On the performance of robust plug-in detectors using M-estimators

Gordana Drašković[a], Arnaud Breloy[b], Frédéric Pascal[a]

[a] *L2S - CentraleSupélec - CNRS - Université Paris-Sud, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette Cedex, France*
[b] *LEME (EA 4416), University Paris Nanterre, 50 rue de Sevres, 92410 Ville d'Avray, France*

Abstract

This paper provides an original asymptotic analysis of robust adaptive detectors performance in the context of non-Gaussian observations. We focus on a single-steering case in homogeneous environment and analyze the properties of different adaptive detectors such as Adaptive (Normalized) Matched Filter (AMF/ANMF), Kelly's GLRT, and Rao test when an estimator of the covariance matrix is plugged in. When the noise distribution turns to be non-Gaussian, the detectors relying on the Sample Covariance Matrix (SCM) can perform poorly and an interesting alternative is the use of M-estimators. In this context, we show that, from Complex Elliptically Symmetric (CES) samples, the distribution of a detection statistic built with M-estimators can be accurately approximated by the one of the same statistic built from the SCM of an equivalent Gaussian setting. The loss due to this approximation is theoretically derived and shown to be negligible in most cases. This explicit equivalent statistic is especially interesting since it allows to tune robust plug-in detectors with well established results from the Gaussian detection framework. Furthermore, this approach provides new insights on the robust estimation tools behavior. Finally, some simulations illustrate the interest of the proposed results.

Keywords: Robust detection, M-estimation, CES distributions, Wishart

Email addresses: gordana.draskovic@l2s.centralesupelec.fr (Gordana Drašković), abreloy@parisnanterre.fr (Arnaud Breloy), frederic.pascal@l2s.centralesupelec.fr (Frédéric Pascal)

distribution, performance analysis

1. Introduction

Adaptive detection of signals embedded in a disturbance is a ubiquitous problem in statistical signal processing. Usually, the problem consists in deciding if a signal of known steering vector is present in a tested sample (primary data), while some signal-free samples (secondary data) are available to estimate the statistical parameters of the noise. This problem has been extensively studied in the context of Gaussian distributed noise [1, 2, 3, 4, 5]. Several decision statistics have been proposed, such as the Generalized Likelihood Ratio Test (GLRT) (Kelly's detector) [2], the Adaptive Matched Filter (AMF or 2-step GLRT) [6], its normalized counterpart (Adaptive Normalized Matched Filter-ANMF or Adaptive Cosine Estimator-ACE) [7, 8], and the Rao test [9]. The associated detectors have been characterized in terms of probability of detection (P_d) and probability of false alarm (P_{fa}), as well as constant false alarm rate (CFAR) properties (see e.g., [10] and references therein) and performance in mismatched scenarios [11]. As core component, these statistics involve the Sample Covariance Matrix (SCM) [12] in their construction. Since this estimator of the noise covariance matrix is sensitive to heavy-tailed distributed samples, this family of Gaussian detectors can exhibit poor performance in non-Gaussian environments.

In order to account for non-Gaussian data, an interesting alternative is brought by the framework of complex elliptically symmetric (CES) distributions [13]. This family of distributions generalizes the multivariate Gaussian one by replacing the exponential shape with an arbitrary function, called density generator. This added flexibility allows to encompass a large panel of well-known heavy tailed distributions, such as Student-t, K-, Weibull, and Generalized Gaussian. An important subclass of CES distributions are the compound-Gaussian (CG) distributions, also referred as to Spherically Invariant Random Vectors (SIRV), that have been employed for radar clutter modeling [14, 15, 16]. Detection pro-

cedures assuming CES/CG distributed samples have been proposed and widely studied in the literature. First, detection procedures assuming known parameters for the noise can found in [17, 18, 19] among others. Then, adaptive detectors have been derived, generally based on a 2-step GLRT, for different covariance matrix estimators [20, 21, 22]. More recently, various robust covariance matrix estimators have been used for detection purposes [23, 24, 25, 26, 27]. An overview of recent advances in radar detection, including robust detection approaches, can be found in [28].

However, in practice, the density generator of the true underlying distribution is is unknown. In this case, a *robust plug-in detector* can be built as an alternative to the classical Gaussian detector [10] where an M-estimator of the scatter [29] is plugged-in instead of the SCM. M-estimators can be considered as a generalization of MLEs in CES, that do not necessary depend on the actual probability density function (PDF) of the distribution. The study of these robust detection processes is not trivial since M-estimators are expressed as solutions of fixed point equations and are not Wishart distributed (as the SCM). Several asymptotical properties of M-estimators have been derived in [30, 31] and [32]. These works have permitted to analyze the asymptotic properties of robust plug-in detectors [26, 27, 31]. The analysis in these last references is conducted with the standard Gaussian asymptotic regime, i.e. a central limit theorem applied on M-estimators around their expected value.

In this paper, we provide a new and more accurate performance analysis of the asymptotic distribution of robust plug-in detectors in a CES context. This asymptotic analysis is obtained by comparing robust plug-in detectors to a *Gaussian-Core Equivalent Detector* (GCED). The GCED is defined as the same detection statistic, built using a theoretical Wishart equivalent of the used M-estimator, referred to as Gaussian-Core Wishart Equivalent (GCWE). Following the idea from [32] the convergences towards the tests computed with the GCED are derived. Finally, the results for particular cases of robust plug-in detectors are presented with some simulations. The particular cases are obtained for two widely used M-estimators:

1. Tyler's M-estimator: The Tyler's M-estimator [33] is the MLE for CG distributions when the texture is assumed to be an unknown deterministic parameter and is an approximate MLE when is assumed to be a positive random variable [34, 35, 36]. It has the rare property that it is "distribution-free" with respect to the class of (continuous) CES [37].

2. Student's M-estimator: We consider the MLE based on Student's t-distribution. This estimator is widely used in signal processing community, thanks to the degree of freedom (DoF) parameter. This estimator represents a trade-off between the Tyler's M-estimator (DoF=0) and the SCM (DoF$\rightarrow \infty$). Consequently, for small values of DoF, this estimator is robust to outliers.

Practical interest of the results in this work is that one can plug-in M-estimators in any detection statistic adapted to non-Gaussian environments and then, use the distribution of *an equivalent detector*, namely the detector built with the corresponding GCWE to theoretically tune the parameters (such as the detection threshold).

The rest of this paper is organized as follows. Section 2 introduces adaptive detection. Section 3 introduces CES distributions and M-estimators. Then, Section 4 presents the main contribution of the paper. Section 5 validates the theoretical results and explains their application with appropriate comments. Finally, some conclusions are drawn in Section 6.

Vectors (resp. matrices) are denoted by bold-faced lowercase letters (resp. uppercase letters). ^{T} and ^{H} respectively represent the transpose and the Hermitian operator. $\mathcal{N}$ and $\mathcal{CN}$ denote the real and complex normal distributions. $\sim$ means "distributed as", $\stackrel{d}{=}$ stands for "shares the same distribution as", $\stackrel{d}{\rightarrow}$ denotes convergence in distribution and $\otimes$ denotes the Kronecker product. vec is the operator which transforms a matrix $m \times n$ into a vector of length mn, using column-wise concatenation. Moreover, $\mathbf{I}$ is the identity matrix and $\mathbf{0}$ the matrix of zeros with appropriate dimension, while $\mathbf{K}$ is the commutation matrix which transforms vec($\mathbf{A}$) into vec($\mathbf{A}^T$).

2. Detection

2.1. Problem setup

We consider the problem of detecting a known complex signal vector $\mathbf{p}$, called steering vector, from the received data $\mathbf{z} = \alpha\mathbf{p} + \mathbf{c}$, called primary data, where $\mathbf{c}$ is the unobserved complex noise random vector and $\alpha \in \mathbb{C}$ is a complex amplitude modelled as an unknown deterministic parameter. The problem of detecting the signal $\mathbf{p}$ can then be expressed as the following binary hypothesis test

$$\begin{cases} H_0 : \mathbf{z} = \mathbf{c} & \mathbf{z}_i = \mathbf{c}_i, \quad i = 1, \ldots, n \\ H_1 : \mathbf{z} = \alpha\mathbf{p} + \mathbf{c} & \mathbf{z}_i = \mathbf{c}_i, \quad i = 1, \ldots, n \end{cases} \tag{1}$$

where the $\mathbf{c}_i$ are n signal-free independent measurements, traditionally referred to as the secondary data. In order to detect the signal, the value of a detection statistic is compared to a pre-computed threshold that is obtained for a given probability of false-alarm (P_{fa}).

Note that in this paper n stands for the number of observations while p corresponds to the vector dimension.

2.2. Adaptive detectors

2.2.1. Kelly's GLR

Assuming that the primary and secondary data are Gaussian-distributed with an unknown covariance matrix $\mathbf{C}$, the GLRT has been proposed by Kelly in [2]

$$\mathcal{D}_{\mathcal{K}}\left(\widehat{\mathbf{C}}\right) = \frac{\left|\mathbf{p}^H\widehat{\mathbf{C}}^{-1}\mathbf{z}\right|^2}{\left(\mathbf{p}^H\widehat{\mathbf{C}}^{-1}\mathbf{p}\right)\left(1 + \frac{1}{n}\mathbf{z}^H\widehat{\mathbf{C}}^{-1}\mathbf{z}\right)} \underset{H_1}{\overset{H_0}{\lessgtr}} n\lambda \tag{2}$$

where

$$\widehat{\mathbf{C}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^H \tag{3}$$

is the well-known SCM. Moreover Kelly showed that $\mathcal{D}_{\mathcal{K}}$ in [2] is a CFAR detector with respect to the clutter covariance matrix, since the $P_{fa} - \lambda$ relationship depends only n and p. The PDF for the detector has been also derived, for both signal-present and signal-absent cases, together with P_d.

2.2.2. AMF

In [6], Robey proposed another GLRT under the same setup, but considering that the CM $\mathbf{C}$ is known. The adaptive version of the detector is then obtained with the SCM $\widehat{\mathbf{C}}$ (Eq. (3)) and given by

$$\mathcal{D}_{\mathcal{W}}\left(\widehat{\mathbf{C}}\right) = \frac{\left|\mathbf{p}^H\widehat{\mathbf{C}}^{-1}\mathbf{z}\right|^2}{\left(\mathbf{p}^H\widehat{\mathbf{C}}^{-1}\mathbf{p}\right)} \underset{H_1}{\overset{H_0}{\lessgtr}} \lambda. \tag{4}$$

One can note that this statistic does not contain the factor in parentheses, found in the denominator of the Kelly's GLRT given by Eq. (2). This term is computationally demanding for real time systems. However, it can be noted that this term tends to 1 when n is large. The $P_{fa} - \lambda$ relationship and P_d for this statistic are given in [6].

2.2.3. ANMF

ANMF [7] has been derived for partially homogeneous Gaussian noise where the CM is different between the primary and secondary data, i.e. $\mathbf{c} \sim \mathcal{CN}(\mathbf{0}, \alpha\mathbf{C})$ and $\mathbf{c}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{C})$,

$$\mathcal{D}_{\mathcal{H}}\left(\widehat{\mathbf{C}}\right) = \frac{\left|\mathbf{p}^H\widehat{\mathbf{C}}^{-1}\mathbf{z}\right|^2}{\left(\mathbf{p}^H\widehat{\mathbf{C}}^{-1}\mathbf{p}\right)\left(\mathbf{z}^H\widehat{\mathbf{C}}^{-1}\mathbf{z}\right)} \underset{H_1}{\overset{H_0}{\lessgtr}} \lambda. \tag{5}$$

This statistic is scale-invariant and "distribution-free" when computed with the true CM $\mathbf{C}$. The resulting PDFs, $P_{fa} - \lambda$ relationship and P_d for the detector built with true CM $\mathbf{C}$ were derived in [7], while the corresponding quantities for the adaptive version was proposed in [38].

2.2.4. Rao statistic

In 2007, De Maio proposed a new detection statistic based on Rao test [9]

$$\mathcal{D}_{\mathcal{R}}\left(\widehat{\mathbf{C}}\right) = \frac{\left|\mathbf{p}^H\widehat{\mathbf{C}}^{-1}\mathbf{z}\right|^2 / \left(\mathbf{p}^H\widehat{\mathbf{C}}^{-1}\mathbf{p}\right)}{\left(1 + \frac{1}{n}\mathbf{z}^H\widehat{\mathbf{C}}^{-1}\mathbf{z}\right)\left[1 + \frac{1}{n}\mathbf{z}^H\widehat{\mathbf{C}}^{-1}\mathbf{z} + \frac{1}{n}\frac{\left|\mathbf{p}^H\widehat{\mathbf{C}}^{-1}\mathbf{z}\right|^2}{\mathbf{p}^H\widehat{\mathbf{C}}^{-1}\mathbf{p}}\right]} \underset{H_1}{\overset{H_0}{\lessgtr}} n\lambda. \tag{6}$$

This test also ensures the CFAR property and it is invariant to the set of transformation defined in [9]. Note that both terms in denominator tend to unity when n is large.

3. Non-Gaussian context

3.1. Complex Elliptically Symmetric distributions

Let $\mathbf{z}$ be a zero-mean p-dimensional complex circular random vector. The vector $\mathbf{z}$ has a complex elliptically symmetric (CES) distribution, denoted as $\mathcal{CES}(\mathbf{0}, \mathbf{\Sigma}, g_{\mathbf{z}})$, if its PDF can be written as

$$h_{\mathbf{z}}(\mathbf{z}) = C|\mathbf{\Sigma}|^{-1} g_{\mathbf{z}}\left(\mathbf{z}^H \mathbf{\Sigma}^{-1} \mathbf{z}\right) \tag{7}$$

where C is a constant and $g_{\mathbf{z}} : [0, \infty) \to [0, \infty)$ is any function such that Eq. (7) defines a PDF and $\mathbf{\Sigma}$ is the Scatter Matrix (SM). The SM is equal to the CM of $\mathbf{z}$ up to a scale factor[1]. Complex Normal distribution is a particular case of CES distributions in which $g_{\mathbf{z}}(x) = e^{-x}$ and $C = \pi^{-p}$.

Definition 3.1 (Gaussian cores of CES) *As exploited in [32], each* $\mathbf{z} \sim \mathcal{CES}(\mathbf{0}, \mathbf{\Sigma}, g_{\mathbf{z}})$ *has the following stochastic representation:*

$$\mathbf{z} \overset{d}{=} \frac{\sqrt{\mathcal{Q}}}{\|\mathbf{g}\|}\mathbf{A}\mathbf{g} \tag{8}$$

where $\mathbf{g} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$*,* $\mathcal{Q}$ *is a non-negative real random variable, called the modular variate, and* $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}^H$ *is a factorization of* $\mathbf{\Sigma}$*. We refer to* $\mathbf{x} = \mathbf{A}\mathbf{g}$ *as the Gaussian core of* $\mathbf{z}$*.*

3.2. M-estimators

Let $(\mathbf{z}_1, \ldots, \mathbf{z}_n)$ be an n-sample of p-dimensional complex independent vectors with $\mathbf{z}_i \sim \mathcal{CES}(\mathbf{0}, \mathbf{\Sigma}, g_{\mathbf{z}})$. An M-estimator, denoted by $\widehat{\mathbf{\Sigma}}$, is defined by the solution of the following M-estimating equation

$$\widehat{\mathbf{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n} u(\mathbf{z}_i^H \widehat{\mathbf{\Sigma}}^{-1} \mathbf{z}_i)\mathbf{z}_i \mathbf{z}_i^H \tag{9}$$

where u is any real-valued weight function on $[0, \infty)$ that respects the Maronna's conditions given in [29]. An M-estimator given by Eq. (9) tends to the theoretical scatter matrix M-functional $\mathbf{\Sigma}_\sigma = \sigma^{-1}\mathbf{\Sigma}$ where the scalar factor $\sigma > 0$

[1] This is the case if the random vector has a finite second-order moment (see [13] for details)

can be found by solving

$$\mathbb{E}[\psi(\sigma t)] = p \tag{10}$$

where $\psi(\sigma t) = u(\sigma t)\sigma t$ and $t = \mathbf{z}^H \boldsymbol{\Sigma}_\sigma^{-1} \mathbf{z}$. Note the SCM given by Eq. (3) satisfies Eq. (9) for $u(x) = 1$, but does not satisfy Maronna's conditions.

Example 3.1 (Tyler's M-estimator) *Tyler's M-estimator [33] is given as the solution of the following equation*

$$\widehat{\boldsymbol{\Sigma}}_T = \frac{p}{n} \sum_{i=1}^{n} \frac{\mathbf{z}_i \mathbf{z}_i^H}{\mathbf{z}_i^H \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{z}_i}. \tag{11}$$

In order to provide a unique solution, the trace of this equation is usually normalized giving the estimation of so-called shape matrix.

Example 3.2 (Student's M-estimator) *Student's M-estimator is a MLE for Student's t-distribution. It is given as the solution of the following equation*

$$\widehat{\boldsymbol{\Sigma}}_t = \frac{p+d}{n} \sum_{i=1}^{n} \frac{\mathbf{z}_i \mathbf{z}_i^H}{\mathbf{z}_i^H \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{z}_i + d}, \tag{12}$$

where $d > 1$ is number of degrees of freedom (DoF). When $d \to \infty$ the Student's t-distribution leads the Gaussian distribution and the Student's M-estimator tends to the SCM ($u(t) \to 1$). On the other hand, for $d = 0$ Student's M-estimator is equivalent to the Tyler's one.

3.3. Asymptotic characterizations of the M-estimators

Theorem 3.1 (Standard asymptotic (SA) regime) *Let $\widehat{\boldsymbol{\Sigma}}$ be an M-estimator as in* (9) *built from samples drawn as $\mathbf{z} \sim \mathcal{CES}(\mathbf{0}, \boldsymbol{\Sigma}, g_{\mathbf{z}})$. The asymptotic distribution of $\widehat{\boldsymbol{\Sigma}}$ is given by [31, 13] as*

$$\sqrt{n}\mathrm{vec}\left(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_\sigma\right) \xrightarrow{d} \mathcal{GCN}\left(\mathbf{0}, \boldsymbol{\Gamma}, \boldsymbol{\Omega}\right)$$

where the asymptotic covariance and pseudo-covariance matrices are

$$\begin{cases} \boldsymbol{\Gamma} = \vartheta_1 \boldsymbol{\Sigma}_\sigma^T \otimes \boldsymbol{\Sigma}_\sigma + \vartheta_2 \mathrm{vec}\left(\boldsymbol{\Sigma}_\sigma\right) \mathrm{vec}\left(\boldsymbol{\Sigma}_\sigma\right)^H, \\ \boldsymbol{\Omega} = \vartheta_1 \left(\boldsymbol{\Sigma}_\sigma^T \otimes \boldsymbol{\Sigma}_\sigma\right) \mathbf{K} + \vartheta_2 \mathrm{vec}\left(\boldsymbol{\Sigma}_\sigma\right) \mathrm{vec}\left(\boldsymbol{\Sigma}_\sigma\right)^T. \end{cases} \tag{13}$$

The constants $\vartheta_1 > 0$ and $\vartheta_2 > -\vartheta_1/p$ are given by

$$\begin{aligned} \vartheta_1 &= c_M^{-2} a_M p(p+1) \\ \vartheta_2 &= (c_M - p^2)^{-2}(a_M - p^2) - c_M^{-2} a_M (p+1) \end{aligned} \tag{14}$$

where $a_M = E[\psi^2(\sigma\mathcal{Q})]$ and $c_M = E[\psi'(\sigma\mathcal{Q})\sigma\mathcal{Q}] + p^2$, with $\mathcal{Q}$ defined in Eq. (8).

Note that for the SCM $\vartheta_1 = 1$ and $\vartheta_2 = 0$ with $\sigma = 1$.

Definition 3.2 (Gaussian-Core Wishart Equivalent (GCWE)) *Let n measurements $(\mathbf{z}_1, \ldots, \mathbf{z}_n)$ be drawn as $\mathbf{z} \sim \mathcal{CES}(\mathbf{0}, \mathbf{\Sigma}, g_{\mathbf{z}})$ and denote $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ their Gaussian cores as $\mathbf{z}_i \overset{d}{=} \sqrt{\mathcal{Q}_i}/\|\mathbf{g}_i\|\mathbf{x}_i$ (cf. Definition 3.1). Let $\widehat{\mathbf{\Sigma}}$ be an M-estimator built with $(\mathbf{z}_1, \ldots, \mathbf{z}_n)$ using Eq. (9). The SCM built from the Gaussian cores, i.e.*

$$\widehat{\mathbf{C}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^H \tag{15}$$

is referred to as Gaussian-Core Wishart Equivalent (GCWE) of $\widehat{\mathbf{\Sigma}}$. It is important to notice that this matrix can not be computed in practice, but represents a theoretical equivalent.

Theorem 3.2 (Convergence towards GCWE) *Let $\widehat{\mathbf{\Sigma}}$ be an M-estimator as in (9) built from n samples drawn as $\mathbf{z} \sim \mathcal{CES}(\mathbf{0}, \mathbf{\Sigma}, g_{\mathbf{z}})$, and $\widehat{\mathbf{C}}$ be its GCWE from Definition 3.2. The asymptotic distribution of $\sigma\widehat{\mathbf{\Sigma}} - \widehat{\mathbf{C}}$ (with σ in (10)) is given by [32]*

$$\sqrt{n}\mathrm{vec}\left(\sigma\widehat{\mathbf{\Sigma}} - \widehat{\mathbf{C}}\right) \overset{d}{\to} \mathcal{GCN}\left(\mathbf{0}, \widetilde{\mathbf{\Gamma}}, \widetilde{\mathbf{\Omega}}\right) \tag{16}$$

where $\widetilde{\mathbf{\Gamma}}$ and $\widetilde{\mathbf{\Omega}}$ are defined by

$$\begin{aligned} \widetilde{\mathbf{\Gamma}} &= \sigma_1 \mathbf{\Sigma}^T \otimes \mathbf{\Sigma} + \sigma_2 \mathrm{vec}(\mathbf{\Sigma})\mathrm{vec}(\mathbf{\Sigma})^H, \\ \widetilde{\mathbf{\Omega}} &= \sigma_1 \left(\mathbf{\Sigma}^T \otimes \mathbf{\Sigma}\right)\mathbf{K} + \sigma_2 \mathrm{vec}(\mathbf{\Sigma})\mathrm{vec}(\mathbf{\Sigma})^T \end{aligned} \tag{17}$$

with σ_1 and σ_2 given by

$$\begin{aligned} \sigma_1 &= \frac{a_M p(p+1) + c_M(c_M - 2b)}{c_M^2} \\ \sigma_2 &= \frac{a_M - p^2}{(c_M - p^2)^2} - \frac{a_M(p+1)}{c_M^2} + 2\frac{p(c_M - b_M)}{c_M(c_M - p^2)} \end{aligned} \tag{18}$$

where $b_M = E[\psi(\sigma\mathcal{Q})\|\mathbf{g}\|^2]$.

4. Robust detection with plug-in adaptive statistics

This section contains the main contribution of the paper. First, we define robust plug-in detectors as the adaptive detectors (itemized in Section 2) with M-estimators plugged-in instead of the classical SCM. Note that, in this case, the notion of robustness refers to the clutter scatter matrix estimation and not to the classical robustness definition, meaning that the detection statistic distribution is (asymptotically) independent of the data distribution under H_0 [39, 40]. Indeed, the distribution of the detection statistic depends on both the received data distribution and the distribution of the scatter matrix estimator. Among analyzed detectors, only the ANMF has this "distribution-free" property over the class of CES distribution and consequently, is robust in both senses.

Let us now state two definition essential for our study.

Definition 4.1 (Robust plug-in detector) *Consider a set of $n+1$ samples $\{\mathbf{z}, \mathbf{z}_1, \ldots, \mathbf{z}_n\}$ drawn as $\mathbf{z}_i \sim \mathcal{CES}(\mathbf{0}, \mathbf{\Sigma}, g_{\mathbf{z}})$ with the detection problem expressed in (1). Let $\widehat{\mathbf{\Sigma}}$ be an M-estimator as in (9) built from $\{\mathbf{z}_1, \ldots, \mathbf{z}_n\}$ and $\mathcal{D}$ be a decision statistic (either Kelly in (2), AMF in (4), ANMF in (5), or Rao in (6)). The statistic $\mathcal{D}(\widehat{\mathbf{\Sigma}})$, that uses the M-estimator $\widehat{\mathbf{\Sigma}}$ as plug-in instead of the traditional SCM, is referred to as a robust plug-in detector.*

Definition 4.2 (Gaussian-Core Equivalent Detector (GCED)) *Consider a set of $n+1$ samples $\{\mathbf{z}, \mathbf{z}_1, \ldots, \mathbf{z}_n\}$ drawn as $\mathbf{z}_i \sim \mathcal{CES}(\mathbf{0}, \mathbf{\Sigma}, g_{\mathbf{z}})$ with the detection problem expressed in (1). Let $\mathcal{D}(\widehat{\mathbf{\Sigma}})$ be a robust plug-in detector as in Definition 4.1. Let $\widehat{\mathbf{C}}$ be the GCWE of $\widehat{\mathbf{\Sigma}}$ (cf. Definition 3.2). The quantity $\mathcal{D}(\widehat{\mathbf{C}})$ is referred to as Gaussian-Core Equivalent Detector (GCED) of $\mathcal{D}(\widehat{\mathbf{\Sigma}})$.*

In the following, we propose several theorems to characterize the distribution of robust plug-in detectors (as in Definition 4.1) conditionally to the tested sample $\mathbf{z}$. First, the results are obtained for the standard asymptotic, then the convergence towards the GCED is derived. Note that, only the results provided in Theorem 4.2 are not derived in this paper. To emphasize this, we provide

the corresponding reference in the theorem title. The section is concluded by examples for specific M-estimators.

4.1. Standard asymptotic regime

In SA regime, the robust adaptive detector is compared to its non-adaptive form, i.e. detector built using the true SM. For $\mathcal{D}_{\mathcal{K}}$ (Eq. (2)), $\mathcal{D}_{\mathcal{W}}$ (Eq. (4)), or $\mathcal{D}_{\mathcal{R}}$ (Eq. (6)), one obtains the following results.

Theorem 4.1 (SA of the robust AMF, Kelly and Rao) *Let $\mathcal{D}(\widehat{\mathbf{\Sigma}})$ be a robust plug-in detector as in Definition 4.1 with $\mathcal{D} \in \{\mathcal{D}_{\mathcal{K}}, \mathcal{D}_{\mathcal{W}}, \mathcal{D}_{\mathcal{R}}\}$. Conditionally to the distributions of $\mathbf{z}$, the asymptotic distribution of $\mathcal{D}\left(\widehat{\mathbf{\Sigma}}\right)$ is given by*

$$\sqrt{n}\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right) - \mathcal{D}\left(\mathbf{\Sigma}\right)\right)_{\mathbf{z}} \xrightarrow{d} \mathcal{N}\left(0, \vartheta_1\sigma_{\mathcal{X}} + \vartheta_2\mathcal{D}^2\left(\mathbf{\Sigma}\right)\right) \tag{19}$$

with $\sigma_{\mathcal{X}} = \mathcal{D}\left(\mathbf{\Sigma}\right)\left(2\mathbf{z}^H\mathbf{\Sigma}^{-1}\mathbf{z} - \mathcal{D}\left(\mathbf{\Sigma}\right)\right)$ and ϑ_1,ϑ_2 in (14).

Proof 4.1 *See Appendix A.*

Due to the specific scale invariance property, the variance of ANMF has a particular form (depending only on the first scale factor) derived in [26].

Theorem 4.2 (SA of the robust ANMF [26]) *Let $\mathcal{D}_{\mathcal{H}}(\widehat{\mathbf{\Sigma}})$ be the robust ANMF detector as in (5) and Definition 4.1. Conditionally to the distributions of $\mathbf{z}$, the asymptotic distribution of $\mathcal{D}_{\mathcal{H}}\left(\widehat{\mathbf{\Sigma}}\right)$ is given by*

$$\sqrt{n}\left(\mathcal{D}_{\mathcal{H}}\left(\widehat{\mathbf{\Sigma}}\right) - \mathcal{D}_{\mathcal{H}}\left(\mathbf{\Sigma}\right)\right)_{\mathbf{z}} \xrightarrow{d} \mathcal{N}\left(0, \vartheta_1\sigma_{\mathcal{H}}\right) \tag{20}$$

where $\sigma_H = 2\mathcal{D}_{\mathcal{H}}\left(\mathbf{\Sigma}\right)\left(\mathcal{D}_{\mathcal{H}}\left(\mathbf{\Sigma}\right) - 1\right)^2$ and ϑ_1, in (14).

Proof 4.2 *This theorem has been proved in [26].*

4.2. Convergence towards GCED

Theorem 4.3 (GCED of the robust AMF, Kelly and Rao) *Let* $\mathcal{D}(\widehat{\mathbf{\Sigma}})$ *be a robust plug-in detector as in Definition 4.1 with* $\mathcal{D} \in \{\mathcal{D}_{\mathcal{K}}, \mathcal{D}_{\mathcal{W}}, \mathcal{D}_{\mathcal{R}}\}$. *Conditionally to the distributions of* $\mathbf{z}$, *the asymptotic GCED of* $\mathcal{D}\left(\widehat{\mathbf{\Sigma}}\right)$ *is given by*

$$\sqrt{n}\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right) - \mathcal{D}\left(\widehat{\mathbf{C}}\right)\right)_{\mathbf{z}} \xrightarrow{d} \mathcal{N}\left(0, \sigma_1\sigma_{\mathcal{X}} + \sigma_2\mathcal{D}_{\mathcal{W}}^2\left(\mathbf{\Sigma}\right)\right) \tag{21}$$

with $\sigma_{\mathcal{X}} = \mathcal{D}\left(\mathbf{\Sigma}\right)\left(2\mathbf{z}^H\mathbf{\Sigma}^{-1}\mathbf{z} - \mathcal{D}\left(\mathbf{\Sigma}\right)\right)$ *and* σ_1, σ_2 *in* (18).

Proof 4.3 *See Appendix A.*

Theorem 4.4 (GCED of the robust ANMF) *Let* $\mathcal{D}_{\mathcal{H}}(\widehat{\mathbf{\Sigma}})$ *be the robust ANMF detector as in* (5) *and Definition 4.1. Conditionally to the distributions of* $\mathbf{z}$, *the asymptotic distribution of* $\mathcal{D}_{\mathcal{H}}\left(\widehat{\mathbf{\Sigma}}\right)$ *is given by*

$$\sqrt{n}\left(\mathcal{D}_{\mathcal{H}}\left(\widehat{\mathbf{\Sigma}}\right) - \mathcal{D}_{\mathcal{H}}\left(\widehat{\mathbf{C}}\right)\right)_{\mathbf{z}} \xrightarrow{d} \mathcal{N}\left(0, \sigma_1\sigma_{\mathcal{H}}\right) \tag{22}$$

where $\sigma_H = 2\mathcal{D}_{\mathcal{H}}\left(\mathbf{\Sigma}\right)\left(\mathcal{D}_{\mathcal{H}}\left(\mathbf{\Sigma}\right) - 1\right)^2$ *and* σ_1 *in* (18).

Proof 4.4 *The proof follows the same lines as the ones for Theorems 4.1 and 4.3. This is detailed at the end of Appendix A.*

Remark 4.1 *Note that the results in the SA regime depend on the scale factors* ϑ_1 *and* ϑ_2 *given by Eqs.* (14). *In the GCED regime, these are replaced by the* σ_1 *and* σ_2 *(Eq.* (18)*). As shown in [32], and will be demonstrated in the following examples, one has that* σ_1, σ_2 *are significantly smaller than* ϑ_1, ϑ_2, *meaning that the variances in the GCED regime are lower than the ones in the SA regime. This directly implies a robust plug-in detector is better described by the distribution of its GCED, than with its asymptotic properties.*

4.3. Examples

Example 4.1 (Robust plug-in detectors with Tyler's M-estimator) *Robust plug-in detectors built from Tyler's M-estimator in* (11) *are characterized by Theorems 4.1-4.2 (for SA) and Theorems 4.3-4.4 (for GCED) with*

coefficients

$$\begin{cases} \vartheta_1 = (p+1)/p \\ \vartheta_2 = -(p+1)/p^2 \end{cases} \quad \textit{and} \quad \begin{cases} \sigma_1 = 1/p \\ \sigma_2 = (p-1)/p^2 \end{cases} \tag{23}$$

independently from the underlying distribution of the secondary data.

Remark 4.2 *For the Tyler's M-estimator the first scaling factors can be obtained directly from Eqs.* (14) *and* (18) *using $\psi(\sigma\mathcal{Q}) = p$ which leads to $a_M = b_M = c_M = p^2$. This is in agreement with the results obtained in [41, 42]. However, since this estimator does not respect all Maronna's conditions, the result for the second scale factors can not be obtained directly from Eqs.*(14) *and* (18)*, but can be found in [31] and [32].*

Example 4.2 (Robust plug-in detectors with Student's M-estimator) *Robust plug-in detectors built from Student's M-estimator in* (11) *are characterized by Theorems 4.1-4.2 (for SA) and Theorems 4.3-4.4 (for GCED) with coefficients*

$$\begin{cases} \vartheta_1 = (p+d+1)/(p+d) \\ \vartheta_2 = (p+1+d)/(d(p+d)) \end{cases} \quad \textit{and} \quad \begin{cases} \sigma_1 = 1/(p+d) \\ \sigma_2 = (p+1+d)/(d(p+d)) \end{cases} \tag{24}$$

in the matched case, i.e. when the samples follow a t-distribution [13] of degrees of freedom d.

Remark 4.3 *The results for Student's M-estimator can be obtained directly from Eqs.* (14) *and* (18) *computing plugging in the coefficients a_M, b_M and c_M obtained for $\psi(x) = (p+d)x/(x+d)$. It can be noted that when $d \to \infty$ one has that variance between the robust plug-in detector and its GCWE goes to zero as expected.*

5. Simulations

5.1. Validation of theoretical results

The simulations have been carried out with complex zero-mean t-distributed secondary data. The scatter matrix $\mathbf{\Sigma}$ is Toeplitz, i.e. its elements are defined

by $\Sigma_{j,k} = \rho^{|j-k|}, \quad j,k = 1,\ldots,p$. The correlation coefficient ρ is set to 0.5. The DoF parameter d is equal to 2. Empirical means are computed using Monte Carlo runs.

Figure A.1 illustrates the theoretical result of Theorems 4.1 and 4.3. The empirical variances of robust plug-in detectors in SA, i.e. when centering around the detectors obtained with the true SM, are plotted and compared to the theoretical results. Moreover, the empirical variances of the difference between the robust plug-in detectors and their GCEDs are also given together with the corresponding theoretical result. First, one can see that the simulations validate theoretical results. Then, one can note that the variances when comparing to the GCEDs are significantly lower than the ones in the SA, which justifies the proposed approximation of the behavior of robust plug-in detectors with the one of their GCEDs. One observes that the error decreases very fast as the number n of samples increases. Furthermore, simulations show that the approximation is also valid for small n. Indeed, as derived for both Tyler's M-estimator and Student M-estimator (σ_1 respectively in Eqs. (23) and (24)), the asymptotic variance given by Theorem 4.4 is of order $1/p$ (compared to order 1 for the classical SA regime). Roughly speaking, this implies that the asymptotic regime is achieved at the speed $\sqrt{np}$ (compared to order $\sqrt{n}$ for the classical SA regime). Hence, when $p = 10$, as it is the case for results displayed on figure A.1, the case $n = 20$ basically corresponds to an equivalent case with $n' = 200$ samples. This explains why the results are accurate even for small values of n.

On Figure A.2 the corresponding results for the ANMF built with the Tyler's estimator (TyE-ANMF) for $p = 10$ and $p = 100$ are plotted. Once again, one can note that the empirical variance of TyE-ANMF when compared to the GCED is remarkably smaller than the one in SA. In addition, one can notice that this difference is even smaller for higher dimension, which is in agreement with theoretical results.

5.2. Application

In order to demonstrate the application of the results, we will first analyze the robust ANMF. We recall that when the scatter matrix $\mathbf{\Sigma}$ is known, $\mathcal{D}_{\mathcal{H}}\left(\mathbf{\Sigma}\right)$ (called NMF) follows a beta distribution $\beta_{1,p-1}(u)$ under H_0 [7]

$$f_{\beta}(u) = (p-1)(1-u)^{p-2}\mathbf{1}_{[0,u]}(u) \tag{25}$$

where $\mathbf{1}_{[0,u]}(\cdot)$ is the indicator function on $[0,1]$. This results in the following $P_{fa} - \lambda$ relationship

$$Pfa = P\left(\mathcal{D}_{\mathcal{H}}\left(\mathbf{\Sigma}\right) > \lambda | H_0\right) = (1-\lambda)^{p-1} \tag{26}$$

and the probability of detection P_d for a given SNR δ and for a fixed value of the detection threshold λ is given by

$$P_d = P\left(\mathcal{D}_{\mathcal{H}}\left(\mathbf{\Sigma}\right) > \lambda | H_1\right) = 1 - \int_0^{\lambda} \beta_{1,p-1}(u) e^{\delta(u-1)} {}_1F_1(p,1;u\delta)du \tag{27}$$

where ${}_1F_1(\cdot)$ is the complex confluent hypergeometric function. Assuming that instead of $\mathbf{\Sigma}$, the SCM computed with Gaussian-distributed data is plugged in, the PDF of $\mathcal{D}_{\mathcal{H}}\left(\widehat{\mathbf{C}}\right)$ is given by [38]

$$f_{\mathcal{D}_{\mathcal{H}}(\widehat{\mathbf{C}})}(u) = K(1-u)^{a-2} {}_2F_1(a,a;b;u)\mathbf{1}_{[0,u]}(u) \tag{28}$$

with

$$P_{fa} = P\left(\mathcal{D}_{\mathcal{H}}\left(\widehat{\mathbf{C}}\right) > \lambda | H_0\right) = (1-\lambda)^{a-1} {}_2F_1(a,a-1;b-1;\lambda), \tag{29}$$

where $K = \frac{(a-1)(p-1)}{n+1}$, $a = n-p+2$, $b = n+2$ and ${}_2F_1(\cdot)$ is the hypergeometric function [43]. The P_d for a given SNR δ is

$$\begin{aligned} P_d &= P\left(\mathcal{D}_{\mathcal{H}}\left(\widehat{\mathbf{C}}\right) > \lambda | H_1\right) \\ &= 1 - \int_0^1 \frac{e^{-\delta u}}{N} \int_0^{\lambda} u^{a-1} \frac{(1-u)^{p-1}(1-x)^{n-p}}{(1-ux)^a} {}_1F_1\left(a,1;\frac{\delta u x(1-u)}{1-ux}\right) dxdu \end{aligned} \tag{30}$$

where $N = \frac{\Gamma(a-1)\Gamma(p-1)}{\Gamma(n+1)}$.

Since the variance between a robust plug-in detector and its GCED is close to zero, one can conclude that the distribution of the robust plug-in detector can be well-approximated with the one of GCED. Figure A.3 illustrates this point. The histogram on Figure A.3 represents the empirical distribution of t-ANMF that is compared to the theoretical distribution of NMF given by Eq. (25) (red curve) and theoretical distribution of GCWE given by Eq. (28) (green curve). One can note that the red curve mismatches the empirical distribution of t-ANMF, while the green one borders the plot area of the histogram showing that Eq. (28) gives a good approximation of the t-ANMF's behaviour.

Having this in mind, one can analyze the $P_{fa} - \lambda$ relationship for the robust ANMF. On Figure A.4 we observe this relationship for different number of secondary data, $n = 20$ and $n = 100$. The empirical results for t-ANMF and TyE-ANMF have been plotted and compared to the theoretical ones given by Eqs. (26) and (29). The empirical results for the corresponding GCWE-ANMF (that is not available in practice) and the SCM-ANMF built with observed data are also provided. First, one can notice a good match of the empirical distributions of both t-ANMF and TyE-ANMF with the theoretical and empirical distribution of the GCWE-ANMF. This shows that the behavior of robust plug-in detectors is better approximated with the one of the GCWE-ANMF (green curve) than with the corresponding NMF (black curve). Secondly, one can see that this claim is even more obvious for small n since all curves approach when n increases. Finally, one can note that the SCM built with secondary data does not satisfy the relationship Eq. (29) anymore. This is expected since the SCM is calculated with a non-Gaussian data and its performances are remarkably degraded, which highly supports the use of M-estimators in this context.

For a constant P_{fa} we can study the P_d of the robust ANMF for a given SNR δ. Figure A.5 shows the empirical P_d for TyE-ANMF where the detection threshold is computed empirically for $P_{fa} = 0.001$ (TyE-ANMF$_{\text{emp}}$) together with the empirical P_d obtained with the threshold computed using Eq. (29) (TyE-ANMF$_{\text{the}}$). The empirical results have been compared to the corresponding theoretical results for GCED given by Eq. (30) (green curves) for

$n = 20, 50, 200$. Finally, the theoretical results for NMF (Eq. (27)) have been plotted. The figure is revealing in several ways. First, the results for TyE-ANMF$_{\text{emp}}$ match perfectly the ones for TyE-ANMF$_{\text{the}}$, meaning that we obtain the same P_d for empirically and theoretically computed threshold which can significantly reduce the computational cost of the method. In addition, both empirical results coincide with the theoretical one for GCED as stated. Finally, the approximation is valid even for small n approaching the results for NMF when n increases. Taken together, these results suggest that one can use M-estimators to compute the value of detection statistic and pre-compute the detection threshold using the theoretical results for GCED.

Finally, we illustrate that the general results of Theorem 4.3 can also be applied identically: we now consider Kelly's GLR [2] computed with a plug-in Student's M-estimator. Following the same approach as previously, the probability of detection of this detector has been set either by Monte-Carlo simulation, or using the one of its theoretical Gaussian equivalent, given by

$$P_{fa} = P\left(\mathcal{D}_{\mathcal{K}}\left(\widehat{\mathbf{C}}\right) > \lambda | H_0\right) = (1-\lambda)^{n-p+1}. \tag{31}$$

The detection probability of these two detectors detector has then been compared to the one expected in the Gaussian context, given by

$$\begin{aligned} P_d &= P\left(\mathcal{D}_{\mathcal{K}}\left(\widehat{\mathbf{C}}\right) > \lambda | H_1\right) \\ &= 1 - \int_0^1 \frac{e^{-\delta u}}{N} \int_0^{\lambda} u^{n-p+1}(1-u)^{p-2}(1-x)^{n-p}{}_1F_1\left(n-p+2, 1; \delta u x\right) dx du. \end{aligned} \tag{32}$$

Figure A.6 displays the obtained results, and the same general conclusions can be drawn.

6. Conclusion

This paper has shed new light on the statistical properties for the detectors built with M-estimators in CES framework, called robust plug-in detectors.

The robust plug-in detectors have been compared with their GCEDs, i.e. corresponding detectors built with the SCM in the Gaussian context, and it has been shown that the variance of the difference between them is considerably smaller than the variance of the robust plug-in detectors in standard asymptotic regime. This result led us to conclude that we can use M-estimators to compute the value of the detection statistic and then use the theoretical results for GCED for theoretical derivation of the detection threshold.

Acknowledgements

This work has been supported by DIGITEO Grant N°2016-1277D-SERTI and ANR-ASTRID MARGARITA grant (ANR-17-ASTR-0015).

Appendix A. Proof of theorems

In order to prove all theorems, we first provide common steps for all detectors, where the detection test is noted as $\mathcal{D}$. Starting with

$$\begin{aligned}
&\sqrt{n}\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)-\mathcal{D}\left(\widehat{\mathbf{C}}\right)\right)=\sqrt{n}\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)-\mathcal{D}\left(\widehat{\mathbf{C}}\right)+\mathcal{D}\left(\mathbf{\Sigma}\right)\right)\\
&=\left[1,-1\right]\begin{bmatrix}\sqrt{n}\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)\\ \sqrt{n}\left(\mathcal{D}\left(\widehat{\mathbf{C}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)\end{bmatrix}
\end{aligned}$$

one obtains

$$\begin{aligned}
\sigma_{\mathcal{D}}^{(n)} &= nE\left[\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)-\mathcal{D}\left(\widehat{\mathbf{C}}\right)\right)\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)-\mathcal{D}\left(\widehat{\mathbf{C}}\right)\right)^{H}\right]\\
&= \left[1,-1\right]E\left[\left[\begin{bmatrix}\sqrt{n}\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)\\ \sqrt{n}\left(\mathcal{D}\left(\widehat{\mathbf{C}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)\end{bmatrix}\right.\right.\\
&\times \left.\left.\begin{bmatrix}\sqrt{n}\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)\\ \sqrt{n}\left(\mathcal{D}\left(\widehat{\mathbf{C}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)\end{bmatrix}^{H}\right]\begin{bmatrix}1\\-1\end{bmatrix}\\
&= \sigma_{\mathcal{D}1}^{(n)}-2\sigma_{\mathcal{D}2}^{(n)}+\sigma_{\mathcal{D}3}^{(n)}
\end{aligned}$$

with

$$\begin{aligned}
\sigma_{\mathcal{D}1}^{(n)} &= nE\left[\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)^{H}\right] \\
\sigma_{\mathcal{D}2}^{(n)} &= nE\left[\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)\left(\mathcal{D}\left(\widehat{\mathbf{C}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)^{H}\right] \\
\sigma_{\mathcal{D}3}^{(n)} &= nE\left[\left(\mathcal{D}\left(\widehat{\mathbf{C}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)\left(\mathcal{D}\left(\widehat{\mathbf{C}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)^{H}\right].
\end{aligned}$$

A first approximation of $\mathcal{D}(\sigma\widehat{\mathbf{\Sigma}})$ yields

$$\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)\simeq\mathcal{D}\left(\mathbf{\Sigma}\right)+\mathcal{D}'\left(\mathbf{\Sigma}\right)\mathrm{vec}\left(\sigma\widehat{\mathbf{\Sigma}}-\mathbf{\Sigma}\right),$$

where $\mathcal{D}'\left(\mathbf{\Sigma}\right)=\frac{\partial\mathcal{D}(\mathbf{\Sigma})}{\partial\mathrm{vec}(\mathbf{\Sigma})}$, giving the well-known result

$$\sqrt{n}\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)-\mathcal{D}\left(\mathbf{\Sigma}\right)\right)\overset{d}{\rightarrow}\mathcal{N}\left(\mathbf{0},\mathcal{D}'\left(\mathbf{\Sigma}\right)\mathbf{\Gamma}\mathcal{D}'\left(\mathbf{\Sigma}\right)^{H}\right) \tag{A.1}$$

where $\mathbf{\Gamma}$ is given by (13). Analogously, one has

$$\sigma_{\mathcal{D}2}^{(n)}\rightarrow\mathcal{D}'\left(\mathbf{\Sigma}\right)\mathbf{\Pi}\mathcal{D}'\left(\mathbf{\Sigma}\right)^{H} \tag{A.2}$$

where $nE\left[\mathrm{vec}\left(\sigma\widehat{\mathbf{\Sigma}}-\mathbf{\Sigma}\right)\mathrm{vec}\left(\widehat{\mathbf{C}}-\mathbf{\Sigma}\right)^{H}\right]\rightarrow\mathbf{\Pi}$ and finally, it can be easily shown that (see [32] for details)

$$\sqrt{n}\left(\mathcal{D}\left(\sigma\widehat{\mathbf{\Sigma}}\right)-\mathcal{D}\left(\widehat{\mathbf{C}}\right)\right)\overset{d}{\rightarrow}\mathcal{N}\left(\mathbf{0},\mathcal{D}'\left(\mathbf{\Sigma}\right)\widetilde{\mathbf{\Gamma}}\mathcal{D}'\left(\mathbf{\Sigma}\right)^{H}\right) \tag{A.3}$$

with $\widetilde{\mathbf{\Gamma}}$ given by (17).

Using Eqs. (A.1) and (A.3) and deriving the derivatives $\mathcal{D}'\left(\mathbf{\Sigma}\right)$ one can obtain the final results for all detectors.

In order to obtain derivatives we will start with Rao statistic $\mathcal{D}'_{\mathcal{R}}(\mathbf{\Sigma})$ since the derivatives for other detectors can be easily obtained from this one.

Let us first rewrite $\mathcal{D}_{\mathcal{R}}$ as

$$\mathcal{D}_{\mathcal{R}}=\frac{a}{(1+\frac{1}{n}b)(1+\frac{1}{n}b+\frac{1}{n}a)}$$

with $a=\left|\mathbf{p}^{H}\mathbf{\Sigma}^{-1}\mathbf{z}\right|^{2}/\left(\mathbf{p}^{H}\mathbf{\Sigma}^{-1}\mathbf{p}\right)$ and $b=\mathbf{z}^{H}\mathbf{\Sigma}^{-1}\mathbf{z}$. Then, one has

$$\begin{aligned}
\partial\mathcal{D}_{\mathcal{R}} &= \frac{\partial a(1+\frac{1}{n}b)(1+\frac{1}{n}b+\frac{1}{n}a)-a(\frac{1}{n}\partial b)(1+\frac{1}{n}b+\frac{1}{n}a)-a(1+\frac{1}{n}b)(\frac{1}{n}\partial b+\frac{1}{n}\partial a)}{(1+\frac{1}{n}b)^{2}(1+\frac{1}{n}b+\frac{1}{n}a)^{2}} \\
&= \mathcal{D}_{\mathcal{R}}\left(\frac{\partial a}{a}-\frac{\frac{1}{n}\partial b}{1+\frac{1}{n}b}-\frac{\frac{1}{n}\partial b+\frac{1}{n}\partial a}{1+\frac{1}{n}b+\frac{1}{n}a}\right).
\end{aligned}$$

Now

$$\begin{aligned}\partial a &= \frac{\partial\left(\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{z}\right)\left(\mathbf{z}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}\right)\left(\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}\right)}{\left(\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}\right)^2} \\ &- \frac{\left(\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{z}\right)\partial\left(\mathbf{z}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}\right)\left(\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}\right)}{\left(\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}\right)^2} \\ &- \frac{\left(\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{z}\right)\left(\mathbf{z}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}\right)\partial\left(\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}\right)}{\left(\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}\right)^2}.\end{aligned}$$

Using

$$\partial\mathbf{A}^{-1} = -\mathbf{A}^{-1}\partial\mathbf{A}\mathbf{A}^{-1}, \tag{A.4}$$

$$\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\right) = \mathrm{vec}^H\left(\mathbf{A}\right)\mathrm{vec}\left(\mathbf{B}\right), \tag{A.5}$$

$$\mathrm{vec}\left(\mathbf{A}^{-1}\mathbf{B}\mathbf{C}^{-1}\right) = \left(\mathbf{C}^T\otimes\mathbf{A}\right)^{-1}\mathrm{vec}\left(\mathbf{B}\right), \tag{A.6}$$

one can show that

$$\partial\left(\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) = -\partial\left(\mathrm{vec}^H\left(\boldsymbol{\Sigma}\right)\right)\left(\boldsymbol{\Sigma}^T\otimes\boldsymbol{\Sigma}\right)^{-1}\mathrm{vec}\left(\mathbf{z}\mathbf{p}^H\right)$$

and thus

$$\partial a = -\partial\left(\mathrm{vec}^H\left(\boldsymbol{\Sigma}\right)\right)\left(\boldsymbol{\Sigma}^T\otimes\boldsymbol{\Sigma}\right)^{-1}a\left(\frac{\mathbf{z}\mathbf{p}^H}{\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{z}}+\frac{\mathbf{p}\mathbf{z}^H}{\mathbf{z}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}}-\frac{\mathbf{p}\mathbf{p}^H}{\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}}\right)$$

and

$$\partial b = -\partial\left(\mathrm{vec}^H\left(\boldsymbol{\Sigma}\right)\right)\left(\boldsymbol{\Sigma}^T\otimes\boldsymbol{\Sigma}\right)^{-1}\mathrm{vec}\left(\mathbf{z}\mathbf{z}^H\right).$$

335 Analogously,

$$\begin{aligned}\partial\mathcal{D}_{\mathcal{K}} &= \mathcal{D}_{\mathcal{K}}\left(\frac{\partial a}{a}+\frac{\frac{1}{n}\partial b}{1+\frac{1}{n}b}\right), \\ \partial\mathcal{D}_{\mathcal{W}} &= \mathcal{D}_{\mathcal{W}}\frac{\partial a}{a}.\end{aligned}$$

One can note that

$$\begin{aligned}\frac{\frac{1}{n}\partial b}{1+\frac{1}{n}b} &\xrightarrow[n\to+\infty]{} 0, \\ \frac{\frac{1}{n}\partial b+\frac{1}{n}\partial a}{1+\frac{1}{n}b+\frac{1}{n}a} &\xrightarrow[n\to+\infty]{} 0\end{aligned}$$

and

$$\mathcal{D}_{\mathcal{R}} \xrightarrow[n\to+\infty]{} \mathcal{D}_{\mathcal{W}},$$
$$\mathcal{D}_{\mathcal{K}} \xrightarrow[n\to+\infty]{} \mathcal{D}_{\mathcal{W}}.$$

Therefore

$$\mathcal{D}'_{\mathcal{R}} \xrightarrow[n\to+\infty]{} \mathcal{D}'_{\mathcal{W}},$$
$$\mathcal{D}'_{\mathcal{K}} \xrightarrow[n\to+\infty]{} \mathcal{D}'_{\mathcal{W}}.$$

and

$$\mathcal{D}'_{\mathcal{R},\mathcal{K},\mathcal{W}} = -\mathcal{D}_{\mathcal{W}} \left(\frac{\text{vec}^H\left(\mathbf{z}\mathbf{p}^H\right)}{\mathbf{z}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}} + \frac{\text{vec}^H\left(\mathbf{p}\mathbf{z}^H\right)}{\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{z}} - \frac{\text{vec}^H\left(\mathbf{p}\mathbf{p}^H\right)}{\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}} \right) \left(\boldsymbol{\Sigma}^T \otimes \boldsymbol{\Sigma}\right)^{-1} \tag{A.7}$$

Finally, the variance in Eq. (A.1) becomes

$$\begin{aligned} \mathcal{D}'_{\mathcal{W}}\left(\boldsymbol{\Sigma}\right)\boldsymbol{\Gamma}\mathcal{D}'_{\mathcal{W}}\left(\boldsymbol{\Sigma}\right) &= \mathcal{D}^2_{\mathcal{W}} \left(\frac{\text{vec}^H\left(\mathbf{z}\mathbf{p}^H\right)}{\mathbf{z}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}} + \frac{\text{vec}^H\left(\mathbf{p}\mathbf{z}^H\right)}{\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{z}} - \frac{\text{vec}^H\left(\mathbf{p}\mathbf{p}^H\right)}{\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}} \right) \\ &\times \left(\vartheta_1 \left(\boldsymbol{\Sigma}^T \otimes \boldsymbol{\Sigma}\right)^{-1} + \vartheta_2 \text{vec}\left(\boldsymbol{\Sigma}^{-1}\right) \text{vec}\left(\boldsymbol{\Sigma}^{-1}\right)^H \right) \\ &\times \left(\frac{\text{vec}\left(\mathbf{z}\mathbf{p}^H\right)}{\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{z}} + \frac{\text{vec}\left(\mathbf{p}\mathbf{z}^H\right)}{\mathbf{z}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}} - \frac{\text{vec}\left(\mathbf{p}\mathbf{p}^H\right)}{\mathbf{p}^H\boldsymbol{\Sigma}^{-1}\mathbf{p}} \right). \end{aligned} \tag{A.8}$$

Using Eqs. (A.5) and (A.6) one can easily show that

$$\frac{\text{vec}^H(\mathbf{a}\mathbf{b}^H)}{\mathbf{a}^H\boldsymbol{\Sigma}^{-1}\mathbf{b}} \left(\boldsymbol{\Sigma}^T \otimes \boldsymbol{\Sigma}\right)^{-1} \frac{\text{vec}(\mathbf{c}\mathbf{d}^H)}{\mathbf{d}^H\boldsymbol{\Sigma}^{-1}\mathbf{c}} = \frac{\mathbf{a}^H\boldsymbol{\Sigma}^{-1}\mathbf{c}\,\mathbf{d}^H\boldsymbol{\Sigma}^{-1}\mathbf{b}}{\mathbf{a}^H\boldsymbol{\Sigma}^{-1}\mathbf{b}\,\mathbf{d}^H\boldsymbol{\Sigma}^{-1}\mathbf{c}}$$

and

$$\frac{\text{vec}^H(\mathbf{a}\mathbf{b}^H)}{\mathbf{a}^H\boldsymbol{\Sigma}^{-1}\mathbf{b}} \text{vec}\left(\boldsymbol{\Sigma}^{-1}\right) \text{vec}\left(\boldsymbol{\Sigma}^{-1}\right)^H \frac{\text{vec}(\mathbf{c}\mathbf{d}^H)}{\mathbf{d}^H\boldsymbol{\Sigma}^{-1}\mathbf{c}} = 1.$$

After some mathematical manipulations one can obtain the final results from Eq. (19), proving Theorem 4.1. Applying $\widetilde{\boldsymbol{\Gamma}}$ instead of $\boldsymbol{\Gamma}$ in Eq. (A.8), one obtains the results from Eq. (21), proving Theorem 4.3.

Analogously, starting with

$$\mathcal{D}'_{\mathcal{H}} = \mathcal{D}_{\mathcal{H}} \left(\frac{\partial a}{a} - \frac{\partial b}{b} \right) \tag{A.9}$$

one obtains the results from Theorem 4.4 which concludes the proof.

References

[1] E. J. Kelly, Adaptive detection in non-stationary interference. part 1 and part 2, Tech. rep., MASSACHUSETTS INST OF TECH LEXINGTON LINCOLN LAB (1985).

[2] E. J. Kelly, An adaptive detection algorithm, IEEE Transactions on Aerospace and Electronic Systems 22 (2) (1986) 115–127.

[3] E. J. Kelly, Adaptive detection in non-stationary interference, part iii, Massachusetts Institute of Technology, Lincoln Laboratory, Lexington, MA, Tech. Rep 761 (1987).

[4] E. J. Kelly, K. M. Forsythe, Adaptive detection and parameter estimation for multidimensional signal models, Tech. rep., MASSACHUSETTS INST OF TECH LEXINGTON LINCOLN LAB (1989).

[5] E. Kelly, Performance of an adaptive detection algorithm; rejection of unwanted signals, IEEE Transactions on Aerospace and Electronic Systems 25 (2) (1989) 122–133.

[6] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, R. Nitzberg, A cfar adaptive matched filter detector, IEEE Transactions on Aerospace and Electronic Systems 28 (1) (1992) 208–216.

[7] E. Conte, M. Lops, G. Ricci, Asymptotically optimum radar detection in Compound-Gaussian clutter, IEEE Transactions on Aerospace and Electronic Systems 31 (2) (1995) 617–625.

[8] S. Kraut, L. L. Scharf, The CFAR adaptive subspace detector is a scale-invariant GLRT, IEEE Transactions on Signal Processing 47 (9) (1999) 2538–2541.

[9] A. D. Maio, Rao test for adaptive detection in gaussian interference with unknown covariance matrix, IEEE Transactions on Signal Processing 55 (7) (2007) 3577–3584.

[10] J.-P. Ovarlez, F. Pascal, P. Forster, Covariance matrix estimation in SIRV and elliptical processes and their applications in radar detection (2015).

[11] F. Gini, M. V. Greco, A. Farina, P. Lombardo, Optimum and mismatched detection against k-distributed plus gaussian clutter, IEEE Transactions on Aerospace and Electronic Systems 34 (3) (1998) 860–876.

[12] M. Bilodeau, D. Brenner, Theory of Multivariate Statistics, New York, NY, USA:Springer-Verlag, 1999.

[13] E. Ollila, D. E. Tyler, V. Koivunen, H. V. Poor, Complex elliptically symmetric distributions: Survey, new results and applications, Signal Processing, IEEE Transactions on 60 (11) (2012) 5597–5625.

[14] J. Billingsley, A. Farina, F. Gini, M. Greco, L. Verrazzani, Statistical analyses of measured radar ground clutter data, IEEE Trans. on Aero. and Elec. Syst. 35 (2) (1999) 579 – 593.

[15] S. Watts, Radar detection prediction in sea clutter using the compound k-distribution model, IEE Proceeding, Part. F 132 (7) (1985) 613–620.

[16] F. Pascal, J.-P. Ovarlez, P. Forster, P. Larzabal, On a sirv-cfar detector with radar experimentations in impulsive clutter, in: 2006 14th European Signal Processing Conference (EUSIPCO), IEEE, 2006, pp. 1–5.

[17] F. Gini, Sub-optimum coherent radar detection in a mixture of K-distributed and Gaussian clutter, IEEE Proceedings - Radar, Sonar and Navigation 144 (1) (1997) 39–48.

[18] K. J. Sangston, F. Gini, M. V. Greco, A. Farina, Structures for radar detection in compound Gaussian clutter, IEEE Transactions on Aerospace and Electronic Systems 35 (2) (1999) 445–458.

[19] K. J. Sangston, F. Gini, M. S. Greco, Coherent radar target detection in heavy-tailed compound-Gaussian clutter, IEEE Transactions on Aerospace and Electronic Systems 48 (1) (2012) 64–77.

[20] F. Gini, M. Greco, Suboptimum approach to adaptive coherent radar detection in compound-Gaussian clutter, IEEE Transactions on Aerospace and Electronic Systems 35 (3) (1999) 1095–1104.

[21] E. Conte, A. De Maio, Adaptive radar detection of distributed targets in non-Gaussian noise, in: RADAR, IET, Edinburgh, UK, 2002.

[22] E. Conte, A. De Maio, G. Ricci, Covariance matrix estimation for adaptive cfar detection in compound-gaussian clutter, IEEE Transactions on Aerospace and Electronic Systems 38 (2) (2002) 415–426.

[23] O. Besson, Y. Abramovich, Adaptive detection in elliptically distributed noise and under-sampled scenario, IEEE Signal Processing Letters 21 (12) (2014) 1531–1535.

[24] Y. I. Abramovich, O. Besson, Fluctuating target detection in fluctuating k-distributed clutter, IEEE Signal Processing Letters 22 (10) (2015) 1791–1795.

[25] J.-P. Ovarlez, F. Pascal, A. Breloy, Asymptotic detection performance analysis of the robust adaptive normalized matched filter, in: 2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), IEEE, 2015, pp. 137–140.

[26] F. Pascal, J.-P. Ovarlez, Asymptotic Properties of the Robust ANMF, in: IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-15, Brisbane, Australia, 2015, pp. 2594–2598.

[27] F. Pascal, J.-P. Ovarlez, Asymptotic detection performance of the robust ANMF, in: 2015 23rd European Signal Processing Conference (EUSIPCO), IEEE, 2015, pp. 524–528.

[28] A. De Maio, S. Greco, Modern radar detection theory, The Institution of Engineering and Technology, 2016.

[29] R. A. Maronna, Robust M-estimators of multivariate location and scatter, Annals of Statistics 4 (1) (1976) 51–67.

[30] D. E. Tyler, Radial estimates and the test for sphericity, Biometrika 69 (2) (1982) 429.

[31] M. Mahot, F. Pascal, P. Forster, J.-P. Ovarlez, Asymptotic properties of robust complex covariance matrix estimates, IEEE Transactions on Signal Processing 61 (13) (2013) 3348–3356.

[32] G. Draškovic, F. Pascal, New insights into the statistical properties of m-estimators, IEEE Transactions on Signal Processing 66 (16) (2018) 4253–4263.

[33] D. E. Tyler, A distribution-free M-estimator of multivariate scatter, The Annals of Statistics 15 (1) (1987) 234–251.

[34] F. Gini, M. Greco, Covariance matrix estimation for cfar detection in correlated heavy tailed clutter, Signal Processing 82 (12) (2002) 1847–1859.

[35] E. Conte, A. De Maio, G. Ricci, Recursive estimation of the covariance matrix of a Compound-Gaussian process and its application to adaptive CFAR detection, Signal Processing, IEEE Transactions on 50 (8) (2002) 1908–1915.

[36] F. Pascal, Y. Chitour, J.-P. Ovarlez, P. Forster, P. Larzabal, Covariance structure maximum-likelihood estimates in Compound-Gaussian noise: existence and algorithm analysis, Signal Processing, IEEE Transactions on 56 (1) (2008) 34–48.

[37] E. Ollila, D. E. Tyler, Distribution-free detection under complex elliptically symmetric clutter distribution, in: 2012 IEEE 7th Sensor Array and Multichannel Signal Processing Workshop (SAM), IEEE, 2012, pp. 413–416.

[38] F. Pascal, J. P. Ovarlez, P. Forster, P. Larzabal, Constant false alarm rate detection in spherically invariant random processes, in: 2004 12th European Signal Processing Conference (EUSIPCO), 2004, pp. 2143–2146.

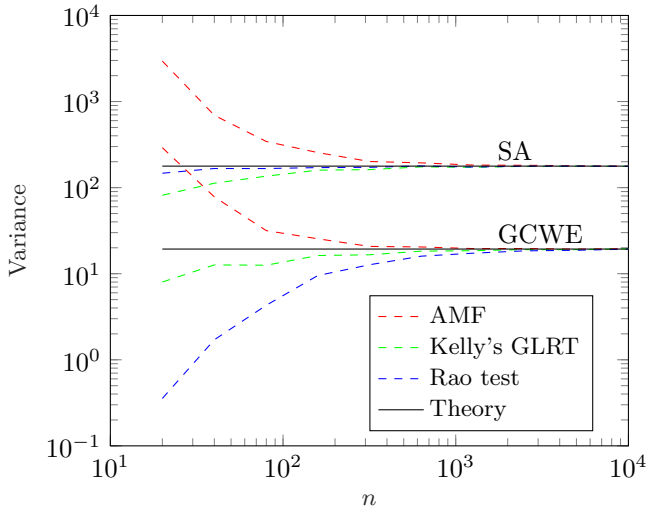
[39] J. T. Kent, Robust properties of likelihood ratio tests, Biometrika 69 (1) (1982) 19–27.

[40] S. Fortunati, M. S. Greco, F. Gini, Asymptotic robustness of kelly's glrt and adaptive matched filter detector under model misspecification, arXiv preprint arXiv:1709.08667 (2017).

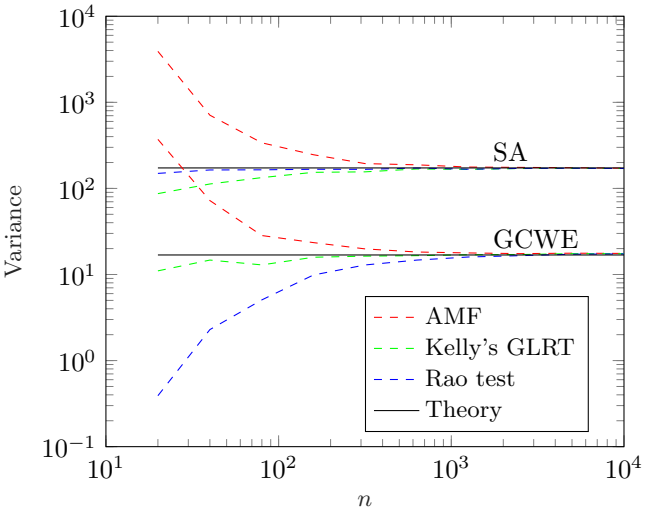
[41] G. Drašković, F. Pascal, New properties for Tyler's covariance matrix estimator, in: 2016 50th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 2016, pp. 820–824.

[42] G. Drašković, F. Pascal, A. Breloy, J.-Y. Tourneret, New asymptotic properties for the robust ANMF, in: IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-17, New Orleans, USA, 2017, pp. 3429–3433.

[43] M. Abramowitz, I. A. Stegun, Handbook of mathematical functions: with formulas, graphs, and mathematical tables, Vol. 55, Courier Corporation, 1964.

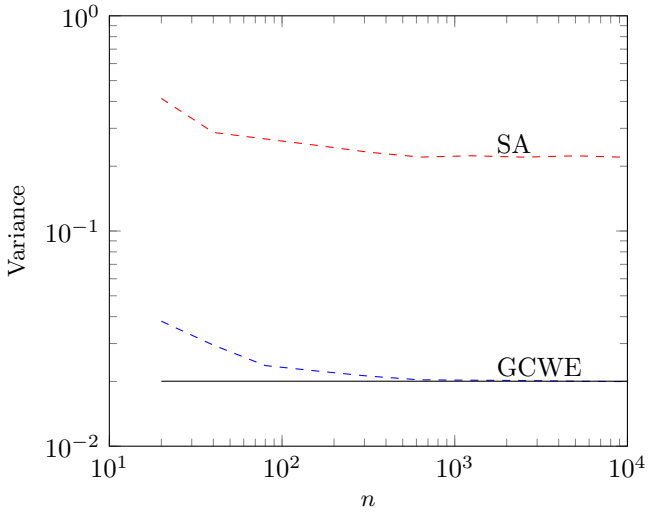


(a) Student's M-estimator

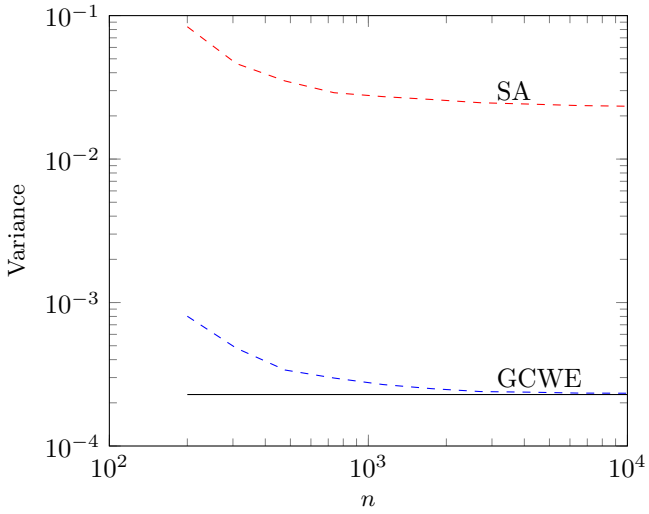


(b) Tyler's M-estimator

Figure A.1: Empirical variances of the robust detectors in standard asymptotic regime and when centered around their GCEDs, compared to the corresponding theoretical results (Examples 4.1 and 4.2); $p = 10$.



(a) $p = 10$



(b) $p = 100$

Figure A.2: Empirical variances of the robust TyE-ANMF in standard asymptotic regime and when centered around its GCED, compared to the theoretical results (Example 4.1).

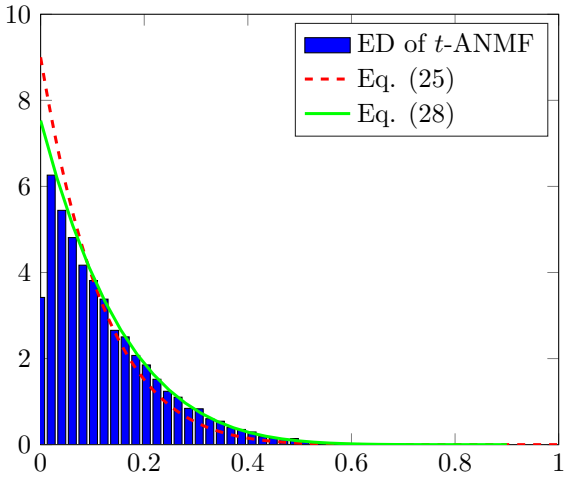


Figure A.3: Empirical distribution of the t-ANMF versus the theoretical distribution of NMF (Eq. (25)) in red (dashed curve) and theoretical approximative distribution (Eq. (28)) in green (solid curve); $p = 10$.

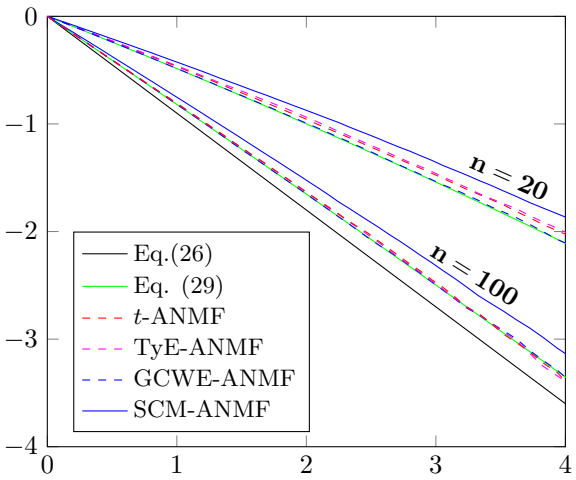


Figure A.4: Comparison between $P_{fa} - \lambda$ relationships for the t-ANMF, TyE-ANMF and SCM-ANMF with the empirical and theoretical results for the GCWE-ANMF (Eq. (29)) and the NMF(Eq. (26)); Student t-distributed data with $d = 2$, $p = 10$.

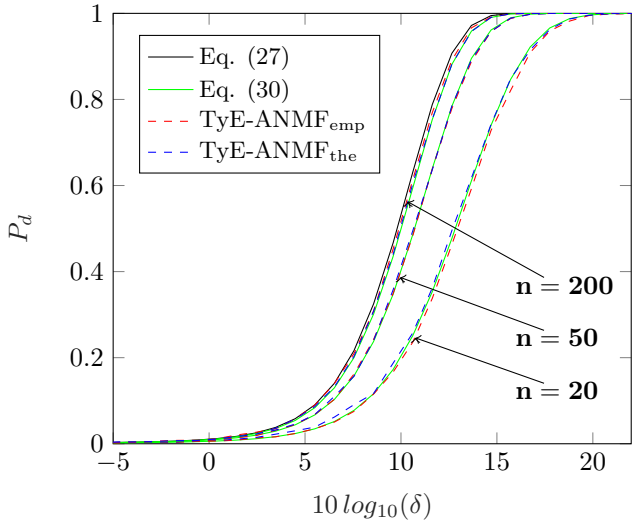


Figure A.5: TyE-ANMF Probability of Detection for $P_{fa} = 0.001$: Comparison of the empirical results for TyE-ANMF obtained for empirically computed threshold (TyE-ANMF$_{\text{emp}}$) and theoretically computed threshold (TyE-ANMF$_{\text{the}}$) using Eq. (29), with the theoretical results for GCED (Eq. (30)) and NMF (Eq. (27)); Student t-distributed data with $d = 2$, $p = 10$.

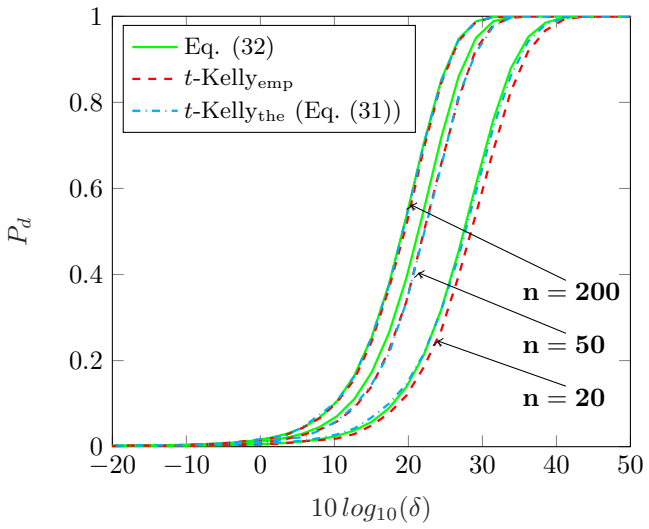


Figure A.6: Probability of Detection of t-Kelly for $P_{fa} = 0.001$: Comparison of the empirical results obtained for empirically computed threshold (t-Kelly$_{\text{emp}}$) and with the theoretically computed threshold using Eq. (31) (t-Kelly$_{\text{the}}$), with the theoretical result for GCED given in Eq. (32); Student t-distributed data with $d = 2$, $p = 10$.