Adaptive Nonzero-Mean Gaussian Detection

Joana Frontera-Pons, Frédéric Pascal, *Senior Member, IEEE*, and Jean-Philippe Ovarlez, *Member, IEEE*

***Abstract*— Classical target detection schemes are usually obtained by deriving the likelihood ratio under Gaussian hypothesis and replacing the unknown background parameters by their estimates. In most applications, interference signals are assumed to be Gaussian with zero mean [or with a known mean vector (MV)] and with an unknown covariance matrix (CM). When the MV is unknown, it has to be jointly estimated with the CM. In this paper, adaptive versions of the classical matched filter (MF) and the normalized MF, as well as two versions of the Kelly detector are first derived and then analyzed for the case where the MV of the background is unknown. More precisely, theoretical closed-form expressions for false alarm (FA) regulation are derived and the constant FA rate property is pursued to allow the detector to be independent of nuisance parameters. Finally, the theoretical contributions are validated through simulations.**

***Index Terms*— Adaptive target detection, false alarm (FA) regulation, nonzero-mean Gaussian distribution.**

I. Introduction

TARGET detection methods have been extensively investigated and analyzed in several signal processing applications and radar processing [1]–[4]. In all these works as well as in several signal processing applications, signals are assumed to be Gaussian with zero mean or with a known mean vector (MV) that can be removed. In such a context, statistical detection theory [5] has led to several well-known algorithms, for instance, the matched filter (MF) and its adaptive versions, the Kelly detector [2] and the adaptive normalized MF (ANMF) [6]. Other interesting approaches based on subspace projection methods have been derived and analyzed in [7]. However, when the MV of the noise background is unknown, these techniques are no longer adapted and improved methods have to be derived by taking into account the MV estimation.

More precisely, this paper deals with the classical adaptive MF (AMF), the Kelly detection test, and the ANMF. These detectors have been derived under Gaussian assumptions and benefit from great popularity in the target detection literature (see [5], [7]). To evaluate the detector performance, the classical process, according to the Neyman–Pearson criterion, is first to regulate the false alarm (FA), by setting a detection threshold for a given probability of FA (PFA). Since the PFA is related to the cumulative distribution function of the detection test, this process is equivalent to the derivation of the detection test statistic. Then, the probability of detection can be evaluated for different signal-to-noise ratios (SNRs). Therefore, keeping the FA rate constant is essential to set a proper detection threshold [8], [9]. The aim is to build a constant FA rate (CFAR) detector that provides detection thresholds that are relatively immune to noise and background variation and allow target detection with a CFAR. The theoretical analysis of CFAR methods for adaptive detectors is a challenging problem, since in adaptive schemes, the statistical distribution of the detectors is not always available in a closed-form expression.

The theoretical contributions of this paper are twofold. First, we derive the expression of each adaptive detector under the Gaussian assumption where both the MV and the covariance matrix (CM) are assumed to be unknown. Then, the exact derivation of the distribution of each proposed detection scheme under null hypothesis, i.e., when no target is supposed to be present, is provided. Thus, through the Gaussian assumption, closed-form expressions for the FA regulation are obtained. This allows one to theoretically set the detection threshold for a given PFA. On the other hand, one major difficulty for the background detection statistic is to assume a tractable model or at least to account for robustness to deviation from the assumed theoretical model in the detection scheme. Since Gaussian assumption is not always fulfilled, alternative robust estimation techniques are proposed in [10]. However, it is essential to observe that the derivations for many results in robust detection framework strongly rely on the results obtained in the Gaussian context. For instance, this is the case in [11], where the derivation of a robust detector distribution is based on its Gaussian counterpart.

One possible application of the detection schemes discussed in this paper is hyperspectral imaging. Hyperspectral sensors measure the radiance of materials within each pixel area at a very large number of contiguous spectral bands and provide image data containing both spatial and spectral information (see [12] and [13] for more details). By exploiting the spectral information, hyperspectral target detection methods can be used to detect targets embedded in the background and that generally cannot be solved by spatial resolution [14]–[17]. Indeed, when the spectral signature of the desired target is known, it can be used as a steering vector similarly to the classical target detection methods studied here [18], [19]. Since hyperspectral data represent the radiation at a large number of wavelengths for each position in an image, they are real and positive. The other approaches found in the literature center the hyperspectral image before applying the detection test, i.e., the global mean of the whole image is removed in a preprocessing stage. This may lead to errors due to the heterogeneity of the background. In order to analyze the

Manuscript received December 9, 2015; revised July 5, 2016; accepted September 26, 2016. Date of publication November 9, 2016; date of current version December 29, 2016.

J. Frontera-Pons is with CEA, Service d'Astrophysique-SEDI, F-91190 Gif-sur-Yvette, France (e-mail: joana.frontera-pons@cea.fr).

F. Pascal is with the Laboratory of Signals and Systems, Centrale-Supélec, 91190 Gif-sur-Yvette, France (e-mail: frederic.pascal@centralesupelec.fr).

J.-P. Ovarlez is with ONERA, DEMR/TSI, 91120 Palaiseau, France (e-mail: jean-philippe.ovarlez@onera.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2016.2619862

proposed techniques in real hyperspectral data, the expressions for FA regulation obtained in this paper should be derived in the real case. Alternatively, the data could be transformed in order to match the hypothesis of complex nonzero mean background model [20]. For instance, using Hilbert transforms is a widely spread method for signal processing in communications [21], [22], which convert real data into complex data without changing the nature of the data. This is beyond the scope of this paper and constitutes the part of perspectives to be further investigated.

This paper is organized as follows. Section II introduces the required background on classical detection techniques as well as the obtention of the adaptive detectors for both unknown MV and CM. Then, Section III provides the main theoretical contributions of this paper by deriving the exact "PFA-threshold" relationship for the AMF, the "plug-in" Kelly detector, and the ANMF under the Gaussian assumption while a generalized version of the Kelly detector is derived. Finally, in Section IV, the theoretical analyses are validated through Monte Carlo simulations. Conclusions and perspectives are drawn in Section V.

II. Background

In the following, vectors (also matrices) are denoted by bold-faced lowercase letters (also uppercase letters). ^{T} and ^{H}, respectively, represent the transpose and the Hermitian operators. $|\mathbf{A}|$ represents the determinant of the matrix $\mathbf{A}$ and $\mathrm{Tr}(\mathbf{A})$ its trace. j is used to denote the unit imaginary number. $\sim$ means "distributed as." $\Gamma(\cdot)$ denotes the gamma function. Eventually, $\Re\{\mathbf{x}\}$ represents the real part of the complex vector $\mathbf{x}$.

After providing the general background in nonzero mean Gaussian detection, this section is devoted to review the expressions of the adaptive detectors.

The problem of detecting a signal corrupted by an additive noise $\mathbf{b}$ in a m-dimensional complex vector $\mathbf{x}$ can be stated as the following binary hypothesis test:

$$\begin{cases} \mathcal{H}_0 : \mathbf{x} = \mathbf{b} & \mathbf{x}_i = \mathbf{b}_i\,,\ i = 1, \dots, N \\ \mathcal{H}_1 : \mathbf{x} = \alpha\mathbf{p} + \mathbf{b}, & \mathbf{x}_i = \mathbf{b}_i\,,\ i = 1, \dots, N \end{cases} \tag{1}$$

where α is an unknown complex scalar amplitude and $\mathbf{p}$ is the steering vector describing the sought signal. Since the background statistics, i.e., the MV and the CM, are assumed to be unknown, they have to be estimated from $\mathbf{x}_i \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, a sequence of N independent and identically distributed (i.i.d.) signal-free secondary data. Then, the adaptive detector is commonly obtained by replacing the unknown parameters by their estimates. In practice, an estimate may be obtained from the pixels surrounding the pixel under test, which play the role of the N i.i.d. signal-free secondary data vectors. The sample size N has to be chosen large enough to ensure the invertibility of the CM and small enough to justify both stationarity and spatial homogeneity. Let us now recall the detectors under interest in this paper.

A. Adaptive Matched Filter

The MF detector is the optimal linear filter for maximizing the SNR in the presence of additive Gaussian noise with known parameters [5]. It corresponds to the generalized likelihood ratio test (GLRT) when the amplitude α of the target to be detected is an unknown parameter.

The MF detection scheme can be written as

$$\Lambda_{MF} = \frac{|\mathbf{p}^H\,\boldsymbol{\Sigma}^{-1}\,(\mathbf{x}-\boldsymbol{\mu})|^2}{(\mathbf{p}^H\,\boldsymbol{\Sigma}^{-1}\,\mathbf{p})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda \tag{2}$$

where $\mathcal{H}_0$ and $\mathcal{H}_1$ denote, respectively, the hypothesis of the absence and the presence of a target to detect and λ is the test threshold. Note that it differs from the classical MF (zero-mean Gaussian noise) by the term $\boldsymbol{\mu}$, the background mean, but without any consequence since $\mathbf{x} - \boldsymbol{\mu} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$. Moreover, the "PFA-threshold" relationship is given by [5]

$$\mathrm{PFA}_{MF} = \exp(-\lambda). \tag{3}$$

The two-step GLRT, called the AMF and denoted $\Lambda^{(N)}_{AMF\,\hat{\boldsymbol{\Sigma}}}$ to underline the dependency with N, is usually built replacing the CM $\boldsymbol{\Sigma}$ by a suitable estimate $\hat{\boldsymbol{\Sigma}}$ obtained from the N secondary data $\{\mathbf{x}_i\}_{i\in[1,N]} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. If we consider a known MV $\boldsymbol{\mu}$, the adaptive version becomes

$$\Lambda^{(N)}_{AMF\,\hat{\boldsymbol{\Sigma}}} = \frac{|\mathbf{p}^H\,\hat{\boldsymbol{\Sigma}}^{-1}\,(\mathbf{x}-\boldsymbol{\mu})|^2}{(\mathbf{p}^H\,\hat{\boldsymbol{\Sigma}}^{-1}\,\mathbf{p})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda. \tag{4}$$

By choosing $\hat{\boldsymbol{\Sigma}} = \hat{\boldsymbol{\Sigma}}_{\mathrm{CSCM}}$, where $\hat{\boldsymbol{\Sigma}}_{\mathrm{CSCM}}$ is the centered sample CM (CSCM) defined in the Appendix, the theoretical "PFA threshold" relationship related to the test given in (1) is given by [3]

$$\mathrm{PFA}_{AMF\,\hat{\boldsymbol{\Sigma}}} = {}_2F_1\left(N-m+1,\ N-m+2;\ N+1;\ -\frac{\lambda}{N}\right) \tag{5}$$

where ${}_2F_1(\cdot)$ is the hypergeometric function [23] defined as

$${}_2F_1(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 \frac{t^{b-1}(1-t)^{c-b-1}}{(1-tz)^a}\,dt. \tag{6}$$

This detector holds the CFAR property, meaning that its FA expression depends only on the dimension of the vector m and the number N of secondary data used for estimation, thus being independent on the noise CM $\boldsymbol{\Sigma}$ and the MV $\boldsymbol{\mu}$.

B. Kelly Detector

The adaptive Kelly detector was derived in [2] and it is based on the GLRT assuming Gaussian distribution. In this case, only the CM $\boldsymbol{\Sigma}$ is unknown and the MV $\boldsymbol{\mu}$ is assumed to be known.

The Kelly detection test is obtained according to

$$\Lambda^{(N)}_{\mathrm{Kelly}\,\hat{\boldsymbol{\Sigma}}_{\mathrm{CSCM}}} \tag{7}$$

$$= \frac{|\mathbf{p}^H\,\hat{\boldsymbol{\Sigma}}^{-1}_{\mathrm{CSCM}}\,(\mathbf{x}-\boldsymbol{\mu})|^2}{\left(\mathbf{p}^H\,\hat{\boldsymbol{\Sigma}}^{-1}_{\mathrm{CSCM}}\mathbf{p}\right)\left(N + (\mathbf{x}-\boldsymbol{\mu})^H\,\hat{\boldsymbol{\Sigma}}^{-1}_{\mathrm{CSCM}}\,(\mathbf{x}-\boldsymbol{\mu})\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_0}{\gtrless}} \lambda. \tag{8}$$

As shown in [2], the PFA for the Kelly test is given by

$$\mathrm{PFA}_{\mathrm{Kelly}} = (1-\lambda)^{N-m+1}. \tag{9}$$

The Kelly detector also satisfies the CFAR property. The AMF (two-step GLRT-based) and the Kelly detector (GLRT-based) have been derived on the same assumptions about the nature of the observations. It is therefore interesting to compare their detection performance for a given PFA. Note that for large values of N, the performances are substantially the same.

C. Adaptive Normalized Matched Filter

The NMF [24] was obtained in Gaussian noise hypothesis, but when considering that the CM is of the form $\sigma^2\,\mathbf{\Sigma}$ with an unknown variance σ^2 but known structure $\mathbf{\Sigma}$. The GLRT leads to

$$\Lambda_{NMF} = \frac{|\mathbf{p}^H\,\mathbf{\Sigma}^{-1}\,(\mathbf{x}-\boldsymbol{\mu})|^2}{(\mathbf{p}^H\,\mathbf{\Sigma}^{-1}\mathbf{p})\,\left((\mathbf{x}-\boldsymbol{\mu})^H\,\mathbf{\Sigma}^{-1}\,(\mathbf{x}-\boldsymbol{\mu})\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda \qquad (10)$$

The PFA–threshold relationship is given by [24]

$$\mathrm{PFA}_{NMF} = (1-\lambda)^{(m-1)}. \qquad (11)$$

The two-step GLRT for this specific covariance structure, referred to as ANMF, is generally obtained when the unknown noise CM $\mathbf{\Sigma}$ is replaced by an estimate [7]

$$\Lambda^{(N)}_{\mathrm{ANMF}\,\hat{\mathbf{\Sigma}}} = \frac{|\mathbf{p}^H\,\hat{\mathbf{\Sigma}}^{-1}\,(\mathbf{x}-\boldsymbol{\mu})|^2}{\left(\mathbf{p}^H\,\hat{\mathbf{\Sigma}}^{-1}\mathbf{p}\right)\,\left((\mathbf{x}-\boldsymbol{\mu})^H\,\hat{\mathbf{\Sigma}}^{-1}\,(\mathbf{x}-\boldsymbol{\mu})\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda. \qquad (12)$$

For the choice for $\hat{\mathbf{\Sigma}} = \hat{\mathbf{\Sigma}}_{\mathrm{CSCM}}$, the PFA follows [7]:

$$\mathrm{PFA}_{\mathrm{ANMF}\,\hat{\mathbf{\Sigma}}_{\mathrm{CSCM}}} = (1-\lambda)^{a-1}\,{}_2F_1(a, a-1; b-1; \lambda) \qquad (13)$$

where $a = N - m + 2$ and $b = N + 2$.

III. Main Results

In this section, let us now assume that the MV $\boldsymbol{\mu}$ is an unknown parameter and let us derive the new corresponding detection schemes. Then, using standard calculus on Wishart distributions, recapped in the Appendix, the distributions of each detection test are provided.

A. Adaptive Matched Filter Detector

When both CM $\mathbf{\Sigma}$ and MV $\boldsymbol{\mu}$ are unknown, the two-step GLRT procedure consists in replacing them by their estimates $\hat{\mathbf{\Sigma}}$ and $\hat{\boldsymbol{\mu}}$ built from the N secondary data $\{\mathbf{x}_i\}_{i\in[1,N]}$ in (2) leading to the AMF detector of the following form:

$$\Lambda^{(N)}_{AMF\,\hat{\mathbf{\Sigma}},\hat{\boldsymbol{\mu}}} = \frac{|\mathbf{p}^H\,\hat{\mathbf{\Sigma}}^{-1}\,(\mathbf{x}-\hat{\boldsymbol{\mu}})|^2}{(\mathbf{p}^H\,\hat{\mathbf{\Sigma}}^{-1}\,\mathbf{p})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda \qquad (14)$$

where the notation $\Lambda^{(N)}_{AMF\,\hat{\mathbf{\Sigma}},\hat{\boldsymbol{\mu}}}$ is used to stress now the dependency on the estimated MV $\hat{\boldsymbol{\mu}}$. Under the Gaussian assumption, and for the particular maximum likelihood estimate (MLE) choice $\hat{\mathbf{\Sigma}} = \hat{\mathbf{\Sigma}}_{\mathrm{SCM}}$ and $\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}_{SMV}$ defined in the Appendix, the distribution of this detection test is given in the next proposition, through its PFA.

Proposition 1: Under Gaussian assumptions, the theoretical relationship between the PFA and the threshold λ is given by

$$\mathrm{PFA}_{AMF\,\hat{\mathbf{\Sigma}},\hat{\boldsymbol{\mu}}} = {}_2F_1\left(N-m,\ N-m+1;\ N;\ -\frac{\lambda}{N+1}\right) \qquad (15)$$

where $\hat{\mathbf{\Sigma}} = \hat{\mathbf{\Sigma}}_{\mathrm{SCM}}$ and $\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}_{SMV}$.

Before turning into the proof, let us comment on this result.

1) Interestingly, this detector also holds the CFAR property in the sense that its FA expression depends only on the dimension m and on the number of secondary data N, but not on the noise parameters $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$. Note that the only effect of estimating the mean is the loss of 1 DOF and the modification of the threshold compared with (5). Obviously, the impact of these modifications decreases as the number of secondary data N used to estimate the unknown parameters increases.
2) Moreover, the result has been obtained when using the MLEs of the unknown parameters, but the proof can be easily extended to other CM estimators such as $\hat{\mathbf{\Sigma}} = \dfrac{1}{N-1}\displaystyle\sum_{i=1}^{N}(\mathbf{x}_i-\hat{\boldsymbol{\mu}})(\mathbf{x}_i-\hat{\boldsymbol{\mu}})^H$, which is the unbiased CM estimate.

Proof: For simplicity matters, the following notations are used: $\hat{\mathbf{\Sigma}} = \hat{\mathbf{\Sigma}}_{\mathrm{SCM}}$ and $\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}_{SMV}$.

Since the derivation of the PFA is done under hypothesis $\mathcal{H}_0$, let us set $\{\mathbf{x}_i\}_{i\in[1,N]} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{\Sigma})$ and $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{\Sigma})$, where all these vectors are independent. Now, let us denote

$$\hat{\mathbf{W}}_{N-1} = \sum_{i=1}^{N}(\mathbf{x}_i-\hat{\boldsymbol{\mu}})(\mathbf{x}_i-\hat{\boldsymbol{\mu}})^H = N\,\hat{\mathbf{\Sigma}}_{\mathrm{SCM}}. \qquad (16)$$

Note that as an application of the Cochran theorem (see [25]), one has

$$\hat{\mathbf{W}}_{N-1} \overset{\text{dist.}}{=} \sum_{i=1}^{N-1}(\mathbf{x}_i-\boldsymbol{\mu})(\mathbf{x}_i-\boldsymbol{\mu})^H = (N-1)\,\hat{\mathbf{\Sigma}}_{\mathrm{CSCM}} \qquad (17)$$

where $\overset{\text{dist.}}{=}$ means *is equal in distribution to*.

Since $\hat{\boldsymbol{\mu}} \sim \mathcal{CN}\left(\boldsymbol{\mu}, \dfrac{1}{N}\mathbf{\Sigma}\right)$, one has $\mathbf{x}-\hat{\boldsymbol{\mu}} \sim \mathcal{CN}\left(\mathbf{0}, \dfrac{N+1}{N}\mathbf{\Sigma}\right)$. This can be equivalently rewritten as

$$\sqrt{N/(N+1)}(\mathbf{x}-\hat{\boldsymbol{\mu}}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}). \qquad (18)$$

Now, let us set $\mathbf{y} = (N/(N+1))^{(1/2)}\left(\mathbf{x}-\hat{\boldsymbol{\mu}}\right)$ with $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$.

As we jointly estimate the mean and the CM, 1 DOF is lost, compared with the case when only the CM is unknown.

Let us now consider the classical AMF test (i.e., $\boldsymbol{\mu}$ known) built from $N-1$ secondary data, rewritten in terms of $\hat{\mathbf{W}}_{N-1}$

$$\Lambda^{(N-1)}_{AMF\ \hat{\mathbf{\Sigma}}} = (N-1)\,\frac{|\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{y}|^2}{(\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{p})}, \qquad (19)$$

where $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ and whose "PFA-threshold" relationship is given by (5), where N is replaced by $N-1$.

Now, for the joint estimation problem, the AMF can be rewritten as

$$\begin{aligned}\Lambda^{(N)}_{AMF\,\hat{\mathbf{\Sigma}},\hat{\boldsymbol{\mu}}} &= N\,\frac{|\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,(\mathbf{x}-\hat{\boldsymbol{\mu}})|^2}{(\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{p})} && (20)\\ &= N\,\frac{N+1}{N}\,\frac{|\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{y}|^2}{(\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{p})} && (21)\\ &\overset{\text{dist.}}{=} \frac{(N+1)}{(N-1)}\,\Lambda^{(N-1)}_{AMF\ \hat{\mathbf{\Sigma}}} && (22)\end{aligned}$$

where $(\mathbf{x}-\hat{\boldsymbol{\mu}})$ has been replaced by $\sqrt{N+1/N}\,\mathbf{y}$ with $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$, as previously stated.

Hence, one can determine the FA relationship

$$\mathrm{PFA}_{AMF\,\hat{\boldsymbol{\Sigma}},\hat{\boldsymbol{\mu}}} = \mathbb{P}\left(\Lambda^{(N)}_{AMF\,\hat{\boldsymbol{\Sigma}},\hat{\boldsymbol{\mu}}} > \lambda|\mathcal{H}_0\right) \tag{23}$$

$$= \mathbb{P}\left(\frac{(N+1)}{(N-1)}\Lambda^{(N-1)}_{AMF\,\hat{\boldsymbol{\Sigma}}} > \lambda|\mathcal{H}_0\right) \tag{24}$$

$$= \mathbb{P}\left(\Lambda^{(N-1)}_{AMF\,\hat{\boldsymbol{\Sigma}}} > \lambda'|\mathcal{H}_0\right) \tag{25}$$

where $\lambda' = ((N-1)/(N+1))\lambda$, which leads to the conclusion. ■

B. Kelly Detector

The exact GLRT Kelly detector for both unknown MV $\boldsymbol{\mu}$ and CM $\boldsymbol{\Sigma}$ is now derived since it does not correspond to the Kelly detector given in (7) in which an estimate of the mean is simply plugged (two-step GLRT). Following the same lines as in [2], we now assume that both the MV and the CM are unknown. The likelihood functions under $\mathcal{H}_0$ and $\mathcal{H}_1$ are given by

$$f_i(\mathbf{x}) = \left(\frac{1}{\pi^m|\boldsymbol{\Sigma}|}\exp\left[-\mathrm{Tr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{T}_i\right)\right]\right)^{N+1} \tag{26}$$

for $i \in \{0, 1\}$, where

$$(N+1)\,\mathbf{T}_0 = (\mathbf{x}-\boldsymbol{\mu}_0)(\mathbf{x}-\boldsymbol{\mu}_0)^H + \sum_{i=1}^{N}(\mathbf{x}_i-\boldsymbol{\mu}_0)(\mathbf{x}_i-\boldsymbol{\mu}_0)^H \tag{27}$$

$$(N+1)\,\mathbf{T}_1 = (\mathbf{x}-\alpha\,\mathbf{p}-\boldsymbol{\mu}_1)(\mathbf{x}-\alpha\,\mathbf{p}-\boldsymbol{\mu}_1)^H + \sum_{i=1}^{N}(\mathbf{x}_i-\boldsymbol{\mu}_1)(\mathbf{x}_i-\boldsymbol{\mu}_1)^H \tag{28}$$

and

$$\boldsymbol{\mu}_0 = \frac{1}{N+1}\left(\mathbf{x}+\sum_{i=1}^{N}\mathbf{x}_i\right) \tag{29}$$

$$\boldsymbol{\mu}_1 = \frac{1}{N+1}\left(\mathbf{x}-\alpha\,\mathbf{p}+\sum_{i=1}^{N}\mathbf{x}_i\right). \tag{30}$$

Under $\mathcal{H}_0$ and $\mathcal{H}_1$, the maxima are achieved at

$$\max_{\boldsymbol{\Sigma},\boldsymbol{\mu}} f_i = \left(\frac{1}{(\pi e)^m|\mathbf{T}_i|}\right)^{N+1} \quad \text{for } i = 0, 1 \tag{31}$$

and taking the $(N+1)$th root, one obtains the following statistic:

$$L(\alpha) = \frac{|\mathbf{T}_0|}{|\mathbf{T}_1|} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta. \tag{32}$$

Then, as this statistic still depends on the unknown amplitude α of the signal, it has to be maximized with respect to α, which is equivalent to minimize $\mathbf{T}_1$ with respect to α. A way to do this is to introduce the following sample CM:

$$\mathbf{S}_0 = \sum_{i=1}^{N}(\mathbf{x}_i-\boldsymbol{\mu}_0)(\mathbf{x}_i-\boldsymbol{\mu}_0)^H. \tag{33}$$

Then, $(N+1)\,|\mathbf{T}_0|$ can be written as

$$(N+1)\,|\mathbf{T}_0| = |\mathbf{S}_0|\,\left(1+(\mathbf{x}-\boldsymbol{\mu}_0)^H\,\mathbf{S}_0^{-1}\,(\mathbf{x}-\boldsymbol{\mu}_0)\right). \tag{34}$$

In the same way, after some manipulations, $(N+1)\,|\mathbf{T}_1|$ becomes

$$\begin{aligned}(N+1)\,|\mathbf{T}_1| &= |\mathbf{S}_0|\,\left(\sum_{i=1}^{N}(\mathbf{x}_i-\boldsymbol{\mu}_1)^H\,\mathbf{S}_0^{-1}\,(\mathbf{x}_i-\boldsymbol{\mu}_1)\right.\\ &\quad\left.+\,(\mathbf{x}-\alpha\,\mathbf{p}-\boldsymbol{\mu}_1)^H\,\mathbf{S}_0^{-1}(\mathbf{x}-\alpha\,\mathbf{p}-\boldsymbol{\mu}_1)\right)\\ &= |\mathbf{S}_0|\,(A+B).\end{aligned} \tag{35}$$

Now, let us rewrite the two terms A and B to separate the terms involving α. By recalling that $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_0-((1/(N+1))\alpha\,\mathbf{p}$, one obtains

$$\begin{aligned}A &= 1+\frac{N\,|\alpha|^2}{(N+1)^2}\mathbf{p}^H\,\mathbf{S}_0^{-1}\,\mathbf{p}+\frac{2}{N+1}\\ &\quad\times\Re\left\{\bar{\alpha}\,\mathbf{p}^H\,\mathbf{S}_0^{-1}\,\sum_{i=1}^{N}(\mathbf{x}_i-\boldsymbol{\mu}_0)\right\}\end{aligned} \tag{36}$$

$$\begin{aligned}B &= (\mathbf{x}-\boldsymbol{\mu}_0)^H\,\mathbf{S}_0^{-1}\,(\mathbf{x}-\boldsymbol{\mu}_0)+\frac{N^2\,|\alpha|^2}{(N+1)^2}\mathbf{p}^H\,\mathbf{S}_0^{-1}\,\mathbf{p}\\ &\quad-\frac{2N}{N+1}\Re\left\{\bar{\alpha}\,\mathbf{p}^H\,\mathbf{S}_0^{-1}\,(\mathbf{x}-\boldsymbol{\mu}_0)\right\}.\end{aligned} \tag{37}$$

Observing that $\mathbf{x}-\boldsymbol{\mu}_0 = -\sum_{i=1}^{N}(\mathbf{x}_i-\boldsymbol{\mu}_0)$ and then rearranging the expression of $(N+1)|\mathbf{T}_1|$, one has

$$\begin{aligned}\frac{(N+1)\,|\mathbf{T}_1|}{|\mathbf{S}_0|} &= \frac{(N+1)\,|\mathbf{T}_0|}{|\mathbf{S}_0|}+\frac{N\,|\alpha|^2}{(N+1)}\mathbf{p}^H\,\mathbf{S}_0^{-1}\,\mathbf{p}\\ &\quad-2\,\Re\left\{\bar{\alpha}\,\mathbf{p}^H\,\mathbf{S}_0^{-1}\,(\mathbf{x}-\boldsymbol{\mu}_0)\right\}.\end{aligned} \tag{38}$$

Now, the term depending on α can be rewritten as follows:

$$\begin{aligned}&\frac{N}{(N+1)}\mathbf{p}^H\,\mathbf{S}_0^{-1}\,\mathbf{p}\left|\alpha-\frac{N+1}{N}\frac{\mathbf{p}^H\,\mathbf{S}_0^{-1}\,(\mathbf{x}-\boldsymbol{\mu}_0)}{\mathbf{p}^H\,\mathbf{S}_0^{-1}\,\mathbf{p}}\right|^2\\ &\quad-\frac{N+1}{N}\frac{\left|\mathbf{p}^H\,\mathbf{S}_0^{-1}\,(\mathbf{x}-\boldsymbol{\mu}_0)\right|^2}{\mathbf{p}^H\,\mathbf{S}_0^{-1}\,\mathbf{p}}.\end{aligned} \tag{39}$$

Minimizing $|\mathbf{T}_1|$ with respect to α is equivalent to cancelling the square term in the previous equation. Thus, the GLRT can now be written according to the following definition.

Definition 2 (Generalized Kelly Detector): Under Gaussian assumptions, the extension of Kelly's test when both the MV and the CM of the background are unknown takes the following form:

$$\Lambda = \frac{\beta(N)\,\left|\mathbf{p}^H\mathbf{S}_0^{-1}(\mathbf{x}-\boldsymbol{\mu}_0)\right|^2}{(\mathbf{p}^H\mathbf{S}_0^{-1}\mathbf{p})\,\left(1+(\mathbf{x}-\boldsymbol{\mu}_0)^H\,\mathbf{S}_0^{-1}\,(\mathbf{x}-\boldsymbol{\mu}_0)\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda \tag{40}$$

where $\beta(N) = ((N+1)/N)$ and

$$\mathbf{S}_0 = \sum_{i=1}^{N}(\mathbf{x}_i-\boldsymbol{\mu}_0)(\mathbf{x}_i-\boldsymbol{\mu}_0)^H$$

$$\boldsymbol{\mu}_0 = \frac{1}{N+1}\left(\mathbf{x}+\sum_{i=1}^{N}\mathbf{x}_i\right).$$

Let us now comment on this new detector. One can observe that both the CM $\mathbf{S}_0$ and mean $\boldsymbol{\mu}_0$ estimates depend on the data $\mathbf{x}$ under test, which is not the case in other classical detectors where the unknown parameters are estimated from signal-free secondary data. Consequently, $\mathbf{S}_0$ and $\mathbf{x}-\boldsymbol{\mu}_0$ are not independent. Moreover, the CM estimate $\mathbf{S}_0$ is not Wishart distributed due to the nonstandard mean estimate $\boldsymbol{\mu}_0$. Thus, the derivation of this ratio distribution is very difficult.

As for previous detector, it would be intuitive to think that the proposed test behaves as the classical Kelly's test but for $N-1$ degrees of freedom. To prove that, let us first rewrite (40) as follows:

$$\Lambda = \frac{\left|\mathbf{p}^H\,\mathbf{S}_0^{-1}\,\mathbf{y}\right|^2}{\left(\mathbf{p}^H\,\mathbf{S}_0^{-1}\,\mathbf{p}\right)\,\left(1+\dfrac{N}{N+1}\mathbf{y}^H\,\mathbf{S}_0^{-1}\,\mathbf{y}\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda \qquad (41)$$

where we use

$$\begin{aligned}(\mathbf{x}-\boldsymbol{\mu}_0) &= \frac{N}{N+1}(\mathbf{x}-\hat{\boldsymbol{\mu}}_{SMV})\\ \hat{\boldsymbol{\mu}}_{SMV} &= \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\\ \mathbf{y} &= \sqrt{\frac{N}{N+1}}\,(\mathbf{x}-\hat{\boldsymbol{\mu}}_{SMV}) \sim \mathcal{CN}(\mathbf{0},\boldsymbol{\Sigma}).\end{aligned}$$

Now, let us set $\mathbf{S}_0^{(i)} = \sum_{i=1}^{N}\left(\mathbf{x}_i-\boldsymbol{\mu}_0^{(i)}\right)\left(\mathbf{x}_i-\boldsymbol{\mu}_0^{(i)}\right)^H$, where $\boldsymbol{\mu}_0^{(i)} = (1/N)\left(\sum_{j\neq i}^{N}\mathbf{x}_j+\mathbf{x}\right)$. Then, the test becomes

$$\frac{\dfrac{N+1}{N}\left|\mathbf{p}^H\,\mathbf{S}_0^{(i)^{-1}}\,\left(\mathbf{x}-\hat{\boldsymbol{\mu}}_{SMV}\right)\right|^2}{\left(\mathbf{p}^H\,\mathbf{S}_0^{(i)^{-1}}\,\mathbf{p}\right)\,\left(1+\left(\mathbf{x}-\hat{\boldsymbol{\mu}}_{SMV}\right)^H\,\mathbf{S}_0^{(i)^{-1}}\,\left(\mathbf{x}-\hat{\boldsymbol{\mu}}_{SMV}\right)\right)}. \qquad (42)$$

One can observe that each $\mathbf{x}_i$ (including $\mathbf{x}$) plays the same role, and thus the distribution of this test is the same for every permutation of the $(N+1)$-sample $(\mathbf{x},\mathbf{x}_1,\ldots,\mathbf{x}_N)$. However, the dependency between the CM estimate and the data under test $\mathbf{x}$ still remains.

To fill this gap, another way of taking advantage of the Kelly's detector when the MV is unknown can be to use the classical scheme recalled in (7) and to plug the classical estimator of the mean, based only on the N secondary data, i.e., $\hat{\boldsymbol{\mu}}_{SMV} = (1/N)\sum_{i=1}^{N}\mathbf{x}_i$. This leads to the two-step GLRT Kelly's detector

$$\begin{aligned}&\Lambda^{(N)}_{\text{Kelly}\,\hat{\boldsymbol{\Sigma}}_{\text{SCM}},\hat{\boldsymbol{\mu}}_{SMV}}\\ &= \frac{|\mathbf{p}^H\hat{\boldsymbol{\Sigma}}_{\text{SCM}}^{-1}(\mathbf{x}-\hat{\boldsymbol{\mu}}_{SMV})|^2}{\left(\mathbf{p}^H\hat{\boldsymbol{\Sigma}}_{\text{SCM}}^{-1}\mathbf{p}\right)\left(N+(\mathbf{x}-\hat{\boldsymbol{\mu}}_{SMV})^H\hat{\boldsymbol{\Sigma}}_{\text{SCM}}^{-1}(\mathbf{x}-\hat{\boldsymbol{\mu}}_{SMV})\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda.\end{aligned} \qquad (43)$$

In this case, the distribution can be derived. This is the purpose of the following proposition.

Proposition 3: The theoretical relationship between the PFA and the threshold is given by

$$\begin{aligned}\text{PFA}_{\text{Kelly}\,\hat{\boldsymbol{\Sigma}}_{\text{SCM}},\hat{\boldsymbol{\mu}}_{SMV}} &= \frac{\Gamma(N)}{\Gamma(N-m+1)\,\Gamma(m-1)}\\ &\times\int_0^1\left[1+\frac{\lambda}{1-\lambda}\left(1-\frac{u}{N+1}\right)\right]^{m-N}u^{N-m}(1-u)^{m-2}\,du.\end{aligned} \qquad (44)$$

Proof: The detection test rewritten with $\hat{\boldsymbol{\Sigma}}_{\text{SCM}}^{-1} = N\,\hat{\mathbf{W}}_{N-1}^{-1}$ becomes

$$\Lambda^{(N)}_{\text{Kelly}\,\hat{\boldsymbol{\Sigma}},\hat{\boldsymbol{\mu}}} = \frac{N^2\,\left|\mathbf{p}^H\hat{\mathbf{W}}_{N-1}^{-1}(\mathbf{x}-\hat{\boldsymbol{\mu}})\right|^2}{N\,\left(\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{p}\right)\,\left(N+N\,\mathbf{y}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\left(\mathbf{x}-\hat{\boldsymbol{\mu}}\right)\right)} \qquad (45)$$

and replacing $(\mathbf{x}-\hat{\boldsymbol{\mu}})$ by $((N+1)/N)^{(1/2)}\,\mathbf{y}$, one obtains

$$\begin{aligned}\Lambda^{(N)}_{\text{Kelly}\,\hat{\boldsymbol{\Sigma}},\hat{\boldsymbol{\mu}}} &= \frac{\dfrac{N+1}{N}N^2\,\left|\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{y}\right|^2}{N\,\left(\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{p}\right)\,\left(N+\dfrac{N+1}{N}N\,\mathbf{y}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{y}\right)} \qquad (46)\\ &= \frac{\left|\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{y}\right|^2}{\left(\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{p}\right)\,\left(\dfrac{N}{N+1}+\mathbf{y}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{y}\right)} \qquad (47)\end{aligned}$$

with $\mathbf{y}\sim\mathcal{CN}(\mathbf{0},\boldsymbol{\Sigma})$.

The classical Kelly detector obtained when the MV is known is recalled here, built with $N-1$ zero-mean Gaussian data, and written with $\hat{\mathbf{W}}_{N-1}$

$$\Lambda^{(N-1)}_{\text{Kelly}\,\hat{\boldsymbol{\Sigma}}} = \frac{\left|\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{y}\right|^2}{\left(\mathbf{p}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{p}\right)\,\left(1+\mathbf{y}^H\,\hat{\mathbf{W}}_{N-1}^{-1}\,\mathbf{y}\right)}. \qquad (48)$$

It is worth pointing out that the term $N/(N+1)$ resulting from the mean estimation in $\Lambda^{(N)}_{\text{Kelly}\,\hat{\boldsymbol{\Sigma}},\hat{\boldsymbol{\mu}}}$ does not appear in the classical Kelly detector (48). This fact prevents from relating the two expressions. Thus, a proof similar to Proposition 1 is not feasible.

According to [4] and [7], an equivalent *LR* can be expressed as

$$\widehat{\kappa}^2 = \frac{\Lambda^{(N)}_{\text{Kelly}\,\hat{\boldsymbol{\Sigma}},\hat{\boldsymbol{\mu}}}}{1-\Lambda^{(N)}_{\text{Kelly}\,\hat{\boldsymbol{\Sigma}},\hat{\boldsymbol{\mu}}}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{\lambda}{1-\lambda}. \qquad (49)$$

Following the same development proposed in [7], the statistic $\widehat{\kappa}^2$ can be identified as the ratio θ/β between two independent scalar random variables θ and β. For this particular development of Kelly distribution with noncentered data, the scalar random variable β is found to have the same distribution as the function $1-u/(N+1)$, where u is a random variable following a complex central beta distribution with $N-m+1$ and $m-1$ degrees of freedom:

$$u\sim f_u(u) = \frac{\Gamma(N)}{\Gamma(N-m+1)\,\Gamma(m-1)}u^{N-m}\,(1-u)^{m-2} \qquad (50)$$

whereas the PDF of the variable θ is distributed according to the complex F-distribution with 1 and $N-m$ degrees of freedom scaled by $1/(N-m)$

$$\theta\sim f_\theta(\theta) = (N-m)\,(1+\theta)^{m-N-1}. \qquad (51)$$

One can now derive the cumulative density function of the Kelly test as

$$\mathbb{P}\left(\Lambda^{(N)}_{\text{Kelly}\,\hat{\boldsymbol{\Sigma}},\hat{\boldsymbol{\mu}}} \leq \lambda\right) = \mathbb{P}\left(\widehat{\kappa}^2 \leq \frac{\lambda}{1-\lambda}\right) = \mathbb{P}\left(\theta \leq \beta\,\frac{\lambda}{1-\lambda}\right) \quad (52)$$

$$= \int_0^1 \left[\int_0^{\frac{\lambda}{1-\lambda}(1-u/(N+1))} f_\theta(v)\,dv\right] f_u(u)\,du. \quad (53)$$

Solving the integral one obtains the "PFA-threshold" relationship

$$\text{PFA}_{\text{Kelly}\,\hat{\boldsymbol{\Sigma}},\hat{\boldsymbol{\mu}}} = \frac{\Gamma(N)}{\Gamma(N-m+1)\,\Gamma(m-1)} \times \int_0^1 \left[1+\frac{\lambda}{1-\lambda}\left(1-\frac{u}{N+1}\right)\right]^{m-N} \times u^{N-m}(1-u)^{m-2}\,du. \quad (54)$$

However, the final expression cannot be further simplified to a closed-form expression as those obtained for the other detectors. ■

C. Adaptive Normalized Matched Filter

Similarly, the ANMF for both MV and CM estimation becomes

$$\Lambda_{\text{ANMF}\,\hat{\boldsymbol{\Sigma}},\hat{\mu}} = \frac{|\mathbf{p}^H\,\hat{\boldsymbol{\Sigma}}^{-1}\,(\mathbf{x}-\hat{\boldsymbol{\mu}})|^2}{(\mathbf{p}^H\,\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{p})\left((\mathbf{x}-\hat{\boldsymbol{\mu}})^H\,\hat{\boldsymbol{\Sigma}}^{-1}\,(\mathbf{x}-\hat{\boldsymbol{\mu}})\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda. \quad (55)$$

Proposition 4: The theoretical relationship between the PFA and the threshold is given by

$$\text{PFA}_{\text{ANMF}\,\hat{\boldsymbol{\Sigma}},\hat{\boldsymbol{\mu}}} = (1-\lambda)^{a-1}\,{}_2F_1\,(a, a-1; b-1; \lambda) \quad (56)$$

where $a = (N-1) - m + 2$, $b = (N-1) + 2$ and $\hat{\boldsymbol{\Sigma}} = \hat{\boldsymbol{\Sigma}}_{\text{SCM}}$ and $\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}_{SMV}$.

Proof: The proof is similar to the proof of Proposition 1. The main difference is due to the normalization term $(\mathbf{x}-\hat{\boldsymbol{\mu}})^H\,\hat{\boldsymbol{\Sigma}}^{-1}\,(\mathbf{x}-\hat{\boldsymbol{\mu}})$. Indeed, the correction factor $N/(N-1)$ appears both at the numerator and at the denominator, and consequently, it disappears. The same argument is also true for the factor N that arises from the CM estimates, i.e., since the detector is homogeneous of degree 0 in terms of CM estimates (i.e., $\Lambda_{\text{ANMF}\,\hat{\boldsymbol{\Sigma}},\hat{\boldsymbol{\mu}}} = \Lambda_{\text{ANMF}_\gamma\,\hat{\boldsymbol{\Sigma}},\hat{\boldsymbol{\mu}}}$ for any real γ), this scalar also disappears. Thus, the distribution of the ANMF with an estimate of the mean is exactly the same as in (13) where N is replaced by $N-1$. ■

IV. Experimental Results

In this section, we validate the theoretical analysis on simulated data. The experiments were conducted on $m = 5$ dimensional Gaussian vectors, for different values of N, the number of secondary data and the computations have been made through 10^6 Monte Carlo trials. The true covariance is chosen as a Toeplitz matrix whose entries are $\Sigma_{i,j} = \rho^{|i-j|}$ and where $\rho = 0.4$. The MV is arbitrarily set to have all entries equal to $(3+4j)$.

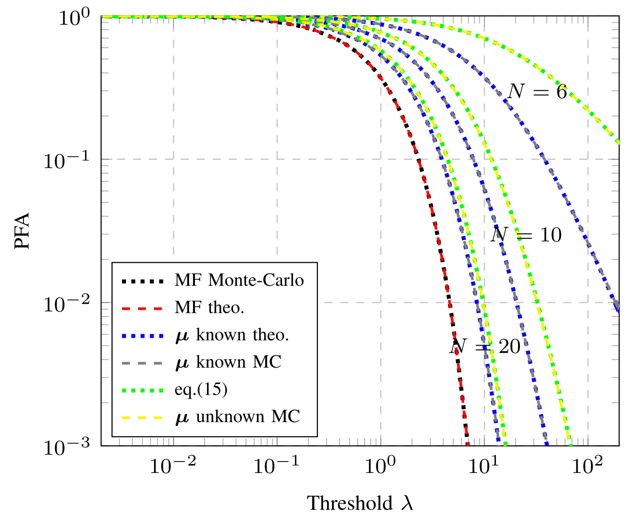

Fig. 1. PFA versus threshold for the AMF for different values of N [$m = 5$, $\mathbf{p} = [1, \ldots, 1]^T$, $\rho = 0.4$, and $\boldsymbol{\mu} = (4+3\,j)\,\mathbf{p}$] when (a) $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are known (MF) (red and black curves), (b) only $\boldsymbol{\mu}$ is known (gray and blue curves), and (c) Proposition 1: both $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are unknown (yellow and green curves).

A. False Alarm Regulation With Simulated Data

The FA regulation is presented for previous detection schemes having a closed-form expression, i.e., for all except the generalized Kelly detector. Fig. 1 shows the FA regulation for the MF, the AMF when only the CM is unknown, and the AMF for both CM and MV unknown. The steering vector used for the simulations is the unity vector $\mathbf{p} = [1, \ldots, 1]^T$ without loss of generality as all the PDFs are found to be independent of the steering vector $\mathbf{p}$. The perfect agreement of the green and yellow curves illustrates the results of Proposition 1. Moreover, remark that when N increases, both AMFs get closer to each other and they approach the known parameter case MF.

Figs. 2 and 3 present the FA regulation for the Kelly detector and the ANMF, respectively, under the Gaussian assumption. For clarity purposes, the results are displayed in terms of the threshold $\eta = (1-\lambda)^{-(N+1)}$ for adaptive Kelly detectors and $\eta = (1-\lambda)^{-m}$ for ANMF and NMF detectors, respectively, and a logarithmic scale is used. This validates the results of Proposition 3 and 4 for the SCM-SMV.

Remark that the derived relationships given by (15) and (56) are quite similar to those for which the mean is known. However, as illustrated in Figs. 1 and 3, there is an important difference for small values of N. It is worth pointing out that the theoretical "PFA-threshold" relationships presented above depend only on the size of the vectors m and the number of secondary data used to estimate the parameters N. Thus, the detector outcome will not depend on the true value of the CM or the MV. These three detectors hold the CFAR property with respect to the background parameters. However, their distribution strongly relies on the underlying distribution of the background, i.e., if Gaussian assumption is not fulfilled, the "PFA-threshold" relationship will divert from the theoretical results derived in this paper.

B. Performance Evaluation

The four detection schemes are compared in terms of probability of detection. The experiments were conducted to

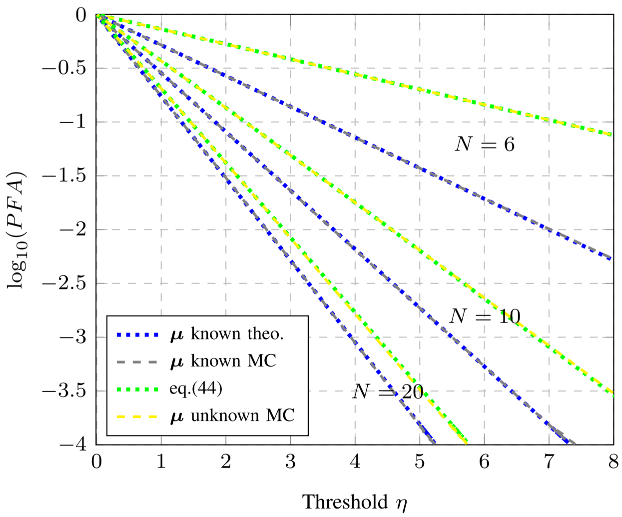

Fig. 2. PFA versus threshold for the "plug-in" Kelly detector for different values of N [$m = 5$, $\mathbf{p} = [1, \ldots, 1]^T$, $\rho = 0.4$, and $\boldsymbol{\mu} = (4 + 3\,j)\,\mathbf{p}$] when (a) only $\boldsymbol{\mu}$ is known (gray and blue curves) and (b) Proposition 3: both $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are unknown (yellow and green curves).

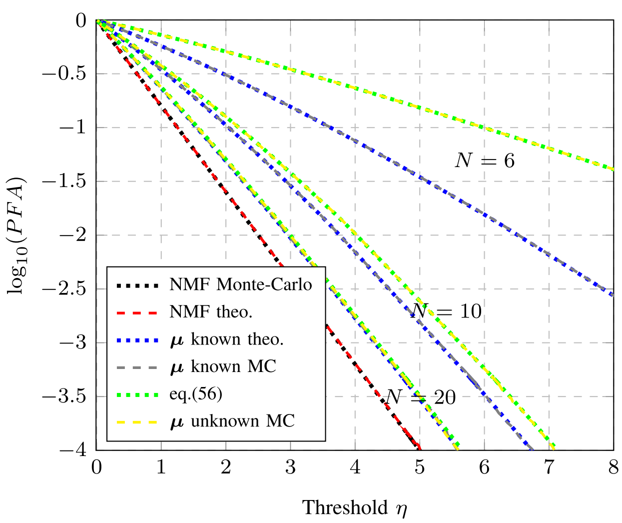

Fig. 3. PFA versus threshold for the ANMF for different values of N [$m = 5$, $\mathbf{p} = [1, \ldots, 1]^T$, $\rho = 0.4$, and $\boldsymbol{\mu} = (4+3\,j)\,\mathbf{p}$] when (a) $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are known (NMF) (red and black curves), (b) only $\boldsymbol{\mu}$ is known (gray and blue curves), and (c) Proposition 4: both $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are unknown (yellow and green curves).

detect a vector $\alpha\,\mathbf{p}$ embedded in Gaussian noise with the same distribution parameters than for FA regulation. The Monte-Carlo simulation was set for dimensions $m = 5$ and $N = 10$ and for the probability of FA $PFA = 10^{-3}$. Then, the threshold λ has been adjusted according to the FA regulation relative to each detectors (AMF, ANMF, two-step GLRT Kelly, Generalized Kelly). Fig. 4 presents the detection probability versus the SNR defined as $\alpha^2\,\mathbf{p}^H\,\boldsymbol{\Sigma}^{-1}\,\mathbf{p}$ with the known steering vector $\mathbf{p} = [1, \ldots, 1]^T$. The detectors delivering the best performance results are the Kelly detectors ("two-step GLRT" and generalized). Actually, these detectors lead to a very similar performance with a small improvement of the generalized (also "two-step GLRT") one at low (also high) SNR. As expected, the AMF and the ANMF require a higher SNR to achieve same performance.

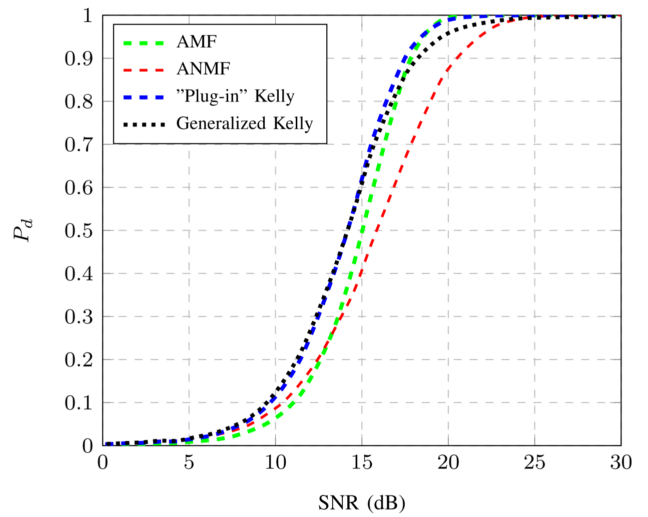

Fig. 4. Probability of detection for $PFA = 10^{-3}$ corresponding to different values of SNR $= \alpha^2\,\mathbf{p}\,\boldsymbol{\Sigma}^{-1}\,\mathbf{p}$ in the Gaussian case ($m = 5$, $N = 20$, $\mathbf{p} = [1, \ldots, 1]^T$, and $\rho = 0.4$).

V. Conclusion

Four adaptive detection schemes, the AMF, Kelly detectors with a "two-step GLRT", and generalized versions and the ANMF, have been analyzed in the case where both the CM and the MV are unknown and need to be estimated. In this context, theoretical closed-form expressions for FA regulation have been derived under the Gaussian assumptions for the SCM-SMV estimates for three detection schemes. The resulting "PFA-threshold" expressions highlight the FA rate constantness of these detectors since they only depend on the size and the number of data but not on the unknown parameters. Finally, the theoretical analysis has been validated through Monte Carlo simulations and the performances of the detectors have been compared in terms of probability of detection. This paper finds its purpose in signal processing methods for which both MV and CM are unknown. Specifically, the proposed methods could be applied for hyperspectral target detection.

Appendix

A m-dimensional vector $\mathbf{x} = \mathbf{u} + j\,\mathbf{v}$ has a complex normal distribution with mean $\boldsymbol{\mu}$ and CM $\boldsymbol{\Sigma} = E[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^H]$, denoted by $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, if $\mathbf{z} = (\mathbf{u}^T, \mathbf{v}^T)^T \in \mathbb{R}^{2m}$ has a normal distribution [26]. If $\operatorname{rank}(\boldsymbol{\Sigma}) = m$, the probability density function exists and is of the form

$$f_{\mathbf{x}}(\mathbf{x}) = \pi^{-m}|\boldsymbol{\Sigma}|^{-1}\exp\{-(\mathbf{x} - \boldsymbol{\mu})^H\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\}. \quad (57)$$

The resulting MLEs are the well-known SCM and SMV defined as

$$\hat{\boldsymbol{\mu}}_{SMV} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i \quad \hat{\boldsymbol{\Sigma}}_{\text{SCM}} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \hat{\boldsymbol{\mu}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^H \quad (58)$$

where the $\mathbf{x}_i$ values are i.i.d. $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Further, we shall denote the CSCM as

$$\hat{\boldsymbol{\Sigma}}_{\text{CSCM}} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^H. \quad (59)$$

Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be an i.i.d. N sample, where $\mathbf{x}_i \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Let us define $\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}_{SMV}$ and $\hat{\mathbf{W}} = N\,\hat{\boldsymbol{\Sigma}}_{\text{SCM}}$ referred to as a Wishart matrix. Thus, one finds (see [27] for the real case) that the following conditions hold true.

1) $\hat{\boldsymbol{\mu}}$ and $\hat{\mathbf{W}}$ are independently distributed.
2) $\hat{\boldsymbol{\mu}} \sim \mathcal{CN}(\boldsymbol{\mu}, (1/N)\boldsymbol{\Sigma})$.
3) $\hat{\mathbf{W}} \sim \mathcal{CW}(N-1, \boldsymbol{\Sigma})$ is Whishart distributed with $N-1$ degrees of freedom.

REFERENCES

[1] F. Gini, A. Farina, and M. Greco, "Selected list of references on radar signal processing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 1, pp. 329–359, Jan. 2001.

[2] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, no. 2, pp. 115–127, Mar. 1986.

[3] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.

[4] C. D. Richmond, "Performance of a class of adaptive detection algorithms in nonhomogeneous environments," *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1248–1262, May 2000.

[5] S. M. Kay, *Fundamentals of Statistical Signal Processing—Detection Theory*, vol. 2. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.

[6] 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, Sep. 1999.

[7] S. Kraut, L. L. Scharf, and L. T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 1–16, Jan. 2001.

[8] F. Gini and M. V. Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter," *Signal Process.*, vol. 82, no. 12, pp. 1847–1859, Dec. 2002.

[9] 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, pp. 1908–1915, Aug. 2002.

[10] J. Frontera-Pons, M. Mahot, J. P. Ovarlez, F. Pascal, and J. Chanussot, "Robust detection using M-estimators for hyperspectral imaging," in *Proc. Workshop Hyperspectral Image Signal Process.*, Jun. 2012, pp. 1–4.

[11] F. Pascal, J.-P. Ovarlez, P. Forster, and P. Larzabal, "On a SIRV-CFAR detector with radar experimentations in impulsive clutter," in *Proc. 14th Eur. Signal Process. Conf.*, Sep. 2006, pp. 1–5.

[12] C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, vol. 1. Springer, 2003.

[13] D. Manolakis, D. Marden, and G. A. Shaw, "Hyperspectral image processing for automatic target detection applications," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 79–116, 2003.

[14] D. Manolakis and G. S. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.

[15] S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 25, no. 7, pp. 5–28, Jul. 2010.

[16] R. M. Willett, M. F. Duarte, M. A. Davenport, and R. G. Baraniuk, "Sparsity and structure in hyperspectral imaging: Sensing, reconstruction, and target detection," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 116–126, Jan. 2014.

[17] H. Kwon and N. M. Nasrabadi, "Kernel matched subspace detectors for hyperspectral target detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 2, pp. 178–194, Feb. 2006.

[18] D. Manolakis, R. Lockwood, T. Cooley, and J. Jacobson, "Is there a best hyperspectral detection algorithm?" in *Proc. SPIE Defense, Secur., Sens.*, 2009, p. 733402.

[19] D. Manolakis, E. Truslow, M. Pieper, T. Cooley, M. Brueggeman, and S. Lipson, "The remarkable success of adaptive cosine estimator in hyperspectral target detection," in *Proc. SPIE Defense, Secur., Sens.*, 2013, p. 874302.

[20] J. Frontera-Pons, F. Pascal, and J. P. Ovarlez, "False-alarm regulation for target detection in hyperspectral imaging," in *Proc. IEEE 5th Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process. (CAMSAP)*, Dec. 2013, pp. 161–164.

[21] J. S. Bendat and A. G. Piersol, "The Hilbert transform," in *Random Data: Analysis and Measurement Procedures*, 4th ed. 1995, pp. 473–503.

[22] V. Cizek, "Discrete Hilbert transform," *IEEE Trans. Audio Electroacoust.*, vol. 18, no. 4, pp. 340–343, Dec. 1970.

[23] M. E. Abramowitz *et al.*, *Handbook of mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, vol. 55. New York, NY, USA: Dover, 1964.

[24] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2146–2157, Aug. 1994.

[25] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. New York, NY, USA: Wiley, 1984.

[26] A. van den Bos, "The multivariate complex normal distribution—A generalization," *IEEE Trans. Inf. Theory*, vol. 41, no. 2, pp. 537–539, Mar. 1995.

[27] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*. London, U.K.: Chapman & Hall, 2000.

Joana Frontera-Pons was born in Mallorca, Spain, in 1986. She received the Telecommunications Engineering Degree from the Technical University of Catalonia, Barcelona, Spain, in 2011, and the Ph.D. degree in signal processing from Supélec, Gif-sur-Yvette, France, in 2014.

Since 2015, she has been a Post-Doctoral Researcher with the CosmoStat team at CEA Saclay, Gif-sur-Yvette. Her current research interests include estimation and target detection in statistical signal processing and hyperspectral imaging.

Frédéric Pascal (S'04–M'06–SM'14) received the master's degree (Hons.) (probabilities, statistics and applications: signal, image, and networks) in applied statistics from University Paris VII - Jussieu, Paris, France, in 2003, the Ph.D. degree in signal processing from University Paris X - Nanterre, Nanterre, France, in 2006, and the Research Directorship Habilitation thesis in signal processing from the University of Paris-Sud, Orsay, France, in 2012.

From 2006 to 2008, he held a post-doctoral position with the Signal Processing and Information Team, SATIE Laboratory, CNRS, École Normale Supérieure de Cachan, Cachan, France. From 2008 to 2011, he was an Assistant Professor with SONDRA, CentraleSupelec, Gif-sur-Yvette, France, where he was an Associate Professor from 2012 to 2013. From 2013 to 2014, he was a Visiting Associate Professor with the Electrical and Computer Engineering Department, National University of Singapore, Singapore. Since 2014, he has been a Full Professor with the L2S laboratory, CentraleSupelec. His current research interests include estimation, detection and classification for statistical signal processing, and applications in radar and image processing.

Dr. Pascal has been a member of the IEEE Signal Processing Society SAM Technical Committee since 2015. He has been the Chair of the EURASIP SAT in Theoretical and Methodological Trends in Signal Processing since 2015 and He has been serving as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING since 2015 and the *EURASIP Journal on Advances in Signal Processing* since 2015.

Jean-Philippe Ovarlez (M'06) was born in Denain, France, in 1963. He received the Engineering degree from Ecole Supérieure d'Electronique Automatique et Informatique, Paris, France, and the Diplôme d'Etudes Approfondies degree in signal processing from the University of Paris XI, Orsay, France, in 1987, the Ph.D. degree in physics from the University of Paris VI, Paris, France, in 1992, the Research Directorship Habilitation thesis in signal processing from the University of Paris-Sud, Orsay, France, and his qualification to the University Professor position.

In 1992, he joined the French Aerospace Laboratory, Electromagnetic and Radar Division, Office National d'Etudes et de Recherches Aérospatiales (ONERA), Palaiseau, France, where he is currently a Chief Scientist and a member of the Scientific Committee of the ONERA Physics Branch, Palaiseau. Since 2008, he has been with the Centrale-Supélec SONDRA Laboratory, In Charge of Signal Processing activities supervision. His current research interests include statistical signal processing for radar and SAR applications, such as time-frequency, imaging, detection, and parameters estimation.

Dr. Ovarlez was a member of the Special Area Team in Theoretical and Methodological Trends in Signal Processing, the EURASIP, and a Treasurer of the IEEE GRSS French Chapter in 2016.