Optimal Design of the Adaptive Normalized Matched Filter Detector using Regularized Tyler Estimators

Abla Kammoun, Romain Couillet, Frédéric Pascal, Mohamed-Slim Alouini

Abstract

This article addresses improvements on the design of the adaptive normalized matched filter (ANMF) for radar detection. It is well-acknowledged that the estimation of the noise-clutter covariance matrix is a fundamental step in adaptive radar detection. In this paper, we consider regularized estimation methods which force by construction the eigenvalues of the covariance estimates to be greater than a positive regularization parameter ρ. This makes them more suitable for high dimensional problems with a limited number of secondary data samples than traditional sample covariance estimates. The motivation behind this work is to understand the effect and properly set the value of ρ that improves estimate conditioning while maintaining a low estimation bias. More specifically, we consider the design of the ANMF detector for two kinds of regularized estimators, namely the regularized sample covariance matrix (RSCM), the regularized Tyler estimator (RTE). The rationale behind this choice is that the RTE is efficient in mitigating the degradation caused by the presence of impulsive noises while inducing little loss when the noise is Gaussian. Based on asymptotic results brought by recent tools from random matrix theory, we propose a design for the regularization parameter that maximizes the asymptotic detection probability under constant asymptotic false alarm rates. Provided Simulations support the efficiency of the proposed method, illustrating its gain over conventional settings of the regularization parameter.

Index Terms

Regularized Tyler's estimator, Adaptive Normalized Mached Filter, robust detection, Random Matrix Theory, Optimal design.

I. Introduction

The estimation of covariance matrices is of fundamental importance for space-time adaptive processing (STAP) which underlies the design of radar systems [1], [2]. In radar detection for instance, it is well-acknowledged that a

A. Kammoun and M.S. Alouini are with the Computer, Electrical, and Mathematical Sciences and Engineering (CEMSE) Division, KAUST, Thuwal, Makkah Province, Saudi Arabia (e-mail: abla.kammoun@kaust.edu.sa, slim.alouini@kaust.edu.sa)

R. Couillet and F. Pascal are with Laboratoire des Signaux et Systèmes (L2S, UMR CNRS 8506) CentraleSupélec-CNRS-Université Paris-Sud, 91192 Gif-sur-Yvette, France (e-mail: romain.couillet@centralesupelec.fr, frederic.pascal@centralesupelec.fr)

Couillet's work is supported by the ANR Project RMT4GRAPH (ANR-14-CE28-0006).

sufficiently accurate covariance matrix is key to enhancing the detection performance (see [3], [4] and references therein). In order to support a possible deficiency in samples (number of samples less than their dimensions), several approaches have been proposed, among which we distinguish methods leveraging structural information on the covariance matrix [5], [6], [7] and regularized covariance matrix estimation methods. [8]. Our focus in this paper is on regularized estimation methods. One regularized estimation method is given by the use of the regularized sample covariance matrix (RSCM). The RSCM fundamentally originates from the diagonal loading approach which can be traced back to the works of Abramovich and Carlson [8], [9]. Similar to the sample covariance matrix (SCM), the RSCM exhibits poor performances when observations contain outliers. This latter scenario is often modeled by assuming that observations are drawn from complex elliptical symmetric distributions (CES), originally introduced by Kelker [10]. In this case, the SCM is no longer the maximum likelihood estimator for the covariance matrix.

A class of covariance estimators coined robust estimators of covariance matrices have been proposed by Huber, and Maronna [11], [12], [13], and extended more recently by Ollila to the complex case [3], [4], [14]. Such estimators based on a generalization of maximum likelihood estimators of the scatter matrix for CES distributions, were shown to exhibit good robustness to the presence of outliers. Similar to the Gaussian case, the regularization technique has been applied to the robust Tyler estimator [15], yielding the so-called regularized Tyler estimator (RTE). A major feature of the RTE is that it exists even when the number of samples is less than their dimensions, which is in contrast to most conventional robust estimation methods. The existence of RTE as well as the convergence of the associated recursive algorithm for computing it have recently been established in several works [16], [17], [18], [19]. Unlike the RSCM, the RTE, as a derivative of the robust Tyler's estimator, is resilient to the presence of outliers. This makes it more suitable to radar applications, for which experimental evidence rules out Gaussian models for the sea clutter [20], [21], [22], [23].

In regularized estimation methods, a recurrent question that naturally arises is how to set the regularization parameter. This question has essentially been investigated in [24], [25] for the RSCM and in [18], [26] for the RTE. Although yielding different expressions, these works have the common denominator of being merely based on a distance minimization between the RTE or the SCM and the true covariance matrix. It is thus not clear whether these choices will allow for good performances when applied to detection problems. We consider in this work the design of the adaptive normalized matched filter (ANMF) for radar detection. Firstly introduced by [27] and analyzed in [28], [29], [30], this scheme was shown to enjoy the interesting features of constant false alarm property with respect to the clutter power and covariance matrix. This detector is obtained by replacing in the test statistic of the normalized matched filter (NMF) the covariance matrix by a given estimate [29], computed based on secondary data, i.e., n signal free independent and identically distributed (i.i.d.) observations. Of interest in this work are the cases where the RSCM or the RTE are used in place of the unknown covariance matrix. We will consider first the scenario where the detector operates over Gaussian correlated clutters and thus uses the RSCM as a replacement for the unknown covariance matrix. This scheme will be referred to as ANMF-RSCM. We particularly characterize the behavior of its corresponding false alarm and detection probabilities under the asymptotic regime in which the number of secondary data samples and their dimensions grow simultaneously to infinity. To this end, advanced tools from random matrix theory, established recently in [31] are extensively used. In a second part, we consider the case in which the clutter is drawn from heavy tailed distributions. It is thus natural to assume that the detector uses the RTE, since the RSCM is vulnerable to the presence of outliers and may provide poor performances. By exploiting recent results on the asymptotic behavior of the RTE estimator [31], [32], we prove that, up to a certain change of variable, the ANMF-RTE is asymptotically equivalent to the ANMF-RSCM when operating over Gaussian clutters. This argues in favor of the role of the RTE to retrieve the Gaussian performances while operating over heavy-tailed distributed clutters. Finally, we carry out a set of distinct experiments in order to illustrate the superiority of the proposed design to some of the recent settings of the regularization parameter [25], [18]. Additionally, we study using simulations the impact of the clutter statistics on the detection probabilities for different SNR values. Part of the results of the paper, which concern the use of the RTE in a non-Gaussian clutter, have been published without proof in [33].

The remainder of the paper is organized as follows. In the first section, we introduce the considered problem. Then, we propose an optimal design approach for the ANMF-RTE and the ANMF-RSCM. Finally, we illustrate using simulations the gain of the proposed design method over conventional settings of the regularization parameter.

Notations: Throughout this paper, we depict vectors in lowercase boldface letters and matrices in uppercase boldface letters. The notation $(.)^*$ stands for the transpose conjugate while $\mathrm{tr}(.)$ and $(.)^{-1}$ are the trace and inverse operators. The notation $\|.\|$ stands for the Euclidean norm for vectors and for spectral norm for matrices. The arrow $\xrightarrow{\text{a.s.}}$ designates almost sure convergence. The statement $X \triangleq Y$ defines the new notation X as being equal to Y.

II. Problem Statement

We consider the problem of detecting a complex signal vector $\mathbf{p}$ corrupted by an additive noise as:

$$\mathbf{y} = \alpha\mathbf{p}+\mathbf{x}$$

where $\mathbf{y} \in \mathbb{C}^N$ represents the vector received by an N-dimensional array of sensors, $\mathbf{x}$ stands for the noise clutter and α is a complex scalar modeling the unknown target amplitude. The signal detection problem is phrased as the following binary hypothesis test:

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

Several models for the clutter $\mathbf{x}$ have been proposed. Among them, we distinguish the class of CES random variates which encompass most of the commonly encountered random models, including the standard Gaussian distribution, the K-distribution, the Weibull distribution and many others [3]. A CES distributed random variable is given by:

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

where τ is a positive scalar random variable called the *texture*, $\mathbf{C}_N$ is the covariance matrixand $\tilde{\mathbf{w}}$ is an N-dimensional vector independent of τ, zero-mean unitarily invariant with norm $\|\tilde{\mathbf{w}}\| = \sqrt{N}$. The quantity $\mathbf{C}_N^{\frac{1}{2}}\tilde{\mathbf{w}}$ is referred to as *speckle*. The design of an appropriate statistic to the above hypothesis test depends on the amount of knowledge

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2017.2766538, IEEE Transactions on Aerospace and Electronic Systems

that is available to the detector. If the clutter is Gaussian with known covariance matrix $\mathbf{C}_N$ while α is unknown, the Generalized Likelihood Ratio (GLRT) for the detection problem in (1) results in the following test statistic:

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

which corresponds to the square-root statistic of the ANMF detector. The statistic T_N has been derived independently by several works, thereby leading the corresponding detector to have many alternative names: the constant false alarm (CFAR) matched subspace detector (MSD) [34], the normalized matched filter (NMF) [35], or the Linear Quadratic GLRT (GLRT-LQ) [36]. Although optimality of T_N in regards of the GLRT principle was established under Gaussian clutters, its use for non-Gaussian clutters is appropriate, knowing that it is provably asymptotically optimal as N becomes increasingly large, under the setting of compound-Gaussian distributed clutters.

Since the covariance matrix $\mathbf{C}_N$ is unknown in practice, a popular approach consists in replacing in T_N the unknown covariance matrix $\mathbf{C}_N$ by an estimate built on signal free i.i.d. observations $\mathbf{x}_1, \cdots, \mathbf{x}_n$, termed secondary data. The resulting detector is called the adaptive normalized matched filter (ANMF). Several concurrent estimators of $\mathbf{C}_N$ can be used. The most popular one is the traditional sample covariance matrix (SCM) given by:

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

which corresponds to the unconstrained Maximum-Likelihood estimator (MLE) if the clutter is Gaussian distributed. However, in some scenarios where the available number of observations n is of the same order or smaller than N, the SCM, being ill-conditioned, will not lead to accurate detection results[1]. A practical approach that has received considerable attention is to regularize the SCM, thereby yielding the regularized SCM (RSCM) given by:

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

where the parameter $\rho \in [0,1]$ serves to give more or less importance to the sample covariance matrix $\widehat{\mathbf{R}}_N$ depending on the available number of samples. The ANMF that uses the RSCM as a plug-in estimator of $\mathbf{C}_N$ will be referred to as ANMF-RSCM.

The RSCM, fundamentally relying on SCM, is vulnerable to the presence of outliers, and as such is not efficient when the samples are drawn from heavy tailed non-Gaussian distributions. A standard alternative to the RSCM is constituted by the class of robust-covariance estimators, known for their resilience to atypical observations. The robust estimator that will be considered in this work was defined in [17] as the unique solution $\hat{\mathbf{C}}_N(\rho)$ to:

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

with $\rho \in \left(\max\left(0, 1-\frac{n}{N}\right), 1\right]$. This estimator corresponds to a hybrid robust-shrinkage estimator reminding Tyler's M-estimator of scale [15] and Ledoit-Wolf's shrinkage estimator [24]. We will thus refer to it as the Regularized-Tyler Estimator (RTE). Besides its robustness, the RTE has many interesting features. First, it is well-suited to situations

[1]Traditionally, it is assumed that $2N$ observations are required to ensure good performances of the sub-optimal filtering, i.e., a 3 dB loss of the output SNR compared to optimal filtering [37].

where $c_N \triangleq \frac{N}{n}$ is large while standard robust covariance estimates are ill-conditioned or even undefined if $N > n$. By varying the regularization parameter ρ, one can move from the unbiased Tyler-estimator [38] ($\rho = 0$) to the identity matrix ($\rho = 1$) which represents a crude guess for the unknown covariance $\mathbf{C}_N$. Its relation to the Tyler's estimator has recently been reported in [18] by viewing it as the solution of a penalized M-estimation cost function. We will denote by ANMF-RTE the ANMF detector that uses the RTE instead of the unknown covariance matrix.

Upon replacing in T_N the unknown covariance matrix by a regularized estimate, be it the SCM or the RTE, the question of how should the regularization parameter ρ be set naturally arises. Recent previous works dealing with this issue propose to set ρ in such a way as to minimize a certain mean-squared-error between $\hat{\mathbf{C}}_N$ and $\mathbf{C}_N$ [18], [39]. While easy-to-compute estimates of these values of ρ were provided, one of the major criticism to these choices is that they are performed regardless of the application under consideration. In particular, a more relevant choice to the application under study consists in selecting the values of ρ that maximize the probability of detection while keeping fixed the FAPs. These values will be considered as optimal in regards of radar detection applications.

To this end, one needs to characterize the distribution of $\widehat{T}_N^{\mathrm{RSCM}}(\rho)$ and $\widehat{T}_N^{\mathrm{RTE}}(\rho)$ given by:

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

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

under hypotheses H_0 and H_1. For fixed N and n, this is not an easy task and in our opinion would not lead, if ever feasible, to easy-to-compute expressions for the optimal values of ρ. In this paper, we relax this restrictive assumption by considering the case where N and n go to infinity with $\frac{N}{n} \to c \in (0, \infty)$. This in particular allows to leverage the recent results of [31] that will be reviewed in Section IV-A.

III. Optimal design of the ANMF-RSCM detector: Gaussian clutter case

In this section, we consider the case of a Gaussian clutter. In other words, we assume that all the secondary data $\mathbf{x}_1, \cdots, \mathbf{x}_n$ are drawn from Gaussian distribution with zero-mean and covariance $\mathbf{C}_N$. Prior to introducing the results about the false alarm and detection probabilities, we shall introduce the following assumptions and notations:

Assumption 1. *For* $i \in \{1, \cdots, n\}$, $\mathbf{x}_i = \mathbf{C}_N^{\frac{1}{2}} \mathbf{w}_i$ *with:*

- $\mathbf{w}_1, \cdots, \mathbf{w}_n$ *are* $N \times 1$ *independent standard complex Gaussian random vectors with zero-mean and covariance* $\mathbf{I}_N$*,*
- $\mathbf{C}_N \in \mathbb{C}^{N \times N}$ *is such that* $\limsup \|\mathbf{C}_N\| < \infty$ *and* $\frac{1}{N} \operatorname{tr} \mathbf{C}_N = 1$*,*
- $\liminf_N \frac{1}{N} \mathbf{p}^* \mathbf{C}_N \mathbf{p} > 0$.

Note that the normalization $\frac{1}{N} \operatorname{tr} \mathbf{C}_N = 1$ is not a restricting constraint since the statistics under study are invariant to the scaling of $\mathbf{C}_N$. The last item in Assumption 1 is required for technical purposes in order to ensure that the considered statistic exhibits fluctuations under H_0 and H_1. In practice, this assumption implies that the steering vector does not lie in the null space of the covariance matrix $\mathbf{C}_N$.

For $\rho \in (0,1]$, we define the RSCM as in (2) and call its corresponding statistic $\widehat{T}_N^{RSCM}$. In order to pave the way towards an optimal setting of the regularization coefficient ρ, we need to characterize the asymptotic false alarm and detection probabilities under the assumptions that $c_N \triangleq \frac{N}{n} \to c$. That is, provided H_0 or H_1 is the actual scenario, ($\mathbf{y}=\mathbf{x}$ or $\mathbf{y}=\alpha\mathbf{p}+\mathbf{x}$), we shall evaluate the probabilities $\mathbb{P}\left[\widehat{T}_N^{\text{RSCM}} > \Gamma | H_0\right]$ and $\mathbb{P}\left[\widehat{T}_N^{\text{RSCM}} > \Gamma | H_1\right]$ for $\Gamma > 0$. Before going further, we need to stress that some extra assumptions on the order of magnitude of α and Γ with respect to N should be made to avoid getting trivial results. Indeed, it appears that under H_0, the random quantities $\frac{1}{\sqrt{N}}\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\frac{\mathbf{p}}{\|\mathbf{p}\|}$, $\frac{1}{N}\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{y}$, and $\mathbf{p}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\frac{\mathbf{p}}{\|\mathbf{p}\|^2}$ are standard objects in random matrix theory, which converge almost surely to their means when both N and n grow to infinity with the same pace[40]. As a result, since $\frac{1}{\sqrt{N}}\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{p} \xrightarrow{\text{a.s.}} 0$, $\widehat{T}_N^{\text{RSCM}} \xrightarrow{\text{a.s.}} 0$ for all $\Gamma > 0$, which does not allow to infer much information about the FAP. It turns out that the proper scaling of Γ should be $\Gamma = N^{-\frac{1}{2}}r$ for some fixed $r > 0$, an assumption already considered in [31]. Similarly, one can see that under H_1, the presence of a signal component in $\mathbf{y}$ causes $\widehat{T}_N^{\text{RSCM}}$ to converge almost surely to some positive constant if α does not vary with N. Therefore, for $\Gamma = N^{-\frac{1}{2}}r$, $\mathbb{P}\left[\widehat{T}_N^{\text{RSCM}} > \Gamma | H_1\right] \to 1$. In order to avoid this trivial statement, we shall assume that $\alpha = N^{-\frac{1}{2}}a$ for some fixed $a > 0$ with $\|\mathbf{p}\|^2 = N$. In practice, this indicates that a low level of the SNR, defined as the ratio between the gain of the signal of interest and the power of the clutter, is considered, since:

$$\text{SNR} = \frac{\alpha^2\mathbf{p}^*\mathbf{p}}{\mathbb{E}\left[\mathbf{x}^*\mathbf{x}\right]} = \frac{a^2}{N}.$$

Our analysis is based on results from random matrix theory. To facilitate the exposition of our results, we need to introduce some key quantities, frequently introduced in the framework of this theory.

Denote for $z \in \mathbb{C}\backslash\mathbb{R}_+$ by $m_N(z)$ the unique complex solution to:

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

that satisfies $\Im(z)\Im(m_N(z)) \geq 0$ or unique positive if $z < 0$. The existence and uniqueness of $m_N(z)$ follows from standard results of random matrix theory [41]. It is a deterministic quantity, which can be computed easily for each z using fixed-point iterations. In our case, it helps characterize the asymptotic behavior of the empirical spectral measure of the random matrix [2] $(1-\rho)\frac{1}{n}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^*$. Define also for $\kappa > 0$, $\mathcal{R}_\kappa^{\text{SCM}}$ as:

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

[2]Let $\hat{\nu}_N = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}$ be the empirical spectral measure of the random matrix $\frac{1-\rho}{n}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^*$ with $\lambda_1, \cdots, \lambda_N$ the eigenvalues of $\frac{1-\rho}{n}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^*$. Denote by $\hat{m}_N(z)$ its Stieltjes transform given by $\hat{m}_N(z) = \int (t-z)^{-1}\hat{\nu}_N(dt) = \frac{1}{N}\sum_{i=1}^N \frac{1}{\lambda_i - z}$. Then, quantity $m_N(z)$ is the Stieljes transform of a certain deterministic measure μ_N, (i.e, $m_N(z) = \int (t-z)^{-1}\mu_N(dt)$) which approximates in the almost sure sense $\hat{m}_N(z)$ (i.e., $\hat{m}_N(z) - m_N(z) \xrightarrow{\text{a.s.}} 0$.).

With these notations at hand, we are now ready to analyze the asymptotic behaviour of the false alarm and detection probabilities. The proof for the following Theorem will not be provided since, as we shall see in Section IV, it follows directly by applying the same approach used in [31].

Theorem 1 (False alarm probability,[31]). *As $N, n \to \infty$ with $c_N \to c \in (0, \infty)$,*

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

where:

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

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

The uniformity over ρ of the convergence result in Theorem 1 is essential in the sequel. It obviously implies the pointwise convergence for each $\rho > 0$ but, more importantly, it will allow us to handle the convergence of the FAP when random values of the regularization parameter are considered. This feature becomes all the more interesting knowing that the detector is required to set the regularization parameter based on random received secondary data. Note that, for technical issues, a set of the form $[0, \kappa)$, where $\kappa > 0$ is as small as desired but fixed, has to be discarded from the uniform convergence region. The result of Theorem 1 provides an analytical expression for the FAP. Since this expression depends on the unknown covariance matrix, it is of practical interest to provide a consistent estimate for it:

Proposition 2 [31]). *For $\rho \in (0, 1)$, define*

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

and let $\hat{\sigma}^2_{N,SCM}(1) = \lim_{\rho \uparrow 1} \hat{\sigma}^2_{N,SCM}(\rho) = \frac{\mathbf{p}^* \widehat{\mathbf{R}}_N \mathbf{p}}{\operatorname{tr} \widehat{\mathbf{R}}_N}$. Then, we have, for any $\kappa > 0$, as $N, n \to \infty$ with $c_N \to c$:

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

and

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

The proof of Proposition 2 follows along the same lines as that of Proposition 1 in [31] and is therefore omitted. We will now derive the asymptotic equivalent for $\mathbb{P}\left[\widehat{T}_N^{\mathrm{RSCM}}(\rho) > \frac{r}{\sqrt{N}} | H_1 \right]$, where under H_1 the received vector $\mathbf{y}$ is supposed to be given by:

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

with $\mathbf{x}$ distributed as the $\mathbf{x}_i$'s in Assumption 1. The results that will be presented in the sequel constitute the major contribution of the present work.

Theorem 3 (Detection probability). *As $N, n \to \infty$ with $c_N \to c$, we have for any $\kappa > 0$*

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

where Q_1 is the Marcum Q-function[3] *while $\sigma^2_{N,\mathrm{SCM}}$ is given in Theorem 1 and $g_{SCM}(\mathbf{p})$ is given by:*

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

Proof. See Appendix A. □

According to Theorem 1 and Theorem 3, $\widehat{T}_N^{\mathrm{RSCM}}(\rho)$ behaves differently depending on whether a signal is present or not. In particular, under H_0, $\sqrt{N}\widehat{T}_N^{\mathrm{RSCM}}(\rho)$ behaves like a Rayleigh distributed random variate with parameter $\sigma_{N,\mathrm{SCM}}(\rho)$ while it becomes well-approximated under H_1 by a Rice distributed random variable with parameters $g_{\mathrm{SCM}}(\mathbf{p})$ and $\sigma_{N,\mathrm{SCM}}(\rho)$. It is worth mentioning that in the theory of radar detection, false alarm and detection probabilities take in general much more involved expressions. It is only when white Gaussian noises are considered that they coincide with Rayleigh and Rice distributions.

It seems that the striking simplicity of our results inhere in the double averaging effect resulting from the considered asymptotic regime. Such a simplification has not been for instance met when considering the classical regime of n tending to infinity while N is fixed [42]. We will now discuss the choice of the regularization parameter ρ and the threshold r. In accordance with the theory of radar detection, we aim at setting ρ and r in such a way to keep the asymptotic FAP equal to a fixed value η while maximizing the asymptotic probability of detection. From Theorem 1, one can easily see that the values of r and ρ that provide an asymptotic FAP equal to η should satisfy:

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

From these choices, we have to take those values that maximize the asymptotic detection which is given, according to Theorem 3, by:

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

The second argument of Q_1 should be kept fixed in order to ensure the required asymptotic FAP. As the Marcum-Q function increases with respect to the first argument, the optimization of the detection probability boils down to

[3] $Q_1(a,b) = \int_b^{+\infty} x \exp\left(-\frac{x^2+a^2}{2}\right) I_0(ax)dx$ where I_0 is the zero-th order modified Bessel function of the first kind.

considering the following values of ρ:

$$\rho \in \operatorname*{argmax}_{\rho \in \mathcal{R}_\kappa^{\mathrm{SCM}}} \{f_{\mathrm{SCM}}(\rho)\} \tag{6}$$

where:

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

and $\operatorname{argmax}_{\rho \in \mathcal{R}_\kappa^{\mathrm{SCM}}} \{f_{\mathrm{SCM}}(\rho)\}$ denotes the set of ρ maximizing $f_{\mathrm{SCM}}(\rho)$.

Let ρ_{SCM}^* be among the values satisfying (6). The maximal asymptotic detection probability that can be obtained while satisfying an asymptotic false alarm probability equal to η is thus given by:

$$P_{d,\mathrm{SCM}} = Q_1\left(\sqrt{2}a\sqrt{f_{\mathrm{SCM}}(\rho_{\mathrm{SCM}}^*)}, \frac{r^*}{\sigma_{N,\mathrm{SCM}}(\rho_{\mathrm{SCM}}^*)}\right) \tag{7}$$

where

$$r_{\mathrm{SCM}}^* = \sigma_{N,\mathrm{SCM}}(\rho_{\mathrm{SCM}}^*)\sqrt{-2\log\eta}.$$

However, the optimization of $f_{\mathrm{SCM}}(\rho)$ is not possible in practice, since the expression of $f_{SCM}(\rho)$ features the covariance matrix $\mathbf{C}_N$ which is unknown to the detector. Acquiring a consistent estimate of $f_{\mathrm{SCM}}(\rho)$ based on the available $\widehat{\mathbf{R}}_N$ is thus mandatory. This is the goal of the following Proposition.

Proposition 4. *For $\rho \in (0,1)$, define $\hat{f}_{\mathrm{SCM}}(\rho)$ as:*

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

and let $\hat{f}_{\mathrm{SCM}}(1) \triangleq \lim_{\rho\uparrow 1}\hat{f}_{\mathrm{SCM}}(\rho) = \frac{N}{\mathbf{p}^*\widehat{\mathbf{R}}_N\mathbf{p}}$. Then, we have as $N,n\to\infty$ with $c_N \to c$:

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

where we recall that $\mathcal{R}_\kappa^{\mathrm{SCM}} = [\kappa, 1]$.

Proof. See Appendix B. □

Since the results in Proposition 4 and Theorem 3 are uniform in ρ, we have the following corollary:

Corollary 5 [31]**).** *Let $\hat{f}_{SCM}(\rho)$ be defined as in Proposition 4. Define $\hat{\rho}_N^*$ as any value satisfying:*

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

Then, for every $r > 0$, as $N,n\to\infty$ with $c_N\to c$:

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

Proof. The proof is similar to that of Corollary 1 of [31] and is thus omitted. □

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2017.2766538, IEEE Transactions on Aerospace and Electronic Systems

From Corollary 5, the following design procedure leads to optimal performance detection results:

- First, setting the regularization parameter to one of the values maximizing $\hat{f}_{\rm SCM}(\rho)$:

$$\hat{\rho}_N^* \in \underset{\rho \in \mathcal{R}_\kappa^{\rm SCM}}{\operatorname{argmax}} \left\{ \hat{f}_{\rm SCM}(\rho) \right\} \tag{8}$$

- Second, selecting the threshold $\hat{r}$ as:

$$\hat{r} = \hat{\sigma}_{N,{\rm SCM}}(\hat{\rho}_N^*) \sqrt{-2\log\eta} \tag{9}$$

It is worth mentioning that, from Corollary 5, the above design strategy is asymptotically optimal in the sense that it should guarantee for N and n large enough close-to-optimal detection performances.

IV. Optimal design of the ANMF-RTE: Non-Gaussian clutter

This section discusses the design of the ANMF-RTE detector in the case where the clutter is non-Gaussian. In particular, we assume that the secondary observations satisfy the following assumptions:

Assumption 2. *For $i \in \{1, \cdots, n\}$, $\mathbf{x}_i = \sqrt{\tau_i}\mathbf{C}_N^{\frac{1}{2}}\mathbf{w}_i = \sqrt{\tau_i}\mathbf{z}_i$ where*

- *$\mathbf{w}_1, \cdots, \mathbf{w}_n$ are $N \times 1$ independent unitarily invariant complex zero-mean random vectors with $\|\mathbf{w}_i\|^2 = N$,*
- *$\mathbf{C}_N \in \mathbb{C}^{N\times N}$ is such that $\limsup \|C_N\| < \infty$ and $\frac{1}{N}\operatorname{tr}\mathbf{C}_N = 1$.*
- *$\tau_i > 0$ are independent of $\mathbf{w}_i$.*
- *$\liminf \frac{1}{N}\mathbf{p}^*\mathbf{C}_N\mathbf{p} > 0$.*

The random model described in Assumption 2 is that of CES distributions which encompass a wide range of observation distributions obtained for different settings of the statistics of τ_i. Prior to stating our main findings, we shall first review some recent results concerning the asymptotic behaviour of the RTE in the asymptotic regime.

A. Background

This section reviews the recent results in [31] about the asymptotic behaviour of the RTE estimator.

Recall that the RTE is defined, for $\rho \in \left(\max\left\{0, 1-\frac{n}{N}\right\}, 1\right]$, as the unique solution to the following equation:

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

The study of the asymptotic behaviour of robust-covariance estimators is much more challenging than that of the traditional sample covariance matrices. The main reasons are twofold: first, robust estimators of covariance matrices do not have closed-form expressions and, second, the dependence between the outer-products involved in their expressions is non-linear. All this does clearly not allow for the use of standard random matrix analysis. In order to study this class of estimators, new technical tools have been developed by Couillet *et al.* [43], [32], [44]. The important advantage of these techniques is that they suggest to replace robust estimators by asymptotically equivalent random matrices for which many results from random matrix theory are applicable. In particular, the RTE estimator defined above has been studied in [31] and has been shown to behave in the regime where $N, n \to \infty$ in such a way that $c_N \to c \in (0, \infty)$ similar to $\hat{\mathbf{S}}_N(\rho)$ given by:

$$\hat{\mathbf{S}}_N(\rho) = \frac{1}{\gamma_N(\rho)} \frac{1-\rho}{1-(1-\rho)c_N} \frac{1}{n} \sum_{i=1}^{n} \mathbf{z}_i \mathbf{z}_i^* + \rho \mathbf{I}_N, \tag{10}$$

where $\gamma_N(\rho)$ is the unique solution to:

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

More specifically, the following theorem applies:

Theorem 6 ([32])**.** *For any $\kappa > 0$ small, define $\mathcal{R}_\kappa^{\mathrm{RTE}} \triangleq \left[\kappa + \max(0, 1-c^{-1}), 1\right]$. Then, as $N, n \to \infty$ with $c_N \to c \in (0, \infty)$, we have:*

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

Theorem 6 establishes a convergence in the operator norm of the difference $\hat{\mathbf{C}}_N(\rho) - \hat{\mathbf{S}}_N(\rho)$. This result allows one to transfer the asymptotic first order analysis of many functionals of $\hat{\mathbf{C}}_N(\rho)$ to $\hat{\mathbf{S}}_N(\rho)$. However, when it comes to the study of fluctuations, this result is of little help. Indeed, although Theorem 6 can be easily refined as

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

for each $\epsilon > 0$, the above convergence does not suffice to obtain the convergence of most of the commonly used functionals which involve fluctuations of order $N^{-\frac{1}{2}}$ or N^{-1} (e.g. quadratic forms of $\hat{\mathbf{C}}_N(\rho)$ or linear statistics of the eigenvalues of $\hat{\mathbf{C}}_N(\rho)$). While a further refinement of the above convergence seems to be out of reach, it has recently been established in [31] that the fluctuations of special functionals can be proved to be much faster. This results from an averaging effect which cancels out terms fluctuating at lower speed. In particular, bilinear forms of the type $\mathbf{a}^* \hat{\mathbf{C}}_N^k(\rho) \mathbf{b}$ were studied in [31], where the following proposition was proved:

Proposition 7 ([31])**.** *Let $\mathbf{a}, \mathbf{b} \in \mathbb{C}^N$ with $\|\mathbf{a}\| = \|\mathbf{b}\| = 1$ deterministic or random independent of $\mathbf{x}_1, \cdots, \mathbf{x}_n$. Then, as $N, n \to \infty$, with $c_N \to c \in (0, \infty)$, for any $\epsilon > 0$ and every $k \in \mathbb{Z}$,*

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

where $\mathcal{R}_\kappa^{RTE}$ is defined as in Theorem 6, where $k \in \mathbb{Z}$ in any power of the matrices $\hat{\mathbf{C}}_N$ and $\hat{\mathbf{S}}_N$.

Some important consequences of Proposition 7 need to be stated. First, we shall recall that, while the crude study of the random variates $\mathbf{a}^* \hat{\mathbf{C}}_N^k(\rho) \mathbf{b}$ seems to be intractable, quadratic forms of the type $\mathbf{a}^* \hat{\mathbf{S}}_N^k(\rho) \mathbf{b}$ are well-understood objects whose behavior can be studied using standard tools from random matrix theory [45]. It is thus interesting to transfer the study of the fluctuations of $\mathbf{a}^* \hat{\mathbf{C}}_N^k(\rho) \mathbf{b}$ to $\mathbf{a}^* \hat{\mathbf{S}}_N^k(\rho) \mathbf{b}$. Proposition 7 achieves this goal by taking $\epsilon < \frac{1}{2}$. Not only does it entail that $\mathbf{a}^* \hat{\mathbf{C}}_N^k(\rho) \mathbf{b}$ fluctuates at the order of $N^{-\frac{1}{2}}$ (since so does $\mathbf{a}^* \hat{\mathbf{S}}_N^k(\rho) \mathbf{b}$) but also it allows one to prove that $\mathbf{a}^* \hat{\mathbf{C}}_N^k(\rho) \mathbf{b}$ and $\mathbf{a}^* \hat{\mathbf{S}}_N^k(\rho) \mathbf{b}$ exhibit asymptotically the same fluctuations. Similar to [31], our concern will be rather focused on the case $k=-1$. In the next section, we will show how this result can be exploited in order to derive the receiver operating characteristic (ROC) of the ANMF-RTE detector.

B. Optimal design of the ANMF-RTE detector

As explained above, in order to allow for an optimal design of the ANMF-RTE detector, one needs to characterize the distribution of $\widehat{T}_N^{\mathrm{RTE}}(\rho)$ under hypotheses H_0 and H_1. Using Proposition 7, we know that the statistic $\widehat{T}_N^{\mathrm{RTE}}(\rho)$ which cannot be handled directly, has the same fluctuations as $\widetilde{T}_N^{\mathrm{RTE}}(\rho)$ obtained by replacing $\hat{\mathbf{C}}_N(\rho)$ by $\hat{\mathbf{S}}_N(\rho)$. That is:

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

where $\hat{\mathbf{S}}_N(\rho)$ is given by (10).

Let $\tilde{\rho}=\rho\left(\rho+\frac{1}{\gamma_N(\rho)}\frac{1-\rho}{1-(1-\rho)c}\right)^{-1}$. Then, $\hat{\mathbf{S}}_N(\rho)=\rho\tilde{\rho}^{-1}\widehat{\mathbf{R}}_N(\tilde{\rho})$, where, with a slight abuse of notation, we denote by $\widehat{\mathbf{R}}_N(\rho)$ the matrix $(1-\rho)\frac{1}{n}\sum_{i=1}^n\mathbf{z}_i\mathbf{z}_i^*+\rho\mathbf{I}_N$. Since $\widetilde{T}_N^{\mathrm{RTE}}(\rho)$ remains unchanged after scaling of $\hat{\mathbf{S}}_N(\rho)$ and $\mathbf{y}$, we also have:

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

where $\tau=1$ under H_0. It turns out that, conditionally to τ, the fluctuations of the robust statistic $\widehat{T}_N^{\mathrm{RTE}}(\rho)$ under H_0 or H_1 are the same as those obtained in Theorem 1 and Theorem 3 once a is replaced by $\frac{a}{\sqrt{\tau}}$ and ρ by $\tilde{\rho}$[4]. As a consequence, we have the following results:

Theorem 8 (False alarm probability, [31])**.** *As $N,n\to\infty$ with $c_N\to c\in(0,\infty)$,*

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

where $\rho\mapsto\tilde{\rho}$ is the aforementioned mapping and

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

with $\mathbf{Q}_N(\tilde{\rho})\triangleq(\mathbf{I}_N+(1-\tilde{\rho})m(-\tilde{\rho})\mathbf{C}_N)^{-1}$.

Theorem 9 (Detection probability)**.** *As $N,n\to\infty$ with $c_N\to c\in(0,\infty)$,*

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

[4]Note that vector $\mathbf{y}$ can be assumed to be Gaussian without impacting the asymptotic distributions of $\sqrt{N}\hat{T}_N^{RTE}$ under H_0 and H_1.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2017.2766538, IEEE Transactions on Aerospace and Electronic Systems

where the expectation is taken over the distribution of τ, $\sigma_{N,\mathrm{RTE}}(\rho)$ has the same expression as in Theorem 8 and

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

and Q_1 is the Marcum Q-function.

Proof. Since the fluctuations of the robust statistic $\widehat{T}_N^{RTE}(\rho)$ is the same as that of $\widehat{T}_N^{RSCM}(\tilde{\rho})$ when a is replaced by $\frac{a}{\sqrt{\tau}}$, we have for any fixed τ,

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

The result thus follows by noticing the following inequality

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

and resorting to the dominated convergence theorem. □

Let $f_{\mathrm{RTE}}(\rho)$ be given by:

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

The detection probability is then maximized for ρ being set to:

$$\rho \in \underset{\rho\in\mathcal{R}_\kappa}{\overset{\mathrm{RTE}}{\operatorname{argmax}}}\left\{f_{\mathrm{RTE}}\right\} \tag{11}$$

Let ρ_{RTE}^* be among the values satisfying (11). The maximal asymptotic detection probability that can be obtained while keeping the asymptotic false alarm probability equal to η is thus given by:

$$P_{d,\mathrm{RTE}} = \mathbb{E}\left[Q_1\left(\frac{\sqrt{2}a}{\sqrt{\tau}}\sqrt{f_{RTE}(\rho_{RTE}^*)}, \frac{r_{RTE}^*}{\sigma_{N,\mathrm{RTE}}(\rho_{RTE}^*)}\right)\right], \tag{12}$$

where $r_{\mathrm{RTE}}^* = \sigma_{N,\mathrm{RTE}}(\rho_{RTE}^*)\sqrt{-2\log\eta}$. Similar to the Gaussian case, we need to build consistent estimates for $\sigma_{N,\mathrm{RTE}}^2(\rho)$ and $f_{\mathrm{RTE}}(\rho)$ given by: A consistent estimate for $\sigma_{N,\mathrm{RTE}}^2(\rho)$ was provided in [31]:

Proposition 10 (Proposition 1 in [31])**.** *For $\rho \in \left(\max(\left\{0, 1-c_N^{-1}\right\}, 1\right)$. Define,*

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

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2017.2766538, IEEE Transactions on Aerospace and Electronic Systems

and let $\hat{\sigma}^2_{N,\mathrm{RTE}}(1) \triangleq \lim_{\rho\uparrow 1} \hat{\sigma}^2_N(\rho)$. *Then, we have as* $N, n \to \infty$ *with* $c_N \to c$*:*

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

Similar to the Gaussian clutter case, acquiring a consistent estimate for $f_{\mathrm{RTE}}(\rho)$ is mandatory for our design. We thus prove the following Proposition:

Proposition 11. *For* $\rho \in \left(\max\left\{0, 1-c_N^{-1}\right\}, 1\right)$, *let*

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

and $\hat{f}_{\mathrm{RTE}} \triangleq \lim_{\rho\uparrow 1} \hat{f}_{\mathrm{RTE}}(\rho)$. *Then, we have, as* $N, n \to \infty$ *with* $c_N \to c$*:*

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

Proof. The proof follows by first replacing $\widehat{\mathbf{R}}_N^{-1}(\rho)$ by $\widehat{\mathbf{R}}_N^{-1}(\tilde{\rho})$ and ρ by $\tilde{\rho}$ in the results of Proposition 4 and using the convergences [31]:

$$\sup_{\rho\in\mathcal{R}_\kappa^{RTE}} \left\|\frac{\hat{\mathbf{C}}_N(\rho)}{\frac{1}{N}\operatorname{tr}\hat{\mathbf{C}}_N(\rho)} - \widehat{\mathbf{R}}_N(\tilde{\rho})\right\| \xrightarrow{\text{a.s.}} 0.$$

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

and

$$\left|\frac{\rho}{\tilde{\rho}} - \frac{1}{N}\operatorname{tr}\hat{\mathbf{C}}_N(\rho)\right| \xrightarrow{\text{a.s.}} 0.$$

□

Since the results in Proposition 11 and Theorem 9 are uniform in ρ, we have the following corollary:

Corollary 12. *Let* $\hat{f}_{\mathrm{RTE}}(\rho)$ *be defined as in Proposition 4. Define* $\hat{\rho}_N^*$ *as any value satisfying:*

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

Then, for every $r > 0$, *as* $N, n \to \infty$ *with* $c_N \to c$*:*

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

Using the same reasoning as the one followed in the Gaussian clutter case, we propose the following design strategy:

- First, set the regularization parameter to one of the values maximizing $\hat{f}_{RTE}(\rho)$:

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

- Second, set the threshold to $\hat{r}$

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

where η is the required FAP.

Similar to the design strategy for Gaussian settings, the above design strategy is asymptotically optimal in the sense that it should guarantee for N and n large enough close-to-optimal detection performances.

V. Numerical results

In this section, we present some simulations in order to validate our results and to discuss some interesting implications. When non-Gaussian clutters are considered, we assume that $\mathbf{x}_i$ are drawn from a K- distribution with zero mean, covariance $\mathbf{C}_N$ and shape ν.

A. False alarm level

From a practical standpoint, acquiring closed-form approximations for the FAP is of paramount importance, since it enables the designer to adjust the threshold to the value that keeps the false alarm probability below a constant. Most of the previous works have failed so far to address this issue, mainly because of the difficulty of acquiring good approximations for the FAP. The only work that attempts to handle the threshold setting for the ANMF-RTE is due to Ollila *et al.* [18], in which the authors propose to consider the classical expression of the false alarm probability obtained when $\hat{\mathbf{C}}_N$ coincides with the true covariance matrix:

$$\hat{r}_{\text{ollila,approx}} = 1-\eta^{\frac{1}{N-1}},$$

with η being the target false alarm rate. Obviously, this is a coarse approximation and leads to high errors, especially in high dimensional settings. In order to illustrate the error caused by the approximation of [18], we provide in Tables I and II , the empirical false alarm probabilities averaged over $100\,000$ realizations of the statistic $\widehat{T}_N^{\text{RTE}}(\rho)$ when ρ is set to the same value proposed by Ollila and given by [18, Equation (19)] :

$$\hat{\rho}_{\text{ollila}} = \frac{N\,\text{tr}\left(\widehat{\mathbf{C}}_0\right)-1}{N\,\text{tr}\left(\widehat{\mathbf{C}}_0\right)-1+n(N+1)\left(N^{-1}\,\text{tr}\left(\widehat{\mathbf{C}}_0^{-2}\right)-1\right)} \tag{13}$$

with $\widehat{\mathbf{C}}_0$ being the conventional Tyler estimator, and the threshold is adjusted to either $\hat{r}_{\text{ollila,approx}}$ or to the proposed setting given by:

$$\hat{r}_{\text{ollila}} = \hat{\sigma}_{N,RTE}(\hat{\rho}_{\text{ollila}})\sqrt{-2\log\eta}.$$

with $\eta = 0.01$ being the target false alarm. The results of this experiment were obtained when the array steering vector $\mathbf{p}$ is the vector of all ones and matrix $\mathbf{C}_N$ given by:

$$[\mathbf{C}_N]_{i,j} = \begin{cases} b^{j-i} & i \leq j \\ \left(b^{i-j}\right)^* & i > j \end{cases}, \qquad |b| \in (0,1)\,, \tag{14}$$

with $b = 0.5$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2017.2766538, IEEE Transactions on Aerospace and Electronic Systems

The numerical findings of Tables I and II show that the threshold setting proposed by [18] is highly inaccurate tending to largely overestimate the FAP, while the proposed setting always ensures the FAP to be close or, in the worst case, below the target FAP.

Finally, it is worth mentioning that although the proposed setting of the regularization parameter does not have an explicit closed-form expression, it has a slightly higher complexity than the one proposed by Ollila. The reason is that, while both settings require solving the fixed point equations involved by the RTE or the Tyler estimator, our proposed setting additionally requires to re-evaluate the RTE at every iteration of the numerical algorithm maximizing $\hat{f}_{RTE}(\rho)$.

N	n	Empirical FAP (Ollila's threshold setting)	Empirical FAP (Proposed threshold setting)
8	16	0.0723	0.0001
16	32	0.0831	0.0026
32	64	0.0857	0.0055
64	128	0.0901	0.0075
128	256	0.0903	0.0082
256	512	0.0908	0.0092

TABLE I: Empirical FAP obtained by Ollila's threshold setting and the proposed setting for a target $FAP = 0.01$, $n = 2N$, $\mathbf{p} = [1, \cdots, 1]^T$, $b = 0.5$

N	n	Empirical FAP (Ollila's threshold setting)	Empirical FAP (Proposed threshold setting)
8	8	0.1270	0.0015
16	16	0.1494	0.0030
32	32	0.1570	0.0052
64	64	0.1623	0.0078
128	128	0.1466	0.0087
256	256	0.1644	0.0097
512	512	0.1646	0.0097

TABLE II: Empirical FAP obtained by Ollila's threshold setting and the proposed setting for a target $FAP = 0.01$, $n = N$, $\mathbf{p} = [1, \cdots, 1]^T$, $b = 0.5$

B. ROC curves for Gaussian and non-Gaussian clutters

From this point onwards, we assume that the array steering vector is given by:

$$\mathbf{p} = \mathbf{a}(f_d) \otimes \mathbf{a}(f_s)$$

where f_d and f_s denote respectively the normalized target Doppler frequency and the target spatial frequency [46], and $\mathbf{a}(f_d) \in \mathbb{C}^{N_p}$ and $\mathbf{a}(f_s) \in \mathbb{C}^{N_a}$ are the temporel and spatial array steering vectors given by:

$$\left[\mathbf{a}_{N_p}(f_d)\right]_k = \exp\left(\jmath 2\pi(k-1) f_d\right), k = 1, \cdots, N_p$$

$$\left[\mathbf{a}_{N_a}(f_s)\right]_\ell = \exp\left(\jmath 2\pi(\ell-1) f_s\right), \ell = 1, \cdots, N_a.$$

with N_p and N_a being the number of pulses and that of sensors. Then $N = N_p N_a$. We assume that the clutter covariance matrix $\mathbf{C}_N$ is given by:

$$\mathbf{C}_N = \alpha \left(\mathbf{I}_N + \sum_{i=1}^{N_\phi} \sigma^2 \mathbf{A}_{N_p}(f_{d_i}, f_{s_i}) \mathbf{A}_{N_p}^H(f_{d_i}, f_{s_i}) \right)$$

where N_ϕ represents the total number of scatterers, f_{d_i} and f_{s_i} their corresponding doppler and spatial frequencies and α is a noramlizing factor and $\mathbf{A}_{N_p}(f_{d_i}, f_{s_i}) = \mathbf{a}_{N_p}(f_{d_i}) \otimes \mathbf{a}_{N_a}(f_{s_i})$. In this section, we carry out Monte Carlo simulations using $50\,000$ runs in order to represent the Receiver Operating Characteristics (ROC). For each run, we generate the primary signal $\mathbf{y} = \alpha \mathbf{p} + \mathbf{x}$ and the secondary data $\mathbf{x}_1, \cdots, \mathbf{x}_n$. This experiment is carried out for both Gaussian and non-Gaussian clutters when $N_a = 4$, $N_p = 32$, $f_s = 0.5$ $f_d = 0.2$ with n taking possibly the values 64 or 128. For Gaussian clutters, we compare the detection performances of the proposed detector based on our setting with that of those proposed by Chen in [25] and corresponding to the following settings of the regularization parameter:

$$\hat{\rho}_{\text{chen}} = \min \left(\frac{\frac{n-2}{n} \operatorname{tr}\left(\widehat{\mathbf{R}}_N^2\right) + \operatorname{tr}^2\left(\widehat{\mathbf{R}}_N\right)}{(n+2)\left[\operatorname{tr}\left(\widehat{\mathbf{R}}_N^2\right) - \frac{\operatorname{tr}^2\left(\widehat{\mathbf{R}}_N\right)}{N}\right]}, 1 \right),$$

$$\hat{\rho}_{\text{chen,OAS}} = \min \left(\frac{1 - \frac{2}{N} \operatorname{tr}\left(\widehat{\mathbf{R}}_N^2\right) + \operatorname{tr}^2\left(\widehat{\mathbf{R}}_N\right)}{\frac{n-1}{N}\left[\operatorname{tr}\left(\widehat{\mathbf{R}}_N^2\right) - \frac{\operatorname{tr}^2\left(\widehat{\mathbf{R}}_N\right)}{N}\right]}, 1 \right)$$

For each detector, we set the threshold to the value that achieves the required asymptotic false alarm probability and compute the detection probability by evaluating the frequency of realizations over which the corresponding statistic is greater than the threshold. This experiment is carried out for $a = 0.35$ and $a = 0.45$. Fig. 1 and Fig. 2 illustrate the obtained results. We observe from these figures the superiority of the proposed regularization setting to both settings of [25], and this for all the considered configurations. We also note that the empirical detection probability matches that expected by theory and given by (7).

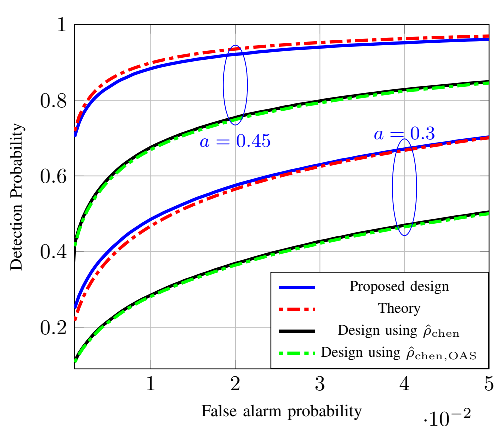

Fig. 1: ROC curves for Gaussian clutters when $N = 128$ ($N_a = 4$, $N_p = 32$), $n = 64$, $f_d = 0.2$, $f_s = 0.5$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2017.2766538, IEEE Transactions on Aerospace and Electronic Systems

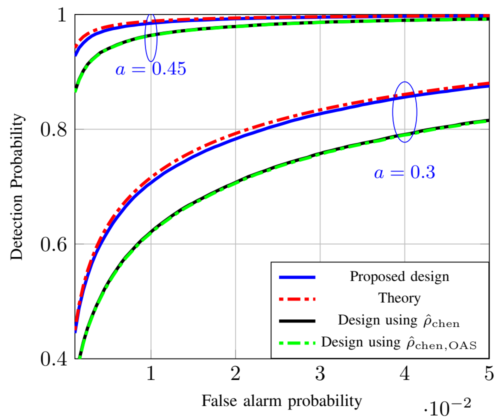

Fig. 2: ROC curves for Gaussian clutters when $N = 128\,(N_a = 4, N_p = 32)$, $n = 128$, $f_d = 0.2$, $f_s = 0.5$.

The analogue experiment is carried out in the case of non-Gaussian clutters following K-distribution with zero mean and shape $\nu = 4.5$ [18] when $N = 128$ and n taking the values 64 and 128. We compare in this case the detection performances with those obtained by the setting of the regularization parameter of Ollila described in (13). Fig. 3 and Fig. 4 illustrate the obtained result. Again, we note the superiority of the proposed detector for all considered configurations, despite the appearing discrepancy between empirical and theoretical results. This discrepancy is due to the number of antennas and of samples being too low for the asymptotic results to be accurate. In order to validate the provided theoretical results, we represent in Fig. 5 the probability of detection of the proposed detector for higher values of N and n, particularly, $N = 320$ and $n = 320$. Clearly, this figure illustrates the improvement in accuracy of our results.

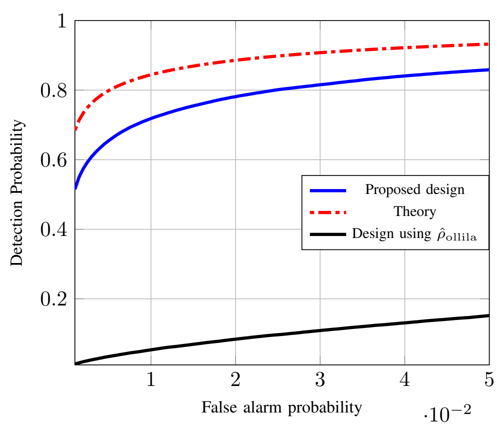

Fig. 3: ROC curves for non Gaussian clutters when $N = 128\,(N_a = 4, N_p = 32)$, $n = 128$, $f_s = 0.5$, $f_d = 0.2$, $a = 0.3$

C. Impact of the distribution shape ν

In a last experiment, we investigate the interplay between a and the distribution shape ν. To this end, we assess the detection probability of the proposed ANMF-RTE with respect to a when operating over K-distributed clutters with different shape values ν. Fig. 6 illustrates the obtained results and when the false alarm probability is fixed to

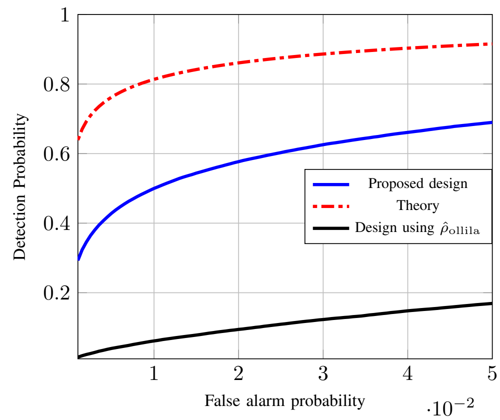

Fig. 4: ROC curves for non Gaussian clutters when $N = 128$ ($N_a = 4$, $N_p = 32$), $n = 64$, $f_s = 0.5$, $f_d = 0.2$, $a = 0.3$

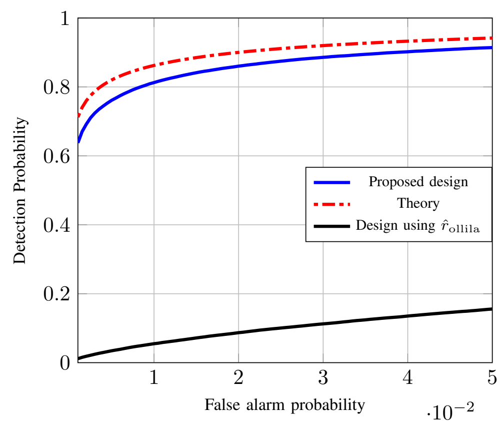

Fig. 5: ROC curves for non-Gaussian clutters when $N = 320$ ($N_a = 10$, $N_p = 32$), $n = 320$, $f_s = 0.5$, $f_d = 0.2$, $a = 0.3$

10^{-3}. We note that for small values of a, higher detection probabilities are achieved when the distribution of the clutter is heavy-tailed (small ν), whereas the opposite occurs for large values of a. In order to explain this change in behavior, we must recall that heavy-tailed clutters (small ν) are characterized by a higher number of occurrences of τ in the vicinity of zero and at the same time by more frequent realizations of large values of τ. If a is small, the improvement in detection performances achieved by heavy-tailed clutters is attributed to the artificial increase in SNR over realizations of small values of τ. As a increases, the power of the signal of interest is high enough so that the effect of realizations with large values of τ becomes dominant. The latter, which are more frequent for small values of ν, are characterized by high levels of noises, thereby entailing a degradation of detection performances.

VI. Conclusion

In this paper, we address the setting of the regularization parameter when the RSCM or the RTE are used in the ANMF detector statistic as a replacement of the unknown covariance matrix, thereby yielding the schemes ANMF-RSCM and ANMF-RTE. One major bottleneck, toward determining the regularization parameter that optimizes the performances of the ANMF detector, is linked to the difficulty to clearly characterize the distribution of the ANMF statistics under the cases of presence or absence of a signal of interest (H_1 and H_0). In order to deal with this issue,

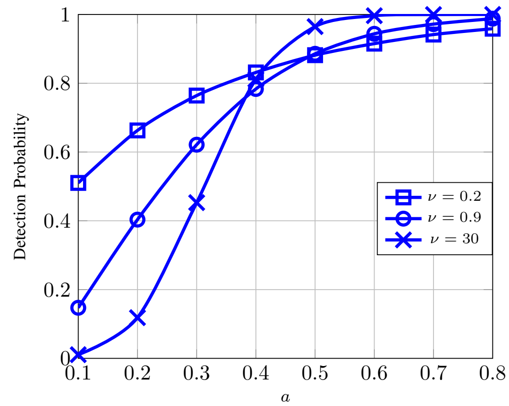

Fig. 6: Detection probability with respect to a, for Non-Gaussian clutters when $N = 128$ ($N_a = 4, N_p = 32$), $n = 128$, $f_s = 0.5$ and $f_d = 0.2$.

we considered the regime under which the number of samples and their dimensions grow large simultaneously. Based on tools from random matrix theory along with recent asymptotic results on the behaviour of the RTE, we derived the asymptotic distribution of the ANMF detector under hypothesis H_0 and H_1. The obtained results have allowed us to propose an (asymptotically) optimal design of the regularization parameter that maximizes the detection probability while keeping fixed the false alarm probability through an appropriate tuning of the threshold value. Simulations results clearly illustrated the gain of our method over previously proposed empirical settings of the regularization coefficient. As a notable outcome of our work, our approach can be applied to the analysis of other involved radar detection schemes using regularized robust-covariance estimates, providing an efficient way to properly set the regularization parameter. One major advantage of our approach is that, contrary to first intuitions, it leads to simple closed-form expressions that can be easily implemented in practice. This may seem on the onset quite surprising given that the handling of the classical regime where n grows to infinity with N fixed has been shown to be delicate. As a matter of fact, it has thus far been considered only for the non-regularized Tyler estimator where intricate expressions in the form of integrals were obtained [42]. Building the bridge between both approaches is an open question that deserves investigation.

Appendix A
Proof of Theorem 3

The proof of Theorem 3 consists of two steps. First, we study the asymptotic behaviour of the detection probability for fixed ρ. Then, by a similar argument to the one considered in [31], we establish the uniformity of the result over the considered set of ρ. Assume that the received signal vector $\mathbf{y}$ is given by:

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

with $\|\mathbf{p}\|^2 = N$ and let us write $\sqrt{N}T_N^{RSCM}(\rho)$ as:

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

A close inspection of the expression of $\sqrt{N}\widehat{T}_N^{RSCM}(\rho)$ reveals that the fluctuations will be governed by the numerator $\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\frac{\mathbf{p}}{\sqrt{N}}$ since, from classical results of random matrix theory, we know that quantities in the denominator exhibit a deterministic behaviour, being well-approximated by some deterministic quantities. In effect,

$$\frac{1}{N}\mathbf{p}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{p}-\frac{1}{\rho N}\mathbf{p}^*\mathbf{Q}_N(\rho)\mathbf{p}\xrightarrow{\text{a.s.}}0, \tag{15}$$

while:

$$\frac{1}{N}\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{y}-\frac{1}{N\rho}\operatorname{tr}\mathbf{C}_N\mathbf{Q}_N(\rho)\xrightarrow{\text{a.s.}}0. \tag{16}$$

The first convergence (15) follows from Theorem 1.1 of [47] whereas the second one is obtained by observing that, because of the low-SNR hypothesis:

$$\frac{1}{N}\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{y}-\frac{1}{N}\mathbf{x}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{x}\xrightarrow{\text{a.s.}}0$$

and then using the well-known convergence result [48]:

$$\frac{1}{N}\mathbf{x}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{x}-\frac{1}{N\rho}\operatorname{tr}\mathbf{C}_N\mathbf{Q}_N(\rho)\xrightarrow{\text{a.s.}}0.$$

We will now deal with the fluctuations of the numerator. We have:

$$\sqrt{N}\left|\frac{1}{N}\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}\mathbf{p}\right|=\left|\frac{a\mathbf{p}^*}{N}\widehat{\mathbf{R}}_N^{-1}\mathbf{p}+\mathbf{x}^*\widehat{\mathbf{R}}_N^{-1}\frac{\mathbf{p}}{\sqrt{N}}\right|.$$

Arguing in a similar way to that in (15), we know that the quantity $\frac{1}{N}a\mathbf{p}^*\widehat{\mathbf{R}}_N^{-1}\mathbf{p}$ does not fluctuate and converges to:

$$\frac{1}{N}a\mathbf{p}^*\widehat{\mathbf{R}}_N^{-1}\mathbf{p}-\frac{a}{N\rho}\mathbf{p}^*\mathbf{Q}_N(\rho)\mathbf{p}\xrightarrow{\text{a.s.}}0,$$

while, from [31]:

$$\left[\frac{1}{\sqrt{N}}\Re(\mathbf{x}^*\widehat{\mathbf{R}}_N^{-1}\mathbf{p}),\frac{1}{\sqrt{N}}\Im(\mathbf{x}^*\widehat{\mathbf{R}}_N^{-1}\mathbf{p})\right]^{\mathrm{T}} -\sqrt{\frac{1}{2\rho^2N}\frac{\mathbf{p}^*\mathbf{C}_N\mathbf{Q}_N^2(\rho)\mathbf{p}}{\left(1-cm(-\rho)^2(1-\rho)^2\frac{1}{N}\operatorname{tr}\mathbf{C}_N^2\mathbf{Q}_N^2(\rho)\right)}}Z'=o_p(1)$$

for some $Z'\sim\mathcal{N}(\mathbf{0},\mathbf{I}_2)$.

Let $\mathbf{r}=\left[\frac{1}{\sqrt{N}}\Re(\mathbf{x}^*\widehat{\mathbf{R}}_N^{-1}\mathbf{p}),\frac{1}{\sqrt{N}}\Im(\mathbf{x}^*\widehat{\mathbf{R}}_N^{-1}\mathbf{p})\right]^{\mathrm{T}}$. Denote by Υ_N and ω_N the quantities:

$$\Upsilon_N=\sqrt{\frac{1}{2\rho^2N}\frac{\mathbf{p}^*\mathbf{C}_N\mathbf{Q}_N^2(\rho)\mathbf{p}}{\left(1-cm(-\rho)^2(1-\rho)^2\frac{1}{N}\operatorname{tr}\mathbf{C}_N^2\mathbf{Q}_N^2(\rho)\right)}}$$

$$\omega_N=\frac{a}{N\rho}\mathbf{p}^*\mathbf{Q}_N(\rho)\mathbf{p}$$

Recall the following distance between probability distributions:

$$\beta\left(\mathbb{P},\tilde{\mathbb{P}}\right)=\sup\left\{\left|\int fd\mathbb{P}-\int fd\tilde{\mathbb{P}}\right|,\|f\|_{BL}\le 1\right\}$$

where $\|f\|_{BL}=\|f\|_{Lip}+\|f\|_\infty$, $\|f\|_{Lip}$ being the Lipschitz norm and $\|.\|_\infty$, the supremum norm [49]. Assume for the moment that $\limsup\Upsilon_N<\infty$. The proof for this statement will be provided later. Then, from Theorem 11.7.1 in [49],

$$\beta\left(\mathcal{L}\left(\mathbf{r}+\begin{bmatrix}\frac{1}{N}a\mathbf{p}^*\widehat{\mathbf{R}}_N^{-1}\mathbf{p}\\0\end{bmatrix}\right),\mathcal{L}\left(\Upsilon_NZ'+\begin{bmatrix}\omega_N\\0\end{bmatrix}\right)\right)\to 0.$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2017.2766538, IEEE Transactions on Aerospace and Electronic Systems

where $\mathcal{L}(X)$ stands for the probability distribution of X. As $\mathbb{R}^2 \to \mathbb{R}, d \mapsto \|d\|$ is Lipschitz, we have for any Lipschitz function g in real variables,

$$\mathbb{E}g\left(\left\|\mathbf{r}+\begin{bmatrix}\frac{1}{N}a\mathbf{p}^*\widehat{\mathbf{R}}_N^{-1}\mathbf{p}\\ 0\end{bmatrix}\right\|\right)-\mathbb{E}g\left(\left\|\Upsilon_N Z^{'}+\begin{bmatrix}\omega_N\\ 0\end{bmatrix}\right\|\right)\to 0.$$

This shows in particular that $\sqrt{N}\left|\frac{1}{N}\mathbf{y}^*\widehat{\mathbf{R}}_N^{-1}\mathbf{p}\right|$ behaves asymptotically as a Rice random variable with location $\frac{a}{N\rho}\mathbf{p}^*\mathbf{Q}_N(\rho)\mathbf{p}$ and scale $\sqrt{\frac{1}{2\rho^2 N}\frac{\mathbf{p}^*\mathbf{C}_N\mathbf{Q}_N^2(\rho)\mathbf{p}}{\left(1-cm^2(-\rho)(1-\rho)^2\frac{1}{N}\operatorname{tr}\mathbf{C}_N^2\mathbf{Q}_N^2(\rho)\right)}}$. Using this result along with Slutsky Lemma, we conclude that under H_1, $\hat{T}_N^{RSCM}(\rho)$ is also asymptotically equivalent to a Rice random variate but with location $\frac{a}{\sqrt{N}}\frac{\sqrt{\mathbf{p}^*\mathbf{Q}_N(\rho)\mathbf{p}}}{\sqrt{\frac{1}{N}\operatorname{tr}\mathbf{C}_N\mathbf{Q}_N(\rho)}}$ and scale $\sigma_{N,SCM}$. We therefore get, for a fixed ρ,

$$\left|\mathbb{P}\left[\widehat{T}_N^{SCM}>\frac{r}{\sqrt{N}}|H_1\right]-Q_1\left(g_{SCM}(\mathbf{p}),\frac{r}{\sigma_{N,SCM}}\right)\right|\xrightarrow{\text{a.s.}}0$$

The generalization of this result to uniform convergence across $\rho\in\mathcal{R}_\kappa^{SCM}$ can be derived along the same steps as in [31]. We now provide details about the control of the $\limsup\Upsilon_N$. For that, one needs to check that:

$$\liminf 1-cm^2(-\rho)(1-\rho)^2\frac{1}{N}\operatorname{tr}\mathbf{C}_N^2\mathbf{Q}_N^2(\rho)>0.$$

The proof hinges on the observation that this term naturally appears when computing the derivative of $m(z)$ with respect to z at $z=-\rho$. Simple calculations reveal that:

$$m'(z)=\left(-z+\frac{c(1-\rho)}{N}\operatorname{tr}\mathbf{C}_N\mathbf{Q}_N(z)\right)^{-2}$$
$$\times\left(1-cm^2(z)(1-\rho)^2\frac{1}{N}\operatorname{tr}\mathbf{C}_N^2\mathbf{Q}_N^2(z)\right)^{-1}.$$

It suffices thus to show that $m'(-\rho)$ is bounded. As m is a Stieltjes transform of some positive probability measure μ, it can be written as:

$$m'(-\rho)=\int\frac{\mu(dx)}{(x+\rho)^2}\le\frac{1}{\kappa^2}$$

which ends the proof.

Appendix B
Proof of proposition 4

For ease of notation, we denote by $f(\rho)$ and $\hat{f}(\rho)$, the quantities $f_{SCM}(\rho)$ and $\hat{f}_{SCM}(\rho)$. It is easy to see that $\hat{f}(\rho)$ and $f(\rho)$ converges to an undetermined form as $\rho\uparrow 1$. Set $\hat{f}(\rho)\triangleq\frac{\hat{h}(\rho)}{\hat{g}(\rho)}$ with $\hat{g}$ and $\hat{h}$ being given by:

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

The handling of the values of ρ approaching 1 can be performed using the l'Hopital's rule.

The idea of the proof is to treat seperately the values of ρ in the interval $[\kappa,1-\kappa]$ and those in $[1-\kappa,1]$ for some κ small enough. In order to allow for a setting of κ that is independent from N, we need to prove that:

$$\lim_{\rho\uparrow 1}\limsup_N\left|\hat{f}(\rho)-\frac{h_N^{'}(1)}{g_N^{'}(1)}\right|=0. \tag{17}$$

To this end, a uniform variant of the l'Hopital's rule is essential. This variant is stated in the following Lemma:

Lemma 13. *Let* $f_N(\rho)=\frac{h_N(\rho)}{g_N(\rho)}$ *with* h_N *and* g_N *being defined in the interval* $\rho\in[0,1]$. *Assume that* $h_N(1)=g_N(1)=0$ *while* $\liminf_N \frac{dg_N}{d\rho}\Big|_{\rho=1}>0$, $\limsup_N \frac{dg_N}{d\rho}\Big|_{\rho=1}<+\infty$ *and* $\limsup_N \frac{dh_N}{d\rho}\Big|_{\rho=1}<+\infty$. *Assume also that the second derivatives of* h_N *and* g_N *are uniformly bounded in* N, *that is:*

$$\sup_{\rho\in[0,1]}\limsup_N \left|h_N^{''}(\rho)\right|<+\infty$$

$$\sup_{\rho\in[0,1]}\limsup_N \left|g_N^{''}(\rho)\right|<+\infty$$

Then,

$$\lim_{\rho\to 1}\limsup_N \left|\frac{h_N(\rho)}{g_N(\rho)}-\frac{h_N^{'}(1)}{g_N^{'}(1)}\right|\to 0.$$

Proof. The proof relies on the Tylor expansion of h_N and g_N in the vicinity of 1, which asserts that for any $\rho\in[0,1]$ there exist ξ_1 and ξ_2 satisfying:

$$h_N(\rho)=h_N^{'}(1)(\rho-1)+(\rho-1)^2h_N^{''}(\xi_1)$$

$$g_N(\rho)=g_N^{'}(1)(\rho-1)+(\rho-1)^2g_N^{''}(\xi_2)$$

We therefore have,

$$\begin{aligned}&\limsup_N \left|\frac{h_N(\rho)}{g_N(\rho)}-\frac{h_N^{'}(1)}{g_N^{'}(1)}\right|\\ &=\limsup_N \left|\frac{h_N^{'}(1)+(\rho-1)h_N^{''}(\xi_1)}{g_N^{'}(1)+(\rho-1)g_N^{''}(\xi_2)}-\frac{h_N^{'}(1)}{g_N^{'}(1)}\right|\\ &=\left|\frac{(\rho-1)h_N^{''}(\xi_1)g_N^{'}(1)-(\rho-1)h_N^{'}(1)g_N^{''}(\xi_2)}{g_N^{'}(1)\left(g_N^{'}(1)+(\rho-1)g_N^{''}(\xi_2)\right)}\right|\\ &\leq|\rho-1|\frac{\limsup_N |h_N^{''}(\xi_1)g_N^{'}(1)|+\limsup_N |h_N^{'}(\xi_1)g_N^{''}(\xi_2)|}{\liminf\left|g_N^{'}(1)\right|^2}\end{aligned}$$

Tending ρ to 1 establishes the desired result. □

Obviously functions $\hat{h}(\rho)$ and $\hat{g}(\rho)$ satisfiy the assumptions of Lemma 13. Applying l'Hopital's rule and using the differentiation rules $\frac{d}{d\rho}\widehat{\mathbf{R}}_N^{-1}(\rho)=-\widehat{\mathbf{R}}_N^{-2}(\rho)\left(-\widehat{\mathbf{R}}_N+\mathbf{I}\right)$ and $\frac{d}{d\rho}\widehat{\mathbf{R}}_N^{-2}(\rho)=-2\widehat{\mathbf{R}}_N^{-3}(\rho)\left(-\widehat{\mathbf{R}}_N+\mathbf{I}\right)$, we finally prove:

$$\lim_{\rho\uparrow 1}\limsup_N \left|\hat{f}(\rho)-\frac{N}{\mathbf{p}^*\widehat{\mathbf{R}}_N\mathbf{p}}\right|=0. \tag{18}$$

Now, using the fact $\frac{1}{N}\mathbf{p}^*\widehat{\mathbf{R}}_N\mathbf{p}-\frac{1}{N}\mathbf{p}^*\mathbf{C}_N\mathbf{p}\xrightarrow{\text{a.s.}}0$ in conjunction to the last item in Assumption 1, we get:

$$\lim_{\rho\uparrow 1}\limsup_N \left|\hat{f}(\rho)-\frac{N}{\mathbf{p}^*\mathbf{C}_N\mathbf{p}}\right|\xrightarrow{\text{a.s.}}0. \tag{19}$$

On the other hand, a careful analysis of the behaviour of $f(\rho)$ near 1 reveals similarly that:

$$\lim_{\rho\uparrow 1}\limsup_N \left|f(\rho)-\frac{N}{\mathbf{p}^*\mathbf{C}_N\mathbf{p}}\right|\to 0 \tag{20}$$

Combining (19) with (20), we finally obtain:

$$\lim_{\rho\uparrow 1}\limsup_{N}\left|\hat{f}(\rho)-f(\rho)\right|\to 0$$

It then suffices to prove Proposition 4 on $\mathcal{R}_\kappa \triangleq [\kappa, 1-\ell]$. To this end, we need to recall the following relations satisfied by $m_N(-\rho)$: $m_N(-\rho)=\frac{1-c}{\rho}+\frac{c}{\rho}\frac{1}{N}\operatorname{tr}\mathbf{Q}_N(\rho)$ and $m_N(-\rho)=\left(\rho+c(1-\rho)\frac{1}{N}\operatorname{tr}\mathbf{C}_N\mathbf{Q}_N(\rho)\right)^{-1}$ Combining these relations, we therefore get:

$$\frac{1}{N}\operatorname{tr}\mathbf{C}_N\mathbf{Q}_N(\rho)=\frac{\rho\left(1-\frac{1}{N}\operatorname{tr}\mathbf{Q}_N(\rho)\right)}{(1-c)(1-\rho)(1-c+\frac{c}{N}\operatorname{tr}\mathbf{Q}_N(\rho))}$$

The result thus follows by using Proposition 2 and noticing, in the same way as in [31], that:

$$\sup_{\rho\in[\kappa,1-\ell]}\left|\frac{1}{N}\operatorname{tr}\mathbf{Q}_N-\frac{1}{\rho}\operatorname{tr}\widehat{\mathbf{R}}_N^{-1}(\rho)\right|\xrightarrow{\text{a.s.}}0,$$

$$\sup_{\rho\in[\kappa,1-\ell]}\left|\frac{\mathbf{p}^*}{\sqrt{N}}\mathbf{Q}_N\mathbf{p}-\frac{1}{\rho\sqrt{N}}\mathbf{p}^*\widehat{\mathbf{R}}_N^{-1}(\rho)\mathbf{p}\right|\xrightarrow{\text{a.s.}}0$$

and

$$\sup_{\rho\in[\kappa,1-\ell]}\left|\frac{\frac{1}{N}\mathbf{p}^*\mathbf{C}_N\mathbf{Q}_N^2(\rho)\mathbf{p}}{1-cm^2(-\rho)(1-\rho)^2\frac{1}{N}\operatorname{tr}\mathbf{C}_N^2\mathbf{Q}_N^2(\rho)}\right. \tag{21}$$

$$\left.-\frac{\frac{1}{N}\left(\mathbf{p}^*\widehat{\mathbf{R}}_N^{-1}\mathbf{p}-\rho\mathbf{p}^*\widehat{\mathbf{R}}_N^{-2}\mathbf{p}\right)}{(1-\rho)\left(\frac{1-\rho}{c}+c\frac{1}{N}\operatorname{tr}\widehat{\mathbf{R}}_N^{-1}(\rho)\right)}\right|\xrightarrow{\text{a.s.}}0. \tag{22}$$

References

[1] J. Ward, "Space-time Adaptive Processing for Airborne Radar," Tech. Rep., Lincoln Lab, MIT, Dec. 1994.

[2] R. Klemm, *Principles of Space-Time Adaptive Processing*, IET, 2002.

[3] E. Olilla, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex Elliptically Symmetric Distributions: Survey, New Results and Applications," *IEEE Trans. Signal Process.*, vol. 60, no. 11, Nov. 2012.

[4] M. Mahot, F. Pascal, P. Forster, and J. P. Ovalez, "Asymptotic Properties of Robust Complex Covariance Matrix Estimates," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3348–3356, July 2013.

[5] M. Steiner and K. Gerlach, "Fast converging adaptive processor or a structured covariance matrix," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 4, pp. 1115–1126, October 2000.

[6] A. Aubry, A. De Maio, L. Pallotta, and A. Farina, "Maximum Likelihood Estimation of a Structured Covariance Matrix with a Condition Number Constraint," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 3004–3020, June 2012.

[7] B. Kang, V. Monga, and M. Rangaswamy, "Rank-Constrained Maximum Likelihood Estimation of Structured Covariance Matrices," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 1, pp. 501–516, Jan. 2014.

[8] Y. Abramovich, "Controlled Method for Adaptive Optimization of Filters Using the Criterion of Maximum SNR," *Radio Eng. Electron. Phys*, vol. 26, no. 3, pp. 87–95, 1981.

[9] B. D. Carlson, "Covariance Matrix Estimation Errors and Diagonal Loading in Adaptive Arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 397–401, 1988.

[10] D. Kelker, "Distribution Theory of Spherical Distributions and a Location Scale Parameter Generalization," *Sankhyā: The Indian Journal of Statistics, Series A*, vol. 32, no. 4, pp. 419–430, Dec. 1970.

[11] P. J. Huber, "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964.

[12] P. J. Huber, "The 1972 wald lecture robust statistics : A review," *Annals of Mathematical Statistics*, vol. 43, no. 4, pp. 1041–1067, August 1972.

[13] R. A. Maronna, "Robust M-estimators of multivariate location and scatter," *Annals of Statistics*, vol. 4, no. 1, pp. 51–67, January 1976.

[14] F. Pascal, Y. Chitour, J.P. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound Gaussian noise: existence and algorithm analysis," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 34–48, Jan. 2008.

[15] D. E. Tyler, "A Distribution-Free M-estimator of Multivariate Scatter," *The Annals of Statistics*, vol. 15, no. 1, pp. 234–251, 1987.

[16] Y. Chitour and F. Pascal, "Exact maximum likelihood estimates for SIRV covariance matrix: existence and algorithm analysis," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4563–4573, Oct. 2008.

[17] F. Pascal, Y. Chitour, and Y. Quek, "Generalized Robust Shrinkage Estimator and Its Application to STAP Detection Problem ," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5640–5651, 2014.

[18] E. Ollila and D. E. Tyler, "Regularized M-estimators of Scatter Matrix," *IEEE Trans. Signal Process.*, vol. 62, no. 22, Nov. 2014.

[19] Y. Sun, P. Babu, and D. P. Palomar, "Regularized Tyler's Scatter Estimator: Existence, Uniqueness and Algorithms," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 5143–5156, Oct. 2014.

[20] K D Ward, "Compound representation of high resolution sea clutter," *Electronics Letters*, vol. 17, no. 16, pp. 561–563, August 1981.

[21] S Watts, "Radar detection prediction in sea clutter using the compound k-distribution model," *IEE Proceeding, Part. F*, vol. 132, no. 7, pp. 613–620, December 1985.

[22] T Nohara and S Haykin, "Canada east coast trials and the k-distribution," *IEE Proceeding, Part. F*, vol. 138, no. 2, pp. 82–88, 1991.

[23] J B Billingsley, "Ground clutter measurements for surface-sited radar," Tech. Rep. 780, MIT, February 1993.

[24] O. Ledoit and M. Wolf, "A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices," *Journal of Multivariate Analysis*, vol. 88, no. 2, pp. 365–411, 2004.

[25] Y. Chen, A. Wiesel, Y. C. Eldar, and A. O. Hero, "Shrinkage Algorithms for MMSE Covariance Estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 10, Oct. 2010.

[26] Y. Chen, A. Wiesel, and O. A. Hero, "Robust Shrinkage Estimation of High Dimensional Covariance Matrices," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4907–4107, 2011.

[27] E. Conte, M. Lops, and G. Ricci, "Asymptotically Optimum Radar Detection in Compound-Gaussian Clutter," *IEEE Trans. Aerosp. Electron. Syst.*, pp. 617–625, Apr. 1995.

[28] J. Liu, Z-J. Zhang, Y. Yang, and H. Liu, "A CFAR Adaptive Subspace Detector for First-Order or Second-Order Gaussian Signals Based on a Single Observation," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5126–5140, 2011.

[29] S. Kraut and L.L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2538–2541, 1999.

[30] F Pascal, J-P Ovarlez, P Forster, and P Larzabal, "Constant false alarm rate detection in spherically invariant random processes," in *Proc. of the European Signal Processing Conf., EUSIPCO-04*, Vienna, Sept. 2004, pp. 2143–2146.

[31] R. Couillet, A. Kammoun, and F. Pascal, "Second Order Statistics of Robust Estimators of Scatter. Application to GLRT Detection for Elliptical Signals," *Journal of Multivariate Analysis*, vol. 143, pp. 249–274, Jan. 2016.

[32] R. Couillet and M. McKay, "Large Dimensional Analysis and Optimization of Robust Shrinkage Covariance Matrix Estimators," *Journal of Multivariate Analysis*, vol. 31, pp. 99–120, 2014.

[33] A. Kammoun, R. Couillet, F. Pascal, and M.-S. Alouini, "Optimal Adaptive Normalized Matched Filter for Large Antenna Arrays," in *Proceedings of the IEEE Statistical Signal Processing Workshop*, June 2016.

[34] L. L. Scharf and B. Friedlander, "Matched Subspace Detectors," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2146–2157, 1994.

[35] E. Conte, A. D. Maio, and G. Ricci, "Recursive Estimation of the Covariance Matrix of a Compound-Gaussian Process and Its Application to Adaptive CFAR Detection," *IEEE Trans. Signal Process.*, vol. 50, no. 8, Aug. 2002.

[36] F. Gini and M. Greco, "Covariance Matrix Estimation for CFAR Detection in Correlated Heavy-Tailed Clutter," *IEEE Trans. Signal Process.*, vol. 82, no. 12, pp. 1847–1859, Dec. 2002.

[37] I. Reed, J. Mallett, and L. Brennan, "Rapid Convergence Rate in Adaptive Arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 10, no. 6, pp. 853–863, Nov. 1974.

[38] F. Pascal and P. Forster and J. P. Ovarlez and P. Larzabal, "Performance Analysis of Covariance Matrix Estimate in Impulsive Noise," *IEEE Trans. Signal Process.*, vol. 56, no. 6, 2008.

[39] Y. I. Abramovich and O. Besson, "Regularized Covariance Matrix Estimation in Complex Elliptically Symmetric Distributions Using the Expected Likelihood Approach: Part 1: The Over-Sampled Case," *IEEE Trans. Signal Process.*, vol. 61, no. 23, Oct. 2013.

[40] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear Precoding in MISO broadcast channels with limited feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509–4537, July 2012.

[41] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, 1995.

[42] F. Pascal and J. P. Ovarlez, "Asymptotic Properties of the Robust ANMF," *ICASSP*, 2015.

[43] R. Couillet, F. Pascal, and J. W. Silverstein, "The random matrix regime of Maronna's M-estimator with elliptically distributed samples," *Journal of Multivariate Analysis*, pp. 56–78, July 2015, .

[44] R. Couillet, F. Pascal, and J. W. Silverstein, "Robust Estimates of Covariance Matrices in the Large Dimensional Regime," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7269–7278, Oct. 2014.

[45] A. Kammoun, M. Kharouf, W. Hachem, and J. Najim, "A Central Limit Theorem for the SINR at the LMMSE Estimator Output for Large-Dimensional Signals," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5048–5063, Nov. 2009.

[46] F. Bandiera, D. Orlando, and G. Ricci, *Advanced Radar Detection Schemes under Mismatched Signal Models*, Synthesis Lectures on Signal Processing N, 8, New Yorl: Morgan and Claypool, 2009.

[47] W. Hachem and P. Loubaton and J. Najim and P. Vallet, "On bilinear forms based on the resolvent of large random matrices," *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques*, vol. 49, no. 1, pp. 36–63, 2013.

[48] W. Hachem, Ph. Loubaton, and J. Najim, "Deterministic Equivalents for Certain Functionals of Large Random Matrices," *Annals of Applied Probability*, vol. 17, no. 3, pp. 875–930, 2007.

[49] R. M. Dudley, *Real Analysis and Probability*, Cambridge University Press, 2004.

Abla Kammoun was born in Sfax, Tunisia. She received the engineering degree in signal and systems from the Tunisia Polytechnic School, La Marsa, and the Master's degree and the Ph.D. degree in digital communications from Telecom Paris Tech [then Ecole Nationale Suprieure des Tlcommunications (ENST)]. From June 2010 to April 2012, she has been a Postdoctoral Researcher in the TSI Department, Telecom Paris Tech. Then she has been at Suplec at the Alcatel-Lucent Chair on Flexible Radio until December 2013. Currently, she is a research scientist at KAUST university. Her research interests include performance analysis, random matrix theory, and semi-blind channel estimation. She is the recepient of the SAM 2014 second Prize Best paper award.

Romain Couillet received his MSc in Mobile Communications at the Eurecom Institute and his MSc in Communication Systems in Telecom ParisTech, France in 2007. From 2007 to 2010, he worked with ST-Ericsson as an Algorithm Development Engineer on the Long Term Evolution Advanced project, where he prepared his PhD with Supelec, France, which he graduated in November 2010. He is currently a full professor in the LSS laboratory at CentraleSuplec, France. His research topics are in random matrix theory applied to statistics, machine learning, signal processing, and wireless communications. He is an IEEE Senior Member. In 2015, he received the HDR title from University ParisSud. He is the recipient of the 2013 CNRS Bronze Medal in the section "science of information and its interactions", of the 2013 IEEE ComSoc Outstanding Young Researcher Award (EMEA Region), of the 2011 EEA/GdR ISIS/GRETSI best PhD thesis award, and of the Valuetools 2008 best student paper award.

Frédéric Pascal (S04, M06, SM14)received the Master's degree ("Probabilities, Statistics and Applications: Signal, Image et Networks") with merit, in Applied Statistics from University Paris VII - Jussieu, Paris, France, in 2003. Then, he received the Ph.D. degree of Signal Processing, from University Paris X Nanterre in 2006. From November 2006 to February 2008, he made a post-doctoral position in the Signal Processing and Information team of the laboratory SATIE, CNRS, ENS Cachan, France. Between March 2008 and December 2011 (resp. Jan. 2012 Dec. 2013), he was an Assistant Professor (resp. Associate Professor) in SONDRA at CentraleSupelec.In 2012, he obtaineda Research Directorship Habilitation (HDR) thesis in Signal Processing from the University of Paris- Sud.Between Aug. 2013 and Aug. 2014, he was a Visiting Associate Professor in the ECE department at the National University of Singapore. From January 2014, Frederic Pascal is a full Professor in the L2S laboratory at CentraleSupelec. From Jan. 2017, he is the head of the signals and Statistics group of L2S. From 2015 to 2017, Frederic Pascal was the Chair of the EURASIP SAT in Theoretical and Methodological Trends in Signal Processing (TMTSP) and he is a member of the IEEE Signal Processing Society SAM technical committee (Jan. 2015-present). Frederic Pascal serves as an Associate Editor for the IEEE Transactions on Signal Processing (2015-present) and for the EURASIP Journal on Advances in Signal Processing (2015-present). His research interests contain estimation, detection and classification for statistical signal processing and applications in radar and image processing.

Mohamed-Slim Alouini was born in Tunis, Tunisia. He received the Ph.D. degree in Electrical Engineering from the California Institute of Technology (Caltech), Pasadena, CA, USA, in 1998. He served as a faculty member in the University of Minnesota, Minneapolis, MN, USA, then in the Texas A&M University at Qatar, Education City, Doha, Qatar before joining King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia as a Professor of Electrical Engineering in 2009. His current research interests include the modeling, design, and performance analysis of wireless communication systems.