HAL
open science

Théorie des matrices aléatoires robustes et applications à la détection radar 1

Frédéric Pascal, Abla Kammoun

To cite this version:

Frédéric Pascal, Abla Kammoun. Théorie des matrices aléatoires robustes et applications à la détection radar 1. Traitement du Signal, 2016, Théorie des Matrices Aléatoires et Applications, 33 (2-3), pp.321-349. 10.3166/ts.33.321-349 . hal-01378253

HAL Id: hal-01378253

https://hal.science/hal-01378253v1

Submitted on 26 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Théorie des matrices aléatoires robustes et applications à la détection radar [1]

Frédéric PASCAL [1], Abla KAMMOUN [2]

1. *Laboratoire des Signaux et Systèmes (L2S, UMR CNRS 8506)*
   CentraleSupélec-CNRS-Université Paris-Sud, 91192 Gif-sur-Yvette, France
   frederic.pascal@centralesupelec.fr
2. *Computer, Electrical, and Mathematical Sciences and Engineering (CEMSE) Division, KAUST, Thuwal, Makkah Province, Saudi Arabia*
   abla.kammoun@kaust.edu.sa

RÉSUMÉ. *Cet article présente de récents résultats issus de la combinaison entre la théorie des matrices aléatoire et la théorie de l'estimation robuste appliquées à des problèmes de détection en radar. Plus précisément, afin de pallier les problèmes de grande dimension des données, nous nous intéressons à une version régularisée de l'estimateur de matrice de covariance de Tyler (Tyler, 1987 ; Pascal, Chitour* et al.*, 2008). Nous montrons ainsi grâce à l'analyse statistique de ce dernier, i.e. l'étude de son comportement au premier et second ordre en régime de grande dimension [2], qu'un détecteur optimal (au sens de la maximisation des performances de détection et/ou de régulation de fausses alarmes) peut être construit. Enfin, des simulations Monte-Carlo montrent la pertinence de cette approche avec la comparaison aux méthodes traditionnellement utilisées.*

ABSTRACT. *This article presents recent results obtained from both Random Matrix Theory and Robust Estimation Theory, and applied to radar detection problems. More precisely, to answer the problem of high dimensional data, we focus on a regularized version of the Tyler's covariance matrix estimator (Tyler, 1987 ; Pascal, Chitour* et al.*, 2008). Thus, it is shown thanks to the statistical analysis of this estimator, i.e. first and second-order behavior in high dimensional regime [3], that an optimal design of a robust detector, namely the adaptive normalized matched filter (ANMF) can be derived. The optimality considered in this paper refers to the maximisation (resp. minimization) of the detection probability (resp. probability of false alarm). Finally, Monte-Carlo simulations are conducted to highlight the improvement brought by the proposed approach compared to classical techniques of the literature.*

1. Soutenu par l'institut iCODE, projet de recherche de l'Idex Paris-Saclay

2. c'est à dire lorsque le nombre d'observations n et leur dimension N tendent vers l'infini à la même vitesse : $N/n \to c \in (0, 1]$.

3. This means that both the number of observations n and their dimension N tend to infinity at the same rate : $N/n \to c \in (0, 1]$.

MOTS-CLÉS : *Matrices aléatoires, estimation robuste, régularisation, détection radar, ANMF*

KEYWORDS: *Random matrix theory, robust estimation theory, regularization, radar detection, ANMF*

DOI:10.3166/RIA.29.1-29 © 2015 Lavoisier

1. Introduction

L'estimation de matrices de covariance ou plus généralement de matrices de dispersion est un problème d'importance fondamentale pour de nombreuses applications en traitement du signal et en traitement radar. On peut citer par exemple les problèmes de détection adaptative, les problèmes d'estimation de directions d'arrivées ou encore certains problèmes de classification. Dans les applications radar spécifiques, on peut citer les traitements spatio-temporels adaptatifs (Space-Time Adaptive Processing, STAP) qui sont évoqués pour certaines configurations particulières des systèmes radar (J. Ward, 1994 ; Klemm, 2002 ; Bidon *et al.*, 2011). Ainsi, en détection radar (STAP ou classique), il est essentiel de construire un estimateur suffisamment précis de la matrice de covariance afin d'obtenir de bonnes performances en détection (voir par exemple (Olilla *et al.*, 2012 ; Mahot *et al.*, 2013) et les références associées). Les applications évoquées font souvent intervenir des observations de grande dimension et en pratique, il n'est pas toujours possible d'obtenir un nombre suffisant d'échantillons pour estimer la matrice de covariance. En effet, les données d'observations doivent être indépendantes et identiquement distribuées et ces hypothèses restreignent généralement la taille du support de l'échantillon d'apprentissage. Pour palier ce problème, i.e. lorsque la dimension des observations est supérieure aux nombre d'observations, des méthodes de régularisation pour l'estimation de matrice de covariance on été proposées (Y. Abramovich, 1981). Une approche classique parmi ces méthodes de régularisation est donnée par la matrice de covariance empirique régularisée (Regularized Sample Covariance Matrix, RSCM). Cette procédure d'estimation peut être vue comme la continuité de la méthode de "diagonal loading", introduit par les travaux d'Abramovich et de Carlson (Y. Abramovich, 1981 ; Carlson, 1988).

Selon la nature statistique du fouillis ou bruit additive, deux méthodes de régularisation peuvent être distinguées : la RSCM et l'estimateur régularisé de Tyler (regularized Tyler estimator, RTE) (Chen *et al.*, 2011 ; Pascal *et al.*, 2014). Ces deux estimateurs sont directement issus des méthodes de régularisation ou "shrinkage" de Ledoit-Wolf (Ledoit, Wolf, 2004).

Directement lié à la matrice de covariance empirique (sample covariance matrix, SCM), le RSCM conserve les principales propriétés de la SCM, notamment les performances limitées dans le cas d'observations hétérogènes, impulsives et/ou non-gaussiennes, mais encore dans le cas où des données aberrantes (ou outliers) sont présentes dans les observations. Ainsi, afin de travailler dans un cadre général, incluant les précédents scénarios, le fouillis est souvent modélisé par des distributions complexes elliptiques symétriques (complex elliptical symmetric distributions, CES), introduit par Kelker (Kelker, 1970). Dans ce contexte, la manière d'estimer la matrice de covariance ou plus généralement la matrice de dispersion [4] (scatter matrix) est plus difficile à mettre en œuvre. Ces problèmes d'estimation ont été abordés dans le

4. Contrairement à la matrice de covariance qui n'est définie que lorsque le moment d'ordre 2 de la distribution existe, la matrice de dispersion est toujours définie et reflète la structure de la distribution.

cadre de la théorie de l'estimation robuste par des auteurs comme Huber, Hampel et Maronna (Huber, 1964 ; 1972 ; Maronna, 1976) et étendus plus récemment dans le cas complexe (Olilla *et al.*, 2012 ; Mahot *et al.*, 2013 ; Pascal, Chitour *et al.*, 2008). Il est important de remarquer que souvent ces techniques d'estimation requièrent la résolution d'une équation de point fixe faisant intervenir l'inverse le matrice estimée. Une des conséquences est la nécessité d'avoir plus d'observations n que la taille des observations N pour estimer la matrice. C'est principalement pour cette raison que les techniques de régularisation ont été introduites afin d'obtenir l'existence de la matrice estimée dans les cas où peu d'observations sont disponibles. L'étude du RTE d'un point de vue algorithmique a été traitée dans différents articles (Chitour, Pascal, 2008 ; Pascal *et al.*, 2014 ; Ollila, Tyler, 2014 ; Sun *et al.*, 2014). Contrairement à la RSCM, le RTE est par construction un estimateur robuste puisqu'il est issu de l'estimateur de Tyler aussi appelé estimateur du point fixe (Tyler, 1987 ; Pascal, Chitour *et al.*, 2008). Cet estimateur, le RTE, est donc adapté au cas d'un faible nombre de données pour l'estimation de la matrice de covariance mais aussi dans les cas de données aberrantes ou hétérogènes. Ce type d'observations a été mis en évidence dans de nombreuses analyses expérimentales du fouillis des données radar (K. D. Ward, 1981 ; Watts, 1985 ; Nohara, Haykin, 1991 ; Billingsley, 1993).

Le principal challenge dans l'utilisation de ces méthodes régularisées consiste à trouver une méthode pour optimiser le paramètre de régularisation. Evidemment, cette optimisation va dépendre du critère à maximiser (ou de manière équivalente à minimiser). Ces problèmes ont été essentiellement investigués dans (Ledoit, Wolf, 2004 ; Chen *et al.*, 2010) pour la RSCM et dans (Ollila, Tyler, 2014 ; Chen *et al.*, 2011) pour le RTE. Bien qu'obtenant des expressions différentes pour le paramètre de régularisation, ces méthodes reposent toute sur la minimisation d'une distance entre le RTE ou la RSCM et la vraie matrice de covariance. Ce sont des méthodes oracles, qui ne permettent pas en pratique d'être implémentées, puisque la vraie matrice de covariance est inconnue. De plus, ce critère d'optimisation n'est pas forcément adapté pour l'optimisation d'autres critères, tels que les performances de détection. Dans cet article, nous nous intéressons à l'optimisation, et en conséquence au design optimal, d'un détecteur robuste, l'adaptive normalized matched filter (ANMF). Introduit par (Conte *et al.*, 1995) et analysé dans (Liu *et al.*, 2011 ; Kraut, Scharf, 1999 ; Pascal *et al.*, 2004), cette méthode présente l'avantage de conserver un taux de fausse alarme constant (propriété TFAC, ou Constant False Alarm Rate, CFAR) par rapport à la puissance du bruit et à sa matrice de covariance (Pascal *et al.*, 2006). Ce détecteur est obtenu en remplaçant dans le filtre adapté normalisé (normalized matched filter, NMF) la vraie matrice de covariance par un estimateur, qui est calculé à l'aide de n données secondaires, supposées indépendantes et identiquement distribuées (IID). Dans cet article, les estimateurs utilisés dans le ANMF sont le RTE et la RSCM. Ainsi, dans un premier temps, nous considérons le cas classique de bruit additif gaussien et nous étudions les performances obtenues avec le RSCM. Ce cas sera appelé le ANMF-RSCM. La principale contribution de cette partie consiste à l'analyse du paramètre de régularisation afin d'optimiser les performances de détection ainsi que les performances de régulation de fausse alarme. Cette analyse doit être basée sur l'étude statistique

de l'estimateur RSCM. Lorsque les dimensions n et N sont fixées, il semble difficile de résoudre ce problème d'optimisation. Dans cet article, nous considérons donc le régime asymptotique dans lequel les deux dimensions tendent vers l'infini dans un même régime, i.e. $N/n \to c \in (0, 1]$. Ceci permet en particulier d'utiliser certains outils classiques provenant de la théorie des grandes matrices aléatoires. Ainsi, le premier objectif de ces travaux est de trouver le paramètre de régularisation optimal qui maximise la probabilité de détection du détecteur résultant pour toute probabilité de fausse alarme.

Dans un second temps, nous considérons que le bruit n'est plus distribué selon une distribution gaussienne classique mais peut être distribué selon des distributions à queue lourde. C'est le contexte des distributions CES. Dans ce cas, il est naturel de considérer des estimateurs de matrice de covariance robuste, i.e. dans cet article le RTE puisque le RSCM n'est pas adapté aux distributions impulsives et/ou à la présence d'outliers. Le détecteur résultant sera appelé le ANMF-RTE. A l'aide de l'étude du comportement asymptotique du RTE, nous montrerons notamment que le comportement asymptotique du ANMF-RTE est équivalent à celui du ANMF-RSCM. Ainsi, l'utilisation conjointe de la théorie des grandes matrices aléatoires et de la théorie de l'estimation robuste permet d'étendre les performances obtenues dans le cas gaussien à des scénarios plus complexes, faisant intervenir des distributions impulsives ou des outliers dans les observations.

Cet article est organisé de la façon suivante. La première partie présente une introduction du problème considéré. Ensuite, une seconde partie contient les principales contributions théoriques, notamment le design optimal du détecteur ANMF construit avec les différents estimateurs régularisés de matrice de covariance. Enfin, ces résultats seront illustrés au travers de simulations Monte-Carlo qui compareront la méthode proposée pour le design optimal du paramètre de régularisation avec les méthodes existantes dans la littérature.

Notations : Les vecteurs sont représentés en caractères gras minuscules et les matrices en lettres capitales grasses. La notation $(.)^*$ représente l'opérateur transposé-conjugué tandis que $\mathrm{tr}(.)$ et $(.)^{-1}$ sont respectivement les opérateurs trace et inverse. La notation $\|.\|$ sera utilisée pour la norme euclidienne dans le cas de vecteurs et la norme spectrale dans le cas de matrice. $\xrightarrow{\text{p.s.}}$ représente la convergence presque sure. La relation $X \triangleq Y$ permet de définir la nouvelle variable X comme étant égale à Y.

2. Etat de l'art et problèmes étudiés

En détection radar, le principal problème consiste généralement à détecter, dans un vecteur d'observation $\mathbf{y}$ de dimension N un vecteur $\alpha\,\mathbf{p}$, d'amplitude complexe α généralement inconnue et de *steering vector* $\mathbf{p}$ caractérisant les paramètres $\boldsymbol{\theta}$ inconnus de la cible (position, vitesse, réponse polarimétrique, direction, etc.). Ce signal d'intérêt est généralement corrompu par un bruit additif de fouillis $\mathbf{x}$ caractérisé par un vecteur de paramètres inconnus (moyenne, variance, covariance, paramètres de forme, etc.). Dans l'exemple du traitement adaptatif spatio-temporel, la dimension

des observations correspond à la concaténation des M impulsions émises par m antennes, $N = M\,m$, et peut être relativement élevée. Ainsi, le problème de détection se formalise généralement par un test d'hypothèses binaires :

$$\begin{cases} H_1 : & \mathbf{y} = \alpha\mathbf{p} + \mathbf{x} \\ H_0 : & \mathbf{y} = \mathbf{x}. \end{cases} \tag{1}$$

Sous les deux hypothèses H_0 et H_1, on suppose que n observations indépendantes et identiquement distribuées (i.i.d.) sont disponibles. Ces observations servent à estimer les paramètres inconnus du bruit. Dans le domaine du radar, ces observations sont appelées les données secondaires.

Plusieurs modèles statistiques ont été proposés pour le fouillis $\mathbf{x}$. Parmi eux, la famille des distributions complexes elliptiques symétriques (Complex Elliptically Symmetric, CES) s'est imposée comme une modélisation naturelle. Ceci peut s'expliquer par l'interprétation physique de ce modèle. En effet, cette classe de distribution peut être interprétée comme des distributions gaussiennes de variance inconnue et aléatoire. Ainsi, cette variance aléatoire peut correspondre à la variation temporelle (ou spatiale) des observations. Cela permet d'introduire une hétérogénéité. De nombreuses distributions multivariées appartiennent à cette familles de CES. On peut citer par exemple les distributions gaussiennes multivariées, les distributions gaussiennes généralisées, la K-distribution, les distributions de Student-t multivariées parmi de nombreuses autres. Une analyse complète de ces modèles peut être trouvée dans (Olilla *et al.*, 2012).

Formellement, un vecteur elliptiquement distribué est donné par :

$$\mathbf{x} = \sqrt{\tau}\mathbf{C}_N^{\frac{1}{2}}\tilde{\mathbf{w}}$$

où τ est une variable (scalaire) aléatoire positive appelée la *texture*, $\mathbf{C}_N$ est la matrice de covariance de $\mathbf{x}$ [5] et $\tilde{\mathbf{w}}$ est un vecteur aléatoire de dimension N, indépendant de τ, centré et distribué uniformément sur l'hypersphère de dimension N. $\tilde{\mathbf{w}}$ est donc unitairement invariant de norme $\|\tilde{\mathbf{w}}\| = \sqrt{N}$. $\mathbf{C}_N^{\frac{1}{2}}\tilde{\mathbf{w}}$ est appelé le *speckle*. Ceci correspond, par opposition à la texture, au fond statistique dans les problèmes de traitement d'image.

La construction d'un test statistique approprié au problème 1 va dépendre de la quantité d'information disponible, notamment les paramètres du bruit additif. Ainsi, si le bruit est gaussien de matrice de covariance $a\,\mathbf{C}_N$ avec $\mathbf{C}_N$ connue et a et $alpha$ inconnus, alors le test du rapport de vraisemblance généralisé (Generalized Likelihood Ratio, GLRT) conduit à la statistique de test suivante :

$$T_N = \frac{\left|\mathbf{y}^*\mathbf{C}_N^{-1}\mathbf{p}\right|}{\sqrt{\mathbf{y}^*\mathbf{C}_N^{-1}\mathbf{y}\mathbf{p}^*\mathbf{C}_N^{-1}\mathbf{p}}}$$

5. Il est important de noter que cette matrice de covariance n'existe pas toujours (uniquement dans le cas où $\mathbf{x}$ possède un moment d'ordre 2 fini. Cependant, cette expression existe toujours, avec $\mathbf{C}_N$ qui correspond à la matrice de dispersion (dans le cas d'un moment d'ordre 2 fini, la matrice de covariance est égale à la matrice de dispersion à un facteur près.

Ceci correspond à la racine carrée de la statistique du test du détecteur appelé filtre adapté normalizé (Normalized Matched Filter) qui a été trouvé de différentes façons par plusieurs auteurs (Picinbono, Vezzosi, 1970 ; Scharf, Lytle, 1971), puis par (Scharf, Friedlander, 1994 ; Conte *et al.*, 2002 ; Gini, Greco, 2002). Il est intéressant de remarquer que sous H_0, cette statistique ne dépend pas de la texture, ce qui lui confère la propriété dite de “taux de fausse alarme constant" (TFAC) par rapport à la texture (texture Constant False Alarm Rate, texture-CFAR). Ainsi, T_N ne dépend que de la matrice de covariance $\mathbf{C}_N$ à un facteur près. Comme évoqué précédemment, cette statistique existe toujours même dans le cas d'un moment d'ordre 2 infini puisqu'elle peut s'écrire comme fonction de la matrice de dispersion uniquement.

En pratique, la matrice de covariance $\mathbf{C}_N$ est inconnue et doit être estimée. L'approche consiste à remplacer la matrice de covariance dans le test par un estimateur, construit à partir des données secondaires supposées i.i.d. $\mathbf{x}_1, \cdots, \mathbf{x}_n$. Ceci conduit à une version dite adaptive du détecteur précédent. Ce détecteur sera appelé ANMF (Adaptive Normalized Matched Filter) dans la suite de cet article. Les principaux challenges consitent maintenant à trouver les estimateurs de matrice de covariance qui conduiront aux meilleurs performances du test de détection. En pratique, les critères d'évaluation des performances considérés sont la probabilité de fausse alarme, la probabilité de détection, la robustesse du détecteur (par rapport à une erreur de modèle, un ou plusieurs outliers, des données manquantes...).

Plusieurs estimateurs de $\mathbf{C}_N$ peuvent être utilisés. Le plus naturel, et un des plus étudiés dans la littérature est la matrice de covariance empirique (SCM) donnée par :

$$\widehat{\mathbf{R}}_N = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^* .$$

Cet estimateur est l'estimateur du Maximum de Vraisemblance (Maximum Likelihood Estimator, MLE) dans le cas d'un bruit additif gaussien, et aussi l'estimateur obtenu par la méthode des moments. Cependant, en pratique, comme dans le cas du STAP, la dimension N des observations peut être grande, i.e. du même ordre de grandeur ou plus grande que le nombre n d'observations, ce qui conduit à un estimateur mal conditionné, voire non-inversible, qui sera moins précis et entraînera des dégradations dans les performances de détection [6]. Ainsi, une alternative a été de considérer une version régularisée de la SCM, notamment la RSCM définie par :

$$\widehat{\mathbf{R}}_N(\rho) = (1-\rho)\widehat{\mathbf{R}}_N + \rho \mathbf{I}_N , \tag{2}$$

où le paramètre $\rho \in [0, 1]$, appelé paramètre de régularisation, est utilisé pour donner plus ou moins d'importance aux échantillons, selon le nombre d'observations. Dans la suite, l'ANMF dans lequel le RSCM est utilisé sera appelé le ANMF-RSCM.

6. Traditionnellement, on suppose que $2N$ observations sont nécessaires afin d'obtenir des performances raisonnables en termes de détection, i.e. une perte de 3 dB sur le rapport signal-à-bruit (RSB ou Signal-to-Noise Ratio, SNR) de sortie comparé au filtre optimal (Reed *et al.*, 1974).

D'autre part, bien que cette régularisation soit intéressante dans le cas où la dimension est grande, il reste néanmoins très dépendant des performances générées par la SCM. Ainsi, en termes de robustesse, cet estimateur ne sera pas adapté au cas où les observations suivent des distributions à queue lourde, ou encore dans le cas où des outliers sont présents dans les observations. Pour palier ces problèmes, des estimateurs issus de la théorie de l'estimation robuste, les M-estimateurs on été introduits dans les années 70's avec les travaux de Huber et Maronna (Huber, 1964 ; 1972 ; Maronna, 1976), puis plus récemment l'estimateur de Tyler ou estimateur du point fixe (Tyler, 1987 ; Pascal, Chitour *et al.*, 2008). L'estimateur de Tyler peut être vu comme l'estimateur limite des M-estimateurs en ce sens qu'il est le plus robuste par rapport aux critères de robustesse classiques tels que le point de rupture, la fonction d'influence ou encore le biais asymptotique. Il est de plus indépendant de la distribution utilisée dès lors que cette dernière appartient à la famille des CES. Ces propriétés le rendent particulièrement intéressant. De plus, tous ces estimateurs ont été récemment analysés dans le régime des grandes dimensions (Couillet, Pascal, Silverstein, 2014a ; 2014b ; Zhang *et al.*, 2014). Dans ces travaux, afin de répondre au problème de grande dimension et simultanément de robustesse, nous nous sommes intéressé à l'estimateur défini dans (Pascal *et al.*, 2014) comme l'unique solution $\hat{\mathbf{C}}_N(\rho)$ de l'équation de point fixe suivante :

$$\hat{\mathbf{C}}_N(\rho) = (1-\rho)\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{x}_i\mathbf{x}_i^*}{\frac{1}{N}\mathbf{x}_i^*\hat{\mathbf{C}}_N^{-1}(\rho)\mathbf{x}_i} + \rho\mathbf{I}_N. \tag{3}$$

avec $\rho \in \left(\max\left(0, 1-\frac{n}{N}\right), 1\right]$.

Cet estimateur sera appelé l'estimateur de Tyler régularisé (Regularized-Tyler Estimator, RTE). Comme indiqué précédemment, cet estimateur possède de très bonnes propriétés en termes de robustesse, mais il présente d'autres aspects très intéressants. Tout d'abord, il est adapté aux situations où $c_N \triangleq \frac{N}{n}$ est grand tandis que les estimateurs robustes classiques de matrice de covariance seront mal conditionnés ou bien pas du tout définis si $N > n$. De plus, le paramètre de régularisation ρ offre un degré de liberté supplémentaire : cet estimateur varie due l'estimateur sans biais de Tyler lorsque $\rho = 0$ (Pascal, Forster *et al.*, 2008) jusqu'à la matrice identité lorsque $\rho = 1$. Le détecteur résultant sera appelé le ANMF-RTE.

Une fois ce détecteur adaptatif construit, i.e. lorsque la vraie matrice de covariance a été remplacée par un estimateur le principal problème est de caractériser les performances de ce détecteur, dans ce cas l'ANMF-RTE. La robustesse de ce détecteur est évidemment assurée par construction du détecteur et de l'estimateur de matrice de covariance. En ce qui concerne les performances en détection, elles reposent sur la distribution de la statistique de test. A distance finie, i.e. lorsque N et n sont fixés, ceci paraît impossible puisque aucune expression analytique du RTE n'est disponible. Ce comportement statistique est en effet indispensable pour régler l'erreur de première espèce ou probabilité de fausse alarme. Ainsi, un premier problème consiste à analyser le détecteur en fonction de ρ afin de trouver la valeur de ρ qui maximiserait les performances du ANMF-RTE. De récents travaux se sont focalisés sur l'optimisation de différents critères par rapport à ρ, comme l'erreur quadratique moyenne (Mean

Square Error, MSE) dans (Ollila, Tyler, 2014 ; Y. I. Abramovich, Besson, 2013). Bien que les résultats soient attractifs dans la façon de calculer le ρ optimal, il semble que cette approche ne peut être optimale, vu que l'application étudiée n'est pas prise en compte. Ainsi, dans cet article, on s'intéresse aux valeurs de ρ qui, soit minimisent la probabilité de fausse alarme, soit maximisent la probabilité de détection (pour une probabilité de fausse alarme fixée). Ce deuxième cas semble être le plus pertinent pour un problème de détection.

Afin d'obtenir ces résultats, il est nécessaire de caractériser la distribution de $\widehat{T}_N^{\mathrm{RSCM}}(\rho)$ et de $\widehat{T}_N^{\mathrm{RTE}}(\rho)$ donnés par :

$$\widehat{T}_N^{\mathrm{RSCM}}(\rho) = \frac{\left|\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{p}\right|}{\sqrt{\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{y}}\sqrt{\mathbf{p}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{p}}} \tag{4}$$

$$\widehat{T}_N^{\mathrm{RTE}}(\rho) = \frac{\left|\mathbf{y}^*\widehat{\mathbf{C}}_N^{-1}(\rho)\mathbf{p}\right|}{\sqrt{\mathbf{y}^*\widehat{\mathbf{C}}_N^{-1}(\rho)\mathbf{y}}\sqrt{\mathbf{p}^*\widehat{\mathbf{C}}_N^{-1}(\rho)\mathbf{p}}} \tag{5}$$

sous les hypothèses H_0 et H_1. L'intérêt de considérer ces deux statistiques de test est de pouvoir traiter les cas classiques (e.g. observations gaussiennes homogènes) mais aussi les cas non-standards (données hétérogènes, impulsives, données manquantes, outliers...).

Pour traiter ce problème, nous considérons dans cet article le cas asymptotique dans lequel N et n tendent vers l'infini à la même vitesse, i.e. avec $\frac{N}{n} \to c \in (0, \infty)$. Ceci a permis d'obtenir les résultats dans le cas de la minimisation de la probabilité de fausse alarme (Couillet, Kammoun, Pascal, 2014). La partie suivante présente les résultats dans le cas gaussien.

3. Design optimal du ANMF-RSCM dans le cas d'un bruit gaussien

Dans cette partie, les données secondaires $\mathbf{x}_1, \cdots, \mathbf{x}_n$ sont supposées i.i.d. gaussiennes centrées et de matrice de covariance $\mathbf{C}_N$.

Pour $\rho \in (0, 1]$, la RSCM est définie par l'équation (2) et la statistique de test correspondante sera notée $\widehat{T}_N^{RSCM}$. Afin d'optimiser les performances de ce détecteur, par rapport au paramètre ρ, il est indispensable de caractériser, asymptotiquement, les probabilités de de fausse alarme et de détection, sous les hypothèses $c_N \triangleq \frac{N}{n} \to c$, lorsque N et n tendent vers l'infini. Plus précisément, il s'agit d'évaluer les probabilités $\mathbb{P}\left[\widehat{T}_N^{\mathrm{RSCM}} > \Gamma | H_0\right]$ et $\mathbb{P}\left[\widehat{T}_N^{\mathrm{RSCM}} > \Gamma | H_1\right]$ pour un certain seuil de détection $\Gamma > 0$, selon l'hypothèse, i.e. H_0 ou H_1, ou de manière équivalente $\mathbf{y} = \mathbf{x}$ or $\mathbf{y} = \alpha\mathbf{p} + \mathbf{x}$.

Avant de calculer les performances asymptotiques, il est intéressant de remarquer que l'analyse en régime de grande dimension entraîne des résultats différents

de l'approche classique de statistique inférentielle. En effet, selon les ordres de grandeur de α et Γ, par rapport à N, des résultats triviaux peuvent être obtenus. C'est le cas par exemple sous H_0, puisque les quantités $\frac{1}{\sqrt{N}}\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\frac{\mathbf{p}}{\|\mathbf{p}\|}$, $\frac{1}{N}\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{y}$ et $\mathbf{p}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\frac{\mathbf{p}}{\|\mathbf{p}\|^2}$ sont des objets classiques dans la théorie des matrices aléatoires qui convergent vers leur espérance lorsque N et n tendent vers l'infini à la même vitesse (Wagner *et al.*, 2012). Ainsi, comme $\frac{1}{\sqrt{N}}\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{p} \xrightarrow{\text{p.s.}} 0$, $\widehat{T}_N^{\text{RSCM}} \xrightarrow{\text{p.s.}} 0$ pour tout $\Gamma > 0$, ce qui a comme conséquence, que dans le régime asymptotique en grande dimension, la probabilité de fausse alarme tend vers zéro pour toute valeur du seuil de détection Γ. Ainsi, le seuil de détection doit être normalisé de la façon suivante : $\Gamma = N^{-\frac{1}{2}}r$ pour une valeur de r fixée. Cette hypothèse a aussi été utilisée dans (Couillet, Kammoun, Pascal, 2014). De façon similaire, sous l'hypothèse H_1, la présence d'un signal d'intérêt entraîne la convergence presque sûre de $\widehat{T}_N^{\text{RSCM}}$ vers une constante positive si α ne varie pas avec N. Ainsi, pour $\Gamma = N^{-\frac{1}{2}}r$, $\mathbb{P}\left[\widehat{T}_N^{\text{RSCM}} > \Gamma | H_1\right] \to 1$, ce qui conduit comme précédemment à un cas trivial. Pour éviter ce problème, définissons $\alpha = N^{-\frac{1}{2}}a$ pour $a > 0$ et fixé avec $\|\mathbf{p}\| = N$. En pratique, ceci signifie que l'exploitation de la dimension des observations, lorsqu'elle est suffisamment grande, permet d'obtenir de (bons) résultats aussi dans des régimes de faibles SNR.

Afin de dérivées les propriétés asymptotiques de ces probabilités, plusieurs hypothèses sont nécessaires :

SUPPOSITION 1. — Pour $i \in \{1, \cdots, n\}$, $\mathbf{x}_i = \mathbf{C}_N^{\frac{1}{2}}\mathbf{w}_i$ avec

– $\mathbf{w}_1, \cdots, \mathbf{w}_n$ sont des vecteurs aléatoires indépendants centrés de taille N et de matrice de covariance $\mathbf{I}_N$,

– $\mathbf{C}_N \in \mathbb{C}^{N\times N}$ est tel quelim sup $\|\mathbf{C}_N\| < \infty$ et $\frac{1}{N}\operatorname{tr}\mathbf{C}_N = 1$,

– $\liminf_N \frac{1}{N}\mathbf{p}^*\mathbf{C}_N\mathbf{p} > 0$.

□

Il est intéressant de noter que $\frac{1}{N}\operatorname{tr}\mathbf{C}_N = 1$ n'est pas une contrainte restrictive puisque les statistiques étudiées dans cet article sont invariante à un changement d'échelle sur $\mathbf{C}_N$. De plus, le dernier point des hypothèses 1 est nécessaire pour assurer les fluctuations des statistiques sous H_0 et H_1.

Pour $z \in \mathbb{C}\backslash\mathbb{R}_+$, soit $m_N(z)$ l'unique solution complexe de :

$$m_N(z) = (-z + c_N(1-\rho) \times \frac{1}{N}\operatorname{tr}\mathbf{C}_N\left(\mathbf{I}_N + (1-\rho)m_N(z)\mathbf{C}_N\right)^{-1}\Big)^{-1}$$

qui satisfasse $\Im(z)\Im(m_N(z)) \geq 0$ ou l'unique positive si $z < 0$. L'existence et l'unicité sont directement obtenus à partir des résultats de (Silverstein, Bai, 1995). Cette quantité est déterministe et peut être calculée simplement pour chaque z à l'aide de l'algorithme de point fixe. Ceci permet notamment de caractériser le comportement

asymptotique de la mesure spectrale empirique de la matrice $(1-\rho)\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i^*$. Définissons aussi pour $\kappa > 0$, $\mathcal{R}_\kappa^{\mathrm{SCM}}$ comme:

$$\mathcal{R}_\kappa^{\mathrm{SCM}} \triangleq [\kappa, 1].$$

Après ces définitions préliminaires, nous pouvons dériver le comportement asymptotique des probabilités de fausse alarme et de détection.

THÉORÈME 2 (Probabilité de fausse alarme). — *Lorsque* $N, n \to \infty$ *avec* $c_N \to c \in (0, \infty)$,

$$\sup_{\rho \in \mathcal{R}_\kappa^{\mathrm{SCM}}} \left| \mathbb{P}\left[\widehat{T}_N^{\mathrm{RSCM}}(\rho) > \frac{r}{\sqrt{N}} | H_0\right] - e^{-\frac{r^2}{2\sigma^2_{N,\mathrm{SCM}}(\rho)}} \right| \xrightarrow{\text{p.s.}} 0$$

où

$$\sigma^2_{N,\mathrm{SCM}}(\rho) \triangleq \frac{1}{2}\frac{\mathbf{p}^*\mathbf{C}_N\mathbf{Q}_N^2(\rho)\mathbf{p}}{\mathbf{p}^*\mathbf{Q}_N(\rho)\mathbf{p}\frac{1}{N}\operatorname{tr}\mathbf{C}_N\mathbf{Q}_N(\rho)} \times \frac{1}{1 - c(1-\rho)^2 m_N(-\rho)^2 \frac{1}{N}\operatorname{tr}\mathbf{C}_N^2\mathbf{Q}_N^2(\rho)}$$

et $\mathbf{Q}_N(\rho) \triangleq (\mathbf{I}_N + (1-\rho)m_N(-\rho)\mathbf{C}_N)^{-1}$.

Quelques remarques sur ce théorème sont essentielles:

REMARQUE 3. —

– Dans le régime de grande dimension, $\widehat{T}_N^{\mathrm{RSCM}}(\rho)$ suit asymptotiquement une distribution de Rayleigh. C'est un résultat très simple.

– L'uniformité dur ρ obtenue dans le théorème 2 est essentielle pour la suite, notamment pour obtenir le ρ optimal qui minimise la probabilité de fausse alarme.

– Cette uniformité dans la convergence implique évidemment la convergence pour toutes les valeurs de $\rho > 0$ mais il permet aussi de gérer la convergence de la probabilité de fausse alarme lorsque les valeurs du paramètre de régularisation sont aléatoires. Ceci est intéressant puisque le détecteur est réglé à partir d'observations aléatoires qui permettent d'estimer les paramètres.

– Notons aussi que l'ensemble $[0, \kappa)$ pour $\kappa > 0$ aussi petit que l'on souhaite est retiré de la région de convergence uniforme pour des raisons techniques apparaissant dans les preuves.

□

Le théorème 2 fournit une expression analytique de la relation entre la probabilité de fausse alarme, généralement fixée par l'utilisateur selon les applications considérées, et le seuil de détection. Cette relation est évidemment fonction du paramètre de régularisation. Ainsi, l'optimisation, i.e. la minimisation de la probabilité de fausse alarme pourra être effectuée grâce à ce paramètre. Notons qu'il existe une infinité de valeurs du seuil de détection pour une probabilité de fausse alarme donnée, ceci grâce au paramètre de régularisation qui joue le rôle d'un degré de liberté supplémentaire.

Cependant, dans le théorème 2, cette expression analytique dépend de la matrice de covariance qui est inconnue en pratique. Il est donc essentiel d'avoir un estimateur consistant de cette quantité. C'est l'objet de la proposition suivante:

PROPOSITION 4. — *Pour $\rho \in (0,1)$, soit*

$$\hat{\sigma}^2_{N,\mathrm{SCM}}(\rho) = \frac{1}{2} \frac{1 - \rho \frac{\mathbf{p}^* \widehat{\mathbf{R}}_N^{-2}(\rho)\mathbf{p}}{\mathbf{p}^* \widehat{\mathbf{R}}_N^{-1}\mathbf{p}}}{\left(1 - c_N + \frac{c_N \rho}{N} \operatorname{tr} \widehat{\mathbf{R}}_N^{-1}(\rho)\right)\left(1 - \frac{\rho}{N} \operatorname{tr} \widehat{\mathbf{R}}_N^{-1}(\rho)\right)}$$

et $\hat{\sigma}^2_{N,SCM}(1) = \lim_{\rho \uparrow 1} \hat{\sigma}^2_{N,SCM}(\rho) = \frac{\mathbf{p}^* \widehat{\mathbf{R}}_N \mathbf{p}}{\operatorname{tr} \widehat{\mathbf{R}}_N}$. Alors, pour tout $\kappa > 0$, on a le résultat suivant

$$\sup_{\rho \in \mathcal{R}_\kappa^{\mathrm{SCM}}} \left| \hat{\sigma}^2_{N,\mathrm{SCM}}(\rho) - \sigma^2_{N,\mathrm{SCM}}(\rho) \right| \xrightarrow{\text{p.s.}} 0.$$

La preuve de la Proposition 4 est similaire à celle de la Proposition 1 dans (Couillet, Kammoun, Pascal, 2014).

Maintenant que l'erreur de première espèce est contrôlée, nous pouvons nous intéresser au comportement de la probabilité de détection en fonction du paramètre de régularisation, afin d'optimiser les performances de détection. Il est important de noter que c'est l'utilisation conjointe des performances calculées sous H_0 et celle sous H_1 qui permettront le design optimal du détecteur. Autrement dit, nous allons dériver l'équivalent asymptotique de $\mathbb{P}\left[\widehat{T}_N^{\mathrm{RSCM}}(\rho) > \frac{r}{\sqrt{N}} | H_1\right]$, sous l'hypothèse H_1, i.e. lorsque le signal reçu est de la forme::

$$H_1 : \quad \mathbf{y} = \frac{a}{\sqrt{N}} \mathbf{p} + \mathbf{x}$$

où $\mathbf{x}$ suit la même distribution que les $\mathbf{x}_i$'s dans l'hypothèse 1.

Le théorème suivant donne le comportement sous H_1:

THÉORÈME 5 (Detection probability). — *Lorsque $N, n \to \infty$ avec $c_N \to c$, on a pour tout $\kappa > 0$*

$$\sup_{\rho \in \mathcal{R}_\kappa^{\mathrm{SCM}}} \left| \mathbb{P}\left[\widehat{T}_N^{\mathrm{RSCM}}(\rho) > \frac{r}{\sqrt{N}} | H_1\right] - Q_1\left(g_{\mathrm{SCM}}(\mathbf{p}), \frac{r}{\sigma_{N,\mathrm{SCM}}(\rho)}\right) \right| \xrightarrow{\text{p.s.}} 0.$$

où Q_1 est la Marcum Q-fonction[7]*, où $\sigma_{N,\mathrm{SCM}}$ est donné dans le théorème Theorem 2 et où $g(\mathbf{p})$ est donné par:*

$$g_{SCM}(\mathbf{p}) = \frac{\sqrt{1 - c(1-\rho)^2 m(-\rho)^2 \frac{1}{N} \operatorname{tr} \mathbf{C}_N^2 \mathbf{Q}_N^2(\rho)}}{\sqrt{\mathbf{p}^* \mathbf{C}_N \mathbf{Q}_N^2(\rho) \mathbf{p}}} \times \sqrt{\frac{2}{N}} a |\mathbf{p}^* \mathbf{Q}_N(\rho) \mathbf{p}|.$$

PREUVE 6. — La preuve est fournit dans (Kammoun *et al.*, 2015. arXiv:1501.06027). ■

Les théorèmes 2 et 5 montrent que $\widehat{T}_N^{\mathrm{RSCM}}(\rho)$ se comporte différemment selon les hypothèses H_0 et H_1, i.e. si un signal d'intérêt est présent ou non. En particulier, sous H_0, $\widehat{T}_N^{\mathrm{RSCM}}(\rho)$ suit une loi de Rayleigh de paramètre $\sigma_{N,\mathrm{SCM}}(\rho)$ tandis que sous H_1, cette statistique suit une loi de Rice de paramètres g_{SCM} et $\sigma_{N,\mathrm{SCM}}(\rho)$.

Ces comportements asymptotiques sont assez intéressants puisque les distributions obtenues sont simples et bien connues, ce qui n'est pas le cas dans le régime de convergence classique, i.e. lorsque N est fixe et que seul $n \to \infty$. Ainsi, dans des cas plus simples, i.e. pour des expressions de détecteurs et/ou d'estimateurs plus simples, les distributions obtenues peuvent être très compliquées, voire sans expression analytique (voir par exemple (Pascal, Ovarlez, 2015 ; Kraut *et al.*, 2001).

Le problème maintenant consiste à traiter le problème d'optimisation par rapport au paramètre de régularisation ρ, qui est naturellement lié au seuil de détection r. Selon la théorie de la détection radar, l'objectif est de régler les paramètres ρ et r afin d'assurer une probabilité de fausse alarme asymptotique fixée, égale à η afin ensuite de maximiser la probabilité de détection. D'après le théorème 2, pour obtenir une probabilité de fausse alarme égale à η, les paramètres ρ et r doivent vérifier :

$$\frac{r}{\sigma_{N,\mathrm{SCM}}(\rho)} = \sqrt{-2 \log \eta}.$$

Il est important de noter que, grâce à ce degré de liberté supplémentaire qu'est le paramètre de régularisation, ils existe une infinité de seuil de détection r qui permettent d'obtenir une probabilité de fausse alarme donnée. Ainsi, il est possible grâce au théorème 5, de maximiser la probabilité de détection asymptotique en trouvant parmi les différentes valeurs de ρ, l'argument maximum qui vérifie

$$Q_1\left(g_{\mathrm{SCM}}(\mathbf{p}), \frac{r}{\sigma_{N,SCM}(\rho)}\right).$$

7. $Q_1(a,b) = \int_b^{+\infty} x \exp\left(-\frac{x^2+a^2}{2}\right) I_0(ax)dx$ avec I_0 la fonction modifiée de Bessel de première espèce à l'ordre 0.

Le deuxième argument de Q_1 doit être fixe afin d'assurer la probabilité de fausse alarme désirée. Ensuite, puisque la Marcum Q-fonction est croissante par rapport au premier argument, l'optimisation de la probabilité de détection se traduit de la façon suivante :

$$\rho \in \operatorname{argmax} \{f_{\mathrm{SCM}}(\rho)\}$$

où

$$f_{\mathrm{SCM}}(\rho) \triangleq \frac{1}{2a^2} g^2_{\mathrm{SCM}}(\mathbf{p})$$

Cependant, comme dans le cas de $\sigma^2_{N,\mathrm{SCM}}(\rho)$, l'optimisation de $f_{\mathrm{SCM}}(\rho)$ n'est pas possible en pratique puisque l'expression de $f_{SCM}(\rho)$ dépend de paramètres inconnus, notamment la matrice de covariance $\mathbf{C}_N$. Il est donc indispensable d'avoir un estimateur consistant de $f_{\mathrm{SCM}}(\rho)$, basé sur $\widehat{\mathbf{R}}_N$. C'est l'objet de la proposition suivante :

PROPOSITION 7. — *Pour $\rho \in (0,1)$, soit $\hat{f}_{\mathrm{SCM}}(\rho)$ tel que*

$$\hat{f}_{\mathrm{SCM}}(\rho) = \frac{\left(\mathbf{p}^* \widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{p}\right)^2 (1-\rho)\left(1 - c + \frac{c}{N}\rho \operatorname{tr} \widehat{\mathbf{R}}_N^{-1}(\rho)\right)^2}{\mathbf{p}^* \widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{p} - \rho \mathbf{p}^* \widehat{\mathbf{R}}_N^{-2}(\rho)\mathbf{p}}$$

et soit $\hat{f}_{\mathrm{SCM}}(1) \triangleq \lim_{\rho \uparrow 1} \hat{f}_{\mathrm{SCM}}(\rho) = \frac{N}{\mathbf{p}^* \widehat{\mathbf{R}}_N \mathbf{p}}$. Alors, on a la propriété suivante

$$\sup_{\rho \in \mathcal{R}_\kappa^{\mathrm{SCM}}} \left| \hat{f}_{\mathrm{SCM}}(\rho) - f_{\mathrm{SCM}}(\rho) \right| \xrightarrow{\text{p.s.}} 0.$$

PREUVE 8. — La preuve est fournit dans (Kammoun *et al.*, 2015. arXiv:1501.06027). ■

Maintenant, comme les résultats de la proposition 7 et du théorème 5 sont uniforme sur ρ, nous avons le corollaire suivant :

COROLLAIRE 9. — *Soit $\hat{f}_{SCM}(\rho)$ défini dans la proposition 7. Soit $\hat{\rho}_N^*$ n'importe quelle valeur satisfaisant*

$$\hat{\rho}_N^* \in \operatorname*{argmax}_{\rho \in \mathcal{R}_\kappa^{SCM}} \left\{ \hat{f}_{\mathrm{SCM}}(\rho) \right\}.$$

Alors, pour chaque $r > 0$,

$$\begin{aligned} &\mathbb{P}\left(\sqrt{N} T_N(\hat{\rho}_N^*) > r | H_1\right) \\ &- \max_{\rho \in \mathcal{R}_\kappa} \left\{ \mathbb{P}\left(\sqrt{N} T_N(\rho) > r | H_1\right) \right\} \xrightarrow{\text{p.s.}} 0. \end{aligned}$$

PREUVE 10. — La preuve est similaire à celle dans (Couillet, Kammoun, Pascal, 2014). ■

Le corollaire 9 permet maintenant de proposer une façon de construire le détecteur optimal (au sens de la maximisation de sa probabilité de détection), de la façon suivante :

– Tout d'abord, choisissons une des valeurs du paramètre de régularisation qui maximise $\hat{f}_{\mathrm{SCM}}(\rho)$:

$$\hat{\rho}_N^* \in \underset{\rho \in \mathcal{R}_\kappa}{\operatorname{argmax}} \left\{ \hat{f}_{\mathrm{SCM}}(\rho) \right\} \tag{6}$$

– Ensuite, pour assurer une probabilité de fausse alarme asymptotique fixée et égale à η, prenons le seuil de détection $\hat{r}$, tel que :

$$\hat{r} = \hat{\sigma}_{N,\mathrm{SCM}}(\hat{\rho}_N^*)\sqrt{-2\log\eta} \tag{7}$$

Cette procédure permet de construire le détecteur optimal avec la RSCM. Notons que cette procédure peut être appliquer au cas plus simple de minimisation de la probabilité de fausse alarme. Ces développements adaptés à un bruit gaussien ne sont plus valables lorsque le bruit devient non-gaussien, hétérogène, impulsif ou encore dans le cas où des outliers seraient présents dans les données d'apprentissage. Ceci fait l'objet de la section suivante.

4. Design optimal du ANMF-RTE : bruit non-gaussien

Cette section traite donc du cas du détecteur ANMF-RTE pour du bruit non-gaussien. L'approche est similaire à celle présentée précédemment si ce n'est que l'analyse de l'estimateur RTE de la matrice de covariance est bien plus complexe, principalement du à la définition de cet estimateur au travers d'une équation de point fixe. Ceci induit une forte dépendance entre toutes les observations et rend impossible les approches classique utilisées dans la théorie des matrices aléatoires en grande dimension.

Pour le bruit non-gaussien, nous supposons que les observations vérifient les conditions suivantes :

SUPPOSITION 11. — Pour $i \in \{1, \cdots, n\}$, $\mathbf{x}_i = \sqrt{\tau_i}\mathbf{C}_N^{\frac{1}{2}}\mathbf{w}_i = \sqrt{\tau_i}\mathbf{z}_i$ où

– $\mathbf{w}_1, \cdots, \mathbf{w}_n$ sont des vecteurs complexes de taille $N \times 1$, indépendants, unitairement invariants, centrés et vérifiant $\|\mathbf{w}_i\|^2 = N$,

– $\mathbf{C}_N \in \mathbb{C}^{N\times N}$ est telle que $\limsup \|C_N\| < \infty$ and $\frac{1}{N}\operatorname{tr}\mathbf{C}_N = 1$.

– $\tau_i > 0$ sont indépendants de $\mathbf{w}_i$.

– $\liminf \frac{1}{N}\mathbf{p}^*\mathbf{C}_N\mathbf{p} > 0$.

□

Le modèle décrit dans les hypothèses 5.2 est ~~exactement~~ celui des distributions CES qui contient une infinité de distributions connues, selon la distribution des τ_i. Afin de conserver un cadre complètement général, la loi des τ_i ne sera pas spécifiée. Mais en particulier, le cas précédent de la loi gaussienne peut être simplement obtenue dans ce contexte en posant $\tau_i = 1$ pour tout i.

Avant de proposer le design optimal du détecteur ANMF-RTE, nous avons besoin de résultats préliminaires, notamment sur le comportement asymptotique de l'estimateur RTE.

4.1. Préliminaires

Cette section présentent de récents résultats prouvés dans (Couillet, Kammoun, Pascal, 2014) sur le comportement asymptotique du RTE.

Rappelons que le RTE est défini pour $\rho \in \left(\max\left\{0, 1-\frac{n}{N}\right\}, 1\right]$, comme l'unique solution de l'équation de point fixe suivante :

$$\hat{\mathbf{C}}_N(\rho) = (1-\rho)\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{x}_i\mathbf{x}_i^*}{\frac{1}{N}\mathbf{x}_i^*\hat{\mathbf{C}}_N^{-1}(\rho)\mathbf{x}_i} + \rho\mathbf{I}_N.$$

Comme dit précédemment, l'analyse des estimateurs robustes de la matrice de covariance (ou plus généralement de la matrice de dispersion) est un challenge bien plus compliqué que dans le cas gaussien. Une des principales raisons est qu'il n'existe pas d'expression analytique de l'estimateur, autre que l'équation de point fixe. De plus, la dépendance entre les échantillons est non-linéaires, ce qui ne permet pas l'utilisation d'outils standards de la théorie des matrices aléatoires. Afin de répondre à ces problèmes, de récents travaux développés par Couillet et al. (Couillet, Pascal, Silverstein, 2014a ; Couillet, McKay, 2014 ; Couillet, Pascal, Silverstein, 2014b) ont proposés de nouvelles approches. Le principe consiste à remplacer ces estimateurs robustes par des estimateurs équivalents (asymptotiquement) plus simples, et pour lesquels les techniques standards de matrice aléatoire peuvent s'appliquer. En particulier, le RTE défini précédemment a été étudié dans (Couillet, Kammoun, Pascal, 2014). il a été montré dans cet article que le RTE se comporte asymptotiquement, i.e. dans le régime où $N, n \to \infty$ tels que $c_N \to c \in (0, \infty)$, comme la matrice aléatoire $\hat{\mathbf{S}}_N(\rho)$ donnée par :

$$\hat{\mathbf{S}}_N(\rho) = \frac{1}{\gamma_N(\rho)}\frac{1-\rho}{1-(1-\rho)c_N}\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^* + \rho\mathbf{I}_N, \tag{8}$$

où $\gamma_N(\rho)$ est l'unique solution de

$$1 = \int \frac{t}{\gamma_N(\rho)\rho + (1-\rho)t}\nu_N(dt).$$

Notons que c'est estimateur est très simple puisque que c'est une matrice de covariance empirique classique dans lequel un coefficient (scalaire) est ajouté. De plus, ce coefficient dépend d'une équation de point fixe en dimension 1 et non d'une équation de point fixe en dimension N comme c'est le cas pour l'estimateur RTE.

Plus précisément, nous avons le théorème suivant qui lie les deux estimateurs :

THÉORÈME 12. — *Pour tout $\kappa > 0$ petit, soit $\mathcal{R}_\kappa^{\mathrm{RTE}} \triangleq \left[\kappa + \max(0, 1 - c^{-1}), 1\right]$. Alors, lorsque $N, n \to \infty$ avec $c_N \to c \in (0, \infty)$, on a :*

$$\sup_{\rho \in \mathcal{R}_\kappa^{\mathrm{RTE}}} \left\| \hat{\mathbf{C}}_N(\rho) - \hat{\mathbf{S}}_N(\rho) \right\| \xrightarrow{\text{p.s.}} 0.$$

Le théorème 12 établit la convergence en norme de la différence $\hat{\mathbf{C}}_N(\rho) - \hat{\mathbf{S}}_N(\rho)$. Ce résultat permet de transférer certaines propriétés de $\hat{\mathbf{S}}_N(\rho)$ à l'estimateur RTE, notamment l'analyse asymptotique au premier ordre de nombreuses fonctionnelles de $\hat{\mathbf{C}}_N(\rho)$. Cependant, si comme dans la section précédente, nous nous intéressons aux fluctuations, ce résultat n'est pas suffisant. En effet, le théorème 12 peut être simplement formuler par :

$$\sup_{\rho \in \mathcal{R}_\kappa^{RTE}} N^{\frac{1}{2}-\epsilon} \left\| \hat{\mathbf{C}}_N(\rho) - \hat{\mathbf{S}}_N(\rho) \right\| \xrightarrow{\text{p.s.}} 0.$$

Pour chaque $\epsilon > 0$, cette convergence n'est pas suffisante pour obtenir la convergence de fonctions classiques dont les fluctuations sont de l'ordre $N^{-\frac{1}{2}}$ ou N^{-1}. C'est le cas par exemple des formes quadratiques de $\hat{\mathbf{C}}_N(\rho)$ qui sont utilisées dans l'ANMF-RTE ou bien aussi des statistiques linéaires des valeurs propres de $\hat{\mathbf{C}}_N(\rho)$. De plus, bien qu'il semble qu'il soit impossible d'améliorer la vitesse de la convergence ci-dessus, il a été récemment montré dans (Couillet, Kammoun, Pascal, 2014) que les fluctuations de certaines fonctions pouvaient être plus rapides. Ceci est principalement du à l'effet de moyennage qui élimine les termes fluctuant à vitesse lente. En particulier, les formes bilinéaires du type $\mathbf{a}^* \hat{\mathbf{C}}_N^k(\rho) \mathbf{b}$ ont été analysées dans (Couillet, Kammoun, Pascal, 2014), dans lequel la proposition suivante a été montrée :

PROPOSITION 13. — *Soit $\mathbf{a}, \mathbf{b} \in \mathbb{C}^N$ avec $\|\mathbf{a}\| = \|\mathbf{b}\| = 1$ déterministe ou aléatoire, indépendant de $\mathbf{x}_1, \cdots, \mathbf{x}_n$. Alors, lorsque $N, n \to \infty$, avec $c_N \to c \in (0, \infty)$, pour tout $\epsilon > 0$ et pour tout $k \in \mathbb{Z}$,*

$$\sup_{\rho \in \mathcal{R}_\kappa^{RTE}} N^{1-\epsilon} \left| \mathbf{a}^* \hat{\mathbf{C}}_N^k(\rho) \mathbf{b} - \mathbf{a}^* \hat{\mathbf{S}}_N^k(\rho) \mathbf{b} \right| \xrightarrow{\text{p.s.}} 0.$$

où $\mathcal{R}_\kappa$ est défini dans le théorème 12, $k \in \mathbb{Z}$ et représente n'importe quel exposant des matrices $\hat{\mathbf{C}}_N$ et $\hat{\mathbf{S}}_N$.

Cette proposition a des des implications importantes. En effet, comme vu précédemment l'analyse des objets aléatoires de la forme $\mathbf{a}^* \hat{\mathbf{C}}_N^k(\rho) \mathbf{b}$ semble très compliquée voire impossible. Cependant, les formes quadratiques du type $\mathbf{a}^* \hat{\mathbf{S}}_N^k(\rho) \mathbf{b}$ sont bien connues en matrice aléatoire et ont été largement étudiées. Ainsi, il est essentiel de pouvoir transférer les propriétés de $\mathbf{a}^* \hat{\mathbf{S}}_N^k(\rho) \mathbf{b}$ à $\mathbf{a}^* \hat{\mathbf{C}}_N^k(\rho) \mathbf{b}$. La proposition 13 permet ce transfert de propriétés en prenant $\epsilon < \frac{1}{2}$. Non seulement cette proposition implique que $\mathbf{a}^* \hat{\mathbf{C}}_N^k(\rho) \mathbf{b}$ fluctue à l'ordre $N^{-\frac{1}{2}}$ (puisque c'est le cas pour $\mathbf{a}^* \hat{\mathbf{S}}_N^k(\rho) \mathbf{b}$) mais aussi elle prouve que $\mathbf{a}^* \hat{\mathbf{C}}_N^k(\rho) \mathbf{b}$ et $\mathbf{a}^* \hat{\mathbf{S}}_N^k(\rho) \mathbf{b}$ ont asymptotiquement les mêmes fluctuations. Comme dans (Couillet, Kammoun, Pascal, 2014), nous nous intéressons

au cas $k = -1$. La partie suivante présente l'utilisation de ces résultats dans le cas de l'optimisation des performances du ANMF-RTE, notamment au travers du critère "receiver operating characteristic" (ROC).

4.2. Design optimal du détecteur ANMF-RTE

Comme précédemment, pour construire le détecteur ANMF-RTE optimal, nous avons besoin de caractériser la distribution (asymptotique, dans le régime considéré) de $\widehat{T}_N^{\mathrm{RTE}}(\rho)$ sous les hypothèses H_0 et H_1. En utilisant la proposition 13, nous savons que la statistique $\widehat{T}_N^{\mathrm{RTE}}(\rho)$ ne peut être traitée directement, mais qu'elle a les mêmes fluctuations que la statistique $\widetilde{T}_N^{\mathrm{RTE}}(\rho)$ obtenue en remplaçant $\hat{\mathbf{C}}_N(\rho)$ par $\hat{\mathbf{S}}_N(\rho)$:

$$\widetilde{T}_N^{\mathrm{RTE}}(\rho) = \frac{\left|\mathbf{y}^*\hat{\mathbf{S}}_N^{-1}(\rho)\mathbf{p}\right|}{\sqrt{\mathbf{p}^*\hat{\mathbf{S}}_N^{-1}(\rho)\mathbf{p}}\sqrt{\mathbf{y}^*\hat{\mathbf{S}}_N^{-1}(\rho)\mathbf{y}}}$$

où $\hat{\mathbf{S}}_N(\rho)$ est donnée par l'équation (8).

Soit, maintenant, $\underline{\rho} = \rho\left(\rho + \frac{1}{\gamma_N(\rho)}\frac{1-\rho}{1-(1-\rho)c}\right)^{-1}$. Alors, $\hat{\mathbf{S}}_N(\rho) = \rho\underline{\rho}^{-1}\widehat{\mathbf{R}}_N(\underline{\rho})$, où, avec un léger abus de notation, nous noterons $\widehat{\mathbf{R}}_N(\rho)$ la matrice $(1-\rho)\frac{1}{n}\sum_{i=1}^n \mathbf{z}_i\mathbf{z}_i^* + \rho\mathbf{I}_N$. Puisque $\widetilde{T}_N^{\mathrm{RTE}}(\rho)$ est invariant, dans le sens où $\widetilde{T}_N^{\mathrm{RTE}}(\rho)$ ne change pas si $\hat{\mathbf{S}}_N(\rho)$ est multipliée par un scalaire, idem pour $\mathbf{y}$, on a aussi :

$$\widetilde{T}_N^{\mathrm{RTE}}(\rho) = \frac{\left|\frac{1}{\sqrt{\tau}}\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\underline{\rho})\mathbf{p}\right|}{\sqrt{\mathbf{p}^*\widehat{\mathbf{R}}_N^{-1}(\underline{\rho})\mathbf{p}}\sqrt{\frac{1}{\tau}\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\underline{\rho})\mathbf{y}}}$$

où $\tau = 1$ sous H_0.

Ainsi, conditionnellement à τ, les fluctuations de la statistique robuste $\widehat{T}_N^{\mathrm{RTE}}(\rho)$ sous H_0 ou sous H_1 sont les mêmes que celles obtenues dans les théorèmes 2 et 5 lorsque a est remplacée par $\frac{a}{\sqrt{\tau}}$ et ρ par $\underline{\rho}$[8]. Ainsi, nous avons les résultats suivants concernant les probabilités asymptotiques de fausse alarme et de détection :

THÉORÈME 14. — *Probabilité de fausse alarme, (Couillet, Kammoun, Pascal, 2014) As $N, n \to \infty$ with $c_N \to c \in (0, \infty)$,*

$$\sup_{\rho\in\mathcal{R}_\kappa^{\mathrm{RTE}}}\left|\mathbb{P}\left[\widehat{T}_N^{\mathrm{RTE}}(\rho) > \frac{r}{\sqrt{N}}|H_0\right] - e^{-\frac{r^2}{2\sigma^2_{N,\mathrm{RTE}}(\rho)}}\right| \to 0,$$

8. Notons que le vecteur aléatoire $\mathbf{y}$ peut être supposé gaussien sans que cela ne change les distributions asymptotiques de $\sqrt{N}\widehat{T}_N^{RTE}$ sous H_0 et H_1.

où $\rho \mapsto \underline{\rho}$ est défini précédemment et

$$\sigma^2_{N,\mathrm{RTE}}(\rho) \triangleq \frac{1}{2}\frac{\mathbf{p}^*\mathbf{C}_N\mathbf{Q}_N^2(\underline{\rho})\mathbf{p}}{\mathbf{p}^*\mathbf{Q}_N(\underline{\rho})\mathbf{p}\frac{1}{N}\operatorname{tr}\mathbf{C}_N\mathbf{Q}_N(\underline{\rho})} \times \frac{1}{\left(1-c(1-\underline{\rho})^2 m(-\underline{\rho})^2\frac{1}{N}\operatorname{tr}\mathbf{C}_N^2\mathbf{Q}_N^2(\underline{\rho})\right)}$$

avec $\mathbf{Q}_N(\underline{\rho}) \triangleq \left(\mathbf{I}_N + (1-\underline{\rho})m(-\underline{\rho})\mathbf{C}_N\right)^{-1}$.

THÉORÈME 15 (Probabilité de détection). — *As $N, n \to \infty$ with $c_N \to c \in (0,\infty)$,*

$$\sup_{\rho\in\mathcal{R}_\kappa^{\mathrm{RTE}}}\left|\mathbb{P}\left[\widehat{T}_N^{\mathrm{RTE}}(\rho) > \frac{r}{\sqrt{N}}|H_1\right] - \mathbb{E}\left[Q_1\left(g_{\mathrm{RTE}}(\mathbf{p}), \frac{r}{\sigma_{N,\mathrm{RTE}}(\rho)}\right)\right]\right| \to 0,$$

où l'espérance est preise sous la distribution de τ et où $\sigma_{N,\mathrm{RTE}}(\rho)$ a la même expression que dans le théorème 14 et

$$g_{\mathrm{RTE}}(\mathbf{p}) = \frac{\sqrt{1-c(1-\underline{\rho})^2 m(-\underline{\rho})\frac{1}{N}\operatorname{tr}\mathbf{C}_N^2\mathbf{Q}_N^2(\underline{\rho})}}{\sqrt{\mathbf{p}^*\mathbf{C}_N\mathbf{Q}_N^2(\underline{\rho})\mathbf{p}}} \times \sqrt{\frac{2}{N\tau}}a\left|\mathbf{p}^*\mathbf{Q}_N(\underline{\rho})\mathbf{p}\right|.$$

et Q_1 est la Marcum Q-fonction.

PREUVE 16. — La preuve est donnée dans (Kammoun *et al.*, 2015. arXiv:1501.06027). ■

Maintenant que les performances asymptotiques ont été obtenues, de façon uniforme par rapport au paramètre de régularisation ρ, et comme dans le cas gaussien, il est nécessaire, afin d'obtenir un détecteur optimal, de construire des estimateurs consistants des quantités faisant intervenir des paramètres inconnus, i.e. $\sigma^2_{N,\mathrm{RTE}}(\rho)$ et $f_{\mathrm{RTE}}(\rho)$ donné par :

$$f_{\mathrm{RTE}}(\rho) = \frac{\tau}{2a^2}g^2_{\mathrm{RTE}}(\mathbf{p})$$

Un estimateur consistant de $\sigma^2_{N,\mathrm{RTE}}(\rho)$ est donné dans (Couillet, Kammoun, Pascal, 2014) et fait l'objet de la proposition suivante :

PROPOSITION 17. — *Proposition 1 dans (Couillet, Kammoun, Pascal, 2014)*
Pour $\rho \in \left(\max(\{0, 1-c_N^{-1}\}, 1\right)$. Soit

$$\hat{\sigma}^2_{N,\mathrm{RTE}}(\rho) = \frac{1}{2}\frac{1-\rho\frac{\mathbf{p}^*\hat{\mathbf{C}}_N^{-2}(\rho)\mathbf{p}}{\mathbf{p}^*\hat{\mathbf{C}}_N^{-1}(\rho)\mathbf{p}}}{(1-c_N+c_N\rho)(1-\rho)}$$

et soit $\hat{\sigma}^2_{N,\mathrm{RTE}}(1) \triangleq \lim_{\rho\uparrow 1} \hat{\sigma}^2_N(\rho)$. *Alors, on a :*

$$\sup_{\rho\in\mathcal{R}^{\mathrm{RTE}}_\kappa} \left|\sigma^2_{N,\mathrm{RTE}}(\rho) - \hat{\sigma}^2_{N,\mathrm{RTE}}(\rho)\right| \xrightarrow{\text{p.s.}} 0.$$

De façon analogue, un estimateur consistant $f_{\mathrm{RTE}}(\rho)$ est donné par la proposition suivante :

PROPOSITION 18. — *Pour* $\rho \in \left(\max\left\{0, 1 - c_N^{-1}\right\}, 1\right)$, *soit*

$$\hat{f}_{\mathrm{RTE}}(\rho) = \left(\mathbf{p}^*\hat{\mathbf{C}}_N^{-1}(\rho)\mathbf{p}\right)^2 \left(\frac{1}{N}\operatorname{tr}\hat{\mathbf{C}}_N(\rho) - \rho\right)$$
$$\times \frac{(1 - c_N + c_N\rho)^2}{\mathbf{p}^*\hat{\mathbf{C}}_N^{-1}(\rho)\mathbf{p} - \rho\mathbf{p}^*\hat{\mathbf{C}}_N^{-2}(\rho)\mathbf{p}}$$

et soit $\hat{f}_{\mathrm{RTE}} \triangleq \lim_{\rho\uparrow 1} \hat{f}_{\mathrm{RTE}}(\rho)$. *Alors, on a :*

$$\sup_{\rho\in\mathcal{R}^{\mathrm{RTE}}_\kappa} \left|\hat{f}_{\mathrm{RTE}}(\rho) - f_{\mathrm{RTE}}(\rho)\right| \xrightarrow{\text{p.s.}} 0.$$

PREUVE 19. — La preuve est donnée dans (Kammoun *et al.*, 2015. arXiv:1501.06027). ■

Puisque les résultats de la proposition 18 et du théorème 15 sont uniforme en ρ, nous avons le corollaire suivant :

COROLLAIRE 20. — *Soit* $\hat{f}_{\mathrm{RTE}}(\rho)$ *défini dans la proposition 7, Soit* $\hat{\rho}^*_N$ *satisfaisant*

$$\hat{\rho}^*_N \in \arg\max_{\rho\in\mathcal{R}^{\mathrm{RTE}}_\kappa} \left\{\hat{f}_{\mathrm{RTE}}(\rho)\right\}.$$

Alors, pour tout $r > 0$,

$$\mathbb{P}\left(\sqrt{N}T_N(\hat{\rho}^*_N) > r|H_1\right)$$
$$- \max_{\rho\in\mathcal{R}^{\mathrm{RTE}}_\kappa} \left\{\mathbb{P}\left(\sqrt{N}T_N(\rho) > r\right)|H_1\right\} \xrightarrow{\text{p.s.}} 0.$$

Ainsi, de façon similaire au cas gaussien, il est maintenant possible de construire un détecteur, robuste cette fois-ci, et optimal (encore au sens de l'optimisation des performances de détection) :

– out d'abord, choisissons une des valeurs du paramètre de régularisation qui maximise $\hat{f}_{RTE}(\rho)$:

$$\hat{\rho}^*_N \in \arg\max_{\rho\in\mathcal{R}^{RTE}_\kappa} \left\{\hat{f}_{RTE}(\rho)\right\};$$

– Ensuite, pour assurer une probabilité de fausse alarme asymptotique fixée et égale à η, prenons le seuil de détection $\hat{r}$, tel que :

$$\hat{r} = \hat{\sigma}_{N,RTE}(\hat{\rho}^*_N)\sqrt{-2\log\eta}$$

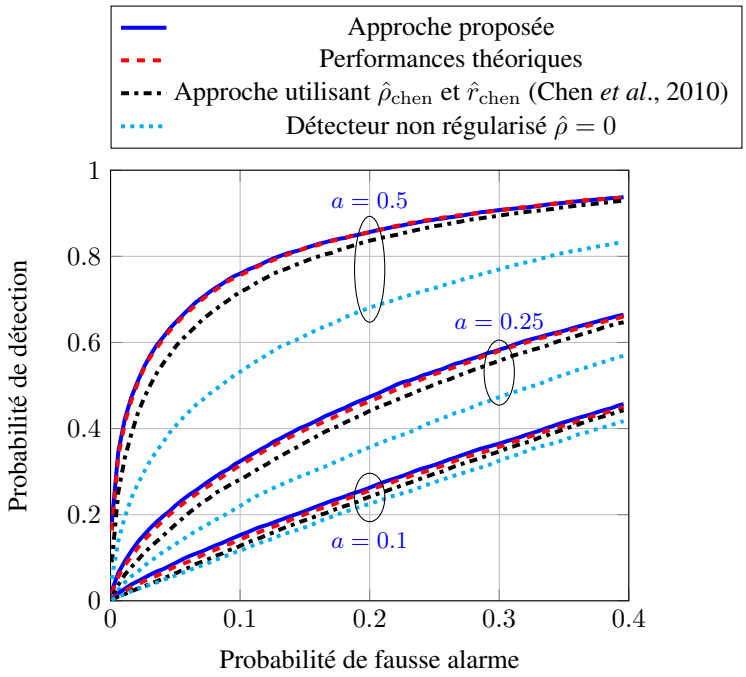

FIGURE *1. Courbes ROC des détecteurs ANMF-RSCM pour* $a = 0.1, 0.25, 0.5$, $\mathbf{p} = \mathbf{a}(\theta)$ *avec* $\theta = 20^o$, $N = 30$, $n = 60$ *: cas gaussien*

5. Simulations

5.1. Bruit gaussien

Dans une première simulation, nous considérons le cas où le bruit aditif est gaussien, de matrice de covariance $\mathbf{C}_N$ ayant une structure de Toeplitz :

$$[\mathbf{C}_N]_{i,j} = \begin{cases} b^{j-i} & i \leq j \\ \left(b^{i-j}\right)^* & i > j \end{cases}, \qquad |b| \in]0, 1[\,, \tag{9}$$

où $b = 0.96\jmath$ $N = 30$ et $n = 60$. Le vecteur directionnel (steering vector) $\mathbf{p}$ est donné par

$$\mathbf{p} = \mathbf{a}(\theta) \tag{10}$$

où $\theta \mapsto [\mathbf{a}(\theta)]_k = e^{-\jmath\pi k \sin(\theta)}$. Pour cette simulation $\theta = 20^o$.

Pour chaque tirage Monte Carlo, les données simulées correspondent à $\mathbf{y} = \alpha\mathbf{p}+\mathbf{x}$ pour la donnée primaire (celle qui est testée dans le détecteur) et à $\mathbf{y}_1 = \mathbf{x}_1, \cdots, \mathbf{y}_n =$

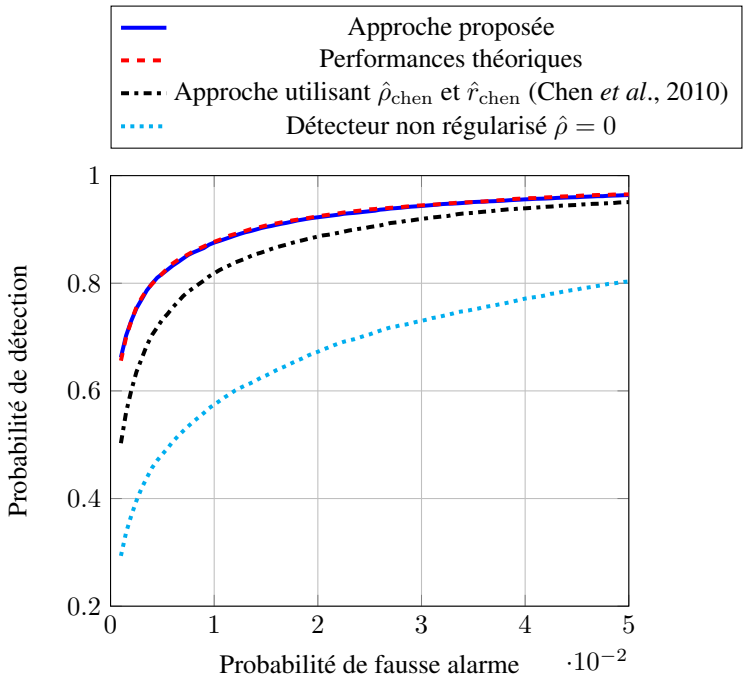

FIGURE 2. *Courbes ROC des détecteurs ANMF-RSCM pour* $a = 0.9$, $\mathbf{p} = \mathbf{a}(\theta)$ *avec* $\theta = 20^o$, $N = 30$, $n = 60$ *: cas gaussien*

$\mathbf{x}_n$ pour les données secondaires qui servent à estimer la matrice de covariance $\hat{\mathbf{R}}_N(\hat{\rho}_N^*)$ et le paramètre $\hat{\rho}_N^*$. En particulier, les valeurs du paramètre de régularisation et du seuil de détection sont déterminées en utilisant les équations (6) et (7). A partir des résultats numériques, nous avons constaté que la fonction $\hat{f}_{\text{SCM}}(\rho)$ était unimodale et ainsi, le maximum peut être obtenu simplement par des méthodes classiques.

Afin d'évaluer les performances du détecteur proposé, i.e. le ANMF-RSCM, nous avons comparé ce dernier avec deux détecteurs ANMF-RSCM pour lesquels le paramètre de régularisation est obtenu de façon différente. Le premier, et le plus pertinent, est celui obtenu avec le paramètre de régularisation dérivé par Chen et al. (Chen *et al.*, 2010, Equation (19)). Il sera noté $\hat{\rho}_{\text{chen}}$. Le second détecteur est celui obtenu lorsqu'il n'y a pas de régularisation, i.e. lorsque $\rho = 0$. Enfin, afin d'assurer la probabilité de fausse alarme voulue, nous supposons que pour le détecteur obtenu par Chen, le seuil de détection $\hat{r}_{\text{chen}}$ est donné par $\hat{r}_{\text{chen}} = \hat{\sigma}_{N,SCM}(\hat{\rho}_{\text{chen}})\sqrt{-2\log\eta}$. Pour l'ANMF non régularisé, la loi exacte (à distance finie) sous H_0, la relation entre la probabilité de fausse alarme et le seuil de détection de ce détecteur est utilisée. Elle est donnée par l'équation (11) dans (Pascal *et al.*, 2004).

La figure 1 représente les courbes ROC, i.e. la probabilité de détection en fonction de la probabilité de fausse alarme, pour le détecteur proposé et les deux autres détecteurs. De plus, les performances théoriques données dans cet article sont aussi représentées sur la figure 1. Les détecteurs sont tracés pour différentes valeurs de $a = \alpha\sqrt{N}$, $a = 0.1, 0.25, 0.5$. Notons que ce paramètre a correspond, de façon indirecte à cause de la renormalisation par $\sqrt{N}$ au Rapport Signal-à-Bruit (RSB).

Tout d'abord, on peut remarquer que pour toutes les valeurs du RSB, il y a un parfaite adéquation entre les performances théoriques établies dans cet article et les simulations Monte Carlo (courbes bleues et rouges pointillées). De plus, on peut aussi constater que pour toutes les valeurs de RSB, le design proposé pour le détecteur ANMF-RTE présente les meilleurs performances comparés aux autres approches.

D'autre part, afin d'illustrer le gain apporté par notre méthode dans les cas intéressants en pratique, i.e. pour des faibles valeurs de la probabilité de fausse alarme, nous avons représenté sur la figure 2 les courbes ROC obtenues pour $a = 0.8$ et pour des probabilités de fausse alarme allant de 10^{-3} à 0.05. Le gain en termes de probabilité de détection va de 15 à 40% pour les faibles valeurs de probabilité de fausse alarme.

5.2. Fouillis non gaussien

Les deuxièmes simulations Monte Carlo s'intéressent aux performances de l'approche proposée dans le cas non gaussien. pour cette simulation, le bruit additif suit une K-distribution, centrée de matrice de covariance $\mathbf{C}_N$ et de paramètre de forme $\nu = 0.5$ (Olilla *et al.*, 2012). La matrice de covariance possède la même structure que celle utilisée dans les simulations précédentes, donnée par l'équation (9), avec $b = 0.96\jmath$ mais avec $N = 30$ et $n = 60$. Notons que les K-distributions correspondent bien aux modèles présentés dans les hypothèses puisqu'une K-distribution est le produit entre un vecteur complexe centré gaussien et une variable τ suivant une loi Gamma, de paramètres ν et $1/\nu$ (afin d'assurer une espérance égale à 1).

Comme dans le cas gaussien, notre approche est comparée à celle d'Ollila and Tyler, qui proposent une méthode pour trouver le paramètre de régularisation optimal (Ollila, Tyler, 2014, Equation (19)). Notons $\hat{\rho}_{\text{Ollila}}$ et $\hat{r}_{\text{Ollila}}$ le paramètre de régularisation et le seuil de détection correspondants. De plus, en accord avec les résultats présentés dans cet article, le seuil de détection $\hat{r}_{\text{Ollila}}$ doit être fixé de la façon suivante, $\hat{r}_{\text{ollila}} = \hat{\sigma}_{N,RTE}(\hat{\rho}_{\text{ollila}})\sqrt{-2\log\eta}$ afin de garantir la probabilité de fausse alarme voulue. Les résultats sont présentés sur la figure 3.

A nouveau, on peut remarquer que pour tous les RSB étudiés, l'approche proposée présente un gain par rapport à la méthode développée par Ollila et al. dans (Ollila, Tyler, 2014). De plus, comme dans les simulations pour le cas gaussien, on peut noter que le gain augment lorsque a augmente mais avec une pente plus faible.

Enfin, dans les dernières simulations, l'impact de a et du paramètre de forme ν de la K-distribution sont analysés. La figure 4 représente la probabilité de détection en fonction de a pour une probabilité de fausse alarme fixée à 0.05 et pour différentes

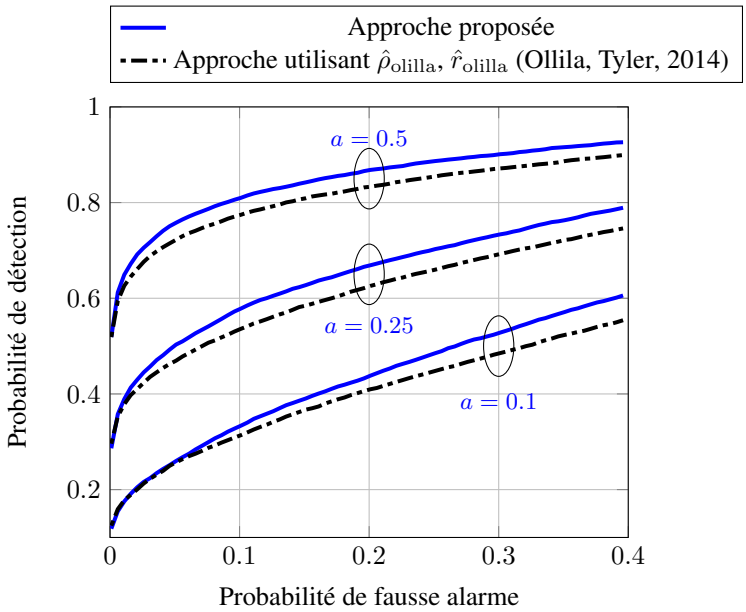

FIGURE *3. Courbes ROC des détecteurs ANMF-RTE pour* $a = 0.1, 0.25, 0.5$, $\mathbf{p} = \mathbf{a}(\theta)$ *avec* $\theta = 20^o$, $N = 30$, $n = 60$*: cas non gaussien, bruit K-distribué*

valeurs du paramètre ν. Notons que ce paramètre de forme ν contrôle l'impulsivité de la distribution : lorsque ν est petit, la K-distribution sera très impulsive (loi à queue lourde) tandis que lorsque $\nu \to \infty$, la K-distribution tend vers la distribution gaussienne.

Sur la figure 4, on peut constater que pour les faibles valeurs de a, les meilleures probabilité de détection sont obtenues pour des distributions à queue (très) lourde tandis que le comportement opposé est observé pour des valeurs élevées de a. Cette différence de comportement s'explique par le fait que dans le cas de distributions à queue lourde (ν petit), il y a beaucoup de réalisations de τ très proches de zéro tandis qu'il y a peu de valeurs "non nulles" mais très élevées. Ainsi, l'augmentation de la probabilité de détection peut s'expliquer par le fait que virtuellement le RSB augmente car il y a un nombre important de faibles valeurs de τ.

6. Conclusion

Dans cet article, nous sous sommes intéressés au paramètre de régularisation lorsque les estimateurs RSM et RTE sont utilisés dans la statistique du détecteur ANMF à la place de la vraie matrice de covariance, inconnue en pratique. Ceci permet de

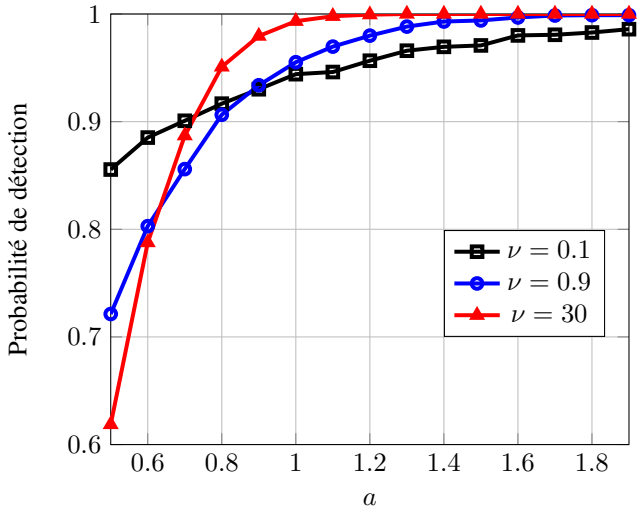

FIGURE *4. Probabilité de détection en fonction de* a, $\mathbf{p} = \mathbf{a}(\theta)$ *avec* $\theta = 20^o$, $N = 30$, $n = 60$, $\nu = 0.1, 0.9, 30$*: cas non gaussien, bruit K-distribué, probabilité de fausse alarme* =5%.

construire les détecteurs ANMF-RSCM et ANMF-RTE. Les modèles statistiques étudiés pour le bruit sont restés très général puisque le cas standard gaussien a d'abord été analysé avant que les résultats soient étendus aux cas des distributions complexes elliptiques. Dans un contexte de grande dimension, i.e. lorsque la dimension et le nombre des données tendent vers l'infini à la même vitesse, cet article a présenté l'analyse statistique complète de ces deux détecteurs sous les deux hypothèses de test, H_0 et H_1 en dérivant les distributions asymptotiques des probabilités de fausse alarme et de détection. Il est intéressant de remarquer que les lois obtenues sont classiques dans la théorie de la détection et très simples à mettre en œuvre : ce sont, sous H_0 une loi de Rayleigh et sous H_1 une loi de Rice. Cette analyse s'est basée sur les comportements asymptotiques des deux estimateurs régularisés. Ensuite, à l'aide de ces résultats, une approche a été proposée pour trouver les paramètres de régularisation optimaux, dans les cas gaussien puis non gaussien, dans le sens où les performances de détection résultantes sont optimisées. Plus précisément, les probabilités de détection associées à ces paramètres de régularisation optimaux des détecteurs ANMF-RSCM et ANMF-RTE sont maximisées. Ces travaux permettent dans des scénarios appliqués difficiles (données hétérogènes, impulsives, données manquantes et/ou aberrantes) d'avoir une procédure de construction d'un détecteur robuste et optimal, l'ANMF-RTE. Les résultats de approche ont été comparés aux différentes méthodes présentes dans la littérature, et cela a illustré une amélioration des performances de détection dans tous

les scénarios étudiés. L'un des avantages majeurs de cette approche est la simplicité d'implémentation de la méthode puisque les expressions analytiques obtenues sont parfaitement classiques. Ceci est d'ailleurs surprenant, au vu de la complexité des résultats obtenus dans le cas où seul le nombre d'observations tend vers l'infini et où la dimension des observations reste fixe. Dans ces cas, les solutions obtenues sont sous forme intégrale, voir par exemple (Kraut *et al*., 2001 ; Pascal, Ovarlez, 2015). Le lien entre ces deux approches reste un problème ouvert, même si quelques pistes ont été récemment explorées (Kammoun *et al*., 2015).

7. Remerciements

Les travaux de F. Pascal ont été partiellement soutenu par l'institut iCODE, projet de recherche de l'Idex Paris-Saclay.

Bibliographie

Abramovich Y. (1981). Controlled Method for Adaptive Optimization of Filters Using the Criterion of Maximum SNR. *Radio Eng. Electron. Phys*, vol. 26, n° 3, p. 87–95.

Abramovich Y. I., Besson O. (2013, octobre). Regularized Covariance Matrix Estimation in Complex Elliptically Symmetric Distributions Using the Expected Likelihood Approach: Part 1: The Over-Sampled Case. , vol. 61, n° 23.

Bidon S., Marcos S., Ovarlez J.-P. (2011). Numéro Spécial : Traitements Spatio-Temporels Adaptatifs (STAP). *Traitement du Signal*, vol. 28, n° 1-2. Consulté sur http://hal-supelec .archives-ouvertes.fr/hal-00659845

Billingsley J. B. (1993, February). *Ground clutter measurements for surface-sited radar*. Rapport technique n° 780. MIT.

Carlson B. D. (1988). Covariance Matrix Estimation Errors and Diagonal Loading in Adaptive Arrays. , vol. 24, n° 4, p. 397–401.

Chen Y., Wiesel A., Eldar Y. C., Hero A. O. (2010, octobre). Shrinkage Algorithms for MMSE Covariance Estimation. , vol. 58, n° 10.

Chen Y., Wiesel A., Hero O. A. (2011). Robust Shrinkage Estimation of High Dimensional Covariance Matrices. , vol. 59, n° 9, p. 4907–4107.

Chitour Y., Pascal F. (2008, octobre). Exact maximum likelihood estimates for SIRV covariance matrix: existence and algorithm analysis. *IEEE Transactions on Signal Processing*, vol. 56, n° 10, p. 4563–4573.

Conte E., Lops M., Ricci G. (1995, avril). Asymptotically Optimum Radar Detection in Compound-Gaussian Clutter. , p. 617-625.

Conte E., Maio A. D., Ricci G. (2002, août). Recursive Estimation of the Covariance Matrix of a Compound-Gaussian Process and Its Application to Adaptive CFAR Detection. , vol. 50, n° 8.

Couillet R., Kammoun A., Pascal F. (2014). Second Order Statistics of Robust Estimators of Scatter. Application to GLRT Detection for Elliptical Signals. *Submitted to Journal of Multivariate Analysis*. (http://arxiv.org/abs/1410.0817)

Couillet R., McKay M. (2014). Large Dimensional Analysis and Optimization of Robust Shrinkage Covariance Matrix Estimators. *Journal of Multivariate Analysis*, vol. 31, p. 99-120.

Couillet R., Pascal F., Silverstein J. W. (2014a). The random matrix regime of Maronna's M-estimator with elliptically distributed samples. *(to appear in) Elsevier JMVA*. (http://arxiv.org/abs/1311.7034)

Couillet R., Pascal F., Silverstein J. W. (2014b, octobre). Robust Estimates of Covariance Matrices in the Large Dimensional Regime. , vol. 60, n° 11, p. 7269–7278.

Gini F., Greco M. (2002, décembre). Covariance Matrix Estimation for CFAR Detection in Correlated Heavy-Tailed Clutter. , vol. 82, n° 12, p. 1847–1859.

Huber P. J. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, vol. 35, n° 1, p. 73–101.

Huber P. J. (1972, August). The 1972 wald lecture robust statistics : A review. *Annals of Mathematical Statistics*, vol. 43, n° 4, p. 1041-1067.

Kammoun A., Couillet R., Pascal F., Alouini M.-S. (2015). Convergence and fluctuations of regularized tyler estimators. *arXiv preprint arXiv:1504.01252*.

Kammoun A., Couillet R., Pascal F., Alouini M.-S. (2015. arXiv:1501.06027). Optimal design of the adaptive normalized matched filter detector. *IEEE Trans on Sig. Proc. (Submitted)*.

Kelker D. (1970, décembre). Distribution Theory of Spherical Distributions and a Location Scale Parameter Generalization. *Sankhyā: The Indian Journal of Statistics, Series A*, vol. 32, n° 4, p. 419–430.

Klemm R. (2002). *Principles of Space-Time Adaptive Processing*. IET.

Kraut S., Scharf L. (1999). The CFAR adaptive subspace detector is a scale-invariant GLRT. , vol. 47, n° 9, p. 2538–2541.

Kraut S., Scharf L., McWhorter L. (2001, january). Adaptive subspace detectors. *IEEE Trans. on Sig. Proc.*, vol. 49, n° 1, p. 1-16.

Ledoit O., Wolf M. (2004). A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices. *Journal of Multivariate Analysis*, vol. 88, n° 2, p. 365-411.

Liu J., Zhang Z.-J., Yang Y., Liu H. (2011). A CFAR Adaptive Subspace Detector for First-Order or Second-Order Gaussian Signals Based on a Single Observation. , vol. 59, n° 11, p. 5126-5140.

Mahot M., Pascal F., Forster P., Ovarlez J. P. (2013, juillet). Asymptotic Properties of Robust Complex Covariance Matrix Estimates. *IEEE Transactions on Signal Processing*, vol. 61, n° 13, p. 3348–3356.

Maronna R. A. (1976, January). Robust M-estimators of multivariate location and scatter. *Annals of Statistics*, vol. 4, n° 1, p. 51-67.

Nohara T., Haykin S. (1991). Canada east coast trials and the k-distribution. *IEE Proceeding, Part. F*, vol. 138, n° 2, p. 82-88.

Olilla E., Tyler D. E., Koivunen V., Poor H. V. (2012, novembre). Complex Elliptically Symmetric Distributions: Survey, New Results and Applications. , vol. 60, n° 11.

Ollila E., Tyler D. E. (2014, novembre). Regularized M-estimators of Scatter Matrix. , vol. 62, n° 22.

Pascal F., Chitour Y., Ovarlez J., Forster P., Larzabal P. (2008, janvier). Covariance structure maximum-likelihood estimates in compound Gaussian noise: existence and algorithm analysis. , vol. 56, n° 1, p. 34–48.

Pascal F., Chitour Y., Quek Y. (2014). Generalized Robust Shrinkage Esitmator and Its Application to STAP Detection Problem . *IEEE Transactions on Signal Processing*, vol. 62, n° 21, p. 5640-5651.

Pascal F., Forster P., Ovarlez J., Larzabal P. (2008). Performance Analysis of Covariance Matrix Estimate in Impulsive Noise. *IEEE Transactions on Signal Processing*, vol. 56, n° 6.

Pascal F., Ovarlez J. P. (2015). Asymptotic Properties of the Robust ANMF. *ICASSP*.

Pascal F., Ovarlez J.-P., Forster P., Larzabal P. (2004, septembre). Constant false alarm rate detection in spherically invariant random processes. In *Proc. of the european signal processing conf., eusipco-04*, p. 2143-2146. Vienna.

Pascal F., Ovarlez J.-P., Forster P., Larzabal P. (2006, septembre). On a sirv-cfar detector with radar experimentations in impulsive noise. In *Proc. of the european signal processing conf., eusipco-06.* Florence, Italy.

Picinbono B., Vezzosi G. (1970). Détection d'un signal certain dans un bruit non stationnaire et non gaussien. *Annales des Télécomm.*, vol. 25, p. 433-439.

Reed I., Mallett J., Brennan L. (1974, novembre). Rapid Convergence Rate in Adaptive Arrays. , vol. 10, n° 6, p. 853–863.

Scharf L. L., Friedlander B. (1994). Matched Subspace Detectors. , vol. 42, n° 8, p. 2146-2157.

Scharf L. L., Lytle D. W. (1971, July). Signal detection in Gaussian noise of unknown level: an invariance application. *IEEE Trans.-IT*, vol. 17, p. 404-411.

Silverstein J. W., Bai Z. D. (1995). On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *Journal of Multivariate Analysis*, vol. 54, n° 2, p. 175-192.

Sun Y., Babu P., Palomar D. P. (2014, octobre). Regularized Tyler's Scatter Estimator: Existence, Uniqueness and Algorithms. , vol. 62, n° 19, p. 5143–5156.

Tyler D. E. (1987). A Distribution-Free M-estimator of Multivariate Scatter. *The Annals of Statistics*, vol. 15, n° 1, p. 234–251.

Wagner S., Couillet R., Debbah M., Slock D. T. M. (2012, juillet). Large system analysis of linear Precoding in MISO broadcast channels with limited feedback. , vol. 58, n° 7, p. 4509-4537.

Ward J. (1994, décembre). *Space-time Adaptive Processing for Airborne Radar*. Rapport technique. Lincoln Lab, MIT.

Ward K. D. (1981, August). Compound representation of high resolution sea clutter. *Electronics Letters*, vol. 17, n° 16, p. 561-563.

Watts S. (1985, December). Radar detection prediction in sea clutter using the compound k-distribution model. *IEE Proceeding, Part. F*, vol. 132, n° 7, p. 613-620.

Zhang T., Cheng X., Singer A. (2014). Marchenko-pastur law for tyler's and maronna's m-estimators. *arXiv preprint arXiv:1401.3424*.