Single-carrier MIMO systems for frequency selective propagation channels in presence of interference

Sonja Hiltunen

UNIVERSITÉ PARIS-EST

École doctorale MSTIC

Mathématiques et Sciences et Techniques de l'Information et de la Communication

Thése de doctorat

Spécialité: Signal, Image, Automatique

Présentée par:

Sonja Hiltunen

Étude des systèmes MIMO pour émetteurs mono-porteuses dans le contexte de canaux sélectifs en fréquence, en présence d'interférences

Analysis of MIMO systems for single-carrier transmitters in frequency-selective channels, in presence of interference

Soutenue le 17 décembre 2015 devant les membres du jury:

Président	Prof. Karim Abed-Meraim	Polytech Orléans
Rapporteur	Prof. Frédéric Pascal	Centrale-Supélec
Rapporteur	Prof. Jean-Pierre Cances	ENSIL Limoges
Examinateur	DR CSNR Walid Hachem	CNRS LTCI, Telecom ParisTech
Directeur de thèse	Prof. Philippe Loubaton	Université Paris Est, LIGM
Co-directeur de thèse	Prof. Pascal Chevalier	CNAM, Thales Communications

Acknowledgements

E quindi uscimmo a riveder le stelle.

I thank my family who always manages to get me in a good mood, simply by being there, always. Without you, the transition to France would have been unbearable. Thank you my mother Leena Seger, my sister Jenni Hiltunen, and my brother Viktor Hiltunen.

I thank my supervisor at Marne-la-Vallée, Philippe Loubaton, for always being availabe, enthusiastic, and showing that failure is not the end of the world. Without you, I probably would've ended up in a very bad place, due to stress. I would also like to thank Jean-Christophe Pesquet for his excellent advise in many topics. My PhD at Université Paris Est has been an opportunity to get to know many amazing people, and I will never forget the years that I've spent here. You are too many to be mentioned all by name, but a special thank you goes to Francesco Dolce, Feriel Abboud and Zakaria Chemli.

I thank Thales and CIFRE for giving me the opportunity to see a large French enterprise in action. Thank you Pascal Chevalier for showing me another world, and keeping me from sinking too deep into academia. Thank you Florian Dupuy for introducing me to the world of Thales, especially useful since you had followed the same route as the one I was starting on.

I thank CNAM for accepting me as part of their team, and especially the people there who became my second family, and made my life in France less tough in the beginning. Thank you especially to Ali Kabalan and Wosen Eshetu Kassa.

I thank of course KTH for giving me a solid scientific background to succeed my PhD. My years as an engineer student in KTH we unforgettable. I especially thank Markus Flierl who got me in touch with Télécom Paristech and France to begin with. One of the most important role-models for me for all these years has been Imadur Rahman at Ericsson, who has followed my progress from an insecure master student to a fully grown PhD. I thank you for your patience with me.

The trip has been long and the cost has been high. A long tale, like a tall tower, must be built a stone at a time - Stephen King

Summary

This thesis is dedicated to the study of $(K \times M)$ Multiple Input Multiple Output (MIMO) synchronization in frequency selective channels, in presence of interference.

Time synchronization of MIMO systems have been strongly studied in the last fifteen years, but most of the existing techniques assume an absence of interference. The current most powerful statistics robust to interference seems to be the one based on the Generalized Likelihood Ratio Test (GLRT). We therefore study the behaviour of the multi-antenna GLRT statistics η_{GLRT} of a known signal corrupted by a multi-path deterministic channel in additive white Gaussian noise with unknown spatial covariance. However, for complexity reasons, it is not always considered realistic for practical situations. An often-quoted and less complex alternative is the MMSE statistics η_{MMSE} which, on the other hand, performs worse than the GLRT. A part of this work has thus been devoted to showing that there exist non-GLRT statistics that are less complex to implement than the η_{GLRT}, while having similar performance. We also propose other approaches aiming at lowering the complexity of η_{GLRT}. We introduce alternative expressions of η_{GLRT}, and do its determinant computation explicitly for $K = 2$, showing that this allows a direct comparison between η_{GLRT} and η_{MMSE}. We introduce two new low-complexity statistics, η_{GLRT0} and η_{GLRT1}, and show that their performance is very close to that of η_{GLRT} for a wide range of parameter choices. Furthermore, we perform a comparative parameter analysis, taking into consideration the noise type, channel type, the number of transmit and receive antennas, and the orthogonality of the synchronization sequence. To further reduce complexity, a powerful procedure of computation rate reduction of the data correlation matrix is proposed. Lastly, the problem of optimization of the number of transmit antennas K for time synchronization has been investigated. showing, for high SNR, increasing performance with K as long as KM does not become greater than 8. All these approaches have been presented with the goal of optimizing the performance-complexity tradeoff.

Another aspect of MIMO synchronization studied in thesis is asymptotic analysis of the same GLRT, but for signals of high dimension. Due to the development of sensor networks

and acquisition devices, it has become common to be faced with multivariate signals of high dimension. Very often, the sample size that can be used in practice in order to perform statistical inference cannot be much larger than the signal dimension. In this context, it is well established that a number of fundamental existing statistical signal processing methods fail. It is therefore of crucial importance to revisit certain classical problems in the high-dimensional signals setting. We address the case where the number of sensors M and the number of samples N of the training sequence converge towards ∞ at the same rate. When the number of paths L does not scale with M and N, we establish that η_{GLRT} has a Gaussian behaviour with asymptotic mean $L \log \frac{1}{1-M/N}$ and variance $\frac{L}{N} \frac{M/N}{1-M/N}$. This is in contrast with the standard asymptotic regime $N \to +\infty$ and M fixed where η_{GLRT} has a χ^2 behaviour. Under hypothesis H_1, η_{GLRT} still has a Gaussian behaviour. The corresponding asymptotic mean and variance are obtained as the sum of the asymptotic mean and variance in the standard regime $N \to +\infty$ and M fixed, and $L \log \frac{1}{1-M/N}$ and $\frac{L}{N} \frac{M/N}{1-M/N}$ respectively, i.e. the asymptotic mean and variance under H_0. We also consider the case where the number of paths L converges towards ∞ at the same rate as M and N. Using known results of concerning the behaviour of linear statistics of the eigenvalues of large F-matrices, we deduce that in the regime where L, M, N converge to ∞ at the same rate, η_{GLRT} still has a Gaussian behaviour under H_0, but with a different mean and variance. The analysis of η_{GLRT} under H_1 when L, M, N converge to ∞ needs to establish a central limit theorem for linear statistics of the eigenvalues of large non zero-mean F-matrices, a difficult task that will be addressed in a future work. Motivated by the results obtained in the case where L remains finite, we propose to approximate the asymptotic distribution of η_{GLRT} by a Gaussian distribution whose mean and variance are the sum of the asymptotic mean and variance under H_0 when $L \to +\infty$ with the asymptotic mean and variance under H_1 in the standard regime $N \to +\infty$ and M fixed. Numerical experiments show that the Gaussian approximation corresponding to the standard regime $N \to +\infty$ and M fixed completely fails as soon as $\frac{M}{N}$ is not small enough. The large system approximations provide better results when $\frac{M}{N}$ increases, while also allowing to capture the actual performance for small values of $\frac{M}{N}$. We also observe that, for finite values of L, M, N, the Gaussian approximation obtained in the regime L, M, N converge towards ∞ is more accurate than the approximation in which L is fixed. We trace the ROC curves obtained through the limiting distributions, and note that the ROC curves that are obtained using the former large system approximation are accurate approximations of the empirical ones in a reasonable range of P_{FA} and P_{ND}.

Keywords*: MIMO, GLRT, Time Synchronization, Interference, Multichannel detection, asymptotic analysis, random matrix theory.*

Resumé

Cette thèse est consacrée à l'étude de la synchronisation des systèmes de communication multi-antennes (systèmes MIMO) en présence d'interférences.

La synchronisation temporelle des systèmes MIMO a été abondemment étudiée dans les quinze dernières années, mais la plupart des techniques existantes supposent que le bruit est blanc temporellement et spatialement, ce qui ne permet pas de modéliser la présence d'interférence. Nous considérons donc le cas de bruits blancs temporellement mais pas spatialement, dont la matrice de covariance spatiale est inconnue. En formulant le problème de l'estimation de l'instant de synchronisation comme un test d'hypothèses, nous aboutissons naturellement au test du rapport de vraisemblance généralisé (GLRT) qui donne lieu à la comparaison avec un seuil d'une statistique de test η_{GLRT}. Cependant, pour des raisons de complexité, l'utilisation de cette statistique n'est pas toujours considérée comme réaliste. La première partie de ce travail a donc été consacrée à mettre en évidence des tests alternatifs moins complexes à mettre en oeuvre, tout en ayant des performances similaires. Une analyse comparative exhaustive, prenant en considération le bruit et l'interférence, le type de canal, le nombre d'antennes en émission et en réception, et l'orthogonalité de la séquence de synchronisation est réalisée. Enfin, nous étudions le problème de l'optimisation du nombre d'antennes en émission K pour la synchronisation temporelle, montrant que pour un RSB élevé, les performances augmentent avec K dès que le produit de K avec le nombre d'antennes de réception M n'est pas supérieur à 8.

Le deuxième aspect de ce travail est une analyse statistique de η_{GLRT} dans le cas où le nombre d'antennes de réception M est élevé. Dans ce contexte, la taille de la séquence d'apprentissage N est du même ordre de grandeur que M, et cela conduit naturellement à étudier le comportement de η_{GLRT} dans le régime asymptotique des grands systèmes $M \to +\infty$, $N \to +\infty$ de telle sorte que $\frac{M}{N}$ tende vers une constante non nulle. Nous considérons le cadre applicatif d'un système muni d'une unique antenne d'émission et d'un canal à trajets multiples, qui est formellement identique à celui d'un système MIMO dont le nombre d'antennes d'émissions correspondrait au nombre de trajets. Lorsque le nombre de trajets L est beaucoup plus faible que N et M, nous

établissons que η_{GLRT} a un comportement Gaussien avec l'esperance asymptotique $L \log \frac{1}{1-M/N}$ et la variance $\frac{L}{N} \frac{M/N}{1-M/N}$. Ceci est en contraste avec le régime asymptotique standard $N \to +\infty$ et M et L fixe où η_{GLRT} a un comportement χ^2. Sous l'hypothèse H_1, η_{GLRT} a aussi un comportement gaussien. L'espérance et la variance asymptotique correspondantes sont obtenues comme la somme de l'esperance et la variance asymptotique dans le régime standard $N \to +\infty$ et M, L fixe, et $L \log \frac{1}{1-M/N}$ et $\frac{L}{N} \frac{M/N}{1-M/N}$ respectivement, soit l'esperance et la variance asymptotique sous H_0. Nous considérons également le cas où le nombre de trajets L tend vers ∞ à la même vitesse que M et N. Nous utilisons des résultats connus concernant le comportement des statistiques linéaires des valeurs propres des grandes F matrices, et déduisons que dans le régime où L, M, N tendent vers ∞ à la même vitesse, η_{GLRT} a encore un comportement Gaussien sous H_0, mais avec une esperance et variance différentes. L'analyse de η_{GLRT} sous H_1 lorsque N, M, L convergent vers $+\infty$ nécessite l'établissement d'un théorème central limite pour les statistiques linéaires des valeurs propres de matrices F de moyennes non-nulles, une tâche difficile. Motivé par les résultats obtenus dans le cas où L reste fini, nous proposons d'approximer la distribution asymptotique de η_{GLRT} par une distribution Gaussienne dont l'esperance et la variance sont la somme de la l'esperance et la variance asymptotique sous H_0 quand $L \to +\infty$ avec l'esperance et la variance asymptotique sous H_1 dans le régime classique $N \to +\infty$ et M fixé. Des simulations numériques permettent de comparer les courbes ROC des différents approximants avec des courbes ROC empiriques. Les résultats montrent que l'approximation Gaussienne correspondante au régime classique $N \to +\infty$ et M fixé échoue complètement dès que $\frac{M}{N}$ n'est pas assez petit. Nos approximants de grandes dimensions fournissent de meilleurs résultats quand $\frac{M}{N}$ augmente, tout en permettant de capturer la performance réelle pour les petites valeurs de $\frac{M}{N}$. Nous observons également que l'approximation Gaussienne obtenue pour le régime òu L, M, N tendent vers ∞ donne des résultats plus proches de la réalité que ceux qui sont fournis par le régime où L est fixé.

***Mots-clés**: MIMO, GLRT, synchronisation temporelle, interférences, Détection multivoie, analyse asymptotique, théorie des matrices aléatoires.*

Publications

During the thesis, one journal paper has been published:

- S. Hiltunen, P. Loubaton, P. Chevalier, *Large system analysis of a GLRT for detection with large sensor arrays in temporally white noise*, IEEE Trans. Sig. Process., Vol 63, No. 20, pp. 5409-5423, Oct. 2015

and one journal paper will be submitted:

- S. Hiltunen, P. Chevalier, P. Loubaton, *New insights into time synchronization of MIMO systems with interference*

The journal paper that has not been submitted yet corresponds to chapter 2, whereas the IEEE journal corresponds to chapter 3. Moreover, the following articles have been presented at conferences:

- S. Hiltunen, P. Chevalier, P. Loubaton, *New insights into time synchronization of MIMO systems with interference*, Proceedings EUSIPCO 2015, Nice, France, Aug. 31 - Sep. 4 2015, pp. 1396-1400
- S. Hiltunen, P. Loubaton and P. Chevalier, *Asymptotic analysis of a GLRT for detection with large sensor arrays*, Proceedings EUSIPCO 2014, Lisbon, Portugal, Sep. 1-5, 2014, pp. 2160-2164

Furthermore, a patent has been filed during this thesis, corresponding to the statistics introduced in chapter 2.

- P. Chevalier, F. Pipon, S. Hiltunen, *Procédé et dispositif de synchronisation MIMO en présence d'interférences*, Thales Brevet No. 1501813

List of notations

General notations

For a complex matrix $\mathbf{A}$, we denote by $\mathbf{A}^T$ and $\mathbf{A}^*$ its transpose and its conjugate transpose, and by $\mathrm{Tr}(\mathbf{A})$ and $\|\mathbf{A}\|$ its trace and spectral norm. $\mathbf{I}$ will represent the identity matrix and $\mathbf{e}_n$ will refer to a vector having all its components equal to 0 except the n-th which is equal to 1. If we need to be precise, we denote by $\mathbf{I_K}$ the $(K \times K)$ identity matrix. We denote by $\det(\mathbf{A})$ its. determinant, and by $\mathrm{Diag}(\mathbf{A})$ the diagonal matrix $\mathbf{D}$ where the elements (i,i) of $\mathbf{D}$ are the elements (i,i) of $\mathbf{A}$, and the other elements are zero. $\mathbf{A}_k$ or $\mathbf{a}(k)$ are used to denote its k:th column, and $A_{i,j}$ the element on its i:th row and j:th column

For a complex vector $\mathbf{a}$, we denote by $\mathrm{Diag}(\mathbf{a})$ the diagonal matrix $\mathbf{D}$ where element (i,i) of $\mathbf{D}$ is the element i of $\mathbf{a}$ and the other elements are zero.

The real normal distribution with mean m and variance σ^2 is denoted $\mathcal{N}_{\mathbb{R}}(m, \sigma^2)$. A complex random variable $Z = X + i\,Y$ follows the distribution $\mathcal{N}_{\mathbb{C}}(\alpha + i\,\beta, \sigma^2)$ if X and Y are independent with respective distributions $\mathcal{N}_{\mathbb{R}}(\alpha, \frac{\sigma^2}{2})$ and $\mathcal{N}_{\mathbb{R}}(\beta, \frac{\sigma^2}{2})$.

For a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ and a random variable X, we write

$$X_n \to X \; a.s. \text{ and } X_n \to_{\mathcal{D}} X$$

when X_n converges almost surely and in distribution, respectively, to X when $n \to +\infty$. Finally, if $(a_n)_{n\in\mathbb{N}}$ is a sequence of positive real numbers, $X_n = o_P(a_n)$ will stand for the convergence of $(X_n/a_n)_{n\in\mathbb{N}}$ to 0 in probability, and $X_n = \mathcal{O}_P(a_n)$ denotes boundedness in probability (i.e. tightness) of the sequence $(X_n/a_n)_{n\in\mathbb{N}}$.

Paper-specific notation

- K : Number of transmit antennas
- M : Number of receive antennas
- N : Synchronization sequence length
- L : Number of paths in the frequency selective channel
- η : A synchronization statistics
- $\mathbf{Y}$: Received signal matrix
- $\mathbf{H}$: Channel matrix
- $\mathbf{S}$: Synchronization sequence matrix
- $\mathbf{V}$: Noise matrix
- $\mathbf{R}$: Noise covariance matrix $\mathbf{R} = \mathbb{E}(\mathbf{v}_n \mathbf{v}_n^*)$
- σ^2 : Noise power
- $|\alpha|_{i,j}$: Spatial correlation between deterministic channels i and j
- $|\rho|_{i,j}$: Temporal correlation between the synchronization sequences sent from transmit antennas i and j
- θ_i : Direction of arrival of the signal from transmit antenna i, in degrees
- $\mathbf{h}_i(\theta_i)$: Deterministic channel i with its direction of arrival in degrees

Correlation matrices used in this paper

For the reader's convenience, Table 1 gathers all the correlation matrices and vectors used in this paper.

Notation	Interpretation	Definition
$\hat{\mathbf{R}}_{yy}$	Empirical correlation of the received sequence.	$\frac{1}{N}\sum_{k=0}^{N-1}\mathbf{Y}_k\mathbf{Y}_k^* = \frac{1}{N}\mathbf{Y}\mathbf{Y}^*$
$\hat{\pi}_y$	Power of the received sequence.	$\frac{1}{N}\sum_{k=0}^{N-1}\mathbf{Y}_k^*\mathbf{Y}_k = \mathrm{Tr}(\hat{\mathbf{R}}_{yy})$
$\hat{\pi}_{y_j}$	Power of received sequence on receive antenna j.	$\frac{1}{N}\sum_{k=0}^{N-1}Y_{j,k}Y_{j,k}^* = (\hat{\mathbf{R}}_{yy})_{j,j}$
$\hat{\mathbf{R}}_{ys}$	Intercorrelation between the received sequence and the useful samples.	$\frac{1}{N}\sum_{k=0}^{N-1}\mathbf{Y}_k\mathbf{S}_k^* = \frac{\mathbf{Y}\mathbf{S}^*}{N}$
$\hat{\mathbf{r}}_{ys_i}$	Intercorrelation between the received signal and the useful samples associated to transmit antenna i.	$\frac{1}{N}\sum_{k=0}^{N-1}\mathbf{Y}_k S_{i,k}^* = \frac{\mathbf{Y}(\mathbf{S}^T)_i^*}{N}$
$\hat{\pi}_{y_j s_i}$	Intercorrelation between the received signal at receive antenna j and the useful samples associated to transmit antenna i.	$\frac{1}{N}\sum_{k=0}^{N-1}Y_{j,k}S_{i,k}^* = \frac{(\mathbf{Y}^T)_j(\mathbf{S}^T)_i^*}{N}$
$\mathbf{R}_{ss}$	Correlation matrix of the useful samples.	$\frac{1}{N}\sum_{k=0}^{N-1}\mathbf{S}_k\mathbf{S}_k^* = \frac{\mathbf{S}\mathbf{S}^*}{N}$
$\pi_{s_{ij}}$	Intercorrelation between useful samples associated to transmit antennas i and j.	$\frac{1}{N}\sum_{k=0}^{N-1}S_{i,k}S_{j,k}^* = \frac{(\mathbf{S}^T)_i(\mathbf{S}^T)_j^*}{N} = (\mathbf{R}_{ss})_{i,j}$

Table 1: Paper-specific notations

Contents

*

Chapter 1

Introduction

This thesis has been carried out within the framework of a CIFRE convention between Thales Communications, Université Paris Est and CNAM, and is dedicated to the study of Multiple Input Multiple Output (MIMO) synchronization in frequency selective channels, in the presence of interference.

The first area of research is the optimization of system parameters in MIMO synchronization (especially optimizing the number of transmit antennas) as well as an investigation of the complexity-performance trade-off in the design of such systems. More specifically, Generalized Likelihood Ratio Tests (GLRTs) are studied, with the goal of simplifying the complexity of synchronization.

The second area of research is an asymptotic analysis of the same GLRT, using random matrix methods. The goal of the asymptotic analysis is to propose asymptotic distribution of the synchronization statistics with different large system assumptions.

1.1 MIMO in synchronization: Friend or foe?

1.1.1 Introduction: Evolution and advantages of MIMO

During the past decades, there has been an explosion in the number of receiver and transmitter architectures aiming at taking advantage of several antennas in transmission and reception in wireless systems. Wireless systems experience different challenges than wireline systems, and the main challenges are multipath and fading channels, which make it challenging to achieve very high data rates and good error performance. Multiple-input-multiple-output (MIMO) has been

hailed as a breakthrough in wireless communication, and has been proposed as a solution to this so-called bottleneck of wireless communications.

The possible gain obtained from using multiple antennas can be divided into three main categories:

1. *Beamforming gain*, obtained by steering the signal in a certain direction, and nulling certain directions. Beamforming, a classical multi-antenna technique, can either be done at transmission or reception.

2. *Diversity gain*, which is obtained by sending the information through channels with different characteristics. For example, in order to take advantage of the different fading characteristics of the channels experienced by the independent antennas, we can send the same information from all transmit antennas. Diversity gain increases the robustness of the system by eliminating fading. In a certain sense, it converts the channel from a fading to a non-fading one, and thus increases the reliability of transmission.

3. *Multiplexing gain*, specific to a MIMO system, is obtained by spatial multiplexing. The transmitted data is divided into independent data streams which can be decoded at the receiver. The separability at reception makes use of rich multipath, which makes the channel spatially selective. Instead of seeing it as a problem, spatial multiplexing actually exploits multipath. The main goal of spatial multiplexing is to maximize transmission rate, as opposed to transmit/receive diversity where the main goal is to increase reliability.

However, when wireless communication was mainly used for voice-based systems, the practical interest for MIMO systems was weak; there was no real need for very high data rates. Due to the difficult characteristics of the wireless channel, rate-intense applications such as high-rate video streaming has traditionally been done with wireline technology. Furthermore, the interest in MIMO was low due to practical problems such as increased hardware costs and challenges in implementation, especially in the handsets. Nevertheless, with the recent upswing in smartphones and other mobile devices, we have been obliged to solve the problem of increasing the data rate even for wireless applications. On the other hand, now that the handsets have become more sophisticated, the constraints on the size of the MIMO system are less severe, and it has become practically feasible to have multiple antennas on both sides of the link. MIMO is thus not only a theoretical consideration, and it has indeed been successfully implemented in several well-established standards. Some examples include CDMA which uses Alamouti space-time-codes, MIMO transmissions in WLAN, and MIMO in the 3GPP LTE standard.

What are the alternatives to MIMO, and why are they not enough for high data-rate wireless communication? A conventional approach to wireless communication is the single-input-single-

output (SISO) system. A SISO system, however, is unable to reject interference, or take advantage of spatial diversity. One approach to solving the problem of fading or multipath in a SIMO system is using channel coding combined with interleaving, to achieve time diversity. But this only works up to a certain point, and cannot produce very high data rates. Traditional multi-antenna systems which have been around for decades have several antennas in either transmission or reception, generally at the base station. These systems use mainly beamforming or spatial diversity at either emission or reception. They can also perform interference reduction by taking advantage of several receive antennas, through for example receive beamforming. In theory, high data-rate communication can be implemented with a SIMO system, but this requires very high bandwidth, which is not easy to obtain since bandwidth is an expensive resource. In short, despite its advantages, a SIMO system is also unfeasible for Gigabit internet.

The basis of a MIMO system, on the other hand, is having multiple spatially distributed antennas at both sides of the link. This transform the system from a vector system to a matrix system, and adds additional degrees of freedom through the so-called *spatial dimension* which can then be exploited to perform spatial multiplexing. A MIMO system also has a *joint* transmit and receive diversity gain. Theoretical results show that the maximum bit rate of a $(K \times M)$ MIMO system grows linearly with $\min(K, M)$. As a summary, the benefits of smart antennas are retained, and new benefits are added.

Nevertheless, as with any system, the gains mentioned for a multi-antenna system do not come for free. Some of the challenges that exist in MIMO synchronization are summarized in the next subsection.

1.1.2 Challenges in MIMO synchronization

An important challenge in MIMO synchronization is that there exists a performance-complexity trade-off; when the number of antennas increases, so does the complexity of the algorithms. From a complexity point of view, it is thus appropriate to ask the question whether we should add extra antennas for the specific system and application under study. Complexity is of course also a problem from the hardware cost and compatibility point of view. Increasing the number of antennas in a system is an important system decision, and the advantages and possible disadvantages should be carefully weighed.

Another important consideration in MIMO systems is the assumptions on interference of the system. Point-to-point systems can be a reasonable assumption in certain applications, for example the IEEE 802.11n standard, since it is designed to ensure that its short range links do not suffer from interference. But in general, MIMO cellular systems are interference

limited, suffering from internal and other cell interference. As the number of interfering streams increases, the interference cannot be suppressed by spatial signal processing, and is treated as noise. Therefore, as opposed to what one may believe, adding transmit antennas in the system can actually decrease the throughput at low SINR.

An additional important system feature is the channel characteristics. A common, often simplified, assumption is that of uncorrelated antenna arrays and Rayleigh fading channels. To fully cover the problem, there is a need for investigating the performance of MIMO systems even for deterministic channels, and understanding which system parameters to choose according to a suitable channel model.

Considering the above challenges, the following section outlines the state of the art in MIMO synchronization, with the goal of validating the need for our study.

1.1.3 State of the art in MIMO synchronization

Time and frequency synchronization of MIMO systems have been strongly studied in the last fifteen years, mainly in the context of direct-sequence coded division multiple access (DS-CDMA) and orthogonal frequency division multiplex (OFDM) links. Both coarse and fine time synchronization jointly with frequency offset estimation and compensation have been analyzed, and many techniques have been proposed either for time-frequency synchronization [1–4] or time synchronization only [5–9]. Nevertheless, most of these techniques assume an absence of interference. The scarce papers of the literature dealing with MIMO synchronization in the presence of interference correspond to [2, 7–9]. However, [2] and [7] only consider the problem of MIMO synchronization in the presence of multi-user interference (MUI), and [8] seems to be the only paper dealing with MIMO synchronization in the presence of interference of any kind, such as hostile jammers. In [8], several statistics are proposed for time synchronization for both flat fading and frequency selective fading channels. Despite the numerous existing algorithms for MIMO time synchronization, many important questions about their optimality, performance and complexity have arisen.

First of all, none of the receivers developed for MIMO synchronization in the absence of interference [1, 3–6, 10–27] has been developed through a GLRT approach in the general case of an SC non-DS-CDMA link and arbitrary potentially non-orthogonal sequences. It is well-known [28], however, that contrary to likelihood ratio test (LRT) statistics, GLRT statistics may be suboptimal from a detection point of view. With this in mind, one may wonder if a non-GLRT statistics can be equal or better than existing GLRT statistics.

While the GLRT statistics in presence of interference has already been presented [8], questions regarding its practical usefulness arise. In [8], two statistics for flat fading channels, robust to interferences, are derived from an MMSE and a GLRT approach respectively. The GLRT statistics, called η_{GLRT} in the following, assumes unknown, Gaussian, spatially colored and temporally white total noise. We note that the GLRT receiver may be very costly to implement, for a large number of antennas in particular, since for a $(K \times M)$ MIMO system, it requires both an $(M \times M)$ matrix inversion and an $(M \times M)$ or $(K \times K)$ determinant computation at each tested sample position. An alternative to η_{GLRT} is the MMSE statistics proposed in [8]. However, the MMSE approach will be shown to be very sensitive to training sequence correlation, which may limit its practical use in this context.

Before going on to describe our proposition to resolve the above problems, we will introduce the considered system model, and briefly recall the concept of synchronization as a hypothesis testing problem. We will also introduce the concept of Generalized Likelihood Ratio Test (GLRT).

1.2 Model and problem formulation

1.2.1 Hypotheses

We consider a $(K \times M)$ MIMO radiocommunication link with K and M narrow-band antennas at transmission and reception respectively, and denote by $\mathbf{s}(k)$ the $(K \times 1)$ complex synchronization sequence, known by the receiver, transmitted on the transmit antennas at time k. Assuming a perfect time and frequency synchronization, the vector $\mathbf{y}(k)$ of the complex envelopes of the signals at the output of the M receive antennas at time k can be written as

$$\mathbf{y}(k) = \sum_{l=0}^{L-1} \mathbf{H}_l \mathbf{s}(k-l) + \mathbf{v}(k) \tag{1.1}$$

Here, $\mathbf{H}_l$ and l are the $(K \times M)$ channel matrix of the path l, and $\mathbf{v}(k)$ is the sampled total noise vector, which contains the potential contribution of MUI interferences, jammers and background noise. The samples $\mathbf{v}(k)$ are assumed to be zero-mean, temporally white, i.i.d, circular, and Gaussian with covariance matrix $\mathbf{R} = \mathbb{E}[\mathbf{v}(k)\mathbf{v}(k)^*]$.

Let us denote by $\mathbf{Y}$ and $\mathbf{V}$ the $(M \times N)$ observation and total noise matrices $\mathbf{Y} = [\mathbf{y}(1), \ldots, \mathbf{y}(N)]$ and $\mathbf{V} = [\mathbf{v}(1), \ldots, \mathbf{v}(N)]$ respectively. Let us denote by $\mathbf{S}$ the $(KL \times N)$

synchronization sequence matrix

$$\mathbf{S} = \begin{pmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \dots & \mathbf{s}_N \\ \mathbf{0} & \mathbf{s}_1 & \dots & \mathbf{s}_{N-1} \\ & \ddots & & \\ \mathbf{0} & & \ddots & \mathbf{s}_{N-L+1} \end{pmatrix} \tag{1.2}$$

where $[\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N]$ is the matrix of synchronization sequences on the first path, and let $\mathbf{H} = [\mathbf{H}_0, \mathbf{H}_1, \dots, \mathbf{H}_{L-1}]$ be the frequency selective channel matrix. With these definitions, we obtain the matrix model

$$\mathbf{Y} = \mathbf{HS} + \mathbf{V} \tag{1.3}$$

Note that for frequency selective channels, a priori information about the maximal number of paths L is required. What is more interesting is that this model can be used for any other application that can be written in the above matrix form, and is not limited to synchronization.

1.2.2 Synchronization as hypothesis testing

The problem of time synchronization of a MIMO link may be viewed as a detection problem with three hypotheses. The first hypothesis (H_1) is that the signal matrix $\mathbf{S}$ is perfectly aligned in time with the observation matrix $\mathbf{Y}$, and corresponds to model (1.3), or

$$\mathrm{H}_1 : \mathbf{Y} = \mathbf{HS} + \mathbf{V} \tag{1.4}$$

The second hypothesis (H_0) is that there is no signal in the observation matrix $\mathbf{Y}$, and corresponds to model (4.5) given by

$$\mathrm{H}_0 : \mathbf{Y} = \mathbf{V} \tag{1.5}$$

The third hypothesis is the intermediate hypothesis, which supposes that a part of the received signal is only noise, while another part contains the useful signal. Let us define the $(KL \times N)$ partial synchronization sequence matrix as

$$\mathbf{S}_p = [\mathbf{0}_{(KL \times A)}, \mathbf{S}_1], \tag{1.6}$$

where $\mathbf{S}_1$ is a $(KL \times (N - A))$ matrix with $A < N$ which contains a subset of the synchronization sequence matrix defined in (4.2). The intermediate hypothesis is then given by

$$\mathrm{H}_\mathrm{p} : \mathbf{Y} = \mathbf{HS}_p + \mathbf{V} \tag{1.7}$$

The hypothesis H_p is important, since especially in a frequency selective channel there may be several good detections, arising from only a part of the signal being present.

Nevertheless, to simplify the problem, we assume a binary hypothesis testing model with only H_0 and H_1. The problem of synchronization as detection then consists in elaborating a statistical test η which is a function of the observations $\mathbf{Y}$, and in comparing its value to a threshold s. If the threshold is exceeded, detection is considered. The threshold is used to choose between H_0 and H_1, by

$$\eta \underset{H_0}{\overset{H_1}{\gtrless}} s \tag{1.8}$$

If the threshold is exceeded several times, a decision rule must be applied to choose the moment of detection.

1.2.3 Generalized Likelihood Ratio Test (GLRT)

According to the Neyman-Pearson theory of detection, the optimal statistical test for the detection of matrix $\mathbf{S}$ from matrix $\mathbf{Y}$ is the likelihood ratio test (LRT), which consists in comparing the function $LRT \triangleq p_{H_1}(\mathbf{Y}|\mathbf{H},\mathbf{S},\mathbf{R})/p_{H_0}(\mathbf{Y}|\mathbf{R})$ to a threshold, where $p_{H_i}(\mathbf{Y}|\ldots)$ $(i=0,1)$ is the conditional probability density of $\mathbf{Y}$ under H_i. For our noise model, the expression of the LRT takes the form

$$\text{LRT} = \frac{\prod_{k=1}^{N} p_{H_1}(\mathbf{y}(k) \mid \mathbf{s}(k),\mathbf{H},\mathbf{R})}{\prod_{k=1}^{N} p_{H_0}(\mathbf{y}(k) \mid \mathbf{R})}. \tag{1.9}$$

For our Gaussian model, the probability density functions take the form

$$p_{H_0}(\mathbf{y}(k) \mid \mathbf{R}) = \frac{1}{\pi^M \det(\mathbf{R})} e^{-\mathbf{y}(k)^* \mathbf{R}^{-1} \mathbf{y}(k)} \tag{1.10}$$

and

$$p_{H_1}(\mathbf{y}(k) \mid \mathbf{s}(k),\mathbf{H},\mathbf{R}) = \frac{1}{\pi^M \det(\mathbf{R})} e^{-(\mathbf{y}(k)-\mathbf{H}\mathbf{s}(k))^* \mathbf{R}^{-1} (\mathbf{y}(k)-\mathbf{H}\mathbf{s}(k))}, \tag{1.11}$$

As, in practice, $\mathbf{R}$, $\mathbf{H}$ or one of the two are unknown, they have to be replaced in (1.9) by their maximum likelihood (ML) estimates under each of the two hypotheses H_1 and H_0, which gives rise to the Generalized Likehood Ratio Test (GLRT). In the case where both $\mathbf{R}$ and $\mathbf{H}$ are unknown, the GLRT test statistic becomes

$$\eta_{\text{GLRT}} = \frac{\underset{\mathbf{R},\mathbf{H}}{\arg\max} \prod_{k=1}^{N} p_{H_1}(\mathbf{y}(k) \mid \mathbf{s}(k),\mathbf{H},\mathbf{R})}{\underset{\mathbf{R}}{\arg\max} \prod_{k=1}^{N} p_{H_0}(\mathbf{y}(k) \mid \mathbf{R})}. \tag{1.12}$$

1.2.4 Performance of a binary hypothesis test

The performance of a synchronization test statistic η is characterized by the probability of non-detection under H_1 (P_{ND}) for a given false alarm probability (P_{FA}), which is the probability

of exceeding the threshold s under H_0. More formally,

$$\begin{aligned} \mathrm{P}_{\mathrm{FA}} &= P(\eta > t \mid \mathrm{H}_0) \\ \mathrm{P}_{\mathrm{ND}} &= P(\eta < t \mid \mathrm{H}_1) \end{aligned} \tag{1.13}$$

Figure 1.1 illustrates the concept of P_{FA} and P_{ND} for the case where both H_0 and H_1 correspond to statistics that follow a Gaussian distribution, for a $\mathrm{P}_{\mathrm{FA}} = 10^{-3}$. In our work, the threshold s

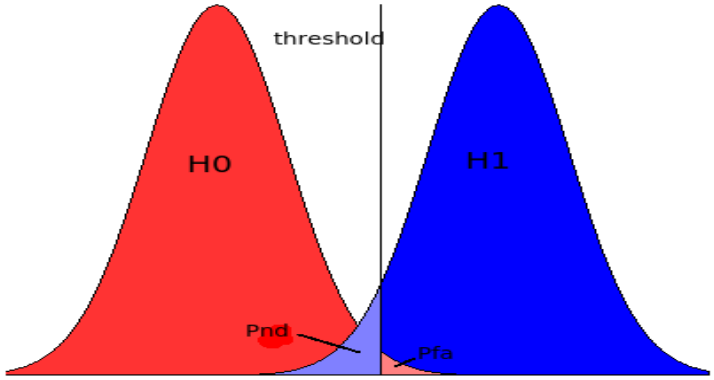


Figure 1.1: P_{ND} and P_{FA} in binary hypothesis testing

for a given P_{FA} will be set empirically by generating a large number of samples under H_0. This threshold will then be used in the simulations to determine the P_{ND}.

1.3 Objective

The objective of this PhD thesis is to obtain responses to some of the open problems in synchronization in MIMO systems, namely:

1. Choice of synchronization algorithms for MIMO systems, and analysis of the complexity-performance trade-off.
2. Possible advantages of MIMO over SIMO, and investigation of the need for transition from SIMO to MIMO in the synchronization context.
3. Characterization of the asymptotic behavior of MIMO synchronization statistics.
4. Prediction of performance of MIMO synchronization.

1.4 Parameter optimization and complexity reduction for time synchronization

1.4.1 Introduction to the problem

The first part of the PhD thesis, fully described in chapter 2, concerns parameter optimization and complexity reduction of MIMO synchronization. As mentioned in subsection 1.1.3, the current most powerful receiver is based on a generalized likelihood ratio test (GLRT), which assumes unknown, circular, temporally white and spatially colored Gaussian noise. Nevertheless, the computational complexity of this statistics is higher than its non-GLRT counterparts, which, unfortunately, do not perform as well in most cases. As complexity is an important issue for practical implementations and may be prohibitive for a large number of antennas, the purpose of this study is to propose several ways of decreasing the complexity of the GLR test while retaining its performance.

1.4.2 Synchronization statistics studied in the thesis

A first straightforward way of optimizing performance, and if applicable, decreasing the complexity, of MIMO synchronization is based on applying the knowledge we have about the system. If we know that the noise is spatially and temporally independent, we may use the following GLRTs (derived in Appendices 2.6 and 2.7), optimized for either equal or unequal noise power at each receiver antenna:

- Both $\mathbf{R}$ and $\mathbf{H}$ are unknown, but $\mathbf{R}$ is known to be in the form $\mathbf{R} = \sigma^2\mathbf{I}$ (spatially white noise with equal noise power at all receive antennas):

$$\eta_{\text{GLRT,we}} = \frac{\underset{\sigma^2,\mathbf{H}}{\arg\max} \prod_{k=1}^{N} p_{H_1}(\mathbf{y}(k) \mid \mathbf{s}(k), \mathbf{H}, \sigma^2\mathbf{I})}{\underset{\sigma^2}{\arg\max} \prod_{k=1}^{N} p_{H_0}(\mathbf{y}(k) \mid \sigma^2\mathbf{I})}. \tag{1.14}$$

- Both $\mathbf{R}$ and $\mathbf{H}$ are unknown, but $\mathbf{R}$ is known to be in the form $\widetilde{\mathbf{R}} = \text{Diag}[\sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2]$ (spatially white noise with unequal powers at the different receive antennas):

$$\eta_{\text{GLRT,wu}} = \frac{\underset{\sigma_1^2,\ldots,\sigma_M^2,\mathbf{H}}{\arg\max} \prod_{k=1}^{N} p_{H_1}(\mathbf{y}(k) \mid \mathbf{s}(k), \mathbf{H}, \widetilde{\mathbf{R}})}{\underset{\sigma_1^2,\ldots,\sigma_M^2}{\arg\max} \prod_{k=1}^{N} p_{H_0}(\mathbf{y}(k) \mid \widetilde{\mathbf{R}})} \tag{1.15}$$

Nevertheless, if both $\mathbf{R}$ and the channel matrix $\mathbf{H}$ are unknown, the GLRT synchronization statistics is calculated using (1.18), which is the starting point of both chapters 2 and 3. This GLRT statistics, already present in the literature [8], is given by

$$\eta'_{\text{GLRT}} = \det[\mathbf{I}_N - \mathbf{S}^*(\mathbf{S}\mathbf{S}^*)^{-1}\mathbf{S}\mathbf{Y}^*(\mathbf{Y}\mathbf{Y}^*)^{-1}\mathbf{Y}]^{-N}. \tag{1.16}$$

A sufficient statistic is the log-likelihood ratio $\eta_{\text{GLRT}} = \log(\eta'_{\text{GLRT}})/N$. Using $\det(\mathbf{I} - \mathbf{AB}) = \det(\mathbf{I} - \mathbf{BA})$, we have the statistics in the form that we will use from now on:

$$\eta_{\text{GLRT}} = -\log\det\left[\mathbf{I}_L - \frac{\mathbf{S}\mathbf{Y}^*}{N}\left(\frac{\mathbf{Y}\mathbf{Y}^*}{N}\right)^{-1}\frac{\mathbf{Y}\mathbf{S}^*}{N}\left(\frac{\mathbf{S}\mathbf{S}^*}{N}\right)^{-1}\right] \tag{1.17}$$

$$= -\log\det\left(\mathbf{I}_M - \hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^*\right) \tag{1.18}$$

The GLR is generally considered an optimal test, but it can be seen from equation (1.18) that its complexity quickly gets prohibitive when K, M, N increase in size. To compute η_{GLRT} we have to, at each time index, compute and invert a $(M \times M)$ matrix $\hat{\mathbf{R}}_{yy}$, and compute a determinant of a square matrix of size $\min(M, K, N)$.

An alternative to the GLRT statistics is the MMSE statistics, which is also robust to interferences. The MMSE statistics minimizes the LS error between the known synchronization sequence $\mathbf{S}$ and its LS estimation from a spatial filtering of the data $\mathbf{Y}$, and is given by

$$\eta_{\text{MMSE}} = \frac{\text{Tr}(\hat{\mathbf{R}}_{ys}^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys})}{\text{Tr}(\mathbf{R}_{ss})} \tag{1.19}$$

It is easy to see that η_{MMSE} is less computationally complex than η_{GLRT}. It still requires a matrix inversion at each time index, but does not necessitate a determinant computation. However, as will be shown in chapter 2, the MMSE becomes suboptimal for non-orthogonal training sequences.

A third interesting statistics robust to interference is the GLRT test statistics optimized for the theoretical case where $\mathbf{R}$ is known and $\mathbf{H}$ is unknown.

$$\eta_{\text{GLRT,kn}} = \frac{\arg\max_{\mathbf{H}} \prod_{k=1}^{M} p_{H_1}(\mathbf{y}(k) \mid \mathbf{s}(k), \mathbf{H}, \mathbf{R})}{\prod_{k=1}^{M} p_{H_0}(\mathbf{y}(k) \mid \mathbf{R})} \tag{1.20}$$

$$= \text{Tr}\left[\mathbf{R}^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^*\right]. \tag{1.21}$$

This is of course an entirely theoretical criterion, since it is practically impossible to know the exact noise covariance matrix $\mathbf{R}$. The interest in this criterion stems from the idea that $\mathbf{R}$ may be replaced with an estimate. The performance of the obtained non-GLRT statistics can then be compared with that of the GLRT statistics η_{GLRT}.

1.4.3 Complexity reduction and parameter optimization of MIMO synchronization

With these considerations as a starting point, we therefore propose in chapter 2 several ways of optimizing and reducing the complexity of MIMO synchronization, summarized as follows:

1. Introduction of explicit expressions for GLRT synchronization in the absence of interference, for two types of spatially white noise.
2. Introduction of alternative expressions of η_{GLRT} optimized for presence of interference, where the determinant computation is done explicitly for $K = 2$. These expressions are mainly useful since they allow for a direct comparison of η_{GLRT} with η_{MMSE}.
3. Introduction of two new low-complexity statistics, η_{GLRT0} and η_{GLRT1}, based on the GLRT optimized for the theoretical case where the covariance of the noise $\mathbf{R}$ is known.
4. Proposition of a procedure of computation rate reduction of the data correlation matrix.
5. Investigation of the problem of optimization of the number of transmit antennas for time synchronization.

Table 1.1 summarizes the different GLRT or non-GLRT statistics that are studied or introduced in this thesis.

Name	**Interpretation**	**Sufficient statistics**
η_{GLRT}	GLRT statistics for unknown $\mathbf{R}$ and $\mathbf{H}$	$-\log\det\left(\mathbf{I}_M - \hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*}\right)$
$\eta_{\text{GLRT,we}}$	GLRT statistics for spatially white noise with $\mathbf{R} = \sigma^2\mathbf{I}_M$	$\frac{\text{Tr}(\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*})}{\text{Tr}(\hat{\mathbf{R}}_{yy})}$
$\eta_{\text{GLRT,wu}}$	GLRT statistics for spatially white noise with $\mathbf{R} = \text{Diag}[\sigma_1^2,\ldots,\sigma_M^2]\mathbf{I}_M$	$\frac{\prod_{n=1}^{M}(\hat{\mathbf{R}}_{yy})_{n,n}}{\prod_{n=1}^{M}(\hat{\mathbf{R}}_{yy}-\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*})_{n,n}}$
$\eta_{\text{GLRT,kn}}$	GLRT statistics for known $\mathbf{R}$ and unknown $\mathbf{H}$	$\text{Tr}\left[\mathbf{R}^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*}\right]$
η_{GLRT0}	Non-GLRT statistics based on $\eta_{\text{GLRT,kn}}$, but with estimated $\mathbf{R} = \hat{\mathbf{R}}_{yy}$	$\text{Tr}\left[\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*}\right]$
η_{GLRT1}	Non-GLRT statistics based on $\eta_{\text{GLRT,kn}}$, but with estimated $\mathbf{R} = \hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*}$	$\text{Tr}\left[(\hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*})^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*}\right]$

Table 1.1: Synchronization statistics studied in this thesis

1.5 Large system analysis using random matrix methods

1.5.1 Introduction to the problem

The second part of the PhD, treated in chapter 3, addresses the behavior of a classical multi-antenna GLRT that is able to detect the presence of a known signal corrupted by a multi-path propagation channel and by an additive temporally white Gaussian noise with unknown spatial covariance matrix. The chapter is focused on the case where the number of sensors M and possibly also the number of paths L is large, and of the same order of magnitude as the sample size N. This context is modeled by two large system asymptotic regimes. The first is $M \to +\infty$, $N \to +\infty$ in such a way that $M/N \to c$ for $c \in (0, +\infty)$. The second asymptotic regime under study is $N, M, L \to +\infty$, in such a way that $M/N \to c$ for $c \in (0, +\infty)$, and $L/N \to d$ for $d \in (0, +\infty)$. The purpose of this chapter is to study the behavior of a GLRT statistics in these regimes, and to show that the corresponding theoretical analysis allows one to accurately predict the performance of the test when M,N and possibly L are of the same order of magnitude.

1.5.2 Why large system approximation?

To see that the problem studied in chapter 3 is not only a theoretical consideration, consider as an example today's multi-antenna systems. The sample size that can in practice be used to perform statistical inference cannot be much larger than the signal dimension, which corresponds to the number of receive antennas.

We therefore distinguish between two kinds of limiting results: The so-called *classical limiting problem*, where only one of the dimensions is assumed to be large, and the *large dimensional limiting problem*, where several dimensions are assumed to go to infinity. It can be expected that classical limiting analysis performs poorly when several dimensions are of the same order of magnitude. One of the purposes of this study is therefore to show that classical limiting results are no longer valid for large dimensional systems.

A possible way of approaching the problem of asymptotic analysis in high dimension is random matrix theory (RMT). The underlying motivation for using this framework is the non-obvious behavior of large random matrices. As an introductory example, let us study the empirical covariance matrix of a high-dimensional observation matrix. Given N $(M \times 1)$ observation vectors $(\mathbf{y}_n)_{n=1,\ldots,N}$, we want to study their empirical covariance matrix, given by

$$\mathbf{R}_N = \frac{1}{N} \sum_{i=1}^{N} \mathbf{y}_i \mathbf{y}_i^*. \tag{1.22}$$

If M is fixed as $N \to \infty$, $\mathbf{R}_N \to \mathbf{R} = \mathbb{E}[|\mathbf{y}_1 \mathbf{y}_1^*|]$, and $\|\mathbf{R} - \mathbf{R}_N\| \to \mathbf{0}$ for any matrix norm [29]. We have that the empirical covariance matrix $\mathbf{R}_N$ is a consistent estimator of the population covariance matrix $\mathbf{R}$. However, in many practical applications, the number of available observations N is of the same order of magnitude as M. In this case, $\|\mathbf{R} - \mathbf{R}_N\|$ can be far from zero even when N is large. For example, if $M > N$, the sample covariance matrix is rank-deficient (while $\mathbf{R}$ can be of full rank), and is therefore not a good approximation of $\mathbf{R}$. If N and M are both large compared to 1, but of the same order of magnitude, it has been shown that the empirical eigenvalue distribution of $\mathbf{R}_N$ is different from the eigenvalue distribution of $\mathbf{R}$. For example, if $\mathbf{R} = \sigma^2 \mathbf{I}_M$, the histogram of the eigenvalues of $\mathbf{R}_N$ tends to approach the probability density of the so-called Marcenko-Pastur distribution [30]. This is in contrast with the case where $N >> M$, where the eigenvalues of $\mathbf{R}_N$ are concentrated around σ^2.

What is infinity? As mentioned before, the asymptotic analysis is done for the case where the dimensions N, K (and possibly also L) $\to \infty$. However, the results obtained with these methods are applicable even in the small dimensional case. In other words: For theoretical convenience, the dimensions are assumed large, but it can be shown by simulation that the results are valid even for small dimensions.

1.5.3 Preliminaries

In chapter 3, we perform an asymptotic analysis to study the limit distribution of the GLRT synchronization statistics studied in chapter 2. We study the case of a single transmit antenna ($K = 1$) sending a length N known synchronization sequence through a frequency selective deterministic channel $\mathbf{H}$ with L paths. We remind the reader that a sufficient statistic is the log-likelihood ratio $\log(\eta'_{\text{GLRT}})/N$, which we here and in chapter 3 call η_N, given by

$$\eta_N = -\log\det\left[\mathbf{I}_L - \frac{\mathbf{S}\mathbf{Y}^*}{N}\left(\frac{\mathbf{Y}\mathbf{Y}^*}{N}\right)^{-1}\frac{\mathbf{Y}\mathbf{S}^*}{N}\left(\frac{\mathbf{S}\mathbf{S}^*}{N}\right)^{-1}\right], \tag{1.23}$$

Before starting the derivation of the asymptotic, we remark that it is possible to assume without restriction that $\frac{\mathbf{S}\mathbf{S}^*}{N} = \mathbf{I}_L$ and that $\mathbb{E}(\mathbf{v}_n \mathbf{v}_n^*) = \sigma^2 \mathbf{I}$, i.e. $\tilde{\mathbf{R}}$ is reduced to the identity matrix. If this is not the case, we denote by $\tilde{\mathbf{S}}$ the matrix

$$\tilde{\mathbf{S}} = \left(\frac{\mathbf{S}\mathbf{S}^*}{N}\right)^{-1/2}\mathbf{S} \tag{1.24}$$

and by $\tilde{\mathbf{Y}}$ and $\tilde{\mathbf{V}}$ the whitened observation and noise matrices

$$\begin{aligned}\tilde{\mathbf{Y}} =& \tilde{\mathbf{R}}^{-1/2}\mathbf{Y},\\ \tilde{\mathbf{V}} =& \tilde{\mathbf{R}}^{-1/2}\mathbf{V}\end{aligned} \tag{1.25}$$

It is clear that $\frac{\tilde{\mathbf{S}}\tilde{\mathbf{S}}^*}{N} = \mathbf{I}_L$ and that $\mathbb{E}(\tilde{\mathbf{v}}_n\tilde{\mathbf{v}}_n^*) = \sigma^2\mathbf{I}$. Moreover, under H_0, it holds that $\tilde{\mathbf{Y}} = \tilde{\mathbf{V}}$, while under H_1, $\tilde{\mathbf{Y}} = \tilde{\mathbf{H}}\tilde{\mathbf{S}} + \tilde{\mathbf{V}}$ where the channel matrix $\tilde{\mathbf{H}}$ is defined by

$$\tilde{\mathbf{H}} = \tilde{\mathbf{R}}^{-1/2}\,\mathbf{H}\,(\mathbf{S}\mathbf{S}^*/N)^{1/2} \tag{1.26}$$

Finally, it holds that the statistics η_N can also be written as

$$\eta_N = -\log\det\left[\mathbf{I}_L - \frac{\tilde{\mathbf{S}}\tilde{\mathbf{Y}}^*}{N}\left(\frac{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^*}{N}\right)^{-1}\frac{\tilde{\mathbf{Y}}\tilde{\mathbf{S}}^*}{N}\right] \tag{1.27}$$

This shows that it is possible to replace $\mathbf{S}$, $\tilde{\mathbf{R}}$ and $\mathbf{H}$ by $\tilde{\mathbf{S}}$, $\mathbf{I}$, and $\tilde{\mathbf{H}}$ without modifying the value of the statistics η_N. The results will be valid replacing $\mathbf{H}$ with $\tilde{\mathbf{H}}$, and we can without restriction make the calculations with the assumption that

$$\frac{\mathbf{S}\mathbf{S}^*}{N} = \mathbf{I}_L,\ \tilde{\mathbf{R}} = \mathbf{I}_M \tag{1.28}$$

A second preliminary step is introducing auxiliary variables that we call $\mathbf{V}_1$ and $\mathbf{V}_2$. We denote by $\mathbf{W}$ a $(N-L)\times N$ matrix for which the matrix $\mathbf{\Theta} = (\mathbf{W}^T, \frac{\mathbf{S}^T}{\sqrt{N}})^T$ is unitary and define the $M\times(N-L)$ and $M\times L$ matrices $\mathbf{V}_1$ and $\mathbf{V}_2$ by

$$(\mathbf{V}_1, \mathbf{V}_2) = \mathbf{V}\mathbf{\Theta}^* = (\mathbf{V}\mathbf{W}^*, \mathbf{V}\frac{\mathbf{S}^*}{\sqrt{N}}) \tag{1.29}$$

It is clear that $\mathbf{V}_1$ and $\mathbf{V}_2$ are complex Gaussian random matrices with independent identically distributed $\mathcal{N}_{\mathbb{C}}(0,\sigma^2)$ entries, and that the entries of $\mathbf{V}_1$ and $\mathbf{V}_2$ are mutually independent. We notice that since $N > M + L$, the matrix $\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}$ is invertible almost surely. We can now express the statistics η_N in terms of $\mathbf{V}_1$ and $\mathbf{V}_2$. Under H_0, it is shown in chapter 3 that the statistics can be written as

$$\eta_N = \log\det\left(\mathbf{I}_L + \mathbf{V}_2^*/\sqrt{N}\ (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\ \mathbf{V}_2/\sqrt{N}\right) \tag{1.30}$$

and similarly under H_1 as

$$\eta_N = \log\det\left(\mathbf{I}_L + \Big(\mathbf{V}_2 + \mathbf{H}\Big)^*/\sqrt{N}\ (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\ \Big(\mathbf{V}_2 + \mathbf{H}\Big)/\sqrt{N}\right). \tag{1.31}$$

The key point here is the independence between $\mathbf{V}_1$ and $\mathbf{V}_2$, which can be exploited to simplify the calculations of the limit distributions.

Note that the asymptotic of this kind of statistics have been studied in the past, but in the classical regime, defined by

$$\textit{Asymptotic regime (a)}\text{:}\ \begin{cases} L, M \text{ is fixed} \\ N \to \infty \end{cases}$$

When the size of M and L increase, the assumptions that $M << N$ and/or $L << N$ are no longer valid, which is the case for several practical applications. Therefore, we consider two new asymptotic regimes:

$$\textit{Asymptotic regime (b)}: \begin{cases} L \text{ is fixed} \\ N, M \to \infty \\ M/N = c_N \to c,\ 0 < c < 1 \end{cases}$$

$$\textit{Asymptotic regime (c)}: \begin{cases} N, M, L \to \infty \\ M/N = c_N \to c,\ 0 < c < 1 \\ L/N = d_N \to d,\ 0 < d < 1 \\ c + d < 1 \end{cases}$$

1.5.4 Example of asymptotic analysis: Expected value of η_N under H_0

To get an idea of the differences between the asymptotic analyses in the asymptotic regimes, we will present as an example the asymptotic analysis to obtain the expected value of η_N under H_0, for the asymptotic regimes (a) and (b).

Let us start with the analysis in the classical regime (a). We use (1.30) and remark that when $N \to +\infty$ and M and L remain fixed, the matrices $\mathbf{V}_1\mathbf{V}_1^*/N$ and $\frac{1}{N}\mathbf{V}_2^*\left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\mathbf{V}_2$ converge a.s. towards $\sigma^2\mathbf{I}$ and the zero matrix respectively. Moreover,

$$\frac{1}{N}\mathbf{V}_2^*\left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\mathbf{V}_2 = \frac{1}{\sigma^2}\mathbf{V}_2^*\mathbf{V}_2/N + o_P(\frac{1}{N}) \tag{1.32}$$

and a standard second order expansion of η_N leads to

$$\eta_N = \frac{1}{\sigma^2}\mathrm{Tr}\left(\mathbf{V}_2^*\mathbf{V}_2/N\right) + o_P(\frac{1}{N}) \tag{1.33}$$

This implies immediately that the limit distribution of $N\,\eta_N$ is a chi-squared distribution with $2ML$ degrees of freedom. Informally, this implies that $\mathbb{E}(\eta_N) \simeq L\frac{M}{N}$ and $\mathrm{Var}(\eta_N) \simeq \frac{L}{N}\frac{M}{N}$.

The analysis of η_N in the asymptotic regime (b) differs deeply from the analysis in the standard asymptotic regime (a). In particular, it is no longer true that the empirical covariance matrix $\mathbf{V}_1\mathbf{V}_1^*/N$ converges in the spectral norm sense towards $\sigma^2\mathbf{I}$. This is due to the fact that the number of entries of this $M \times M$ matrix is of the same order of magnitude as the number of available scalar observations. We also note that for any deterministic $M \times M$ matrix $\mathbf{A}$, the diagonal entries of the $L \times L$ matrix $\frac{1}{N}\mathbf{V}_2^*\mathbf{A}\mathbf{V}_2$ converge towards 0 when $N \to +\infty$ and M remains fixed, while this does not hold when M and N are of the same order of magnitude.

We now want to calculate the asymptotic expected value of η_N in the asymptotic region (b), and show that

$$\eta_N - L \log\left(\frac{1}{1-c_N}\right) \to 0 \, a.s. \tag{1.34}$$

First, let us denote by $\mathbf{F}_N$ the $L \times L$ matrix

$$\mathbf{F}_N = \mathbf{V}_2^*/\sqrt{N} \; (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1} \; \mathbf{V}_2/\sqrt{N}. \tag{1.35}$$

and remark that under H_0, (1.30) can be written as

$$\eta_N = \log\det\left(\mathbf{I}_L + \mathbf{F}_N\right) \tag{1.36}$$

As L does not increase with M and N, it is sufficient to establish that

$$\mathbf{F}_N - \frac{c_N}{1-c_N}\,\mathbf{I}_L \to 0 \; a.s. \tag{1.37}$$

Our approach is based on the observation that if $\mathbf{A}_N$ is a $M \times M$ deterministic Hermitian matrix verifying $\sup_N \|\mathbf{A}_N\| < a < +\infty$, then

$$\mathbb{E}_{\mathbf{V}_2}\Big| \left(\mathbf{V}_2^*/\sqrt{N}\,\mathbf{A}_N\mathbf{V}_2/\sqrt{N}\right)_{k,l} - \frac{\sigma^2}{N}\mathrm{Tr}(\mathbf{A}_N)\,\delta(k-l)\Big|^4 \; \leq \frac{C(a)}{N^2} \tag{1.38}$$

where $C(a)$ is a constant term depending on a, and where $\mathbb{E}_{\mathbf{V}_2}$ represents the mathematical expectation operator w.r.t. $\mathbf{V}_2$. This is a consequence of Proposition 4 in Appendix 3.7. Assume for the moment that there exists a deterministic constant a such that

$$\| \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1} \| \leq a \tag{1.39}$$

for each N greater than a non random integer N_0. Then, as $\mathbf{V}_1$ and $\mathbf{V}_2$ are independent, it is possible to use (4.38) for $\mathbf{A}_N = \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}$ and to take the mathematical expectation w.r.t. $\mathbf{V}_1$ of (4.38) to obtain that

$$\mathbb{E}\left|\left(\mathbf{F}_N\right)_{k,l} - \frac{\sigma^2}{N}\mathrm{Tr}\left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\,\delta(k-l)\right|^4 \leq \frac{C(a)}{N^2} \tag{1.40}$$

for each $N > N_0$, and, using the Borel-Cantelli lemma, that

$$\mathbf{F}_N - \frac{\sigma^2}{N}\mathrm{Tr}\left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\,\mathbf{I}_L \to 0 \; a.s. \tag{1.41}$$

To conclude, we use known results related to the almost sure convergence of the eigenvalue distribution of matrix $\mathbf{V}_1\mathbf{V}_1^*/N$ towards the so-called Marcenko-Pastur distribution (see Eq. (3.77) in Appendix 3.7) which imply that

$$\frac{1}{N}\,\mathrm{Tr}\Big(\mathbf{V}_1\mathbf{V}_1^*/N\Big)^{-1} - \frac{c_N}{\sigma^2(1-c_N)} \to 0 \tag{1.42}$$

almost surely. This, in conjunction with (4.41), leads to (1.37) and eventually to (1.34).

However, there does not exist a deterministic constant a satisfying (3.36) for each N greater than a non-random integer. In order to solve this issue, it is sufficient to replace matrix $(\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}$ by a convenient regularized version. The details of the regularization will be given in chapter 3

1.5.5 Limit distributions

The main results of this work are the limit distributions summarized in Table 1.2. Note that the limit distribution for region (c) needs to establish a central limit theorem for linear statistics of the eigenvalues of large non-zero-mean F-matrices, a difficult task that has not been addressed in this work. Instead, in this work, we have proposed an approximation which will be described in detail in chapter 3.

Asymptotic regime	**Distribution under** H_0	**Distribution under** H_1
(a) Classical, $N \to \infty$	$\eta_N \sim \frac{1}{N}\chi^2_{2ML}$ $\Big(\mathbb{E}[\eta_N] = Lc_N$ $\mathrm{Var}[\eta_N] = Lc_N \cdot \frac{1}{N}\Big)$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\Big(\log\det\Big(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^*}{\sigma^2}\Big), \frac{\kappa_1}{N}\Big)$
(b) Proposed, $M, N \to \infty$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\Big(L\log\frac{1}{1-c_N}, \frac{Lc_N}{1-c_N}\cdot\frac{1}{N}\Big)$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\Big(L\log\frac{1}{1-c_N} + \log\det\Big(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^*}{\sigma^2}\Big),$ $\frac{\kappa_1}{N} + \frac{Lc_N}{1-c_N}\cdot\frac{1}{N}\Big)$
(c) Proposed, $L, M, N \to \infty$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\Big(\tilde{\eta}_N, \tilde{\delta}_N\Big)$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\Big(\tilde{\eta}_N + \log\det\Big(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^*}{\sigma^2}\Big), \frac{\kappa_1}{N} + \tilde{\delta}_N\Big)$

Table 1.2: Asymptotic distributions of η_N for the three asymptotic regimes, under H_0 *and* H_1

Chapter 2

Parameter optimization for time synchronization of multi-antenna systems

2.1 Introduction

Two decades ago MIMO systems, which use multiple antennas at both transmitter and receiver, were developed to increase the throughput (bit rate) and reliability of communications over fading channels through spatial multiplexing [31, 32] and space-time coding (STC) [33, 34] at transmission, without the need of increasing the receiver bandwidth. This powerful technology has been adopted in several wireless standards such as IEEE 802.11n, IEEE 802.16 [35], LTE [36] or LTE-Advanced [37] in particular. Nevertheless, as wireless spectrum is an expensive resource, increasing network capacity without requiring additional bandwidth is a great challenge for wireless networks.

This has motivated the development of Interference Mitigation (IM) techniques at both the transmitter and the receiver, allowing several MIMO links to share the same spectral resources without impacting the transmission quality of each link. IM techniques at transmission require informations at transmission about the propagation channels of the links sharing the same resources. On the contrary, IM at reception may also be implemented for MIMO schemes which do not require any information at transmission about the propagation channels. This has motivated the development of Interference Cancellation techniques at reception for MIMO systems using the V-BLAST scheme [38] or STC ([39, 40] and references herein) in particular.

However, in order to be efficient, all these MIMO schemes require a preliminary step of time and frequency synchronization which has to be also robust to interferences. Time and frequency synchronization of MIMO systems have been strongly studied in the last fifteen years, mainly in the context of direct-sequence coded division multiple access (DS-CDMA) and orthogonal frequency division multiplex (OFDM) links. Both coarse and fine time synchronization jointly with frequency offset estimation and compensation have been analyzed, and many techniques have been proposed either for time-frequency synchronization [1–4] or time synchronization only [5–9]. Nevertheless, most of these techniques assume an absence of interference. The scarce papers of the literature dealing with MIMO synchronization in the presence of interference correspond to [2, 7–9]. However, [2] and [7] only consider the problem of MIMO synchronization in the presence of multi-users interference (MUI), and [8] seems to be the only paper dealing with MIMO synchronization in the presence of interference of any kind, such as hostile jammers. In [8], several statistics are proposed for time synchronization for both flat fading and frequency selective fading channels. Despite the numerous existing statistics for MIMO time synchronization, many important questions about their optimality, performance and complexity have arisen.

First of all, none of the statistics developed for MIMO systems in the absence of interference [1, 10–17] [3–6, 18–27] has been developed through a GLRT approach in the general case of a SC non DS-CDMA link and arbitrary potentially non-orthogonal sequences. Although it is well-known [28] that, contrary to likelihood ratio test (LRT) statistics, GLRT statistics are suboptimal from a detection point of view, one may wonder if a GLRT statistics may be better than existing ones in this context. To adress this question, we derive in this paper the GLRT statistics for absence of interference, which assumes, Gaussian, temporally and spatially white noise. In the following, this statistics will be called $\eta_{\mathrm{GLRT,we}}$. We will point out the links between $\eta_{\mathrm{GLRT,we}}$ and the statistics in the state of the art, and investigate its performance in absence of interference through numerical simulation.

While the GLRT statistics in presence of interference has already been presented [8], questions regarding its practical usefulness arise. In [8], two statistics for flat fading channels, robust to interferences, are derived from an MMSE and a GLRT approach respectively. The GLRT statistics, called η_{GLRT} in the following, assumes unknown, Gaussian, spatially colored and temporally white total noise. We note that this GLRT statistics may be very costly to implement, for a large number of antennas in particular, since for a $(K \times M)$ MIMO system, it requires both an $(M \times M)$ matrix inversion and an $(M \times M)$ or $(K \times K)$ determinant computation at each tested sample position. An alternative to η_{GLRT} is the MMSE statistics proposed in [8]. However, the MMSE statistics is shown in this paper to be very sensitive to training sequence

correlation, which may limit its practical use in this context.

In order to optimize the performance-complexity trade-off, we start by the analysis in the case of fixed number of antennas, and propose several ways to decrease the complexity of η_{GLRT} while retaining its performance. The first way concerns MIMO systems with $K = 2$ transmit antennas and consists in computing explicitly the determinant appearing in the formulation of η_{GLRT}, allowing in particular a direct comparison with both the MMSE test proposed in [8] and SIMO statistics [41, 42]. As the direct computation of the determinant is complicated for $K > 2$, a second proposition to decrease synchronization complexity consists in introducing two new low-complexity MIMO statistics which are robust to interference. These two new statistics, η_{GLRT0} and η_{GLRT1}, correspond to two estimates of the GLRT statistics in known, Gaussian, spatially correlated and temporally white total noise, called $\eta_{\mathrm{GLRT,kn}}$. For stationary interference, a third way of decreasing the complexity of the statistics is to compute and inverse at a lower rate the data correlation matrix appearing in the expressions of the statistics by computing it on an observation interval greater than the synchronization sequence length. Finally, we want to optimize the number of transmit antennas used for synchronization for a given number of receive antennas and for given kinds of propagation channels. Note that such a problem has been preliminarily investigated in [6] in the DS-CDMA context and in [43] for precoded synchronization schemes. The goal is to enlighten in particular the conditions under which it becomes sub-optimal to implement MIMO synchronization with respect to SIMO synchronization [41, 42], without and with interference.

In order to corroborate our results, we perform numerical simulations. For the sake of comparison, we introduce some of the statistics in the literature for the absence of interference, and show that the gain in performance for using the statistics optimized for this scenario is not very large. In both presence and absence of interference, we show that our proposed statistics perform well compared with η_{GLRT}. We also show that the computation rate decrease does not decrease the performance by too much, and that its complexity gains are significant, especially for large values of K, M. We also show by simulation that the optimal number of transmit antennas is dependent on the number of antennas under study.

This chapter is organized as follows: Section 2.2 is devoted to the current state of the art in time synchronization with interference, with an introduction to the MMSE and GLRT statistics. In section 2.3, we assume that the number of transmit and receive antennas is fixed, and present ways to reduce the complexity while retaining the performance. We derive the simplified form of η_{GLRT}, and the expressions of the two new low-complexity non-GLRT statistics η_{GLRT0} and

η_{GLRT1}. We also perform a comparative performance analysis of the above-mentioned statistics for different system parameters. We then go on to describe how to perform the computational rate decrease, and show its performance by simulation. We finish the section with a complexity analysis of the presented statistics for different kinds of systems. In Section 2.4, we describe how to optimize the number of transmit antennas, and show by simulation the optimal number of antennas for absence and presence of interference, for different channel types. Section 2.5 concludes the chapter.

2.2 Statistics optimized for the presence of interference

In this section, we briefly recall the GLRT and MMSE statistics introduced in [8].

2.2.1 η_{GLRT}: GLRT synchronization statistics

To model the presence of interference, we assume zero-mean, i.i.d, temporally white and spatially colored, circular, Gaussian samples $\mathbf{v}(k)$ whose covariance matrix is given by $\mathbf{R}$. If both $\mathbf{R}$ and the channel matrix $\mathbf{H}$ are unknown, the GLRT is given by equation (1.18), and gives rise to the GLRT statistics η_{GLRT}. The derivation of this synchronization statistics is summarized in Appendix 2.8, but has also been shown in [8] to be

$$\eta_{\text{GLRT}} = \det[\mathbf{I}_N - \hat{\mathbf{P}}_s\hat{\mathbf{P}}_y]^{-N} \tag{2.1}$$

where $\hat{\mathbf{P}}_s$ and $\hat{\mathbf{P}}_y$ are $(N \times N)$ matrices corresponding to the projection operators onto the row spaces spanned by $\mathbf{S}$ and $\mathbf{Y}$ respectively, defined by $\hat{\mathbf{P}}_s \triangleq \mathbf{S}^*(\mathbf{S}\mathbf{S}^*)^{-1}\mathbf{S}$ and $\hat{\mathbf{P}}_y \triangleq \mathbf{Y}^*(\mathbf{Y}\mathbf{Y}^*)^{-1}\mathbf{Y}$.

From (2.1), it is clear that a sufficient statistic is given by the log-likelihood

$$\eta_{\text{GLRT}} = -N \log\det[\mathbf{I}_N - \hat{\mathbf{P}}_s\hat{\mathbf{P}}_y]. \tag{2.2}$$

It is straighforward to show that (2.2) can also be written as

$$\begin{aligned}\eta_{\text{GLRT}} &= -N \log\det[\mathbf{I}_M - \hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^*] \\ &= -N \log\det[\mathbf{I}_K - \mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys}]\end{aligned} \tag{2.3}$$

Expression (2.3), not presented in [8], requires that $\mathbf{R}_{ss}$ is invertible, which is assumed in the following. This is only possible if $K \leq N$, which is then assumed in the following. Under these assumptions, expressions (2.2), (2.3) show that, at each tested sample position, η_{GLRT} requires the computation of at least a $(M \times M)$ matrix inversion, $\hat{\mathbf{R}}_{yy}^{-1}$, and the determinant of a $(P \times P)$

matrix where $P = \mathrm{Inf}(N, M, K)$, which may be prohibitive for large N and a large number of antennas.

In the particular case of a SIMO system ($K = 1$), $\hat{\mathbf{R}}_{ys}$ reduces to the vector $\hat{\mathbf{r}}_{ys_1}$, $\mathbf{R}_{ss}$ reduces to the scalar π_{s_1} and we deduce from (2.3) that a sufficient statistic for the η_{GLRT} is given by

$$\eta_{\mathrm{GLRT,SIMO}} = \frac{\hat{\mathbf{r}}_{ys_1}^{*} \hat{\mathbf{R}}_{yy}^{-1} \hat{\mathbf{r}}_{ys_1}}{\pi_{s_1}}. \tag{2.4}$$

2.2.2 η_{MMSE}: MMSE synchronization statistics

MMSE time synchronization consists in finding the sample position which minimizes the least square (LS) error between the known sampled vectors $\mathbf{s}(k)$ and their LS estimation from a spatial filtering of the data $\mathbf{y}(k)(1 \leq k \leq N)$. After elementary computations, we obtain a sufficient statistic for MMSE synchronization, given by [8]

$$\eta_{\mathrm{MMSE}} = \frac{\mathrm{Tr}(\hat{\mathbf{R}}_{ys}^{*} \hat{\mathbf{R}}_{yy}^{-1} \hat{\mathbf{R}}_{ys})}{\mathrm{Tr}(\mathbf{R}_{ss})} = \frac{\sum_{i=1}^{K} \hat{\mathbf{r}}_{ys_i}^{*} \hat{\mathbf{R}}_{yy}^{-1} \hat{\mathbf{r}}_{ys_i}}{\sum_{m=1}^{K} \pi_{s_m}} \tag{2.5}$$

Comparing (2.5) to (2.4), we deduce that, to within a constant, η_{MMSE} corresponds to the weighted sum of K SIMO GLRT statistics, each associated with a different transmit antenna. The computation of the MMSE statistics requires a $(M \times M)$ matrix inversion at each tested sample position but no determinant computation, which is less complex than the η_{GLRT} computation. For SIMO links ($K = 1$), η_{MMSE} and η_{GLRT} coincide, but this is a priori no longer true for $K > 1$ as will be shown in the next section.

2.3 Optimization of the synchronization for fixed K, M

2.3.1 Direct expression of η_{GLRT} for K=2

For $K = 2$, the determinant (2.3) can be easily computed and the result allows a direct comparison with the MMSE statistics (2.5).

For this purpose, we define the following vectors $\mathbf{u}_1 \triangleq \hat{\mathbf{R}}_{yy}^{-1/2} \hat{\mathbf{r}}_{ys_1}$, $\mathbf{u}_2 \triangleq \hat{\mathbf{R}}_{yy}^{-1/2} \hat{\mathbf{r}}_{ys_2}$, $\mathbf{v}_1 \triangleq \hat{\mathbf{R}}_{yy}^{-1/2} \hat{\mathbf{h}}_1$ and $\mathbf{v}_2 \triangleq \hat{\mathbf{R}}_{yy}^{-1/2} \hat{\mathbf{h}}_2$, where $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2] = \hat{\mathbf{R}}_{ys} \mathbf{R}_{ss}^{-1}$ is the ML estimate of $\mathbf{H}$ under H_1. Inserting these vectors into (2.2) for $K = 2$, we obtain

$$\eta_{\mathrm{GLRT,s}} = \det(\mathbf{I}_M - \mathbf{u}_1 \mathbf{v}_1^{*} - \mathbf{u}_2 \mathbf{v}_2^{*})^N \triangleq \det(\mathbf{I}_M - \mathbf{B})^N \tag{2.6}$$

where $\mathbf{B} \triangleq \mathbf{u}_1 \mathbf{v}_1^{*} - \mathbf{u}_2 \mathbf{v}_2^{*}$ is a rank 2 matrix such that $\mathrm{span}(\mathbf{B}) = \mathrm{span}(\mathbf{u}_1, \mathbf{u}_2)$. It is well-known that $\det[\mathbf{I}_M - \mathbf{B}] = (1-\lambda_1)(1-\lambda_2) = 1 - S + \Pi$, where λ_1 and λ_2 are the two non-zero eigenvalues

of $\mathbf{B}$ and where $S \triangleq \lambda_1 + \lambda_2$ and $\Pi \triangleq \lambda_1\lambda_2$. A straightforward computation of λ_1 and λ_2 from $\mathbf{B}$ gives

$$\begin{aligned} 1 - S + \Pi =& 1 - (\mathbf{v}_1^*\mathbf{u}_1 + \mathbf{v}_2^*\mathbf{u}_2) \\ &+ (\mathbf{v}_1^*\mathbf{u}_1)(\mathbf{v}_2^*\mathbf{u}_2) - (\mathbf{v}_1^*\mathbf{u}_2)(\mathbf{v}_2^*\mathbf{u}_1) \\ \triangleq& 1 - \eta_{\text{GLRT}} \end{aligned} \tag{2.7}$$

Using (2.7) into (2.6) and the definitions of $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{v}_1$ and $\mathbf{v}_2$, we deduce that a sufficient statistic for (2.6) is given by

$$\begin{aligned} \eta_{\text{GLRT,s}} =& \hat{\mathbf{h}}_1^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{r}}_{ys_1} - (\hat{\mathbf{h}}_1^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{r}}_{ys_1})(\hat{\mathbf{h}}_2^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{r}}_{ys_2}) \\ &+ \hat{\mathbf{h}}_2^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{r}}_{ys_2} + (\hat{\mathbf{h}}_1^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{r}}_{ys_2})(\hat{\mathbf{h}}_2^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{r}}_{ys_1}) \end{aligned} \tag{2.8}$$

For orthogonal sequences, we know that $\hat{\mathbf{h}}_i = \pi_{s_{ii}}^{-1}\hat{\mathbf{r}}_{ys_i} (1 \leq i \leq 2)$. Inserting this into (2.8), we deduce that for orthogonal sequences and $K = 2$, η_{GLRT} becomes

$$\begin{aligned} \eta_{\text{GLRT,s}} =& \frac{\hat{\mathbf{r}}_{ys_1}^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{r}}_{ys_1}}{\pi_{s_1}} + \frac{\hat{\mathbf{r}}_{ys_2}^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{r}}_{ys_2}}{\pi_{s_2}} \\ &- \frac{\hat{\mathbf{r}}_{ys_1}^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{r}}_{ys_1}\hat{\mathbf{r}}_{ys_2}^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{r}}_{ys_2}}{\pi_{s_1}\pi_{s_2}} + \frac{\|\hat{\mathbf{r}}_{ys_1}^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{r}}_{ys_2}\|^2}{\pi_{s_1}\pi_{s_2}}. \end{aligned} \tag{2.9}$$

For $K = 1$ (SIMO case), this further reduces to (2.4). Expressions (2.9) and (2.4) show that η_{GLRT} for $K = 2$ equals to the sum of 2 SIMO synchronization statistics, and additional terms that depend on the correlation between the training sequences. Moreover, comparing (2.5) for $K = 2$ with (2.8) and (2.9), we deduce that for MIMO links ($K > 1$), η_{MMSE} no longer corresponds to η_{GLRT}, not even for orthogonal synchronization sequences having the same power, which was not obvious a priori.

2.3.2 New low-complexity statistics

The direct computation of the determinant (2.3) is not so straightforward for $K > 2$ while the MMSE statistics (2.5) has been shown in [8] to become sub-optimal for non-orthogonal synchronization sequences. In this context, a second way of decreasing the complexity of η_{GLRT} for arbitrary values of K while trying to keep its performance is to develop alternative statistics. To this aim, it seems natural to think that non-GLRT statistics corresponding to good estimates $\eta_{\text{GLRT,kn}}$, the GLRT statistics in known total noise, have good chances of approaching the performance of η_{GLRT}. For this reason, in this section, we introduce $\eta_{\text{GLRT,kn}}$ and propose two new low-complexity statistics corresponding to two different estimates of this test.

Introduction to $\eta_{\text{GLRT,kn}}$

The statistics $\eta_{\text{GLRT,kn}}$ is obtained by considering expression (1.9) and assuming an unknown channel matrix $\mathbf{H}$ and a zero-mean, i.i.d, circular, Gaussian, temporally white and spatially colored total noise whose covariance matrix $\mathbf{R}$ is assumed to be known. Replacing in (1.9) $\mathbf{H}$ by its ML estimate $\hat{\mathbf{H}} = \hat{\mathbf{R}}_{ys}\hat{\mathbf{R}}_{ss}^{-1}$, valid for any K, generates $\eta_{\text{GLRT,kn}}$. It is shown in Appendix 2.9 that a sufficient statistic is given by

$$\eta_{\text{GLRT,kn}} = \text{Tr}\left[\mathbf{R}_{ss}^{-1}\ \hat{\mathbf{R}}_{ys}^{*}\ \mathbf{R}^{-1}\ \hat{\mathbf{R}}_{ys}\right]. \tag{2.10}$$

In the particular case of K orthogonal synchronization sequences, expression (2.10) reduces to

$$\eta_{\text{GLRT,kn}} = \sum_{i=1}^{K} \frac{\hat{\mathbf{r}}_{ys_i}^{*}\mathbf{R}^{-1}\hat{\mathbf{r}}_{ys_i}}{\pi_{s_{ii}}}. \tag{2.11}$$

A similar problem has been considered in [44] in the context of MIMO radar. Note that $\eta_{\text{GLRT,kn}}$ does not require any determinant computation, and corresponds, for orthogonal sequences, to the sum of K SIMO $\eta_{\text{GLRT,kn}}$ statistics, each associated to a different transmit antenna. Unfortunately, it cannot be used in practice since $\mathbf{R}$ is unknown. It can however be estimated by replacing $\mathbf{R}$ by an estimate, which is done in the following sections.

η_{GLRT0}: Estimated $\eta_{\text{GLRT,kn}}$ statistics under H_0

A first possibility to derive from (2.10) a low-complexity statistics useful in practice is to replace in (2.10) the matrix $\mathbf{R}$ by its ML estimate under H_0, $\hat{\mathbf{R}}_0$. It is easy to show that $\hat{\mathbf{R}}_0 = \hat{\mathbf{R}}_{yy}$, which gives rise to the estimated $\eta_{\text{GLRT,kn}}$ statistics under H_0, defined by

$$\eta_{\text{GLRT0}} = \text{Tr}\left[\mathbf{R}_{ss}^{-1}\ \hat{\mathbf{R}}_{ys}^{*}\ \hat{\mathbf{R}}_{yy}^{-1}\ \hat{\mathbf{R}}_{ys}\right]. \tag{2.12}$$

In the particular case of K orthogonal synchronization sequences, expression (2.12) reduces to

$$\eta_{\text{GLRT0}} = \sum_{i=1}^{K} \frac{\hat{\mathbf{r}}_{ys_i}^{*}\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{r}}_{ys_i}}{\pi_{s_{ii}}}, \tag{2.13}$$

which corresponds, to within a constant, to (2.5). In this case, the MMSE statistics can be interpreted as an estimate of $\eta_{\text{GLRT,kn}}$ under H_0. Otherwise, the MMSE statistics has no link with η_{GLRT0}.

η_{GLRT1}: Estimated $\eta_{\text{GLRT,kn}}$ statistics under H_1

A second possibility to derive from (2.10) a low-complexity statistics useful in practice is to replace in (2.10) the matrix $\mathbf{R}$ by its ML estimate under H_1, $\hat{\mathbf{R}}_1$. It is easy to show that $\hat{\mathbf{R}}_1$ is

defined by

$$\hat{\mathbf{R}}_1 = \hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*} \tag{2.14}$$

In (2.14), the estimated contributions of the transmitted synchronization sequences have been removed from $\mathbf{y}$. This gives rise to the estimated $\eta_{\mathrm{GLRT,kn}}$ statistics under H_1, defined by

$$\eta_{\mathrm{GLRT1}} = \mathrm{Tr}\left[\mathbf{R}_{ss}^{-1}\ \hat{\mathbf{R}}_{ys}^{*}\ \hat{\mathbf{R}}_1^{-1}\ \hat{\mathbf{R}}_{ys}\right]. \tag{2.15}$$

In the particular case of K orthogonal synchronization sequences, expression (2.15) reduces to

$$\eta_{\mathrm{GLRT1}} = \sum_{i=1}^{K} \frac{\hat{\mathbf{r}}_{ys_i}^{*}\hat{\mathbf{R}}_1^{-1}\hat{\mathbf{r}}_{ys_i}}{\pi_{s_{ii}}}. \tag{2.16}$$

Comparative performance analysis of MIMO synchronization statistics

We present in this section a comparative performance analysis of the MIMO statistics introduced in section 2.3, both in the absence and in the presence of interference. This analysis allows in particular to enlighten the practical interest of the new statistics introduced in this paper (η_{GLRT0}, η_{GLRT1}) with respect to η_{GLRT}, which has been presented as best statistics in [8]. For this purpose we consider in this section $K \times M$ MIMO links for which the transmitting and the receiving antennas are omnidirectional.

Assumptions For all simulations, the false alarm rate is $\mathrm{P_{FA}} = 10^{-3}$ and the figures are built from 10^5 independent realizations.

- *Channel models:* Two kinds of channel matrices $\mathbf{H}$, corresponding to deterministic and random channel matrices, are considered.

  For the deterministic case, a possible model is a line-of-sight physical MIMO channel, with angles of departure θ_t and angles of arrival θ_r. Assuming that the antennas in both the transmitting and receiving end are spaced close to each other, the MIMO channel can be written as

  $$\mathbf{H} = ae^{-\frac{2\pi j d}{\lambda_c}}\mathbf{e}_r(\theta_r)\ \mathbf{e}_t^{H}(\theta_t)$$

  where a is the attenuation along the line-of-sight path, assumed to be the same for all antenna pairs, and d the distance between transmit antenna one and receive antenna one. In our simulations, we have $a = 1$. The vectors $\mathbf{e}_r(\theta_r)$ and $\mathbf{e}_t(\theta_t)$ are the spatial signatures for the receiver and transmitter respectively, defined by

  $$\begin{aligned}\mathbf{e}_r(\theta_r) =& [1, e^{-2\pi j\Delta_r\theta_r}, \ldots, e^{-2\pi j(N-1)\Delta_r\theta_r}]^T\\ \mathbf{e}_t(\theta_t) =& [1, e^{-2\pi j\Delta_t\theta_t}, \ldots, e^{-2\pi j(M-1)\Delta_t\theta_t}]^T\end{aligned}$$

where λ_c is the carrier wavelength, Δ_r the normalized receive antenna separation, and Δ_t the normalized transmit antenna separation. In the following, we assume that $\Delta_t\lambda_c$ and $\Delta_r\lambda_c$ are both 0.5.

In this work, we consider the scenario where the transmit antennas are assumed to be distributed in space or well-separated from each other. In this scenario, the columns of the channel matrix can be separated into K columns as $\mathbf{H} = [\mathbf{h}_1(\theta_1), \ldots, \mathbf{h}_K(\theta_K)]$. Each column $\mathbf{h}_k(\theta_k)$ $(1 \leq k \leq K)$ is the contribution of transmit antenna k on the receive antenna array, and is defined by

$$\mathbf{h}_k(\theta_k) = a_k e^{-\frac{2j\pi d_k}{\lambda_c}} \mathbf{e}_r(\theta_k). \tag{2.17}$$

a_k is the attenuation along path k, and d_k the distance between transmit antenna k and the first receive antenna. Again, we set a_k to 1. Note that each column of the channel matrix can alternatively be seen as a separate SIMO system on the receive antenna array.

To quantify the degree of separation of the transmit antennas in our system, we also define the spatial correlation coefficient α_{ij} between the channels $\mathbf{h}_i(\theta_i)$ and $\mathbf{h}_j(\theta_j)$ as

$$|\alpha_{ij}(\theta_i, \theta_j)|^2 = \frac{|(\mathbf{h}_i(\theta_i)^*\mathbf{h}_j(\theta_j))|^2}{(\mathbf{h}_i(\theta_i)^*\mathbf{h}_i(\theta_i))(\mathbf{h}_j(\theta_j)^*\mathbf{h}_j(\theta_j))} \tag{2.18}$$

and note that $0 \leq |\alpha_{ij}| \leq 1$.

For the random channel, the coefficients $H_{i,j}$ $(1 \leq i \leq M), (1 \leq j \leq K)$ of $\mathbf{H}$ are assumed to be zero-mean i.i.d circular Gaussian variables such that $\mathbb{E}[|H_{i,j}|^2] = 1$, which modelizes a Rayleigh flat fading model with a maximal diversity.

In all cases, the receiving array is assumed to be a uniformly spaced linear array of M antennas, whose interelement spacing is equal to half a wavelength.

- *Synchronization sequences:* Each synchronization sequence is composed of N samples. The synchronization sequences have the same power ($\pi_{s_{ii}} \triangleq \pi_s, 1 \leq i \leq K$) and are normalized such that the signal to thermal noise ratio per receive antenna, defined by SNR$\triangleq K\pi_s/\sigma^2$, may be arbitrarily chosen.

  To create orthogonal training sequences, we use Zadoff-Chu sequences. More specifically, the rows of the training sequence matrix $\mathbf{S}$ are chosen as cyclic shifts of a Zadoff-Chu sequence of length N [45]. Due to the autocorrelation properties of Zadoff-Chu sequences, we have that $\mathbf{S}\mathbf{S}^*/N = \mathbf{I}_K$ if $K \leq N$.

  To artificially create non-orthogonal sequences, the beginning of the synchronization sequence is set to be the same for all transmit antennas. For example, to create a

correlation of approximately 0.25 for $K = 2$ and $N = 32$,

$$\mathbf{S} = \begin{pmatrix} \mathbf{s}_{rand} & \mathbf{s}_{1,rand} \\ \mathbf{s}_{rand} & \mathbf{s}_{2,rand} \end{pmatrix} \tag{2.19}$$

where $\mathbf{s}_{rand}$ is a length $0.25N = 8$ random QPSK sequence, same for both antennas, and $\mathbf{s}_{1,rand}$ and $\mathbf{s}_{2,rand}$ are random QPSK sequences. Finally, to quantify the exact correlation, we define the temporal correlation between synchronization sequences from transmit antenna i and j by

$$|\rho_s[i,j]| \triangleq \sqrt{\frac{|\pi_{s_{ij}}|^2}{\pi_{s_{ii}}\pi_{s_{jj}}}} \tag{2.20}$$

In our simulations, if not otherwise stated, $N = 32$.

- *Total noise:* In the absence of interference, the total noise is assumed to be composed of a background noise such that

$$\mathbf{v}(k) = \mathbf{n}(k). \tag{2.21}$$

Here, $\mathbf{n}(k)$ is the sampled background noise vector, assumed to be zero-mean, Gaussian, SO circular, spatially and temporally white with a mean power per receive antenna equal to σ^2.

In the presence of interference, we also have the contribution of one single antenna interference such that

$$\mathbf{v}(k) = j_I(k)\mathbf{h}_I + \mathbf{n}(k). \tag{2.22}$$

where $\mathbf{h}_I$ the channel vector of the interference.

For the deterministic channel, the interference channel vector is defined by equation (2.17) as $\mathbf{h}_I(\theta_I)$, where θ_I is the direction of arrival of the interference on the receive antenna array.

For the random channel, the components $h_I[i]$ $(1 \leq i \leq M)$ are zero-mean i.i.d circular Gaussian variables verifying $\mathbb{E}[|h_I[i]|^2] = 1$.

$j_I(k)$ are the complex samples of the interference, assumed to be QPSK and such that $\pi_I \triangleq \mathbb{E}[|j_I(k)|^2]$ is the input mean power of the interference per antenna. The interference to noise ratio per receive antenna, defined by $\text{INR} \triangleq \pi_I/\sigma^2$, is chosen such that $\text{INR}_{dB} = \text{SNR}_{dB} + 15$ dB.

Synchronization statistics optimized for absence of interference Several statistical tests for time synchronization of MIMO links in the absence of interference have been proposed in the literature [1, 3–6, 10–27], mainly for DS-CDMA and OFDM links. Some of them may be

also used for non DS-CDMA SC links. To enrich the analysis of this subsection, we will introduce two GLRT test statistics in absence of interference, and go on to describe two of the statistics found in the literature.

- It is shown in Appendix 2.6 that the GLRT sufficient statistics optimized for the assumption that $\mathbf{R} = \sigma^2\mathbf{I}$ where $\mathbf{R}$ and $\mathbf{H}$ are unknown is given by

$$\eta_{\text{GLRT,we}} = \frac{\text{Tr}(\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*})}{\text{Tr}(\hat{\mathbf{R}}_{yy})} \tag{2.23}$$

  In the particular case of K decorrelated (or orthogonal) training sequences, expression (2.23) reduces to

$$\eta_{\text{GLRT,we}} = \sum_{i=1}^{K} \frac{\hat{\mathbf{r}}_{ys_i}^{*}\hat{\mathbf{r}}_{ys_i}}{\hat{\pi}_y\pi_{s_{ii}}}. \tag{2.24}$$

  Further, as shown in [41, 42], the $\eta_{\text{GLRT,we}}$ statistic for a SIMO system is given by

$$\eta_{\text{GLRT,we}} = \frac{\hat{\mathbf{r}}_{ys}^{*}\hat{\mathbf{r}}_{ys}}{\hat{\pi}_y\pi_s}.$$

  and we see that (2.24) corresponds to the sum of K SIMO $\eta_{\text{GLRT,we}}$ statistics, each associated to a transmit antenna.

- It is shown in Appendix 2.7 that the GLRT sufficient statistics optimized for a correlation matrix in the form

$$\mathbf{R} = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_M^2 \end{pmatrix} \tag{2.25}$$

  with unknown parameters $\mathbf{H}$ and $[\sigma_1^2, \sigma_2^2, \dots, \sigma_M^2]$, where σ_n^2 is the power of the noise received on receive antenna n, is given by

$$\eta_{\text{GLRT,wu}} = \frac{\prod_{n=1}^{M}(\hat{\mathbf{R}}_{yy})_{n,n}}{\prod_{n=1}^{M}(\hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*})_{n,n}} \tag{2.26}$$

- *Mody's Test:* One of the reference tests for time synchronization of MIMO links without interference is the one proposed in [1,11,14] for OFDM links. It assumes orthogonal training sequences, and may also be used for SC links. It consists in testing the sum, over the K training sequences, of the normalized square modulus of the correlation between the sequences and each component, $\mathbf{y}_n(k), 1 \le n \le M$ of $\mathbf{y}(k)$. The test can then be written as

$$\eta_{\text{MODY}} \triangleq \sup_{n} \Big\{ \sum_{i=1}^{K} \frac{|\hat{\pi}_{y_n s_i}|^2}{\hat{\pi}_{y_n}} \Big\}. \tag{2.27}$$

Comparing (2.27) to (2.23) and (2.24), we deduce that the Mody's test corresponds (to within a constant) to the GLRT1 if and only if the link is a multi-input single-output (MISO) link ($M = 1$) with orthogonal training sequences.

- *Correlation test:* Another reference test for time synchronization of MIMO links without interference is the Correlation test proposed in [8] for SC links. It makes no assumptions on the training sequences. It consists in testing the normalized Frobenius norm squared of the $(M \times K)$ matrix $N\hat{\mathbf{R}}_{ys} = N[\hat{\mathbf{r}}_{ys_1}, \ldots, \hat{\mathbf{r}}_{ys_K}]$, which contains all the correlations between the K training sequences and the M components of $\mathbf{Y}$. It can be written as:

$$\eta_{\mathrm{COR}} = \frac{\mathrm{Tr}(\hat{\mathbf{R}}_{ys}\hat{\mathbf{R}}_{ys}^{*})}{\mathrm{Tr}(\hat{\mathbf{R}}_{yy})\,\mathrm{Tr}(\mathbf{R}_{ss})} = \sum_{i=1}^{K} \frac{\hat{\mathbf{r}}_{ys_i}^{*}\hat{\mathbf{r}}_{ys_i}}{\hat{\pi}_y \sum_{m=1}^{K} \pi_{s_m}} \tag{2.28}$$

Comparing (2.28) with (2.23) and (2.24), we deduce that for an arbitrary $(K \times M)$ MIMO link, the correlation test corresponds (to within a constant) to $\eta_{\mathrm{GLRT,we}}$ if and only if the training sequences are orthogonal and have the same power.

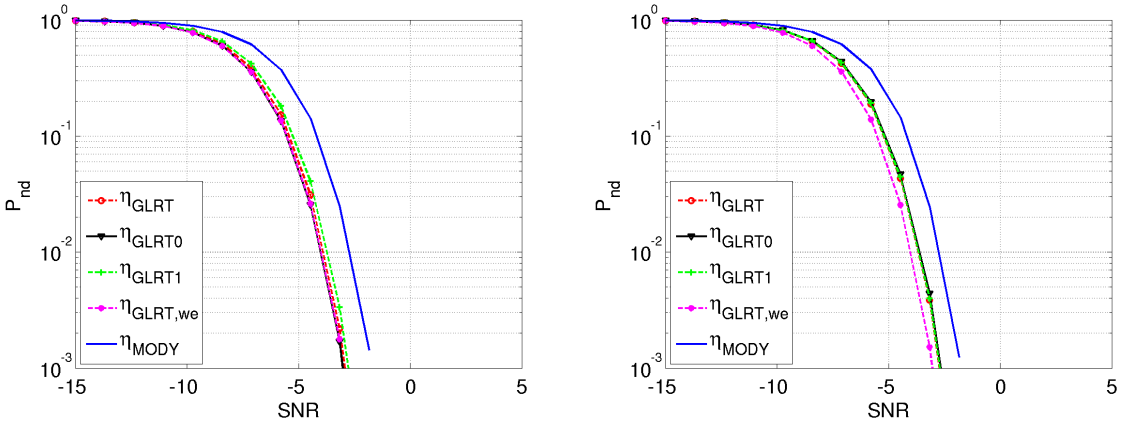
Performance in the absence of interference

We assume in this section an absence of interference such that $\mathbf{R} = \sigma^2 \mathbf{I}_N$, and consider both orthogonal and non-orthogonal training sequences.

For orthogonal sequences having the same power, the statistical test η_{COR} is equivalent to $\eta_{\text{GLRT,we}}$ whereas η_{MMSE} and η_{GLRT0} are equivalent. We thus only consider in this case, $\eta_{\text{GLRT,we}}$, η_{MODY}, η_{GLRT}, η_{GLRT0} and η_{GLRT1}. Under these asumptions, Figures 2.1a and 2.1b show, for a (2×2) MIMO link, the variations of the non-detection probability as a function of the SNR per receive antenna at the output of the previous statistics for a deterministic channel matrix $\mathbf{H}$. The vector $\mathbf{h}_1$ is associated with a direction of arrival (DOA) of $\theta_1 = 0°$, which is orthogonal to the line array, whereas $\mathbf{h}_2$ corresponds to a DOA such that $|\alpha_{12}|^2 = 0$ (2.1a) or $|\alpha_{12}|^2 = 0.6$ (2.1b. respectively. Figures 2.2a and 2.2b show the same variations but for a random channel matrix $\mathbf{H}$ of dimension (2×2) (2.2a) and (4×4) (2.2a).

For the deterministic channel, we note that for all spatial correlations, the performance stays approximately the same. The relative performance of $\eta_{\text{GLRT,we}}$ improves when the spatial correlation increases. We also note that there is no large difference in performance between the statistics under investigation, except for η_{MODY} which performs slightly worse than the others.

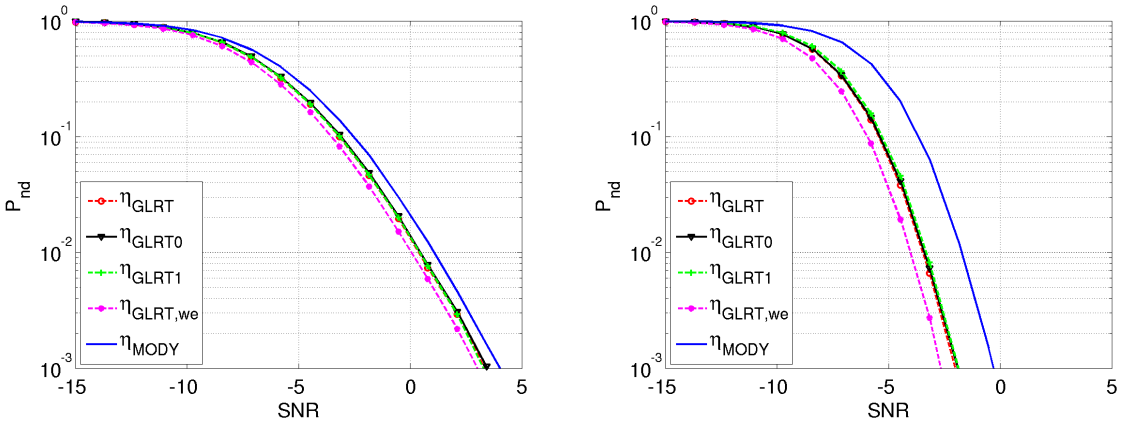
For the random channel, we see that $\eta_{\text{GLRT,we}}$ performs slightly better than all the other statistics. η_{GLRT}, η_{GLRT0} and η_{GLRT1} are almost equivalent in performance for both $M = 2$ and $M = 4$. Moreover, we note that the performance of η_{MODY} deteriorates when M increases.



(a) $|\alpha_{12}|^2 = 0$ (b) $|\alpha_{12}|^2 = 0.6$

Figure 2.1: P_{ND} *as a function of SNR,* $N = 32$*,* $K = M = 2$*,* $\text{P}_{\text{FA}} = 10^{-3}$*, No interference, Orthogonal sequences, Deterministic channel*

For non-orthogonal sequences, we must consider $\eta_{\text{GLRT,we}}$, η_{MODY}, η_{COR}, η_{GLRT}, η_{MMSE},



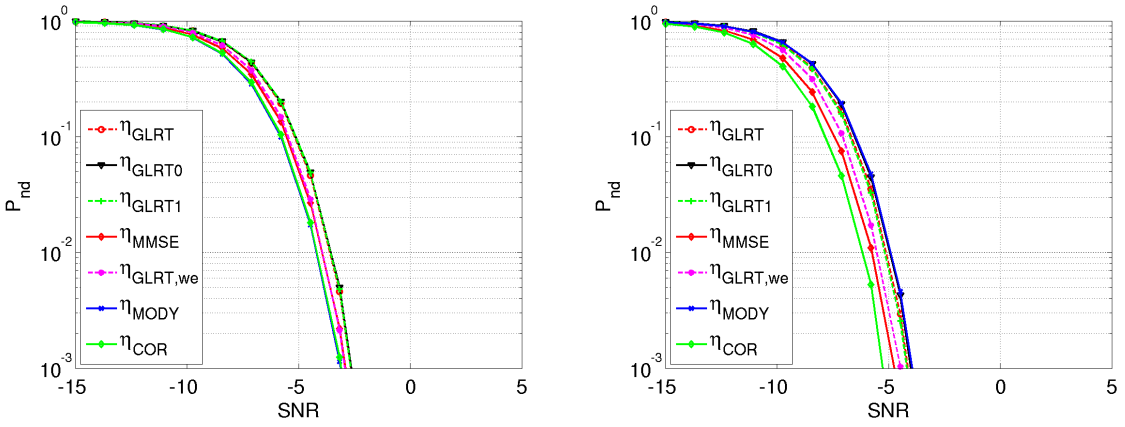
(a) 2×2 (b) 4×4

Figure 2.2: $\mathrm{P_{ND}}$ *as a function of SNR,* $N = 32$, $\mathrm{P_{FA}} = 10^{-3}$, *No interference, Orthogonal sequences, Random channel*

η_{GLRT0} and η_{GLRT1} which are a priori not equivalent. Under these assumptions, Figures 2.3 and 2.4 consider the same scenarios as Figures 2.1 and 2.2 respectively but for non-orthogonal training sequences, such that the temporal correlation ρ between the two synchronization sequences is either 0 or 0.6.

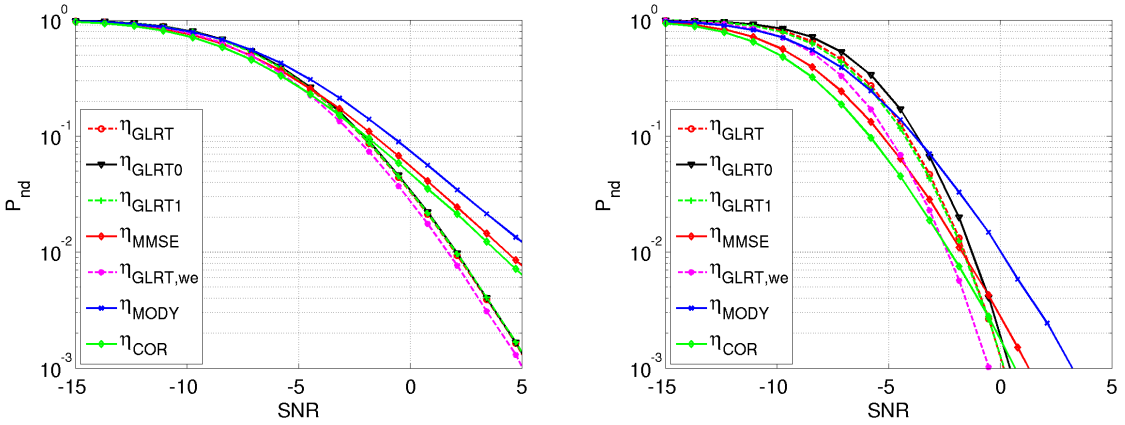
In this scenario, we see that for the deterministic channel model, $\eta_{\mathrm{GLRT,we}}$ is not the best statistics, and performs systematically worse than η_{COR} and η_{MMSE}. This illustrates the fact that, despite a widespread misconception, the GLRT is not necessarily an optimal test when it comes to performance. We also note that whatever the spatial correlation, the difference between η_{COR} and η_{GLRT} is less than 2dB. Further, for the statistics optimized for presence of interference, η_{MMSE} has the best performance. We also note that in the deterministic channel, the performance of η_{GLRT}, η_{GLRT0} and η_{GLRT1} is almost equal.

For the random channel, for both the 2×2 and 4×4 case, we notice that there exists a limit SNR, for which $\eta_{\mathrm{GLRT,we}}$ is always better than η_{COR}. For the low SNR regime, however, η_{COR} performs better. When the number of antennas goes from 2×2 to 4×4, we see that η_{GLRT0} drops in performance, compared with η_{GLRT} and η_{GLRT1}. We also note that the gain in performance for using the $\eta_{\mathrm{GLRT,we}}$ increases when the number of antennas increases. In the 2×2 case, it is favourable to use η_{GLRT}, since the difference in performance is small, and since it is more robust to potential interferences, as we will see in subsection 2.3.2.



(a) $|\alpha_{12}|^2 = 0$ (b) $|\alpha_{12}|^2 = 0.6$

Figure 2.3: $\mathrm{P_{ND}}$ *as a function of SNR,* $N = 32$, $K = M = 2$, $\mathrm{P_{FA}} = 10^{-3}$, *No interference, Non-orthogonal sequences, Deterministic channel*



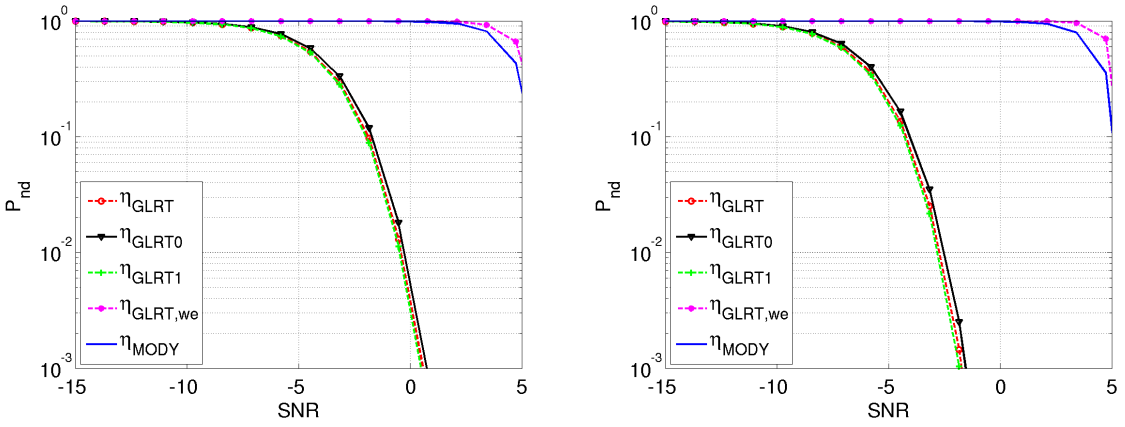
(a) 2×2 (b) 4×4

Figure 2.4: $\mathrm{P_{ND}}$ *as a function of SNR,* $N = 32$, $\mathrm{P_{FA}} = 10^{-3}$, *No interference, Non-orthogonal sequences, Random channel*

Performance in the presence of interference

In this subsection, we assume the presence of interference, represented by the noise model of equation (2.22) such that $\mathbf{R} = \pi_I \mathbf{h}_I \mathbf{h}_I^* + \sigma^2 \mathbf{I}_M$, and consider both orthogonal and non-orthogonal training sequences.

For orthogonal sequences having the same power, η_{MMSE} and η_{GLRT0} are equivalent, and we therefore only consider η_{GLRT}, η_{GLRT0} and η_{GLRT1}. Under the same assumptions as before,



(a) $|\alpha_{12}|^2 = 0, |\alpha_{1I}|^2 = 0.262, |\alpha_{2I}|^2 = 0.728$ (b) $|\alpha_{12}|^2 = 0.6, |\alpha_{1I}|^2 = 0.262, |\alpha_{2I}|^2 = 0.022$

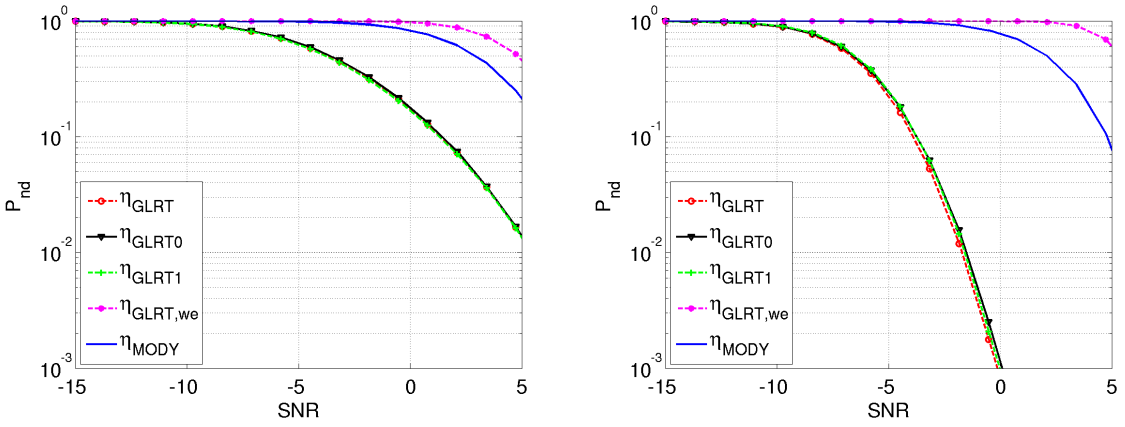
Figure 2.5: P_{ND} *as a function of SNR,* $N = 32$, $K = M = 2$, $P_{FA} = 10^{-3}$, *Interference, Orthogonal sequences, Deterministic channel*

Figures 2.5 and 2.6 consider the same scenarios and show the same variations as Figures 2.1 and 2.2 respectively but in the presence of one interference. For Figure 2.5, the vector $\mathbf{h}_I$ is associated with the DOA $\theta_I = 20°$. This means means that for figure 2.5a, $|\alpha_{1I}|^2 = 0.262$ and $|\alpha_{2I}|^2 = 0.728$, whereas for 2.5b, $|\alpha_{1I}|^2 = 0.262$ and $|\alpha_{2I}|^2 = 0.022$ For Figure 2.6, $\mathbf{h}_I$ follows our random model.

We start by stating the obvious fact that both $\eta_{\text{GLRT,we}}$ and η_{MODY} fail in this scenario, since they are optimized for spatially white noise. We also note that for orthogonal training sequences, for both deterministic and random channels, there is hardly any difference in performance between η_{GLRT}, η_{GLRT0} and η_{GLRT1}, and by extension, the least complex statistics should be used.

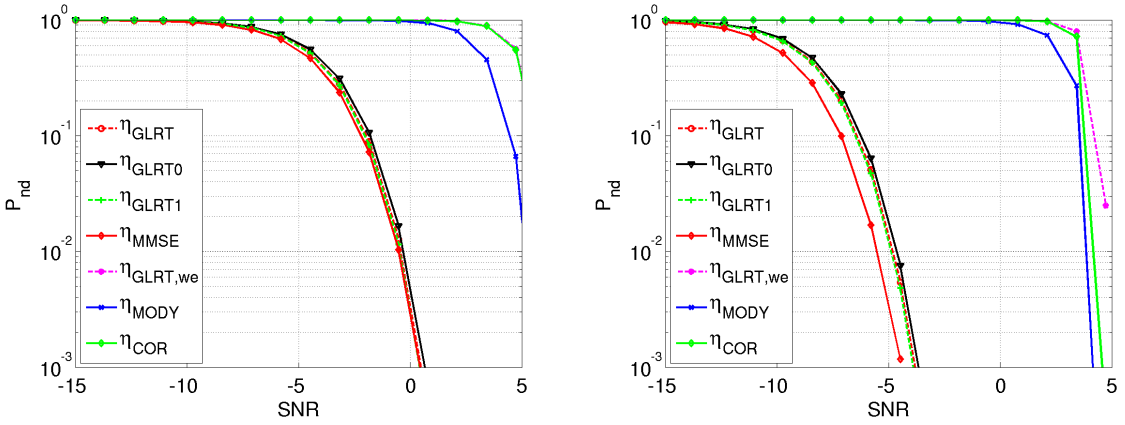
For non-orthogonal sequences, we must consider the η_{MMSE} statistics in addition to the previous ones. Figures 2.7 and 2.8 consider the same scenarios and as Figures 2.3 and 2.4 respectively but in the presence of one interference. For Figure 2.7, the vector $\mathbf{h}_I$ is associated with the DOA $\theta_I = 20°$, whereas for Figure 2.8, $\mathbf{h}_I$ is random.

For the deterministic channel, we note that η_{MMSE} is the best statistics, and also that it performs relatively better with increasing spatial correlation. For the random channel, however, η_{GLRT}, η_{GLRT0} and η_{GLRT1} perform better. For the 4×4 system, We see again that there exists a limit SNR below which η_{MMSE} performs better than the other statistics. We also see that when the number of antennas increases, η_{GLRT0} performs slightly worse than η_{GLRT} and η_{GLRT1}.



(a) 2×2 (b) 4×4

Figure 2.6: $\mathrm{P_{ND}}$ *as a function of SNR,* $N = 32$, $\mathrm{P_{FA}} = 10^{-3}$, *Interference, Orthogonal sequences, Random channel*

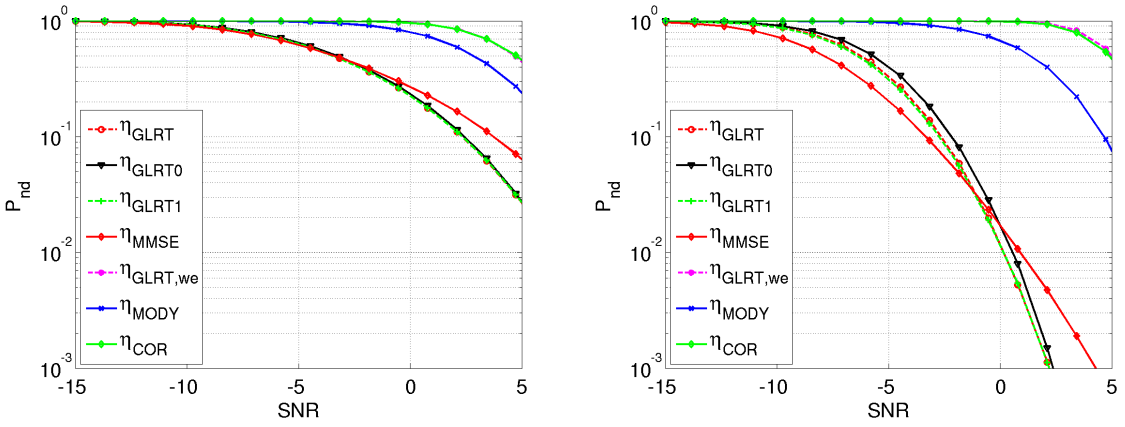


(a) $|\alpha_{12}|^2 = 0, |\alpha_{1I}|^2 = 0.262, |\alpha_{2I}|^2 = 0.728$ (b) $|\alpha_{12}|^2 = 0.6, |\alpha_{1I}|^2 = 0.262, |\alpha_{2I}|^2 = 0.022$

Figure 2.7: $\mathrm{P_{ND}}$ *as a function of SNR,* $N = 32$, $K = M = 2$, $\mathrm{P_{FA}} = 10^{-3}$, *Interference, Non-orthogonal sequences, Deterministic channel*

2.3.3 Computation rate decrease of $\hat{\mathbf{R}}_{yy}$

At each tested sample position, the use of η_{GLRT}, η_{MMSE}, η_{GLRT0} and η_{GLRT1} requires both the computation and inversion of the $(M \times M)$ $\hat{\mathbf{R}}_{yy}$ matrix defined from N observation samples. Clearly, the resulting computation rate may become very costly for high values of M. This is the motivation for our proposed method of decreasing the complexity of η_{GLRT} and η_{GLRT0}, which is based on decreasing the rate of computation and inversion of $\hat{\mathbf{R}}_{yy}$ by a factor $\beta > 1$.



(a) 2×2 (b) 4×4

Figure 2.8: P_{ND} *as a function of SNR,* $N = 32$, $P_{FA} = 10^{-3}$, *Interference, Non-orthogonal sequences, Random channel*

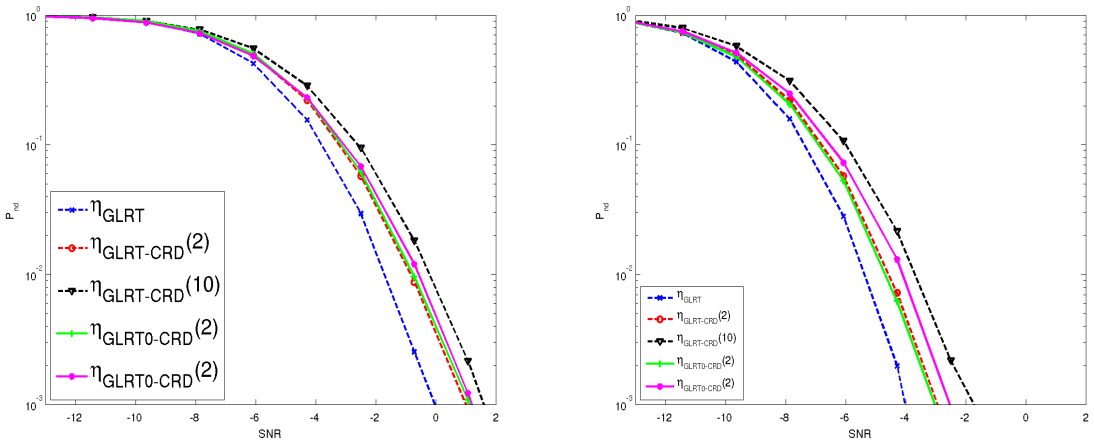
Implementation of computation rate decrease

The principle is to build an $(M \times N')$ observation matrix $\mathbf{Y}' = [\mathbf{y}(1), \ldots, \mathbf{y}(N')]$ from N' observation samples such that $N' > N$ and, for the $\beta = N' - N + 1$ tested sample positions l $(1 \leq l \leq \beta)$, to replace in η_{GLRT} or η_{GLRT0} the $\hat{\mathbf{R}}_{yy}$ matrix by the $\hat{\mathbf{R}}'_{yy}$ matrix defined by $\hat{\mathbf{R}}'_{yy} = \mathbf{Y}'\mathbf{Y}'^{*}/N'$. Note that $N' - N$ samples are now data samples instead of synchronization samples. As the data samples associated with different antennas are uncorrelated, this strategy to decrease the complexity of η_{GLRT} and η_{GLRT0} is only valid for orthogonal synchronization sequences. Further, this strategy requires constant values of $\mathbf{H}$ and $\mathbf{R}$ over N' samples, which may limit the value of N'. However, it allows to compute and inverse only one $(M \times M)$ matrix per set of β tested sample positions, hence a gain of β in the matrix computation and inversion. Note that this strategy cannot be applied to η_{GLRT1} since the computation of $\hat{\mathbf{R}}_1$ requires an update of $\hat{\mathbf{R}}_{ys}$ at each time sample.

Performance analysis

Figure 2.9 shows, for $(K, M) = (4, 4)$ and $(2, 8)$, $N'/N = 2$ and 10, the variations of P_{ND} as a function of the SNR per receive antenna at the output of η_{GLRT}, $\eta_{GLRT-CRD}$ and $\eta_{GLRT0-CRD}$, where S-CRD means statistics S with a computation rate decrease. Note an increasing performance degradation of $\eta_{GLRT-CRD}$ and $\eta_{GLRT0-CRD}$ with respect to η_{GLRT} (equivalent in this case to η_{GLRT0}) as N'/N increases, while remaining lower than 1dB for $N'/N = 2$, which enlightens the interest of $\eta_{GLRT-CRD}$ and $\eta_{GLRT0-CRD}$. Note also a better performance of $\eta_{GLRT0-CRD}$ with

respect to $\eta_{\text{GLRT}-\text{CRD}}$ for $N'/N = 10$, showing a better robustness of the former.



(a) 4×4 (b) 2×8

Figure 2.9: Performance of statistics with computation rate decrease

2.3.4 Complexity analysis

In order to understand the relative complexities of the MIMO synchronization statistics presented in this chapter, this subsection presents a complexity analysis of the considered statistics. An outline of the complexity of some common matrix operations is found in Appendix 2.10. Here, we go on to outline the approximate number of complex operations to compute each of the statistics.

To simplify the analysis, and since the statistics with computation rate decrease are only applicable for orthogonal synchronization sequences, we perform this analysis assuming that $\mathbf{R}_{ss}$ is diagonal, i.e. that the sequences are orthogonal. Further, we note that by defining $\mathbf{G}_M = \hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys}\hat{\mathbf{R}}_{ys}^{*}$ and $\mathbf{G}_K = \hat{\mathbf{R}}_{ys}^{*}\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys}$ we can rewrite η_{GLRT} as

$$\eta_{\text{GLRT}} = \det(\mathbf{I}_M - \mathbf{G}_M) = \det(\mathbf{I}_K - \mathbf{G}_K) \tag{2.29}$$

and η_{GLRT0} as

$$\eta_{\text{GLRT0}} = \text{Tr}(\mathbf{G}_M) = \text{Tr}(\mathbf{G}_K). \tag{2.30}$$

Thus, the computation of both statistics requires the computation of either $\mathbf{G}_M$ or $\mathbf{G}_K$. Depending of the values of M, K, we can therefore choose which statistics to use in order to reduce complexity. For the statistics with computation rate decrease, $\eta_{\text{GLRT}-\text{CRD}}$ and $\eta_{\text{GLRT0}-\text{CRD}}$, the calculation of $\hat{\mathbf{R}}_{yy}$ and its inversion is done less often, which makes it less complex.

	Trace / Det	Inverse	Matrix products
$\eta_{\text{GLRT}-\text{K}}$	$\frac{2}{3}K^3+2K-1$	$\frac{8}{3}M^3$	$MK(2N-1)+\frac{M^2+M}{2}(2N-1)+MK(2M-1)+K^2(2M-1)$
$\eta_{\text{GLRT0}-\text{K}}$	$K-1$	$\frac{8}{3}M^3$	$MK(2N-1)+\frac{M^2+M}{2}(2N-1)+MK(2M-1)+K^2(2M-1)$
$\eta_{\text{GLRT}-\text{CRD}-\text{K}}$	$\frac{2}{3}K^3+2K-1$	$\frac{8}{3\beta}M^3$	$MK(2N-1)+\frac{M^2+M}{2}(2N-\frac{1}{\beta})+MK(2M-1)+K^2(2M-1)$
$\eta_{\text{GLRT0}-\text{CRD}-\text{K}}$	$K-1$	$\frac{8}{3\beta}M^3$	$MK(2N-1)+\frac{M^2+M}{2}(2K-\frac{1}{\beta})+MK(2M-1)+K^2(2M-1)$

Table 2.1: Number of complex operations required for computing the considered statistics

To illustrate the relative complexities, Table 2.1 summarizes the number of complex operations required to compute the considered statistics. Further, figure 2.10 shows, for $N=32$, $N'/N=10$, $K=2$ and $K=8$, the number of complex operations required by η_{GLRT}, $\eta_{\text{GLRT}-\text{CRD}}$, η_{GLRT0} and $\eta_{\text{GLRT0}-\text{CRD}}$ as a function of M. Note the increasing complexity with K and M of all the statistics and the great interest of optimizing K for both performance and complexity reasons. Note the lower complexities of the statistics with CRD and the lowest complexity of $\eta_{\text{GLRT0}-\text{CRD}}$. Note finally the lower complexity of η_{GLRT0} with respect to η_{GLRT} as K increases.

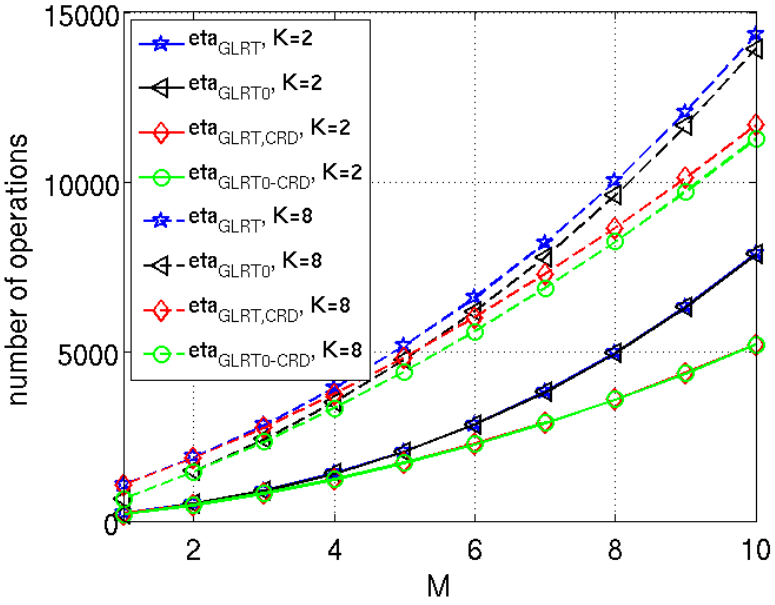


Figure 2.10: Number of operations as a function of M, $N'=10N=320$

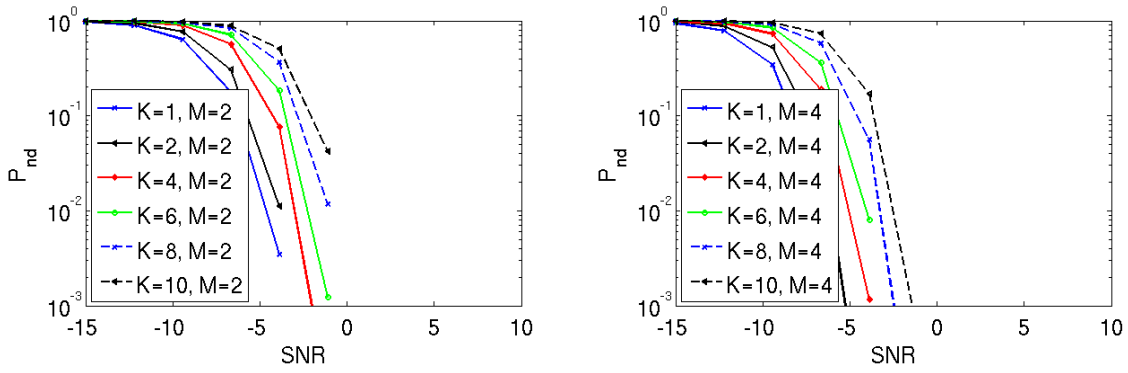
2.4 Optimization of K for a fixed M and statistics

As the complexity of all MIMO synchronization statistics increases with the number of transmit antennas K, it is important in practice to wonder whether this parameter can be optimized for synchronization purposes. We first note that in the presence of one interference, to reject interferences, at least two receive antennas are required. In this section, we show, for orthogonal sequences of $N = 32$ QPSK samples, $M = 2$ and different values of K, the variations of $\mathrm{P_{ND}}$ as a function of the SNR per receive antenna at the output of η_{GLRT}. The figures in presence of interference are performed for for INR = SNR + 15 dB. Similar results are obtained for η_{GLRT0} and η_{GLRT1}.

2.4.1 Deterministic channel

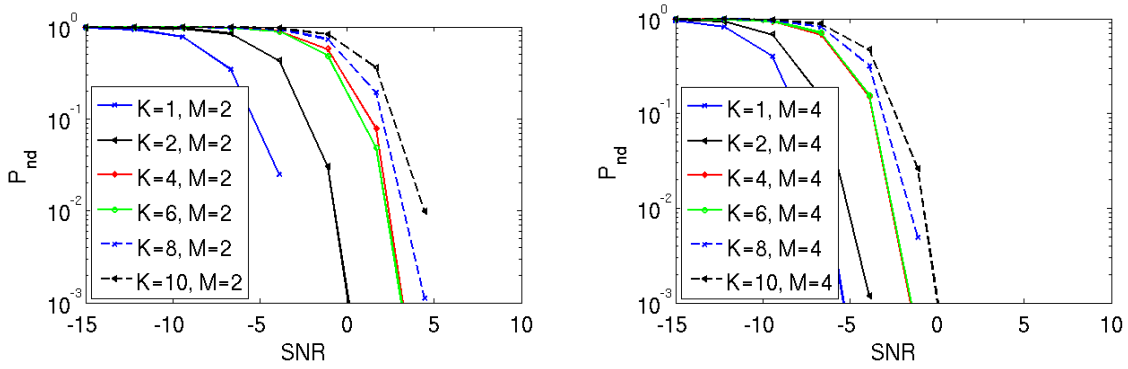
Adding transmit antennas in the deterministic case does not improve the performance. For a deterministic channel, since there is no fading to begin with, additional transmit antennas do not create any diversity gain. Additionally, they can be seen as additional interference, which deteriorates the performance.

Figures 2.11 and 2.12 illustrate this effect, for absence and presence of interference, respectively. Adding transmit antennas while keeping the number of receive antennas constant leads to a performance loss.



(a) $M = 2$ (b) $M = 4$

Figure 2.11: Determistic channel, absence of interference

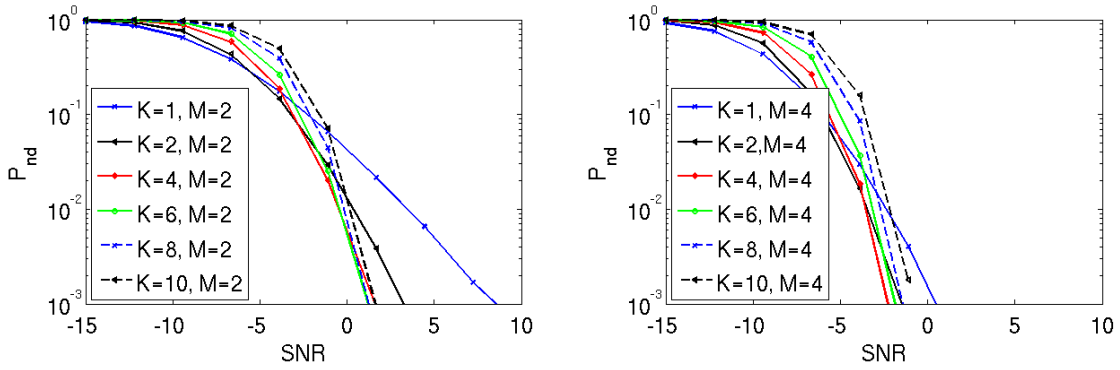


(a) $M = 2$ (b) $M = 4$

Figure 2.12: Deterministic channel, presence of interference.

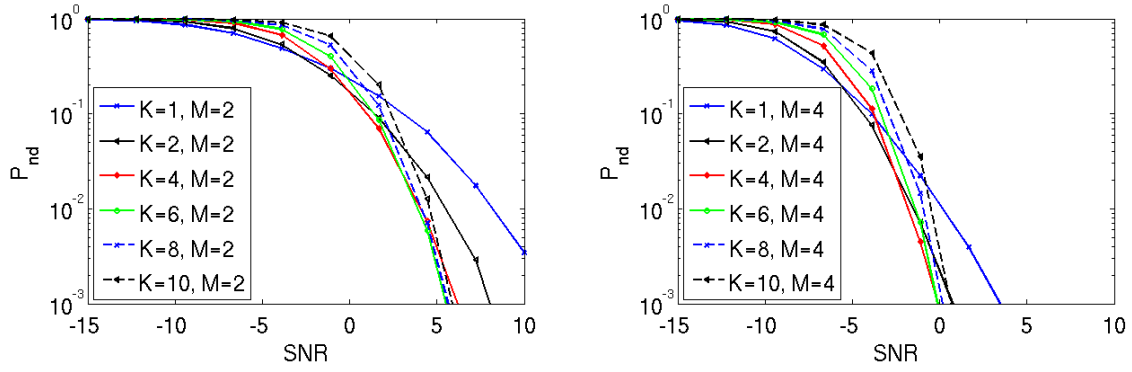
2.4.2 Random channel

For fading channels, for a given value of M and at least for high SNR, increasing K while transmitting the same global power increases the spatial diversity order of the MIMO system. However, increasing K also increases the number of transmitted sequences and thus the amount of interference at reception. A compromize between diversity and interferences should then be found. Figures 2.13 and 2.14 show the performance for varying K and M for absence and presence of interference, respectively. In both absence and presence of interference, for low SNR, we note the optimality of the SIMO scheme for synchronization, proving in this case that the dominant limitation parameter are the internal interferences from the additional transmit antennas. On the contrary at high SNR, we note the sub-optimality of the SIMO scheme due to fading and increasing performance with K as long as $KM \leq 8$, due to an increase of the system diversity order up to 8. For this value, the fading has practically disappeared. Above a system diversity order of 8, the increase in the diversity gain is very weak while the interference level increases, hence the non-increasing or even decreasing performance with increasing K. Similar results are obtained for other values of M.



(a) $M = 2$ (b) $M = 4$

Figure 2.13: Random channel, absence of interference



(a) $M = 2$ (b) $M = 4$

Figure 2.14: Random channel, presence of interference.

2.5 Conclusion

In this chapter, new insights into the time synchronization of $K \times M$ MIMO systems, corrupted by interferences of any kind, have been given. Several schemes aiming at reducing the complexity of η_{GLRT}, the GLRT statistics optimized for presence of temporally white interference [8], have been proposed. Alternative expressions of η_{GLRT} have been introduced and the determinant computation has been done explicitly for $K = 2$, allowing a direct comparison of η_{GLRT} with η_{MMSE}. Two new low-complexity statistics, η_{GLRT0} and η_{GLRT1}, have been introduced, and have been shown to have a performance very close to that of η_{GLRT} for a wide range of parameter choices. A comparative parameter analysis has been performed, taking into consideration the noise type, channel type, the number of transmit and receive antennas, and the orthogonality of the

synchronization sequence. To further reduce complexity, a powerful procedure of computation rate reduction of the data correlation matrix has been proposed for orthogonal sequences. Furthermore, the problem of optimization of K for time synchronization has been investigated. showing, for high SNR, increasing performance with K as long as KM does not become greater than 8. All these results have been presented with the goal of optimizing the performance-complexity tradeoff.

2.6 Appendix: Derivation of $\eta_{\mathbf{GLRT,we}}$

It is shown in this appendix that expression (2.23) is a sufficient statistic for the GLRT detection of the known synchronization sequence matrix $\mathbf{S}$ from the observation matrix $\mathbf{Y}$, assuming zero-mean, stationary, i.i.d, spatially white, circular Gaussian samples $\mathbf{v}(k)$ $(0 \leq k \leq K-1)$, with correlation matrix in the form $\mathbf{R} = \sigma^2\mathbf{I}$, and unknown parameters $\mathbf{H}$ and σ^2. To this aim, let us first compute the ML estimates of $(\mathbf{H}, \sigma^2)$ under H_1 and of σ^2 under H_0 respectively. Using (1.11), the Log-likelihood, $\log(L_1)$, of $(\mathbf{H}, \sigma^2)$ under H_1, observing $\mathbf{Y}$, can be written as

$$\log(L_1) = -MN\log(\pi) - MN\log(\sigma^2) - \sum_{k=0}^{N-1}((\mathbf{y}(k) - \mathbf{H}\mathbf{s}(k))^*(\mathbf{y}(k) - \mathbf{H}\mathbf{s}(k)))/\sigma^2 \tag{2.31}$$

Derivating this expression with respect to σ^2 and setting the result to zero, we obtain the ML estimate $\hat{\sigma}_1^2$ of σ^2 under H_1, given by

$$\hat{\sigma}_1^2 = \frac{1}{MN}\sum_{k=0}^{N-1}(\mathbf{y}(k) - \mathbf{H}\mathbf{s}(k))^*(\mathbf{y}(k) - \mathbf{H}\mathbf{s}(k)) \tag{2.32}$$

In a similar way, it is easy to show that the ML estimate $\hat{\sigma}_0^2$ of σ^2 under H_0 is given by

$$\hat{\sigma}_0^2 = \frac{1}{MN}\sum_{k=0}^{N-1}\mathbf{y}(k)^H\mathbf{y}(k) = \frac{1}{M}\,\mathrm{Tr}(\hat{\mathbf{R}}_{yy}) \tag{2.33}$$

Inserting (2.32) into (2.31) , we get the ML estimate, $\hat{\mathbf{H}}$, of $\mathbf{H}$ which maximizes (2.31) as

$$\hat{\mathbf{H}} = \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1} \tag{2.34}$$

Developing (2.32), using (2.34) and taking the trace, it is straightforward to show that $\hat{\sigma}_1^2$ takes the form

$$\hat{\sigma}_1^2 = \frac{1}{M}\,\mathrm{Tr}(\hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^*). \tag{2.35}$$

Replacing in (1.9) $(\mathbf{H}, \sigma^2)$ by $(\hat{\mathbf{H}}, \hat{\sigma}_1^2)$ under H_1 and σ^2 by $\hat{\sigma}_0^2$ under H_0, we obtain the GLRT test, given by

$$\eta_{\mathrm{GLRT,we}} = \Big(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2}\Big)^{NM} \tag{2.36}$$

Inserting the expressions for $\hat{\sigma}_0^2$ and $\hat{\sigma}_1^2$ into (2.36), we deduce that

$$\eta_{\mathrm{GLRT,we}} = \Big(\frac{\mathrm{Tr}(\hat{\mathbf{R}}_{yy})}{\mathrm{Tr}(\hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^*)}\Big)^{NM} \tag{2.37}$$

and that a sufficient statistic for the problem is given by

$$\eta_{\mathrm{GLRT,we}} = \frac{\mathrm{Tr}(\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^*)}{\mathrm{Tr}(\hat{\mathbf{R}}_{yy})} \tag{2.38}$$

2.7 Appendix: Derivation of $\eta_{\mathbf{GLRT,wu}}$

It is shown in this Appendix that expression (1.15) is a sufficient statistic for the GLRT detection of the known synchronization sequence matrix $\mathbf{S}$ from the observation matrix $\mathbf{Y}$, assuming zero-mean, stationary, i.i.d, spatially white, circular Gaussian samples $\mathbf{v}(k)$ $(0 \leq k \leq N-1)$, and a correlation matrix in the form

$$\mathbf{R} = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_M^2 \end{pmatrix} \tag{2.39}$$

with unknown parameters $\mathbf{H}$ and $\mathbf{R} = \text{Diag}[\sigma_1^2, \sigma_2^2, \dots, \sigma_M^2]$, where σ_n^2 is the power of the noise received on receive antenna n.

Let us first compute the ML estimates of $(\mathbf{H}, \mathbf{R})$ under H_1 and of $\mathbf{R}$ under H_0 respectively. Since the probability density function is known, the Log-likelihood, $\log(L_1)$, of $(\mathbf{H}, \mathbf{R})$ under H_1, observing $\mathbf{Y}$, can be written as

$$\log(L_1) = -MN\log(\pi) - N\log\Big(\prod_{p=1}^{M}\sigma_p^2\Big) + \sum_{k=0}^{N-1}(\mathbf{Y}_k - \mathbf{H}\mathbf{S}_k)^*\mathbf{R}^{-1}(\mathbf{Y}_k - \mathbf{H}\mathbf{S}_k) \tag{2.40}$$

The derivative of this expression with respect to σ_n^2 is

$$\frac{\partial \log(L_1)}{\partial \sigma_n^2} = -\frac{N}{\sigma_n^2} + \sum_{k=0}^{N-1}(\mathbf{Y}_k - \mathbf{H}\mathbf{S}_k)^* \begin{pmatrix} \ddots & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_n^4} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \ddots \end{pmatrix} (\mathbf{Y}_k - \mathbf{H}\mathbf{S}_k) \tag{2.41}$$

$$= -\frac{N}{\sigma_n^2} + \frac{1}{\sigma_n^4}\sum_{k=0}^{N-1}(\mathbf{Y}_k - \mathbf{H}\mathbf{S}_k)_n^*(\mathbf{Y}_k - \mathbf{H}\mathbf{S}_k)_n \tag{2.42}$$

setting the result to zero, we obtain that the ML estimate $\hat{\sigma}_{1,n}^2$ of σ_n^2 under H_1 is given by

$$\hat{\sigma}_{1,n}^2 = \frac{1}{N}\sum_{k=0}^{N-1}(\mathbf{Y}_k - \mathbf{H}\mathbf{S}_k)_n^*(\mathbf{Y}_k - \mathbf{H}\mathbf{S}_k)_n. \tag{2.43}$$

To get the ML estimate, $\hat{\mathbf{H}}$, of $\mathbf{H}$ which maximizes (2.40), is again given by

$$\hat{\mathbf{H}} = \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}. \tag{2.44}$$

Developing (2.43) and using (2.44), $\hat{\sigma}^2_{1,n}$ takes the form

$$\hat{\sigma}^2_{1,n} = (\hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}^*_{ys})_{n,n} \tag{2.45}$$

In a similar way, it is easy to show that the ML estimate $\hat{\sigma}^2_{0,n}$ of σ^2_n under H_0 is given by

$$\hat{\sigma}^2_{0,n} = \frac{1}{N}\sum_{k=0}^{N-1} Y^*_{n,k}Y_{n,k} = (\hat{\mathbf{R}}_{yy})_{n,n} \tag{2.46}$$

Replacing in (1.9) $(\mathbf{H}, \mathbf{R})$ by $(\hat{\mathbf{H}}, \mathrm{Diag}[\hat{\sigma}^2_{1,1}, \ldots, \hat{\sigma}^2_{1,M}])$ under H_1 and $\mathbf{R}$ by $\mathrm{Diag}[\hat{\sigma}^2_{0,1}, \ldots, \hat{\sigma}^2_{0,M}])$ under H_0, we obtain the GLRT test, given by

$$\eta_{\mathrm{GLRT,wu}} = \frac{\prod_{n=1}^{M}(\hat{\mathbf{R}}_{yy})_{n,n}}{\prod_{n=1}^{M}(\hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}^*_{ys})_{n,n}}.$$

2.8 Appendix: Derivation of $\eta_{\mathbf{GLRT}}$

It is shown in this appendix that expression (2.2) is a sufficient statistic for the GLRT detection of the known synchronization sequence matrix $\mathbf{S}$ from the observation matrix $\mathbf{Y}$, assuming zero-mean, stationary, i.i.d, spatially colored, circular, Gaussian samples $\mathbf{v}(k)(0 \leq k \leq N-1)$, and unknown parameters $\mathbf{H}$ and $\mathbf{R}$. To this aim, let us first compute the ML estimates of ($\mathbf{H}$,$\mathbf{R}$) under H_1 and of $\mathbf{R}$ under H_0 respectively. Using (1.11), the Log-likelihood, $\log(L_1)$, of ($\mathbf{H}$,$\mathbf{R}$) under H_1, observing $\mathbf{Y}$, can be written as

$$\log(L_1) = -MN\log(\pi) - N\log(\det(\mathbf{R})) \tag{2.47}$$

$$- \sum_{k=0}^{N-1} (\mathbf{y}(k) - \mathbf{H}\mathbf{s}(k))^* \mathbf{R}^{-1} (\mathbf{y}(k) - \mathbf{H}\mathbf{s}(k)) \tag{2.48}$$

It is then straightforward to show that the ML estimate, $\hat{\mathbf{H}}$, of $\mathbf{H}$ i.e. the $\hat{\mathbf{H}}$ which maximizes (2.48), is still given by (2.34). Moreover, it is well-known [22, 23] that the ML estimate, $\hat{\mathbf{R}}_1$, of $\mathbf{R}$ under H_1, i.e. the matrix $\mathbf{R}$ which maximizes (2.48) is given by

$$\hat{\mathbf{R}}_1 = \frac{1}{N} \sum_{k=0}^{N-1} (\mathbf{y}(k) - \mathbf{H}\mathbf{s}(k))(\mathbf{y}(k) - \mathbf{H}\mathbf{s}(k))^* \tag{2.49}$$

Developing (2.49) and using (2.34), it is straightforward to show that $\hat{\mathbf{R}}_1$ takes the form $\hat{\mathbf{R}}_1 = \hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^*$. In a similar way, it is straightforward to show that the ML estimate $\hat{\mathbf{R}}_0$ of $\mathbf{R}$ under H_0 is given by

$$\hat{\mathbf{R}}_0 = \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{y}(k)\mathbf{y}(k)^* = \hat{\mathbf{R}}_{yy} \tag{2.50}$$

Replacing in (1.9) $(\mathbf{H}, \mathbf{R})$ by $(\hat{\mathbf{H}}, \hat{\mathbf{R}}_1)$ under H_1 and $\mathbf{R}$ by $\hat{\mathbf{R}}_0$ under H_0 and using (1.10) and (1.11), we obtain the GLRT test, given by

$$\eta_{\text{GLRT}} = \Big(\frac{\det(\hat{\mathbf{R}}_0)}{\det(\hat{\mathbf{R}}_1)}\Big)^N \tag{2.51}$$

Inserting the expressions for $\hat{\mathbf{R}}_0$ and $\hat{\mathbf{R}}_1$ into (2.51), we deduce that a sufficient statistic for the previous problem is given by

$$\eta_{\text{GLRT}} = \Big(\frac{\det(\hat{\mathbf{R}}_1)}{\det(\hat{\mathbf{R}}_0)}\Big)^{-N}$$

Taking the logarithm, we have

$$\eta_{\text{GLRT}} = -N\log\det\left(\mathbf{I}_M - \hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^*\right) \tag{2.52}$$

which corresponds to equation (2.2).

2.9 Appendix: Derivation of $\eta_{\mathbf{GLRT,kn}}$

It is shown in this appendix that expression (2.10) is a sufficient statistic for the GLRT detection of the known synchronization sequence matrix $\mathbf{S}$ from the observation matrix $\mathbf{Y}$ assuming zero-mean, stationary, i.i.d, spatially colored, circular, Gaussian samples $\mathbf{v}(k)$ $(0 \leq k \leq N-1)$ with a known covariance matrix $\mathbf{R}$, and an unknown channel matrix $\mathbf{H}$. The ML estimate, $\hat{\mathbf{H}}$, of $\mathbf{H}$ under H_1 is still given by (2.34). Using (1.10) and (1.11) and taking the logarithm of (1.9), we obtain a sufficient statistic for our problem as

$$\begin{aligned}\mathrm{GLRT3} = \sum_{k=0}^{N-1} \Big(& \mathbf{s}(k)^* \mathbf{H}^* \mathbf{R}^{-1} \mathbf{s}(k) \mathbf{H} - \mathbf{y}(k)^* \mathbf{R}^{-1} \mathbf{H} \mathbf{s}(k) \\ & - \mathbf{s}(k)^* \mathbf{H}^* \mathbf{R}^{-1} \mathbf{y}(k) \Big) \end{aligned} \tag{2.53}$$

Replacing in (2.53) the channel matrix $\mathbf{H}$ by its ML estimate $\hat{\mathbf{H}} = \hat{\mathbf{R}}_{ys} \mathbf{R}_{ss}^{-1}$, and taking the trace, we deduce that the sufficient statistic GLRT3 is defined by (2.10).

2.10 Appendix: Complexity of common matrix operations

In this Appendix, we remind the reader of the complexity of some common matrix operations, which are used as a basis of our computations of complexity, used for Table 2.1.

LU decomposition

Consider the $(N \times N)$ matrix $\mathbf{A}$ and its LU-decomposition

$$\mathbf{A} = \mathbf{LU} \tag{2.54}$$

where $\mathbf{L}$ is a lower-diagonal matrix and $\mathbf{U}$ an upper-diagonal matrix. It is well known that the cost of performing this decomposition is approximately $\frac{2}{3}N^3$.

Determinant

A common way of calculating the determinant is using the LU-decomposition, since the decomposition into upper and lower diagonal matrices simplifies the computation. We can easily see that

$$\det(\mathbf{A}) = \det(\mathbf{LU}) = \det(\mathbf{L})\det(\mathbf{U}) \tag{2.55}$$

$$= \prod_{i=1}^{N} L_{ii} \prod_{i=1}^{N} U_{ii} \tag{2.56}$$

The total number of operations for calculating the determinant using LU-decomposition is $\frac{2}{3}N^3 + 2(N-1) + 1$.

Matrix inverse

To simplify the calculation of the inverse, we start by LU-decomposition of the matrix $\mathbf{A}$ that we want to invert. We then calculate the inverse of the matrix

$$\mathbf{A}^{-1} = (\mathbf{LU})^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1} \tag{2.57}$$

The calculation of the inverse then requires solving a system of equations of N equations. The cost of solving this system of equations is $2N^3$. The total cost of inverting a matrix is thus $\frac{2}{3}N^3 + 2N^3 = \frac{8}{3}N^3$.

Matrix product

Each element of a matrix product $\mathbf{C} = \mathbf{AB}$ where $\mathbf{A}$ is a $(N \times K)$ matrix and $\mathbf{B}$ a $(K \times M)$ matrix requires K multiplications and $K-1$ additions. Therefore, to compute the NM elements, we have a computational complexity of $NM(2K-1)$. If $\mathbf{A}$ and $\mathbf{B}$ are both $(N \times N)$, the cost is $N^2(2N-1) = 2N^3 - N^2$.

Product of a matrix with its conjugate transpose

When calculating a matrix product of the type $\mathbf{C} = \mathbf{YY}^H$, some of the products are the same up to conjugation, and we can save on the number of multiplications. In fact, we only need to calculate the terms on the diagonal and the diagonals below it, and take their conjugate to fill the remaining elements. If $\mathbf{Y}$ is $N \times K$, we need to compute $\sum_{i=1}^{N} i = \frac{N^2+N}{2}$ elements. To calculate each element of $\mathbf{C}$, we still need to perform $K + (K-1) = 2K-1$ multiplications and additions per element. This results in a total computational complexity of $\frac{N^2+N}{2}(2K-1)$.

Chapter 3

Large system analysis of a GLRT for detection with large sensor arrays in temporally white noise

3.1 Introduction

Due to the spectacular development of sensor networks and acquisition devices, it has become common to be faced with multivariate signals of high dimension. Very often, the sample size that can be used in practice in order to perform statistical inference cannot be much larger than the signal dimension. In this context, it is well established that a number of fundamental existing statistical signal processing methods fail. It is therefore of crucial importance to revisit certain classical problems in the high-dimensional signals setting. Previous works in this direction include e.g. [46] and [47] in source localization using a subspace method, or [48], [49], [50], [51] in the context of unsupervised detection.

In the present chapter, we address the problem of detecting the presence of a known signal using a large array of sensors. We assume that the observations are corrupted by a temporally white, but spatially correlated (with unknown spatial covariance matrix) additive complex Gaussian noise, and study the generalized likelihood ratio test (GLRT). Although our results can be used in more general situations, we focus on the detection of a known synchronization sequence transmitted by a single transmitter in an unknown multipath propagation channel. The behaviour of the GLRT in this context has been extensively addressed in previous works, but for the low dimensional signal case (see e.g. [52], [8], [53], [54], [55], [56], [9]). The asymptotic

behaviour of the relevant statistics has thus been studied in the past, but it has been assumed that the number of samples of the training sequence N converges towards $+\infty$ while the number of sensors M remains fixed. This is a regime which in practice makes sense when $M << N$. When the number of sensors M is large, this regime is however often unrealistic, since in order to avoid wasting resources, the size N of the training sequence is usually chosen of the same order of magnitude as M. Therefore, we consider in this chapter the asymptotic regime in which both M and N converge towards ∞ at the same rate.

We consider both the case where the number of paths L remains fixed, and the case where L converges towards ∞ at the same rate as M and N. When L is fixed, we prove that the GLRT statistics η_N converges under hypothesis H_0 towards a Gaussian distribution with mean $L \log \frac{1}{1-M/N}$ and variance $\frac{L}{N} \frac{M/N}{1-M/N}$. This is in contrast with the standard asymptotic regime $N \to +\infty$ and M fixed in which the distribution of η_N converges towards a χ^2 distribution. Under hypothesis H_1, we prove that η_N has a similar behaviour than in the standard asymptotic regime $N \to +\infty$ and M fixed, except that the terms $L \log \frac{1}{1-M/N}$ and $\frac{L}{N} \frac{M/N}{1-M/N}$ are added to the asymptotic mean and the asymptotic variance, respectively. When L converges towards ∞ at the same rate as M and N, we use existing results (see [57] and [58]) characterizing the behaviour of linear statistics of the eigenvalues of large multivariate F–matrices, and infer that the distribution of η_N under H_0 is also asymptotically Gaussian. The asymptotic mean converges towards ∞ at the same rate as L, M, N while the asymptotic variance is a $\mathcal{O}(1)$ term. The asymptotic behaviour of η_N under hypothesis H_1 when L scales with M, N is not covered by the existing literature. The derivation of the corresponding new mathematical results would need an extensive work that is not in the scope of the present chapter. We rather propose a pragmatic approximate distribution for η_N, motivated by the additive structure of its asymptotic mean and variance in the regime where L is fixed.

We evaluate the accuracy of the various Gaussian approximations by numerical simulations, by comparing the asymptotic means and variances with their empirical counterparts evaluated by Monte-Carlo simulations. Further, we compare the ROC curves corresponding to the various approximations with the empirical ones. The numerical results show that the standard approximations obtained when $N \to +\infty$ and M is fixed completely fail if $\frac{M}{N}$ is greater than $\frac{1}{8}$. The large system approximations corresponding to a fixed L and $L \to +\infty$ appear reliable for small values of $\frac{M}{N}$, and, of course, for larger values of $\frac{M}{N}$. For the values of L, M, N that are considered, the approximations obtained in the regime $L \to +\infty$ at the same rate as M and N appear to be the most accurate, and the corresponding ROC-curves are shown to be good approximations of the empirical ones. Therefore, the proposed Gaussian approximations allow to reliably predict the performance of the GLRT when the number of array elements is large.

This chapter is organized as follows. In section 3.2 we recall the expression of the statistics η_N corresponding to the GLRT, and explain that, in order to study η_N, assuming that the additive noise is spatially white and that the training sequence matrix is orthogonal is not a restriction. In section 3.3, we recall the asymptotic behaviour of η_N in the traditional asymptotic regime $N \to +\infty$ and M fixed. The main results of this chapter, concerning the asymptotic behaviour of η_N in the regime M, N converge towards ∞ at the same rate, are presented in section 3.4. In this section, we only give outlines of the proofs, while providing the remaining technical details in Appendices. Section 3.5 is devoted to the numerical results, and section 3.6 concludes the chapter.

3.2 Presentation of the problem.

In the following, we assume that a single transmitter sends a known synchronization sequence $(s_n)_{n=1,\dots,N}$ through a fixed channel with L paths, and that the corresponding signal is received on a receiver with M sensors. The received M-dimensional signal is denoted by $(\mathbf{y}_n)_{n=1,\dots,N}$. When the transmitter and the receiver are perfectly synchronized, $\mathbf{y}_n$ is assumed to be given for each $n = 1, \dots, N$ by

$$\mathbf{y}_n = \sum_{l=0}^{L-1} \mathbf{h}_l s_{n-l} + \mathbf{v}_n \tag{3.1}$$

where $(\mathbf{v}_n)_{n\in\mathbb{Z}}$ is an additive independent identically distributed complex Gaussian noise verifying

$$\begin{aligned} \mathbb{E}(\mathbf{v}_n) &= 0 \\ \mathbb{E}(\mathbf{v}_n \mathbf{v}_n^T) &= 0 \\ \mathbb{E}(\mathbf{v}_n \mathbf{v}_n^*) &= \mathbf{R} = \sigma^2 \tilde{\mathbf{R}} \end{aligned} \tag{3.2}$$

where $\mathbf{R} > 0$ and $\frac{1}{M}\mathrm{Tr}(\tilde{\mathbf{R}}) = 1$. Denoting by $\mathbf{H}$ the $M \times L$ matrix $\mathbf{H} = (\mathbf{h}_0, \dots, \mathbf{h}_{L-1})$, the received signal matrix $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_N)$ in the presence of a useful signal can be written as

$$\mathbf{Y} = \mathbf{HS} + \mathbf{V} \tag{3.3}$$

where $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_N)$ and where $\mathbf{S}$ represents the known signal matrix. We assume from now on that the size N of the training sequence satisfies $N > M + L$. We remark that the forthcoming results are valid as soon as the matrix collecting the observations can be written as in Eq. (3.3). In particular, by appropriately modifying the matrices $\mathbf{H}$ and $\mathbf{S}$, this system model can equivalently be used for a link with multiple transmit antennas.

Furthermore, in the absence of a useful signal, the received signal matrix is given by

$$\mathbf{Y} = \mathbf{V}. \tag{3.4}$$

In this chapter, we study the classical problem of testing the hypothesis H_1 characterized by Equation (3.3) against the hypothesis H_0 defined by equation (3.4), in the aim of testing whether there is a useful signal present in the received signal. The hypotheses are

$$\begin{aligned} \mathrm{H}_0 &: \mathbf{Y} = \mathbf{V} \\ \mathrm{H}_1 &: \mathbf{Y} = \mathbf{HS} + \mathbf{V}, \end{aligned} \tag{3.5}$$

where we assume from now on that $\mathbf{H}$ and $\mathbf{R}$ are unknown at the receiver side. In the following, we will review the expression of the corresponding generalized maximum likelihood test (GLRT) derived in [8]. The generalized likelihood ratio r_N is, as mentioned in chapter 1, defined by [55]

$$r_N = \frac{\max\limits_{\mathbf{R},\mathbf{H}} p_{H_1}(\mathbf{Y} \mid \mathbf{S}, \mathbf{H}, \mathbf{R})}{\max\limits_{\mathbf{R}} p_{H_0}(\mathbf{Y} \mid \mathbf{R})} \tag{3.6}$$

We recall that the probability density functions are given by

$$\begin{aligned} p_{H_0}(\mathbf{Y} \mid \mathbf{R}) =& \frac{1}{\pi^{NM} \left(\det(\mathbf{R})\right)^N} e^{-\mathrm{Tr}[\mathbf{Y}^* \mathbf{R}^{-1} \mathbf{Y}]} \\ p_{H_1}(\mathbf{Y} \mid \mathbf{S}, \mathbf{H}, \mathbf{R}) =& \frac{1}{\pi^{NM} \left(\det(\mathbf{R})\right)^N} e^{-\mathrm{Tr}[(\mathbf{Y}-\mathbf{HS})^* \mathbf{R}^{-1} (\mathbf{Y}-\mathbf{HS})]}. \end{aligned} \tag{3.7}$$

The first step to calculate r_N is to determine $\hat{\mathbf{R}}_1$ and $\hat{\mathbf{H}}$, the $\mathbf{R}$ and $\mathbf{H}$ that maximize the numerator, and $\hat{\mathbf{R}}_0$, the $\mathbf{R}$ that maximizes the denominator, of equation (1.18). Straightforward calculations show that $\hat{\mathbf{H}} = \frac{\mathbf{YS}^*}{N}(\frac{\mathbf{SS}^*}{N})^{-1}$ and $\hat{\mathbf{R}}_1 = \frac{\mathbf{YY}^*}{N} - (\frac{\mathbf{YS}^*}{N})(\frac{\mathbf{SS}^*}{N})^{-1}(\frac{\mathbf{SY}^*}{N})$. Similarly, $\hat{\mathbf{R}}_0$ is given by $\hat{\mathbf{R}}_0 = \frac{\mathbf{YY}^*}{N}$.

Inserting these estimates into equation (1.18) leads to $r_N = \left(\det(\hat{\mathbf{R}}_1 \hat{\mathbf{R}}_0^{-1})\right)^{-N}$. Therefore, the log-likelihood ratio η_N, defined by $\eta_N = \frac{\log r_N}{N}$, is given by

$$\eta_N = -\log\det\left[\mathbf{I}_M - \hat{\mathbf{R}}_0^{-1/2} \frac{\mathbf{YS}^*}{N} \left(\frac{\mathbf{SS}^*}{N}\right)^{-1} \frac{\mathbf{SY}^*}{N} \hat{\mathbf{R}}_0^{-1/2}\right] \tag{3.8}$$

or, using the identity $\det(\mathbf{I} - \mathbf{AB}) = \det(\mathbf{I} - \mathbf{BA})$, by

$$\eta_N = -\log\det\left[\mathbf{I}_L - \mathbf{T}_N\right] \tag{3.9}$$

where $\mathbf{T}_N$ is the $L \times L$ matrix defined by

$$\mathbf{T}_N = \left(\frac{\mathbf{SS}^*}{N}\right)^{-1/2} \frac{\mathbf{SY}^*}{N} \left(\frac{\mathbf{YY}^*}{N}\right)^{-1} \frac{\mathbf{YS}^*}{N} \left(\frac{\mathbf{SS}^*}{N}\right)^{-1/2} \tag{3.10}$$

The generalized maximum likelihood test consists then in comparing η_N to a threshold.

In order to study the behaviour of the test in Eq. (3.9), we study the limit distribution of η_N under each hypothesis. For this, we remark that it is possible to assume without restriction that $\frac{\mathbf{S}\mathbf{S}^*}{N} = \mathbf{I}_L$ is verified and that $\mathbb{E}(\mathbf{v}_n\mathbf{v}_n^*) = \sigma^2\mathbf{I}$, i.e. $\tilde{\mathbf{R}}$ is reduced to the identity matrix. If this is not the case, we denote by $\tilde{\mathbf{S}}$ the matrix

$$\tilde{\mathbf{S}} = \left(\frac{\mathbf{S}\mathbf{S}^*}{N}\right)^{-1/2} \mathbf{S} \tag{3.11}$$

and by $\tilde{\mathbf{Y}}$ and $\tilde{\mathbf{V}}$ the whitened observation and noise matrices

$$\begin{aligned} \tilde{\mathbf{Y}} =& \tilde{\mathbf{R}}^{-1/2}\,\mathbf{Y}, \\ \tilde{\mathbf{V}} =& \tilde{\mathbf{R}}^{-1/2}\,\mathbf{V} \end{aligned} \tag{3.12}$$

It is clear that $\frac{\tilde{\mathbf{S}}\tilde{\mathbf{S}}^*}{N} = \mathbf{I}_L$ and that $\mathbb{E}(\tilde{\mathbf{v}}_n\tilde{\mathbf{v}}_n^*) = \sigma^2\mathbf{I}$. Moreover, under H_0, it holds that $\tilde{\mathbf{Y}} = \tilde{\mathbf{V}}$, while under H_1, $\tilde{\mathbf{Y}} = \tilde{\mathbf{H}}\tilde{\mathbf{S}} + \tilde{\mathbf{V}}$ where the channel matrix $\tilde{\mathbf{H}}$ is defined by

$$\tilde{\mathbf{H}} = \tilde{\mathbf{R}}^{-1/2}\,\mathbf{H}\,(\mathbf{S}\mathbf{S}^*/N)^{1/2} \tag{3.13}$$

Finally, it holds that the statistics η_N can also be written as

$$\eta_N = -\log\det\left[\mathbf{I}_L - \frac{\tilde{\mathbf{S}}\tilde{\mathbf{Y}}^*}{N}\left(\frac{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^*}{N}\right)^{-1}\frac{\tilde{\mathbf{Y}}\tilde{\mathbf{S}}^*}{N}\right] \tag{3.14}$$

This shows that it is possible to replace $\mathbf{S}$, $\tilde{\mathbf{R}}$ and $\mathbf{H}$ by $\tilde{\mathbf{S}}$, $\mathbf{I}$, and $\tilde{\mathbf{H}}$ without modifying the value of statistics η_N. Therefore, without restriction, we assume from now on that

$$\frac{\mathbf{S}\mathbf{S}^*}{N} = \mathbf{I}_L,\ \tilde{\mathbf{R}} = \mathbf{I}_M \tag{3.15}$$

In the following, we denote by $\mathbf{W}$ a $(N-L)\times N$ matrix for which the matrix $\mathbf{\Theta} = (\mathbf{W}^T, \frac{\mathbf{S}^T}{\sqrt{N}})^T$ is unitary and define the $M\times(N-L)$ and $M\times L$ matrices $\mathbf{V}_1$ and $\mathbf{V}_2$ by

$$(\mathbf{V}_1, \mathbf{V}_2) = \mathbf{V}\mathbf{\Theta}^* = (\mathbf{V}\mathbf{W}^*, \mathbf{V}\frac{\mathbf{S}^*}{\sqrt{N}}) \tag{3.16}$$

It is clear that $\mathbf{V}_1$ and $\mathbf{V}_2$ are complex Gaussian random matrices with independent identically distributed $\mathcal{N}_{\mathbb{C}}(0,\sigma^2)$ entries, and that the entries of $\mathbf{V}_1$ and $\mathbf{V}_2$ are mutually independent. We notice that since $N > M + L$, the matrix $\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}$ is invertible almost surely. We now express the statistics η_N in terms of $\mathbf{V}_1$ and $\mathbf{V}_2$. We observe that

$$\frac{\mathbf{V}\mathbf{V}^*}{N} = \frac{\mathbf{V}_1\mathbf{V}_1^*}{N} + \frac{\mathbf{V}_2\mathbf{V}_2^*}{N} \tag{3.17}$$

and that

$$\frac{\mathbf{V}\mathbf{S}^*}{N} = \frac{1}{\sqrt{N}} (\mathbf{V}_1, \mathbf{V}_2) \begin{pmatrix} \mathbf{W} \\ \frac{\mathbf{S}}{\sqrt{N}} \end{pmatrix} \frac{\mathbf{S}^*}{\sqrt{N}} \tag{3.18}$$

coincides with $\frac{\mathbf{V}_2}{\sqrt{N}}$ because $\mathbf{W}\frac{\mathbf{S}^*}{\sqrt{N}} = 0$. Therefore, under hypothesis H_0, η_N can be written as

$$\eta_N = -\log\det\left(\mathbf{I} - \frac{\mathbf{V}_2^*}{\sqrt{N}}\left(\frac{\mathbf{V}_1\mathbf{V}_1^*}{N} + \frac{\mathbf{V}_2\mathbf{V}_2^*}{N}\right)^{-1}\frac{\mathbf{V}_2}{\sqrt{N}}\right) \tag{3.19}$$

Using the identity

$$\mathbf{A}^*(\mathbf{B}\mathbf{B}^* + \mathbf{A}\mathbf{A}^*)^{-1} = \mathbf{A}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{A}\left(\mathbf{I} + \mathbf{A}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{A}\right)^{-1} \tag{3.20}$$

we obtain that, under hypothesis H_0, η_N can be written as

$$\eta_N = \log\det\left(\mathbf{I}_L + \mathbf{V}_2^*/\sqrt{N}\ (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\ \mathbf{V}_2/\sqrt{N}\right) \tag{3.21}$$

Similarly, it is easy to check that, under H_1, η_N is given by

$$\eta_N = \log\det\left(\mathbf{I}_L + \mathbf{G}_N\right) \tag{3.22}$$

where the matrix $\mathbf{G}_N$ is defined by

$$\mathbf{G}_N = \left(\mathbf{H} + \mathbf{V}_2/\sqrt{N}\right)^*\ (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\ \left(\mathbf{H} + \mathbf{V}_2/\sqrt{N}\right) \tag{3.23}$$

3.3 Standard asymptotic analysis of η_N.

In order to give a better understanding of the similarities and differences with the more complicated case where M and N converge towards $+\infty$ at the same rate, we first recall some standard results concerning the asymptotic distribution of η_N under H_0 and H_1 when $N \to +\infty$ but M remains fixed.

3.3.1 Hypothesis H_0.

A general result concerning the GLRT, known as Wilk's theorem (see e.g. [55], [59] Chapter 8-5), implies that $N\eta_N$ converges in distribution towards a χ^2 distribution with $2ML$ degrees of freedom. For the reader's convenience, we provide an informal justification of this claim. We use

(3.21) and remark that when $N \to +\infty$ and M and L remain fixed, the matrices $\mathbf{V}_1\mathbf{V}_1^*/N$ and $\frac{1}{N}\mathbf{V}_2^*\left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\mathbf{V}_2$ converge a.s. towards $\sigma^2\mathbf{I}$ and the zero matrix respectively. Moreover,

$$\frac{1}{N}\mathbf{V}_2^*\left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\mathbf{V}_2 = \frac{1}{\sigma^2}\mathbf{V}_2^*\mathbf{V}_2/N + o_P(\frac{1}{N}) \tag{3.24}$$

and a standard second order expansion of η_N leads to

$$\eta_N = \frac{1}{\sigma^2}\mathrm{Tr}\left(\mathbf{V}_2^*\mathbf{V}_2/N\right) + o_P(\frac{1}{N}) \tag{3.25}$$

This implies immediately that the limit distribution of $N\,\eta_N$ is a chi-squared distribution with $2ML$ degrees of freedom. Informally, this implies that $\mathbb{E}(\eta_N) \simeq L\frac{M}{N}$ and $\mathrm{Var}(\eta_N) \simeq \frac{L}{N}\frac{M}{N}$.

3.3.2 Hypothesis H_1.

Under hypothesis H_1, η_N is given by (3.22). When $N \to +\infty$ and M and L remain fixed, the matrix $\mathbf{V}_1\mathbf{V}_1^*/N$ converges a.s. towards $\sigma^2\mathbf{I}$ and it is easily seen that

$$\begin{aligned}\eta_N = \log\det\Big(\mathbf{I}+\frac{\mathbf{H}\mathbf{H}^*}{\sigma^2}\Big)+ \\ \mathrm{Tr}\Big[\Big(\mathbf{I}+\frac{\mathbf{H}\mathbf{H}^*}{\sigma^2}\Big)^{-1}\mathbf{\Delta}_N\Big] + \mathcal{O}_P(1/N)\end{aligned} \tag{3.26}$$

where the matrix $\mathbf{\Delta}_N$ is given by

$$\mathbf{\Delta}_N = \mathbf{H}^*\mathbf{\Upsilon}_N\mathbf{H} + \frac{1}{\sigma^2}\left(\frac{\mathbf{V}_2^*}{\sqrt{N}}\mathbf{H} + \mathbf{H}^*\frac{\mathbf{V}_2}{\sqrt{N}}\right) \tag{3.27}$$

with $\mathbf{\Upsilon}_N = \left(\mathbf{V}_1\mathbf{V}_1/N\right)^{-1} - \mathbf{I}/\sigma^2$. Standard calculations show that

$$\sqrt{N}\left(\eta_N - \log\det\left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^*}{\sigma^2}\right)\right) \to \mathcal{N}(0,\kappa_1) \tag{3.28}$$

where κ_1 is given by

$$\kappa_1 = \mathrm{Tr}\left[\mathbf{I} - \left(\mathbf{I} + \frac{\mathbf{H}^*\mathbf{H}}{\sigma^2}\right)^{-2}\right] \tag{3.29}$$

Note that in [55] and [9], the asymptotic distribution of η_N is studied under the assumption that the entries of the matrix $\mathbf{H}$ are $\mathcal{O}(\frac{1}{\sqrt{N}})$ terms. In that context, η_N behaves as a non-central χ^2 distribution.

3.4 Main results.

In this section, we present the main results of this chapter related to the asymptotic behaviour of η_N when M and N converge towards ∞ at the same rate. The analysis of η_N in the asymptotic

regime M and N converge towards ∞ at the same rate differs deeply from the standard regime studied in section 3.3. In particular, it is no longer true that the empirical covariance matrix $\mathbf{V}_1\mathbf{V}_1^*/N$ converges in the spectral norm sense towards $\sigma^2\mathbf{I}$. This, of course, is due to the fact that the number of entries of this $M \times M$ matrix is of the same order of magnitude than the number of available scalar observations (i.e. $M(N-L) = \mathcal{O}(MN)$). We also note that for any deterministic $M \times M$ matrix $\mathbf{A}$, the diagonal entries of the $L \times L$ matrix $\frac{1}{N}\mathbf{V}_2^*\mathbf{A}\mathbf{V}_2$ converge towards 0 when $N \to +\infty$ and M remains fixed, while this does not hold when M and N are of the same order of magnitude (see Proposition 4 in Appendix 3.7). It turns out that the asymptotic regime where M and N converge towards ∞ at the same rate is more complicated than the conventional regime of section 3.3. As the proofs of the following theorems are rather technical, we just provide in this section the outlines of the approaches that are used to establish them. The detailed proofs are given in the Appendix 3.8.

3.4.1 Asymptotic behaviour of η_N when the number of paths L remains fixed when M and N increase.

All along this section, we assume that:

Assumption 1.
- *M and N converge towards $+\infty$ in such a way that $c_N = \frac{M}{N} < 1 - \frac{L}{N}$ converges towards c, where $0 < c < 1$*
- *the number of paths L remains fixed when M and N increase.*

In the asymptotic regime defined by Assumption 1, M can be interpreted as a function $M(N)$ of N. Therefore, M-dimensional vectors or matrices where one of the dimensions is M will be indexed by N in the following. Moreover, in order to simplify the exposition, $N \to +\infty$ should be interpreted in this section as the asymptotic regime defined by Assumption 1.

As M is growing, we have to be precise with how the power of the useful signal component $\mathbf{HS}$ is normalized. In the following, we assume that the norms of vectors $(\mathbf{h}_l)_{l=0,\dots,L-1}$ remain bounded when the number of sensors M increases. This implies that the signal to noise ratio at the output of the matched filter $\mathbf{S}^*\mathbf{H}^*\mathbf{Y}/\sqrt{N}$, i.e. $\mathrm{Tr}\left((\mathbf{H}^*\mathbf{H})^2\right) / \left(\sigma^2\mathrm{Tr}(\mathbf{H}^*\mathbf{H})\right)$, is a $\mathcal{O}(1)$ term in our asymptotic regime. We mention however that the received signal to noise ratio $\mathrm{Tr}(\mathbf{H}^*\mathbf{H})/(M\sigma^2)$ converges towards 0 at rate $\frac{1}{N}$ when N increases.

Asymptotic behaviour of η_N under hypothesis H_0

Under hypothesis H_0, the following theorem holds.

Theorem 1. *It holds that*

$$\eta_N - L\,\log\left(\frac{1}{1-c_N}\right) \to 0\, a.s. \tag{3.30}$$

and that

$$\frac{\sqrt{N}}{\sqrt{\frac{Lc_N}{1-c_N}}}\left(\eta_N - L\,\log\left(\frac{1}{1-c_N}\right)\right) \to_{\mathcal{D}} \mathcal{N}_{\mathbb{R}}(0,1) \tag{3.31}$$

Informally, Theorem 1 leads to $\mathbb{E}(\eta_N) \simeq -L\log(1-c_N)$ and $\mathrm{Var}(\eta_N) \simeq \frac{L}{N}\frac{c_N}{1-c_N}$. We recall that if M is fixed, $N\eta_N$ behaves like a χ^2 distribution with $2ML$ degrees of freedom. In that context, $\mathbb{E}(\eta_N) \simeq Lc_N$ and $\mathrm{var}(\eta_N) \simeq \frac{L}{N}c_N$. Therefore, the behaviour of η_N in the two asymptotic regimes deeply differ. However, if $c_N \to 0$, $-\log(1-c_N) \simeq c_N$, and the asymptotic means and variances of η_N tend to coincide.

Outline of the proof. We denote by $\mathbf{F}_N$ the $L \times L$ matrix

$$\mathbf{F}_N = \mathbf{V}_2^*/\sqrt{N}\ (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\ \mathbf{V}_2/\sqrt{N} \tag{3.32}$$

and remark that under H_0, (3.21) leads to

$$\eta_N = \log\det\left(\mathbf{I}_L + \mathbf{F}_N\right) \tag{3.33}$$

First step: proof of (3.30). As L does not increase with M and N, it is sufficient to establish that

$$\mathbf{F}_N - \frac{c_N}{1-c_N}\,\mathbf{I}_L \to 0\ a.s. \tag{3.34}$$

Our approach is based on the observation that if $\mathbf{A}_N$ is a $M \times M$ deterministic Hermitian matrix verifying $\sup_N \|\mathbf{A}_N\| < a < +\infty$, then,

$$\mathbb{E}_{\mathbf{V}_2}\Big|\left(\mathbf{V}_2^*/\sqrt{N}\,\mathbf{A}_N\mathbf{V}_2/\sqrt{N}\right)_{k,l} - \frac{\sigma^2}{N}\mathrm{Tr}(\mathbf{A}_N)\,\delta(k-l)\Big|^4 \ \le \frac{C(a)}{N^2} \tag{3.35}$$

where $C(a)$ is a constant term depending on a, and where $\mathbb{E}_{\mathbf{V}_2}$ represents the mathematical expectation operator w.r.t. $\mathbf{V}_2$. This is a consequence of Proposition 4 in the Appendix 3.7. Assume for the moment that there exists a deterministic constant a such that

$$\|\,(\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\,\| \le a \tag{3.36}$$

for each N greater than a non random integer N_0. Then, as $\mathbf{V}_1$ and $\mathbf{V}_2$ are independent, it is possible to use (3.35) for $\mathbf{A}_N = (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}$ and to take the mathematical expectation w.r.t. $\mathbf{V}_1$ of (3.35) to obtain that

$$\mathbb{E}\left|(\mathbf{F}_N)_{k,l} - \frac{\sigma^2}{N}\mathrm{Tr}\,(\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\ \delta(k-l)\right|^4 \le \frac{C(a)}{N^2} \tag{3.37}$$

for each $N > N_0$, and, using the Borel-Cantelli lemma, that

$$\mathbf{F}_N - \frac{\sigma^2}{N}\mathrm{Tr}\left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\mathbf{I}_L \to 0 \; a.s. \tag{3.38}$$

In order to conclude, we use known results related to the almost sure convergence of the eigenvalue distribution of matrix $\mathbf{V}_1\mathbf{V}_1^*/N$ towards the so-called Marcenko-Pastur distribution (see Eq. (3.77) in the Appendix 3.7) which imply that

$$\frac{1}{N}\mathrm{Tr}\Big(\mathbf{V}_1\mathbf{V}_1^*/N\Big)^{-1} - \frac{c_N}{\sigma^2(1-c_N)} \to 0 \tag{3.39}$$

almost surely. This, in conjunction with (3.38), leads to (3.34) and eventually to (3.30).

However, there does not exist a deterministic constant a satisfying (3.36) for each N greater than a non random integer. In order to solve this issue, it is sufficient to replace matrix $(\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}$ by a convenient regularized version. It is well known (see Proposition 1 in the Appendix 3.7) that the smallest and the largest eigenvalue of $\mathbf{V}_1\mathbf{V}_1^*/N$ converge almost surely towards $\sigma^2(1-\sqrt{c})^2 > 0$ and $\sigma^2(1+\sqrt{c})^2$ respectively. This implies that if $\mathcal{E}_N$ is the event defined by

$$\begin{aligned}\mathcal{E}_N = \{&\text{one of the eigenvalues of } \mathbf{V}_1\mathbf{V}_1^*/N \text{ escapes from} \\ &[\sigma^2(1-\sqrt{c})^2 - \epsilon, \sigma^2(1+\sqrt{c})^2 + \epsilon]\}\end{aligned} \tag{3.40}$$

(where ϵ is chosen such that $\sigma^2(1-\sqrt{c})^2 - \epsilon > 0$) then, almost surely, for N larger than a random integer, it holds that $\mathbb{1}_{\mathcal{E}_N^c} = 1$. Therefore, almost surely, for N large enough, it holds that $\eta_N = \eta_N \mathbb{1}_{\mathcal{E}_N^c}$. These two random variables thus share the same almost sure asymptotic behaviour. Moreover, it is clear that $\eta_N \mathbb{1}_{\mathcal{E}_N^c}$ coincides with $\log\det(\mathbf{I}+\mathbf{F}_N\mathbb{1}_{\mathcal{E}_N^c})$. In order to study the almost sure behaviour of $\eta_N\mathbb{1}_{\mathcal{E}_N^c}$, it is thus sufficient to evaluate the behaviour of matrix $\mathbf{F}_N\mathbb{1}_{\mathcal{E}_N^c}$, which has the same expression than $\mathbf{F}_N$, except that matrix $(\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}$ is replaced by $(\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\,\mathbb{1}_{\mathcal{E}_N^c}$. The latter matrix verifies

$$\left\|(\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\,\mathbb{1}_{\mathcal{E}_N^c}\right\| \leq \frac{1}{\sigma^2(1-\sqrt{c})^2-\epsilon} \tag{3.41}$$

for each integer N almost surely. Therefore, the regularized matrix $(\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\,\mathbb{1}_{\mathcal{E}_N^c}$ satisfies (3.36) almost surely for each integer N for $a = \frac{1}{\sigma^2(1-\sqrt{c})^2-\epsilon}$. This immediately leads to the conclusion that $\mathbf{F}_N\,\mathbb{1}_{\mathcal{E}_N^c}$ has the same almost sure behaviour than $\frac{c_N}{1-c_N}\mathbf{I}_L\mathbb{1}_{\mathcal{E}_N^c}$, or equivalently that $\frac{c_N}{1-c_N}\mathbf{I}_L$. This, in turn, implies (3.30).

Second step: proof of (3.31). As $\eta_N = \eta_N\mathbb{1}_{\mathcal{E}_N^c}$ almost surely for N large enough, the asymptotic distributions of $\sqrt{N}[\eta_N - L\log(\frac{1}{1-c_N})]$ and $\sqrt{N}[\eta_N\mathbb{1}_{\mathcal{E}_N^c} - L\log(\frac{1}{1-c_N})]$ coincide. We thus study the latter sequence of random variables because the presence of the regularization factor $\mathbb{1}_{\mathcal{E}_N^c}$ allows to simplify a lot the derivations.

A standard second order expansion of $\log\det(\mathbf{I}+\mathbf{F}_N\mathbb{1}_{\mathcal{E}_N^c})$ leads to

$$\sqrt{N}[\eta_N\mathbb{1}_{\mathcal{E}_N^c} - L\log(\frac{1}{1-c_N})] = \\ (1-c_N)\sqrt{N}\Big(\mathrm{Tr}(\mathbf{F}_N\mathbb{1}_{\mathcal{E}_N^c} - \frac{c_N}{1-c_N}\mathbf{I})\Big) + o_P(1) \tag{3.42}$$

It is thus sufficient to evaluate the asymptotic behaviour of the characteristic function $\psi_{N,0}$ of random variable $\beta_{0,N} = (1-c_N)\sqrt{N}\left(\mathrm{Tr}(\mathbf{F}_N\mathbb{1}_{\mathcal{E}_N^c} - \frac{c_N}{1-c_N}\mathbf{I})\right)$ defined by $\psi_{N,0}(u) = \mathbb{E}(e^{iu\beta_{N,0}})$. For this, we first evaluate $\mathbb{E}_{\mathbf{V}_2}(e^{iu\beta_{N,0}})$, and using Proposition 2 and Proposition 4 in Appendix 3.7, we establish that $\mathbb{E}_{\mathbf{V}_2}(e^{iu\beta_{N,0}})$ has the same asymptotic behaviour as

$$\exp\left[-\frac{u^2}{2}\,\sigma^4\,L\,(1-c_N)^2c_N\,\frac{1}{M}\mathrm{Tr}\left(\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}\right)^{-2}\mathbb{1}_{\mathcal{E}_N^c}\right] \tag{3.43}$$

It is known that $\frac{1}{M}\mathrm{Tr}\left(\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}\right)^{-2}$ behaves almost surely as $\frac{1}{\sigma^4(1-c_N)^3}$ (see Eq. (3.78) in the Appendix 3.7). From this, we obtain immediately that

$$\psi_{N,0}(u) - \exp\left(-\frac{u^2}{2}\,\frac{Lc_N}{1-c_N}\right) \;\to\; 0 \tag{3.44}$$

for each u, which, in turn, establishes (3.31).

Asymptotic behaviour of η_N under hypothesis H_1

The behaviour of η_N under hypothesis H_1 is given by the following result.

Theorem 2. *It holds that*

$$\eta_N - \overline{\eta}_{N,1} \to 0 \; a.s. \tag{3.45}$$

where $\overline{\eta}_{N,1}$ is defined by

$$\overline{\eta}_{N,1} = L\log\frac{1}{1-c_N} + \log\det\left(\mathbf{I}+\mathbf{H}^*\mathbf{H}/\sigma^2\right) \tag{3.46}$$

Moreover,

$$\frac{\sqrt{N}}{\left(\frac{Lc_N}{1-c_N}+\kappa_1\right)^{1/2}}(\eta_N - \overline{\eta}_{N,1}) \to_{\mathcal{D}} \mathcal{N}_{\mathbb{R}}(0,1) \tag{3.47}$$

where κ_1 is defined by (3.29).

Remark 1. *Interestingly, it is seen that the asymptotic mean and variance of η_N are equal to the sum of the asymptotic mean and variance of η_N in the standard regime $N\to+\infty$ and M fixed, with the extra terms $L\log\left(\frac{1}{1-c_N}\right)$ and $\frac{Lc_N}{N(1-c_N)}$, which coincide with the asymptotic mean and variance of η_N under H_0.*

Outline of the proof. We recall that, under H_1, η_N is given by (3.22). As in the proof of Theorem 1, it is sufficient to study the regularized statistics $\eta_N \mathbb{1}_{\mathcal{E}_N^c}$ which is also equal to

$$\eta_N \mathbb{1}_{\mathcal{E}_N^c} = \log\det\left(\mathbf{I}_L + \mathbb{1}_{\mathcal{E}_N^c}\,\mathbf{G}_N\right) \tag{3.48}$$

First step: proof of (3.45). In order to evaluate the almost sure behaviour of $\eta_N \mathbb{1}_{\mathcal{E}_N^c}$, we expand $\mathbf{G}_N \mathbb{1}_{\mathcal{E}_N^c}$ as

$$\begin{aligned}\mathbf{G}_N \mathbb{1}_{\mathcal{E}_N^c} = \mathbf{H}^* \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1} \mathbf{H}\, \mathbb{1}_{\mathcal{E}_N^c} + \mathbf{F}_N \mathbb{1}_{\mathcal{E}_N^c} + \\ (\mathbf{V}_2/\sqrt{N})^* \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1} \mathbf{H}\, \mathbb{1}_{\mathcal{E}_N^c} + \\ \mathbf{H}^* \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1} \left(\mathbf{V}_2/\sqrt{N}\right) \mathbb{1}_{\mathcal{E}_N^c}\end{aligned} \tag{3.49}$$

The first term of the righthandside of (3.49) is known to behave as $\frac{\mathbf{H}^*\mathbf{H}}{\sigma^2(1-c_N)}$ (see (3.81) in Appendix 3.7) while the independence between $\mathbf{V}_1$ and $\mathbf{V}_2$ implies that the third and the fourth terms converge almost surely towards the zero matrix. This is because the fourth-order moments w.r.t. $\mathbf{V}_2$ of their entries are $\mathcal{O}(\frac{1}{N^2})$ terms.
Second step: proof of (3.47). Using a standard second order expansion, we obtain immediately that

$$\sqrt{N}\left(\eta_N \mathbb{1}_{\mathcal{E}_N^c} - \overline{\eta}_{N,1}\right) = \sqrt{N}\,\mathrm{Tr}\left(\mathbf{D}_N \mathbf{\Delta}_N\right) + o_P(1) \tag{3.50}$$

where $\mathbf{\Delta}_N$ and $\mathbf{D}_N$ are defined by

$$\mathbf{\Delta}_N = \mathbf{G}_N \mathbb{1}_{\mathcal{E}_N^c} - \left(\frac{\mathbf{H}^*\mathbf{H}}{\sigma^2(1-c_N)} + \frac{c_N}{1-c_N}\mathbf{I}\right) \tag{3.51}$$

and

$$\mathbf{D}_N = (1-c_N)(\mathbf{I}_L + \mathbf{H}^*\mathbf{H}/\sigma^2)^{-1} \tag{3.52}$$

In order to establish (3.47), it is therefore sufficient to evaluate the asymptotic behaviour of the characteristic function $\psi_{N,1}$ of random variable $\beta_{N,1} = \sqrt{N}\,\mathrm{Tr}\left(\mathbf{D}_N\mathbf{\Delta}_N\right)$. We define κ_N and ω_N by

$$\kappa_N = \mathrm{Tr}\left(\mathbf{C}_N \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\right) \tag{3.53}$$

and

$$\begin{aligned}\omega_N = \mathrm{Tr}\left[\mathbf{D}_N \mathbf{F}_N \mathbb{1}_{\mathcal{E}_N^c}\right] + \\ \mathrm{Tr}\left[\mathbf{D}_N (\mathbf{V}_2/\sqrt{N})^* \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1} \mathbf{H}\, \mathbb{1}_{\mathcal{E}_N^c}\right] + \\ \mathrm{Tr}\left[\mathbf{D}_N\, \mathbf{H}^* \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1} \left(\mathbf{V}_2/\sqrt{N}\right) \mathbb{1}_{\mathcal{E}_N^c}\right]\end{aligned} \tag{3.54}$$

where $\mathbf{C}_N$ the $M \times M$ matrix given by

$$\mathbf{C}_N = (1 - c_N)\mathbf{H}(\mathbf{I}_L + \mathbf{H}^*\mathbf{H}/\sigma^2)^{-1}\mathbf{H}^* \tag{3.55}$$

Then, $\beta_{N,1}$ can be written as

$$\begin{aligned} \beta_{N,1} &= \sqrt{N}\left(\kappa_N - \frac{\mathrm{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)}\right) + \\ &\quad \sqrt{N}\left(\omega_N - \frac{c_N}{1-c_N}\mathrm{Tr}(\mathbf{D}_N)\right) \end{aligned} \tag{3.56}$$

Using the equation above as well as Proposition 2 and Proposition 4 from Appendix 3.7, we establish that $\mathbb{E}_{\mathbf{V}_2}(e^{iu\beta_{N,1}})$ behaves as

$$\exp\left(iu\sqrt{N}\left(\kappa_N - \frac{\mathrm{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)}\right)\right)\exp(-\frac{u^2}{2}\zeta) \tag{3.57}$$

where $\zeta = \frac{c_N}{(1-c_N)^3}\mathrm{Tr}(\mathbf{D}_N^2) + 2\frac{c_N}{(1-c_N)}\mathrm{Tr}(\mathbf{D}_N^2\mathbf{H}^*\mathbf{H})$. In order to obtain the limiting behaviour of $\psi_{N,1}(u)$, it is thus sufficient to evaluate the limit of

$$\mathbb{E}_{\mathbf{V}_1}\left[\exp\left(iu\sqrt{N}\left(\kappa_N - \frac{\mathrm{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)}\right)\right)\right] \tag{3.58}$$

This technical point is addressed in Appendix 3.7, Proposition 3.

Remark 2. *It is useful to recall that the expression of the asymptotic mean and variance of η_N provided in Theorem 2 assumes that $\tilde{\mathbf{R}} = \mathbf{I}$ and that $\frac{\mathbf{S}\mathbf{S}^*}{N} = \mathbf{I}$. If this is not the case, we have to replace $\mathbf{H}$ by $\tilde{\mathbf{R}}^{-1/2}\mathbf{H}\,(\mathbf{S}\mathbf{S}^*/N)^{1/2}$ in Theorem 2.*

Remark 3. *We note that Theorem 2 allows to quantify the influence of an overdetermination of L on the asymptotic distribution of η_N under H_1. This analysis is interesting from a practical point of view, since it is not always possible to know the exact number of paths and their delays. If L is overestimated, i.e. if the true number of paths is $L_1 < L$, then, matrix $\mathbf{H}$ can be written as $\mathbf{H} = (\mathbf{H}_1, 0)$. We also denote by $\mathbf{S}_1$ and $\mathbf{S}_2$ the $L_1 \times N$ and $(L - L_1) \times N$ matrices such that $\mathbf{S} = \left(\mathbf{S}_1^T, \mathbf{S}_2^T\right)^T$. It is easy to check that the second term of $\overline{\eta}_{N,1}$, i.e.*

$$\log\det\left(\mathbf{I}_L + (\mathbf{S}\mathbf{S}^*/N)^{1/2}\mathbf{H}^*\tilde{\mathbf{R}}^{-1}\mathbf{H}(\mathbf{S}\mathbf{S}^*/N)^{1/2}\right) \tag{3.59}$$

coincides with

$$\log\det\left(\mathbf{I}_{L_1} + (\mathbf{S}_1\mathbf{S}_1^*/N)^{1/2}\mathbf{H}_1^*\tilde{\mathbf{R}}^{-1}\mathbf{H}_1(\mathbf{S}_1\mathbf{S}_1^*/N)^{1/2}\right) \tag{3.60}$$

and is thus not affected by the overdetermination of L. Therefore, choosing $L > L_1$ increases $\overline{\eta}_{N,1}$ by the factor $(L - L_1)\log\left(\frac{1}{1-c_N}\right)$. As for the asymptotic variance, it is also easy to verify that κ_1 is not affected by the overdetermination of the number of paths, and that the asymptotic variance is increased by the factor $(L - L_1)\frac{c_N}{1-c_N}$. It is interesting to notice that the standard asymptotic analysis of subsection 3.3.2 does not allow to predict any influence of the overdetermination of L on the asymptotic distribution of η_N.

3.4.2 Asymptotic behaviour of η_N when the number of paths L converges towards ∞ at the same rate as M and N.

The asymptotic regime considered in section 3.4.1 is relevant when the number of paths L is much smaller than M and N. This hypothesis may however be restrictive, so that it is of potential interest to study the following regime:

Assumption 2. *L, M and N converge towards $+\infty$ in such a way that $c_N = \frac{M}{N}$ and $d_N = \frac{L}{N}$ converge towards c and d, where $0 < c + d < 1$*

As explained below in Paragraph 3.4.2, the behaviour of η_N under H_0 in this regime is a consequence of existing results. The behaviour of η_N under H_1 is however not covered by the existing litterature. The derivation of the corresponding new mathematical results needs extensive work that is not in the scope of the present chapter. Motivated by the additive structure of the asymptotic mean and variance of η_N under H_1 under assumption 1, we propose in Paragraph 3.4.2 a pragmatic Gaussian approximation of the distribution of η_N under H_1

Asymptotic behaviour of η_N under hypothesis H_0

Theorem 3. *We define $\tilde{\eta}_N$ by*

$$\begin{aligned}\tilde{\eta}_N = &- N((1-c_N)\log(1-c_N) \\ &+ (1-d_N)\log(1-d_N)) \\ &+ N(1-c_N-d_N)\log(1-c_N-d_N)\end{aligned} \tag{3.61}$$

and $\tilde{\delta}_N$ by

$$\tilde{\delta}_N = -\log\left(\frac{2\sqrt{a_N^2 - b_N^2}}{a_N + \sqrt{a_N^2 - b_N^2}}\right) \tag{3.62}$$

where

$$a_N = \left(1 - \frac{c_N}{1-d_N}\right)^2 + \frac{d_N}{1-d_N}\left(1 + \frac{c_N(1-c_N)}{d_N(1-d_N)}\right) \tag{3.63}$$

$$b_N = 2\frac{d_N}{1-d_N}\sqrt{\frac{c_N(1-c_N)}{d_N(1-d_N)}} \tag{3.64}$$

Then, it holds that $\mathbb{E}(\eta_N) = \tilde{\eta}_N + \mathcal{O}(\frac{1}{N})$ and that

$$\frac{1}{\sqrt{\delta_N}}(\eta_N - \tilde{\eta}_N) \to_{\mathcal{D}} \mathcal{N}_{\mathbb{R}}(0,1) \tag{3.65}$$

Justification. The eigenvalues of $\mathbf{F}_N$ coincide with the non-zero eigenvalues of $(\mathbf{V}_2\mathbf{V}_2^*)/N\ (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}$. Therefore, η_N appears a linear statistics of the eigenvalues of this matrix. $(\mathbf{V}_2\mathbf{V}_2^*)/N\ (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}$ is a multivariate F–matrix. The asymptotic behaviour of the empirical eigenvalue distribution of this kind of random matrix as well as the corresponding central limit theorems are well established (see e.g. Theorem 4-10 and Theorem 9-14 in [57] as well as [58]) when the dimensions of V_1 and V_2 converge towards $+\infty$ at the same rate. Theorem 3 follows from these results.

Remark 4. *We notice that the results of Theorem 3 differ deeply from the results of Theorem 1. We first remark that $\tilde{\eta}_N$, and thus $\mathbb{E}(\eta_N)$, converge towards ∞ at the same rate that L, M, N. Moreover, $\eta_N - \mathbb{E}(\eta_N)$ is an $\mathcal{O}_P(1)$ term under assumption 2, while it is an $\mathcal{O}_P(\frac{1}{\sqrt{N}})$ term when L does not scale with M, N. However, it is possible to informally obtain the expressions of the asymptotic mean and variance of η_N in Theorem 1 from (3.61) and (3.62). For this, we remark that a first order expansion w.r.t. $d_N = \frac{L}{N}$ of $\tilde{\eta}_N$ and $\tilde{\delta}_N$ leads to*

$$\tilde{\eta}_N = L\left(\log(\frac{1}{1-c_N}) + \mathcal{O}(L/N)\right) \tag{3.66}$$

and to

$$\tilde{\delta}_N = \frac{L}{N}\frac{c_N}{1-c_N} + \mathcal{O}\left((L/N)^2\right) \tag{3.67}$$

which, of course, is in accordance with Theorem 1.

Asymptotic behaviour of η_N under hypothesis H_1

Under H_1, η_N is a linear statistics of the eigenvalues of matrix

$$\left(\mathbf{H} + \mathbf{V}_2/\sqrt{N}\right)\left(\mathbf{H} + \mathbf{V}_2/\sqrt{N}\right)^* (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1} \tag{3.68}$$

To the best of our knowledge, the asymptotic behaviour of the linear statistics of the eigenvalues of this matrix has not yet been studied in the asymptotic regime where L, M, N converge towards ∞ at the same rate. It is rather easy to evaluate an approximation of the empirical mean of η_N under H_1 using the results of [60]. However, to establish the asymptotic gaussianity of η_N and the expression of the corresponding variance, we need to establish a central limit theorem for linear statistics of the eigenvalues of non-zero mean large F-matrices. This needs an important work that is not in the scope of the present chapter, which is why we propose the following pragmatic approximation of the distribution of η_N.

Claim 1. *It is relevant to approximate the distribution of η_N under H_1 by a real Gaussian distribution with mean $\tilde{\eta}_N + \log\det\left(\mathbf{I} + \mathbf{H}^*\mathbf{H}/\sigma^2\right)$ and variance $\tilde{\delta}_N + \kappa_1/N$.*

Justification of Claim 1. As mentioned in Remark 1, when $M, N \to \infty$ and L is fixed, under H_1, the asymptotic mean $\overline{\eta}_{N,1}$ is the sum of the asymptotic mean under H_0 given by (3.30) and the second term $\log\det\left(\mathbf{I} + \mathbf{H}^*\mathbf{H}/\sigma^2\right)$. Thus, in the regime where $N, M, L \to \infty$, it seems reasonable to approximate the asymptotic mean of η_N by the sum of $\tilde{\eta}_N$ defined by (3.61) with the second term $\log\det\left(\mathbf{I} + \mathbf{H}^*\mathbf{H}/\sigma^2\right)$. We can reason similarly with the variance. The asymptotic variance under H_1, (3.47), is the sum of the asymptotic variance under H_0, outlined in Theorem 1, and the extra term $\frac{\kappa_1}{N}$. Therefore, the asymptotic variance under H_1 in the regime where $N, M, L \to \infty$ can be approximated by the asymptotic variance under H_0 for the same regime, plus the extra term $\frac{\kappa_1}{N}$. The results provided by this approximation are evaluated numerically in section 3.5.

For the reader's convenience, the main results of this chapter are summarized in Table 3.1, where $\tilde{\delta}_N$ is given by equation (3.62), κ_1 by equation (3.29) and $\tilde{\eta}_N$ by equation (3.61).

Assumption on parameters	Distribution under H_0	Distribution under H_1
(a) Classical, $N \to \infty$	$\eta_N \sim \frac{1}{N}\chi^2_{2ML}$ $\left(\mathbb{E}[\eta_N] = Lc_N, \mathrm{Var}[\eta_N] = Lc_N \cdot \frac{1}{N}\right)$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\left(\log\det\left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^*}{\sigma^2}\right), \frac{\kappa_1}{N}\right)$
(b) Proposed, $M, N \to \infty$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\left(L\log\frac{1}{1-c_N}, \frac{Lc_N}{1-c_N} \cdot \frac{1}{N}\right)$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\left(L\log\frac{1}{1-c_N} + \log\det\left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^*}{\sigma^2}\right), \frac{\kappa_1}{N} + \frac{Lc_N}{1-c_N}\cdot\frac{1}{N}\right)$
(c) Proposed, $L, M, N \to \infty$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\left(\tilde{\eta}_N, \tilde{\delta}_N\right)$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\left(\tilde{\eta}_N + \log\det\left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^*}{\sigma^2}\right), \frac{\kappa_1}{N} + \tilde{\delta}_N\right)$

Table 3.1: Asymptotic distribution of η_N for different assumptions, under H_0 and H_1

3.5 Numerical results.

In this section, we validate the relevance of the Gaussian approximations of section 3.4. In our numerical experiments, we have calculated the asymptotic expected values and variances as well as their empirical counterparts, evaluated by Monte Carlo simulations with 100.000 trials. In this section, to refer to the different approximations, we use the (a), (b) and (c) defined in table 3.1.

The fixed channel $\mathbf{H}$ is equal to $\mathbf{H} = \frac{1}{\left(\mathrm{Tr}(\overline{\mathbf{H}}\overline{\mathbf{H}}^*)\right)^{1/2}}\overline{\mathbf{H}}$ where $\overline{\mathbf{H}}$ is a realization of a $M \times L$ Gaussian random matrix with i.i.d. $\mathcal{N}_c(0, \frac{1}{M})$ entries. We remark that $\mathrm{Tr}(\mathbf{H}\mathbf{H}^*) = 1$.

The rows of the training sequence matrix $\mathbf{S}$ are chosen as cyclic shifts of a Zadoff-Chu sequence of length N [45]. Due to the autocorrelation properties of Zadoff-Chu sequences, designed so that the correlation between any shift of the sequence with itself is zero, we have $\mathbf{S}\mathbf{S}^*/N = \mathbf{I}_L$ if $L \leq N$.

3.5.1 Influence of $c_N = \frac{M}{N}$ on the asymptotic means and variances.

We first evaluate the behaviour of the means and variances of the three Gaussian approximations in terms of $c_N = \frac{M}{N}$. We only show the results for the asymptotic variance under H_1, but note that the results are similar for the expected values and under hypothesis H_0. Figure 1 compares the theoretical variances with the empirical variances obtained by simulation, under hypothesis H_1, as a function of c_N, the ratio between M and N. In this simulation, $M = 10$, $L = 5$ and $N = 20, 40, 60, 80, 160, 320$. When c_N is small, the three approximations (a), (b) and (c) give the same variance, as expected, and are very close to the empirical variance. When $c_N \geq \frac{1}{8}$, the assumption that M is small compared to N is no longer valid, and the classical asymptotic analysis (a) fails. The two large system approximations (b) and (c) provide similar results when $c_N \leq \frac{1}{4}$, i.e. when $N = 40$, or equivalently when $\frac{L}{N} \leq \frac{1}{8}$. However, when $N = 20$, i.e. $\frac{L}{N} = \frac{1}{4}$, (c), the approximation corresponding to the regime where L, M, N converge towards ∞ leads to a much more accurate prediction of the empirical variance. We remark that the approximation (c) is also reliable for rather small values of L, M, N, i.e. $L = 5, M = 10, N = 20$. We also remark that the regimes (b) and (c) where M, N are of the same order of magnitude capture the actual performance even when c_N is small, which, by extension, implies that the standard asymptotic analysis (a) always performs worse compared to the two large system approximations. If N, M increase while c_N stays the same, the results will be even closer to the theoretical values, since the number of samples is larger.

In the simulations that follow, we will use $c_N = 1/2$ with $N = 300$, $M = 150$ and $L = 10$, if not otherwise stated.

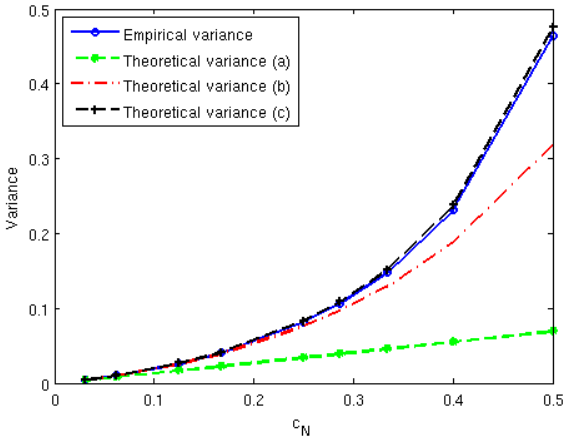
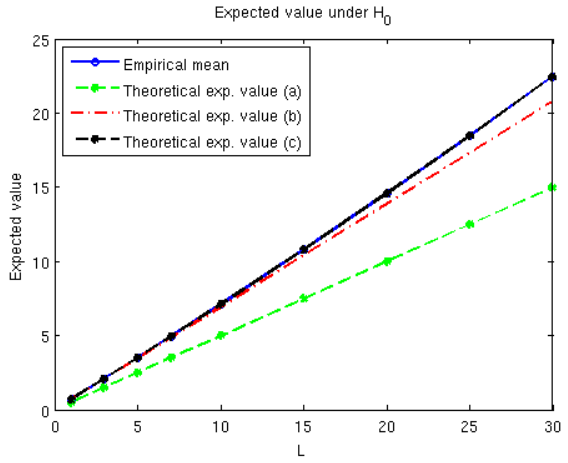
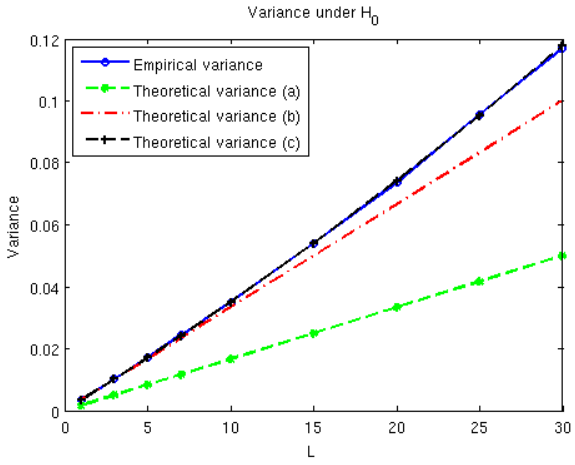


Figure 3.1: Proposed asymptotic analysis with standard asymptotic analysis

3.5.2 Comparison of the asymptotic means and variances of the approximations of η_N under H_0

We first compare in figures 2 and 3 the asymptotic expected values and variances with the empirical ones when L increases from $L = 1$ to $L = 30$ while $M = 150$ and $N = 300$, i.e. $c_N = 1/2$. The figures show that the standard asymptotic analysis of section 3.3 completely fails for all values of L. This is expected, given the value of $\frac{M}{N}$. As L increases, the assumption that L is small becomes increasingly invalid, and the only model that functions well in this regime is the model (c). This is valid both for the expected value and variance, and the theoretical values are very close to their empirical counterparts. We remark that the approximation (c), valid when $L \to +\infty$, also allows to capture the actual empirical performance when L is small.



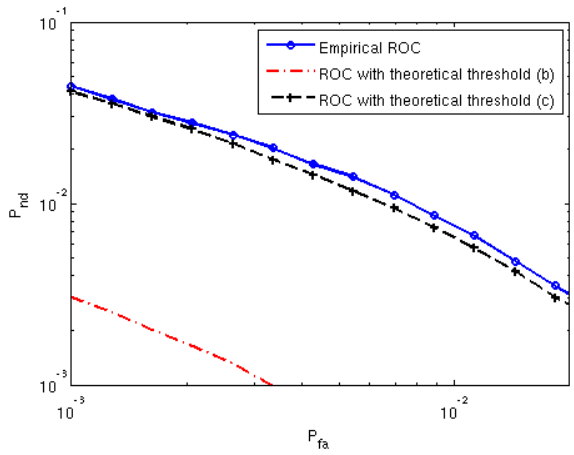


3.5.3 Validation of asymptotic distribution under H_0

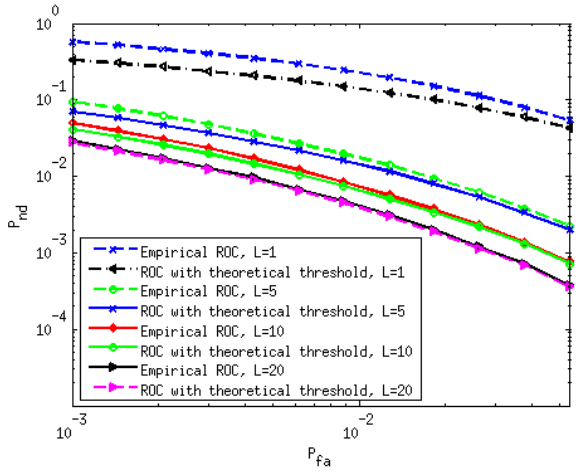
Although the expected values and variances can be very accurate, this does not necessarily mean that the empirical distribution is Gaussian. Therefore, we need to validate also the distribution under H_0. The asymptotic distribution under H_0 can be validated by analyzing its accuracy when calculating a threshold used to obtain ROC-curves. Note that this analysis also shows the applicability of the results for a practical case of timing synchronization.

We calculate the ROC curves in two different ways. The first is the ROC curve calculated empirically. We determine a threshold s from the empirical distribution under H_0 which gives a given probability of false alarm as $P_{fa} = \mathbb{P}(\eta_N > s)$. Its corresponding probability of non-detection, P_{nd}, is then obtained as the probability that the empirical values of the synchronization statistics under H_1 pass this threshold. The other ROC-curves are obtained by calculating the threshold s from the asymptotic Gaussian distributions under H_0, and using this theoretical threshold to calculate the P_{nd} from the empirical distribution under H_1.

Figure 4 shows the ROC-curves obtained with the approaches mentioned above when $L = 10, M = 150, N = 300$. Since the standard asymptotic analysis (a) gives very bad results, its results are omitted. It is clear that ROC-curve obtained by using the asymptotic distribution (b), obtained with the assumption that L is small, differs greatly from the results from the approximation (c), even for this relatively small value of L. This is because the theoretical threshold depends greatly on the expected value, and if it is not precisely evaluated, it gives erroneous results. In (c), the model where $N, M, L \to \infty$, the expected value and variance are very close to their empirical counterparts, and the resulting threshold can be used to precisely predict the synchronization performance for the set of parameters used when $P_{fa} \geq 10^{-3}$ and $P_{nd} > 10^{-3}$. Figure 5 shows, for the regime (c), the ROC curves obtained with the theoretical



threshold, together with the empirical results. In the figure, L goes from 1 to 20, while $M = 15L$ goes from 15 to 300 and $N = 30L$ goes from 30 to 600. It is seen that when the three parameters grow, the distance between the theoretical and empirical ROC curves decreases.

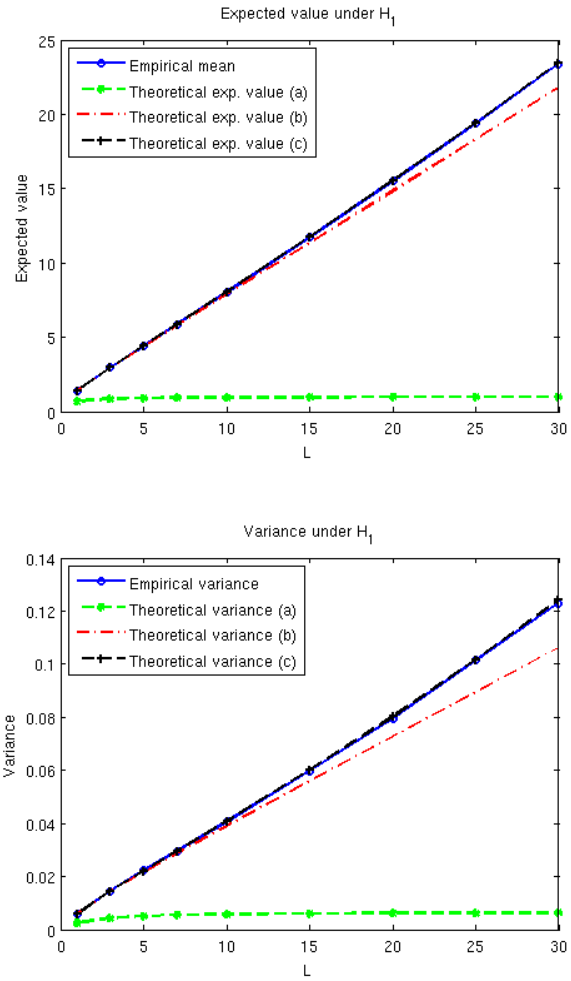


3.5.4 Comparison of the asymptotic means and variances of the approximations of η_N under H_1.

In this section, we will proceed to validate the expected value and variance under H_1.

Figures 6 and 7 validate the asymptotic expected values and variances under H_1. Similarly to hypothesis H_0, the theoretical expected values and variances are poorly evaluated using the standard asymptotic analysis (a). We note that the asymptotic expected values deduced for the regime (c) are very close to the empirical expected values and variances. For an L sufficiently

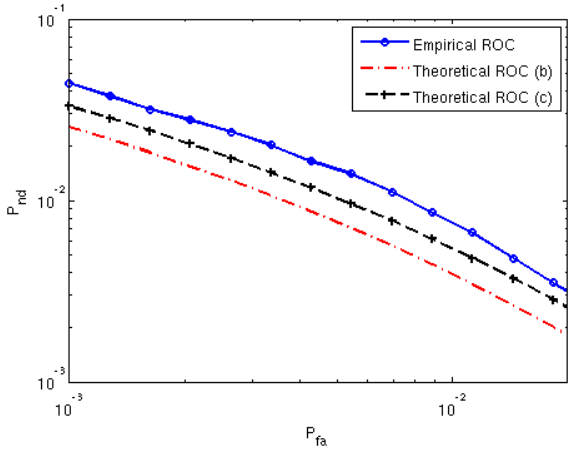
small, however, also the regime (b) gives asymptotic expected values and variances that are close to their empirical counterparts.



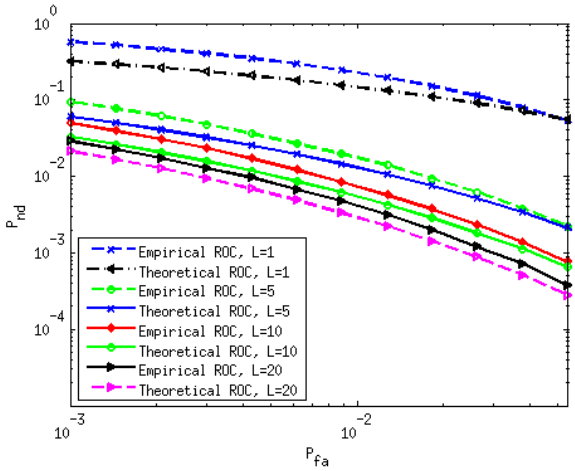
3.5.5 Validation of asymptotic distribution under H_1

To validate the asymptotic distributions under H_1, we calculate theoretical ROC-curves using both asymptotic distributions. For each P_{fa}, a threshold s is calculated from the theoretical Gaussian distribution under H_0. This threshold is then used to calculate the P_{nd} from the theoretical Gaussian distribution under H_1, using $P_{nd} = 1 - \mathbb{P}_{H_1}(\eta_N > s)$. Figure 8 shows these theoretical ROC curves plotted together with the empirical ROC curve. Here, $L = 10, M = 150$ and $N = 300$. It is seen that the approximation corresponding to the regime $N, M, L \to \infty$ provides, as in the context of hypothesis H_0, a more accurate theoretical ROC curve. It is seen that the ROC curve associated with the regime small L (b) is closer from the empirical

ROC curve than in the context of hypothesis H_0. This is because the corresponding asymptotic means are, for both H_0 and H_1, less than the actual empirical means. These two errors tend to compensate in the theoretical ROC curves (b), which explains why the theoretical ROC curve (b) of figure 8 is more accurate than the corresponding ROC curve of figure 4, for small L.



We now evaluate the behaviour of the ROC curves when N, M, L grow at the same rate. In figure 9, L goes from 1 to 20, while $M = 15L$ goes from 15 to 300 and $N = 30L$ goes from 30 to 600. The results show that as N, M, L grow proportionally, the theoretical results tend to approach the empirical values, but that, in contrast with the context of figure 5, a residual error remains. It would be interesting to evaluate more accurately the asymptotic behaviour of η_N under H_1 in the regime $L \to +\infty$, and to check if the residual error tends to diminish. However, as mentioned in Paragraph 3.4.2, this needs to establish a central limit theorem for linear statistics of the eigenvalues of non zero mean large F-matrices, which is a non trivial task.



3.6 Conclusion.

In this chapter, we have studied the behaviour of the multi-antenna GLR detection test statistics η_N of a known signal corrupted by a multi-path deterministic channel and an additive white Gaussian noise with unknown spatial covariance. We have addressed the case where the number of sensors M and the number of samples N of the training sequence converge towards ∞ at the same rate. When the number of paths L does not scale with M and N, we have established that η_N has a Gaussian behaviour with asymptotic mean $L \log \frac{1}{1-M/N}$ and variance $\frac{L}{N} \frac{M/N}{1-M/N}$. This is in contrast with the standard regime $N \to +\infty$ and M fixed where η_N has a χ^2 behaviour. Under hypothesis H_1, η_N has still a Gaussian behaviour. The corresponding asymptotic mean and variance are obtained as the sum of the asymptotic mean and variance in the standard regime $N \to +\infty$ and M fixed, and $L \log \frac{1}{1-M/N}$ and $\frac{L}{N} \frac{M/N}{1-M/N}$ respectively, i.e. the asymptotic mean and variance under H_0. We have also considered the case where the number of paths L converges towards ∞ at the same rate as M and N. Using known results of [57] and [58], concerning the behaviour of linear statistics of the eigenvalues of large F-matrices, we have deduced that in the regime where L, M, N converge to ∞ at the same rate, η_N still has a Gaussian behaviour under H_0, but with a different mean and variance. The analysis of η_N under H_1 when L, M, N converge to ∞ needs to establish a central limit theorem for linear statistics of the eigenvalues of large non zero-mean F-matrices, a difficult task that we will address in a future work. Motivated by the results obtained in the case where L remains finite, we have proposed to approximate the asymptotic distribution of η_N by a Gaussian distribution whose mean and variance are the sum of the asymptotic mean and variance under H_0 when $L \to +\infty$ with the asymptotic mean and variance under H_1 in the standard regime $N \to +\infty$ and M fixed. Numerical experiments have shown that the Gaussian approximation corresponding to the standard regime $N \to +\infty$ and M fixed completely fails as soon as $\frac{M}{N}$ is not small enough. The large system approximations provide better results when $\frac{M}{N}$ increases, while also allowing to capture the actual performance for small values of $\frac{M}{N}$. We have also observed that, for finite values of L, M, N, the Gaussian approximation obtained in the regime L, M, N converge towards ∞ is more accurate than the approximation in which L is fixed. In particular, the ROC curves that are obtained using the former large system approximation are accurate approximations of the empirical ones in a reasonable range of P_{fa}, P_{nd}. We therefore believe that our results can be used to reliably predict the performance of the GLRT, and that the tools that are developed in this chapter are useful in the context of large antenna arrays.

3.7 Appendix: Useful technical results.

In this appendix, we provide some useful technical results concerning the behaviour of certain large random matrices. In the remainder of this appendix, $\mathbf{\Sigma}_N$ represents a $M \times N$ matrix with $\mathcal{N}_{\mathbb{C}}(0, \frac{\sigma^2}{N})$ i.i.d. elements. We of course assume in this section that M and N both converge towards $+\infty$ in such a way that $c_N = \frac{M}{N} < 1$ converges towards $c < 1$. In the following, we give some results concerning the behaviour of the eigenvalues $\hat{\lambda}_{1,N} \leq \hat{\lambda}_{2,N} \ldots \leq \hat{\lambda}_{M,N}$ of the matrix $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ as well as on its resolvent $\mathbf{Q}_N(z)$ defined for $z \in \mathbb{C} - \mathbb{R}^+$ by

$$\mathbf{Q}_N(z) = (\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* - z\mathbf{I}_M)^{-1} \tag{3.69}$$

We first state the following classical result (see e.g. [57], Theorem 5.11).

Proposition 1. *When $N \to +\infty$, $\hat{\lambda}_{1,N}$ converges almost surely towards $\sigma^2(1-\sqrt{c})^2$ while $\hat{\lambda}_{M,N}$ converges a.s. to $\sigma^2(1+\sqrt{c})^2$.*

In the following, we denote by $\mathcal{I}_\epsilon$ the interval defined by

$$\mathcal{I}_\epsilon = [\sigma^2(1-\sqrt{c})^2 - \epsilon, \sigma^2(1+\sqrt{c})^2 + \epsilon] \tag{3.70}$$

(with ϵ chosen in such a way that $\sigma^2(1-\sqrt{c})^2 - \epsilon > 0$) and by $\mathcal{E}_N$ the event defined by

$$\mathcal{E}_N = \{\text{one of the } (\hat{\lambda}_{k,N})_{k=1,\ldots,M} \text{ escapes from } \mathcal{I}_\epsilon\} \tag{3.71}$$

and remark that the almost sure convergence of $\hat{\lambda}_{1,N}$ and $\hat{\lambda}_{M,N}$ implies that

$$\begin{aligned} \mathbb{1}_{\mathcal{E}_N^c} = 1 \text{ almost surely for each } N \\ \text{larger than a random integer} \end{aligned} \tag{3.72}$$

Proposition 1 implies that the resolvent $\mathbf{Q}_N(z)$ is almost surely defined on $\mathbb{C} - \mathcal{I}_\epsilon$ for N large enough, and in particular for $z = 0$.

Another important property is the almost sure convergence of the empirical eigenvalue distribution $\hat{\mu}_N = \frac{1}{M}\sum_{k=1}^M \delta_{\hat{\lambda}_{k,N}}$ of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ towards the Marcenko-Pastur distribution (see e.g. [57] and [30] and the references therein). Formally, this means that the Stieltjes transform $\hat{m}_N(z)$ of $\hat{\mu}_N$ defined by

$$\hat{m}_N(z) = \int_{\mathbb{R}} \frac{d\hat{\mu}_N(\lambda)}{\lambda - z} = \frac{1}{M}\mathrm{Tr}\left(\mathbf{Q}_N(z)\right) \tag{3.73}$$

satisfies

$$\lim_{N \to +\infty} (\hat{m}_N(z) - m_{c_N}(z)) = 0 \tag{3.74}$$

almost surely for each $z \in \mathbb{C} - \mathbb{R}^+$ (and uniformly on each compact subset of $\mathbb{C} - \mathbb{R}^+$), where $m_{c_N}(z)$ represents the Stieltjes transform of the Marcenko-Pastur distribution of parameter c_N, denoted by μ_{c_N} in the following. $m_{c_N}(z)$ satisfies the following fundamental equation

$$m_{c_N}(z) = \frac{1}{-z\left(1 + \sigma^2 c_N m_{c_N}(z)\right) + \sigma^2(1 - c_N)} \tag{3.75}$$

for each $z \in \mathbb{C}$. μ_{c_N} is known to be absolutely continuous, its support is the interval $[\sigma^2(1 - \sqrt{c_N})^2, \sigma^2(1 + \sqrt{c_N})^2]$, and its density is given by

$$\frac{\sqrt{\left(x - x_{c_N}^-\right)\left(x_{c_N}^+ - x\right)}}{2\sigma^2 c_N \pi x} \mathbb{1}_{[x_{c_N}^-, x_{c_N}^+]}(x). \tag{3.76}$$

with $x_{c_N}^- = \sigma^2(1 - \sqrt{c_N})^2$ and $x_{c_N}^+ = \sigma^2(1 + \sqrt{c_N})^2$. As μ_{c_N} is supported by $[\sigma^2(1 - \sqrt{c_N})^2, \sigma^2(1 + \sqrt{c_N})^2]$, the almost sure convergence (3.74) holds not only on $\mathbb{C} - \mathbb{R}^+$, but also for each $z \in \mathbb{C} - \mathcal{I}_\epsilon$. In particular, (3.74) is valid for $z = 0$. Solving the equation (3.75) for $z = 0$ leads immediately to $m_{c_N}(0) = \frac{1}{\sigma^2(1 - c_N)}$, and to

$$\lim_{N \to +\infty} \frac{1}{M} \mathrm{Tr}\left(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*\right)^{-1} - \frac{1}{\sigma^2(1 - c_N)} = 0 \tag{3.77}$$

almost surely. Taking the derivative of (3.74) w.r.t. z at $z = 0$, and using that $m'_{c_N}(0) = \frac{1}{\sigma^4(1 - c_N)^3}$, we also obtain that

$$\lim_{N \to +\infty} \frac{1}{M} \mathrm{Tr}\left(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*\right)^{-2} - \frac{1}{\sigma^4(1 - c_N)^3} = 0 \tag{3.78}$$

almost surely. Moreover, it is possible to specify the convergence speed in (3.77) and (3.78). The following proposition is a direct consequence of Theorem 9.10 in [57].

Proposition 2. *It holds that*

$$\frac{1}{M} \mathrm{Tr}\left(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*\right)^{-1} - \frac{1}{\sigma^2(1 - c_N)} = \mathcal{O}_P(\frac{1}{N}) \tag{3.79}$$

$$\frac{1}{M} \mathrm{Tr}\left(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*\right)^{-2} - \frac{1}{\sigma^4(1 - c_N)^3} = \mathcal{O}_P(\frac{1}{N}) \tag{3.80}$$

Theorem 9.10 in [57] implies that the left hand side of (3.79), renormalized by N, converges in distribution towards a Gaussian distribution, which, in turn, leads to (3.79). (3.80) holds for the same reason.

Remark 5. *As $c_N \to c$, the previous results of course imply that $\frac{1}{M} \mathrm{Tr}\left(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*\right)^{-1}$ (resp. $\frac{1}{M} \mathrm{Tr}\left(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*\right)^{-2}$) converge towards $\frac{1}{\sigma^2(1-c)}$ (resp. $\frac{1}{\sigma^4(1-c)^3}$). However, the rate of convergence is not a $\mathcal{O}_P(\frac{1}{N})$ term if the convergence speed of c_N towards c is less than $\mathcal{O}(\frac{1}{N})$. Therefore, it is more relevant to approximate the left hand sides of (3.79) and (3.80) by $\frac{1}{\sigma^2(1-c_N)}$ and $\frac{1}{\sigma^4(1-c_N)^3}$.*

The above results allow to characterize the asymptotic behaviour of the normalized trace of $(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*)^{-1}$ and $(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*)^{-2}$. However, it is also useful to obtain similar results on the bilinear forms of these matrices.

Proposition 3. *We consider two deterministic M-dimensional unit norm vectors $\mathbf{u}_N$ and $\mathbf{v}_N$. Then, it holds that*

$$\lim_{N\to+\infty} \mathbf{u}_N^* \left(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*\right)^{-1} \mathbf{v}_N - \frac{\mathbf{u}_N^*\mathbf{v}_N}{\sigma^2(1-c_N)} = 0 \tag{3.81}$$

and that

$$\lim_{N\to+\infty} \mathbf{u}_N^* \left(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*\right)^{-2} \mathbf{v}_N - \frac{\mathbf{u}_N^*\mathbf{v}_N}{\sigma^4(1-c_N)^3} = 0 \tag{3.82}$$

almost surely. Moreover,

$$\mathbf{u}_N^* \left(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*\right)^{-1} \mathbf{v}_N - \frac{\mathbf{u}_N^*\mathbf{v}_N}{\sigma^2(1-c_N)} = \mathcal{O}_P(\frac{1}{\sqrt{N}}) \tag{3.83}$$

Finally, if $\mathbf{C}_N$ is a positive $M\times M$ matrix such that $\mathrm{Rank}(\mathbf{C}_N) = K$ is independent of N, and satisfying for each N $0 < d_1 \leq \mathrm{Tr}(\mathbf{C}_N^2) < d_2 < \infty$ for some constants d_1 and d_2, then, we consider the sequence of random variables $(\kappa_N)_{N\geq 1}$ defined by

$$\kappa_N = \mathrm{Tr}\left(\mathbf{C}_N(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*)^{-1}\right) \tag{3.84}$$

Define by θ_N the term

$$\theta_N = \frac{\mathrm{Tr}(\mathbf{C}_N^2)}{\sigma^4(1-c_N)^3} \tag{3.85}$$

Then, it holds that

$$\mathbb{E}\left[\exp\left(iu\sqrt{N}\left(\kappa_N - \frac{\mathrm{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)}\right)\right)\right] - \exp\left(-\frac{\theta_N u^2}{2}\right) \to 0 \tag{3.86}$$

for each $u\in\mathbb{R}$, and that

$$\frac{\sqrt{N}}{\sqrt{\theta_N}}\left(\kappa_N - \frac{\mathrm{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)}\right) \to_{\mathcal{D}} \mathcal{N}(0,1) \tag{3.87}$$

The almost sure convergence result (3.81) is well known (see e.g. [61] in the context of a more general matrix model), while (3.82) can be established by differentiating the behaviour of the bilinear forms of $\mathbf{Q}_N(z)$ w.r.t. z. Moreover, (3.83) is a consequence of (3.87) used for the rank 1 matrix $\mathbf{C}_N = \mathbf{v}_N\mathbf{u}_N^*$. (3.86) and (3.87) are new and need to be established.

A technical difficulty appears in the present context because we consider the resolvent of the matrix $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ at $z=0$ while in previous works, z is supposed to belong to $\mathbb{C}-\mathbb{R}^+$. To solve

this issue, we use the regularization technic introduced in a more general context in [62]. For the proof, we refer the reader to Appendix 3.9.

We finish this appendix by a standard result whose proof is omitted.

Proposition 4. *We consider a $M \times L$ random matrix $\mathbf{\Gamma}_N$ with $\mathcal{N}_{\mathbb{C}}(0, \frac{\sigma^2}{N})$ i.i.d. entries, as well as the following deterministic matrices: $\mathbf{A}_N$ is $M \times M$ and hermitian, $\mathbf{B}_N$ is $M \times L$ and satisfies $\sup_N \|\mathbf{B}_N\| < +\infty$ while $\mathbf{D}_N$ is a positive $L \times L$ matrix and also verifies $\sup_N \|\mathbf{D}_N\| < +\infty$. Then, if $(\omega_N)_{N \geq 1}$ represents the sequence of random variables defined by*

$$\omega_N = \mathrm{Tr}\left[\mathbf{D}_N \left(\mathbf{\Gamma}_N^* \mathbf{A}_N \mathbf{\Gamma}_N + \mathbf{\Gamma}_N^* \mathbf{B}_N + \mathbf{B}_N^* \mathbf{\Gamma}_N\right)\right] \tag{3.88}$$

it holds that

$$\mathbb{E}(\omega_N) = \sigma^2 \frac{1}{N} \mathrm{Tr}(\mathbf{A}_N)\, \mathrm{Tr}(\mathbf{D}_N), \tag{3.89}$$

$$\mathrm{Var}(\omega_N) = \frac{1}{N} \zeta_N$$

where ζ_N is defined by

$$\zeta_N = \sigma^4 \frac{1}{N} \mathrm{Tr}(\mathbf{A}_N^2) \mathrm{Tr}(\mathbf{D}_N^2) + 2\sigma^2 \frac{1}{N} \mathrm{Tr}\left(\mathbf{D}_N^2 \mathbf{B}_N^* \mathbf{B}_N\right) \tag{3.90}$$

Moreover,

$$\begin{aligned} \mathbb{E}\left|\omega_N - \mathbb{E}(\omega_N)\right|^4 \leq & \frac{a_1}{N^2} + \frac{a_2}{N^2} \left(\frac{1}{N} \mathrm{Tr}(\mathbf{A}_N^2)\right)^2 \\ & + \frac{a_3}{N^3} \frac{1}{N} \mathrm{Tr}(\mathbf{A}_N^8) \end{aligned} \tag{3.91}$$

where a_1, a_2, a_3 are constant terms depending on $L, \sup_N \|\mathbf{B}_N\|$ and $\sup_N \|\mathbf{D}_N\|$. Finally, if $\limsup_N \zeta_N < +\infty$, it holds that

$$\mathbb{E}\left(\exp i\, u \sqrt{N} \left(\omega_N - \mathbb{E}(\omega_N)\right)\right) - e^{-\frac{u^2 \zeta_N}{2}} \to 0 \tag{3.92}$$

for each $u \in \mathbb{R}$.

3.8 Appendix: Proofs of Theorems 1 and 2

Proof of Theorem 1. In order to establish Theorem 1, we use the results of Appendix 3.7 for the matrix $\mathbf{\Sigma}_N = \frac{1}{\sqrt{N}} \mathbf{V}_1$. We note that $\frac{1}{\sqrt{N}} \mathbf{V}_1$ is a $M \times (N - L)$ matrix while the results of Appendix 3.7 have been presented in the context of a $M \times N$ matrix. In principle, it should be necessary to exchange N by $N - L$ in Propositions 1 to 3. However, $c_N - \frac{M}{N-L} = \mathcal{O}(\frac{1}{N})$, so that it possible to use the results of the above propositions without exchanging N by $N - L$.

We first verify (3.30). For this, we introduce the event $\mathcal{E}_N$ defined by (3.71). We first remark that $\eta_N - \eta_N \mathbb{1}_{\mathcal{E}_N^c} \to 0, a.s.$ It is thus sufficient to study the behaviour of $\eta_N \mathbb{1}_{\mathcal{E}_N^c}$ which is also equal to

$$\eta_N \mathbb{1}_{\mathcal{E}_N^c} = \log\det\left(\mathbf{I} + \mathbf{F}_N \mathbb{1}_{\mathcal{E}_N^c}\right) \tag{3.93}$$

We now study the behaviour of each entry (k,l) of matrix $\mathbb{1}_{\mathcal{E}_N^c}\mathbf{F}_N$. For this, we use Proposition 4 for $\mathbf{D}_N = \mathbf{e}_k\mathbf{e}_l^T$, $\mathbf{\Gamma}_N = \frac{\mathbf{V}_2}{\sqrt{N}}$ and $\mathbf{A}_N = \mathbb{1}_{\mathcal{E}_N^c}\left(\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}\right)^{-1}$. $\mathbf{A}_N$ is of course not deterministic, but as $\mathbf{V}_2$ and $\mathbf{V}_1$ are independent, it is possible to use the results of Proposition 4 by replacing the mathematical expectation operator by the mathematical expectation operator $\mathbb{E}_{\mathbf{V}_2}$ w.r.t. $\mathbf{V}_2$. We note that the present matrix $\mathbf{A}_N$ verifies

$$\mathbf{A}_N \leq \frac{\mathbf{I}}{\sigma^2(1-\sqrt{c})^2 - \epsilon} \tag{3.94}$$

because $\mathbb{1}_{\mathcal{E}_N^c} \neq 0$ implies that all the eigenvalues of $\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}$ belong to $\mathcal{I}_\epsilon = [\sigma^2(1-\sqrt{c})^2 - \epsilon, \sigma^2(1+\sqrt{c})^2 + \epsilon]$. Therefore, (3.91) immediately implies that

$$\mathbb{E}_{\mathbf{V}_2}\left|\mathbf{F}_{N,k,l}\mathbb{1}_{\mathcal{E}_N^c} - \mathbb{E}_{\mathbf{V}_2}\left(\mathbf{F}_{N,k,l}\mathbb{1}_{\mathcal{E}_N^c}\right)\right|^4 \leq \frac{a}{N^2} \tag{3.95}$$

where a is a deterministic constant. Taking the mathematical expectation of the above inequality w.r.t. $\mathbf{V}_1$, and using the Borel-Cantelli Lemma lead to

$$\mathbf{F}_{N,k,l}\mathbb{1}_{\mathcal{E}_N^c} - \mathbb{E}_{\mathbf{V}_2}\left(\mathbf{F}_{N,k,l}\mathbb{1}_{\mathcal{E}_N^c}\right) \to 0 \ a.s. \tag{3.96}$$

or equivalently, to

$$\mathbf{F}_{N,k,l}\mathbb{1}_{\mathcal{E}_N^c} - \delta(k-l)\,\sigma^2 c_N\,\frac{1}{M}\mathrm{Tr}\left(\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}\right)^{-1} \to 0 \ a.s. \tag{3.97}$$

(3.77) implies that $\mathbf{F}_{N,k,l}\mathbb{1}_{\mathcal{E}_N^c} - \delta(k-l)\,\frac{c_N}{1-c_N} \to 0$ almost surely, or equivalently that

$$\mathbf{F}_N - \frac{c_N}{1-c_N}\mathbf{I} \to 0 \ a.s. \tag{3.98}$$

This eventually leads to (3.30).

We now establish (3.31). For this, we first remark that (3.72) implies that $\eta_N = \eta_N\mathbb{1}_{\mathcal{E}_N^c} + \mathcal{O}_P(\frac{1}{N^p})$ for each integer p. Therefore, the asymptotic behaviour of the distribution of the left hand side of (3.31) is not modified if η_N is replaced by $\eta_N\mathbb{1}_{\mathcal{E}_N^c}$ given by (3.93). We denote by $\mathbf{\Delta}_N$ the matrix defined by

$$\mathbf{\Delta}_N = \mathbf{F}_N\mathbb{1}_{\mathcal{E}_N^c} - \frac{c_N}{1-c_N}\mathbf{I} \tag{3.99}$$

We first prove that $\mathbf{\Delta}_N = \mathcal{O}_P(\frac{1}{\sqrt{N}})$. For this, we express $\mathbf{\Delta}_N$ as

$$\mathbf{\Delta}_N = \left(\mathbf{F}_N \mathbb{1}_{\mathcal{E}_N^c} - \sigma^2 c_N \frac{1}{M}\mathrm{Tr}\left(\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}\right)^{-1}\mathbb{1}_{\mathcal{E}_N^c}\mathbf{I}\right) + \sigma^2 c_N \frac{1}{M}\mathrm{Tr}\left(\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}\right)^{-1}\mathbb{1}_{\mathcal{E}_N^c}\mathbf{I} - \frac{c_N}{1-c_N}\mathbf{I} \quad (3.100)$$

The first term of the right hand side of (3.100) is $\mathcal{O}_P(\frac{1}{\sqrt{N}})$ because the fourth-order moments of its entries are $\mathcal{O}(\frac{1}{N^2})$ terms. As for the second term, (3.79) implies that it is a $\mathcal{O}_P(\frac{1}{N})$. A standard second order expansion of $\log\det(\mathbf{I} + \mathbf{F}_N\mathbb{1}_{\mathcal{E}_N^c})$ leads to

$$\eta_N \mathbb{1}_{\mathcal{E}_N^c} = L\log\frac{1}{1-c_N} + (1-c_N)\mathrm{Tr}(\mathbf{\Delta}_N) + \mathcal{O}_P(\frac{1}{N}) \quad (3.101)$$

Therefore, it holds that

$$\sqrt{N}\left(\eta_N\mathbb{1}_{\mathcal{E}_N^c} - L\log\frac{1}{1-c_N}\right) = \sqrt{N}(1-c_N)\mathrm{Tr}(\mathbf{\Delta}_N) + \mathcal{O}_P(\frac{1}{\sqrt{N}}), \quad (3.102)$$

or, using (3.100), that

$$\sqrt{N}\left(\eta_N\mathbb{1}_{\mathcal{E}_N^c} - L\log\frac{1}{1-c_N}\right) = \sqrt{N}(1-c_N)\mathrm{Tr}\Big(\mathbf{F}_N\mathbb{1}_{\mathcal{E}_N^c} - \sigma^2 c_N\frac{1}{M}\mathrm{Tr}\left(\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}\right)^{-1}\mathbb{1}_{\mathcal{E}_N^c}\Big) + \mathcal{O}_P(\frac{1}{\sqrt{N}}) \quad (3.103)$$

As

$$\mathbb{E}_{\mathbf{V}_2}\left(\mathrm{Tr}\left(\mathbf{F}_N\mathbb{1}_{\mathcal{E}_N^c}\right)\right) = \sigma^2 c_N\frac{1}{M}\mathrm{Tr}\left(\left(\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}\right)^{-1}\mathbb{1}_{\mathcal{E}_N^c}\right), \quad (3.104)$$

Proposition 4 used for $\mathbf{A}_N = \left(\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}\right)^{-1}\mathbb{1}_{\mathcal{E}_N^c}$, $\mathbf{B}_N = 0$ and $\mathbf{D}_N = (1-c_N)\mathbf{I}$ leads to

$$\mathbb{E}_{\mathbf{V}_2}\left(\exp iu\sqrt{N}\left(\eta_N - L\log\frac{1}{1-c_N}\right)\right) - \exp\left[-\frac{u^2}{2}\sigma^4 L(1-c_N)^2 c_N\frac{1}{M}\mathrm{Tr}\left(\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}\right)^{-2}\mathbb{1}_{\mathcal{E}_N^c}\right] \to 0 \quad (3.105)$$

a.s. for each $u \in \mathbb{R}$. (3.78) and the dominated convergence theorem finally implies that

$$\mathbb{E}\left(\exp iu\sqrt{N}\left(\eta_N - L\log\frac{1}{1-c_N}\right)\right) - \exp\left[-\frac{u^2}{2}\frac{Lc_N}{1-c_N}\right] \to 0 \quad (3.106)$$

This establishes (3.31).

Proof of Theorem 2 We recall that, under H_1, η_N is given by (3.22). As in the proof of Theorem 1, it is sufficient to study the regularized statistics $\eta_N \mathbb{1}_{\mathcal{E}_N^c}$ which is also equal to

$$\eta_N \mathbb{1}_{\mathcal{E}_N^c} = \log\det\left(\mathbf{I}_L + \mathbb{1}_{\mathcal{E}_N^c}\,\mathbf{G}_N\right) \tag{3.107}$$

In order to evaluate the almost sure behaviour of $\eta_N \mathbb{1}_{\mathcal{E}_N^c}$, we expand $\mathbf{G}_N \mathbb{1}_{\mathcal{E}_N^c}$ as

$$\begin{aligned}\mathbf{G}_N \mathbb{1}_{\mathcal{E}_N^c} = \mathbf{H}^* \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1} \mathbf{H}\, \mathbb{1}_{\mathcal{E}_N^c} + \mathbf{F}_N \mathbb{1}_{\mathcal{E}_N^c} + \\ (\mathbf{V}_2/\sqrt{N})^* \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1} \mathbf{H}\, \mathbb{1}_{\mathcal{E}_N^c} + \\ \mathbf{H}^* \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1} (\mathbf{V}_2/\sqrt{N})\, \mathbb{1}_{\mathcal{E}_N^c}\end{aligned} \tag{3.108}$$

By (3.81), the first term of the right hand side of (3.108) behaves almost surely as $\frac{\mathbf{H}^*\mathbf{H}}{\sigma^2(1-c_N)}$, while it has been shown before that the second term converges a.s. towards $\frac{c_N}{1-c_N}\mathbf{I}$. To address the behaviour of entry (k,l) of the sum of the third and the fourth terms, we use Proposition 4 for $\mathbf{\Gamma}_N = \frac{\mathbf{V}_2}{\sqrt{N}}$, $\mathbf{A}_N = 0$, $\mathbf{B}_N = \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\mathbf{H}\, \mathbb{1}_{\mathcal{E}_N^c}$ and $\mathbf{D}_N = \mathbf{e}_k\mathbf{e}_l^T$. (3.91) implies that entry (k,l) converges almost surely towards 0. Therefore, we have proved that

$$\mathbf{G}_N - \left(\frac{\mathbf{H}^*\mathbf{H}}{\sigma^2(1-c_N)} + \frac{c_N}{1-c_N}\mathbf{I}\right) \to \mathbf{0}\ a.s. \tag{3.109}$$

from which (3.45) follows immediately.

The proof of (3.47) is similar to the proof of (3.31), thus we do not provide all the details. We replace η_N by $\eta_N \mathbb{1}_{\mathcal{E}_N^c}$, and remark that the matrix $\mathbf{\Delta}_N$, given by

$$\mathbf{\Delta}_N = \mathbf{G}_N \mathbb{1}_{\mathcal{E}_N^c} - \left(\frac{\mathbf{H}^*\mathbf{H}}{\sigma^2(1-c_N)} + \frac{c_N}{1-c_N}\mathbf{I}\right) \tag{3.110}$$

verifies $\mathbf{\Delta}_N = \mathcal{O}_P(\frac{1}{\sqrt{N}})$. To check this, it is sufficient to use the expansion (3.49), and to recognize that:

- by (3.83),

$$\mathbf{H}^* \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1} \mathbf{H}\, \mathbb{1}_{\mathcal{E}_N^c} - \frac{\mathbf{H}^*\mathbf{H}}{\sigma^2(1-c_N)} = \mathcal{O}_P(\frac{1}{\sqrt{N}}), \tag{3.111}$$

- by Proposition 4 and (3.91),

$$\begin{aligned}(\mathbf{V}_2/\sqrt{N})^* \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1} \mathbf{H}\, \mathbb{1}_{\mathcal{E}_N^c} + \\ \mathbf{H}^* \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1} (\mathbf{V}_2/\sqrt{N})\, \mathbb{1}_{\mathcal{E}_N^c} = \mathcal{O}_P(\frac{1}{\sqrt{N}})\end{aligned} \tag{3.112}$$

- it has been shown before that

$$\mathbf{F}_N \mathbb{1}_{\mathcal{E}_N^c} - \frac{c_N}{1-c_N}\mathbf{I} = \mathcal{O}_P(\frac{1}{\sqrt{N}}). \tag{3.113}$$

Using a standard linearization of $\log\det(\mathbf{I} + \mathbf{G}_N \mathbb{1}_{\mathcal{E}_N^c})$, this implies that

$$\eta_N \mathbb{1}_{\mathcal{E}_N^c} - \overline{\eta}_{N,1} = \mathrm{Tr}\left(\mathbf{D}_N \mathbf{\Delta}_N\right) + \mathcal{O}_P(1/N) \tag{3.114}$$

where $\mathbf{D}_N$ is the $L \times L$ matrix given by

$$\mathbf{D}_N = (1-c_N)(\mathbf{I}_L + \mathbf{H}^*\mathbf{H}/\sigma^2)^{-1} \tag{3.115}$$

We define κ_N and ω_N by

$$\kappa_N = \mathrm{Tr}\left(\mathbf{C}_N \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\right) \tag{3.116}$$

and

$$\begin{aligned}\omega_N = \mathrm{Tr}\left[\mathbf{D}_N\mathbf{F}_N\mathbb{1}_{\mathcal{E}_N^c}\right] + \\ \mathrm{Tr}\left[\mathbf{D}_N(\mathbf{V}_2/\sqrt{N})^* \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\mathbf{H}\,\mathbb{1}_{\mathcal{E}_N^c}\right] + \\ \mathrm{Tr}\left[\mathbf{D}_N\,\mathbf{H}^*\left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\left(\mathbf{V}_2/\sqrt{N}\right)\mathbb{1}_{\mathcal{E}_N^c}\right]\end{aligned} \tag{3.117}$$

where $\mathbf{C}_N$ the $M \times M$ matrix given by

$$\mathbf{C}_N = (1-c_N)\mathbf{H}(\mathbf{I}_L + \mathbf{H}^*\mathbf{H}/\sigma^2)^{-1}\mathbf{H}^* \tag{3.118}$$

Using (3.114), we obtain that

$$\begin{aligned}\eta_N \mathbb{1}_{\mathcal{E}_N^c} - \overline{\eta}_{N,1} = \kappa_N - \frac{\mathrm{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)} \; + \\ \omega_N - \frac{c_N}{1-c_N}\mathrm{Tr}(\mathbf{D}_N) + \mathcal{O}_P(\frac{1}{N})\end{aligned} \tag{3.119}$$

We also remark that (3.79) used for $\mathbf{\Sigma}_N = \frac{1}{\sqrt{N}}\mathbf{V}_1$ implies that

$$\omega_N - \mathbb{E}_{\mathbf{V}_2}(\omega_N) = \omega_N - \frac{c_N}{1-c_N}\mathrm{Tr}(\mathbf{D}_N) + \mathcal{O}_P(\frac{1}{N}) \tag{3.120}$$

Therefore, it holds that

$$\sqrt{N}\left(\eta_N \mathbb{1}_{\mathcal{E}_N^c} - \overline{\eta}_{N,1}\right) = \sqrt{N}\left(\mathrm{Tr}\left(\mathbf{D}_N\mathbf{\Delta}_N\right)\right) \tag{3.121}$$

can be written as

$$\sqrt{N}\left(\eta_N \mathbb{1}_{\mathcal{E}_N^c} - \overline{\eta}_{N,1}\right) = \sqrt{N}\left(\kappa_N - \frac{\mathrm{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)}\right) + \sqrt{N}\left(\omega_N - \mathbb{E}_{\mathbf{V}_2}(\omega_N)\right) + \mathcal{O}_P(\frac{1}{\sqrt{N}}) \quad (3.122)$$

We denote by ζ_N the term

$$\zeta_N = \sigma^4 \frac{1}{N}\mathrm{Tr}\left((\mathbf{V}_1\mathbf{V}_1^*/N)^{-2}\,\mathbb{1}_{\mathcal{E}_N^c}\right)\,\mathrm{Tr}(\mathbf{D}_N^2) + 2\sigma^2\frac{1}{N}\mathrm{Tr}\left(\mathbf{D}_N^2\mathbf{H}^*(\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\mathbf{H}\,\mathbb{1}_{\mathcal{E}_N^c}\right) \quad (3.123)$$

We use Proposition 4 and (3.92) for $\mathbf{\Gamma}_N = \mathbf{V}_2/\sqrt{N}$, $\mathbf{A}_N = (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\,\mathbb{1}_{\mathcal{E}_N^c}$ and $\mathbf{B}_N = (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\,\mathbf{H}\,\mathbb{1}_{\mathcal{E}_N^c}$, and obtain that

$$\mathbb{E}_{\mathbf{V}_2}\left[\exp\left(iu\sqrt{N}\left(\eta_N\mathbb{1}_{\mathcal{E}_N^c} - \overline{\eta}_{N,1}\right)\right)\right] - \exp\left(iu\sqrt{N}\left(\kappa_N - \frac{\mathrm{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)}\right)\right)\exp(-\frac{u^2}{2}\zeta_N) \to 0 \quad (3.124)$$

a.s. ζ_N has almost surely the same behavior as ζ given by

$$\zeta = \frac{c_N}{(1-c_N)^3}\mathrm{Tr}(\mathbf{D}_N^2) + 2\frac{c_N}{(1-c_N)}\mathrm{Tr}(\mathbf{D}_N^2\mathbf{H}^*\mathbf{H}) \quad (3.125)$$

which implies that

$$\exp(-\frac{u^2}{2}\zeta_N) - \exp(-\frac{u^2}{2}\zeta) \to 0\ a.s. \quad (3.126)$$

Therefore, taking the mathematical expectation of (3.124) w.r.t $\mathbf{V}_1$ and using the dominated convergence theorem as well as (3.86), lead, after some calculations, to

$$\mathbb{E}\left[\exp\left(iu\sqrt{N}(\eta_N - \overline{\eta}_{N,1})\right)\right] - \exp\left[-\frac{u^2}{2}\left(\frac{Lc_N}{1-c_N} + \kappa_1\right)\right] \to 0 \quad (3.127)$$

for each u. As $\inf_N(\frac{Lc_N}{1-c_N} + \kappa_1) > 0$, (3.47) follows from (3.127) (see Proposition 6 in [63]).

3.9 Appendix: Proof of (3.87)

To establish (3.87), we follow the approach of [63] which is based on the joint use of the integration by parts formula and of the Poincaré-Nash inequality (see section III-B of [63]). However, the

approach of [63] allows to manage functionals of the resolvent $\mathbf{Q}_N(z)$ for $z \in \mathbb{C} - \mathbb{R}^+$. For this, the inequality $\|\mathbf{Q}_N(z)\| \leq \frac{1}{\text{dist}(z,\mathbb{R}^+)}$ plays a fundamental role. For $z = 0$, $\|\mathbf{Q}_N(0)\|$ coincides with $\frac{1}{\lambda_{1,N}}$ which is not upper-bounded by a deterministic positive constant for N greater than a non random integer. This issue was solved before using the regularization term $\mathbb{1}_{\mathcal{E}_N^c}$. However, the use of the integration by parts formula and the Poincaré-Nash inequality needs to consider smooth enough functions of $\mathbf{\Sigma}_N$. Motivated by [62], we consider the regularization term χ_N defined by

$$\chi_N = \det\left[\phi\left(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*\right)\right] \tag{3.128}$$

where ϕ is a smooth function such that

$$\phi(\lambda) = 1 \text{ if } \lambda \in \mathcal{I}_\epsilon = [\sigma^2(1-\sqrt{c})^2 - \epsilon, \sigma^2(1+\sqrt{c})^2 + \epsilon]$$

$$\phi(\lambda) = 0 \text{ if } \lambda \in [\sigma^2(1-\sqrt{c})^2 - 2\epsilon, \sigma^2(1+\sqrt{c})^2 + 2\epsilon]^c$$

$$\phi \in [0,1] \text{ elsewhere}$$

In the following, we need to use the following property: for each $\epsilon > 0$, it holds that

$$P\left(\mathcal{E}_N\right) = \mathcal{O}(\frac{1}{N^p}) \tag{3.129}$$

where $\mathcal{E}_N$ is defined by (3.71). Property (3.129) is not mentioned in Theorem 5.11 of [57] which addresses the non Gaussian case. However, (3.129) follows directly from Gaussian concentration arguments.

It is clear that

$$\left(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*\right)^{-1} \chi_N \leq \frac{\mathbf{I}}{\sigma^2((1-\sqrt{c})^2 - 2\epsilon)} \tag{3.130}$$

Lemma 3-9 of [62] also implies that, considered as a function of the entries of $\mathbf{\Sigma}_N$, χ_N is continuously differentiable. Moreover, it follows from Proposition 1 that almost surely, for N large enough, $\chi_N = 1$ and $\kappa_N = \kappa_N \chi_N$. Therefore, it holds that $\kappa_N \chi_N = \kappa_N + \mathcal{O}_P(\frac{1}{N^p})$, and that

$$\sqrt{N}\left(\kappa_N - \frac{\text{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)}\right) = \sqrt{N}\left(\kappa_N \chi_N - \frac{\text{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)}\right) + \mathcal{O}_P(\frac{1}{N^p}) \tag{3.131}$$

for each $p \in \mathbb{N}$. In order to establish (3.86), it is thus sufficient to prove that

$$\mathbb{E}\left[\exp\left(iu\sqrt{N}\left(\kappa_N \chi_N - \frac{\text{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)}\right)\right)\right] - \exp\left(-\frac{\theta_N u^2}{2}\right) \to 0 \tag{3.132}$$

for each u. To obtain (3.87), we remark that, as $\inf_N \theta_N > 0$, it follows from (3.132) that

$$\frac{\sqrt{N}}{\sqrt{\theta_N}} \left(\kappa_N \chi_N - \frac{\mathrm{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)} \right) \to_{\mathcal{D}} \mathcal{N}_{\mathbb{R}}(0,1)$$

(see Proposition 6 in [63]). (3.87) eventually appears as a consequence of (3.131).

The above regularization trick thus allows to replace the matrix $(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*)^{-1}$ by $(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*)^{-1} \chi_N$, which verifies (3.130). In order to establish (3.132), it is sufficient to prove that

$$\mathbb{E}(\kappa_N \chi_N) - \frac{\mathrm{Tr}(\mathbf{C}_N)}{\sigma^2(1-c_N)} = \mathrm{o}(\frac{1}{\sqrt{N}}) \tag{3.133}$$

and that

$$\mathbb{E}\left[\exp\left(iu\sqrt{N} \left(\kappa_N \chi_N - \mathbb{E}(\kappa_N \chi_N) \right) \right) \right] - \exp\left(-\frac{\theta_N u^2}{2} \right) \to 0 \tag{3.134}$$

for each u.

In the rest of this section, to simplify the notations, we omit to write the dependence on N of the various terms $\mathbf{\Sigma}_N$, $\mathbf{Q}_N(0)$, χ_N..., and denote them by $\mathbf{\Sigma}, \mathbf{Q}(0), \chi, \ldots$. However, we keep the notation c_N, in order to avoid confusion between c_N and c. Furthermore, the matrix $\mathbf{Q}(0)$ is denoted by $\mathbf{Q}$. If x is a random variable, x° represents the zero mean variable $x^\circ = x - \mathbb{E}(x)$. In the following, we denote by δ the random variable defined by

$$\delta = \sqrt{N}\, \kappa\, \chi$$

and by $\psi^\circ(u)$ the characteristic function of δ° defined by

$$\psi^\circ(u) = \mathbb{E}\left(\exp iu\delta^\circ \right)$$

We first establish the following Proposition.

Proposition 5. *It holds that*

$$\left(\psi^\circ(u)\right)' = -u\, \mathbb{E}\left(\mathrm{Tr}(\mathbf{C}^2 \mathbf{Q}^2 \chi) \right) \psi^\circ(u) + \mathcal{O}(\frac{1}{\sqrt{N}}) \tag{3.135}$$

where $'$ represents the derivative w.r.t. the variable u.

Proof. We consider the characteristic function $\psi(u)$ of δ, and evaluate

$$\psi'(u) = i\sqrt{N} \mathbb{E}\left(\mathrm{Tr}(\mathbf{Q}\mathbf{C}\chi) e^{iu\delta} \right)$$

We remark that $\mathbf{Q}\mathbf{\Sigma}\mathbf{\Sigma}^* = \mathbf{I}$ so that

$$\mathbb{E}\left(\mathbf{Q}\mathbf{\Sigma}\mathbf{\Sigma}^*\chi e^{iu\delta}\right) = \mathbb{E}(\chi e^{iu\delta})\mathbf{I}$$

We claim that

$$\mathbb{E}(\chi\, e^{iu\delta}) = \psi(u) + \mathcal{O}(\frac{1}{N^p}) \tag{3.136}$$

for each p. We remark that

$$\left|\mathbb{E}\left(e^{iu\delta}(1-\chi)\right)\right| \leq 1 - \mathbb{E}(\chi)$$

We recall that the event $\mathcal{E}$ is defined by (3.71) and that $P(\mathcal{E}) = \mathcal{O}(\frac{1}{N^p})$ for each p (see (3.129)). $\mathbb{I}_{\mathcal{E}^c} \leq \chi$ leads to $1 - \mathbb{E}(\chi) \leq P(\mathcal{E})$. This justifies (3.136). Therefore, it holds that

$$\mathbb{E}\left(\mathbf{Q}\mathbf{\Sigma}\mathbf{\Sigma}^*\chi\, e^{iu\delta}\right) = \left(\psi(u) + \mathcal{O}(\frac{1}{N^p})\right)\mathbf{I} \tag{3.137}$$

for each p. We now evaluate each entry of the left hand side of (3.137) using the integration by parts formula. For this, we denote by $(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N)$ the columns of $\mathbf{\Sigma}$. It holds that

$$(\mathbf{Q}\mathbf{\Sigma}\mathbf{\Sigma}^*)_{r,s} = \sum_{j=1}^{N} (\mathbf{Q}\boldsymbol{\xi}_j)_r \overline{\mathbf{\Sigma}}_{s,j}$$

and that

$$\mathbb{E}\left[(\mathbf{Q}\boldsymbol{\xi}_j)_r \overline{\mathbf{\Sigma}}_{s,j}\chi\, e^{iu\delta}\right] = \sum_{t=1}^{M} \mathbb{E}\left(\mathbf{Q}_{r,t}\mathbf{\Sigma}_{t,j}\overline{\mathbf{\Sigma}}_{s,j}\chi\, e^{iu\delta}\right)$$

The integration by parts formula leads to

$$\mathbb{E}\left(\mathbf{Q}_{r,t}\overline{\mathbf{\Sigma}}_{s,j}\chi\, e^{iu\delta}\mathbf{\Sigma}_{t,j}\right) = \frac{\sigma^2}{N}\mathbb{E}\left[\frac{\partial\left(\mathbf{Q}_{r,t}\overline{\mathbf{\Sigma}}_{s,j}\chi\, e^{iu\delta}\right)}{\partial\overline{\mathbf{\Sigma}}_{t,j}}\right]$$

After some algebra, we obtain that

$$\begin{aligned}\mathbb{E}\left(\mathbf{Q}_{r,t}\overline{\mathbf{\Sigma}}_{s,j}\chi\, e^{iu\delta}\mathbf{\Sigma}_{t,j}\right) = \frac{\sigma^2}{N}\mathbb{E}\left(\mathbf{Q}_{r,t}\chi\, e^{iu\delta}\right)\,\delta(t=s) \\ - \frac{\sigma^2}{N}\mathbb{E}\left((\mathbf{Q}\boldsymbol{\xi}_j)_r\,\mathbf{Q}_{t,t}\,\overline{\mathbf{\Sigma}}_{s,j}\,\chi\, e^{iu\delta}\right) \\ - \frac{i\sigma^2 u}{\sqrt{N}}\mathbb{E}\left(\mathbf{Q}_{r,t}\,(\mathbf{Q}\mathbf{C}\mathbf{Q}\boldsymbol{\xi}_j)_t\,\overline{\mathbf{\Sigma}}_{s,j}\,\chi\, e^{iu\delta}\right) \\ + \frac{\sigma^2}{N}\mathbb{E}\left(\mathbf{Q}_{r,t}\,\overline{\mathbf{\Sigma}}_{s,j}\,e^{iu\delta}\,\frac{\partial\chi}{\partial\overline{\mathbf{\Sigma}}_{t,j}}\right)\end{aligned} \tag{3.138}$$

We now need to study more precisely the properties of the derivative of χ w.r.t. $\overline{\mathbf{\Sigma}}_{t,j}$. For this, we give the following Lemma

Lemma 1. *We denote by $\mathcal{A}$ the event:*

$$\mathcal{A} = \{\textit{one of the } \hat{\lambda}_{k,N} \textit{ escapes from } \mathcal{I}_\epsilon\} \cap \{(\hat{\lambda}_{l,N})_{l=1,\ldots,M} \in \mathrm{supp}(\phi)\} \quad (3.139)$$

Then, it holds that

$$\frac{\partial \chi}{\partial \overline{\mathbf{\Sigma}}_{t,j}} = 0 \textit{ on } \mathcal{A}^c \quad (3.140)$$

and that

$$\mathbb{E}\left|\frac{\partial \chi}{\partial \overline{\mathbf{\Sigma}}_{i,j}}\right|^2 = \mathcal{O}(\frac{1}{N^p}) \quad (3.141)$$

for each p.

Proof. Lemma 1 follows directly from Lemma 3.9 of [62] and from the calculations in the proof of Proposition 3.3 of [62].

Lemma 1 implies that the last term of (3.138) is $\mathcal{O}(\frac{1}{N^p})$ for each p. To check this, we remark that

$$\mathbb{E}\left(\mathbf{Q}_{r,t}\,\overline{\mathbf{\Sigma}}_{s,j}\,e^{iu\delta}\,\frac{\partial \chi}{\partial \overline{\mathbf{\Sigma}}_{t,j}}\right) = \mathbb{E}\left(\mathbf{Q}_{r,t}\,\overline{\mathbf{\Sigma}}_{s,j}\,e^{iu\delta}\,\mathbb{1}_{\mathcal{A}}\frac{\partial \chi}{\partial \overline{\mathbf{\Sigma}}_{t,j}}\right)$$

The Schwartz inequality leads to

$$|\mathbb{E}\Big(\mathbf{Q}_{r,t}\,\overline{\mathbf{\Sigma}}_{s,j}\,e^{iu\delta}\,\mathbb{1}_{\mathcal{A}}\frac{\partial \chi}{\partial \overline{\mathbf{\Sigma}}_{t,j}}\Big)|^2 \leq \mathbb{E}\Big(|\mathbf{Q}_{r,t}\,\overline{\mathbf{\Sigma}}_{s,j}|^2\,\mathbb{1}_{\mathcal{A}}\Big)\mathbb{E}\left|\frac{\partial \chi}{\partial \overline{\mathbf{\Sigma}}_{i,j}}\right|^2$$

On event $\mathcal{A}$, all the eigenvalues of $\mathbf{\Sigma}\mathbf{\Sigma}^*$ belong to $[\sigma^2(1-\sqrt{c})^2 - 2\epsilon, \sigma^2(1+\sqrt{c})^2 + 2\epsilon]$. Therefore, $|\mathbf{Q}_{r,t}\,\mathbb{1}_{\mathcal{A}}$ is bounded and (3.141) implies that the last term of (3.138) is $\mathcal{O}(\frac{1}{N^p})$ for each p. Summing (3.138) over t, we obtain that

$$\begin{aligned}\mathbb{E}\left((\mathbf{Q}\boldsymbol{\xi}_j)_r\overline{\mathbf{\Sigma}}_{s,j}\,\chi\,e^{iu\delta}\right) = \frac{\sigma^2}{N}&\mathbb{E}\left(\mathbf{Q}_{r,s}\chi\,e^{iu\delta}\right) \\ &- \sigma^2 c_N\,\mathbb{E}\left(\hat{m}(0)\,(\mathbf{Q}\boldsymbol{\xi}_j)_r\,\overline{\mathbf{\Sigma}}_{s,j}\,\chi\,e^{iu\delta}\right) \\ &- \frac{i\sigma^2 u}{\sqrt{N}}\mathbb{E}\left((\mathbf{Q}^2\mathbf{C}\mathbf{Q}\boldsymbol{\xi}_j)_r\,\overline{\mathbf{\Sigma}}_{s,j}\,\chi\,e^{iu\delta}\right) + \mathcal{O}(\frac{1}{N^p}) \quad (3.142)\end{aligned}$$

where we recall that $\hat{m}(0) = \frac{1}{M}\mathrm{Tr}(\mathbf{Q})$ represents the Stieltjes transform of the empirical eigenvalue distribution $\hat{\mu}$ of $\mathbf{\Sigma}\mathbf{\Sigma}^*$ at $z = 0$. Using that $(1-\chi) \leq \mathbb{1}_{\mathcal{E}}$, it is easy to check that for each p, it

holds that

$$\mathbb{E}\Big((\mathbf{Q}\boldsymbol{\xi}_j)_r\,\hat{m}(0)\,\overline{\boldsymbol{\Sigma}}_{s,j}\,\chi\,e^{iu\delta}\Big) = \mathbb{E}\left((\mathbf{Q}\boldsymbol{\xi}_j)_r\,\hat{m}(0)\,\overline{\boldsymbol{\Sigma}}_{s,j}\,\chi^2\,e^{iu\delta}\right) + \mathcal{O}(\frac{1}{N^p})$$

We denote by β the term $\beta = \hat{m}(0)\chi$, and express β as $\beta = \alpha + \beta^\circ$. Replacing χ by χ^2 in the second term of the right hand side of (3.142) and plugging $\beta = \alpha + \beta^\circ$ into (3.142), we obtain that immediately that

$$\begin{aligned}\mathbb{E}\left((\mathbf{Q}\boldsymbol{\xi}_j)_r\overline{\boldsymbol{\Sigma}}_{s,j}\,\chi\,e^{iu\delta}\right) &= \frac{\sigma^2}{N(1+\sigma^2 c_N\alpha)}\mathbb{E}\left(\mathbf{Q}_{r,s}\chi\,e^{iu\delta}\right)\\ &- \frac{i\sigma^2 u}{\sqrt{N}(1+\sigma^2 c_N\alpha)}\mathbb{E}\left((\mathbf{Q}^2\mathbf{C}\mathbf{Q}\boldsymbol{\xi}_j)_r\,\overline{\boldsymbol{\Sigma}}_{s,j}\,\chi\,e^{iu\delta}\right)\\ &- \frac{\sigma^2 c_N}{1+\sigma^2 c_N\alpha}\mathbb{E}\left(\beta^\circ\,(\mathbf{Q}\boldsymbol{\xi}_j)_r\,\overline{\boldsymbol{\Sigma}}_{s,j}\,\chi\,e^{iu\delta}\right) + \mathcal{O}(\frac{1}{N^p}) \quad (3.143)\end{aligned}$$

Summing over j, we get that

$$\begin{aligned}\mathbb{E}\left((\mathbf{Q}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^*)_{r,s}\,\chi\,e^{iu\delta}\right) &= \frac{\sigma^2}{1+\sigma^2 c_N\alpha}\mathbb{E}\left(\mathbf{Q}_{r,s}\chi\,e^{iu\delta}\right)\\ &- \frac{i\sigma^2 u}{\sqrt{N}(1+\sigma^2 c_N\alpha)}\mathbb{E}\left((\mathbf{Q}^2\mathbf{C}\mathbf{Q}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^*)_{r,s}\,\chi\,e^{iu\delta}\right)\\ &- \frac{\sigma^2 c_N}{1+\sigma^2 c_N\alpha}\mathbb{E}\left(\beta^\circ(\mathbf{Q}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^*)_{r,s}\,\chi\,e^{iu\delta}\right) + \mathcal{O}(\frac{1}{N^p}) \quad (3.144)\end{aligned}$$

or, using that $\mathbf{Q}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^* = \mathbf{I}$,

$$\begin{aligned}\mathbb{E}\left(\chi\,e^{iu\delta}\right)\,\delta(r=s) &= \frac{\sigma^2}{1+\sigma^2 c_N\alpha}\mathbb{E}\left(\mathbf{Q}_{r,s}\chi\,e^{iu\delta}\right)\\ &- \frac{i\sigma^2 u}{\sqrt{N}(1+\sigma^2 c_N\alpha)}\mathbb{E}\left((\mathbf{Q}^2\mathbf{C})_{r,s}\,\chi\,e^{iu\delta}\right)\\ &- \frac{\sigma^2 c_N}{1+\sigma^2 c_N\alpha}\mathbb{E}\left(\beta^\circ\,\chi\,e^{iu\delta}\right)\,\delta(r=s) + \mathcal{O}(\frac{1}{N^p}) \quad (3.145)\end{aligned}$$

In order to evaluate α, we take $u = 0$ and sum over $r = s$ in (3.145), and obtain that

$$\alpha = \frac{1}{\sigma^2(1-c_N)} + \frac{1}{1-c_N}\mathbb{E}\,(\beta^\circ\,\chi) + \mathcal{O}(\frac{1}{N^p})$$

$\mathbb{E}\,(\beta^\circ\,\chi)$ coincides with $\mathbb{E}\,(\beta^\circ\,\chi^\circ)$. Using (3.141), the Poincaré-Nash inequality leads immediately to $\mathbb{E}\,((\chi^\circ)^2) = \mathcal{O}(\frac{1}{N^p})$, and to

$$\alpha = \frac{1}{\sigma^2(1-c_N)} + \mathcal{O}(\frac{1}{N^p}) \quad (3.146)$$

for each p. As a consequence, we also get that

$$\mathbb{E}(\mathbf{Q}_{r,s}\,\chi) = \frac{1}{\sigma^2(1-c_N)}\,\delta(r=s) + \mathcal{O}(\frac{1}{N^p}) \tag{3.147}$$

We now use (3.145) in order to evaluate $\mathbb{E}\left((\mathbf{Q}_{r,s}\chi)^\circ\,\chi\,e^{iu\delta}\right)$. For this, we first establish that the use of (3.130) and of the Poincaré-Nash inequality implies that

$$\mathrm{Var}(\beta) = \mathbb{E}\left((\beta^\circ)^2\right) = \mathcal{O}(\frac{1}{N^2}) \tag{3.148}$$

To check this, we use the Poincaré-Nash inequality:

$$\mathrm{Var}(\beta) \leq \frac{\sigma^2}{N}\mathbb{E}\left(\sum_{i,j}\left|\frac{\partial\beta}{\partial\mathbf{\Sigma}_{i,j}}\right|^2 + \left|\frac{\partial\beta}{\partial\overline{\mathbf{\Sigma}}_{i,j}}\right|^2\right)$$

We just evaluate the terms corresponding to the derivatives with respect to the terms $(\overline{\mathbf{\Sigma}}_{i,j})_{i=1,\ldots,M,j=1,\ldots,N}$. It is easily seen that

$$\frac{\partial\beta}{\partial\overline{\mathbf{\Sigma}}_{i,j}} = -\frac{1}{M}(\mathbf{e}_i^T\mathbf{Q}^2\boldsymbol{\xi}_j)\,\chi + \frac{1}{M}\mathrm{Tr}(\mathbf{Q})\frac{\partial\chi}{\partial\overline{\mathbf{\Sigma}}_{i,j}}$$

Therefore, it holds that

$$\left|\frac{\partial\beta}{\partial\overline{\mathbf{\Sigma}}_{i,j}}\right|^2 \leq 2\frac{1}{M^2}\boldsymbol{\xi}_j^*\mathbf{Q}^2\mathbf{e}_i\mathbf{e}_i^T\mathbf{Q}^2\boldsymbol{\xi}_j\,\chi^2 + 2\frac{1}{M}\mathrm{Tr}(\mathbf{Q})\left|\frac{\partial\chi}{\partial\overline{\mathbf{\Sigma}}_{i,j}}\right|^2$$

Using the identity $\mathbf{Q}\mathbf{\Sigma}\mathbf{\Sigma}^* = \mathbf{I}$ as well that $\frac{\partial\chi}{\partial\overline{\mathbf{\Sigma}}_{i,j}} = \mathbb{1}_{\mathcal{A}}\frac{\partial\chi}{\partial\overline{\mathbf{\Sigma}}_{i,j}}$ (see (3.140)), we obtain that

$$\frac{\sigma^2}{N}\sum_{i,j}\left(\mathbb{E}\left|\frac{\partial\beta}{\partial\overline{\mathbf{\Sigma}}_{i,j}}\right|^2\right) \leq 2\sigma^2\frac{1}{MN}\mathbb{E}\left(\frac{1}{M}\mathrm{Tr}(\mathbf{Q}^3)\,\chi\right) + 2\frac{\sigma^2}{N}\mathbb{E}\left(\frac{1}{M}\mathrm{Tr}(\mathbf{Q})\,\mathbb{1}_{\mathcal{A}}\sum_{i,j}\left|\frac{\partial\chi}{\partial\overline{\mathbf{\Sigma}}_{i,j}}\right|^2\right)$$

On the set $\mathcal{A}$, the eigenvalues of $\mathbf{\Sigma}\mathbf{\Sigma}^*$ are located into $[\sigma^2(1-\sqrt{c})^2 - 2\epsilon, \sigma^2(1+\sqrt{c})^2 + 2\epsilon]$. Therefore, we get that

$$\frac{1}{M}\mathrm{Tr}(\mathbf{Q})\,\mathbb{1}_{\mathcal{A}} \leq \frac{1}{\sigma^2(1-\sqrt{c})^2 - 2\epsilon}$$

Using (3.141), we obtain that

$$2\frac{\sigma^2}{N}\mathbb{E}\left(\frac{1}{M}\mathrm{Tr}(\mathbf{Q})\,\mathbb{1}_{\mathcal{A}}\sum_{i,j}\left|\frac{\partial\chi}{\partial\overline{\mathbf{\Sigma}}_{i,j}}\right|^2\right) = \mathcal{O}(\frac{1}{N^p})$$

for each p. Moreover, (3.130) implies that

$$\frac{1}{M}\mathrm{Tr}(\mathbf{Q}^3)\,\chi \leq \frac{1}{\left(\sigma^2(1-\sqrt{c})^2 - 2\epsilon\right)^3}$$

and that

$$2\sigma^2\frac{1}{MN}\mathbb{E}\left(\frac{1}{M}\mathrm{Tr}(\mathbf{Q}^3)\,\chi\right)=\mathcal{O}(\frac{1}{N^2})$$

This establishes (3.148).

Therefore, the Schwartz inequality leads to $\mathbb{E}\left(\beta^\circ\,\chi\,e^{iu\delta}\right)=\mathcal{O}(\frac{1}{N})$. Writing $\mathbb{E}\left(\mathbf{Q}_{r,s}\chi\,e^{iu\delta}\right)$ as

$$\mathbb{E}\left(\mathbf{Q}_{r,s}\chi\,e^{iu\delta}\right)=\mathbb{E}\left(\mathbf{Q}_{r,s}\chi^2\,e^{iu\delta}\right)+\mathcal{O}(\frac{1}{N^p})=$$
$$\mathbb{E}(\mathbf{Q}_{r,s}\chi)\mathbb{E}(\chi\,e^{iu\delta})+\mathbb{E}\left((\mathbf{Q}_{r,s}\chi)^\circ\chi\,e^{iu\delta}\right)+\mathcal{O}(\frac{1}{N^p})=$$
$$\mathbb{E}(\mathbf{Q}_{r,s}\chi)\mathbb{E}(\chi\,e^{iu\delta})+\mathbb{E}\left((\mathbf{Q}_{r,s}\chi)^\circ\,e^{iu\delta}\right)+\mathcal{O}(\frac{1}{N^p})$$

(3.146), (3.147) and (3.145) lead to

$$\mathbb{E}\left((\mathbf{Q}_{r,s}\chi)^\circ\,e^{iu\delta}\right)=\frac{iu}{\sqrt{N}}\,\mathbb{E}\left((\mathbf{Q}^2\mathbf{C})_{r,s}\,\chi\,e^{iu\delta}\right)+\mathcal{O}(\frac{1}{N}) \tag{3.149}$$

or equivalently to

$$\mathbb{E}\left(\delta^\circ\,e^{iu\delta}\right)=iu\,\mathbb{E}\left(\mathrm{Tr}(\mathbf{Q}^2\mathbf{C}^2)\,\chi\,e^{iu\delta}\right)+\mathcal{O}(\frac{1}{\sqrt{N}})$$

Using the Nash-Poincaré inequality, it can be checked that

$$\mathrm{Var}\left(\mathrm{Tr}(\mathbf{Q}^2\mathbf{C}^2)\,\chi\right)=\mathcal{O}(\frac{1}{N})$$

Therefore, the Schwartz inequality leads to

$$\mathbb{E}\left(\mathrm{Tr}(\mathbf{Q}^2\mathbf{C}^2)\,\chi\,e^{iu\delta}\right)=\mathbb{E}\left(\mathrm{Tr}(\mathbf{Q}^2\mathbf{C}^2)\,\chi\right)\mathbb{E}(e^{iu\delta})+\mathcal{O}(\frac{1}{\sqrt{N}})$$

and we get that

$$\mathbb{E}\left(\delta^\circ\,e^{iu\delta}\right)=iu\,\mathbb{E}\left(\mathrm{Tr}(\mathbf{Q}^2\mathbf{C}^2)\,\chi\right)\mathbb{E}(e^{iu\delta})+\mathcal{O}(\frac{1}{\sqrt{N}}) \tag{3.150}$$

Plugging $\delta=\delta^\circ+\mathbb{E}(\delta)$ into (3.150) eventually leads to

$$\mathbb{E}\left(\delta^\circ\,e^{iu\delta^\circ}\right)=iu\,\mathbb{E}\left(\mathrm{Tr}(\mathbf{Q}^2\mathbf{C}^2)\,\chi\right)\mathbb{E}(e^{iu\delta^\circ})+\mathcal{O}(\frac{1}{\sqrt{N}}) \tag{3.151}$$

which is equivalent to (3.135). This, in turn, establishes Proposition 5.

We now complete the proof of (3.134). We integrate (3.135), and obtain that

$$\psi^\circ(u)=\exp\left[-\frac{u^2}{2}\,\mathbb{E}\left(\mathrm{Tr}(\mathbf{Q}^2\mathbf{C}^2\,\chi)\right)\right]+\mathcal{O}(\frac{1}{\sqrt{N}})$$

(see section V-C of [63] for more details). (3.82) implies that

$$\mathrm{Tr}(\mathbf{Q}^2\mathbf{C}^2) - \frac{\mathrm{Tr}(\mathbf{C}^2)}{\sigma^4(1-c_N)^3} \to 0\, a.s.$$

As $\mathrm{Tr}(\mathbf{Q}^2\mathbf{C}^2)\,\chi - \mathrm{Tr}(\mathbf{Q}^2\mathbf{C}^2)$ also converges to 0 almost surely, we obtain that

$$\mathrm{Tr}(\mathbf{Q}^2\mathbf{C}^2)\chi - \frac{\mathrm{Tr}(\mathbf{C}^2)}{\sigma^4(1-c_N)^3} \to 0\, a.s.$$

As matrix $\mathbf{Q}^2\chi$ is bounded and $\sup_N \mathrm{Tr}(\mathbf{C}^2) < +\infty$, it is possible to use the Lebesgue dominated convergence theorem and to conclude that

$$\mathbb{E}\left(\mathrm{Tr}(\mathbf{Q}^2\mathbf{C}^2)\chi\right) - \frac{\mathrm{Tr}(\mathbf{C}^2)}{\sigma^4(1-c_N)^3} \to 0$$

This proves (3.134).

It remains to establish (3.133). For this, we use (3.147), and obtain that

$$\mathbb{E}\left(\mathrm{Tr}(\mathbf{QC})\,\chi\right) - \frac{\mathrm{Tr}(\mathbf{C})}{\sigma^2(1-c_N)} = \mathcal{O}(\frac{1}{N^p})$$

for each p. This, of course, implies (3.133).

Chapter 4

Conclusion

The focus of this PhD thesis has been MIMO synchronization for frequency selective channels in the presence of interference. During the course of these three years, we have made numerous discoveries of great practical and theoretical interest concerning the subject.

The starting point of the study has been two GLRT statistics for different noise characteristics. The two tests in question are $\eta_{\text{GLRT,we}}$, optimized for temporally and spatially independent noise, and η_{GLRT}, optimized for temporally independent but possibly spatially dependent noise. As opposed to popular belief, we first make the point that the GLRTs are not always optimal from a performance point of view, and there are some situations where they perform worse than the correlation-based or MMSE test statistics. Furthermore, we show by simulation that η_{GLRT} is preferrable if the noise characteristics are unknown, since the gain of $\eta_{\text{GLRT,we}}$ over η_{GLRT} even in spatially independent noise is small.

We then go on to remark that η_{GLRT} is often discarded in the literature, due to its high implementation complexity. A part of this work has thus been devoted to showing that there exist non-GLRT statistics that are less complex to implement than the η_{GLRT}, while having similar performance. In the literature, the MMSE statistics is often quoted as the second-best alternative to the GLR test. The MMSE performs worse than the GLRT, but is still robust to interference, and has lower implementation complexity. We show that the complexities of the new non-GLRT statistics are similar to the complexity of the MMSE statistics, and show by numerical simulations that the MMSE is inferior to the new statistics.

In addition to defining new synchronization statistics, we have investigated the possibility of lowering the complexity of the existing ones. We have formulated a new, suboptimal, way of performing synchronization where only a part of the statistics is updated at each time index, by using the same empirical covariance matrix for several time indices. Even though this is a

suboptimal solution and entails a slight performance loss, its use can be envisioned when system resources are limited. The approaches for complexity reduction described above are based on viewing the system parameters as fixed, and optimizing the synchronization for those parameters.

Another equally valid approach is viewing the system as flexible, with free choice of whether or not to use all available resources, such as transmit antennas. Given a MIMO system, we can choose whether it is necessary to use all transmit antennas, or if we can save resources, such as transmit power, by only transmitting on a subset of the available antennas. To complete the analysis, we have thus included a section which concerns the optimization of the number of transmit antennas in MIMO synchronization. We have shown that depending on the channel type, SNR and number of receive antennas M, increasing the number of transmit antennas K for MIMO synchronization is only useful up to a certain $K = K_{opt}$, above which the performance decreases.

In addition to the more applied contributions described above, this thesis also presents an important theoretical contribution through the study of the limiting distributions of η_{GLRT}. We have investigated the asymptotic performance of a η_{GLRT} in two asymptotic regimes that require that the analysis is done using large random matrix methods. The asymptotic analysis shows that in both asymptotic regions, the GLRT statistics follows a Gaussian distribution both under H_0 and H_1. We have shown that this asymptotic performance is much more accurate than the so-called classical asymptotic analysis, since the classical asymptotic analysis assumes that M, the number of receive antennas, and L, the number of channel paths, are of negligible size when compared to N, the length of the training sequence.

These results have then been used to trace theoretical ROC curves, and shown to be very close to the emprical ROC curves, especially for large P_{FA}. These final results can be seen as the link between the first and second parts of the PhD thesis. The theoretical results can be used to perform the parameter optimization theoretically, without the need of simulations. For example, the question of optimization of the number of transmit antennas could be evaluated by using these asymptotic distributions. Further, since the tools for calculating the limit distribution of η_{GLRT} are established, they can easily be adapted to calculate the asymptotics of other similar statistics.

Résumé long

4.1 Synchronisation MIMO: ami ou ennemi?

4.1.1 Introduction: Évolution et avantages de MIMO

Au cours de ces dernières décennies, de nombreuses architectures pour l'émission et la réception de systèmes sans fil ont été développées. Ces architectures, qui visent à tirer profit de plusieurs antennes (d'émission et/ou de réception), ont eu plusieurs défis à relever, défis que n'ont pas les systèmes filaires classiques. Parmi ces défis, les trajets multiples et l'atténuation du canal de transmission ont rendu difficile la montée en débit ainsi que la diminution du taux d'erreur. L'arrivée des systèmes Multiple-Input Multiple-Output (MIMO) a alors été vue comme une révolution dans le monde des télécommunications sans fil car elle promettait une solution à la congestion des communications sans fil.

Le gain apporté par les antennes multiples peut être divisé en trois grandes catégories:

1. *Le gain de beamforming (formation de faisceaux).* , Il est obtenu en orientant la puissance du signal dans une direction donnée (et en l'annulant dans les autres directions). Le beamforming, une technique multi-antennes classique, peut être utilisé soit à la transmission soit à la réception.

2. *Le gain de diversité.* Il est obtenu en envoyant l'information à travers plusieurs canaux de transmission aux caractéristiques différentes. Par exemple, afin de tirer profit des différentes caractéristiques d'évanouissement (fading) des canaux vues par différentes antennes, il est possible d'envoyer la même information à toutes les antennes d'émission. Le gain de diversité augmente ainsi la robustesse du système en éliminant les évanouissements. Dans un certain sens, il permet de convertir un canal avec fading en un canal sans fading.

3. *Le gain de multiplexage.* Ce dernier, spécifique aux systèmes MIMO, est obtenu par une technique de multiplexage spatial. Les données transmises sont donc divisées en plusieurs

flux indépendants qui devront être décodés au niveau du récepteur. La séparabilité à la réception fait usage de riche multi-trajets, ce qui rend le canal spatialement sélectif. Au lieu de voir les trajets multiples comme un problème, le multiplexage spatial les exploite. L'objectif principal du multiplexage spatial est de maximiser la vitesse de transmission, un objectif nettement différent de celui de la diversité de transmission / réception dont l'objectif principal est d'augmenter la fiabilité.

Lorsque les communications sans fil étaient principalement utilisées pour transmettre de la voix, l'intérêt des systèmes MIMO restait faible. Il n'y avait pas besoin de débits élevés, ou plutôt, en raison des caractéristiques difficiles des canaux de transmission sans fil, les applications haut débit telles que la vidéo étaient généralement transmises au moyen de technologies filaires. L'intérêt du MIMO était également faible en raison des problèmes pratiques rencontrés, entre autres le coût et les défis matériels dans la mise en œuvre, notamment dans les équipements côté utilisateur. Néanmoins, avec l'utilisation croissante des téléphones et autres appareils mobiles, il a été nécessaire de résoudre le problème de l'augmentation du débit, même pour les applications sans fil. D'autre part, maintenant que les téléphones sont devenus plus sophistiqués, les contraintes sur la taille des systèmes MIMO sont moins sévères, et il est devenu possible d'avoir plusieurs antennes sur les deux côtés du lien. Le MIMO n'est donc plus seulement qu'une considération théorique; il a en effet été mis en œuvre avec succès dans plusieurs standards bien établis. Quelques exemples incluent le CDMA qui utilise des codes espace-temps Alamouti et le beamforming MIMO dans le standard 3GPP LTE.

Quelles sont les alternatives au MIMO, et pourquoi ne sont-elles pas pertinentes pour les communications sans fil à haut débit ? Une approche classique est le système Single-Input Single-Output (SISO). Mais les systèmes SISO ne peuvent pas rejeter les interférences, ni profiter de la diversité spatiale. Les systèmes multi-antennes traditionnels, présents depuis des décennies, ont plusieurs antennes, soit en émission (SIMO) soit en réception (MISO), et généralement au niveau de la station de base. Ces systèmes utilisent principalement la diversité spatiale, ou bien le beamforming (en émission ou en réception). Ils peuvent ainsi effectuer la réduction d'interférence en tirant profit de plusieurs antennes en réception grâce au beamforming. En théorie, des communications à très haut débit pourraient être mises en œuvre avec un système SIMO, mais cela nécessiterait une bande passante très élevée, ce qui n'est pas possible dans la pratique puisque la bande passante est une ressource coûteuse. Bref, malgré tous ses avantages, un système SIMO n'est pas envisageable pour mettre en place une liaison Gigabit Internet.

La base d'un système MIMO est d'avoir de multiples antennes réparties dans l'espace sur les deux côtés du lien. Cela transforme la représentation du système : on passe d'un système vectoriel à un système matriciel. Cela ajoute également des degrés de liberté supplémentaires

puisque la dimension spatiale peut être exploitée pour effectuer du multiplexage spatial. Enfin, un système MIMO dispose aussi d'un gain de diversité en émission et réception. Ainsi, des résultats théoriques montrent que le débit maximum d'un système MIMO $(K \times M)$ croît linéairement avec $\min(K, M)$. En résumé, les avantages des antennes intelligentes sont conservés et de nouveaux avantages sont ajoutés.

Néanmoins, comme dans tout système, ce qu'on gagne d'un côté, on le perd de l'autre. Certains des défis présents dans la synchronisation MIMO sont résumés dans le paragraphe suivant.

4.1.2 Défis de la synchronisation MIMO

Un défi important dans la synchronisation MIMO est de trouver le bon compromis performance/complexité. En effet, lorsque le nombre d'antennes augmente, la complexité des algorithmes augmente en même temps. Il convient donc de se demander si l'ajout d'antennes supplémentaires dans un système est bien adapté à l'application visée. La complexité est donc pour partie un problème de point de vue, mais aussi de coût du matériel et de la compatibilité. Autrement dit, augmenter le nombre d'antennes dans un système est une décision importante, où chaque avantage et chaque inconvénient doivent être soigneusement pondérés.

Une autre considération importante dans les systèmes MIMO est les hypothèses faites sur les interférences présentes dans le système. Les systèmes point-à-point peuvent être une hypothèse raisonnable dans certaines applications, par exemple la norme IEEE 802.11n, car ils sont conçus pour que des liaisons à courte portée ne souffrent pas d'interférences. Mais en général, les systèmes MIMO cellulaires sont limités par des interférences, soit par des interférences provenant de la cellule elle-même, soit par des interférences externes. Comme le nombre d'interférences augmente, elles ne peuvent pas être supprimées par un traitement du signal spatial, et sont donc traitées comme du bruit. Par conséquent, contrairement à ce que l'on pourrait croire, le fait d'ajouter des antennes en émission peut effectivement diminuer le débit à faible SNR.

Une caractéristique supplémentaire du système est les caractéristiques du canal. Une hypothèse commune est de supposer que le canal est un canal de Rayleigh. Pour traiter complètement le problème, il est intéressant d'étudier les performances des systèmes MIMO même pour les canaux déterministes, et comprendre quels paramètres du système choisir selon le modèle de canal approprié.

Considérant les défis ci-dessus, la section suivante résume l'état de l'art de la synchronisation MIMO. Cela a pour but de valider la nécessité de cette étude.

4.1.3 Etat de l'art en synchronisation MIMO

La synchronisation en temps et en fréquence des systèmes MIMO a été fortement étudiée au cours de ces quinze dernières années, principalement dans le cadre de liens DS-CDMA et OFDM.

OFDM. La synchronisation grossière et fine, ainsi que la compensation du décalage en fréquence, ont été analysées, puis de nombreuses techniques ont été proposées, soit pour la synchronisation temps-fréquence [1–4] soit pour la synchronisation temporelle [5–9]. Néanmoins, la plupart de ces techniques supposent l'absence d'interférences. Les rares articles de la littérature qui traitent de la synchronisation MIMO en présence d'interférences sont [2, 7–9]. Cependant, [2] et [7] envisagent le problème de la synchronisation MIMO en présence d'interférences multi-utilisateurs (MUI). [8] semble être le seul article traitant de synchronisation MIMO en présence d'interférences de tout types, par exemple des brouilleurs hostiles. Dans [8], plusieurs statistiques sont proposées pour la synchronisation temporelle, à la fois pour du fading plat et du fading sélectif en fréquence.

Malgré les nombreux algorithmes existants pour la synchronisation temporelle des systèmes MIMO, de nombreuses questions importantes restent non résolues, notamment au sujet de leur optimalité, leurs performances et leur complexité.

Tout d'abord, aucun des récepteurs mis au point pour la synchronisation MIMO dans un environnement sans interférence [1, 3–6, 10–27] n'a été développé en optant une approche GLRT dans le cas général d'un lien mono-porteuse non DS-CDMA avec des séquences d'apprentissage potentiellement non orthogonales. De plus, il est bien connu [28] que, contrairement au test de rapport de vraisemblance (LRT), les statistiques GLRT sont sous-optimales du point de vue de la détection. En gardant cela à l'esprit, on peut se demander si une statistique non-GLRT peut avoir des performances égales ou meilleures que les statistiques GLRT existantes.

Alors que la statistique GLRT dans un environnement en présence d'interférences a déjà été présentée [8], des questions concernant son utilité pratique se posent encore. Dans [8] , deux statistiques pour les canaux à fading plat, robustes aux interférences, sont dérivées, respectivement d'une approche MMSE et d'une approche GLRT. La statistique GLRT, qui sera appelée η_{GLRT} par la suite, assume un bruit total inconnu, gaussien, coloré spatialement et blanc temporellement. Notons d'abord que cette statistique peut être très coûteuse à mettre en œuvre, en particulier pour un nombre important d'antennes. En effet, pour un système MIMO $(K \times M)$, il faut pour chaque indice de l'échantillon, inverser une matrice $(M \times M)$ et calculer un déterminant $(M \times M)$ ou $(K \times K)$. Une alternative à η_{GLRT} est la statistique MMSE proposée dans [8]. Cependant, cette approche se montre sensible à la corrélation des séquences d'apprentissage, ce qui peut

limiter son utilisation pratique.

Avant de décrire notre proposition pour résoudre les problèmes ci-dessus, nous allons introduire le modèle du système considéré, et brièvement rappeler le concept de synchronisation en tant que problème de test d'hypothèse. Nous allons également introduire le concept du test du rapport de vraisemblance généralisé (GLRT).

4.2 Modèle et problème

4.2.1 Hypothèses

Nous considérons un lien de radiocommunications MIMO ($K \times M$) avec K antennes à bande étroite à l'émission et M à la réception. Nous notons $\mathbf{S}(k)$ la séquence de synchronisation complexe ($K \times 1$). Celle-ci, transmise sur les antennes d'émission à l'instant k, est connue du récepteur. En supposant la synchronisation parfaite, le vecteur, $\mathbf{y}(k)$ des enveloppes complexes des signaux à la sortie des M antennes de réception à l'instant k peut être écrit comme suit :

$$\mathbf{y}(k) = \sum_{l=0}^{L-1} \mathbf{s}(k)\mathbf{H}_l(k-l) + \mathbf{v}(k) \tag{4.1}$$

Ici, $\mathbf{H}_l$ est la matrice ($K \times M$) du canal le long du chemin l, et $\mathbf{v}(k)$ le vecteur du bruit total échantillonné, qui contient la contribution potentielle des interférences et du bruit de fond. Les échantillons $\mathbf{v}(k)$ sont supposés être de moyenne nulle, blancs temporellement, et gaussiens avec une matrice de covariance $\mathbf{R} = \mathbb{E}[\mathbf{v}(k)\mathbf{v}(k)^*]$.

Notons $\mathbf{Y}$ et $\mathbf{V}$ les matrices ($M \times N$) d'observation et de bruit total avec, $\mathbf{Y} = [\mathbf{y}(1), \ldots, \mathbf{y}(N)]$ et $\mathbf{V} = [\mathbf{v}(1), \ldots, \mathbf{v}(N)]$, $\mathbf{S}$ la matrice ($KL \times N$) des séquences de synchronisation

$$\mathbf{S} = \begin{pmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \ldots & \mathbf{s}_N \\ \mathbf{0} & \mathbf{s}_1 & \ldots & \mathbf{s}_{N-1} \\ & \ddots & & \\ & & & \\ \mathbf{0} & & \ddots & \mathbf{s}_{N-L+1} \end{pmatrix} \tag{4.2}$$

où $[\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_N]$ est la matrice de synchronisation sur le premier chemin, et $\mathbf{H} = [\mathbf{H}_0, \mathbf{H}_1, \ldots, \mathbf{H}_{L-1}]$ la matrice de canal sélectif en fréquence. Avec ces définitions, nous obtenons le modèle matriciel suivant :

$$\mathbf{Y} = \mathbf{HS} + \mathbf{V} \tag{4.3}$$

Notons que pour des canaux sélectifs en fréquence, une information sur le nombre maximal de chemins L est a priori requise. Ce qui est intéressant ici c'est que ce modèle peut être appliqué à

toutes les applications pouvant s'écrire sous la forme matricielle ci-dessus, il ne se limite donc pas à la synchronisation.

4.2.2 Synchronisation comme test d'hypothèse

Le problème de la synchronisation temporelle d'une liaison MIMO peut être considéré comme un problème de détection à trois hypothèses. La première hypothèse (H_1) est que la matrice du signal $\mathbf{S}$ est parfaitement alignée dans le temps avec la matrice d'observation $\mathbf{Y}$. Elle correspond au modèle (1.3), où

$$H_1 : \mathbf{Y} = \mathbf{HS} + \mathbf{V} \tag{4.4}$$

La seconde hypothèse (H_0) est qu'il n'y a pas de signal dans la matrice d'observation $\mathbf{Y}$. Cela correspond au modèle (4.5) donné par :

$$H_0 : \mathbf{Y} = \mathbf{V} \tag{4.5}$$

La troisième hypothèse est une hypothèse intermédiaire. Elle suppose qu'une partie du signal reçu est uniquement composée du bruit, tandis que l'autre partie contient une fraction du signal utile. Définissons la matrice $(KL \times N)$ de la séquence de synchronisation partielle

$$\mathbf{S}_p = [\mathbf{0}_{(KL\times A)}, \mathbf{S}_1], \tag{4.6}$$

où $\mathbf{S}_1$ est une matrice $(KL \times (N - A))$, $A < N$, qui contient une sous-partie de la matrice de la séquence de synchronisation définie en (4.2). L'hypothèse intermédiaire est alors donnée par :

$$H_p : \mathbf{Y} = \mathbf{HS}_p + \mathbf{V} \tag{4.7}$$

L'hypothèse H_p n'est pas négligeable, en particulier dans un canal sélectif en fréquence, où plusieurs bonnes détections peuvent provenir de la présence de seulement une partie du signal.

Néanmoins, pour simplifier le problème, nous considérons une hypothèse binaire avec seulement H_0 et H_1. Le problème de la synchronisation comme détection consiste alors à élaborer un test statistique η en fonction des observations $\mathbf{Y}$, et à le comparer à un seuil s. Si le seuil est dépassé, la détection est alors validée. Plus précisément, le seuil est utilisé pour choisir soit H_0 soit H_1, avec :

$$\eta \underset{H_0}{\overset{H_1}{\gtrless}} s \tag{4.8}$$

Si le seuil est dépassé à plusieurs reprises, une règle de décision doit être appliquée afin de choisir le bon moment de détection.

4.2.3 Test du rapport de vraisemblance généralisé (GLRT)

D'après la théorie de détection de Neyman-Pearson, la statistique optimale pour la détection d'une matrice $\mathbf{S}$ de la matrice $\mathbf{Y}$ est le test du rapport de vraisemblance (Likelihood Ratio Test, LRT), qui consiste à comparer la fonction $LRT \triangleq p_{H_1}(\mathbf{Y}|\mathbf{H},\mathbf{S},\mathbf{R})/p_{H_0}(\mathbf{Y}|\mathbf{R})$ à un seuil, où $p_{h_i}(\mathbf{Y}|\ldots)$ $(i=0,1)$ est la densité de probabilité conditionnelle de $\mathbf{Y}$ sachant H_i. Pour notre modèle de bruit, l'expression de la LRT prend la forme

$$\mathrm{LRT} = \frac{\prod_{k=1}^{N} p_{H_1}(\mathbf{Y}(k) \mid \mathbf{S}(k),\mathbf{H},\mathbf{R})}{\prod_{k=1}^{N} p_{H_0}(\mathbf{Y}(k) \mid \mathbf{R})}. \tag{4.9}$$

Pour notre modèle gaussien, les fonctions de densité de probabilité prennent la forme

$$p_{H_0}(\mathbf{Y}(k) \mid \mathbf{R}) = \frac{1}{\pi^M \det(\mathbf{R})} e^{-\mathbf{Y}(k)^*\mathbf{R}^{-1}\mathbf{Y}(k)} \tag{4.10}$$

et

$$p_{H_1}(\mathbf{Y}(k) \mid \mathbf{S}(k),\mathbf{H},\mathbf{R}) = \frac{1}{\pi^M \det(\mathbf{R})} e^{-(\mathbf{Y}(k)-\mathbf{HS}(k))^*\mathbf{R}^{-1}(\mathbf{Y}(k)-\mathbf{HS}(k))}, \tag{4.11}$$

Comme dans la pratique les matrices $\mathbf{R}$, $\mathbf{H}$ ou les deux sont inconnues, elles doivent être remplacées dans (4.9) par leur estimation Maximum Likelihood (ML) sous chacune des deux hypothèses H_1 et H_0, ce qui donne lieu au test du rapport de vraisemblance generalisé (GLRT). Dans le cas où $\mathbf{R}$ et $\mathbf{H}$ sont tous deux inconnus, la statistique GLRT devient

$$\eta_{\mathrm{GLRT}} = \frac{\arg\max_{\mathbf{R},\mathbf{H}} \prod_{k=1}^{N} p_{H_1}(\mathbf{Y}(k) \mid \mathbf{S}(k),\mathbf{H},\mathbf{R})}{\arg\max_{\mathbf{R}} \prod_{k=1}^{N} p_{H_0}(\mathbf{X}(k) \mid \mathbf{R})}. \tag{4.12}$$

4.2.4 Performance d'un test d'hypothèse binaire

La performance d'un critère de synchronisation η est caractérisée par sa probabilité de non-détection sous H_1 ($\mathrm{P_{ND}}$) pour une probabilité de fausse alarme ($\mathrm{P_{FA}}$), c'est-à-dire la probabilité de dépasser le seuil s sous H_0, fixe. Plus formellement,

$$\begin{aligned} \mathrm{P_{FA}} &= P(\eta > t \mid \mathrm{H}_0) \\ \mathrm{P_{ND}} &= P(\eta < t \mid \mathrm{H}_1) \end{aligned} \tag{4.13}$$

La figure 4.1 illustre le concept de $\mathrm{P_{FA}}$ et $\mathrm{P_{ND}}$ pour le cas où η sous H_0 et H_1 suit une distribution gaussienne, pour un $\mathrm{P_{FA}} = 10^{-3}$. Dans notre travail, le seuil s pour une $\mathrm{P_{FA}}$ donnée sera fixé de manière empirique en générant un grand nombre d'échantillons sous H_0. Ce seuil sera ensuite utilisé dans les simulations pour déterminer la $\mathrm{P_{ND}}$.

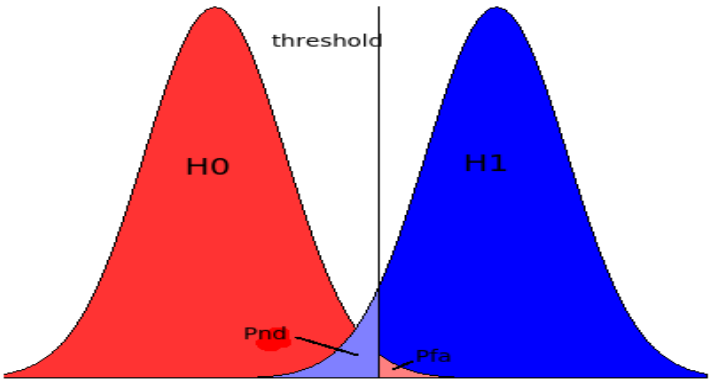


Figure 4.1: P_{ND} et P_{FA} pour un test d'hypothèses binaire

4.3 Objectif

L'objectif de cette thèse est d'obtenir les réponses à certains problèmes ouverts sur la synchronisation MIMO:

1. Choix des algorithmes de synchronisation pour les systèmes MIMO, et analyse du compromis complexité/performance.
2. Les avantages possibles du MIMO face au SIMO, et enquête sur la nécessité de passer de systèmes SIMO à des systèmes MIMO dans un contexte de synchronisation.
3. Caractérisation du comportement asymptotique de la synchronisation MIMO.
4. Prédiction des performances de la synchronisation MIMO.

4.4 Optimisation des paramètres et réduction de la complexité de la synchronisation temporelle

4.4.1 Introduction au problème

La première partie de la thèse, décrit en détail dans le chapitre 2, cherche à optimiser les paramètres de la synchronisation MIMO et à en réduire sa complexité. Comme mentionné au paragraphe 4.1.3, le récepteur le plus puissant aujourd'hui est celui basé sur le test du rapport de vraisemblance généralisé (GLRT). On suppose alors que le bruit est inconnu, gaussien, temporellement blanc et spatialement coloré. Néanmoins, la complexité de cette statistique est supérieure à ses homologues non GLRT, qui, malheureusement, ne fonctionnent pas aussi bien dans la plupart des cas. Comme la complexité est une question majeure lors des implémentations pratiques, et qu'elle peut être démesurée pour un grand nombre d'antennes, le but de cette étude est de proposer plusieurs façons de diminuer la complexité du test GLR tout en conservant sa performance.

4.4.2 Les statistiques de synchronisation étudiées dans la thèse

Une première façon simple d'optimiser les performances de la synchronisation MIMO, et dans le cas échéant d'en diminuer la complexité, consiste à s'appuyer sur les connaissances que nous avons du système. Si nous savons que le bruit est spatialement et temporellement indépendant, nous pouvons alors utiliser les GLRT suivants (Voir les appendices 2.6 et 2.7 pour la preuve), optimisés pour une puissance de bruit égale ou inégale à chaque antenne du récepteur :

- $\mathbf{R}$ et $\mathbf{H}$ sont toutes deux inconnues, mais l'on sait tout de même que $\mathbf{R}$ est sous la forme $\mathbf{R} = \sigma^2\mathbf{I}$ (bruit spatialement blanc avec une puissance du bruit identique sur toutes les antennes de réception) :

$$\eta_{\text{GLRT, nous}} = \frac{\underset{\sigma^2,\mathbf{H}}{\arg\max} \prod_{k=1}^{N} p_{H_1}(\mathbf{Y}(k) \mid \mathbf{S}(k), \mathbf{H}, \sigma^2\mathbf{I})}{\underset{\sigma^2}{\arg\max} \prod_{k=1}^{N} p_{H_0}(\mathbf{Y}(k) \mid \sigma^2\mathbf{I})}. \tag{4.14}$$

- $\mathbf{R}$ et $\mathbf{H}$ sont inconnues, mais l'on sait tout de même que $\mathbf{R}$ est sous la forme $\widetilde{\mathbf{R}} = \text{Diag}[\sigma_1^2, \sigma_2^2, \ldots, \Sigma_M^2]$ (bruit spatialement blanc avec une puissance de bruit différente sur

chaque antenne de réception) :

$$\eta_{\text{GLRT, wu}} = \frac{\underset{\sigma_1^2,\ldots,\sigma_M^2,\mathbf{H}}{\arg\max} \prod_{k=1}^{N} p_{H_1}(\mathbf{Y}(k) \mid \mathbf{S}(k), \mathbf{H}, \widetilde{\mathbf{R}})}{\underset{\sigma_1^2,\ldots,\sigma_M^2}{\arg\max} \prod_{k=1}^{N} p_{H_0}(\mathbf{Y}(k) \mid \widetilde{\mathbf{R}})} \tag{4.15}$$

Néanmoins, si $\mathbf{R}$ et la matrice de canal $\mathbf{H}$ sont inconnues, la statistique de synchronisation GLRT est calculée en utilisant (4.18), qui est le point de départ des chapitres 2 et 3. Cette statistique GLRT, déjà présente dans la littérature [8], est donnée par :

$$\eta'_{\text{GLRT}} = \det[\mathbf{I}_N - \mathbf{S}^*(fatSFats^*)^{-1}\mathbf{S}\mathbf{Y}^*(\mathbf{Y}\mathbf{Y}^*)^{-1}\mathbf{Y}]^{-N}. \tag{4.16}$$

Une statistique suffisante est le logarithme du rapport de vraisemblance, $\eta_{\text{GLRT}} = \log(\eta_{\text{GLRT}})/N$. En utilisant $\det(\mathbf{I} - \mathbf{AB}) = \det(\mathbf{I} - \mathbf{BA})$, nous avons la statistique sous la forme que nous allons utiliser à partir de maintenant :

$$\eta_{\text{GLRT}} = -\log\det\Big[\mathbf{I}_L - \frac{\mathbf{S}\mathbf{Y}^*}{N}\Big(\frac{\mathbf{Y}\mathbf{Y}^*}{N}\Big)^{-1}\frac{\mathbf{y}\mathbf{S}^*}{N}\Big(\frac{\mathbf{S}\mathbf{S}^*}{N}\Big)^{-1}\Big] \tag{4.17}$$

$$= -\log\det\Big(\mathbf{I}_M - \hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^*\Big) \tag{4.18}$$

Le GLRT est généralement considéré comme un test optimal, mais d'après l'équation (1.18) on peut voir que sa complexité devient rapidement démeusurée quand K, M et N deviennent grands. Pour calculer η_{GLRT}, nous devons, pour chaque indice, calculer et inverser une matrice $(M \times M)$ $\hat{\mathbf{R}}_{yy}$, et calculer le déterminant d'une matrice carrée de taille $\min(M, K, N)$.

Une alternative à la statistique GLRT est la statistique MMSE, qui est aussi robuste aux interférences. La statistique MMSE minimise l'erreur LS entre la séquence de synchronisation connue $\mathbf{S}$ et son estimation LS obtenue par filtrage spatial des données $\mathbf{Y}$. Le critère MMSE est donné par

$$\eta_{\text{MMSE}} = \frac{\text{Tr}(\hat{\mathbf{R}}_{ys}^*\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys})}{\text{Tr}(\mathbf{R}_{ss})} \tag{4.19}$$

Il est facile de voir que η_{MMSE} est moins complexe à calculer que η_{GLRT}. Il faut encore une inversion de matrice à chaque indice de temps, mais nous n'avons pas de déterminant à calculer. Cependant, comme il sera montré dans le chapitre 2, la MMSE devient sous-optimale pour des séquences d'apprentissage non-orthogonales.

Une troisième statistique intéressante, robuste aux interférences, est la statistique GLRT optimisée pour le cas théorique où $\mathbf{R}$ est connue et $\mathbf{H}$ est inconnue.

$$\eta_{\text{GLRT, kn}} = \frac{\underset{\mathbf{H}}{\arg\max} \prod_{k=1}^{M} p_{H_1}(\mathbf{y}(k) \mid \mathbf{S}(k), \mathbf{H}, \mathbf{R})}{\prod_{k=1}^{M} p_{H_0}(\mathbf{y}(k) \mid \mathbf{R})} \tag{4.20}$$

$$= \text{Tr}\Big[\mathbf{R}^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^*\Big]. \tag{4.21}$$

Ceci est bien sûr un critère purement théorique puisqu'il est impossible en pratique de connaître la matrice de covariance du bruit $\mathbf{R}$. L'intérêt de ce critère découle de l'idée que la matrice $\mathbf{R}$ peut être remplacée par son estimation. Les performances des statistiques non GLRT obtenues peuvent ensuite être comparées avec celles de la GLRT η_{GLRT}.

4.4.3 Réduction de la complexité et optimisation des paramètres de synchronisation MIMO

Avec ces considérations comme point de départ, nous proposons dans le chapitre 2 plusieurs moyens d'optimiser et de réduire la complexité de la synchronisation MIMO, résumé comme suit:

1. Introduction des expressions explicites pour la synchronisation GLRT en l'absence d'interférence, pour deux types de bruit spatialement blanc.
2. Introduction d'une expression alternative de η_{GLRT} optimisée en présence d'interférences, où le calcul du déterminant est fait explicitement pour $K = 2$. Ces expressions sont particulièrement utiles car elles permettent la comparaison directe de η_{GLRT} avec η_{MMSE}.
3. Introduction de deux nouvelles statistiques de faible complexité, η_{GLRT0} et η_{GLRT1}, basé sur le GLRT optimisé pour le cas théorique où la covariance du bruit $\mathbf{R}$ est connue.
4. Proposition d'une procédure de réduction du taux de calcul de la matrice de corrélation du signal reçu.
5. Étude du problème de l'optimisation du nombre d'antennes d'émission pour la synchronisation temporelle.

Le tableau 4.1 résume les différentes statistiques, GLRT ou non GLRT, qui sont étudiées ou introduites dans cette thèse.

Nom	**Interprétation**	**statistiques suffisantes**
η_{GLRT}	statistiques GLRT pour $\mathbf{R}$ et $\mathbf{H}$ inconnues	$-\log\det\left(\mathbf{I}_M - \hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*}\right)$
$\eta_{\text{GLRT, nous}}$	statistiques GLRT pour bruit spatialement blanc avec $\mathbf{R} = \sigma^2\mathbf{I}_M$	$\frac{\text{Tr}(\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*})}{\text{Tr}(\hat{\mathbf{R}}_{yy})}$
$\eta_{\text{GLRT, wu}}$	statistiques GLRT pour bruit spatialement blanc avec $\mathbf{R} = \text{Diag}[\sigma_1^2,\ldots,\sigma_M^2]\mathbf{I}_M$	$\frac{\prod_{n=1}^{M}(\hat{\mathbf{R}}_{yy})_{n,n}}{\prod_{n=1}^{M}(\hat{\mathbf{R}}_{yy}-\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*})_{n,n}}$
$\eta_{\text{GLRT, kn}}$	statistiques GLRT pour $\mathbf{R}$ connue et $\mathbf{H}$ inconnue	$\text{Tr}\left[\mathbf{R}^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*}\right]$
η_{GLRT0}	Statistiques non-GLRT basé sur $\eta_{\text{GLRT, kn}}$, avec $\mathbf{R} = \hat{\mathbf{R}}_{yy}$.	$\text{Tr}\left[\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*}\right]$
η_{GLRT1}	Statistiques non-GLRT basé sur $\eta_{\text{GLRT, kn}}$, avec $\mathbf{R} = \hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*}$	$\text{Tr}\left[(\hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*})^{-1}\hat{\mathbf{R}}_{ys}\mathbf{R}_{ss}^{-1}\hat{\mathbf{R}}_{ys}^{*}\right]$

Table 4.1: statistiques de synchronisation étudiées dans cette thèse

4.5 Analyse du grand système avec matrices aléatoires

4.5.1 Introduction au problème

La deuxième partie de la thèse, traité dans le chapitre 3, traite le comportement d'un critère GLRT MIMO qui permet de détecter la présence d'un signal connu corrompu par un canal de propagation par trajets multiples et d'un bruit additif gaussian, blanc temporellement, avec une matrice de covariance spatiale inconnue. Le chapitre se concentre sur le scenario où le nombre de capteurs M et éventuellement le nombre de chemins L est grand, et du même ordre de grandeur que le nombre d'échantillons N. Ce contexte est modélisé par deux régimes asymptotiques. Le premier est $M \to +\infty$, $N \to +\infty$ de telle sorte que $M/N \to c$ pour $C \in (0, +\infty)$. Le second régime asymptotique sous étude est $N, M, L \to +\infty$, de telle sorte que $M/N \to c$ pour c $\in (0, +\infty)$, et $L/N \to d$ pour $d \in (0, +\infty)$. Le but de ce chapitre est d'étudier le comportement de la statistique GLRT dans ces régimes, et de montrer que l'analyse théorique correspondant permet de prédire avec précision les performances du test lorsque M, N et peut-être L sont du même ordre de grandeur.

4.5.2 Pourquoi l'approximation du grand système?

Pour voir que le problème étudié dans le chapitre 3 n'est pas simplement une considération théorique, considérons comme exemple les systèmes multi-antennes d'aujourd'hui. La norme

d'échantillonnage qui peut en pratique être utilisée pour effectuer l'inférence statistique ne peut pas être beaucoup plus grande que la dimension du signal, ce qui correspond au nombre d'antennes en réception.

Nous avons donc une distinction entre deux types de résultats limitatifs: Le soi-disant *problème de limite classique*, où seulement une des dimensions est supposée grande, et le *problème de limite en grande dimension*, où plusieurs dimensions sont supposées tendre vers l'infini. On peut supposer que l'analyse classique fonctionne mal lorsque plusieurs dimensions sont du même ordre de grandeur. L'un des objectifs de cette étude est donc de montrer que les résultats limitatifs classiques ne sont plus valables pour les systèmes en grande dimension.

Une façon d'aborder le problème d'analyse asymptotique en grande dimension est la théorie des matrices aléatoires (Random matrix theory, RMT). La motivation sous-jacente de l'utilisation de cet outil est le comportement non évident des grandes matrices aléatoires. À titre d'exemple, étudions la matrice de covariance empirique d'une matrice d'observations des grandes dimensions. Soit $\mathbf{Y}_n)_{n=1,\ldots,N}$ les N $(M \times 1)$ vecteurs d'observation, nous voulons étudier la matrice de covariance empirique donnée par

$$\mathbf{R}_N = \frac{1}{N}\sum_{i=1}^{N} \mathbf{Y}_i \mathbf{Y}_i^*. \tag{4.22}$$

Si M est fixé quand $N \to \infty$, $\mathbf{R}_N \to \mathbf{R} = \mathbb{E}[|\mathbf{Y}_1\mathbf{Y}_1^*|]$, et $\|\mathbf{R} - \mathbf{R}_N\| \Rightarrow \mathbf{0}$ pour toute norme matricielle [29]. Cette matrice de covariance empirique $\mathbf{R}_N$ est un estimateur consistant de matrice de covariance de la population $\mathbf{R}$. Cependant, dans de nombreuses applications pratiques, le nombre d'observations disponibles N à le même ordre de grandeur que M. Dans ce cas, $|\mathbf{R}-\mathbf{R}_N\|$ peut être loin de zéro, même quand N est grand. Par exemple, si $M > N$, $\mathbf{R}_N$ est de rang incomplet (alors que $\mathbf{R}$ peut être de plein rang), il n'est donc pas une bonne approximation de $\mathbf{R}$. Si N et M sont tous grands par rapport à 1, mais de la même ordre de grandeur, il a été démontré que la distribution empirique des valeurs propres $\mathbf{R}_N$ est différente de la distribution des valeurs propres de $\mathbf{R}$. Par exemple, si $\mathbf{R} = \sigma^2\mathbf{I}_M$, l'histogramme des valeurs propres de $\mathbf{R}_N$ tend à se rapprocher de la densité de probabilité de la distribution Marcenko-Pastur [30]. Ceci est en contraste avec le cas $N >> M$, où les valeurs propres de $\mathbf{R}_N$ sont concentrées autour de σ^2.

Qu'est-ce que l'infini? Comme mentionné précédemment, l'analyse asymptotique est effectuée pour le cas où les dimensions N, K (Et peut-être aussi L) $\to \infty$. Cependant, les résultats obtenus avec ces méthodes sont applicables même avec de petites dimensions. En d'autres termes: pour plus de commodité théorique, les dimensions sont supposées grandes, mais il peut être démontré par simulation, que les résultats sont valables même pour les petites dimensions.

4.5.3 Préliminaires

Dans le chapitre 3, nous allons effectuer l'analyse asymptotique pour étudier la distribution limite de la statistique de synchronisation GLRT étudiée au chapitre 2. Nous étudions le cas où une antenne d'émission ($K = 1$) envoie une séquence de synchronisation connue de longueur N par un canal sélectif en fréquence $\mathbf{H}$ muni de L chemins. Une statistique suffisante du critère GLRT est $eta_N = \log(\eta'_{\mathrm{GLRT}})/N$, défini par

$$\eta_N = -\log\det\left[\mathbf{I}_L - \frac{\mathbf{S}\mathbf{Y}^*}{N}\left(\frac{\mathbf{Y}\mathbf{Y}^*}{N}\right)^{-1}\frac{\mathbf{y}\mathbf{S}^*}{N}\left(\frac{\mathbf{S}\mathbf{S}^*}{N}\right)^{-1}\right], \tag{4.23}$$

Avant de démarrer le calcul de l'analyse asymptotique, on remarque qu'il est possible d'assumer sans restriction que $\frac{\mathbf{S}\mathbf{S}^*}{N} = \mathbf{I}_L$ et que $\mathbb{E}(\mathbf{v}_n\mathbf{v}_n^*) = \Sigma^2\mathbf{I}$, c'est-à-dire que $\tilde{\mathbf{R}}$ est réduite à la matrice identité. Si ce n'est pas le cas, on note alors $\tilde{\mathbf{S}}$ la matrice

$$\tilde{\mathbf{S}} = \left(\frac{\mathbf{S}\mathbf{S}^*}{N}\right)^{-1/2}\mathbf{S} \tag{4.24}$$

et par $\tilde{\mathbf{Y}}$ et $\tilde{\mathbf{V}}$ les matrices d'observation et de bruit blanchies.

$$\begin{aligned}\tilde{\mathbf{Y}} &= \tilde{\mathbf{R}}^{-1/2},\mathbf{Y},\\ \tilde{\mathbf{V}} &= \tilde{\mathbf{R}}^{-1/2}\mathbf{V}\end{aligned} \tag{4.25}$$

Il est clair que $\frac{\tilde{\mathbf{S}}\tilde{\mathbf{S}}^*}{N} = \mathbf{I}_L$ et que $\mathbb{E}(\tilde{\mathbf{v}}_n\tilde{\mathbf{v}}_n^*) = \sigma^2\mathbf{I}$. De plus, sous H_0, $\tilde{\mathbf{Y}} = \tilde{\mathbf{V}}$, tandis que sous H_1, $\tilde{\mathbf{Y}} = \tilde{\mathbf{H}}\tilde{\mathbf{S}} + \tilde{\mathbf{V}}$ où la matrice de canal $\tilde{\mathbf{H}}$ est définie par

$$\tilde{\mathbf{H}} = \tilde{\mathbf{R}}^{-1/2}\,\mathbf{H}\,(\mathbf{S}\mathbf{S}^*/N)^{1/2} \tag{4.26}$$

Enfin, les statistiques η_N peuvent également être écrits comme

$$\eta_N = \log\det\Big[\mathbf{I}_L - \frac{\tilde{\mathbf{S}}\tilde{\mathbf{Y}}^*}{N}\Big(\frac{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^*}{N}\Big)^{-1}\frac{\tilde{\mathbf{Y}}\tilde{\mathbf{S}}^*}{N}\Big] \tag{4.27}$$

Cela montre qu'il est possible de remplacer $\mathbf{S}$, $\tilde{\mathbf{R}}$ et $\mathbf{H}$ par $\tilde{\mathbf{S}}$, $\mathbf{I}$ et $\tilde{\mathbf{H}}$ sans modifier la valeur de la statistique η_N. Les résultats sont valables en remplaçant $\mathbf{H}$ avec $\tilde{\mathbf{H}}$, et nous pouvons sans restriction faire les calculs avec l'hypothèse que

$$\frac{\mathbf{S}\mathbf{S}^*}{N} = \mathbf{I}_L,\ \tilde{\mathbf{R}} = \mathbf{I}_M \tag{4.28}$$

Une deuxième étape préliminaire est l'introduction des variables auxiliaires que nous appelons $\mathbf{V}_1$ et $\mathbf{V}_2$. On note $\mathbf{W}$ la matrice $(N-L)\times N$ pour laquelle la matrice $\boldsymbol{\Theta} = (\mathbf{W}^T, \frac{\mathbf{S}^T}{\sqrt{N}})^T$ est unitaire et on définit les matrices $\mathbf{V}_1$ et $\mathbf{V}_2$ aux tailles respectives $M\times(N-L)$ et $M\times L$ par

$$(\mathbf{V}_1, \mathbf{V}_2) = \mathbf{V}\boldsymbol{\theta}^* = (\mathbf{V}\mathbf{W}^*, \mathbf{V}\frac{\mathbf{S}^*}{\sqrt{N}}) \tag{4.29}$$

Il est clair que $\mathbf{V}_1$ et $\mathbf{V}_2$ sont des matrices aléatoires gaussiennes complexes aux entrées $\mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ iid, et que les entrées de $\mathbf{V}_1$ et $\mathbf{V}_2$ sont mutuellement indépendantes. Nous remarquons que si $N > M + L$, la matrice $\frac{\mathbf{V}_1\mathbf{V}_1^*}{N}$ est presque sûrement inversible. Nous pouvons maintenant exprimer la statistique η_N en termes de $\mathbf{V}_1$ et $\mathbf{V}_2$. Sous H_0, il est montré dans le chapitre 3 que les statistiques peuvent être écrites comme

$$\eta_N = \log\det\Big(\mathbf{I}_L + \mathbf{V}_2^*/\sqrt{N}\Big(\mathbf{V}_1\mathbf{V}_1^*/N\Big)^{-1}\mathbf{V}_2/\sqrt{N}\Big) \tag{4.30}$$

et de même sous H_1, où

$$\eta_N = \log\det\left(\mathbf{I}_L + \Big(\mathbf{V}_2 + \mathbf{H}\Big)^*/\sqrt{N}\ (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}\ \Big(\mathbf{V}_2 + \mathbf{H}\Big)/\sqrt{N}\right). \tag{4.31}$$

Le point clé ici est l'indépendance entre $\mathbf{V}_1$ et $\mathbf{V}_2$, qui peut être exploitée pour simplifier les calculs des distributions limites.

Noton que les distributions asymptotiques de ce genre de statistiques ont été étudiées par le passé, mais dans le régime classique, défini par

1. *régime asymptotique (a)*: $\begin{cases} L, M \ \text{ est fixé} \\ N \to \infty \end{cases}$

Lorsque la taille de M et L augmentent, l'hypothèse que $M << N$ et / ou $L << N$ ne sont plus valables, ce qui est le cas pour de nombreuses applications pratiques. Par conséquent, nous considérons deux nouveaux régimes asymptotiques:

1. *régime asymptotique (b)*: $\begin{cases} L \ \text{ est fixé} \\ N, M \to \infty \\ M/N = c_n \to c,\ 0 < c < 1 \end{cases}$
2. *régime asymptotique (c)*: $\begin{cases} N, M, L \to \infty \\ M/N = c_n \to c,\ 0 < c < 1 \\ L/N = d_N \to d,\ 0 < d < 1 \\ C + d < 1 \end{cases}$

4.5.4 Exemple d'analyse asymptotique: l'espérance de η_N sous de H_0

Pour avoir une idée des différences qu'ont les régimes en ce qui concerne leur analyse asymptotique, nous présenterons comme exemple l'analyse asymptotique qui permet d'obtenir l'espérance de η_N sous H_0, pour les régimes asymptotiques (a) et (b).

Commençons par l'analyse dans le régime classique (a). Nous utilisons (4.30) et remarquons que lorsque $N \to +\infty$ et que M et L restent fixes, les matrices $\mathbf{V}_1\mathbf{V}_1^*/N$ et $\frac{1}{N}\mathbf{V}_2^*\left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\mathbf{V}_2$ convergent respectivement vers $\sigma^2\mathbf{I}$ et la matrice zéro. De plus,

$$\frac{1}{N}\mathbf{V}_2^*\left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\mathbf{V}_2 = \frac{1}{\sigma^2}\mathbf{V}_2^*\mathbf{V}_2/N + o_P(\frac{1}{N}) \tag{4.32}$$

et une expansion du second ordre standard η_N conduit à

$$\eta_N = \frac{1}{\sigma^2}\mathrm{Tr}\left(\mathbf{V}_2^*\mathbf{V}_2/N\right) + o_P(\frac{1}{N}) \tag{4.33}$$

Cela implique immédiatement que la distribution limite d'η_N est une distribution chi-carré avec $2ML$ degrés de liberté. De manière informel, cela implique que $\mathbb{E}(\eta_N) \simeq L\frac{M}{N}$ et $\mathrm{Var}(\eta_N) \simeq \frac{L}{N}\frac{M}{N}$.

L'analyse d'η_N dans le régime asymptotique (b) diffère profondément de l'analyse dans le régime asymptotique standard (a). En particulier, il n'est plus vrai que la matrice de covariance empirique $\mathbf{V}_1\mathbf{V}_1^*/N$ converge dans le sens de la norme spectrale vers $\sigma^2\mathbf{I}$. Cela en raison que le nombre d'entrées de cette matrice $M \times M$ est du même ordre de grandeur que le nombre d'observations scalaires disponibles. Nous notons également que pour toutes matrices deterministes $M \times M$ $\mathbf{A}$, les entrées diagonales de la matrice $L \times L$ $\frac{1}{N}\mathbf{V}_2^*\mathbf{A}\mathbf{V}_2$ convergent vers 0 lorsque $N \to +\infty$ et M reste fixe, ce qui n'est pas le cas lorsque M et N ont le même ordre de grandeur.

Nous voulons maintenant calculer l'espérance asymptotique d'η_N dans la région asymptotique (b), et montrer que

$$\eta_N - L\log\left(\frac{1}{1}c_n\right) \to 0 \ p.s. \tag{4.34}$$

Tout d'abord, notons $\mathbf{F}_N$ la matrice $L \times L$

$$\mathbf{F}_N = \mathbf{V}_2^*/\sqrt{N}\ \left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^{-1}\ \mathbf{V}_2/\sqrt{N}. \tag{4.35}$$

et remarquons que, sous H_0, (4.30) peut être écrit comme

$$\eta_N = \log\det\Big(\mathbf{I}_L + \mathbf{F}_N\Big) \tag{4.36}$$

Comme L n'augmente pas avec M et N, il suffit d'établir que

$$\mathbf{F}_N - \frac{c_n}{1-c_n}\,\mathbf{I}_L \to 0 \ p.s. \tag{4.37}$$

Notre approche est basée sur l'observation que si $\mathbf{A}_N$ est une matrice hermitienne déterministe $M \times M$ vérifiant $\sup_N \|\mathbf{A}_N\| < A < +\infty$, alors

$$\mathbb{E}_{\mathbf{V}_2}\Big|\Big(\mathbf{V}_2^*/\sqrt{N}\,\mathbf{A}_N\mathbf{V}_2/\sqrt{N}\Big)_{k,l} - \frac{\sigma^2}{N}\mathrm{Tr}(\mathbf{A}_N)\,\Delta(k-l)\Big|^4 \ \leq \frac{C(a)}{N^2} \tag{4.38}$$

où $C(a)$ est un terme constant en fonction de a, et où $\mathbb{E}_{\mathbf{V}_2}$ représente l'espérance mathématique par rapport à $\mathbf{V}_2$. Ceci est une conséquence de la proposition 4 dans l'Appendix 3.7. Supposons pour le moment qu'il existe une constante a déterministe telle que

$$\|\Big(\mathbf{V}_1\mathbf{V}_1^*/N\Big)^{-1}\| \leq a \tag{4.39}$$

pour chaque N supérieur à un entier non-aléatoire N_0. Ensuite, comme $\mathbf{V}_1$ et $\mathbf{V}_2$ sont indépendants, il est possible d'utiliser (4.38) pour $\mathbf{A}_N = (\mathbf{V}_1\mathbf{V}_1^*/N)^{-1}$ et de prendre l'espérance mathématique par rapport à $\mathbf{V}_1$ (4.38) pour obtenir que

$$\mathbb{E}\left|\Big(\mathbf{F}_N\Big)_{k,l} - \frac{\sigma^2}{N}\mathrm{Tr}\Big(\mathbf{V}_1\mathbf{V}_1^*/N\Big)^{-1}\delta(kl)\right|^4 \leq \frac{C(a)}{N^2} \tag{4.40}$$

pour chaque $N > N_0$ et en utilisant le lemme de Borel-Cantelli, que

$$\mathbf{F}_N - \frac{\sigma^2}{N}\mathrm{Tr}\left(\mathbf{V}_1\mathbf{V}_1^*/N\right)^1\ \mathbf{I}_L \to 0\ p.s. \tag{4.41}$$

Pour conclure, nous utilisons les résultats connus liés à la convergence presque sûre de la distribution des valeurs propres de la matrice $\mathbf{V}_1\mathbf{V}_1^*/N$ pour la distribution Marcenko-Pastur (voir l'équation (3.77) dans l'Annexe 3.7) qui impliquent presque sûrement que

$$\frac{1}{N}\mathrm{Tr}\Big(\mathbf{V}_1\mathbf{V}_1^*/N\Big)^{-1} - \frac{c_n}{\sigma^2(1-c_n)} \to 0 \tag{4.42}$$

Ceci, en liaison avec (4.41), conduit à (4.37) et éventuellement à (1.34).

Cependant, il n'existe qu'une constante a déterministe satisfaisant (4.39) pour chaque N supérieur à un nombre entier non aléatoire. Afin de résoudre ce problème, il suffit de remplacer la matrice $\Big(\mathbf{V}_1\mathbf{V}_1^*/N\Big)^{-1}$ par une version régularisée. Les détails de la régularisation seront donnés dans le chapitre 3

4.5.5 Distributions limites

Les principaux résultats de ces travaux sont les distributions limites résumées dans le tableau 4.2. Noton que la distribution limite pour la région (c) doit établir un théorème de la limite centrale pour une statistique linéaire des valeurs propres d'une matrice F (F-matrix) de moyenne non-nulle, une tâche difficile qui n'a pas été abordée dans ce travail. Au lieu de cela, nous avons proposé une approximation qui sera décrite en détail dans le chapitre 3.

Régime asymptotique	**Distribution sous** H_0	**Distribution sous** H_1
(a) Classique, $N \to \infty$	$\eta_N \sim \frac{1}{N}\chi^2_{2ML}$ $\Big(\mathbb{E}[\eta_N] = Lc_N$ $\mathrm{Var}[\eta_N] = Lc_N \cdot \frac{1}{N}\Big)$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\Big(\log\det\Big(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^*}{\sigma^2}\Big), \frac{\kappa_1}{N}\Big)$
(b) Proposé, $M, N \to \infty$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\Big(L\log\frac{1}{1}c_n, \frac{Lc_N}{1-c_n} \cdot \frac{1}{N}\Big)$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\Big(L\log\frac{1}{1}c_n + \log\det\Big(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^*}{\sigma^2}\Big),$ $\frac{\kappa_1}{N} + \frac{Lc_N}{1-c_n} \cdot \frac{1}{N}\Big)$
(c) Proposé, $L, M, N \to \infty$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\Big(\tilde{\eta}_N, \tilde{\delta}_N\Big)$	$\eta_N \sim \mathcal{N}_{\mathbb{R}}\Big(\tilde{\eta}_N + \log\det\Big(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^*}{\sigma^2}\Big), \frac{\kappa_1}{N} + \tilde{\delta}_N\Big)$

Table 4.2: Distributions asymptotiques de η_N pour les trois régimes asymptotiques, sous H_0 *et* H_1

Bibliography

[1] A.N. Mody and G.L. Stuber, "Synchronization for mimo ofdm systems," *Proc. IEEE Global Telecommunications Conf. GLOBECOM*, vol. 1, pp. 509–513, 2001.

[2] A. Saemi, J. Cances, and V. Meghdadi, "Synchronization algorithms for mimo ofdma systems," *IEEE Trans. Wireless. Commun.*, vol. 6, no. 12, pp. 4441–4451, 2007.

[3] C.L Wang and H.C Wang, "Optimized joint fine timing synchronization and channel estimation for mimo systems," *IEEE Trans. on Communications*, vol. 59, no. 4, pp. 1089–1098, 2011.

[4] M. Marey, O.A. Dobre, and R. Inkol, "A novel blind block timing and frequency synchronization algorithm for alamouti stbc," *IEEE Comm. Letters*, vol. 17, no. 3, pp. 569–572, 2013.

[5] Y. Wu, S. Chan, and E. Serpedin, "Symbol timing estimation in space-time coding systems based on orthogonal training sequences," *IEEE Trans. On Wireless Communications*, vol. 4, no. 2, pp. 603–613, 2005.

[6] S.H Won and L. Hanzo, "Non-coherent code acquisition in the multiple transmit/multiple receive antenna aided single- and multi-carrier ds-cdma downlink," *IEEE Trans. On Wireless Communications*, vol. 6, no. 11, pp. 3864–3869, 2007.

[7] T. Tang and R. Heath, "A space-time receiver with joint synchronization and interference cancellation in asynchronous mimo-ofdm systems," *IEEE Trans. Vehicular Technology*, vol. 57, no. 5, pp. 2991–3005, 2008.

[8] D. W. Bliss and P. A. Parker, "Temporal synchronization of mimo wireless communication in the presence of interference," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1794–1806, 2010.

[9] Y. Zhou, E.Serpedin, K. Qarage, and O. Dobre, "On the performance of generalized likelihood ratio test for data-aided timing synchronization of mimo systems," *Proc. International Conf. on Communications (ICC'12), Bucharest*, pp. 43–46, 2012.

[10] A. Naguib, V. Tarokh, N. Seshadri, and A.R. Calderbank, "A space-time coding modem for high data rate wireless communications," *IEEE Journal on Selected areas in Communications*, vol. 16, no. 8, pp. 1459–1478, 1998.

[11] A.N. Mody and G.L. Stuber, "Receiver implementation for a mimo ofdm system," *Proc. IEEE Global Telecommunications Conf. GLOBECOM*, vol. 1, pp. 716–720, 2002.

[12] Y. Zhang and S.L. Miller, "Code acquisition in transmit diversity ds-cdma systems," *IEEE Trans. On Communication*, vol. 51, no. 8, pp. 1378–1388, 2003.

[13] T.C.W. Schenk A. van Zelst, "Implementation of a mimo ofdm-based wireless lan system," *IEEE Trans. On Signal Processing*, vol. 52, no. 2, pp. 483–494, 2004.

[14] G.L. Stuber, J.R. Barry, S.W. McLaughlin, Y. Li, M.A. Ingram, and T.G. Pratt, "Broadband mimo-ofdm wireless communications," *Proc. IEEE*, vol. 92, no. 2, pp. 271–294, 2004.

[15] A. Saemi, V. Meghdadi, J.P. Cances, M.J Syed, G. Ferre, and J.M Dumas, "Fine timing and frequency synchronization for mimo system," *Proc. IST Mobile and Wireless Communications Summit, Dresden, Germany*, June 2005.

[16] E. Zhou, X. Zhang, H. Zhao, and W. Wang, "Synchronization algorithms for mimo ofdm systems," *Proc. Wireless Commun. and Networking Conf.*, pp. 18–22, 2005.

[17] Z. Ma, X. Wu, and W. Zhu, "An ici-free synchronization algorithm in mimo-ofdm system," *Proc. Int. Symposium Wireless Pervasive Computing (ISWPC'07)*, pp. 242–246, Feb. 2007.

[18] L. Qi and H. Bo, "Joint timing synchronization and frequency-offset acquisition algorithm for mimo ofdm systems," *Journal of Systems Engineering and Electronics*, vol. 20, no. 3, pp. 470–478, 2009.

[19] N. Han, N. Du, and Y. Ma, "Research of time-frequency synchronization in mimo-ofdm system," *Proc. IEEE Symposium on Electrical and Electronics Engineering (EEESYM'12)*, pp. 555–558, 2012.

[20] T.L. Kung and K.K. Parhi, "Optimized joint timing synchronization and channel estimation for communications systems with multiple transmit antennas," *EURASIP Journal on Advances in Signal Processing*, vol. 2013:139, pp. 1–12, 2013.

[21] N.D. Long and H. Park, "Joint fine time synchronization and channel estimation for mimo-ofdm wlan," *Proc. IEEE Intelligent Signal Processing and Communication Systems Conference (ISPACS'04)*, pp. 463–467, 2004.

[22] A. Saemi, V. Meghdadi, J.P. Cances, M.R. Zahabi, and J.M Dumas, "Ml mimo-ofdm time synchronization/channel estimation for unknown frequency selective fading channels," *Proc. International Symposium ELMAR'06, Zadar, Croatia*, pp. 255–258, June 2006.

[23] K. Rajawat and A. Chaturvedi, "A low complexity symbol timing estimator for mimo systems using two samples per symbol," *IEEE Comm. Letters*, vol. 10, no. 7, pp. 525–527, 2006.

[24] D. Wang and J. Zhang, "Timing synchronization for mimo-ofdm wlan systems," *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, pp. 1177–1182, 2007.

[25] S.H Won and L. Hanzo, "Analysis of serial-search-based code acquisition in the multiple-transmit/multiple-receive-antenna-aided ds-cdma downlink," *IEEE Trans. On Vehicular Technology*, vol. 57, no. 2, pp. 1032–1039, 2008.

[26] S.H Won and L. Hanzo, "Non-coherent and differentially coherent code acquisition in mimo assisted ds-cdma multi-path downlink scenatios," *IEEE Trans. On Wireless Communications*, vol. 7, no. 5, pp. 1585–1593, 2008.

[27] S.H Won and L. Hanzo, "Initial and post-initial code acquisition in the noncoherent multiple-input/multiple-output-aided ds-cdma downlink," *IEEE Trans. On Vehicular Technology*, vol. 58, no. 5, pp. 2322–2330, 2009.

[28] H.L Van Trees, *Detection, Estimation and Modulation Theory – Part I*, John Wiley and Sons, 1968.

[29] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*, Cambridge University Press, 1st edition, 2011.

[30] L.A. Pastur and M. Shcherbina, *Eigenvalue Distribution of Large Random Matrices*, Mathematical Surveys and Monographs, Providence: American Mathematical Society, 2011.

[31] G.J. Foschini and M.J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, 1998.

[32] E.I. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. Comm.*, vol. 10, pp. 585–596, 1999.

[33] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, 1998.

[34] V. Tarokh, H. Jafarkhani, and A.R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.

[35] A. Stamoulis A.R. Calderbank S.N Diggavi, N. Al-Dhahir, "Great expectations: The value of spatial diversity to wireless networks," *Proc. IEEE*, vol. 92, no. 2, pp. 219–270, 2004.

[36] D. Astely, E. Dahlman, A. Furuskar, Y. Jading, M. Lindstrom, and S. Parkvall, "Lte: The evolution of mobile broadband," *IEEE Communications Magazine*, pp. 44–51, april 2009.

[37] D. Bai, C. Park, J. Lee, H. Nguyen, J. Singh, A. Gupta, Z. Pi, T. Kim, C. Lim, M.G. Kim, and I. Kang, "Lte-advanced modem design: Challenges and perspectives," *IEEE Communications Magazine*, pp. 178–186, february 2009.

[38] G.J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, pp. 41–59, 1996.

[39] A. Naguib, N. Seshadri, and A.R. Calderbank, "Space-time coding and signal processing for high data rate wireless communications," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 76–92, 2000.

[40] F. Dupuy P. Chevalier, "Widely linear alamouti receivers for the reception of real-valued signals corrupted by interferences – the alamouti-saic/maic concept," *IEEE Trans. Signal Processing*, vol. 59, no. 7, pp. 3339–3354, 2011.

[41] L.E. Brennan and I.S. Reed, "An adaptive array signal processing algorithm for communications," *IEEE Trans. Aerosp. Electronic Systems*, vol. 18, no. 1, pp. 124–130, 1982.

[42] D.M. Duglos and R.A. Scholtz, "Acquisition of spread spectrum signals by an adaptive array," *IEEE Trans. Acou. Speech. Signal Proc.*, vol. 37, no. 8, pp. 1253–1270, 1989.

[43] T. Koivisto and V. Koivunen, "Diversity transmission of synchronization sequences in mimo systems," *IEEE Trans. On Wireless Comm.*, vol. 11, no. 11, pp. 4048–4057, 2012.

[44] A. De Maio, M. Lops, and L. Venturino, "Diversity-integration trade-offs in mimo detection," *IEEE Trans. Signal Proc.*, vol. 56, no. 10, 2008.

[45] D. C. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 531–532, 1972.

[46] X. Mestre and M.A Lagunas, "Modified subspace algorithms for doa estimation with large arrays," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 598–614, 2008.

[47] P. Vallet, P. Loubaton, and X. Mestre, "Improved subspace estimation for multivariate observations of high dimension: the deterministic signals case," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1043–1068, 2012.

[48] P. Bianchi, M. Debbah, M. Maeda, and J. Najim, "Performance of statistical tests for single source detection using random matrix theory," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2400–2419, 2011.

[49] S. Kritchman and B. Nadler, "Non-parametric detection of the number of signals, hypothesis testing and random matrix theory," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3930–3941, 2009.

[50] R.R. Nadakuditi and J.W Silverstein, "Fundamental limit of sample generalized eigenvalue based detection of signals in noise using relatively few signal-bearing and noise-only samples," *IEEE Journal on Selected Topics in Signal Process.*, vol. 4, no. 3, pp. 468–480, 2010.

[51] R.R. Nadakuditi and A. Edelman, "Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2625–2637, 2008.

[52] D. Astely, A. Jakobsson, and A.L. Swindlehurst, "Burst synchronization on frequency selective channels with co-channel interference using an antenna array," *Proc. IEEE 49th Veh. Technol. Conf.*, vol. 49, pp. 2363–2367, 1999.

[53] A. Dogandzic and A. Nehorai, "Finite-length mimo equalization using canonical correlation analysis," *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 984–989, 2002.

[54] Y. Jiang, P. Stoica, and J. Li, "Array signal processing in the known waveform and steering vector case," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 23–35, 2004.

[55] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*, Prentice Hall, 1st edition, 1998.

[56] M. Viberg, P. Stoica, and B. Ottersten, "Maximum likelihood array processing in spatially correlated noise fields using parameterized signals," *IEEE Trans. Signal Process.*, vol. 45, no. 4, pp. 996–1004, 1997.

[57] Z. Bai and J.W. Silverstein, *Spectral analysis of large dimensional random matrices*, Springer Series in Statistics, 2nd edition, 2010.

[58] S. Zheng, "Central limit theorems for linear spectral statistics of large dimensional f-matrices," *Ann. Inst. Henri Poincaré*, , no. 2, pp. 444–476, 2012.

[59] G.A. Young and R.L. Smith, *Essentials of statistical inference*, Cambridge Series in Statistical and Probabilistic Mathematics, 2005.

[60] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton, and J. Najim, "On the capacity achieving transmit covariance matrices for mimo rician channels: an asymptotic approach," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1048–1069, 2010.

[61] W. Hachen, P. Loubaton, J. Najim, X. Mestre, and P. Vallet, "On bilinear forms based on the resolvent of large random matrices," *Annales Inst. Henri Poincaré-Probabilités et Statistiques*, vol. 49, no. 1, pp. 36–63, 2013.

[62] W. Hachem, P. Loubaton, J. Najim, X. Mestre, and P. Vallet, "Large information plus noise random matrix models and consistent subspace estimation in large sensor networks," *to appear in Random Matrices, Theory and Applications (RMTA), vol. 1, no. 2 (2012), also available on Arxiv (arXiv:1106.5119)*, 2012.

[63] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur, "A new approach for capacity analysis of large dimensional multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3987–4004, 2008.