Fluctuations de fonctionnelles spectrales de grandes matrices aléatoires et applications aux communications numériques.

Malika Kharouf

THÈSE DE DOCTORAT DE

TÉLÉCOM PARISTECH ET UNIVÉRSITÉ HASSAN II AIN CHOCK

SPÉCIALITÉ

SIGNAL ET IMAGE

PRÉSENTÉE PAR

MALIKA KHAROUF

FLUCTUATIONS DE FONCTIONNELLES SPECTRALES DE GRANDES MATRICES ALÉATOIRES ET APPLICATIONS AUX COMMUNICATIONS NUMÉRIQUES

SOUTENUE LE 19 JUIN 2010 DEVANT LE JURY COMPOSÉ DE

ERIC MOULINES	PRESIDENT DU JURY
DRISS ABOUTAJDINE	RAPPORTEUR
DJALIL CHAFAI	RAPPORTEUR
AHMED ELKHARROUBI	DIRECTEUR DE THÈSE
WALID HACHEM	DIRECTEUR DE THÈSE
JAMAL NAJIM	DIRECTEUR DE THÈSE

Remerciements

C'est avec un grand plaisir que je dédie cette page à toutes les personnes qui ont contribuées de près ou de loin à la réussite de ma thèse.

Les premières personnes que je tiens à remercier sont Walid Hachem et Jamal Najim, mes directeurs de thèse. Je les remercie pour l'aide scientifique précieuse, pour leur disponibilité et pour leurs conseils avisés. Qu'ils trouvent ici mes profonds remerciements et reconnaissance.

Je suis également reconnaissante à Ahmed Elkharroubi d'avoir assuré mon encadrement au sein du département de Mathématiques et Informatique à l'Université Hassan II Ain Chock.

Je tiens à exprimer mes remerciements aux membres du jury, qui ont accépté d'évaluer mon travail de thèse.

Mes vifs remerciements vont à Eric Moulines d'avoir accépté de présider le jury de ma thèse. Je présente également toute ma reconnaissance aux professeurs Driss Aboutajdine, Djalil Chafai et Philippe Loubaton pour l'intérêt qu'ils ont porté à ma thèse et d'en être les rapporteurs.

J'ai eu l'occasion de collaborer avec Jack Silverstein, je le remercie pour les discussions très enrichissantes et je souhaite que l'achèvement de ce travail soit le debut d'une collaboration sans cesse grandissante. Je remercie également Abla Kammoun pour sa précieuse collaboration.

Mes remerciements vont également à tous les membres du laboratoire LTCI/ TSI à Télécom ParisTech ainsi qu'à tous les membres du département de Mathématiques et Informatique de l'Université Hassan II, faculté des sciences Ain Chock. Merci pour m'avoir accuelli tout au long de ces trois années de thèse.

Une dédicace spéciale à tous les gens que j'ai eu le plaisir de côtoyer durant ces quelques années de thèse aussi bien à Télécom ParisTech qu'à l'Université Hassan II Casablanca. Je les remercie pour les nombreuses discussions stimulantes, pour leur soutient moral et pour leur amitié.

Je tiens à remercier également tous mes collègues à l'Université Paris 10 Nanterre. Enfin, je tiens à remercier ma famille pour son assistance aussi bien matérielle que morale et pour sa patience avec moi tout au long de ces longues années d'études.

CONTENTS

CHAPTER 1

Introduction

La théorie des matrices aléatoires présente un ensemble d'outils mathématiques efficaces pour l'étude de performances des systèmes de communications numériques. L'objectif de cette thèse est de développer des résultats analytiques basés sur la théorie des matrices aléatoires pour étudier les fluctuations de quelques indicateurs de performance pour les systèmes de communications sans fil, en particulier, les sysèmes multi-antennes MIMO (pour Multiple Input Multiple Output) et les systèmes de codage des transmissions CDMA (pour Code Division Multiple Access). Plus précisemment, nous étudions les fluctuations du rapport signal sur bruit (SINR, pour Signal Interference plus Noise Ratio), indice de performance mesuré à la sortie d'un recepteur linéaire minimisant l'erreur quadratique des symboles estimés (LMMSE pour Linear Minimum Mean Squared Error) pour les transmissions par la technique CDMA. Le SINR pouvant s'écrire comme une valeur prise par une forme quadratique associée à une matrice aléatoire sur un vecteur aléatoire, son étude analytique fait donc appel à la théorie des matrices aléatoires.

La compréhension de la loi limite des fluctuations du SINR permet de comprendre le comportement d'autres indices de performance comme le taux d'erreur binaire et la probabilité de dépassement. Nous nous intéressons à l'étude de ces deux indices pour un modèle gaussien dont les corrélations mutuelles entre les émetteurs et les récepteurs sont prises en comptes.

La loi limite de l'information mutuelle d'un canal de Rice fait l'objet de notre travail présenté au chapitre 4) de cette thèse.

Ce chapitre introductif comprend trois parties. Dans la première, nous rappelons brièvement les principaux résultats du premier et du second ordre de fonctionnelles spectrales pour quelques modèles de matrices aléatoires. La deuxième partie est dédiée au contexte applicatif de nos travaux théoriques. Un bref aperçu des différents résultats et contributions développés tout au long de cette thèse sont présentés dans la troisième partie.

1 Comportement global du spectre de grandes matrices aléatoires

La théorie des grandes matrices aléatoires s'intéresse, entre autres, aux propriétés macroscopiques du spectre des matrices aléatoires, telles que le comportement asymptotique global du spectre, le comportement asymptotique des valeurs propres extrêmes, la loi jointe des valeurs propres, etc.

Cette théorie a connu un grand succès dans différentes branches de la physique théorique et des mathématiques. Une des raisons du succès de la théorie des matrices aléatoires est sa propriété d'universalité: le comportement asymptotique du spectre est indépendant de la distribution initiale des entrées de la matrice aléatoire en question.

Ce constat a été réalisé par Wigner en 1958 lorsqu'il a abordé l'étude spectrale des grandes matrices aléatoires pour résoudre des problèmes de la mécanique quantique. Wigner étudia le modèle dit du GUE (Gaussian Unitary Ensemble), et son théorème affirme que la limite du spectre des matrices GUE, quand la taille de la matrice, tends vers l'infini, est une loi déterministe (loi du demi-cercle).

Ce résultat a été étendu par plusieurs mathématiciens pour d'autres modèles de matrices aléatoires. Citons entre autres modèles, les matrices de Wigner à entrées indépendantes non identiquement distribuées, les matrices de Wishart, les matrices de Gram (cf [53]).

Le régime au bord du spectre corrobore ce constat d'universalité. En fait, Tracy et Widom ([77], 2002) ainsi que Soshnikov ([76], 1999) ont démontré, entre autres, que la convergence en distribution de la plus grande valeur propre d'une matrice de Wigner converge, en un certain sens, vers la loi de Tracy-Widom.

Ces propriétés ont fait, entre autres, de la théorie des matrices aléatoires, aux yeux des mathématiciens et des physiciens, un outil prometteur pour la résolution des problèmes théoriques aussi bien que pratiques.

1.1 Résultats du premier ordre: Lois des Grands Nombres

Considérons une matrice aléatoire Hermitienne $\mathbf{Y}_n$, de dimensions $n \times n$ et soient $\lambda_{n,1}, \ldots, \lambda_{n,n}$ ses n valeurs propres. Une des grandes questions de la théorie des matrices aléatoires est d'étudier le comportement asymptotique de la loi spectrale

$$\mu_n(d\lambda, w) = \frac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_{n,j}(w)}(d\lambda).$$

Après les travaux de Wigner et Dyson sur les matrices de Wigner et les travaux de Marcenko et Pastur sur les matrices de Gram, le comportement asymptotique de la mesure spectrale a été largement étudié pour une grande classe de modèles de grandes matrices aléatoires, citons, entre autres, [4, 11, 26, 31, 47] et récemment [6, 35], etc.

Ces travaux ont été faits selon deux stratégies:

Existence d'une loi limite:

La première stratégie consiste à établir l'existence d'une loi limite (déterministe) approximant la loi spectrale μ_n de la matrice aléatoire étudiée.

Marcenko et Pastur [52] ont étudié un modèle de matrice de covariance empirique, dans lequel la matrice $\mathbf{Y}_n$ est donnée par:

$$\mathbf{Y}_n \quad = \quad \mathbf{X}_n \mathbf{T}_n \mathbf{X}_n^* + \mathbf{A}_n, \tag{1.1}$$

où $\mathbf{T}_n$ ($N \times N$ diagonale), $\mathbf{X}_n$ ($N \times n$) et $\mathbf{A}_n$ ($N \times N$ hermitienne) sont indépendantes. Ils ont montré que, quand μ_{T_n}, la mesure spectrale de $\mathbf{T}_n$, converge vers une loi déterministe μ_T, et quand, μ_{A_n} converge vaguement vers une loi $\mathcal{A}$, alors, la loi spectrale μ_{Y_n} converge en probabilité vers une loi déterministe $\mathbb{P}_{MP}$ caratérisé par sa transformée de Stieltjes f_{MP} qui vérifie l'équation suivante:

$$f_{MP}(z) = m_{\mathcal{A}} \left(z - c \int \frac{\tau \mu_T(d\tau)}{1 + \tau f_{MP}(z)} \right), \quad z \in \mathbb{C}^+ = \{z \in \mathbb{C} | \Im(z) > 0\},$$

avec $m_{\mathcal{A}}$ la transformée de Stieltjes de la loi μ_{A_n} et c est tel que $\frac{N}{n} \xrightarrow[n\to\infty]{} c$. Ce résultat a été généralisé dans d'autres travaux ([4, 31, 41, 87, 89, 90]). Dans [4] par exemple, Bai et Silverstein ont montré une convergence presque sûre sous l'hypothèse que la matrice T_n soit diagonale et que sa loi limite soit atteinte par une convergence presque sûre.

Les modèles de matrices de Gram non centrés a suscité également beaucoup d'intérêt. Dozier et Silverstein ont étudié le modèle information-plus-bruit suivant:

$$\mathbf{\Sigma}_n \quad = \quad \mathbf{Y}_n + \mathbf{A}_n. \tag{1.2}$$

Les entrées de la matrice $\mathbf{Y}_n$ sont supposées independantes et identiquement distribuées (i.i.d.). La matrice $\mathbf{A}_n$ est independante de $\mathbf{Y}_n$ et telle que la distribution empirique de $\mathbf{A}_n\mathbf{A}_n^*$ converge vers une loi limite déterministe. Dans ce cas, la convergence presque sûre et en distribution de la mesure spectrale de la matrice $\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$ a été prouvée. Cette mesure est caractérisée par sa transformée de Stieltjes définie à partir d'une certaine équation fonctionnelle.
Dans la même direction, Hachem *et al.* [34] prouvent le même résultat dans le cas où la matrice $\mathbf{A}_n$ est déterministe pseudo-diagonale et la matrice $\mathbf{Y}_n$ est aléatoire dont les entrées sont indépendantes mais non identiquement distribuées (cas à profil de variance).

Existence d'approximants déterministes:

Dans le cas des matrices de Gram non centrées $\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$, avec,

$$\mathbf{\Sigma}_n = \mathbf{Y}_n + \mathbf{A}_n,$$

la convergence de la mesure spectrale de $\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$ n'est pas toujours garantie vu la difficulté de prouver l'existence d'une loi limite de $\mathbf{A}_n\mathbf{A}_n^*$ pour quelques modèles. Une approche alternative consiste à montrer l'existence d'une suite de mesures déterministes $(\pi_n)_n$ approximant la suite $\mu_{\Sigma_n\Sigma_n^*}$. Girko, à qui revient cette approche, a remarqué que la transformée de Stieltjes de la mesure spectrale $\mu_{\Sigma_n\Sigma_n^*}$ est égale à la trace normalisée de la matrice résolvante $(\mathbf{\Sigma}_n\mathbf{\Sigma}_n^* - zI_n)^{-1}$. L'idée consiste donc à montrer que les entrées de la fonction complexe matricielle résolvante $(\mathbf{\Sigma}_n\mathbf{\Sigma}_n^* - zI_n)^{-1}$ ont le même comportement asymptotique que les entrées d'une fonctionnelle matricielle complexe déterministe $\mathbf{T}_n(z)$. Cette fonction matricielle est caractérisée par un système de $(n + N)$ équations couplées et vérifie:

$$\frac{1}{n}\mathrm{Tr}\,(\mathbf{\Sigma}_n\mathbf{\Sigma}_n^* - zI_n)^{-1} - \frac{1}{n}\mathrm{Tr}\mathbf{T}_n(z) \xrightarrow[n\to\infty, n/N\to c\in(0,\infty)]{} 0$$

avec I_n la matrice identité de taille n. En remarquant que la trace normalisée de la matrice $\mathbf{T}_n(z)$ est une transformée de Stieltjes d'une mesure π_n, cela montre l'existence des mesures déterministes approximants la suite $\mu_{\Sigma_n\Sigma_n^*}$.

Girko [26] a montré ce résultat dans le cas où les entrées de $\mathbf{\Sigma}_n$ sont données par

$$\mathbf{\Sigma}_{ij}^n = \frac{\sigma_{ij}}{\sqrt{n}}\mathbf{X}_{ij}^n + \mathbf{A}_{ij}^n, \tag{1.3}$$

avec $(\sigma_{ij}^n)_{ij}$ une famille de réels, et les $\mathbf{X}_{ij}^n$ sont des variables al'eatoires i.i.d. Les lignes et les colonnes de la matrice $\mathbf{A}_n = (\mathbf{A}_{ij}^n)_{ij}$ sont supposées avoir des normes $\mathbb{L}_1$ finies.
Motivés par des applications aux communications numériques dans lesquelles cette condition sur les entrées de la matrice $\mathbf{A}_n$ n'est pas réalisable, Hachem *et al.* [35] ont montré l'existence de cette fonctionnelle matricielle $\mathbf{T}_n$ pour un modèle de matrice de Gram non centré avec profil de variance. La condition sur la matrice $\mathbf{A}_n$ suppose seulement la bornitude des normes euclidiennes des lignes et des colonnes.

Un cas particulier du modèle à profil de variance étudié dans [35] auquel nous nous sommes particulièrement intéressés est donné par

$$\mathbf{\Sigma}_n = \frac{1}{\sqrt{n}}\mathbf{D}_n\mathbf{X}_n\tilde{\mathbf{D}} + \mathbf{A}_n,$$

avec $\mathbf{D}_n$ et $\tilde{\mathbf{D}}_n$ sont diagonales réelles.
Dans ce cas, le système fondamental définissant la fonction matricielle $\mathbf{T}_n(z)$ se réduit à un système de deux équations donné par:

$$\begin{cases} \delta(z) = \frac{1}{n}\mathrm{Tr}\left(\mathbf{D}_n\left(-z(I_N + \mathbf{D}_n\tilde{\delta}) + \mathbf{A}_n(I_n + \tilde{\mathbf{D}}_n\delta)^{-1}\tilde{\mathbf{A}}_n^*\right)^{-1}\right), \\ \tilde{\delta}(z) = \frac{1}{n}\mathrm{Tr}\left(\tilde{\mathbf{D}}_n\left(-z(I_n + \tilde{\mathbf{D}}_n\delta) + \mathbf{A}_n^*(I_N + \mathbf{D}_n\tilde{\delta})^{-1}\mathbf{A}_n\right)^{-1}\right) \end{cases}$$

et la matrice $\mathbf{T}_n(z)$ est donnée par:

$$\mathbf{T}_n(z) = \left(-z(I_N + \mathbf{D}_n\tilde{\delta}) + \mathbf{A}_n(I_n + \tilde{\mathbf{D}}_n\delta)^{-1}\tilde{\mathbf{A}}_n^*\right)^{-1}.$$

1.2 Résultats de fluctuations: Théorème de la Limite Central (TLC)

Fluctuations de statistiques spectrales linéaires:

Un prolongement naturel de cette étape où est étudié des résultats décrivant le comportement asymptotique des mesures spectrales de grandes matrices aléatoires, est d'étudier les fluctuations autour de ces limites/ approximants déterministes.
La litterature mathématique sur l'étude des fluctuations de statistiques linéaires spectrales montre bien le caractère gaussien de ces fluctuations pour un grand nombre de modèles de matrices aléatoires.

Les résultats type TLC pour les fonctionnelles spectrales données par:

$$\chi_n(g) = \sum_{i=1}^{N} g(\lambda_{n,i}),$$

avec g fonction test définie sur $\mathbb{R}$, les $\lambda_{n,i}$ étant les valeurs propres de la matrice aléatoire étudiée, remontent à Arharov [2] qui a étudié les fluctuations des traces normalisées de puissances de matrices de covariance empiriques d'entrées i.i.d. Gaussiennes ($g(\lambda_{n,i}) = \lambda_{n,i}^{\nu}$, ν un entier). Ce travail a été généralisé par Jonsson [41] pour un modèle non Gaussien. Jonsson s'est basé sur la méthode des moments et sur un argument combinatoire. Dans ces deux travaux, la normalité asymptotique est prouvé sans fournir une expression explicite de la variance.

Basé sur la méthode de la transformée de Stieltjes et la technique des martingales (REFORM pour REsolvent, FORmula and Martingale), Girko [28] a fourni une expression explicite de la variance pour ce modèle de matrices de Gram pour la fonctionnelle: $g(\lambda) = (\lambda - z)^{-1}$, avec $\Im(z) \neq 0$.

Bai et Silverstein se sont intéressés dans [3], à l'étude des fluctuations du vecteur aléatoire $\left(\int g_1(\lambda)\mu_{Y_nY_n^*}(d\lambda), \ldots, \int g_k(\lambda)\mu_{Y_nY_n^*}(d\lambda)\right)$ pour des g_i appartiennent à une grande famille de fonctions analytiques et cela pour un modèle de matrice de Gram séparable à gauche, i.e.:

$$\mathbf{Y}_n = \frac{1}{\sqrt{n}}\mathbf{T}^{1/2}\mathbf{X}_n.$$

Sous l'hypothèse que le moment d'ordre 4 des entrées de la matrice X_n soit égale au moment 4 gaussien (soit 3 dans le cas réel et 2 dans le cas complexe), Bai et Silverstein ont prouvé que ce vecteur converge faiblement vers un vecteur Gaussien dont le vecteur des espérances et la matrices des covariances sont donnés explicitement. En plus d'un argument de tension de la suite $\left(\int g_1(\lambda)\mu_{Y_nY_n^*}(d\lambda), \ldots, \int g_k(\lambda)\mu_{Y_nY_n^*}(d\lambda)\right)_n$, la preuve se base essentiellement sur la méthode des martingales.

L'approche REFORM est utilisée pour étudier un cas unidimensionnel. Dans leur travail [37], Hachem *et al.* ont étudié pour un modèle de matrice de Gram avec profil de variance les fluctuations de $\chi_n(g)$ pour la fonctionnelle $g : \lambda \mapsto \log(\lambda + \rho)$, avec ρ un réel non-négatif. L'étude portait donc sur les fluctuations de la statistique spectrale suivante:

$$\mathcal{I}_n(\rho) \quad = \quad \int \log(\lambda + \rho)\mu_{Y_nY_n^*}(d\lambda) = \frac{1}{N}\log\det\left(\mathbf{Y}_n\mathbf{Y}_n^* + \rho I_N\right) \tag{1.4}$$

Cette statistique, très populaire dans la théorie de l'information, domaine motivant ce travail, représente l'information mutuelle entre le vecteur émis et le vecteur reçu dans le cadre des systèmes de transmission à entrées multiples et à sorties multiples. La matrice $\mathbf{Y}_n$, qui représente le canal de transmission, est supposée à profil de variance dont les entrées sont données par $Y_{ij} = \frac{\sigma_{ij}(n)}{\sqrt{n}}X_{ij}^n$, où, $(\sigma_{ij}(n); 1 \leq i \leq N, 1 \leq j \leq n)$ est une suite réelle, les X_{ij} étant i.i.d centrées.

Pour étudier les fluctuations de $\mathcal{I}_n(\rho)$ autour de son approximant déterministe $V_n(\rho)$ déjà fourni par les mêmes auteurs dans [35], l'approche consiste à étudier dans un premier temps les fluctuations de la variable $\mathcal{I}_n(\rho)$ autour de son espérance $\mathbb{E}\mathcal{I}_n(\rho)$. La variance fournie comporte un terme additif proportionnelle au quatrième cumulant des X_{ij}. La présence d'un terme proportionnel au quatrième cumulant de la variable X_{11} qui est dû à la non gaussianité de celle-ci, a été déjà mis en évidence par Khorunzhy, Khoruzhenko et Pastur [47] ainsi que Anderson et Zeitouni [1]. Dans un deuxième temps, l'étude asymptotique du biais entre $\mathbb{E}\mathcal{I}_n(\rho)$ et son équivalent déterministe $V_n(\rho)$ est menée.

Le cas gaussien a été traité dans [36] par des téchniques différentes, fondées sur la nature gaussiennes des entrées. Dans ce travail, le modèle de Gram-Kronecker suivant est étudié:

$$Y_n = \frac{1}{\sqrt{n}} D_n^{1/2} X_n \tilde{D}_n^{1/2},$$

avec D_n et $\tilde{D}_n$ sont deux matrices déterministes diagonales et les entrées de la matrice X_n sont i.i.d. gaussiennes. Dans ce travail, l'étude des fluctuations de la statistique $\mathcal{I}_n(\rho)$ donnée par (1.4) se base sur deux outils importants dans le cas gaussien: l'inégalité de Nash-Poincaré et la formule d'intégration par partie. L'expression de la variance dans ce cas rejoigne l'expression trouvée dans [37] avec un biais nul.

Basés sur la méthode des répliques, Moustakas *et al.* [55] ont montré que la distribution asymptotique de la fonctionelle $\mathcal{I}_n(\rho)$ est gaussienne dont les paramètres (espérance et variance) sont donnés explicitement. L'approche adoptée ici consiste à calculer les moments de la statistique en question en se basant sur la méthode des répliques et de montrer que, asymptotiquement, seuls le premier et le second moments ne sont pas négligeables, ce qui est caractèristique, en un certain sens, de la gaussianité.
Les résultats oubtenus, dans le cas gaussien, par la méthode des répliques se révèlent pertinents, mais l'incovenient majeur de cette méthode réside dans le fait que ses hyporhèses ne sont pas justifiées de façon mathématiquement rigoureuse.

Taricco [78] a généralisé ce résultat en utilisant également la méthode des répliques pour un modèle séparable non centré,

$$Y_n = \frac{1}{\sqrt{n}} D_n X_n \tilde{D}_n + A_n,$$

avec X_n à entrées i.i.d. gaussiennes et A_n une matrice déterministe.

L'étude de fluctuations pour des fonctionnelles spectrales des matrices de Wigner a également fait l'objet de plusieurs travaux, citons, entre autres, $[1, 14, 47, 74, 75]$. Pastur et Lytova [51] ont étudié les fluctuations de fonctionnelles spectrales linéaires pour des matrices de Wigner et des matrices de covariances empiriques. Le but de leur travail est de développer des outils permettant d'étendre la normalité asymptotique établie pour des modèles Gaussiens au cas des modèles non Gaussiens.

Fluctuations des formes quadratiques aléatoires:

L'étude des fluctuations des formes quadratiques aléatoires a suscité également beaucoup d'intérêt vu leur importance dans les applications (voir partie 2. pour des applications aux communications numériques). On s'intéresse aux formes quadratiques de type:

$$\beta(\rho) \quad = \quad y^* \left(\mathbf{Y}\mathbf{Y}^* + \rho I_N\right)^{-1} y, \tag{1.5}$$

où y et $\mathbf{Y}$ sont respectivement un vecteur et une matrice aléatoires indépendants et ρ un réel positif.

L'étude de ces formes quadratiques a été faite suivant différentes méthodes.

Dans [81], l'étude de cette forme quadratique repose sur un traitement direct à la fois des valeurs et des vecteurs propres de la matrice $\mathbf{Y}\mathbf{Y}^*$. Tse et Zeitouni ont étudié un modèle i.i.d centré pour la forme quadratique $\beta(\rho)$. Ce travail se base sur les travaux de

Silverstein ([66–70]) qui étudient le comportement asymptotique des vecteurs propres de quelques matrices de Gram. En faisant la décomposition spectrale de la résolvante de la matrice $\mathbf{YY}^*$, (1.5) devient:

$$\beta(\rho) = y^*\mathbf{O}\mathrm{diag}\left(\frac{1}{\lambda_i+\rho}\right)_{i=1}^{N}\mathbf{O}^*y$$

avec $\mathbf{O}$ la matrice unitaire des vecteurs propres de la matrice résolvante $Q(\rho) = (\mathbf{YY}^* + \rho I_N)^{-1}$, et les λ_i sont les valeurs propres de $\mathbf{YY}^*$.
Si on suppose que la matrice unitaire $\mathbf{O}$ est asymptotiquement Haar-distribuée, alors, le processus $Z_n(t)$ défini par,

$$Z_n(t) = \frac{1}{2n}\sum_{i=1}^{[nt]}(v_i^2 - \frac{1}{n})$$

avec $v_i = \mathbf{O}^*y$, vérifie la convergence suivante dans l'espace $\mathbf{D}[0,1]$ des fonctions continues à droite, ayant une limite à gauche:

$$(\mathbf{Z}_n(t))_{t\in[0,1]} \xrightarrow{\mathcal{D}} \left(\mathbf{W}_n^0 + \frac{\eta}{\sqrt{2}}t\right)$$

où $\mathbf{W}_n^0$ est un pont Brownien et η une variable normale.
En prenant comme élément de $\mathbf{D}[0,1]$ la fonction de répartition de la loi spectrale de $\mathbf{YY}^*$, on obtient la normalité asymptotique de la statistique $\beta(\rho)$, lorsqu'elle est centrée et bien normalisée.

Pan, Guo et Zhou [58] ont montré la normalité asymptotique de cette forme quadratique aléatoire pour un modèle séparable à droite. L'approche utilisée consiste à conditionner par rapport à la matrice $\mathbf{Y}$ et utiliser le résultat de l'article de Gotze et Tikhomirov [30] étudiant les fluctuations d'une forme quadratique aléatoire basée sur une matrice déterministe et déduisant la normalité asymptotique en montrant que la vitesse de convergence de la fonction de répartition de la variable forme quadratique vers celle de la loi normale est controlée par la norme de la plus petite valeur propre de la matrice $\mathbf{YY}^*$.

2 Matrices aléatoires et communications numériques sans fil

Systèmes à entrées multiples et à sorties multiples

Ces deux dernières décennies ont été témoins d'une renaissance dans la théorie de l'information de Shannon notamment pour les systèmes de communications sans fil. Dans une course pour l'amélioration des technologies de transmission de l'information, Foschini, des Bell Labs utilisa une technique permettant d'accroître les debits de transmission par l'emploi de plusieurs antennes à la fois à l'émission et à la réception (figure 1.1).

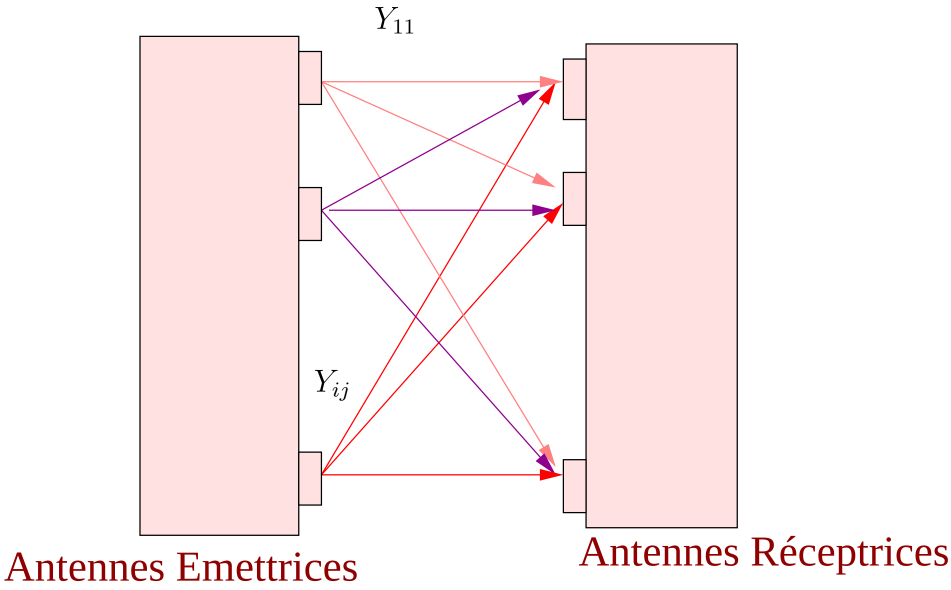


Figure 1.1: Représentation MIMO.

Gain en diversité

Cette technique de communication à entrées multiples et à sorties multiples appelée MIMO pour Multiple-Input Multiple-Output, consiste à transmettre plusieurs répliques du même signal à plusieurs récepteurs. Cela permet un gain matriciel dans le sens où chaque récepteur reçoit plusieurs copies du même signal envoyé par des transmetteurs différents et donc avec des atténuations différentes. Il est donc possible qu'au moins un des signaux reçus ne soit pas atténué, ce qui rend possible une transmission de bonne qualité.

Représentation matricielle d'un système MIMO. Si l'on considère, à un instant donné n, un signal reçu $\mathbf{r}_n$ et le signal émis $\mathbf{t}_n$, le système MIMO à N émetteurs et n récepteurs peut être décrit par le système linéaire suivant:

$$\mathbf{r}_n = \mathbf{Y}_n \mathbf{t}_n + \mathbf{b}_n \tag{1.6}$$

où $\mathbf{Y}_n$ est la $N \times n$ matrice modélisant le canal de transmission, dont les entrées représentent les gains entre les antennes de transmission et les antennes de réception, et le N-dimensionnel vecteur $\mathbf{b}_n$ correspond au bruit pertubant le signal émis.

Les multiples versions du signal qui peuvent provoquer des interférences constructives et destructives entres elles sont, entres autres, des causes d'altération du canal de transmission. Il est donc légitime de supposer que, à chaque utilisation du canal, la matrice $\mathbf{Y}_n$ est une réalisation d'une matrice aléatoire. Cela dit, le système (1.6) sera donc caractérisé par la distribution de la matrice aléatoire dont les réalisations $\mathbf{Y}_n$ traduisent les caractéristiques du canal. Le canal de tansmission peut connaître deux scénarios:

Régime d'évanouissement rapide: Fast fading environment.

Ce cas de figure se présente quand la réponse du canal change rapidement durant la période de transmission. Cet évanouissement est dû, par exemple, aux réflexions du signal à des objets proches. Dans ce cas, chaque transmission correspond à une nouvelle réalisation du canal.

Régime d'évanouissement lent: Slow fading environment.

L'évanouissement lent d'un canal est dû aux phénomènes de masquages et d'ombrage qui peuvent se présenter entre l'émetteur et le récepteur. Dans ce cas, le canal peut être considéré comme constant pendant la période d'utilisation.

L'évaluation des performances des canaux MIMO se fait à travers l'étude de ses indices de performances tels que la capacité du canal de transmission, la probabilité de dépassement d'un seuil donné pour l'information mutuelle, le taux d'erreur en sortie d'un récepteur, le rapport signal sur bruit..
L'étude mathématique de ces indicateurs tire profil du fait que la plupart de ces indicateurs s'expriment comme des fonctionnelles spectrales de la matrice-canal.

2.1 Information mutuelle dans un système multi-antennes

Dans un système multi-antennes, l'information mutuelle entre le signal transmis $\mathbf{t}_n$ et le signal reçu $\mathbf{r}_n$ est donnée par:

$$\begin{aligned} \frac{1}{N} I(\mathbf{t}_n, \mathbf{r}_n | \mathbf{Y}_n) &= \frac{1}{N} \log\det\left(\mathbf{I}_N + \rho \mathbf{Y}_n \mathbf{Y}_n^*\right) && (1.7) \\ &= \int_0^{\infty} \log(1+\rho z) d\mu_{\mathbf{Y}_n \mathbf{Y}_n^*}(z) && (1.8) \end{aligned}$$

avec $\mu_{Y_n Y_n^*}$ la mesure spectrale des valeurs prores de $\mathbf{Y}_n \mathbf{Y}_n^*$ et ρ représentant le rapport signal sur bruit donné par:

$$\rho = \frac{N\mathbb{E}||\mathbf{t}_n||^2}{K\mathbb{E}||\mathbf{b}_n||^2}$$

Capacité du canal dans un régime d'évanouissement rapide

Dans le cas d'évanouissement rapide, les changements du canal se traduisent par des réalisations indépendantes de la matrice aléatoire modélisant le canal. On s'intéresse donc à l'information mutuelle "moyenne" entre $\mathbf{r}_n$ et $\mathbf{t}_n$. Dans ce cas, à chaque unité de temps de transmission n on a une réalisation du canal $\mathbf{Y}_n$, et l'information mutuelle entre $\mathbf{r}_n$ et vecteur transmis $\mathbf{t}_n$ de matrice de covariance $\mathbf{Q}_n$ sera donc donnée par:

$$\mathbb{E} \log\det\left(\mathbf{I}_N + \rho\mathbf{Y}_n\mathbf{Q}_n\mathbf{Y}_n^*\right)$$

où ρ représente la variance du bruit $\mathbf{b}_n$.

Dans ce cas, la capacité ergodique du canal qui représente le maximum de l'information pouvant transiter à travers le canal est déterminée comme le maximum de l'information mutuelle entre $\mathbf{t}_n$ et $\mathbf{r}_n$, sous certaines contraintes sur la matrice de covariance de $\mathbf{t}_n$. Plus précisemment, la capacité ergodique est donnée par:

$$\sup_{Q\geq 0, \frac{1}{n}\mathrm{Tr}Q\leq 1} \mathbb{E} \log\det\left(\mathbf{I}_N + \rho\mathbf{Y}_n\mathbf{Q}_n\mathbf{Y}_n^*\right).$$

où $\mathrm{Tr}Q$ est la trace de la matrice $\mathbf{Q}$.

Capacité du canal sous un régime d'évanouissement lent

Dans un régime d'évanouissement lent, comme la réalisation du canal peut persister pendant que plusieurs messages peuvent être transmis, l'information mutuelle sera donnée par:

$$\log\det\left(\mathbf{I}_N + \rho\mathbf{Y}_n\mathbf{Q}_n\mathbf{Y}_n^*\right).$$

Dans ce cas, un indice pouvant mesurer la pertinence du choix de la réalisation du canal est la probabilité de dépassement (outage probability). Cet indice correspond à la probabilité que la capacité instantanée du canal de transmission soit inférieure ou égale au nombre de bits transmis par utilisation du canal (rendement de la transmission).

Analyse mathématique de l'information mutuelle

Telatar [80] et Foschini [22] sont les premiers qui ont considéré le scenario d'une matrice-canal $\mathbf{Y}_n$ aléatoire. Dans leurs travaux, le modèle étudié est celui de Rayleigh: une matrice-canal d'entrées qui s'écrivent sous la forme $Y_{n,k} = \alpha_{n,k} exp(j\theta_{n,k})$, où $(\alpha_{n,k})_{n,k}$ est une suite de variables aléatoires independantes (v.a.i.) suivant une loi de Rayleigh, et $(\theta_{n,k})_{n,k}$ est une suite de v.a.i. distribuées selon une loi uniforme, autrement, les entrées sont complexes i.i.d gaussiennes dont la partie réelle et la partie imaginaire de chaque entrée sont indépendantes centrées et de variance 1/2. Dans ce cas, il est possible de trouver une expression explicite de l'information mutuelle $I(\mathbf{t}_n, \mathbf{r}_n|\mathbf{Y}_n)$ mais cette expression reste tout de même peu exploitable vu la difficulté de pouvoir en tirer des informations sur l'influence des paramètres du canal sur sa performance. Telatar [80] a donc procédé à la recherche des équivalents de cette statistique en espérant qu'ils soient plus facile à calculer et à interpréter.

En faisant l'hypothèse que les trajets entre chaque antenne d'émission et de réception sont indépendants, Telatar a prouvé que la capacité théorique du canal MIMO croit linéairement avec le minimum des nombres d'antennes à l'émission ou à la reception. la figure 1.2 confirme le fait que la capacité augmente avec le nombre d'antennes pour des SNR croissants.

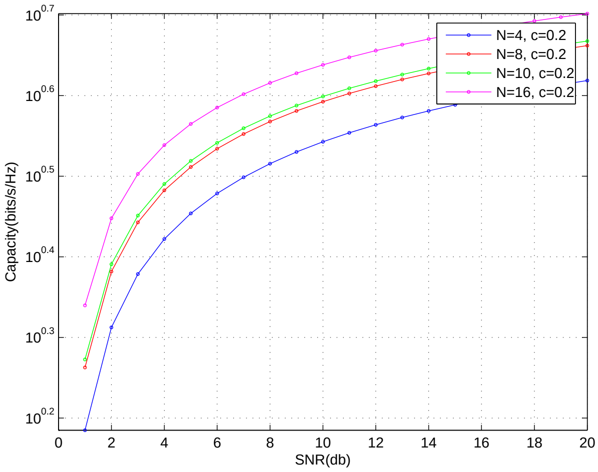


Figure 1.2: Capacité en fonction du SNR.

En effet, en se basant sur le fameux résultat de Marcenko-Pastur [52] sur le comportement asymptotique de la mesure spectrale des matrices de Gram, Telatar a montré que l'information mutuelle converge, quand le nombre d'antennes d'émission et de réception tendent vers l'infini au même rythme ($\frac{N}{n} \to c > 0$), vers une quantité déterministe $V(\rho, c)$ qui ne dépends que du rapport signal sur bruit ρ et de la constante c. La dépendance de cet approximant des paramètres du canal est beaucoup plus explicite et son implémentation informatique est moins coûteuse.

Les résultats très encourageants de ces deux travaux ont motivé le développement de plusieurs travaux applicant la théorie des matrices aléatoires pour la résolution des problèmes de la théorie de l'information.

Dans la littérature, on trouve deux grands scénarios pour le canal de transmission:

2.2 Modèles de Rayleigh

Les modèles de Rayleigh modélisent les canaux de transmission lorsque celle ci se rend au récepteur en passant par des réflexions et des échos. D'un point de vue mathématique, on parle des modèles centrés dans lesquels les entrées de la matrice-canal sont modélisées par des variables aléatoires centrées.

Modèles de Kronecker centrés

Ce sont les modèles centrés qui supposent des corrélations à l'émission et/ou à la réception entre les antennes tout en supposant la non corrélations entre les antennes émettrices et celles réceptrices. Ces modèles sont donnés par:

$$\mathbf{Y} = \mathbf{D}^{1/2}\mathbf{X}\tilde{\mathbf{D}}^{1/2}$$

Dans le but de comprendre l'impact de ces corrélations sur l'information mutuelle, ce modèle a fait l'objet de plusieurs travaux , citons à titre d'exemple [16, 36, 55, 56, 84].

Ces travaux établissent une approximation $V_n(\rho, c)$ de l'information mutuelle $\mathcal{I}_n(\rho) = \log\det(I_N + \rho Y_n Y_n^*)$ dans le régime asymptotique où le nombre d'antennes à l'émission et le nombre d'antennes à la réception tendent vers l'infini au même rythme.

Considérons par exemple le travail de Hachem *et al.* [36]. Le modèle de Kronecker centré Gaussien, où les matrices de corrélations $\mathbf{D}_n$ et $\tilde{\mathbf{D}}_n$ sont supposées diagonales et les entrés de $\mathbf{X}$ sont gaussiennes est étudié dans ce travail. Il a été prouvé que l'information mutuelle $\mathcal{I}_n(\rho)$ admet un équivalent déterministe $\mathbf{V}_n$ dont l'expression est donnée par

$$\mathbf{V}_n(\rho, c) = \log\det\left(\mathbf{I}_n + \rho\delta_n(\rho)\tilde{\mathbf{D}}_n\right) + \log\det\left(\mathbf{I}_N + \rho\tilde{\delta}_n(\rho)\mathbf{D}_n\right) - n\rho\delta_n(\rho)\tilde{\delta}_n(\rho)$$

avec $(\delta_n(\rho), \tilde{\delta}_n(\rho))$ solution du système (1.1). Il a été prouvé également que la vitesse de convergence de $\mathbb{E}\mathcal{I}_n(\rho)$ vers cet approximant est inversement proportionnel au nombre de récepteurs/ émetteurs.

Dans [16], Chuah et *al.* étudient le comportement asymptotique de l'information mutuelle et de la capacité des systm̀es multi-antennes (MEA: Multiple-Element Arrays). Deux cas ont été étudiés: le cas i.i.d sans corrélations et le cas gaussien avec corrélations.

L'étude des fluctuations de l'information mutuelle autour de son approximant déterministe présente un grand intérêt pratique dans le sens où, en plus de son rôle classique de mesurer la pertinence de l'approximant déterministe, elle permet aussi de déterminer la probabilité de dépassement (outage probability) correspondante à la probabilité que la capacité instantanée du canal de transmission soit inférieure ou égale au nombre de bits transmis par utilisation canal (rendement de transmission).

Moustakas *et al.* [55] ont étudié les fluctuations de la variable information mutuelle pour un canal de Rayleigh dans le cas de présence de corrélations entre les antennes. Dans un scénario d'évanouissement lent, les auteurs ont montré la normalité asymptotique des fluctuations de l'information mutuelle en se basant sur la méthode des répliques, et ont montré, par simulations, que le régime asymptotique peut être atteint pour un nombre réaliste d'antennes.

Modèles centrés généraux

Dans ce cas, on permet des corrélations entre les antennes des deux côtés en plus des corrélations mutuelles dans chaque côté, on a donc à traiter des modèles de type:

$$\mathbf{Y}_{ij}^n = \frac{\sigma_{ij}(n)}{\sqrt{n}}\mathbf{X}_{ij}^n$$

Pour ces modèles, citons [3, 37]. Dans ces travaux, le comportement asymptotique gaussien de l'information mutuelle est prouvé pour des modèles généraux pas forcément gaussiens.

2.3 Modèles de Rice

Lorsque l'on suppose l'existence d'un trajet direct entre le transmetteur et le récepteur, on dit qu'on parle de ligne de vue (L.O.S pour Line Of Sight). On désigne ainsi la possibilté de voir directement le récepteur à partir du transmetteur. Lorsqu'une ligne de vue existe, on parle généralement des modèles de Rice, ou encore des modèles Information-plus-Bruit. Dans ce cas, on rajoute un terme déterministe décrivant le gain dû au trajet direct. Ces modèles sont donnés sous la forme:

$$\mathbf{\Sigma}_n = \mathbf{Y}_n + \mathbf{A}_n.$$

Ce modèle a fait l'objet de plusieurs travaux. Citons, entre autres, [17, 37, 50], et dans le cas gaussien, citons les travaux de Hachem et *al* [36], Taricco [78, 79], etc.

L'information mutuelle ainsi que la capacité du canal de transmission ont fait l'objet de plusieurs études. Les résultats mathématiques présentés dans la première section étudiant la fonctionnelle $\log\det$ montrent que, pour la plupart de ces modèles, la distribution de l'information mutuelle peut être caractérisée par la donnée des approximants de ses paramètres, qui sont relativement simples à calculer et à interpréter par rapport aux vrais paramètres.

2.4 Rapport Signal à Interférence plus Bruit

Une des techniques multi-antennes permettant d'augmenter la capacité des systèmes de communications est l'accès multiple par division de codes CDMA (Code Division Multiple Access). La technique CDMA est un mode d'accès multiple dans lequel chaque usager est caractérisé par une séquence codée permettant de restituer le signal qu'il a émis ou celui qui lui est destiné. Le système (1.6) modélise cette technique en supposant que $\mathbf{t}_n$ représente le vecteur des codes transmis.

Le but dans un système CDMA est de pouvoir estimer les symboles transmis à partir du vecteur reçu. Un des estimateurs les plus populaires est l'estimateur linéaire de Wiener, ou encore l'estimateur LMMSE, pour Linear Minimum Mean Squared Error. Supposons qu'on cherche à estimer le premier symbol t_1. L'estimateur LMMSE estime $\bar{t}_1 = \mathbf{g}^*\mathbf{r}$ tel que le vecteur $\mathbf{g}$ est le vecteur minimisant la quantité: $\mathbb{E}|\mathbf{g}^*\mathbf{r} - t_1|^2$. La performance de l'estimateur LMMSE est souvent évaluée en terme du Rapport Signal à Interference-Plus-Bruit (RSIB) mesuré à la sortie du récepteur. Le RSIB est donné par:

$$RSIB = \frac{|\mathbf{g}^*\mathbf{y}|^2}{\mathbb{E}|\mathbf{g}^*\mathbf{r}_{in}|^2} \tag{1.9}$$

avec $\mathbf{y}$ représente la colonne correspondante à l'utilisateur d'intérêt (le premier dans ce cas) et $\mathbf{r}_{in}$ représente le signal inutile (interférence plus bruit).

Après quelques manipulations matricielles, il est possible de montrer que le RSIB peut s'écrire sous la forme :

$$RSIB = \mathbf{y}^* \left(\mathbf{Y_1}\mathbf{Y_1}^* + \rho I_N\right)^{-1} \mathbf{y}$$

avec $\mathbf{Y_1}$ la matrice résultante en éliminant la première colonne de la matrice $\mathbf{Y}$ et ρ est tel que $\mathbb{E}\mathbf{b}\mathbf{b}^* = \rho I_N$.

Cette expression étant peu exploitable en raison de la complexité d'inverser des matrices de taille $N \times N$ ainsi que la difficulté d'y voir l'influence des paramètres du modèle sur le

RSIB, il est donc important de chercher des approximants relativement simple à interpréter et à calculer.

L'étude mathématique du RSIB s'inscrit dans le cadre de l'étude des formes quadratiques aléatoires.

2.5 Taux d'Erreur et Probabilité de Dépassement

La performance de l'indice Rapport Signal à Interférence plus Bruit est étudiée à partir de la probabilité de dépassement qui mesure le rendement de la transmission. Un autre indice pour mesurer la qualité de la transmission de même importance que la capacité est le taux d'erreur ou BER (pour Bit Error Rate). C'est une mesure qui s'appuie sur le ratio de bits faux. La distribution du RSIB peut être utiliser pour prédire le BER dans un canal de transmission. En effet, si on considère un recepteur LMMSE, il est prové que pour certains types de modulations, le BER peut être caractérisé par la distribution du RSIB, β. Soit,

$$BER = \frac{1}{2\pi}\mathbb{E}\int_{\sqrt{\beta}}^{\infty} e^{-t^2/2}dt$$

où l'espérance ici est prise selon la loi de β.
Li *et al.* [49] ont proposé d'approximer le BER en supposant que le RSIB suit une loi Gamma généralisée de paramètres (α, b, ξ). L'idée est de trouver des approximants des paramètres de la loi Gamma généralisée et donc d'exprimer le BER en fonction de la fonction génératrice des moments du RSIB. Soit,

$$BER = \frac{1}{\pi}\int_0^{\frac{1}{\pi}} \mathbf{M}\left(-\frac{1}{2sin^2\phi}\right) d\phi.$$

3 Contributions de la thèse

3.1 Etude des fluctuations des formes quadratiques aléatoires

Considérons la forme quadratique aléatoire suivante:

$$\beta_K(\rho) \quad = \quad \mathbf{y}^* \left(\mathbf{Y}_K\mathbf{Y}_K^* + \rho\mathbf{I}_K\right)\mathbf{y} \tag{1.10}$$

avec, $\mathbf{Y}_K$ est une matrice aléatoire de dimension $N \times K$ dont les entrées sont données par $Y_{nk} = \frac{\sigma_{nk}}{\sqrt{n}}X_{nk}$, où $(\sigma_{nk})_{nk}$ est une suite de réels positifs et X_{nk} des variables aléatoires i.i.d. centrées. Le vecteur aléatoire $\mathbf{y}$ est indépendant de la matrice $\mathbf{Y}$.

Nous avons établi un TLC pour les formes quadratiques de la forme (1.10), dont nous avons fourni des expressions explicites de la moyenne et de la variance. Notre approche consiste à montrer la normalité asymptotique en se basant sur la méthode REFORM. Nous avons mis en évidence que les fluctuations de cette forme quadratique proviennent essentiellement du caractère aléatoire du vecteur $\mathbf{y}$ et donc le caractère aléatoire des vecteurs propres de la matrice de Gram $\mathbf{YY}^*$ n'intervient pas dans les fluctuations de la forme quadratique.
Notons que notre résultat est d'ordre non-asymptotique, i.e. l'expression de la variance

dépend de n. Nous pensons que l'intérêt de ce résultat est double: d'une part, nous supposons peu d'hypothèses sur la suite profils de variance, ce qui rend le résultat plus général, et d'autre part, la donnée des suites dépendant de n jouant le rôle de la variance dans le TLC facilite son implémentation informatique et son calcul pratique.
Ce travail présenté dans le chapitre (2) fait l'objet de l'article [45] publié dans IEEE Information Theory.

3.2 Contribution analytique et numérique pour le taux d'erreur et la probabilité de dépassement

La compréhension des fluctuations du SINR permet d'étudier le comportement d'autres indices de performances comme le taux d'erreur (BER pour Bit Error Rate) et la probabilité de dépassement (outage probability).

IL a été prouvé que les fluctuations du SINR sont asymptotiquement gaussiennes. Cependant, cette approximation n'est pas efficace pour l'étude du BER et de la probabilité de dépassement vu que la loi normale permet des valeurs négatives. Li et *al.* [49] proposent une approximation par la loi Gamma généralisée. La loi Gamma généralisée étant positive et admet un moment d'ordre trois non nul, ce qui est important pour le calcul du BER et de l' *outage probability.* Cette approximation se révèle pertinente même pour les systèmes de faibles dimensions.

Dans notre travail présenté dans le chapitre (3), nous adoptons cette approche pour un modèle gaussien séparable. Nous calculons les trois premiers moments du SINR en utilisant des techniques basées sur la nature gaussienne des entrées de la matrice-canal.

Ce travail a fait l'objet de l'article [45] publié dans IEEE Information Theory.

3.3 Etude des fluctuations de la fonctionnelle spectrale logdet

Le TLC pour les fonctionnelles spectrales linéaires a fait l'objet de plusieurs travaux. Notre contribution présentée dans le chapitre (4), consiste à étabir un TLC pour la fonctionnelle

$$\mathcal{I}_n(\rho) \quad = \quad \frac{1}{N}\sum_{i=1}^{N}\log\left(\lambda_i^n+\rho\right), \tag{1.11}$$

avec λ_i^n sont les valeurs propres d'une matrice de Gram $\mathbf{\Sigma}_n\mathbf{\Sigma}_n^*$. Nous considérons le modèle matriciel non centré suivant:

$$\mathbf{\Sigma}_n \quad = \quad \frac{1}{\sqrt{n}}\mathbf{D}_n^{1/2}\mathbf{X}_n\tilde{\mathbf{D}}_n^{1/2}+\mathbf{A}_n.$$

Les matrices $\mathbf{D}_n$ et $\tilde{\mathbf{D}}_n$ sont positives, diagonales. La matrice $\mathbf{X}_n$ est complexe dont les entrées sont i.i.d. centrées et réduites, et la matrice $\mathbf{A}_n$ est déterministe. Nous montrons que la statistique $\mathcal{I}_n$, lorsqu'elle est centrée et bien normalisée, vérifie un TLC. Pour cette fin, nous adoptons l'approche REFORM, nous développons des outils mathématiques pour vérifier les conditions du TLC pour les martingales (cf, Billingsley [9]). Nous montrons que la variance dépend à la fois des valeurs et des vecteurs propres de la matrice de centrage $\mathbf{A}_n$.
L'intérêt applicatif de ce travail est l'étude des fluctuations de l'information mutuelle dans le cas des modèles de Rice.

Afin détudier les fluctuations de la variable information mutuelle autour de son approximant déterministe $\mathbf{V}_n$, l'étude du comportement asymptotique du biais qui apparait naturellement en remplaçant $\mathbb{E}\mathcal{I}_n(\rho)$ par $\mathbf{V}_n$ est nécessaire. Ce point n'est pas abordé dans cette thèse, est en cours de réalisation.

Ce travail étudiant les fluctuations de l'information mutuelle dans le cas Rice, est en cours de finalisation pour lequel l'essentiel des résultats mathématiques a été établi (cf. chapitre 4), et donnera lieu à la rédaction d'un article avec W. Hachem, J. Najim et J. W. Silverstein.

4 Liste de publications

Publications dans des revues internationales à comité de lecture.

1. "A Central Limit Theorem for the SINR at the LMMSE estimator output for large dimensional signals". IEEE Inf. Theory, Vol. 55(11), nov. 2009. Avec A. Kammoun, W. Hachem et J. Najim.

2. "BER and Outage Probability approximations for LMMSE detectors on correlated MIMO channels". IEEE Inf. Theory, Vol. 55(10), oct. 2009. Avec A. Kammoun, W. Hachem et J. Najim.

Publications dans des actes de conférences internationales.

3. "Outage probability approximation for the Wiener Filter SINR in MIMO systems". IEEE Workshop on Signal Processing Advances in Wireless Communications SPAWC 2008. Avec A. Kammoun, W. Hachem et J. Najim.

4. "Fluctuations of the SNR at the Wiener Filter Output for Large Dimensional Signals". IEEE Workshop on Signal Processing Advances in Wireless Communications SPAWC 2008. Avec A. Kammoun, W. Hachem et J. Najim.

5. "On the Fluctuations of the Mutual Information for Non Centered MIMO Channels: The Non Gaussian Case". IEEE Workshop on Signal Processing Advances in Wireless Communications SPAWC 2010. Avec A. Kammoun, W. Hachem, J. Najim et A. Elkharroubi.

Article en cours de préparation

6. "A CLT for Information-Theoretic Statistics of non-Centred Gram Random Matrices", avec W. Hachem, J. Najim et J. W. Silverstein.

CHAPTER 2

Central Limit Theorem for quadratic forms

Contents

The material of this chapter is the article entitled "A Central Limit Theorem for the SINR LMMSE Estimator Output for Large Dimensional Signals" [45] published in IEEE Information Theory revue.

1 Introduction

The most of the schemes of multi-user and multi-access communication systems such as Multiple Input Multiple Output (MIMO) systems and Code Division Multiple Access (CDMA) are modeled as a linear random system:

$$\mathbf{r} = \mathbf{\Sigma s} + \mathbf{n} \tag{2.1}$$

the N dimensional random vector $\mathbf{r} \in \mathbb{C}^N$ represents the received signal, the $K+1$ multi-dimensional transmitted signal is given by a random vector $\mathbf{s}$ and satisfying $\mathbb{E}\mathbf{ss}^* = \mathbf{I}_{K+1}$, $\mathbf{\Sigma}$ is the channel matrix and $\mathbf{n}$ is an independent Additive White Gaussian Noise (AWGN) with covariance matrix $\mathbb{E}\mathbf{nn}^* = \rho \mathbf{I}_N$ whose variance $\rho > 0$ is known.
The choice of the random character of the channel matrix is justified by the random fluctuating nature of the channel transmission. The theory of large random matrices is a powerful mathematical tool widely used to adress problems in multidimensional wireless communications and signal processing. This chapter examines the mathematical properties of a quantity fundamental in analyzing the performance of the Linear Minimum Mean

Squared Error (LMMSE) estimator for multidimensional signals in the large dimension regime. The output Signal to Interference-Noise Ratio (SINR) associated with a given user k is typically used as a measure for evaluating the performance of the LMMSE estimator. Without loss of generality, we suppose that we are interested by the first user $k = 0$. For user 1 and under the LMMSE estimator, The transmitted signal s_0 is estimated by $\hat{s}_0 = \mathbf{g}^*\mathbf{r}$ where the $N \times 1$ vector $\mathbf{g}$ minimizes the quadratic error $\mathbb{E}|\hat{s}_0 - s_0|^2$ and maximizes the 1's output SINR β_K given by:

$$\beta_K = \mathbf{y}^* \left(\mathbf{Y}\mathbf{Y}^* + \rho \mathbf{I}_N\right)^{-1} \mathbf{y} \ . \tag{2.2}$$

where the $N \times 1$ vector $\mathbf{y}$ and the $N \times K$ matrix $\mathbf{Y}$ derived from the following decomposition of the channel matrix: $\mathbf{\Sigma} = [\mathbf{y}\ \mathbf{Y}]$. Large random matrix theory shows that, when the dimension of the received and transmitted signals go to infinity with the same rate, the SINR β_K converges, in some sense, to an explicit deterministic quantity $\overline{\beta}_K$. Beyond the convergence of the SINR, a natural practical and theoretical problem concerns the study of the distribution of its fluctuations.

In this chapter, we consider the following statistical model:

$$\mathbf{\Sigma} = \left(\Sigma_{nk}\right)_{n=1,k=0}^{N,K} = \left(\frac{\sigma_{nk}}{\sqrt{K}} W_{nk}\right)_{n=1,k=0}^{N,K} \tag{2.3}$$

where the complex random variables W_{nk} are i.i.d. with $\mathbb{E}W_{nk} = 0$, $\mathbb{E}W_{nk}^2 = 0$ and $\mathbb{E}|W_{nk}|^2 = 1$ and where $(\sigma_{nk}^2;\ 1 \le n \le N;\ 0 \le k \le K)$ is an array of real numbers. Due to the fact that $\mathbb{E}|\Sigma_{nk}|^2 = \frac{\sigma_{nk}^2}{K}$, the array (σ_{nk}^2) is referred to as a variance profile.

The literature. The asymptotic first order results of quadratique forms described by the model (2.3) have been studied in various works (see, e.g. [4, 26]). Applications in the field of wireless communications can be found in e.g. [15] in the separable case, and in [83] in the general variance profile case.

Concerning the CLT for $\beta_K - \overline{\beta}_K$, only some particular cases of the general model (2.3) have been considered in the literature among which the i.i.d. case is studied in [81] which is based in the result of [70] pertaining to the asymptotic behavior of the eigenvectors of $\mathbf{Y}\mathbf{Y}^*$.

In this chapter, we establish a Central Limit Theorem for a large class of random matrices $\mathbf{\Sigma}$. We prove that there exists a sequence $\theta_K^2 = \mathcal{O}(1)$ such that $\frac{\sqrt{K}}{\theta_K}(\beta_K - \overline{\beta}_K)$ converges in distribution to the standard normal law $\mathcal{N}(0,1)$ in the asymptotic regime.

In section 2, we recall some results concerning the asymptotic behavior of quadratic forms. Our mean contribution, the CLT is given in section 3. In section 4, we provide some applications in the field of wireless communications and numerical illustrations.

2 First Order Results: Deterministic Approximations of Random Quadratic Forms

The aim of this section is to give a short outline of the existing first ordre results. These results are given in both cases, the general case (with a general variance profile) and the separable case which is of great importance in the applications.

The model Consider the quadratic form (2.2):

$$\beta_K = \mathbf{y}^* \left(\mathbf{Y}\mathbf{Y}^* + \rho \mathbf{I}_N \right)^{-1} \mathbf{y}$$

where the sequence of matrices $\mathbf{\Sigma}(K) = [\mathbf{y}(K)\ \mathbf{Y}(K)]$ is given by

$$\mathbf{\Sigma}(K) \;=\; (\Sigma_{nk}(K))_{n=1,k=0}^{N,K} \;=\; \left(\frac{\sigma_{nk}(K)}{\sqrt{K}} W_{nk} \right)_{n=1,k=0}^{N,K} .$$

Let us state the main assumptions:

A1 The complex random variables $(W_{nk};\ n \geq 1,\ k \geq 0)$ are i.i.d. with $\mathbb{E}W_{10} = 0$, $\mathbb{E}W_{10}^2 = 0$, $\mathbb{E}|W_{10}|^2 = 1$ and $\mathbb{E}|W_{10}|^8 < \infty$.

A2 There exists a real number $\sigma_{\max} < \infty$ such that

$$\sup_{K \geq 1} \max_{\substack{1 \leq n \leq N \\ 0 \leq k \leq K}} |\sigma_{nk}(K)| \leq \sigma_{\max} .$$

Let $(a_m; 1 \leq m \leq M)$ be complex numbers, then $\mathrm{diag}(a_m; 1 \leq m \leq M)$ refers to the $M \times M$ diagonal matrix whose diagonal elements are the a_m's. If $\mathbf{A} = (a_{ij})$ is a square matrix, then $\mathrm{diag}(\mathbf{A})$ refers to the matrix $\mathrm{diag}(a_{ii})$. Consider the following diagonal matrices based on the variance profile along the columns and the rows of $\mathbf{\Sigma}$:

$$\begin{array}{rcll} \mathbf{D}_k(K) & = & \mathrm{diag}(\sigma^2_{1k}(K), \cdots, \sigma^2_{Nk}(K)), & 0 \leq k \leq K \\ \widetilde{\mathbf{D}}_n(K) & = & \mathrm{diag}(\sigma^2_{n1}(K), \cdots, \sigma^2_{nK}(K)), & 1 \leq n \leq N. \end{array} \tag{2.4}$$

A3 The variance profile satisfies

$$\liminf_{K \geq 1} \min_{0 \leq k \leq K} \frac{1}{K} \mathrm{Tr}\, \mathbf{D}_k(K) > 0 .$$

Since $\mathbb{E}|W_{10}|^2 = 1$, one has $\mathbb{E}|W_{10}|^4 \geq 1$. The following is needed:

A4 At least one of the following conditions is satisfied:

$$\mathbb{E}|W_{10}|^4 > 1 \quad \text{or} \quad \liminf_{K} \frac{1}{K^2} \mathrm{Tr} \left(\mathbf{D}_0(K) \sum_{k=1}^{K} \mathbf{D}_k(K) \right) > 0 .$$

Remark.
If needed, one can attenuate the assumption on the eighth moment in **A1**. For instance, one can adapt without difficulty the proofs of our results to the case where $\mathbb{E}|W_{10}|^{4+\epsilon} < \infty$ for $\varepsilon > 0$. We assumed $\mathbb{E}|W_{10}|^8 < \infty$ because at some places we rely on results of [37] which are stated with the assumption on the eighth moment.
Assumption **A3** is technical. It has already appeared in [35].
Assumption **A4** is necessary to get a non-vanishing variance θ_K^2 in Theorem 4.

2.1 Mathematical tools

Stieltjes transform and resolvent Stieltjes transforms of probability measures and resolvent of hermitian matrices play a fundamental role in our approach. Let us begin by the following definitions.

Definition: *Stieltjes transform*
Let μ be a probability measure over $\mathbb{R}$. Its Stieltjes transform f is defined as

$$f(z) = \int_{\mathbb{R}} \frac{\mu(d\lambda)}{\lambda - z}, \quad z \in \mathbb{C}/\text{supp}(\mu),$$

where, $supp(\mu)$ refers to the support of measure μ. We shall denote by $\mathcal{S}(\mathbb{R}^+)$ the set of Stieltjes transforms of probability measures with support in $\mathbb{R}^+$.
The following proposition presents the main properties of the Stieltjes transforms.

Proposition 1 *The following properties hold true.*

1. *Let z be a complex number. Let f be a Stieltjes transform of a probability measure μ, then:*
   a) f is analytic over $\mathbb{C}/supp(\mu)$, and
   b) Let $d(z, \mathbb{R}^+)$ refers to the distance of z from the set of positive real numbers $\mathbb{R}^+$. Then, if $f \in \mathcal{S}(\mathbb{R}^+)$, we have, $|f(z)| \leq \frac{1}{d(z,\mathbb{R}^+)}$.

2. *Let $\mathbb{P}_n$ and $\mathbb{P}$ be probability measures over $\mathbb{R}$ and denote by f_n and f their Stieltjes transforms. Then,*

$$\left(\forall z \in \mathbb{C}^+, f_n(z) \xrightarrow[n\to\infty]{} f(z)\right) \quad \Longrightarrow \quad \mathbb{P}_n \xrightarrow[n\to\infty]{\mathcal{D}} \mathbb{P}.$$

   where $\mathbb{D}$ stands for the convergence in distribution.

Definition: *The resolvent matrix*
Let $\mathbf{A}$ be an $N \times N$ hermitian matrix. The complexe matrix fonction $\mathbf{Q}(z)$ defined as

$$\mathbf{Q}(z) = (\mathbf{A} - z\mathbf{I}_N)^{-1}, \quad z \in \mathbb{C} - \mathbb{R}$$

represents the resolvent of $\mathbf{A}$.

The following proposition illustrates the very close link between Stieltjes transform of the empirical distribution of the eigenvalues of a given matrix and the resolvent of this matrix. The second item of this proposition gives an upper bound for the spectrale norms of the resolvent matrices.

Proposition 2 *Let $\mathbf{Q}(z)$ be the resolvent of a hermitian matrix $\mathbf{A}$. Then,*

1. *The function $f_n(z) = \frac{1}{N} Tr\mathbf{Q}(z)$ is the Stieltjes transform of the empirical distribution of the eigenvalues of $\mathbf{A}$.*

2. $||\mathbf{Q}(z)|| \leq \frac{1}{d(z,\mathbb{R}^+)}$, *for every* $z \in \mathbb{C} - \mathbb{R}$.

The following lemma which reproduces [5, Lemma 2.7] will be used throughout this work. It characterizes the asymptotic behavior of an important class of quadratic forms:

Lemma 2.1 *Let* $\mathbf{x} = [X_1, \ldots, X_N]^t$ *be a* $N \times 1$ *vector where the* X_n *are centered i.i.d. complex random variables with unit variance. Let* $\mathbf{A}$ *be a deterministic* $N \times N$ *complex matrix. Then, for any* $p \geq 2$*, there exists a constant* C_p *depending on* p *only such that*

$$\mathbb{E}\left|\frac{1}{N}\mathbf{x}^*\mathbf{A}\mathbf{x} - \frac{1}{N}\mathrm{Tr}\,(\mathbf{A})\right|^p \leq \frac{C_p}{N^p}\left(\left(\mathbb{E}|X_1|^4\mathrm{Tr}\,(\mathbf{A}\mathbf{A}^*)\right)^{p/2} + \mathbb{E}|X_1|^{2p}\mathrm{Tr}\,\left((\mathbf{A}\mathbf{A}^*)^{p/2}\right)\right) . \tag{2.5}$$

Noticing that $\mathrm{Tr}\,(\mathbf{A}\mathbf{A}^*) \leq N\|\mathbf{A}\|^2$ and that $\mathrm{Tr}\,\left((\mathbf{A}\mathbf{A}^*)^{p/2}\right) \leq N\|\mathbf{A}\|^p$, we obtain the simpler inequality

$$\mathbb{E}\left|\frac{1}{N}\mathbf{x}^*\mathbf{A}\mathbf{x} - \frac{1}{N}\mathrm{Tr}\,(\mathbf{A})\right|^p \leq \frac{C_p}{N^{p/2}}\|\mathbf{A}\|^p\left(\left(\mathbb{E}|X_1|^4\right)^{p/2} + \mathbb{E}|X_1|^{2p}\right) \tag{2.6}$$

which is useful in case one has bounds on $\|\mathbf{A}\|$.

2.2 Deterministic approximations of random quadratic forms

Denote by $\mathbf{Q}_K(z)$ and $\widetilde{\mathbf{Q}}_K(z)$ the resolvents of $\mathbf{Y}(K)\mathbf{Y}(K)^*$ and $\mathbf{Y}(K)^*\mathbf{Y}(K)$ respectively, that is the $N \times N$ and $K \times K$ matrices defined by:

$$\mathbf{Q}_K(z) = (\mathbf{Y}(K)\mathbf{Y}(K)^* - z\mathbf{I}_N)^{-1} \qquad \text{and} \qquad \widetilde{\mathbf{Q}}_K(z) = (\mathbf{Y}(K)^*\mathbf{Y}(K) - z\mathbf{I}_K)^{-1} .$$

It is known [26, 35] that there exists a deterministic diagonal $N \times N$ matrix function $\mathbf{T}(z)$ that approximates the resolvent $\mathbf{Q}(z)$ in the following sense: Given a test matrix $\mathbf{S}$ with bounded spectral norm, the quantity $\frac{1}{K}\mathrm{Tr}\;\mathbf{S}(\mathbf{Q}(z) - \mathbf{T}(z))$ converges a.s. to zero as $K \to \infty$. It is also known that the approximation $\overline{\beta}_K$ of the quadratic form β_K is simply related to $\mathbf{T}(z)$ (cf. Theorem 2). As we shall see, matrix $\mathbf{T}(z)$ also plays a fundamental role in the second order result (Theorem 4).

In the following theorem, we recall the definition and some of the main properties of $\mathbf{T}(z)$.

Theorem 1 *The following hold true:*

1. *[35, Theorem 2.4] Let* $(\sigma^2_{nk}(K);\ 1 \leq n \leq N;\ 1 \leq k \leq K)$ *be a sequence of arrays of real numbers and consider the matrices* $\mathbf{D}_k(K)$ *and* $\widetilde{\mathbf{D}}_n(K)$ *defined in* (2.4)*. The system of* $N + K$ *functional equations*

$$\begin{cases} t_{n,K}(z) &= \dfrac{-1}{z\left(1 + \frac{1}{K}\mathrm{Tr}\,(\widetilde{\mathbf{D}}_n(K)\widetilde{\mathbf{T}}_K(z))\right)}, \quad 1 \leq n \leq N \\ \tilde{t}_{k,K}(z) &= \dfrac{-1}{z\left(1 + \frac{1}{K}\mathrm{Tr}\,(\mathbf{D}_k(K)\mathbf{T}_K(z))\right)}, \quad 1 \leq k \leq K \end{cases} \tag{2.7}$$

where

$$\mathbf{T}_K(z) = diag(t_{1,K}(z), \ldots, t_{N,K}(z)), \quad \widetilde{\mathbf{T}}_K(z) = diag(\tilde{t}_{1,K}(z), \ldots, \tilde{t}_{K,K}(z))$$

admits a unique solution $(\mathbf{T}, \widetilde{\mathbf{T}})$ *among the diagonal matrices for which the* $t_{n,K}$*'s and the* $\tilde{t}_{k,K}$*'s belong to class* $\mathcal{S}$*. Moreover, functions* $t_{n,K}(z)$ *and* $\tilde{t}_{k,K}(z)$ *admit an analytical continuation over* $\mathbb{C} - \mathbb{R}_+$ *which is real and positive for* $z \in (-\infty, 0)$*.*

2. *[35, Theorem 2.5] Assume that Assumptions* **A1** *and* **A2** *hold true. Consider the sequence of random matrices* $\mathbf{Y}(K)\mathbf{Y}(K)^*$ *where* $\mathbf{Y}$ *has dimensions* $N \times K$ *and whose entries are given by* $Y_{nk} = \frac{\sigma_{nk}}{\sqrt{K}} W_{nk}$*. For every sequence* $\mathbf{S}_K$ *of* $N \times N$ *diagonal matrices and every sequence* $\widetilde{\mathbf{S}}_K$ *of* $K \times K$ *diagonal matrices with*

$$\sup_K \max\left(\|\mathbf{S}_K\|, \|\widetilde{\mathbf{S}}_K\| \right) < \infty\,,$$

the following limits hold true almost surely:

$$\begin{aligned} \lim_{K\to\infty} \frac{1}{K} \mathrm{Tr}\ \mathbf{S}_K \left(\mathbf{Q}_K(z) - \mathbf{T}_K(z) \right) &= 0, \qquad \forall z \in \mathbb{C} - \mathbb{R}_+, \\ \lim_{K\to\infty} \frac{1}{K} \mathrm{Tr}\ \widetilde{\mathbf{S}}_K \left(\widetilde{\mathbf{Q}}_K(z) - \widetilde{\mathbf{T}}_K(z) \right) &= 0, \qquad \forall z \in \mathbb{C} - \mathbb{R}_+\ . \end{aligned}$$

Using Theorem 1 and Lemma 2.1, we are in position to characterize the asymptotic behavior of the quadratic form β_K given by (2.2). We begin by rewriting β_K as

$$\beta_K = \frac{1}{K} \mathbf{w}_0^* \mathbf{D}_0^{1/2} \left(\mathbf{Y}\mathbf{Y}^* + \rho \mathbf{I}_N \right)^{-1} \mathbf{D}_0^{1/2} \mathbf{w}_0 = \frac{1}{K} \mathbf{w}_0^* \mathbf{D}_0^{1/2} \mathbf{Q}(-\rho) \mathbf{D}_0^{1/2} \mathbf{w}_0 \tag{2.8}$$

where the $N \times 1$ vector $\mathbf{w}_0$ is given by $\mathbf{w}_0 = [W_{10}, \ldots, W_{N0}]^t$ and the diagonal matrix $\mathbf{D}_0$ is given by (2.4). Recall that $\mathbf{w}_0$ and $\mathbf{Q}$ are independent and that $\|\mathbf{D}_0\| \le \sigma^2_{\max}$ by **A2**. Furthermore, one can easily notice that $\|\mathbf{Q}(-\rho)\| = \|(\mathbf{Y}\mathbf{Y}^* + \rho \mathbf{I})^{-1}\| \le 1/\rho$.

Denote by $\mathbb{E}_{\mathbf{Q}}$ the conditional expectation with respect to $\mathbf{Q}$, i.e. $\mathbb{E}_{\mathbf{Q}} = \mathbb{E}(\,\cdot\,\|\mathbf{Q})$. From Inequality (2.6), there exists a constant $C > 0$ for which

$$\begin{aligned} \mathbb{E}\mathbb{E}_{\mathbf{Q}} \left| \beta_K - \frac{1}{K} \mathrm{Tr}\, \mathbf{D}_0 \mathbf{Q}(-\rho) \right|^4 &\le \frac{C}{K^2} \left(\frac{N}{K} \right)^2 \mathbb{E}\|\mathbf{D}_0 \mathbf{Q}\|^4 \left((\mathbb{E}|W_{10}|^4)^2 + \mathbb{E}|W_{10}|^8 \right) \\ &\le \frac{C}{K^2} \left(\frac{N}{K} \right)^2 \left(\frac{\sigma^2_{\max}}{\rho} \right)^4 \left((\mathbb{E}|W_{10}|^4)^2 + \mathbb{E}|W_{10}|^8 \right) \\ &= \mathcal{O}\left(\frac{1}{K^2} \right)\,. \end{aligned}$$

By the Borel-Cantelli Lemma, we therefore have

$$\beta_K - \frac{1}{K} \mathrm{Tr}\left(\mathbf{D}_0 \mathbf{Q}(-\rho) \right) \xrightarrow[K\to\infty]{} 0 \quad \text{a.s.}$$

Using this result, simply apply Theorem 1–(2) with $\mathbf{S} = \mathbf{D}_0$ (recall that $\|\mathbf{D}_0\| \le \sigma^2_{\max}$) to obtain:

Theorem 2 *Let* $\overline{\beta}_K = \frac{1}{K} \mathrm{Tr}\left(\mathbf{D}_0(K) \mathbf{T}_K(-\rho) \right)$ *where* $\mathbf{T}_K$ *is given by Theorem 1–(1). Assume* **A1** *and* **A2***. Then*

$$\beta_K - \overline{\beta}_K \xrightarrow[K\to\infty]{} 0 \quad a.s.$$

The deterministic approximation in the separable case

In the separable case $\sigma_{nk}(K) = d_n(K)\tilde{d}_k(K)$, matrices $\mathbf{D}_k(K)$ and $\widetilde{\mathbf{D}}_n(K)$ are written as $\mathbf{D}_k(K) = \tilde{d}_k(K)\mathbf{D}(K)$ and $\widetilde{\mathbf{D}}_n(K) = d_n(K)\widetilde{\mathbf{D}}(K)$ where $\mathbf{D}(K)$ and $\widetilde{\mathbf{D}}(K)$ are the diagonal matrices

$$\mathbf{D}(K) = \mathrm{diag}(d_1(K), \ldots, d_N(K)), \quad \widetilde{\mathbf{D}}(K) = \mathrm{diag}(\tilde{d}_1(K), \ldots, \tilde{d}_K(K)) \ . \tag{2.9}$$

and one can check that the system of $N + K$ equations leading to $\mathbf{T}_K$ and $\widetilde{\mathbf{T}}_K$ simplifies into a system of two equations, and Theorem 1 takes the following form:

Proposition 3 *[35, Sec. 3.2]*

1. *Assume $\sigma^2_{nk}(K) = d_n(K)\tilde{d}_k(K)$. Given $\rho > 0$, the system of two equations*

$$\begin{cases} \delta_K(\rho) &= \frac{1}{K}\mathrm{Tr}\left(\mathbf{D}\left(\rho(\mathbf{I}_N + \tilde{\delta}_K(\rho)\mathbf{D})\right)^{-1}\right) \\ \tilde{\delta}_K(\rho) &= \frac{1}{K}\mathrm{Tr}\left(\widetilde{\mathbf{D}}\left(\rho(\mathbf{I}_K + \delta_K(\rho)\widetilde{\mathbf{D}})\right)^{-1}\right) \end{cases} \tag{2.10}$$

   where $\mathbf{D}$ and $\widetilde{\mathbf{D}}$ are given by (2.9) *admits a unique solution $(\delta_K(\rho), \tilde{\delta}_K(\rho))$. Moreover, in this case matrices $\mathbf{T}(-\rho)$ and $\widetilde{\mathbf{T}}(-\rho)$ provided by Theorem 1–(1) coincide with*

$$\mathbf{T}(-\rho) = \frac{1}{\rho}(\mathbf{I} + \tilde{\delta}(\rho)\mathbf{D})^{-1} \quad \textit{and} \quad \widetilde{\mathbf{T}}(-\rho) = \frac{1}{\rho}(\mathbf{I} + \delta(\rho)\widetilde{\mathbf{D}})^{-1} \ . \tag{2.11}$$

2. *Assume that* **A1** *and* **A2** *hold true. Let matrices $\mathbf{S}_K$ and $\widetilde{\mathbf{S}}_K$ be as in Theorem 1–(2). Then, almost surely*

$$\frac{1}{K}\mathrm{Tr}\left(\mathbf{S}_K\left(\mathbf{Q}_K(-\rho) - \mathbf{T}_K(-\rho)\right)\right) \to 0, \textit{and} \quad \frac{1}{K}\mathrm{Tr}\left(\widetilde{\mathbf{S}}_K\left(\widetilde{\mathbf{Q}}_K(-\rho) - \widetilde{\mathbf{T}}_K(-\rho)\right)\right) \to 0,$$

   as $K \to \infty$.

With these equations we can adapt the result of Theorem 2 to the separable case. Notice that $\mathbf{D}_0 = \tilde{d}_0\mathbf{D}$ and that $\delta(\rho)$ given by the system (2.10) coincides with $\frac{1}{K}\mathrm{Tr}\,(\mathbf{DT})$, hence

Proposition 4 *Assume that $\sigma^2_{nk}(K) = d_n(K)\tilde{d}_k(K)$, and that* **A1** *and* **A2** *hold true. Then*

$$\frac{\beta_K}{\tilde{d}_0} - \delta_K(\rho) \xrightarrow[K\to\infty]{} 0 \quad a.s.$$

where $\delta_K(\rho)$ is given by Proposition 3–(1).

3 Second Order Results: Central Limit Theorem for Quadratic Forms

Our principal theoretical contribution to the study of the fluctuations of quadratic forms is presented in this section. Our approach is based on the decomposition of the quadratic form into a sum of martingale differences and on the use of the CLT for martingales [9].

Let us begin with some mathematical preliminaries.

3.1 Preliminaries

The following CLT for martingales is the key tool to study the asymptotic behavior of β_K:

Theorem 3 *[9] Let $X_{N,K}, X_{N-1,K}, \ldots, X_{1,K}$ be a martingale difference sequence with respect to the increasing filtration $\mathcal{G}_{N,K}, \ldots, \mathcal{G}_{1,K}$. Assume that there exists a sequence of real positive numbers s_K^2 such that*

$$\frac{1}{s_K^2}\sum_{n=1}^{N}\mathbb{E}\left[X_{n,K}^2 \| \mathcal{G}_{n+1,K}\right] \xrightarrow[K\to\infty]{} 1 \tag{2.12}$$

in probability. Assume further that the Lyapunov condition holds:

$$\exists \alpha > 0, \quad \frac{1}{s_K^{2(1+\alpha)}}\sum_{n=1}^{N}\mathbb{E}\left|X_{n,K}\right|^{2+\alpha} \xrightarrow[K\to\infty]{} 0 \ , \tag{2.13}$$

Then $s_K^{-1}\sum_{n=1}^{N} X_{n,K}$ converges in distribution to $\mathcal{N}(0,1)$ as $K \to \infty$.

Remark 1 *This theorem is proved in [9], gathering Theorem 35.12 (which is expressed under the weaker Lindeberg condition) together with the arguments of Section 27 (where it is proved that Lyapunov's condition implies Lindeberg's condition).*

The following inequality will be of help to check Lyapunov's condition.

Lemma 3.1 (Burkholder's inequality) *Let X_k be a complex martingale difference sequence with respect to the increasing sequence of σ–fields $\mathcal{F}_k$. Then for $p \geq 2$, there exists a constant C_p for which*

$$\mathbb{E}\left|\sum_k X_k\right|^p \leq C_p\left(\mathbb{E}\left(\sum_k \mathbb{E}\left[|X_k|^2 \| \mathcal{F}_{k-1}\right]\right)^{p/2} + \mathbb{E}\sum_k |X_k|^p\right) .$$

The following lemma gathers useful matrix results, whose proofs can be found in [39]:

Lemma 3.2 *Assume* $\mathbf{X} = [x_{ij}]_{i,j=1}^N$ *and* $\mathbf{Y}$ *are complex* $N \times N$ *matrices. Then*

1. *For every* $i, j \le N$, $|x_{ij}| \le \|\mathbf{X}\|$. *In particular,* $\|diag(\mathbf{X})\| \le \|\mathbf{X}\|$.
2. $\|\mathbf{XY}\| \le \|\mathbf{X}\| \, \|\mathbf{Y}\|$.
3. *For* $\rho > 0$, *the resolvent* $(\mathbf{XX}^* + \rho \mathbf{I})^{-1}$ *satisfies* $\|(\mathbf{XX}^* + \rho\mathbf{I})^{-1}\| \le \rho^{-1}$.
4. *If* $\mathbf{Y}$ *is Hermitian nonnegative, then* $|\mathrm{Tr}\,(\mathbf{XY})| \le \|\mathbf{X}\| \mathrm{Tr}\,(\mathbf{Y})$.

Let $\mathbf{X} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^*$ be a spectral decomposition of $\mathbf{X}$ where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ is the matrix of singular values of $\mathbf{X}$. For a real $p \ge 1$, the Schatten ℓ_p-norm of $\mathbf{X}$ is defined as $\|\mathbf{X}\|_p = (\sum \lambda_i^p)^{1/p}$. The following bound over the Schatten ℓ_p-norm of a triangular matrix will be of help (for a proof, see [7], [57, page 278]):

Lemma 3.3 *Let* $\mathbf{X} = [x_{ij}]_{i,j=1}^N$ *be a* $N \times N$ *complex matrix and let* $\widetilde{\mathbf{X}} = [x_{ij}\mathbf{1}_{i>j}]_{i,j=1}^N$ *be the strictly lower triangular matrix extracted from* $\mathbf{X}$. *Then for every* $p \ge 1$, *there exists a constant* C_p *depending on* p *only such that*

$$\|\widetilde{\mathbf{X}}\|_p \le C_p \|\mathbf{X}\|_p \ .$$

The following lemma lists some properties of the resolvent $\mathbf{Q}$ and the deterministic approximation matrix $\mathbf{T}$.

Lemma 3.4 *The following facts hold true:*

1. *Assume* **A2**. *Consider matrices* $\mathbf{T}_K(-\rho) = diag(t_1(-\rho), \dots, t_N(-\rho))$ *defined by Theorem 1–(1). Then for every* $1 \le n \le N$,

$$\frac{1}{\rho + \sigma^2_{\max}} \le t_n(-\rho) \le \frac{1}{\rho} \ . \tag{2.14}$$

2. *Assume in addition* **A1** *and* **A3**. *Let* $\mathbf{Q}_K(-\rho) = (\mathbf{YY}^* + \rho\mathbf{I})^{-1}$ *and let matrices* $\mathbf{S}_K$ *be as in the statement of Theorem 1–(2). Then*

$$\sup_K \mathbb{E}\,|\mathrm{Tr}\ \mathbf{S}_K(\mathbf{Q}_K - \mathbf{T}_K)|^2 < \infty \ . \tag{2.15}$$

Proof Let us establish (2.14). The lower bound immediately follows from the representation

$$t_n = \frac{1}{\rho + \frac{1}{K}\sum_{k=1}^K \frac{\sigma^2_{nk}}{1 + \frac{1}{K}\sum_{\ell=1}^N \sigma^2_{\ell k} t_\ell}} \overset{(a)}{\ge} \frac{1}{\rho + \sigma^2_{\max}}$$

where (a) follows from **A2** and $t_\ell(-\rho) \ge 0$. The upper bound requires an extra argument: As proved in [35, Theorem 2.4], the t_n's are Stieltjes transforms of probability measures supported by $\mathbb{R}_+$, i.e. there exists a probability measure μ_n over $\mathbb{R}_+$ such that $t_n(z) = \int \frac{\mu_n(dt)}{t-z}$. Thus

$$t_n(-\rho) = \int_0^\infty \frac{\mu_n(dt)}{t+\rho} \leq \frac{1}{\rho} ,$$

and (2.14) is proved.

We now briefly justify (2.15). We have $\mathbb{E}\,|\mathrm{Tr}\, \mathbf{S}(\mathbf{Q}-\mathbf{T})|^2 = \mathbb{E}\,|\mathrm{Tr}\, \mathbf{S}(\mathbf{Q}-\mathbb{E}\mathbf{Q})|^2 + |\mathrm{Tr}\, \mathbf{S}(\mathbb{E}\mathbf{Q}-\mathbf{T})|^2$. In [37, Lemma 6.3] it is stated that $\sup_K \mathbb{E}\,|\mathrm{Tr}\, \mathbf{S}(\mathbf{Q}-\mathbb{E}\mathbf{Q})|^2 < \infty$. Furthermore, in the proof of [37, Theorem 3.3] it is shown that $\sup_K K\|\mathbb{E}\mathbf{Q}-\mathbf{T}\| < \infty$, hence $|\mathrm{Tr}\, \mathbf{S}(\mathbb{E}\mathbf{Q}-\mathbf{T})| \leq K\|\mathbf{S}(\mathbb{E}\mathbf{Q}-\mathbf{T})\| \leq K\|\mathbb{E}\mathbf{Q}-\mathbf{T}\|\|\mathbf{S}\| < \infty$ by Lemma 3.2–(2). The result follows.

3.2 The main results: Central Limit Theorem for quadratique forms

The main result is given in the following theorem.

Theorem 4 *1. Assume that* **A2**, **A3** *and* **A4** *hold true. Let* $\mathbf{A}_K$ *and* $\mathbf{\Delta}_K$ *be the* $K\times K$ *matrices*

$$\begin{aligned} \mathbf{A}_K &= \left[\frac{1}{K} \frac{\frac{1}{K}\mathrm{Tr}\, \mathbf{D}_\ell \mathbf{D}_m \mathbf{T}(-\rho)^2}{\left(1+\frac{1}{K}\mathrm{Tr}\, \mathbf{D}_\ell \mathbf{T}(-\rho)\right)^2} \right]_{\ell,m=1}^K \quad \text{and} \\ \mathbf{\Delta}_K &= \mathit{diag}\left(\left(1+\frac{1}{K}\mathrm{Tr}\, \mathbf{D}_\ell \mathbf{T}(-\rho)\right)^2 ; 1\leq \ell \leq K \right) , \end{aligned} \tag{2.16}$$

where $\mathbf{T}$ *is defined in Theorem 1–(1). Let* $\mathbf{g}_K$ *be the* $K\times 1$ *vector*

$$\mathbf{g}_K = \left[\frac{1}{K}\mathrm{Tr}\, \mathbf{D}_0\mathbf{D}_1\mathbf{T}(-\rho)^2, \cdots, \frac{1}{K}\mathrm{Tr}\, \mathbf{D}_0\mathbf{D}_K\mathbf{T}(-\rho)^2 \right]^t .$$

Then the sequence of real numbers

$$\theta_K^2 = \frac{1}{K}\mathbf{g}_K^t(\mathbf{I}_K - \mathbf{A}_K)^{-1}\mathbf{\Delta}_K^{-1}\mathbf{g}_K + (\mathbb{E}|W_{10}|^4 - 1)\frac{1}{K}\mathrm{Tr}\, \mathbf{D}_0^2\mathbf{T}(-\rho)^2 \tag{2.17}$$

is well defined and furthermore

$$0 < \liminf_K \theta_K^2 \leq \limsup_K \theta_K^2 < \infty .$$

2. Assume in addition **A1**. *Then the sequence* $\beta_K = \mathbf{y}^*(\mathbf{Y}\mathbf{Y}^* + \rho\mathbf{I})^{-1}\mathbf{y}$ *satisfies*

$$\frac{\sqrt{K}}{\theta_K}\left(\beta_K - \overline{\beta}_K\right) \xrightarrow[K\to\infty]{} \mathcal{N}(0,1)$$

in distribution where $\overline{\beta}_K = \frac{1}{K}\mathrm{Tr}\, \mathbf{D}_0\mathbf{T}_K$ *is defined in the statement of Theorem 2.*

Remark 2 (On the achievability of the minimum of the variance) *As* $\mathbb{E}|W_{10}|^2 = 1$, *one clearly has* $\mathbb{E}|W_{10}|^4 - 1 \geq 0$ *with equality if and only if* $|W_{10}| = 1$ *with probability one. Moreover, we shall prove in the sequel that* $\liminf_K \frac{1}{K}\mathbf{D}_0(K)\mathbf{T}_K^2 > 0$. *Therefore* $(\mathbb{E}|W_{10}|^4 - 1)\frac{1}{K}\mathrm{Tr}\, \mathbf{D}_0^2\mathbf{T}^2$ *is nonnegative, and is zero if and only if* $|W_{10}| = 1$ *with probability one. As a consequence,* θ_K^2 *is minimum with respect to the distribution of the* W_{nk} *if and only if these random variables have their values on the unit circle.*

3.3 Proof of the main theorem

The following lemma, which directly follows from [Lemma 5.2 and Proposition 5.5] [37], states some important properties of the matrices $\mathbf{A}_K$ defined in the statement of Theorem 4. In the remainder of this chapter, $C = C(\rho, \sigma_{\max}^2, \liminf \frac{N}{K}, \sup \frac{N}{K}) < \infty$ denotes a positive constant whose value may change from line to line.

Lemma 3.5 *Assume* **A2** *and* **A3**. *Consider matrices* $\mathbf{A}_K$ *defined by* (2.16). *Then the following facts hold true:*

1. *Matrix* $\mathbf{I}_K - \mathbf{A}_K$ *is invertible, and* $(\mathbf{I}_K - \mathbf{A}_K)^{-1} \succcurlyeq \mathbf{0}$.
2. *Element* (k,k) *of the inverse satisfies* $\left[(\mathbf{I}_K - \mathbf{A}_K)^{-1}\right]_{k,k} \geq 1$ *for every* $1 \leq k \leq K$.
3. *The maximum row sum norm of the inverse satisfies* $\limsup_K ||(\mathbf{I}_K - \mathbf{A}_K)^{-1}||_\infty < \infty$.

Proof of Theorem 4–(1)

Due to Lemma 3.5–(1), θ_K^2 is well defined. Let us prove that $\limsup_K \theta_K^2 < \infty$. The first term of the right-hand side of (2.17) satisfies

$$\frac{1}{K}\mathbf{g}^t(\mathbf{I}_K - \mathbf{A}_K)^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} \leq \|\mathbf{g}\|_\infty \|(\mathbf{I}_K - \mathbf{A}_K)^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g}\|_\infty$$
$$\leq \|\mathbf{g}\|_\infty ||(\mathbf{I}_K - \mathbf{A}_K)^{-1}||_\infty \|\boldsymbol{\Delta}^{-1}\mathbf{g}\|_\infty \leq \|\mathbf{g}\|_\infty^2 ||(\mathbf{I}_K - \mathbf{A}_K)^{-1}||_\infty \quad (2.18)$$

due to $||\boldsymbol{\Delta}^{-1}||_\infty \leq 1$. Recall that $\|\mathbf{T}\| \leq \rho^{-1}$ by Lemma 3.4–(1). Therefore, any element of $\mathbf{g}$ satisfies

$$\frac{1}{K}\operatorname{Tr}\mathbf{D}_0\mathbf{D}_k\mathbf{T}^2 \leq \frac{N}{K}\|\mathbf{D}_0\|\|\mathbf{D}_k\|\|\mathbf{T}\|^2 \leq \frac{N}{K}\frac{\sigma_{\max}^4}{\rho^2} \quad (2.19)$$

by **A2**, hence $\sup_K \|\mathbf{g}\| \leq C$. From Lemma 3.5–(3) and (2.18), we then obtain

$$\limsup_K \frac{1}{K}\mathbf{g}^t(\mathbf{I}_K - \mathbf{A}_K)^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} \leq C. \quad (2.20)$$

We can prove similarly that the second term in the right-hand side of (2.17) satisfies $\sup_K((\mathbb{E}|W_{10}|^4 - 1)\frac{1}{K}\operatorname{Tr}\mathbf{D}_0^2\mathbf{T}(-\rho)^2) \leq C$. Hence $\limsup_K \theta_K^2 < \infty$.

Let us prove that $\liminf_K \theta_K^2 > 0$. We have

$$\begin{aligned}
\frac{1}{K}\mathbf{g}^t(\mathbf{I}_K - \mathbf{A}_K)^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} &\stackrel{(a)}{\geq} \frac{1}{K}\mathbf{g}^t \mathrm{diag}\left((\mathbf{I}_K - \mathbf{A}_K)^{-1}\right)\boldsymbol{\Delta}^{-1}\mathbf{g} \\
&\stackrel{(b)}{\geq} \frac{1}{\left(1+\frac{N}{K}\frac{\sigma_{\max}^2}{\rho}\right)^2}\frac{1}{K}\sum_{k=1}^K\left(\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_0\mathbf{D}_k\mathbf{T}^2\right)^2 \\
&\stackrel{(c)}{\geq} \frac{1}{\left(1+\frac{N}{K}\frac{\sigma_{\max}^2}{\rho}\right)^2}\left(\frac{1}{K^2}\mathrm{Tr}\ \mathbf{D}_0\left(\sum_{k=1}^K\mathbf{D}_k\right)\mathbf{T}^2\right)^2 \\
&\stackrel{(d)}{\geq} \frac{1}{\left(1+\frac{N}{K}\frac{\sigma_{\max}^2}{\rho}\right)^2(\rho+\sigma_{\max}^2)^4}\left(\frac{1}{K^2}\mathrm{Tr}\ \mathbf{D}_0\sum_{k=1}^K\mathbf{D}_k\right)^2 \\
&\geq C\left(\frac{1}{K^2}\mathrm{Tr}\ \mathbf{D}_0\sum_{k=1}^K\mathbf{D}_k\right)^2 ,
\end{aligned}$$

where (a) follows from the fact that $(\mathbf{I}_K - \mathbf{A}_K)^{-1} \succcurlyeq \mathbf{0}$ (Lemma 3.5–(1), and the straightforward inequalities $\boldsymbol{\Delta}^{-1} \succcurlyeq \mathbf{0}$ and $\mathbf{g} \succcurlyeq \mathbf{0}$), (b) follows from Lemma 3.5–(2) and $\|\boldsymbol{\Delta}\| \leq (1+\frac{N}{K}\frac{\sigma_{\max}^2}{\rho})^2$, (c) follows from the elementary inequality $n^{-1}\sum x_i^2 \geq (n^{-1}\sum x_i)^2$, and (d) is due to Lemma 3.4–(1). Similar derivations yield:

$$(\mathbb{E}|W_{10}|^4-1)\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_0^2\mathbf{T} \quad\geq\quad \frac{\mathbb{E}|W_{10}|^4-1}{(\rho+\sigma_{\max}^2)^2}\left(\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_0\right)^2 \quad\geq\quad C(\mathbb{E}|W_{10}|^4-1)$$

by **A3**. Therefore, if **A4** holds true, then $\liminf_K \theta_K^2 > 0$ and Theorem 4–(1) is proved.

Proof of Theorem 4–(2)

Recall that the quadratic form β_K is given by Equation (2.8). The random variable $\frac{\sqrt{K}}{\theta_K}(\beta_K - \overline{\beta}_K)$ can therefore be decomposed as

$$\begin{aligned}
\frac{\sqrt{K}}{\theta_K}(\beta_K - \overline{\beta}_K) &= \frac{1}{\sqrt{K}\theta_K}\left(\mathbf{w}_0^*\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0^{1/2}\mathbf{w}_0 - \mathrm{Tr}\,(\mathbf{D}_0\mathbf{Q})\right) + \frac{1}{\sqrt{K}\theta_K}\left(\mathrm{Tr}\,(\mathbf{D}_0(\mathbf{Q}-\mathbf{T}))\right) \\
&= U_{1,K} + U_{2,K}\ .
\end{aligned} \tag{2.21}$$

Thanks to Lemma 3.4–(2) and to the fact that $\liminf_K \theta_K^2 > 0$, we have $\mathbb{E}U_{K,2}^2 < CK^{-1}$ which implies that $U_{K,2} \to 0$ in probability as $K \to \infty$. Hence, in order to conclude that

$$\frac{\sqrt{K}}{\theta_K}(\beta_K - \overline{\beta}_K) \xrightarrow[K\to\infty]{} \mathcal{N}(0,1) \quad \text{in distribution ,}$$

it is sufficient by Slutsky's theorem to prove that $U_{1,K} \to \mathcal{N}(0,1)$ in distribution. The remainder of the section is devoted to this point.

Remark 3 *Decomposition* (2.21) *and the convergence to zero (in probability) of* $U_{2,K}$ *yield the following interpretation: The fluctuations of* $\sqrt{K}(\beta_K - \overline{\beta}_K)$ *are mainly due to the* *fluctuations of vector* $\mathbf{w}_0$*. Indeed the contribution of the fluctuations*[1] *of* $\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_0\mathbf{Q}$*, due to the random nature of* $\mathbf{Y}$*, is negligible.*

Denote by $\mathbb{E}_n$ the conditional expectation $\mathbb{E}_n[\,\cdot\,] = \mathbb{E}[\,\cdot \parallel W_{n,0}, W_{n+1,0}, \dots, W_{N,0}, \mathbf{Y}]$. Put $\mathbb{E}_{N+1}[\,\cdot\,] = \mathbb{E}[\,\cdot \parallel \mathbf{Y}]$ and note that $\mathbb{E}_{N+1}(\mathbf{w}_0^*\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0^{1/2}\mathbf{w}_0) = \mathrm{Tr}\,\mathbf{D}_0\mathbf{Q}$. With these notations at hand, we have:

$$U_{1,K} = \frac{1}{\theta_K}\sum_{n=1}^{N}(\mathbb{E}_n - \mathbb{E}_{n+1})\frac{\mathbf{w}_0^*\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0^{1/2}\mathbf{w}_0}{\sqrt{K}} \triangleq \frac{1}{\theta_K}\sum_{n=1}^{N} Z_{n,K}\ . \tag{2.22}$$

Consider the increasing sequence of $\sigma-$fields

$$\mathcal{F}_{N,K} = \sigma(W_{N,0}, \mathbf{Y})\ ,\quad \cdots\ ,\quad \mathcal{F}_{1,K} = \sigma(W_{1,0}, \cdots, W_{N,0}, \mathbf{Y})\ .$$

Then the random variable $Z_{n,K}$ is integrable and measurable with respect to $\mathcal{F}_{n,K}$; moreover it readily satisfies $\mathbb{E}_{n+1}Z_{n,K} = 0$. In particular, the sequence $(Z_{N,K}, \dots, Z_{1,K})$ is a martingale difference sequence with respect to $(\mathcal{F}_{N,K}, \cdots, \mathcal{F}_{1,K})$.

In order to prove that

$$U_{1,K} = \frac{1}{\theta_K}\sum_{n=1}^{N} Z_{n,K} \xrightarrow[K\to\infty]{} \mathcal{N}(0,1) \quad \text{in distribution}\ , \tag{2.23}$$

we shall apply Theorem 3 to the sum $\frac{1}{\theta_K}\sum_{n=1}^{N} Z_{n,K}$ and the filtration $(\mathcal{F}_{n,K})$. The proof is carried out into four steps:

Step 1 We first establish Lyapunov's condition. Due to the fact that $\liminf_K \theta_K^2 > 0$, we only need to show that

$$\exists\ \alpha > 0,\quad \sum_{n=1}^{N} \mathbb{E}|Z_{n,K}|^{2+\alpha} \xrightarrow[K\to\infty]{} 0\ . \tag{2.24}$$

Step 2 We prove that $V_K = \sum_{n=1}^{N} \mathbb{E}_{n+1}Z_{n,K}^2$ satisfies

$$V_K - \left(\frac{\left(\mathbb{E}|W_{10}|^4 - 2\right)}{K}\mathrm{Tr}\,\left(\mathbf{D}_0^2(\mathrm{diag}(\mathbf{Q}))^2\right) + \frac{1}{K}\mathrm{Tr}\,(\mathbf{D}_0\mathbf{Q}\mathbf{D}_0\mathbf{Q})\right) \xrightarrow[K\to\infty]{} 0 \quad \text{in probability}\ . \tag{2.25}$$

[1]In fact, one may prove that the fluctuation of $\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_0(\mathbf{Q} - \mathbf{T})$ are of order K, i.e. $\mathrm{Tr}\,\mathbf{D}_0(\mathbf{Q} - \mathbf{T})$ asymptotically behaves as a Gaussian random variable. Such a speed of fluctuations already appears in [37], when studying the fluctuations of the mutual information.

Step 3 We first show that

$$\frac{1}{K}\operatorname{Tr}\mathbf{D}_0^2(\operatorname{diag}(\mathbf{Q}))^2 - \frac{1}{K}\operatorname{Tr}\mathbf{D}_0^2\mathbf{T}^2 \xrightarrow[K\to\infty]{} 0 \quad \text{in probability.} \tag{2.26}$$

In order to study the asymptotic behavior of $\frac{1}{K}\operatorname{Tr}(\mathbf{D}_0\mathbf{Q}\mathbf{D}_0\mathbf{Q})$, we introduce the random variables $U_\ell = \frac{1}{K}\operatorname{Tr}(\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{Q})$ for $0 \le \ell \le K$ (the one of interest being U_0). We then prove that the U_ℓ's satisfy the following system of equations:

$$U_\ell = \sum_{k=1}^{K} c_{\ell k} U_k + \frac{1}{K}\operatorname{Tr}\mathbf{D}_0\mathbf{D}_\ell\mathbf{T}^2 + \epsilon_\ell, \quad 0 \le \ell \le K, \tag{2.27}$$

where

$$c_{\ell k} = \frac{1}{K}\frac{\frac{1}{K}\operatorname{Tr}\mathbf{D}_\ell\mathbf{D}_k\mathbf{T}(-\rho)^2}{\left(1+\frac{1}{K}\operatorname{Tr}\mathbf{D}_k\mathbf{T}(-\rho)\right)^2}, \quad 0 \le \ell \le K,\ 1 \le k \le K \tag{2.28}$$

and the perturbations ϵ_ℓ satisfy $\mathbb{E}|\epsilon_\ell| \le CK^{-\frac{1}{2}}$ where we recall that C is independent of ℓ.

Step 4 We prove that $U_0 = \frac{1}{K}\operatorname{Tr}\mathbf{D}_0\mathbf{Q}\mathbf{D}_0\mathbf{Q}$ satisfies

$$U_0 = \frac{1}{K}\operatorname{Tr}\mathbf{D}_0^2\mathbf{T}^2 + \frac{1}{K}\mathbf{g}^t\left(\mathbf{I}-\mathbf{A}\right)^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} + \epsilon \tag{2.29}$$

with $\mathbb{E}|\epsilon| \le CK^{-\frac{1}{2}}$. This equation combined with (2.25) and (2.26) yields $\sum_n \mathbb{E}_{n+1}Z_{n,K}^2 - \theta_K^2 \to 0$ in probability. As $\liminf_K \theta_K^2 > 0$, this implies $\frac{1}{\theta_K}\sum_n \mathbb{E}_{n+1}Z_{n,K}^2 \to 1$ in probability, which proves (2.23) and thus ends the proof of Theorem 4.

Write $\mathbf{B} = [b_{ij}]_{i,j=1}^N = \mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0^{1/2}$ and recall from (2.22) that $Z_{n,K} = \frac{1}{\sqrt{K}}(\mathbb{E}_n - \mathbb{E}_{n+1})\mathbf{w}_0^*\mathbf{B}\mathbf{w}_0$. We have

$$\mathbb{E}_n\mathbf{w}_0^*\mathbf{B}\mathbf{w}_0 = \sum_{\ell=1}^{n-1} b_{\ell\ell} + \sum_{\ell_1,\ell_2=n}^{N} W_{\ell_1 0}^* W_{\ell_2 0} b_{\ell_1\ell_2}\ .$$

Hence

$$Z_{n,K} = \frac{1}{\sqrt{K}}\left(\left(|W_{n0}|^2-1\right)b_{nn} + W_{n0}^*\sum_{\ell=n+1}^{N} W_{\ell 0}b_{n\ell} + W_{n0}\sum_{\ell=n+1}^{N} W_{\ell 0}^* b_{\ell n}\right)\ . \tag{2.30}$$

Step 1: Validation of the Lyapunov condition Recall Assumption **A1**. Eq. (2.30) yields:

$$\begin{aligned}
|Z_{n,K}|^4 &\le \frac{1}{K^2}\left(\frac{|W_{n0}|^2+1}{\rho\sigma_{\max}^2} + 2\left|W_{n0}\sum_{\ell=n+1}^{N} W_{\ell 0}b_{n\ell}\right|\right)^4 \\
&\le \frac{2^3}{K^2}\left(\left(\frac{|W_{n0}|^2+1}{\rho\sigma_{\max}^2}\right)^4 + 2^4\left|W_{n0}\sum_{\ell=n+1}^{N} W_{\ell 0}b_{n\ell}\right|^4\right) \qquad (2.31)
\end{aligned}$$

where we use the fact that $|b_{nn}| \le (\rho\sigma^2_{\max})^{-1}$ (cf. Lemma 3.2–(1)) and the convexity of $x \mapsto x^4$. Due to Assumption **A1**, we have:

$$\mathbb{E}\left(|W_{n0}|^2+1\right)^4 \le 2^3\left(\mathbb{E}|W_{n0}|^8+1\right) < \infty \ . \tag{2.32}$$

Considering the second term at the right-hand side of (2.31), we write

$$\begin{aligned}
\mathbb{E}\left|W_{n0}\sum_{\ell=n+1}^{N} W_{\ell 0}b_{n\ell}\right|^4 &= \mathbb{E}\,|W_{n0}|^4\,\mathbb{E}\left|\sum_{\ell=n+1}^{N} W_{\ell 0}b_{n\ell}\right|^4 ,\\
&\overset{(a)}{\le} C\left(\mathbb{E}\left(\sum_{\ell=n+1}^{N}(\mathbb{E}|W_{\ell 0}|^2)|b_{n\ell}|^2\right)^2+\sum_{\ell=n+1}^{N}(\mathbb{E}|W_{\ell 0}|^4)(\mathbb{E}|b_{n\ell}|^4)\right) ,\\
&\overset{(b)}{\le} C\left(\mathbb{E}\left(\sum_{\ell=n+1}^{N}|b_{n\ell}|^2\right)^2+\sum_{\ell=n+1}^{N}\mathbb{E}|b_{n\ell}|^2\right) ,
\end{aligned}$$

where (a) follows from Lemma 3.1 (Burkholder's inequality), the filtration being $\mathcal{F}_{N,K},\dots,$ $\mathcal{F}_{n+1,K}$ and (b) follows from the bound $|b_{n\ell}|^4 \le |b_{n\ell}|^2 \max|b_{n\ell}|^2 \le |b_{n\ell}|^2(\sigma^2_{\max}\rho^{-1})^2$ (cf. Lemma 3.2–(1)). Now, notice that

$$\sum_{\ell=n+1}^{N}|b_{n\ell}|^2 < \sum_{\ell=1}^{N}|b_{n\ell}|^2 = \left[\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0\mathbf{Q}\mathbf{D}_0^{1/2}\right]_{nn} \le \|\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0\mathbf{Q}\mathbf{D}_0^{1/2}\| \le \frac{\sigma^4_{\max}}{\rho^2} \ .$$

This yields $\mathbb{E}|W_{n0}\sum_{\ell=n+1}^{N}W_{\ell 0}b_{n\ell}|^4 \le C$. Gathering this result with (2.32), getting back to (2.31), taking the expectation and summing up finally yields:

$$\sum_{n=1}^{N}\mathbb{E}|Z_{n,K}|^4 \ \le\ \frac{C}{K} \xrightarrow[K\to\infty]{} 0$$

which establishes Lyapunov's condition (2.24) with $\alpha = 2$.

Step 2: Proof of (2.25) Eq. (2.30) yields:

$$\begin{aligned}
\mathbb{E}_{n+1}Z^2_{n,K} \;=\; \frac{1}{K}\Bigg(&\left(\mathbb{E}|W_{10}|^4-1\right)b^2_{nn}+\mathbb{E}_{n+1}\left(W^*_{n0}\sum_{\ell=n+1}^{N}W_{\ell 0}b_{n\ell}+W_{n0}\sum_{\ell=n+1}^{N}W^*_{\ell 0}b_{\ell n}\right)^2\\
&+2b_{nn}\left(\mathbb{E}\,W^*_{10}|W_{10}|^2\right)\sum_{\ell=n+1}^{N}W_{\ell 0}b_{n\ell}+2b_{nn}\left(\mathbb{E}\,W_{10}|W_{10}|^2\right)\sum_{\ell=n+1}^{N}W^*_{\ell 0}b_{\ell n}\Bigg) \ .
\end{aligned}$$

Note that the second term of the right-hand side writes:

$$\mathbb{E}_{n+1}\left(W^*_{n0}\sum_{\ell=n+1}^{N}W_{\ell 0}b_{n\ell}+W_{n0}\sum_{\ell=n+1}^{N}W^*_{\ell 0}b_{\ell n}\right)^2 = 2\sum_{\ell_1,\ell_2=n+1}^{N}W_{\ell_1 0}W^*_{\ell_2 0}b_{n\ell_1}b_{\ell_2 n} \ .$$

Therefore, $V_K = \sum_{n=1}^N \mathbb{E}_{n+1} Z_{n,K}^2$ writes:

$$V_K = \frac{\left(\mathbb{E}|W_{10}|^4 - 1\right)}{K} \sum_{n=1}^N b_{nn}^2 + \frac{2}{K} \sum_{n=1}^N \sum_{\ell_1,\ell_2=n+1}^N W_{\ell_1 0} W_{\ell_2 0}^* b_{n\ell_1} b_{\ell_2 n}$$
$$+ \frac{2}{K} \Re \left(\left(\mathbb{E}\, W_{10}^* |W_{10}|^2\right) \sum_{n=1}^N b_{nn} \sum_{\ell=n+1}^N W_{\ell 0} b_{n\ell} \right) ,$$

where $\Re$ denotes the real part of a complex number. We introduce the following notations:

$$\mathbf{R} = (r_{ij})_{i,j=1}^N \stackrel{\triangle}{=} (b_{ij} \mathbf{1}_{i>j})_{i,j=1}^N \qquad \text{and} \qquad \Gamma_K = \frac{1}{K} \sum_{n=1}^N b_{nn} \sum_{\ell=n+1}^N W_{\ell 0} b_{n\ell} .$$

Note in particular that $\mathbf{R}$ is the strictly lower triangular matrix extracted from $\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0^{1/2}$. We can now rewrite V_K as:

$$V_K = \frac{\left(\mathbb{E}|W_{10}|^4 - 1\right)}{K} \mathrm{Tr}\, \left(\mathbf{D}_0^2 (\mathrm{diag}(\mathbf{Q}))^2\right) + \frac{2}{K} \mathbf{w}_0^* \mathbf{R}\mathbf{R}^* \mathbf{w}_0 + 2\Re\left(\Gamma_K \mathbb{E} W_{10}^* |W_{10}|^2\right) . \quad (2.33)$$

We now prove that the third term of the right-hand side vanishes, and find an asymptotic equivalent for the second one. Using Lemma 3.2, we have:

$$\begin{aligned}
\mathbb{E}_{N+1}|\Gamma_K|^2 &= \frac{1}{K^2} \sum_{n,m=1}^N b_{nn} b_{mm} \sum_{\ell=1}^N b_{n\ell} b_{m\ell}^* \mathbf{1}_{\ell>n} \mathbf{1}_{\ell>m} = \frac{1}{K^2} \mathrm{Tr}\, (\mathrm{diag}(\mathbf{B}) \mathbf{R}^* \mathbf{R} \mathrm{diag}(\mathbf{B})) \\
&= \frac{1}{K^2} \mathrm{Tr}\, \left(\mathbf{D}_0^{1/2} \mathrm{diag}(\mathbf{Q}) \mathbf{D}_0^{1/2} \mathbf{R}^* \mathbf{R} \mathbf{D}_0^{1/2} \mathrm{diag}(\mathbf{Q}) \mathbf{D}_0^{1/2}\right) \\
&\le \frac{1}{K^2} \|\mathbf{D}_0\|^2 \|\mathbf{Q}\|^2 \mathrm{Tr}\,(\mathbf{R}^*\mathbf{R}) \quad \le \quad \frac{1}{K^2} \|\mathbf{D}_0\|^2 \|\mathbf{Q}\|^2 \mathrm{Tr}\,(\mathbf{B}^2) \quad \le \quad \frac{1}{K^2} \|\mathbf{D}_0\|^4 \|\mathbf{Q}\|^2 \mathrm{Tr}\,(\mathbf{Q}^2) \\
&\le \frac{1}{K} \|\mathbf{D}_0\|^2 \|\mathbf{Q}\|^4 \quad \le \quad \frac{1}{K} \frac{\sigma_{\max}^4}{\rho^4} \quad \xrightarrow[K\to\infty]{} \quad 0 .
\end{aligned}$$

In particular, $\mathbb{E}|\Gamma_K|^2 \to 0$ and

$$\Re\left(\left(\mathbb{E} W_{10}^* |W_{10}|^2\right) \Gamma_K\right) \xrightarrow[K\to\infty]{} 0 \quad \text{in probability} . \quad (2.34)$$

Consider now the second term of the right-hand side of Eq. (2.33). We prove that:

$$\frac{1}{K} \mathbf{w}_0^* \mathbf{R}\mathbf{R}^* \mathbf{w}_0 - \frac{1}{K} \mathrm{Tr}\,(\mathbf{R}\mathbf{R}^*) \xrightarrow[K\to\infty]{} 0 \quad \text{in probability.} \quad (2.35)$$

By Lemma 2.1 (Ineq. (2.5)), we have

$$\mathbb{E} \left(\frac{1}{K} \mathbf{w}_0^* \mathbf{R}\mathbf{R}^* \mathbf{w}_0 - \frac{1}{K} \mathrm{Tr}\,(\mathbf{R}\mathbf{R}^*) \right)^2 \le \frac{C}{K^2} (\mathbb{E}|W_{10}|^4) \mathrm{Tr}\,(\mathbf{R}\mathbf{R}^*\mathbf{R}\mathbf{R}^*) .$$

Notice that $\mathrm{Tr}\,(\mathbf{R}\mathbf{R}^*\mathbf{R}\mathbf{R}^*) = \|\mathbf{R}\|_4^4$ where $\|\mathbf{R}\|_4$ is the Schatten ℓ_4-norm of $\mathbf{R}$. Using Lemma 3.3, we have:

$$\|\mathbf{R}\|_4^4 \le C \|\mathbf{D}_0^{1/2} \mathbf{Q} \mathbf{D}_0^{1/2}\|_4^4 \le NC \|\mathbf{D}_0^{1/2} \mathbf{Q} \mathbf{D}_0^{1/2}\|^4 \le N \frac{C\sigma_{\max}^8}{\rho^4} .$$

Therefore,

$$\mathbb{E}\left(\frac{1}{K}\mathbf{w}_0^*\mathbf{R}\mathbf{R}^*\mathbf{w}_0 - \frac{1}{K}\mathrm{Tr}\,(\mathbf{R}\mathbf{R}^*)\right)^2 \le C\frac{N}{K^2} \xrightarrow[K\to\infty]{} 0$$

which implies (2.35). Now, due to the fact that $\mathbf{B} = \mathbf{B}^*$, we have

$$\begin{aligned}
\frac{2}{K}\mathrm{Tr}\ \mathbf{R}\mathbf{R}^* &= \frac{2}{K}\sum_{n=1}^{N}\sum_{\ell=n+1}^{N}|b_{n\ell}|^2 \\
&= \frac{1}{K}\sum_{n,\ell=1}^{N}|b_{n\ell}|^2 - \frac{1}{K}\sum_{n=1}^{N}|b_{nn}|^2 \\
&= \frac{1}{K}\mathrm{Tr}\ \mathbf{D}_0\mathbf{Q}\mathbf{D}_0\mathbf{Q} - \frac{1}{K}\mathrm{Tr}\ \mathbf{D}_0^2(\mathrm{diag}(\mathbf{Q}))^2
\end{aligned} \tag{2.36}$$

Gathering (2.33–2.36), we obtain (2.25). Step 2 is proved.

Step 3: Proof of (2.26) **and** (2.27) We begin with some identities. Write $\mathbf{Q}(z) = [q_{ij}(z)]_{i,j=1}^N$ and $\widetilde{\mathbf{Q}}(z) = [\tilde{q}_{ij}(z)]_{i,j=1}^K$. Denote by $\mathbf{y}_k$ the column number k of $\mathbf{Y}$ and by ξ_n the row number n of $\mathbf{Y}$. Denote by $\mathbf{Y}^k$ the matrix that remains after deleting column k from $\mathbf{Y}$ and by $\mathbf{Y}_n$ the matrix that remains after deleting row n from $\mathbf{Y}$. Finally, write $\mathbf{Q}_k(z) = (\mathbf{Y}^k\mathbf{Y}^{k*} - z\mathbf{I})^{-1}$ and $\widetilde{\mathbf{Q}}_n(z) = (\mathbf{Y}_n^*\mathbf{Y}_n - z\mathbf{I})^{-1}$. The following formulas can be established easily (see for instance [39, §0.7.3. and §0.7.4]):

$$q_{nn}(-\rho) = \frac{1}{\rho(1+\xi_n\widetilde{\mathbf{Q}}_n(-\rho)\xi_n^*)}, \quad \tilde{q}_{kk}(-\rho) = \frac{1}{\rho(1+\mathbf{y}_k^*\mathbf{Q}_k(-\rho)\mathbf{y}_k)}, \tag{2.37}$$

$$\mathbf{Q} = \mathbf{Q}_k - \frac{\mathbf{Q}_k\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_k}{1+\mathbf{y}_k^*\mathbf{Q}_k\mathbf{y}_k} \tag{2.38}$$

Lemma 3.6 *The following hold true:*

1. *(*Rank one perturbation inequality*) The resolvent* $\mathbf{Q}_k(-\rho)$ *satisfies* $|\mathrm{Tr}\,\mathbf{A}(\mathbf{Q}-\mathbf{Q}_k)| \le \|\mathbf{A}\|/\rho$ *for any* $N\times N$ *matrix* $\mathbf{A}$.

2. *Let Assumptions* **A1**–**A3** *hold. Then,*

$$\max_{1\le n\le N}\mathbb{E}(q_{nn}(-\rho) - t_n(-\rho))^2 \le \frac{C}{K}\ . \tag{2.39}$$

The same conclusion holds true if q_{nn} *and* t_n *are replaced with* $\tilde{q}_{kk}$ *and* $\tilde{t}_k$ *respectively.*

Proof 1 *The proof of Part 1 can be found in [37, Proof of Lemma 6.3] (see also [4, Lemma 2.6]). Let us prove Part 2. We have from Equations* (2.7) *and* (2.37)

$$\begin{aligned}
|q_{nn}(-\rho) - t_n(-\rho)| &= \frac{1}{\rho(1+\frac{1}{K}\mathrm{Tr}\,\widetilde{\mathbf{D}}_n\widetilde{\mathbf{T}})(1+\xi_n\widetilde{\mathbf{Q}}_n\xi_n^*)}\left|\xi_n\widetilde{\mathbf{Q}}_n\xi_n^* - \frac{1}{K}\mathrm{Tr}\,\widetilde{\mathbf{D}}_n\widetilde{\mathbf{T}}\right| \\
&\le \frac{1}{\rho}\left|\xi_n\widetilde{\mathbf{Q}}_n\xi_n^* - \frac{1}{K}\mathrm{Tr}\,\widetilde{\mathbf{D}}_n\widetilde{\mathbf{T}}\right|\ .
\end{aligned}$$

Hence,

$$\mathbb{E}(q_{nn}-t_n)^2 \quad \leq \quad \frac{2}{\rho}\mathbb{E}\left(\xi_n\widetilde{\mathbf{Q}}_n\xi_n^* - \frac{1}{K}\mathrm{Tr}\,\widetilde{\mathbf{D}}_n\widetilde{\mathbf{Q}}\right)^2 + \frac{2}{\rho K^2}\mathbb{E}\left(\mathrm{Tr}\,\widetilde{\mathbf{D}}_n(\widetilde{\mathbf{Q}}-\widetilde{\mathbf{T}})\right)^2 \quad \leq \quad \frac{C}{K}$$

by Lemma 2.1 and Lemma 3.4–(2), which proves (2.39).

We are now in position to prove (2.26). First, notice that:

$$\begin{aligned}\mathbb{E}\left|q_{nn}^2-t_n^2\right| &= \mathbb{E}\left|q_{nn}-t_n\right|(q_{nn}+t_n)\\ &\leq \sqrt{\mathbb{E}(q_{nn}-t_n)^2}\sqrt{\mathbb{E}(q_{nn}+t_n)^2} \quad \leq \quad \frac{2}{\rho}\sqrt{\mathbb{E}(q_{nn}-t_n)^2}\ . \end{aligned}\tag{2.40}$$

Now,

$$\begin{aligned}\frac{1}{K}\mathbb{E}\left|\mathrm{Tr}\ \mathbf{D}_0^2(\mathrm{diag}(\mathbf{Q})^2-\mathbf{T}^2)\right| &\leq \frac{1}{K}\sum_{n=1}^{N}\sigma_{0,n}^4\mathbb{E}\left|q_{nn}^2-t_n^2\right| \quad \leq \quad \frac{\sigma_{\max}^4 N}{K}\max_{1\leq n\leq N}\mathbb{E}\left|q_{nn}^2-t_n^2\right|\\ &\leq \frac{2\sigma_{\max}^4 N}{\rho K}\sqrt{\max_{1\leq n\leq N}\mathbb{E}(q_{nn}-t_n)^2} \quad \xrightarrow[K\to\infty]{} \quad 0\ ,\end{aligned}$$

where the last inequality follows from (2.40) together with Lemma 3.6–(2). Convergence (2.26) is established.

We now establish the system of equations (2.27). Our starting point is the identity

$$\mathbf{Q} = \mathbf{T} + \mathbf{T}(\mathbf{T}^{-1}-\mathbf{Q}^{-1})\mathbf{Q} = \mathbf{T} + \frac{\rho}{K}\mathbf{T}\,\mathrm{diag}(\mathrm{Tr}\,\widetilde{\mathbf{D}}_1\widetilde{\mathbf{T}},\ldots,\mathrm{Tr}\,\widetilde{\mathbf{D}}_N\widetilde{\mathbf{T}})\mathbf{Q} - \mathbf{T}\mathbf{Y}\mathbf{Y}^*\mathbf{Q}\ .$$

Using this identity, we develop $U_\ell = \frac{1}{K}\mathrm{Tr}\,\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{Q}$ as

$$\begin{aligned}U_\ell &= \frac{1}{K}\mathrm{Tr}\,\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T} + \frac{\rho}{K^2}\mathrm{Tr}\,\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathrm{diag}(\mathrm{Tr}\,\widetilde{\mathbf{D}}_1\widetilde{\mathbf{T}},\ldots,\mathrm{Tr}\,\widetilde{\mathbf{D}}_N\widetilde{\mathbf{T}})\mathbf{Q} - \frac{1}{K}\mathrm{Tr}\,\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{Y}\mathbf{Y}^*\mathbf{Q}\\ &\stackrel{\triangle}{=} X_1+X_2-X_3\ .\end{aligned}\tag{2.41}$$

Lemma 3.4–(2) with $\mathbf{S}=\mathbf{D}_0\mathbf{D}_\ell\mathbf{T}$ yields:

$$X_1 = \frac{1}{K}\mathrm{Tr}\,\mathbf{D}_0\mathbf{D}_\ell\mathbf{T}^2+\epsilon_1 \tag{2.42}$$

where $\mathbb{E}|\epsilon_1|\leq\sqrt{\mathbb{E}\epsilon_1^2}\leq C/K$. Consider now the term $X_3=\frac{1}{K}\sum_{k=1}^K\mathrm{Tr}\,\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}$. Using (2.37) and (2.38), we have

$$\mathbf{y}_k^*\mathbf{Q} = \left(1-\frac{\mathbf{y}_k^*\mathbf{Q}\mathbf{y}_k}{1+\mathbf{y}_k^*\mathbf{Q}\mathbf{y}_k}\right)\mathbf{y}_k^*\mathbf{Q}_k = \rho\,\tilde{q}_{kk}\,\mathbf{y}_k^*\mathbf{Q}_k\ .$$

Hence

$$\begin{aligned}X_3 &= \frac{\rho}{K}\sum_{k=1}^K\tilde{q}_{kk}\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k\\ &= \frac{\rho}{K}\sum_{k=1}^K\tilde{t}_k\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k + \frac{\rho}{K}\sum_{k=1}^K(\tilde{q}_{kk}-\tilde{t}_k)\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k\\ &\stackrel{\triangle}{=} X_3'+\epsilon_2\ .\end{aligned}\tag{2.43}$$

By Cauchy-Schwartz inequality,

$$\mathbb{E}|\epsilon_2| \le \frac{\rho}{K}\sum_{k=1}^{K}\sqrt{\mathbb{E}(\tilde{q}_{kk}-\tilde{t}_k)^2}\sqrt{\mathbb{E}(\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k)^2}\ .$$

We have $\mathbb{E}(\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k)^2 \le \sigma_{\max}^8\rho^{-6}\mathbb{E}\|\mathbf{y}_k\|^4 \le C$. Using in addition Lemma 3.6–(2), we obtain

$$\mathbb{E}|\epsilon_2| \le \frac{C}{\sqrt{K}}\ .$$

Consider X_3'. From (2.37) and (2.38), we have $\mathbf{Q} = \mathbf{Q}_k - \rho\tilde{q}_{kk}\mathbf{Q}_k\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_k$. Hence, we can develop X_3' as

$$\begin{aligned} X_3' &= \frac{\rho}{K}\sum_{k=1}^{K}\tilde{t}_k\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k - \frac{\rho^2}{K}\sum_{k=1}^{K}\tilde{t}_k\tilde{q}_{kk}\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k \\ &\triangleq X_4 + X_5\ . \end{aligned} \tag{2.44}$$

Consider X_4. Notice that $\mathbf{y}_k$ and $\mathbf{Q}_k$ are independent. Therefore, by Lemma 2.1, we obtain

$$\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k = \frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T} + \epsilon_3 = \frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T} + \epsilon_3 + \epsilon_4$$

where $\mathbb{E}\epsilon_3^2 < CK^{-1}$ by Ineq. (2.6). Applying twice Lemma 3.6–(1) to $\epsilon_4 = \frac{1}{K}(\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T} - \mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T})$ yields $|\epsilon_4| < CK^{-1}$.
Note in addition that $\sum\tilde{t}_k\mathbf{D}_k = \mathrm{diag}(\mathrm{Tr}\,\widetilde{\mathbf{D}}_1\widetilde{\mathbf{T}},\ldots,\mathrm{Tr}\,\widetilde{\mathbf{D}}_N\widetilde{\mathbf{T}})$. Thus, we obtain

$$\begin{aligned} X_4 &= \frac{\rho}{K^2}\mathrm{Tr}\left(\sum_{k=1}^{K}\tilde{t}_k\mathbf{D}_k\right)\mathbf{Q}\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T} + \epsilon_5 \\ &= X_2 + \epsilon_5\ , \end{aligned} \tag{2.45}$$

where $\epsilon_5 = \epsilon_3 + \epsilon_4$, which yields $\mathbb{E}|\epsilon_5| \le CK^{-\frac{1}{2}}$.
We now turn to X_5. First introduce the following random variable:

$$\epsilon_6 = \tilde{t}_k\tilde{q}_{kk}\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k - \tilde{t}_k\tilde{q}_{kk}\left(\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\right)\left(\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\right)$$

Then

$$\begin{aligned} |\epsilon_6| \le \frac{1}{\rho^2}\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{y}_k&\left|\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k - \frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\right| \\ &+ \frac{1}{\rho^2}\left|\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{y}_k^* - \frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\right|\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T} \end{aligned}$$

and one can prove that $\mathbb{E}|\epsilon_6| < CK^{-\frac{1}{2}}$ with help of Lemma 2.1, together with Cauchy-Schwarz inequality. In addition, we can prove with the help of Lemma 3.6 that:

$$\begin{aligned} \tilde{t}_k\tilde{q}_{kk}\left(\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\right)\left(\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\right) &= \tilde{t}_k^2\left(\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}\mathbf{D}_0\mathbf{Q}\right)\left(\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\right) + \epsilon_7 \\ &= \tilde{t}_k^2\left(\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}\mathbf{D}_0\mathbf{Q}\right)\left(\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{D}_\ell\mathbf{T}^2\right) + \epsilon_7 + \epsilon_8 \end{aligned}$$

where ϵ_7 and ϵ_8 are random variables satisfying $\mathbb{E}|\epsilon_7| < CK^{-\frac{1}{2}}$ by Lemma 3.6, and $\max_{k,\ell} \mathbb{E}|\epsilon_8| \le \max_{k,\ell} \sqrt{\mathbb{E}|\epsilon_8|^2} \le CK^{-\frac{1}{2}}$ by Lemma 3.4–(2). Using the fact that $\rho^2 \tilde{t}_k^2 = (1 + \frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{T})^{-2}$, we end up with

$$X_5 = -\frac{\rho^2}{K}\sum_{k=1}^{K} \tilde{t}_k^2 \left(\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{Q}\mathbf{D}_0\mathbf{Q}\right)\left(\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{D}_\ell\mathbf{T}^2\right) + \epsilon_9 = -\sum_{k=1}^{K} c_{\ell k}U_k + \epsilon_9 \qquad (2.46)$$

where $c_{\ell k}$ is given by (2.28), and where $\mathbb{E}|\epsilon_9| < CK^{-\frac{1}{2}}$.

Plugging Eq. (2.42)–(2.46) into (2.41), we end up with $U_\ell = \sum_{k=1}^{K} c_{\ell k}U_k + \frac{1}{K}\mathrm{Tr}\,\mathbf{D}_0\mathbf{D}_\ell\mathbf{T}^2 + \epsilon$ with $\mathbb{E}|\epsilon| < CK^{-\frac{1}{2}}$. Step 3 is established.

Step 4 : Proof of (2.29) Define the following $(K+1)\times 1$ vectors:

$$\mathbf{u} = [U_k]_{k=0}^K, \ \mathbf{d} = \left[\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_0\mathbf{D}_k\mathbf{T}^2\right]_{k=0}^K, \ \epsilon = [\epsilon_k]_{k=0}^K ,$$

where the U_k's and ϵ_k's are defined in (2.27). Recall the definition of the $c_{\ell k}$'s for $0 \le \ell \le K$ and $1 \le k \le K$, define $c_{\ell 0} = 0$ for $0 \le \ell \le K$ and consider the $(K+1)\times(K+1)$ matrix $\mathbf{C} = [c_{\ell k}]_{\ell,k=0}^K$.

With these notations, System (2.27) writes

$$(\mathbf{I}_{K+1} - \mathbf{C})\,\mathbf{u} = \mathbf{d} + \epsilon\ . \qquad (2.47)$$

Let $\alpha = \frac{1}{K}\mathrm{Tr}\ \mathbf{D}_0^2\mathbf{T}^2$ and $\beta = (1 + \frac{1}{K}\mathrm{Tr}\ \mathbf{D}_0\mathbf{T})^2$. We have in particular

$$\mathbf{d} = \begin{bmatrix} \alpha \\ \mathbf{g} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & \frac{1}{K}\mathbf{g}^T\boldsymbol{\Delta}^{-1} \\ \mathbf{0} & \mathbf{A}^T \end{bmatrix}$$

(recall that $\mathbf{A}$, $\boldsymbol{\Delta}$ and $\mathbf{g}$ are defined in the statement of Theorem 4).
Consider a square matrix $\mathbf{X}$ which first column is equal to $[1, 0, \dots, 0]^t$, and partition $\mathbf{X}$ as $\mathbf{X} = \begin{bmatrix} 1 & \mathbf{x}_{01}^T \\ \mathbf{0} & \mathbf{X}_{11} \end{bmatrix}$. Recall that the inverse of $\mathbf{X}$ exists if and only if $\mathbf{X}_{11}^{-1}$ exists, and in this case the first row $[\mathbf{X}^{-1}]_0$ of $\mathbf{X}^{-1}$ is given by

$$\left[\mathbf{X}^{-1}\right]_0 = \left[1 \quad -\mathbf{x}_{01}^T\mathbf{X}_{11}^{-1}\right]$$

(see for instance [39]). We now apply these results to the system (2.47). Due to (2.47), U_0 can be expressed as

$$U_0 = [(\mathbf{I} - \mathbf{C})^{-1}]_0(\mathbf{d} + \epsilon)\ .$$

By Lemma 3.5–(1), $(\mathbf{I}_K - \mathbf{A}^T)^{-1}$ exists hence $(\mathbf{I} - \mathbf{C})^{-1}$ exists,

$$\left[(\mathbf{I}_{K+1} - \mathbf{C})^{-1}\right]_0 = \left[1 \quad \frac{1}{K}\mathbf{g}^T\boldsymbol{\Delta}^{-1}(\mathbf{I}_K - \mathbf{A}^T)^{-1}\right] ,$$

and

$$U_0 = \alpha + \frac{1}{K}\mathbf{g}^T\boldsymbol{\Delta}^{-1}\left(\mathbf{I} - \mathbf{A}^T\right)^{-1}\mathbf{g} + \epsilon_0 + \frac{1}{K}\mathbf{g}^T\boldsymbol{\Delta}^{-1}\left(\mathbf{I} - \mathbf{A}^T\right)^{-1}\epsilon'$$

with $\epsilon' = [\epsilon_1, \dots, \epsilon_K]^T$. Gathering the estimates of the previous part together with the fact that $\|\mathbb{E}\epsilon\|_\infty \le CK^{-\frac{1}{2}}$, we get (2.29). Step 4 is established, so is Theorem 4.

Separable case

In the separable case, $\theta_K^2 = \tilde{d}_0^2 \Omega_K^2$ where Ω_K^2 is given by the following corollary.

Corollary 1 *Assume that* **A2** *is satisfied and that* $\sigma_{nk}^2 = d_n \tilde{d}_k$. *Assume moreover that*

$$\min\left(\liminf_K \frac{1}{K}\mathrm{Tr}\,(\mathbf{D}(K)), \liminf_K \frac{1}{K}\mathrm{Tr}\,(\widetilde{\mathbf{D}}(K))\right) > 0 \tag{2.48}$$

where $\mathbf{D}$ *and* $\widetilde{\mathbf{D}}$ *are given by* (2.9). *Let* $\gamma = \frac{1}{K}\mathrm{Tr}\,\mathbf{D}^2\mathbf{T}^2$ *and* $\tilde{\gamma} = \frac{1}{K}\mathrm{Tr}\,\widetilde{\mathbf{D}}^2\widetilde{\mathbf{T}}^2$. *Then the sequence*

$$\Omega_K^2 = \gamma\left(\frac{\rho^2\gamma\tilde{\gamma}}{1-\rho^2\gamma\tilde{\gamma}} + \left(\mathbb{E}|W_{10}|^4 - 1\right)\right) \tag{2.49}$$

satisfies $0 < \liminf_K \Omega_K^2 \le \limsup_K \Omega_K^2 < \infty$. *If, in addition,* **A1** *holds true, then:*

$$\frac{\sqrt{K}}{\Omega_K}\left(\frac{\beta_K}{\tilde{d}_0} - \delta_K\right) \xrightarrow[K\to\infty]{} \mathcal{N}(0,1)$$

in distribution.

Remark 4 *Condition* (2.48) *is the counterpart of Assumption* **A3** *in the case of a separable variance profile and suffices to establish* $0 < \liminf_K(1-\rho^2\gamma\tilde{\gamma}) \le \limsup_K(1-\rho^2\gamma\tilde{\gamma}) < 1$ *(see for instance [36]), hence the fact that* $0 < \liminf_K \Omega_K^2 \le \limsup_K \Omega_K^2 < \infty$. *The remainder of the proof of Corollary 1 is postponed to Appendix 3.3.*

Recall that in the separable case, $\mathbf{D}_k = \tilde{d}_k\mathbf{D}$ and $\widetilde{\mathbf{D}}_n = d_n\widetilde{\mathbf{D}}$. Let $\tilde{\mathbf{d}}$ be the $K\times 1$ vector $\tilde{\mathbf{d}} = [\tilde{d}_k]_{k=1}^K$. In the separable case, Eq. (2.17) is written

$$\frac{\theta^2}{\tilde{d}_0^2} = \frac{1}{K\tilde{d}_0^2}\mathbf{g}^T(\mathbf{I}-\mathbf{A})^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} + \gamma(\mathbb{E}|W_{10}|^4 - 1)\ , \tag{2.50}$$

where γ is defined in statement of the corollary. Here, vector $\mathbf{g}$ and matrix $\mathbf{A}$ are given by

$$\mathbf{g} = \gamma\tilde{d}_0\tilde{\mathbf{d}} \quad \text{and} \quad \mathbf{A} = \left[\frac{1}{K}\frac{\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_\ell\mathbf{D}_m\mathbf{T}^2}{\left(1+\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_\ell\mathbf{T}\right)^2}\right]_{\ell,m=1}^K = \frac{\gamma}{K}\boldsymbol{\Delta}^{-1}\tilde{\mathbf{d}}\tilde{\mathbf{d}}^T\ .$$

By the matrix inversion lemma [39], we have

$$\begin{aligned}
\frac{1}{K\tilde{d}_0^2}\mathbf{g}^T(\mathbf{I}-\mathbf{A})^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} &= \frac{\gamma^2}{K}\tilde{\mathbf{d}}^T\left(\boldsymbol{\Delta} - \frac{\gamma}{K}\tilde{\mathbf{d}}\tilde{\mathbf{d}}^T\right)^{-1}\tilde{\mathbf{d}} \\
&= \frac{\gamma^2}{K}\tilde{\mathbf{d}}^T\left(\boldsymbol{\Delta}^{-1} + \frac{\gamma}{K}\frac{1}{1-\frac{\gamma}{K}\tilde{\mathbf{d}}^T\boldsymbol{\Delta}^{-1}\tilde{\mathbf{d}}}\boldsymbol{\Delta}^{-1}\tilde{\mathbf{d}}\tilde{\mathbf{d}}^T\boldsymbol{\Delta}^{-1}\right)\tilde{\mathbf{d}}\ .
\end{aligned}$$

Noticing that

$$\frac{1}{K}\tilde{\mathbf{d}}^T\boldsymbol{\Delta}^{-1}\tilde{\mathbf{d}} = \frac{1}{K}\sum_{k=1}^K\frac{\tilde{d}_k^2}{\left(1+\frac{1}{K}\mathrm{Tr}\,\mathbf{D}_k\mathbf{T}\right)^2} = \frac{\rho^2}{K}\sum_{k=1}^K\tilde{d}_k^2\tilde{t}_k^2 = \rho^2\tilde{\gamma}\ ,$$

we obtain

$$\frac{1}{K\tilde{d}_0^2}\mathbf{g}^T(\mathbf{I}-\mathbf{A})^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} = \gamma\frac{\rho^2\gamma\tilde{\gamma}}{1-\rho^2\gamma\tilde{\gamma}}\ .$$

Plugging this equation into (2.50), we obtain (2.49).

4 Applicative Contexts and Simulations

4.1 Applicative contexts.

The aim of this part is to present some applicative contexts in the field of wireless communications where the channel is described by the models studied in this work.

- Multiple antenna transmissions with $K+1$ distant sources sending their signals toward an array of N antennas. The corresponding transmission model is $\mathbf{r} = \mathbf{\Xi}\mathbf{s} + \mathbf{n}$ where $\mathbf{\Xi} = \frac{1}{\sqrt{K}}\mathbf{H}\mathbf{P}^{1/2}$, matrix $\mathbf{H}$ is a $N \times (K+1)$ random matrix with complex Gaussian elements representing the radio channel, $\mathbf{P} = \mathrm{diag}(p_0, \ldots, p_K)$ is the (deterministic) matrix of the powers given to the different sources, and $\mathbf{n}$ is the usual AWGN satisfying $\mathbb{E}\mathbf{n}\mathbf{n}^* = \rho \mathbf{I}_N$. Write $\mathbf{H} = [\mathbf{h}_0 \cdots \mathbf{h}_K]$, and assume that the columns $\mathbf{h}_k$ are independent, which is realistic when the sources are distant one from another. Let $\mathbf{C}_k$ be the covariance matrix $\mathbf{C}_k = \mathbb{E}\mathbf{h}_k\mathbf{h}_k^*$ and let $\mathbf{C}_k = \mathbf{U}_k\mathbf{\Lambda}_k\mathbf{U}_k$ be a spectral decomposition of $\mathbf{C}_k$ where $\mathbf{\Lambda}_k = \mathrm{diag}(\lambda_{nk};\ 1 \le n \le N)$ is the matrix of eigenvalues. Assume now that the eigenvector matrices $\mathbf{U}_0, \ldots, \mathbf{U}_K$ are all equal (to some matrix $\mathbf{U}$, for instance), a case considered in e.g., [48] (note that sometimes they are all identified with the Fourier $N \times N$ matrix [62]). Let $\mathbf{\Sigma} = \mathbf{U}^*\mathbf{\Xi}$. Then matrix $\mathbf{\Sigma}$ is described by the statistical model (2.3) where the W_{nk} are standard Gaussian i.i.d., and $\sigma_{nk}^2 = \lambda_{nk}p_k$. If we partition $\mathbf{\Xi}$ as $\mathbf{\Xi} = [\mathbf{x}\ \mathbf{X}]$ similarly to the partition $\mathbf{\Sigma} = [\mathbf{y}\ \mathbf{Y}]$ above, then the SINR β at the output of the LMMSE estimator for the first element of vector $\mathbf{s}$ in the transmission model $\mathbf{r} = \mathbf{\Xi}\mathbf{s} + \mathbf{n}$ is

  $$\beta = \mathbf{x}^* \left(\mathbf{X}\mathbf{X}^* + \rho \mathbf{I}_N\right)^{-1} \mathbf{x} = \mathbf{y}^* \left(\mathbf{Y}\mathbf{Y}^* + \rho \mathbf{I}_N\right)^{-1} \mathbf{y}$$

  due to the fact that $\mathbf{U}$ is a unitary matrix. Therefore, the problem of LMMSE SINR convergence for this MIMO model is a particular case of the general problem of convergence of the right-hand member of (2.2) for model (2.3).

  It is also worth to say a few words about the separable case in this context. If we assume that $\mathbf{\Lambda}_0 = \cdots = \mathbf{\Lambda}_K$ and these matrices are equal to $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$, then the model for $\mathbf{H}$ is the well-known Kronecker model with correlations at reception [64]. In this case,

  $$\mathbf{\Sigma} = \mathbf{U}^*\mathbf{\Xi} = \frac{1}{\sqrt{K}}\mathbf{U}^*\mathbf{H}\mathbf{P}^{1/2} = \frac{1}{\sqrt{K}}\mathbf{\Lambda}^{1/2}\mathbf{W}\mathbf{P}^{1/2} \tag{2.51}$$

  where $\mathbf{W}$ is a random matrix with iid standard Gaussian elements. This model coincides with the separable variance profile model with $d_n = \lambda_n$ and $\tilde{d}_k = p_k$.

- CDMA transmissions on flat fading channels. Here N is the spreading factor, $K+1$ is the number of users, and

  $$\mathbf{\Sigma} = \mathbf{V}\mathbf{P}^{1/2} \tag{2.52}$$

  where $\mathbf{V}$ is the $N \times (K+1)$ signature matrix assumed here to have random i.i.d. elements with mean zero and variance N^{-1}, and where $\mathbf{P} = \mathrm{diag}(p_0, \ldots, p_K)$ is the users powers matrix. In this case, the variance profile is separable with $d_n = 1$ and $\tilde{d}_k = \frac{K}{N}p_k$. Note that elements of $\mathbf{V}$ are not Gaussian in general.

- Cellular MC-CDMA transmissions on frequency selective channels. In the uplink direction, the matrix $\mathbf{\Sigma}$ is written as:

$$\mathbf{\Sigma} = [\mathbf{H}_0\mathbf{v}_0 \ \cdots \ \mathbf{H}_{K+1}\mathbf{v}_{K+1}] \ , \tag{2.53}$$

where $\mathbf{H}_k = \mathrm{diag}(h_k(\exp(2\imath\pi(n-1)/N); \ 1 \le n \le N)$ is the radio channel matrix of user k ($\imath = \sqrt{-1}$) in the discrete Fourier domain (here N is the number of frequency bins) and $\mathbf{V} = [\mathbf{v}_0, \cdots, \mathbf{v}_K]$ is the $N \times (K+1)$ signature matrix with i.i.d. elements as in the CDMA case above. Modeling this time the channel transfer functions as deterministic functions, we have $\sigma^2_{nk} = \frac{K}{N}|h_k(\exp(2\imath\pi(n-1)/N))|^2$.
In the downlink direction, we have

$$\mathbf{\Sigma} = \mathbf{HVP}^{1/2} \tag{2.54}$$

where $\mathbf{H} = \mathrm{diag}(h(\exp(2\imath\pi(n-1)/N); \ 1 \le n \le N)$ is the radio channel matrix in the discrete Fourier domain, the $N \times (K+1)$ signature matrix $\mathbf{V}$ is as above, and $\mathbf{P} = \mathrm{diag}(p_0, \ldots, p_K)$ is the matrix of the powers given to the different users. Model (2.54) coincides with the separable variance profile model with $d_n = \frac{K}{N}|h(\exp(2\imath\pi(n-1)/N))|^2$ and $d_k = p_k$.

4.2 Simulations and numerical results

The general (non necessarily separable) case

In this part, the accuracy of the Gaussian approximation is verified by simulation. In order to validate the results of Theorems 2 and 4 for practical values of K, we consider the example of a MC-CDMA transmission in the uplink direction. We recall that K is the number of interfering users in this context. In the simulation, the discrete time channel impulse response of user k is represented by the vector with $L = 5$ coefficients $\mathbf{g}_k = [g_{k,0}, \ldots, g_{k,L-1}]^t$. In the simulations, these vectors are generated pseudo-randomly according to the complex multivariate Gaussian law $\mathcal{CN}(0, 1/L\mathbf{I}_L)$. Setting the number of frequency bins to N, the channel matrix $\mathbf{H}_k$ for user k in the frequency domain (see Eq. (2.53)) is $\mathbf{H}_k = \mathrm{diag}(h_k(\exp(2\imath\pi(n-1)/N); 1 \le n \le N)$ where $h_k(z) = \frac{\sqrt{P_k}}{\|\mathbf{g}_k\|}\sum_{l=0}^{L-1} g_{k,l}z^{-l}$, the norm $\|\mathbf{g}_k\|$ is the Euclidean norm of $\mathbf{g}_k$ and P_k is the power received from user k. Concerning the distribution of the user powers P_k, we assume that these are arranged into five power classes with powers $P, 2P, 4P, 8P$ and $16P$ with relative frequencies given by Table 5. The user of interest (User 0) is assumed to belong to Class 1. Finally, we assume that

Table 2.1: Power classes and relative frequencies

Class	1	2	3	4	5
Power	P	$2P$	$4P$	$8P$	$16P$
Relative frequency	1/8	1/4	1/4	1/8	1/4

the number K of interfering users is set to $K = N/2$.
In table 4.2, we present the corresponding values of the SINR normalized MSE with respect

to K. This table proves that the SINR normalized MSE $K\mathbb{E}\left(\beta_K - \overline{\beta}_K\right)^2/\theta_K^2$ is closed to one for values of K as small as $K = 8$. We precise here that the SNR $\frac{P}{\rho}$ for the user of interest is fixed to 10dB. In the table 4.2, we fixe $K = 64$ and we study the evolution of

Table 2.2: SINR normalized MSE vs K (SNR $= 10\ dB$)

K	8	16	32	64	128	256
$K\mathbb{E}(\beta_K - \overline{\beta}_K)^2/\theta_K^2$	0.9761	0.9845	1.0464	1.0187	1.0127	0.9919

SINR normalized MSE with respect to the input SNR $\frac{P}{\rho}$.

Figure 2.1 shows the histogram of $\sqrt{K}(\beta_K - \bar{\beta}_K)/\theta_K$ for $N = 16$ and $N = 64$. This figure

Table 2.3: SINR normalized MSE vs SNR ($K = 64$)

SNR	0	5	10	15	20	25	30
$K\mathbb{E}(\beta_K - \overline{\beta}_K)^2/\theta_K^2$	1.0283	1.0294	1.0373	1.0358	1.0347	1.0348	1.0350

gives an idea of the similarity between the distribution of $\sqrt{K}(\beta_K - \bar{\beta}_K)/\theta_K$ and $\mathcal{N}(0,1)$. More precisely, Figure 2.2 quantifies this similarity through a Quantile-Quantile plot.

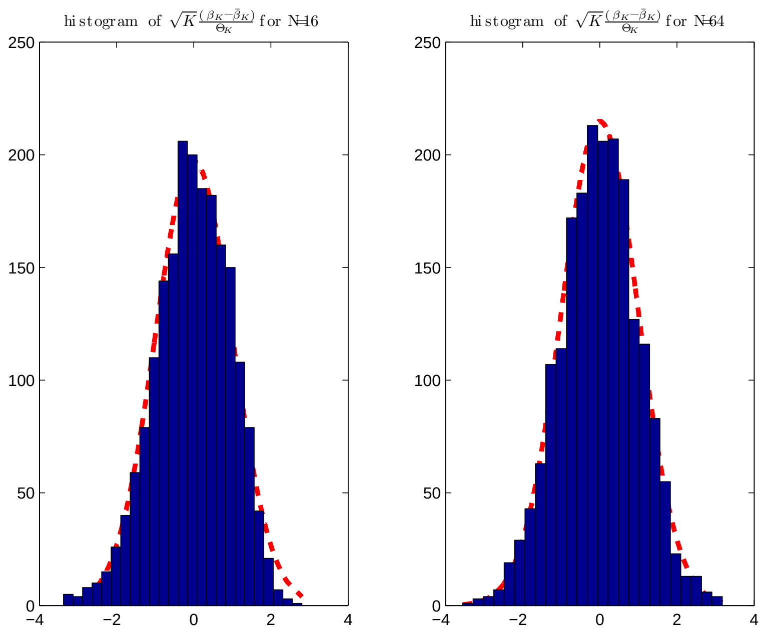


Figure 2.1: Histogram of $\sqrt{K}(\beta_K - \bar{\beta}_K)$ for $N = 16$ and $N = 64$.

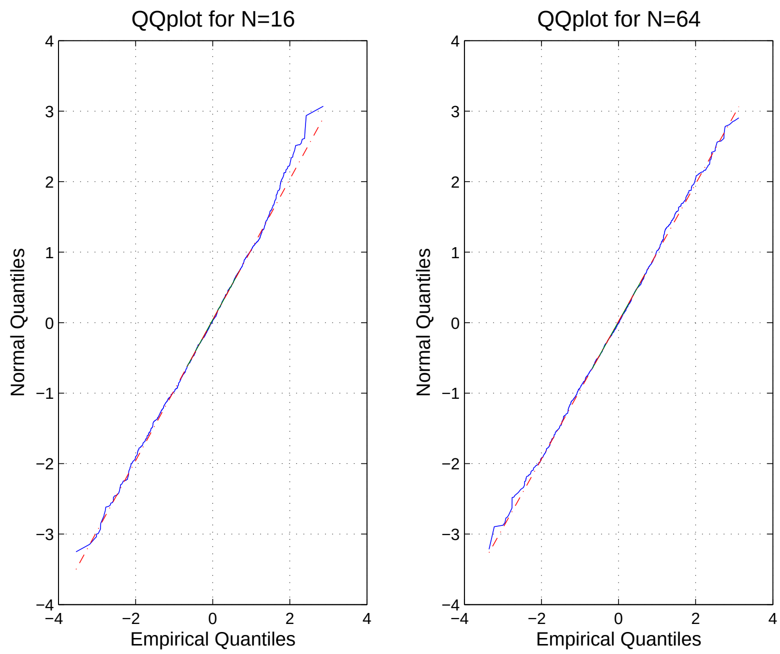


Figure 2.2: Q-Q plot for $\sqrt{K}(\beta_K - \bar{\beta}_K)$, $N = 16$ and $N = 64$; dash doted line is the 45 degree line.

The separable case

In order to test the results of Proposition 4 and Corollary 1, we consider the following multiple antenna (MIMO) model with exponentially decaying correlation at reception:

$$\mathbf{\Sigma} = \frac{1}{\sqrt{K}}\mathbf{\Psi}^{1/2}\mathbf{W}\mathbf{P}^{1/2}$$

where $\mathbf{\Psi} = [a^{m-n}]_{m,n=0}^{N-1}$ with $0 < a < 1$ is the covariance matrix that accounts for the correlations at the receiver side, $P = \mathrm{diag}\,(p_0, \cdots, p_K)$ is the matrix of the powers given to the different sources and W is a $N \times (K+1)$ matrix with Gaussian standard iid elements. Let $\mathbf{P_u}$ denote the vector containing the powers of the interfering sources. We set $\mathbf{P_u}$ (up to a permutation of its elements) to:

$$\mathbf{P_u} = \begin{cases} [4P \;\; 5P] \quad \text{if} \quad K = 2 \\ [P \;\; P \;\; 2P \;\; 4P] \quad \text{if} \quad K = 4 \\ [P \;\; P \;\; 2P \;\; 2P \;\; 2P \;\; 4P \;\; 4P \;\; 4P \;\; 8P \;\; 16P \;\; 16P \;\; 16P\,] \text{ if } K = 12\,. \end{cases}$$

For $K = 2^p$ with $3 \leq p \leq 7$, we assume that the powers of the interfering sources are arranged into 5 classes as in Table 5. We set the SNR P/ρ to 10 dB and a to 0.1. We investigate in this section the accuracy of the Gaussian approximation in terms of the outage probability. In Fig.2.3, we compare the empirical 1% outage SINR with the one predicted by the Central Limit Theorem. We note that the Gaussian approximation tends to under estimate the 1% outage SINR. We also note that it has a good accuracy for small values of α and for enough large values of N ($N \geq 64$).

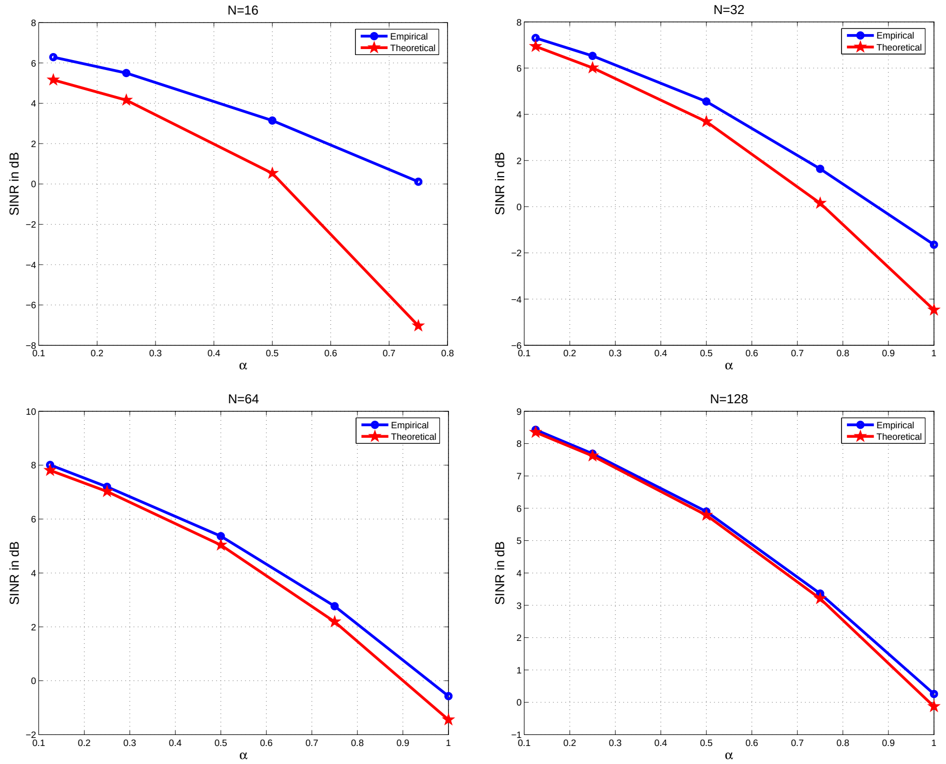


Figure 2.3: Theoretical and empirical 1% outage SINR

CHAPTER 3

Statistical Distribution of the SINR for the MMSE Receiver Correlated MIMO Channels

Contents

This chapter corresponds to the article "BER and Outage Probability Approximations for LMMSE Detectors on Correlated MIMO Channels" published in IEEE Information Theory Journal.

1 Introduction

In this chapter we study the statistical distribution of the Signal to Interference-plus-Noise Ratio for the Minimum Mean Square Error receiver in MIMO wireless communications (for systems with small dimensions). The channel model is assumed to be (receive) correlated Rayleigh with unequal powers.

We consider the following linear model:

$$\mathbf{r} = \mathbf{\Sigma}\mathbf{s} + \mathbf{n},$$

where $\mathbf{s} = [s_0, \cdots, s_K]^{\mathrm{T}}$ is the transmitted complex vector signal with size $K+1$ satisfying $\mathbb{E}\mathbf{s}\mathbf{s}^* = \mathbf{I}_{K+1}$, and $\mathbf{\Sigma}$ is the $N \times (K+1)$ channel matrix. We assume that the matrix $\mathbf{\Sigma}$ writes as

$$\mathbf{\Sigma} = \frac{1}{\sqrt{K}}\mathbf{\Psi}^{\frac{1}{2}}\mathbf{W}\mathbf{P}^{\frac{1}{2}},$$

where $\mathbf{\Psi}$ is a $N \times N$ Hermitian nonnegative receiver correlation matrix which is assumed to be nonrandom, $\mathbf{P} = \mathrm{diag}\,(p_0, \cdots, p_K)$ is the deterministic matrix of the powers allocated to the different users and $\mathbf{W} = [\mathbf{w}_0, \cdots, \mathbf{w}_K]$ ($\mathbf{w}_k$ being the kth column) is a $N \times (K+1)$ complex Gaussian matrix with centered unit variance (standard) independent and identically distributed (i.i.d) entries. The engineering goal is to estimate the transmitted symbol s_k for each user. Assume that the receiver has already acquired the knowledge of the channel matrix. For user k, the Linear Minimum Mean Squar Error LMMSE estimator generates an output in a form $\mathbf{g}^*\mathbf{r}$ where $\mathbf{g}$ minimizes the following mean-squared error

$$\mathbb{E}\left|\mathbf{g}^*\mathbf{r} - s_k\right|^2$$

Without loss of generality, the first stream is assumed (k=0). It is well known that the relevant performance measure of the LMMSE estimator is the SINR of the estimate symbol. For our model where the matrix channel $\mathbf{\Sigma}$ is given by (3.1), the SINR of the first user is given by:

$$\beta_k = \frac{p_0}{K}\mathbf{w}_0^*\mathbf{\Psi}^{1/2}\left(\frac{1}{K}\mathbf{\Psi}^{1/2}\mathbf{W}_0\mathbf{P}_0\mathbf{W}_0^*\mathbf{\Psi}^{1/2} + \rho I_N\right)^{-1}\mathbf{\Psi}^{1/2}\mathbf{w}_0$$

where $\mathbf{W}_0 = [w_1, \ldots, w_K]$ and $\mathbf{P}_0 = \mathrm{diag}\,(p_1, \ldots, p_K)$.

The study of the SINR allows also the study of another performance indices such that the Bit Error Rate (BER) and the outage probability. Based on Random Matrix Theory and on the gaussian character of the entries of the channel matrix, we derive closed-form expressions for the first three moments. Using the generalized Gamma approximation, we provide closed-form expressions for the BER and numerical approximations for the outage probability.

2 Bit Error Rate and Outage Probability approximations

2.1 Generalised Gamma distribution

Recall that if a random variable X follows a generalized gamma distribution $G(\alpha, b, \xi)$, where α and b are respectively referred to as the shape and scale parameters, then:

$$\mathbb{E}X = \alpha b, \quad \mathrm{var}(X) = \alpha b^2 \quad \text{and} \quad \mathbb{E}(X - \mathbb{E}X)^3 = (\xi + 1)\alpha b^3 .$$

The probability density function (pdf) of the generalized Gamma distribution with parameters (α, b, ξ) does not have a closed form expression but its moment generating function (MGF) writes [23]:

$$\mathrm{MGF}(s) = \begin{cases} \exp(\frac{\alpha}{\xi-1}(1-(1-b\xi s)^{\frac{\xi-1}{\xi}})) & \text{if } \xi > 1, \\ (1-sb)^{-\alpha}, \;\; s < \frac{1}{b} & \text{if } \xi = 1, \\ \exp(\frac{\alpha}{1-\xi}((1-b\xi s)^{\frac{\xi-1}{\xi}} - 1)) & \text{if } \xi > 1. \end{cases}$$

2.2 BER approximation

Under QPSK constellations with Gray encoding and assuming that the noise at the LMMSE output is Gaussian, the BER is given by:

$$\mathrm{BER} = \mathbb{E}Q(\sqrt{\beta_K})$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2}\, dt$ and the expectation is taken over the distribution of the SNR β_K. Based on the asymptotic normality of the SNR, [81] and [61] proposed to use the limiting BER value given by:

$$\mathrm{BER} = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{\overline{\beta}_K}}^{\infty} e^{-t^2/2} dt,$$

where $\overline{\beta}_K$ denotes an asymptotic deterministic approximation of the first moment of β_K. It was shown however in [49] that this expression is inaccurate since a Gaussian random variable allows negative values and has a zero third moment while the output SNR is always positive and has a non-zero third moment for finite system dimensions. To overcome these difficulties, Li *et al.* [49] approximate the BER by considering first that the SNR follows a Gamma distribution with scale α and shape b, these parameters being tuned by equating the first two moments of the Gamma distribution with the first two asymptotic moments of the SNR. However, the third asymptotic moment was shown to be different from the third moment of the Gamma distribution which only depends on the scale α and shape b. In light of this consideration, Li *et al.* [49] refine this approximation and consider that the SNR follows a generalized Gamma distribution which is adjusted by assuming that its first three moments equate the first three asymptotic moments of the SNR. As expected, this approximation has proved to be more accurate than the Gamma approximation, and so will be the one considered in this paper. Next, we briefly review this technique, which we will rely on to provide accurate approximations for the BER and outage probability.

Let $\mathbb{E}_\infty(\beta_K)$, $\mathrm{var}_\infty(\beta_K)$ and $\mathrm{S}_\infty(\beta_K)$ denote respectively the deterministic approximations of the asymptotic central moments of β_K. Then, the parameters ξ, α and b are determined by solving:

$$\mathbb{E}_\infty(\beta_K) = \alpha b, \quad \mathrm{var}_\infty(\beta_K) = \alpha b^2 \quad \text{and} \quad \mathrm{S}_\infty(\beta_K) = (\xi+1)\alpha b^3,$$

thus giving the following values:

$$\alpha = \frac{(\mathbb{E}_\infty(\beta_K))^2}{\mathrm{var}_\infty(\beta_K)}, \quad \beta = \frac{\mathrm{var}_\infty(\beta_K)}{\mathbb{E}_\infty(\beta_K)} \quad \text{and} \quad \xi = \frac{\mathrm{S}_\infty(\beta_K)\mathbb{E}_\infty(\beta_K)}{(\mathrm{var}_\infty(\beta_K))^2} - 1.$$

Using the MGF, one can evaluate the BER by using the following relation [73], that holds for QPSK constellation:

$$\mathrm{BER} = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \mathrm{MGF}\left(-\frac{1}{2\sin^2\phi}\right) d\phi. \tag{3.1}$$

Note that similar expressions for the BER exist for other constellations and can be derived by plugging the following identity involving the function $Q(x)$ [73]:

$$Q(x) = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \exp\left(-\frac{x^2}{2\sin^2\theta}\right) d\theta$$

into the BER expression.

2.3 Outage probability approximation

Only the moment generation function (MGF) has a closed form expression. Knowing the MGF, one can compute numerically the cumulative distribution function by applying the saddle point approximation technique [12]. Denote by $K(y) = \log(\mathrm{MGF}(y))$ the cumulative generating function, by y the threshold SNR and by t_y the solution of $K'(t_y) = y$. Let w_0 and u_0 be given by: $w_0 = \mathrm{sign}(t_y)\sqrt{2\left(t_y y - K(t_y)\right)}$ and $u_0 = t_y\sqrt{K''(t_y)}$. The saddle point approximate of the outage probability is given by:

$$\mathrm{P}_{out} = \Phi(w_0) + \phi(w_0)\left(\frac{1}{w_0} - \frac{1}{u_0}\right), \tag{3.2}$$

where $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\,dt$ and $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ denote respectively the standard normal cumulative distribution function and probability distribution function.

So far, we have presented the technique that will be used in simulations for the evaluation of the BER and outage probability. This technique is heavily based on the computation of the three first asymptotic moments of the SNR β_K, an issue that is handled in the next section.

3 Asymptotic moments

3.1 Assumptions

In the following, we assume that both K and N go to $+\infty$, their ratio being bounded below and above as follows:

$$0 \ < \ \ell^- = \liminf \frac{K}{N} \ \leq \ \ell^+ = \limsup \frac{K}{N} \ < \ +\infty \ .$$

In the sequel, the notation $K \to \infty$ will refer to this asymptotic regime. Recall the expression of the SINR given by (3.3).

$$\beta_k \ = \ \frac{p_0}{K}\mathbf{w}_0^*\mathbf{\Psi}^{1/2}\left(\frac{1}{K}\mathbf{\Psi}^{1/2}\mathbf{W}_0\mathbf{P}_0\mathbf{W}_0^*\mathbf{\Psi}^{1/2} + \rho I_N\right)^{-1}\mathbf{\Psi}^{1/2}\mathbf{w}_0$$

Let $\mathbf{\Psi} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ be a spectral decomposition of $\mathbf{\Psi}$. Then, β_K writes:

$$\begin{aligned}
\beta_K \ &= \ \frac{p_0}{K}\mathbf{w}_0^*\mathbf{U}\mathbf{D}^{\frac{1}{2}}\left(\frac{1}{K}\mathbf{D}^{\frac{1}{2}}\mathbf{U}^*\widetilde{\mathbf{W}}\widetilde{\mathbf{P}}\widetilde{\mathbf{W}}^*\mathbf{U}\mathbf{D}^{\frac{1}{2}} + \rho\mathbf{I}_N\right)^{-1}\mathbf{D}^{\frac{1}{2}}\mathbf{U}^*\mathbf{w}_0 \ , \\
&= \ \frac{p_0}{K}\mathbf{z}^*\mathbf{D}^{\frac{1}{2}}\left(\frac{1}{K}\mathbf{D}^{\frac{1}{2}}\mathbf{Z}\tilde{\mathbf{D}}\mathbf{Z}^*\mathbf{D}^{\frac{1}{2}} + \rho\mathbf{I}\right)^{-1}\mathbf{D}^{\frac{1}{2}}\mathbf{z}, \\
&= \ \frac{p_0}{\rho K}\mathbf{z}^*\mathbf{D}^{\frac{1}{2}}\left(\frac{1}{K\rho}\mathbf{D}^{\frac{1}{2}}\mathbf{Z}\tilde{\mathbf{D}}\mathbf{Z}^*\mathbf{D}^{\frac{1}{2}} + \mathbf{I}\right)^{-1}\mathbf{D}^{\frac{1}{2}}\mathbf{z}
\end{aligned}$$

where: $\mathbf{z} = \mathbf{U}^*\mathbf{w}_0$ (resp. $\mathbf{Z} = \mathbf{U}^*\widetilde{\mathbf{W}}$) is a $N \times 1$ vector with complex independent standard Gaussian entries (resp. $N \times K$ matrix with independent Gaussian entries). We will frequently write $\mathbf{D}_K$ and $\widetilde{\mathbf{D}}_K$ to emphasize the dependence in K, but may drop the subscript K as well. Assume the following mild conditions:

Assumption A-1 *There exist real numbers $d_{\max} < \infty$ and $\tilde{d}_{\max} < \infty$ such that:*

$$\sup_K \|\mathbf{D}_K\| \leq d_{\max} \quad \text{and} \quad \sup_K \|\widetilde{\mathbf{D}}_K\| \leq \tilde{d}_{\max},$$

where $\|\mathbf{D}_K\|$ and $\|\widetilde{\mathbf{D}}_K\|$ are the spectral norms of $\mathbf{D}_K$ and $\widetilde{\mathbf{D}}_K$.

Assumption A-2 *The normalized traces of $\mathbf{D}_K$ and $\widetilde{\mathbf{D}}_K$ satisfy:*

$$\inf_K \frac{1}{K}\mathrm{Tr}(\mathbf{D}_K) > 0 \quad \text{and} \quad \inf_K \frac{1}{K}\mathrm{Tr}(\widetilde{\mathbf{D}}_K) > 0.$$

3.2 Asymptotic moments computation

In this section, we provide closed form expressions for the first three asymptotic moments. We shall first introduce some deterministic quantities that are used for the computation of the first, second and third asymptotic moments.

Proposition 5 *(cf. [36]) For every integer K and any $t > 0$, the system of equations in $(\delta, \tilde{\delta})$*

$$\begin{cases} \delta_K &= \frac{1}{K}\mathrm{Tr}\mathbf{D}_K\left(\mathbf{I} + t\tilde{\delta}_K\mathbf{D_K}\right)^{-1}, \\ \tilde{\delta}_K &= \frac{1}{K}\mathrm{Tr}\widetilde{\mathbf{D}}_K\left(\mathbf{I} + t\delta_K\widetilde{\mathbf{D}}_K\right)^{-1}, \end{cases}$$

admits a unique solution $\left(\delta_K(t), \tilde{\delta}_K(t)\right)$ satisfying $\delta_K(t) > 0$, $\tilde{\delta}_K(t) > 0$.

Let $\mathbf{T}$ and $\widetilde{\mathbf{T}}$ be the $N \times N$ and $K \times K$ diagonal matrices defined by:

$$\mathbf{T} = \left(\mathbf{I} + t\tilde{\delta}_K\mathbf{D}\right)^{-1} \quad \text{and} \quad \widetilde{\mathbf{T}} = \left(\mathbf{I} + t\delta_K\widetilde{\mathbf{D}}\right)^{-1}.$$

Note that in particular: $\delta = \frac{1}{K}\mathrm{Tr}\mathbf{DT}$ and $\tilde{\delta} = \frac{1}{K}\mathrm{Tr}\widetilde{\mathbf{D}}\widetilde{\mathbf{T}}$. Define also γ and $\tilde{\gamma}$ as $\gamma = \frac{1}{K}\mathrm{Tr}\mathbf{D}^2\mathbf{T}^2$ and $\tilde{\gamma} = \frac{1}{K}\mathrm{Tr}\widetilde{\mathbf{D}}^2\widetilde{\mathbf{T}}^2$. Finally, replace t by $\frac{1}{\rho}$ and introduce the following deterministic quantities:

$$\begin{aligned} \Omega_K^2 &= \frac{\gamma}{\rho^2}\left(\frac{\gamma\tilde{\gamma}}{\rho^2 - \gamma\tilde{\gamma}} + 1\right), \\ \nu_K &= \frac{2\rho^3}{K\left(\rho^2 - \gamma\tilde{\gamma}\right)^3}\left[\mathrm{Tr}\mathbf{D}^3\mathbf{T}^3 - \frac{\gamma^3}{\rho^3}\mathrm{Tr}\widetilde{\mathbf{D}}^3\widetilde{\mathbf{T}}^3\right]. \end{aligned}$$

As usual, the notation $\alpha_K = \mathcal{O}(\beta_K)$ means that $\alpha_K(\beta_K)^{-1}$ is uniformly bounded as $K \to \infty$. Then, the first three asymptotic moments are given by the following theorem:

Theorem 5 *Assuming that the matrices $\mathbf{D}$ and $\widetilde{\mathbf{D}}$ satisfy the conditions stated in 1 and 2, then the following convergences hold true:*

1. *First asymptotic moment [43, 45]:*

$$\frac{\delta_K}{\rho} = \mathcal{O}(1) \quad \text{and} \quad \mathbb{E}\left(\frac{\beta_K}{p_0}\right) - \frac{\delta_K}{\rho} \xrightarrow[K\to\infty]{} 0,$$

2. *Second asymptotic moment [43, 45]:*

$$\Omega_K = \mathcal{O}(1) \quad and \quad K\mathbb{E}\left(\frac{\beta_K}{p_0} - \mathbb{E}\left(\frac{\beta_K}{p_0}\right)\right)^2 - \Omega_K^2 \xrightarrow[K\to\infty]{} 0,$$

3. *Third asymptotic moment:*

$$\nu_K = \mathcal{O}(1) \quad and \quad K^2\mathbb{E}\left(\frac{\beta_K}{p_0} - \mathbb{E}\left(\frac{\beta_K}{p_0}\right)\right)^3 - \nu_K \xrightarrow[K\to\infty]{} 0.$$

The two first items of the theorem are proved in [45]

4 Proof of the main theorem

In the sequel, we shall heavily rely on the results and techniques developed in [36]. In the sequel, $\mathbf{D}$ and $\widetilde{\mathbf{D}}$ are respectively $N \times N$ and $K \times K$ diagonal matrices which satisfy 1 and 2, $\mathbf{Z}$ is a $N \times K$ matrix whose entries are i.i.d. standard complex Gaussian, $\mathbf{X}$ is a $N \times K$ matrix defined by:

$$\mathbf{X} = \mathbf{D}^{\frac{1}{2}}\mathbf{Z}\widetilde{\mathbf{D}}^{\frac{1}{2}} .$$

We shall often write $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_K]$ where the $\mathbf{x}_j$'s are $\mathbf{X}$'s columns. We recall hereafter the mathematical tools that will be of constant use in the sequel.

4.1 Notations

Define the resolvant matrix $\mathbf{H}$ by:

$$\mathbf{H} = \left(\frac{t}{K}\mathbf{D}^{\frac{1}{2}}\mathbf{Z}\widetilde{\mathbf{D}}\mathbf{Z}^*\mathbf{D}^{\frac{1}{2}} + \mathbf{I}_N\right)^{-1} = \left(\frac{t}{K}\mathbf{X}\mathbf{X}^* + \mathbf{I}_N\right)^{-1} .$$

We introduce the following intermediate quantities:

$$\beta(t) = \frac{1}{K}\mathrm{Tr}(\mathbf{D}\mathbf{H}), \quad \alpha(t) = \frac{1}{K}\mathrm{Tr}(\mathbf{D}\mathbb{E}\mathbf{H}) \quad \text{and} \quad \overset{o}{\beta} = \beta - \alpha .$$

Matrix $\widetilde{\mathbf{R}}(t) = \mathrm{diag}\,(\tilde{r}_1, \cdots, \tilde{r}_K)$ is a $K \times K$ diagonal matrix defined by:

$$\widetilde{\mathbf{R}}(t) = \left(\mathbf{I} + t\alpha(t)\widetilde{\mathbf{D}}_K\right)^{-1} .$$

Let $\tilde{\alpha} = \frac{1}{K}\mathrm{Tr}(\widetilde{\mathbf{D}}\widetilde{\mathbf{R}})$. Then, matrix $\mathbf{R}(t) = \mathrm{diag}\,(r_1, \cdots, r_N)$ is a $N \times N$ matrix defined by:

$$\mathbf{R}(t) = (\mathbf{I} + t\tilde{\alpha}(t)\mathbf{D})^{-1} .$$

4.2 Mathematical Tools

The results below, of constant use in the proof of Theorem 5, can be found in [36].

4.2.1 Differentiation formulas

$$\frac{\partial H_{pq}}{\partial X_{ij}} = -\frac{t}{K}\left[\mathbf{X}^*\mathbf{H}\right]_{jq} H_{pi} = -\frac{t}{K}\left[\mathbf{x}_j^*\mathbf{H}\right]_q H_{pi}. \quad (3.3)$$

$$\frac{\partial H_{pq}}{\partial \overline{X_{ij}}} = -\frac{t}{K}\left[\mathbf{H}\mathbf{X}\right]_{pj} H_{iq} = -\frac{t}{K}\left[\mathbf{H}\mathbf{x}_j\right]_p H_{iq} \quad (3.4)$$

4.2.2 Integration by parts formula for Gaussian functionals

Let Φ be a $\mathcal{C}^1$ complex function polynomially bounded together with its derivatives, then:

$$\mathbb{E}\left[X_{ij}\Phi(\mathbf{X})\right] = d_i\tilde{d}_j\mathbb{E}\left(\frac{\partial\Phi(\mathbf{X})}{\partial\overline{X_{ij}}}\right). \quad (3.5)$$

4.2.3 Poincaré-Nash inequality

Let $\mathbf{X}$ and Φ be as above, then:

$$\mathrm{Var}(\Phi(\mathbf{X})) \leq \sum_{i=1}^{N}\sum_{j=1}^{K} d_i\tilde{d}_j\mathbb{E}\left[\left|\frac{\partial\Phi(\mathbf{X})}{\partial X_{ij}}\right|^2 + \left|\frac{\partial\Phi(\mathbf{X})}{\partial\overline{X_{ij}}}\right|^2\right]. \quad (3.6)$$

4.2.4 Deterministic approximations and various estimations

Proposition 6 *(cf. [36]) Let $(\mathbf{A}_K)$ and $(\mathbf{B}_K)$ be two sequences of respectively $N\times N$ and $K\times K$ diagonal deterministic matrices whose spectral norm are uniformly bounded in K, then the following hold true:*

$$\frac{1}{K}\mathrm{Tr}(\mathbf{A}\mathbf{R}) = \frac{1}{K}\mathrm{Tr}(\mathbf{A}\mathbf{T}) + \mathcal{O}(K^{-2}), \qquad \frac{1}{K}\mathrm{Tr}(\mathbf{B}\widetilde{\mathbf{R}}) = \frac{1}{K}\mathrm{Tr}(\mathbf{B}\widetilde{\mathbf{T}}) + \mathcal{O}(K^{-2}).$$

Proposition 7 *(cf. [36]) Let $(\mathbf{A}_K)$, $(\mathbf{B}_K)$ and $(\mathbf{C}_K)$ be three sequences of $N\times N$, $K\times K$ and $N\times N$ diagonal deterministic matrices whose spectral norm are uniformly bounded in K. Consider the following functions:*

$$\Phi(\mathbf{X}) = \frac{1}{K}\mathrm{Tr}\left(\mathbf{A}\mathbf{H}\frac{\mathbf{X}\mathbf{B}\mathbf{X}^*}{K}\right), \quad \Psi(\mathbf{X}) = \frac{1}{K}\mathrm{Tr}\left(\mathbf{A}\mathbf{H}\mathbf{D}\mathbf{H}\frac{\mathbf{X}\mathbf{B}\mathbf{X}^*}{K}\right).$$

Then,

1. *the following estimations hold true:*

$$\mathrm{var}\,\Phi(\mathbf{X}),\ \mathrm{var}\,\Psi(\mathbf{X}),\ \mathrm{var}(\beta) \quad \textit{and} \quad \mathrm{var}\left(\frac{1}{K}\mathrm{Tr}\mathbf{A}\mathbf{H}\mathbf{C}\mathbf{H}\right) \quad \textit{are} \quad \mathcal{O}(K^{-2})\,.$$

2. the following approximations hold true:

$$\mathbb{E}\left[\Phi(\mathbf{X})\right] = \frac{1}{K}\mathrm{Tr}\left(\widetilde{\mathbf{D}}\widetilde{\mathbf{T}}\mathbf{B}\right)\frac{1}{K}\mathrm{Tr}\left(\mathbf{ADT}\right) + \mathcal{O}(K^{-2}), \tag{3.7}$$

$$\mathbb{E}\left[\Psi(\mathbf{X})\right] = \frac{1}{1-t^2\gamma\tilde{\gamma}}\left(\frac{1}{K^2}\mathrm{Tr}\left(\widetilde{\mathbf{D}}\widetilde{\mathbf{T}}\mathbf{B}\right)\mathrm{Tr}(\mathbf{AD}^2\mathbf{T}^2) - \right. \tag{3.8}$$

$$\left. \frac{t\gamma}{K^2}\mathrm{Tr}\left(\widetilde{\mathbf{D}}^2\widetilde{\mathbf{T}}^2\mathbf{B}\right)\mathrm{Tr}(\mathbf{ADT})\right) + \mathcal{O}(K^{-2}), \tag{3.9}$$

$$\mathbb{E}\frac{1}{K}\mathrm{Tr}\left[\mathbf{AHDH}\right] = \frac{1}{1-t^2\gamma\tilde{\gamma}}\frac{1}{K}\mathrm{Tr}(\mathbf{ADT}^2) + \mathcal{O}(K^{-2}). \tag{3.10}$$

Proofs of Propositions 6 and 7 are essentially provided in [36]. In the same vein, the following proposition will be needed.

Proposition 8 *Let $(\mathbf{A}_K)$, $(\mathbf{B}_K)$ and $(\mathbf{C}_K)$ be three sequences of $N \times N$, $K \times K$ and $N \times N$ diagonal deterministic matrices whose spectral norm are uniformly bounded in K. Consider the following function:*

$$\varphi(\mathbf{X}) = \frac{1}{K}\mathrm{Tr}\left[\mathbf{CHAHAH}\frac{\mathbf{XBX}^*}{K}\right].$$

Then $\mathrm{var}\,\varphi(\mathbf{X}) = \mathcal{O}(K^{-2})$ *and* $\mathrm{var}\left(\frac{1}{K}\mathrm{Tr}\mathbf{AHAHAH}\right) = \mathcal{O}(K^{-2})$.

Proof 2 *The proof mainly relies on Poincaré-Nash inequality. Using the Poincaré-Nash inequality, we have:*

$$\mathrm{var}(\varphi(\mathbf{X})) \le \sum_{i=1}^{N}\sum_{j=1}^{K} d_i\tilde{d}_j\mathbb{E}\left|\frac{\partial\varphi}{\partial X_{ij}}\right|^2 + \sum_{i=1}^{N}\sum_{j=1}^{K} d_i\tilde{d}_j\mathbb{E}\left|\frac{\partial\varphi}{\partial \overline{X_{ij}}}\right|^2.$$

We only deal with the first term of the last inequality (the second term can be handled similarly). We have $\varphi(\mathbf{X}) = \frac{1}{K^2}\sum_{p,r,s,t=1}^{N}\sum_{u=1}^{K} c_{pp}H_{pr}A_{rr}H_{rs}A_{ss}H_{st}X_{tu}B_{uu}X_{pu}^*$. After straightforward calculations using the differentiation formula (3.3)*, we get that:*

$$\frac{\partial\varphi}{\partial X_{ij}} = \phi_{ij}^{(1)} + \phi_{ij}^{(2)} + \phi_{ij}^{(3)} + \phi_{ij}^{(4)},$$

where:

$$\phi_{ij}^{(1)} = -\frac{t}{K^3}\left[\mathbf{X}^*\mathbf{HAHAHXBX}^*\mathbf{CH}\right]_{ji}, \quad \phi_{ij}^{(2)} = -\frac{t}{K^3}\left[\mathbf{X}^*\mathbf{HAHXBX}^*\mathbf{CHAH}\right]_{ji},$$

$$\phi_{ij}^{(3)} = -\frac{t}{K^3}\left[\mathbf{X}^*\mathbf{HXBX}^*\mathbf{CHAHAH}\right]_{ji}, \quad \phi_{ij}^{(4)} = \frac{1}{K^2}\left[\mathbf{BX}^*\mathbf{CHAHAH}\right]_{ji}.$$

Hence, $\left|\frac{\partial\varphi}{\partial X_{ij}}\right|^2 \leq 4\left(\left|\phi_{ij}^{(1)}\right|^2 + \left|\phi_{ij}^{(2)}\right|^2 + \left|\phi_{ij}^{(3)}\right|^2 + \left|\phi_{ij}^{(4)}\right|^2\right)$ *and*

$$\begin{aligned}\sum_{i=1}^{N}\sum_{j=1}^{K} d_i\tilde{d}_j\mathbb{E}\left[\left|\frac{\partial\varphi}{\partial X_{ij}}\right|^2\right] \leq\; & \frac{4t^2}{K^6}\mathbb{E}\mathrm{Tr}\left(\mathbf{DHCXBX^*HAHAHX\widetilde{D}X^*HAHAHXBX^*CH}\right)\\ & +\frac{4t^2}{K^6}\mathbb{E}\mathrm{Tr}\left(\mathbf{DHAHCXBX^*HAHX\widetilde{D}X^*HAHXBX^*CHAH}\right)\\ & +\frac{4t^2}{K^6}\mathbb{E}\mathrm{Tr}\left(\mathbf{DHAHAHCXBX^*HX\widetilde{D}X^*HXBX^*CHAHAH}\right)\\ & +\frac{4}{K^4}\mathbb{E}\mathrm{Tr}\left(\mathbf{DHAHAHCXB\widetilde{D}BX^*CHAHAH}\right).\end{aligned}$$

We only prove that the first term of the right hand side is of order K^{-2}; the other terms being handled similarly. Using Cauchy-Schwartz inequality, we get:

$$\begin{aligned}4\sum_{i=1}^{N}\sum_{j=1}^{K} d_i\tilde{d}_j\mathbb{E}\left|\phi_{ij}^1\right|^2 &\leq \frac{4t^2 d_{\max}\|\mathbf{H}\|^2\|\mathbf{C}\|^2}{K^6}\mathbb{E}\mathrm{Tr}\left((\mathbf{HA})^2\,\mathbf{HX\widetilde{D}X^*H}\,(\mathbf{AH})^2\,(\mathbf{XBX^*})^2\right),\\ &\leq \frac{4t^2}{K^6}d_{\max}\|\mathbf{H}\|^2\|\mathbf{C}\|^2\left(\mathbb{E}\mathrm{Tr}\,(\mathbf{HA})^2\,\mathbf{HX\widetilde{D}X^*H}\,(\mathbf{AH})^2\right.\\ &\quad \left.(\mathbf{HA})^2\,\mathbf{HX\widetilde{D}X^*H}\,(\mathbf{AH})^2\right)^{\frac{1}{2}}\times\left(\mathbb{E}\mathrm{Tr}\,(\mathbf{XBX^*})^4\right)^{\frac{1}{2}}\\ &\leq \frac{4t^2}{K^2}d_{\max}\|\mathbf{H}\|^8\|\mathbf{C}\|^2\|\mathbf{A}\|^4\sqrt{\mathbb{E}\frac{1}{K}\left(\frac{\mathbf{X\widetilde{D}X^*}}{K}\right)^2}\sqrt{\mathbb{E}\frac{1}{K}\left(\frac{\mathbf{XBX^*}}{K}\right)^4},\end{aligned}$$

where the first inequality follows by using the fact that $|\mathrm{Tr}\mathbf{AB}| \leq \|\mathbf{B}\|\mathrm{Tr}\,(\mathbf{A})$*,* $\mathbf{A}$ *being hermitian non-negative matrix and the second follows by applyig twice Cauchy-Schwartz inequalities:* $\mathrm{Tr}\,(\mathbf{AB}) \leq \sqrt{\mathrm{Tr}\,(\mathbf{AA^*})}\sqrt{\mathrm{Tr}\,(\mathbf{BB^*})}$ *and* $\mathbb{E}XY \leq \sqrt{\mathbb{E}X^2}\sqrt{\mathbb{E}Y^2}$*. We end up the proof of the first statement by using the fact that* $\frac{1}{K}\mathbb{E}\left[\frac{1}{K}\mathrm{Tr}\left(\frac{1}{K}\mathbf{XB}_K\mathbf{X^*}\right)^n\right]$ *is uniformly bounded in K whenever $\mathbf{B}_K$ is a sequence of diagonal matrices with uniformly bounded spectral norm and n is a given integer.*

The second statement follows from the resolvent identity:

$$\frac{1}{K}\mathrm{Tr}\mathbf{AHAHAH} = \frac{1}{K}\mathrm{Tr}\mathbf{AHAHA} - \frac{t}{K}\mathrm{Tr}\mathbf{AHAHAHXX^*}.$$

According to the first part of the proposition,

$$\mathrm{var}\left(\frac{1}{K}\mathrm{Tr}\mathbf{AHAHAHXX^*}\right) = \mathcal{O}(K^{-2}).$$

Now, $\mathrm{Tr}\mathbf{AHAHA} = \mathrm{Tr}\mathbf{A^2HAH}$ *and* $\mathrm{var}\frac{1}{K}\mathrm{Tr}\mathbf{A^2HAH} = \mathcal{O}(K^{-2})$ *by Proposition 7-1). Hence, applying inequality* $\mathrm{var}(X+Y) \leq \mathrm{var}(X) + \mathrm{var}(Y) + 2\sqrt{\mathrm{var}(X)\mathrm{var}(Y)}$ *yields the desired result. Proof of Proposition 8 is completed.*

Remark 5 *One can note that the third asymptotic moment is of order $\mathcal{O}(K^{-2})$. This is in accordance with the asymptotic normality of the SNR, where the third moment of $\sqrt{K}(\beta_K - \mathbb{E}(\beta_K))$ will eventually vanish, as this quantity becomes closer to a Gaussian random variable. However, its value remains significant for small dimension systems.*

We are now in position to complete the proof of Theorem 5. Using the notations of [36], the SNR writes:

$$\beta_K = \frac{tp_0}{K}\mathbf{z}^*\mathbf{D}^{\frac{1}{2}}\mathbf{H}(\mathbf{t})\mathbf{D}^{\frac{1}{2}}\mathbf{z},$$

where $t = \frac{1}{\rho}$. Hence, the third moment is given by:

$$\begin{aligned}
\mathbb{E}\left(\beta_K - \mathbb{E}\beta_K\right)^3 &= \frac{(tp_0)^3}{K^3}\mathbb{E}\left(\mathbf{z}^*\mathbf{D}^{\frac{1}{2}}\mathbf{H}\mathbf{D}^{\frac{1}{2}}\mathbf{z} - \mathbb{E}\mathrm{Tr}\mathbf{D}\mathbf{H}\right)^3, \\
&= \frac{(tp_0)^3}{K^3}\mathbb{E}\left(\mathbf{z}^*\mathbf{D}^{\frac{1}{2}}\mathbf{H}\mathbf{D}^{\frac{1}{2}}\mathbf{z} - \mathrm{Tr}\mathbf{D}\mathbf{H} + \mathrm{Tr}\mathbf{D}\mathbf{H} - \mathbb{E}\mathrm{Tr}\mathbf{D}\mathbf{H}\right)^3, \\
&= \frac{(tp_0)^3}{K^3}\left[\mathbb{E}\left(\mathbf{z}^*\mathbf{D}^{\frac{1}{2}}\mathbf{H}\mathbf{D}^{\frac{1}{2}}\mathbf{z} - \mathrm{Tr}\mathbf{D}\mathbf{H}\right)^3 + 3\mathbb{E}\left(\mathbf{z}^*\mathbf{D}^{\frac{1}{2}}\mathbf{H}\mathbf{D}^{\frac{1}{2}}\mathbf{z} - \mathrm{Tr}\mathbf{D}\mathbf{H}\right)^2\right. \\
&\quad \times\left(\mathrm{Tr}\mathbf{D}\mathbf{H} - \mathbb{E}\mathrm{Tr}\mathbf{D}\mathbf{H}\right) + 3\mathbb{E}\left(\mathbf{z}^*\mathbf{D}^{\frac{1}{2}}\mathbf{H}\mathbf{D}^{\frac{1}{2}}\mathbf{z} - \mathrm{Tr}\mathbf{D}\mathbf{H}\right)\left(\mathrm{Tr}\mathbf{D}\mathbf{H} - \mathbb{E}\mathrm{Tr}\mathbf{D}\mathbf{H}\right)^2 \\
&\quad \left.+\mathbb{E}\left(\mathrm{Tr}\mathbf{D}\mathbf{H} - \mathbb{E}\mathrm{Tr}\mathbf{D}\mathbf{H}\right)^3\right], \\
&= \frac{(tp_0)^3}{K^3}\left[\mathbb{E}\left(\mathbf{z}^*\mathbf{D}^{\frac{1}{2}}\mathbf{H}\mathbf{D}^{\frac{1}{2}}\mathbf{z} - \mathrm{Tr}\mathbf{D}\mathbf{H}\right)^3 + 3\mathbb{E}\left(\mathbf{z}^*\mathbf{D}^{\frac{1}{2}}\mathbf{H}\mathbf{D}^{\frac{1}{2}}\mathbf{z} - \mathrm{Tr}\mathbf{D}\mathbf{H}\right)^2\right. \\
&\quad \left.\times\left(\mathrm{Tr}\mathbf{D}\mathbf{H} - \mathbb{E}\mathrm{Tr}\mathbf{D}\mathbf{H}\right) + \mathbb{E}\left(\mathrm{Tr}\mathbf{D}\mathbf{H} - \mathbb{E}\mathrm{Tr}\mathbf{D}\mathbf{H}\right)^3\right] \qquad (3.11)
\end{aligned}$$

In order to deal with the first term of the right-hand side of (3.11), notice that if $\mathbf{M}$ is a deterministic matrix and $\mathbf{x}$ is a standard Gaussian vector, then:

$$\mathbb{E}\left(\mathbf{x}^*\mathbf{M}\mathbf{x} - \mathrm{Tr}\mathbf{M}\right)^3 = \mathrm{Tr}(\mathbf{M}^3)\mathbb{E}\left(|x_1|^2 - 1\right)^3$$

(such an identity can be easily proved by considering the spectral decomposition of $\mathbf{M}$). Hence,

$$\begin{aligned}
\mathbb{E}\left(\mathbf{z}^*\mathbf{D}^{\frac{1}{2}}\mathbf{H}\mathbf{D}^{\frac{1}{2}}\mathbf{z} - \mathrm{Tr}\mathbf{D}\mathbf{H}\right)^3 &= \mathbb{E}\mathrm{Tr}\left(\mathbf{D}\mathbf{H}\right)^3\mathbb{E}\left(|Z_{11}|^2 - 1\right)^3, \\
&= 2\mathbb{E}\mathrm{Tr}\left(\mathbf{D}\mathbf{H}\mathbf{D}\mathbf{H}\mathbf{D}\mathbf{H}\right).
\end{aligned}$$

The second term of the right-hand side of (3.11) is uniformly bounded in K. Indeed:

$$\begin{aligned}
3\mathbb{E}\left(\mathbf{z}^*\mathbf{D}^{\frac{1}{2}}\mathbf{H}\mathbf{D}^{\frac{1}{2}}\mathbf{z} - \mathrm{Tr}(\mathbf{D}\mathbf{H})\right)^2 &= 3\mathbb{E}\left(|Z_{11}|^2 - 1\right)^2\mathrm{Tr}\mathbf{D}\mathbf{H}\mathbf{D}\mathbf{H}\left(\mathrm{Tr}\mathbf{D}\mathbf{H} - \mathbb{E}\mathrm{Tr}\mathbf{D}\mathbf{H}\right), \\
&\leq 3\sqrt{\mathrm{var}\left(\mathrm{Tr}\mathbf{D}\mathbf{H}\mathbf{D}\mathbf{H}\right)}\sqrt{\mathrm{var}\left(\mathrm{Tr}\mathbf{D}\mathbf{H}\right)}
\end{aligned}$$

which is $\mathcal{O}(1)$ according to Proposition 7. It remains to deal with $\mathbb{E}\left(\mathrm{Tr}\mathbf{D}\mathbf{H} - \mathbb{E}\mathrm{Tr}\mathbf{D}\mathbf{H}\right)^3$, which can be proved to be uniformly bounded in K using concentration results for the spectral measure of random matrices [32] (see also [49, eq.(86)-(87)], where details are provided). Consequently, we end up with the following approximation:

$$K^2\mathbb{E}\left(\beta_K - \mathbb{E}\beta_K\right)^3 = \frac{(tp_0)^3}{K}\mathbb{E}\left(|Z_{11}|^2 - 1\right)^3\mathbb{E}\mathrm{Tr}\mathbf{D}\mathbf{H}\mathbf{D}\mathbf{H}\mathbf{D}\mathbf{H} + \mathcal{O}\left(K^{-1}\right)$$

which is deterministic but still depends on the distribution of the entries via the expectation operator $\mathbb{E}$. The rest of the proof is devoted to provide a deterministic approximation of $\mathbb{E}\mathrm{Tr}\left(\mathbf{D}\mathbf{H}\mathbf{D}\mathbf{H}\mathbf{D}\mathbf{H}\right)$ depending on γ, $\tilde{\gamma}$, $\mathbf{T}$ and $\widetilde{\mathbf{T}}$.

Note that $\mathbf{H} = \mathbf{I} - \frac{t}{K}\mathbf{H}\mathbf{X}\mathbf{X}^*$, thus:

$$\begin{aligned}
[\mathbf{HDHDH}]_{pp} &= [\mathbf{HDHD}]_{pp} - t\left[\mathbf{HDHDH}\frac{\mathbf{XX}^*}{K}\right]_{pp}, \\
&= [\mathbf{HDHD}]_{pp} - \frac{t}{K}\sum_{j=1}^{K}[\mathbf{HDHDHx}_j]_p\overline{X_{pj}}. \qquad (3.12)
\end{aligned}$$

Let us deal with the second term of (3.12). We have:

$$\mathbb{E}\frac{1}{K}[\mathbf{HDHDHx}_j]_p\overline{X_{pj}} = \frac{1}{K}\sum_{k=1}^{N}\mathbb{E}\left([\mathbf{HDHDH}]_{pk}X_{kj}\overline{X_{pj}}\right).$$

Using the integration by part formula (3.5), we get:

$$\begin{aligned}
\mathbb{E}[\mathbf{HDHDHx}_j]_p\overline{X_{pj}} &= \sum_{k=1}^{N} d_k\tilde{d}_j\delta(p-k)\mathbb{E}[\mathbf{HDHDH}]_{pk} \\
&\quad + \sum_{k=1}^{N} d_k\tilde{d}_j\mathbb{E}\left[\overline{X_{pj}}\sum_{\ell,m=1}^{N}\frac{\partial\left[H_{p\ell}d_\ell d_m H_{\ell m}H_{mk}\right]}{\partial\overline{X_{kj}}}\right], \\
&= d_p\tilde{d}_j\mathbb{E}[\mathbf{HDHDH}]_{pp} - \frac{t}{K}\sum_{k,\ell,m=1}^{N} d_k\tilde{d}_j d_m d_\ell\mathbb{E}\left[\overline{X_{pj}}[\mathbf{Hx}_j]_p H_{k\ell}H_{\ell m}H_{mk}\right] \\
&\quad - \frac{t}{K}\sum_{k,\ell,m=1}^{N} d_k\tilde{d}_j d_m d_\ell\mathbb{E}\left[\overline{X_{pj}}H_{p\ell}[\mathbf{Hx}_j]_\ell H_{km}H_{mk}\right] \\
&\quad - \frac{t}{K}\sum_{k,\ell,m=1}^{N} d_k\tilde{d}_j d_m d_\ell\mathbb{E}\left[H_{p\ell}H_{\ell m}[\mathbf{Hx}_j]_m H_{kk}\right]. \\
&= d_p\tilde{d}_j\mathbb{E}[\mathbf{HDHDH}]_{pp} - \frac{t}{K}\tilde{d}_j\mathbb{E}\left[[\mathbf{Hx}_j]_p\overline{X_{pj}}\mathrm{Tr}(\mathbf{DHDHDH})\right] \\
&\quad - \frac{t}{K}\tilde{d}_j\mathbb{E}\left[[\mathbf{HDHx}_j]_p\overline{X_{pj}}\mathrm{Tr}(\mathbf{DHDH})\right] \\
&\quad - \frac{t}{K}\tilde{d}_j\mathbb{E}\left[[\mathbf{HDHDHx}_j]_p\overline{X_{pj}}\mathrm{Tr}(\mathbf{DH})\right].
\end{aligned}$$

Substituting in the last term $\frac{1}{K}\mathrm{Tr}\mathbf{DH} = \overset{o}{\beta} + \alpha$ where $\overset{o}{\beta} = \beta - \alpha$, we get:

$$\begin{aligned}
\mathbb{E}[\mathbf{HDHDHx}_j]_p\overline{X_{pj}} &= d_p\tilde{d}_j\mathbb{E}[\mathbf{HDHDH}]_{pp} - \frac{t}{K}\tilde{d}_j\mathbb{E}\left[[\mathbf{Hx}_j]_p\overline{X_{pj}}\mathrm{Tr}(\mathbf{DHDHDH})\right] \\
&\quad - \frac{t}{K}\tilde{d}_j\mathbb{E}\left[[\mathbf{HDHx}_j]_p\overline{X_{pj}}\mathrm{Tr}(\mathbf{DHDH})\right] - t\tilde{d}_j\mathbb{E}\left[[\mathbf{HDHDHx}_j]_p\overline{X_{pj}}\overset{o}{\beta}\right] \\
&\quad - t\tilde{d}_j\mathbb{E}\left[[\mathbf{HDHDHx}_j]_p\overline{X_{pj}}\right]\alpha.
\end{aligned}$$

Therefore, we have:

$$\begin{aligned}
&\left(1 + t\alpha\tilde{d}_j\right)\mathbb{E}\left[[\mathbf{HDHDHx}_j]_p\overline{X_{pj}}\right] \\
=\;& d_p\tilde{d}_j\mathbb{E}\left[\mathbf{HDHDH}\right]_{pp} - \frac{t}{K}\mathbb{E}\left[[\mathbf{Hx}_j]_p\overline{X_{pj}}\tilde{d}_j\mathrm{Tr}\left[\mathbf{DHDHDH}\right]\right] \\
&-\frac{t}{K}\tilde{d}_j\mathbb{E}\left[[\mathbf{HDHx}_j]_p\overline{X_{pj}}\mathrm{Tr}\left[\mathbf{DHDH}\right]\right] - t\tilde{d}_j\mathbb{E}\left[[\mathbf{HDHDHx}_j]_p\overline{X_{pj}}\overset{o}{\beta}\right].
\end{aligned}$$

Multiplying the right hand and the left hand sides by $\tilde{r}_j = \frac{1}{1+t\alpha\tilde{d}_j}$, we get:

$$\begin{aligned}
\mathbb{E}\left[\mathbf{HDHDHx}_j\right]_p\overline{X_{pj}} = \tilde{r}_j d_p\tilde{d}_j\mathbb{E}\left[\mathbf{HDHDH}\right]_{pp} - \frac{t}{K}\tilde{r}_j\mathbb{E}\left[[\mathbf{Hx}_j]_p\overline{X_{pj}}\tilde{d}_j\mathrm{Tr}\left[\mathbf{DHDHDH}\right]\right] \\
-\frac{t}{K}\tilde{d}_j\tilde{r}_j\mathbb{E}\left[[\mathbf{HDHx}_j]_p\overline{X_{pj}}\mathrm{Tr}\left[\mathbf{DHDH}\right]\right] - t\tilde{d}_j\tilde{r}_j\mathbb{E}\left[[\mathbf{HDHDHx}_j]_p\overline{X_{pj}}\overset{o}{\beta}\right].
\end{aligned} \tag{3.13}$$

Plugging (3.13) into (3.12), we obtain:

$$\begin{aligned}
\mathbb{E}\left[\mathbf{HDHDH}\right]_{pp} =\;& \mathbb{E}\left[\mathbf{HDHD}\right]_{pp} - \sum_{j=1}^{K}\frac{t}{K}\tilde{r}_j d_p\tilde{d}_j\mathbb{E}\left[\mathbf{HDHDH}\right]_{pp} \\
&+\frac{t^2}{K^2}\sum_{j=1}^{K}\tilde{r}_j\mathbb{E}\left[\mathbf{Hx}_j\right]_p\overline{X_{pj}}\tilde{d}_j\mathrm{Tr}\left[\mathbf{DHDHDH}\right] \\
&+\frac{t^2}{K^2}\sum_{j=1}^{K}\tilde{d}_j\tilde{r}_j\mathbb{E}\left[\mathbf{HDHx}_j\right]_p\overline{X_{p,j}}\mathrm{Tr}\left[\mathbf{DHDH}\right] + \frac{t}{K}\sum_{j=1}^{K}\tilde{d}_j\tilde{r}_j\mathbb{E}\left[\mathbf{HDHDHx}_j\right]_p\overline{X_{p,j}}\overset{o}{\beta}, \\
=\;& \mathbb{E}\left[\mathbf{HDHD}\right]_{pp} - t\tilde{\alpha}d_p\mathbb{E}\left[\mathbf{HDHDH}\right]_{pp} + \frac{t^2}{K^2}\mathbb{E}\mathrm{Tr}(\mathbf{DHDHDH})\left[\mathbf{HX\widetilde{R}\widetilde{D}X}^*\right]_{pp} \\
&+\frac{t^2}{K^2}\mathbb{E}\mathrm{Tr}\left[\mathbf{DHDH}\right]\left[\mathbf{HDHX\widetilde{D}\widetilde{R}X}^*\right]_{pp} + \frac{t^2}{K}\mathbb{E}\overset{o}{\beta}\left[\mathbf{HDHDHX\widetilde{D}\widetilde{R}X}^*\right]_{pp}.
\end{aligned}$$

Hence,

$$\begin{aligned}
(1 + t\tilde{\alpha}d_p)\mathbb{E}\left[\mathbf{HDHDH}\right]_{pp} = \mathbb{E}\left[\mathbf{HDHD}\right]_{pp} + \frac{t^2}{K^2}\mathbb{E}\mathrm{Tr}\left[\mathbf{DHDHDH}\right]\left[\mathbf{HX\widetilde{R}\widetilde{D}X}^*\right]_{pp} \\
+\frac{t^2}{K^2}\mathbb{E}\mathrm{Tr}\left[\mathbf{DHDH}\right]\left[\mathbf{HDHX\widetilde{D}\widetilde{R}X}^*\right]_{pp} + \frac{t^2}{K}\mathbb{E}\overset{o}{\beta}\left[\mathbf{HDHDHX\widetilde{D}\widetilde{R}X}^*\right]_{pp}.
\end{aligned}$$

Multiplying the left and right hand sides by $r_p = \frac{1}{1+t\tilde{\alpha}d_p}$, we get:

$$\begin{aligned}
\mathbb{E}\left[\mathbf{HDHDH}\right]_{pp} = r_p\mathbb{E}\left[\mathbf{HDHD}\right]_{pp} + \frac{t^2}{K^2}r_p\mathbb{E}\mathrm{Tr}\left[\mathbf{DHDHDH}\right]\left[\mathbf{HX\widetilde{R}\widetilde{D}X}^*\right]_{pp} \\
+\frac{t^2}{K^2}r_p\mathbb{E}\mathrm{Tr}\left[\mathbf{DHDH}\right]\left[\mathbf{HDHX\widetilde{D}\widetilde{R}X}^*\right]_{pp} + \frac{t^2}{K}r_p\mathbb{E}\overset{o}{\beta}\left[\mathbf{HDHDHX\widetilde{D}\widetilde{R}X}^*\right]_{pp}.
\end{aligned} \tag{3.14}$$

Multiplying by d_p, summing over p and dividing by K, we obtain:

$$\begin{aligned}
\mathbb{E}\frac{1}{K}\mathrm{Tr}\left[\mathbf{DHDHDH}\right] &= \mathbb{E}\frac{1}{K}\sum_{p=1}^{K} d_p\left[\mathbf{HDHDH}\right]_{pp}, \\
&= \frac{1}{K}\sum_{p=1}^{K} r_p d_p \mathbb{E}\left[\mathbf{HDHD}\right]_{pp} + \frac{t^2}{K^3}\mathbb{E}\,\mathrm{Tr}\left(\mathbf{DHDHDH}\right)\mathrm{Tr}\left(\mathbf{DRHX}\widetilde{\mathbf{R}}\widetilde{\mathbf{D}}\mathbf{X}^*\right) \\
&\quad + \frac{t^2}{K^3}\mathbb{E}\mathrm{Tr}\left(\mathbf{DHDH}\right)\mathrm{Tr}\left(\mathbf{DRHDHX}\widetilde{\mathbf{D}}\widetilde{\mathbf{R}}\mathbf{X}^*\right) \\
&\quad + \frac{t^2}{K^2}\mathbb{E}\ \overset{o}{\beta}\ \mathrm{Tr}\left(\mathbf{DRHDHDHX}\widetilde{\mathbf{D}}\widetilde{\mathbf{R}}\mathbf{X}^*\right), \\
&\triangleq \chi_1 + \chi_2 + \chi_3 + \chi_4, \qquad (3.15)
\end{aligned}$$

where:

$$\begin{aligned}
\chi_1 &= \frac{1}{K}\mathbb{E}\mathrm{Tr}\left(\mathbf{DRHDHD}\right), \\
\chi_2 &= \frac{t^2}{K}\mathbb{E}\mathrm{Tr}\left(\mathbf{DHDHDH}\right)\frac{1}{K}\mathrm{Tr}\left(\mathbf{DRH}\frac{\mathbf{X}\widetilde{\mathbf{D}}\widetilde{\mathbf{R}}\mathbf{X}^*}{K}\right), \\
\chi_3 &= \frac{t^2}{K}\mathbb{E}\mathrm{Tr}\left(\mathbf{DHDH}\right)\frac{1}{K}\mathrm{Tr}\left(\mathbf{DRHDH}\frac{\mathbf{X}\widetilde{\mathbf{D}}\widetilde{\mathbf{R}}\mathbf{X}^*}{K}\right), \\
\chi_4 &= \frac{t^2}{K}\mathbb{E}\ \overset{o}{\beta}\ \mathrm{Tr}\left(\mathbf{DRHDHDH}\frac{\mathbf{X}\widetilde{\mathbf{D}}\widetilde{\mathbf{R}}\mathbf{X}^*}{K}\right).
\end{aligned}$$

According to Proposition 7, $\mathrm{var}\frac{1}{K}\mathrm{Tr}\left(\mathbf{DRHDHDH}\frac{\mathbf{X}\widetilde{\mathbf{D}}\widetilde{\mathbf{R}}\mathbf{X}^*}{K}\right)$ is of order $\mathcal{O}(K^{-2})$. Similarly, $\mathrm{var}(\beta) = \mathcal{O}(K^{-2})$. Hence, using Cauchy-Schwartz inequality, we get the estimation $\chi_4 = \mathcal{O}(K^{-2})$. It remains to work out the expressions involved in χ_1, χ_2 and χ_3 by removing the terms with expectation and replacing them with deterministic equivalents.

Since $\mathrm{var}\frac{1}{K}\mathrm{Tr}\left(\mathbf{DRH}\frac{\mathbf{X}\widetilde{\mathbf{D}}\widetilde{\mathbf{R}}\mathbf{X}^*}{K}\right) = \mathcal{O}(K^{-2})$ by Proposition 7 and $\mathrm{var}(\frac{1}{K}\mathrm{Tr}\mathbf{DHDHDH}) = \mathcal{O}(K^{-2})$ by Proposition 8, we have:

$$\begin{aligned}
\chi_2 &= \frac{t^2}{K}\mathbb{E}\mathrm{Tr}\left(\mathbf{DHDHDH}\right)\mathbb{E}\left(\frac{1}{K}\mathrm{Tr}\left[\mathbf{DRH}\frac{\mathbf{X}\widetilde{\mathbf{D}}\widetilde{\mathbf{R}}\mathbf{X}^*}{K}\right]\right) + \mathcal{O}(K^{-2}), \\
&\overset{(a)}{=} \frac{t^2}{K}\mathbb{E}\mathrm{Tr}\left(\mathbf{DHDHDH}\right)\frac{1}{K}\mathrm{Tr}\left(\widetilde{\mathbf{D}}\widetilde{\mathbf{T}}\widetilde{\mathbf{D}}\widetilde{\mathbf{R}}\right)\frac{1}{K}\mathrm{Tr}\left(\mathbf{DRDT}\right) + \mathcal{O}(K^{-2}), \\
&\overset{(b)}{=} \frac{t^2}{K}\mathbb{E}\mathrm{Tr}\left(\mathbf{DHDHDH}\right)\gamma\tilde{\gamma} + \mathcal{O}(K^{-2}). \qquad (3.16)
\end{aligned}$$

where (a) follows from Proposition 7-2) and (b), from Proposition 6. Similar arguments yield:

$$\begin{aligned}
\chi_3 &= \frac{t^2}{K}\mathbb{E}\mathrm{Tr}\left(\mathbf{DHDH}\right)\mathbb{E}\left(\frac{1}{K}\mathrm{Tr}\left[\mathbf{DRHDH}\frac{\mathbf{X}\widetilde{\mathbf{D}}\widetilde{\mathbf{R}}\mathbf{X}^*}{\mathbf{K}}\right]\right) + \mathcal{O}(K^{-2}), \\
&= \frac{t^2\gamma}{(1-t^2\gamma\tilde{\gamma})^2}\left[\frac{1}{K}\mathrm{Tr}\left(\widetilde{\mathbf{D}}\widetilde{\mathbf{T}}\widetilde{\mathbf{D}}\widetilde{\mathbf{R}}\right)\frac{1}{K}\mathrm{Tr}\left(\mathbf{DRD}^2\mathbf{T}^2\right)\right. \\
&\quad \left. -\frac{t\gamma}{K}\mathrm{Tr}\left(\widetilde{\mathbf{D}}^2\widetilde{\mathbf{T}}^2\widetilde{\mathbf{D}}\widetilde{\mathbf{R}}\right)\frac{1}{K}\mathrm{Tr}(\mathbf{DRDT})\right] + \mathcal{O}(K^{-2})\ , \\
&= \frac{t^2\gamma}{(1-t^2\gamma\tilde{\gamma})^2}\left[\frac{\tilde{\gamma}}{K}\mathrm{Tr}(\mathbf{D}^3\mathbf{T}^3) - \frac{t\gamma^2}{K}\mathrm{Tr}(\widetilde{\mathbf{D}}^3\widetilde{\mathbf{T}}^3)\right] + \mathcal{O}(K^{-2}) \qquad (3.17)
\end{aligned}$$

and

$$\begin{aligned}
\chi_1 &= \frac{1}{1-t^2\gamma\tilde{\gamma}}\frac{1}{K}\mathrm{Tr}\left(\mathbf{D}^2\mathbf{RDT}^2\right) + \mathcal{O}(K^{-2}) \\
&= \frac{1}{1-t^2\gamma\tilde{\gamma}}\frac{1}{K}\mathrm{Tr}(\mathbf{D}^3\mathbf{T}^3) + \mathcal{O}(K^{-2}). \qquad (3.18)
\end{aligned}$$

Plugging (3.17), (3.16) and (3.18) into (3.15), we obtain:

$$\frac{1}{K}\mathbb{E}\mathrm{Tr}(\mathbf{DHDHDH}) = \frac{1}{K(1-t^2\gamma\tilde{\gamma})^3}\mathrm{Tr}\mathbf{D}^3\mathbf{T}^3 - \frac{t^3\gamma^3}{K(1-t^2\gamma\tilde{\gamma})^3}\mathrm{Tr}\widetilde{\mathbf{T}}^3\widetilde{\mathbf{D}}^3 + \mathcal{O}(K^{-2}).$$

Hence,

$$\begin{aligned}
K^2\mathbb{E}\left(\frac{\beta_K}{p_0} - \mathbb{E}\frac{\beta_K}{p_0}\right)^3 &= \frac{\rho^3}{K\left(\rho^2-\gamma\tilde{\gamma}\right)^3}\left[\mathrm{Tr}\mathbf{D}^3\mathbf{T}^3 - \frac{\gamma^3}{\rho^3}\mathrm{Tr}\widetilde{\mathbf{D}}^3\widetilde{\mathbf{T}}^3\right]\mathbb{E}\left(|Z_{11}|^2-1\right)^3 + \mathcal{O}\left(\frac{1}{K}\right), \\
&= \frac{2\rho^3}{K\left(\rho^2-\gamma\tilde{\gamma}\right)^3}\left[\mathrm{Tr}\mathbf{D}^3\mathbf{T}^3 - \frac{\gamma^3}{\rho^3}\mathrm{Tr}\widetilde{\mathbf{D}}^3\widetilde{\mathbf{T}}^3\right] + \mathcal{O}\left(\frac{1}{K}\right).
\end{aligned}$$

The fact that $\nu_K = \frac{2\rho^3}{K(\rho^2-\gamma\tilde{\gamma})^3}\left[\mathrm{Tr}\mathbf{D}^3\mathbf{T}^3 - \frac{\gamma^3}{\rho^3}\mathrm{Tr}\widetilde{\mathbf{D}}^3\widetilde{\mathbf{T}}^3\right]$ is of order $\mathcal{O}(1)$ is straightforward and its proof is omitted. Proof of Theorem 5 is completed.

5 Simulation results

In our simulations, we consider a MIMO system in the uplink direction. The base station is equipped with N receiving antennas and detects the symbols transmitted by a particular user in the presence of K interfering users. We assume that the correlation matrix $\mathbf{\Psi}$ is given by $\mathbf{\Psi}(i,j) = \sqrt{\frac{K}{N}}a^{|i-j|}$ with $0 \leq a < 1$. Recall that $\widetilde{\mathbf{P}}$ is the matrix of the interfering users' powers. We set $\widetilde{\mathbf{P}}$ (up to a permutation of its diagonal elements) to:

$$\widetilde{\mathbf{P}} = \begin{cases} \mathrm{diag}([4P \quad 5P]) & \text{if} \quad K=2 \\ \mathrm{diag}([P \quad P \quad 2P \quad 4P]) & \text{if} \quad K=4 \end{cases},$$

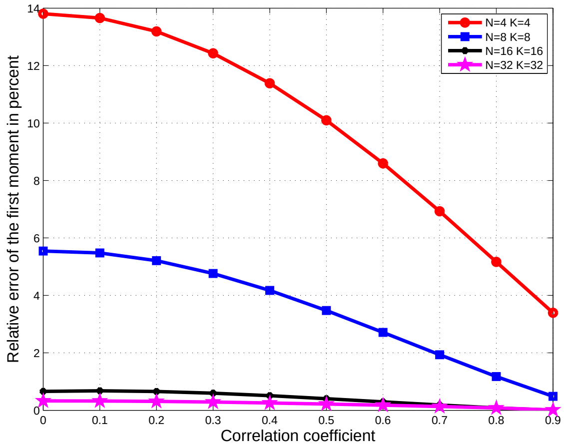
Table 3.1: Power classes and relative frequencies

Class	1	2	3	4	5
Power	P	$2P$	$4P$	$8P$	$16P$
Relative frequency	1/8	1/4	1/4	1/8	1/4

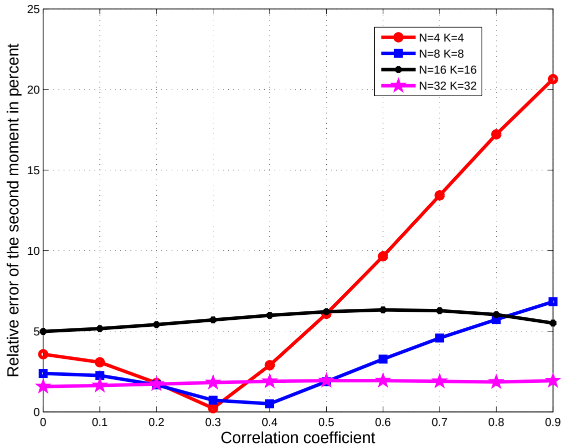
where P is the power of the user of interest. For $K = 2^p$ with $3 \leq p \leq 5$, we assume that the powers of the interfering sources are arranged into five classes as in Table 5. We investigate the impact of the correlation coefficient a on the accuracy of the asymptotic moments when the input SNR is set to 15dB for $N = K$ (Fig. 3.1) and $N = 2K$ (Fig. 3.2). In these figures, the relative error on the estimated first three moments $\frac{|\mu_\infty - \mu|}{\mu}$ (μ_∞ and μ denote respectively the asymptotic and empirical moment) is depicted with respect to the correlation coefficient a. These simulations show that when the number of antennas is small, the asymptotic approximation of the second and third moments degrades for large correlation coefficients (a close to one). Despite these discrepancies for a close to 1, simulations show that the BER and the outage probability are well approximated even for small system dimensions.

Indeed, Figure 3.3 shows the evolution of the empirical BER and the theoretical BER predicted by (3.1) versus the input SNR for different values of a, K and N. In Figure 3.4, the saddle point approximate of the outage probability given by (3.2) is compared with the empirical one. In both Figures 3.3 and 3.4, 2000 channel realizations have been considered, and in Fig. 3.4, the input SNR has been set to 15 dB. These figures show that even for small system dimensions, the BER is well approximated for a wide range of SNR values. For high SNR values, the proposed approximation tends to underestimate the bit error rate. A possible reason might be that the first three moments are not sufficient to estimate accurately the bit error rate (BER). To get a more accurate bit error rate value, one should go beyond these moments and take into account the values of higher order moments.

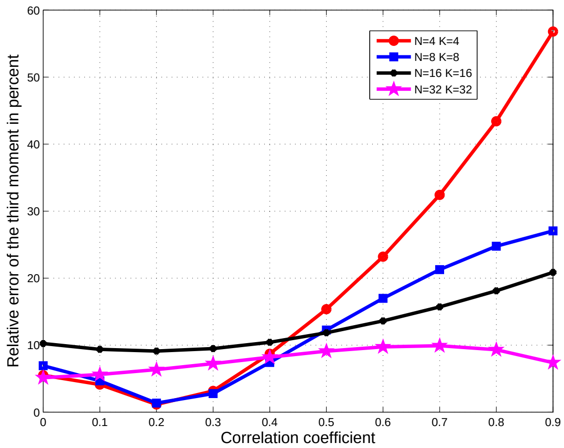
The outage probability is also well approximated except for small values of the SNR threshold that are likely to be in the tail of the asymptotic distribution.



(a) First moment of the SNR

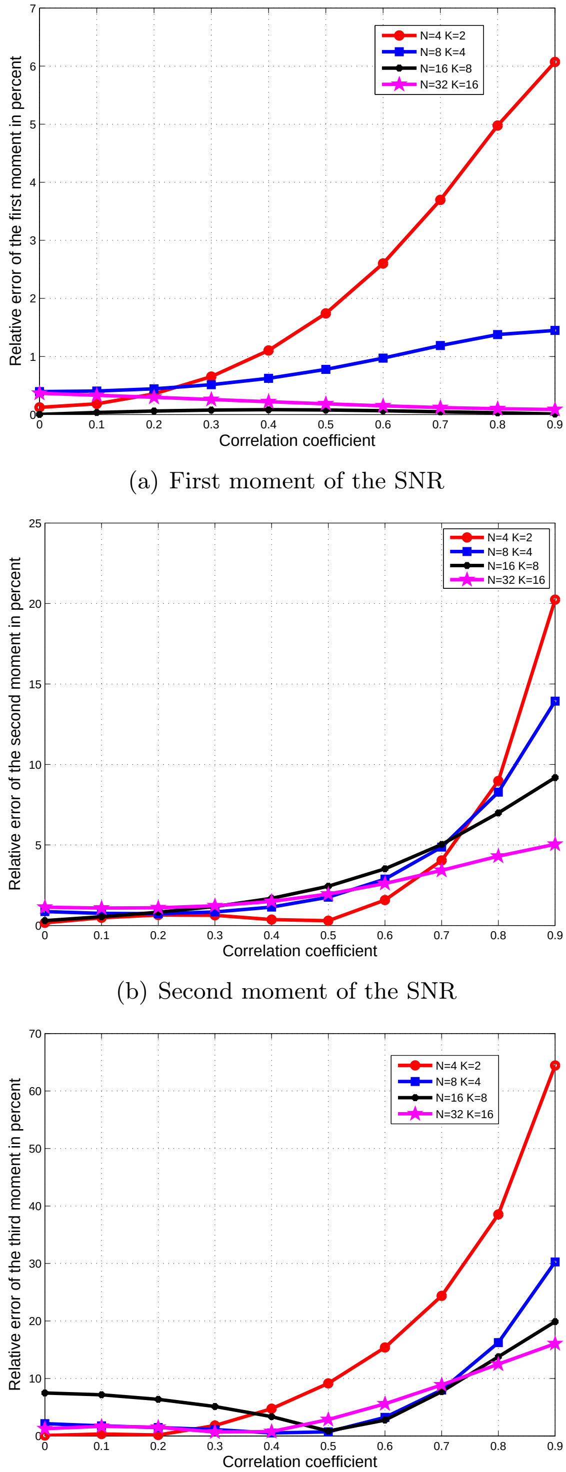


(b) Second moment of the SNR



(c) Third moment of the SNR

Figure 3.1: Absolute value of the relative error when $N = K$

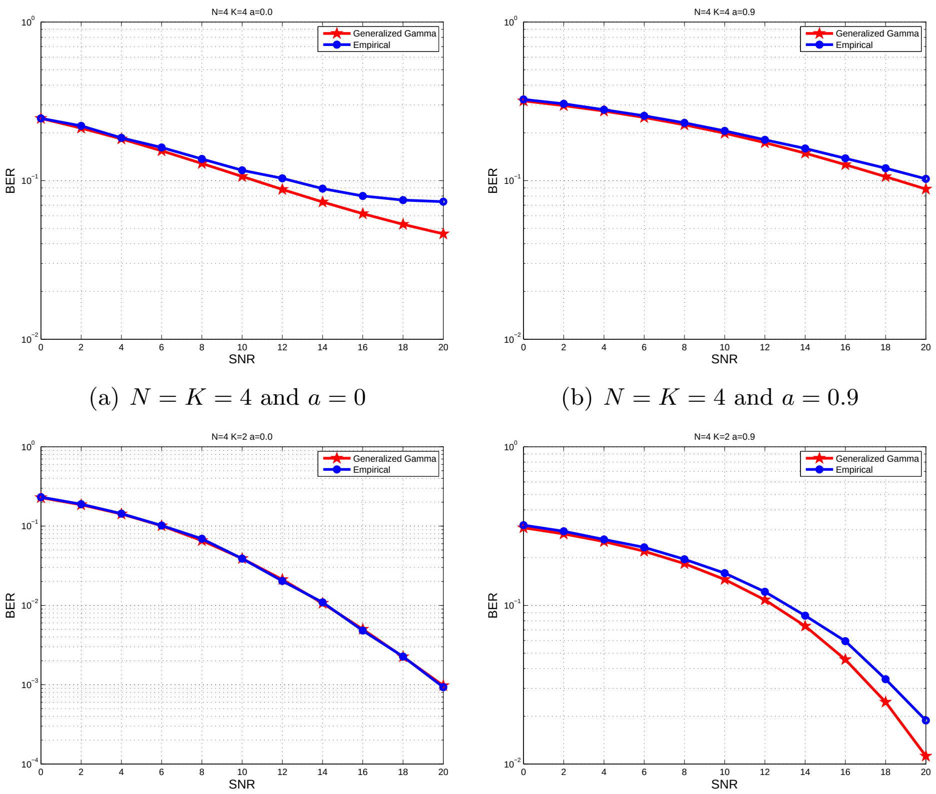


(a) First moment of the SNR

(b) Second moment of the SNR

(c) Third moment of the SNR

Figure 3.2: Absolute value of the relative error when $N = 2K$



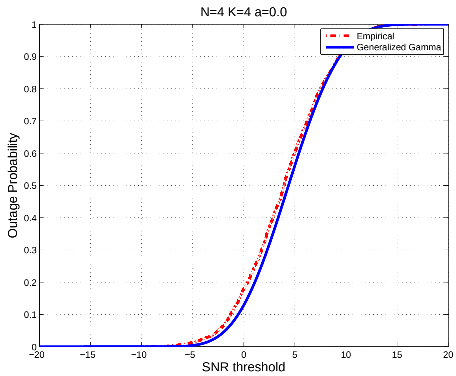
(a) $N = K = 4$ and $a = 0$

(b) $N = K = 4$ and $a = 0.9$

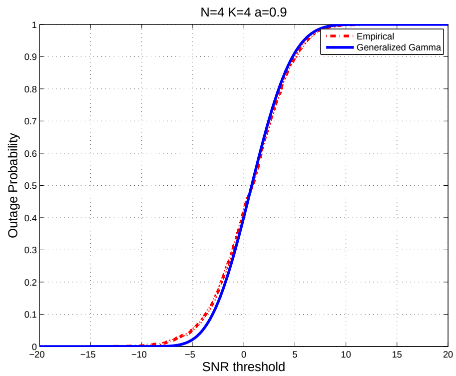
(c) $N = 2K = 4$ and $a = 0$

(d) $N = 2K = 4$ and $a = 0.9$

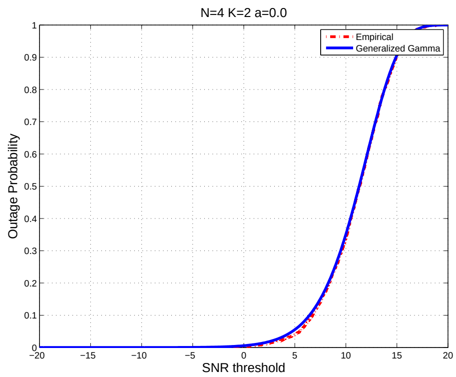
Figure 3.3: BER vs input SNR



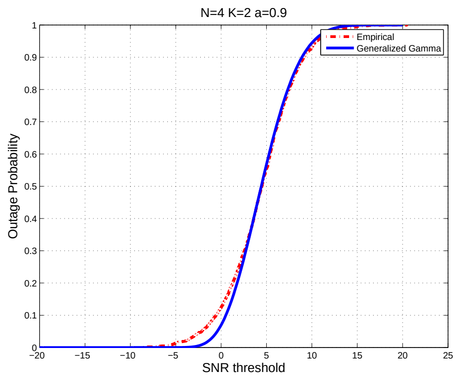
(a) $N = K = 4$ and $a = 0$



(b) $N = K = 4$ and $a = 0.9$



(c) $N = 2K = 4$ and $a = 0$



(d) $N = 2K = 4$ and $a = 0.9$

Figure 3.4: Outage Probability vs SNR threshold

CHAPTER 4

A CLT for Information-Theoretic Statistics of non-Centred Gram Random Matrices

Contents

1 Introduction

This chapter provides a CLT for Information-Theoretic Statistics of non-centred Gram matrices. This study is ongoing work.

The model, the statistics, and the literature

Consider a $N \times n$ random matrix $\Sigma_n = (\xi_{ij}^n)$ where:

$$\Sigma_n = \frac{1}{\sqrt{n}} D_n^{\frac{1}{2}} X_n \tilde{D}_n^{\frac{1}{2}} + A_n \ , \tag{4.1}$$

where $A_n = (a_{ij}^n)$ is a deterministic $N \times n$ matrix with uniformily bounded spectral norm, D_n and $\tilde{D}_n$ are diagonal deterministic matrices with nonnegative entries, with respective dimensions $N \times N$ and $n \times n$; $X_n = (X_{ij})$ is a $N \times n$ matrix with the entries X_{ij}'s being centered, independent and identically distributed (i.i.d.) random variables with unit variance $\mathbb{E}|X_{ij}|^2 = 1$ and finite 16^{th} moment, Consider the following linear statistics of the eigenvalues:

$$\mathcal{I}_n(\rho) = \frac{1}{N} \log\det\left(\Sigma_n \Sigma_n^* + \rho I_N\right) = \frac{1}{N} \sum_{i=1}^{N} \log(\lambda_i + \rho)$$

where I_N is the $N \times N$ identity matrix, $\rho > 0$ is a given parameter and the λ_i's are the eigenvalues of matrix $\Sigma_n \Sigma_n^*$ (Σ_n^* stands for the Hermitian adjoint of Σ_n). This functional, known as the mutual information for multiple antenna radio channels, is fundamental in wireless communication as it caracterizes the performance of a (coherent) communication over a wireless Multiple-Input Multiple-Output (MIMO) channel with gain matrix Σ_n. Channels with non-centered gain matrix $\Sigma_n = n^{-1/2} D_n^{1/2} X_n \tilde{D}_n^{1/2} + A_n$ are known as Rician channels and the deterministic matrix A_n accounts for a so-called line-of-sight component, D_n and $\tilde{D}_n$ for the correlations respectively at the receiving and emitting sides.

Since the seminal work of Telatar [80], the study of the mutual information $\mathcal{I}_n(\rho)$ of a MIMO channel (and other performance indicators) in the regime where the dimensions of the gain matrix grow to infinity at the same pace has turned to be extremely fruitful. However, Rician channels have been comparatively less studied from this point of view, as their analysis is more difficult due to the presence of the deterministic matrix A_n. First order results can be found in Girko [24, 25]; Dozier and Silverstein [18, 19] established convergence results for the spectral measure; and the systematic study of the convergence of $\mathcal{I}_n(\rho)$ for a correlated Rician channel has been undertaken by Hachem et al. in [20, 35], etc.

Fluctuations for particular linear statistics (and general classes of linear statistics) of large random matrices have been widely studied: CLTs for Wigner matrices can be traced back to Girko [28] (see also [29]). Results for this class of matrices have also been obtained by Khorunzhy et al. [47], Boutet de Monvel and Khorunzhy [11], Johansson [40], Sinai and Sochnikov [74], Soshnikov [75], Anderson and Zeitouni [1], Chatterjee [14], Lytova and Pastur [51], etc. The case of Gram matrices has been studied in Arharov [2], Jonsson [41], Bai and Silverstein [3], Hachem et al. [37], etc.

Fluctuation results dedicated to wireless communication applications have been developed in the centred case ($\mathbf{A}_n = 0$) by Debbah and Müller [17] and Tulino and Verdù [85], Hachem *et al.* [36] for gaussian entries and [37]. Other fluctuation results either based on the replica method or on saddle-point analysis have been developed by Moustakas, Sengupta *et al.* [55, 63]. In the Rician case, based on the replica method, Taricco, provided an asymptotic formula for the variance of the mutual information statistic in [78, 79].

Presentation of the results

The purpose of this article is to establish a Central Limit Theorem (CLT) for $\mathcal{I}_n(\rho)$ in the following regime

$$N, n \to \infty \quad \text{and} \quad 0 < \liminf \frac{N}{n} \le \limsup \frac{N}{n} < \infty\ , \tag{4.2}$$

(simply denoted by $N, n \to \infty$ in the sequel) under mild assumptions for matrices X_n, A_n, D_n and $\tilde{D}_n$.

Fundamental equations, deterministic equivalents

We collect here resuls from [35]. The following system of equations

$$\left\{\begin{array}{lcl} \delta_n(z) &=& \frac{1}{n}\mathrm{Tr}\, D_n\left(-z(I_N+\tilde{\delta}_n(z)D_n)+A_n(I_n+\delta_n(z)\tilde{D}_n)^{-1}A_n^*\right)^{-1} \\ \tilde{\delta}_n(z) &=& \frac{1}{n}\mathrm{Tr}\, \tilde{D}_n\left(-z(I_n+\delta_n(z)\tilde{D}_n)+A_n^*(I_N+\tilde{\delta}_n(z)D_n)^{-1}A_n\right)^{-1} \end{array}\right. , \quad z\in\mathbb{C}-\mathbb{R}^+ \tag{4.3}$$

admits a unique solution $(\delta_n, \tilde{\delta}_n)$ in the class of Stieltjes transforms of nonnegative measures[1] with support in $\mathbb{R}^+$. Matrices $T_n(z)$ and $\tilde{T}_n(z)$ defined by

$$\left\{\begin{array}{lcl} T_n(z) &=& \left(-z(I_N+\tilde{\delta}_n(z)D_n)+A_n(I_n+\delta_n\tilde{D}_n)^{-1}A_n^*\right)^{-1} \\ \tilde{T}_n(z) &=& \left(-z(I_n+\delta_n(z)\tilde{D}_n)+A_n^*(I_N+\tilde{\delta}_nD_n)^{-1}A_n\right)^{-1} \end{array}\right. \tag{4.4}$$

are approximations of the resolvent $Q_n(z) = (\Sigma_n\Sigma_n^* - zI_N)^{-1}$ and the co-resolvent $\tilde{Q}_n(z) = (\Sigma_n^*\Sigma_n - zI_N)^{-1}$ in the sense that ($\xrightarrow{a.s.}$ stands for the almost sure convergence):

$$\frac{1}{N}\mathrm{Tr}\ (Q_n(z) - T_n(z)) \xrightarrow[N,n\to\infty]{a.s.} 0\ ,$$

which readily gives a deterministic approximation of the Stieltjes transform $N^{-1}\mathrm{Tr}\, Q_n(z)$ of the spectral measure of $\Sigma_n\Sigma_n^*$ in terms of T_n (and similarly for $\tilde{Q}_n$ and $\tilde{T}_n$). Also proved in [38] is the convergence of bilinear forms

$$u_n^*(Q_n(z) - T_n(z))v_n \xrightarrow[N,n\to\infty]{a.s.} 0\ , \tag{4.5}$$

where (u_n) and (v_n) are sequences of $N\times 1$ deterministic vectors with uniformily bounded euclidian norm, which complements the picture of T_n approximating Q_n.

Matrices $T_n = (t_{ij}\ ; 1\le i,j\le N)$ and $\tilde{T}_n = (\tilde{t}_{ij}\ ; 1\le i,j\le n)$ will play a fundamental role in the sequel and enable us to express a deterministic equivalent to $\mathbb{E}\mathcal{I}_n(\rho)$. Define $V_n(\rho)$ by:

$$V_n(\rho) = \frac{1}{N}\log\det\left(\rho(I_N+\tilde{\delta}_nD_n)I_N + A_n(I_n+\delta_n\tilde{D}_n)^{-1}A_n^*\right) + \frac{1}{N}\log(I_n+\delta_n\tilde{D}_n) - \frac{\rho n}{N}\delta_n\tilde{\delta}_n\ , \tag{4.6}$$

where δ_n and $\tilde{\delta}_n$ are evaluated at $z=-\rho$. Then the difference $\mathbb{E}\,\mathcal{I}_n(\rho) - V_n(\rho)$ goes to zero as $N,n\to\infty$.

The general approach

The approach developed in this article is conceptually simple. The quantity $\mathcal{I}_n(\rho) - \mathbb{E}\mathcal{I}_n(\rho)$ is decomposed into a sum of increments of martingale; we then rely on a central limit theorem for martingales, and in order to identify the variance, systematically approximate

[1] In fact, δ_n is the Stieltjes transform of a measure with total mass equal to $n^{-1}\mathrm{Tr}D_n$ while $\tilde{\delta}_n$ is the Stieltjes transform of a measure with total mass equal to $n^{-1}\mathrm{Tr}\tilde{D}_n$.

random quantities by their deterministic counterparts as the size of the vectors and the matrices goes to infinity. The martingale method which is used to establish the fluctuations of $\mathcal{I}_n(\rho)$ can be traced back to Girko's REFORM (REsolvent, FORmula and Martingale) method (see [28, 29]) and is close to the one developed in [3].

In this study, the fact that matrix Σ_n is non-centered ($\mathbb{E}\,\Sigma_n = A_n$) raises specific issues, from a different nature than those addressed in close-by results [1, 3, 37], etc. These issues arise from the presence in the computations of bilinear forms $u_n^* Q_n(z)\, v_n$ where at least one of the vectors u_n or v_n is deterministic. Often, the deterministic vector is related to the columns of matrix A_n, and has to be dealt with in such a way that the assumption over the spectral norm of A_n is exploited. These issues are partly circumvented by the convergence result (4.5) established in [38].

The fluctuations

In every case where the fluctuations of the mutual information have been studied, the variance of $N\left(\mathcal{I}_n(\rho) - V_n(\rho)\right)$ always proved to take a (somehow unexpected) remarkably simple closed-form expression (see for instance [55, 78, 85] and in a more mathematical flavour [36, 37]). The same phenomenon again occurs for the matrix model Σ_n under consideration. Drop the subscripts N, n and let

$$\gamma = \frac{1}{n}\mathrm{Tr}\, DTDT\ ,\ \tilde{\gamma} = \frac{1}{n}\mathrm{Tr}\, \tilde{D}\tilde{T}\tilde{D}\tilde{T}\ , \tag{4.7}$$

Let $\kappa = \mathbb{E}|X_{ij}|^4 - 2$ and denote by

$$\Theta_n^2 = -\log\left(\left(1 - \frac{1}{n}\mathrm{Tr}\, D^{\frac{1}{2}} T A (I + \tilde{\delta} D)^{-1} \tilde{D} (I + \tilde{\delta} D)^{-1} A^* T D^{\frac{1}{2}}\right)^2 - \rho^2 \gamma\tilde{\gamma}\right) + \kappa\, \frac{\rho^2}{n^2} \sum_i d_i^2 t_{ii}^2 \sum_j \tilde{d}_j^2 \tilde{t}_{jj}^2 \tag{4.8}$$

where $d_i = [D_n]_{ii}$, $\tilde{d}_j = [\tilde{D}_n]_{jj}$, and all the needed quantities are evaluated at $z = -\rho$. The CLT then expresses as:

$$\frac{N}{\Theta_n}\left(\mathcal{I}_n - \mathbb{E}\mathcal{I}_n\right) \xrightarrow[N,n\to\infty]{\mathcal{D}} \mathcal{N}(0,1)\ ,$$

where $\xrightarrow{\mathcal{D}}$ stands for the convergence in distribution.

2 The Central Limit Theorem for $\mathcal{I}_n(\rho)$

The indicator function of the set $\mathcal{A}$ will be denoted by $\mathbf{1}_{\mathcal{A}}(x)$, its cardinality by $\#\mathcal{A}$. As usual, $\mathbb{R}^+ = \{x \in \mathbb{R}\ :\ x \geq 0\}$ and $\mathbb{C}^+ = \{z \in \mathbb{C}\ :\ \Im(z) > 0\}$ and $\mathbf{i} = \sqrt{-1}$; if $z \in \mathbb{C}$, then $\bar{z}$ stands for its complex conjugate. Denote by $\xrightarrow{\mathcal{P}}$ the convergence in probability of random variables and by $\xrightarrow{\mathcal{D}}$ the convergence in distribution of probability measures. Denote by $\mathrm{diag}(a_i;\ 1 \leq i \leq k)$ the $k \times k$ diagonal matrix whose diagonal entries are the a_i's. Element (i, j) of matrix M will be either denoted m_{ij} or $[M]_{ij}$ depending on the notational context. if M is a $n \times n$ square matrix, $\mathrm{diag}(M) = \mathrm{diag}(m_{ii}; 1 \leq i \leq n)$. Denote by M^T the matrix transpose of M, by M^* its Hermitian adjoint, by $\mathrm{Tr}\,(M)$ its trace and $\det(M)$ its determinant (if M is square). When dealing with vectors, $\|\cdot\|$ will refer to the Euclidean norm, and $\|\cdot\|_\infty$, to the max (or ℓ_∞) norm. In the case of matrices, $\|\cdot\|$ will refer to the spectral norm.

Recall that $\Sigma_n = n^{-1/2} D_n^{1/2} X_n \tilde{D}_n^{1/2} + A_n$; denote $D_n = \mathrm{diag}(d_i,\ 1 \le i \le N)$ and $\tilde{D}_n = \mathrm{diag}(\tilde{d}_j,\ 1 \le j \le n)$. When no confusion can occur, we shall often drop subscripts and superscripts n for readability.

Assumption A-1 *The random variables $(X_{ij}^n\ ;\ 1 \le i \le N, 1 \le j \le n\,, n \ge 1)$ are complex, independent and identically distributed. They satisfy*

$$\mathbb{E} X_{ij}^n = 0, \quad \mathbb{E}|X_{ij}^n|^2 = 1 \quad \text{and} \quad \mathbb{E}|X_{ij}^n|^{16} < \infty\ .$$

Assumption A-2 *The random variables X_{ij} satisfy the following circularity condition,*

$$\mathbb{E} X_{ij}^p \bar{X}_{ij}^q = 0 \quad \textit{for} \quad p \ne q, \quad p,q \ge 0, \quad \forall i,j \tag{4.9}$$

Assumption A-3 *The family of deterministic $N \times n$ matrices $(A_n, n \ge 1)$ is uniformly bounded for the spectral norm:*

$$\lambda_{\max} = \sup_{n \ge 1} \|A_n\| < \infty\ .$$

Assumption A-4 *The families of real deterministic $N \times N$ and $n \times n$ matrices (D_n) and $(\tilde{D}_n)$ are diagonal with non-negative diagonal elements, and are bounded for the spectral norm as $N, n \to \infty$:*

$$\boldsymbol{d}_{\max} = \sup_{n \ge 1} \|D_n\| < \infty \quad \textit{and} \quad \tilde{\boldsymbol{d}}_{\max} = \sup_{n \ge 1} \|\tilde{D}_n\| < \infty\ .$$

Moreover,

$$\boldsymbol{d}_{\min} = \inf_N \frac{1}{N} \mathrm{Tr}\, D_n > 0 \quad \textit{and} \quad \tilde{\boldsymbol{d}}_{\min} = \inf_n \frac{1}{n} \mathrm{Tr}\, \tilde{D}_n > 0\ .$$

We can now state the main theorem of the article.

Theorem 2.1 (The CLT) *Consider the $N \times n$ matrix $\Sigma_n = n^{-1/2} D_n^{1/2} X_n \tilde{D}_n^{1/2} + A_n$ and assume that* **A-1**, **A-3** *and* **A-4** *hold true. Recall the definitions of $\tilde{\delta}$ given by* (4.3), *T and $\tilde{T}$ given by* (4.4), *and γ and $\tilde{\gamma}$ given by* (4.7). *Let $\rho > 0$. All the considered quantities are evaluated at $z = -\rho$. Then:*

1. *the quantity*

$$\Theta_n^2 = -\log\left(\left(1 - \frac{1}{n} \mathrm{Tr}\, D^{\frac{1}{2}} T A (I + \delta \tilde{D})^{-1} \tilde{D} (I + \delta \tilde{D})^{-1} A^* T D^{\frac{1}{2}}\right)^2 - \rho^2 \gamma \tilde{\gamma}\right) + \kappa \frac{\rho^2}{n^2} \sum_i d_i^2 t_{ii}^2 \sum_j \tilde{d}_j^2 \tilde{t}_{jj}^2$$

   is well-defined and satisfies

$$0 \quad < \quad \liminf_n \Theta_n^2 \quad \le \quad \limsup_n \Theta_n^2 \quad < \quad \infty$$

   as $N, n \to \infty$.

2. *Consider the random variable* $\mathcal{I}_n(\rho) = N^{-1}\log\det\left(\Sigma_n\Sigma_n^* + \rho I_N\right)$, *then the following convergence holds true:*

$$\frac{N}{\Theta_n}\left(\mathcal{I}_n - \mathbb{E}\mathcal{I}_n\right) \xrightarrow[N,n\to\infty]{\mathcal{D}} \mathcal{N}(0,1)\ ,$$

3 Controls over the varaince Θ_n^2

3.1 Controls over Θ_n^2

The following estimates will be usefull.

Proposition 3.1 *Assume that the setting of Theorem 2.1 holds true. Then:*

1. *The quantities δ and $\tilde{\delta}$, evaluated at $z = -\rho$, satisfy,*

$$\begin{aligned}
\delta_{\min} &= \frac{\boldsymbol{d}_{\min}}{\rho + \boldsymbol{d}_{\max}\tilde{\boldsymbol{d}}_{\max} + \|a_{\max}\|^2} \le \delta \le \frac{l^+}{\rho}\boldsymbol{d}_{\max} = \delta_{\max}, \quad \textit{and} \\
\tilde{\delta}_{\min} &= \frac{\tilde{\boldsymbol{d}}_{\min}}{\rho + \tilde{\boldsymbol{d}}_{\max}\boldsymbol{d}_{\max} + \|a_{\max}\|^2} \le \tilde{\delta} \le \frac{\tilde{\boldsymbol{d}}_{\max}}{\rho} = \tilde{\delta}_{\max}.
\end{aligned}$$

2. *The quantities γ and $\tilde{\gamma}$, evaluated at $z = -\rho$, satisfy,*

$$\begin{aligned}
\frac{l^-\boldsymbol{d}_{\min}^2}{\rho^2\left(1 + \boldsymbol{d}_{\max}\tilde{\delta}_{\max} + \rho^{-1}\|a_{\max}\|^2\right)^2} &\le \gamma \le \frac{l^+}{\rho^2}, \quad \textit{and} \\
\frac{\tilde{\boldsymbol{d}}_{\min}^2}{\rho^2\left(1 + \tilde{\boldsymbol{d}}_{\max}\delta_{\max} + \rho^{-1}\|a_{\max}\|^2\right)^2} &\le \tilde{\gamma} \le \frac{1}{\rho^2}.
\end{aligned}$$

3. *The quantities $\underline{\gamma}$ and $\underline{\tilde{\gamma}}$, evaluated at $z = -\rho$, satisfy,*

We are now in position to prove the first part of Theorem 2.1.

Proof 3 (Proof of Theorem 2.1-(1)) *We first prove that:* $\liminf \Delta_n > 0$, *where,*

$$\Delta_n = \left(1 - \frac{\rho^2}{n}\mathrm{Tr}\, D^{\frac{1}{2}}TA\tilde{\Psi}\tilde{D}\tilde{\Psi}A^*TD^{\frac{1}{2}}\right)^2 - \rho^2\gamma\tilde{\gamma}$$

Proof of $\liminf \Delta_n > 0$. *We have,*

$$\rho\tilde{\Psi}\tilde{D} = \left(I_n + \delta\tilde{D}\right)^{-1}\tilde{D} = \frac{1}{\delta}I_n - \frac{1}{\delta}\left(I_n + \delta\tilde{D}\right)^{-1} \le \frac{1}{\delta}I_n.$$

Then,

$$1 - \frac{\rho^2}{n}\mathrm{Tr}\ TA\tilde{\Psi}\tilde{D}\tilde{\Psi}A^*TD \ge 1 - \frac{\rho}{\delta n}\mathrm{Tr}\ TA\tilde{\Psi}A^*TD.$$

Moreover, we have, $TA\tilde{\Psi}A^* = \frac{1}{\rho}I_N + \frac{1}{\rho}T\Psi^{-1}$*, then we obtain,*

$$\begin{aligned}
1 - \frac{\rho^2}{n}\mathrm{Tr}\; TA\tilde{\Psi}\tilde{D}\tilde{\Psi}A^*TD &\geq 1 - \frac{1}{\delta n}\mathrm{Tr}\; TD + \frac{1}{\delta n}\mathrm{Tr}\; T\Psi^{-1}TD \\
&\geq 1 - \frac{1}{\delta n}\mathrm{Tr}\; TD + \frac{\rho}{\delta n}\mathrm{Tr}\; T^2D + \frac{\rho\tilde{\delta}}{\delta n}\mathrm{Tr}\; TDTD \\
&= 1 - \frac{\delta}{\delta} + \frac{\rho}{\delta n}\mathrm{Tr}\; T^2D + \rho\frac{\tilde{\delta}}{\delta}\gamma \\
&= \rho\frac{\tilde{\delta}}{\delta}\gamma + \frac{\rho}{\delta n}\mathrm{Tr}\; T^2D.
\end{aligned}$$

On the other hand, one can prove that: $\frac{1}{n}\mathrm{Tr}\; TA\tilde{\Psi}\tilde{D}\tilde{\Psi}A^*TD = \frac{1}{n}\mathrm{Tr}\; \tilde{T}A^*\Psi D\Psi A\tilde{T}\tilde{D}$*. Then, the same kind of arguments can be used to prove that:*

$$1 - \frac{\rho^2}{n}\mathrm{Tr}\; TA\tilde{\Psi}\tilde{D}\tilde{\Psi}A^*TD \geq \rho\frac{\delta}{\tilde{\delta}}\tilde{\gamma} + \frac{\rho}{\tilde{\delta} n}\mathrm{Tr}\; \tilde{T}^2\tilde{D}.$$

We therefore have,

$$\left(1 - \frac{\rho^2}{n}\mathrm{Tr}\; TA\tilde{\Psi}\tilde{D}\tilde{\Psi}A^*TD\right)^2 \geq \rho^2\gamma\tilde{\gamma} + \varsigma_n.$$

where $\varsigma_n = \rho\frac{\gamma}{\delta n}\mathrm{Tr}\,\tilde{T}^2\tilde{D} + \rho\frac{\tilde{\gamma}}{\tilde{\delta} n}\mathrm{Tr}\,T^2D + \frac{\rho}{\delta n}\mathrm{Tr}\; T^2D\frac{\rho}{\tilde{\delta} n}\mathrm{Tr}\; \tilde{T}^2\tilde{D}$*. Due to proposition 3.1-(1), we can see easily that* $\liminf_n \varsigma_n > 0$*, which ends the proof of* $\liminf_n \Delta_n > 0$.

Proof of $\liminf_n \Theta_n^2 > 0$.

From facts $\frac{1}{n}\mathrm{Tr}\,S^2\frac{1}{n}\mathrm{Tr}\,\tilde{S}^2 \leq \gamma\tilde{\gamma}$*, and* $\mathbb{E}|X_{11}|^4 - 2 \geq -1$*, the following inequalities are justified,*

$$\begin{aligned}
\Theta_n^2 &= -\log\Delta_n + \rho^2\frac{(\mathbb{E}|X_{11}|^4 - 2)}{n}\mathrm{Tr}\,S^2\frac{1}{n}\mathrm{Tr}\,\tilde{S}^2 \\
&\geq -\log\left(\left(1 - \frac{\rho^2}{n}\mathrm{Tr}\; D^{1/2}TA\tilde{\Psi}\tilde{D}\tilde{\Psi}A^*TD^{1/2}\right)^2 - \rho^2\gamma\tilde{\gamma}\right) - \rho^2\gamma\tilde{\gamma} \\
&\geq -\log\left(1 - \rho^2\gamma\tilde{\gamma}\right) - \rho^2\gamma\tilde{\gamma}.
\end{aligned}$$

The function $x \to -\log(1-x) - x$ *is increasing on* $[0,1]$ *and takes the value zero at zero.Therefore, it is sufficient to prove that* $\liminf_n \rho^2\gamma\tilde{\gamma}$ *is bounded away from zero to conclude; this readily follows from the previous calculations. The lower bound is proved.*

Proof of $\limsup_n \Theta_n^2 < \infty$.

The fact that Δ_n *is nonnegative and bounded away from zero together with the following upper bound:*

$$\frac{1}{n}TrS^2\frac{1}{n}Tr\tilde{S}^2 \leq \gamma\tilde{\gamma} \leq \frac{l^+}{\rho^4} < \infty,$$

can garantee the upper boundness of Θ_n^2*, and the proof of theorem 2.1-(1) is done.*

3.2 Notations and classical results

Denote by Y the $N \times n$ matrix $n^{-1/2}D^{1/2}X\tilde{D}^{1/2}$; by (η_j), (a_j) and (y_j) the columns of matrices Σ, A and Y. Denote by Σ_j, A_j, Y_j and $\tilde{D}_j$, the matrices Σ, A, Y and $\tilde{D}$ where column j has been removed. The associated resolvent is $Q_j(z) = (\Sigma_j\Sigma_j^* - zI_N)^{-1}$. Denote by $\mathbb{E}_j$ the conditional expectation with respect to the σ-field $\mathcal{F}_j$ generated by the vectors $(y_\ell,\ 1 \le \ell \le j)$. By convention, $\mathbb{E}_0 = \mathbb{E}$.

We introduce here intermediate quantities of constant use in the rest of this chapter.

$$\tilde{b}_j(z) = \frac{1}{-z\left(1 + a_j^*Q_ja_j + \frac{\tilde{d}_j}{n}\operatorname{Tr} DQ_j\right)}\ , \quad 1 \le j \le n\ , \tag{4.10}$$

$$\tilde{c}_j(z) = \frac{1}{-z\left(1 + a_j^*\mathbb{E}Q_ja_j + \frac{\tilde{d}_j}{n}\operatorname{Tr} D\mathbb{E}Q\right)}\ , \quad 1 \le j \le n\ , \tag{4.11}$$

$$\begin{aligned} e_j(z) &= \eta_j^*Q_j(z)\eta_j - \left(\frac{\tilde{d}_j}{n}\operatorname{Tr} DQ_j(z) + a_j^*Q_j(z)a_j\right) \\ &= \left(y_j^*Q_j(z)y_j - \frac{\tilde{d}_j}{n}\operatorname{Tr} DQ_j(z)\right) + a_j^*Q_j(z)y_j + y_j^*Q_j(z)a_j. \end{aligned} \tag{4.12}$$

Using the well-known caracterization of Stieltjes transforms (see for instance [35, Proposition 2.2-(2)]), one can easily prove that $\tilde{b}_j$ is the Stieltjes transform of a probability measure. In particular $|\tilde{b}_j(z)| \le (\operatorname{dist}(z, \mathbb{R}^+))^{-1}$ for $z \in \mathbb{C} - \mathbb{R}^+$. The same estimate holds for $\tilde{c}_j$. We also introduce the following matrix:

$$C(z) = \left(-z\left(I_N + \tilde{\alpha}D\right) + A\left(I_n + \alpha\tilde{D}\right)^{-1}A^*\right)^{-1}\ , \tag{4.13}$$

where, $\alpha = \frac{1}{n}\operatorname{Tr} D\mathbb{E}Q$ and $\tilde{\alpha} = \frac{1}{n}\operatorname{Tr}\tilde{D}\mathbb{E}\tilde{Q}$.

We also remind classical identities of constant use in the sequel. The first one expresses the diagonal elements of the co-resolvent; the second one is a usefull combination of Woodbury's identity and rank one perturbation identities.

Diagonal elements of the co-resolvent

$$\tilde{q}_{jj}(z) = -\frac{1}{z(1 + \eta_j^*Q_j(z)\eta_j)}\ , \tag{4.14}$$

$$Q(z) = Q_j(z) + z\tilde{q}_{jj}(z)Q_j(z)\eta_j\eta_j^*Q_j(z)\ . \tag{4.15}$$

Note that:

$$\tilde{q}_{jj} = \tilde{b}_j + z\tilde{q}_{jj}\tilde{b}_je_j\ . \tag{4.16}$$

A usefull consequence of (4.15) is:

$$\eta_j^*Q(z) = \frac{\eta_j^*Q_j(z)}{1 + \eta_j^*Q_j(z)\eta_j} = -z\tilde{q}_{jj}(z)\eta_j^*Q_j(z)\ . \tag{4.17}$$

Diagonal elements of matrix $\tilde{T}$

Define the $N \times N$ matrix $\mathcal{T}_j$ as

$$\mathcal{T}_j = \left(-z(I_N + \tilde{\delta} D) + A_j (I_{n-1} + \delta \tilde{D}_j)^{-1} A_j^* \right)^{-1} , \tag{4.18}$$

where δ and $\tilde{\delta}$ are defined in (4.3). Notice the difference between $\mathcal{T}_j$ and T; however, $\mathcal{T}_j$ naturally pops up when expressing the diagonal elements of $\tilde{T}$. Indeed after simple algebra, we obtain:

$$\tilde{t}_{jj}(z) = -\frac{1}{z\left(1 + a_j^* \mathcal{T}_j(z) a_j + \frac{\tilde{d}_j}{n} \operatorname{Tr} DT(z)\right)} . \tag{4.19}$$

We have also the following identity:

$$-z\tilde{t}_{\ell\ell} a_\ell^* \mathcal{T}_\ell b = \frac{a_\ell^* T b}{1+\delta}, \tag{4.20}$$

where, b is a given $N \times 1$ vector.

3.3 Important estimates

Lemma 3.2 *Let $\boldsymbol{x} = (x_1, \cdots, x_n)$ be a $n \times 1$ vector where the x_i are centered i.i.d. complex random variables with unit variance. Denote by $\boldsymbol{y} = n^{-1/2}\boldsymbol{x}$ and let $M = (m_{ij})$ and $P = (p_{ij})$ be a $n \times n$ deterministic complex matrices and D is a $N \times N$ diagonal nonnegative matrix, then:*

1. *(Bai and Silverstein, Lemma 2.7 in [5]) For any $p \geq 2$, there exists a constant K_p for which*

$$\mathbb{E}|\boldsymbol{y}^* M \boldsymbol{y} - \frac{1}{n}\operatorname{Tr} M|^p \leq \frac{K_p}{n^{p/2}} \left(\left(\mathbb{E}|x_1|^4 \frac{\operatorname{Tr} MM^*}{n} \right)^{p/2} + \mathbb{E}|x_1|^{2p} \frac{\operatorname{Tr}(MM^*)^{p/2}}{n^{p/2}} \right) .$$

2. *Let $\boldsymbol{u}$ be a deterministic $n \times 1$ vector. Then:*

$$\begin{aligned} \mathbb{E}\Big[&\Big((D^{1/2}\boldsymbol{y} + \boldsymbol{u})^* M (D^{1/2}\boldsymbol{y} + \boldsymbol{u}) - \mathbb{E}|y_1|^2 \operatorname{Tr} DM - u^* M u \Big) \\ &\qquad \Big((D^{1/2}\boldsymbol{y} + \boldsymbol{u})^* P (D^{1/2}\boldsymbol{y} + \boldsymbol{u}) - \mathbb{E}|y_1|^2 \operatorname{Tr} DP - \boldsymbol{u}^* P \boldsymbol{u} \Big) \Big] \\ &= (\mathbb{E}|y_1|^2)^2 \operatorname{Tr} MDPD + \mathbb{E}|y_1|^2 \left(\boldsymbol{u}^* MDP \boldsymbol{u} + \boldsymbol{u}^* PDM \boldsymbol{u} \right) \\ &\qquad + \kappa \sum_{i=1}^{N} d_{ii}^2 m_{ii} p_{ii} \end{aligned}$$

where $\kappa = \mathbb{E}|y_1|^4 - 2(\mathbb{E}|y_1|^2)^2$.

Remark 3.1 *Using Lemma 3.2-(1), one can prove that:*

$$\mathbb{E}\left(|e_j|^p \mid y_k, k \neq j \right) = \mathcal{O}(n^{-p/2}) . \tag{4.21}$$

which readily implies that $\mathbb{E}|e_j|^p = \mathcal{O}(n^{-p/2})$.

Let M be a sequence of $N \times N$ deterministic matrices with bounded spectral norm. If $p \geq 2$, then:

$$\mathbb{E}|y_1^* M y_1 - \frac{1}{N}\mathrm{Tr}\, M|^p \;\leq\; \frac{K}{n^{p/2}} \; . \tag{4.22}$$

Proposition 3.3 *Assume that the setting of Theorem 2.1 holds true. Let u_n be a deterministic complex $N \times 1$ vector uniformily bounded for the euclidian norm: $\sup_{n\geq 1}\|u_n\| < \infty$. Then, for every $z \in \mathbb{C} - \mathbb{R}^+$, the following estimates hold true:*

$$\sum_{j=1}^{n} \mathbb{E}|u_n^* Q_j a_j|^2 \;\leq\; K_1(z) \quad \textit{and} \quad \mathbb{E}\left(\sum_{j=1}^{n} \mathbb{E}_j |u_n^* Q_j a_j|^2\right)^2 \;\leq\; K_2(z) \; ,$$

where $K_{1,2}(z) < \infty$ do not depend on n, N.

Lemma 3.4 *Control of $\frac{1}{n}\mathrm{Tr}\, D\,(T - \mathbb{E}Q)$ in $\mathbb{R}_-^*$.*
For all $\rho \in \mathbb{R}_+^*$, we have:

$$|\frac{1}{n}\mathrm{Tr}\, D(T - \mathbb{E}Q)| \leq \frac{K}{\sqrt{n}}$$

Lemma 3.5 *Assume that the setting of Theorem 2.1 holds true. Let u_n and v_n be deterministic complex $N \times 1$ vectors uniformily bounded for the euclidian norm: $\sup_{n\geq 1}\max(\|u_n\|, \|v_n\|) < \infty$. Then, for every $z \in \mathbb{C} - \mathbb{R}^+$,*

$$\mathbb{E}|u_n^*(Q(z) - T(z))v_n|^4 \leq \frac{K(z)}{n^2} \; ,$$

where $K(z)$ does not depend on n, N.

Remark 3.2 *Of course, the counterpart of this lemma for the co-resolvent $\tilde{Q}$ and matrix $\tilde{T}$ holds true. In particular taking the vectors u_n and v_n equals to the jth canonical vector and applying Cauchy-Schwarz inequality yield the following estimate:*

$$\mathbb{E}|\tilde{q}_{jj} - \tilde{t}_{jj}|^2 = \mathcal{O}(n^{-1}). \tag{4.23}$$

Proposition 3.6 *Assume that the setting of Theorem 2.1 holds true. Then the following estimates hold*

$$\mathbb{E}|\tilde{q}_{jj} - \tilde{c}_j|^2 \;=\; \mathcal{O}(n^{-1}) \; , \tag{4.24}$$
$$\mathbb{E}|\tilde{q}_{jj} - \tilde{b}_j|^2 \;=\; \mathcal{O}(n^{-1}) \; , \tag{4.25}$$

Lemma 3.7 *The resolvents Q and the perturbed resolvent Q_j satisfy:*

$$|\mathrm{Tr}\; A\,(Q - Q_j)\,| \leq \frac{\|A\|}{|\Im(z)|}$$

for any $N \times N$ bounded spectral norm matrix A.

4 Decomposition of $\mathcal{I}_n - \mathbb{E}\mathcal{I}_n$, Cumulant term in the variance

The main tool we shall rely on to establish the Central Limit theorem is the following Central limit theorem for martingales which can essentially be found in [9]:

Theorem 4.1 (CLT for martingales, Th. 35.12 in [9]) *Let $\gamma_1^{(n)}, \gamma_2^{(n)}, \dots, \gamma_n^{(n)}$ be a martingale difference sequence with respect to the increasing filtration $\mathcal{F}_1^{(n)}, \dots, \mathcal{F}_n^{(n)}$. Assume that there exists a sequence of real positive numbers Υ_n^2 such that*

$$\frac{1}{\Upsilon_n^2}\sum_{j=1}^{n} \mathbb{E}_{j-1}\gamma_j^{(n)2} \xrightarrow[n\to\infty]{\mathcal{P}} 1 \ . \tag{4.26}$$

Assume further that the Lyapounov condition ([9, Section 27]) holds true:

$$\exists \delta > 0, \quad \frac{1}{\Upsilon_n^{2(1+\delta)}}\sum_{j=1}^{n} \mathbb{E}|\gamma_j^{(n)}|^{2+\delta} \xrightarrow[n\to\infty]{} 0 \ .$$

Then $\Upsilon_n^{-1}\sum_{j=1}^{n}\gamma_j^{(n)}$ converges in distribution to $\mathcal{N}(0,1)$.

Remark 4.1 *Note that if moreover* $\liminf_n \Upsilon_n^2 > 0$*, it is sufficient to prove:*

$$\sum_{j=1}^{n} \mathbb{E}_{j-1}\gamma_j^{(n)2} - \Upsilon_n^2 \xrightarrow[n\to\infty]{\mathcal{P}} 0 \ , \tag{4.27}$$

instead of (4.26).

4.1 Decomposition of $\mathcal{I}_n - \mathbb{E}\mathcal{I}_n$ as a sum of increments of martingale

Denote by

$$\Gamma_j = \frac{\eta_j^* Q_j \eta_j - \left(\frac{\tilde{d}_j}{n}\operatorname{Tr} DQ_j + a_j^* Q_j a_j\right)}{1 + \frac{\tilde{d}_j}{n}\operatorname{Tr} DQ_j + a_j^* Q_j a_j} \ .$$

With this notation at hand, the decomposition of $\mathcal{I}_n - \mathbb{E}\mathcal{I}_n$ as

$$\mathcal{I}_n - \mathbb{E}\mathcal{I}_n = -\sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j-1})\log(1+\Gamma_j) \tag{4.28}$$

follows verbatim from [37, Section 6.2]. Moreover, it is a matter of bookeeping to establish the following (cf. [37, Section 6.4]):

$$\sum_{j=1}^{n} \mathbb{E}_{j-1}\left[(\mathbb{E}_j - \mathbb{E}_{j-1})\log(1+\Gamma_j)\right]^2 - \sum_{j=1}^{n}\mathbb{E}_{j-1}(\mathbb{E}_j\Gamma_j)^2 \xrightarrow[N,n\to\infty]{\mathcal{P}} 0 \ . \tag{4.29}$$

Hence, the details are omitted. In view of Theorems 2.1-(1), 4.1 and equations (4.27), (4.28) and (4.29), the CLT will be established if one proves that the Lyapounov condition and the following convergence hold true:

$$\exists \delta > 0, \quad \frac{1}{\Theta_n^{2(1+\delta)}} \sum_{j=1}^{n} \mathbb{E}|\mathbb{E}_j \Gamma_j|^{2+\delta} \quad \xrightarrow[n\to\infty]{} \quad 0\ , \tag{4.30}$$

$$\text{and} \qquad \sum_{j=1}^{n} \mathbb{E}_{j-1}(\mathbb{E}_j \Gamma_j)^2 - \Theta_n^2 \quad \xrightarrow[N,n\to\infty]{\mathcal{P}} \quad 0\ , \tag{4.31}$$

where Θ_n^2 is defined in Theorem 2.1.

4.2 Further decomposition of $\mathcal{I}_n - \mathbb{E}\mathcal{I}_n$

Notice that $\mathbb{E}_{j-1}(\mathbb{E}_j\Gamma_j)^2 = \mathbb{E}_{j-1}(\mathbb{E}_j z\tilde{b}_j e_j)^2$. We prove hereafter that

$$\sum_{j=1}^{n} \mathbb{E}_{j-1}(\mathbb{E}_j z\tilde{b}_j e_j)^2 - \sum_{j=1}^{n} z^2 \tilde{t}_{jj}^2 \mathbb{E}_{j-1}(\mathbb{E}_j e_j)^2 \xrightarrow[N,n\to\infty]{\mathcal{P}} 0\ . \tag{4.32}$$

Using the triangular inequality together with estimates (4.23) and (4.25) yields that $\mathbb{E}|\tilde{b}_j - \tilde{t}_{jj}|^2 = \mathcal{O}(n^{-1})$. Now this estimate, together with (4.21), readily implies that:

$$\mathbb{E}|\mathbb{E}_{j-1}(\mathbb{E}_j z\tilde{b}_j e_j)^2 - z^2\tilde{t}_{jj}^2 \mathbb{E}_{j-1}(\mathbb{E}_j e_j)^2| = \mathcal{O}(n^{-3/2})\ ,$$

hence (4.32). Using the identity in Lemma 3.2-(2), we develop the quantity $\mathbb{E}_{j-1}(\mathbb{E}_j e_j)^2$:

$$\begin{aligned}\sum_{j=1}^{n} z^2\tilde{t}_{jj}^2 \mathbb{E}_{j-1}(\mathbb{E}_j e_j)^2 &= \frac{\kappa}{n}\sum_{j=1}^{n} z^2\tilde{t}_{jj}^2 \mathbb{E}_{j-1}\left(\frac{1}{n}\sum_{\ell\ell}^{N}(\mathbb{E}_j q_{\ell\ell}^{(j)})^2\right) + \frac{1}{n}\sum_{j=1}^{n} z^2\tilde{t}_{jj}^2\left(\frac{1}{n}\mathrm{Tr}\,(\mathbb{E}_j Q_j)^2 + 2\mathbb{E}_j a_j^*(\mathbb{E}_j Q_j)^2 a_j\right)\\ &\triangleq \sum_{j=1}^{n}\chi_{1j} + \sum_{j=1}^{n}\chi_{2j}\end{aligned} \tag{4.33}$$

4.3 Computation of the cumulant term of the variance:

Write,

$$\begin{aligned}&\frac{1}{n}\sum_{\ell=1}^{N}(\mathbb{E}_{j-1}[Q_j]_{\ell\ell})^2 - \frac{1}{n}\sum_{\ell=1}^{N}(\mathbb{E}_{j-1}[Q_j]_{\ell\ell})t_{\ell\ell} = \\ &\qquad \frac{1}{n}\sum_{\ell=1}^{N}(\mathbb{E}_{j-1}[Q_j]_{\ell\ell})(\mathbb{E}_{j-1}[Q_j]_{\ell\ell} - \mathbb{E}_{j-1}q_{\ell\ell}) + \frac{1}{n}\sum_{\ell=1}^{N}(\mathbb{E}_{j-1}[Q_j]_{\ell\ell})(\mathbb{E}_{j-1}q_{\ell\ell} - t_{\ell\ell})\end{aligned}$$

The first term is a deterministic $\mathcal{O}(n^{-1})$ by the rank one perturbation inequality. The convergence to zero in probability of the second term can easily be handled by Lemma 3.5. Hence,

$$\sum_{j=1}^{n}\chi_{1j} - \frac{\kappa z^2}{n}\sum_{j=1}^{n}\tilde{t}_{jj}^2\left(\frac{1}{n}\sum_{\ell=1}^{N}(\mathbb{E}_{j-1}[Q_j]_{\ell\ell})t_{\ell\ell}\right) \xrightarrow[N,n\to\infty]{\mathcal{P}} 0\ .$$

Iterating the same arguments, we can replace the remaining term $\mathbb{E}_{j-1}[Q_j]_{\ell\ell}$ by $t_{\ell\ell}$ and finally obtain:

$$\sum_{j=1}^{n}\chi_{1j} - \frac{\kappa z^2}{n^2}\sum_{j=1}^{n}\sum_{\ell=1}^{N}\tilde{t}_{jj}^2 t_{\ell\ell}^2 \xrightarrow[N,n\to\infty]{\mathcal{P}} 0\ .$$

5 Identification of the variance as Θ_n^2

In this part, we handle the term $\sum_j^n \chi_{2j}$. For this end, we shall use the result of the following lemma which says that we can replace $\sum_j^n \chi_{2j}$ by its expectation. In the sequel, we take $z = -\rho$.

Lemma 5.1 *For any $N \times 1$ vector a with bounded euclidean norm, we have,*

$$\max_j var(a^*(\mathbb{E}_j Q)^2 a) \leq \frac{K}{\sqrt{n}}.$$

Moreover,

$$\max_j var\left(\frac{1}{n}\mathrm{Tr}\ (\mathbb{E}_j Q)^2\right) \leq \frac{K}{n}.$$

5.1 Study of the gaussian part of the variance

The aim of this part is to prove the following convergence:

$$\sum_{j=1}^n \chi_{2j} + \log(\Delta_n) \xrightarrow[n\to\infty]{\mathcal{P}} 0 \tag{4.34}$$

where, $\Delta_n = \left(1 - \frac{\rho^2}{n}\mathrm{Tr}\ D^{\frac{1}{2}}TA\tilde{\Psi}\tilde{D}\tilde{\Psi}A^*TD^{\frac{1}{2}}\right)^2 - \rho^2\gamma\tilde{\gamma}$. The proof of this convergence will be carried out in three steps:

1. Thanks to the rank one perturbation result, we can deal with $\phi_j = \frac{1}{n}\mathrm{Tr}\, D\mathbb{E}_j QDQ$ instead of $\frac{1}{n}\mathrm{Tr}\, D\mathbb{E}_j Q_j DQ_j$. Let $\theta_{kj} = a_k^*(\mathbb{E}_j Q_k)Q_k a_k$. We prove that:

$$\mathbb{E}(\phi_j) = \gamma + \alpha_j\mathbb{E}(\phi_j) + \frac{\gamma}{n}\sum_{k=1}^n \rho^2\tilde{t}_{kk}^2\tilde{d}_k\mathbb{E}(\theta_{kj}) + \varepsilon_j, \tag{4.35}$$

   where, $\alpha_j = \frac{1}{n}\sum_{\ell=1}^j \rho^2\tilde{\psi}_\ell^2\tilde{d}_\ell a_\ell^* TDTa_\ell + \frac{\gamma}{n}\sum_{\ell=1}^j \rho^2\tilde{t}_{\ell\ell}^2\tilde{d}_\ell^2$ and $\varepsilon_j \xrightarrow[n\to\infty]{\mathcal{P}} 0$

2. We introduce the intermediate quantities: $\zeta_{kj} = a_k^*(\mathbb{E}_j Q)DQa_k$. We then prove the following equations:

$$\mathbb{E}(\zeta_{kj}) = a_k^*TDTa_k + \beta_{kj}\mathbb{E}(\phi_j) + \frac{a_k^*TDTa_k}{n}\sum_{\ell=1}^j \rho^2\tilde{t}_{\ell\ell}^2\tilde{d}_\ell\mathbb{E}(\theta_{\ell j}) + \varepsilon_{kj}, \tag{4.36}$$

$$\rho^2\tilde{t}_{kk}^2\tilde{d}_k\mathbb{E}(\theta_{kj}) = \frac{\tilde{d}_k}{\left(1+\tilde{d}_k\delta\right)^2}\mathbb{E}(\zeta_{kj}) - \frac{\tilde{d}_k}{\left(1+\tilde{d}_k\delta\right)^2}\left(\frac{\tilde{d}_k a_k^*Ta_k}{1+\tilde{d}_k\delta}\right)^2\mathbb{E}(\psi_j) + \varepsilon_{kj}, \tag{4.37}$$

   where,

$$\beta_{kj} = \sum_{l=1}^j \rho^3\tilde{\psi}_l\tilde{t}_{\ell\ell}^2\tilde{d}_l^2 a_k^*\mathcal{T}_l a_l a_l^*\mathbb{E}Q_l a_k + a_k^*TDTa_k\frac{\rho^2}{n}\sum_{l=1}^j \tilde{d}_l^2\tilde{t}_{\ell\ell}^2 + \sum_{l=1}^j \rho^2\tilde{t}_{\ell\ell}^2\tilde{d}_l^2 a_l^*\mathcal{T}_l a_k a_k^*Ta_l$$

   and ε_{kj} converges in probability to zero.

3. We finally prove (4.34).

Proof of (4.35)

Recall that,

$$Q = (\Sigma\Sigma^* + \rho I_N)^{-1} \quad \text{and} \quad T = \left(\Psi^{-1} + \rho A\tilde{\Psi}A^*\right)^{-1}.$$

The resolvent identity expressed by: $Q = T + T\left(T^{-1} - Q^{-1}\right)Q$ gives,

$$Q = T + \frac{\rho}{n}(\mathrm{Tr}\ \tilde{D}\tilde{T})TDQ + \rho TA\tilde{\Psi}A^*Q - T\Sigma\Sigma^*Q.$$

Plugging this in the expression of ϕ_j we obtain,

$$\begin{aligned}
\phi_j &= \frac{1}{n}\mathrm{Tr}\ DTDQ + (\frac{\rho}{n}\mathrm{Tr}\ \tilde{D}\tilde{T})\frac{1}{n}\mathrm{Tr}\ DTD(\mathbb{E}_jQ)DQ \\
&\quad + \frac{\rho}{n}\mathrm{Tr}\ DTA\tilde{\Psi}A^*(\mathbb{E}_jQ)DQ - \frac{1}{n}\mathrm{Tr}\ DT\mathbb{E}_j(\Sigma\Sigma^*Q)DQ \\
&= \chi_1 + \chi_2 + \chi_3 + \chi_4.
\end{aligned}$$

Treatment of the term χ_1.
We have : $\chi_1 = \frac{1}{n}\mathrm{Tr}\ (D\mathbf{T})^2 + \varepsilon_n$, where $\mathbb{E}|\varepsilon_n| = \mathbb{E}|\frac{1}{n}\mathrm{Tr}\ D\mathbf{T}D\,(Q - \mathbf{T})\,| \leq \frac{K}{\sqrt{n}}$ by lemma (3.4).
Treatment of the term χ_3. Using identity (4.15), χ_3 verifies:

$$\begin{aligned}
\chi_3 &= \frac{\rho}{n}\mathrm{Tr}\ DTA\tilde{\Psi}A^*(\mathbb{E}_jQ)DQ \\
&= \frac{\rho}{n}\mathrm{Tr}\ DTA\tilde{\Psi}A^*(\mathbb{E}_jQ_\ell)DQ - \frac{\rho^2}{n}\mathrm{Tr}\ DTA\tilde{\Psi}A^*\mathbb{E}_j(\tilde{q}_{\ell\ell}Q_\ell\eta_\ell\eta_\ell^*Q_\ell)DQ \\
&= \frac{\rho}{n}\sum_{\ell=1}^{n}\tilde{\psi}_\ell a_\ell^*(\mathbb{E}_jQ_\ell)DQDTa_\ell - \frac{\rho^2}{n}\sum_{\ell=1}^{n}\tilde{t}_{\ell\ell}\tilde{\psi}_\ell a_\ell^*\mathbb{E}_j(Q_\ell\eta_\ell\eta_\ell^*Q_\ell)DQDTa_\ell + \varepsilon_{3,1} \\
&= X_1 + X_2 + \varepsilon_{3,1},
\end{aligned}$$

where $\varepsilon_{3,1} = \frac{\rho^2}{n}\sum_{\ell=1}^{n}\tilde{t}_{\ell\ell}\tilde{\psi}_\ell a_\ell^*\mathbb{E}_j((\tilde{q}_{\ell\ell} - \tilde{t}_{\ell\ell})Q_\ell\eta_\ell\eta_\ell^*Q_\ell)DQD\mathbf{T}a_\ell$.
Let us dealing with term X_2. We have:

$$\begin{aligned}
X_2 &= -\frac{\rho^2}{n}\sum_{l=1}^{n}\tilde{t}_{\ell\ell}\tilde{\psi}_l a_l^*\mathbb{E}_j(Q_l a_l\eta_l^*Q_l)DQDTa_l + \varepsilon_{3,2} \\
&= -\frac{\rho^2}{n}\sum_{l=1}^{n}\tilde{t}_{\ell\ell}\tilde{\psi}_l a_l^*\mathcal{T}_l a_l\mathbb{E}_j(\eta_l^*Q_l)DQDTa_l + \varepsilon_{3,3} \\
&= -\frac{\rho^2}{n}\sum_{l=1}^{n}\tilde{t}_{\ell\ell}\tilde{\psi}_l a_l^*\mathcal{T}_l a_l a_l^*\mathbb{E}_j(Q_l)DQDTa_l - \frac{\rho^2}{n}\sum_{\ell=1}^{j}\tilde{t}_{\ell\ell}\tilde{\psi}_\ell a_\ell^*\mathcal{T}_\ell a_\ell y_\ell^*\mathbb{E}_j(Q_\ell)DQ_\ell DTa_\ell \\
&\quad + \frac{\rho^3}{n}\sum_{\ell=1}^{j}\tilde{t}_{\ell\ell}^2\tilde{\psi}_\ell a_\ell^*\mathcal{T}_\ell a_\ell y_\ell^*\mathbb{E}_j(Q_\ell)DQ_\ell\eta_\ell\eta_\ell^*Q_\ell DTa_\ell + \varepsilon_{3,3} \\
&= -\frac{\rho^2}{n}\sum_{\ell=1}^{n}\tilde{t}_{\ell\ell}\tilde{\psi}_\ell a_\ell^*\mathcal{T}_\ell a_\ell a_\ell^*\mathbb{E}_j(Q_\ell)DQDTa_\ell + \frac{\rho^3}{n}\sum_{\ell=1}^{j}\tilde{t}_{\ell\ell}^2\tilde{\psi}_\ell a_\ell^*\mathcal{T}_\ell a_\ell y_\ell^*\mathbb{E}_j(Q_\ell)DQ_\ell y_\ell a_\ell^*Q_\ell DTa_\ell + \varepsilon_{3,4} \\
&= -\frac{\rho^2}{n}\sum_{\ell=1}^{n}\tilde{t}_{\ell\ell}\tilde{\psi}_\ell a_\ell^*\mathcal{T}_\ell a_\ell a_\ell^*\mathbb{E}_j(Q_\ell)DQDTa_\ell + \phi_j\,\frac{\rho^3}{n}\sum_{\ell=1}^{j}\tilde{d}_\ell\tilde{t}_{\ell\ell}^2\tilde{\psi}_\ell a_\ell^*\mathcal{T}_\ell a_\ell a_\ell^*\mathcal{T}_\ell DTa_\ell + \varepsilon_{3,5}.
\end{aligned}$$

Management of epsilons:

1. We have:

$$
\begin{aligned}
|\varepsilon_{3,2}| &= |\frac{\rho^2}{n}\sum_{\ell=1}^{n}\tilde{t}_{\ell\ell}\tilde{\psi}_\ell a_\ell^*\mathbb{E}_j(Q_\ell y_\ell \eta_\ell^* Q_\ell)DQDTa_\ell| \\
&\le K\mathbb{E}|\mathbb{E}_j(a_\ell^* Q_\ell y_\ell \eta_\ell^* Q_\ell)DQDTa_\ell| \le K\mathbb{E}^{1/2}|a_\ell^* Q_\ell y_\ell|^2\mathbb{E}^{1/2}\|\eta_\ell^* Q_\ell\|^2 \le \frac{K}{\sqrt{n}}
\end{aligned}
$$

   the last inequality follows from lemma (3.2-1).

2. Thanks to lemma (3.5), we have, $\mathbb{E}|\varepsilon_{3,3}| \le \frac{K}{\sqrt{n}}$.

3. We have:

$$
\begin{aligned}
|\mathbb{E}\varepsilon_{3,4}| &= |\mathbb{E}\left(\frac{\rho^3}{n}\sum_{\ell=1}^{j}\tilde{t}_{\ell\ell}^2\tilde{\psi}_\ell a_\ell^*\mathcal{T}_\ell a_\ell y_\ell^*\mathbb{E}_j(Q_\ell)DQ_\ell\eta_\ell\eta_\ell^*Q_\ell DTa_\ell + \varepsilon_{3,3}\right)| \\
&\le K\mathbb{E}|y_\ell^*\mathbb{E}_j(Q_\ell)DQ_\ell a_\ell y_\ell^* Q_\ell DTa_\ell| + \mathbb{E}|\varepsilon_{3,3}| \le \frac{K}{\sqrt{n}}
\end{aligned}
$$

4. Finally, we have, $\varepsilon_{3,5} = \frac{\rho^3}{n}\sum_{\ell=1}^{j}\tilde{t}_{\ell\ell}^2\tilde{\psi}_\ell a_\ell^*\mathcal{T}_\ell a_\ell\left(y_\ell^*\mathbb{E}_j(Q_\ell)DQ_\ell y_\ell - \phi_j\right)a_\ell^*Q_\ell DTa_\ell + \varepsilon_{3,4}$, and identity (4.20) together with lemma (3.2-1) guarantee that $\mathbb{E}|\varepsilon_{3,5}| \le \frac{K}{\sqrt{n}}$.

Then χ_3 becomes:

$$
\begin{aligned}
\chi_3 &= \frac{\rho}{n}\sum_{\ell=1}^{n}\tilde{\psi}_\ell a_\ell^*(\mathbb{E}_jQ_\ell)DQDTa_\ell - \frac{\rho^2}{n}\sum_{\ell=1}^{n}\tilde{t}_{\ell\ell}\tilde{\psi}_\ell a_\ell^*\mathcal{T}_\ell a_\ell a_\ell^*\mathbb{E}_j(Q_\ell)DQDTa_\ell \\
&\quad +\phi_j\frac{\rho^3}{n}\sum_{\ell=1}^{j}\tilde{d}_\ell\tilde{\psi}_\ell^3 a_\ell^*Ta_\ell a_\ell^*TDTa_\ell + \varepsilon_n \\
&= \frac{\rho}{n}\sum_{\ell=1}^{n}\tilde{\psi}_\ell\left(1-\rho\tilde{t}_{\ell\ell}a_\ell^*\mathcal{T}_\ell a_\ell\right)a_\ell^*(\mathbb{E}_jQ_\ell)DQDTa_\ell + \phi_j\frac{\rho^3}{n}\sum_{\ell=1}^{j}\tilde{d}_\ell\tilde{\psi}_\ell^3 a_\ell^*Ta_\ell a_\ell^*TDTa_\ell + \varepsilon_{3j},
\end{aligned}
$$

where $\max_j \mathbb{E}|\varepsilon_{3j}| \le \frac{K}{\sqrt{n}}$. Using identity (4.20), one can easily prove that: $\rho\tilde{\psi}_\ell\left(1-\rho\tilde{t}_{\ell\ell}a_\ell^*\mathcal{T}_\ell a_\ell\right) = \rho\tilde{t}_{\ell\ell}$. We therefore obtain,

$$
\chi_3 = \frac{1}{n}\sum_{\ell=1}^{n}\rho\tilde{t}_{\ell\ell}a_\ell^*(\mathbb{E}_jQ_\ell)DQD\mathbf{T}a_\ell + \phi_j\frac{\rho^3}{n}\sum_{\ell=1}^{j}\tilde{d}_\ell\tilde{\psi}_\ell^3 a_\ell^*Ta_\ell a_\ell^*TDTa_\ell + \varepsilon_{3j}.
$$

Treatment of term χ_4.

Using identity $\eta_l^* Q = \rho \tilde{q}_{\ell\ell} \eta_l^* Q_l$, we have,

$$
\begin{aligned}
\chi_4 &= -\frac{1}{n} \mathrm{Tr}\ DT \mathbb{E}_j (\Sigma\Sigma^* Q) DQ \\
&= -\frac{1}{n} \mathrm{Tr}\ DT \sum_{\ell=1}^{n} \mathbb{E}_j (\eta_\ell \eta_\ell^* Q) DQ \\
&= -\frac{1}{n} \mathrm{Tr}\ DT \sum_{\ell=1}^{n} \rho \tilde{t}_{\ell\ell} \mathbb{E}_j (y_\ell y_\ell^* Q_\ell) DQ - \frac{1}{n} \mathrm{Tr}\ DT \sum_{\ell=1}^{n} \rho \tilde{t}_{\ell\ell} \mathbb{E}_j (y_\ell a_\ell^* Q_\ell) DQ \\
&\quad - \frac{1}{n} \mathrm{Tr}\ DT \sum_{\ell=1}^{n} \rho \tilde{t}_{\ell\ell} \mathbb{E}_j (a_\ell y_\ell^* Q_\ell) DQ - \frac{1}{n} \mathrm{Tr}\ DT \sum_{\ell=1}^{n} \rho \tilde{t}_{\ell\ell} \mathbb{E}_j (a_\ell a_\ell^* Q_\ell) DQ + \varepsilon_j \\
&= X_1 + X_2 + X_3 + X_4 + \varepsilon_j,
\end{aligned}
$$

where $\max_j \mathbb{E}|\varepsilon_j| \leq \frac{K}{\sqrt{n}}$.
We begin by dealing with the first term X_1. We have,

$$
\begin{aligned}
X_1 &= -\frac{1}{n} \mathrm{Tr}\ DT \sum_{\ell=1}^{n} \rho \tilde{t}_{\ell\ell} \mathbb{E}_j (y_\ell y_\ell^* Q_\ell) DQ \\
&= -\frac{1}{n} \mathrm{Tr}\ DT \sum_{\ell=1}^{j} \rho \tilde{t}_{\ell\ell} y_\ell y_\ell^* \mathbb{E}_j (Q_\ell) DQ - \frac{1}{n} \sum_{\ell=j+1}^{n} \rho \tilde{t}_{\ell\ell} \tilde{d}_\ell \frac{1}{n} \mathrm{Tr}\ DTD(\mathbb{E}_j Q_\ell) DQ \\
&= U_1 - \frac{1}{n} \mathrm{Tr}\ DTD(\mathbb{E}_j Q) DQ \frac{1}{n} \sum_{\ell=j+1}^{n} \rho \tilde{t}_{\ell\ell} \tilde{d}_\ell + \varepsilon_{4,2},
\end{aligned}
$$

where, $\mathbb{E}|\varepsilon_{4,2}| \leq \frac{K}{\sqrt{n}}$ from the rank one perturbation result.
Furthermore,

$$
\begin{aligned}
U_1 &= -\frac{1}{n} \mathrm{Tr}\ DT \sum_{\ell=1}^{j} \rho \tilde{t}_{\ell\ell} y_\ell y_\ell^* (\mathbb{E}_j Q_\ell) DQ_\ell + \frac{1}{n} \mathrm{Tr}\ DT \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 y_\ell y_\ell^* (\mathbb{E}_j Q_\ell) DQ_\ell \eta_\ell \eta_\ell^* Q_\ell + \varepsilon_{4,3} \\
&= -\frac{1}{n} \mathrm{Tr}\ DTD(\mathbb{E}_j Q) DQ \frac{1}{n} \sum_{\ell=1}^{j} \rho \tilde{t}_{\ell\ell} \tilde{d}_\ell + \frac{1}{n} \mathrm{Tr}\ DT \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 y_\ell y_\ell^* (\mathbb{E}_j Q_\ell) DQ_\ell \eta_\ell \eta_\ell^* Q_\ell + \varepsilon_{4,4} \\
&= -\frac{1}{n} \mathrm{Tr}\ DTD(\mathbb{E}_j Q) DQ \frac{1}{n} \sum_{\ell=1}^{j} \rho \tilde{t}_{\ell\ell} \tilde{d}_\ell + \phi_j \frac{1}{n} \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell^2 \frac{1}{n} \mathrm{Tr}\ DQ_\ell DT + \varepsilon_{4,5} \\
&= -\frac{1}{n} \mathrm{Tr}\ DTD(\mathbb{E}_j Q) DQ \frac{1}{n} \sum_{\ell=1}^{j} \rho \tilde{t}_{\ell\ell} \tilde{d}_\ell + \phi_j \frac{\gamma}{n} \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell^2 + \varepsilon_{4,6},
\end{aligned}
$$

where,

1. $\mathbb{E}|\varepsilon_{4,3}|^2 \leq \frac{K}{\sqrt{n}}$ thanks to identity (4.23),

2. Lemma (3.2-1) and the rank one perturbation result justify that: $\mathbb{E}|\varepsilon_{4,4}|^3 \leq \frac{K}{n}$,

3. The following majorations:

$$\mathbb{E}|\frac{1}{n}\text{Tr } DT\sum_{l=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2 y_\ell y_\ell^*(\mathbb{E}_j Q_\ell)DQ_\ell\eta_\ell a_\ell^* Q_\ell| \leq \frac{K}{n^{1/2}}$$
$$\mathbb{E}|\frac{1}{n}\text{Tr } DT\sum_{l=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2 y_\ell y_\ell^*(\mathbb{E}_j Q_\ell)DQ_\ell a_\ell\eta_\ell^* Q_\ell| \leq \frac{K}{n^{1/2}},$$

   yield from lemma (3.2-1), and imply $\mathbb{E}|\varepsilon_{4,5}| \leq \frac{K}{n^{1/2}}$.

4. $\mathbb{E}|\varepsilon_{4,6}| \leq \frac{K}{n}$ which yields from the rank one perturbation result.

Consequently,

$$X_1 = -\frac{\rho\tilde{\delta}}{n}\text{Tr } DTD(\mathbb{E}_j Q)DQ + \phi_j\,\frac{\gamma}{n}\sum_{l=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2\tilde{d}_{\ell\ell}^2 + \varepsilon_j.$$

Treatment of X_2. We have,

$$\begin{aligned}
X_2 &= -\frac{1}{n}\text{Tr } DT\sum_{l=1}^{n}\rho\tilde{t}_{\ell\ell}\mathbb{E}_j(y_\ell a_\ell^* Q_\ell)DQ \\
&= -\frac{1}{n}\text{Tr } DT\sum_{l=1}^{j}\rho\tilde{t}_{\ell\ell}y_\ell a_\ell^*\mathbb{E}_j(Q_\ell)DQ_\ell + \frac{1}{n}\text{Tr } DT\sum_{l=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2 y_\ell a_\ell^*\mathbb{E}_j(Q_\ell)DQ_\ell\eta_\ell\eta_\ell^* Q_\ell + \varepsilon_{4,21} \\
&= \frac{\gamma}{n}\frac{1}{n}\sum_{\ell=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2\tilde{d}_\ell\theta_{\ell j} + \varepsilon_{4,22},
\end{aligned}$$

where, $\mathbb{E}|\varepsilon_{4,22}| \leq \frac{K}{n^{1/2}}$ from,

- $\mathbb{E}|\frac{1}{n}\text{Tr } DT\sum_{\ell=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2 y_\ell a_\ell^*\mathbb{E}_j(Q_\ell)DQ_\ell y_\ell\eta_\ell^* Q_\ell| \leq \frac{K}{n^{1/2}}$,
- $\mathbb{E}|\frac{1}{n}\text{Tr } DT\sum_{\ell=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2 y_\ell a_\ell^*\mathbb{E}_j(Q_\ell)DQ_\ell a_\ell a_\ell^* Q_\ell| \leq \frac{K}{n^{1/2}}$,

Treatment of X_3. We can apply the same arguments as previously to justify the following computations:

$$X_3 = -\frac{1}{n}\text{Tr } DT\sum_{\ell=1}^{j}\rho\tilde{t}_{\ell\ell}a_\ell y_\ell^*\mathbb{E}_j(Q_\ell)DQ_\ell + \frac{1}{n}\text{Tr } DT\sum_{\ell=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2 a_\ell y_\ell^*\mathbb{E}_j Q_\ell DQ_\ell\eta_\ell\eta_\ell^* Q_\ell + \varepsilon_{4,31},$$

where $\mathbb{E}|\varepsilon_{4,31}| \leq \frac{K}{n^{1/2}}$. Then,

$$\begin{aligned}
X_3 &= \phi_j\frac{1}{n}\sum_{\ell=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2\tilde{d}_\ell a_\ell^* Q_\ell DTa_\ell + \varepsilon_{4,32} \\
&= \phi_j\,\frac{1}{n}\sum_{\ell=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2\tilde{d}_\ell a_\ell^*\mathcal{T}_\ell DTa_\ell + \varepsilon_{4,33},
\end{aligned}$$

where $\mathbb{E}|\varepsilon_{4,32}| \leq \frac{K}{\sqrt{n}}$ and $\mathbb{E}|\varepsilon_{4,33}| \leq \frac{K}{\sqrt{n}}$.
On the other hand, since $\tilde{t}_{\ell\ell} a_\ell^* T_\ell D T a_\ell = \tilde{\psi}_\ell a_\ell^* T D T a_\ell$, then, we have:

$$X_3 = \phi_j \frac{1}{n} \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell} \tilde{d}_\ell \tilde{\psi}_\ell a_\ell^* T D T a_\ell + \varepsilon_{4,33}.$$

Consequently, χ_4 verifies:

$$\begin{aligned}
\chi_4 &= -\frac{\rho\tilde{\delta}}{n} \mathrm{Tr}\ DTD(\mathbb{E}_j Q) DQ + \phi_j \frac{1}{n} \left(\gamma \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell^2 + \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell} \tilde{d}_\ell \tilde{\psi}_\ell a_\ell^* T D T a_\ell \right) \\
&\quad + \frac{\gamma}{n} \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell \theta_{\ell j} - \frac{1}{n} \sum_{l=1}^{n} \rho \tilde{t}_{\ell\ell} a_\ell^* (\mathbb{E}_j Q_\ell) DQDT a_\ell + \varepsilon_4,
\end{aligned}$$

where $\mathbb{E}|\varepsilon_4| \leq \frac{K}{\sqrt{n}}$.

Result for ϕ_j. We have the following identity:

$$\begin{aligned}
\mathbb{E}\phi_j &= \gamma + \frac{\gamma}{n} \sum_{l=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell \mathbb{E}\theta_{\ell j} \\
&\quad + \mathbb{E}\phi_j \left(\frac{1}{n} \sum_{l=1}^{j} \rho^2 \tilde{\psi}_\ell^2 \tilde{d}_\ell a_\ell^* T D T a_\ell + \frac{\gamma}{n} \sum_{l=1}^{j} \rho^2 \tilde{d}_\ell^2 \tilde{t}_{\ell\ell}^2 \right) + \varepsilon_j,
\end{aligned} \tag{4.38}$$

---

where, $|\varepsilon_j| \leq \frac{K}{\sqrt{n}}$.

Proof of 4.36.

Let ζ_{kj} be defined for all $1 \leq k, j \leq n$ by: $\zeta_{kj} = a_k^* (\mathbb{E}_j Q) DQ a_k$.
The identity of the resolvent gives,

$$Q = T + \frac{\rho}{n} (\mathrm{Tr}\ \tilde{D}\tilde{T}) TDQ + \rho T A \tilde{\Psi} A^* Q - T \Sigma \Sigma^* Q,$$

we then have,

$$\begin{aligned}
\zeta_{kj} &= a_k^* T D Q a_k + \frac{\rho\tilde{\delta}}{n} a_k^* T D (\mathbb{E}_j Q) DQ a_k + \rho a_k^* T A \tilde{\Psi} A^* (\mathbb{E}_j Q) DQ a_k - a_k^* T \mathbb{E}_j (\Sigma\Sigma^* Q) DQ a_k \\
&= \chi_1 + \chi_2 + \chi_3 + \chi_4.
\end{aligned}$$

Treatment of χ_1. Thanks to the asymptotic behavior of the individual elements of the resolvent Q (lemma 3.5), χ_1 can be written as,

$$\chi_1 = a_k^* T D Q a_k = a_k^* T D T a_k + \varepsilon_{jk,1},$$

where $\mathbb{E}|\varepsilon_{jk,1}|^2 = \mathbb{E}|a_k^* TD(Q-T)a_k|^2 \leq \frac{K}{n}$.

Treatment of χ_3. We have,

$$\begin{aligned}
\chi_3 &= \rho a_k^* TA\tilde{\Psi}A^*(\mathbb{E}_j Q)DQa_k \\
&= \sum_{\ell=1}^n \rho\tilde{\psi}_\ell a_k^* Ta_\ell a_\ell^*(\mathbb{E}_j Q)DQa_k \\
&= \sum_{\ell=1}^n \rho\tilde{\psi}_\ell a_k^* Ta_\ell a_\ell^*(\mathbb{E}_j Q_\ell)DQa_k - \sum_{\ell=1}^n \rho^2\tilde{\psi}_\ell a_k^* Ta_\ell a_\ell^*(\mathbb{E}_j \tilde{q}_{\ell\ell} Q_\ell \eta_\ell \eta_\ell^* Q_\ell)DQa_k \\
&= X_1 + X_2. \qquad (4.39)
\end{aligned}$$

Let us begin with X_2. We start by substituting $\tilde{q}_{\ell\ell}$ by $\tilde{t}_{\ell\ell}$. We have,

$$X_2 = -\sum_{\ell=1}^n \rho^2 \tilde{t}_{\ell\ell}\tilde{\psi}_\ell a_k^* Ta_\ell a_\ell^*(\mathbb{E}_j Q_\ell \eta_\ell \eta_\ell^* Q_\ell)DQa_k + \varepsilon_{jk,21}.$$

Writing $Q_\ell(z) = (\Sigma\Sigma^* - zI_N - \eta_\ell\eta_\ell^*)^{-1}$ and using the inversion formula for small-rank perturbation of a matrix [39, Section 0.7.4], we end up with:

$$Q_\ell = Q + \frac{Q\eta_\ell\eta_\ell^* Q}{1-\eta_\ell^* Q\eta_\ell} = Q + (1+\eta_\ell^* Q_\ell \eta_\ell)Q\eta_\ell\eta_\ell^* Q\ . \qquad (4.40)$$

Hence,

$$\begin{aligned}
\varepsilon_{jk,21} &= -\sum_{\ell=1}^n \rho^2\tilde{\psi}_\ell a_k^* Ta_\ell a_\ell^* \mathbb{E}_j\left((\tilde{q}_{\ell\ell}-\tilde{t}_{\ell\ell})Q_\ell\eta_\ell\eta_\ell^* Q\right)DQa_k \\
&\quad -\sum_{\ell=1}^n \rho^2\tilde{\psi}_\ell a_k^* Ta_\ell a_\ell^*(\mathbb{E}_j(\tilde{q}_{\ell\ell}-\tilde{t}_{\ell\ell})\,(1+\eta_\ell^* Q_l\eta_\ell)Q_\ell\eta_\ell\eta_\ell^* Q\eta_\ell\eta_\ell^* Q)\,DQa_k \\
&= \varepsilon_{jk,21}^1 + \varepsilon_{jk,21}^2,
\end{aligned}$$

and we have:

$$\begin{aligned}
\mathbb{E}|\varepsilon_{jk,21}^1| &= \mathbb{E}|\sum_{\ell=1}^n \rho^2\tilde{\psi}_\ell a_k^* Ta_\ell a_\ell^*(\mathbb{E}_j\tilde{q}_{\ell\ell}Q_\ell\eta_\ell\eta_\ell^* Q)DQa_k| \\
&= \mathbb{E}|\rho^2\mathbb{E}_j\left(a_k^* TA\mathrm{diag}\left((\tilde{q}_{\ell\ell}-\tilde{t}_{\ell\ell})\tilde{\psi}_\ell a_\ell^* Q_\ell\eta_\ell\right)\Sigma^* Q\right)DQa_k| \\
&\leq \mathbb{E}|\max_\ell\left((\tilde{q}_{\ell\ell}-\tilde{t}_{\ell\ell})\tilde{\psi}_\ell a_\ell^* Q_\ell\eta_\ell\right)\|a_k^* TA\Sigma^* Q\|\|DQa_k\|| \\
&\leq K\mathbb{E}\max_\ell\left((\tilde{q}_{\ell\ell}-\tilde{t}_{\ell\ell})\tilde{\psi}_\ell a_\ell^* Q_\ell\eta_\ell\right)\|\Sigma^* Q\|
\end{aligned}$$

Consider a singular value decomposition of the matrix Σ with singular values σ_i. We then have: $\|\Sigma^* Q\| = \max_{i=1:N}\|\frac{\sigma_i}{\sigma_i^2+\rho}\| \leq K$. On the other hand, we have,

$$\begin{aligned}
\mathbb{E}\max_\ell|\tilde{q}_{\ell\ell}-\tilde{t}_{\ell\ell}| &= n\mathbb{E}\max_\ell\frac{1}{n}|\tilde{q}_{\ell\ell}-\tilde{t}_{\ell\ell}| = n\int_{\mathbb{R}^+}\mathbb{P}\left(\max_\ell\frac{1}{n}|\tilde{q}_{\ell\ell}-\tilde{t}_{\ell\ell}| \geq x\right)dx \\
&\leq n^2\int_{\mathbb{R}^+}\max_\ell\mathbb{P}\left(\frac{1}{n}|\tilde{q}_{\ell\ell}-\tilde{t}_{\ell\ell}|\geq x\right)dx = n^2\int_{\mathbb{R}^+}\max_\ell\mathbb{P}\left(|\tilde{q}_{\ell\ell}-\tilde{t}_{\ell\ell}|\geq nx\right)dx \\
&\leq n^2\int_{\mathbb{R}^+}\frac{\mathbb{E}|\tilde{q}_{\ell\ell}-\tilde{t}_{\ell\ell}|^2}{(nx)^2}dx \leq \frac{K}{n^{1/2}}
\end{aligned}$$

where the last inequality follows from lemma (3.5). We therefore have: $\mathbb{E}|\varepsilon^1_{jk,21}| \leq \frac{K}{n^{1/2}}$.
Similarely, we prove that $\mathbb{E}|\varepsilon^2_{jk,21}| \leq \frac{K}{n^{1/2}}$.
By the same kind of argument, it follows that:

$$\begin{aligned} X_2 &= -\sum_{\ell=1}^{n} \rho^2 \tilde{t}_{\ell\ell}\tilde{\psi}_\ell a_k^* T a_\ell a_\ell^* \mathcal{T}_\ell a_\ell a_\ell^* \mathbb{E}_j(Q_\ell) D Q a_k - \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}\tilde{\psi}_\ell a_k^* T a_\ell a_\ell^* \mathcal{T}_\ell a_\ell y_\ell^* \mathbb{E}_j(Q_\ell) D Q a_k + \varepsilon_{jk} \\ &= X_{21} + X_{22} + \varepsilon_{jk} \end{aligned}$$

where: $\mathbb{E}|\varepsilon_{jk}| \leq \frac{K}{n^{1/2}}$.

Woodbury's lemma (identity (4.15)) applied to the resolvent matrix on X_{22} with standard computations yield,

$$\begin{aligned} X_{22} &= -\sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}\tilde{\psi}_\ell a_k^* T a_\ell a_\ell^* \mathcal{T}_\ell a_\ell y_\ell^* \mathbb{E}_j(Q_\ell) D Q_\ell a_k + \sum_{\ell=1}^{j} \rho^3 \tilde{t}_{\ell\ell}^2 \tilde{\psi}_\ell a_k^* T a_\ell a_\ell^* \mathcal{T}_\ell a_\ell y_\ell^* \mathbb{E}_j(Q_\ell) D Q_\ell \eta_\ell \eta_\ell^* Q_\ell a_k + \varepsilon_{jk} \\ &= -\sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}\tilde{\psi}_\ell a_k^* T a_\ell a_\ell^* \mathcal{T}_\ell a_\ell y_\ell^* \mathbb{E}_j(Q_\ell) D Q_\ell a_k + \sum_{\ell=1}^{j} \rho^3 \tilde{t}_{\ell\ell}^2 \tilde{\psi}_\ell a_k^* T a_\ell a_\ell^* \mathcal{T}_\ell a_\ell y_\ell^* \mathbb{E}_j(Q_\ell) D Q_\ell y_\ell a_\ell^* Q_\ell a_k + \varepsilon_{jk} \end{aligned}$$

where ε_{jk} converges to zero in probability, and we have,

$$\mathbb{E} X_{22} = \mathbb{E}\phi_j \sum_{\ell=1}^{j} \rho^3 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell \tilde{\psi}_\ell a_k^* T a_\ell a_\ell^* \mathcal{T}_\ell a_\ell a_\ell^* \mathbb{E} Q_\ell a_k + \mathbb{E}\varepsilon_{jk} + \mathbb{E}\tilde{\varepsilon}_{jk},$$

where,

$$\begin{aligned} \tilde{\varepsilon}_{jk} &= \sum_{\ell=1}^{j} \rho^3 \tilde{t}_{\ell\ell}^2 \tilde{\psi}_\ell a_k^* T a_\ell a_\ell^* \mathcal{T}_\ell a_\ell \left(y_\ell^* \mathbb{E}_j(Q_\ell) D Q_\ell y_\ell - \tilde{d}_\ell \phi_j \right) a_\ell^* Q_\ell a_k \\ &\quad + \sum_{\ell=1}^{j} \rho^3 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell \tilde{\psi}_\ell a_k^* T a_\ell a_\ell^* \mathcal{T}_\ell a_\ell \left(\phi_j - \mathbb{E}\phi_j \right) a_\ell^* Q_\ell a_k + \mathbb{E}\phi_j \sum_{\ell=1}^{j} \rho^3 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell \tilde{\psi}_\ell a_k^* T a_\ell a_\ell^* \mathcal{T}_\ell a_\ell a_\ell^* \left(Q_\ell - \mathbb{E} Q_\ell \right) a_k \\ &= \tilde{\varepsilon}^1_{jk} + \tilde{\varepsilon}^2_{jk} + \tilde{\varepsilon}^3_{jk} \end{aligned}$$

where,

$$\begin{aligned} \mathbb{E}|\tilde{\varepsilon}^1_{jk}| &\leq K \sum_{\ell=1}^{j} |a_k^* T a_\ell| \mathbb{E}^{1/2} |a_\ell^* Q_\ell a_k|^2 \mathbb{E}^{1/2} |y_\ell^* \mathbb{E}_j(Q_\ell) D Q_\ell y_\ell - \frac{\tilde{d}_\ell}{n} \mathrm{Tr}\, D Q_\ell D Q_\ell|^2 \\ &\quad + K \sum_{\ell=1}^{j} |a_k^* T a_\ell| \mathbb{E}^{1/2} |a_\ell^* Q_\ell a_k|^2 \mathbb{E}^{1/2} |\frac{\tilde{d}_\ell}{n} \mathrm{Tr}\, D Q_\ell D Q_\ell - \tilde{d}_\ell \phi_j|^2 \\ &\leq \frac{K}{n^{1/2}} \left(\sum_{\ell=1}^{j} |a_k^* T a_\ell a_\ell^*|^2 \right)^{1/2} \left(\mathbb{E}|a_\ell^* Q_\ell a_k|^2 \right)^{1/2} \leq \frac{K}{n^{1/2}} \end{aligned}$$

For ε_{jk}^2, we have,

$$\begin{aligned}
\mathbb{E}|\tilde{\varepsilon}_{jk}^2| &\leq K\sum_{\ell=1}^{j}|a_k^*Ta_\ell|\mathbb{E}^{1/2}|\phi_j-\mathbb{E}\phi_j|^2\mathbb{E}^{1/2}|a_\ell^*Q_\ell a_k|^2 \\
&\overset{(a)}{\leq} \frac{K}{\sqrt{n}}\left(\sum_{\ell=1}^{j}|a_k^*Ta_\ell|^2\right)^{1/2}\left(\sum_{\ell=1}^{j}|a_\ell^*Q_\ell a_k|^2\right)^{1/2} \leq \frac{K}{\sqrt{n}},
\end{aligned}$$

where (a) follows from lemma (5.1). The last term ε_{jk}^3 has a null expectation.

Finally, after remarking that $\rho\tilde{t}_{ll}\tilde{\psi}_l a_l^*\mathcal{T}_l a_l = \tilde{\psi}_l - \tilde{t}_{ll}$, plugging the expression of $\mathbb{E}X_2$ into (4.39) gives,

$$\mathbb{E}\chi_3 = \sum_{\ell=1}^{n}\rho\tilde{t}_{\ell\ell}a_k^*Ta_\ell a_\ell^*\mathbb{E}((\mathbb{E}_jQ_\ell)DQ)a_k + \mathbb{E}(\psi_j)\sum_{\ell=1}^{j}\rho^3\tilde{\psi}_\ell\tilde{t}_{\ell\ell}^2\tilde{d}_\ell a_\ell^*\mathcal{T}_\ell a_\ell a_\ell^*(\mathbb{E}Q_\ell)a_k + \varepsilon_{jk},$$

where, ε_{jk} converges to zero in probability.

Treatment of χ_4. We have,

$$\begin{aligned}
\chi_4 &= -a_k^*T\mathbb{E}_j(\Sigma\Sigma^*Q)DQa_k \overset{(a)}{=} -\sum_{\ell=1}^{n}\rho a_k^*T\mathbb{E}_j(\tilde{q}_{\ell\ell}\eta_\ell\eta_\ell^*Q_\ell)DQa_k \\
&\overset{(b)}{=} -\sum_{\ell=1}^{n}\rho\tilde{t}_{\ell\ell}a_k^*Ta_\ell a_\ell^*\mathbb{E}_j(Q_\ell)DQa_k - \sum_{\ell=1}^{n}\rho\tilde{t}_{\ell\ell}a_k^*T\mathbb{E}_j(y_\ell y_\ell^*Q_\ell)DQa_k \\
&\quad -\sum_{\ell=1}^{n}\rho\tilde{t}_{\ell\ell}a_k^*Ta_\ell\mathbb{E}_j(y_\ell^*Q_\ell)DQa_k - \sum_{\ell=1}^{n}\rho\tilde{t}_{\ell\ell}a_k^*T\mathbb{E}_j(y_\ell a_\ell^*Q_\ell)DQa_k + \varepsilon_{jk} \\
&= X_1 + X_2 + X_3 + X_4 + \varepsilon_{jk}.
\end{aligned}$$

where (a) follows from identity (4.17) and (b) from lemma (5.1), and ε_{jk} converges to zero.

Using Woodbury's lemma and the rank one perturbation result (lemma (3.7)), X_2 satisfies:

$$\begin{aligned}
X_2 &= -\sum_{\ell=1}^{n}\rho\tilde{t}_{\ell\ell}a_k^*T\mathbb{E}_j(y_\ell y_\ell^*Q_\ell)DQ_\ell a_k + \sum_{\ell=1}^{n}\rho^2\tilde{t}_{\ell\ell}\tilde{q}_{\ell\ell}a_k^*T\mathbb{E}_j(y_\ell y_\ell^*Q_\ell)DQ_\ell\eta_{\ell\ell}\eta_{\ell\ell}^*Q_\ell a_k \\
&= -\frac{\rho\tilde{\delta}}{n}a_k^*TD(\mathbb{E}_jQ)DQa_k + \tilde{X}_2,
\end{aligned}$$

with,

$$
\begin{aligned}
\tilde{X}_2 &= \sum_{\ell=1}^{n} \rho^2 \tilde{t}_{\ell\ell}\tilde{q}_{\ell\ell} a_k^* T \mathbb{E}_j(y_\ell y_\ell^* Q_\ell) D Q_\ell \eta_{\ell\ell}\eta_{\ell\ell}^* Q_\ell a_k \\
&= \sum_{\ell=1}^{n} \rho^2 \tilde{t}_{\ell\ell}^2 a_k^* T \mathbb{E}_j(y_\ell y_\ell^* Q_\ell) D Q_\ell y_\ell y_\ell^* Q_\ell a_k + \varepsilon_{jk}^1 \\
&= \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 a_k^* T y_\ell y_\ell^* \mathbb{E}_j(Q_\ell) D Q_\ell y_\ell y_\ell^* Q_\ell a_k + \frac{1}{n}\sum_{\ell=j+1}^{n} \rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell a_k^* T D \mathbb{E}_j(Q_\ell) D Q_\ell y_\ell y_\ell^* Q_\ell a_k + \varepsilon_{jk}^1 \\
&= \phi_j \, \frac{\rho^2}{n} \sum_{\ell=1}^{j} \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell^2 a_k^* T D Q_\ell a_k + \varepsilon_{jk}^2 \\
&= \phi_j \, a_k^* T D T a_k \frac{\rho^2}{n} \sum_{\ell=1}^{j} \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell^2 + \varepsilon_{jk}^3,
\end{aligned}
$$

ε_{jk}^i, for $i = 1, 2, 3$, can be treated using standard previous tools and we can prove that they converge to zero in probability.

Similarely, we have,

$$
\begin{aligned}
X_3 &= \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 a_k^* T a_\ell y_\ell^* \mathbb{E}_j(Q_\ell) D Q_\ell y_\ell a_\ell^* Q_\ell a_k + \varepsilon_{jk}^1 \\
&= \phi_j \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell a_k^* T a_\ell a_\ell^* Q_\ell a_k + \varepsilon_{jk}^2,
\end{aligned}
$$

and,

$$
\begin{aligned}
X_4 &= \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 a_k^* T y_\ell a_\ell^* \mathbb{E}_j(Q_\ell) D Q_\ell a_\ell y_\ell^* Q_\ell a_k + \varepsilon_{jk}^1 \\
&= \frac{1}{n} \sum_{l=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell a_k^* T D Q_\ell a_k \theta_{\ell\ell} + \varepsilon_{jk}^2 \\
&= \frac{a_k^* T D T a_k}{n} \sum_{l=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell \theta_{\ell\ell} + \varepsilon_{jk}^3
\end{aligned}
$$

where ε_{jk}^i converge to zero in probability for $i = 1, 2$ and 3.
We therefore obtain,

$$
\begin{aligned}
\mathbb{E}(\chi_4) &= -\sum_{\ell=1}^{n} \rho \tilde{t}_{\ell\ell} a_\ell a_\ell^* \mathbb{E}\left(\mathbb{E}_j(Q_\ell) D Q\right) a_k - \frac{\rho\tilde{\delta}}{n} a_k^* T D \mathbb{E}\left(\mathbb{E}_j(Q) D Q\right) a_k \\
&\quad + \mathbb{E}(\phi_j) a_k^* T D T a_k \frac{\rho^2}{n} \sum_{\ell=1}^{j} \tilde{t}_{\ell\ell}^2 \tilde{d}_{\ell\ell}^2 + \mathbb{E}(\phi_j) \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell a_k^* T a_\ell a_\ell^* \mathbb{E}(Q_\ell) a_k \\
&\quad + \frac{a_k^* T D T a_k}{n} \sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell \mathbb{E}(\theta_{\ell j}) + \varepsilon_{jk}
\end{aligned}
$$

Consequently, $\mathbb{E}(\zeta_{kj})$ verifies:

$$\begin{aligned}\mathbb{E}(\zeta_{kj}) &= a_k^* TDTa_k + \frac{a_k^* TDTa_k}{n}\sum_{\ell=1}^{j}\rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell \mathbb{E}(\theta_{\ell j}) \\ & +\mathbb{E}(\phi_j)\left(\sum_{\ell=1}^{j}\rho^3\tilde{\psi}_\ell \tilde{t}_{\ell\ell}^2\tilde{d}_\ell a_k^*\mathcal{T}_\ell a_\ell a_\ell^*(\mathbb{E}Q_\ell)a_k + \frac{\rho^2 a_k^* TDTa_k}{n}\sum_{\ell=1}^{j}\tilde{t}_{\ell\ell}^2\tilde{d}_\ell^2 + \right. \\ & \left. \sum_{\ell=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2\tilde{d}_\ell a_k^* T a_\ell a_\ell^*\mathbb{E}Q_\ell a_k\right) + \varepsilon_{kj}.\end{aligned} \tag{4.41}$$

where ε_{jk} converges to zero in probability.

Proof of (4.37)

Using Woodbury's identity, we obtain,

$$\begin{aligned}\zeta_{kj} &= a_k^*(\mathbb{E}_j Q)Qa_k \\ &= a_k^*(\mathbb{E}_j Q_k)Q_k a_k - \rho a_k^*(\mathbb{E}_j \tilde{q}_{kk} Q_k \eta_k^* \eta_k Q_k)Q_k a_k \\ & \quad -\rho\tilde{q}_{kk}a_k^*(\mathbb{E}_j Q_k)Q_k\eta_k^*\eta_k Q_k a_k + \rho^2\tilde{q}_{kk}a_k^*(\mathbb{E}_j\tilde{q}_{kk}Q_k\eta_k^*\eta_k Q_k)Q_k\eta_k^*\eta_k Q_k a_k \\ &= \theta_{kj} - \rho\tilde{t}_{kk}a_k^*\mathcal{T}_k a_k\theta_{kj} - \rho\tilde{t}_{kk}a_k^*\mathcal{T}_k a_k\theta_{kj} + \rho^2\tilde{t}_{kk}^2\tilde{d}_k\left(a_k^*\mathcal{T}_k a_k\right)^2\phi_j + \rho^2\tilde{t}_{kk}^2(a_k^*\mathcal{T}_k a_k)^2\theta_{kj} + \epsilon_{kj},\end{aligned}$$

where ϵ_{kj} converges to zero in probability which can be proved by using standard arguments. We therefore have,

$$\mathbb{E}\zeta_{kj} = \left(1-\rho\tilde{t}_{kk}a_k^*\mathcal{T}_k a_k\right)^2\mathbb{E}\theta_{kj} + \rho^2\tilde{t}_{kk}^2\tilde{d}_k(a_k^*\mathcal{T}_k a_k)^2\mathbb{E}\phi_j + \epsilon_{kj},$$

or again,

$$\mathbb{E}(\zeta_{kj}) = \rho^2\tilde{t}_{kk}^2(1+\tilde{d}_k\delta)^2\mathbb{E}(\theta_{kj}) + \left(\frac{\tilde{d}_k^{1/2}a_k^* Ta_k}{1+\tilde{d}_k\delta}\right)^2\mathbb{E}(\phi_j) + \epsilon_{kj}, \tag{4.42}$$

because $\rho\tilde{t}_{kk}a_k^*\mathcal{T}_k a_k = \rho\tilde{\psi}_k a_k^* Ta_k$.

Proof of (4.34)

System with unknown parametrs ϕ_j and θ_{jk}:

Introduce the following notation: $\varphi_{kj} = \rho^2\tilde{d}_k\tilde{t}_{kk}^2\theta_{kj}$, then from equations (4.35), (4.36) and (4.37), ϕ_j and φ_{kj}satisfy the following system for all $1 \le j,k \le n$:

$$\begin{cases}\mathbb{E}(\phi_j) = \gamma + \alpha_j\mathbb{E}(\phi_j) + \frac{\gamma}{n}\sum_{\ell=1}^{n}\mathbb{E}(\varphi_{\ell j}) + \epsilon_j \\ \mathbb{E}(\varphi_{kj}) = \mu_k + \beta_{kj}\mathbb{E}(\phi_j) + \mu_k\frac{1}{n}\sum_{\ell=1}^{j}\mathbb{E}(\varphi_{\ell j}) + \epsilon_{kj},\end{cases}$$

where,

$$\begin{aligned}
\gamma &= \frac{1}{n}\mathrm{Tr}\ (DT)^2, \\
\alpha_j &= \frac{1}{n}\sum_{\ell=1}^{j}\rho^2\tilde{\psi}_\ell^2\tilde{d}_\ell a_\ell^* TDTa_\ell + \frac{\gamma}{n}\sum_{\ell=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2\tilde{d}_l^2, \\
\mu_k &= \rho^2\tilde{\psi}_k^2\tilde{d}_k a_k^* TDTa_k, \\
\beta_{kj} &= \rho^2\tilde{\psi}_k^2\tilde{d}_k\left(\frac{a_k^*TDTa_k}{n}\sum_{\ell=1}^{j}\rho^2\tilde{t}_{\ell\ell}^2\tilde{d}_\ell^2 + \sum_{\ell=1}^{j}\rho^2\tilde{\psi}_\ell\tilde{t}_{\ell\ell}\tilde{d}_\ell a_k^*Ta_\ell a_\ell^*\mathbb{E}Q_\ell a_k - \rho^2\tilde{\psi}_k^2\tilde{d}_k(a_k^*Ta_k)^2\right).
\end{aligned}$$

Let,

$$\Upsilon_j = \mathbb{E}\left[\phi_j, \varphi_{jj}, \varphi_{j-1,j}, \ldots, \varphi_{1j}\right]^T \quad \forall 1 \le j \le n.$$

Then, we have,

$$\Upsilon_j = \Gamma_j\Upsilon_j + B_j$$

with, $\Gamma_j = \begin{pmatrix} \alpha_j & \frac{\gamma}{n} & \ldots & \frac{\gamma}{n} \\ \beta_{jj} & \frac{\mu_j}{n} & \ldots & \frac{\mu_j}{n} \\ \beta_{j-1j} & \frac{\mu_{j-1}}{n} & \ldots & \frac{\mu_{j-1}}{n} \\ . & . & \ldots & . \\ \beta_{1j} & \frac{\mu_1}{n} & \ldots & \frac{\mu_1}{n} \end{pmatrix}$ and $B_j = \begin{pmatrix} \gamma \\ \mu_j \\ \mu_{j-1} \\ . \\ \mu_1 \end{pmatrix}$.

We can easily solve this system because matrix Γ_j is rank two: $\Gamma_j = UV^T$, where:
$U = \begin{pmatrix} u_1 & u_2 \end{pmatrix} = \begin{pmatrix} \alpha_j & \frac{\gamma}{n} \\ \beta_{jj} & \frac{\mu_j}{n} \\ \beta_{j-1j} & \frac{\mu_{j-1}}{n} \\ . & . \\ \beta_{1j} & \frac{\mu_1}{n} \end{pmatrix}$ and $V = \begin{pmatrix} v_1 & v_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ . & \\ 0 & 1 \end{pmatrix}$.

From Woodbury's identity we have: $(I_{j+1} - \Gamma_j)^{-1} = I_{j+1} + U\left(I_2 - V^TU\right)^{-1}V^T$.
Let us develop:

$$I_2 - V^TU = \begin{pmatrix} 1-\alpha_j & -\frac{\gamma}{n} \\ -\sum_{k=1}^{j}\beta_{kj} & 1 - \frac{1}{n}\sum_{k=1}^{j}\mu_k \end{pmatrix}$$

The determinant of this matrix is:

$$\Delta_j = (1-\alpha_j)\left(1 - \frac{1}{n}\sum_{k=1}^{j}\mu_k\right) - \frac{\gamma}{n}\sum_{k=1}^{j}\beta_{kj}.$$

By developing, we obtain,
$$\mathbb{E}\begin{pmatrix} \phi_j \\ \tilde{\varphi}_{jj} \end{pmatrix} = \begin{pmatrix} \gamma \\ \mu_j \end{pmatrix} + \frac{1}{\Delta_j}\begin{pmatrix} \gamma\alpha_j + \frac{\gamma^2}{n}\sum_{k=1}^{n}\beta_{kj} + \frac{\gamma(1-\alpha_j)}{n}\sum_{k=1}^{n}\mu_k \\ \gamma\beta_{jj} + \mu_j\left(\frac{\gamma}{n}\sum_{k=1}^{n}\beta_{kj} + \frac{1-\alpha_j}{n}\sum_{k=1}^{j}\mu_k\right) \end{pmatrix} + \epsilon_j,$$
where ε_j is an 2-dimensional random vector whose entries converge to zero in probability.

We therefore have,

$$\begin{aligned}
\sum_{j=1}^{n} \chi_{2j} &= \frac{1}{n}\sum_{j=1}^{n}\left(\rho^2 \tilde{t}_{jj}^2 \tilde{d}_j^2 \mathbb{E}(\phi_j) + 2\mathbb{E}\tilde{\varphi}_{jj}\right) + \varepsilon_n \\
&= \frac{\gamma}{n}\sum_{j=1}^{n}\rho^2 \tilde{t}_{jj}^2 \tilde{d}_j^2 + \frac{1}{n}\sum_{j=1}^{n}\frac{\rho^2 \tilde{t}_{jj}^2 \tilde{d}_j^2}{\delta_j}\left(\gamma\alpha_j + \frac{\gamma^2}{n}\sum_{k=1}^{j}\beta_{kj} + \frac{\gamma(1-\alpha_j)}{n}\sum_{k=1}^{j}\mu_k\right) \\
&\quad + \frac{2}{n}\sum_{j=1}^{n}\mu_j + \frac{2}{n}\sum_{j=1}^{n}\frac{1}{\Delta_j}\left(\gamma\beta_{jj} + \mu_j\left(\frac{\gamma}{n}\sum_{k=1}^{j}\beta_{kj} + \frac{1-\alpha_j}{n}\sum_{k=1}^{j}\mu_j\right)\right) + \varepsilon_n.
\end{aligned}$$

Recall that:

$$\beta_{kj} = \rho^2 \tilde{\psi}_k^2 \tilde{d}_k \left(a_k^* T D T a_k \frac{1}{n}\sum_{\ell=1}^{j}\rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell^2 + \sum_{\ell=1}^{j}\rho^2 \tilde{\psi}_\ell \tilde{t}_{\ell\ell} \tilde{d}_\ell a_k^* T a_\ell a_\ell^* \mathbb{E} Q_\ell a_k + \rho^2 \tilde{\psi}_k^2 \tilde{d}_k (a_k^* T a_k)^2\right).$$

It is now possible to replace $\tilde{\psi}_\ell \tilde{t}_{\ell\ell} a_\ell^* \mathbb{E} Q_\ell a_k$ by $a_\ell^* T a_k$ in the formula of $\frac{1}{n}\sum_{k=1}^{n}\beta_{jj}$ and in $\frac{1}{n}\sum_{j=1}^{n}\frac{1}{n}\sum_{k=1}^{j}\beta_{kj}$, and β_{kj} becomes,

$$\beta_{kj} = \rho^2 \tilde{\psi}_k^2 \tilde{d}_k \left(a_k^* T D T a_k \frac{1}{n}\sum_{\ell=1}^{j}\rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell^2 + \sum_{\ell=1}^{j}\rho^2 \tilde{\psi}_\ell^2 \tilde{d}_\ell a_k^* T a_\ell a_\ell^* T a_k + \rho^2 \tilde{\psi}_k^2 \tilde{d}_k (a_k^* T a_k)^2\right).$$

Now, we proceed to prove the following approximation:

$$\sum_{j=1}^{n} = -\log(\Delta_n) + \mathcal{O}(\frac{1}{\sqrt{n}}),$$

where, $\Delta_n = \left(1 - \frac{\rho^2}{n}\operatorname{Tr} D^{1/2} T A \tilde{\Psi} \tilde{D} \tilde{\Psi} A^* T D^{1/2}\right)^2 - \rho^2 \gamma\tilde{\gamma}$.

This approximation will be carried out following two steps:

1. In the first step, we prove that:

$$\sum_{j=1}^{n}\mathbb{E}\chi_{2j} = \frac{1}{n}\sum_{j=1}^{n}\frac{\rho^2\gamma \tilde{t}_{jj}^2 \tilde{d}_j^4 + 2\gamma\beta_{jj} + 2\mu_j(1-\alpha_j)}{\Delta_j}$$

2. In the second step, we prove that:

$$\sum_{j=1}^{n}\mathbb{E}\chi_{2j} = -\log\Delta_n + \mathcal{O}(\frac{1}{n}).$$

3. The third step is dedicated to show that:

$$\Delta_n = \left(1 - \frac{\rho^2}{n}\operatorname{Tr} D^{1/2} T A \tilde{\Psi} \tilde{D} \tilde{\Psi} A^* T D^{1/2}\right)^2 - \rho^2\gamma\tilde{\gamma}.$$

Proof of step 1. We start by handling the second term in the right-hand side in eqrefeq-chi3.

$$
\begin{aligned}
&\frac{1}{n}\sum_{j=1}^{n}\frac{\rho^2\tilde{t}_{jj}^2\tilde{d}_j^2}{\Delta_j}\left(\gamma\alpha_j+\frac{\gamma^2}{n}\sum_{k=1}^{j}\beta_{kj}+\frac{\gamma(1-\alpha_j)}{n}\sum_{k=1}^{j}\mu_k\right)\\
=\;&\frac{\gamma}{n}\sum_{j=1}^{n}\rho^2\tilde{t}_{jj}^2\tilde{d}_j^2\frac{\alpha_j+\frac{\gamma}{n}\sum_{k=1}^{j}\beta_{kj}+\frac{1-\alpha_j}{n}\sum_{k=1}^{j}\mu_k}{(1-\alpha_j)\left(1-\frac{1}{n}\sum_{k=1}^{j}\mu_k\right)-\frac{\gamma}{n}\sum_{k=1}^{j}\beta_{kj}}\\
=\;&\frac{\gamma}{n}\sum_{j=1}^{n}\rho^2\tilde{t}_{jj}^2\tilde{d}_j^2\frac{\alpha_j\left(1-\frac{1}{n}\sum_{k=1}^{j}\mu_j\right)+(\frac{1}{n}\sum_{k=1}^{j}\mu_j-1)+1+\frac{\gamma}{n}\sum_{k=1}^{j}\beta_{kj}}{(1-\alpha_j)\left(1-\frac{1}{n}\sum_{k=1}^{j}\mu_k\right)-\frac{\gamma}{n}\sum_{k=1}^{j}\beta_{kj}}\\
=\;&-\frac{\gamma}{n}\sum_{j=1}^{n}\rho^2\tilde{t}_{jj}^2\tilde{d}_j^2+\frac{\gamma}{n}\sum_{j=1}^{n}\rho^2\frac{\tilde{d}_{jj}^2\tilde{d}_j^2}{\Delta_j}
\end{aligned}
$$

The fourth one verifies:

$$
\begin{aligned}
&\frac{2}{n}\sum_{j=1}^{n}\frac{1}{\Delta_j}\left(\gamma\beta_{jj}+\mu_j\left(\frac{\gamma}{n}\sum_{k=1}^{j}\beta_{kj}+\frac{1-\alpha_j}{n}\sum_{k=1}^{j}\mu_j\right)\right)\\
=\;&\frac{2}{n}\sum_{j=1}^{n}\frac{\gamma\beta_{jj}+\mu_j\left(\frac{\gamma}{n}\sum_{k=1}^{j}\beta_{kj}+\frac{1-\alpha_j}{n}\sum_{k=1}^{j}\mu_j\right)}{(1-\alpha_j)\left(1-\frac{1}{n}\sum_{k=1}^{j}\mu_k\right)-\frac{\gamma}{n}\sum_{k=1}^{j}\beta_{kj}}\\
=\;&-\frac{2}{n}\sum_{j=1}^{n}\frac{\mu_j\left((1-\alpha_j)\left(1-\frac{1}{n}\sum_{k=1}^{j}\mu_k\right)-\frac{\gamma}{n}\sum_{k=1}^{j}\beta_{kj}\right)-(1-\alpha_j)\mu_j-\gamma\beta_{jj}}{(1-\alpha_j)\left(1-\frac{1}{n}\sum_{k=1}^{j}\mu_k\right)-\frac{\gamma}{n}\sum_{k=1}^{j}\beta_{kj}}\\
=\;&-\frac{2}{n}\sum_{j=1}^{n}\mu_j+\frac{2}{n}\sum_{j=1}^{n}\frac{\gamma\beta_{jj}+\mu_j(1-\alpha_j)}{\Delta_j}
\end{aligned}
$$

Therefore, $\sum_{j=1}^{n}\mathbb{E}\chi_{2j}$ becomes:

$$
\sum_{j=1}^{n}\mathbb{E}\chi_{2j}=\frac{1}{n}\sum_{j=1}^{n}\frac{\rho^2\tilde{d}_j^2\tilde{t}_{jj}^2\gamma+2\gamma\beta_{jj}+2\mu_j(1-\alpha_j)}{\Delta_j}.
$$

and step 1) is done.

Proof of step 2. The aim of this step is to establish the following,

$$
\sum_{j=1}^{n}\mathbb{E}\chi_{2j}=-\log\Delta_n+\mathcal{O}(\frac{1}{n}).
$$

To this end, we begin by proving that:

$$
\frac{1}{n}\left(\rho^2\tilde{d}_j^2\tilde{t}_{jj}^2\gamma+2\gamma\beta_{jj}+2\mu_j(1-\alpha_j)\right)=-\left(\Delta_j-\Delta_{j-1}\right)+\mathcal{O}(\frac{1}{n^2}).
$$

Consider the following notations:

$$
G_j=\frac{1}{n}\sum_{k=1}^{j}\rho^2\tilde{\psi}_k^2\tilde{d}_k\sum_{\ell=1}^{j}\rho^2\tilde{d}_\ell\tilde{\psi}_\ell^2a_k^*Ta_\ell a_\ell^*Ta_k-\frac{1}{n}\sum_{k=1}^{j}\rho^4\tilde{\psi}_k^4\tilde{d}_k^2(a_k^*Ta_k)^2
$$

$$M_j = \frac{1}{n}\sum_{k=1}^{j} \tilde{t}_{kk}^2 \tilde{d}_k^2 \quad \text{and} \quad F_j = \frac{1}{n}\sum_{k=1}^{j} \rho^2 \tilde{\psi}_k^2 \tilde{d}_k a_k^* T D T a_k = \frac{1}{n}\sum_{k=1}^{j} \mu_k.$$

Then, $\frac{1}{n}\sum_{k=1}^{j} \beta_{kj}$ becomes $\frac{1}{n}\sum_{k=1}^{j} \beta_{kj} = G_j + \rho^2 M_j F_j$.
On the other hand, we have,

$$\alpha_j = \frac{1}{n}\sum_{\ell=1}^{j} \rho^2 \tilde{d}_\ell^2 \tilde{\psi}_\ell^2 a_\ell^* T D T a_\ell + \frac{\gamma}{n}\sum_{\ell=1}^{j} \rho^2 \tilde{d}_l^2 \tilde{t}_{ll}^2 = F_j + \rho^2\gamma M_j.$$

Then,

$$\begin{aligned}
\Delta_j &= (1 - F_j - \rho^2\gamma M_j)(1 - F_j) - \gamma G_j - \rho^2\gamma M_j F_j \\
&= (1 - F_j)^2 - \rho^2\gamma M_j + \rho^2\gamma M_j F_j - \gamma G_j - \rho^2\gamma M_j F_j = (1 - F_j)^2 - \rho^2\gamma M_j - \gamma G_j,
\end{aligned}$$

which implies:

$$\Delta_j - \Delta_{j-1} = \left((1 - F_j)^2 - (1 - F_{j-1})^2\right) - \rho^2\gamma(M_j - M_{j-1}) - \gamma(G_j - G_{j-1})$$

and we have, $M_j - M_{j-1} = \frac{1}{n}\tilde{d}_j^2 \tilde{t}_{jj}^2$ and

$$(1 - F_j)^2 - (1 - F_{j-1})^2 = (1 - F_j)^2 - (1 - F_j + \frac{1}{n}\mu_j)^2 = -\frac{2}{n}\mu_j(1 - F_j) - \frac{\mu_j^2}{n^2},$$

and,

$$\begin{aligned}
G_j - G_{j-1} &= -\frac{\rho^2}{n}\tilde{\psi}_j^2 \tilde{d}_j \mu_j + \frac{\rho^4}{n}\tilde{\psi}_j^2 \tilde{d}_j a_j^* T(\sum_{\ell=1}^{j} \tilde{d}_\ell \tilde{\psi}_\ell^2 a_\ell a_\ell^*) T a_j \\
&\quad + \frac{\rho^2 \tilde{d}_j \tilde{\psi}_j^2}{n} a_j^* T \left(\sum_{\ell=1}^{j} \rho^2 \tilde{\psi}_\ell^2 \tilde{d}_\ell a_\ell a_\ell^* \right) T a_j + \frac{\rho^2}{n}\tilde{\psi}_j^2 (a_j^* T a_j)^2.
\end{aligned}$$

Then, $\Delta_j - \Delta_{j-1}$ can be written as:

$$\begin{aligned}
\Delta_j - \Delta_{j-1} &= -\frac{2}{n}\mu_j(1 - F_j) - \frac{\mu_j^2}{n^2} - \rho^2\frac{\gamma}{n}\tilde{d}_j^2 \tilde{t}_{jj}^2 + \frac{\rho^2}{n}\tilde{d}_j \tilde{\psi}_j^2 \mu_j \\
&\quad - \frac{2\rho^4}{n}\tilde{d}_j \tilde{\psi}_j^2 \gamma a_j^* T(\sum_{\ell=1}^{j} \tilde{d}_\ell \tilde{\psi}_\ell^2 a_\ell a_\ell^*) T a_j + \frac{\rho^2}{n}\tilde{d}_j^2 \tilde{\psi}_j^2 (a_j^* T a_j)^2.
\end{aligned}$$

Let us now evaluate the quantity: $\frac{1}{n}\Big(\rho^2 \tilde{d}_j^2 \tilde{t}_{jj}^2 \gamma + 2\gamma\beta_{jj} + 2\mu_j(1 - \alpha_j)\Big)$.
We have: $\alpha_j = \rho\gamma M_j + F_j$ and

$$\beta_{jj} = \frac{\mu_j}{n}\sum_{\ell=1}^{j} \rho^2 \tilde{t}_{\ell\ell}^2 \tilde{d}_\ell^2 + \rho^4 \tilde{d}_j \tilde{\psi}_j^2 a_j^* T(\sum_{\ell=1}^{j} \tilde{d}_\ell \psi_\ell^2 a_\ell a_\ell^*) T a_j + \rho^2 \tilde{\psi}_j^2 \tilde{d}_j^2 (a_j^* T a_j)^2.$$

We then have,

$$
\begin{aligned}
& \frac{1}{n}\left(\rho^2 \tilde{d}_j^2 \tilde{t}_{jj}^2 \gamma + 2\gamma\beta_{jj} + 2\mu_j(1-\alpha_j)\right) \\
= \; & \frac{\rho^2 \tilde{d}_j^2 \tilde{t}_{jj}^2 \gamma}{n} + \frac{2\gamma\mu_j}{n} M_j + \frac{2\rho \tilde{d}_j \tilde{\psi}_j^2 \gamma}{n} a_j^* T (\sum_{\ell=1}^{j} \tilde{\psi}_\ell \tilde{d}_\ell a_\ell a_\ell^*) T a_j + \frac{2\rho^2 \tilde{d}_j^2 \tilde{\psi}_j^2 \gamma}{n} (a_j^* T a_j)^2 \\
& + \frac{2\mu_j}{n}(1 - F_j) - \frac{2\mu_j \gamma}{n} M_j + \rho^2 \tilde{\psi}_j^2 \tilde{d}_j^2 (a_j^* T a_j)^2 \\
= \; & -(\Delta_j - \Delta_{j-1}) - \frac{\mu_j^2}{n^2} = -(\Delta_j - \Delta_{j-1}) + \mathcal{O}(\frac{1}{n^2}).
\end{aligned}
$$

Finally, we have,

$$
\sum_{j=1}^{n} \mathbb{E}\chi_{2j} = -\frac{1}{n} \sum_{j=1}^{n} \frac{\Delta_j - \Delta_{j-1}}{\Delta_j} + \mathcal{O}(\frac{1}{n}) = -\log(\Delta_n) + \mathcal{O}(\frac{1}{n}).
$$

Proof of Step 3.
The aim of this part is to prove that:

$$
\Delta_n = \left(1 - \frac{\rho^2}{n} \operatorname{Tr} D^{1/2} T A \tilde{\Psi} \tilde{D} \tilde{\Psi} A^* T D^{1/2}\right)^2 - \rho^2 \gamma \tilde{\gamma}.
$$

Recall that,

$$
\begin{aligned}
F_n &= \frac{\rho^2}{n} \mathrm{Tr} D^{1/2} T A \tilde{\Psi} \tilde{D} \tilde{\Psi} A^* T D^{1/2}, \\
M_n &= \frac{1}{n} \sum_{k=1}^{n} \tilde{t}_{kk}^2 \tilde{d}_k^2, \\
G_n &= \frac{\rho^4}{n} \sum_{k=1}^{n} \tilde{d}_k \tilde{\psi}_k^2 a_k^* T \left(\sum_{l=1}^{n} \tilde{d}_l^2 \tilde{\psi}_l^2 a_l a_l^*\right) T a_k - \frac{1}{n} \sum_{k=1}^{n} \rho^4 \tilde{d}_k^2 \tilde{\psi}_k^4 (a_k^* T a_k)^2 \\
\Delta_n &= (1 - F_n)^2 - \rho^2 \gamma M_n - \gamma G_n.
\end{aligned}
$$

It remains then to prove that:

$$
\rho^2 M_n + G_n = \rho^2 \tilde{\gamma}. \tag{4.43}
$$

We have,

$$
\begin{aligned}
\rho^2 M_n + G_n &= \frac{\rho^2}{n} \operatorname{Tr} \tilde{S}^2 + \frac{\rho^4}{n} \sum_{k=1}^{n} \tilde{d}_k \tilde{\psi}_k^2 a_k^* T \left(\sum_{\ell=1}^{n} \tilde{d}_\ell \tilde{\psi}_\ell^2 a_\ell a_\ell^*\right) T a_k - \frac{1}{n} \sum_{k=1}^{n} \rho^4 \tilde{d}_k^2 \tilde{\psi}_k^4 (a_k^* T a_k)^2 \\
&= \frac{\rho^4}{n} \mathrm{Tr} T A \tilde{\Psi} \tilde{D} \tilde{\Psi} A^* T A \tilde{\Psi} \tilde{D} \tilde{\Psi} A^* + \frac{\rho^2}{n} \sum_{k=1}^{n} \tilde{d}_k^2 \tilde{\psi}_k \tilde{t}_{kk} - \frac{\rho^3}{n} \sum_{k=1}^{n} \tilde{d}_k^2 \tilde{\psi}_k^3 a_k^* T a_k \\
&= \frac{\rho^4}{n} \mathrm{Tr} T A \tilde{\Psi} \tilde{D} \tilde{\Psi} A^* T A \tilde{\Psi} \tilde{D} \tilde{\Psi} A^* + \frac{\rho^2}{n} \mathrm{Tr} \tilde{D} \tilde{T} \tilde{D} \tilde{\Psi} - \frac{\rho^3}{n} \mathrm{Tr} \tilde{D} \tilde{\Psi} \tilde{D} \tilde{\Psi} A^* T A \tilde{\Psi}.
\end{aligned} \tag{4.44}
$$

On the other hand, we have:

$$\begin{aligned}
\tilde{\gamma} &= \frac{1}{n}\operatorname{Tr}\tilde{D}\tilde{T}\tilde{D}\tilde{T} \overset{(a)}{=} \frac{1}{n}\operatorname{Tr}\tilde{D}\tilde{T}\tilde{D}\left(\tilde{\Psi}-\rho\tilde{\Psi}A^*TA\tilde{\Psi}\right) \\
&= \frac{1}{n}\operatorname{Tr}\tilde{D}\tilde{T}\tilde{D}\tilde{\Psi} - \frac{\rho}{n}\operatorname{Tr}\tilde{D}\tilde{T}\tilde{D}\tilde{\Psi}A^*TA\tilde{\Psi} \\
&\overset{(b)}{=} \frac{1}{n}\operatorname{Tr}\tilde{D}\tilde{T}\tilde{D}\tilde{\Psi} - \frac{\rho}{n}\operatorname{Tr}\tilde{D}\tilde{\Psi}\tilde{D}\tilde{\Psi}A^*TA\tilde{\Psi} + \frac{\rho^2}{n}\operatorname{Tr}\tilde{D}\tilde{\Psi}A^*TA\tilde{\Psi}\tilde{D}\tilde{\Psi}A^*TA\tilde{\Psi}
\end{aligned} \tag{4.45}$$

where (a) and (b) follow from Woodbury's lemma which ensure that $\tilde{T}=\tilde{\Psi}-\rho\tilde{\Psi}A^*TA\tilde{\Psi}$.

Identities (4.44) and (4.45) imply: $\rho^2 M_n + G_n = \rho^2\tilde{\gamma}$.
Consequently,

$$\Delta_n = \left(1-\frac{\rho^2}{n}\operatorname{Tr}D^{1/2}TA\tilde{\Psi}\tilde{D}\tilde{\Psi}A^*TD^{1/2}\right)^2 - \rho^2\gamma\tilde{\gamma}.$$

which gives the desired result.

CHAPTER 5

Appendices

Proof of Proposition 3.1

1. We have, $\tilde{\delta} = \frac{1}{n}\operatorname{Tr}\tilde{D}\tilde{T} = \frac{1}{n}\sum_{j=1}^{n}\frac{d_j}{\rho\left(1+\tilde{d}_j\delta+a_j^*\mathcal{T}_j a_j\right)} \leq \frac{\tilde{\mathbf{d}}_{\max}}{\rho} = \delta_{\max}$, and the upper bound for $\tilde{\delta}$ is done. On the other hand, since,
$$1+\tilde{d}_j\delta+a_j^*\mathcal{T}_j a_j \leq 1+\frac{\tilde{\mathbf{d}}_{\max}\mathbf{d}_{\max}}{\rho}+\frac{\|a_{\max}\|^2}{\rho},$$
this implies that: $\tilde{\delta} \geq \frac{\tilde{\mathbf{d}}_{\min}}{\rho+\tilde{\mathbf{d}}_{\max}\mathbf{d}_{\max}+\|a_{\max}\|^2} = \tilde{\delta}_{\min}$.
Similarly, we can prove that:
$$\delta_{\min} = \frac{\mathbf{d}_{\min}}{\rho+\mathbf{d}_{\max}\tilde{\mathbf{d}}_{\max}+\|a_{\max}\|^2} \leq \delta \leq \frac{l^+}{\rho}\mathbf{d}_{\max} = \delta_{\max}.$$

2. We turn now to the quantities γ and $\tilde{\gamma}$. We have,
$$\begin{aligned}
\tilde{\gamma} &= \frac{1}{n}\operatorname{Tr}\tilde{D}\tilde{T}\tilde{D}\tilde{T} \\
&\overset{(a)}{=} \frac{1}{n}\operatorname{Tr}\tilde{D}\tilde{T}\tilde{D}\tilde{\Psi} - \frac{\rho}{n}\operatorname{Tr}\tilde{D}\tilde{T}\tilde{D}\tilde{\Psi}A^*TA\tilde{\Psi} \\
&= \frac{1}{n}\sum_{j=1}^{n}\tilde{d}_j^2\tilde{\psi}_j\tilde{t}_{jj} - \frac{\rho}{n}\sum_{j=1}^{n}\tilde{d}_j^2\tilde{\psi}_j^2\tilde{t}_{jj}a_j^*Ta_j \\
&= \frac{1}{n}\sum_{j=1}^{n}\tilde{d}_j^2\tilde{\psi}_j\tilde{t}_{jj}\left\{1-\rho\tilde{\psi}_j a_j^*Ta_j\right\} \\
&\overset{(b)}{=} \frac{1}{n}\sum_{j=1}^{n}\frac{\tilde{d}_j^2(1+\tilde{d}_j\delta)}{\rho^2\left(1+\tilde{d}_j\delta+a_j^*\mathcal{T}_j a_j\right)^2\left(1+\delta\tilde{d}_j\right)} \\
&\geq \frac{\tilde{\mathbf{d}}_{\min}^2(1+\tilde{\delta}_{\min})}{\rho^2\left(1+\tilde{\delta}_{\max}+\|a_{\max}\|^2/\rho\right)^2\left(1+\delta_{\max}\tilde{\mathbf{d}}_{\max}\right)}
\end{aligned}$$

(a) follows from the identity: $\tilde{T} = \tilde{\Psi} - \rho\tilde{\Psi}A^*TA\tilde{\Psi}$, and (b) from the identity $\tilde{\psi}_j a_j^* T a_j = \tilde{t}_{jj} a_j^* \mathcal{T}_j a_j$ and the expression of $\tilde{t}$. On the other hand, it is clear that $\tilde{\gamma} \leq \frac{1}{\rho^2}$.

Analogously, denote by $\mathbf{a}^i$ the row number i of the matrix A, and $\tilde{\mathcal{T}}_i = \left(\tilde{\Psi}^{-1} + \rho A^{i*}\Psi^i A^i\right)^{-1}$, where $\Psi^i = \mathrm{diag}(\psi_l)_{l=1,l\neq i}^N$. Then, we have,

$$
\begin{aligned}
\tilde{\gamma} &= \frac{1}{n}\mathrm{Tr}\, DTDT \\
&= \frac{1}{n}\mathrm{Tr}\, DTD\Psi - \frac{\rho}{n}\mathrm{Tr}\, DTD\Psi A\tilde{T}A^*\Psi \\
&= \frac{1}{n}\sum_{i=1}^N d_i^2\psi_i t_{ii} - \frac{\rho}{n}\sum_{i=1}^N d_i^2\psi_i^2 t_{ii}\mathbf{a}^i\tilde{T}\mathbf{a}^{i*} \\
&= \frac{1}{n}\sum_{i=1}^N d_i^2\psi_i t_{ii}\left\{1 - \rho\psi_i\mathbf{a}^i\tilde{T}\mathbf{a}^{i*}\right\} \\
&\overset{(a)}{=} \frac{1}{n}\sum_{i=1}^N d_i^2\psi_i t_{ii}\left\{1 - \rho t_{ii}\mathbf{a}^i\tilde{\mathcal{T}}_i\mathbf{a}^{i*}\right\} \\
&= \frac{1}{n}\sum_{i=1}^N d_i^2\psi_i t_{ii}\frac{1+\delta_i}{1+\delta_i+\mathbf{a}^i\tilde{\mathcal{T}}_i\mathbf{a}^{*i}} \\
&\geq \frac{l^-\mathbf{d}_{\min}^2(1+\delta_{\min})}{\rho^2\left(1+\delta_{\max}+\|a_{\max}\|^2/\rho\right)^2\left(1+\tilde{\delta}_{\max}\mathbf{d}_{\max}^2\right)}.
\end{aligned}
$$

where, (a) follows from the fact that $t_{ii} = \frac{1}{\rho}\left(1 + d_i\tilde{\delta} + \mathbf{a}^i\tilde{\mathcal{T}}_i\mathbf{a}^{*i}\right)^{-1}$ which follows from Woodbury's lemma. On the other hand, one can easily see that $\gamma \leq \frac{l^+}{\rho^2}$.

Proof of Proposition 3.3

Note that if we replace matrix Q_j by Q in the statement of Proposition 3.3, the proof is straightforward. Indeed:

$$
\sum_{j=1}^n |u^*Qa_j|^2 = u^*QAA^*Q^*u \leq \frac{\|u\|^2\|A\|^2}{d(z,\mathbb{R}^+)^2} < \infty\ .
$$

If $z \in \mathbb{C}-\mathbb{R}^+$, then so does $1+\eta_j^*Q_j(z)\eta_j$. In particular $(1+\eta_j^*Q_j(z)\eta_j)^{-1}$ does not vanish. Using Eq. (4.17), we obtain

$$
1-\eta_j^*Q\eta_j = 1-\eta_j^*Q_j\eta_j + \frac{(\eta_j^*Q_j\eta_j)^2}{1+\eta_j^*Q_j(z)\eta_j} = \frac{1}{1+\eta_j^*Q_j(z)\eta_j}\ .
$$

Writing $Q_j(z) = (\Sigma\Sigma^* - zI_N - \eta_j\eta_j^*)^{-1}$ and using the inversion formula for small-rank perturbation of a matrix [39, Section 0.7.4], we end up with:

$$
Q_j = Q + \frac{Q\eta_j\eta_j^*Q}{1-\eta_j^*Q\eta_j} = Q + (1+\eta_j^*Q_j\eta_j)Q\eta_j\eta_j^*Q\ . \tag{5.1}
$$

Hence,

$$\begin{aligned}\sum_{j=1}^{n} \mathbb{E}|u^* Q_j A_j|^2 &\leq 2\mathbb{E}u^2 QAA^*Q^*u + 2\mathbb{E}u^*Q\Sigma \operatorname{diag}(\alpha_j)\Sigma^* Q^* u \\ &\leq K + 2\|u\|^2 \mathbb{E}\left(\|Q\Sigma\|^2 \max_j \alpha_j\right) ,\end{aligned}$$

where $\alpha_j = |1 + \eta_j^* Q_j \eta_j|^2 |\eta_j^* Q_j a_j|^2$. Considering a sigular value decomposition of Σ, with singular values (σ_i), we easily obtain:

$$\|Q\Sigma\| = \max_{1\leq i\leq N} \left(\frac{\sigma_i}{|\sigma_i^2 - z|}\right) \leq K .$$

As η_j is a column of Σ, we also obtain:

$$|\eta_j^* Q_j a_j| \leq \|a_j\| \, \|\eta_j^* Q\| \leq \|a_j^*\| \, \|\Sigma^* Q\| \leq K .$$

We finally end up with:

$$\begin{aligned}\sum_{j=1}^{n} \mathbb{E}|u^* Q_j a_j|^2 &\leq K\left(1 + \mathbb{E}\left(\max_{1\leq j\leq n} |1 + \eta_j^* Q_j \eta_j|^2\right) + 1\right) \\ &\leq K\left(1 + \mathbb{E}\left(\max_{1\leq j\leq n} \|\eta_j\|^4\right) + 1\right) \quad \leq \quad K\left(1 + \mathbb{E}\left(\max_{1\leq j\leq n} \|y_j\|^4\right) + 1\right) .\end{aligned}$$

Let $\chi_j = \|y_j\|^2 - Nn^{-1}$, then:

$$\mathbb{P}\left(\max_j \chi_j^2 \geq \lambda\right) \leq n\mathbb{P}\left(\chi_1^2 \geq \lambda\right) \leq n \times \frac{\mathbb{E}\chi_1^{2+2\varepsilon}}{\lambda^{1+\varepsilon}} \overset{(a)}{\leq} \frac{K}{n^{\varepsilon}\lambda^{1+\varepsilon}} ,$$

where (a) follows from (4.22). As a consequence, we have $\mathbb{E}\left(\max_j \chi_j^2\right) \leq K$. It remains to notice that $\|y_j\|^4 \leq 2\chi_j^2 + 2N^2 n^{-2}$ to establish the first estimate.

Consider now the sum $\sum \mathbb{E}_j |u^* Q_j a_j|^2$. Using decomposition (5.1) yields:

$$\sum_{j=1}^{n} \mathbb{E}_j |u^* Q_j a_j|^2 \leq 2\sum_{j=1}^{n} \mathbb{E}_j |u^* Q a_j|^2 + 2\sum_{j=1}^{n} \mathbb{E}_j |(1 + \eta_j^* Q_j \eta_j) u^* Q \eta_j \eta_j^* Q a_j|^2 .$$

Consider first

$$\begin{aligned}\mathbb{E}\left(\sum_{j=1}^{n} \mathbb{E}_j |u^* Q a_j|^2\right)^2 &= \sum_{j_1, j_2 = 1}^{n} \mathbb{E}\left(\mathbb{E}_{j_1} |u^* Q a_{j_1}|^2 \mathbb{E}_{j_2} |u^* Q a_{j_2}|^2\right) \\ &\leq 2\sum_{j_1 \leq j_2} \mathbb{E}\left(\mathbb{E}_{j_1} |u^* Q a_{j_1}|^2 \mathbb{E}_{j_2} |u^* Q a_{j_2}|^2\right) \\ &= 2\sum_{j_1=1}^{n} \mathbb{E}\left(\mathbb{E}_{j_1} |u^* Q a_{j_1}|^2 \sum_{j_2=j_1}^{n} \mathbb{E}_{j_1} |u^* Q a_{j_2}|^2\right) .\end{aligned}$$

As $\sum_{j_2=j_1}^{n}\mathbb{E}_{j_1}|u^*Qa_{j_2}|^2$ is bounded by the first part of the proposition, we get:

$$\mathbb{E}\left(\sum_{j=1}^{n}\mathbb{E}_j|u^*Qa_j|^2\right)^2 \leq K\sum_{j_1=1}^{n}\mathbb{E}|u^*Qa_{j_1}|^2,$$

which is again uniformily bounded by the first part of the proposition. In order to prove

$$\mathbb{E}\left(\sum_{j=1}^{n}\mathbb{E}_j|(1+\eta_j^*Q_j\eta_j)u^*Q\eta_j\eta_j^*Qa_j|^2\right)^2 \leq K < \infty\ ,$$

we use the same ideas as previously, that is develop the square and rely on the results established in the first part of the proposition. This concludes the proof of Proposition 3.3.

Proof of lemma (3.4).

The aim of this lemma is to prove that for all $\rho \in \mathbb{R}_+^*$, we have :$\left|\frac{1}{n}\mathrm{Tr}\mathbf{D}(\mathbf{T}(-\rho)-\mathbb{E}\mathbf{Q}(-\rho))\right| \leq \frac{K}{\sqrt{n}}$.

Let us firstly introduce the following deterministic matrices:

$$\mathbf{R} = \left(\mathbf{W}^{-1}+\rho\mathbf{A}\tilde{\mathbf{W}}\mathbf{A}^*\right)^{-1} \quad \text{and} \quad \tilde{\mathbf{R}} = \left(\tilde{\mathbf{W}}^{-1}+\rho\mathbf{A}^*\mathbf{W}\mathbf{A}\right)^{-1},$$

where,

$$\mathbf{W} = diag(w_i)_{i=1}^{N} = diag\left(\frac{1}{\rho\left(1+\frac{d_i^2}{n}\mathrm{Tr}\tilde{\mathbf{D}}\mathbb{E}\tilde{\mathbf{Q}}\right)}\right)_{i=1}^{N}$$

and,

$$\tilde{\mathbf{W}} = diag(\tilde{w}_j)_{j=1}^{n} = diag\left(\frac{1}{\rho\left(1+\frac{\tilde{d}_j}{n}\mathrm{Tr}\mathbf{D}\mathbb{E}\mathbf{Q}\right)}\right)_{j=1}^{n}.$$

We have,

$$\frac{1}{n}\mathrm{Tr}\mathbf{D}\left(\mathbf{T}(-\rho)-\mathbb{E}\mathbf{Q}(-\rho)\right) = \frac{1}{n}\mathrm{Tr}\mathbf{D}\left(\mathbf{T}(-\rho)-\mathbf{R}(-\rho)\right)+\frac{1}{n}\mathrm{Tr}\mathbf{D}\left(\mathbf{R}(-\rho)-\mathbb{E}\mathbf{Q}(-\rho)\right).$$

Then, it remains to show that:

$$\left|\frac{1}{n}\mathrm{Tr}\mathbf{D}\left(\mathbf{T}(-\rho)-\mathbf{R}(-\rho)\right)\right| \leq \frac{K}{\sqrt{n}} \quad \text{and} \quad \left|\frac{1}{n}\mathrm{Tr}\mathbf{D}\left(\mathbf{R}(-\rho)-\mathbb{E}\mathbf{Q}(-\rho)\right)\right| \leq \frac{K}{\sqrt{n}}.$$

Let us begin with $\frac{1}{n}\mathrm{Tr}\mathbf{D}\left(\mathbf{R}(-\rho)-\mathbb{E}\mathbf{Q}(-\rho)\right)$. We have:

$$\frac{1}{n}\mathrm{Tr}\mathbf{D}\left(\mathbf{R}(-\rho)-\mathbb{E}\mathbf{Q}(-\rho)\right) = \frac{1}{n}\sum_{i=1}^{n}e_i^*\mathbf{D}\left(\mathbf{R}(-\rho)-\mathbb{E}\mathbf{Q}(-\rho)\right)e_i.$$

To prove $\left|\frac{1}{n}\mathrm{Tr}\mathbf{D}\left(\mathbf{R}(-\rho)-\mathbb{E}\mathbf{Q}(-\rho)\right)\right| \leq \frac{K}{\sqrt{n}}$, we shall prove that, for all sequences of deterministic vectors $(a_n)_n$ and $(b_n)_n$ such that $\sup_n \max(||a_n||,||b_n||) < \infty$, we have, $|a_n\left(\mathbb{E}\mathbf{Q}(-\rho)-\mathbf{R}(-\rho)\right)b_n| \leq \frac{K}{\sqrt{n}}$.
We have,

$$\begin{aligned}
a^*\left(\mathbb{E}\mathbf{Q}-\mathbf{R}\right)b &= a^*\mathbb{E}\left(\mathbf{R}(\mathbf{R}^{-1}-\mathbf{Q}^{-1})\mathbf{Q}\right)b \\
&= a^*\mathbb{E}\left(\mathbf{R}\left(W^{-1}+\rho A\tilde{W}A^*-\Sigma\Sigma^*-\rho I_N\right)\mathbf{Q}\right)b \\
&= a^*\mathbb{E}\left(\mathbf{R}\left(\rho I_N+(\frac{\rho}{n}\mathrm{Tr}\tilde{\mathbf{D}}\mathbb{E}\tilde{\mathbf{Q}})\mathbf{D}+\rho A\tilde{W}A^*-\Sigma\Sigma^*-\rho I_N\right)\mathbf{Q}\right)b \\
&= a^*\mathbb{E}\left(\mathbf{R}(\frac{\rho}{n}\mathrm{Tr}(\tilde{\mathbf{D}}\mathbb{E}\tilde{\mathbf{Q}})\mathbf{D}+\rho A\tilde{W}A^*)\mathbf{Q}\right)b - a^*\mathbf{R}\mathbb{E}(\Sigma\Sigma^*\mathbf{Q})b \\
&= X_1+X_2.
\end{aligned}$$

Let us begin with term X_2: We have,

$$\begin{aligned}
X_2 &= -a^*\mathbf{R}\sum_{j=1}^n\mathbb{E}(\tilde{q}_{jj}\eta_j\eta_j^*\mathbf{Q}_j)b \\
&= -a^*\mathbf{R}\sum_{j=1}^n\mathbb{E}(\tilde{q}_{jj}y_jy_j^*\mathbf{Q}_j)b - a^*\mathbf{R}\sum_{j=1}^n\mathbb{E}(\tilde{q}_{jj}a_ja_j^*\mathbf{Q}_j)b \\
&\quad -a^*\mathbf{R}\sum_{j=1}^n\mathbb{E}(\tilde{q}_{jj}y_ja_j^*\mathbf{Q}_j)b - a^*\mathbf{R}\sum_{j=1}^n\mathbb{E}(\tilde{q}_{jj}a_jy_j^*\mathbf{Q}_j)b \\
&= X_3+X_4+X_5+X_6.
\end{aligned}$$

We shall prove that $X_3+X_4=-X_1+\epsilon$, where $|\epsilon_1| \leq \frac{K}{\sqrt{n}}$.
Actually, we have:

$$\begin{aligned}
X_3 &= -\rho a^*\mathbf{R}\sum_{j=1}^n\mathbb{E}\left(\tilde{q}_{jj}y_jy_j^*\mathbf{Q}_j\right)b \\
&= -\rho a^*\mathbf{R}\sum_{j=1}^n\mathbb{E}(\tilde{q}_{jj})\mathbb{E}\left(y_jy_j^*\mathbf{Q}_j\right)b - \rho a^*\mathbf{R}\sum_{j=1}^n\mathbb{E}\left((\tilde{q}_{jj}-\mathbb{E}\tilde{q}_{jj})y_jy_j^*\mathbf{Q}_j\right)b \\
&= -\rho a^*\mathbf{R}\mathbf{D}\mathbb{E}\mathbf{Q}b\frac{1}{n}\sum_{j=1}^n\tilde{d}_j\mathbb{E}\tilde{q}_{jj} - \rho a^*\mathbf{R}\sum_{j=1}^n\mathbb{E}\left((\tilde{q}_{jj}-\mathbb{E}\tilde{q}_{jj})y_jy_j^*\mathbf{Q}_j\right)b \\
&= -\rho a^*\mathbf{R}\mathbf{D}\mathbb{E}\mathbf{Q}b\frac{1}{n}\mathrm{Tr}\tilde{\mathbf{D}}\mathbb{E}\tilde{\mathbf{Q}}+\epsilon_1,
\end{aligned}$$

where,

$$\mathbb{E}|\epsilon_1|^2 \leq K\sum_{j=1}^n\mathbb{E}^{1/2}|\tilde{q}_{jj}-\mathbb{E}\tilde{q}_{jj}|^2\mathbb{E}^{1/2}|y_j^*\mathbf{Q}_jba^*\mathbf{R}y_j|^2$$

On the other hand, we have,

$$\begin{aligned}
\mathbb{E}|\tilde{q}_{jj}-\mathbb{E}\tilde{q}_{jj}|^2 &\leq \mathbb{E}|\tilde{q}_{jj}\mathbb{E}\tilde{q}_{jj}\left((\tilde{q}_{jj}^{-1}-(1+\mathbb{E}(\eta_j^*\mathbf{Q}_j\eta_j))^{-1})+((1+\mathbb{E}(\eta_j^*\mathbf{Q}_j\eta_j))^{-1}-\mathbb{E}\tilde{q}_{jj}^{-1})\right)|^2 \\
&\leq K\mathbb{E}|e_j|^2 \leq \frac{K}{\sqrt{n}}
\end{aligned}$$

which implies that: $\mathbb{E}|\epsilon|^2 \leq \frac{K}{\sqrt{n}}$. X_4 satisfies:

$$\begin{aligned}
X_4 &= -\rho a^* \mathbf{R} \sum_{j=1}^{n} \mathbb{E}(\tilde{q}_{jj} a_j a_j^* \mathbf{Q}_j) b \\
&= -\rho a^* \mathbf{R} \sum_{j=1}^{n} \mathbb{E}(\tilde{q}_{jj}) a_j a_j^* \mathbb{E}(\mathbf{Q}_j) b - \rho a^* \mathbf{R} \sum_{j=1}^{n} \mathbb{E}((\tilde{q}_{jj} - \mathbb{E}\tilde{q}_{jj}) a_j a_j^* \mathbf{Q}_j) b \\
&= -\rho a^* \mathbf{R} A \tilde{W} A^* \mathbb{E}\mathbf{Q} b + \epsilon_2,
\end{aligned}$$

with, $|\epsilon_2| \leq \frac{K}{\sqrt{n}}$. We therefore have: $X_3 + X_4 = -X_1 + \epsilon$, with $|\epsilon| \leq \frac{K}{\sqrt{n}}$.
Let us now deal with term X_5. We have,

$$X_5 = -\rho a^* \mathbf{R} \sum_{j=1}^{n} \mathbb{E}(\tilde{q}_{jj} y_j a_j^* \mathbf{Q}_j) b = 0 - \rho a^* \mathbf{R} \sum_{j=1}^{n} \mathbb{E}\left((\tilde{q}_{jj} - \mathbb{E}\tilde{q}_{jj}) y_j a_j^* \mathbf{Q}_j b\right).$$

Then,

$$|X_5| \leq \sum_{j=1}^{n} \mathbb{E}^{1/4} |\rho(\tilde{q}_{jj} - \mathbb{E}\tilde{q}_{jj})|^4 \mathbb{E}^{1/4} |a^* \mathbf{R} y_j|^4 \mathbb{E}^{1/4} |a_j^* \mathbf{Q}_j b|^2 \leq \frac{K}{\sqrt{n}}.$$

Similarly, we prove that $|X_6| \leq \frac{K}{\sqrt{n}}$, and $|a^*(\mathbb{E}\mathbf{Q} - \mathbf{R})b| \leq \frac{K}{\sqrt{n}}$ is done.
This yields, in particular, for $a_n = b_n = e_i$, and for all $\rho \in \mathbb{R}_+^*$,

$$\left| \frac{1}{n} \mathrm{Tr}\left(\mathbb{E}\mathbf{Q}(-\rho) - \mathbf{R}(-\rho)\right) \right| \leq \left| \frac{1}{n} \sum_{i=1}^{N} e_i^* \mathbf{D} \left(\mathbb{E}\mathbf{Q}(-\rho) - \mathbf{R}(-\rho)\right) e_i \right| \leq \frac{N}{n} \frac{K}{\sqrt{n}} \leq \frac{K}{\sqrt{n}}. \tag{5.2}$$

We turn now to the term: $\left|\frac{1}{n}\mathrm{Tr}\mathbf{D}\left(\mathbf{R}(-\rho) - \mathbf{T}(-\rho)\right)\right|$. We have,

$$\begin{aligned}
\frac{1}{n}\mathrm{Tr}\mathbf{D}\left(\mathbf{R} - \mathbf{T}\right) &= \frac{1}{n}\mathrm{Tr}\mathbf{D}\mathbf{T}\left(\mathbf{T}^{-1} - \mathbf{R}^{-1}\right)\mathbf{R} \\
&= \frac{1}{n}\mathrm{Tr}\mathbf{D}\mathbf{T}\Bigg(\rho\left(I_N + (\frac{1}{n}\mathrm{Tr}\tilde{\mathbf{D}}\tilde{\mathbf{T}})\mathbf{D}\right) + \rho A \tilde{\Psi} A^* \\
&\quad -\rho\left(I_N + (\frac{1}{n}\mathrm{Tr}\tilde{\mathbf{D}}\mathbb{E}\tilde{\mathbf{Q}})\mathbf{D}\right) - \rho A \tilde{W} A^*\Bigg)\mathbf{R} \\
&= \left(\frac{1}{n}\mathrm{Tr}\tilde{\mathbf{D}}(\tilde{\mathbf{T}} - \mathbb{E}\tilde{\mathbf{Q}})\right)\frac{\rho}{n}\mathrm{Tr}\mathbf{D}\mathbf{T}\mathbf{D}\mathbf{R} \\
&\quad + \left(\frac{1}{n}\mathrm{Tr}\mathbf{D}(\mathbb{E}\mathbf{Q} - \mathbf{T})\right)\frac{\rho^2}{n}\mathrm{Tr}\mathbf{D}\mathbf{T}A\tilde{W}\tilde{\mathbf{D}}\tilde{\Psi}A^*\mathbf{R}.
\end{aligned}$$

Similarly, we can prove,

$$\frac{1}{n}\mathrm{Tr}\tilde{\mathbf{D}}(\tilde{\mathbf{R}} - \tilde{\mathbf{T}}) = \frac{\rho}{n}\mathrm{Tr}\tilde{\mathbf{D}}\tilde{\mathbf{T}}\tilde{\mathbf{D}}\tilde{\mathbf{R}}\left(\frac{1}{n}\mathrm{Tr}\mathbf{D}(\mathbf{T} - \mathbb{E}\mathbf{Q})\right) + \frac{\rho^2}{n}\mathrm{Tr}\tilde{\mathbf{D}}\tilde{\mathbf{T}}A^*W\mathbf{D}\Psi A\tilde{\mathbf{R}}\left(\frac{1}{n}\mathrm{Tr}\tilde{\mathbf{D}}(\mathbb{E}\tilde{\mathbf{Q}} - \tilde{\mathbf{T}})\right).$$

Further, we have,

$$\begin{cases} \frac{1}{n}\mathrm{Tr}\mathbf{D}(\mathbb{E}\mathbf{Q}-\mathbf{T})=\frac{1}{n}\mathrm{Tr}\mathbf{D}(\mathbb{E}\mathbf{Q}-\mathbf{R})+\frac{1}{n}\mathrm{Tr}\mathbf{D}(\mathbf{R}-\mathbf{T}) \\ \frac{1}{n}\mathrm{Tr}\tilde{\mathbf{D}}(\tilde{\mathbf{T}}-\mathbb{E}\tilde{\mathbf{Q}})=\frac{1}{n}\mathrm{Tr}\tilde{\mathbf{D}}(\mathbb{E}\tilde{\mathbf{Q}}-\tilde{\mathbf{R}})+\frac{1}{n}\mathrm{Tr}\tilde{\mathbf{D}}(\tilde{\mathbf{R}}-\tilde{\mathbf{T}}) \end{cases}$$

We then obtain the following linear system,

$$\begin{aligned} \mathbf{U} &= \begin{pmatrix} \frac{1}{n}\mathrm{Tr}\mathbf{D}(\mathbf{R}-\mathbf{T}) \\ \frac{1}{n}\mathrm{Tr}\tilde{\mathbf{D}}(\tilde{\mathbf{R}}-\tilde{\mathbf{T}}) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\rho^2}{n}\mathrm{Tr}\mathbf{D}\mathbf{T}A\tilde{W}\tilde{\mathbf{D}}\tilde{\Psi}A^*\mathbf{R} & -\frac{\rho}{n}\mathrm{Tr}\mathbf{R}\mathbf{D}\mathbf{T}\mathbf{D} \\ -\frac{\rho}{n}\mathrm{Tr}\tilde{\mathbf{D}}\tilde{\mathbf{T}}\tilde{\mathbf{D}}\tilde{\mathbf{R}} & \frac{\rho^2}{n}\mathrm{Tr}\tilde{\mathbf{D}}\tilde{\mathbf{T}}A^*W\mathbf{D}\Psi A\tilde{\mathbf{R}} \end{pmatrix}\mathbf{U} + \begin{pmatrix} \frac{1}{n}\mathrm{Tr}\mathbf{D}(\mathbb{E}\mathbf{Q}-\mathbf{R}) \\ \frac{1}{n}\mathrm{Tr}\tilde{\mathbf{D}}(\mathbb{E}\tilde{\mathbf{Q}}-\tilde{\mathbf{R}}) \end{pmatrix} \\ &= \mathbf{M}\mathbf{U}+\epsilon. \end{aligned}$$

(5.2) ensures that,

$$|\epsilon| = \left|\left| \begin{pmatrix} \frac{1}{n}\mathrm{Tr}\mathbf{D}(\mathbb{E}\mathbf{Q}-\mathbf{R}) \\ \frac{1}{n}\mathrm{Tr}\tilde{\mathbf{D}}(\mathbb{E}\tilde{\mathbf{Q}}-\tilde{\mathbf{R}}) \end{pmatrix} \right|\right| \leq \frac{K}{\sqrt{n}}$$

Then, to obtain the desired result, it remains to show that $\mathbf{M}$ is invertible. This yield from the fact that

$$\liminf_n \det\left(I_2 - \begin{pmatrix} \frac{\rho^2}{n}\mathrm{Tr}\mathbf{D}\mathbf{T}A\tilde{\Psi}\tilde{\mathbf{D}}\tilde{\Psi}A^*\mathbf{T} & -\frac{\rho}{n}\mathrm{Tr}(\mathbf{D}\mathbf{T})^2 \\ -\frac{\rho}{n}\mathrm{Tr}(\tilde{\mathbf{D}}\tilde{\mathbf{T}})^2 & \frac{\rho^2}{n}\mathrm{Tr}\tilde{\mathbf{D}}\tilde{\mathbf{T}}A^*\Psi\mathbf{D}\Psi A\tilde{\mathbf{T}} \end{pmatrix} \right) > 0$$

i.e. $\liminf_n \left(\left(1-\frac{\rho^2}{n}\mathrm{Tr}\mathbf{D}\mathbf{T}A\tilde{\Psi}\tilde{\mathbf{D}}\tilde{\Psi}A^*\mathbf{T}\right)\left(1-\frac{\rho^2}{n}\mathrm{Tr}\tilde{\mathbf{D}}\tilde{\mathbf{T}}A^*\Psi\mathbf{D}\Psi A\tilde{\mathbf{T}}\right)-\rho^2\gamma\tilde{\gamma}\right) \geq K > 0$, which is proved in theorem (2.1-(1)).

Proof of lemma (3.3)

We write $a_n^*\left(\mathbf{Q}-\mathbf{T}\right)b_n = a_n^*\left(\mathbf{Q}-\mathbb{E}\mathbf{Q}\right)b_n + a_n^*\left(\mathbb{E}\mathbf{Q}-\mathbf{T}\right)b_n$.
Il is proved in the pevious lemma that: $|a_n^*\left(\mathbb{E}\mathbf{Q}-\mathbf{T}\right)b_n| \leq \frac{K}{\sqrt{n}}$. Then it remains to prove that $\mathbb{E}|a^*(\mathbf{Q}-\mathbb{E}\mathbf{Q})b|^2 \leq \frac{K}{n}$.
Denote by $\mathbb{E}_j(.)$ the conditional expectation given by $\mathbb{E}_j(.) = \mathbb{E}(./\sigma(y_1,\ldots,y_j))$. Let $\mathbb{E}_0(.) = \mathbb{E}(.)$. As it may be observed, $a^*(\mathbf{Q}-\mathbb{E}\mathbf{Q})b$ can be written as:

$$a^*(\mathbf{Q}-\mathbb{E}\mathbf{Q})b = \sum_{j=1}^{n} a^*(\mathbb{E}_j-\mathbb{E}_{j-1})(\mathbf{Q}-\mathbf{Q}_j)b = -\rho\sum_{j=1}^{n} a^*(\mathbb{E}_j-\mathbb{E}_{j-1})\left(\tilde{q}_{jj}\mathbf{Q}_j\eta_j\eta_j^*\mathbf{Q}_j\right)b \tag{5.3}$$

Recall that

$$\tilde{q}_{jj} = \frac{1}{\rho\left(1+\frac{\tilde{d}_j}{n}x_j^*\mathbf{D}^{1/2}\mathbf{Q}_j\mathbf{D}^{1/2}x_j + y_j^*\mathbf{Q}_j a_j + a_j^*\mathbf{Q}_j y_j + a_j^*\mathbf{Q}_j a_j\right)}$$

and,

$$\tilde{b}_j = \frac{1}{\rho\left(1+\frac{\tilde{d}_j}{n}\mathrm{Tr}\mathbf{D}\mathbf{Q}_j + a_j^*\mathbf{Q}_j a_j\right)}.$$

Since, $\mathbf{Q}-\mathbf{Q}_j=-\rho\tilde{q}_{jj}\mathbf{Q}_j\eta_j\eta_j^*\mathbf{Q}_j$ and $\tilde{q}_{jj}=\tilde{b}_j-\rho\tilde{q}_{jj}\tilde{b}_je_j$ with $e_j=y_j^*\mathbf{Q}_jy_j-\frac{\tilde{d}_j}{n}\mathrm{Tr}\mathbf{D}\mathbf{Q}_j+y_j^*\mathbf{Q}_ja_j+a_j^*\mathbf{Q}_jy_j$, then, identity (5.3) becomes:

$$\begin{aligned}
a^*(\mathbf{Q}-\mathbb{E}\mathbf{Q})b &= -\rho\sum_{j=1}^{n}a^*(\mathbb{E}_j-\mathbb{E}_{j-1})\left(\tilde{b}_j\mathbf{Q}_j\eta_j\eta_j^*\mathbf{Q}_j\right)b \\
&\quad +\rho^2\sum_{j=1}^{n}a^*(\mathbb{E}_j-\mathbb{E}_{j-1})\left(\tilde{q}_{jj}\tilde{b}_je_j\mathbf{Q}_j\eta_j\eta_j^*\mathbf{Q}_j\right)b=X_1+X_2
\end{aligned}$$

We begin by the study of term X_1. We have,

$$\begin{aligned}
\mathbb{E}|X_1|^2 &\leq \rho^2\sum_{j=1}^{n}\mathbb{E}\left|(\mathbb{E}_j-\mathbb{E}_{j-1})\tilde{b}_ja^*\mathbf{Q}_j\eta_j\eta_j^*\mathbf{Q}_jb\right|^2 \\
&\overset{(a)}{\leq} \rho^2\sum_{j=1}^{n}\mathbb{E}\left|\mathbb{E}_j\tilde{b}_j\left(y_j^*\mathbf{Q}_jba^*\mathbf{Q}_jy_j-\frac{\tilde{d}_j}{n}\mathrm{Tr}(\mathbf{D}\mathbf{Q}_jba^*\mathbf{Q}_j)-a_j^*\mathbf{Q}_jba^*\mathbf{Q}_jy_j-y_j^*\mathbf{Q}_jba^*\mathbf{Q}_ja_j\right)\right|^2 \\
&\leq \frac{2|\rho|^2\tilde{d}_{max}^4}{|\rho|^2}\sum_{j=1}^{n}\mathbb{E}\left|\frac{1}{n}x_j^*\mathbf{D}^{1/2}\mathbf{Q}_jba^*\mathbf{Q}_j\mathbf{D}^{1/2}x_j-\frac{1}{n}\mathrm{Tr}(\mathbf{D}\mathbf{Q}_jba^*\mathbf{Q}_j)\right|^2 \\
&\quad +\frac{4|\rho|^2\tilde{d}_{max}^2}{|\rho|^2}\sum_{j=1}^{n}\mathbb{E}\left|\frac{1}{\sqrt{n}}x_j^*\mathbf{D}^{1/2}\mathbf{Q}_jba^*\mathbf{Q}_ja_j\right|^2\leq\frac{K}{n}
\end{aligned}$$

where (a) follows from facts:

1. The indenpence between y_j and $\mathcal{F}_{j-1}$ and the fact that $\mathbb{E}_j\mathbf{Q}_j=\mathbb{E}_j\mathbf{Q}_{j-1}$ imply:

$$\mathbb{E}_{j-1}\left(\tilde{b}_ja^*\mathbf{Q}_j\eta_j\eta_j^*\mathbf{Q}_jb\right)=\frac{\tilde{d}_j}{n}\mathrm{Tr}\mathbf{D}\left(\mathbb{E}_j\tilde{b}_j\mathbf{Q}_jba^*\mathbf{Q}_j\right)+a_j^*\left(\mathbb{E}_j\tilde{b}_j\mathbf{Q}_jba^*\mathbf{Q}_j\right)a_j,$$

2. The $\mathcal{F}_j$-measurability of y_j implies:

$$\begin{aligned}
&\mathbb{E}_j\left(\tilde{b}_ja^*\mathbf{Q}_j\eta_j\eta_j^*\mathbf{Q}_jb\right)=y_j^*\left(\mathbb{E}\tilde{b}_j\mathbf{Q}_jba^*\mathbf{Q}_j\right)y_j \\
&+a_j^*\left(\mathbb{E}_j\tilde{b}_j\mathbf{Q}_jba^*\mathbf{Q}_j\right)a_j+a_j^*\left(\mathbb{E}\tilde{b}_j\mathbf{Q}_jba^*\mathbf{Q}_j\right)y_j+y_j^*\left(\mathbb{E}\tilde{b}_j\mathbf{Q}_jba^*\mathbf{Q}_j\right)a_j.
\end{aligned}$$

Treatment of $\mathbb{E}|X_2|^2$: We have:

$$\begin{aligned}
X_2 &= \rho^2\sum_{j=1}^{n}(\mathbb{E}_j-\mathbb{E}_{j-1})\,\tilde{b}_j^2e_ja^*\mathbf{Q}_j\eta_j\eta_j^*\mathbf{Q}_jb-\rho^3\sum_{j=1}^{n}(\mathbb{E}_j-\mathbb{E}_{j-1})\,\tilde{b}_j^2\tilde{q}_{jj}e_j^2a^*\mathbf{Q}_j\eta_j\eta_j^*\mathbf{Q}_jb \\
&= X_3+X_4
\end{aligned}$$

Let us begin with X_3. We have,

$$\begin{aligned}
X_3 &= \rho^2 \sum_{j=1}^{n} \left(\mathbb{E}_j - \mathbb{E}_{j-1}\right) \tilde{b}_j^2 e_j a^* \mathbf{Q}_j \eta_j \eta_j^* \mathbf{Q}_j b \\
&= \rho^2 \sum_{j=1}^{n} \left(\mathbb{E}_j - \mathbb{E}_{j-1}\right) \tilde{b}_j^2 e_j a^* \mathbf{Q}_j y y_j^* \mathbf{Q}_j b + \rho^2 \sum_{j=1}^{n} \left(\mathbb{E}_j - \mathbb{E}_{j-1}\right) \tilde{b}_j^2 e_j a^* \mathbf{Q}_j a_j y_j^* \mathbf{Q}_j b \\
&\quad + \rho^2 \sum_{j=1}^{n} \left(\mathbb{E}_j - \mathbb{E}_{j-1}\right) \tilde{b}_j^2 e_j a^* \mathbf{Q}_j y_j a_j^* \mathbf{Q}_j b + \rho^2 \sum_{j=1}^{n} \left(\mathbb{E}_j - \mathbb{E}_{j-1}\right) \tilde{b}_j^2 e_j a^* \mathbf{Q}_j a_j a_j^* \mathbf{Q}_j b \\
&= X_{31} + X_{32} + X_{33} + X_{34}
\end{aligned}$$

One can remark that all these quantities are a sum of martingale difference sequences. Then, we have:

$$\begin{aligned}
\mathbb{E}|X_3|^2 \leq \rho^2 \sum_{j=1}^{n} \Bigg(&\left|\left(\mathbb{E}_j - \mathbb{E}_{j-1}\right) \tilde{b}_j^2 e_j a^* \mathbf{Q}_j y y_j^* \mathbf{Q}_j b\right|^2 + \left|\left(\mathbb{E}_j - \mathbb{E}_{j-1}\right) \tilde{b}_j^2 e_j a^* \mathbf{Q}_j a_j y_j^* \mathbf{Q}_j b\right|^2 \\
&+ \left|\left(\mathbb{E}_j - \mathbb{E}_{j-1}\right) \tilde{b}_j^2 e_j a^* \mathbf{Q}_j y_j a_j^* \mathbf{Q}_j b\right|^2 + \left|\left(\mathbb{E}_j - \mathbb{E}_{j-1}\right) \tilde{b}_j^2 e_j a^* \mathbf{Q}_j a_j a_j^* \mathbf{Q}_j b\right|^2 \Bigg)
\end{aligned}$$

Jensen's inequality and Cauchy-Schwarz's inequality ensure the following,

$$\mathbb{E}|X_{31}|^2 \leq K \sum_{j=1}^{n} \mathbb{E}^{1/2}|e_j|^4 \mathbb{E}^{1/2}|y_j^* \mathbf{Q}_j b a^* \mathbf{Q}_j y_j|^4 \leq K \sum_{j=1}^{n} \frac{K}{n}\frac{K}{n} \leq \frac{K}{n}$$

$$\mathbb{E}|X_{32}|^2 \leq K \sum_{j=1}^{n} \mathbb{E}^{1/2}|a^* \mathbf{Q}_j a_j|^4 \mathbb{E}^{1/4}|y_j^* \mathbf{Q}_j b|^8 \mathbb{E}^{1/4}|e_j|^8 \leq K\sqrt{n}\frac{K}{n}\frac{K}{n} \leq \frac{K}{n^{3/2}}$$

Similarly, we prove that: $\mathbb{E}|X_{33}|^2 \leq \frac{K}{n^{3/2}}$.
Treatment of the term X_{34}.

Let us denote by M_j the matrix given by $M_j = \tilde{b}_j^2 (a_j^* \mathbf{Q}_j b a \mathbf{Q}_j a_j) \mathbf{Q}_j$. We then have:

$$\begin{aligned}
\mathbb{E}|X_{34}|^2 &= \rho^4 \sum_{j=1}^{n} \mathbb{E}\left|\left(\mathbb{E}_j - \mathbb{E}_{j-1}\right) \tilde{b}_j^2 e_j a^* \mathbf{Q}_j a_j a_j^* \mathbf{Q}_j b\right|^2 \leq \rho^4 \sum_{j=1}^{n} \mathbb{E}\left|\tilde{b}_j^2 e_j a^* \mathbf{Q}_j a_j a_j^* \mathbf{Q}_j b\right|^2 \\
&= \rho^4 \sum_{j=1}^{n} \mathbb{E}\left|\tilde{b}_j^2 a^* \mathbf{Q}_j a_j a_j^* \mathbf{Q}_j b \left(y_j^* \mathbf{Q}_j y_j - \frac{\tilde{d}_j}{n} \mathrm{Tr} \mathbf{D} \mathbf{Q}_j + y_j^* \mathbf{Q}_j a_j + a_j^* \mathbf{Q}_j y_j \right)\right|^2 \\
&= \rho^4 \sum_{j=1}^{n} \mathbb{E}\left|\eta_j^* M_j \eta_j - \frac{\tilde{d}_j}{n} \mathrm{Tr} \mathbf{D} M_j + y_j^* M_j a_j + a_j^* \mathbf{Q}_j y_j\right|^2
\end{aligned}$$

Since y_j and $\mathbf{Q}_j$ are independent, lemma (1) implies the following:

$$
\begin{aligned}
& \mathbb{E}_{\mathbf{Q}_j}\left|\eta_j^* M_j \eta_j - \frac{\tilde{d}_j}{n}\mathrm{Tr}\mathbf{D}M_j + y_j^* M_j a_j + a_j^*\mathbf{Q}_j y_j\right|^2 \\
\leq \; & \frac{\rho^4}{n}\sum_{j=1}^{n}\left(\frac{\tilde{d}_j^2}{n}\mathrm{Tr}(diag((\mathbf{D}M_j)))^2 + \frac{\tilde{d}_j^2}{n}\mathrm{Tr}\mathbf{D}M_j\mathbf{D}M_j + 2\tilde{d}_j a_j^* M_j \mathbf{D} M_j a_j\right) \\
\leq \; & \frac{K}{n}\sum_{j=1}^{n}\mathbb{E}|a^*\mathbf{Q}_j a_j|^2 \leq \frac{K}{n}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\mathbb{E}|X_{34}|^2 &= \rho^4\sum_{j=1}^{n}\mathbb{E}\left|\eta_j^* M_j \eta_j - \frac{\tilde{d}_j}{n}\mathrm{Tr}\mathbf{D}M_j + y_j^* M_j a_j + a_j^*\mathbf{Q}_j y_j\right|^2 \\
&= \rho^4\sum_{j=1}^{n}\mathbb{E}\mathbb{E}_{\mathbf{Q}_j}\left|\eta_j^* M_j \eta_j - \frac{\tilde{d}_j}{n}\mathrm{Tr}\mathbf{D}M_j + y_j^* M_j a_j + a_j^*\mathbf{Q}_j y_j\right|^2 \leq \frac{K}{n}.
\end{aligned}
$$

Treatment of the term X_4.

We have,

$$
\begin{aligned}
\mathbb{E}|X_4|^2 &\leq K\sum_{j=1}^{n}\mathbb{E}^{1/2}|e_j|^8\mathbb{E}^{1/4}|a^*\mathbf{Q}_j y_j|^8\mathbb{E}^{1/4}|\eta_j^*\mathbf{Q}_j a|^8 + K\sum_{j=1}^{n}\mathbb{E}^{1/2}|e_j|^8\mathbb{E}^{1/2}|a_j^*\mathbf{Q}_j b a^*\mathbf{Q}_j a_j|^4 \\
&\leq \frac{K}{n^2} + \frac{K}{n^{3/2}} \leq \frac{K}{n^{3/2}}.
\end{aligned}
$$

This ends the proof of $\mathbb{E}|a^*(\mathbf{Q}-\mathbb{E}\mathbf{Q})b|^2 \leq \frac{K}{n}$, then, the proof of lemma (3.3).

BIBLIOGRAPHY

[1] G. W. Anderson and O. Zeitouni, "A CLT for a band matrix model," *Probab. Theory Related Fields* 134 283-338. MR2222385

[2] L. V. Arharov, "Limit theorems for the characteristic roots of a sample covariance matrix," *Dokl. Akad. Nauk SSSR.* 199 994-997. 1971. MR0309171

[3] Z. D. Bai and J. W. Silverstein, "CLT of linear spectral statistics of large dimensional sample covariace matrices," *Annals of Probability* . 32(1A) (2004), pp.553-605.

[4] Z. D. Bai and J. W. Silverstein, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, 54(2), pp.175-192, 1995.

[5] Z.D. Bai,and J.W. Silverstein "No eigenvalues outside the support of the limiting spectral distribution of large dimensional random matrices," *Annals of Probability.*, 26(1) pp. 316-345, 1998.

[6] Z. Bai and J.W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices, Sciences Press, Beijing*, 2006.

[7] R.J. Bhansali, L. Giraitis, and P.S. Kokoszka, "Convergence of quadratic forms with nonvanishing diagonal," *Stat. Probab. Letters*, vol. 77, pp. 726–734, 2007.

[8] E. Biglieri, G. Caire, and G. Taricco, "CDMA system design through asymptotic analysis," *IEEE Trans. on Communications*, vol. 48, no. 11, pp. 1882–1896, Nov. 2000.

[9] P. Billingsley, *Probability and Measure*, *John Wiley, 3rd edition*, 1995.

[10] A. Boutet de Monvel, A. Khorunzhy and V. Vasilchuk, *Limiting eigenvalue distribution of random matrices with correlated entries*, *Markov Process. Related Fields*, 2(4): 607-636, 1996.

[11] A. Boutet de Monvel, A. Khorunzhy, *Limit theorems for random matrices*, *Markov Process. Related Fields*, 4(2): 175-197, 1998.

[12] R. W. Butler and A. T. A. Wood, “Saddlepoint approximation for moment generating functions of truncated random variables,” *The Annales of Statistics,*, vol. 32, no. 6, pp. 2712-2730, 2004.

[13] D. Chafai, “Circular law for non-centred random matrices,” *Journal of Theoretical Probability*, 2010.

[14] S. Chatterjee, “Fluctuations of eigenvakues and second order Poincaré inequalities,” *Probab. Theory Related Fields*, 143 1-40. MR2449121. 2007

[15] J.-M. Chaufray, W. Hachem, and Ph. Loubaton, “Asymptotic Analysis of Optimum and Sub-Optimum CDMA Downlink MMSE Receivers,” *IEEE Trans. on Information Theory*, vol. 50, no. 11, pp. 2620–2638, Nov. 2004.

[16] C. N. Chuah, D.N.C. Tse, J.M. Kahn, and R.A. Valenzuela, “Capacity Scaling in MIMO Wireless Systems Under Correlated Fading,” *IEEE Trans. on Information Theory*, Vol. 48, no. 3, pp. 637-650, Mar. 2002.

[17] M. Debbah and R. Müller, “On the limiting behavior of directional MIMO models,” *Processing of the Sixth Baiona Workshop On Signal Processing in Communication*, Spain, Septembre 2003.

[18] R. B. Dozier and J. W. Silverstein, “Analysis of the Limiting Spectral Distribution of Large Dimensional Information-Plus-Noise Type Matrices,” *Journal of Multivariate Analysis.*, 98(4) (2007), pp. 678-694.

[19] 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) (2007), pp. 678-694.

[20] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton and J. Najim, “On the capacity achieving covariance matrix for Rician MIMO channels: an asymptotic approach,” *IEEE Inf. Theory,* Vol. 56 (3), mars. 2010.

[21] J. Evans and D.N.C Tse, “Large system performance of linear multiuser receivers in multipath fading channels,” *IEEE Trans. on Information Theory*, vol. 46, no. 6, pp. 2059–2078, Sept. 2000.

[22] G.J. Foschini “Layered Space-Time Architecture for Wireless Communication in a Fading Environment when using Multiple Antennas,” *Bell Labs Technical Journal*, Vol. 1, no. 2, pp. 41–59. 1996

[23] H. U. Gerber, “From the generalized gamma to the generalized negative binomial distribution,” *Insurance: Mathematics and Economics*, vol. 10, no. 4, pp. 303-309, 1992.

[24] V. L. Girko, “Theory of Stochastic Canonical Equations.I.” “Mathematics and Its Applications 535. Kluwer Academic, Dordrecht 2001.

[25] V. L. Girko, “Theory of Stochastic Canonical Equations.II.” “Mathematics and Its Applications 535. Kluwer Academic, Dordrecht 2001.

[26] V. L. Girko, "Theory of random determinants," *Mathematics and its Applications (Soviet Series)*, Kluwer Academic Publishers Group, Dordrecht, 1990.

[27] V. L. Girko, "Convergence rate of the expected spectral functions of symetric random matrices equal to $o(n^{-1/2})$," *Random Operator and Stochastic Equations.*, Vol. 6, pp. 359-408, 1998.

[28] V. L. Girko, "Sluchainye Matristsy. Izdat. Ob [Random matrices]," *ed. Visca Skola*, pri Kiev. Gosudarstv. Univ. Kiev. MR1080966.

[29] V. L. Girko, "Thirty years of the central resolvent law and three laws on the 1/n expansion for resolvent of random matrices" "Random Oper. Stochastic Equations, 11(2):167-212, 2003.

[30] F. Gotze and A. N. Tikhomirov, "Asymptotic distribution of quadratic forms and applications," *J. Theoret. Probab.*, 15 424-475. MR1898815 2003.

[31] U. Grenander and J.W. Silverstein, "Spectral analysis of networks with random topologies," *SIAM J. Appl. Math.* 32(2):499-519, 1977.

[32] A. Guillonnet and O. Zeitouni, "Concentration of the Spectral Measure for Large Matrices," *Electronic Communications in Probability*,pp. 119-136, 2000.

[33] W. Hachem, P. Loubaton, and J. Najim, "The empirical distribution of the eigenvalues of a Gram matrix with a given variance profile," *Ann. Inst. H. Poincaré Prob. Statist.*, 42(6):649-670, 2006.

[34] W. Hachem, P. Loubaton, and J. Najim, "The empirical distribution of the eigenvalues of a Gram matrix with a given variance profile," *Ann. Inst. H. Poincaré Prob. Statist.*, 42(6):649-670, 2006.

[35] W. Hachem, P. Loubaton, and J. Najim, "Deterministic Equivalents for Certain Functionals of Large Random Matrices," *Ann. Appl. Probab.*, vol. 17, no. 3, pp. 875–930, 2007.

[36] W. Hachem, O. Khorunzhy, P. Loubaton, J. Najim and L. Pastur "A new approach for capacity analysis of large dimensional multi-atenna channels," *IEEE If. Theory.*, Vol. 54(9), sept. 2008.

[37] W. Hachem, Ph. Loubaton, and J. Najim, "A CLT For Information-Theoretic Statistics of Gram Random Matrices with a Given Variance Profile," *Ann. Appl. Probab.* (18) 2008, no. 6, 2071–2130.

[38] W. Hachem, Ph. Loubaton, J. Najim and P. Vallet, "On bilinear forms based on the resolvent of large random martices," *Technical report,* Télécom Paristech and Université Paris-Est, may 2010. arXiv: 1004.3848, available at http://front .math.ucdavis.edu/1004.3848.

[39] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge Univ. Press, 1994.

[40] K. Johansson, "On fluctuations of eigenvalues of random Hermitian matrices," *Duck Math.* J., 91(1):151-204, 1998.

[41] D. Jonsson, "Some limit theorems for the eigenvalues of a sample covariance matrix," *J. Multivariate Anal.* 12: 1-38.

[42] 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," *IEEE Inf. Theory*, vol. 55(11), nov. 2009.

[43] A. Kammoun, M. Kharouf, W. Hachem and J. Najim, "Fluctuations of the SNR at the Wiener Filter Output for Large Dimensional Signals," *IEEE Worshop on Signal Processing Advances in Wireless Communications*, 2008.

[44] A. Kammoun, M. Kharouf, W. Hachem and J. Najim, "Outage probability approximation for the Wiener Filter SINR in MIMO systems," *IEEE Worshop on Signal Processing Advances in Wireless Communications*, 2008.

[45] A. Kammoun, M. Kharouf, W. Hachem and J. Najim, "BER and Outage Probability Approximations for LMMSE Detectors on Correlated MIMO Channels," *IEEE Transactions on Information Theory*, Vol. 55, No. 10, pages 4386-4397, October 2009.

[46] A. Kammoun, M. Kharouf, W. Hachem J. Najim and A. Elkharroubi, "On the Fluctuations of the Mutual Information for non-Centred MIMO Channels: The non-Gaussian Case," *IEEE Workshop on Signal Processing Advances in Wireless Communications*, Marrakech, Morocco, 2010.

[47] A. M. Khorunzhy, B. A. Khoruzhenko and L. A. Pastur, "On the asymptotic properties of large random matrices with independent entries ," *Journal Math. Phys.* (1996).arXiv:cond-mat/9606174.

[48] J.H. Kotecha and A.M. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Trans. on Signal Processing*, vol. 52, no. 2, pp. 546–557, Feb. 2004.

[49] P. LI, D. Paul, R. Narasimhan and J. Cioffi, "On the Distribution of SINR for the MMSE MIMO Receiver and Performance Analysis," *IEEE Inf. Theory*, vol. 52, no. 1, pp. 271-286, Junuary 2006.

[50] L. Li, A.M. Tulino, and S. Verdu, "Design of reduced-rank MMSE multiuser detectors using random matrix theory , *IEEE Trans. on Information Theory*, vol. 50, no. 6, pp. 986-1008, June 2004.

[51] A. Lytova, L. Pastur, "Central limit theorem for linear eigenvalue statistics of random matrices with independent entries., *Avaible on Arxiv.*, 2009.

[52] V. A. Marcenko, L. A. Pastur "Distribution of eigenvalues in certain sets of rendom matrices," *Mat. SB (N. S.)*, 72 (114), 507-536. 1967.

[53] M. L. Mehta, "Random matrices," *Academic Press Inc. Boston, MA, second edition*, 1991.

[54] A.L. Moustakas and S.H. Simon, "On the outage capacity of correlated multiple-path MIMO channels," *IEEE Trans. on Information Theory*, vol. 53, no. 11, pp. 3887–3903, Nov. 2007.

[55] 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 Analysis," *IEEE Transactions on Information Theory* . Vol. 49, no.10, October 2003.

[56] R.R. Müller, "A Random Matrix Model of Communication Via Antenna Arrays," *IEEE Trans. on Information Theory*, Vol. 48, no. 9, pp. 2495-2506, Sept. 2002.

[57] N.K. Nikolski, *Operators, Functions and Systems: An Easy Reading. Vol. 2: Model Operators and Systems*, Mathematical Surveys and Monographs. AMS, 2002.

[58] G.M. Pan, M.H. Guo and W. Zhou, "Asymptotic distributions of the signal-to-interference ratios of LMMSE detection in multiuser communications, *Annals of Probability*, Vol. 17, No. 1, pp. 181-206, 2007.

[59] M.J.M. Peacock, I.B. Collings, and M.L. Honig, "Asymptotic spectral efficiency of multiuser multisignature CDMA in frequency-selective channels," *IEEE Trans. on Information Theory*, vol. 52, no. 3, pp. 1113–1129, Mar. 2006.

[60] W. Phoel and M.L. Honig, "Performance of coded DS-CDMA with pilot-assisted channel estimation and linear interference suppression," *IEEE Trans. on Communications*, vol. 50, no. 5, pp. 822–832, May 2002.

[61] H. V. Poor and S. Verdu, "Probability of error in MMSE multiuser detection, *IEEE Trans. on Information Theory*, vol. 43, no. 3, pp. 858-871, May. 1997.

[62] A.M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. on Signal Processing*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.

[63] A. M. Sengupta and P.P. Mitra, "Capacity of multivariate channels with multiplicative noise: random matrix techniques and large-n expansions for full transfer matrices," *Available at http://arxiv.org/abs/physics/0010081*, 2000.

[64] Shiu D.-S., G.J. Foschini, M.J. Gans, and J.M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. on Communications*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[65] D. Shlyakhtenko, "Random Gaussian band matrices and freeness with amalgamation," *Internat. Math. Res. Notices*, , no. 20, pp. 1013–1025, 1996.

[66] J. W. Silverstein, "Describing the behavior of random matrices using sequences of measures on orthogonal groups, *SIAM J. Math. Anal.*, 12 274-281. 1989.

[67] J. W. Silverstein, "Describing the behavior of eigenvectors of random matrices using sequences of measures on orthogonal groups, *SIAM Journal on Mathematical Analysis*, 12(2) pp. 274-281. 1981.

[68] J. W. Silverstein, "Some limit theorems on the eigenvectors of large dimensional sample covariance matrices, *J. Multivariate Anal.*, 15 295-324. 1984.

[69] J. W. Silverstein, "On the eigenvectors of large dimensional sample covariance matrices, *J. Multivariate Anal.*, 30 1-16. 1989.

[70] J. W. Silverstein, "Weak convergence of random functions defined by the eigenvectors of sample covariance matrices, *Annals of Probability*, 18(3) (1990), pp. 1174-1194.

[71] J. W. Silverstein, "The spectral radii and norms of large dimensional non-central random matrices, *Stochastic Models*, 10(3) (1994), pp. 525-532.

[72] J. W. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices, *Journal of Multivariate Analysis*, 54(2) (1995), pp. 295-309.

[73] M. K. Simon, M. S. Alouini, "Digital Communication Over Fading Channels, *Wiley Series in Telecommunications and Signal Processing*, second edition, 2005.

[74] Y. G. Sinai, A. B. Soshnikov, "Central limit theorem for traces of large random symmetric matrices with independent matrix elements, *Bol. Soc. Brasil. Mat. (N. S.)*, 29 1-24. MR1620151.

[75] A. Soshnikov, "The central limit theorem for local linear statistics in classical compact groups and related combinatorial identities," *Ann. Probab.*, 28(3): 1353-1370, 2000.

[76] A. Soshnikov, "Universality at the edge of the spectrum in Wigner random matrices," 1999.

[77] C.A. Tracy and H. Widom, "Distribution functions for largest eigenvalues and their applications," *Proceedings of the International Congress of Mathematicians*, Vol. I, (Beijing, 2002) (Beijing) Higher Ed. Press, pp. 587-596, 2002.

[78] G. Taricco, "Asymptotic mutual information statistics of separately correlated rician fading MIMO channels, *IEEE Trans. on Information Theory*, vol. 45, no. 2, pp. 641-657, Mar. 1999.

[79] G. Taricco, "On the capacity of separately-correlated MIMO Rician fading channels, *Proc. 49th Annual IEEE Globcom Conference*, 2006, San Francisco (CA), USA, Novembre 2006.

[80] I.E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," *European Transactions on Telecommunications*, Vol. 10, no. 6, pp. 585–595, Nov.-Dec. 1999.

[81] D.N.C Tse and O. Zeitouni, "Linear multiuser receivers in random environments, *IEEE Trans. on Information Theory*, vol. 46, no. 1, pp. 171-188, Jan. 2000.

[82] D.N.C Tse and S. Hanly, "Linear multi-user receiver: Effective interference, effective bandwidth and user capacity, *IEEE Trans. on Information Theory*, vol. 45, no. 2, pp. 641-657, Mar. 1999.

[83] A.M. Tulino, L. Li, and S. Verdú, “Spectral Efficiency of Multicarrier CDMA,” *IEEE Trans. on Information Theory*, vol. 51, no. 2, pp. 479–505, Feb. 2005.

[84] A.M. Tulino, A. Lozano, and S. Verdú, “Impact of Antenna Correlation on the Capacity of Multiantenna Channels,” *IEEE Trans. on Information Theory*, Vol. 51, no. 7, pp. 2491-2509, July 2005.

[85] A. Tulino and S. Verdu, “Random matrix theory and wireless communications,” *Foundations and Trends in Communications and Information Theory* 1-182, 2007.

[86] S. Verdu and Sh. Shamai, “Spectral efficiency of CDMA with random spreading, *IEEE Trans. on Information Theory*, vol. 45, no. 2, pp. 622-640, Mar. 1999.

[87] K.W. Wachter, “The limiting empirical measure of multiple discriminant ratios,” *Ann. Stat.* 8: 937-957, 1980.

[88] E. P. Wigner, “On the distributions of the roots of certain symmetric matrices,” *Ann. Math.*, 67, 325-327. 1958.

[89] Y.Q. Yin, “Limiting spectral distribution for a class of random matrices,” *J. Multivariate Anal.* 20: 50-68, 1986.

[90] Y.Q. Yin and P.R. Krishnaiah, “A limit theorem for the eigenvalues of product of two random matrices,” *J. Multivariate Anal.* 13: 489-507, 1983.