
Tutoriel sur les matrices aléatoires – Introduction aux équivalents déterministes

Axel Müller¹, Mérouane Debbah²

*Mathematical and Algorithmic Sciences Lab
France Research Center, Huawei Technologies Co. Ltd.
Arcs de Seine Bâtiment A, 20 Quai du Point du Jour
92100 Boulogne-Billancourt, France
{axel.mueller,merouane.debbah}@huawei.com*

RÉSUMÉ. Dans cet article, nous décrivons une approche analytique des Matrices Aléatoires avec des applications en Télécommunications. L'approche en question utilise les équivalents déterministes et permet le design des réseaux du futur quand le nombre d'utilisateurs, le nombre de stations de base et le nombre d'antennes augmente au même rythme. Nous étudions un exemple particulier de détermination de débit dans les réseaux sans fils pour lequel l'équivalent déterministe est calculé étape par étape. Nous décrivons également l'ensemble des concepts, lemmes et outils de base issus des matrices aléatoires nécessaires à la bonne compréhension de la théorie.

ABSTRACT. In the following, we provide a tutorial on the practical application of random matrix theory (RMT) in communication problems, in order to facilitate the utilization of the analytic RMT approach for interested researchers. To this end, we first state the necessary basic theoretical concepts, lemmas and tools from RMT. After this, we build intuition, confidence, and insight into RMT concepts and their applications, by putting the introduced theoretical results into a tutorial like context. To familiarize the reader with the introduced tools, we rely on an example of a step by step derivation of the deterministic equivalent for a non-trivial rate problem. Thus, we provide the theory needed to soundly use the framework of RMT in the design of future large scale (w.r.t. the numbers of users, base stations and antennas) communication networks.

MOTS-CLÉS : Théorie des Matrices aléatoires, Equivalents Déterministes, MIMO, Télécommunications.

KEYWORDS: Random Matrix Theory, Deterministic Equivalents, MIMO, Telecommunication.

DOI:10.3166/TS.x.1-27 © 2015 Lavoisier

Extended Abstract (in English)

In order to satisfy increasing throughput and other performance demands, future networks will need to adopt new approaches of network design, which ultimately result in very large systems (w.r.t. the numbers of users, base stations (BSs) and antennas). Most of the analytic tools used today in the communications community were originally developed for the analysis of point-to-point communication or small MIMO systems. Therefore, it is not a surprise that they often fail to provide meaningful insight into possible solutions for a new era of large dense heterogeneous multi user multi cell communication systems. New tools, which are adapted to the large nature of the systems, need to be developed and applied to give insight into and find the right combination of future system design solutions. Fortunately, the mathematical tool of large scale random matrix theory (RMT) has matured in recent years, to a point where it is of excellent use in this task. Publications documenting this maturing process (e.g., (Bai, Silverstein, 1998b ; Tulino, Verdú, 2004 ; Hachem, Khorunzhy *et al.*, 2008 ; Bai, Silverstein, 2009 ; Couillet, Debbah, 2011 ; Hoydis, 2012)) are numerous and they are not necessarily adapted to helping interested researchers to enter the field. This article introduces the theory needed to soundly use the framework of RMT, especially the technique known as *deterministic equivalent* (DE) approach, in the form of a tutorial and adheres to a predominately pedagogical style.

DEs are mathematical objects that are able to provide accurate deterministic approximations of random quantities, which represent important system performance indicators in cellular networks. For example, the capacity of large dimensional multi antenna channels. We define the DE of a sequence of random quantities as follows:

DÉFINITION 1 (Deterministic Equivalent (Hachem *et al.*, 2007)). — *The deterministic equivalent of a sequence of random complex values $(X_n)_{n \geq 1}$ is a deterministic sequence $(\bar{X}_n)_{n \geq 1}$, which approximates X_n such that*

$$X_n - \bar{X}_n \xrightarrow[n \rightarrow +\infty]{a.s.} 0$$

where *a.s.* is taken to mean almost sure convergence.

Quite often the quantity X_n is going to be a functional of the resolvent of a Hermitian matrix. Also, even relatively simple problems can result in a DE \bar{X}_n , which is not guaranteed to converge itself. Yet, it is possible to deterministically calculate a DE \bar{X}_n . Intuitively, one can remark that \bar{X}_n is still “contains” the factor n , even as $n \rightarrow \infty$. In fact, the DE gives us an approximation for each value of N , which becomes more precise for increasing N . The realizations of the random variable, almost surely (*a.s.*) fall within an increasingly narrow bound around the DE. Thus, DEs tend to be much more accurate for finite (and even small) system dimensions than the classical limits. Assume we analyse a random quantity involving a random matrix $\mathbf{X} \in \mathbb{C}^{M \times K}$. RMT handles cases in which both dimensions (M, K) grow large, while classic limit approximations (e.g., the strong law of large numbers) can only treat the case where M grows large. As a consequence, RMT results exploit more degrees of freedom than classical approaches and, thus, usually far outperform them w.r.t. convergence speed.

For example, even an 8×8 matrix offers already up to 64 degrees of freedom, which mostly leads to quite acceptable convergence. In general, RMT achieves impressive convergence rates for linear functionals of eigenvalues, e.g., for central limit theorems in $1/M$ (i.e., $M(X_M - \bar{X}_M) \rightarrow \mathcal{N}(0, 1)$) (Hachem, Loubaton, Najim, 2008, Remark 3.1) and for expectations in $1/M^2$ (i.e., $\mathbb{E}[X_M] = \bar{X}_M + \mathcal{O}(1/M^2)$) (Pastur, Shcherbina, 2011)(Vallet *et al.*, 2010, Proposition 4), when the random quantities are complex Gaussian distributed. Quadratic forms are usually slower, for example central limit theorems in $1/\sqrt{M}$ (Kammoun *et al.*, 2009, Theorem 3) and expectations in $1/M^{\frac{3}{2}}$ (Rubio *et al.*, 2012, Lemma 1). The well-known RMT Lemmas 15, 17, 20 and 22 are of frequent use in finding such DEs in the first place. Practical usage of DEs is very intuitive (i.e., summation and multiplication behave as one would expect), as is shown by the results stated in Theorems 12, 13, and Lemma 14.

1. Introduction

Afin de satisfaire les demandes croissantes de débits et autre performances, il est communément admis que les futurs réseaux de télécommunication devront adopter des approches complémentaires pour le design des réseaux, à savoir la densification (réseaux à petites cellule), les systèmes d'antennes larges (Massive Multiple Input Multiple Output (MIMO)), les réseaux hétérogènes et la coordination/coopération (réseaux MIMO). Tous ces nouveaux paradigmes impliquent des réseaux très larges en ce qui concerne le nombre d'utilisateurs, le nombre de stations de base et le nombre d'antennes en général. La plupart des outils analytiques utilisés aujourd'hui ont été développés à l'origine pour l'analyse des communications point à point ou pour des petits systèmes MIMO. Ce n'est donc pas surprenant que ces outils ne permettent pas de comprendre les paramètres d'intérêt des réseaux hétérogènes multi-utilisateurs et multi-cellulaires. De nouveaux outils, adaptés à la nature large des systèmes, doivent être développés et appliqués à ces réseaux afin de trouver les configurations optimales. Ces dernières années, un effort considérable a été fait dans ce domaine et la théorie pour le design de ces systèmes est maintenant assez mûre. Nous pouvons citer en particulier les travaux de (Bai, Silverstein, 1998b ; Tulino, Verdú, 2004 ; Hachem, Khorunzhy *et al.*, 2008 ; Bai, Silverstein, 2009 ; Couillet, Debbah, 2011 ; Hoydis, 2012)) qui permettront au lecteur averti de connaître le domaine mais ne seront pas adaptés pour des chercheurs nouveaux dans le domaine.

Dans cet article, nous abordons le cadre général de la Théorie des Matrices Aléatoires, en particulier l'approche des *Equivalents Déterministes*. Afin de familiariser le lecteur avec les outils introduits, nous considérons l'exemple non-trivial de la détermination du rapport signal à interférence plus bruit (Signal to Interference plus Noise Ratio, SINR) pour lequel nous expliquons étape par étape l'obtention des Equivalents Déterministes pour différentes structures de réception.

La plupart des concepts introduits dans cet article ont été abordés dans des travaux précédents (par exemple, (Couillet, Debbah, 2011 ; Hoydis, 2012)). Nous nous distinguons de ces travaux par une approche plus pédagogique, à l'instar de (Müller,

2014). En particulier, nous aidons le lecteur à mieux comprendre les transitions des différentes étapes permettant d'obtenir le résultat. Nous n'insistons pas sur les détails d'ordre purement mathématiques lorsqu'ils ne sont pas essentiels. Cet article est par conséquent adapté aux lecteurs non-experts du domaine intéressés par l'utilisation de l'approche analytique de la théorie des matrices aléatoires et de ses applications en télécommunications.

2. La Transformée de Stieltjes

Le point de départ dans l'introduction de la théorie des matrices aléatoires débute avec la définition de la Transformée de Stieltjes. Ceci est principalement dû à l'histoire de la discipline avec les travaux de Marčenko et Pastur qui ont utilisé cette approche (Marčenko, Pastur, 1967) pour déterminer la distribution des valeurs propres de certaines matrices aléatoires. D'autres travaux ont suivi (par exemple, (Telatar, 1999; Tse, Hanly, 1999; Verdú, Shamai, 1999)), étendu (par exemple, (Silverstein, Bai, 1995; Tulino, Verdú, 2004)) ou se sont développés autour (par exemple, (Hachem *et al.*, 2007)) de cette approche dans le contexte des systèmes de télécommunications¹. Nous verrons que cette approche est très adaptée au contexte des télécommunication en raison du lien intime entre Transformée de Stieltjes et SINR.

Commençons tout d'abord par certaines définitions:

DÉFINITION 2. — *Pour une mesure μ qui assigne une mesure finie à un ensemble borné sur \mathbb{R} , nous notons:*

$$F_\mu(x) = \mu((-\infty, x]).$$

Si μ est une mesure de probabilité, alors F_μ est appelée la fonction de répartition (cumulative distribution function (cdf)).

Nous définissons maintenant la Transformée de Stieltjes d'une mesure par:

DÉFINITION 3 (Transformée de Stieltjes). — *Soit μ une mesure finie non négative avec pour support $\text{supp}(\mu) \subset \mathbb{R}$, $\mu(\mathbb{R}) < \infty$, et F_μ définie comme dans la définition 2. La Transformée de Stieltjes $m(z)$ de μ est définie $\forall z \in \mathbb{C} \setminus \text{supp}(\mu)$ comme*

$$m(z) = \int_{\mathbb{R}} \frac{1}{\lambda - z} \mu(d\lambda) = \int_{\mathbb{R}} \frac{1}{\lambda - z} dF_\mu(\lambda). \quad (1)$$

La seconde égalité (1) n'est pas d'emblée évidente et Billingsley (Billingsley, 1995) nous invite à mieux cerner les différences de notation entre $\int f(x)\mu(dx)$ et $\int f(x)dF_\mu(x)$ ².

1. Le travail de Marčenko et Pastur sur le spectre des matrices aléatoires a été précédé par ceux de Wigner (Wigner, 1958).

2. Le lecteur intéressé est invité à étudier (Billingsley, 1995, (17.22)ff.) concernant les distinctions subtiles entre l'intégrale de Riemann-Stieltjes et l'intégrale de Lebesgue-Stieltjes.

A noter que la littérature utilise $\int_{\mathbb{R}} \frac{1}{\lambda-z} d\mu(\lambda)$ comme une notation alternative de (1)³.

Les propriétés suivantes permettent de retrouver μ quand seulement la Transformée de Stieltjes $m(z)$ est connue. Ces résultats sont disponibles dans les références suivantes (Hachem *et al.*, 2007) ou (Krein, Nudelman, 1977).

PROPRIÉTÉ 4. — Soit $m(z)$ la Transformée de Stieltjes d'une mesure finie μ sur \mathbb{R} . alors ,

$$(i) \mu(\mathbb{R}) = \lim_{y \rightarrow \infty} -\mathbf{i}ym(\mathbf{i}y),$$

(ii) $\mu([a, b]) = \lim_{y \rightarrow 0^+} \frac{1}{\pi} \int_a^b \Im\{m(x + \mathbf{i}y)\} dx$, si a, b sont des points de continuité de μ .

Nous continuons avec la définition de la mesure de probabilité empirique des valeurs propres d'une matrice Hermitienne \mathbf{X} .

DÉFINITION 5 (Mesure Empirique de Probabilité des Valeurs Propres). — Soit $\mathbf{X} \in \mathbb{C}^{N \times N}$ une matrice Hermitienne avec des valeurs propres réelles $\lambda_1, \dots, \lambda_N$. La mesure empirique de probabilité $\mu_{\mathbf{X}}$ des valeurs propres de \mathbf{X} est défini par

$$\mu_{\mathbf{X}}(A) = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i(\mathbf{X}) \in A}.$$

Les notations équivalentes $\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i(\mathbf{X})}(A)$ et $\frac{1}{N} \sum_{i=1}^N \mathbb{1}_A(\lambda_i(\mathbf{X}))$ sont également souvent utilisées dans la littérature. Cette mesure constitue une mesure de comptage normalisée.

Nous définissons sa fonction de répartition correspondante (selon la définition 2) comme

DÉFINITION 6 (Répartition Empirique du Spectre (Empirical Spectral Distribution (e.s.d.))). — Soit $\mu_{\mathbf{X}}(a)$ la mesure de probabilité empirique des valeurs propres \mathbf{X} définie selon la définition 5. La fonction de répartition empirique $F^{\mathbf{X}}(x)$ des valeurs propres de \mathbf{X} est donnée par:

$$F^{\mathbf{X}}(x) = \mu_{\mathbf{X}}((-\infty, x]) = \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{\{\lambda_i(\mathbf{X}) \leq x\}}.$$

Nous verrons dans les lignes suivantes l'intérêt de la Transformée de Stieltjes appliquée à la répartition empirique du spectre, qui apparait pour le moment comme une manière compliquée d'obtenir la distribution des valeurs propres. En effet, cette transformée permet comme le souligne (Tao, 2012) d'obtenir l'information spectrale de manière compacte et d'être aisément manipulée par les méthodes d'analyse complexe.

3. Cette pratique répandue vient de l'utilisation de la relation $\int dF(x) = \int F(dx)$ par différents auteurs.

Afin de montrer le lien entre la Transformée de Stieltjes et le spectre des matrices Hermitienne, nous introduisons la notion de résolvante \mathbf{Q} d'une matrice Hermitienne \mathbf{X} :

$$\mathbf{Q}(z) = (\mathbf{X} - z\mathbf{I}_M)^{-1} .$$

Ou plus généralement:

DÉFINITION 7 (Notation pour la Résolvante). — *La résolvante \mathbf{Q}_M de la matrice $\mathbf{A}_M \in \mathbb{C}^{M \times M}$ est définie par*

$$\mathbf{Q}_M(z) = (\mathbf{A}_M - z\mathbf{I})^{-1} .$$

Elle est définie pour tout $z \in \mathbb{C}$ différent des valeurs propres de \mathbf{A}_M .

La résolvante est un concept central en théorie spectrale. En particulier, elle indique les valeurs propres de \mathbf{X} en définissant le support de la variable complexe scalaire z .

En reprenant la définition de la Transformée de Stieltjes et en lui associant la mesure de probabilité empirique $\mu_{\mathbf{X}}$ de la définition 5, on peut directement obtenir:

$$m_{\mu_{\mathbf{X}}}(z) = \int_{\mathbb{R}} \frac{1}{\lambda - z} \mu_{\mathbf{X}}(d\lambda) = \frac{1}{M} \sum_{i=1}^M \frac{1}{\lambda_i(\mathbf{X}) - z} .$$

Ce qui permet d'obtenir:

$$\begin{aligned} m_{\mu_{\mathbf{X}}}(z) &= \frac{1}{M} \text{tr} \text{diag} \left(\frac{1}{\lambda_1(\mathbf{X}) - z}, \dots, \frac{1}{\lambda_M(\mathbf{X}) - z} \right) \\ &= \frac{1}{M} \text{tr} \left\{ [\text{diag}(\lambda_1(\mathbf{X}), \dots, \lambda_M(\mathbf{X})) - z\mathbf{I}_M]^{-1} \right\} \\ &\triangleq \frac{1}{M} \text{tr} \left[(\mathbf{\Lambda} - z\mathbf{I}_M)^{-1} \right] \end{aligned}$$

pour toute matrice unitaire $\mathbf{U} \in \mathbb{C}^{M \times M}$

$$m_{\mu_{\mathbf{X}}}(z) = \frac{1}{M} \text{tr} \left[(\mathbf{\Lambda} \mathbf{U} \mathbf{U}^H - z \mathbf{U} \mathbf{U}^H)^{-1} \right] = \frac{1}{M} \text{tr} \left[(\mathbf{U} \mathbf{\Lambda} \mathbf{U}^H - z \mathbf{I}_M)^{-1} \right]$$

Si maintenant \mathbf{U} est choisi pour contenir les vecteurs propres de la matrices Hermitienne \mathbf{X} , nous obtenons finalement

$$m_{\mu_{\mathbf{X}}}(z) = \frac{1}{M} \text{tr} \left[\underbrace{(\mathbf{X} - z\mathbf{I}_M)^{-1}}_{\text{Resolvent of } \mathbf{X}} \right]. \quad (2)$$

Pour des questions de simplicité, nous utiliserons la notation suivante $m_{\mu_{\mathbf{X}}}(z)$ à la place de $m_{\mathbf{X}}(z)$ lorsque cela ne pose pas de problèmes de compréhension.

3. Equivalents Déterministes

Nous allons maintenant introduire un des concepts les plus importants en théorie des matrices aléatoires (du moins lorsque des résultats du premier ordre sont concernés) pour l'analyse des systèmes de télécommunications sans fils, à savoir la définition des *Equivalents Déterministes*.

Afin de définir les Equivalents Déterministes, il est nécessaire d'introduire le concept de convergence presque sûre d'une séquence de variables aléatoires:

DÉFINITION 8 (Convergence Presque Sûre). — *La séquence de variables $(X_n)_{n \geq 1}$ converge presque sûrement vers X , si*

$$P \left(\limsup_{n \rightarrow \infty} |X_n - X| = 0 \right) = 1.$$

La notation suivante est utilisée $X_n \xrightarrow[n \rightarrow \infty]{a.s.} X$ ou $X_n \xrightarrow{a.s.} X$, si le contexte ne porte pas à confusion.

Nous définissons l'Equivalent Déterministe d'une séquence de quantités aléatoires de la manière suivante:

DÉFINITION 9 (Equivalent déterministe). — *L'Equivalent Déterministe d'une séquence de valeurs aléatoires complexes $(X_n)_{n \geq 1}$ est une séquence déterministe $(\bar{X}_n)_{n \geq 1}$, qui approxime X_n de telle manière que*

$$X_n - \bar{X}_n \xrightarrow[n \rightarrow +\infty]{a.s.} 0.$$

Les Equivalents Déterministes furent initialement proposés sous cette forme par Hachem *et al.* in (Hachem *et al.*, 2007 ; Hachem, Khorunzhy *et al.*, 2008). Les Equivalents Déterministes permettent d'obtenir des approximations déterministes précises d'indicateurs de performance d'importants systèmes de réseaux cellulaires. Un exemple connu est la capacité des réseaux multi-antennes larges.

Souvent, la quantité X_n sera une fonctionnelle de la résolvante d'une matrice Hermitienne. C'est le cas par exemple de la trace normalisée, qui d'après (2) est la Transformée de Stieltjes d'une mesure de probabilité. Cependant, nous sommes souvent intéressés à des formes plus compliquées liées à des propriétés spectrales. La quantité X_n se concentrera autour de \bar{X}_n dans le régime des grandes valeurs de n et si \bar{X}_n a une limite, nous obtenons même une convergence presque sûre. Il est à noter que même des problèmes simples peuvent aboutir à des Equivalents Déterministes \bar{X}_n qui ne convergent pas forcément. Dans ce cas, il est tout de même possible de calculer de manière déterministe \bar{X}_n .

Dans les applications pratiques des Equivalents Déterministes, les termes "limite presque sûre" et "Approximations large échelle" sont souvent interchangeables.

REMARQUE 10 (Limite Presque Sûre). — Si une séquence de variables complexes aléatoires $(X_n)_{n \geq 1}$ converge presque sûrement vers une quantité simple déterministe \bar{X} ,

$$X_n \xrightarrow[n \rightarrow +\infty]{\text{a.s.}} \bar{X}$$

alors nous appelons cette quantité \bar{X} la "limite presque sûre" de X_n . Souvent, la notation suivante est utilisée $\lim X_n = \bar{X}$. \square

REMARQUE 11 (Approximations Large Echelle). — Si un Equivalent Déterministe est utilisé comme approximation à n fini, nous appellerons cela alors une *approximation large échelle*. \square

Nous insistons sur le fait que même si les concepts de Transformée de Stieltjes et Equivalent Déterministe sont introduits en parallèle, il s'agit de concepts complètement différents. La Transformée de Stieltjes est un outil (précis et non-asymptotique) permettant l'analyse du spectre de matrices et de son analyse complexe via la distribution empirique du spectre. L'Equivalent Déterministe est une approximation déterministe d'une séquence de quantités aléatoires, représentant souvent un indicateur de performance d'un problème défini par des quantités aléatoires. Cependant, il se trouve que les Equivalents Déterministes et les Transformées de Stieltjes sont souvent faciles à trouver et beaucoup d'indicateurs de performances s'expriment sous la forme de Transformées de Stieltjes

Les théorèmes et Lemmes suivants permettront de manipuler les équivalents déterministes.

Un théorème très pratique, connu en anglais sous le nom de Continuous Mapping theorem établit qu'une fonction f continue appliquée à une variable qui converge vers X convergera vers $f(X)$ pour tous les modes de convergence.

THÉORÈME 12 (Théorème du prolongement par continuité (Van der Vaart, 2000, Théorème 2.3)). — Soit $(X_n)_{n \geq 1}$ une séquence de variables aléatoires réelles et soit $f : \mathbb{R} \mapsto \mathbb{R}$ continue en tout point de l'ensemble A avec $P(X \in A) = 1$, pour une variable aléatoire X . Alors si $X_n \xrightarrow{\text{a.s.}} X$, cela implique que $f(X_n) \xrightarrow{\text{a.s.}} f(X)$.

Ce théorème formule le fait qu'une fonction d'un Equivalent Déterministe se comporte comme il faut pour les valeurs qu'elle approxime.

Dans certains cas, on peut prouver que $X_n \xrightarrow{\text{a.s.}} X$ mais nous sommes aussi intéressés par le fait que $(X_n)_{n \geq 1}$ converge en moyenne vers X , à savoir, $\lim_n \mathbb{E} [|X_n - X|] = 0$. Cela est une convergence plus faible, qui peut aussi convenir, et peut être obtenue par le théorème de la convergence dominée.

THÉORÈME 13 (Théorème de la Convergence Dominée (Billingsley, 1995, Theorem 16.4)). — Soit $(f_n)_{n \geq 1}$ une séquence de fonctions mesurables réelles telle que la limite $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ existe. Supposons qu'il existe une fonction intégrable

$g : \mathbb{R} \mapsto [0, \infty]$ avec $|f_n(x)| \leq g(x)$ pour chaque $x \in \mathbb{R}$. Alors si f est intégrable, nous avons f_n pour chaque n , et

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} f_n d\mu = \int_{\mathbb{R}} f d\mu.$$

L'argument standard pour montrer que la convergence presque sûre de l'Equivalent Déterministe induit la convergence en moyenne peut être formulé de la manière suivante: Considérons les fonctions $f_n = |X_n - X|$ pour tout n . Comme $X_n \xrightarrow{\text{a.s.}} X$, cela implique que $f_n \xrightarrow{\text{a.s.}} f = 0$. Si on peut montrer que $f_n \leq g$ et $\mathbb{E}[g] < \infty$, alors le théorème de la convergence dominée permet de montrer que $\lim_{n \rightarrow \infty} \mathbb{E}[|X_n - X|] = 0$.

Par exemple, les Transformées de Stieltjes sont bornées par $1/|z|$ pour des mesures à support réel, à savoir la mesure empirique de probabilité des valeurs propres dans la définition 5. Ainsi, les Transformées de Stieltjes de cette mesure sont des fonctions bornées, ce qui permet de conclure sur la convergence en moyenne à partir de la convergence de la Transformée de Stieltjes.

Le dernier Lemme est important lorsque nous considérons des produits ou ratios des Equivalents Déterministes.

LEMME 14. — (Peacock et al., 2008, Lemme 1) Soient $(a_n)_{n \geq 1}$ et $(b_n)_{n \geq 1}$ deux séquences de variables complexes aléatoires. Soient $(\bar{a}_n)_{n \geq 1}$ et $(\bar{b}_n)_{n \geq 1}$ deux séquences déterministes de quantités complexes. Supposons que $a_n - \bar{a}_n \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0$ et $b_n - \bar{b}_n \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0$.

(i) Si $|a_n|, |\bar{b}_n|$ et/ou $|\bar{a}_n|, |b_n|$ sont presque sûrement bornées⁴, alors

$$a_n b_n - \bar{a}_n \bar{b}_n \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0.$$

(ii) Si $|a_n|, |\bar{b}_n|^{-1}$ et/ou $|\bar{a}_n|, |b_n|^{-1}$ sont presque sûrement bornées, alors

$$a_n / b_n - \bar{a}_n / \bar{b}_n \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0.$$

Ce Lemme nous permet d'utiliser une approche "mix and match" ou "divide and conquer" pour calculer des Equivalents Déterministes qui concernent des produits ou sommes. Afin d'être plus précis, les théorèmes 12 et 13, combinés avec le Lemme 14, permettent de trouver directement les Equivalents Déterministes de certaines fonctions continues du SINR alors que seuls les Equivalents Déterministes de l'interférence et de la puissance du signal ont été calculés.

4. Toutes les quantités x_n sont telles que $\limsup |x_n| < \infty$ avec une probabilité 1.

4. Outils et Lemmes communs pour la Théorie des Matrices Aléatoires

Avant d'avancer dans les calculs relatifs à la Théorie des Matrices Aléatoires, nous avons besoin de quelques outils et lemmes standard qui seront très utiles par la suite.

LEMME 15 (Identités Matricielles). — Soient \mathbf{A} , \mathbf{B} des matrices complexes inversibles et \mathbf{C} une matrice complexe rectangulaire. Nous avons alors les relations suivantes:

Identité de Woodbury:

$$(\mathbf{A} + \mathbf{C}\mathbf{B}\mathbf{C}^H)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{B}^{-1} + \mathbf{C}^H\mathbf{A}^{-1}\mathbf{C})^{-1}\mathbf{C}^H\mathbf{A}^{-1}. \quad (3)$$

Identité de Searl:

$$(\mathbf{I} + \mathbf{A}\mathbf{B})^{-1}\mathbf{A} = \mathbf{A}(\mathbf{I} + \mathbf{B}\mathbf{A})^{-1}. \quad (4)$$

Identité de la Résolvante:

$$\mathbf{A}^{-1} - \mathbf{B}^{-1} = -\mathbf{A}^{-1}(\mathbf{A} - \mathbf{B})\mathbf{B}^{-1} = \mathbf{A}^{-1}(\mathbf{B} - \mathbf{A})\mathbf{B}^{-1}. \quad (5)$$

Un des lemmes les plus important dans la Théorie des Matrices Aléatoire concerne le *Lemme de la Trace*. Il concerne la convergence de formes quadratiques et fut introduit par (Bai, Silverstein, 1998a). Dans la suite, nous continuons à considérer des séquences de matrices et vecteurs aléatoires avec des dimensions qui augmentent à savoir, $(\mathbf{A}_M)_{M \geq 1} \in \mathbb{C}^{M \times M}$ et $(\mathbf{x}_M)_{M \geq 1} \in \mathbb{C}^M$ ou $(\mathbf{y}_M)_{M \geq 1} \in \mathbb{C}^M$. Cependant, afin de faciliter la lecture, nous utiliserons les abréviations suivantes $(\mathbf{A}_M)_{M \geq 1}$ pour \mathbf{A}_M ou même comme \mathbf{A} , si le contexte n'est pas ambigu.

LEMME 16 (Résultats Préliminaires sur le Lemme de la Trace (Bai, Silverstein, 2009, Lemma B.26)). — Soit $\mathbf{A} \in \mathbb{C}^{M \times M}$ déterministe et $\mathbf{x} = [x_1 \dots x_M]^T \in \mathbb{C}^M$ un vecteur aléatoire avec des composantes indépendantes. Supposons que $\mathbb{E}[x_i] = 0$, $\mathbb{E}[|x_i|^2] = 1$, et $\mathbb{E}[|x_i|^\ell] \leq v_\ell < \infty$ pour chaque $\ell \leq 2p$. Alors, pour chaque $p \geq 1$,

$$\mathbb{E}[|\mathbf{x}^H \mathbf{A} \mathbf{x} - \text{tr} \mathbf{A}|^p] \leq C_p (\text{tr} \mathbf{A} \mathbf{A}^H)^{\frac{p}{2}} \left(v_{\frac{p}{4}} + v_{2p} \right)$$

où C_p est une constante qui ne dépend que de p .

LEMME 17 (Lemme de la Trace (Bai, Silverstein, 1998a)). — Soit $\mathbf{x}_M = [x_1, \dots, x_M]^T$ un vecteur $M \times 1$ où x_m sont des variables complexes Gaussiennes i.i.d de variance unité. Soit \mathbf{A}_M une matrice $M \times M$ indépendante de \mathbf{x}_M . Si en plus, nous avons $\limsup_M \|\mathbf{A}_M\|_2 < \infty$, alors nous avons le résultat suivant:

$$\frac{1}{M} \mathbf{x}^H \mathbf{A}_M \mathbf{x} - \frac{1}{M} \text{tr}(\mathbf{A}_M) \xrightarrow[M \rightarrow +\infty]{a.s.} 0. \quad (6)$$

PREUVE 18. — Une conséquence du lemme 16 est que pour tout $p \geq 2$, il existe une constante C_p , qui ne dépend que de p , telle que

$$\mathbb{E}_{\mathbf{x}_M} \left[\left| \frac{1}{M} \mathbf{x}_M^H \mathbf{A}_M \mathbf{x}_M - \frac{1}{M} \operatorname{tr}(\mathbf{A}_M) \right|^p \right] \leq \frac{C_p}{M^p} \left((\mathbb{E}|x_m|^4 \operatorname{tr}(\mathbf{A} \mathbf{A}^H))^{p/2} + \mathbb{E}|x_m|^{2p} \operatorname{tr}(\mathbf{A} \mathbf{A}^H)^{p/2} \right)$$

où l'espérance concerne la distribution \mathbf{x}_M . Si en plus $\limsup_M \|\mathbf{A}\|_2 < \infty$ avec le fait que $\operatorname{tr}(\mathbf{A} \mathbf{A}^H) \leq M \|\mathbf{A}\|_2^2$ et $\operatorname{tr}(\mathbf{A} \mathbf{A}^H)^{p/2} \leq M \|\mathbf{A}\|_2^p$, nous obtenons alors l'inégalité plus simple suivante:

$$\mathbb{E}_{\mathbf{x}_M} \left[\left| \frac{1}{M} \mathbf{x}_M^H \mathbf{A}_M \mathbf{x}_M - \frac{1}{M} \operatorname{tr}(\mathbf{A}_M) \right|^p \right] \leq \frac{C'_p \|\mathbf{A}\|_2^p}{M^{p/2}}$$

où $C'_p = C_p \left((\mathbb{E}[|x_m|^4])^{p/2} + \mathbb{E}[|x_m|^{2p}] \right)$. En choisissant $p = 4$ (l'ordre du moment adéquate existe) et en utilisant l'inégalité de Markov (Billingsley, 1995, Equation (5.31)), nous avons:

$$P \left(\left| \frac{1}{M} \mathbf{x}_M^H \mathbf{A}_M \mathbf{x}_M - \frac{1}{M} \operatorname{tr}(\mathbf{A}_M) \right| > \varepsilon \right) \leq \frac{1}{\varepsilon^4} \frac{C'_4 \|\mathbf{A}\|_2^4}{M^2}.$$

En observant que cette quantité est sommable,

$$\sum_M P \left(\left| \frac{1}{M} \mathbf{x}_M^H \mathbf{A}_M \mathbf{x}_M - \frac{1}{M} \operatorname{tr}(\mathbf{A}_M) \right| > \varepsilon \right) < +\infty$$

nous pouvons appliquer le premier lemme de Borel-Cantelli (Billingsley, 1995, Theorem 4.3) pour obtenir le résultat suivant de convergence presque sûre:

$$\frac{1}{M} \mathbf{x}^H \mathbf{A}_M \mathbf{x} - \frac{1}{M} \operatorname{tr}(\mathbf{A}_M) \xrightarrow[M \rightarrow +\infty]{a.s.} 0.$$

■

D'autres versions de ce résultat existent, qui sont adaptées aux variantes spécifiques des hypothèses et du problème. Par exemple, (Wagner *et al.*, 2012, Lemma 4) montre que $\limsup_M \|\mathbf{A}\|_2 < \infty$, doit être vrai presque sûrement et les hypothèses sur la contrainte \mathbf{x}_M i.i.d. peuvent être remplacées par une simple indépendance (voir Lemme 16 et (Hoydis, 2012)).

Un complément naturel du Lemme 17 est le lemme suivant,

LEMME 19 ((Couillet, Debbah, 2011, Lemma 3.7)). — Soit \mathbf{A}_M défini comme au Lemme 17, à savoir $\limsup_M \|\mathbf{A}\|_2 < \infty$, et $\mathbf{x}_M, \mathbf{y}_M$ aléatoires, mutuellement indépendants avec des composantes complexes gaussiennes de moyenne nulle et de variance 1. Alors, pour tout $p \geq 2$, nous avons

$$\mathbb{E} \left[\left| \frac{1}{M} \mathbf{y}_M^H \mathbf{A}_M \mathbf{x}_M \right|^p \right] = O(M^{-p/2}).$$

En particulier,

$$\frac{1}{M} \mathbf{y}_M^H \mathbf{A}_M \mathbf{x}_M \xrightarrow[M \rightarrow +\infty]{a.s.} 0. \quad (7)$$

Ce Lemme montre que les quantités aléatoires similaires à des formes quadratiques tendent vers zéro asymptotiquement.

Nous avons vu que les Lemmes précédents requièrent une indépendance statistique entre la matrice et les vecteurs en question. Cela n'est souvent pas le cas. Par conséquent, les deux Lemmes suivants d'inversion de matrice peuvent être utilisés pour retirer les colonnes interférentes de la matrice. Cela sera effectivement le cas pour les matrices de Gram, à savoir des matrices de la forme $\mathbf{X}\mathbf{X}^H = \sum_m \mathbf{x}_m \mathbf{x}_m^H$, pour $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_M] \in \mathbb{C}^{M \times M}$.

LEMME 20 (Lemme I d'inversion de Matrice (Henderson, Searle, 1981 ; Silverstein, Bai, 1995, Lemma 2.2)). — (Henderson, Searle, 1981) Soit \mathbf{A} une matrice $M \times M$ inversible et $\mathbf{x} \in \mathbb{C}^M$, $c \in \mathbb{C}$ tel que $\mathbf{A} + c\mathbf{x}\mathbf{x}^H$ est inversible. Alors, en utilisant (3), nous avons

$$\mathbf{x}^H (\mathbf{A} + c\mathbf{x}\mathbf{x}^H)^{-1} = \frac{\mathbf{x}^H \mathbf{A}^{-1}}{1 + c\mathbf{x}^H \mathbf{A}^{-1} \mathbf{x}} \quad (8)$$

et

$$(\mathbf{A} + c\mathbf{x}\mathbf{x}^H)^{-1} \mathbf{x} = \frac{\mathbf{A}^{-1} \mathbf{x}}{1 + c\mathbf{x}^H \mathbf{A}^{-1} \mathbf{x}}. \quad (9)$$

LEMME 21 (Lemme II d'inversion de Matrice). — En utilisant les mêmes définitions que le Lemme 20 et en combinant ce Lemme avec (5), nous avons la relation suivante:

$$(\mathbf{A} + c\mathbf{x}\mathbf{x}^H)^{-1} = \mathbf{A}^{-1} - \frac{c\mathbf{A}^{-1} \mathbf{x}\mathbf{x}^H \mathbf{A}^{-1}}{1 + c\mathbf{x}^H \mathbf{A}^{-1} \mathbf{x}}. \quad (10)$$

Le Lemme suivant de perturbation de rang 1 est très utile dans le cas où nous devons utiliser le lemme d'inversion pour retirer une dépendance statistique avant d'utiliser le Lemme de la Trace. Il montre en particulier que l'Equivalent Déterministe est valide pour la matrice complète (voir par exemple (15)).

LEMME 22 (Lemme de Perturbation de Rang 1 (Bai, Silverstein, 2007, Lemma 2.1)). — Soit $z \in \mathbb{C} \setminus \mathbb{R}^+$, $\mathbf{A} \in \mathbb{C}^{M \times M}$, $\mathbf{B} \in \mathbb{C}^{M \times M}$ avec \mathbf{B} Hermitienne définie non-négative et $\mathbf{x} \in \mathbb{C}^M$. Alors

$$\left| \text{tr} [\mathbf{A} ((\mathbf{B} - z\mathbf{I}_M)^{-1} - (\mathbf{B} + \mathbf{x}\mathbf{x}^H - z\mathbf{I}_M)^{-1})] \right| \leq \frac{\|\mathbf{A}\|_2}{\text{dist}(z, \mathbb{R}^+)}$$

où $\text{dist}()$ est la distance Euclidienne. Si $z \in \mathbb{R}^-$ et $\limsup_M \|\mathbf{A}\|_2 < \infty$, alors nous avons:

$$\frac{1}{M} \left| \text{tr} [\mathbf{A} ((\mathbf{B} - z\mathbf{I}_M)^{-1} - (\mathbf{B} + \mathbf{x}\mathbf{x}^H - z\mathbf{I}_M)^{-1})] \right| \leq \frac{1}{M} \frac{\|\mathbf{A}\|_2}{|z|} \xrightarrow[M \rightarrow \infty]{} 0.$$

Nous remarquerons que la variable z correspond à l'inverse du rapport signal sur bruit (Signal to Noise Ratio, SNR) dans les problèmes de communications qui utilisent le précodage linéaire régularisé. Ceci explique partiellement la détérioration des performances des approximations à fort SNR.

Dans (Couillet, Debbah, 2011, Lemma 14.3), nous pouvons également trouver une variante du Lemme 22 pour $z = 0$, sous l'hypothèse que la plus petite valeur propre de la matrice Hermitienne B est bornée loin de zéro pour tout M , $\liminf_{M \rightarrow \infty} \lambda_{\min}(B) > 0$:

$$\frac{1}{M} \operatorname{tr} \mathbf{A} \mathbf{B}^{-1} - \frac{1}{M} \operatorname{tr} \mathbf{A} (\mathbf{B} + \mathbf{x} \mathbf{x}^H)^{-1} \xrightarrow[M \rightarrow \infty]{a.s.} 0.$$

5. Application de la Théorie des Matrices Aléatoires aux Télécommunications

Dans cette partie, nous motivons l'usage de la Théorie des Matrices Aléatoires et discutons de l'applicabilité des résultats pour l'analyse des systèmes de télécommunication avancés.

5.1. Les Avantages de l'Analyse à Large Dimension

Avant d'aller plus loin, penchons nous sur les raisons qui motivent l'utilisation de l'analyse à large dimension. Les Systèmes sans fils deviennent de plus en plus compliqués et nous avons besoin d'outils pour simplifier leur analyse. L'approche standard aujourd'hui est d'utiliser des simulations Monte-Carlo (MC) qui ne permettent malheureusement pas de comprendre les paramètres qui influent les performances du système. L'utilisation des équivalents déterministes offre certains avantages par rapport à une approche MC. Tout d'abord, comme les Equivalents Déterministes ne contiennent plus d'aléas, il est possible de simplifier l'analyse et la compréhension des différents paramètres en jeu. Cela est le cas quand les performances du système sont déterminées par différents paramètres agissant de manière non linéaire. La solution Equivalent Déterministe donne les interactions et relations directes entre les paramètres et les performances du système, ce qui est impossible avec l'analyse MC. De plus, la formulation analytique des Equivalents Déterministes permet une optimisation directe en utilisant des outils mathématiques classiques.

Une question pertinente est de savoir pourquoi l'analyse n'est pas conduite en dimension finie. Une réponse simple et rapide est que l'analyse est souvent compliquée pour donner des résultats utiles et solubles. Un exemple typique est le suivant⁵:

Définissons un système Multi-Utilisateur Multiple Input Multiple Output (MIMO) en voie montante, dans lequel la station de base a une unité centrale de traitement et M antennes distribuées. Nous considérons K utilisateurs mono-antennes qui transmettent en même temps et sur la même fréquence, en utilisant des signaux gaussiens

5. Cette exemple est tiré de (Hoydis, 2010).

$x_i \sim \mathcal{CN}(0, P)$ qui forment un vecteur de symboles $\mathbf{x} = [x_1, \dots, x_K]$. Nous supposons que $P = \mathcal{O}(1/K)$, de telle manière que la puissance reste bornée pour un nombre croissant de terminaux. Pour le modèle de canal, nous considérons un canal de Rayleigh $h_{i,j} \sim \mathcal{CN}(0, v_{i,j})$, $1 \leq i \leq M$, $1 \leq j \leq K$. En d'autres termes, la matrice de canal résultante \mathbf{H} a un profil de variance $\mathbf{V} = \{v_{i,j}\}$, $1 \leq i \leq M$, $1 \leq j \leq K$. En considérant un bruit Gaussien blanc au récepteur et sans traitement de signal au récepteur, nous avons la relation entrée-sortie suivante:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}.$$

La première question qui se pose concerne la capacité du canal d'un tel système. Nous savons à partir des travaux de Telatar (Telatar, 1999, Théorème 2) que dans le cas d'un canal Gaussien normal i.i.d. (à savoir, $v_{i,j} = 1 \forall i, j$), l'information mutuelle ergodique par antenne de réception est donnée par⁶

$$C_M^{iid} = E_{\mathbf{H}} \left[\frac{1}{M} \log \det (\mathbf{I}_M + P\mathbf{H}\mathbf{H}^H) \right] = \int_0^\infty \log(1 + P\lambda) f(\lambda) d\lambda$$

où $f(\lambda)$ est la fonction de densité de probabilité des valeurs propres non-ordonnées λ de la matrice de Wishart $\mathbf{H}\mathbf{H}^H$ et donnée par

$$f(\lambda) = \frac{M-K}{M} \delta(\lambda) + \frac{K}{M} \frac{1}{K} \sum_{k=0}^{K-1} \frac{k!}{(k+M-K)!} [L_k^{M-K}(\lambda)]^2 \lambda^{M-K} e^{-\lambda}.$$

Ici, $L_k^{M-K}(\lambda)$ est un polynôme de Laguerre associé d'ordre k :

$$L_k^N(\lambda) = \frac{\lambda^{-N} e^\lambda}{k!} \frac{d^k}{d\lambda^k} (e^{-\lambda} \lambda^{k+N}) = \sum_{l=0}^k (-1)^l \frac{(k+N)!}{(k-l)!(N+l)!} \lambda^l.$$

(Dohler, 2003, Eq. (2.38) et (2.45)) décrit une manière de calculer de manière explicite l'intégrale dans l'équation de la capacité. Dans le cas où $N = K = 2$, nous avons:

$$f(\lambda) = \frac{1}{2} \sum_{k=0}^1 [L_k^0(\lambda)]^2 e^{-\lambda} = \frac{1}{2} [1 + (1 - \lambda^2)] e^{-\lambda}.$$

En notant que $L_0^0 = 1$ et $L_1^0 = 1 - \lambda$, nous obtenons:

$$C_2^{iid} = \frac{1}{2} \int_0^\infty \log(1 + P\lambda) [1 + (1 - \lambda^2)] e^{-\lambda} d\lambda \quad (11)$$

$$= \frac{1}{2} - \frac{1}{2P} + \left(1 + \frac{1}{2P^2} e^{1/P} E_1(1/P) \right) \quad (12)$$

6. Dans le cas de canaux Gaussiens, la transmission d'un signal Gaussien de moyenne nulle et de covariance $\frac{PK}{K} \mathbf{I}_K$ maximise l'information mutuelle, ce qui est équivalent à la capacité (Telatar, 1999, Théorème 1).

où $E_1(z) = \int_1^\infty \frac{e^{-tz}}{t} dt$. En résumé, il est possible de calculer une forme explicite de l'information mutuelle ergodique pour des systèmes simples de canaux sans profils de variance. Cependant, les résultats en question n'offrent que très peu de conclusions. Ainsi, il est difficile de prédire l'influence de P dans (12)

De plus, l'approche à dimension finie ne marche plus lorsque nous sortons du cadre des hypothèses idéales. Le cas où les composantes de la matrice ne sont plus gaussiennes devient insoluble. Même si nous considérons un profil de variance simple comme

$$\mathbf{V} = \begin{pmatrix} 1 & \alpha \\ \alpha & 1 \end{pmatrix} \quad (13)$$

Trouver l'information mutuelle ergodique devient très difficile.

En d'autres termes, même pour les systèmes simples, l'approche à dimension finie n'est pas d'un grand intérêt. Par contre, en utilisant l'approche en grande dimension, nous pouvons aisément traiter le cas du profil de variance arbitraire. En utilisant Hachem et al. (Hachem *et al.*, 2007), nous avons le théorème suivant:

THÉORÈME 23 (Capacité avec un profil de variance (Hachem *et al.*, 2007, Théorème 4.1) (aussi (Hoydis *et al.*, 2011) et (Hachem, Khorunzhy *et al.*, 2008, Théorème 1))). — Soit $M, K \rightarrow \infty$ tel que $0 < \frac{K}{M} < \infty$ et $v_{i,j} < v_{max} < \infty, \forall i, j$. Alors pour le modèle considéré dans la section 5.1, nous avons $C_M - \bar{C}_M \xrightarrow{\text{a.s.}} 0$, où

$$\bar{C}_M = \frac{1}{M} \sum_{i=1}^K \log(1 + \delta_j) - \frac{1}{M} \sum_{i=1}^M \log\left(\frac{1}{PK} e_i\right) - \frac{1}{M} \sum_{j=1}^K \frac{\delta_j}{1 + \delta_j}$$

où $\delta_j = \frac{1}{K} \sum_{l=1}^M v_{l,j} e_l$ for $j = 1, \dots, K$ et e_i for $i = 1, \dots, M$ est donné comme l'unique solution positive aux systèmes de M équations implicites

$$e_i = \left(\frac{1}{PK} + \frac{1}{K} \sum_{j=1}^K \frac{v_{i,j}}{1 + \frac{1}{K} \sum_{l=1}^M v_{l,j} e_l} \right)^{-1}.$$

Bien qu'il apparaisse obscur à première vue, le Théorème 23 offre un certain nombre d'atouts. Il permet en autres une optimisation très simple. Dans tous les cas, des Equivalents Déterministes de ce type sont les seules formulations déterministes connues pour la capacité du canal pour un profil de variance général.

5.2. Précisions des Résultats

La plupart des publications autour des techniques des grands systèmes utilisent des approches très pragmatiques en ce qui concerne la robustesse et la précision des

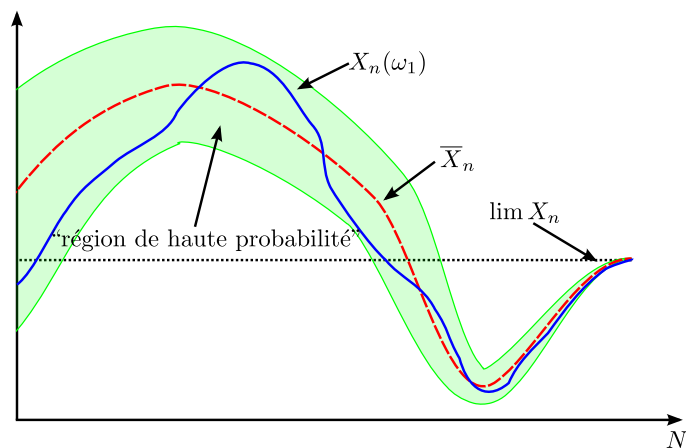


FIGURE 1. Comparaison qualitative de l'Equivalent Déterministe avec le calcul des limites classiques et du cas d'une seule réalisation.

résultats pour des systèmes de dimensions finies. Les publications illustrent les théorèmes par quelques simulations qui comparent les solutions analytiques à des courbes obtenues par des simulations Monte Carlo. Dans la figure 1, nous montrons les implications des deux approches. La figure montre une réalisation d'une séquence aléatoire de variable $X_n(\omega_1)$, qui représente un indicateur de performance système (aléatoire par exemple par rapport aux réalisations du canal) qui dépend de la dimension du système N . En prenant la limite par rapport à la dimension du système, nous ne pouvons qu'obtenir $\lim_{n \rightarrow \infty} X_n$. A l'inverse, l'approche par Equivalent Déterministe (\bar{X}_n) nous fournit plus d'information. Intuitivement, on peut remarquer que \bar{X}_n est encore indexée par n , même quand $n \rightarrow \infty$. En fait, l'Equivalent Déterministe fournit une approximation pour chaque valeur de N , qui devient plus précise pour des N croissants. Les réalisations de la variable aléatoire sont presque sûrement autour d'une région étroite autour de l'Equivalent Déterministe (voir "région de haute probabilité" dans la figure). De plus, l'approche par Equivalent Déterministe permet des approximations de séquences aléatoires qui peuvent ne pas du tout converger (contrairement au cas illustré dans la Figure 1), ce est impossible en utilisant les limites classiques. Ainsi, les Equivalents Déterministes sont en général plus précis pour des systèmes à dimension finie (même petit) que les limites classiques. Lorsque l'on considère les statistiques de premier ordre, les résultats de Monte Carlo et ceux des Equivalents Déterministes sont en accord pour n de l'ordre de 10.

Il faut également signaler un autre point important. Supposons que nous analysons une quantité aléatoire telle qu'une matrice $\mathbf{X} \in \mathbb{C}^{M \times K}$. La théorie des matrices aléa-

toires se projette dans le cas où (M, K) sont grands, alors que les approximations de limites classiques (la loi des grands nombres par exemple) ne peut traiter que le cas où M est grand. Par conséquent, la théorie des Matrices Aléatoires exploite plus de degrés de liberté que les approches classiques et assure par conséquent une convergence plus rapide. Ainsi, une matrice de taille 8×8 offre déjà 64 degrés de liberté, ce qui aboutit à une convergence acceptable.

En général, la théorie des matrices aléatoires assure une vitesse de convergence assez impressionnante en ce qui concerne des fonctionnelles linéaires de valeurs propres. C'est le cas par exemple pour des théorèmes centraux limites avec une vitesse en $1/M$ ($M(X_M - \bar{X}_M) \rightarrow \mathcal{N}(0, 1)$) (Hachem, Loubaton, Najim, 2008, Remarque 3.1)(Pastur, Shcherbina, 2011) et la moyenne en $1/M^2$ (à savoir, $\mathbb{E}[X_M] = \bar{X}_M + \mathcal{O}(1/M^2)$) (Hachem, Khorunzhy *et al.*, 2008, Preuve de Théorème 1)(Pastur, Shcherbina, 2011) (Vallet *et al.*, 2010, Proposition 4), lorsque les quantités aléatoires sont des Gaussiennes complexes. Les formes quadratiques sont plus lentes: le cas des théorèmes centraux limites sont en $1/\sqrt{M}$ (Kammoun *et al.*, 2009, Théorème 3)(Rubio *et al.*, 2012, Théorèmes 1&2) et la moyenne en $1/M^{\frac{3}{2}}$ (Rubio *et al.*, 2012, Lemme 1).

5.3. Equivalents Déterministes et Problèmes de Télécommunications

Nous avons déjà montré la connexion entre la Transformée de Stieltjes et les traces de résolvante dans (2). Maintenant, nous voulons donner un rapide mais détaillé aperçu de l'importance de la trace de la résolvante, en particulier pour les questions relatives au SINR et comment cela est lié aux théorèmes des Equivalents Déterministes. L'exemple suivant est basé sur (Couillet, Debbah, 2009) et (Couillet, Debbah, 2011), où un système CDMA avec un détecteur MMSE est traité.

En considérant un système en liaison montante MIMO Multi-utilisateur avec K utilisateurs mono-antennes transmettant vers une station de base unique équipée de M antennes et utilisant un récepteur de type MMSE. L'évanouissement entre le canal de l'utilisateur k et la station de base est modélisé selon la distribution $\mathbf{h}_k \sim \mathcal{CN}(0, \frac{1}{M}\mathbf{I}_M) \in \mathbb{C}^M$ avec $\mathbb{E}\|\mathbf{h}_k\|_2 = 1$. Nous incluons l'impact des effets du masquage et perte de trajet dans une variable d_k .

Sous l'hypothèse d'un signal transmis Gaussien $s_k \sim \mathcal{CN}(0, 1)$ et en considérant le bruit en réception \mathbf{n} comme un bruit additif Gaussien de moyenne zéro et de variance σ^2 , nous aboutissons au modèle de transmission suivant:

$$\mathbf{y} = \sum_{k=1}^K d_k \mathbf{h}_k s_k + \mathbf{n} = \mathbf{H}\mathbf{D}\mathbf{s} + \mathbf{n}$$

où $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K] \in \mathbb{C}^{M \times K}$, $\mathbf{s} = [s_1, \dots, s_K] \in \mathbb{C}^K$ et $\mathbf{D} = \text{diag}(d_1, \dots, d_K) \in \mathbb{C}^{K \times K}$.

Le récepteur linéaire pour chaque utilisateur est donné par:

$$\mathbf{r}_k^H = \mathbf{h}_k^H (\mathbf{H}\mathbf{D}^2\mathbf{H}^H + \sigma^2\mathbf{I}_M)^{-1} .$$

Ainsi, le rapport signal à interférence plus bruit pour l'utilisateur k est donné par:

$$\begin{aligned} \text{SINR}_k &= \frac{\mathbb{E}_{\mathbf{s}} |\mathbf{r}_k^H d_k \mathbf{h}_k s_k|^2}{\mathbb{E}_{\mathbf{s}, \mathbf{n}} |\sum_{j \neq k} \mathbf{r}_k^H d_j \mathbf{h}_j s_j + \mathbf{r}_k^H \mathbf{n}|^2} \\ &= \frac{|d_k|^2 \mathbf{r}_k^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{r}_k}{\sum_{j \neq k} |d_j|^2 \mathbf{r}_k^H \mathbf{h}_j \mathbf{h}_j^H \mathbf{r}_k + \mathbf{r}_k^H \sigma^2 \mathbf{r}_k} \\ &= \frac{|d_k|^2 \mathbf{r}_k^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{r}_k}{\mathbf{r}_k^H (\mathbf{H}\mathbf{D}^2\mathbf{H}^H - |d_k|^2 \mathbf{h}_k \mathbf{h}_k^H) \mathbf{r}_k + \mathbf{r}_k^H \sigma^2 \mathbf{r}_k} \\ &= \frac{|d_k|^2 \mathbf{r}_k^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{r}_k}{\mathbf{r}_k^H (\mathbf{H}\mathbf{D}^2\mathbf{H}^H + \sigma^2 \mathbf{I}_M - \sigma^2 \mathbf{I}_M - |d_k|^2 \mathbf{h}_k \mathbf{h}_k^H) \mathbf{r}_k + \mathbf{r}_k^H \sigma^2 \mathbf{r}_k} \\ &= \frac{|d_k|^2 \mathbf{r}_k^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{r}_k}{\mathbf{r}_k^H (\mathbf{H}\mathbf{D}^2\mathbf{H}^H + \sigma^2 \mathbf{I}_M) \mathbf{r}_k - \mathbf{r}_k^H \sigma^2 \mathbf{r}_k - |d_k|^2 \mathbf{r}_k^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{r}_k + \mathbf{r}_k^H \sigma^2 \mathbf{r}_k} . \end{aligned}$$

En prenant en compte l'ensemble des termes qui s'annulent, ainsi que ceux présent dans la définition de \mathbf{r}_k^H , nous avons:

$$\text{SINR}_k = \frac{|d_k|^2 \mathbf{r}_k^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{r}_k}{\mathbf{x}^H \mathbf{r}_k - |d_k|^2 \mathbf{r}_k^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{r}_k} = \frac{|d_k|^2 \mathbf{r}_k^H \mathbf{h}_k}{1 - |d_k|^2 \mathbf{r}_k^H \mathbf{h}_k}$$

Finalement, en réintroduisant la définition de \mathbf{r}_k^H partout et en appliquant le Lemme 20, nous avons

$$\text{SINR}_k = |d_k|^2 \mathbf{h}^H (\mathbf{H}\mathbf{D}^2\mathbf{H}^H - |d_k|^2 \mathbf{h}_k \mathbf{h}_k^H + \sigma^2 \mathbf{I}_M)^{-1} \mathbf{h}_k . \quad (14)$$

Jusqu'à maintenant, le calcul n'a utilisé aucun concept venant de la théorie des Matrices Aléatoires. Nous pouvons simplifier l'expression du SINR dans (14) dans l'hypothèse où $M \rightarrow \infty$, avec $0 < M/K = c < \infty$. Nous commençons par rappeler le Lemme 17 et le Lemme 22 afin d'obtenir:

$$\text{SINR}_k - \frac{1}{M} |d_k|^2 \text{tr} (\mathbf{H}\mathbf{D}^2\mathbf{H}^H + \sigma^2 \mathbf{I}_M)^{-1} \xrightarrow[M \rightarrow +\infty]{a.s.} 0 . \quad (15)$$

Rappelons la définition de la transformée de Stieltjes (Définition 3) ainsi que la mesure de comptage normalisée des valeurs propres de la matrice $\mathbf{H}\mathbf{D}^2\mathbf{H}^H$, $m_{\mathbf{H}\mathbf{D}^2\mathbf{H}^H}(z) = \text{tr} (\mathbf{H}\mathbf{D}^2\mathbf{H}^H - z\mathbf{I}_M)^{-1}$ (voir la définition 5). Ceci nous permet d'écrire (15) comme

$$\text{SINR}_k - |h_k|^2 m_{\mathbf{X}\mathbf{D}^2\mathbf{X}^H}(-\sigma^2) \xrightarrow[M \rightarrow +\infty]{a.s.} 0 .$$

N'oublions pas que la Transformée de Stieltjes $m_{\mathbf{X}\mathbf{D}^2\mathbf{X}^H}(-\sigma^2)$ représente toujours une quantité aléatoire.

Il est possible de trouver un Equivalent Déterministe de cette expression, que nous démontrerons dans la sous-section 5.4, et aboutit au théorème 24.

Une application directe nous donne

$$m_{\mathbf{HD}^2\mathbf{H}^{\mathbf{H}}}(-\sigma^2) - \overline{m}_{\mathbf{HD}^2\mathbf{H}^{\mathbf{H}}}(-\sigma^2) \xrightarrow[M \rightarrow +\infty]{a.s.} 0.$$

avec

$$\overline{m}_{\mathbf{HD}^2\mathbf{H}^{\mathbf{H}}}(-\sigma^2) = \left(\sigma^2 + c \sum_{i=1}^K \frac{|d_i|^2}{1 + |d_i|^2 \overline{m}_{\mathbf{HD}^2\mathbf{H}^{\mathbf{H}}}(-\sigma^2)} \right)^{-1}$$

Cette formulation est déterministe par rapport aux composantes de \mathbf{H} , mais conditionnellement aux composantes de \mathbf{D} , qui représentent les coefficients de masquage et de perte de trajet d_k .

5.4. Calcul d'Equivalents Déterministes

Nous continuons cette partie en appliquant les outils et concepts introduits pour le calcul d'Equivalents Déterministes. Nous essayons d'expliquer chaque étape du calcul, même ceux qui paraissent évidents. Cet exemple illustre également la méthodologie à l'origine de la détermination de l'Equivalent Déterministe⁷. Cependant, le but principal ici est de fournir une application intéressante des outils et Lemmes introduits précédemment. Cet exemple fournit une explication intuitive du calcul de l'Equivalent Déterministe. Les calculs suivants s'inspirent de (Couillet *et al.*, 2011). Les aspects trop techniques ou qui n'aident pas à la compréhension ne sont pas considérés. Pour les détails techniques, nous invitons le lecteur à consulter (Couillet *et al.*, 2011), ou la version réduite (Couillet, 2010).

THÉORÈME 24. — Soient $\mathbf{T} \in \mathbb{C}^{M \times K}$ une matrice diagonale définie non-négative et $\mathbf{R} \in \mathbb{C}^{M \times M}$ une matrice définie non-négative, ayant une norme spectrale bornée, $\limsup_{K \rightarrow \infty} \|\mathbf{T}\| = \limsup_{K \rightarrow \infty} \lambda_{\max}(\mathbf{T}) < \infty$ et $\limsup_{K \rightarrow \infty} \|\mathbf{R}\| < \infty$. Soit $\mathbf{X} \in \mathbb{C}^{M \times K}$ une matrice dont les éléments sont distribués selon $\mathcal{CN}(0, \frac{1}{K})$. Soit également $\mathbf{B} = \mathbf{R}^{\frac{1}{2}} \mathbf{X} \mathbf{T} \mathbf{X}^{\mathbf{H}} \mathbf{R}^{\frac{1}{2}}$.

Soit $M, K \rightarrow \infty$, tel que $M/K \rightarrow c$, où c est une constante bornée, $0 < c < \infty$. Dans ce cas, le résultat suivant est obtenu:

$$\frac{1}{M} \text{tr} \left[(\mathbf{B} - z \mathbf{I}_M)^{-1} \right] - \overline{m}_M(z) \xrightarrow[M, K \rightarrow +\infty]{a.s.} 0$$

7. La méthode utilisée ci-après est connue sous le nom de "l'approche Bai-Silverstein", selon les étapes décrites dans (Silverstein, Bai, 1995). Il existe d'autres techniques de preuves, en particulier "l'approche Pastur", qui utilise des méthodes Gaussiennes (Pastur, 1999 ; Dupuy, Loubaton, 2010) et qui est considérée plus puissante mais moins évidente.

où $z \in \mathbb{C} \setminus \mathbb{R}^+$ et \bar{m}_M sont donnés par

$$\bar{m}_M = \frac{1}{M} \text{tr} (\mathbf{R} e(z) - z \mathbf{I}_M)^{-1}$$

qui inclut de trouver l'unique solution positive de l'équation au point fixe.

$$e(z) = \frac{1}{K} \sum_{i=1}^K \frac{t_i}{1 + t_i c \frac{1}{M} \text{tr} \mathbf{R} (\mathbf{R} e(z) - z \mathbf{I}_M)^{-1}}.$$

Il n'est pas évident d'obtenir un tel théorème. En suivant l'approche de Bai-Silverstein, nous commençons par faire une supposition sur la forme générale du résultat (voir la remarque 25 pour motiver ce choix):

$$\begin{aligned} & \frac{1}{M} \text{tr} (\mathbf{B} - z \mathbf{I}_M)^{-1} - \text{Equivalent Déterministe} \xrightarrow[M, K \rightarrow +\infty]{a.s.} 0 \\ \xrightarrow[\text{éclairée}]{\text{supposition}} & \frac{1}{M} \text{tr} (\mathbf{B} - z \mathbf{I}_M)^{-1} - \frac{1}{M} \text{tr} (\mathbf{R} e(z) - z \mathbf{I}_M)^{-1} \xrightarrow[M, K \rightarrow +\infty]{a.s.} 0. \end{aligned}$$

Le but principal est de trouver maintenant une formulation pour $e(z)$, qui ne dépend pas des quantités aléatoires et qui assure une convergence presque sûre.

En utilisant l'identité de la résolvante (5), nous obtenons rapidement:

$$\begin{aligned} & \frac{1}{M} \text{tr} (\mathbf{B} - z \mathbf{I}_M)^{-1} - \frac{1}{M} \text{tr} (\mathbf{R} e(z) - z \mathbf{I}_M)^{-1} \\ & \stackrel{(5)}{=} \frac{1}{M} \text{tr} \left[(\mathbf{B} - z \mathbf{I}_M)^{-1} (e(z) \mathbf{R} - \mathbf{B} - z \mathbf{I}_M + z \mathbf{I}_M) (\mathbf{R} e(z) - z \mathbf{I}_M)^{-1} \right] \\ & = \frac{1}{M} \text{tr} \left[(\mathbf{B} - z \mathbf{I}_M)^{-1} \left(e(z) \mathbf{R} - \mathbf{R}^{\frac{1}{2}} \mathbf{X} \mathbf{T} \mathbf{X}^H \mathbf{R}^{\frac{1}{2}} \right) (\mathbf{R} e(z) - z \mathbf{I}_M)^{-1} \right] \\ & = \frac{1}{M} \text{tr} \left[(\mathbf{B} - z \mathbf{I}_M)^{-1} \mathbf{R}^{\frac{1}{2}} (e(z) - \mathbf{X} \mathbf{T} \mathbf{X}^H) \mathbf{R}^{\frac{1}{2}} (\mathbf{R} e(z) - z \mathbf{I}_M)^{-1} \right] \\ & = \frac{1}{M} \text{tr} \left[(\mathbf{B} - z \mathbf{I}_M)^{-1} \mathbf{R} (\mathbf{R} e(z) - z \mathbf{I}_M)^{-1} \right] e(z) \\ & \quad - \frac{1}{M} \text{tr} \left[(\mathbf{B} - z \mathbf{I}_M)^{-1} \mathbf{R}^{\frac{1}{2}} \mathbf{X} \mathbf{T} \mathbf{X}^H \mathbf{R}^{\frac{1}{2}} (\mathbf{R} e(z) - z \mathbf{I}_M)^{-1} \right]. \end{aligned}$$

En se rappelant que $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_K]$ and $\mathbf{T} = \text{diag}(t_1, \dots, t_K)$, nous avons $\mathbf{X} \mathbf{T} \mathbf{X}^H = \sum_{i=1}^K t_i \mathbf{x}_i \mathbf{x}_i^H$. Ainsi, nous pouvons faire sortir cette somme.

$$\begin{aligned} & = \frac{1}{M} \text{tr} \left[(\mathbf{B} - z \mathbf{I}_M)^{-1} \mathbf{R} (\mathbf{R} e(z) - z \mathbf{I}_M)^{-1} \right] e(z) \\ & \quad - \frac{1}{M} \sum_{i=1}^K t_i \text{tr} \left[\underbrace{(\mathbf{B} - z \mathbf{I}_M)^{-1} \mathbf{R}^{\frac{1}{2}} \mathbf{x}_i}_{\bar{\mathbf{x}}_i} \underbrace{\mathbf{x}_i^H \mathbf{R}^{\frac{1}{2}} (\mathbf{R} e(z) - z \mathbf{I}_M)^{-1}}_{\bar{\mathbf{x}}_i^H} \right]. \end{aligned}$$

Comme l'argument de la trace est un scalaire, il est possible d'enlever l'opérateur et nous obtenons:

$$\begin{aligned} &= \frac{1}{M} \text{tr} \left[(\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R} (\mathbf{R}e(z) - z\mathbf{I}_M)^{-1} \right] e(z) \\ &\quad - \frac{1}{M} \sum_{i=1}^K t_i \mathbf{x}_i^{\text{H}} \underbrace{\mathbf{R}^{\frac{1}{2}} (\mathbf{R}e(z) - z\mathbf{I}_M)^{-1} (\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R}^{\frac{1}{2}} \mathbf{x}_i}_{\triangleq \tilde{\mathbf{A}}} . \end{aligned}$$

A ce niveau, nous serions tentés d'appliquer le Lemme de la Trace (Lemme 17) directement à la forme $\mathbf{x}_i^{\text{H}} \tilde{\mathbf{A}} \mathbf{x}_i$ mais il faut d'abord vérifier les pré-requis. En particulier, nous devons être sûrs que \mathbf{x}_i est statistiquement indépendant de $\tilde{\mathbf{A}}$, ce qui est possible seulement si \mathbf{x}_i est statistiquement indépendant de \mathbf{B} . Ceci n'est clairement pas le cas, comme le montre le résultat suivant:

$$\mathbf{B} = \mathbf{R}^{\frac{1}{2}} \mathbf{X} \mathbf{T} \mathbf{X}^{\text{H}} \mathbf{R}^{\frac{1}{2}} = \sum_{j=1}^K t_j \mathbf{R}^{\frac{1}{2}} \mathbf{x}_j \mathbf{x}_j^{\text{H}} \mathbf{R}^{\frac{1}{2}} = \sum_{j \neq i}^K t_j \mathbf{R}^{\frac{1}{2}} \mathbf{x}_j \mathbf{x}_j^{\text{H}} \mathbf{R}^{\frac{1}{2}} + t_i \mathbf{R}^{\frac{1}{2}} \mathbf{x}_i \mathbf{x}_i^{\text{H}} \mathbf{R}^{\frac{1}{2}} .$$

Ainsi, nous devons d'abord appliquer le Lemme 20 qui permet d'éliminer la partie dépendante. Comme précédemment, il est possible de découper l'équation en:

$$(\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R}^{\frac{1}{2}} \mathbf{x}_i = \underbrace{\left(\sum_{j \neq i}^K t_j \mathbf{R}^{\frac{1}{2}} \mathbf{x}_j \mathbf{x}_j^{\text{H}} \mathbf{R}^{\frac{1}{2}} - z\mathbf{I}_M + \underbrace{t_i}_{\tilde{c}} \underbrace{\mathbf{R}^{\frac{1}{2}} \mathbf{x}_i}_{\tilde{\mathbf{x}}_i} \underbrace{\mathbf{x}_i^{\text{H}} \mathbf{R}^{\frac{1}{2}}}_{\tilde{\mathbf{x}}_i^{\text{H}}} \right)^{-1}}_{\tilde{\mathbf{A}}} \underbrace{\mathbf{R}^{\frac{1}{2}} \mathbf{x}_i}_{\tilde{\mathbf{x}}_i}$$

et appliquer le Lemme d'inversion matricielle, ce qui donne:

$$\begin{aligned} &\frac{1}{M} \text{tr} (\mathbf{B} - z\mathbf{I}_M)^{-1} - \frac{1}{M} \text{tr} (\mathbf{R}e(z) - z\mathbf{I}_M)^{-1} \\ &= \frac{1}{M} \text{tr} \left[(\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R} (\mathbf{R}e(z) - z\mathbf{I}_M)^{-1} \right] e(z) \\ &\quad - \frac{1}{M} \sum_{i=1}^K t_i \frac{\overbrace{\mathbf{x}_i^{\text{H}} \mathbf{R}^{\frac{1}{2}} (\mathbf{R}e(z) - z\mathbf{I}_M)^{-1} \left(\sum_{j \neq i}^K t_j \mathbf{R}^{\frac{1}{2}} \mathbf{x}_j \mathbf{x}_j^{\text{H}} \mathbf{R}^{\frac{1}{2}} - z\mathbf{I}_M \right)^{-1} \mathbf{R}^{\frac{1}{2}} \mathbf{x}_i}_{\tilde{\mathbf{A}}}}{1 + t_i \mathbf{x}_i^{\text{H}} \mathbf{R}^{\frac{1}{2}} \left(\sum_{j \neq i}^K t_j \mathbf{R}^{\frac{1}{2}} \mathbf{x}_j \mathbf{x}_j^{\text{H}} \mathbf{R}^{\frac{1}{2}} - z\mathbf{I}_M \right)^{-1} \mathbf{R}^{\frac{1}{2}} \mathbf{x}_i} . \end{aligned}$$

Maintenant, nous voyons que $\tilde{\mathbf{A}}$ est statistiquement indépendant de \mathbf{x}_i et nous pouvons appliquer le Lemme de la Trace (Lemme 17) au numérateur et au dénominateur. Ce qui nous donne la convergence $\mathbf{x}_i^{\text{H}} \tilde{\mathbf{A}} \mathbf{x}_i - \frac{1}{K} \text{tr}(\tilde{\mathbf{A}}) \xrightarrow[K \rightarrow +\infty]{a.s.} 0$. Nous pouvons remarquer que les étapes suivantes ne sont valides que dans le sens presque sûr et seulement pour le cas du régime matriciel large. Dans la partie suivante, nous

utilisons abusivement la notation “ \approx ” pour exprimer cette restriction quand cela est nécessaire.

$$\begin{aligned} & \frac{1}{M} \operatorname{tr} (\mathbf{B} - z\mathbf{I}_M)^{-1} - \frac{1}{M} \operatorname{tr} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1} \\ & \approx \frac{1}{M} \operatorname{tr} \left[(\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1} \right] e(z) \\ & \quad - \frac{1}{M} \sum_{i=1}^K t_i \frac{\frac{1}{K} \operatorname{tr} \mathbf{R}^{\frac{1}{2}} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1} \left(\sum_{j \neq i}^K t_j \mathbf{R}^{\frac{1}{2}} \mathbf{x}_j \mathbf{x}_j^H \mathbf{R}^{\frac{1}{2}} - z\mathbf{I}_M \right)^{-1} \mathbf{R}^{\frac{1}{2}}}{1 + t_i \frac{1}{K} \operatorname{tr} \mathbf{R}^{\frac{1}{2}} \left(\sum_{j \neq i}^K t_j \mathbf{R}^{\frac{1}{2}} \mathbf{x}_j \mathbf{x}_j^H \mathbf{R}^{\frac{1}{2}} - z\mathbf{I}_M \right)^{-1} \mathbf{R}^{\frac{1}{2}}}. \end{aligned}$$

En utilisant le Lemme de perturbation de rang 1 (Lemme 22), nous savons que:

$$\operatorname{tr} \mathbf{R}^{\frac{1}{2}} \left(\sum_{j \neq i}^K t_j \mathbf{R}^{\frac{1}{2}} \mathbf{x}_j \mathbf{x}_j^H \mathbf{R}^{\frac{1}{2}} - z\mathbf{I}_M \right)^{-1} \mathbf{R}^{\frac{1}{2}}$$

converge presque sûrement vers

$$\operatorname{tr} \mathbf{R}^{\frac{1}{2}} \left(\sum_{j=1}^K t_j \mathbf{R}^{\frac{1}{2}} \mathbf{x}_j \mathbf{x}_j^H \mathbf{R}^{\frac{1}{2}} - z\mathbf{I}_M \right)^{-1} \mathbf{R}^{\frac{1}{2}} = \operatorname{tr} \mathbf{R}^{\frac{1}{2}} (\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R}^{\frac{1}{2}}.$$

Par conséquent, il est possible d’écrire:

$$\begin{aligned} & \frac{1}{M} \operatorname{tr} (\mathbf{B} - z\mathbf{I}_M)^{-1} - \frac{1}{M} \operatorname{tr} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1} \\ & \approx \frac{1}{M} \operatorname{tr} \left[(\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1} \right] e(z) \\ & \quad - \frac{1}{M} \sum_{i=1}^K t_i \frac{\frac{1}{K} \operatorname{tr} \mathbf{R}^{\frac{1}{2}} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1} (\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R}^{\frac{1}{2}}}{1 + t_i \frac{1}{K} \operatorname{tr} \mathbf{R}^{\frac{1}{2}} (\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R}^{\frac{1}{2}}} \\ & = \frac{1}{M} \operatorname{tr} \left[(\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1} \right] e(z) \\ & \quad - \frac{1}{K} \sum_{i=1}^K t_i \frac{\frac{1}{M} \operatorname{tr} (\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1}}{1 + t_i \frac{M}{K} \frac{1}{M} \operatorname{tr} \mathbf{R}^{\frac{1}{2}} (\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R}^{\frac{1}{2}}} \end{aligned}$$

REMARQUE 25 (Conjecture). — Ce n’est qu’à ce stade que nous pouvons voir que notre conjecture est judicieuse. Ce choix résulte de la forme suivante qui montre que $\operatorname{tr} [\mathbf{A}_j] e(z) - \operatorname{tr} [\mathbf{A}_j] x(z)$, où $x(z)$ est un candidat potentiel de notre Equivalent Déterministe. Trouver le bon Equivalent Déterministe résulte de plusieurs essais et erreurs. \square

En compilant tous les termes, nous avons finalement:

$$\begin{aligned} & \frac{1}{M} \operatorname{tr} (\mathbf{B} - z\mathbf{I}_M)^{-1} - \frac{1}{M} \operatorname{tr} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1} \\ &= \overbrace{\frac{1}{M} \operatorname{tr} \left[(\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1} \right]}^{\text{Borné puisque } \mathbf{R} \text{ est bornée.}} \times \\ & \quad \underbrace{\left[e(z) - \frac{1}{K} \sum_{i=1}^K \frac{t_i}{1 + t_i c \frac{1}{M} \operatorname{tr} \mathbf{R} (\mathbf{B} - z\mathbf{I}_M)^{-1}} \right]}_{\xrightarrow{!} 0}. \end{aligned}$$

Ainsi, nous pouvons voir qu'en choisissant $e(z)$ de telle manière que le terme multiplicatif à droite devient "0", le résultat escompté peut être obtenu. Cependant, un tel résultat contiendrait encore de l'aléatoire. De plus, l'expression dans le dénominateur $\frac{1}{M} \operatorname{tr} \mathbf{R} (\mathbf{B} - z\mathbf{I}_M)^{-1}$ diffère du résultat escompté. Si nous avions $\frac{1}{M} \operatorname{tr} (\mathbf{B} - z\mathbf{I}_M)^{-1}$, nous pourrions trouver une expression déterministe pour notre problème original. Au contraire, nous avons créé un nouveau terme qui doit être évalué. Cela va être fait dans le paragraphe suivant.

Pour résoudre ce problème, nous devons recommencer depuis le début. Cette fois-ci, nous débutons par le problème complémentaire $\frac{1}{M} \operatorname{tr} \mathbf{R} (\mathbf{B} - z\mathbf{I}_M)^{-1}$ et "conjecturons" la solution complémentaire $\frac{1}{M} \operatorname{tr} \mathbf{R} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1}$. Ceci nous donne une solution complémentaire qui, comme nous le verrons, combinée au premier résultat nous permettra d'avoir une solution analytique.

En reprenant les mêmes étapes que précédemment, nous avons:

$$\begin{aligned} & \frac{1}{M} \operatorname{tr} \mathbf{R} (\mathbf{B} - z\mathbf{I}_M)^{-1} - \frac{1}{M} \operatorname{tr} \mathbf{R} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1} \\ & \quad \vdots \\ & \approx \frac{1}{M} \operatorname{tr} \left[\mathbf{R} (\mathbf{B} - z\mathbf{I}_M)^{-1} \mathbf{R} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1} \right] \\ & \quad \left[e(z) - \frac{1}{K} \sum_{i=1}^K \frac{t_i}{1 + t_i c \frac{1}{M} \operatorname{tr} \mathbf{R} (\mathbf{B} - z\mathbf{I}_M)^{-1}} \right]. \end{aligned}$$

Maintenant, nous choisissons finalement:

$$e(z) = \frac{1}{K} \sum_{i=1}^K \frac{t_i}{1 + t_i c \frac{1}{M} \operatorname{tr} \mathbf{R} (\mathbf{B} - z\mathbf{I}_M)^{-1}}$$

ce qui résout simultanément ce cas et notre conjecture précédente.

A partir de la seconde "conjecture", nous observons également que (pour le choix particulier de $e(z)$), nous avons $\frac{1}{M} \operatorname{tr} \mathbf{R} (\mathbf{B} - z\mathbf{I}_M)^{-1} - \frac{1}{M} \operatorname{tr} \mathbf{R} (\mathbf{R} e(z) - z\mathbf{I}_M)^{-1}$

$\xrightarrow{\text{a.s.}} 0$. A partir de cette observation, un autre choix $e(z)$, qui sera finalement déterministe est:

$$e(z) = \frac{1}{K} \sum_{i=1}^K \frac{t_i}{1 + t_i c \frac{1}{M} \text{tr} \mathbf{R} (\mathbf{R} e(z) - z \mathbf{I}_M)^{-1}}$$

qui est une équation du point fixe, qui peut se résoudre itérativement⁸. Maintenant, nous devons montrer que l'équation du point fixe a une solution unique et que l'algorithme du point fixe converge. Cela peut être fait aisément en utilisant le *cadre des fonctions d'interférence standard* de (Yates, 1995), comme cela a été montré ici (Hoydis, 2012, Théorèmes 22, 23, 24).

6. Conclusion

Dans cet article, nous avons fourni une introduction aux applications pratiques de la théorie des matrices aléatoires (en particulier les Equivalents Déterministes) pour l'analyse des réseaux de télécommunications. Dans cette optique, nous avons d'abord établi les concepts théoriques de base, les lemmes et les outils dans ce domaine. Ces concepts ont ensuite été appliqués dans différents calculs, permettant d'illustrer la pertinence de l'approche pour des problèmes de communications simples puis des modèles plus élaborés. Bien que la théorie des matrices aléatoires trouve d'immenses applications, il est important de mentionner certaines limitations. Par exemple, nous devons avoir à l'esprit les mauvaises performances pour des résolvantes "mal conditionnées" ou encore la possible assez faible vitesse de convergence des équivalents déterministes vers leur quantité aléatoire. Cependant, comme cela a été démontré dans plusieurs travaux, la théorie des matrices aléatoires reste une approche robuste et pratique afin de permettre une abstraction de problèmes complexes par des équations qui ne dépendent au final que de très peu de paramètres.

Remerciements

Ce travail a été soutenu par l'ERC Starting Grant 305123 MORE (Advanced Mathematical Tools for Complex Network Engineering).

Bibliographie

- Bai Z. D., Silverstein J. W. (1998a). No Eigenvalues outside the Support of the Limiting Spectral Distribution of Large-dimensional Random Matrices. *The Annals of Applied Probability*, vol. 26, p. 316-345.
- Bai Z. D., Silverstein J. W. (1998b, janvier). No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Large Dimensional Sample Covariance Matrices. *Annals of Probability*, vol. 26, n° 1, p. 316-345.

8. Nous reconnaissons le fait que la dernière étape de cet exemple est plus intuitive que rigoureuse.

- Bai Z. D., Silverstein J. W. (2007). On the Signal-to-Interference Ratio of CDMA Systems in Wireless Communications. *The Annals of Applied Probability*, vol. 17, p. 81-101.
- Bai Z. D., Silverstein J. W. (2009). *Spectral Analysis of Large Dimensional Random Matrices* (2nd edition éd.). Springer Series in Statistics.
- Billingsley P. (1995). *Probability and Measure* (Third éd.). Hoboken, NJ, John Wiley & Sons, Inc.
- Couillet R. (2010). *Application of Random Matrix Theory to Future Wireless Flexible Networks*. Thèse de doctorat non publiée, Supélec.
- Couillet R., Debbah M. (2009). Uplink Capacity of Self-Organizing Clustered Orthogonal CDMA Networks in Flat Fading Channels. In *Information theory workshop, 2009. itw 2009. iee*. Taormina, Sicily.
- Couillet R., Debbah M. (2011). *Random Matrix Methods for Wireless Communications*. New York, NY, USA, Cambridge University Press. (First Edition)
- Couillet R., Debbah M., Silverstein J. W. (2011). A Deterministic Equivalent for the Analysis of Correlated MIMO Multiple Access Channels. *IEEE Transactions on Information Theory*, vol. 57, n° 6, p. 3493–3514.
- Dohler M. (2003). *Virtual Antenna Arrays*. Thèse de doctorat non publiée, University of London, Strand, London WC2R 2LS.
- Dupuy F., Loubaton P. (2010). On the Capacity Achieving Covariance Matrix for Frequency Selective MIMO Channels Using the Asymptotic Approach. *IEEE Transactions on Information Theory*. Consulté sur <http://arxiv.org/abs/1001.3102>
- Hachem W., Khorunzhy O., Loubaton P., Najim J., Pastur L. A. (2008, Sept.). A New Approach for Capacity Analysis of Large Dimensional Multi-Antenna Channels. *IEEE Transactions on Information Theory*, vol. 54, n° 9, p. 3987-4004.
- Hachem W., Loubaton P., Najim J. (2007). Deterministic Equivalents for Certain Functionals of Large Random Matrices. *Annals of Applied Probability*, vol. 17, n° 3, p. 875-930.
- Hachem W., Loubaton P., Najim J. (2008, décembre). A CLT for Information Theoretic Statistics of Gram Random Matrices with a Given Variance Profile. *Annals of Probability*, vol. 18, n° 6, p. 2071-2130.
- Henderson H. V., Searle S. R. (1981). On deriving the inverse of a sum of matrices. *Siam Review*, vol. 23, n° 1, p. 53–60.
- Hoydis J. (2010). *Channel Modeling for Cooperative Multi-Cell Systems*. (Unpublished Lecture Notes)
- Hoydis J. (2012). *Random Matrix Methods for Advanced Communication Systems*. Thèse de doctorat non publiée, Supélec, Alcatel-Lucent Chair on Flexible Radio.
- Hoydis J., Kobayashi M., Debbah M. (2011). Optimal Channel Training in Uplink Network MIMO Systems. *IEEE Transactions on Signal Processing*, vol. 59, n° 6, p. 2824-2833.
- Kammoun A., Kharouf M., Hachem W., Najim J. (2009, Nov.). A Central Limit Theorem for the SINR at the LMMSE Estimator Output for Large-Dimensional Signals. *IEEE Transactions on Information Theory*, vol. 55, n° 11, p. 5048–5063.
- Krein M. K., Nudelman A. A. (1977). *The Markov Moment Problem and Extremal Problems*. Providence, RI, USA, American Mathematical Society.

- Marčenko V. A., Pastur L. A. (1967, avril). Distributions of Eigenvalues for Some Sets of Random Matrices. *Math USSR-Sbornik*, vol. 1, n° 4, p. 457-483.
- Müller A. (2014). *Random Matrix Analysis of Future Multi Cell MU-MIMO Networks*. Thèse de doctorat non publiée, SUPÉLEC.
- Pastur L. A. (1999). A Simple Approach to Global Regime of Random Matrix Theory. In *Mathematical Results in Statistical Mechanics*, p. 429-454. World Scientific Publishing.
- Pastur L. A., Shcherbina M. (2011). *Eigenvalue distribution of large random matrices* n° 171. American Mathematical Soc.
- Peacock M. J. M., Collings I. B., Honig M. L. (2008). Eigenvalue Distributions of Sums and Products of Large Random Matrices via Incremental Matrix Expansions. *IEEE Transactions on Information Theory*, vol. 54, n° 5, p. 2123-2138.
- Rubio F., Mestre X., Hachem W. (2012). A CLT on the SNR of Diagonally Loaded MVDR Filters. *Signal Processing, IEEE Transactions on*, vol. 60, n° 8, p. 4178-4195.
- Silverstein J. W., Bai Z. D. (1995). On the Empirical Distribution of Eigenvalues of a Class of Large Dimensional Random Matrices. *Journal of Multivariate Analysis*, vol. 54, n° 2, p. 175-192.
- Tao T. (2012). *Topics in Random Matrix Theory* (vol. 132). American Mathematical Soc.
- Telatar E. (1999). Capacity of Multi-antenna Gaussian Channels. *European Transactions on Telecommunications*, vol. 10, n° 6, p. 585-595. (Also, 1995 in Technical Memorandum, Bell Laboratories, Lucent Technologies)
- Tse D. N. C., Hanly S. V. (1999, février). Linear Multiuser Receivers: Effective Interference, Effective Bandwidth and User Capacity. *IEEE Transactions on Information Theory*, vol. 45, n° 2, p. 641-657.
- Tulino A., Verdú S. (2004). Random Matrix Theory and Wireless Communications. *Foundations and Trends in Communications and Information Theory*, vol. 1, n° 1, p. 1-182.
- Vallet P., Loubaton P., Mestre X. (2010). Improved Subspace Estimation for Multivariate Observations of High Dimension: The Deterministic Signals Case. *IEEE Transactions on Information Theory*. Consulté sur <http://arxiv.org/abs/1002.3234>
- Van der Vaart A. W. (2000). *Asymptotic Statistics*. New York, Cambridge University Press.
- Verdú S., Shamai S. (1999, février). Spectral Efficiency of CDMA with Random Spreading. *IEEE Transactions on Information Theory*, vol. 45, n° 2, p. 622-640.
- Wagner S., Couillet R., Debbah M., Slock D. (2012, July). Large System Analysis of Linear Precoding in MISO Broadcast Channels with Limited Feedback. *IEEE Transactions on Information Theory*, vol. 58, n° 7, p. 4509-4537.
- Wigner E. (1958, mars). On the Distribution of Roots of Certain Symmetric Matrices. *The Annals of Mathematics*, vol. 67, n° 2, p. 325-327.
- Yates R. D. (1995). A Framework for Uplink Power Control in Cellular Radio Systems. *IEEE Journal on Selected Areas in Communications*, vol. 13, n° 7, p. 1341-1347.

Biographie

Axel Müller. En 2010, Axel Müller a reçu son M.Sc. en ingénierie de communications à l'université d'Ulm (Allemagne). De 2011 à 2014, il a préparé son Ph.D. en télécommunication à Centrale-Supélec, Paris, sur le sujet "Analyse des réseaux multi-cellulaires multi-utilisateurs futurs par la théorie des matrices aléatoires". Son directeur de thèse était professeur Mérouane Debbah et son co-encadrant était professeur Romain Couillet. Depuis 2014, Axel Müller est ingénieur de recherche au centre de R&D de Huawei en France ou il travaille sur des problèmes concertants 4.5G, 5G et 6G.

Mérouane Debbah. Mérouane Debbah est un ancien élève de l'Ecole Normale Supérieure de Cachan. Depuis 1999, il a travaillé pour les laboratoires Motorola à Saclay, le Centre de Recherche en Télécoms de Vienne et le département Communications Mobiles de l'Institut Eurecom (Sophia Antipolis). Depuis 2007 il est professeur titulaire à CentraleSupélec. De 2007 à 2014, il était directeur de la chaire Alcatel-Lucent sur la Radio Flexible. Depuis 2014, il est vice-président du centre de R&D de Huawei en France, et également directeur du laboratoire mathématiques et algorithmique. Mérouane Debbah est IEEE Fellow et WWRF Fellow.