Adaptive detection of a Gaussian signal in Gaussian noise

Olivier Besson, Eric Chaumette, François Vincent

Adaptive detection of a Gaussian signal in Gaussian noise

Olivier Besson, Eric Chaumette and François Vincent
University of Toulouse, ISAE-Supaéro
Department of Electronics, Optronics and Signal
10 Avenue Edouard Belin, 31055 Toulouse France

***Abstract*—Adaptive detection of a Swerling I-II type target in Gaussian noise with unknown covariance matrix is addressed in this paper. The most celebrated approach to this problem is Kelly's generalized likelihood ratio test (GLRT), derived under the hypothesis of deterministic target amplitudes. While this conditional model is ubiquitous, we investigate here the equivalent GLR approach for an unconditional model where the target amplitudes are treated as Gaussian random variables at the design of the detector. The GLRT is derived which is shown to be the product of Kelly's GLRT and a corrective, data dependent, term. Numerical simulations are provided to compare the two approaches.**

I. Problem formulation

Thirty years ago, in a series of technical reports and papers now became classic references [1]–[5], Kelly thoroughly investigated the problem of detecting a signal of interest (SoI) buried in Gaussian noise with unknown covariance matrix. This problem can be formulated as the following composite binary hypothesis test

$$\begin{aligned} H_0 &: \begin{cases} \boldsymbol{x}_{t_p} = \boldsymbol{n}_{t_p}; t_p = 1, \cdots, T_p \\ \boldsymbol{y}_{t_s} = \boldsymbol{n}_{t_s};\ t_s = 1, \cdots, T_s \end{cases} \\ H_1 &: \begin{cases} \boldsymbol{x}_{t_p} = \alpha_{t_p}\boldsymbol{v} + \boldsymbol{n}_{t_p}; t_p = 1, \cdots, T_p \\ \boldsymbol{y}_{t_s} = \boldsymbol{n}_{t_s};\ t_s = 1, \cdots, T_s \end{cases} \end{aligned} \tag{1}$$

where $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_{T_p} \end{bmatrix} \in \mathbb{C}^{M \times T_p}$ stands for the observation matrix where the presence of a signal of interest is sought. The latter has (unit-norm) known signature $\boldsymbol{v}$ and its complex amplitude is α_{t_p}. The data matrix $\boldsymbol{X}$ is often referred to as the primary data. $\boldsymbol{n}_{t_p}$ corresponds to the additive noise, which is assumed to be zero-mean, complex Gaussian distributed with unknown positive definite covariance matrix $\boldsymbol{R} \in \mathbb{C}^{M \times M}$, which we denote as $\boldsymbol{n}_{t_p} \sim \mathbb{C}\mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$. Additionally, it is assumed that T_s snapshots $\boldsymbol{y}_{t_s}$ are available, which contain noise only, i.e., $\boldsymbol{y}_{t_s}$ are independent, zero-mean complex Gaussian vectors drawn from $\boldsymbol{y}_{t_s} \sim \mathbb{C}\mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$.

The problem in (1) arises in many fields of engineering, and is particularly important for radar applications. In the latter case, the matrix $\boldsymbol{X}$ corresponds to the radar returns at the range cells under test (CUT), $\boldsymbol{v}$ is the target space or time or space-time signature, and $\boldsymbol{y}_{t_s}$ are radar data collected in range cells in the vicinity of the CUT [4]. The most usual case corresponds to a single CUT for which $T_p = 1$, while the case $T_p > 1$ is related to the detection of a range spread target [6] or to the detection over multiple coherent processing intervals. The reference approach to solve this problem is Kelly's generalized likelihood ratio test (GLRT) [1], [4], which was obtained under the assumption that α_{t_p} are unknown deterministic quantities. Kelly's GLRT takes the following form:

$$GLR_c^{1/T_t} = \frac{|\boldsymbol{I} + \boldsymbol{X}^H \boldsymbol{S}_y^{-1} \boldsymbol{X}|}{|\boldsymbol{I} + \boldsymbol{X}^H \boldsymbol{S}_y^{-1/2} \boldsymbol{P}^{\perp}_{\boldsymbol{S}_y^{-1/2}\boldsymbol{v}} \boldsymbol{S}_y^{-1/2} \boldsymbol{X}|} \tag{2}$$

where $T_t = T_p + T_s$, $\boldsymbol{S}_y = \boldsymbol{Y}\boldsymbol{Y}^H$ is T_s times the sample covariance matrix of the secondary data $\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{y}_1 & \cdots & \boldsymbol{y}_{t_s} \end{bmatrix}$, $\boldsymbol{P}^{\perp}_{\boldsymbol{S}_y^{-1/2}\boldsymbol{v}}$ denotes the orthogonal projector onto the subspace orthogonal to $\boldsymbol{S}_y^{-1/2}\boldsymbol{v}$, and $|.|$ stands for the determinant of a matrix. Kelly provided a detailed statistical analysis of this detector both in the case of matched or mismatched signature [2], [5]. Under the same assumption and in the case $T_p = 1$, Robey *et al.* derived the adaptive matched filter in [7]. This is indeed a two step GLRT where at the first step $\boldsymbol{R}$ is assumed to be known (and the GLR is derived from $\boldsymbol{X}$ only), and at the second step $T_s^{-1}\boldsymbol{S}_y$ is substituted for $\boldsymbol{R}$.

Surprisingly enough, considering α_{t_p} as a random variable has received little attention, and the quasi totality of recent studies followed the lead of [4] and considered α_{t_p} as deterministic parameters. Although the literature on the topic cannot be browsed exhaustively, we are not aware of references that would address detection of a Gaussian signal in colored noise with unknown covariance matrix (while the case of white noise has been examined thoroughly). In [8], detection of an arbitrary Gaussian signal is addressed but this signal is not aligned on a known signature. The advantages of a "conditional" model are that 1)one does not formulate any assumption on the amplitude statistics and 2)derivations involve a simple linear least-squares problem with respect to α_{t_p}. One drawback might be that the number of unknowns grows with T_p and, therefore, an unconditional model is worthy of investigation. Moreover, a stochastic assumption for α_{t_p} makes sense. Indeed, we assume herein that α_{t_p} are independent and drawn from a complex Gaussian distribution with zero mean and unknown variance P, i.e., $\alpha_{t_p} \sim \mathbb{C}\mathcal{N}(0, P)$, which complies with the widely accepted Swerling I-II target model [9], [10]. The problem in (1) can thus be re-formulated as

$$\begin{aligned} H_0 &: \boldsymbol{X} \sim \mathbb{C}\mathcal{N}\left(\boldsymbol{0}, \boldsymbol{R}, \boldsymbol{I}_{T_p}\right); \boldsymbol{Y} \sim \mathbb{C}\mathcal{N}\left(\boldsymbol{0}, \boldsymbol{R}, \boldsymbol{I}_{T_s}\right) \\ H_1 &: \boldsymbol{X} \sim \mathbb{C}\mathcal{N}\left(\boldsymbol{0}, \boldsymbol{R} + P\boldsymbol{v}\boldsymbol{v}^H, \boldsymbol{I}_{T_p}\right); \boldsymbol{Y} \sim \mathbb{C}\mathcal{N}\left(\boldsymbol{0}, \boldsymbol{R}, \boldsymbol{I}_{T_s}\right). \end{aligned} \tag{3}$$

The main difference with the deterministic approach is that, under H_1, the SoI is embedded in the covariance matrix of $\boldsymbol{X}$ instead of in its mean value.

The aim of this paper is to provide answers to the following questions:

1) is it possible to derive the GLRT for the problem in (3)?
2) if so, does it result in any improvement compared to (2)?

II. Generalized likelihood ratio test

In this section, we derive the GLRT for the problem described in (3) and relate it to Kelly's GLRT in the deterministic case. Since both P and $\boldsymbol{R}$ are unknown, the GLR in this case writes

$$\frac{\max\limits_{P,\boldsymbol{R}} p_1\left(\boldsymbol{X},\boldsymbol{Y}\right)}{\max\limits_{\boldsymbol{R}} p_0\left(\boldsymbol{X},\boldsymbol{Y}\right)} \tag{4}$$

where $p_\ell(\boldsymbol{X},\boldsymbol{X})$ is the probability density function of the observations under hypothesis H_ℓ.

Under H_0 the p.d.f. of the observations is given by

$$p_0\left(\boldsymbol{X},\boldsymbol{Y}\right) \propto |\boldsymbol{R}|^{-T_t}\mathrm{etr}\left\{-\boldsymbol{R}^{-1}\left(\boldsymbol{S}_y+\boldsymbol{X}\boldsymbol{X}^H\right)\right\} \tag{5}$$

where $\mathrm{etr}\left\{.\right\}$ stands the exponential of the trace of a matrix and $\propto$ means proportional to. In this case, it is well known that the maximum of $p_0\left(\boldsymbol{X},\boldsymbol{Y}\right)$ is achieved for $\boldsymbol{R}=T_t^{-1}\left(\boldsymbol{S}_y+\boldsymbol{X}\boldsymbol{X}^H\right)$ and is thus given by

$$\max_{\boldsymbol{R}} p_0\left(\boldsymbol{X},\boldsymbol{Y}\right) \propto |\boldsymbol{S}_y+\boldsymbol{X}\boldsymbol{X}^H|^{-T_t}. \tag{6}$$

Under H_1, let $\boldsymbol{V}=[\boldsymbol{v}\quad \boldsymbol{V}_\perp]$ be a unitary matrix, with $\boldsymbol{V}_\perp$ a basis for the subspace orthogonal to $\boldsymbol{v}$, i.e., $\boldsymbol{V}_\perp^H\boldsymbol{v}=\boldsymbol{0}$ and $\boldsymbol{V}_\perp^H\boldsymbol{V}_\perp=\boldsymbol{I}_{M-1}$. This transformation brings $\boldsymbol{v}$ to $\boldsymbol{V}^H\boldsymbol{v}=\boldsymbol{e}_1=[1\quad 0\quad \cdots\quad 0]^T$. Let us define the transformed data $\tilde{\boldsymbol{X}}=\boldsymbol{V}^H\boldsymbol{X}=\begin{bmatrix}\tilde{\boldsymbol{X}}_1\\ \tilde{\boldsymbol{X}}_2\end{bmatrix}$ and $\tilde{\boldsymbol{Y}}=\boldsymbol{V}^H\boldsymbol{X}=\begin{bmatrix}\tilde{\boldsymbol{Y}}_1\\ \tilde{\boldsymbol{Y}}_2\end{bmatrix}$, and transformed covariance matrix $\tilde{\boldsymbol{R}}=\boldsymbol{V}^H\boldsymbol{R}\boldsymbol{V}$. The joint p.d.f. of $\boldsymbol{X}$ and $\boldsymbol{Y}$ can be expressed as

$$\begin{aligned} p_1\left(\boldsymbol{X},\boldsymbol{Y}\right) &\propto |\tilde{\boldsymbol{R}}|^{-T_s}|\tilde{\boldsymbol{R}}+P\boldsymbol{e}_1\boldsymbol{e}_1^H|^{-T_p}\\ &\times \mathrm{etr}\left\{-\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{Y}}\tilde{\boldsymbol{Y}}^H\right\}\mathrm{etr}\left\{-(\tilde{\boldsymbol{R}}+P\boldsymbol{e}_1\boldsymbol{e}_1^H)^{-1}\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^H\right\}. \end{aligned} \tag{7}$$

Let us decompose $\tilde{\boldsymbol{R}}$ as

$$\tilde{\boldsymbol{R}}=\begin{pmatrix}\tilde{R}_{11} & \tilde{\boldsymbol{R}}_{12}\\ \tilde{\boldsymbol{R}}_{21} & \tilde{\boldsymbol{R}}_{22}\end{pmatrix} \tag{8}$$

and let $\tilde{R}_{1.2}=\tilde{R}_{11}-\tilde{\boldsymbol{R}}_{12}\tilde{\boldsymbol{R}}_{22}^{-1}\tilde{\boldsymbol{R}}_{21}$ and $\boldsymbol{\beta}=\tilde{\boldsymbol{R}}_{22}^{-1}\tilde{\boldsymbol{R}}_{21}$. Observe that $\tilde{\boldsymbol{R}}$ can be equivalently parametrized by $(\tilde{R}_{11},\,\tilde{\boldsymbol{R}}_{21},\,\tilde{\boldsymbol{R}}_{22})$ or $(\tilde{R}_{1.2},\,\boldsymbol{\beta},\,\tilde{\boldsymbol{R}}_{22})$. Using the facts that $|\tilde{\boldsymbol{R}}|=\tilde{R}_{1.2}|\tilde{\boldsymbol{R}}_{22}|$ and

$$\tilde{\boldsymbol{R}}^{-1}=\tilde{R}_{1.2}^{-1}\begin{pmatrix}1 & -\boldsymbol{\beta}^H\\ -\boldsymbol{\beta} & \boldsymbol{\beta}\boldsymbol{\beta}^H\end{pmatrix}+\begin{pmatrix}0 & \boldsymbol{0}\\ \boldsymbol{0} & \tilde{\boldsymbol{R}}_{22}^{-1}\end{pmatrix} \tag{9}$$

one can rewrite (7) as

$$\begin{aligned} p_1\left(\boldsymbol{X},\boldsymbol{Y}\right) &\propto |\tilde{\boldsymbol{R}}_{22}|^{-T_t}\tilde{R}_{1.2}^{-T_s}\left(P+\tilde{R}_{1.2}\right)^{-T_p}\\ &\times \mathrm{etr}\left\{-\tilde{\boldsymbol{R}}_{22}^{-1}\left(\tilde{\boldsymbol{Y}}_2\tilde{\boldsymbol{Y}}_2^H+\tilde{\boldsymbol{X}}_2\tilde{\boldsymbol{X}}_2^H\right)\right\}\\ &\times \exp\left\{-\begin{bmatrix}1 & -\boldsymbol{\beta}^H\end{bmatrix}\tilde{\boldsymbol{A}}\begin{bmatrix}1\\ -\boldsymbol{\beta}\end{bmatrix}\right\} \end{aligned} \tag{10}$$

where we temporarily define

$$\tilde{\boldsymbol{A}}=\tilde{R}_{1.2}^{-1}\tilde{\boldsymbol{S}}_y+\left(P+\tilde{R}_{1.2}\right)^{-1}\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^H. \tag{11}$$

Since

$$\begin{aligned} \begin{bmatrix}1 & -\boldsymbol{\beta}^H\end{bmatrix}\tilde{\boldsymbol{A}}\begin{bmatrix}1\\ -\boldsymbol{\beta}\end{bmatrix} &= \left(\boldsymbol{\beta}-\tilde{\boldsymbol{A}}_{22}^{-1}\tilde{\boldsymbol{A}}_{21}\right)^H\tilde{\boldsymbol{A}}_{22}\left(\boldsymbol{\beta}-\tilde{\boldsymbol{A}}_{22}^{-1}\tilde{\boldsymbol{A}}_{21}\right)\\ &+\tilde{A}_{11}-\tilde{\boldsymbol{A}}_{12}\tilde{\boldsymbol{A}}_{22}^{-1}\tilde{\boldsymbol{A}}_{21} \end{aligned} \tag{12}$$

it follows that

$$\begin{aligned} p_1\left(\boldsymbol{X},\boldsymbol{Y}\right) &\propto |\tilde{\boldsymbol{R}}_{22}|^{-T_t}\mathrm{etr}\left\{-\tilde{\boldsymbol{R}}_{22}^{-1}\left(\tilde{\boldsymbol{Y}}_2\tilde{\boldsymbol{Y}}_2^H+\tilde{\boldsymbol{X}}_2\tilde{\boldsymbol{X}}_2^H\right)\right\}\\ &\times \exp\left\{-\left(\boldsymbol{\beta}-\tilde{\boldsymbol{A}}_{22}^{-1}\tilde{\boldsymbol{A}}_{21}\right)^H\tilde{\boldsymbol{A}}_{22}\left(\boldsymbol{\beta}-\tilde{\boldsymbol{A}}_{22}^{-1}\tilde{\boldsymbol{A}}_{21}\right)\right\}\\ &\times \tilde{R}_{1.2}^{-T_s}\left(P+\tilde{R}_{1.2}\right)^{-T_p}\exp\left\{-\tilde{A}_{1.2}\right\}. \end{aligned} \tag{13}$$

Clearly, $p_1\left(\boldsymbol{X},\boldsymbol{Y}\right)$ is maximized for $\tilde{\boldsymbol{R}}_{22}=T_t^{-1}\left(\tilde{\boldsymbol{Y}}_2\tilde{\boldsymbol{Y}}_2^H+\tilde{\boldsymbol{X}}_2\tilde{\boldsymbol{X}}_2^H\right)$, $\boldsymbol{\beta}=\tilde{\boldsymbol{A}}_{22}^{-1}\tilde{\boldsymbol{A}}_{21}$, which results in

$$\begin{aligned} \max_{\tilde{\boldsymbol{R}}_{22},\boldsymbol{\beta}} p_1\left(\boldsymbol{X},\boldsymbol{Y}\right) &\propto |\tilde{\boldsymbol{Y}}_2\tilde{\boldsymbol{Y}}_2^H+\tilde{\boldsymbol{X}}_2\tilde{\boldsymbol{X}}_2^H|^{-T_t}\\ &\times \tilde{R}_{1.2}^{-T_s}\left(P+\tilde{R}_{1.2}\right)^{-T_p}\exp\left\{-\tilde{A}_{1.2}\right\}. \end{aligned} \tag{14}$$

Next, observe that $\tilde{A}_{1.2}^{-1}$ is the upper-left corner of $\tilde{\boldsymbol{A}}^{-1}$ and the latter is given by

$$\begin{aligned} \tilde{\boldsymbol{A}}^{-1} &= \tilde{R}_{1.2}\left[\tilde{\boldsymbol{S}}_y+(1+P\tilde{R}_{1.2}^{-1})^{-1}\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^H\right]^{-1}\\ &= \tilde{R}_{1.2}\boldsymbol{V}^H\left[\boldsymbol{S}_y+(1+P\tilde{R}_{1.2}^{-1})^{-1}\boldsymbol{X}\boldsymbol{X}^H\right]^{-1}\boldsymbol{V}. \end{aligned} \tag{15}$$

It ensues that

$$\tilde{A}_{1.2}^{-1}=\tilde{R}_{1.2}\boldsymbol{v}^H\left[\boldsymbol{S}_y+(1+P\tilde{R}_{1.2}^{-1})^{-1}\boldsymbol{X}\boldsymbol{X}^H\right]^{-1}\boldsymbol{v}. \tag{16}$$

For the sake of notational convenience, let us introduce $a=\tilde{R}_{1.2}$ and $b=P\tilde{R}_{1.2}^{-1}$. Observe that $b=P\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}$ is tantamount the signal to noise ratio at the output of the optimal filter $\boldsymbol{R}^{-1}\boldsymbol{v}$. Then, one can rewrite (14) as

$$\begin{aligned} &\max_{\tilde{\boldsymbol{R}}_{22},\boldsymbol{\beta}} p_1\left(\boldsymbol{X},\boldsymbol{Y}\right) \propto |\boldsymbol{V}_\perp^H\left(\boldsymbol{S}_y+\boldsymbol{X}\boldsymbol{X}^H\right)\boldsymbol{V}_\perp|^{-T_t}\\ &\times a^{-T_t}\left(1+b\right)^{-T_p}\\ &\times \exp\left\{-a^{-1}\left[\boldsymbol{v}^H\left(\boldsymbol{S}_y+(1+b)^{-1}\boldsymbol{X}\boldsymbol{X}^H\right)^{-1}\boldsymbol{v}\right]^{-1}\right\}. \end{aligned} \tag{17}$$

Using the readily verified facts that

$$\max_a a^{-T_t}\exp\left\{-\xi^{-1}a^{-1}\right\}=\left(\frac{e}{T_t}\right)^{-T_t}\xi^{T_t} \tag{18}$$

along with

$$\begin{aligned} |\boldsymbol{V}_\perp^H\left(\boldsymbol{S}_y+\boldsymbol{X}\boldsymbol{X}^H\right)\boldsymbol{V}_\perp| &= \left(\boldsymbol{v}^H\boldsymbol{S}_y^{-1}\boldsymbol{v}\right)\\ &\times |\boldsymbol{S}_y||\boldsymbol{I}+\boldsymbol{X}^H\boldsymbol{S}_y^{-1/2}\boldsymbol{P}^\perp_{\boldsymbol{S}_y^{-1/2}\boldsymbol{v}}\boldsymbol{S}_y^{-1/2}\boldsymbol{X}| \end{aligned} \tag{19}$$

we get that

$$\max_{\tilde{\boldsymbol{R}}_{22},\boldsymbol{\beta},a} p_1(\boldsymbol{X},\boldsymbol{Y}) \propto |\boldsymbol{S}_y|^{-T_t} \times |\boldsymbol{I}+\boldsymbol{X}^H\boldsymbol{S}_y^{-1/2}\boldsymbol{P}^{\perp}_{\boldsymbol{S}_y^{-1/2}\boldsymbol{v}}\boldsymbol{S}_y^{-1/2}\boldsymbol{X}|^{-T_t} \times (1+b)^{-T_p}\left[\frac{\boldsymbol{v}^H\left(\boldsymbol{S}_y+(1+b)^{-1}\boldsymbol{X}\boldsymbol{X}^H\right)^{-1}\boldsymbol{v}}{\boldsymbol{v}^H\boldsymbol{S}_y^{-1}\boldsymbol{v}}\right]^{T_t}. \quad (20)$$

It is possible to show that the term in the last line can be equivalently written as

$$\frac{\boldsymbol{v}^H\left(\boldsymbol{S}_y+(1+b)^{-1}\boldsymbol{X}\boldsymbol{X}^H\right)^{-1}\boldsymbol{v}}{\boldsymbol{v}^H\boldsymbol{S}_y^{-1}\boldsymbol{v}} = \frac{|\boldsymbol{I}+(1+b)^{-1}\boldsymbol{X}^H\boldsymbol{S}_y^{-1/2}\boldsymbol{P}^{\perp}_{\boldsymbol{S}_y^{-1/2}\boldsymbol{v}}\boldsymbol{S}_y^{-1/2}\boldsymbol{X}|}{|\boldsymbol{I}+(1+b)^{-1}\boldsymbol{X}^H\boldsymbol{S}_y^{-1}\boldsymbol{X}|}. \quad (21)$$

Finally, the GLR for Gaussian signals is given by

$$\begin{aligned} GLR_u^{1/T_t} &= \frac{|\boldsymbol{I}+\boldsymbol{X}^H\boldsymbol{S}_y^{-1}\boldsymbol{X}|}{|\boldsymbol{I}+\boldsymbol{X}^H\boldsymbol{S}_y^{-1/2}\boldsymbol{P}^{\perp}_{\boldsymbol{S}_y^{-1/2}\boldsymbol{v}}\boldsymbol{S}_y^{-1/2}\boldsymbol{X}|} \times \max_b \frac{|\boldsymbol{I}+(1+b)^{-1}\boldsymbol{X}^H\boldsymbol{S}_y^{-1/2}\boldsymbol{P}^{\perp}_{\boldsymbol{S}_y^{-1/2}\boldsymbol{v}}\boldsymbol{S}_y^{-1/2}\boldsymbol{X}|}{(1+b)^{T_p/T_t}|\boldsymbol{I}+(1+b)^{-1}\boldsymbol{X}^H\boldsymbol{S}_y^{-1}\boldsymbol{X}|} \\ &= \frac{\boldsymbol{v}^H\boldsymbol{S}_y^{-1}\boldsymbol{v}}{\boldsymbol{v}^H\left(\boldsymbol{S}_y+\boldsymbol{X}\boldsymbol{X}^H\right)^{-1}\boldsymbol{v}} \times \max_b \frac{\boldsymbol{v}^H\left(\boldsymbol{S}_y+(1+b)^{-1}\boldsymbol{X}\boldsymbol{X}^H\right)^{-1}\boldsymbol{v}}{(1+b)^{T_p/T_t}\left(\boldsymbol{v}^H\boldsymbol{S}_y^{-1}\boldsymbol{v}\right)}. \end{aligned} \quad (22)$$

The first term of the product is recognized as Kelly's test statistic, i.e., the GLR for deterministic amplitudes α_{t_p}. The second term (which is always lower than one) is a corrective term due to the fact that now α_{t_p} are considered as Gaussian distributed random variables.

Remark 1. Since the above GLR involves the same quantities as Kelly's GLR, it follows that is has a constant false alarm rate with respect to $\boldsymbol{R}$, i.e., its distribution under H_0 is independent of $\boldsymbol{R}$.

Remark 2. The new detector involves additional computations compared to Kelly's detector due to the need to solve the optimization problem in (22). However, the extra cost is not that large. Let us define $\eta = (1+b)^{-1} \in [0,1]$ and $\boldsymbol{S}_{xy} = \boldsymbol{S}_y + \boldsymbol{X}\boldsymbol{X}^H$. Then, if the determinant form is employed, one can make use of the fact that $|\boldsymbol{I}+\eta\boldsymbol{M}| = \prod[1+\eta\lambda_j(\boldsymbol{M})]$ where $\lambda_j(\boldsymbol{M})$ are the eigenvalues of $\boldsymbol{M}$, to efficiently compute the function to be maximized with respect to η. Likewise, if the second form of the detector is used, one can notice that

$$\begin{aligned} f(\eta) &= \boldsymbol{v}^H\left(\boldsymbol{S}_y+\eta\boldsymbol{X}\boldsymbol{X}^H\right)^{-1}\boldsymbol{v} \\ &= \boldsymbol{v}^H\left(\boldsymbol{S}_{xy}+(\eta-1)\boldsymbol{X}\boldsymbol{X}^H\right)^{-1}\boldsymbol{v} \\ &= \boldsymbol{v}^H\boldsymbol{S}_{xy}^{-1}\boldsymbol{v} - (\eta-1)\boldsymbol{v}^H\boldsymbol{S}_{xy}^{-1}\boldsymbol{X} \\ &\quad \left[\boldsymbol{I}_{T_p}+(\eta-1)\boldsymbol{X}^H\boldsymbol{S}_{xy}^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^H\boldsymbol{S}_{xy}^{-1}\boldsymbol{v} \end{aligned} \quad (23)$$

which can be used, e.g., to compute efficiently $f(\eta)$ over a grid of values of η and solve the optimization problem.

III. Numerical simulations

We now provide numerical illustrations of the performance of the new detector and compare it with Kelly's GLRT. We consider a radar scenario with $M = 16$ pulses. The SoI signature is given by $\boldsymbol{v} = \begin{bmatrix}1 & e^{i2\pi f_s} & \dots & e^{i2\pi(M-1)f_s}\end{bmatrix}^T$ with $f_s = 0.09$. The noise vectors $\boldsymbol{n}_{t_p}$ and $\boldsymbol{n}_{t_s}$ include both thermal noise and clutter components, which are assumed to be uncorrelated so that $\boldsymbol{R} = \boldsymbol{R}_c + \sigma_n^2\boldsymbol{I}$. The clutter covariance matrix is selected as $[\boldsymbol{R}_c]_{m_1,m_2} \propto \exp\left\{-2\pi^2\sigma_f^2(m_1-m_2)^2\right\}$ with $\sigma_f^2 = 0.01$. The clutter to white noise ratio (CWNR) is defined as $CWNR = \mathrm{Tr}\{\boldsymbol{R}_c\}/\mathrm{Tr}\{\sigma_n^2\boldsymbol{I}\}$ and is set to $CWNR = 20$dB in the simulations. The signal to noise ratio is defined as $SNR = P\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}$. The probability of false alarm is set to $P_{fa} = 10^{-3}$.

In Figures 1-2 we provide an excerpt of the results obtained. The main conclusions are the following. When $T_p = 1$, the two detectors provide the same probability of detection, whatever T_s. Differences can only be observed when $T_p = 4, 8$, and T_s is small, typically $T_s = M+1$, see Figures 1-2. In this case, the new detector provides improvement compared to Kelly's detector. Otherwise, even when $T_p = 4, 8$ and $T_s = 2M$, the two detectors behave the same. Note that we did not observe scenarios where the new detector would perform worst than Kelly's GLRT.

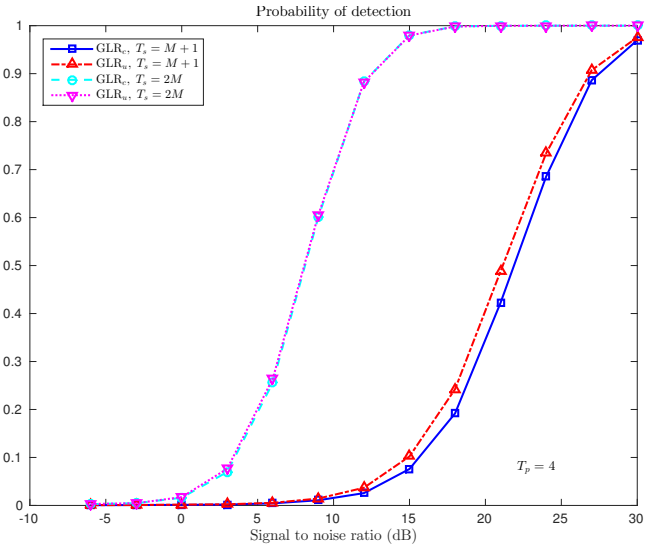

Fig. 1. Probability of detection versus SNR. $M = 16$ and $T_p = 4$.

IV. Conclusions

In this paper, we revisited the classical problem of detecting a signal of interest in colored Gaussian noise with unknown covariance matrix. The chief systematic approach is to follow Kelly's lead and use the GLRT based on deterministic signal amplitudes (conditional model). Herein, we took a different path and investigated whether it was possible to derive the GLRT assuming Gaussian signal amplitudes (unconditional model). It proved to be possible and an expression for the GLRT was derived, which bears some resemblance with Kelly's GLRT. The new detector was shown to improve over Kelly's only when the number of primary data is not small

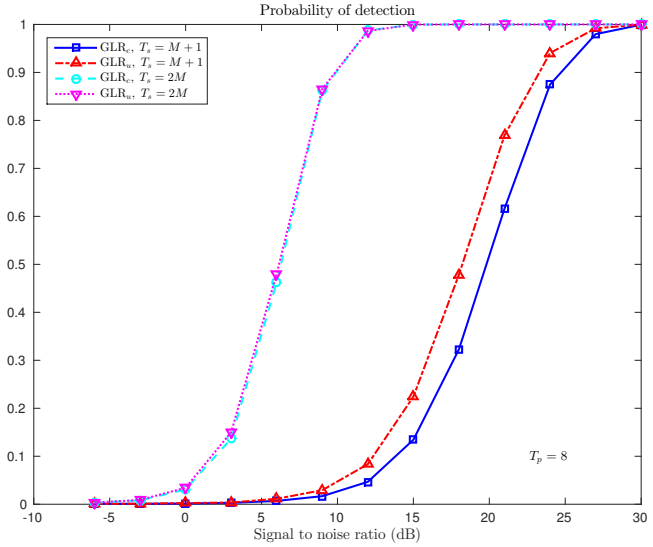

Fig. 2. Probability of detection versus SNR. $M = 16$ and $T_p = 8$.

while the number of secondary data is small. Otherwise the two detectors offer the same performance.

ACKNOWLEDGMENT

This work is partly supported by DGA/MRIS under grant no. 2015.60.0090.00.470.75.01.

APPENDIX

Similarly to [7], let us investigate a two-step approach where, at the first step, we assume that $\boldsymbol{R}$ is known. Then, one has

$$p_1(\boldsymbol{X}) \propto |\boldsymbol{R} + P\boldsymbol{v}\boldsymbol{v}^H|^{-T_p} \text{etr}\left\{-\boldsymbol{X}^H\left(\boldsymbol{R} + P\boldsymbol{v}\boldsymbol{v}^H\right)^{-1}\boldsymbol{X}\right\}$$
$$= p_0(\boldsymbol{X})\left(1 + P\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}\right)^{-T_p} \exp\left\{\frac{P\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{S}_x\boldsymbol{R}^{-1}\boldsymbol{v}}{1 + P\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}}\right\}. \quad (24)$$

where $\boldsymbol{S}_x = \boldsymbol{X}\boldsymbol{X}^H$. Let $u = \boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}$ and $v = \boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{S}_x\boldsymbol{R}^{-1}\boldsymbol{v}$. Some simple calculations enable one to prove that

$$\max_P \frac{p_1(\boldsymbol{X})}{p_0(\boldsymbol{X})} = \begin{cases} 1 & v \le uT_p \\ \left(\frac{v}{uT_p}\right)^{-T_p} \exp\left\{T_p\left(\frac{v}{uT_p} - 1\right)\right\} & v > uT_p. \end{cases} \quad (25)$$

Let $g(x) = x^{-T_p}\exp\{(x-1)T_p\}$ and $u(.)$ denote the unit-step function, i.e., $u(x) = 1$ for $x > 0$, and 0 if $x < 0$. Then, the GLRT for known $\boldsymbol{R}$ is given by

$$GLRT_{|\boldsymbol{R}}(\boldsymbol{X}) = 1 + \left[g\left(\frac{\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{S}_x\boldsymbol{R}^{-1}\boldsymbol{v}}{(\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v})T_p}\right) - 1\right] \times u\left(\frac{\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{S}_x\boldsymbol{R}^{-1}\boldsymbol{v}}{(\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v})T_p} - 1\right). \quad (26)$$

In order to make the detector adaptive, à la AMF, $T_s^{-1}\boldsymbol{S}_y$ should be substituted for $\boldsymbol{R}$ in the previous equation.

REFERENCES

[1] E. J. Kelly, "Adaptive detection in non-stationary interference, Part I and Part II," Massachusetts Institute of Technology, Lincoln Laboratory, Lexington, MA, Tech. Rep. 724, June 1985.

[2] ——, "Adaptive detection in non-stationary interference, Part III," Massachusetts Institute of Technology, Lincoln Laboratory, Lexington, MA, Tech. Rep. 761, 24 August 1987.

[3] E. J. Kelly and K. M. Forsythe, "Adaptive detection and parameter estimation for multidimensional signal models," Massachusetts Institute of Technology, Lincoln Laboratory, Lexington, MA, Tech. Rep. 848, April 1989.

[4] E. J. Kelly, "An adaptive detection algorithm," *IEEE Transactions Aerospace Electronic Systems*, vol. 22, no. 1, pp. 115–127, March 1986.

[5] ——, "Performance of an adaptive detection algorithm; rejection of unwanted signals," *IEEE Transactions Aerospace Electronic Systems*, vol. 25, no. 2, pp. 122–133, April 1989.

[6] E. Conte, A. De Maio, and G. Ricci, "GLRT-based adaptive detection algorithm for range-spread targets," *IEEE Transactions Signal Processing*, vol. 49, no. 7, pp. 1336–1348, July 2001.

[7] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Transactions Aerospace Electronic Systems*, vol. 28, no. 1, pp. 208–216, January 1992.

[8] R. S. Raghavan, H. F. Qiu, and D. J. McLaughlin, "CFAR detection in clutter with unknown correlation properties," *IEEE Transactions Aerospace Electronic Systems*, vol. 31, no. 2, pp. 647–657, April 1995.

[9] P. Swerling, "Probability of detection for fluctuating targets," *IRE Transactions Information Theory*, vol. 6, no. 2, pp. 269–308, April 1960.

[10] ——, "Radar probability of detection for some additional fluctuating target cases," *IEEE Transactions Aerospace Electronic Systems*, vol. 33, no. 2, pp. 698–709, April 1997.