DOI: 10.1109/TSP.2015.2451117
To cite this article: F. Bandiera, O. Besson, A. Coluccia and G. Ricci, "ABORT-like detectors: a Bayesian approach," in IEEE Transactions on Signal Processing, vol. 63, no. 19, pp. 5274-5284, June 30, 2015, doi: 10.1109/TSP.2015.2451117.

ABORT-Like Detectors: A Bayesian Approach

Francesco Bandiera, *Member, IEEE*, Olivier Besson, *Senior Member, IEEE*, Angelo Coluccia, *Member, IEEE*, and Giuseppe Ricci, *Senior Member, IEEE*

***Abstract*—In this paper, we deal with the problem of adaptive radar detection of point-like targets in presence of noise with unknown spectral properties. As customary, we assume that a set of data sharing the same properties of the noise in the cell under test is available. To cope with a limited number of training data, a Bayesian framework is adopted at the design stage. In order to come up with detectors with good rejection capabilities, the possible presence of a fictitious signal under the null hypothesis is modeled probabilistically, as opposite to the conventional ABORT-like approach. Several detectors are devised for the problem at hand, with different complexities. The performance assessment, conducted by means of Monte Carlo simulations, reveals that a good trade-off between detection power and selectivity can be achieved, even assuming a limited number of training data.**

***Index Terms*—Adaptive detection, Bayesian estimation, orthogonal rejection.**

I. Introduction

RESEARCH activity in the area of adaptive radar detection of targets embedded in Gaussian or non-Gaussian disturbance has received a great attention in the last decades. Most of the published works follow the lead of the seminal paper by Kelly [1], where the generalized likelihood ratio test (GLRT) is used to conceive an adaptive decision scheme capable of detecting coherent pulse trains in presence of Gaussian disturbance with unknown spectral properties. Other classical examples of detection algorithms include the adaptive matched filter (AMF) [2] and the adaptive normalized matched filter [3], also known as adaptive coherence estimator [4]. More generally, such problem has been extensively studied and a large number of solutions are available in the open literature, see [5] for a list of references.

Since most of the previously quoted solutions suppose exact knowledge of the signal array response vector, they may experience a performance degradation when the actual steering vector is not aligned with the nominal one. Therefore, it is of interest to take into account possible steering vector mismatches at the design stage of the detector. To this end, coherent returns can be roughly classified into mainlobe and sidelobe signals [6], [7]

- a sidelobe signal (sidelobe interferer) is a coherent signal from a "direction significantly different" to that in which the radar system is steered: it can be due, for example, to a strong target located in a sidelobe direction or to the retransmission of a modulated signal (for the purpose of degrading the reception of the signal of interest);
- a mainlobe target is instead a coherent signal backscattered from the nominal direction or a direction slightly different from the nominal one as a consequence of an imperfect modeling of the nominal steering vector, where the mismatch may be due to multipath propagation, array calibration uncertainties, beampointing errors, etc.

Thus, the effectiveness of the detector depends on its ability to detect the presence of what is classified as mainlobe target, limiting the number of "false alarms" due to sidelobe interferers.

An effective approach to deal with this problem has been introduced in [6], where the adaptive beamformer orthogonal rejection test (ABORT) is proposed; such a detector takes into account rejection capabilities at the design stage. The idea of the ABORT is to modify the null hypothesis, which usually states that data under test contains noise only, so that it possibly contains a fictitious signal which, in some way, is orthogonal to the assumed target's signature. Doing so, if a mismatched signal is present, the detector will be less inclined to declare a detection, as the null hypothesis will be more plausible than in the case where, under the null hypothesis, the test vector contains noise only. The extension of this idea to the case of signals belonging to known subspaces of the observables has been dealt with in [7], as a possible means to maintain an acceptable detection loss for slightly mismatched mainlobe targets. It is important to observe that in the detector proposed in [6] the fictitious signal is assumed to be orthogonal to the nominal steering vector in the "quasi-whitened" space, i.e., after whitening of the data through the sample covariance matrix computed over the secondary data set. The same approach is also proposed in [7], where the quasi-whitening transformation is presented as a way to face with the absence of knowledge about the interference subspace. Following the above mentioned approach, in [8] it is assumed that the useful and the fictitious signals are orthogonal in the "truly" whitened observation space, i.e., after whitening with the true noise covariance matrix. The resulting detector, called the whitened-ABORT (W-ABORT), exhibits good selectivity properties at the price of a performance loss for matched signals.

Mainlobe and/or sidelobe targets have also been represented using the tools of subspace detection in [9]–[13] or constraining the corresponding steering vector to belong to a cone [14]–[16].

The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Hongbin Li.

F. Bandiera, A. Coluccia, and G. Ricci are with the Dipartimento di Ingegneria dell'Innovazione, Università del Salento, 73100 Lecce, Italy (e-mail: francesco.bandiera@unisalento.it; angelo.coluccia@unisalento.it; giuseppe.ricci@unisalento.it).

O. Besson is with the Départment Electronique Optronique Signal, Université de Toulouse, Toulouse 31055, France (e-mail: olivier.besson@isae-supaero.fr).

Another class of receivers designed to deal with the problem of steering vector mismatches is that of the so-called two-stage detectors, i.e., detection architectures composed by cascading two detectors with different behaviors: the overall one declares the presence of a target if the cell under test survive both detection thresholdings, see [17]–[23]. Furthermore, the detector proposed in [24] is a special case of the more general class of tunable (possibly space-time) detectors which have been shown to be an effective means to attack detection of mainlobe targets or rejection of sidelobe ones [24], [25].

Most of the above papers rely on the assumption that a set of secondary data, namely radar returns free of signals components, but sharing certain properties of the noise in the data under test, is available. Such secondary data are usually used to come up with fully-adaptive detection schemes. However, it has been evidenced that the homogeneous assumption for the secondary data is an idealized situation [26], [27], and that non-homogeneous environments are more commonly encountered. This fact calls, in turn, for the necessity to design adaptive algorithms capable of working with a number of secondary data as small as possible. To this aim, the tools of Bayesian estimation can be adopted, which allow to obtain general and flexible models, without very restrictive assumptions [28]–[35]; however, the quoted papers mainly deal with the problem of modeling the noise characteristics rather than the signal ones. [36] is a first attempt to introduce randomness in the fictitious signal.

In this paper, we use Bayesian statistics to address the problem of adaptive radar detection of point-like targets under steering vector mismatches. More in detail, we extend the ABORT framework of [6], [8] to the case where also the useful and the fictitious signals (in addition to the disturbance covariance matrix) are modeled as random quantities with some preassigned *a priori* distribution. We treat both the case where the useful and the interfering signals are orthogonal in the original space and the case where they are orthogonal in the whitened observation space. We propose *ad hoc* detection strategies based upon the likelihood ratio principle in which the unknown quantities are estimated according to several criteria. The performance assessment, carried out by Monte Carlo simulation, shows that some of the proposed solutions can provide good selectivity properties while retaining an optimal performance for matched signals. Remarkably, they are especially attractive when the number of secondary data is very small and the conventional detectors cannot be used.

The remainder of the paper is organized as follows. Section II is devoted to the problem formulation while Section III addresses the design of the proposed detectors. Section IV presents the performance assessment and, finally, Section V contains the conclusions of the work.

II. Problem Formulation

We assume the radar is equipped with a linear array formed by M_a antennas that collect M_t samples from the range cell under test (CUT). The signal received from the CUT could be noise only or a noisy version of the signal backscattered by the target that we model as a coherent target echo. As customary, we suppose that a set of secondary data, free of signal components, but sharing the same statistical properties of the remaining disturbance in the CUT, is available. Secondary data are usually chosen as range cells surrounding the CUT in order to preserve the homogeneity assumption.

We denote by $\boldsymbol{x} \in \mathbb{C}^{M\times 1}$, $M = M_a \times M_t$, the vector containing the returns from the CUT and by $\boldsymbol{x}_t \in \mathbb{C}^{M\times 1}$, $t = 1, \ldots, T$, the secondary data. The problem of detecting the possible presence of a coherent return from a given cell (in range, doppler, and azimuth) is formulated in terms of the following hypothesis testing problem

$$H_0 : \begin{cases} \boldsymbol{x} = \boldsymbol{u} + \boldsymbol{n} \\ \boldsymbol{x}_t = \boldsymbol{n}_t; \quad t = 1, \ldots, T \end{cases}$$
$$H_1 : \begin{cases} \boldsymbol{x} = \alpha \boldsymbol{v} + \boldsymbol{n} \\ \boldsymbol{x}_t = \boldsymbol{n}_t; \quad t = 1, \ldots, T \end{cases} \tag{1}$$

where

- $\boldsymbol{v} \in \mathbb{C}^{M\times 1}$ stands for the signal of interest (known steering vector of the target) and α denotes its amplitude, modeled as a complex normal random variable, i.e., $\alpha \sim \mathcal{CN}(0, \sigma_\alpha^2)$, with σ_α^2 possibly very large.
- $\boldsymbol{n}$ and $\boldsymbol{n}_t \in \mathbb{C}^{M\times 1}$ are conditionally independent and identically distributed random vectors drawn from a complex Gaussian distribution with zero mean and unknown, positive definite covariance matrix $\boldsymbol{R} \in \mathbb{C}^{M\times M}$. More precisely, $\boldsymbol{n}, \boldsymbol{n}_t | \boldsymbol{R} \sim \mathcal{CN}_M(\boldsymbol{0}, \boldsymbol{R})$. As to $\boldsymbol{R}$, it is assumed to follow a (complex) inverse Wishart distribution with ν ($\nu > M$) degrees of freedom and mean $\bar{\boldsymbol{R}}$, i.e.,

$$\pi(\boldsymbol{R}) = \frac{\det((\nu - M)\bar{\boldsymbol{R}})^\nu}{\tilde{\Gamma}_M(\nu)\det(\boldsymbol{R})^{\nu+M}} \mathrm{etr}\left\{-(\nu - M)\boldsymbol{R}^{-1}\bar{\boldsymbol{R}}\right\} \tag{2}$$

where $\det(\cdot)$ and $\mathrm{etr}\{\cdot\}$ denote the determinant and the exponential of the trace of the matrix between braces, respectively, and $\tilde{\Gamma}_M(\nu)$ is given by

$$\tilde{\Gamma}_M(\nu) = \pi^{\frac{M(M-1)}{2}} \prod_{m=1}^{M} \Gamma(\nu - m + 1)$$

with $\Gamma(\cdot)$ denoting the Euler's Gamma function.
Note that an inverse Wishart distribution for $\boldsymbol{R}$ is a conjugate distribution with respect to the Gaussian model for $\boldsymbol{n}, \boldsymbol{n}_t$, which will make derivation of posterior distributions relatively easy. Such a choice is also a convenient way to introduce colored loading in a Bayesian framework. In the following we assume however that $\bar{\boldsymbol{R}} = \mu \boldsymbol{I}_M$, i.e., a non-informative prior, which corresponds to diagonal loading (DL). This has the advantage of not requiring knowledge of the mean of $\boldsymbol{R}$, and at the same time DL is known to be very effective in low sample support [37]. For brevity we will use the notation $\boldsymbol{R} \sim \mathcal{CW}^{-1}(\nu, (\nu - M)\mu \boldsymbol{I}_M)$.
- As anticipated in Section I, $\boldsymbol{u}$ is introduced to enhance the selectivity of the detector. In fact, it makes the detector less inclined to declare the presence of the nominal signal (H_1 hypothesis) in case of mismatches. $\boldsymbol{u}$ is modeled as an unknown random vector taking on values either orthogonal to $\boldsymbol{v}$ ($\boldsymbol{v}^H \boldsymbol{u} = 0$) or conditionally orthogonal to $\boldsymbol{v}$ in the

whitened space (i.e., $\boldsymbol{v}^H \boldsymbol{R}^{-1} \boldsymbol{u} = 0$), where ^{H} denotes conjugate transpose.

III. Detector Designs

In this section, we propose several detectors that implement a ratio of conditional likelihoods where the conditioning (random) quantities are replaced by proper Bayesian estimates. We consider both $\boldsymbol{u}$ orthogonal to $\boldsymbol{v}$ and $\boldsymbol{R}^{-1/2}\boldsymbol{u}$ orthogonal to $\boldsymbol{R}^{-1/2}\boldsymbol{v}$; for the former case several alternatives are devised.

A. Case of $\boldsymbol{u}$ Orthogonal to $\boldsymbol{v}$

Let us first consider the problem under H_0 and denote by $\boldsymbol{V}_\perp$ an $M \times (M-1)$ matrix whose columns form an orthonormal basis for the subspace orthogonal to $\boldsymbol{v}$, i.e., $\boldsymbol{V}_\perp^H \boldsymbol{V}_\perp = \boldsymbol{I}_{M-1}$ and $\boldsymbol{V}_\perp^H \boldsymbol{v} = \boldsymbol{0}$, with $\boldsymbol{I}_N$ the identity matrix of order N. Since $\boldsymbol{u} \perp \boldsymbol{v}$, it follows that $\boldsymbol{u} = \boldsymbol{V}_\perp \boldsymbol{b}$ for some $\boldsymbol{b} \in \mathbb{C}^{(M-1)\times 1}$. The joint likelihood function of $\boldsymbol{x}, \boldsymbol{X} = [\boldsymbol{x}_1 \cdots \boldsymbol{x}_T]$ is thus given by

$$p_0(\boldsymbol{x}, \boldsymbol{X}|\boldsymbol{b}, \boldsymbol{R}) = \frac{1}{\pi^{M(T+1)}} \det(\boldsymbol{R})^{-(T+1)} \times \exp\left\{-(\boldsymbol{x} - \boldsymbol{V}_\perp \boldsymbol{b})^H \boldsymbol{R}^{-1} (\boldsymbol{x} - \boldsymbol{V}_\perp \boldsymbol{b})\right\} \operatorname{etr}\left\{-\boldsymbol{R}^{-1}\boldsymbol{S}\right\} \quad (3)$$

where $\boldsymbol{S} = \boldsymbol{X}\boldsymbol{X}^H$.

We assume that $\boldsymbol{b}$ follows a Bernoulli-Gaussian distribution i.e., its probability density function (pdf) is given by

$$\pi(\boldsymbol{b}) = (1-p)\delta(\|\boldsymbol{b}\|) + p\pi^{-(M-1)}\sigma_\beta^{-2(M-1)} \exp\left\{-\sigma_\beta^{-2}\boldsymbol{b}^H\boldsymbol{b}\right\} \quad (4)$$

where $\|\cdot\|$ is the Euclidean norm, $\delta(\cdot)$ is the Dirac delta, p stands for the probability that $\boldsymbol{b} \neq \boldsymbol{0}$, and σ_β^2 is possibly very large. We denote this distribution as $\boldsymbol{b} \sim \mathrm{BG}(p, \boldsymbol{0}, \sigma_\beta^2 \boldsymbol{I}_{M-1})$. An equivalent and possibly more convenient way to represent $\boldsymbol{b}$ is to write $\boldsymbol{b} = \iota\boldsymbol{\beta}$ where $\iota \in \{0,1\}$ is a binary variable, and $\boldsymbol{\beta} \in \mathbb{C}^{(M-1)\times 1}$. From (4), ι follows a Bernoulli distribution with parameter p, i.e., $\mathrm{Prob}[\iota = 1] = p$: we denote this distribution as $\iota \sim \mathrm{Ber}(p)$. With some slight abuse, we will write $\pi(\iota) = (1-p)^{1-\iota}\, p^\iota$, keeping in mind that ι is binary. Moreover, ι is independent of $\boldsymbol{\beta}$: the latter is Gaussian distributed, with mean zero and covariance matrix $\sigma_\beta^2 \boldsymbol{I}_{M-1}$, i.e., $\boldsymbol{\beta} \sim \mathcal{CN}_{M-1}(\boldsymbol{0}, \sigma_\beta^2 \boldsymbol{I}_{M-1})$. We will indifferently use the parametrization in terms of $\boldsymbol{b}$ or in terms of $(\iota, \boldsymbol{\beta})$, depending on which one is more convenient.

We will investigate three approaches. In the first one, joint estimation of $\boldsymbol{b}$ (or ι and $\boldsymbol{\beta}$) and $\boldsymbol{R}$ will be conducted, while in a second approach, we will marginalize with respect to $\boldsymbol{R}$, leaving estimation of $\boldsymbol{b}$ only. The third method entails an approach *à la AMF* where detection is carried out on $\boldsymbol{x}$ only, assuming $\boldsymbol{R}$ is known: an estimate of $\boldsymbol{R}$ based on $\boldsymbol{X}$ only is then substituted for $\boldsymbol{R}$ in the detection statistic.

1) Joint Estimation of $\boldsymbol{b}$ and $\boldsymbol{R}$: First observe that under H_0 the joint posterior distribution of ι, $\boldsymbol{\beta}$, and $\boldsymbol{R}$ is given by

$$\begin{aligned} p_0(\iota, \boldsymbol{\beta}, \boldsymbol{R}|\boldsymbol{x}, \boldsymbol{X}) &\propto p_0(\boldsymbol{x}, \boldsymbol{X}|\iota, \boldsymbol{\beta}, \boldsymbol{R})\pi(\iota)\pi(\boldsymbol{\beta})\pi(\boldsymbol{R}) \\ &\propto (1-p)^{1-\iota}p^\iota \exp\left\{-\sigma_\beta^{-2}\boldsymbol{\beta}^H\boldsymbol{\beta}\right\} \det \boldsymbol{R}^{-(\nu+M+T+1)} \\ &\quad \times \exp\left\{-(\boldsymbol{x} - \iota\boldsymbol{V}_\perp\boldsymbol{\beta})^H \boldsymbol{R}^{-1} (\boldsymbol{x} - \iota\boldsymbol{V}_\perp\boldsymbol{\beta})\right\} \operatorname{etr}\left\{-\boldsymbol{R}^{-1}\boldsymbol{S}\right\} \\ &\quad \times \operatorname{etr}\left\{-(\nu - M)\mu\boldsymbol{R}^{-1}\right\} \end{aligned} \quad (5)$$

where $\propto$ denotes proportional to.

Let us now derive the conditional posterior distributions of ι, $\boldsymbol{\beta}$, and $\boldsymbol{R}$. Since

$$\begin{aligned} T_0 &= (\boldsymbol{x} - \iota\boldsymbol{V}_\perp\boldsymbol{\beta})^H \boldsymbol{R}^{-1} (\boldsymbol{x} - \iota\boldsymbol{V}_\perp\boldsymbol{\beta}) + \sigma_\beta^{-2}\boldsymbol{\beta}^H\boldsymbol{\beta} \\ &= \boldsymbol{x}^H\boldsymbol{R}^{-1}\boldsymbol{x} - \iota\boldsymbol{\beta}^H\boldsymbol{V}_\perp^H\boldsymbol{R}^{-1}\boldsymbol{x} - \iota\boldsymbol{x}^H\boldsymbol{R}^{-1}\boldsymbol{V}_\perp\boldsymbol{\beta} \\ &\quad + \iota^2\boldsymbol{\beta}^H\boldsymbol{V}_\perp^H\boldsymbol{R}^{-1}\boldsymbol{V}_\perp\boldsymbol{\beta} + \sigma_\beta^{-2}\boldsymbol{\beta}^H\boldsymbol{\beta} \\ &= \boldsymbol{x}^H\boldsymbol{R}^{-1}\boldsymbol{x} - \iota^2\boldsymbol{x}^H\boldsymbol{R}^{-1}\boldsymbol{V}_\perp\boldsymbol{\Sigma}_\beta\left(\iota, \boldsymbol{R}^{-1}\right)\boldsymbol{V}_\perp^H\boldsymbol{R}^{-1}\boldsymbol{x} \\ &\quad + \left[\boldsymbol{\beta} - \boldsymbol{\mu}_\beta\left(\iota, \boldsymbol{R}^{-1}, \boldsymbol{x}\right)\right]^H \boldsymbol{\Sigma}_\beta^{-1}\left(\iota, \boldsymbol{R}^{-1}\right) \\ &\quad \times \left[\boldsymbol{\beta} - \boldsymbol{\mu}_\beta\left(\iota, \boldsymbol{R}^{-1}, \boldsymbol{x}\right)\right] \end{aligned} \quad (6)$$

with

$$\boldsymbol{\Sigma}_\beta\left(\iota, \boldsymbol{R}^{-1}\right) = \left[\iota^2\boldsymbol{V}_\perp^H\boldsymbol{R}^{-1}\boldsymbol{V}_\perp + \sigma_\beta^{-2}\boldsymbol{I}_{M-1}\right]^{-1} \quad (7a)$$

$$\boldsymbol{\mu}_\beta\left(\iota, \boldsymbol{R}^{-1}, \boldsymbol{x}\right) = \iota\boldsymbol{\Sigma}_\beta\left(\iota, \boldsymbol{R}^{-1}\right)\boldsymbol{V}_\perp^H\boldsymbol{R}^{-1}\boldsymbol{x}. \quad (7b)$$

It ensues that the posterior distribution of $\boldsymbol{\beta}|\iota, \boldsymbol{R}, \boldsymbol{x}, \boldsymbol{X}$ is Gaussian

$$\boldsymbol{\beta}|\iota, \boldsymbol{R}, \boldsymbol{x}, \boldsymbol{X} \sim \mathcal{CN}_{M-1}\left(\boldsymbol{\mu}_\beta\left(\iota, \boldsymbol{R}^{-1}, \boldsymbol{x}\right), \boldsymbol{\Sigma}_\beta\left(\iota, \boldsymbol{R}^{-1}\right)\right). \quad (8)$$

Similarly

$$p_0(\iota|\boldsymbol{\beta}, \boldsymbol{R}, \boldsymbol{x}, \boldsymbol{X}) \propto (1-p)^{1-\iota}p^\iota \times \exp\left\{\iota\left[2\mathrm{Re}(\boldsymbol{\beta}^H\boldsymbol{V}_\perp^H\boldsymbol{R}^{-1}\boldsymbol{x}) - \boldsymbol{\beta}^H\boldsymbol{V}_\perp^H\boldsymbol{R}^{-1}\boldsymbol{V}_\perp\boldsymbol{\beta}\right]\right\} \quad (9)$$

which implies that

$$\iota|\boldsymbol{\beta}, \boldsymbol{R}, \boldsymbol{x}, \boldsymbol{X} \sim \mathrm{Ber}\left(\mathrm{Prob}\left[\iota = 1|\boldsymbol{\beta}, \boldsymbol{R}, \boldsymbol{x}, \boldsymbol{X}\right]\right) \quad (10)$$

with $\mathrm{Prob}[\iota = 1|\boldsymbol{\beta}, \boldsymbol{R}, \boldsymbol{x}, \boldsymbol{X}]$ given in (11), shown at the bottom of the page. Finally, we have

$$p_0(\boldsymbol{R}|\iota, \boldsymbol{\beta}, \boldsymbol{x}, \boldsymbol{X}) \propto \det(\boldsymbol{R})^{-(\nu+M+T+1)} \operatorname{etr}\left\{-\boldsymbol{R}^{-1}\boldsymbol{M}(\iota, \boldsymbol{\beta}, \boldsymbol{x}, \boldsymbol{X})\right\} \quad (12)$$

with

$$\boldsymbol{M}(\iota, \boldsymbol{\beta}, \boldsymbol{x}, \boldsymbol{X}) = (\nu - M)\mu\boldsymbol{I}_M + \boldsymbol{S} + (\boldsymbol{x} - \iota\boldsymbol{V}_\perp\boldsymbol{\beta})(\boldsymbol{x} - \iota\boldsymbol{V}_\perp\boldsymbol{\beta})^H. \quad (13)$$

This conditional posterior distribution is a complex inverse Wishart distribution with $\nu + T + 1$ degrees of freedom and parameter matrix $\boldsymbol{M}(\iota, \boldsymbol{\beta}, \boldsymbol{x}, \boldsymbol{X})$.

$$\mathrm{Prob}\left[\iota = 1|\boldsymbol{\beta}, \boldsymbol{R}, \boldsymbol{x}, \boldsymbol{X}\right] = \frac{p\exp\left\{2\mathrm{Re}(\boldsymbol{\beta}^H\boldsymbol{V}_\perp^H\boldsymbol{R}^{-1}\boldsymbol{x}) - \boldsymbol{\beta}^H\boldsymbol{V}_\perp^H\boldsymbol{R}^{-1}\boldsymbol{V}_\perp\boldsymbol{\beta}\right\}}{(1-p) + p\exp\left\{2\mathrm{Re}(\boldsymbol{\beta}^H\boldsymbol{V}_\perp^H\boldsymbol{R}^{-1}\boldsymbol{x}) - \boldsymbol{\beta}^H\boldsymbol{V}_\perp^H\boldsymbol{R}^{-1}\boldsymbol{V}_\perp\boldsymbol{\beta}\right\}}. \quad (11)$$

Therefore, the conditional distributions $p_0(\boldsymbol{\beta}|\iota, \boldsymbol{R}, \boldsymbol{x}, \boldsymbol{X})$, $p_0(\iota|\boldsymbol{\beta}, \boldsymbol{R}, \boldsymbol{x}, \boldsymbol{X})$, and $p_0(\boldsymbol{R}|\iota, \boldsymbol{\beta}, \boldsymbol{x}, \boldsymbol{X})$ all belong to known families and are easy to simulate. This suggests the use of a Gibbs sampler which enables one to obtain the minimum mean-square error (MMSE) estimates of ι, $\boldsymbol{\beta}$, and $\boldsymbol{R}$. The Gibbs sampler is summarized in Table I, where it is shown how to sample ι, $\boldsymbol{\beta}$, and $\boldsymbol{R}$, thus obtaining the sequences $\iota(n)$, $\boldsymbol{\beta}(n)$, and $\boldsymbol{R}(n)$, respectively, using the previously introduced posterior distributions. The MMSE estimates can be obtained by averaging the N_r samples of $\boldsymbol{b}(n) = \iota(n)\boldsymbol{\beta}(n)$ and $\boldsymbol{R}(n)$ after the so-called burning period of duration N_{bi}, i.e.,

$$\hat{\boldsymbol{b}}_{\text{mmse}} = \frac{1}{N_r} \sum_{n=N_{bi}+1}^{N_{bi}+N_r} \boldsymbol{b}(n)$$

and

$$\hat{\boldsymbol{R}}^0_{\text{mmse}} = \frac{1}{N_r} \sum_{n=N_{bi}+1}^{N_{bi}+N_r} \boldsymbol{R}(n).$$

Under H_1, the likelihood function is given by

$$p_1(\boldsymbol{x}, \boldsymbol{X}|\alpha, \boldsymbol{R}) \propto \det(\boldsymbol{R})^{-(T+1)} \text{etr}\left\{-\boldsymbol{R}^{-1}\boldsymbol{S}\right\} \times \exp\left\{-\left(\boldsymbol{x}-\alpha\boldsymbol{v}\right)^H \boldsymbol{R}^{-1}\left(\boldsymbol{x}-\alpha\boldsymbol{v}\right)\right\}. \quad (14)$$

The posterior distribution of α and $\boldsymbol{R}$ is then

$$\begin{aligned} p_1(\alpha, \boldsymbol{R}|\boldsymbol{x}, \boldsymbol{X}) &\propto p_1(\boldsymbol{x}, \boldsymbol{X}|\alpha, \boldsymbol{R})\pi(\alpha)\pi(\boldsymbol{R}) \\ &\propto \det(\boldsymbol{R})^{-(\nu+M+T+1)} \text{etr}\left\{-\boldsymbol{R}^{-1}\boldsymbol{S}\right\} \\ &\quad \times \text{etr}\left\{-(\nu-M)\mu\boldsymbol{R}^{-1}\right\} \\ &\quad \times \exp\left\{-\left(\boldsymbol{x}-\alpha\boldsymbol{v}\right)^H \boldsymbol{R}^{-1}\left(\boldsymbol{x}-\alpha\boldsymbol{v}\right) - \sigma_\alpha^{-2}|\alpha|^2\right\}. \end{aligned} \quad (15)$$

Since

$$\begin{aligned} T_1 &= \left(\boldsymbol{x}-\alpha\boldsymbol{v}\right)^H \boldsymbol{R}^{-1}\left(\boldsymbol{x}-\alpha\boldsymbol{v}\right) + \sigma_\alpha^{-2}|\alpha|^2 \\ &= \left(\sigma_\alpha^{-2} + \boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}\right)\left|\alpha - \frac{\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{x}}{\left(\sigma_\alpha^{-2} + \boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}\right)}\right|^2 \\ &\quad + \boldsymbol{x}^H\boldsymbol{R}^{-1}\boldsymbol{x} - \frac{|\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{x}|^2}{\left(\sigma_\alpha^{-2} + \boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}\right)} \end{aligned} \quad (16)$$

it follows that

TABLE I
GIBBS SAMPLER FOR ESTIMATION OF ι, $\boldsymbol{\beta}$, AND $\boldsymbol{R}$

Require: initial values $\boldsymbol{R}(0)$, $\boldsymbol{\beta}(0)$
1: **for** $n = 1, \cdots, N_{bi} + N_r$ **do**
2: sample $\iota(n)$ from Ber $\left(\text{Prob}\left[\iota = 1|\boldsymbol{\beta}(n-1), \boldsymbol{R}(n-1), \boldsymbol{x}, \boldsymbol{X}\right]\right)$;
3: sample $\boldsymbol{R}(n)$ from $p\left(\boldsymbol{R}|\iota(n), \boldsymbol{\beta}(n-1), \boldsymbol{x}, \boldsymbol{X}\right)$;
4: sample $\boldsymbol{\beta}(n)$ from $p\left(\boldsymbol{\beta}|\iota(n), \boldsymbol{R}(n), \boldsymbol{x}, \boldsymbol{X}\right)$;
5: compute $\boldsymbol{b}(n) = \iota(n) \times \boldsymbol{\beta}(n)$;
6: **end for**
Ensure: sequence of random variables $\boldsymbol{b}(n)$, $\boldsymbol{R}(n)$.

$$\alpha|\boldsymbol{R}, \boldsymbol{x}, \boldsymbol{X} \sim \mathcal{CN}\left(\frac{\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{x}}{\sigma_\alpha^{-2} + \boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}}, \left(\sigma_\alpha^{-2} + \boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}\right)^{-1}\right). \quad (17)$$

Accordingly, it is straightforward to show that

$$\boldsymbol{R}|\alpha, \boldsymbol{x}, \boldsymbol{X} \sim \mathcal{CW}^{-1}\left(\nu + T + 1, \tilde{\boldsymbol{S}} + \left(\boldsymbol{x}-\alpha\boldsymbol{v}\right)\left(\boldsymbol{x}-\alpha\boldsymbol{v}\right)^H\right) \quad (18)$$

where

$$\tilde{\boldsymbol{S}} = (\nu - M)\mu\boldsymbol{I}_M + \boldsymbol{S}. \quad (19)$$

The above posterior distributions allow one to obtain the MMSE estimates of α, $\boldsymbol{R}$ under H_1 by resorting again to a Gibbs sampler; accordingly, the corresponding estimates are denoted by $\hat{\alpha}_{\text{mmse}}$ and $\hat{\boldsymbol{R}}^1_{\text{mmse}}$.

Based on the above computed estimates, the test is given by

$$t(\boldsymbol{x}, \boldsymbol{X}) = \frac{p_1(\boldsymbol{x}, \boldsymbol{X}|\hat{\alpha}_{\text{mmse}}, \hat{\boldsymbol{R}}^1_{\text{mmse}})}{p_0(\boldsymbol{x}, \boldsymbol{X}|\hat{\boldsymbol{b}}_{\text{mmse}}, \hat{\boldsymbol{R}}^0_{\text{mmse}})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad (20)$$

where η is a threshold to be set according to the desired probability of false alarm (P_{fa}). This detector will be referred to in the following as one-step detector with rejection in the orthogonal initial space (1S-OIS).

2) Marginalizing With Respect to $\boldsymbol{R}$*:* An alternative approach consists in marginalizing with respect to the covariance matrix, so as to leave estimation of $\boldsymbol{b}$ only under H_0 and α only under H_1. To be more specific, let us marginalize the likelihood function in (3) with respect to $\boldsymbol{R}$ ultimately obtaining (21), shown at the bottom of the page, where $\tilde{\boldsymbol{S}}$ is given by (19). A straightforward calculation shows that

$$\left(\boldsymbol{x} - \boldsymbol{V}_\perp\boldsymbol{b}\right)^H \tilde{\boldsymbol{S}}^{-1}\left(\boldsymbol{x} - \boldsymbol{V}_\perp\boldsymbol{b}\right) = \frac{|\boldsymbol{x}^H\boldsymbol{v}|^2}{\boldsymbol{v}^H\tilde{\boldsymbol{S}}\boldsymbol{v}} + \left(\boldsymbol{b} - \boldsymbol{b}_0\right)^H\left(\boldsymbol{V}_\perp^H\tilde{\boldsymbol{S}}^{-1}\boldsymbol{V}_\perp\right)\left(\boldsymbol{b} - \boldsymbol{b}_0\right) \quad (22)$$

$$\begin{aligned} p_0(\boldsymbol{x}, \boldsymbol{X}|\iota, \boldsymbol{\beta}) &= \int p_0(\boldsymbol{x}, \boldsymbol{X}|\iota, \boldsymbol{\beta}, \boldsymbol{R})\pi(\boldsymbol{R})d\boldsymbol{R} \\ &\propto \int \det(\boldsymbol{R})^{-(\nu+M+T+1)} \text{etr}\left\{-\boldsymbol{R}^{-1}\left[\boldsymbol{S} + (\nu-M)\mu\boldsymbol{I}_M + \left(\boldsymbol{x} - \iota\boldsymbol{V}_\perp\boldsymbol{\beta}\right)\left(\boldsymbol{x} - \iota\boldsymbol{V}_\perp\boldsymbol{\beta}\right)^H\right]\right\} d\boldsymbol{R} \\ &\propto \det\left[\boldsymbol{S} + (\nu-M)\mu\boldsymbol{I}_M + \left(\boldsymbol{x} - \iota\boldsymbol{V}_\perp\boldsymbol{\beta}\right)\left(\boldsymbol{x} - \iota\boldsymbol{V}_\perp\boldsymbol{\beta}\right)^H\right]^{-(\nu+T+1)} \\ &\propto \left\{\left[1 + \left(\boldsymbol{x} - \iota\boldsymbol{V}_\perp\boldsymbol{\beta}\right)^H \tilde{\boldsymbol{S}}^{-1}\left(\boldsymbol{x} - \iota\boldsymbol{V}_\perp\boldsymbol{\beta}\right)\right]\det(\tilde{\boldsymbol{S}})\right\}^{-(\nu+T+1)} \end{aligned} \quad (21)$$

with $\boldsymbol{b}_0 = (\boldsymbol{V}_\perp^H \tilde{\boldsymbol{S}}^{-1} \boldsymbol{V}_\perp)^{-1} \boldsymbol{V}_\perp^H \tilde{\boldsymbol{S}}^{-1} \boldsymbol{x}$. Therefore,

$$p_0(\boldsymbol{x}, \boldsymbol{X}|\iota, \boldsymbol{\beta}) \propto \left[\left(1 + \frac{|\boldsymbol{x}^H \boldsymbol{v}|^2}{\boldsymbol{v}^H \tilde{\boldsymbol{S}} \boldsymbol{v}}\right) \times \det(\tilde{\boldsymbol{S}}) \right]^{-(\nu+T+1)} \times \left[1 + (\iota\boldsymbol{\beta} - \boldsymbol{b}_0)^H \boldsymbol{\Sigma}^{-1} (\iota\boldsymbol{\beta} - \boldsymbol{b}_0)\right]^{-(\nu+T+1)} \quad (23)$$

with

$$\boldsymbol{\Sigma} = \left(1 + \frac{|\boldsymbol{x}^H \boldsymbol{v}|^2}{\boldsymbol{v}^H \tilde{\boldsymbol{S}} \boldsymbol{v}}\right) \left(\boldsymbol{V}_\perp^H \tilde{\boldsymbol{S}}^{-1} \boldsymbol{V}_\perp\right)^{-1}. \quad (24)$$

The posterior distribution of ι, $\boldsymbol{\beta}$ is thus

$$p_0(\iota, \boldsymbol{\beta}|\boldsymbol{x}, \boldsymbol{X}) \propto (1-p)^{1-\iota} p^\iota \exp\left\{-\sigma_\beta^{-2} \boldsymbol{\beta}^H \boldsymbol{\beta}\right\} \times \left[1 + (\iota\boldsymbol{\beta} - \boldsymbol{b}_0)^H \boldsymbol{\Sigma}^{-1} (\iota\boldsymbol{\beta} - \boldsymbol{b}_0)\right]^{-(\nu+T+1)}. \quad (25)$$

It ensues that the conditional posterior distribution of $\iota|\boldsymbol{\beta}, \boldsymbol{x}, \boldsymbol{X}$ is still Bernoulli

$$\iota|\boldsymbol{\beta}, \boldsymbol{x}, \boldsymbol{X} \sim \text{Ber}\left(\text{Prob}\left[\iota = 1|\boldsymbol{\beta}, \boldsymbol{x}, \boldsymbol{X}\right]\right) \quad (26)$$

with

$$\text{Prob}\left[\iota = 1|\boldsymbol{\beta}, \boldsymbol{x}, \boldsymbol{X}\right] = \frac{p\eta_1}{(1-p)\eta_0 + p\eta_1} \quad (27a)$$

$$\eta_0 = \left[1 + \boldsymbol{b}_0^H \boldsymbol{\Sigma}^{-1} \boldsymbol{b}_0\right]^{-(\nu+T+1)} \quad (27b)$$

$$\eta_1 = \left[1 + (\boldsymbol{\beta} - \boldsymbol{b}_0)^H \boldsymbol{\Sigma}^{-1} (\boldsymbol{\beta} - \boldsymbol{b}_0)\right]^{-(\nu+T+1)}. \quad (27c)$$

As for the conditional posterior distribution of $\boldsymbol{\beta}$, we have from (25)

$$p_0(\boldsymbol{\beta}|\iota, \boldsymbol{x}, \boldsymbol{X}) \propto \exp\left\{-\sigma_\beta^{-2} \boldsymbol{\beta}^H \boldsymbol{\beta}\right\} \times \left[1 + (\iota\boldsymbol{\beta} - \boldsymbol{b}_0)^H \boldsymbol{\Sigma}^{-1} (\iota\boldsymbol{\beta} - \boldsymbol{b}_0)\right]^{-(\nu+T+1)}. \quad (28)$$

As expected, when $\iota = 0$, $\boldsymbol{\beta}|\iota, \boldsymbol{x}, \boldsymbol{X}$ follows its prior Gaussian distribution. However, when $\iota = 1$, one has a product of Gaussian and Student distributions. In order to obtain the MMSE estimates of ι and $\boldsymbol{\beta}$, one could again investigate a Gibbs sampler. For the sampling of $\boldsymbol{\beta}|\iota = 1, \boldsymbol{x}, \boldsymbol{X}$, we suggest a Metropolis-Hastings [38] approach with a proposal distribution that would be a Student distribution, which amounts, in fact, to ignoring the exponential term in (28).

Let us now turn to the H_1 hypothesis. From (14), we can infer that

$$\begin{aligned} p_1(\boldsymbol{x}, \boldsymbol{X}|\alpha) &= \int p_1(\boldsymbol{x}, \boldsymbol{X}|\alpha, \boldsymbol{R}) \pi(\boldsymbol{R}) d\boldsymbol{R} \\ &\propto \int \det(\boldsymbol{R})^{-(\nu+M+T+1)} \\ &\quad \times \text{etr}\left\{-\boldsymbol{R}^{-1}\left[\tilde{\boldsymbol{S}} + (\boldsymbol{x} - \alpha\boldsymbol{v})(\boldsymbol{x} - \alpha\boldsymbol{v})^H\right]\right\} d\boldsymbol{R} \\ &\propto \det\left[\tilde{\boldsymbol{S}} + (\boldsymbol{x} - \alpha\boldsymbol{v})(\boldsymbol{x} - \alpha\boldsymbol{v})^H\right]^{-(\nu+T+1)} \\ &\propto \left\{\left[1 + (\boldsymbol{x} - \alpha\boldsymbol{v})^H \tilde{\boldsymbol{S}}^{-1} (\boldsymbol{x} - \alpha\boldsymbol{v})\right] \det(\tilde{\boldsymbol{S}})\right\}^{-(\nu+T+1)} \\ &\propto g(\boldsymbol{x}, \boldsymbol{X}) \times \left[\left(1 + \sigma_1^{-2} |\alpha - \alpha_0|^2\right) \det(\tilde{\boldsymbol{S}})\right]^{-(\nu+T+1)} \end{aligned} \quad (29)$$

with $\tilde{\boldsymbol{S}}$ given by (19),

$$\alpha_0 = \left(\boldsymbol{v}^H \tilde{\boldsymbol{S}}^{-1} \boldsymbol{x}\right) / \left(\boldsymbol{v}^H \tilde{\boldsymbol{S}}^{-1} \boldsymbol{v}\right) \quad (30a)$$

$$\sigma_1^2 = \frac{1 + \boldsymbol{x}^H \tilde{\boldsymbol{S}}^{-1} \boldsymbol{x} - \frac{|\boldsymbol{v}^H \tilde{\boldsymbol{S}}^{-1} \boldsymbol{x}|^2}{\boldsymbol{v}^H \tilde{\boldsymbol{S}}^{-1} \boldsymbol{v}}}{\boldsymbol{v}^H \tilde{\boldsymbol{S}}^{-1} \boldsymbol{v}} \quad (30b)$$

and $g(\boldsymbol{x}, \boldsymbol{X}) = \left(1 + \boldsymbol{x}^H \tilde{\boldsymbol{S}}^{-1} \boldsymbol{x} - \frac{|\boldsymbol{x}^H \tilde{\boldsymbol{S}}^{-1} \boldsymbol{v}|^2}{\boldsymbol{v}^H \tilde{\boldsymbol{S}}^{-1} \boldsymbol{v}}\right)^{-(\nu+T+1)}$.

The posterior distribution of α is thus

$$p_1(\alpha|\boldsymbol{x}, \boldsymbol{X}) \propto \pi(\alpha) \left[1 + \sigma_1^{-2} |\alpha - \alpha_0|^2\right]^{-(\nu+T+1)}. \quad (31)$$

If one assumes a flat prior for α, i.e., $\pi(\alpha) \propto 1$, then $\alpha|\boldsymbol{x}, \boldsymbol{X}$ follows a Student distribution and the MMSE estimate of α is simply α_0. On the other hand, with a Gaussian prior for α, the posterior is a product of Gaussian and Student distributions. The MMSE estimator cannot be obtained in closed-form but Markov chain Monte Carlo (MCMC) methods have to be looked for, with a Metropolis-Hastings strategy to draw samples from (31). For the sake of simplicity, we assume that σ_α^2 is very large and we simply approximate the MMSE estimate of α by α_0.

Finally, the above estimates can be used in the following test statistic

$$t(\boldsymbol{x}, \boldsymbol{X}) = \frac{p_1(\boldsymbol{x}, \boldsymbol{X}|\hat{\alpha}_{\text{mmse}})}{p_0(\boldsymbol{x}, \boldsymbol{X}|\hat{\boldsymbol{b}}_{\text{mmse}})} \equiv \frac{1 + \left(\boldsymbol{x} - \boldsymbol{V}_\perp \hat{\boldsymbol{b}}_{\text{mmse}}\right)^H \tilde{\boldsymbol{S}}^{-1} \left(\boldsymbol{x} - \boldsymbol{V}_\perp \hat{\boldsymbol{b}}_{\text{mmse}}\right)}{1 + (\boldsymbol{x} - \hat{\alpha}_{\text{mmse}}\boldsymbol{v})^H \tilde{\boldsymbol{S}}^{-1} (\boldsymbol{x} - \hat{\alpha}_{\text{mmse}}\boldsymbol{v})} \quad (32)$$

where $\equiv$ indicates equivalence of the statistics for the hypothesis testing problem, and the last line follows from (21) and (29). This detector will be referred to in the following as marginalized one-step detector with rejection in the orthogonal initial space (M-1S-OIS).

3) A Two-Step Approach: A third possibility which is worth investigating is a 2-step approach, where $\boldsymbol{R}$ is estimated from the secondary data $\boldsymbol{X}$ only. This estimate is then substituted for $\boldsymbol{R}$ in a test statistic build on $\boldsymbol{x}$ only. More precisely, since $p(\boldsymbol{R}|\boldsymbol{X}) \propto |\boldsymbol{R}|^{-(\nu+M+T)} \text{etr}\{-\boldsymbol{R}^{-1}[(\nu - M)\mu \boldsymbol{I}_M + \boldsymbol{S}]\}$, it follows that the MMSE estimate of $\boldsymbol{R}$ from $\boldsymbol{X}$ is simply

$$\hat{\boldsymbol{R}}_{\text{mmse}} = (\nu + T - M)^{-1} \left[(\nu - M)\mu \boldsymbol{I}_M + \boldsymbol{S}\right] \quad (33)$$

which corresponds to a diagonally loaded version of the sample covariance matrix. Let us consider the posterior distribution of $\boldsymbol{b}|\boldsymbol{R}, \boldsymbol{x}$. From

$$p_0(\boldsymbol{x}|\boldsymbol{b}, \boldsymbol{R}) = \frac{1}{\pi^M} \det(\boldsymbol{R})^{-1} \times \exp\left\{-(\boldsymbol{x} - \boldsymbol{V}_\perp \boldsymbol{b})^H \boldsymbol{R}^{-1} (\boldsymbol{x} - \boldsymbol{V}_\perp \boldsymbol{b})\right\}$$

using (4) and (6), it follows that

$$\begin{aligned} p_0(\boldsymbol{b}|\boldsymbol{R}, \boldsymbol{x}) &\propto (1-p)\delta(\|\boldsymbol{b}\|) \\ &\quad + p\sigma_\beta^{-2(M-1)} \det\left(\boldsymbol{\Sigma}_b\left(\boldsymbol{R}^{-1}\right)\right) \\ &\quad \times \exp\left\{\boldsymbol{\mu}_b^H\left(\boldsymbol{R}^{-1}, \boldsymbol{x}\right) \boldsymbol{\Sigma}_b^{-1}\left(\boldsymbol{R}^{-1}\right) \boldsymbol{\mu}_b\left(\boldsymbol{R}^{-1}, \boldsymbol{x}\right)\right\} \\ &\quad \times \mathcal{CN}_{M-1}\left(\boldsymbol{\mu}_b\left(\boldsymbol{R}^{-1}, \boldsymbol{x}\right), \boldsymbol{\Sigma}_b\left(\boldsymbol{R}^{-1}\right)\right) \end{aligned} \quad (34)$$

where

$$\boldsymbol{\Sigma}_b\left(\boldsymbol{R}^{-1}\right) = \left[\boldsymbol{V}_\perp^H \boldsymbol{R}^{-1}\boldsymbol{V}_\perp + \sigma_\beta^{-2}\boldsymbol{I}_{M-1}\right]^{-1} \quad \text{(35a)}$$

$$\boldsymbol{\mu}_b\left(\boldsymbol{R}^{-1},\boldsymbol{x}\right) = \boldsymbol{\Sigma}_b\left(\boldsymbol{R}^{-1}\right)\boldsymbol{V}_\perp^H \boldsymbol{R}^{-1}\boldsymbol{x}. \quad \text{(35b)}$$

Therefore, $\boldsymbol{b}|\boldsymbol{R},\boldsymbol{x}$ follows a Bernoulli-Gaussian distribution, i.e.,

$$\boldsymbol{b}|\boldsymbol{R},\boldsymbol{x} \sim \mathrm{BG}\left(p_1, \boldsymbol{\mu}_b\left(\boldsymbol{R}^{-1},\boldsymbol{x}\right), \boldsymbol{\Sigma}_b\left(\boldsymbol{R}^{-1}\right)\right). \quad (36)$$

With p_1 given by (37), shown at the bottom of the page. As a consequence, the MMSE estimate of $\boldsymbol{b}$, conditioned on $\boldsymbol{R}$, is

$$\mathcal{E}\left\{\boldsymbol{b}|\boldsymbol{R},\boldsymbol{x}\right\} = p_1\boldsymbol{\mu}_b\left(\boldsymbol{R}^{-1},\boldsymbol{x}\right). \quad (38)$$

The MMSE estimate of α (conditioned on $\boldsymbol{R}$) is, see (17)

$$\mathcal{E}\left\{\alpha|\boldsymbol{R},\boldsymbol{x}\right\} = \frac{\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{x}}{\sigma_\alpha^{-2} + \boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}}. \quad (39)$$

These estimates can now be used to construct the following test statistic

$$\begin{aligned} t(\boldsymbol{x},\boldsymbol{X}) &= \frac{p_1\left(\boldsymbol{x}|\mathcal{E}\left\{\alpha|\hat{\boldsymbol{R}}_{\mathrm{mmse}},\boldsymbol{x}\right\}, \hat{\boldsymbol{R}}_{\mathrm{mmse}}\right)}{p_0\left(\boldsymbol{x}|\mathcal{E}\left\{\boldsymbol{b}|\hat{\boldsymbol{R}}_{\mathrm{mmse}},\boldsymbol{x}\right\}, \hat{\boldsymbol{R}}_{\mathrm{mmse}}\right)} \\ &\equiv \left(\boldsymbol{x} - \boldsymbol{V}_\perp\mathcal{E}\left\{\boldsymbol{b}|\hat{\boldsymbol{R}}_{\mathrm{mmse}},\boldsymbol{x}\right\}\right)^H \\ &\quad\times \hat{\boldsymbol{R}}_{\mathrm{mmse}}^{-1}\left(\boldsymbol{x} - \boldsymbol{V}_\perp\mathcal{E}\left\{\boldsymbol{b}|\hat{\boldsymbol{R}}_{\mathrm{mmse}},\boldsymbol{x}\right\}\right) \\ &\quad- \left(\boldsymbol{x} - \mathcal{E}\left\{\alpha|\hat{\boldsymbol{R}}_{\mathrm{mmse}},\boldsymbol{x}\right\}\boldsymbol{v}\right)^H \\ &\quad\times \hat{\boldsymbol{R}}_{\mathrm{mmse}}^{-1}\left(\boldsymbol{x} - \mathcal{E}\left\{\alpha|\hat{\boldsymbol{R}}_{\mathrm{mmse}},\boldsymbol{x}\right\}\boldsymbol{v}\right). \end{aligned} \quad (40)$$

In contrast to the two previous methods, no MCMC simulation method is required here and the resulting detection scheme remains rather simple. This detector will be referred to in the following as two-step detector with rejection in the orthogonal initial space (2S-OIS).

B. Case of $\boldsymbol{R}^{-1/2}\boldsymbol{u}$ Orthogonal to $\boldsymbol{R}^{-1/2}\boldsymbol{v}$

Let us now consider the alternative hypothesis that $\boldsymbol{u}$ is orthogonal to $\boldsymbol{v}$ in the *whitened space*. Since the subspace spanned by $\boldsymbol{R}^{1/2}\boldsymbol{V}_\perp$ is the orthogonal complement of the one spanned by $\boldsymbol{R}^{-1/2}\boldsymbol{v}$, it follows that $\boldsymbol{u} = \boldsymbol{R}\boldsymbol{V}_\perp\boldsymbol{b}$ where $\boldsymbol{b} \in \mathbb{C}^{M-1}$. This case is going to differ from the previous one since, as will be shown shortly: the likelihood function under H_0 will include terms of the form $\mathrm{etr}\{-\boldsymbol{R}^{-1}\boldsymbol{A} - \boldsymbol{R}\boldsymbol{B}\}$, i.e., a combination of terms that appear in both Wishart and inverse Wishart distributions, thus precluding the implementation of a fully Bayesian approach based on the whole data $\boldsymbol{x}$, $\boldsymbol{X}$.

Similarly to the previous section, we again assume that $\boldsymbol{b}$ follows a Bernoulli-Gaussian distribution. Under H_0, the likelihood of $\boldsymbol{x}$ is thus given by

$$\begin{aligned} &p_0(\boldsymbol{x},\boldsymbol{X}|\boldsymbol{b},\boldsymbol{R}) \propto \det(\boldsymbol{R})^{-(T+1)} \\ &\quad\times \exp\left\{-\left(\boldsymbol{x} - \boldsymbol{R}\boldsymbol{V}_\perp\boldsymbol{b}\right)^H\boldsymbol{R}^{-1}\left(\boldsymbol{x} - \boldsymbol{R}\boldsymbol{V}_\perp\boldsymbol{b}\right)\right\}\mathrm{etr}\left\{-\boldsymbol{R}^{-1}\boldsymbol{S}\right\} \end{aligned}$$

with the prior distributions still given by (2) and (4). The posterior distribution of $\boldsymbol{b}$, $\boldsymbol{R}$ is hence

$$\begin{aligned} &p_0(\boldsymbol{b},\boldsymbol{R}|\boldsymbol{x},\boldsymbol{X}) \propto \det(\boldsymbol{R})^{-(\nu+M+T+1)}\mathrm{etr}\left\{-\boldsymbol{R}^{-1}\tilde{\boldsymbol{S}}\right\} \\ &\quad\times\left[(1-p)\exp\left\{-\boldsymbol{x}^H\boldsymbol{R}^{-1}\boldsymbol{x}\right\}\delta(\|\boldsymbol{b}\|)\right. \\ &\quad+ p\pi^{-(M-1)}\sigma_\beta^{-2(M-1)} \\ &\quad\left.\times\exp\left\{-\left(\boldsymbol{x} - \boldsymbol{R}\boldsymbol{V}_\perp\boldsymbol{b}\right)^H\boldsymbol{R}^{-1}\left(\boldsymbol{x} - \boldsymbol{R}\boldsymbol{V}_\perp\boldsymbol{b}\right) - \sigma_\beta^{-2}\boldsymbol{b}^H\boldsymbol{b}\right\}\right] \end{aligned} \quad (41)$$

with $\tilde{\boldsymbol{S}}$ given by (19). In this case, we have

$$\begin{aligned} T_0 &= \left(\boldsymbol{x} - \boldsymbol{R}\boldsymbol{V}_\perp\boldsymbol{b}\right)^H\boldsymbol{R}^{-1}\left(\boldsymbol{x} - \boldsymbol{R}\boldsymbol{V}_\perp\boldsymbol{b}\right) + \sigma_\beta^{-2}\boldsymbol{b}^H\boldsymbol{b} \\ &= \boldsymbol{x}^H\boldsymbol{R}^{-1}\boldsymbol{x} - \boldsymbol{b}^H\boldsymbol{V}_\perp^H\boldsymbol{x} - \boldsymbol{x}^H\boldsymbol{V}_\perp\boldsymbol{b} \\ &\quad+ \boldsymbol{b}^H\boldsymbol{V}_\perp^H\boldsymbol{R}\boldsymbol{V}_\perp\boldsymbol{b} + \sigma_\beta^{-2}\boldsymbol{b}^H\boldsymbol{b} \\ &= \boldsymbol{x}^H\boldsymbol{R}^{-1}\boldsymbol{x} - \boldsymbol{x}^H\boldsymbol{V}_\perp\left(\boldsymbol{V}_\perp^H\boldsymbol{R}\boldsymbol{V}_\perp + \sigma_\beta^{-2}\boldsymbol{I}_{M-1}\right)^{-1}\boldsymbol{V}_\perp^H\boldsymbol{x} \\ &\quad+ \left[\boldsymbol{b} - \boldsymbol{\mu}_b\left(\boldsymbol{R},\boldsymbol{x}\right)\right]^H\boldsymbol{\Sigma}_b^{-1}\left(\boldsymbol{R}\right)\left[\boldsymbol{b} - \boldsymbol{\mu}_b\left(\boldsymbol{R},\boldsymbol{x}\right)\right] \end{aligned} \quad (42)$$

with

$$\boldsymbol{\Sigma}_b\left(\boldsymbol{R}\right) = \left[\boldsymbol{V}_\perp^H\boldsymbol{R}\boldsymbol{V}_\perp + \sigma_\beta^{-2}\boldsymbol{I}_{M-1}\right]^{-1} \quad \text{(43a)}$$

$$\boldsymbol{\mu}_b\left(\boldsymbol{R},\boldsymbol{x}\right) = \boldsymbol{\Sigma}_b\left(\boldsymbol{R}\right)\boldsymbol{V}_\perp^H\boldsymbol{x}. \quad \text{(43b)}$$

The conditional posterior distribution of $\boldsymbol{b}$ is thus

$$\begin{aligned} p_0(\boldsymbol{b}|\boldsymbol{R},\boldsymbol{x},\boldsymbol{X}) &\propto (1-p)\delta(\|\boldsymbol{b}\|) \\ &\quad+ p\sigma_\beta^{-2(M-1)}\det(\boldsymbol{\Sigma}_b\left(\boldsymbol{R}\right)) \\ &\quad\times\exp\left\{\boldsymbol{\mu}_b^H\left(\boldsymbol{R},\boldsymbol{x}\right)\boldsymbol{\Sigma}_b^{-1}\left(\boldsymbol{R}\right)\boldsymbol{\mu}_b\left(\boldsymbol{R},\boldsymbol{x}\right)\right\} \\ &\quad\times\mathcal{CN}\left(\boldsymbol{\mu}_b\left(\boldsymbol{R},\boldsymbol{x}\right),\boldsymbol{\Sigma}_b\left(\boldsymbol{R}\right)\right) \end{aligned} \quad (44)$$

which is recognized as a Bernoulli-Gaussian distribution, i.e.,

$$\boldsymbol{b}|\boldsymbol{R},\boldsymbol{x},\boldsymbol{X} \sim \mathrm{BG}\left(p_1,\boldsymbol{\mu}_b\left(\boldsymbol{R},\boldsymbol{x}\right),\boldsymbol{\Sigma}_b\left(\boldsymbol{R}\right)\right) \quad (45)$$

with now

$$p_1 = \frac{p\sigma_\beta^{-2(M-1)}\det(\boldsymbol{\Sigma}_b\left(\boldsymbol{R}\right))\mathrm{e}^{\boldsymbol{\mu}_b^H\left(\boldsymbol{R},\boldsymbol{x}\right)\boldsymbol{\Sigma}_b^{-1}\left(\boldsymbol{R}\right)\boldsymbol{\mu}_b\left(\boldsymbol{R},\boldsymbol{x}\right)}}{(1-p)+p\sigma_\beta^{-2(M-1)}\det(\boldsymbol{\Sigma}_b\left(\boldsymbol{R}\right))\mathrm{e}^{\boldsymbol{\mu}_b^H\left(\boldsymbol{R},\boldsymbol{x}\right)\boldsymbol{\Sigma}_b^{-1}\left(\boldsymbol{R}\right)\boldsymbol{\mu}_b\left(\boldsymbol{R},\boldsymbol{x}\right)}}. \quad (46)$$

$$p_1 = \frac{p\sigma_\beta^{-2(M-1)}\det(\boldsymbol{\Sigma}_b(\boldsymbol{R}^{-1}))\mathrm{e}^{\boldsymbol{\mu}_b^H\left(\boldsymbol{R}^{-1},\boldsymbol{x}\right)\boldsymbol{\Sigma}_b^{-1}\left(\boldsymbol{R}^{-1}\right)\boldsymbol{\mu}_b\left(\boldsymbol{R}^{-1},\boldsymbol{x}\right)}}{(1-p) + p\sigma_\beta^{-2(M-1)}\det(\boldsymbol{\Sigma}_b(\boldsymbol{R}^{-1}))\mathrm{e}^{\boldsymbol{\mu}_b^H\left(\boldsymbol{R}^{-1},\boldsymbol{x}\right)\boldsymbol{\Sigma}_b^{-1}\left(\boldsymbol{R}^{-1}\right)\boldsymbol{\mu}_b\left(\boldsymbol{R}^{-1},\boldsymbol{x}\right)}} \quad (37)$$

As for the conditional posterior distribution of $\boldsymbol{R}$, one has

$$p_0\left(\boldsymbol{R}|\boldsymbol{b},\boldsymbol{x},\boldsymbol{X}\right) \propto \det \boldsymbol{R}^{-(\nu+M+T+1)} \mathrm{etr}\left\{-\boldsymbol{R}\boldsymbol{V}_{\perp}\boldsymbol{b}\boldsymbol{b}^H\boldsymbol{V}_{\perp}^H\right\} \times \mathrm{etr}\left\{-\boldsymbol{R}^{-1}\left[(\nu-M)\mu\boldsymbol{I}_M+\boldsymbol{S}+\boldsymbol{x}\boldsymbol{x}^H\right]\right\}. \quad (47)$$

In contrast to the case $\boldsymbol{u} \perp \boldsymbol{v}$, the posterior distribution $p_0(\boldsymbol{R}|\boldsymbol{b},\boldsymbol{x},\boldsymbol{X})$ does no longer belong to a known family (due to the terms in $\boldsymbol{R}$ and $\boldsymbol{R}^{-1}$ in the trace) and therefore it does not seem feasible to draw samples from it. Accordingly, marginalization with respect to $\boldsymbol{R}$ is not feasible.

In order to circumvent this problem, we again turn to a 2-step approach *à la AMF*, i.e., i) assume $\boldsymbol{R}$ is known and devise a test from $\boldsymbol{x}$ only and ii) replace $\boldsymbol{R}$ by its MMSE estimate from the training samples. The latter is still given by (33).

Let us now turn to the first step which consists of deriving a test based on $\boldsymbol{x}$ only, assuming that $\boldsymbol{R}$ is known. Under H_0, $p_0(\boldsymbol{b}|\boldsymbol{R},\boldsymbol{x})$ is given by (44). It follows that the MMSE estimate of $\boldsymbol{b}|\boldsymbol{R}$, which we denote as $\hat{\boldsymbol{b}}_{|\boldsymbol{R}}$, is simply

$$\hat{\boldsymbol{b}}_{|\boldsymbol{R}} = p_1\boldsymbol{\mu}_b\left(\boldsymbol{R},\boldsymbol{x}\right). \quad (48)$$

Under H_1, recalling that $\alpha \sim \mathcal{CN}(0,\sigma_\alpha^2)$, the posterior distribution of $\alpha|\boldsymbol{R}$ is still given by (17) and consequently the MMSE of $\alpha|\boldsymbol{R}$ is simply

$$\hat{\alpha}_{|\boldsymbol{R}} = \frac{\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{x}}{\sigma_\alpha^{-2}+\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}}. \quad (49)$$

For the sake of simplicity, we will set $\sigma_\alpha^{-2} = 0$ in the previous equation, which amounts to a flat prior for α. The decision statistic takes the following form

$$t(\boldsymbol{x},\boldsymbol{X}) = \frac{p_1(\boldsymbol{x}|\hat{\alpha}_{|\hat{\boldsymbol{R}}_{\mathrm{mmse}}},\hat{\boldsymbol{R}}_{\mathrm{mmse}})}{p_0(\boldsymbol{x}|\hat{\boldsymbol{b}}_{|\hat{\boldsymbol{R}}_{\mathrm{mmse}}},\hat{\boldsymbol{R}}_{\mathrm{mmse}})} \equiv \left(\boldsymbol{x}-\hat{\boldsymbol{R}}_{\mathrm{mmse}}\boldsymbol{V}_{\perp}\hat{\boldsymbol{b}}_{|\hat{\boldsymbol{R}}_{\mathrm{mmse}}}\right)^H\hat{\boldsymbol{R}}_{\mathrm{mmse}}^{-1}\left(\boldsymbol{x}-\hat{\boldsymbol{R}}_{\mathrm{mmse}}\boldsymbol{V}_{\perp}\hat{\boldsymbol{b}}_{|\hat{\boldsymbol{R}}_{\mathrm{mmse}}}\right) - \left(\boldsymbol{x}-\hat{\alpha}_{|\hat{\boldsymbol{R}}_{\mathrm{mmse}}}\boldsymbol{v}\right)^H\hat{\boldsymbol{R}}_{\mathrm{mmse}}^{-1}\left(\boldsymbol{x}-\hat{\alpha}_{|\hat{\boldsymbol{R}}_{\mathrm{mmse}}}\boldsymbol{v}\right). \quad (50)$$

We would like to emphasize that this last approach is quite simple, as the MMSE estimate of $\boldsymbol{R}$ (based on $\boldsymbol{X}$ only) and the MMSE estimates of $\boldsymbol{b}$ or α (conditioned on $\hat{\boldsymbol{R}}_{\mathrm{mmse}}$) are given in *closed-form*. This detector will be referred to in the following as two-step detector with rejection in the orthogonal whitened space (2S-OWS).

IV. Performance Analysis

The performance analysis is conducted by Monte Carlo simulation. We resort to $100/P_{fa}$ independent trials to evaluate the thresholds necessary to ensure a preassigned value of P_{fa} and to 10^3 independent trials to compute the probabilities to decide for H_1 when a useful signal is present (P_d).

We consider a scenario with $M_a = 1$ and $M = M_t = 8$. We generate $\boldsymbol{n}$, $\boldsymbol{n}_t$ as independent random vectors ruled by a complex Gaussian distribution with zero mean and (deterministic) covariance matrix $\boldsymbol{R} = \boldsymbol{C}+\sigma_w^2\boldsymbol{I}_M$, where the clutter covariance matrix $\boldsymbol{C}$ is Gaussian-shaped, namely it is Toeplitz with first row

TABLE II
Main Parameters Used in the Implementation of the Detectors

Parameter	setting
σ_β^2	11
diagonal loading level (DL)	6 dB
μ	$\sigma_w^2 10^{DL/10}$
degrees of freedom ν	$M+T+1$
σ_α^{-2}	0
Number of burn-in iterations (N_{bi})	10
Number of iterations (N_r)	100
Number of moves in the Metropolis-Hastings	5

$\left[1 \; \mathrm{e}^{-\frac{(2\pi\sigma_f)^2}{2}} \; \cdots \; \mathrm{e}^{-\frac{(2\pi\sigma_f(M-1))^2}{2}}\right]$ and $\sigma_f = 0.01$; the power of the thermal noise is $\sigma_w^2 = 0.1$, so that the clutter-to-noise ratio (CNR) is 10 dB.

The steering vector of the target assumes a normalized Doppler frequency $f_\nu = 0.08$, a velocity such that the target competes with noise. The amplitude of the target α is generated as a complex normal random variable; the signal to noise ratio is defined as

$$\mathrm{SNR} = E[|\alpha|^2]\,\boldsymbol{v}^H\boldsymbol{R}^{-1}\boldsymbol{v}$$

and simulation results are reported as a function of the SNR.

The thresholds of the detectors are set to guarantee a chosen P_{fa} by considering a noise-only hypothesis ($\boldsymbol{x} = \boldsymbol{n}$), since we are interested in assessing the performance of the receivers in terms of detection power (for matched steering vector) as well as their selectivity (for mismatched steering vector). The values of the main parameters used to implement the detectors are reported in the Table II.

For comparison purposes we consider some natural competitors for the problem at hand. Obvious references are the Kelly's detector [1] and the W-ABORT [8]. Furthermore, we consider also a generalization of the W-ABORT, hereafter referred to as TW-ABORT, which exhibits improved detection power vs selectivity trade-off, provided its tuning parameter p' is chosen in the appropriate range (see [36]).

The case $T = 2M = 16$ is shown in Fig. 1 for two representative values of the parameter p in its lower and higher region (so as to assess the performance in the whole range omitting redundant intermediate values). As mentioned, for the sake of comparison we report also the Kelly's detector, the W-ABORT, and the TW-ABORT (with $p' = 10^{-4}$).

From Fig. 1 it is apparent that the Bayesian detectors are generally very powerful with superior performance especially at low to intermediate SNR; the 2S-OWS is always better than all other competitors, including the Kelly's detector even at high SNR; the M-1S-OIS conversely shows a performance degradation, more severe for higher p. The comparison confirms the smaller loss of TW-ABORT with respect to the Kelly's detector compared to the W-ABORT.

Lower values of T are critical since for $T < M$ the matrix $\boldsymbol{S}$ becomes singular hence all detectors involving such a matrix cannot be computed; in fact, as T approaches M the performances of the competitors severely degrade. Fig. 2 reports the P_d at the onset of the critical region, i.e., for $T = M+1 = 9$ and

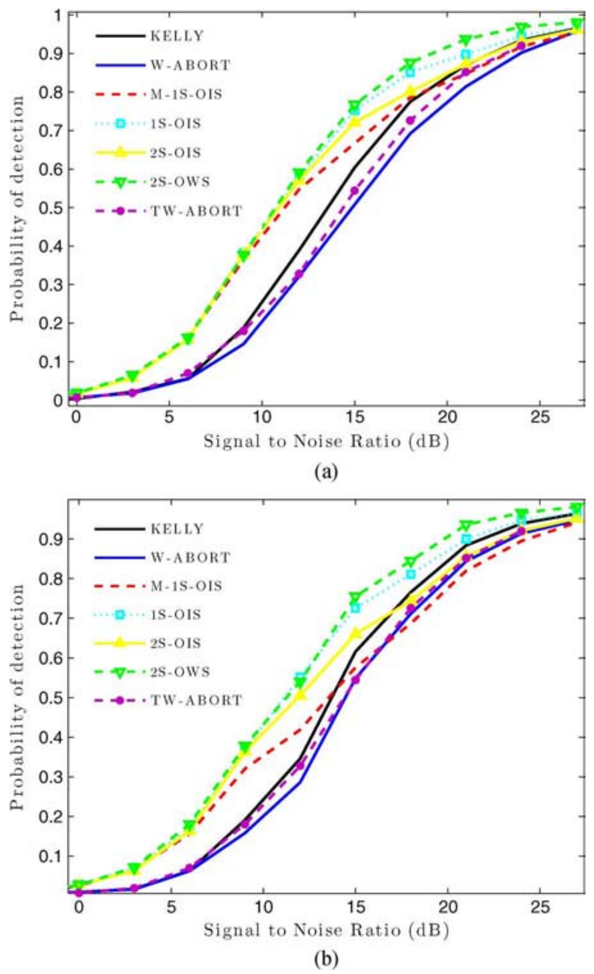

Fig. 1. Probability of detection for $T = 2M$. (a) $p = 0.1$; (b) $p = 0.9$.

$T = M = 8$: it is apparent that, although the classical detectors are still computable,[1] their performance are not acceptable when the number of secondary data is small. It is therefore quite interesting that the performances of the proposed detectors are conversely satisfactory also when a limited number of training data is available, and show a certain robustness to variations of T.

To assess the selectivity of the proposed detectors, we simulated a target with a mismatched steering vector having normalized Doppler frequency $f_\nu + \delta f_\nu$, where $\delta f_\nu = 0.5/M$ (so the matched case analyzed above is trivially obtained for $\delta f_\nu = 0$). Fig. 3 reports the case $T = 2M$: it is apparent that Bayesian detectors, besides being very powerful, can also be more selective than the Kelly's detector; a clear ranking can be identified, with M-1S-OIS the most selective (among the proposed detectors) and 2S-OWS the least selective, comparable to the Kelly's detector. In general, however, none of the detectors is as selective as the W-ABORT. The TW-ABORT confirms an improved power vs selectivity trade-off compared to W-ABORT and Kelly's detector. The relative performances for other values of the parameters are similar, hence figures are omitted. For $T = M$ and $T = M + 1$ these considerations are confirmed,

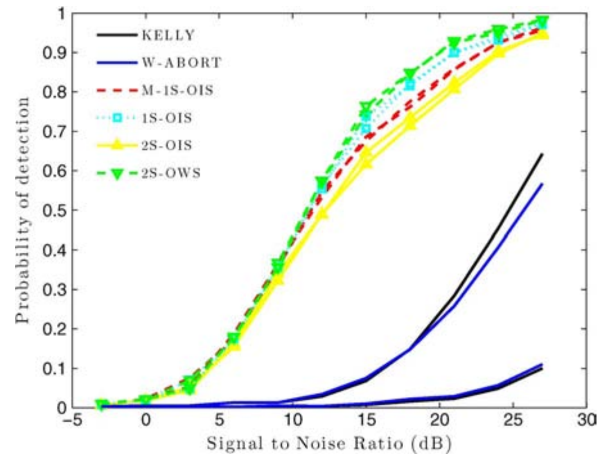

Fig. 2. Probability of detection at the onset of the critical region, for $T = M$ and $T = M + 1$; $p = 0.1$.

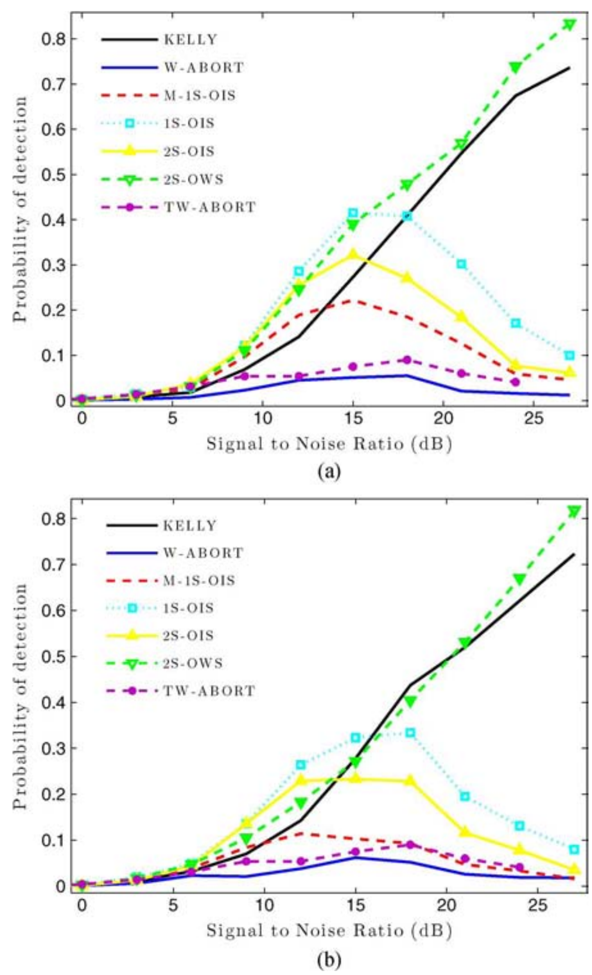

Fig. 3. Probability of detection under mismatched conditions for $T = 2M$ $(\delta f_\nu = 0.5/M)$. (a) $p = 0.1$; (b) $p = 0.9$.

except that the performance of the M-1S-OIS slightly worsens, as shown in Fig. 4 (the curves for $T = M + 1$ are very similar hence omitted).

It is worth noticing that the superior detection power of the 2S-OWS is traded-off with less selectivity compared to the Kelly's detector, while the other proposed detectors are just

[1]Just the Kelly's detector and W-ABORT are reported in order not to burden too much the figure.

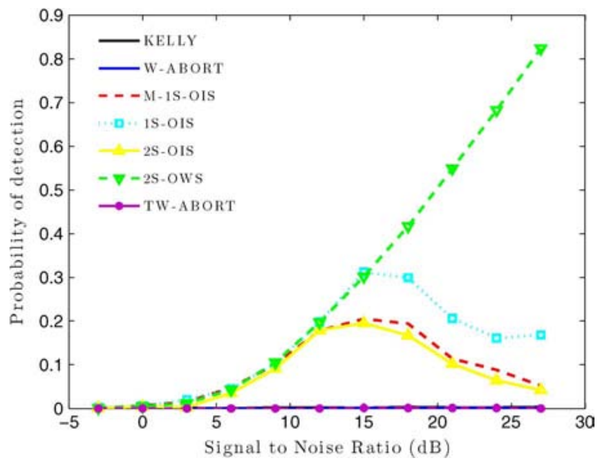

Fig. 4. Probability of detection under mismatched conditions for $T = M$ ($\delta f_\nu = 0.5/M$) and $p = 0.1$.

slightly worse in P_d (compared to the 2S-OWS) but much better in selectivity. As a whole, the analysis suggests that a palette of different trade-offs is possible, with the class of detectors based on the Bayesian approach exhibiting superior performance compared to the classical approaches; this is especially remarkable in case of a limited number of secondary data, when non-Bayesian approaches fail.

We investigated also the sensitivity of the detectors to errors in the parameters assumed for setting the threshold. In particular, we analyzed how the P_{fa} changes when assuming under H_0 a covariance matrix $\boldsymbol{R}$ different from the real one. In fact, the superior performance of the proposed detectors could result, in principle, from a trade-off with the constant false alarm rate (CFAR) property, while the Kelly's detector and the W-ABORT are CFAR. To investigate this point, we set the threshold by using $\sigma_f = \frac{1}{2\pi}\sqrt{-2\log\rho}$, where ρ denotes a chosen one-lag correlation coefficient. Results for $\rho = 0.1$ and $\rho = 0.9$ are shown in Fig. 5, where the value of the true one-lag correlation of the clutter is reported in abscissa. It is apparent that the proposed detectors are quasi-CFAR (with respect to the one-lag correlation coefficient) for $T = 2M = 16$, with fairly limited departure of the curves from the straight line in $P_{fa} = 10^{-3}$ for both sides of the mismatch, i.e., an assumed correlation higher or lower than the true one, respectively. As T reduces, this property becomes weaker for all proposed detectors but for 1S-OIS, which is somewhat more robust to these variations; the case $T = M$ is reported in Fig. 6.

Finally, for the sake of completeness, some results in terms of P_d, selectivity, and P_{fa} sensitivity are reported in Fig. 7 for the case $T = M/2 = 4$, i.e., when the number of training data is very limited. The simulations basically confirm the previous considerations, except for some degradation in the "CFAR property", which as already said is traded-off for P_d whenever conventional approaches are not applicable.

V. Conclusions

In this paper, we have considered the problem of adaptive radar detection of point-like targets in presence of noise with unknown spectral properties. To cope with a limited number of training data, a Bayesian framework has been adopted at the design stage, modeling the noise covariance matrix according to a complex inverse Wishart distribution. In order to come up with detectors with good rejection capabilities, the possible presence of a fictitious signal under the null hypothesis has been modeled probabilistically, as opposite to the conventional ABORT-like approach. We have devised several detectors based on the ratio of conditional distributions under the two hypothesis, where the

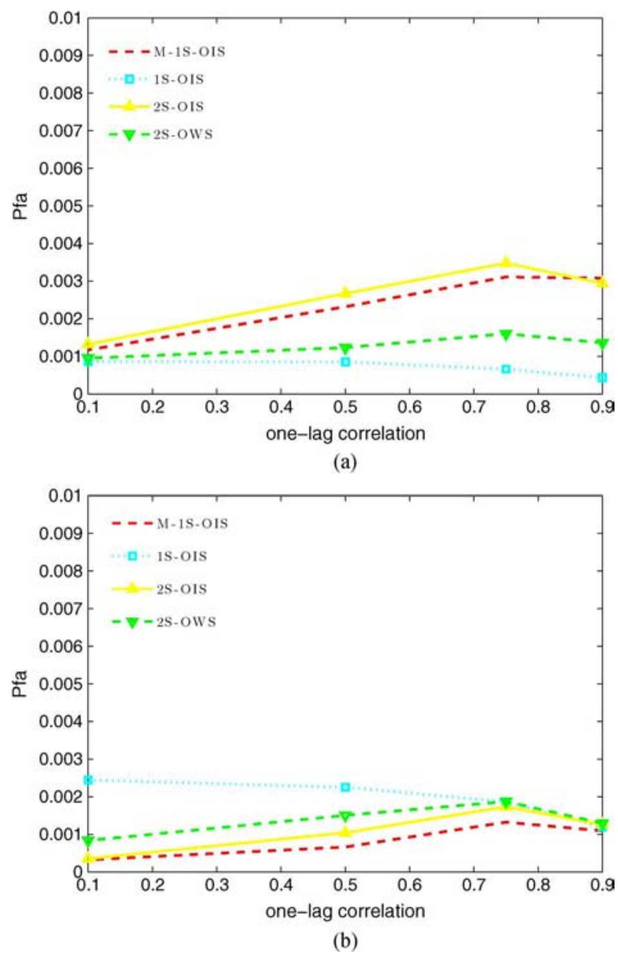

Fig. 5. P_{fa} under mismatched conditions for $T = 2M$ and $p = 0.1$. (a) $\rho = 0.1$; (b) $\rho = 0.9$.

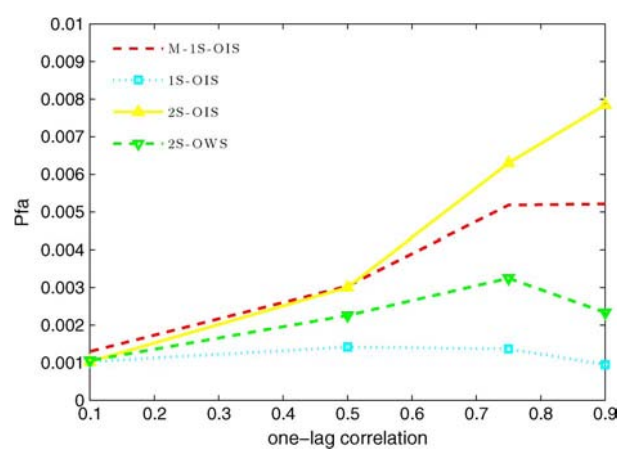

Fig. 6. P_{fa} under mismatched conditions for $T = M$, $p = 0.1$ and $\rho = 0.1$.

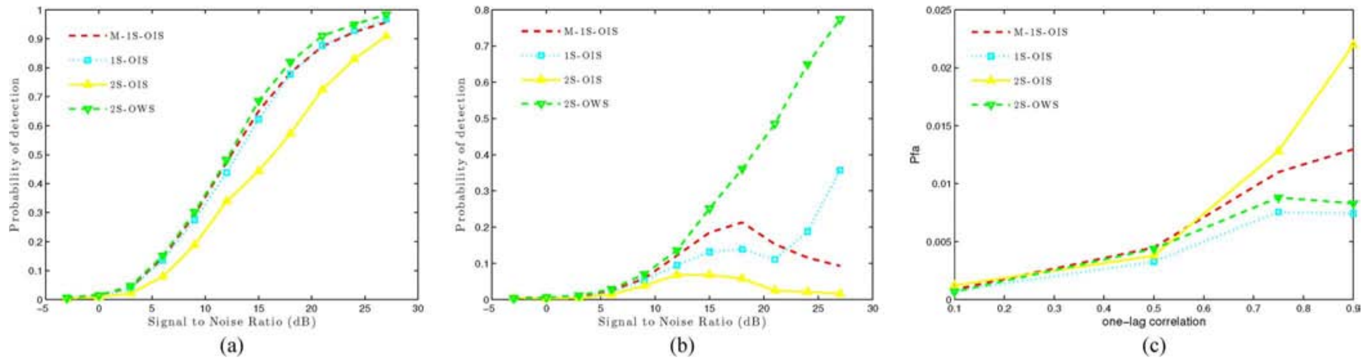

Fig. 7. Probability of detection and sensitivity analysis for $T = M/2$ and $p = 0.1$. (a) $\delta f_\nu = 0$; (b) $\delta f_\nu = 0.5/M$; (c) $\rho = 0.1$.

unknown random parameters are replaced by MMSE estimates based on one-step or two-step procedures. Both the case $\boldsymbol{u}$ orthogonal to $\boldsymbol{v}$ and $\boldsymbol{R}^{-1/2}\boldsymbol{u}$ orthogonal to $\boldsymbol{R}^{-1/2}\boldsymbol{v}$ have been addressed, obtaining four new detectors with different complexity and different behavior in terms of detection power vs selectivity trade-off. Simulation results have shown that satisfactory performance can be obtained even assuming a limited number of training data. Thus, the proposed detectors seem good candidates to conceive augmented schemes that also select training data by discarding possible contaminated returns. The selection processing might be based on the results contained in [39] or specifically designed within the Bayesian framework by following the lead of [40].

References

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

[2] F. C. Robey, D. L. Fuhrman, E. J. Kelly, and R. Nitzberg, “A CFAR adaptive matched filter detector,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 1, pp. 208–216, Jan. 1992.

[3] E. Conte, M. Lops, and G. Ricci, “Asymptotically optimum radar detection in compound gaussian noise,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 2, pp. 617–625, Apr. 1995.

[4] L. L. Scharf and T. McWhorter, “Adaptive matched subspace detectors and adaptive coherence estimators,” in *Proc. 30th Annu. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 3–6, 1996, pp. 1114–1117.

[5] 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.

[6] N. B. Pulsone and C. M. Rader, “Adaptive beamformer orthogonal rejection test,” *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 521–529, Mar. 2001.

[7] G. A. Fabrizio, A. Farina, and M. D. Turley, “Spatial adaptive subspace detection in OTH radar,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1407–1427, Oct. 2003.

[8] 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.

[9] F. Gini and A. Farina, “Matched subspace CFAR detection of hovering helicopters,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 35, no. 4, pp. 1293–1305, Oct. 1999.

[10] K. J. Sangston, F. Gini, M. V. Greco, and A. Farina, “Structures for radar detection in compound gaussian clutter,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 35, no. 2, pp. 445–458, Apr. 1999.

[11] F. Gini and A. Farina, “Vector subspace detection in compound-gaussian clutter. Part I: Survey and new results,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 4, pp. 1295–1311, Oct. 2002.

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

[13] 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.

[14] A. De Maio, “Robust adaptive radar detection in the presence of steering vector mismatches,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 4, pp. 1322–1337, Oct. 2005.

[15] 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. 2006.

[16] 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.

[17] C. D. Richmond, “Performance of the adaptive sidelobe blanker detection algorithm in homogeneous clutter,” *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1235–1247, May 2000.

[18] A. De Maio, “Rao test for adaptive detection in gaussian interference with unknown covariance matrix,” *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3577–3584, Jul. 2007.

[19] M. Greco, F. Gini, and A. Farina, “Radar detection and classification of jamming signals belonging to a cone class,” *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1984–1993, May 2008.

[20] C. D. Richmond, “The theoretical performance of a class of space-time adaptive detection and training strategies for airborne radar,” in *Proc. 32nd Annu. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 1–4, 1998, pp. 1327–1321.

[21] 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.

[22] 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.

[23] F. Bandiera, D. Orlando, and G. Ricci, “Advanced radar detection schemes under mismatched signal models,” in *Synthesis Lectures on Signal Processing*. San Mateo, CA, USA: Morgan & Claypool, March 2009.

[24] 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.

[25] F. Bandiera, D. Orlando, and G. Ricci, “One- and two-stage tunable receivers*,” *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3264–3273, Aug. 2000.

[26] R. Klemm, *Principles of Space-Time Adaptive Processing*, 3rd ed. London, U.K.: Institution of Engineering and Technology, 2002, IEE Radar, Sonar, Navigation and Avionics Series 12.

[27] W. L. Melvin, “Space-time adaptive radar performance in heterogeneous clutter,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 621–633, Apr. 2000.

[28] A. De Maio and A. Farina, “Adaptive radar detection: A Bayesian approach,” in *Proc. Int. Radar Symp. (IRS)*, Krakow, Poland, May 24–26, 2006, pp. 1–4.

[29] A. De Maio, A. Farina, and G. Foglia, “Adaptive radar detection: A Bayesian approach,” in *Proc. IEEE Radar Conf.*, Boston, MA, USA, Apr. 17–20, 2007, pp. 624–629.

[30] O. Besson, J.-Y. Tourneret, and S. Bidon, “Knowledge-aided Bayesian detection in heterogeneous environments,” *IEEE Signal Process. Lett.*, vol. 14, no. 5, pp. 355–358, May 2007.

[31] S. Bidon, O. Besson, and J.-Y. Tourneret, "A Bayesian approach to adaptive detection in non-homogeneous environments," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 205–217, Jan. 2008.
[32] O. Besson, S. Bidon, and J.-Y. Tourneret, "Covariance matrix estimation with heterogeneous samples," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 909–920, Mar. 2008.
[33] F. Bandiera, O. Besson, and G. Ricci, "Knowledge-aided covariance matrix estimation and adaptive detection in compound-gaussian noise," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5390–5396, Oct. 2010.
[34] F. Bandiera, O. Besson, and G. Ricci, "Adaptive detection of distributed targets in compound-gaussian noise without secondary data: A Bayesian approach," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5698–5708, Dec. 2011.
[35] C. Hao, X. Shang, F. Bandiera, and L. Cai, "Bayesian radar detection with orthogonal rejection," *IEICE Trans. Fundamentals Electron., Commun., Comput. Sci.*, vol. E95-A, no. 2, pp. 596–599, Feb. 2012.
[36] A. Coluccia and G. Ricci, "A tunable W-ABORT-like detector with improved detection vs. rejection capabilities trade-off," *IEEE Signal Process. Lett.*, vol. 22, no. 6, pp. 713–717, Jun. 2015.
[37] Y. I. Abramovich, N. K. Spencer, and A. Y. Gorokhov, "Modified GLRT and AMF framework for adaptive detectors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 3, pp. 1017–1051, Jul. 2007.
[38] C. P. Robert and G. Casella, *Monte Carlo statistical methods*, 2nd ed. Berlin, Germany: Springer-Verlag, 2004.
[39] A. Aubry, A. De Maio, L. Pallotta, and A. Farina, "Covariance matrix estimation via geometric barycenters and its application to radar training data selection," *IET Radar, Sonar, Navig.*, vol. 7, no. 6, pp. 600–614, 2013.
[40] O. Besson and S. Bidon, "Adaptive processing with signal contaminated training samples," *IEEE Trans. Signal Process.*, vol. 61, no. 17, pp. 4318–4329, Sep. 2013.

Francesco Bandiera (M'01) was born in Maglie (Lecce), Italy, on March 9, 1974. He received the Dr. Eng. degree in computer engineering and the Ph.D. degree in information engineering from the University of Salento (formerly University of Lecce), Lecce, in 2001 and 2005, respectively.

From July 2001 to February 2002, he was with the University of Sannio, Benevento, Italy, where he was engaged in a research project on mobile cellular telephony. Since December 2004, he has been an Assistant Professor with the Dipartimento di Ingegneria dellInnovazione, University of Salento, where he is engaged in teaching and research. He has held visiting positions with the Department of Electrical, Computer, and Energy Engineering, University of Colorado, Boulder, CO, USA (September 2003–March 2004), and with the Department of Avionics and Systems, Institut Suprieur de lAronautique et de lEspace (ISAE, formerly ENSICA), Toulouse, France (October 2006). His main research interests are in the field of statistical signal processing with focus on radar signal processing, multiuser communications, and pollution detection on the sea surface based upon SAR imagery, and, more recently, localization algorithms.

Olivier Besson (S'90–M'92–SM'04) is a Professor with the Department of Electronics Optronics and Signal of ISAE-Supaéro, Toulouse, France. His research interests are in statistical signal and array processing, estimation, detection, mainly for radar applications.

Angelo Coluccia (M'13) received the Eng. degree in Telecommunication Engineering (summa cum laude) in 2007 and the PhD degree in Information Engineering in 2011, both from the University of Salento, Lecce, Italy.

Former researcher at Forschungszentrum Telekommunikation Wien, Vienna, since 2008, he has been engaged in research projects on traffic analysis, security and anomaly detection in operational cellular networks. He is currently Assistant Professor at the Dipartimento di Ingegneria dell'Innovazione, University of Salento, where he teaches the course of Telecommunication Systems. His research interests are signal processing, communications and wireless networks, in particular cooperative sensing/estimation approaches for localization and other (possibly distributed) applications.

Giuseppe Ricci (M'01–SM'10) was born in Naples, Italy, on February 15, 1964. He received the Dr. degree and the Ph.D. degree, both in Electronic Engineering, from the University of Naples "Federico II" in 1990 and 1994, respectively.

Since 1995 he has been with the University of Salento (formerly University of Lecce) first as an Assistant Professor of Telecommunications and, since 2002, as a Professor. His research interests are in the field of statistical signal processing with emphasis on radar processing, localization algorithms, and CDMA systems. More precisely, he has focused on high-resolution radar clutter modeling, detection of radar signals in Gaussian and non-Gaussian disturbance, oil spill detection from SAR data, track-before-detect algorithms fed by space-time radar data, localization in wireless sensor networks, multiuser detection in overlay CDMA systems, and blind multiuser detection. He has held visiting positions at the University of Colorado at Boulder (CO, USA) in 1997–1998 and in April/May 2001, at the Colorado State University (Fort Collins, CO, USA) in July/September 2003, March 2005, September 2009, and March 2011, at Ensica (Toulouse, France) in March 2006, and at the University of Connecticut (Storrs, CT, USA) in September 2008.