



HAL
open science

Detection of Gaussian Signal Using Adaptively Whitened Data

Olivier Besson

► **To cite this version:**

Olivier Besson. Detection of Gaussian Signal Using Adaptively Whitened Data. IEEE Signal Processing Letters, 2019, 26 (3), pp.430-434. 10.1109/LSP.2019.2893761 . hal-02618454

HAL Id: hal-02618454

<https://hal.science/hal-02618454>

Submitted on 25 May 2020

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

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



Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of some Toulouse researchers and makes it freely available over the web where possible.

This is an author's version published in: <https://oatao.univ-toulouse.fr/25979>

Official URL : <https://doi.org/10.1109/LSP.2019.2893761>

To cite this version :

Besson, Olivier Detection of Gaussian Signal Using Adaptively Whitened Data. (2019) IEEE Signal Processing Letters, 26 (3). 430-434. ISSN 1070-9908

Any correspondence concerning this service should be sent to the repository administrator:

tech-oatao@listes-diff.inp-toulouse.fr

Detection of Gaussian Signal Using Adaptively Whitened Data

Olivier Besson 

Abstract—The adaptive matched filter, like many other adaptive detection schemes, uses in its test statistic the data under test whitened by the sample covariance matrix \mathbf{S} of the training samples. Actually, it is a generalized likelihood ratio test (GLRT) based on the conditional (i.e., for given \mathbf{S}) distribution of the adaptively whitened data. In this letter, we investigate detection of a Gaussian rank-one signal using the marginal (unconditional) distribution of the adaptively whitened data. A first contribution is to derive the latter and to show that it only depends on a scalar parameter, namely the signal to noise ratio. Then, a GLRT is formulated from this unconditional distribution and shown to have the constant false alarm rate property. We show that it bears close resemblance with the plain GLRT based on the whole data set (data under test and training samples). The new detector performs as well as the plain GLRT and even better with multiple cells under test and low training sample support.

Index Terms—Adaptive detection, Cholesky factorization, Gaussian rank-one signals, generalized likelihood ratio test, Wishart matrices.

I. INTRODUCTION

A RECURRENT problem in many applications, including radar, is to detect a signal of interest (target) in the presence of noise (thermal noise, clutter, jamming) whose statistics are generally unknown, and must be learned from target-free training samples which, hopefully, share the same statistical parameters as the noise in the data under test [1]–[3]. This canonical problem was rigorously formulated by Kelly [4] under a Gaussian assumption for the noise. The generalized likelihood ratio test (GLRT) was derived and its performance (probability of false alarm, probability of detection) was thoroughly analyzed. In [5], Robey *et al.* considered the same problem but proposed a two-step approach: first, use a matched filter under the assumption of perfectly known noise covariance matrix, then substitute the latter for the sample covariance matrix of the training samples. The detector is referred to as the adaptive matched filter (AMF), and Kelly’s GLRT and the AMF have become references against which most detectors published later have been systematically compared. While it appears difficult to outperform Kelly’s GLRT in this canonical framework (Gaussian noise, known target signature, noise homogeneity between training samples and data under test), in practice, the latter is

seldom met and therefore a considerable amount of work has been devoted to propose detectors that perform well in more adverse conditions. These include methods able to mitigate the effects of low training sample support see e.g., [6]–[8], or possible mismatches either in the steering vector [9]–[12] or between the covariance matrix of the training set and that of the data under test [13]–[18].

In a recent paper [19], we considered a similar problem as in [4], except that a stochastic model for the target amplitude was used. Indeed, while the latter is assumed to be deterministic in [4], we modeled it as a random vector with Gaussian distribution, which corresponds to a Swerling I-II target model [20], [21]. The problem addressed was to decide between the two following hypotheses

$$\begin{aligned} H_0 : \mathbf{X} &\stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \mathbf{R}, \mathbf{I}); \mathbf{Z} \stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \mathbf{R}, \mathbf{I}) \\ H_1 : \mathbf{X} &\stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \mathbf{R} + P\mathbf{v}\mathbf{v}^H, \mathbf{I}); \mathbf{Z} \stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \mathbf{R}, \mathbf{I}) \end{aligned} \quad (1)$$

where $\mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{\Omega})$ stands for the complex matrix-variate Gaussian distribution [22], [23]. In (1) \mathbf{v} stands for the (known) target signature, P denotes its unknown power and \mathbf{R} is the unknown noise covariance matrix. The GLRT as well as the AMF were derived for this problem. They were shown to depend, as is the case in the large majority of detection schemes, on the adaptively whitened data $\mathbf{S}^{-1/2}\mathbf{X}$ and target signature $\mathbf{S}^{-1/2}\mathbf{v}$ where $\mathbf{S} = \mathbf{Z}\mathbf{Z}^H$ and $\mathbf{S}^{-1/2}$ is the inverse of a square-root of \mathbf{S} . In fact, the AMF proceeds first by deriving the GLRT for the hypotheses testing problem (1) with \mathbf{R} known, which is equivalent to considering

$$\begin{aligned} H_0 : \mathbf{R}^{-1/2}\mathbf{X}|\mathbf{R} &\stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \mathbf{I}, \mathbf{I}) \\ H_1 : \mathbf{R}^{-1/2}\mathbf{X}|\mathbf{R} &\stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \mathbf{I} + P\mathbf{R}^{-1/2}\mathbf{v}\mathbf{v}^H\mathbf{R}^{-1/2}, \mathbf{I}) \end{aligned} \quad (2)$$

Then, \mathbf{R} is substituted for $\hat{\mathbf{R}} = T_s^{-1}\mathbf{S}$ where T_s stands for the number of training samples. This means that the AMF is the GLRT for the following problem

$$\begin{aligned} H_0 : \hat{\mathbf{R}}^{-1/2}\mathbf{X}|\mathbf{S} &\stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \mathbf{I}, \mathbf{I}) \\ H_1 : \hat{\mathbf{R}}^{-1/2}\mathbf{X}|\mathbf{S} &\stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \mathbf{I} + P\hat{\mathbf{R}}^{-1/2}\mathbf{v}\mathbf{v}^H\hat{\mathbf{R}}^{-1/2}, \mathbf{I}) \end{aligned} \quad (3)$$

This interpretation raises two observations. First, the hypotheses testing problem in (3) does not reflect the true statistical reality since $\mathbb{E}\{\hat{\mathbf{R}}^{-1/2}\mathbf{X}\mathbf{X}^H\hat{\mathbf{R}}^{-1/2}|\mathbf{S}\} = \hat{\mathbf{R}}^{-1/2}\mathbf{R}\hat{\mathbf{R}}^{-1/2} + P\hat{\mathbf{R}}^{-1/2}\mathbf{v}\mathbf{v}^H\hat{\mathbf{R}}^{-1/2}$ and $\hat{\mathbf{R}}^{-1/2}\mathbf{R}\hat{\mathbf{R}}^{-1/2} \neq \mathbf{I}_M$. Second, the

The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ashish Pandharipande.

The author is with The Institut Supérieur de l’Aéronautique et de l’Espace, University of Toulouse, Toulouse 31055, France (e-mail: olivier.besson@isae-supaero.fr).

Digital Object Identifier 10.1109/LSP.2019.2893761

AMF implicitly uses the *conditional* distribution of the adaptively whitened data. Therefore, one may wonder why the *marginal* distribution of $\mathbf{S}^{-1/2}\mathbf{X}$ is not used to formulate the hypotheses testing problem. This is what we explore in this letter. More precisely, we derive the GLRT based on the marginal (unconditional) distribution of $\mathbf{S}^{-1/2}\mathbf{X}$. Towards this end, we first derive the latter, which is shown to depend only on $P\mathbf{v}^H\mathbf{R}^{-1}\mathbf{v}$. Then the maximum likelihood estimator of this scalar parameter is derived to obtain the GLRT. The approach proposed here is thus different from the AMF one, but also from the plain GLRT which uses the joint distribution of (\mathbf{X}, \mathbf{Z}) . However, as indicated before, most detection schemes rely on $\mathbf{S}^{-1/2}\mathbf{X}$ and, therefore, it seems meaningful to get its distribution and derive the corresponding GLRT.

II. GENERALIZED LIKELIHOOD RATIO FROM ADAPTIVELY WHITENED DATA

As stated above, we start with the detection problem in (1) where $\mathbf{X} \in \mathbb{C}^{M \times T_p}$ is the data under test, which may or may not contain the target, and $\mathbf{Z} \in \mathbb{C}^{M \times T_s}$ stands for the training samples with $T_s \geq M$, M being the size of the observation space. Our aim is to build a test that depends only on the adaptively whitened data. With no loss of generality, we consider that \mathbf{v} has unit-norm and we let $\mathbf{Q} = [\mathbf{V}_\perp \ \mathbf{v}]$ be a unitary matrix with $\mathbf{V}_\perp^H \mathbf{v} = \mathbf{0}$. Let us consider the transformed data $\tilde{\mathbf{X}} = \mathbf{Q}^H \mathbf{X}$, $\tilde{\mathbf{Z}} = \mathbf{Q}^H \mathbf{Z}$ and transformed covariance matrix $\tilde{\mathbf{R}} = \mathbf{Q}^H \mathbf{R} \mathbf{Q}$ and target signature $\tilde{\mathbf{v}} = \mathbf{Q}^H \mathbf{v} = \mathbf{e}_M = [0 \ \cdots \ 0 \ 1]^T$. Then, (1) is equivalent to

$$\begin{aligned} H_0 : \tilde{\mathbf{X}} &\stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \tilde{\mathbf{R}}, \mathbf{I}); \tilde{\mathbf{Z}} \stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \tilde{\mathbf{R}}, \mathbf{I}) \\ H_1 : \tilde{\mathbf{X}} &\stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \tilde{\mathbf{R}} + P\tilde{\mathbf{v}}\tilde{\mathbf{v}}^H, \mathbf{I}); \tilde{\mathbf{Z}} \stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \tilde{\mathbf{R}}, \mathbf{I}) \end{aligned} \quad (4)$$

Since the square-root of a matrix is not unique, and the distribution of the quasi whitened data depends on this square root, we choose to use the Cholesky factor [24], as it is a widely used method due to its computational simplicity. Let \mathbf{L} be the Cholesky factor of $\tilde{\mathbf{S}} = \tilde{\mathbf{Z}}\tilde{\mathbf{Z}}^H$, which we denote by $\mathbf{L} = \text{chol}(\tilde{\mathbf{S}})$, and let $\mathbf{Y} = \mathbf{L}^{-1}\tilde{\mathbf{X}}$. Our aim is to derive the distribution of \mathbf{Y} under both hypotheses, and subsequently to obtain the GLRT based on \mathbf{Y} .

A. Distribution of Adaptively Whitened Data \mathbf{Y}

We begin with obtaining a stochastic representation of $\mathbf{Y} = \mathbf{L}^{-1}\tilde{\mathbf{X}}$ as a function of known distributions. Let $\mathbf{G} = \text{chol}(\tilde{\mathbf{R}})$ and $\mathbf{F} = \text{chol}(\tilde{\mathbf{R}} + P\mathbf{e}_M\mathbf{e}_M^H)$ denote the Cholesky factors of $\tilde{\mathbf{R}}$ and $\tilde{\mathbf{R}} + P\mathbf{e}_M\mathbf{e}_M^H$. Let us partition \mathbf{G} as $\mathbf{G} = \begin{pmatrix} \mathbf{G}_{11} & \mathbf{0} \\ \mathbf{G}_{21} & G_{22} \end{pmatrix}$. Then, since $\mathbf{F}\mathbf{F}^H = \mathbf{G}\mathbf{G}^H + P\mathbf{e}_M\mathbf{e}_M^H$, we have that

$$\mathbf{F} = \begin{pmatrix} \mathbf{G}_{11} & \mathbf{0} \\ \mathbf{G}_{21} & (G_{22}^2 + P)^{1/2} \end{pmatrix} \quad (5)$$

Additionally, from the fact that $\mathbf{G}\mathbf{G}^H = \mathbf{Q}^H\mathbf{R}\mathbf{Q}$, one can easily obtain that $G_{22}^2 = (\mathbf{v}^H\mathbf{R}^{-1}\mathbf{v})^{-1}$. Finally, from (5), one can

deduce that

$$\begin{aligned} \mathbf{G}^{-1}\mathbf{F} &= \begin{pmatrix} \mathbf{G}_{11}^{-1} & \mathbf{0} \\ -G_{22}^{-1}\mathbf{G}_{21}\mathbf{G}_{11}^{-1} & G_{22}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{G}_{11} & \mathbf{0} \\ \mathbf{G}_{21} & (G_{22}^2 + P)^{1/2} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{I}_{M-1} & \mathbf{0} \\ \mathbf{0} & G_{22}^{-1}(G_{22}^2 + P)^{1/2} \end{pmatrix} \end{aligned} \quad (6)$$

Since $\tilde{\mathbf{Z}} \stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \tilde{\mathbf{R}} = \mathbf{G}\mathbf{G}^H, \mathbf{I})$, we can then write that

$$\tilde{\mathbf{S}} = \tilde{\mathbf{Z}}\tilde{\mathbf{Z}}^H \stackrel{d}{=} \mathbf{G}\mathbf{W}\mathbf{G}^H = \mathbf{G}\mathbf{T}\mathbf{T}^H\mathbf{G}^H \quad (7)$$

where $\mathbf{W} \stackrel{d}{=} \mathbb{C}\mathcal{W}_M(T_s, \mathbf{I})$ and $\mathbf{T} = \text{chol}(\mathbf{W})$. Note that all elements of \mathbf{T} are independent and that $\mathbf{T}_{mm}^2 \stackrel{d}{=} \mathbb{C}\chi_{T_s - m + 1}^2$ and $\mathbf{T}_{ij} \stackrel{d}{=} \mathcal{CN}(0, 1)$ for $i \neq j$ [25]. Now, $\mathbf{G}\mathbf{T}$ is a lower triangular matrix with real positive diagonal elements so that $\mathbf{L} \stackrel{d}{=} \mathbf{G}\mathbf{T}$. Since $\tilde{\mathbf{X}} \stackrel{d}{=} \mathbf{F}\mathbf{N}$ where $\mathbf{N} \stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \mathbf{I}, \mathbf{I})$, it follows that

$$\begin{aligned} \mathbf{Y} &= \mathbf{L}^{-1}\tilde{\mathbf{X}} \\ &\stackrel{d}{=} \mathbf{T}^{-1}\mathbf{G}^{-1}\mathbf{F}\mathbf{N} \\ &\stackrel{d}{=} \mathbf{T}^{-1} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \lambda^{1/2} \end{pmatrix} \mathbf{N} \\ &= \begin{pmatrix} \mathbf{T}_{11}^{-1} & \mathbf{0} \\ -T_{22}^{-1}\mathbf{T}_{21}\mathbf{T}_{11}^{-1} & T_{22}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \lambda^{1/2} \end{pmatrix} \begin{pmatrix} \mathbf{N}_1 \\ \mathbf{N}_2 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{T}_{11}^{-1}\mathbf{N}_1 \\ -T_{22}^{-1}\mathbf{T}_{21}\mathbf{T}_{11}^{-1}\mathbf{N}_1 + \lambda^{1/2}T_{22}^{-1}\mathbf{N}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix} \end{aligned} \quad (8)$$

where $\lambda = 1 + P\mathbf{v}^H\mathbf{R}^{-1}\mathbf{v}$. Equation (8) provides a stochastic representation of the adaptively whitened data \mathbf{Y} . First, note that the distribution of \mathbf{Y} depends on a single scalar parameter λ which is related to the signal to noise ratio at the output of the optimal filter $\mathbf{R}^{-1}\mathbf{v}$. Next, one can observe that $\mathbf{Y}_1 \stackrel{d}{=} \mathbf{T}_{11}^{-1}\mathbf{N}_1$ has the same parameter-free distribution under H_0 and H_1 . Let us now consider the conditional distribution of $\mathbf{Y}_2|\mathbf{Y}_1$. We have that

$$\mathbf{Y}_2|\mathbf{Y}_1, \mathbf{T}_{21}, T_{22} \stackrel{d}{=} \mathcal{CN}(-T_{22}^{-1}\mathbf{T}_{21}\mathbf{Y}_1, \lambda T_{22}^{-2}, \mathbf{I}) \quad (9)$$

and $T_{22}^2 \stackrel{d}{=} \mathbb{C}\chi_{T_s - M + 1}^2$, $\mathbf{T}_{21} \stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Therefore,

$$\begin{aligned} p(\mathbf{Y}_2|\mathbf{Y}_1, \mathbf{T}_{21}, T_{22}) &= \pi^{-T_p} (\lambda T_{22}^{-2})^{-T_p} \\ &\quad \times \text{etr} \left\{ -(\lambda T_{22}^{-2})^{-1} \|\mathbf{Y}_2 + T_{22}^{-1}\mathbf{T}_{21}\mathbf{Y}_1\|^2 \right\} \end{aligned} \quad (10)$$

where $\text{etr}\{\cdot\}$ stands for the exponential of the trace of the matrix between braces. We first marginalize with respect to $\mathbf{T}_{21} \stackrel{d}{=} \mathcal{CN}(\mathbf{0}, \mathbf{I})$. To this end, note that

$$\begin{aligned} &(\lambda T_{22}^{-2})^{-1} \|\mathbf{Y}_2 + T_{22}^{-1}\mathbf{T}_{21}\mathbf{Y}_1\|^2 + \|\mathbf{T}_{21}\|^2 \\ &= T_{22}^2 [(\mathbf{I} + \lambda^{-1}\mathbf{Y}\mathbf{Y}^H)_{2,1} - 1] \\ &\quad + (\mathbf{T}_{21} + \mathbf{M})(\mathbf{I} + \lambda^{-1}\mathbf{Y}_1\mathbf{Y}_1^H)(\mathbf{T}_{21} + \mathbf{M})^H \end{aligned} \quad (11)$$

where

$$\mathbf{M} = \lambda^{-1}T_{22}\mathbf{Y}_2\mathbf{Y}_1^H (\mathbf{I} + \lambda^{-1}\mathbf{Y}_1\mathbf{Y}_1^H)^{-1} \quad (12)$$

and

$$\begin{aligned}
& (\mathbf{I} + \lambda^{-1} \mathbf{Y} \mathbf{Y}^H)_{2,1} \\
&= 1 + \lambda^{-1} \mathbf{Y}_2 \mathbf{Y}_2^H - \lambda^{-2} \mathbf{Y}_2 \mathbf{Y}_1^H (\mathbf{I} + \lambda^{-1} \mathbf{Y}_1 \mathbf{Y}_1^H)^{-1} \mathbf{Y}_1 \mathbf{Y}_2^H \\
&= 1 + \lambda^{-1} \mathbf{Y}_2 (\mathbf{I} + \lambda^{-1} \mathbf{Y}_1^H \mathbf{Y}_1)^{-1} \mathbf{Y}_2^H \quad (13)
\end{aligned}$$

Using the fact that

$$\begin{aligned}
& \int \exp \left\{ -(\mathbf{T}_{21} + \mathbf{M}) (\mathbf{I} + \lambda^{-1} \mathbf{Y}_1 \mathbf{Y}_1^H) (\mathbf{T}_{21} + \mathbf{M})^H \right\} d\mathbf{T}_{21} \\
&= \pi^{M-1} |\mathbf{I} + \lambda^{-1} \mathbf{Y}_1 \mathbf{Y}_1^H|^{-1} \quad (14)
\end{aligned}$$

we obtain

$$\begin{aligned}
p(\mathbf{Y}_2 | \mathbf{Y}_1, T_{22}) &= \pi^{-T_p} (\lambda T_{22}^{-2})^{-T_p} |\mathbf{I} + \lambda^{-1} \mathbf{Y}_1 \mathbf{Y}_1^H|^{-1} \\
&\quad \times \exp \left\{ -T_{22}^2 [(\mathbf{I} + \lambda^{-1} \mathbf{Y} \mathbf{Y}^H)_{2,1} - 1] \right\} \quad (15)
\end{aligned}$$

Next, we need further marginalizing with respect to $T_{22} \stackrel{d}{=} \mathbb{C}\chi_{T_s - M + 1}^2$. It is straightforward to verify that

$$\begin{aligned}
& \int_0^\infty x^{2T_p} e^{-(a-1)x^2} p \sqrt{\mathbb{C}\chi_{T_s - M + 1}^2}(x) dx \\
&= \frac{\Gamma(T_p + L)}{\Gamma(L)} a^{-(T_p + L)} \quad (16)
\end{aligned}$$

where $L = T_s - M + 1$, so that

$$\begin{aligned}
p(\mathbf{Y}_2 | \mathbf{Y}_1) &= \frac{\Gamma(T_p + L)}{\Gamma(L)} \pi^{-T_p} \lambda^{-T_p} \\
&\quad |\mathbf{I} + \lambda^{-1} \mathbf{Y}_1 \mathbf{Y}_1^H|^{-1} (\mathbf{I} + \lambda^{-1} \mathbf{Y} \mathbf{Y}^H)_{2,1}^{-(T_p + L)} \quad (17)
\end{aligned}$$

Finally, the distribution of the adaptively whitened data \mathbf{Y} can be written as

$$\begin{aligned}
p(\mathbf{Y}) &= p(\mathbf{Y}_2 | \mathbf{Y}_1) p(\mathbf{Y}_1) \\
&= \frac{\Gamma(T_p + L)}{\Gamma(L)} \pi^{-T_p} \lambda^{-T_p} p(\mathbf{Y}_1) \\
&\quad |\mathbf{I} + \lambda^{-1} \mathbf{Y}_1 \mathbf{Y}_1^H|^{-1} (\mathbf{I} + \lambda^{-1} \mathbf{Y} \mathbf{Y}^H)_{2,1}^{-(T_p + L)} \quad (18)
\end{aligned}$$

Using the fact that $|\mathbf{I} + \lambda^{-1} \mathbf{Y} \mathbf{Y}^H| = |\mathbf{I} + \lambda^{-1} \mathbf{Y}_1 \mathbf{Y}_1^H| |\mathbf{I} + \lambda^{-1} \mathbf{Y} \mathbf{Y}^H|_{2,1}$ and that $|\mathbf{I} + \lambda^{-1} \mathbf{Y} \mathbf{Y}^H| = |\mathbf{I} + \lambda^{-1} \mathbf{Y}^H \mathbf{Y}|$, (18) can be equivalently written as

$$\begin{aligned}
p(\mathbf{Y}) &= \frac{\Gamma(T_p + L)}{\Gamma(L)} \pi^{-T_p} \lambda^{-T_p} p(\mathbf{Y}_1) |\mathbf{I} + \lambda^{-1} \mathbf{Y}_1^H \mathbf{Y}_1|^{T_p + L - 1} \\
&\quad |\mathbf{I} + \lambda^{-1} \mathbf{Y}^H \mathbf{Y}|^{-(T_p + L)} \quad (19)
\end{aligned}$$

Expressions (18) or (19) are new and provide the distribution of the data after adaptive whitening. It should be observed that this distribution depends only on λ , a scalar parameter, which is related to signal to noise ratio. The distribution under the hypothesis of noise only is obtained by setting $\lambda = 1$.

B. Generalized Likelihood Ratio Based on \mathbf{Y}

Thanks to the previous result, we are now in a position to set a composite hypothesis testing problem based on the unconditional distribution of \mathbf{Y} . More precisely, we consider the two

hypotheses:

$$\begin{aligned}
H_0 : p_0(\mathbf{Y}) &= C p(\mathbf{Y}_1) |\mathbf{I} + \mathbf{Y}_1^H \mathbf{Y}_1|^{T_p + L - 1} |\mathbf{I} + \mathbf{Y}^H \mathbf{Y}|^{-(T_p + L)} \\
H_1 : p_1(\mathbf{Y}) &= C p(\mathbf{Y}_1) \eta^{T_p} |\mathbf{I} + \eta \mathbf{Y}_1^H \mathbf{Y}_1|^{T_p + L - 1} \\
&\quad \times |\mathbf{I} + \eta \mathbf{Y}^H \mathbf{Y}|^{-(T_p + L)} \quad (20)
\end{aligned}$$

where $\eta = \lambda^{-1} = (1 + P \mathbf{v}^H \mathbf{R}^{-1} \mathbf{v})^{-1}$. The generalized likelihood ratio corresponding to (20) is thus given by

$$\begin{aligned}
\text{GLR}(\mathbf{Y}) &= \\
&= \frac{\max_{\eta \in [0,1]} \eta^{T_p} |\mathbf{I} + \eta \mathbf{Y}_1^H \mathbf{Y}_1|^{T_p + L - 1} |\mathbf{I} + \eta \mathbf{Y}^H \mathbf{Y}|^{-(T_p + L)}}{|\mathbf{I} + \mathbf{Y}_1^H \mathbf{Y}_1|^{T_p + L - 1} |\mathbf{I} + \mathbf{Y}^H \mathbf{Y}|^{-(T_p + L)}} \quad (21)
\end{aligned}$$

We would like to offer a few remarks regarding this GLR. First, since the distribution of \mathbf{Y} under H_0 is parameter free, the detector possesses the constant false alarm rate (CFAR) property, i.e., the threshold can be set irrespective of \mathbf{R} . The probability of detection will depend only on signal to noise ratio. Second, it is instructive to relate the present GLR to the plain GLR of [19]. Towards this end, we need to relate \mathbf{Y} to (\mathbf{X}, \mathbf{Z}) . Some straightforward calculations enable to show that $\mathbf{Y}^H \mathbf{Y} = \mathbf{X}^H \mathbf{S}^{-1} \mathbf{X}$ and $\mathbf{Y}_1^H \mathbf{Y}_1 = \mathbf{X}^H \mathbf{S}^{-H/2} \mathbf{P}_{\mathbf{S}^{-1/2} \mathbf{v}}^\perp \mathbf{S}^{-1/2} \mathbf{X}$, so that the plain GLR is

$$\begin{aligned}
\text{GLR}(\mathbf{X}, \mathbf{Z}) &= \\
&= \frac{\max_{\eta \in [0,1]} \eta^{T_p} |\mathbf{I} + \eta \mathbf{Y}_1^H \mathbf{Y}_1|^{T_p + T_s} |\mathbf{I} + \eta \mathbf{Y}^H \mathbf{Y}|^{-(T_p + T_s)}}{|\mathbf{I} + \mathbf{Y}_1^H \mathbf{Y}_1|^{T_p + T_s} |\mathbf{I} + \mathbf{Y}^H \mathbf{Y}|^{-(T_p + T_s)}} \quad (22)
\end{aligned}$$

The GLR (21) using the adaptively whitened data \mathbf{Y} bears a striking resemblance with the GLR (22) using the initial data (\mathbf{X}, \mathbf{Z}) : they use the same statistics, the difference lies in the exponent of the determinants.

Another comment concerns the implementation of (21). Observe that one only needs to solve a 1-D optimization problem over the interval $[0,1]$, similarly to the plain GLR. Moreover, the matrices involved are $T_p \times T_p$ and reduce to scalars for $T_p = 1$. Therefore, obtaining the GLRT is rather simple from the computational point of view.

Let us now consider the usual case of a single vector under test, and show that a *closed-form expression* of the GLR can be obtained. The density under H_1 is, up to the scaling factor $C p(\mathbf{y}_1)$,

$$g(\eta) = \frac{\eta (1 + a\eta)^L}{(1 + b\eta)^{L+1}} \quad (23)$$

where $a = \mathbf{y}_1^H \mathbf{y}_1 = \mathbf{x}^H \mathbf{S}^{-1} \mathbf{x} - (\mathbf{v}^H \mathbf{S}^{-1} \mathbf{v})^{-1} |\mathbf{x}^H \mathbf{S}^{-1} \mathbf{v}|^2$ and $b = \mathbf{x}^H \mathbf{S}^{-1} \mathbf{x}$. It is straightforward to show that

$$g'(\eta) = g(\eta) \frac{1 - \eta_0^{-1} \eta}{\eta (1 + a\eta) (1 + b\eta)} \quad (24)$$

where $\eta_0 = [bL - a(L+1)]^{-1}$. Studying the sign of $g'(\eta)$, it comes that whenever $\eta_0 < 0$ or $\eta_0 > 1$, the solution is $\eta_\star = 1$ and $\text{GLR}(\mathbf{y}) = 1$. On the other hand, for $0 \leq \eta_0 \leq 1$, $\eta_\star = \eta_0$

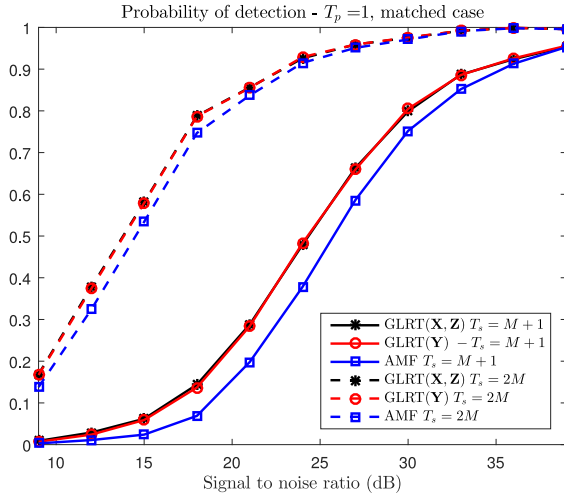


Fig. 1. Probability of detection versus SNR for $T_p = 1$.

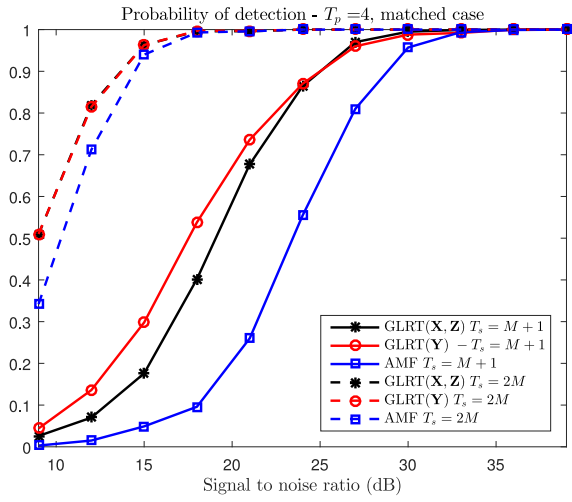


Fig. 2. Probability of detection versus SNR for $T_p = 4$.

and

$$\begin{aligned} \text{GLR}(\mathbf{y}) &= \frac{L^L (1+b)^{L+1}}{(L+1)^{L+1} (1+a)^L (b-a)} \\ &= \frac{L^L (1 + \mathbf{x}^H \mathbf{S}^{-1} \mathbf{x})^{L+1}}{(L+1)^{L+1} \left(1 + \mathbf{x}^H \mathbf{S}^{-1} \mathbf{x} - \frac{|\mathbf{x}^H \mathbf{S}^{-1} \mathbf{v}|^2}{\mathbf{v}^H \mathbf{S}^{-1} \mathbf{v}} \right)^L \frac{|\mathbf{x}^H \mathbf{S}^{-1} \mathbf{v}|^2}{\mathbf{v}^H \mathbf{S}^{-1} \mathbf{v}}} \end{aligned} \quad (25)$$

III. NUMERICAL SIMULATIONS

We consider the same radar scenario as in [19], i.e. a coherent processing interval with $M = 8$ pulses. The target normalized Doppler frequency is $f_d = 0.09$ and its signature $\mathbf{v} = M^{-1/2} [1 e^{i2\pi f_d} \dots e^{i2\pi(M-1)f_d}]^T$. The disturbance comprises thermal noise, with covariance matrix $\sigma^2 \mathbf{I}$, and clutter with covariance matrix $[\mathbf{R}_c]_{m,n} = P_c \exp \left\{ -2\pi^2 \sigma_f^2 (m-n)^2 \right\}$ where $\sigma_f^2 = 0.01$. The clutter to white noise ratio (CNR) is defined as $\text{CNR} = (M\sigma^2)^{-1} \text{Tr}\{\mathbf{R}_c\}$ and is set to $\text{CNR} = 30$ dB. The probability of false alarm is set to $P_{fa} = 10^{-3}$.

In Figures 1-2 we display the probability of detection of the plain GLRT, the AMF and the proposed GLRT, for $T_p = 1$ and

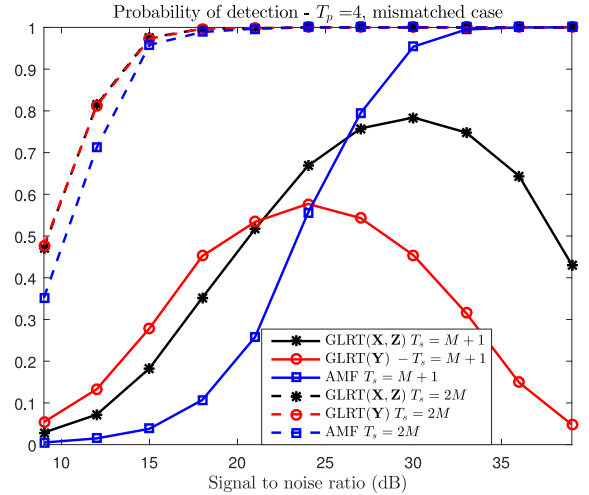


Fig. 3. Probability of detection versus SNR for $T_p = 4$ in the mismatched case.

$T_p = 4$, respectively. First, it can be observed that GLRT(Y) incurs no loss compared to GLRT(X, Z), and is even better in the case of multiple cells under test, especially with $T_s = M + 1$. This means that considering Y in small training sample support does not result in any loss, on the contrary. The second observation is that GLRT(Y) outperforms the AMF, which shows that considering the unconditional distribution of Y enables one to achieve a significant improvement compared to using the conditional distribution.

Finally, we examine in Figure 3 the performance in the presence of a slight mismatch regarding the Doppler frequency: while the detectors assume $f_d = 0.09$, the actual Doppler frequency is $0.09 + 0.1/M$. From observation of this figure, one can observe that the new detector offers an interesting tradeoff between robustness and selectivity. For $T_s = 2M$, when SNR increases the new detector assesses that the actual signature is different from \mathbf{v} and its probability of detection decreases faster than that of the plain GLRT.

To conclude these simulations, the new detector based on the marginal distribution of the adaptively whitened data improves over the plain GLRT in matched conditions and $T_p > 1$, and is a bit more selective in mismatched conditions.

IV. CONCLUSION

In this letter, we started from the fact that the adaptive matched filter is a generalized likelihood ratio test for a composite hypothesis testing problem based on the conditional distribution of the adaptively whitened data $(\mathbf{Z}\mathbf{Z}^H)^{-1/2} \mathbf{X}$, and from the observation that many adaptive detection schemes use this statistic. Therefore, we investigated another path, also different from the conventional plain GLRT based on the whole data set (\mathbf{X}, \mathbf{Z}) , namely GLRT from the unconditional distribution of $(\mathbf{Z}\mathbf{Z}^H)^{-1/2} \mathbf{X}$. The latter was derived and was shown to be parameter free under the null hypothesis and to depend only on the signal to noise ratio under the alternative hypothesis. From this distribution, the GLRT was formulated, and related to the plain GLRT. Numerical simulations showed that the new detector performs at least as well as the plain GLRT and better with multiple cells under test and low training sample support.

REFERENCES

- [1] M. A. Richards, J. A. Scheer, and W. A. Holm, Eds., *Principles of Modern Radar: Basic Principles*, vol. 1. Raleigh, NC, USA: Scitech Publishing, 2010.
- [2] W. L. Melvin and J. A. Scheer, Eds., *Principles of Modern Radar: Advanced Principles*, vol. 2. Raleigh, NC, USA: Scitech Publishing, 2012.
- [3] J. A. Scheer and W. L. Melvin, Eds., *Principles of Modern Radar: Radar Applications*, vol. 3. Raleigh, NC, USA: Scitech Publishing, 2013.
- [4] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, no. 1, pp. 115–127, Mar. 1986.
- [5] 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.
- [6] I. P. Kirsteins and D. W. Tufts, "Adaptive detection using low rank approximation to a data matrix," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 30, no. 1, pp. 55–67, Jan. 1994.
- [7] T. F. Ayoub and A. M. Haimovich, "Modified GLRT detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 3, pp. 810–818, Jul. 2000.
- [8] M. Rangaswamy, F. C. Lin, and K. Gerlach, "Robust adaptive signal processing methods for heterogeneous radar clutter scenarios," *Signal Process.*, vol. 84, no. 9, pp. 1653–1665, Sep. 2004.
- [9] S. Z. Kalson, "An adaptive array detector with mismatched signal rejection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 195–207, Jan. 1992.
- [10] O. Besson, "Detection of a signal in linear subspace with bounded mismatch," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 3, pp. 1131–1139, Jul. 2006.
- [11] F. Bandiera, A. De Maio, and G. Ricci, "Adaptive CFAR radar detection with conic rejection," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2533–2541, Jun. 2007.
- [12] F. Bandiera, D. Orlando, and G. Ricci, *Advanced Radar Detection Schemes Under Mismatched Signal Models* (Synthesis lectures on Signal Processing). San Rafael, CA, USA: Morgan & Claypool, 2009.
- [13] L. L. Scharf and T. McWhorter, "Adaptive matched subspace detectors and adaptive coherence estimators," in *Proc. 30th Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, USA, Nov. 1996, pp. 1114–1117.
- [14] S. Kraut, L. L. Scharf, and R. W. Butler, "The adaptive coherence estimator: A uniformly most powerful invariant adaptive detection statistic," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 427–438, Feb. 2005.
- [15] N. B. Pulsone and M. A. Zatman, "A computationally efficient two-step implementation of the GLRT," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 609–616, Mar. 2000.
- [16] 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.
- [17] E. Conte, A. De Maio, and G. Ricci, "GLRT-based adaptive detection algorithm for range-spread targets," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1336–1348, Jul. 2001.
- [18] F. Bandiera, O. Besson, and G. Ricci, "An ABORT-like detector with improved mismatched signals rejection capabilities," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 14–25, Jan. 2008.
- [19] O. Besson, A. Coluccia, E. Chaumette, G. Ricci, and F. Vincent, "Generalized likelihood ratio test for detection of Gaussian rank-one signals in Gaussian noise with unknown statistics," *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 1082–1092, Feb. 2017.
- [20] P. Swerling, "Probability of detection for fluctuating targets," *IRE Trans. Inf. Theory*, vol. 6, no. 2, pp. 269–308, Apr. 1960.
- [21] P. Swerling, "Radar probability of detection for some additional fluctuating target cases," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, no. 2, pp. 698–709, Apr. 1997.
- [22] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*. London, U.K.: Chapman & Hall, 2000.
- [23] T. Kollo and D. von Rosen, *Advanced Multivariate Statistics With Matrices*. Berlin, Germany: Springer, 2005.
- [24] N. J. Higham, "Cholesky factorization," *Wiley Interdisciplinary Rev., Comput. Statist.*, vol. 1, no. 2, pp. 251–254, 2009.
- [25] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (An introduction)," *Ann. Math. Statist.*, vol. 34, no. 1, pp. 152–177, Mar. 1963.