\pi^2 c_k c_l}\mathbb{E}\left[\oint_{\mathcal{R}_k} z_1 \left(\frac{\underline{m}^{'}(z_1)X(z_1)}{\underline{m}^2(z_1)} - \frac{Y(z_1)}{\underline{m}(z_1)} \right) dz_1 \times \right. \\
&\quad \left. \oint_{\mathcal{R}'_\ell} z_2 \left(\frac{\underline{m}^{'}(z_2)X(z_2)}{\underline{m}^2(z_2)} - \frac{Y(z_2)}{\underline{m}(z_2)} \right) dz_2 \right] \ ,
\end{aligned}$$

where (a) follows from Remark 4.1.2 and enforces the fact that the contours are non-overlapping. Choosing non-overlapping contours will help us to compute the $\Theta_{k\ell}$'s by evaluating contour integrals with no singularities on the contours.

Write :

$$\begin{aligned}
\Theta_{k\ell} &= \mathbb{E}\left(\Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{R}_k)\Upsilon(X, Y, \underline{m}, \underline{m}', \mathcal{R}'_\ell)\right) \ , \\
&\overset{(a)}{=} -\frac{1}{4\pi^2 c_k c_l} \oint_{\mathcal{R}_k} \oint_{\mathcal{R}'_\ell} z_1 z_2 \left(\frac{\underline{m}^{'}(z_1)\underline{m}^{'}(z_2)\kappa(z_1, z_2)}{\underline{m}^2(z_1)\underline{m}^2(z_2)} \right. \\
&\quad \left. - \frac{\underline{m}^{'}(z_1)\partial_2\kappa(z_1, z_2)}{\underline{m}^2(z_1)\underline{m}(z_2)} - \frac{\underline{m}^{'}(z_2)\partial_1\kappa(z_1, z_2)}{\underline{m}(z_1)\underline{m}^2(z_2)} + \frac{\partial^2_{12}\kappa(z_1, z_2)}{\underline{m}(z_1)\underline{m}(z_2)} \right) dz_1 dz_2 \ ,
\end{aligned}$$

where (a) follows from Lemma 4.1.1 and the fact that $\inf_{z\in\mathcal{K}} |\underline{m}(z)| > 0$ together with Fubini's theorem, and $\partial_1, \partial_2, \partial^2_{12}$ respectively stand for $\partial/\partial z_1$, $\partial/\partial z_2$ and $\partial^2/\partial z_1 \partial z_2$. The above double integral is also well-defined as $\kappa(z_1, z_2)$ is well-defined and C^∞-differentiable over $\mathcal{R}_k \times \mathcal{R}'_\ell$. By integration by parts, we obtain

$$\begin{aligned}
&\oint \frac{z_1 z_2 \underline{m}'(z_2)\partial_1\kappa(z_1, z_2)}{\underline{m}(z_1)\underline{m}^2(z_2)} dz_1 \\
&= \oint \left(-\frac{z_2 \underline{m}'(z_2)\kappa(z_1, z_2)}{\underline{m}(z_1)\underline{m}^2(z_2)} + \frac{z_1 z_2 \underline{m}'(z_1)\underline{m}'(z_2)\kappa(z_1, z_2)}{\underline{m}^2(z_1)\underline{m}^2(z_2)} \right) dz_1 \ .
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\oint \frac{z_1 z_2 \underline{m}(z_2)\partial_{12}\kappa(z_1, z_2)}{\underline{m}(z_1)\underline{m}^2(z_2)} dz_1 \\
&= -\oint \frac{z_2 \partial_2\kappa(z_1, z_2)}{\underline{m}(z_1)\underline{m}(z_2)} dz_1 + \oint \frac{z_1 z_2 \underline{m}'(z_1)\partial_2\kappa(z_1, z_2)}{\underline{m}^2(z_1)\underline{m}(z_2)} dz_1 \ .
\end{aligned}$$

Hence

$$\Theta_{k\ell} = -\frac{1}{4\pi^2 c_k c_l}\left\{\oint_{\mathcal{R}_k}\oint_{\mathcal{R}'_\ell}\frac{z_2\underline{m}'(z_2)\kappa(z_1,z_2)}{\underline{m}(z_1)\underline{m}^2(z_2)}dz_1dz_2 - \oint_{\mathcal{R}_k}\oint_{\mathcal{R}'_\ell}\frac{z_2\partial_2\kappa(z_1,z_2)}{\underline{m}(z_1)\underline{m}(z_2)}dz_1dz_2\right\} .$$

Another integration by parts yields

$$\oint\frac{z_2\partial_2\kappa(z_1,z_2)}{\underline{m}(z_1)\underline{m}(z_2)}dz_2 = -\oint\frac{\kappa(z_1,z_2)}{\underline{m}(z_1)\underline{m}(z_2)}dz_2 + \oint\frac{z_2\underline{m}'(z_2)\kappa(z_1,z_2)}{\underline{m}(z_1)\underline{m}^2(z_2)}dz_2 .$$

Finally, we obtain :

$$\Theta_{k\ell} = -\frac{1}{4\pi^2 c_k c_\ell}\oint_{\mathcal{R}_k}\oint_{\mathcal{R}'_\ell}\frac{\kappa(z_1,z_2)}{\underline{m}(z_1)\underline{m}(z_2)}d\,z_1 d\,z_2 ,$$

and (4.1.10) is established.

4.1.5 Estimation of the covariance matrix

Theorem 4.1.3 describes the limiting performance of the estimator of Theorem 4.1.1, with an exact characterization of its variance. Unfortunately, the variance $\mathbf{\Theta}$ depends upon unknown quantities. We provide hereafter consistent estimates $\hat{\mathbf{\Theta}}$ for $\mathbf{\Theta}$ based on the observation $\hat{\mathbf{R}}_N$.

Theorem 4.1.5 *Assume that Assumptions* **A1** *and* **A2** *hold true, and recall the definition of* $\Theta_{k\ell}$ *given in* (4.1.10) *and Theorem 4.2.1. Let* $\hat{\Theta}_{k\ell}$ *be defined by*

$$\hat{\Theta}_{k\ell} = \frac{M^2}{N_k N_\ell}\left[\sum_{(i,j)\in\mathcal{N}_k\times\mathcal{N}_\ell,\ i\neq j}\frac{-1}{(\hat{\mu}_i-\hat{\mu}_j)^2 m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_j)} + \delta_{k\ell}\sum_{i\in\mathcal{N}_k}\left(\frac{m'''_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)}{6m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)^3} - \frac{m''_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)^2}{4m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)^4}\right)\right] , \quad (4.1.14)$$

where $(\mathcal{N}_k)$ *and* $(\hat{\mu}_k)$ *are defined in Theorem 4.1.1, then :*

$$\hat{\Theta}_{k\ell} - \Theta_{k\ell} \xrightarrow{a.s.} 0$$

as $N, M \to \infty$.

Theorem 4.1.5 is useful in practice as one can obtain simultaneously an estimate $\hat{\rho}_k$ of the values of ρ_k as well as an estimation of the degree of confidence for each $\hat{\rho}_k$.

Proof: In view of formula (4.1.10), taking into account the fact that $m_{\underline{\hat{\mathbf{R}}}_N}$ and $m'_{\underline{\hat{\mathbf{R}}}_N}$ are consistent estimates for $\underline{m}$ and $\underline{m}'$, it is natural to define $\hat{\Theta}_{k\ell}$ by replacing the unknown quantities $\underline{m}$ and

$\underline{m}'$ in (4.1.10) by their empirical counterparts $m_{\underline{\hat{\mathbf{R}}}_N}$ and $m'_{\underline{\hat{\mathbf{R}}}_N}$, hence we define $\hat{\Theta}_{k\ell}$ as

$$\begin{aligned}\hat{\Theta}_{k\ell} = -\frac{M^2}{4\pi^2 N_k N_\ell} \oint_{\mathcal{R}_k} \oint_{\mathcal{R}'_\ell} &\left(\frac{m'_{\underline{\hat{\mathbf{R}}}_N}(z_1) m'_{\underline{\hat{\mathbf{R}}}_N}(z_2)}{(m_{\underline{\hat{\mathbf{R}}}_N}(z_1) - m_{\underline{\hat{\mathbf{R}}}_N}(z_2))^2} - \frac{1}{(z_1 - z_2)^2} \right) \\ &\times \frac{1}{m_{\underline{\hat{\mathbf{R}}}_N}(z_1) m_{\underline{\hat{\mathbf{R}}}_N}(z_2)} d\,z_1 d\,z_2 \ .\end{aligned} \tag{4.1.15}$$

The proof of Theorem 4.1.5 now breaks down into two steps : The convergence of $\hat{\Theta}_{k\ell}$ to $\Theta_{k\ell}$, which relies on the definition (4.1.15) of $\hat{\Theta}_{k\ell}$ and on a dominated convergence argument, and the effective computation of the integral in (4.1.15) which relies on Cauchy's residue theorem [61], and yields (4.1.14).

We first address the convergence of $\hat{\Theta}_{k\ell}$ to $\Theta_{k\ell}$. Due to [9, 10], almost surely, the eigenvalues of $\hat{\mathbf{R}}_N$ will eventually belong to any ε-blow-up of the support $\underline{\mathcal{S}}$ of the probability measure associated to $\underline{m}$, *i.e.* the set $\{x \in \mathbb{R} : d(x, \underline{\mathcal{S}}) < \varepsilon\}$. Hence, if ε is small enough, the distance between these eigenvalues and any $z \in \mathcal{K}$ will be eventually uniformly lower-bounded. By [64, Lemma 1], the same holds true for the zeros of $m_{\underline{\hat{\mathbf{R}}}_N}$ (which are real). In particular, this implies that $m_{\underline{\hat{\mathbf{R}}}_N}$ is eventually uniformly lower-bounded on $\mathcal{K}$ (if not, then by compacity, there would exist $z \in \mathcal{K}$ such that $m_{\underline{\hat{\mathbf{R}}}_N}(z) = 0$ which yields a contradiction because all the zeroes of $m_{\underline{\hat{\mathbf{R}}}_N}$ are strictly within any contour). With these arguments at hand, one can easily apply the dominated convergence theorem and conclude that $\hat{\Theta}_{k\ell} \to \Theta_{k\ell}$ a.s.

We now evaluate the integral (4.1.15) by computing the residues of the integrand within $\mathcal{R}_k$ and $\mathcal{R}'_\ell$. There are two cases to discuss depending on whether $k \neq \ell$ and $k = \ell$. Denote by $h(z_1, z_2)$ the integrand in (4.1.15), that is :

$$h(z_1, z_2) = \left(\frac{m'_{\underline{\hat{\mathbf{R}}}_N}(z_1) m'_{\underline{\hat{\mathbf{R}}}_N}(z_2)}{(m_{\underline{\hat{\mathbf{R}}}_N}(z_1) - m_{\underline{\hat{\mathbf{R}}}_N}(z_2))^2} - \frac{1}{(z_1 - z_2)^2} \right) \times \frac{1}{m_{\underline{\hat{\mathbf{R}}}_N}(z_1) m_{\underline{\hat{\mathbf{R}}}_N}(z_2)}. \tag{4.1.16}$$

Note that, when z_2 is fixed, for $z_1 \to \hat{\lambda}_i$,

$$\frac{m'_{\underline{\hat{\mathbf{R}}}_N}(z_1)}{(m_{\underline{\hat{\mathbf{R}}}_N}(z_1) - m_{\underline{\hat{\mathbf{R}}}_N}(z_2))^2} \xrightarrow[z_1 \to \hat{\lambda}_i]{a.s.} M.$$

Then $h(z_1, z_2) \xrightarrow{a.s.} 0$ when $z_1 \to \hat{\lambda}_i$. Same result holds for $z_1 \to 0$. That is to say, $\hat{\lambda}_i$ and 0 are not poles of $h(z_1, z_2)$.

To apply the residue theorem, we first consider the case where $k \neq \ell$.

In this case, the two integration contours are different and never intersect (in particular, z_1 is always different from z_2). Let z_2 be fixed, and denote by $\hat{\mu}_i$ the zeros (labeled in increasing order) of $m_{\underline{\hat{\mathbf{R}}}_N}$, then the computation of the residue $\mathrm{Res}(h(\cdot, z_2), \hat{\mu}_i)$ of $h(\cdot, z_2)$ at a zero $\hat{\mu}_i$ of

$m_{\underline{\hat{\mathbf{R}}}_N}$ which is located within $\mathcal{R}_k$ is straightforward and yields

$$r(z_2) \triangleq \operatorname{Res}(h(\cdot, z_2), \hat{\mu}_i) = \left(\frac{m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)\, m'_{\underline{\hat{\mathbf{R}}}_N}(z_2)}{m^2_{\underline{\hat{\mathbf{R}}}_N}(z_2)} - \frac{1}{(\hat{\mu}_i - z_2)^2} \right) \frac{1}{m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i) m_{\underline{\hat{\mathbf{R}}}_N}(z_2)} \,. \tag{4.1.17}$$

Similarly, if one computes $\operatorname{Res}(r, \hat{\mu}_j)$ at a zero $\hat{\mu}_j$ of $m_{\underline{\hat{\mathbf{R}}}_N}$ located within $\mathcal{R}'_\ell$, one obtains :

$$\operatorname{Res}(r, \hat{\mu}_j) = -\frac{1}{(\hat{\mu}_i - \hat{\mu}_j)^2\, m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i) m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_j)} \,.$$

As stated in the following proposition, let $z_2 \in \mathcal{R}'_\ell \backslash \mathbb{R}$, the set $\mathcal{R}_{z_2} = \{z_1 \in \mathbb{C} : z_1 \neq z_2, m_{\underline{\hat{\mathbf{R}}}_N}(z_1) = m_{\underline{\hat{\mathbf{R}}}_N}(z_2) \neq 0\}$ is eventually empty a.s. for all N, M large, and if $z_2 \in \mathcal{R}'_\ell \cap \mathbb{R}$, this set is not empty, however, the integration with respect to z_2 for this residue is zero because the set $\mathcal{R}'_\ell \cap \mathbb{R}$ only contains two points, hence the residue in this set has not to be counted.

Proposition 4.1.2 *Let Assumptions* **A1** *and* **A2** *hold true, then for $z_2 \in \mathcal{R}'_\ell \backslash \mathbb{R}$, almost surely,*

$$\mathcal{R}_{z_2} = \{z_1 \in \mathbb{C} : z_1 \neq z_2, m_{\underline{\hat{\mathbf{R}}}_N}(z_1) = m_{\underline{\hat{\mathbf{R}}}_N}(z_2) \neq 0\} = \varnothing$$

for all N, n large.

The proof of Proposition 4.1.2 is postponed to Appendix 4.3.3.

It remains to count the number of $\hat{\mu}_i$ within each contour. By [64, Lemma 1], eventually, there are exactly as many $\hat{\mu}_i$ as eigenvalues within each contour, hence the result in the case $k \neq \ell$:

$$\hat{\Theta}_{k\ell} = \frac{M^2}{N_k N_\ell} \sum_{(i,j)\in\mathcal{N}_k\times\mathcal{N}_\ell} -\frac{1}{(\hat{\mu}_i - \hat{\mu}_j)^2 m'_{\underline{\hat{\mathbf{R}}}_N}(\mu_i) m'_{\underline{\hat{\mathbf{R}}}_N}(\mu_j)} \,.$$

We now compute the integral (4.1.15) in the case where $k = \ell$, and begin by the computation of the residues at $\hat{\mu}_i$. The definition (4.1.17) of r and the computation of $\operatorname{Res}(r, \hat{\mu}_j)$ still hold true in the case where $\hat{\mu}_j$ is within $\mathcal{R}_k$ but different from $\hat{\mu}_i$. It remains to compute $\operatorname{Res}(r, \hat{\mu}_i)$. Taking $z_2 \to \mu_i$, we get :

$$\begin{aligned}
\lim_{z_2\to\hat{\mu}_i} (z_2 - \hat{\mu}_i)^3 \left(\frac{1}{m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i) m_{\underline{\hat{\mathbf{R}}}_N}(z_2)(\hat{\mu}_i - z_2)^2} \right) &= \frac{1}{m'^2_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)}, \\
\lim_{z_2\to\hat{\mu}_i} (z_2 - \hat{\mu}_i)^2 \left(\frac{1}{m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i) m_{\underline{\hat{\mathbf{R}}}_N}(z_2)(\hat{\mu}_i - z_2)^2} \frac{1}{m'_{\underline{\hat{\mathbf{R}}}_N}{}^2(\hat{\mu}_i)(z_2 - \hat{\mu}_i)^3} \right) &= -\frac{m''_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)}{2m'_{\underline{\hat{\mathbf{R}}}_N}{}^3(\hat{\mu}_i)} \,.
\end{aligned}$$

Finally,

$$\begin{aligned}
&\lim_{z_2\to\hat{\mu}_i} (z_2 - \hat{\mu}_i) \left(\frac{1}{m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i) m_{\underline{\hat{\mathbf{R}}}_N}(z_2)(\hat{\mu}_i - z_2)^2} - \frac{1}{m'^2_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)(z_2 - \hat{\mu}_i)^3} + \frac{m''_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)}{2m'^3_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)(z_2 - \hat{\mu}_i)^2} \right) \\
&= \frac{m'''_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)}{6m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)^3} - \frac{m''_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)^2}{4m'_{\underline{\hat{\mathbf{R}}}_N}(\hat{\mu}_i)^4}.
\end{aligned}$$

Hence the residue :

$$\mathrm{Res}(r,\hat{\mu}_i) = \frac{m'''_{\underline{\mathbf{R}}_N}(\hat{\mu}_i)}{6m'_{\underline{\mathbf{R}}_N}(\hat{\mu}_i)^3} - \frac{m''_{\underline{\mathbf{R}}_N}(\hat{\mu}_i)^2}{4m'_{\underline{\mathbf{R}}_N}(\hat{\mu}_i)^4}\ .$$

There are two other cases that should be taken into account for the computation of the integral : The set $\mathcal{R}_{z_2}$, and the residue for $z_1 = z_2$. The first case can be handled as before. For $z_1 = z_2$, note that

$$\frac{m'_{\hat{\underline{\mathbf{R}}}_N}(z_1)m'_{\hat{\underline{\mathbf{R}}}_N}(z_2)}{(m_{\hat{\underline{\mathbf{R}}}_N}(z_1) - m_{\hat{\underline{\mathbf{R}}}_N}(z_2))^2} \xrightarrow[z_1\to z_2]{a.s.} 1.$$

$z_1 = z_2$ is not the residue for this term. It remains to compute $\frac{1}{(z_1-z_2)^2}\frac{1}{m_{\hat{\underline{\mathbf{R}}}_N}(z_1)m_{\hat{\underline{\mathbf{R}}}_N}(z_2)}$ for the residue $z_1 = z_2$. The integration by parts formula yields that :

$$\oint \frac{1}{(z_1-z_2)^2}\frac{dz_1}{m_{\hat{\underline{\mathbf{R}}}_N}(z_1)m_{\hat{\underline{\mathbf{R}}}_N}(z_2)} = \oint -\frac{m'_{\hat{\underline{\mathbf{R}}}_N}(z_1)}{(z_1-z_2)}\frac{dz_1}{m^2_{\hat{\underline{\mathbf{R}}}_N}(z_1)m_{\hat{\underline{\mathbf{R}}}_N}(z_2)}.$$

Then the residue for $z_1 = z_2$ is :

$$-\frac{m'_{\hat{\underline{\mathbf{R}}}_N}(z_2)}{m^3_{\hat{\underline{\mathbf{R}}}_N}(z_2)}.$$

This is the derivative function of $\frac{1}{2m^2_{\hat{\underline{\mathbf{R}}}_N}(z_2)}$, then the integration with respect to z_2 is zero.

Finally both have a null contribution, hence the formula :

$$\hat{\Theta}_{kk} = \frac{M^2}{N_k^2}\left[\sum_{(i,j)\in\mathcal{N}_k^2,\ i\neq j}\frac{-1}{(\hat{\mu}_i-\hat{\mu}_j)^2 m'_{\hat{\underline{\mathbf{R}}}_N}(\hat{\mu}_i)m'_{\hat{\underline{\mathbf{R}}}_N}(\hat{\mu}_j)} + \sum_{i\in\mathcal{N}_k}\left(\frac{m'''_{\hat{\underline{\mathbf{R}}}_N}(\hat{\mu}_i)}{6m'_{\hat{\underline{\mathbf{R}}}_N}(\hat{\mu}_i)^3} - \frac{m''_{\hat{\underline{\mathbf{R}}}_N}(\hat{\mu}_i)^2}{4m'_{\hat{\underline{\mathbf{R}}}_N}(\hat{\mu}_i)^4}\right)\right]\ .$$

□

4.1.6 Performance of cognitive radios

We consider the system model introduced in Section 4.1.1. Assuming the spectrum of $\mathbf{R}_M$ allows one to clearly distinguish the successive clusters (as in Figure 4.1), Proposition 4.1.1 enables the detection of primary transmitters and the estimation of their transmit powers $P_1, \ldots, P_K$; this boils down to estimating the largest K eigenvalues of $\mathbf{WPW}^{\mathsf{H}} + \sigma^2\mathbf{I}_N$, i.e. the $P_k + \sigma^2$, and to subtract σ^2 (optionally estimated from the smallest eigenvalue of $\mathbf{WPW}^{\mathsf{H}} + \sigma^2\mathbf{I}_N$ if $n < N$). Call $\hat{P}_k$ the estimate of P_k.

Based on these power estimates, the sensor can determine the optimal coverage for secondary communications that ensures no interference to the primary network. A basic idea for instance is to ensure that the closest primary user, i.e. that with strongest received power, is not interfered. Our interest is then cast on P_K. Now, since the power estimator is imperfect, it is hazardous for

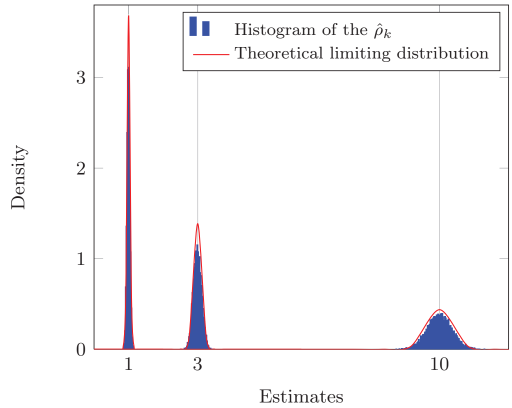

FIGURE 4.3 – Comparison of empirical against theoretical variances, based on Theorem 4.2.1, for three eigenvalues, $\rho_1 = 1$, $\rho_2 = 3$, $\rho_3 = 10$, $N_1 = N_2 = N_3 = 20$, $N = 60$, $M = 600$ and SNR= 20 dB.

the secondary network to state that K has power $\hat{P}_K$ or to add some empirical security margin to $\hat{P}_K$. The results of Section 4.1.4 partially answer this problem.

Theorems 4.1.3 and 4.1.5 enable the secondary sensor to evaluate the accuracy of $\hat{P}_k$. In particular, assume that the cognitive radio protocol allows the secondary network to interfere the primary network with probability q and denote A the value

$$A \triangleq \inf_a \{\Pr(P_K - \hat{P}_K > a) \leq q\}.$$

According to Theorem 4.1.3, for N, M large, A is well approximated by $\hat{\Theta}_{K,K} Q^{-1}(q)$, with Q the Gaussian *Q-function*. If the sensor detects a user with power P_K, estimated by $\hat{P}_K$, $\Pr(\hat{P}_K + A < P_K) < q$ and then it is safe for the secondary network to assume the worst case scenario where user K transmits at power $\hat{P}_K + A \simeq \hat{P}_K + \hat{\Theta}_{K,K} Q^{-1}(q)$.

In Figure 4.3, the performance of Theorem 4.1.3 is compared against $10,000$ Monte Carlo simulations of a scenario of three users, with $n_1 = n_2 = n_3 = 20$, $N = 60$ and $M = 600$. It appears that the limiting distribution is very accurate for these values of N, M. We also performed simulations to obtain empirical estimates $\hat{\Theta}_{k,k}$ of $\Theta_{k,k}$ from Theorem 4.1.5, which suggest that $\bar{\Theta}_{k,k}$ is an accurate estimator as well.

4.2 Estimation of Parametrized Covariance Matrices

This section is based on the paper [101].

Estimating the covariance matrix of a series of independent multivariate observations is a crucial issue in many signal processing applications. A reliable estimate of the covariance matrix is for instance needed in principal component analysis [48], direction of arrival estimation for antenna arrays [80], blind subspace estimation [67], capacity estimation [33], estimation/detection procedures [80, 95], etc.

In the case where the dimension N of the observations is small compared to the number M

of observations, the empirical covariance matrix based on the observations often provides a good estimate for the unknown covariance matrix. This estimate becomes however much less accurate, and even not consistent with the dimension N getting higher (see for instance [65, Theorem 2]).

An interesting theoretical framework for modern estimation of multi-dimensional variables occurs whenever the number of available samples M grows at the same pace as the dimension N of the considered variables. Shifting to this new assumption induces fundamental differences in the behavior of the empirical covariance matrix as analyzed in Mestre's work [65, 64]. Recently, several attempts have been done to address this problem (cf. [65, 64, 54, 4, 51, 57]) using large random matrix theory which proposed powerful tools, mainly spurred by Girko's G-estimators [34], to cope with this new context. In [65, 64], Mestre considers the eigenvalue estimation of a parametrized model of covariance matrices similar to the model we shall study in this article. In [54] and [57], grid-based techniques for inverting the Marčenko-Pastur equation are proposed. In [51], the problem of estimating a specific linear functional of the eigenvalues of an unknown covariance matrix is addressed. In [4], the eigenvalues of an unknown parametrized covariance matrix are estimated by resorting on the empirical moments of the observations. This technique, which goes back to Pisarenko's ideas [76], will be also combined to large random matrix theory here.

We shall consider the case where the dimension of each observation N together with the number of samples M go to infinity at the same pace, *i.e.* their ratio converges to some nonnegative constant $c > 0$. In order to present the contribution provided in this paper, let us describe the model under study.

Consider an $N \times M$ matrix $\mathbf{X}_N = (X_{ij})$ whose entries are independent and identically distributed (i.i.d.) random variables. Let $\mathbf{R}_N$ be an $N \times N$ Hermitian matrix with L (L being fixed and known) distinct eigenvalues $0 < \rho_1 < \cdots < \rho_L$ with respective multiplicities $N_1, \cdots, N_L$ (notice that $\sum_{i=1}^{L} N_i = N$). Consider now

$$\mathbf{Y}_N = \mathbf{R}_N^{1/2} \mathbf{X}_N .$$

The matrix $\mathbf{Y}_N = [\mathbf{y}_1, \cdots, \mathbf{y}_M]$ is the concatenation of M independent observations, where each observation writes $\mathbf{y}_i = \mathbf{R}_N^{1/2} \mathbf{x}_i$ with $\mathbf{X}_N = [\mathbf{x}_1, \cdots, \mathbf{x}_M]$. In particular, the covariance matrix of each observation $\mathbf{y}_i$ is $\mathbf{R}_N = \mathbb{E}\mathbf{y}_i \mathbf{y}_i^H$ (matrix $\mathbf{R}_N$ is sometimes called the population covariance matrix).

In this chapter, we consider the problem of estimating individually the eigenvalues ρ_i as well as their multiplicities N_i in the case where the total number of eigenvalues is fixed and known.

Such a scenario is customary in applications for wireless communications. A relevant example concerns uplink CDMA systems operating over flat fading channels, where users are arranged into L classes, each class corresponding to a distinct power amount. In this case, matrix $\mathbf{Y}_N$ can be modeled as :

$$\mathbf{Y}_N = \mathbf{W}\mathbf{P}^{\frac{1}{2}}\mathbf{X} + \sigma \mathbf{V}_N$$

where $\mathbf{W}$ are $\mathbf{V}_N$ represent respectively the signature matrix assumed to be orthogonal and the noise matrix, while $\mathbf{P}$ is diagonal with diagonal elements taking distinct values among the finite set $\{\rho_1, \cdots, \rho_L\}$. In a decentralized context, where each user selects its own power from this finite set according to a defined control energy strategy, the base station which stands for the receiver can have to estimate the number of users in each class as well as their corresponding powers. Obviously, this problem amounts to estimating the eigenvalues of the theoretical covariance matrix as well as their corresponding multiplicities. Similar scenarios are studied in [100, 99].

Among the proposed parametric techniques, we cite the one developed by Mestre [64] and taken up by Vallet *et al* [93] and Couillet *et al* [26] for more elaborated models. Although being computationally efficient, this technique requires a *separability condition*, namely the assumption that the number of samples is large compared to the dimension of each sample (small limiting ratio $c = \lim \frac{N}{M} > 0$). In such a case, the limiting spectrum of the empirical covariance matrix possesses as many clusters[3] as there are eigenvalues to be estimated, and each eigenvalue can be estimated by a contour integral surrounding the related cluster. Mestre's technique cannot be applied any more in the case where c is larger (which reflects a higher dimension of the observations relatively to the sample dimension). In fact, the dimension of the clusters may grow and neighbouring clusters may merge, violating the one-to-one correspondence between clusters and eigenvalues to be estimated (see for instance Fig. 4.1 and 4.2).

A way to circumvent the separability condition has recently been proposed by Bai, Chen and Yao [4], based on the use of the empirical asymptotic moments :

$$\hat{\alpha}_k = \frac{1}{M}\mathrm{Tr}\left(\mathbf{Y}_N\mathbf{Y}_N\right)^k, k \in \{1, \cdots, 2L\},$$

which can be shown to be a sufficient statistic to estimate $\left(\frac{N_1}{N}, \cdots, \frac{N_L}{N}, \rho_1, \cdots, \rho_L\right)$. Although being robust to separability condition, this technique suffers from numerical difficulties, since the proposed estimator has no closed-form expression and thus should be determined numerically. An interesting contribution, although not directly focused on estimating the covariance of the observations is the work by Rubio and Mestre [78], where an alternative way to estimate the moments

$$\gamma_k = \frac{1}{N}\mathrm{Tr}(\mathbf{R}_N^k),$$

for all $k \in \mathbb{N}$ is proposed, yielding an explicit (yet lengthy) formula.

In this chapter, we improve existing work in several directions : With respect to Mestre's seminal papers [65, 64], we propose a joint estimation of the eigenvalues and their multiplicities, and drop the separability assumption. The proposed estimator is close in spirit to the one developed by Bai *et al.* in [4], although we carefully establish the existence and uniqueness of the estimator, which is not explicit in [4]. Comparisons on the relative numerical efficiency of both procedures is provided in the simulations section. Finally, we study the fluctuations of the estimator and establish a Central Limit Theorem (CLT).

The remainder of this chapter is organized as follows. In Section 4.2.1, the main assumptions are provided and Mestre's estimator [64] is briefly reviewed. In Section 4.2.2, the proposed estimator is described. Its fluctuations are studied in Section 4.2.3, where a CLT is stated. Si-

3. By cluster, we mean a connected component of the support of the limiting probability distribution of the spectrum.

mulations are presented in Section 4.2.4. Finally, the remaining technical details are postponed to the Appendices.

4.2.1 Main assumptions and general background

Consider the model

$$\mathbf{Y}_N = \mathbf{R}_N^{1/2}\mathbf{X}_N,$$

and

$$\hat{\mathbf{R}}_N = \frac{1}{M}\mathbf{Y}_N\mathbf{Y}_N^H.$$

At first, an assumption about the matrix $\mathbf{R}_N$ is needed :

Assumption 4.2.1 *$\mathbf{R}_N$ is an $N \times N$ Hermitian non-negative definite matrix with L (L being fixed and known) distinct eigenvalues $0 < \rho_1 < \cdots < \rho_L$ with respective multiplicities $N_1, \cdots, N_L$ (notice that $\sum_{i=1}^{L} N_i = N$).*

As mentioned earlier, we consider the asymptotic regime where the number of samples M and the dimension N grow to infinity at the same pace, together with the multiplicities of each eigenvalue of $\mathbf{R}_N$.

Assumption 4.2.2 *Let M, N be integers such that :*

$$N, M \to \infty\ , \quad \textit{with} \quad \frac{N}{M} \to c \in (0,\infty)\ , \quad \textit{and} \quad \frac{N_i}{N} \to c_i \in (0,\infty)\ ,\ 1 \le i \le L. \tag{4.2.1}$$

This assumption will be shortly referred to as $N, M \to \infty$.

The following assumption is standard and is sufficient for estimation purposes.

Assumption 4.2.3 *Let $\mathbf{X}_N = (X_{ij})$ be a $N \times M$ matrix whose entries are i.i.d. random variables in $\mathbb{C}$ such that $\mathbb{E}(\mathbf{X}_{1,1}) = 0$, $\mathbb{E}(|\mathbf{X}_{1,1}|^2) = 1$ with finite fourth moment : $\mathbb{E}(|\mathbf{X}_{1,1}|^4) < \infty$.*

Remark 4.2.1 *In order to establish the fluctuations of the estimators, the Gaussianity of the entries of $\mathbf{X}_N$ is needed (although this technical condition may be removed with substantial extra work).*

Assumption 3b : The entries of the $N \times M$ matrix $\mathbf{X}_N = (X_{ij})$ are i.i.d. standard complex Gaussian variables, *i.e.* $X_{ij} = U + \mathbf{i}V$, where U, V are both independent real Gaussian random variables $\mathcal{N}(0, \frac{1}{2})$.

It is well-known in large random matrix theory that under Assumptions 4.2.1, 4.2.2 and 4.2.3, $F^{\hat{\mathbf{R}}_N}$ converges to a limiting probability distribution (cf. Theorem 4.1.2). In Mestre's paper [64], a *separability condition* (cf. Section 4.1.1) is needed in order to derive the estimator of $\mathbf{R}_N$'s eigenvalues :

Assumption 4.2.4 *The support $\mathcal{S}$ of the limiting probability distribution of $F^{\hat{\mathbf{R}}_N}$ is composed of L compact connex disjoint subsets, and not reduced to a singleton. (cf. Section **??** and Fig. 4.1 and 4.2)*

Remark 4.2.2 *Note that when $M < N$, matrix $\hat{\mathbf{R}}_N$ is singular and thus admits $(N-M)$ eigenvalues equal to zero. Hence, the limiting spectrum of $\hat{\mathbf{R}}_N$ has an additional mass in zero with weight $1-\frac{1}{c}$, which will not be considered among the L clusters.*

Recall the notations in Theorem 4.1.2 with the Stieljes transform $m(z)$ (resp. $\underline{m}(z)$) of the LSD of $\hat{\mathbf{R}}_N$ (resp. $\hat{\underline{\mathbf{R}}}_N = \frac{1}{M}\mathbf{Y}_N^H\mathbf{Y}_N$) and the deterministic equivalent $m_N(z)$ (resp. $\underline{m}_N(z)$). In [64], Mestre proposes a novel approach to estimate the eigenvalues $(\rho_k;\ 1\le k\le L)$ of the population covariance matrix based on the observations $\hat{\mathbf{R}}_N$ under the additional Assumption 4.2.4. His approach relies on large random matrix theory and the separability condition presented above plays a major role in the mere definition of the estimators. Recall the estimator (cf. Theorem 4.1.1)

$$\hat{\rho}_k - \rho_k \xrightarrow[M,N\to\infty]{a.s.} 0 ,$$

where

$$\hat{\rho}_k = \frac{n}{N_k}\sum_{m\in\mathcal{N}_k}\left(\hat{\lambda}_m - \hat{\mu}_m\right) .$$

A CLT has been derived in Theorem 4.1.3 for this estimator under the extra assumption that the entries of $\mathbf{X}_N$ are Gaussian. To unify the notation, recall again the theorem :

Theorem 4.2.1 ([99]) *With the same notations as before, under Assumptions 4.2.1, 4.2.2, 3b, 4.2.4 and with known multiplicities $N_1,\cdots,N_L$, then :*

$$(M(\tilde{\rho}_k-\rho_k),\ 1\le k\le L) \xrightarrow[M,N\to\infty]{\mathcal{D}} \mathbf{x}\sim\mathcal{N}_L(0,\mathbf{\Theta}) ,$$

where $\mathcal{N}_L$ refers to a real L-dimensional Gaussian distribution, and $\mathbf{\Theta}$ is a $L\times L$ matrix whose entries $\Theta_{k\ell}$ are given by,

$$\Theta_{k\ell} = -\frac{1}{4\pi^2c^2c_kc_\ell}\oint_{\mathcal{C}_k}\oint_{\mathcal{C}_\ell}\left[\frac{\underline{m}'(z_1)\underline{m}'(z_2)}{(\underline{m}(z_1)-\underline{m}(z_2))^2}-\frac{1}{(z_1-z_2)^2}\right]\frac{1}{\underline{m}(z_1)\underline{m}(z_2)}dz_1dz_2 ,$$

where $\mathcal{C}_k$ (resp. $\mathcal{C}_\ell$) is a closed counterclockwise oriented contour which only contains the k-th cluster (resp. ℓ-th) .

The main objective of this section is to provide estimators for the ρ_k's without relying any more on the separability condition (*i.e.* to remove Assumption 4.2.4). A Central Limit Theorem will be established as well for the proposed estimator. As a by-product, the knowledge of the multiplicities will no longer be needed, and they will be estimated as well.

4.2.2 Estimation of the eigenvalues ρ_i

In this section, we provide a method to estimate consistently the eigenvalues of the population covariance matrix and their multiplicities without the need of the separability condition (cf. Fig. 4.2). Our method is based on the asymptotic evaluation of the moments of the eigenvalues of $\mathbf{R}_N$,

$$\gamma_i \triangleq \frac{1}{N}\mathrm{Tr}\mathbf{R}_N^i = \sum_{k=1}^{L} \frac{N_k}{N}\rho_k^i, \quad 1 \leq i \leq 2L-1. \tag{4.2.2}$$

If $(\widehat{m}_i)_{1\leq i\leq 2L-1}$ are the empirical moments of the sample eigenvalues, then it is well-known that except for $i=1$, γ_i cannot be approximated by $\widehat{m}_i$. Consistent estimators for γ_i are provided in [78], where it has been proved that :

$$\gamma_i - \tilde{\gamma}_i \xrightarrow[N,M\to+\infty]{} 0,$$

where

$$\tilde{\gamma}_i = \sum_{l=1}^{i} \mu_S(l,i)\widehat{m}_l, \tag{4.2.3}$$

$\mu_S(l,i)$ being some given coefficients that depend on the system dimensions and on the empirical moments $\widehat{m}_i$ [78]. An alternative is to use the Stieltjes transform :

Lemma 4.2.1 *Let Assumptions 4.2.1, 4.2.2 and 4.2.3 hold true. Let $\hat{\gamma}_i$ be the real quantities given by :*

$$\begin{cases} \hat{\gamma}_0 &= 1, \\ \hat{\gamma}_1 &= -\frac{M}{2N\mathbf{i}\pi}\oint_{\mathcal{C}} \frac{zm'_{\underline{\hat{\mathbf{R}}}_N}(z)}{m_{\underline{\hat{\mathbf{R}}}_N}(z)}dz, \\ \hat{\gamma}_k &= \frac{M(-1)^k}{2Nk\mathbf{i}\pi}\oint_{\mathcal{C}} \frac{dz}{m^k_{\underline{\hat{\mathbf{R}}}_N}(z)}, \quad \text{for } 2\leq k\leq 2L-1 \end{cases}$$

where $\mathcal{C}$ is a counterclockwise oriented contour which encloses the support $\mathcal{S}$ of the limiting distribution of the eigenvalues of $\underline{\hat{\mathbf{R}}}_N$. Let γ_i be the theoretical moments as given in (4.2.2). Then, for $1\leq i\leq 2L-1$,

$$\hat{\gamma}_i - \gamma_i \xrightarrow[N,M\to\infty]{a.s.} 0\ .$$

The proof of this lemma is postponed to Appendix 4.3.4. While the estimates proposed by [78] are better in practice, estimates $(\hat{\gamma}_i)$ will be of interest in order to establish the Central Limit Theorem, and to obtain a closed-form expression of the asymptotic variance.

An interesting remark is that the map that links the eigenvalues and their multiplicities to their first $2L-1$ moments is invertible. Retrieving the eigenvalues from the estimates of the $2L-1$ moments is thus possible. This is the basic idea on which our method is founded.

The main result is stated as below :

Theorem 4.2.2 *Let Assumptions 4.2.1, 4.2.2, 4.2.3 hold true and let $(\hat{\gamma}_k, 1\leq k\leq 2L-1)$ be* *as in Lemma 4.2.1. Consider the following system of equations :*

$$\begin{cases} \sum_{i=1}^{L} x_i = 1, \\ \sum_{i=1}^{L} x_i y_i^k = \hat{\gamma}_k \quad \textit{for } 1 \leq k \leq 2L-1, \end{cases} \tag{4.2.4}$$

where $(x_i)_{1\leq i\leq L}$ and $(y_i)_{1\leq i\leq L}$ are $2L$ unknown parameters. Then for N, M large enough, the system of equations (4.2.4) has one and only one real solution $(\hat{c}_1, \cdots, \hat{c}_L, \hat{\rho}_1, \cdots, \hat{\rho}_L)$ with $\hat{\rho}_1 < \cdots < \hat{\rho}_L$. Moreover, $(\hat{c}_1, \cdots, \hat{c}_L, \hat{\rho}_1, \cdots, \hat{\rho}_L)$ is a consistent estimator of $(c_1, \cdots, c_L, \rho_1, \cdots, \rho_L)$, i.e.,

$$\hat{c}_\ell - c_\ell \xrightarrow[N,M\to\infty]{a.s.} 0 \quad \textit{and} \quad \hat{\rho}_\ell - \rho_\ell \xrightarrow[N,M\to\infty]{a.s.} 0,$$

with $c_\ell = \lim \frac{N_\ell}{N}$ for $1 \leq \ell \leq L$.

Remark 4.2.3 *The condition of separability is not required in the previous theorem. Moreover, the multiplicities are assumed to be unknown and thus have to be estimated. Fig. 4.2 represents a case where the three clusters are merged into one cluster. In such a situation, the estimator in [64] is biased whereas the proposed one is asymptotically consistent.*

Remark 4.2.4 *We use the estimator proposed in Lemma 4.2.1. However, the proof below does not depend on the estimator of the moments we choose. In fact, for any consistent estimator of the moments γ_i, the above theorem always holds true.*

Proof of Theorem 4.2.2

The proof can be split into two main steps. By using the inverse function theorem, we can prove the almost sure existence of a real solution. Then, the uniqueness is ensured by a matrix inversion argument.

Existence of a real solution of the system

The first task is to show that the system of equations (4.2.4) admits, for N sufficiently large, one real solution $(\hat{c}_1, \cdots, \hat{c}_L, \hat{\rho}_1 \cdots, \hat{\rho}_L)$ satisfying $\hat{\rho}_1 < \hat{\rho}_2 < \cdots < \hat{\rho}_L$. We shall also establish the consistency of the obtained solution. The proof of the existence of a real solution follows in the same way as in [4]. It is merely based on the use of the inverse function theorem which ensures the existence as soon as the Jacobian matrix of the considered transformation is invertible. We recall below the inverse function theorem [56] :

Theorem 4.2.3 ([56]) *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable function. Let $\mathbf{a}$ and $\mathbf{b}$ be vectors of $\mathbb{R}^n$ such that $f(\mathbf{a}) = \mathbf{b}$. If the Jacobian matrix of f at $\mathbf{a}$ is invertible, then there exists a neighborhood U containing $\mathbf{a}$ such that $f : U \to f(U)$ is a diffeomorphism, i.e, for every $\mathbf{y} \in f(U)$ there exists a unique $\mathbf{x}$ such that $f(\mathbf{x}) = \mathbf{y}$. In particular, f is invertible in U.*

Consider the functional f defined as :

$$f(x_1, \cdots, x_L, y_1, \cdots, y_L) = \left(\sum_{\ell=1}^{L} x_\ell \ , \ \sum_{\ell=1}^{L} x_\ell y_\ell \ , \ \cdots \ , \ \sum_{\ell=1}^{L} x_\ell y_\ell^{2L-1} \right).$$

Consider $\mathbf{z} = (x_1, \cdots, x_L, y_1, \cdots, y_L)$ and denote by $\mathbf{c} = (c_1, \cdots, c_L, \rho_1, \cdots, \rho_L)$; we then have :

$$\mathbf{M} \triangleq \frac{\partial f}{\partial \mathbf{z}}\Big|_{\mathbf{z}=\mathbf{c}} = \begin{bmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ \rho_1 & \cdots & \rho_L & c_1 & \cdots & c_L \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \rho_1^{2L-1} & \cdots & \rho_L^{2L-1} & (2L-1)c_1\rho_1^{2L-2} & \cdots & (2L-1)c_L\rho_L^{2L-2} \end{bmatrix}. \tag{4.2.5}$$

As proven in [4, Proposition 1], matrix $\mathbf{M}$ is invertible.

The inverse function theorem then applies. Denote by $\psi_i = \sum_{k=1}^{L} c_k\rho_k^i$ for $0 \leq i \leq 2L-1$. There exists a neighborhood U of $(c_1, \cdots, c_L, \rho_1, \cdots, \rho_L)$ and a neighborhood V of $(\psi_0, \cdots, \psi_{2L-1})$ such that f is a diffeomorphism from U onto V. On the other hand, we have :

$$\hat{\gamma}_i - \gamma_i \xrightarrow{a.s.} 0.$$

As $\gamma_i - \psi_i \to 0$, therefore, almost surely, $(\hat{\gamma}_0, \cdots, \hat{\gamma}_{2L-1}) \in V$ for N and M large enough. Hence, a real solution

$$(\hat{c}_1, \cdots, \hat{c}_L, \hat{\rho}_1, \cdots, \hat{\rho}_L) = f^{-1}(\hat{\gamma}_0, \cdots, \hat{\gamma}_{2L-1}) \in U$$

exists. And by the continuity, one can get easily that :

$$\hat{c}_\ell - c_\ell \xrightarrow[N,M\to\infty]{a.s.} 0 \quad \text{and} \quad \hat{\rho}_\ell - \rho_\ell \xrightarrow[N,M\to\infty]{a.s.} 0 \quad \text{for} \quad 1 \leq \ell \leq L\ .$$

Uniqueness of the solution of the system

Consider the polynomial Q with degree L defined as :

$$\mathrm{Q(X)} = \prod_{\ell=0}^{L}(X - \hat{\rho}_\ell) \triangleq \sum_{\ell=0}^{L} s_\ell \mathrm{X}^\ell$$

where $s_L = 1$. Denote by $\mathbf{s} = [s_0, \cdots, s_{L-1}]^T$. It is clear that $g : (\hat{\rho}_1, \cdots, \hat{\rho}_L) \to \mathbf{s}$ is a homeomorphism. It remains thus to show that vector $\mathbf{s}$ is uniquely determined by $(\hat{\gamma}_0, \cdots, \hat{\gamma}_{2L-1})$.

It is clear that each $\hat{\rho}_k$ is also the zero of the polynomial functions $\mathrm{R}_\ell(X)$ given by :

$$\mathrm{R}_\ell(\mathrm{X}) = \sum_{i=0}^{L} s_i \mathrm{X}^{i+\ell}\ ,$$

where $0 \leq \ell \leq L-1$. In other words, for $1 \leq k \leq L$, we get :

$$\sum_{i=0}^{L} s_i \hat{\rho}_k^{\ell+i} = 0,$$

or equivalently :

$$\sum_{i=0}^{L} s_i \hat{c}_k \hat{\rho}_k^{\ell+i} = 0. \tag{4.2.6}$$

Summing (4.2.6) over k, we obtain :

$$\sum_{i=0}^{L} \hat{\gamma}_{i+\ell} s_i = 0 \ , \tag{4.2.7}$$

for $0 \leq \ell \leq L-1$. Since $s_L = 1$, (4.2.7) becomes :

$$\hat{\gamma}_{L+\ell} + \sum_{i=0}^{L-1} s_i \hat{\gamma}_{i+\ell} = 0 \ , \tag{4.2.8}$$

for $0 \leq \ell \leq L-1$.

Writing (4.2.8) in a matrix form, we get : $\mathbf{\Gamma s} = -\mathbf{b}$, where

$$\mathbf{\Gamma} = \begin{bmatrix} \hat{\gamma}_0 & \hat{\gamma}_1 & \cdots & \hat{\gamma}_{L-1} \\ \hat{\gamma}_1 & \hat{\gamma}_2 & \cdots & \hat{\gamma}_L \\ \vdots & \ddots & \ddots & \vdots \\ \hat{\gamma}_{L-1} & \hat{\gamma}_L & \cdots & \hat{\gamma}_{2L-2} \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} \hat{\gamma}_L \\ \vdots \\ \hat{\gamma}_{2L-1} \end{bmatrix}. \tag{4.2.9}$$

On the other hand, we have $\mathbf{\Gamma} = \mathbf{ADA}^T$, where $\mathbf{D} = \mathrm{diag}(\hat{c}_1, \hat{c}_2, \cdots, \hat{c}_L)$ and

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \hat{\rho}_1 & \hat{\rho}_2 & \cdots & \hat{\rho}_L \\ \vdots & \vdots & & \vdots \\ \hat{\rho}_1^{L-1} & \hat{\rho}_2^{L-1} & \cdots & \hat{\rho}_L^{L-1} \end{bmatrix}. \tag{4.2.10}$$

Then,

$$\det(\mathbf{\Gamma}) = \prod_{k=1}^{L} \hat{c}_k \prod_{1 \leq i < j \leq L} (\hat{\rho}_i - \hat{\rho}_j)^2 > 0.$$

Therefore, the vector $\mathbf{s}$ is then uniquely determined by $\mathbf{\Gamma}$ and $\mathbf{b}$ and is given by :

$$\mathbf{s} = -\mathbf{\Gamma}^{-1}\mathbf{b}.$$

Hence the uniqueness. Proof of Theorem 4.2.2 is completed.

Summary of the main steps of the estimation procedure

The proof of the uniqueness shows that the solutions of the system of equations (4.2.4) can be directly obtained from the estimates of the first $2L-1$ moments. More precisely, the estimation of the eigenvalues and their corresponding multiplicities can be performed through the following steps :

---

1. Set $\hat{\gamma}_0$ to 1. Estimate the first $2L-1$ moments using (4.2.3). Coefficients $\mu_S(l, i)$ are computed using Eq. (46) in [78].
2. Construct the matrix $\mathbf{\Gamma}$ and $\mathbf{b}$ using (4.2.9).

3. `Compute the vector s as` $\mathbf{s} = -\mathbf{\Gamma}^{-1}\mathbf{b}$.
4. `Determine (by using for instance function roots of` MATLAB`) the roots` $\hat{\rho}_1, \cdots, \hat{\rho}_L$ `of the polynomial whose coefficients are given by vector s.`
5. `Construct matrix` $\mathbf{A}$ `as specified by` (4.2.10)`, and vector` $\mathbf{d} = [\hat{\gamma}_0, \cdots, \hat{\gamma}_{L-1}]^{\mathrm{T}}$.
6. `The coefficient estimates` $\hat{\mathbf{c}} = [\hat{c}_1, \cdots, \hat{c}_L]^{\mathrm{T}}$ `are thus given by:`

$$\hat{\mathbf{c}} = \mathbf{A}^{-1}\mathbf{d}.$$

Remark 4.2.5 *Note that while the existence of a real solution is only proven for M and N large enough, the previous algorithm always yield a solution, even for very small dimensions. However, in such scenarios, the validity of the obtained solution is not ensured. In fact, if N and M are not large enough, the moment estimates are not accurate, and the solution of the algorithm may yield complex or negative eigenvalues. This event completely disappears when N and M or only M take higher values. In practice, getting such inadequate solutions should warn that more samples are required.*

4.2.3 Fluctuations of the estimator

In this section, we shall study the fluctuations of the multiplicities and eigenvalues estimators $(\hat{c}_1, \cdots, \hat{c}_L, \hat{\rho}_1, \cdots, \hat{\rho}_L)$ introduced in Theorem 4.2.2. In particular, we establish a Central Limit Theorem for the whole vector in the case where the entries of matrix $\mathbf{X}_N$ are Gaussian.

Theorem 4.2.4 *Let Assumptions 4.2.1, 4.2.2, 3b hold true. Let $(\hat{c}_1, \cdots, \hat{c}_L, \hat{\rho}_1, \cdots, \hat{\rho}_L)$ be the estimators obtained in Theorem 4.2.2. Then*

$$M\left[\hat{c}_1 - \frac{N_1}{N}, \cdots, \hat{c}_L - \frac{N_L}{N}, \hat{\rho}_1 - \rho_1, \cdots, \hat{\rho}_L - \rho_L\right] \xrightarrow[N,M\to\infty]{\mathcal{D}} \mathcal{N}_{2L}(0, \mathbf{\Theta})$$

where $\mathbf{\Theta}$ is a $2L \times 2L$ matrix admitting the decomposition $\mathbf{\Theta} = \mathbf{M}^{-1}\mathbf{W}\mathbf{M}^{-1^T}$ and matrix $\mathbf{M}$ is the Jacobian matrix of f evaluated for $z = c$ and is defined in (4.2.5) and

$$\mathbf{W} = \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix},$$

where $\mathbf{V}$ is a $(2L-1) \times (2L-1)$ matrix whose entries are given by (for $1 \le k, \ell \le 2L-1$) :

$$V_{k,\ell} = -\frac{(-1)^{k+\ell}}{4\pi^2 c^2} \oint_{\mathcal{C}_1} \oint_{\mathcal{C}_2} \left(\frac{\underline{m}'(z_1)\underline{m}'(z_2)}{(\underline{m}(z_1) - \underline{m}(z_2))^2} - \frac{1}{(z_1 - z_2)^2} \right) \times \frac{1}{\underline{m}^k(z_1)\underline{m}^\ell(z_2)} d\,z_1 d\,z_2$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are two closed contours non-overlapping which contain the support $\mathcal{S}$ of F and are counterclockwise oriented.

Proof: The proof relies on the same techniques as developed in [99]. We outline hereafter the main steps and then provide the details.

By Theorem 4.2.2, the estimate vector $(\hat{c}_1, \cdots, \hat{c}_L, \hat{\rho}_1, \cdots, \hat{\rho}_L)$ verifies the following system of equations :

$$\begin{cases} \sum_{i=1}^{L} \hat{c}_i = 1, \\ \sum_{i=1}^{L} \hat{c}_i \hat{\rho}_i = \hat{\gamma}_1, \\ \sum_{i=1}^{L} \hat{c}_i \hat{\rho}_i^k = \hat{\gamma}_k \text{ for } 2 \leq k \leq 2L-1, \end{cases}$$

where the $\hat{\gamma}_i$'s are the moment estimates provided by Lemma 4.2.1.

Using the integral representation of $\sum_{i=1}^{L} c_i \rho_i$ and $\sum_{i=1}^{L} c_i \rho_i^k$ (cf. Section 4.3.4 in the Appendix and Formula (4.3.15)), we get :

$$\begin{cases} \sum_{i=1}^{L} M\left(\hat{c}_i - \frac{N_i}{N}\right) = 0, \\ \sum_{i=1}^{L} M\left(\hat{c}_i\hat{\rho}_i - \frac{N_i}{N}\rho_i\right) = -\frac{M^2}{2N\mathbf{i}\pi} \oint_{\mathcal{C}} z \left(\frac{m'_{\hat{\mathbf{R}}_N}(z)}{m_{\hat{\mathbf{R}}_N}(z)} - \frac{\underline{m}'_N(z)}{\underline{m}_N(z)}\right) dz, \\ \sum_{i=1}^{L} M\left(\hat{c}_i\hat{\rho}_i^k - \frac{N_i}{N}\rho_i^k\right) = \frac{M^2(-1)^k}{2\mathbf{i}(k-1)N\pi} \oint_{\mathcal{C}} \left(\frac{1}{m_{\hat{\mathbf{R}}_N}(z)^{k-1}} - \frac{1}{\underline{m}_N(z)^{k-1}}\right) dz, \ 2 \leq k \leq 2L-1. \end{cases}$$

Denote by $C(\mathcal{C}, \mathbb{C})$ the set of continuous functions from $\mathcal{C}$ to $\mathbb{C}$ endowed with the supremum norm $\|u\|_\infty = \sup_{\mathcal{C}} |u|$. In the same way as in [99], consider the process : $(X_N, X'_N, u_N, u'_N) : \mathcal{C} \to \mathbb{C}$, where

$$\begin{aligned} X_N(z) &= M\left(m_{\hat{\underline{\mathbf{R}}}_N}(z) - \underline{m}_N(z)\right), \\ X'_N(z) &= M\left(m'_{\hat{\underline{\mathbf{R}}}_N}(z) - \underline{m}'_N(z)\right), \\ u_N(z) &= m_{\hat{\underline{\mathbf{R}}}_N}(z), \ u'_N(z) = m'_{\hat{\underline{\mathbf{R}}}_N}(z). \end{aligned}$$

Then, $M \sum_{i=1}^{L} \left(\hat{c}_i\hat{\rho}_i - \frac{N_i}{N}\rho_i\right)$ can be written as :

$$\begin{aligned} M \sum_{i=1}^{L} \left(\hat{c}_i\hat{\rho}_i - \frac{N_i}{N}\rho_i\right) &= -\frac{M}{2\mathbf{i}N\pi} \oint_{\mathcal{C}} z \left(\frac{\underline{m}_N(z)X'_N(z) - u'_N(z)X_N(z)}{\underline{m}_N(z)u_N(z)}\right) dz, \\ &\triangleq \Upsilon_N(X_N, X'_N, u_N, u'_N), \end{aligned}$$

where

$$\Upsilon_N(x, x', u, u') = -\frac{M}{2\mathbf{i}N\pi} \oint_{\mathcal{C}} z \left(\frac{\underline{m}_N(z)x'(z) - u'(z)x(z)}{\underline{m}_N(z)u(z)}\right) dz.$$

On the other hand, using the decomposition $a^k - b^k = (a-b)\sum_{\ell=0}^{k-1} a^\ell b^{k-1-\ell}$, we can prove that :

$$\begin{aligned} \sum_{i=1}^{L} M\left(\hat{c}_i\hat{\rho}_i^k - \frac{N_i}{N}\rho_i^k\right) &= \frac{M^2(-1)^k}{2\mathbf{i}N\pi(k-1)} \oint_{\mathcal{C}} \sum_{\ell=0}^{k-2} - \frac{m_{\hat{\underline{\mathbf{R}}}_N}(z) - \underline{m}_N(z)}{m_{\hat{\underline{\mathbf{R}}}_N}^{\ell+1}(z)\underline{m}_N^{k-1-\ell}(z)} dz \\ &= \frac{M(-1)^{k+1}}{2\mathbf{i}N(k-1)\pi} \oint_{\mathcal{C}} \sum_{\ell=0}^{k-2} X_N(z)u_N(z)^{-\ell-1}\underline{m}_N(z)^{-k+1+\ell} dz \\ &\triangleq \Phi_{N,k}(X_N, u_N), \end{aligned}$$

for $2 \leq k \leq 2L-1$, where

$$\Phi_{N,k}(x,u) = \frac{M(-1)^{k+1}}{2\mathbf{i}N(k-1)\pi} \oint_{\mathcal{C}} \sum_{\ell=0}^{k-2} x(z)u(z)^{-\ell-1}\underline{m}_N(z)^{-k+1+\ell}dz.$$

The main idea of the proof of the theorem lies in the following steps :

1. Prove the convergence of the processes (X_N, X'_N, u_N, u'_N) and (X_N, u_N) over the contour $\mathcal{C}$ by using Bai and Silverstein's theorem [11].
2. Prove the convergence of $[\Upsilon_N(X_N, X'_N, u_N, u'_N), \Phi_{N,2}(X_N, u_N), \cdots, \Phi_{N,L}(X_N, u_N)]^T$ to a Gaussian random vector with the help of the continuous mapping theorem (cf. Theorem 4.2.5).
3. Compute the limiting covariance between $M\sum_{i=1}^L \left(\hat{c}_i\hat{\rho}_i^k - \frac{N_i}{N}\rho_i^k\right)$ and $M\sum_{i=1}^L \left(\hat{c}_i\hat{\rho}_i^\ell - \frac{N_i}{N}\rho_i^\ell\right)$.
4. Conclude by expressing $M\left[\hat{c}_1 - \frac{N_1}{N}, \cdots, \hat{c}_L - \frac{N_L}{N}, \hat{\rho}_1 - \rho_1, \cdots, \hat{\rho}_L - \rho_L\right]^T$ as a linear function of $M[\hat{\gamma}_0 - \gamma_0, \cdots, \hat{\gamma}_{2L-1} - \gamma_{2L-1}]^T$.

Fluctuations of the moments

The next step is to prove the convergence of the vector

$$\left[\Upsilon_N(X_N, X'_N, u_N, u'_N), \Phi_{N,2}(X_N, u_N), \cdots, \Phi_{N,L}(X_N, u_N)\right]^T.$$

The convergence of $\Upsilon_N(X_N, X'_N, u_N, u'_N)$ to a Gaussian random variable has been established in Theorem 4.1.3 where it has been proved that :

$$\Upsilon_N(X_N, X'_N, u_N, u'_N) \xrightarrow[M,N\to\infty]{\mathcal{D}} \Upsilon(X, Y, \underline{m}, \underline{m}')$$

where

$$\Upsilon(x,y,v,w) = \frac{1}{2\mathbf{i}\pi c}\oint_{\mathcal{C}} z\left(\frac{\underline{m}(z)y(z) - w(z)x(z)}{\underline{m}(z)v(z)}\right)dz.$$

The next task is to prove the convergence in distribution of $\Phi_{N,k}(X_N, u_N)$ over the contour $\mathcal{C}$, for $2 \leq k \leq L$. Let $\Phi_k(x,u)$ be defined as :

$$\Phi_k(x,u) = \frac{(-1)^k}{2\mathbf{i}c\pi}\oint_{\mathcal{C}} x(z)u(z)^{-k}dz.$$

We want to show that $\Phi_k(X_N, u_N)$ converges in distribution to a Gaussian vector. The continuous mapping theorem is useful to transform one convergence to another.

Theorem 4.2.5 (cf. [50, Th. 4.27]) *For any metric spaces S_1 and S_2, let ξ, $(\xi_n)_{n\geq1}$ be random elements in S_1 with $\xi_n \xrightarrow[n\to\infty]{\mathcal{D}} \xi$ and consider some measurable mappings f, $(f_n)_{n\geq1} : S_1 \to S_2$ and a measurable set $\Gamma \subset S_1$ with $\xi \in \Gamma$ a.s. such that $f_n(s_n) \to f(s)$ as $s_n \to s \in \Gamma$. Then $f_n(\xi_n) \xrightarrow[n\to\infty]{\mathcal{D}} f(\xi)$.*

Consider the set :

$$\Gamma = \left\{ (x,u) \in C^2\left(\mathcal{C},\mathbb{C}\right), \inf_{\mathcal{C}} |u| > 0 \right\}.$$

Then, since $\inf_{\mathcal{C}} |\underline{m}| > 0$ (see [5, Section 9.12]), the dominated convergence theorem implies that the convergence of $(x_N, y_N) \to (x,y) \in \Gamma$ leads to $\Phi_{N,k}(x_N, y_N) \to \Phi_k(x,y)$. The continuous mapping theorem applies, thus giving :

$$\Phi_{N,k}(X_N, u_N) \xrightarrow[M,N\to\infty]{\mathcal{D}} \Phi_k(X,u).$$

It now remains to prove that the limit law $\Phi_k(X,u)$ is Gaussian. For that, it suffices to notice that the integral can be written as the limit of a finite Riemann sum and that a finite Riemann sum of the elements of a Gaussian random vector is still Gaussian.

The convergence of $\Upsilon_N(X_N, X'_N, u_N, u'_N)$ and $\Phi_{N,k}(X_N, u_N)$ to Gaussian random variables is not sufficient to establish a CLT for the whole vector. It remains to prove that any linear combination of $\left[\Upsilon_N(X_N, X'_N, u_N, u'_N), \Phi_{N,2}(X_N, u_N), \cdots, \Phi_{N,L}(X_N, u_N)\right]^T$ converges toward a Gaussian distribution, which can easily be established in the same way as before. It implies that this vector converges to a Gaussian vector. This ends the proof of the fluctuations of the moments.

Computation of the variance

We now come to the third step. We shall therefore evaluate the quantities :

$$\begin{aligned}
\mathbf{V}_{1,1} &= \mathbb{E}\left[\Upsilon(X,Y,\underline{m},\underline{m}')\Upsilon(X,Y,\underline{m},\underline{m}')\right], \\
\mathbf{V}_{1,k} = \mathbf{V}_{k,1} &= \mathbb{E}\left[\Upsilon(X,Y,\underline{m},\underline{m}')\Phi_k(X,\underline{m})\right], \quad 2 \le k \le L, \\
\mathbf{V}_{k,\ell} &= \mathbb{E}\left[\Phi_k(X,\underline{m})\Phi_\ell(X,\underline{m})\right], \quad 2 \le k,\ell \le 2L-1.
\end{aligned}$$

The details of the calculations are in Appendix 4.3.5 and yield : For $1 \le k,\ell \le 2L-1$

$$\mathbf{V}_{k,\ell} = -\frac{(-1)^{k+\ell}}{4\pi^2 c^2} \oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} \left[\frac{\underline{m}'(z_1)\underline{m}'(z_2)}{(\underline{m}(z_1)-\underline{m}(z_2))^2} - \frac{1}{(z_1-z_2)^2}\right] \frac{1}{\underline{m}^k(z_1)\underline{m}^\ell(z_2)} dz_1 dz_2\,. \tag{4.2.11}$$

Let $\mathbf{w}_M = M\left[\hat{\gamma}_0 - \gamma_0, \cdots, \hat{\gamma}_{2L-1} - \gamma_{2L-1}\right]^T$. We have just proved that the vector $\mathbf{w}_M$ converges asymptotically to :

$$\mathbf{w}_M \xrightarrow[N,M\to+\infty]{\mathcal{D}} \mathcal{N}_{2L}(0,\mathbf{W}),$$

where

$$\mathbf{W} = \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix}$$

and $\mathbf{V}$ is the $(2L-1)\times(2L-1)$ matrix whose entries $V_{k,l}$ are given by (4.2.11).

Remark 4.2.6 *The zeros in the variance simply follow from the fact that $\hat{\gamma}_0 - \gamma_0 = 0$.*

Fluctuations of the eigenvalues estimates

To transfer this convergence to $\mathbf{q}_M \triangleq M\left[\hat{c}_1 - \frac{N_1}{N}, \cdots, \hat{c}_L - \frac{N_L}{N}, \hat{\rho}_1 - \rho_1, \cdots, \hat{\rho}_L - \rho_L\right]^T$, we shall use Slutsky's lemma which is as below :

Lemma 4.2.2 (cf. [28]) *Let $\mathbf{X}_n$, $\mathbf{Y}_n$ be sequences of vector or matrix random elements. If $\mathbf{X}_n$ converges in distribution to a random element $\mathbf{X}$, and $\mathbf{Y}_n$ converges in probability to a constant $\mathbf{C}$, then*

$$\mathbf{Y}_n^{-1}\mathbf{X}_n \xrightarrow{\mathcal{D}} \mathbf{C}^{-1}\mathbf{X}$$

provided that $\mathbf{C}$ is invertible.

We will show that $\mathbf{w}_M$ satisfies the following linear system :

$$\mathbf{w}_M = \hat{\mathbf{M}}_M \mathbf{q}_M \tag{4.2.12}$$

where we will try to find a matrix $\hat{\mathbf{M}}_M$ who converges in probability to $\mathbf{M}$ which is given by (4.2.5).

To this end, let us work out the expression of $w_{k,M}$, the k-th element of $\mathbf{w}_M$.

If $k = 1$, it is easy to see that $w_{1,M} = 0$.

For $k \geq 2$, $w_{k,M}$ is given by :

$$\begin{aligned}
w_{k,M} &= M \sum_{i=1}^{L} \left(\hat{c}_i \hat{\rho}_i^{k-1} - \frac{N_i}{N} \rho_i^{k-1} \right) \\
&= M \sum_{i=1}^{L} \left(\hat{c}_i \hat{\rho}_i^{k-1} - \frac{N_i}{N} \hat{\rho}_i^{k-1} + \frac{N_i}{N} \hat{\rho}_i^{k-1} - \frac{N_i}{N} \rho_i^{k-1} \right) \\
&= M \sum_{i=1}^{L} \left(\left(\hat{c}_i - \frac{N_i}{N} \right) \hat{\rho}_i^{k-1} + \frac{N_i}{N} (\hat{\rho}_i - \rho_i) \sum_{\ell=0}^{k-2} \hat{\rho}_i^{\ell} \rho_i^{k-2-\ell} \right).
\end{aligned}$$

Then define

$$\hat{\mathbf{M}}_M = \begin{pmatrix}
1 & \cdots & 1 & 0 & \cdots & 0 \\
\hat{\rho}_1 & \cdots & \hat{\rho}_L & \frac{N_1}{N} & \cdots & \frac{N_L}{N} \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\hat{\rho}_1^{2L-1} & \cdots & \hat{\rho}_L^{2L-1} & \frac{N_1}{N} \sum_{\ell=0}^{2L-2} \hat{\rho}_1^{\ell} \rho_1^{2L-2-\ell} & \cdots & \frac{N_L}{N} \sum_{\ell=0}^{2L-2} \hat{\rho}_L^{\ell} \rho_L^{2L-2-\ell}
\end{pmatrix}.$$

We can easily check that Eq. (4.2.12) is satisfied and $\hat{\mathbf{M}}_M$ converges in probability to $\mathbf{M}$. It remains to check that $\mathbf{M}$ is invertible. Note that the non-singularity of matrix $\mathbf{M}$ has already been established in Section 4.2.2, where this property was required to prove the existence of an estimator. As a consequence, using Slutsky's lemma, we deduce that :

$$\hat{\mathbf{M}}_M \mathbf{q}_M \xrightarrow[M,N\to+\infty]{\mathcal{D}} \mathcal{N}_{2L}(0, \mathbf{W})$$

and

$$\mathbf{q}_M \xrightarrow[M,N\to+\infty]{\mathcal{D}} \mathcal{N}_{2L}\left(0, \mathbf{M}^{-1}\mathbf{W}(\mathbf{M}^{-1})^T\right).$$

This ends the proof for the fluctuation.

$\square$

4.2.4 Simulations

In this section, we compare the performance of the proposed estimator with Mestre's estimator [64] in Section 4.2.4 ; we then compare the proposed estimator with the estimator proposed by Bai et al. [4] in Section 4.2.4. We finally verify by simulations the accuracy of the Gaussian approximation stated by the CLT in Section 4.2.4.

Comparison with the estimator in [64]

As will be seen below, the separability assumption is compulsory for Mestre's method to be effective. If this assumption holds true, a simple clustering procedure enables to estimate the unknown multiplicities and Mestre's method outperforms our moment estimator (see Fig. 4.4).

If, however, the separability assumption is not met, then it is not clear how to directly estimate (even roughly) the multiplicities ; and even if those were known, Mestre's estimation method has no methodological foundations (as the estimator is not even consistent in this case !) and the computation of Mestre's estimator yields a systematic error (see Fig. 4.5 for instance).

A final remark is in order with respect to the separability assumption : Although it is easy in simulations to generate data fulfilling or violating the separability assumption, it is not an easy task, while facing real data, to decide whether the separability assumption holds true or not. Building such a test remains an open problem, advocating for our procedure by default - unless any extra argument emerges to support a separability assumption. Otherwise stated, the non-separability assumption is much more realistic in practical cases.

In the first experiment, we consider the case where the separability condition holds true. We assume also that the covariance matrix has three different eigenvalues $(\rho_1, \rho_2, \rho_3) = (1, 3, 7)$, which are distributed as : $\frac{N_1}{N} = 0.5, \frac{N_2}{N} = \frac{N_3}{N} = \frac{1}{4}$. The ratio $\frac{N}{M}$ is set to $\frac{30}{200} = \frac{3}{20}$. The separability condition being met, the clusters are well separated so that the multiplicities can be estimated in a heuristic way based on the difference of the ordered eigenvalues. More precisely, an empirical method for estimating the multiplicities consists in the following steps :

- Arrange the eigenvalues of the covariance matrix in increasing order : $\hat{\lambda}_1 \leq \cdots, \leq \hat{\lambda}_N$.
- Take L indexes $i_1, \cdots, i_L$ satisfying :

$$i_1 = \arg\max_i \left(\hat{\lambda}_{i+1} - \hat{\lambda}_i\right) ,$$

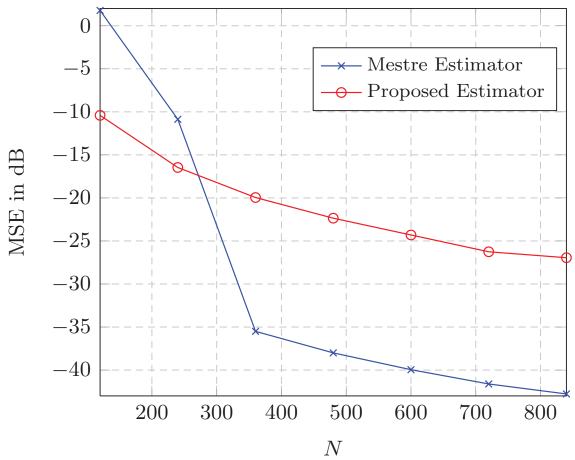

FIGURE 4.4 – Experienced MSE with N when $\frac{N}{M} = \frac{3}{20}$ and $(\rho_1, \rho_2, \rho_3) = (1, 3, 7)$

$$\begin{aligned} i_2 &= \arg\max_{i \neq i_1} \left(\hat{\lambda}_{i+1} - \hat{\lambda}_i\right) \quad , \\ &\vdots \\ i_L &= \arg\max_{i \notin \{i_1, \cdots, i_{L-1}\}} \left(\hat{\lambda}_{i+1} - \hat{\lambda}_i\right) \quad . \end{aligned}$$

– Arrange these indexes in the increasing order : $i_{[1]} \leq \cdots \leq i_{[L]}$. Empirical estimates of the multiplicities are thus given by :

$$\begin{aligned} \hat{N}_1 &= i_{[1]} \\ \hat{N}_2 &= i_{[2]} - i_{[1]} \\ &\vdots \\ \hat{N}_L &= N - i_{[L-1]}. \end{aligned}$$

This empirical method has proved to be efficient in the asymptotic regime. For example, by applying exact separation results from [9, 10], it can be proved that the estimates of the normalized multiplicities $(\hat{N}_k/N)$ are asymptotically consistent.

Fig. 4.4 compares the performance of the Mestre's estimator using the aforementioned method for estimating the multiplicities with that of the proposed estimator, in terms of MSE :

$$\text{MSE} \triangleq \frac{1}{1000} \sum_{k=1}^{1000} \sum_{i=1}^{3} |\hat{\rho}_i^k - \rho_i|^2,$$

In this case, Mestre's estimator outperforms the proposed estimator. This can be attributed to numerical difficulties which will be discussed in the next section.

In the second experiment, we consider the case where the separability condition does not hold. In particular, we assume that the covariance matrix $\mathbf{R}_N$ has three different eigenvalues $(\rho_1, \rho_2, \rho_3) = (1, 2, 3)$, each with the same multiplicity, *i.e.* $\frac{N_1}{N} = \frac{N_2}{N} = \frac{N_3}{N} = \frac{1}{3}$. We also set the

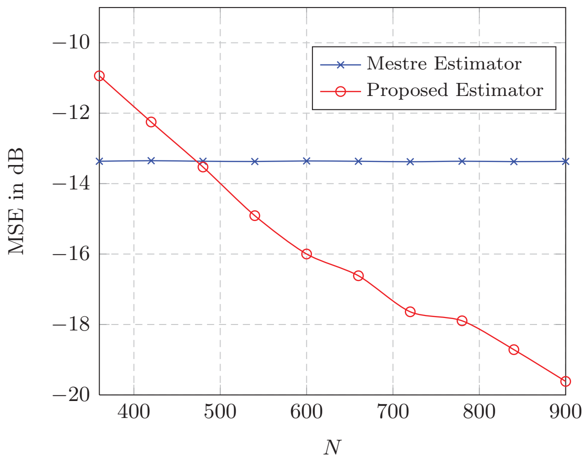

FIGURE 4.5 – Experienced MSE with N when $\frac{N}{M} = \frac{3}{8}$ and $(\rho_1, \rho_2, \rho_3) = (1, 2, 3)$

ratio between the dimension of variables and the number of samples $\frac{N}{M}$ to $3/8$, a ratio which is too high for the separability condition to hold true. We assume for our estimator that the multiplicities are not known, a hypothesis that obviously cannot be used for Mestre's estimator. We thus favour Mestre's estimator by assuming that it knows perfectly the multiplicities. Fig. 4.5 compares the obtained results in terms of MSE : for different values of M and N satisfying a constant ratio $c = N/M = 3/8$ and 1000 realizations. We note that as M and N increase, the estimator in [64] exhibits an error floor, underlying the fact that without the separability assumption, Mestre's estimators are no longer consistent.

Comparison with the method in [4]

The estimator proposed in [4] and our proposed estimator are similar at first sight. The main difference lies in the intermediate quantities which are estimated before estimating the eigenvalues and their multiplicities. While the technique of [4] is based on the numerical computation of the empirical moments $\frac{1}{N}\mathrm{Tr}(\mathbf{Y}_N\mathbf{Y}_N^H)^k$, our technique rather relies on building consistent estimators of the theoretical moments $\frac{1}{N}\mathrm{Tr}\mathbf{R}_N^k$. This difference induces important numerical consequences in the computation of the estimates : In [4], the functional relation between the quantities to be estimated and the empirical moments $\frac{1}{N}\mathrm{Tr}(\mathbf{Y}_N\mathbf{Y}_N^H)^k$ yields a system of equations whose resolution relies on iterative methods (based for instance on the functions `fsolve` or `fminsearch` in MATLAB) which are extremely slow.

On the other hand, the method proposed in this article is based on a bijective system of equations that links the theoretical moments to the eigenvalues and their multiplicities, whose resolution relies on simple computations : A matrix inversion and solving a polynomial (see for instance end of Section 4.2.2).

Simulation results indicate that our algorithm allows a great gain of complexity compared to [4], while keeping the same level of performance. Execution times for one realization are provided in the following table 4.1 for the same simulation setting as the second experiment. Note that unlike our method which exhibits low complexity, the complexity of the method of [4] tends to

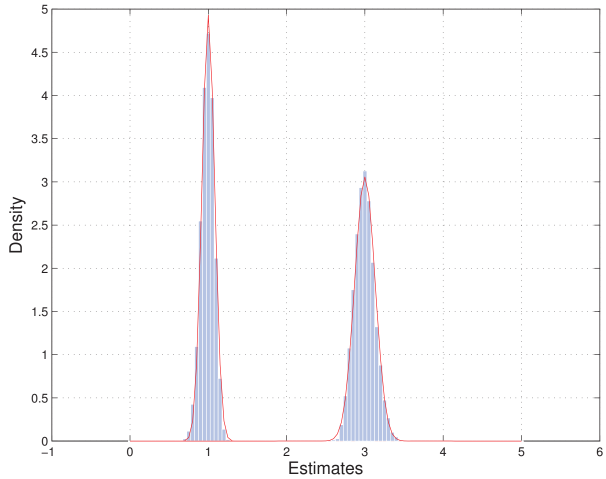

FIGURE 4.6 – Comparison of empirical against theoretical variances for $c_1 = c_2 = 0.5$ and $\rho_1 = 1$ and $\rho_2 = 3$

increase exponentially as the dimensions N and M increase.

N, M	Proposed method	Bai, Chen and Yao's method
$N = 300,\ M = 800$	0.5s	10.68s
$N = 360,\ M = 960$	0.55s	25.57s
$N = 420,\ M = 1120$	0.67s	42.62s

TABLE 4.1 – Execution time to obtain an estimator for one realization (in seconds)

Accuracy of the Gaussian approximation

Finally, we verify by simulations the accuracy of the Gaussian approximation. We consider the case where there are two different eigenvalues $\rho_1 = 1$ and $\rho_2 = 3$ that are uniformly distributed. Unlike the first experiment, we assume that the multiplicities are unknown. We represent in Fig 4.6 the histogram for $\hat{\rho}_1$ and $\hat{\rho}_2$ when $N = 60$ and $M = 120$. We also represent in red line, the corresponding Gaussian distribution. We note that as it was predicted by our derived results, the histogram is similar to that of a Gaussian random variable.

4.3 Appendices

4.3.1 Proof of Proposition 4.1.1

Let us first begin by considerations related to the supports of the probability distributions associated to $m(z)$ and $m_N(z)$. Denote by $\mathcal{S}$ and $\mathcal{S}_N$ these supports and recall that $\mathcal{S}$ is the union of L disjoint clusters : For $a_1 \le b_1 < \cdots < a_L \le b_L$,

$$\mathcal{S} = [a_1, b_1] \cup \cdots \cup [a_L, b_L] \ .$$

The following lemma clarifies the relations between $\mathcal{S}_N$ and $\mathcal{S}$.

Lemma 4.3.1 *Under Assumptions* **A1** *and* **A2**, *for N large enough, the support $\mathcal{S}_N$ of the probability distribution associated to the Stieltjes transform $m_N(z)$ is the union of L clusters : For $a_1^N \le b_1^N < \cdots < a_L^N \le b_L^N$,*

$$\mathcal{S}_N = [a_1^N, b_1^N] \cup \cdots \cup [a_L^N, b_L^N] \ .$$

Moreover, the following convergence holds true :

$$a_\ell^N \xrightarrow[N,M\to\infty]{} a_\ell \ , \quad b_\ell^N \xrightarrow[N,M\to\infty]{} b_\ell \ ,$$

for $1 \le \ell \le L$.

Remark 4.3.1 *If the support $\mathcal{S}_N$ contains zero, (ex : $N > M$), $a_1^N = b_1^N = 0$. By Assumption* **A1**, *the multiplicity N_1 corresponding to zero satisfies $\frac{N_1}{N} \to c_1 > 0$, hence zero is also in the support $\mathcal{S}$. In this case, we will get that $a_1 = b_1 = 0$, and the conclusion still holds true.*

Proof :[Proof of Lemma 4.3.1] Recall the relations :

$$\underline{m}_N(z) = -\left(z - \frac{N}{M}\int \frac{t}{1+t\underline{m}_N} F^{\mathbf{R}_N}(dt)\right)^{-1} \tag{4.3.1}$$

and

$$m_N(z) = \frac{M}{N}\underline{m}_N(z) - \left(1 - \frac{M}{N}\right)\frac{1}{z}. \tag{4.3.2}$$

As the Stieltjes transform of δ_0 (the Dirac mass at 0) is $-\frac{1}{z}$ and $\underline{m}_N(z)$ is a continuous function over $\mathbb{R}_+^*$, for a,b with $0 < a < b$, by the inverse formula of Stieltjes transform, one gets :

$$F_N([a,b]) = \frac{M}{N}\underline{F}_N([a,b]).$$

So it suffices to study the support $\underline{\mathcal{S}}_N$ associated to $\underline{F}_N$.

From the definition of $\underline{m}_N(z)$ (see formula (4.3.1)), we obtain :

$$z_{\mathbf{R}_N}(\underline{m}_N) = -\frac{1}{\underline{m}_N} + \frac{N}{M}\int \frac{t dF^{\mathbf{R_N}}(t)}{1+t\underline{m}_N}.$$

Denote by $B = \{m \in \mathbb{R} : m \neq 0, -m^{-1} \notin \{\rho_1, \cdots, \rho_L\}\}$. From Lemma [84], for a real number x, $x \in \underline{\mathcal{S}}_N^c \iff \underline{m}_x \in B$ and $z'_{\mathbf{R_N}}(\underline{m}_x) = \frac{1}{\underline{m}_x^2} - \frac{N}{M}\int \frac{t^2 dF^{\mathbf{R_N}}(t)}{(1+t\underline{m}_x)^2} > 0$ with $\underline{m}_N(x) = \underline{m}_x$ and $z_{\mathbf{R_N}}(\underline{m}_x) = x$.

Then if $a \in \partial\underline{\mathcal{S}}_N$, $m_a \notin B$ or $z'_{\mathbf{R_N}}(m_a) \le 0$ with $m_a = \underline{m}_N(a)$. Now we will show that $m_a \in B$. With the same argument in [84], $m_a \neq 0$. If $-m_a^{-1} \in S_{F^{\mathbf{R_N}}}$, as $F^{\mathbf{R_N}}$ is discrete, we get that $\lim_{m\to m_a}\int \frac{t^2 dF^{\mathbf{R_N}}(t)}{(1+tm)^2} \longrightarrow \infty$. So on the neighborhood to the left and to the right of m_a, $z'_{\mathbf{R_N}} < 0$ which contradicts [84, Lemma 3.5].

Hence $z'_{\mathbf{R_N}}(m_a) \leq 0$. By the continuity, we get

$$z'_{\mathbf{R_N}}(m_a) = \frac{1}{m_a^2} - \frac{N}{M}\int \frac{t^2 dF^{\mathbf{R}_N}(t)}{(1+tm_a)^2} = 0.$$

This is equivalent to the following equation :

$$z'_{\mathbf{R_N}}(m_a) = \frac{1}{m_a^2} - \frac{1}{M}\sum_{i=1}^{L} N_i \frac{\rho_i^2}{(1+\rho_i m_a)^2} = 0. \tag{4.3.3}$$

By multiplying the common denominator, one gets a polynomial of the degree 2L in m_a. Let us now prove that these 2L roots are real. At first, note that :

$$\frac{1}{m^2} - \frac{N}{M}\int \frac{t^2 dF^{\mathbf{R}_N}(t)}{(1+tm)^2} \xrightarrow[m \to -\frac{1}{\rho_i}]{} -\infty,$$

and

$$z''_{\mathbf{R_N}}(m) = -\frac{2}{m^3} + \frac{N}{M}\int \frac{2t^3 dF^{\mathbf{R}_N}(t)}{(1+tm)^3}.$$

So $z''_{\mathbf{R_N}}(m)$ has one and only one zero in the open set $(-\frac{1}{\rho_i}, -\frac{1}{\rho_{i+1}})$ for $1 \leq i \leq L-1$. Then for $\beta_i \in (-\frac{1}{\rho_i}, -\frac{1}{\rho_{i+1}})$ such that $z''_{\mathbf{R_N}}(\beta_i) = 0$, it suffices to show that $z'_{\mathbf{R_N}}(\beta_i) > 0$ in order to prove that there will be two zeros for $z'_{\mathbf{R_N}}(m)$ in the set $(-\frac{1}{\rho_i}, -\frac{1}{\rho_{i+1}})$. From the separability condition (cf. Assumption **A2**), $\inf_N\{\frac{M}{N} - \Psi_N(i)\} > 0$, and

$$\begin{aligned} z'_{\mathbf{R_N}}\left(-\frac{1}{\alpha_i}\right) &= \alpha_i^2 - \frac{N}{M}\int \frac{t^2 dF^{\mathbf{R_N}}(t)}{(1-\frac{t}{\alpha_i})^2} , \\ &= \alpha_i^2 \left(1 - \frac{1}{M}\sum_{r=1}^{L} N_i \frac{\rho_i^2}{(\alpha_i - \rho_i)^2}\right) > 0 . \end{aligned}$$

Thus we obtain $2(L-1)$ roots. Besides, in the open set $(-\rho_L^{-1}, 0)$,

$$\frac{1}{m^2} - \frac{N}{M}\int \frac{t^2 dF^{\mathbf{R}_N}(t)}{(1+tm)^2} \xrightarrow[m_a \to 0^-]{} +\infty,$$

there exists another root in this set. In the open set $(-\infty, -\rho_1^{-1})$,

$$\frac{1}{m^2} - \frac{N}{M}\int \frac{t^2 dF^{\mathbf{R}_N}(t)}{(1+tm)^2} \xrightarrow[m \to -\infty]{} 0$$

and

$$\frac{1}{m^2} - \frac{N}{M}\int \frac{t^2 dF^{\mathbf{R}_N}(t)}{(1+tm)^2} \underset{m \to -\infty}{\sim} \frac{1}{m^2}(1 - \frac{L}{M}) > 0.$$

Hence the last root in this open set. This proves that $\mathcal{S}_N = [a_1^N, b_1^N] \cup \cdots \cup [a_L^N, b_L^N]$.

To prove $a_\ell^N \xrightarrow[N,M \to \infty]{} a_\ell$ and $b_\ell^N \xrightarrow[N,M \to \infty]{} b_\ell$, note that a_i b_i satisfy the same type of equation by replacing $\frac{N}{M}$ by c and $F^{\mathbf{R_N}}$ by F^R. As $\frac{N}{M} \to c$ and $\frac{K_i}{M} \to c_i$, the roots of Eq. (4.3.3) converge to those of the limiting equation (see [27] for instance). Hence the conclusion.

□

We are now in position to establish the proof of Proposition 4.1.1.

Denote by $\mathcal{S}(\varepsilon)$ the ε-blow-up of $\mathcal{S}$, i.e. $\mathcal{S}(\varepsilon) = \{x \in \mathbb{R},\ d(x, S) < \varepsilon\}$. Let $\varepsilon > 0$ be small enough and consider a smooth function ϕ equal to zero on $\mathcal{S}(\varepsilon/3)$, equal to 1 if $x \notin \mathcal{S}(\varepsilon)$, equal to zero again if $|x| \geq \tau$ (as we shall see, τ will be chosen to be large), and smooth in-between with $0 \leq \phi \leq 1$:

$$\phi(x) = \begin{cases} 0 & \text{if } d(x, \mathcal{S}) < \varepsilon/3 \ , \\ 1 & \text{if } d(x, \mathcal{S}) > \varepsilon \ , |x| \leq \tau - \epsilon \\ 0 & \text{if } |x| > \tau \ . \end{cases}$$

Notice that if $N, M \to \infty$ and N is large enough, then by Lemma 4.3.1, $\phi(x) = 0$ for all $x \in \mathcal{S}_N$. Now if $\mathbf{Z}$ is a $M \times M$ hermitian matrix with spectral decomposition $\mathbf{Z} = \mathbf{U}\,\mathrm{diag}\,(\gamma_i;\ 1 \leq i \leq M))\ \mathbf{U}^H$, where $\mathbf{U}$ is unitary and $(\gamma_i;\ 1 \leq i \leq M) = \mathrm{eig}(\mathbf{Z})$, write $\phi(\mathbf{Z}) = \mathbf{U}\,\mathrm{diag}\,(\phi(\gamma_i);\ 1 \leq i \leq M))\ \mathbf{U}^H$.

We have :

$$\begin{aligned} &\mathbb{P}(\sup_n d(\lambda_n, S) > \varepsilon) \\ &\quad \leq \quad \mathbb{P}(\|\hat{\mathbf{R}}_N\| > \tau - \varepsilon) \ + \ \mathbb{P}(\mathrm{Tr}\,\phi(\hat{\mathbf{R}}_N) \geq 1) \\ &\quad = \quad \mathbb{P}(\|\hat{\mathbf{R}}_N\| > \tau - \varepsilon) \ + \ \mathbb{P}([\mathrm{Tr}\,\phi(\hat{\mathbf{R}}_N)]^p \geq 1) \\ &\quad \overset{(a)}{\leq} \quad \mathbb{P}(\|\hat{\mathbf{R}}_N\| > \tau - \varepsilon) + \mathbb{E}[\mathrm{Tr}\,\phi(\hat{\mathbf{R}}_N)]^p, \end{aligned}$$

for every $p \geq 1$, where (a) follows from Markov's inequality. The fact that $\mathbb{P}(\|\hat{\mathbf{R}}_N\| > \tau) = \mathcal{O}(N^{-\ell})$ for τ large enough and every $\ell \in \mathbb{N}^*$ is well-known (see for instance [5, Section 9.7]). We shall therefore establish estimates over $\mathbb{E}[\mathrm{Tr}\,\phi(\hat{\mathbf{R}}_N)]^p$. Take $p = 2^k$; we prove the following statement by induction : For $k \geq 1$ and for every integer $\beta < 2^k$ and for every smooth function f with compact support whose value on $\mathcal{S}(\varepsilon/3)$ is zero ,

$$\mathbb{E}\left(\mathrm{Tr} f(\hat{\mathbf{R}}_N)\right)^{2^k} \quad = \quad \mathcal{O}\left(\frac{1}{N^\beta}\right) \ .$$

First note that, due to Lemma 4.3.1, $\int_{\mathcal{S}_N} f(\lambda)\, F_N(d\lambda) = 0$ (where F_N is the probability distribution associated to m_N) for N, M large enough ($N, M \to \infty$). A minor modification of [93, Lemma 2] (whose model is slightly different) with the help of [41, Proposition 5] yields that for $N, M \to \infty$ and N large enough, $\mathbb{E}\,\mathrm{Tr}\, f(\hat{\mathbf{R}}_N) = \mathcal{O}(N^{-1})$, and the property is verified for $k = 0$.

Let $k > 0$ be fixed and assume that the result holds true for $\beta < 2^k$. We want to show that $\mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^{(k+1)}} = \mathcal{O}(N^{-2\beta})$. At step $k+1$, the expectation writes :

$$\left|\mathbb{E}[\mathrm{Tr}\, f(\hat{\mathbf{R}}_N)]^{2^{(k+1)}}\right|$$

$$\begin{aligned} &= \left|\mathbb{E}\left([\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^k} + \mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^k} - \mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^k}\right)^2\right| \\ &\leq 2\left(\mathrm{Var}[\mathrm{Tr}\, f(\hat{\mathbf{R}}_N)]^{2^k} + |\mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^k}|^2\right). \end{aligned} \tag{4.3.4}$$

The second term of the right hand side (r.h.s.) of the equation can be handled by the induction hypothesis :

$$\left|\mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^k}\right|^2 = \mathcal{O}\left(\frac{1}{N^{2\beta}}\right).$$

We now rely on Poincaré-Nash inequality (see for instance [41, Section II-B]) to handle the first term of the r.h.s. Applying this inequality, we obtain :

$$\mathrm{Var}\left((\mathrm{Tr} f(\hat{\mathbf{R}}_N))^{2^k}\right) \leq K \sum_{i,j} \mathbb{E}\left[\left|\frac{\partial[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^k}}{\partial Y_{i,j}}\right|^2 + \left|\frac{\partial[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^k}}{\partial \overline{Y}_{i,j}}\right|^2\right], \tag{4.3.5}$$

where K is a constant which does not depend on N, M and which is greater than $\mathbf{R}_N$'s eigenvalues. In order to compute the derivatives of the r.h.s., we rely on [38, Lemma 4.6]. This yields :

$$\begin{aligned} \frac{\partial}{\partial Y_{i,j}}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^k} &= \frac{2^k}{M}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^k-1}[\mathbf{Y}_N^* f'(\hat{\mathbf{R}}_N)]_{j,i}\ , \\ \frac{\partial}{\partial \overline{Y}_{i,j}}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^k} &= \frac{2^k}{M}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^k-1}[f'(\hat{\mathbf{R}}_N)\mathbf{Y}_N]_{i,j}\ . \end{aligned}$$

Plugging these derivatives into (4.3.5), we obtain :

$$\begin{aligned} &\mathrm{Var}(\mathrm{Tr}[f(\hat{\mathbf{R}}_N)]^{2^k}) \\ &\leq \frac{K\,2^{2k+1}}{M^2}\mathbb{E}\left[(\mathrm{Tr} f(\hat{\mathbf{R}}_N))^{(2^{k+1}-2)}\,\mathrm{Tr}(f'(\hat{\mathbf{R}}_N)\mathbf{Y}_N\mathbf{Y}_N^* f'(\hat{\mathbf{R}}_N))\right], \\ &= \frac{K\,2^{2k+1}}{M}\mathbb{E}\left[(\mathrm{Tr} f(\hat{\mathbf{R}}_N))^{(2^{k+1}-2)}\,\mathrm{Tr}(f'(\hat{\mathbf{R}}_N)^2\hat{\mathbf{R}}_N)\right], \\ &\leq \frac{K\,2^{2k+1}}{M}\left|\mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^{k+1}}\right|^{\frac{2^{k+1}-2}{2^{k+1}}} \times \left|\mathbb{E}[\mathrm{Tr} f'(\hat{\mathbf{R}}_N)^2\hat{\mathbf{R}}_N]^{2^k}\right|^{\frac{1}{2^k}}, \end{aligned}$$

where the last inequality is a consequence of Hölder's inequality.

As the function $h(\lambda) = \lambda[f'(\lambda)]^2$ satisfies the induction hypothesis, we have for every $\alpha < 1$:

$$\left|\mathbb{E}\mathrm{Tr}[f'(\hat{\mathbf{R}}_N)^2\hat{\mathbf{R}}_N]^{2^k}\right|^{\frac{1}{2^k}} = \mathcal{O}(N^{-\alpha}).$$

Plugging this estimate into (4.3.4), we obtain :

$$\begin{aligned} &\left|\mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^{(k+1)}}\right| \\ &\leq K\left(\frac{1}{N^{1+\alpha}}|\mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^{(k+1)}}|^{\frac{2^{k+1}-2}{2^{k+1}}}\right) + \mathcal{O}(N^{-2\beta}), \end{aligned} \tag{4.3.6}$$

where K is a constant independent of M, N, k. Notice that inequality (4.3.6) involves twice the quantity of interest $\mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^{(k+1)}}$ that we want to upper bound by $O(N^{-2\beta})$. We shall proceed iteratively.

Notice that $\mathrm{Tr}[f(\hat{\mathbf{R}}_N)] \leq \sup_{x\in\mathbb{R}} |f(x)| \times N$ because f is bounded on $\mathbb{R}$; hence the rough estimate :

$$\mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^{(k+1)}} = \mathcal{O}(N^{2^{k+1}}).$$

Plugging this into (4.3.6) yields :

$$\mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^{(k+1)}} = \mathcal{O}(N^{a_1}) \ ,$$

where $a_0 = 2^{k+1}$ and $a_1 = a_0 \frac{2^{k+1}-2}{2^{k+1}} - (1+\alpha)$. Iterating the procedure, we obtain :

$$\mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^{(k+1)}} = \mathcal{O}\left(N^{a_\ell \vee (-2\beta)}\right) \ ,$$

where $a_\ell = a_{\ell-1} \frac{2^{k+1}-2}{2^{k+1}} - (1+\alpha)$ and $x \vee y$ stands for $\sup(x,y)$. Now, in order to conclude the proof, it remains to prove that i) the sequence (a_ℓ) converges to some limit a_∞, ii) for some well-chosen $\alpha < 1$, $a_\infty \in (-2^{k+1}, -2\beta)$. Write :

$$a_{\ell+1} + 2^k(1+\alpha) = \frac{2^k - 1}{2^k}(a_\ell + 2^k(1+\alpha)) \ ,$$

hence a_ℓ converges to $-2^k(1+\alpha)$ which readily belongs to $(-2^{k+1}, -2\beta)$ for a well-chosen $\alpha \in (0,1)$. Finally $\mathbb{E}[\mathrm{Tr} f(\hat{\mathbf{R}}_N)]^{2^{(k+1)}} = \mathcal{O}(N^{-2\beta})$ which ends the induction.

It remains to apply this estimate to $\mathbb{E}[\mathrm{Tr}\phi(\hat{\mathbf{R}}_N)]^\ell$ in order to get the desired result.

4.3.2 Proof of Lemma 4.1.1

Notice that $X_N(z) = M(m_{\hat{\mathbf{R}}_N} - \underline{m}_N) = \overline{X_N(\overline{z})}$ for $z \in \overline{\mathbb{C}^+}$. So it suffices to verify the arguments for $z \in \mathbb{C}^+$. As $\frac{1}{\hat{\rho}_k - z}$ can converge to infinity if z is close to the real axis, the process $X_N(z)$ might be large when z is close to the real axis. Thus we begin the proof by considering a truncated version of the process X_N. More precisely, let ε_N be a real sequence decreasing to zero satisfying for some $\delta \in]0,1[$:

$$\varepsilon_N \geq N^{-\delta} \ .$$

With the same notations as in Lemma 4.3.1, denote by $\mathcal{S} = [a_1, b_1] \cup \cdots \cup [a_L, b_L]$; and take p_k, q_k such that $b_{k-1} < p_k < a_k$ and $b_k < q_k < a_{k+1}$ for $1 \leq k \leq L$ with conventions $b_0 = 0$ and $a_{L+1} = \infty$, $i.e.$ $[p_k, q_k]$ only contains the k-th cluster. Let $d > 0$. Consider :

$$\begin{aligned}
\mathcal{R}_{k,1} &= \{x + \mathbf{i}d : x \in [p_k, q_k]\} \ , \\
\mathcal{R}_{k,2} &= \left\{p_k + \mathbf{i}v : v \in \left[\frac{\varepsilon_N}{N}, d\right]\right\} \ , \\
\mathcal{R}_{k,3} &= \left\{q_k + \mathbf{i}v : v \in \left[\frac{\varepsilon_N}{N}, d\right]\right\} \ ,
\end{aligned}$$

and let $\tilde{\mathcal{R}}_k = \mathcal{R}_{k,1} \cup \mathcal{R}_{k,2} \cup \mathcal{R}_{k,3}$. The process $\hat{X}_N(\cdot)$ is defined by

$$\hat{X}_N(z) = \begin{cases} X_N(z) & \text{for } z \in \tilde{R}_k, \\ X_N(p_k + \mathbf{i}\frac{\varepsilon_N}{N}) & \text{for } x = p_k, v \in [0, \frac{\varepsilon_N}{N}], \\ X_N(q_k + \mathbf{i}\frac{\varepsilon_N}{N}) & \text{for } x = q_k, v \in [0, \frac{\varepsilon_N}{N}]. \end{cases}$$

This partition of $\tilde{\mathcal{R}}_k$ is identical to that used in [11, Section 1]. With probability one (see [9] and [10]), for all $\epsilon > 0$,

$$\lim_N \sup_{\lambda \in \mathrm{eig}(\hat{\mathbf{R}}_N)} d(\lambda, \mathcal{S}_N) < \epsilon$$

with $d(x, S)$ the Euclidean distance of x to the set S. Notice that :

$$\begin{aligned} \left| \oint_{\mathcal{R}_k} \Big(X_N(z) - \hat{X}_N(z) \Big) dz \right| \leq 2 \int_0^{\frac{\varepsilon_N}{N}} & \left| X_N(p_k + \mathbf{i}x) - \hat{X}_N(p_k + \mathbf{i}\frac{\varepsilon_N}{N}) \right| \\ & + \left| X_N(q_k + \mathbf{i}x) - \hat{X}_N(q_k + \mathbf{i}\frac{\varepsilon_N}{N}) \right| dx. \end{aligned}$$

Furthermore, with probability one, for all N large,

$$\int_0^{\frac{\varepsilon_N}{N}} \left| m_{\underline{\hat{\mathbf{R}}}_N}(p_k + \mathbf{i}x) - m_{\underline{\hat{\mathbf{R}}}_N}(p_k + \mathbf{i}\frac{\varepsilon_N}{N}) \right| dx \leq \frac{\varepsilon_N}{N} (|p_k - a_k|^{-1} + |p_k - b_{k-1}|^{-1}),$$

and

$$\int_0^{\frac{\varepsilon_N}{N}} \left| \underline{m}_N(p_k + \mathbf{i}x) - \underline{m}_N(p_k + \mathbf{i}\frac{\varepsilon_N}{N}) \right| dx \leq 2K \frac{\varepsilon_N}{N}$$

where $K = \sup_{z \in \mathcal{R}_k} |\underline{m}_N(z)|$. Thus, with probability one,

$$\left| \oint \Big(X_N(z) - \hat{X}_N(z) \Big) dz \right| \leq K_1 \varepsilon_N,$$

where K_1 is a constant which does not depend on N, M . A similar result can be achieved for the derivative functions $X'_N(z)$ and $\hat{X}'_N(z)$. One can get :

$$\begin{aligned} \left| \oint_{\mathcal{R}_k} \Big(X'_N(z) - \hat{X}'_N(z) \Big) dz \right| \leq 2 \int_0^{\frac{\varepsilon_N}{N}} & \left| X'_N(p_k + \mathbf{i}x) - \hat{X}'_N(p_k + \mathbf{i}\frac{\varepsilon_N}{N}) \right| \\ & + \left| X'_N(q_k + \mathbf{i}x) - \hat{X}'_N(q_k + \mathbf{i}\frac{\varepsilon_N}{N}) \right| dx. \end{aligned}$$

With probability one, for all N, M large,

$$\int_0^{\frac{\varepsilon_N}{N}} \left| m'_{\underline{\hat{\mathbf{R}}}_N}(p_k + \mathbf{i}x) - m'_{\underline{\hat{\mathbf{R}}}_N}(p_k + \mathbf{i}\frac{\varepsilon_N}{N}) \right| dx \leq \frac{\varepsilon_N}{N} (|p_k - a_k|^{-2} + |p_k - b_{k-1}|^{-2})$$

and

$$\int_0^{\frac{\varepsilon_N}{N}} \left| \underline{m}'_N(p_k + \mathbf{i}x) - \underline{m}'_N(p_k + \mathbf{i}\frac{\varepsilon_N}{N}) \right| dx \leq 2K' \frac{\varepsilon_N}{N}$$

where $K' = \sup_{z\in\mathcal{R}_k} |\underline{m}'_N(z)|$. So with probability one, for all N, M large,

$$\left|\oint_{\mathcal{R}_k} \Big(X_N(z) - \hat{X}_N(z)\Big)dz\right| \leq K_1\varepsilon_N, \tag{4.3.7}$$

$$\left|\oint_{\mathcal{R}_k} \Big(X'_N(z) - \hat{X}'_N(z)\Big)dz\right| \leq K_2\varepsilon_N \tag{4.3.8}$$

for some constants K_1 and K_2. Both terms converge to zero as $N, M \to \infty$. Then, by Slutsky's lemma [28], it suffices to establish the arguments for $\hat{X}_N(z)$ and $\hat{X}'_N(z)$.

As mentioned in Section 4.1.4, there are two conditions to prove (see for instance Billingsley [18, Theorem 13.1]) to establish the convergence in distribution of the process $(\hat{X}_N, \hat{X}'_N)$ to the process (X, Y) over the compact $\mathcal{K}$:

- Finite-dimensional convergence of the process $(\hat{X}_N, \hat{X}'_N)$ over the compact $\mathcal{K}$.
- Tightness on the compact $\mathcal{K}$.

Finite-dimensional convergence

In [11], Bai and Silverstein establish a central limit theorem for $F^{\hat{\mathbf{R}}_N}$ with the complex Gaussian entries X_{ij}. We recall below their main result.

Proposition 4.3.1 (cf. [11]) *With the notations introduced in Section 4.1.1, for $f_1, \ldots, f_p$, analytic on an open region containing $\mathbb{R}$,*

1. $\left(N \int f_i(x)d(F^{\hat{\mathbf{R}}_N} - F_N)(x)\right)_{1\leq i\leq p}$ *forms a tight sequence on N,*
2. $$\left(N \int f_i(x)d(F^{\hat{\mathbf{R}}_N} - F_N)(x)\right)_{1\leq i\leq p} \xrightarrow[N,M\to\infty]{\mathcal{D}} \mathcal{N}_p(0, \mathbf{V}),$$
   where $\mathbf{V} = (V_{ij})$ and
   $$V_{ij} = -\frac{1}{4\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} f_i(z_1)f_j(z_2)v_{ij}(z_1, z_2)dz_1dz_2,$$
   with
   $$v_{ij}(z_1, z_2) = \frac{\underline{m}'(z_1)\underline{m}'(z_2)}{(\underline{m}(z_1) - \underline{m}(z_2))^2} - \frac{1}{(z_1 - z_2)^2}$$
   where the integration is over positively oriented contours $\mathcal{C}_1$ and $\mathcal{C}_2$ which are supposed to be non-overlapping and both circle around the support $\mathcal{S}$.

Now we apply Proposition 4.3.1 to establish the finite-dimensional convergence. For all $z_i \in \mathcal{K}\backslash\mathbb{R}$, note that

$$m_{\hat{\mathbf{R}}_N}(z) - m_N(z) = \frac{1}{2i\pi}\oint_{\mathcal{C}} \frac{1}{x - z}d(F^{\hat{\mathbf{R}}_N} - F_N)(x)$$

with the contour $\mathcal{C}$ which contains the support $\mathcal{S}$. As $X_N(z) = M(m_{\underline{\hat{\mathbf{R}}}_N}(z) - \underline{m}_N(z))$, Proposition 4.3.1 directly implies that for every finite $p \in \mathbb{N}$, the random vector

$$\left(\hat{X}_N(z_1), \hat{X}'_N(z_1), \cdots, \hat{X}_N(z_p), \hat{X}'_N(z_p)\right)$$

converges to a centered Gaussian vector by considering the functions :

$$(f_1(x), f_2(x), \cdots, f_{2p-1}(x), f_{2p}(x))$$

where for all $1 \leq i \leq p$, $f_{2i-1}(x) = \frac{1}{x-z_i}$ and $f_{2i}(x) = \frac{1}{(x-z_i)^2}$. Hence the finite dimensional convergence.

The proof of the tightness is based on Poincaré-Nash inequality (see for instance [93] and [41]). In Appendix 4.3.1, it is proved that for all $\epsilon > 0$ and all $\ell \in \mathbb{N}$,

$$\mathbb{P}\left(\sup_{\lambda \in \mathrm{eig}(\hat{\mathbf{R}}_N)} d(\lambda, \mathcal{S}) > \epsilon \right) = o(N^{-\ell}).$$

Following the same idea as Bai and Silverstein [11, Section 3 and 4], it is indeed a tight sequence. The details of the proof are in Appendix 4.3.2. Thus Lemma 4.1.1 is proved.

Tightness

We will show the tightness of the sequence $M(m_{\underline{\hat{\mathbf{R}}}_N} - \underline{m}_N)$ and $M(m'_{\underline{\hat{\mathbf{R}}}_N} - \underline{m}'_N)$ by using Poincaré-Nash's inequality [41] on the compact $\mathcal{K}$. As the compact $\mathcal{K}$ is the union of $2L$ contours $\mathcal{R}_k$ and $\mathcal{R}'_\ell$, it is sufficient to prove the tightness on every contour $\mathcal{R}_k$ (or equivalently $\tilde{\mathcal{R}}_k$). First, denote by $M(m_{\underline{\hat{\mathbf{R}}}_N}(z) - \underline{m}_N(z)) = X_N^1(z) + X_N^2(z)$ with $X_N^1(z) = M(m_{\underline{\hat{\mathbf{R}}}_N}(z) - \mathbb{E}[m_{\underline{\hat{\mathbf{R}}}_N}(z)])$ and $X_N^2(z) = M(\mathbb{E}[m_{\underline{\hat{\mathbf{R}}}_N}(z)] - \underline{m}_N(z))$.
We now prove tightness based on [18, Theorem 13.1], i.e.

1. Tightness at any point of the contour (here $\tilde{\mathcal{R}}_k$).
2. Satisfaction of the condition

$$\sup_{N, z_1, z_2 \in \tilde{\mathcal{R}}_k} \frac{\mathbb{E}|(\hat{X}_N^1(z_1) - \hat{X}_N^1(z_2))|^2}{|z_1 - z_2|^2} \leq K.$$

Condition 1) is achieved by an immediate application of Proposition 4.3.1. We now verify the second condition.

We evaluate $\frac{\mathbb{E}|(\hat{X}_N^1(z_1) - \hat{X}_N^1(z_2))|^2}{|z_1 - z_2|^2}$. Denote by $\hat{\lambda}'_1 \leq \cdots \leq \hat{\lambda}'_M$, the eigenvalues of $\underline{\hat{\mathbf{R}}}_N$. Note that

$$\begin{aligned} m_{\underline{\hat{\mathbf{R}}}_N}(z_1) - m_{\underline{\hat{\mathbf{R}}}_N}(z_2) &= \frac{z_1 - z_2}{M} \sum_{i=1}^{M} \frac{1}{(\hat{\lambda}'_i - z_1)(\hat{\lambda}'_i - z_2)} \\ &= \frac{z_1 - z_2}{M} \mathrm{Tr}(\mathbf{D}_N^{-1}(z_1) \mathbf{D}_N^{-1}(z_2)) \end{aligned}$$

with $\mathbf{D}_N(z) = \underline{\hat{\mathbf{R}}}_N - z\mathbf{I}_M$. We have

$$\begin{aligned}
&\frac{\partial}{\partial Y_{i,j}}\left(\frac{m_{\underline{\hat{\mathbf{R}}}_N}(z_1) - m_{\underline{\hat{\mathbf{R}}}_N}(z_2)}{z_1 - z_2}\right) \\
&= \frac{\partial}{\partial Y_{i,j}}\mathrm{Tr}(\underline{\hat{\mathbf{R}}}_N - z_1\mathbf{I}_M)^{-1}(\underline{\hat{\mathbf{R}}}_N - z_2\mathbf{I}_M)^{-1} \\
&= \frac{1}{M}\left[-\mathbf{Y}_N^*\mathbf{D}_N^{-2}(z_1)\mathbf{D}_N^{-1}(z_2) - \mathbf{Y}_N^*\mathbf{D}_N^{-1}(z_1)\mathbf{D}_N^{-2}(z_2)\right]_{j,i},
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\partial}{\partial \bar{Y}_{i,j}}\left(\frac{m_{\underline{\hat{\mathbf{R}}}_N}(z_1) - m_{\underline{\hat{\mathbf{R}}}_N}(z_2)}{z_1 - z_2}\right) \\
&= \frac{1}{M}[-\mathbf{D}_N^{-2}(z_1)\mathbf{D}_N^{-1}(z_2)\mathbf{Y}_N - \mathbf{D}_N^{-1}(z_1)\mathbf{D}_N^{-2}(z_2)\mathbf{Y}_N]_{i,j}.
\end{aligned}$$

Then by the Poincaré-Nash inequality and the fact that $\underline{\hat{\mathbf{R}}}_N$ is uniformly bounded in spectral norm almost surely, one gets

$$\begin{aligned}
\frac{\mathbb{E}|\hat{X}_N^1(z_1) - \hat{X}_N^1(z_2)|^2}{|z_1 - z_2|^2} &\leq \frac{C_1}{N}\mathbb{E}\Big[\mathrm{Tr}(\mathbf{L}_{N,1})\Big] \\
&= \frac{C_1}{N}\mathbb{E}(\mathrm{Tr}(\mathbf{L}_{N,1})\mathrm{I}_{\sup_n d(\hat{\lambda}'_n,\underline{\mathcal{S}})\leq\varepsilon}) + \frac{C_1}{N}\mathbb{E}(\mathrm{Tr}(\mathbf{L}_{N,1})\mathrm{I}_{\sup_n d(\hat{\lambda}'_n,\underline{\mathcal{S}})>\varepsilon})
\end{aligned}$$

with

$$\begin{aligned}
\mathbf{L}_{N,1} =& \underline{\hat{\mathbf{R}}}_N\mathbf{D}_N^{-4}(z_1)\mathbf{D}_N^{-2}(z_2) + 2\underline{\hat{\mathbf{R}}}_N\mathbf{D}_N^{-3}(z_1)\mathbf{D}_N^{-3}(z_2) \\
&+ \underline{\hat{\mathbf{R}}}_N\mathbf{D}_N^{-2}(z_1)\mathbf{D}_N^{-4}(z_2)
\end{aligned}$$

and C_1 a constant which does not depend on N or M. For the first term, $\mathrm{Tr}(\mathbf{L}_{N,1})$ is bounded on the set $\sup_n d(\hat{\lambda}'_n, \underline{\mathcal{S}}) \leq \varepsilon$. For the second term, since for all $i \in \mathbb{N}$ and all $z \in \tilde{\mathcal{R}}_k$, $\frac{1}{|\hat{\lambda}'_n - z|^i} \leq \frac{N^i}{\varepsilon_N^i}$, it leads to

$$\sum_{n=1}^{N}\frac{1}{|\hat{\lambda}'_n - z|^i} \leq \frac{N^{i+1}}{\varepsilon_N^i}.$$

Then

$$|\mathrm{Tr}(\mathbf{L}_{N,1})| \leq \mathcal{O}\left(\frac{N^7}{\varepsilon_N^6}\right).$$

As $\mathbb{P}(\sup d(\hat{\lambda}'_n, \underline{\mathcal{S}}) \geq \varepsilon) = \mathbb{P}(\sup d(\hat{\lambda}_n, \mathcal{S}) \geq \varepsilon) = o(N^{-16})$, take $\varepsilon_N = N^{-0.01}$, one obtains

$$\begin{aligned}
\left|\mathbb{E}(\mathrm{Tr}(\mathbf{L}_{N,1})\mathrm{I}_{\sup_n d(\hat{\lambda}'_n,\underline{\mathcal{S}})>\varepsilon})\right| &\leq \mathbb{E}\left|\mathrm{Tr}(\mathbf{L}_{N,1})I_{\sup d(\hat{\lambda}'_n,\underline{\mathcal{S}})>\varepsilon}\right| \\
&\leq \mathcal{O}\left(\frac{N^7}{\epsilon_N^6}\mathbb{P}(\sup d(\hat{\lambda}'_n, \underline{\mathcal{S}}) > \varepsilon)\right) \to 0.
\end{aligned}$$

The second condition to establish the tightness is achieved.

For $\hat{X}_N^2(z)$, following exactly the same method as in [5, Section 9.11], one can prove that $\hat{X}_N^2(z)$ is bounded and forms an equicontinuous family that converges to 0. Hence the tightness for $M(m_{\underline{\hat{\mathbf{R}}}_N}(z) - \underline{m}_N(z))$.

The next step is to prove the tightness of $M(m'_{\hat{\underline{\mathbf{R}}}_N}(z) - \underline{m}'_N(z))$. We have

$$\begin{aligned}
m'_{\hat{\underline{\mathbf{R}}}_N}(z_1) - m'_{\hat{\underline{\mathbf{R}}}_N}(z_2) &= \frac{z_1 - z_2}{M} \sum_{i=1}^{M} \frac{2\hat{\lambda}'_i - z_1 - z_2}{(\hat{\lambda}'_i - z_1)^2(\hat{\lambda}'_i - z_2)^2} \\
&= \frac{z_1 - z_2}{M} \mathrm{Tr}\left(\mathbf{D}_N^{-2}(z_1)\mathbf{D}_N^{-2}(z_2)(\mathbf{D}_N(z_1) + \mathbf{D}_N(z_2))\right).
\end{aligned}$$

Following the same method as derived before, one obtains

$$\begin{aligned}
&\frac{\partial}{\partial Y_{ij}} \mathbf{D}_N^{-1}(z_1)\mathbf{D}_N^{-2}(z_2) \\
&= -\frac{1}{M}\left[\mathbf{Y}_N^*\mathbf{D}_N^{-2}(z_1)\mathbf{D}_N^{-2}(z_2) + 2\mathbf{Y}_N^*\mathbf{D}_N^{-1}(z_1)\mathbf{D}_N^{-3}(z_2)\right]_{j,i},
\end{aligned}$$

and

$$\left|\frac{\partial}{\partial Y_{ij}} \mathrm{Tr}\mathbf{D}_N^{-2}(z_1)\mathbf{D}_N^{-2}(z_2)(\mathbf{D}(z_1) + \mathbf{D}_N(z_2))\right|^2 = \frac{1}{M}\mathrm{Tr}(\mathbf{L_2})$$

with

$$\begin{aligned}
\mathbf{L}_{N,2} = & 4\hat{\underline{\mathbf{R}}}_N\Big(3\mathbf{D}_N^{-4}(z_1)\mathbf{D}_N^{-4}(z_2) + 2\mathbf{D}_N^{-3}(z_1)\mathbf{D}_N^{-5}(z_2) + 2\mathbf{D}_N^{-5}(z_1)\mathbf{D}_N^{-3}(z_2) \\
& + \mathbf{D}_N^{-2}(z_1)\mathbf{D}_N^{-6}(z_2) + \mathbf{D}_N^{-6}(z_1)\mathbf{D}_N^{-2}(z_2)\Big).
\end{aligned}$$

Then Poincaré-Nash inequality yields that

$$\begin{aligned}
\mathrm{Var}\frac{|\hat{X}_N^{1\prime}(z_1) - \hat{X}_N^{1\prime}(z_2)|}{|z_1 - z_2|} \leq & \frac{C_1}{N}\mathbb{E}(\mathrm{Tr}(\mathbf{L}_{N,2})\mathrm{I}_{\sup_n d(\hat{\lambda}'_n, \underline{\mathcal{S}}) \leq \varepsilon}) \\
& + \frac{C_1}{N}\mathbb{E}(\mathrm{Tr}(\mathbf{L}_{N,2})\mathrm{I}_{\sup_n d(\hat{\lambda}'_n, \underline{\mathcal{S}}) > \varepsilon})
\end{aligned}$$

with C_1 the same constant defined as before. The term $\mathrm{Tr}(\mathbf{L}_{N,2})$ is bounded on the set $\sup d(\hat{\lambda}'_n, \underline{\mathcal{S}}) \leq \varepsilon$. For the second term, $|\mathrm{Tr}(\mathbf{L}_{N,2})| \leq \mathcal{O}\left(\frac{N^9}{\varepsilon_N^8}\right)$. As $\mathbb{P}(\sup d(\hat{\lambda}'_n, \underline{\mathcal{S}}) \geq \varepsilon) = o(N^{-16})$ and $\varepsilon_N = N^{-0.01}$,

$$\left|\mathbb{E}(\mathrm{Tr}(\mathbf{L}_{N,2})\mathrm{I}_{\sup_n d(\hat{\lambda}'_n, \underline{\mathcal{S}}) > \varepsilon})\right| \to 0.$$

The proof of the tightness of $\hat{X}_N^{1\prime}(z)$ is achieved as before.

The proof of the tightness is completed if one verifies that $\hat{X}_N^{2\prime}(z)$ for $z \in \tilde{\mathcal{R}}_k$ is bounded and forms an equicontinuous family, and converges to 0. We will use the same method for the process $\hat{X}_N^2(z)$ (see [5, Section 9.11]).

By Formula (9.11.1) in [5, Section 9.11], it is proved that :

$$(\mathbb{E}m_{\hat{\underline{\mathbf{R}}}_N} - \underline{m}_N)\left(1 - \frac{\frac{N}{M}\int \frac{\underline{m}_N t^2 dF^{\mathbf{R_N}}(t)}{(1+t\mathbb{E}m_{\hat{\underline{\mathbf{R}}}_N})(1+t\underline{m}_N)}}{-z + \frac{N}{M}\int \frac{t dF^{\mathbf{R_N}}}{1+t\mathbb{E}m_{\hat{\underline{\mathbf{R}}}_N}} - T_N}\right) = \mathbb{E}m_{\hat{\underline{\mathbf{R}}}_N}\underline{m}_N T_N \tag{4.3.9}$$

where

$$\begin{aligned}
T_N &= \frac{N}{M^2}\sum_{j=1}^{M}\mathbb{E}\beta_j d_j(\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N})^{-1},\\
d_j &= -\mathbf{q}_j^*\mathbf{R}^{1/2}(\hat{\mathbf{R}}_{(j)}-z\mathbf{I})^{-1}(\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}\mathbf{R}+\mathbf{I})^{-1}\mathbf{R}^{1/2}\mathbf{q}_j\\
&+(1/M)\mathrm{Tr}(\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}\mathbf{R}+\mathbf{I})^{-1}\mathbf{R}(\hat{\mathbf{R}}_N-z\mathbf{I})^{-1},\\
\beta_j &= \frac{1}{1+\frac{1}{M}\mathbf{y}_j^*(\hat{\mathbf{R}}_{(j)}-z\mathbf{I})^{-1}\mathbf{y}_j},\\
\mathbf{q}_j &= 1/\sqrt{N}\mathbf{x}_j,\\
\hat{\mathbf{R}}_{(j)} &= \hat{\mathbf{R}}_N-\frac{1}{M}\mathbf{y}_j\mathbf{y}_j^*.
\end{aligned}$$

If one differentiates (4.3.9) with respect to z, the equation becomes

$$\begin{aligned}
&(\mathbb{E}m'_{\underline{\hat{\mathbf{R}}}_N}-\underline{m}'_N)\left(1-\frac{\frac{N}{M}\int\frac{\underline{m}_N t^2 dF^{\mathbf{R_N}}(t)}{(1+t\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N})(1+t\underline{m}_N)}}{-z+\frac{N}{M}\int\frac{tdF^{\mathbf{R_N}}}{1+t\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}}-T_N}\right)\\
&+(\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}-\underline{m}_N)\left(1-\frac{\frac{N}{M}\int\frac{\underline{m}_N t^2 dF^{\mathbf{R_N}}(t)}{(1+t\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N})(1+t\underline{m}_N)}}{-z+\frac{N}{M}\int\frac{tdF^{\mathbf{R_N}}}{1+t\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}}-T_N}\right)'\\
&=\mathbb{E}m'_{\underline{\hat{\mathbf{R}}}_N}\underline{m}_N T_N+\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}\underline{m}'_N T_N+\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}\underline{m}_N T'_N.
\end{aligned}$$

In the work of [5, Section 9.11], it is proved that when N, M tend to infinity,

1. $\sup_{z\in\tilde{\mathcal{R}}_k}|\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}(z)-\underline{m}(z)|\to 0$ and $\sup_{z\in\tilde{\mathcal{R}}_k}|\underline{m}_N(z)-\underline{m}(z)|\to 0$,
2. $\dfrac{\frac{N}{M}\int\frac{t^2\underline{m}_N dF^{\mathbf{R_N}}(t)}{(1+t\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N})(1+t\underline{m}_N)}}{-z+\frac{N}{M}\int\frac{tdF^{\mathbf{R_N}}}{1+t\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}}-T_N}$ converges ,
3. $\hat{X}_N^2(z)\to 0, \quad T_N\to 0.$

   With the same method, one can show easily that
4. $\sup_{z\in\tilde{\mathcal{R}}_k}|\mathbb{E}m'_{\underline{\hat{\mathbf{R}}}_N}(z)-\underline{m}'(z)|\to 0$,
5. $\sup_{z\in\tilde{\mathcal{R}}_k}|\underline{m}'_N(z)-\underline{m}'(z)|\to 0$,
6. $\frac{N}{M}\left(\sum_{j=1}^M\mathbb{E}\beta_j d_j\right)'$ converges.

With these results, it suffices to show that $T'_N\to 0$, and $\hat{X}_N^{2\prime}$ is equicontinuous.

In [5, Section 9.9], they show that for $m,p\in\mathbb{N}$ and a non-random $N\times N$ matrix $\mathbf{A}_k$, $k=1,..,m$ and $\mathbf{B}_\ell$, $\ell=1,..,q$, we have

$$\begin{aligned}
&\left|\mathbb{E}\left(\prod_{k=1}^m\mathbf{r}_t^*\mathbf{A}_k\mathbf{r_t}\prod_{\ell=\mathbf{1}}^{\mathbf{q}}(\mathbf{r_t^*R_\ell r_t}-\mathbf{M^{-1}}\mathrm{Tr}\mathbf{RB}_\ell)\right)\right|\\
&\le KM^{-(1\wedge q)}\prod_{k=1}^m\|\mathbf{A}_k\|\prod_{\ell=1}^q\|\mathbf{B}_\ell\|.
\end{aligned}\tag{4.3.10}$$

We have also that for any positive p,

$$\max_{z\in\hat{\mathcal{R}_k},N} (\mathbb{E}\|\mathbf{D}^{-1}(z)\|^p, \mathbb{E}\|\mathbf{D}_j^{-1}(z)\|^p, \mathbb{E}\|\mathbf{D}_{ij}^{-1}(z)\|^p) \leq K_p \tag{4.3.11}$$

and

$$\sup_{N,M,z\in\tilde{\mathcal{R}}_k} \|(\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}(z)\mathbf{R}+\mathbf{I})^{-1}\| < \infty \tag{4.3.12}$$

where K_p is a constant which depends only on p.

With all these preliminaries, as $T_N \to 0$, by the dominated convergence theorem of derivation, it suffices to show that T_N' is bounded over $\tilde{\mathcal{R}}_k$. In [5, Section 9.11], it is sufficient to show that $(f_M'(z))$ is bounded where

$$\begin{aligned} f_M(z) = \sum_{j=1}^{M} \mathbb{E}\Big[\Big(&\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j - M^{-1}\mathrm{Tr}\mathbf{D}_j^{-1}\mathbf{R}) \\ &\times (\mathbf{r}_j^*\mathbf{D}_j^{-1}\ (\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}\mathbf{R}+\mathbf{I})^{-1}\mathbf{r}_j - M^{-1}\mathrm{Tr}\mathbf{D}_j^{-1}(\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}\mathbf{R}+\mathbf{I})^{-1}\mathbf{R}\Big)\Big]. \end{aligned}$$

With the help of (4.3.10)-(4.3.12), $f_M'(z)$ is indeed bounded in $\tilde{\mathcal{R}}_k$.

Now we will show that $\hat{X}_N^{2\prime}$ is equicontinuous. With the light work as before, it is sufficient to show that $f_M''(z)$ is bounded. Using (4.3.10), in the development of $f_M''(z)$, all terms depend on

$$\frac{1}{M}\mathrm{Tr}(\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}\mathbf{R}_N+\mathbf{I}_N)^{-k_1}\mathbf{D}_1^{-k_2}\mathbf{R}_N^{k_3}\mathbb{E}\mathrm{Tr}\mathbf{D}_1^{-k_4}\mathbf{R}_N^{k_5}$$

with certain $k_1, k_2, k_3, k_4, k_5 \in \mathbb{N}$. Thanks to (4.3.11) and (4.3.12), these terms are all bounded. Therefore, RHS is indeed bounded. This ends the proof of the tightness.

4.3.3 Study of the set $\mathcal{R}_{z_2}$

For z_2 fixed, denote $\mathcal{R}_{z_2} = \{z_1 \in \mathbb{C} : z_1 \neq z_2, m_{\underline{\hat{\mathbf{R}}}_N}(z_1) = m_{\underline{\hat{\mathbf{R}}}_N}(z_2)\}$. We will show that this set is empty a.s. for all N, M large. Suppose that $z_1 \in \mathcal{R}_{z_2}$. We first use [5, Formula (9.11.4)] that

$$\sup_{z\in\mathcal{R}_{z_2}} |\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}(z) - \underline{m}_N(z)| \to 0.$$

By Formula (4.3.1), we get

$$z(\underline{m}_N) = -\frac{1}{\underline{m}_N} + \frac{N}{M}\int \frac{t}{1+t\underline{m}_N}F^{R_N}(dt).$$

As $z_1 \neq z_2$,

$$\underline{m}_N(z_1) \neq \underline{m}_N(z_2).$$

Take $\epsilon = |\underline{m}_N(z_1) - \underline{m}_N(z_2)|$. For N sufficiently high,

$$|\mathbb{E}m_{\underline{\hat{\mathbf{R}}}_N}(z_1) - \underline{m}_N(z_1)| < \frac{\epsilon}{4}$$

and

$$|\mathbb{E}m_{\hat{\underline{\mathbf{R}}}_N}(z_2) - \underline{m}_N(z_2)| < \frac{\epsilon}{4}.$$

Finally

$$\begin{aligned}
&|\mathbb{E}m_{\hat{\underline{\mathbf{R}}}_N}(z_1) - \mathbb{E}m_{\hat{\underline{\mathbf{R}}}_N}(z_2)| \\
&\geq \epsilon - |\mathbb{E}m_{\hat{\underline{\mathbf{R}}}_N}(z_1) - \underline{m}_N(z_1)| - |\mathbb{E}m_{\hat{\underline{\mathbf{R}}}_N}(z_2) - \underline{m}_N(z_2)| \\
&\geq \frac{\epsilon}{2},
\end{aligned}$$

which implies, along with $m_{\hat{\underline{\mathbf{R}}}_N}(z) - \mathbb{E}m_{\hat{\underline{\mathbf{R}}}_N}(z) \xrightarrow{a.s.} 0$ that for all large N, M, $\mathcal{R}_{z_2} = \varnothing$ almost surely.

4.3.4 Proof of lemma 4.2.1

By Cauchy's formula, write :

$$\sum_{k=1}^{L} \frac{N_k}{N}\rho_k^\ell = \frac{1}{2\mathbf{i}\pi N}\oint_\Gamma \sum_{r=1}^{L} \frac{N_r\omega^\ell}{\omega - \rho_r}d\omega,$$

where Γ is a counterclockwise oriented contour that surrounds all the eigenvalues $\{\rho_1, \cdots, \rho_L\}$. Performing the changing variable $\omega = -\frac{1}{\underline{m}_N(z)}$ in the same manner as in [64], we get :

$$\sum_{k=1}^{L} \frac{N_k}{N}\rho_k^\ell = \frac{(-1)^{\ell+1}}{2\mathbf{i}\pi N}\oint_{\mathcal{C}} \sum_{r=1}^{L} \frac{N_r\underline{m}_N'(z)dz}{\underline{m}_N^{\ell+1}(z)\left(\rho_r\underline{m}_N(z)+1\right)},$$

where the contour $\mathcal{C}$ is counterclockwise oriented which contains the whole support $\mathcal{S}$.

From (4.1.7), we can establish that :

$$m_N(z) = -\frac{1}{Nz}\sum_{r=1}^{L}\frac{N_r}{1+\rho_r\underline{m}_N(z)},$$

thus yielding :

$$\sum_{k=1}^{L} \frac{N_k}{N}\rho_k^\ell = \frac{(-1)^\ell}{2\mathbf{i}\pi}\oint_{\mathcal{C}} \frac{z\underline{m}_N'(z)}{\underline{m}_N^{\ell+1}(z)}m_N(z)dz. \tag{4.3.13}$$

Plugging the relation :

$$m_N(z) = \frac{M}{N}\underline{m}_N(z) + \frac{M(1-\frac{N}{M})}{Nz}$$

into (4.3.13), we obtain :

$$\sum_{k=1}^{L} \frac{N_k}{N}\rho_k^\ell = \frac{(-1)^\ell}{2\mathbf{i}\pi}\oint_{\mathcal{C}} \frac{Mz\underline{m}_N'(z)dz}{N\underline{m}_N^{\ell}(z)} + \frac{(-1)^\ell}{2\mathbf{i}\pi}\oint_{\mathcal{C}} \frac{M(1-\frac{N}{M})\underline{m}_N'(z)}{N\underline{m}_N^{\ell+1}(z)}dz. \tag{4.3.14}$$

Since $\frac{\underline{m}'_N(z)}{\underline{m}_N^{\ell+1}(z)}$ is the derivative of $-\frac{1}{\ell \underline{m}_N^{\ell}(z)}$,

$$\oint_{\mathcal{C}} \frac{\underline{m}'_N(z)}{\underline{m}_N^{\ell+1}(z)} dz = 0.$$

The second term on the right hand side of (4.3.14) is then equal to zero. It remains thus to deal with $\oint_{\mathcal{C}} \frac{z\underline{m}'_N(z)}{\underline{m}_N^{\ell}(z)}$. If $\ell \geq 2$, by integration by parts, we obtain :

$$\oint_{\mathcal{C}} \frac{z\underline{m}'_N(z)}{\underline{m}_N^{\ell}(z)} dz = \frac{1}{\ell - 1} \oint_{\mathcal{C}} \frac{dz}{\underline{m}_N^{\ell-1}(z)}.$$

We thus obtain :

$$\sum_{k=1}^{L} \frac{N_k}{N} \rho_k^{\ell} = \frac{M(-1)^{\ell}}{2\mathbf{i}\pi N(\ell-1)} \oint_{\mathcal{C}} \frac{dz}{\underline{m}_N^{\ell-1}(z)}. \tag{4.3.15}$$

This proves that the theoretical moments admit the following integral representation :

$$\begin{aligned} \gamma_1 &= -\frac{M}{2N\mathbf{i}\pi} \oint_{\mathcal{C}} \frac{z\underline{m}'_N(z)}{\underline{m}_N(z)} dz \\ \gamma_l &= \frac{M(-1)^{\ell}}{2\mathbf{i}N(\ell-1)} \oint_{C} \frac{dz}{\underline{m}_N^{l-1}(z)}, \quad 2 \leq \ell \leq 2L-1. \end{aligned}$$

Finally, we show that consistent estimates of γ_i can be obtained by substituting the unknown term $\underline{m}_N(z)$ by its asymptotic equivalent $m_{\underline{\hat{\mathbf{R}}}_N}(z)$. Let $\hat{\gamma}_0, \cdots, \hat{\gamma}_{2L-1}$ the real quantities given by :

$$\begin{aligned} \hat{\gamma}_0 &= 1, \\ \hat{\gamma}_1 &= -\tfrac{M}{2N\mathbf{i}\pi} \oint_{\mathcal{C}} \tfrac{z m'_{\underline{\hat{\mathbf{R}}}_N}(z)}{m_{\underline{\hat{\mathbf{R}}}_N}(z)} dz, \\ &\vdots \\ \hat{\gamma}_{2L-1} &= \tfrac{M(-1)^{2L-1}}{2N(2L-1)\mathbf{i}\pi} \oint_{\mathcal{C}} \tfrac{dz}{m_{\underline{\hat{\mathbf{R}}}_N}^{2L-1}(z)}. \end{aligned}$$

Then, by the dominated convergence theorem and the fact that with probability one [5, Section 9.12], for all N, M large enough,

$$\inf_{z \in \mathcal{C}} |\underline{m}_N(z)| > 0$$

and

$$\inf_{z \in \mathcal{C}} |m_{\underline{\hat{\mathbf{R}}}_N}(z)| > 0,$$

one obtains : for all $k \geq 2$,

$$\left| \int_{\mathcal{C}} \frac{dz}{\underline{m}_N^{k-1}(z)} - \int_{\mathcal{C}} \frac{dz}{m_{\underline{\hat{\mathbf{R}}}_N}^{k-1}(z)} \right| \xrightarrow{a.s.} 0$$

and

$$\left| \int_{\mathcal{C}} \frac{m'_{\underline{\hat{\mathbf{R}}}_N}(z) dz}{m_{\underline{\hat{\mathbf{R}}}_N}(z)} - \int_{\mathcal{C}} \frac{\underline{m}'_N(z) dz}{\underline{m}_N(z)} \right| \xrightarrow{a.s.} 0.$$

Consequently :

$$\hat{\gamma}_i - \gamma_i \xrightarrow[N,M\to\infty]{a.s.} 0.$$

4.3.5 Calculation of the variance

In this section, we will show the calculations of the variance matrix $\mathbf{V}$. The computation of $\mathbf{V}_{1,1}$ has been carried out in Theorem 4.1.3 where it was shown that :

$$\mathbf{V}_{1,1} = -\frac{1}{4\pi^2 c^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2}\left[\frac{\underline{m}'(z_1)\underline{m}'(z_2)}{(\underline{m}(z_1)-\underline{m}(z_2))^2} - \frac{1}{(z_1-z_2)^2}\right]\frac{1}{\underline{m}(z_1)\underline{m}(z_2)}dz_1dz_2,$$

with $\mathcal{C}_1$ and $\mathcal{C}_2$ defined in the theorem. Using the fact that $\inf_{z\in\mathcal{C}}|\underline{m}(z)| > 0$ together with Fubini's theorem, the quantity $\mathbf{V}_{k,\ell}$ for $k\geq 2, \ell\geq 2$, becomes :

$$\mathbf{V}_{k,\ell} = -\frac{(-1)^{k+\ell}}{4\pi^2c^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2}\mathbb{E}\left[X(z_1)X(z_2)\right]\underline{m}^{-k}(z_1)\underline{m}^{-\ell}(z_2)dz_1dz_2.$$

Substituting $\mathbb{E}\left[X(z_1)X(z_2)\right]$ by $\kappa(z_1,z_2)$, we obtain :

$$\mathbf{V}_{k,\ell} = -\frac{(-1)^{k+\ell}}{4\pi^2 c^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2}\left[\frac{\underline{m}'(z_1)\underline{m}'(z_2)}{(\underline{m}(z_1)-\underline{m}(z_2))^2} - \frac{1}{(z_1-z_2)^2}\right]\frac{1}{\underline{m}^k(z_1)\underline{m}^\ell(z_2)}dz_1dz_2.$$

Finally, it remains to compute $\mathbf{V}_{k,1}$. Expanding $\Upsilon(X,Y,\underline{m},\underline{m}')$ and $\Phi_k(X,\underline{m})$, we obtain :

$$\begin{aligned}\mathbf{V}_{k,1} &= -\frac{(-1)^{k+1}}{4\pi^2c^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2}\left[\frac{z_2}{\underline{m}(z_2)\underline{m}^k(z_1)}\mathbb{E}\left[X(z_1)X'(z_2)\right]dz_1dz_2\right.\\ &\qquad\qquad\left. -\frac{\underline{m}'(z_2)}{\underline{m}(z_2)^2\underline{m}^k(z_1)}\mathbb{E}\left[X(z_1)X(z_2)\right]\right]dz_1dz_2\\ &= -\frac{(-1)^{k+1}}{4\pi^2c^2}\left(\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2}\frac{z_2\partial_2\kappa(z_1,z_2)}{\underline{m}(z_2)\underline{m}(z_1)^k}dz_1dz_2 - \oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2}\frac{\underline{m}'(z_2)\kappa(z_1,z_2)}{\underline{m}^2(z_2)\underline{m}^k(z_1)}dz_1dz_2\right).\end{aligned}$$

By integration by parts, we obtain :

$$\oint_{\mathcal{C}_2}\frac{z_2\partial_2\kappa(z_1,z_2)}{\underline{m}(z_2)\underline{m}^k(z_1)}dz_2 = -\oint_{\mathcal{C}_2}\frac{\kappa(z_1,z_2)}{\underline{m}(z_2)\underline{m}^k(z_1)}dz_2 + \oint_{\mathcal{C}_2}\frac{\underline{m}'(z_2)\kappa(z_1,z_2)}{\underline{m}(z_2)^2\underline{m}^k(z_1)}dz_2.$$

Hence,

$$\mathbf{V}_{k,1} = -\frac{(-1)^{k+1}}{4\pi^2c^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2}\frac{\kappa(z_1,z_2)dz_1dz_2}{\underline{m}(z_2)\underline{m}^k(z_1)}.$$

This extends the expression of $\mathbf{V}_{k,\ell}$ for any $k,\ell\in\{1,\cdots,L-1\}$, thus yielding :

$$\mathbf{V}_{k,\ell} = -\frac{(-1)^{k+\ell}}{4\pi^2 c^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2}\left[\frac{\underline{m}'(z_1)\underline{m}'(z_2)}{(\underline{m}(z_1)-\underline{m}(z_2))^2} - \frac{1}{(z_1-z_2)^2}\right]\frac{1}{\underline{m}^k(z_1)\underline{m}^\ell(z_2)}dz_1dz_2. \tag{4.3.16}$$

Bibliographie

[1] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[2] G. W. Anderson and O. Zeitouni. A CLT for a band matrix model. *Probab. Theory Related Fields*, 134(2) :283–338, 2006.

[3] L. Arnold. On the asymptotic distribution of the eigenvalues of random matrices. *Journal of Mathematics and Analytic Applications*, 20 :262–268, 1967.

[4] Z. Bai, J. Chen, and J. Yao. On estimation of the population spectral distribution from a high-dimensional sample covariance matrix. *Australian and New Zealand Journal of Statistics*, 52(4) :423–437, 2010.

[5] Z. Bai and J. W. Silverstein. Spectral Analysis of Large Dimensional Random Matrices. *Springer Series in Statistics*, 2009.

[6] Z. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices.* Springer Series in Statistics. Springer, New York, second edition, 2010.

[7] Z. Bai and J. W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of information-plus-noise type matrices. *Random Matrices Theory Appl.*, 1(1) :1150004, 44, 2012.

[8] Z. Bai, X. Wang, and W. Zhou. Functional CLT for sample covariance matrices. *Bernoulli*, 16(4) :1086–1113, 2010.

[9] Z. D. Bai and J. W. Silverstein. No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Large Dimensional Sample Covariance Matrices. *Annals of Probability*, 26(1) :316–345, Jan. 1998.

[10] Z. D. Bai and J. W. Silverstein. Exact Separation of Eigenvalues of Large Dimensional Sample Covariance Matrices. *The Annals of Probability*, 27(3) :1536–1555, 1999.

[11] Z. D. Bai and J. W. Silverstein. CLT of linear spectral statistics of large dimensional sample covariance matrices. *Annals of Probability*, 32(1A) :553–605, 2004.

[12] Z. Bao, G. M. Pan, and W. Zhou. On the MIMO channel capacity for the general channels. preprint, 2013.

[13] G. Ben Arous and A. Guionnet. Erratum : “Large deviations for Langevin spin glass dynamics”. *Probab. Theory Related Fields*, 103(3) :431, 1995.

[14] F. Benaych-George, A. Guionnet, and C. Male. Central limit theorems for linear statistics of heavy tailed random matrices. *preprint*, 2013.

[15] P. J. Bickel and E. Levina. Regularized estimation of large covariance matrices. *Ann. Statist.*, 36(1) :199–227, 2008.

[16] P. J. Bickel and E. Levina. Regularized estimation of large covariance matrices. *Ann. Statist.*, 36(1) :199–227, 2008.

[17] P. Billingsley. *Convergence of probability measures.* John Wiley & Sons, Inc., Hoboken, NJ, 1968.

[18] P. Billingsley. *Probability and Measure.* John Wiley & Sons, Inc., Hoboken, NJ, third edition, 1995.

[19] T. Cabanal-Duvillard. Fluctuations de la loi empirique de grandes matrices aléatoires. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(3) :373–402, 2001.

[20] T. T. Cai, C.-H. Zhang, and H. H. Zhou. Optimal rates of convergence for covariance matrix estimation. *Ann. Statist.*, 38(4) :2118–2144, 2010.

[21] M. Capitaine. Exact separation phenomenon for the eigenvalues of large information-plus-noise type matrices. application to spiked models. Technical report, CNRS and Institut de mathématiques de Toulouse, available on `arXiv:1301.3940`, 2013.

[22] M. Capitaine, C. Donati-Martin, and D. Féral. The largest eigenvalues of finite rank deformation of large Wigner matrices : convergence and nonuniversality of the fluctuations. *Ann. Probab.*, 37(1) :1–47, 2009.

[23] S. Chatterjee. Fluctuations of eigenvalues and second order Poincaré inequalities. *Probab. Theory Related Fields*, 143(1-2) :1–40, 2009.

[24] R. Couillet and M. Debbah. Free deconvolution for OFDM multicell SNR detection. In *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'08)*, Cannes, France, 2008.

[25] R. Couillet and M. Debbah. *Random matrix methods for wireless communications.* Cambridge University Press, Cambridge, 2011.

[26] R. Couillet, J. Silverstein, Z. Bai, and M. Debbah. Eigen-inference for energy estimation of multiple sources. *Information Theory, IEEE Transactions on*, 57(4) :2420 –2439, april 2011.

[27] F. Cucker and A. Corbalan. An Alternate Proof of the Continuity of the Roots of a Polynomial. *The American Mathematical Monthly*, 96(4) :342–345, Apr. 1989.

[28] A. W. V. der Vaart. *Asymptotic statistics.* Cambridge University Press, New York, 2000.

[29] B. Dozier, J. W. Silverstein, and B. Dozier. Analysis of the Limiting Spectral Distribution of Large Dimensional Information-Plus-Noise Type Matrices. *Journal of Multivariate Analysis*, 98(6) :1099–1122, 2007.

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

[31] R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.

[32] J. Dumont, W. Hachem, P. Loubaton, and J. Najim. On the Capacity Achieving Covariance Matrix for Rician MIMO Channels : An Asymptotic Approach. *IEEE Trans. Inf. Theory*, 56(3) :1048–1069, 2010.

[33] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6(3) :311–335, Mar. 1998.

[34] V. L. Girko. Ten years of general statistical analysis.

[35] V. L. Girko. *Theory of stochastic canonical equations. Vol. I*, volume 535 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2001.

[36] V. L. Girko. *Theory of stochastic canonical equations. Vol. II*, volume 535 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2001.

[37] A. Guionnet. Large deviations upper bounds and central limit theorems for non-commutative functionals of Gaussian large random matrices. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(3) :341–384, 2002.

[38] U. Haagerup and S. Thorbjørnsen. A new application of random matrices : $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ is not a group. *Ann. of Math. (2)*, 162(2) :711–775, 2005.

[39] W. Hachem, M. Kharouf, J. Najim, and J. W. Silverstein. A CLT for information-theoretic statistics of non-centered Gram random matrices. *Random Matrices Theory Appl.*, 1(2) :1150010–1, 50, 2012.

[40] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur. A new approach for mutual information analysis of large dimensional multi-antenna channels. *Information Theory, IEEE Transactions on*, 54(9) :3987 –4004, sept. 2008.

[41] W. Hachem, O. Khorunzhy, P. Loubaton, J. Najim, and L. A. Pastur. A new approach for capacity analysis of large dimensional multi-antenna channels. *IEEE Trans. Inf. Theory*, 54(9), 2008.

[42] 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 Appl.*, 1(2) :1150006, 51, 2012.

[43] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet. A subspace estimator for fixed rank perturbations of large random matrices. *J. Multivariate Anal.*, 114 :427–447, 2013.

[44] W. Hachem, P. Loubaton, and J. Najim. Deterministic Equivalents for Certain Functionals of Large Random Matrices. *Annals of Applied Probability*, 17(3) :875–930, 2007.

[45] W. Hachem, P. Loubaton, and J. Najim. A CLT for Information Theoretic Statistics of Gram Random Matrices with a Given Variance Profile. *Annals of Probability*, 18(6) :2071–2130, Dec. 2008.

[46] W. Hachem, P. Loubaton, J. Najim, and P. Vallet. On bilinear forms based on the resolvent of large random matrices. *to be published in "Annales de l'Institut Henri Poincaré"*, 46(1), 2013.

[47] K. Johansson. On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.*, 91(1) :151–204, 1998.

[48] I. Jolliffe. *Principal Component Analysis.* Springer Verlag, New York, 1986.

[49] D. Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.*, 12(1) :1–38, 1982.

[50] O. Kallenberg. *Foundations of mordern Probability, 2nd edition.* Springer Verlag New York, 2002.

[51] A. Kammoun, R. Couillet, J. Najim, and M. Debbah. Performance of capacity inference methods under colored interference. *submitted to IEEE Trans. Inf. Th.*, 2011. [Online] `http://front.math.ucdavis.edu/1105.5305` .

[52] A. Kammoun, M. Kharouf, W. Hachem, and J. Najim. BER and Outage Probability Approximations for LMMSE Detectors on Correlated MIMO channels. *Information Theory*, 55(10) :4386–4397, 2009.

[53] A. Kammoun, M. Kharouf, W. Hachem, and J. Najim. A central limit theorem for the sinr at the lmmse estimator output for large-dimensional signals. *Information Theory, IEEE Transactions on*, 55(11) :5048 –5063, nov. 2009.

[54] N. E. Karoui. Spectrum estimation for large dimensional covariance matrices using random matrix theory. *Annals of Statistics*, 36(6) :2757–2790, Dec. 2008.

[55] A. M. Khorunzhy, B. A. Khoruzhenko, and L. A. Pastur. Asymptotic properties of large random matrices with independent entries. *J. Math. Phys.*, 37(10) :5033–5060, 1996.

[56] S. G. Krantz. *Function Theory of Several Complex Variables.* AMS Chelsea Publishing, 1992.

[57] O. Ledoit and M. Wolf. Nonlinear Shrinkage Estimation of Large-Dimensional Covariance Matrices . *Institute for Empirical Research in Economics University of Zurich*, October 2010.

[58] P. Loubaton and P. Vallet. Almost sure localization of the eigenvalues in a Gaussian information plus noise model—application to the spiked models. *Electron. J. Probab.*, 16 :no. 70, 1934–1959, 2011.

[59] F. Luo, Y. Yang, J. Zhong, H. Gao, L. Khan, D. K. Thompson, and J. Zhou. Constructing gene co-expression networks and predicting functions of unknown genes by random matrix theory. *BMC bioinformatics*, 8(1) :299, 2007.

[60] A. Lytova and L. Pastur. Central Limit Theorem for linear eigenvalue statistics of random matrices with independent entries. *The Annals of Probability*, 37(5) :1778–1840, 2009.

[61] J. Marsden and M. Hoffman. *Basic Complex Analysis, 3rd ed.* New York : Freeman, 1987.

[62] V. A. Marčenko and L. A. Pastur. Distributions of eigenvalues for some sets of random matrices. *Math USSR-Sbornik*, 1(4) :457–483, Apr. 1967.

[63] M. L. Mehta. *Random matrices.* ELSEVIER, Sand Diego, CA, first edition, 2004.

[64] X. Mestre. Improved estimation of eigenvalues of covariance matrices and their associated subspaces using their sample estimates. *IEEE Trans. Inf. Theory*, 54(11) :5113–5129, Nov. 2008.

[65] X. Mestre. On the asymptotic behavior of the sample estimates of eigenvalues and eigenvectors of covariance matrices. *IEEE Trans. Signal Process.*, 56(11) :5353–5368, Nov. 2008.

[66] X. Mestre and M. Lagunas. Modified Subspace Algorithms for DoA Estimation With Large Arrays. *IEEE Trans. Signal Process.*, 56(2) :598–614, Feb. 2008.

[67] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue. Subspace Methods for the Blind Identification of Multichannel FIR Filters. *IEEE Trans. Signal Process.*, 43(2), Feb. 1995.

[68] A. L. Moustakas, S. H. Simon, and A. M. Sengupta. MIMO capacity through correlated channels in the presence of correlated interferers and noise : A (not so) large N analysis. *IEEE Trans. Inf. Theory*, 49(10) :2545–2561, 2003.

[69] A. L. Moustakas, S. H. Simon, and A. M. Sengupta. MIMO capacity through correlated channels in the presence of correlated interferers and noise : a (not so) large N analysis. *IEEE Trans. Inform. Theory*, 49(10) :2545–2561, 2003. Special issue on space-time transmission, reception, coding and signal processing.

[70] J. Najim. Gaussian fluctuations for linear spectral statistics of large random covariance matrices. *submitted*, 2013.

[71] J. Najim and J. Yao. Fluctuations for linear spectral statistics of large random information-plusnoise type matrices. *preprint*, 2013.

[72] J. Najim and J. Yao. Fluctuations for linear spectral statistics of large random matrices :from analytic to non-analytic functionals. *preprint*, 2013.

[73] G. M. Pan and W. Zhou. Central limit theorem for signal-to-interference ratio of reduced rank linear receiver. *Ann. Appl. Probab.*, 18(3) :1232–1270, 2008.

[74] L. Pastur and M. Shcherbina. *Eigenvalue distribution of large random matrices*, volume 171 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2011.

[75] L. A. Pastur. A simple approach to global regime of random matrix theory. In *Mathematical results in statistical mechanics*, pages 429–454. World Scientific Publishing, 1999.

[76] V. Pisarenko. The retrieval of harmonics from a covariance function. *Geophysics J. Roy. Astron. Soc.*, 33 :347–366, Nov. 1973.

[77] V. Plerous, P. Gopikrishnan, B. Rosenow, L. Amaral, T. Guhr, and H. Stanley. Random matrix approach to cross correlations in financial data. *Phys. Rev. E*, 65(6), June 2002.

[78] F. Rubio and X. Mestre. Generalized Consistent Estimation on Low-rank Krylov Subspaces of Arbitrarily High Dimension. *IEEE Trans. Signal Process.*, 57(10), 2009.

[79] O. . Ryan and M. Debbah. Free deconvolution for signal processing applications. In *Proc. IEEE International Symposium on Information Theory (ISIT'07)*, pages 1846–1850, Nice, France, June 2007.

[80] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3) :276–280, 1986.

[81] H. Schultz. Non-commutative polynomials of independent Gaussian random matrices. The real and symplectic cases. *Probab. Theory Related Fields*, 131(2) :261–309, 2005.

[82] M. Shcherbina. Central limit theorem for linear eigenvalue statistics of the wigner and sample covariance random matrices. *arXiv preprint arXiv :1101.3249*, 2011.

[83] J. W. Silverstein and Z. D. Bai. On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *Journal of Multivariate Analysis*, 54(2) :175–192, 1995.

[84] J. W. Silverstein and S. Choi. Analysis of the limiting spectral distribution of large dimensional random matrices. *Journal of Multivariate Analysis*, 54(2) :295–309, 1995.

[85] J. W. Silverstein and P. L. Combettes. Signal detection via spectral theory of large dimensional random matrices. *IEEE Trans. Signal Process.*, 40(8) :2100–2105, 1992.

[86] T. Tao. *Topics in random matrix theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[87] I. E. Telatar. Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunications*, 10(6) :585–595, Feb. 1999.

[88] C. A. Tracy and H. Widom. On orthogonal and symplectic matrix ensembles. *Communications in Mathematical Physics*, 177(3) :727–754, 1996.

[89] D. N. C. Tse and S. V. Hanly. Multiaccess fading channels. I. Polymatroid structure, optimal resource allocation and throughput capacities. *IEEE Trans. Inf. Theory*, 44(7) :2796–2815, 1998.

[90] P. Vallet, P. Loubatin, and X. Mestre. Improved subspace doa estimation with large arrays : The deterministic signal case. *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'09)*, pages 19–24, 2009.

[91] P. Vallet and P. Loubaton. A G-estimator of the MIMO channel ergodic capacity. In *Proc. IEEE International Symposium on Information Theory (ISIT'09)*, pages 774–778, 2009.

[92] P. Vallet, P. Loubaton, and X. Mestre. Improved subspace DoA estimation methods with large arrays : The deterministic signals case. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'09)*, pages 2137–2140, 2009.

[93] P. Vallet, P. Loubaton, and X. Mestre. Improved subspace estimation for multivariate observations of high dimension : The deterministic signals case. *Information Theory, IEEE Transactions on*, 58(2) :1043 –1068, feb. 2012.

[94] S. Verdù and S. Shamai. Spectral efficiency of CDMA with random spreading. *IEEE Trans. Inf. Theory*, 45(2) :622–640, Feb. 1999.

[95] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Transactions on Signal, Speech and Signal Processing*, 33(2) :387–392, 1985.

[96] E. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *The annal of mathematics*, 62(3) :548–564, Nov. 1955.

[97] E. Wigner. On the distribution of roots of certain symmetric matrices. *The Annals of Mathematics*, 67(2) :325–327, Mar. 1958.

[98] J. Wishart. The generalized product moment distribution in samples from a normal multivariate population. *Biometrika*, 20(1-2) :32–52, Dec. 1928.

[99] J. Yao, R. Couillet, J. Najim, and M. Debbah. Fluctuations of an improved population eigenvalue estimator in sample covariance matrix models. *IEEE Trans. Inform. Theory*, 59(2) :1149–1163, 2013.

[100] J. Yao, R. Couillet, J. Najim, E. Moulines, and M. Debbah. Clt for eigen-inference methods in cognitive radios. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 2980 –2983, may 2011.

[101] J. Yao, A. Kammoun, and J. Najim. Eigenvalue estimation of parameterized covariance matrices of large dimensional data. *IEEE Trans. Signal Process.*, 60(11) :5893–5905, 2012.

[102] Y. Q. Yin, Z. D. Bai, and P. R. Krishnaiah. On the limit of the largest eigenvalue of the large dimensional sample covariance matrix. *Probability Theory and Related Fields*, 78(4) :509–521, 1988.