

A Robust Matched Detector

Jean Jacques Fuchs

Abstract—We address the matched detector problem in the case the signal to be detected is imperfectly known. While in the standard detector the signal is known to lie along a particular direction, we consider the case where this direction is known up to additive white Gaussian noise. This somehow amounts to assuming that the signal lies in a cone the aperture of which depends upon the level of uncertainty. We build the associated generalized likelihood ratio (GLR), analyze its statistical properties, indicate how to set the threshold to achieve a given false alarm rate, and how to predict the associated probability of detection. The so-obtained detector reduces to the conventional one when the uncertainty vanishes and we analyze its behavior when the level of uncertainty, which has to be known a priori, is mis-evaluated. It appears that the sensitivity of the detector is quite low with respect to this kind of errors. More importantly several realistic examples are presented that indicate that the proposed detector remains quite efficient when the true signals are far from being of the assumed model and whatever the model of the uncertainty actually is. It is this robustness that makes the detector valuable.

Index Terms—Detection, generalized likelihood ratio test (GLRT), hypothesis testing, matched filter, robustness, total least squares.

I. PROBLEM STATEMENT

IN MANY applications it is desired to detect the presence of a signal s whose observation y is perturbed by interferences i and noise e

$$y = s + i + e, \quad s = \mu x, \quad i = A\varphi.$$

The signal $s \in R^K$ is the product of the magnitude $\mu \in R$ and a signature $x \in R^K$. The interferences $i \in R^K$ lie in a known m -dimensional subspace of R^K that is spanned by the column vectors of the matrix A and $e \in R^K$ represents the additive broadband noise. Observing y , the problem is to decide if μ is equal to zero or different from zero. This is one of the problems considered in [1] that is known as the matched detector or matched filter detector. If A is perfectly known, the difficulties induced by the presence of the interferences are marginal [2] and we will ignore them in the sequel to simplify the exposition.

We extend this model to the case where x is not exactly known. Indeed exact knowledge of the signature x may seem unrealistic since many factors such as bad calibration, distorted antenna shape, jitter, local scattering, source spreading, or errors in the localization to cite a few, give rise to uncertainties

about x . A difference between the assumed and the actual signatures leads to performance loss unless this possibility has been taken into account in the design of the test.

According to the scenario one considers, there are different solutions. One can, for instance, model the uncertainty to circumvent its effect, identify the signature prior to the detection, somehow combine estimation and detection in a unique scheme, or develop robust approaches, i.e., tests that are somehow immune to uncertainties without requiring to model them precisely.

Robust adaptive beamforming somehow belongs to this last set of solutions. It has attracted a lot of attention recently [3]–[5]. The idea is to design a beamformer whose performance remains acceptable when the actual signature departs from the assumed one. Diagonal loading is efficient for small discrepancies [6]; worst-case approaches that minimize the worst performance for all signatures in a given domain have also been proposed [2], [7], [8]. Less prior information is needed and more flexibility on the uncertainty is offered in the Bayesian approaches where the unknown parameters in x are assumed to be random variables with known probability density functions [5]. This type of approaches was also used, for instance, in [9] in a slightly different context.

Several recent papers consider the case where the signature lies in a known p -dimensional subspace of R^K [10]–[12]. In the single observation case, which is considered here, the optimal detector is usually called a subspace detector and is outlined, for instance, in [1], [13], [14]. It amounts to projecting the observation on the whole subspace rather than just on the one-dimensional signature. In the multiple observations case, when several y 's associated with the same unknown signature x are available, it is possible to estimate the optimal direction in the subspace, and the associated detector is called a matched-direction detector [10]–[12], [15]. A subspace model is a useful way to model uncertainty in numerous situations as, for instance, in radar or sonar signal processing. When the signature is a steering vector and one wants to test for targets on a fixed regular grid [11], [15]–[17], a low dimensional subspace is used to model the steering vector within each cell in the grid. Similarly when the direction of look is not perfectly known, a subspace can be used to model the first few terms in the Taylor series expansion around the expected direction. Another relevant example corresponds to the case of a Ricean channel [18] where there exists some angular spread around the line-of-sight direction. The same idea prevails in the presence of local scattering [19]. Since the performance of these detectors somehow diminishes with p , the dimension of the subspace, it is desired to keep p as small as possible or to reduce it by approximating the subspace by its principal modes [11], [16].

As a matter of fact, it appears that, in general, and certainly so in these last examples, one has a good knowledge about a

Manuscript received December 12, 2006; revised March 27, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Leslie Collins.

The author is with the IRISA/Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France (e-mail: fuchs@irisa.fr).

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

Digital Object Identifier 10.1109/TSP.2007.898786

mean or dominant signature or direction, but there is some uncertainty about it. This idea is not present in the subspace approach where all directions within the subspace become equally important. This idea is at the basis of the model we consider in this contribution. We assume that, besides $y = \mu x + e$, we observe z that satisfies

$$z = x + \alpha n$$

where x is the true unknown signature and αn is a vector of Gaussian noise whose magnitude can be tuned by α , a real number. This model, where the uncertainty upon the signature is modeled by white Gaussian noise, may indeed seem far from any reality. The interest of the detector we will develop assuming this model will therefore depend upon its capacity to handle realistic situations that do not follow this assumed model.

In Section II, we further define the model we use. In Section III, we develop the maximum likelihood estimates that are used to build the generalized likelihood ratio test (GLRT) in Section V. Simulation results are presented in Section VI. In Section IV, we briefly describe the matched subspace detector [1] for easy reference.

II. DATA MODEL

We are given two K -dimensional observation vectors y and z and we assume that they satisfy

$$y = \mu x + e \quad z = x + \alpha n \quad (1)$$

where e and n are independent Gaussian K -dimensional perturbation vectors with covariance matrices $\sigma^2 I_K$; we also denote by e and $n \sim N(0, \sigma^2 I)$. This means that, besides the standard observation y , we have an additional observation z that is a perturbed version of the actual signature x which is unknown. The intensity of the perturbation depends upon the parameter α that has to be tuned by the user, i.e., to be known. Following the concept of uncertainty ratio introduced in [11], we define the discrepancy ratio (DR) associated with model (1) as

$$\text{DR} = \frac{\alpha^2 E(n^T n)}{x^T x} = \frac{K \alpha^2 \sigma^2}{x^T x} \quad (2)$$

that measures the ratio of the average power in the discrepancy to the power in the signature. If one sees $z = x + \alpha n$ as lying in a cone with axis x then DR is roughly the square of the tangent of the half-angle of the aperture of this cone.

The model (1) is difficult to justify in practice and should only be considered as a rough way to model the knowledge one has. It is actually the robustness of the test deduced from (1) with respect to modeling errors that will justify its interest and usefulness.

The assumptions made on e and n correspond to the case where the noise covariances are known up to scale factors. This is the standard assumption unless the covariance structure can be learned from training data that are free from the signal of interest, in which case one talks about adaptive detection [2], [20], [21].

We will further assume that α and σ are known. One may relax the assumption on σ since it is possible to identify this

variance. Roughly speaking one would work with the likelihood ratio instead of the log-likelihood ratio and would end up with Fisher-like distributions instead of chi-square distributions [1], [12], [16]. There is no way to weaken the assumption on α , however. The parameter α is simply not identifiable, i.e., the likelihood function can be made infinite by a proper choice of α when this parameter is assumed unknown [22]. More specifically in (1) there are two degrees of freedom either the two variances of y and z or the set $\{\alpha, \sigma\}$ we introduced and the proposed parametrization is the only one for which one of the two quantities, namely σ , is identifiable. It is actually not too surprising that it is the ratio α that intervenes, i.e., that the amount of uncertainty acceptable on the signature is proportional to the noise level on the observation y .

For fixed or identified σ , it is the choice of α that tunes the DR (2). The impact of a wrong value of α (in case speaking of an exact value makes sense) on the false alarm and detection probabilities will be investigated. *A priori* a large value of DR, or equivalently α , means high uncertainty and additional robustness to false alarm; but if the concern is more on missed detection, it is indeed preferable to underestimate DR. It happens that for reasonable values of DR, say, $\text{DR} \in [0, 1/4]$, a wrong choice of α has negligible effects on the probability of false alarm while there may exist quite a gap between the observed (actual) and the expected probability of detection (see Sections V-D and VI-A-2)).

III. ESTIMATION PROBLEM

A. Maximum Likelihood Estimates

From the model described in (1) and the different assumptions we made, it follows that the quantities to be estimated are μ and x . If we denote by $f_E(e)$ and $f_N(n)$ the probability density functions (pdf) of e and n , which are independent identically distributed Gaussian random vectors, the joint probability density function $f_{Y,Z}(y, z)$ of the observations (y, z) is given by

$$f_{Y,Z}(y, z) = f_E(y - \mu x) f_N\left(\frac{1}{\alpha}(z - x)\right).$$

The opposite of the log-likelihood function of (y, z) , denoted by $\ell(\mu, x)$ is then, up to additive and multiplicative constants, equal to

$$\ell(\mu, x) = \|y - \mu x\|^2 + \frac{1}{\alpha^2} \|z - x\|^2. \quad (3)$$

The maximum likelihood estimates (MLE) are then obtained by minimizing $\ell(\cdot)$ with respect to (w.r.t.) μ and x . This is a nonlinear optimization problem in which one recognizes a total-least-squares (TLS) problem whose solution is known to depend upon a singular value decomposition (SVD) [23]. To put it into the basic TLS setting, we perform a change of variables. Introducing $\rho = \bar{z} - \bar{x}$ and $r = y - \bar{\mu} \bar{x}$ with $\bar{x} = x/\alpha$, $\bar{z} = z/\alpha$ and $\bar{\mu} = \alpha \mu$, and eliminating \bar{x} , the minimization problem becomes equivalent to

$$\min_{\bar{\mu}, \rho, r} \|\rho\|^2 + \|r\|^2 \text{ subject to } y - r = \bar{\mu}(\bar{z} - \rho).$$

We have transformed the initial minimization problem with no constraints and $K + 1$ unknowns into a minimization problem with K constraints and $2K + 1$ unknowns. We further transform this problem. Introducing the two-column matrices $\hat{C} = [\bar{z} \ y]$, $\delta C = [\rho \ r]$, and $C = \hat{C} - \delta C$, the constraint in the above problem then becomes $C[-\bar{\mu} \ 1]^T = 0$, which says that the two columns in C are collinear, i.e., the matrix C is of rank one, hence the new formulation

$$\min_{\delta C} \|\delta C\|_F^2 \text{ subject to } \text{rank}(\hat{C} - \delta C) = 1 \quad (4)$$

where $\|A\|_F^2 = \text{trace}(A^T A)$ denotes the square of the Frobenius norm. This optimization problem amounts to seeking the perturbation matrix δC with the smallest Frobenius norm that, when subtracted from \hat{C} , makes the matrix $C = \hat{C} - \delta C$ rank deficient. The solution of this problem is known [23]. The optimal δC is $\hat{s}_{\min} \hat{u}_{\min} \hat{v}_{\min}^T$, the rank-one matrix associated with the *smallest* singular triplet in the SVD of \hat{C} , and the value of the optimal cost is \hat{s}_{\min}^2 . The MLEs of μ and x follow easily, but since we do not need them, we will not detail them. One can also obtain the statistical properties of these estimates by computing the Fisher information matrix and inverting it [22]. But, since we only require some very specific results that can be obtained by straightforward means, we postpone this analysis until Section V-C where we only develop the parts that are useful for our purpose.

B. Bayesian Point of View

Before we proceed let us show that the model we propose can be given a Bayesian interpretation. More precisely, we show that the cost function (3) can also be attained by using a Bayesian approach with a noninformative prior on the parameter μ and seeking a maximum *a posteriori* estimate. In the Bayesian theory of parameter estimation [13], [24], the unknown parameters are seen as realizations of random variables and are thus endowed with so-called prior distributions. In the unique observation equation $y = \mu x + e$, we thus need to define prior distribution for both x and μ , and we will denote by $f_X(x)$ and $f_\mu(\mu)$ the corresponding pdf's which are assumed to exist. For x there is a choice that corresponds clearly to our previous model (3): x is a realization of a Gaussian random vector with mean z and covariance matrix $\alpha^2 \sigma^2 I$. For μ there is no obvious choice; perhaps the most natural one, since we will develop a test between a simple hypothesis $\mu = 0$ and a composite two-sided alternative $\mu \neq 0$, is to take a Gaussian zero mean prior $N(0, \gamma^2 \sigma^2)$ where γ will allow to tune the confidence one has in the prior information. In the present detection context a small value of γ will bias the estimate of μ towards zero and therefore will bring additional robustness against false alarms but will favor missed detections. We denote by $f_{Y|\mu, X}(y|\mu, x)$ the conditional pdf that generates the data y . The posterior pdf of x and μ conditioned on the observation y is then given by Bayes rule

$$f_{\mu, X|Y}(\mu, x|y) = \frac{f_{Y|\mu, X}(y|\mu, x) f_\mu(\mu) f_X(x)}{\int \int f_{Y|\mu, X}(y|\mu', x') f_\mu(\mu') f_X(x') d\mu' dx'}$$

The maximum *a posteriori* estimates (MAP) of x and μ are then obtained by maximizing this quantity w.r.t. x and μ . Since the denominator is a constant, we seek the maximum of the numerator which is the product of 3 Gaussian densities. The MAP estimates of x and μ are therefore the solution of

$$\min_{x, \mu} \|y - \mu x\|^2 + \frac{1}{\alpha^2} \|z - x\|^2 + \frac{1}{\gamma^2} \mu^2$$

this should be compared to (3). It is possible to show that the global minimum is generically unique. The estimate of μ is the solution of a fifth-order algebraic equation in which it is possible to distinguish the part due to the prior on μ that attracts the solution towards zero. To make the prior distribution on μ noninformative, one lets γ diverge to infinity, and the MAP estimates are then identical to the ML estimates associated with model (1) since the above minimization problem becomes identical to (3).

IV. MATCHED SUBSPACE DETECTOR

For easy reference, let us give a quick summary of what is usually called the matched subspace detector (MSD) [1], [13] where the signature x is known to lie in the range of the (K, p) matrix H with $p < K$. In that case, one observes

$$y = \mu x + e \quad \text{with} \quad x = H\bar{\theta} \quad (5)$$

and wants to decide if $\mu = 0$ or $\mu \neq 0$. The noise e follows $e \sim N(0, \sigma^2 I)$ with σ known. Since there is a scaling ambiguity between μ and $\bar{\theta}$ in (5), we introduce θ and define $\mu x = H\theta$. In order to develop the GLRT, one maximizes the likelihood of the observation vector y , i.e., minimizes w.r.t. θ

$$\ell(\theta) = \|y - H\theta\|^2$$

under both hypotheses. Under $H_0, \theta = 0, y = e$ and the value of the minimum is $\ell(0|H_0) = y^T y$. Under H_1 , the minimum is attained for $\hat{\theta} = H^+ y$ with $H^+ = (H^T H)^{-1} H^T$, the pseudo-inverse of H , and is equal to $\ell(\hat{\theta}|H_1) = y^T (I - H H^+) y$. Combining these results, the scaled logarithmic likelihood ratio, denoted by $L(y)$ becomes

$$\begin{aligned} L(y) &= \frac{1}{\sigma^2} \left\{ \ell(0|H_0) - \ell(\hat{\theta}|H_1) \right\} \\ &= \frac{1}{\sigma^2} y^T H H^+ y = \frac{1}{\sigma^2} y^T P_H y \end{aligned} \quad (6)$$

where P_H denotes the orthogonal projection onto the range space of H . For completeness let us briefly recall the statistical properties of $L(y)$ in (6) under both hypotheses [1].

Under $H_0, y = e$ and $L(y) = (1/\sigma^2) e^T P_H e$ follows a Chi-squared random variable with p degrees of freedom, we denote $L(y|H_0) \sim \chi_p^2$. Under H_1 , its law becomes a noncentral Chi-square with p degrees of freedom and noncentrality parameter $\lambda^2 = (1/\sigma^2) \theta^T H H^T \theta = (1/\sigma^2) \mu^2 x^T x$, i.e., $L(y|H_1) \sim \chi_p^2(\lambda^2)$. When the dimension p of the subspace is equal to 1, the MSD becomes the well known matched detector found in many textbooks [3], [13], [25].

Note that the MSD is *optimal* for the scenario for which it has been built and can thus not be outperformed. By optimal in this context we mean that it is UMPU (i.e., Uniformly the Most Powerful test among those that are Unbiased) [13], [25]. This tells us that, for the same probability of false alarm (P_{FA}), there cannot exist a test that has better probability of detection (P_D) uniformly in θ .

V. DEVELOPMENT OF THE ROBUST MATCHED DETECTOR

A. Generalized Likelihood Ratio Test (GLRT)

We come back to (1). Observing y and z we want to decide between the two hypotheses

$$\begin{aligned} H_0 : \quad & y = e, \quad z = x + \alpha n \\ H_1 : \quad & y = \mu x + e, \quad z = x + \alpha n \end{aligned}$$

i.e., to decide whether $\mu = 0$ or $\mu \neq 0$. We propose to use the generalized likelihood ratio test (GLRT) to take the decision. Since the different noise variances are known, it is convenient to work with twice the logarithmic GLR which we denote by $L(y, z)$. With $\ell(\mu, x)$ defined in (3), one gets

$$L(y, z) = \frac{1}{\sigma^2} \{ \ell(0, \hat{x}|H_0) - \ell(\hat{\mu}, \hat{x}|H_1) \}$$

where $\ell(\hat{\mu}, \hat{x}|H_k)$ is the value of the minimum of $\ell(\cdot)$ under hypothesis k , i.e., when the unknown parameters μ and x are replaced by their ML estimates.

Under H_0 , the value of the minimum of $\ell(0, x)$ with respect to x is $\ell(0, z|H_0) = \|y\|^2$. Under H_1 , we have seen at the end of Section III-A that the value of the minimum is $\ell(\hat{\mu}, \hat{x}|H_1) = \hat{s}_{\min}^2$, denoted by $s_{\min}^2(\hat{C})$, the square of the smallest nonzero singular value of

$$\hat{C} = \begin{bmatrix} z & y \\ \alpha & \end{bmatrix}. \quad (7)$$

This is also the smallest eigenvalue of the order two positive definite matrix $\hat{C}^T \hat{C}$. It follows that the logarithmic GLR, we will use in the test, takes the form

$$L(y, z) = \frac{1}{\sigma^2} \left\{ y^T y - s_{\min}^2(\hat{C}) \right\}. \quad (8)$$

Using (7), it is easy to show that $L(y, z)$ is always positive as it should be. In order to characterize the detection test which will consist in comparing $L(y, z)$ with a threshold η , we need to get the pdf of $L(y, z)$ under both H_0 and H_1 . Since the hypothesis H_0 is simple, the threshold η will be set to yield a given false alarm rate P_{FA} , and the P_D under H_1 is then mainly a function of what we will call the signal-to-noise ratio (SNR) that depends upon the value of the parameter μ .

B. Convergence Towards the Matched Detector

Let us show that, as α decreases to zero, the test value $L(y, z)$ of the robust matched detector (RMD) we propose converges

towards the standard test $L(y)$ in (6), which for $p = 1$ and $H = x$ becomes

$$L(y) = \frac{1}{\sigma^2} \frac{(x^T y)^2}{x^T x}. \quad (9)$$

We observe that as $\alpha \rightarrow 0$, the $K \times 2$ matrix $\hat{C} = [z/\alpha \ y]$ has its largest singular value diverge to infinity while its smallest singular value $s_{\min}(\hat{C})$ converges towards a constant. An easy way to prove this is to introduce the pseudo-inverse \hat{C}^+ of \hat{C} . It is given by

$$\hat{C} = \begin{bmatrix} z/\alpha & y \\ \alpha & \end{bmatrix} \Rightarrow \hat{C}^+ = \begin{bmatrix} \alpha z^+ (I_n - y y^{\perp+}) \\ y^{\perp+} \end{bmatrix}$$

with $y^{\perp} = (I - z z^+) y$

as can be verified. As $\alpha \rightarrow 0$, this $2 \times K$ matrix has its first row that becomes zero while the second row remains constant. This means that one of its singular values becomes zero while the other one, whose value is the inverse of $s_{\min}(\hat{C})$, converges towards the Euclidean norm of this constant row $y^{\perp+}$

$$s_{\min}(\hat{C}) \rightarrow \|y^{\perp+}\|^{-1} = \|y^{\perp}\| = (y^T (I - x x^+) y)^{1/2}$$

where we have further replaced z by x since $z = x$ for $\alpha = 0$. The announced result follows. This is of course not unexpected and tells us that the RMD converges towards the matched detector as the uncertainty on the signature converges to zero.

C. Statistical Analysis and Test Procedure

To evaluate the statistical properties of the GLR $L(y, z)$ in (8) under both H_0 and H_1 , we need to know the pdf of $s_{\min}^2(\hat{C})$. One can decompose \hat{C} in (7) as

$$\hat{C} = C + \Delta = \begin{bmatrix} z & y \\ \alpha & \end{bmatrix} = \begin{bmatrix} x & \mu x \\ \alpha & \end{bmatrix} + [n \ e]$$

where its mean C has rank 1 and all the components in the perturbation Δ are $N(0, \sigma^2)$. The minimal singular value $s_{\min}(\hat{C})$ is thus equal to zero when $\sigma = 0$ and one expects it to be of order σ . A similar remark holds for $s_{\min}^2(\hat{C})$, the smallest eigenvalue of $\hat{C}^T \hat{C}$. In the Appendix we obtain the following expression of the first-order term in its expansion with respect to σ^2

$$s_{\min}^2(\hat{C}) = \frac{1}{1 + \alpha^2 \mu^2} \left(e^T \left(I - \frac{x x^T}{x^T x} \right) e + \alpha^2 \mu^2 n^T \left(I - \frac{x x^T}{x^T x} \right) n \right) + o(\sigma^2). \quad (10)$$

This approximation is valid for small σ and more precisely if $(x^T x)((1/\alpha^2) + \mu^2) \gg K \sigma^2$. Note that this condition is independent of K , the number of observations, since $x^T x = O(K)$ and is more difficult to satisfy under H_0 when $\mu = 0$. It is then linked to the DR (2) and becomes $DR \ll 1$. It is this last condition, valid under both H_0 and H_1 , that we keep for the sequel

$$DR = \frac{K \alpha^2 \sigma^2}{x^T x} \ll 1 \quad \text{and in practice} \quad DR \leq \frac{1}{4}. \quad (11)$$

Replacing $s_{\min}^2(\hat{C})$ by its approximation (10) in $L(y, z)$ (8) one gets

$$L(y, z) = \frac{1}{\sigma^2} \frac{(x^T y)^2}{(x^T x)} + \frac{1}{\sigma^2} \frac{\alpha^2 \mu^2}{1 + \alpha^2 \mu^2} \times \left(e^T \left(I - \frac{xx^T}{x^T x} \right) e - n^T \left(I - \frac{xx^T}{x^T x} \right) n \right) + O(\sigma^2). \quad (12)$$

Under H_1 , the first term of the expansion in (12) is a $\chi_1^2(\lambda^2)$ random variable with noncentrality parameter λ^2 the SNR

$$\text{SNR} = \lambda^2 = \frac{\mu^2(x^T x)}{\sigma^2}. \quad (13)$$

This term is the standard GLR (9), it has mean $1 + \lambda^2$ and variance $2(1 + 2\lambda^2)$. The second term of the expansion in (12), which is correlated with the first one, vanishes with α . It is the weighted difference of two $\chi_{K-1}^2(0)$ random variables and has mean zero and variance $4(K-1)\{(\alpha^2 \mu^2)/(1 + \alpha^2 \mu^2)\}^2$. Provided (11) is satisfied, it follows that the distribution of $L(y, z|H_1)$ admits no simple analytical expression; it is a random variable that is close to a $\chi_1^2(\lambda^2)$ random variable but has a slightly larger variance.

Under H_0 , one has $\mu = 0$, the second and last term of the expansion in (12) vanishes and the next term is therefore needed. Pushing further the analysis in the appendix one gets

$$s_{\min}^2(\hat{C}) = \left(e^T \left(I - \frac{xx^T}{x^T x} \right) e \right) - \frac{\alpha^2}{x^T x} \frac{(x^T e)^2}{x^T x} \times \left(e^T \left(I - \frac{xx^T}{x^T x} \right) e \right) + o(\sigma^4) \quad (14)$$

which yields the following approximation for the GLR under H_0 :

$$L(y, z|H_0) \simeq \frac{1}{\sigma^2} \frac{(x^T e)^2}{x^T x} \left(1 + \frac{\alpha^2 \sigma^2}{x^T x} \frac{1}{\sigma^2} e^T \left(I - \frac{xx^T}{x^T x} \right) e \right) \quad (15)$$

the first term in the approximation is a $\chi_1^2(0)$ random variable. We have factored it and it thus appears that $L(y, z|H_0)$ is approximately a χ_1^2 random variable times another random variable correlated with the first one whose mean, we denote by β , is given as $\beta = 1 + (\alpha^2 \sigma^2 / x^T x)(K-1) \simeq 1 + \text{DR}$. The distribution of $L(y, z|H_0)$ admits thus no simple expression. In the sequel we approximate the tail of the pdf of $L(y, z|H_0)$ by the tail of the pdf of a $\chi_1^2(0)$ random variable multiplied by β , whose analytical expression is not needed to evaluate the threshold, see (16).

It follows that the approximate GLR test we propose decides that H_0 is true when $L(y, z) \leq \eta$ and that H_1 holds otherwise. The threshold η is fixed to achieve a given false alarm probability (P_{FA}) according to

$$P_{\text{FA}} = Pr \left\{ \chi_1^2(0) \geq \frac{\eta}{\beta} \right\} \quad \text{with} \quad \beta = 1 + \frac{\alpha^2 \sigma^2}{x^T x} (K-1). \quad (16)$$

The notation $Pr\{\chi_1^2(\lambda^2) \geq \eta\}$ denotes the probability that a $\chi_1^2(\lambda^2)$ random variable is greater than η .

Note that relation (16) is only a way to fix η and should not be seen as defining the P_{FA} but we will see below in Section VI-A-1) that fixing η in this way is indeed quite valid both as far as the observed P_{FA} 's and the P_D 's are concerned provided (11) is satisfied. We want to highlight at this point of the analysis that, while it is asymptotic and relies on first or second-order expansion, one can check on the data if one is in the asymptotic regime. Indeed condition (11) can be checked on the data since one may replace the unknown x by the observed z , the so-introduced error being negligible.

D. Modeling Error

As already indicated it is the value of α that allows to tune the DR (2) and we now analyze the properties of the test we propose when the value one assumes for α is in error. Since in practice model (1) is quite often an easy way to represent a far more complex reality, DR and α have quite often no real meaning and some sort of robustness of the test w.r.t. them is crucial.

Of course to be able to perform the analysis, we consider (1) to represent the reality and only assume the value of α to be wrong. To fix the notations, we consider that one assumes wrongly that $z = x + \alpha n$ while indeed $z = x + \gamma n$ with γ different from α .

Before we proceed let us note that, under H_1 , one has to distinguish between the *observed* or actual probability of detection (P_D), the *optimal* P_D and the *expected* P_D . The *observed* P_D is the one we estimate through simulations, the *optimal* P_D is the one one would obtain if the test had been designed for the true model and the *expected* P_D is the one one predicts believing the model used is the true one. Under H_0 , the situation is easier; there is just the observed P_{FA} , the expected and optimal P_{FA} being identical and equal to the one chosen by the user when setting the threshold, provided the approximations we made are valid. The aim of the analysis that follows is to evaluate the observed probabilities since we already know how to compute the expected and optimal ones.

The test is of course always built around the GLR (8) where $s_{\min}(\hat{C})$ depends upon both α and γ . To evaluate its properties in the presence of these modeling errors, one goes through the development of the Appendix and finds that one has to replace n by $(\gamma/\alpha)n$ in (10). The sequel of the analysis is similar to the one developed in Section V-C using the approximation (12) or (15). We distinguish now the two possible errors.

1) *Overestimating the Uncertainty*: Let us start with the specific case where $\gamma = 0$ and $\alpha > 0$, i.e., where one assumes wrongly that the signature x is imperfectly known. One applies to the observations y and z the test built around $L(y, z)$ in (8) with $\alpha > 0$ while one should apply this same test with $\alpha = \gamma = 0$ or equivalently with $n = 0$.

Note that the *optimal* statistic in this case is $L(y)$ in (9) which is identical to $L(y, z)$ when $\alpha = 0$.

Under H_0 , since (15) does not depend upon n , it is a valid approximation and the observed P_{FA} should be identical to the expected and optimal one. Indeed, we will set the threshold η to β times the optimal, one but since the same holds for the statistic $L(y, z)$, this introduces no error.

Under H_1 , it follows from (12) with $n = 0$ that

$$L(y, z|H_1) = L(y|H_1) + \frac{1}{\sigma^2} \frac{\alpha^2 \mu^2}{1 + \alpha^2 \mu^2} e^T \left(I - \frac{xx^T}{x^T x} \right) e + O(\sigma^2).$$

Neglecting the higher order term, it appears that $L(y, z|H_1)$ is larger than the optimal statistic $L(y|H_1)$ by the positive quantity

$$\frac{1}{\sigma^2} \frac{\alpha^2 \mu^2}{1 + \alpha^2 \mu^2} e^T \left(I - \frac{xx^T}{x^T x} \right) e \quad (17)$$

since a similar error holds for the threshold. This is a favorable situation but it is impossible to draw precise conclusions. The simulations proposed in Section VI-A-2) (see Fig. 2) show that the loss in performance induced by this modeling error is reasonably small. Of course, since one over-estimates the uncertainty, the expectations are pessimistic and the expected probabilities of detection are smaller than both the optimal and observed ones.

2) *Underestimating the Uncertainty:* We consider now the case where α is smaller than γ . Under H_0 , the statistic $L(y, z)$ that is used and the way the threshold η is fixed are again quite coherent and the observed and expected P_{FA} should be identical. Note, however, that η is smaller than what it should be. Under H_1 , one has from (12)

$$L(y, z|H_1) = \frac{1}{\sigma^2} \frac{(x^T y)^2}{(x^T x)} + \frac{1}{\sigma^2} \frac{\alpha^2 \mu^2}{1 + \alpha^2 \mu^2} \times \left(e^T \left(I - \frac{xx^T}{x^T x} \right) e - \frac{\gamma^2}{\alpha^2} n^T \left(I - \frac{xx^T}{x^T x} \right) n \right) + O(\sigma^2) \quad (18)$$

which tells us that the statistic $L(y, z|H_1)$ is also smaller than what it should be. Again this indicates that the observed P_D should be close to the optimal one.

The simulations proposed in Section VI-A-2) (see Fig. 3) show that the difference between the observed and the optimal P_D 's is surprisingly small. But, of course, since one under-estimates the uncertainty, the expected probability of detection is larger than both the optimal and observed ones.

VI. SIMULATION RESULTS

There are two parts in this section. In the first we essentially check that the theoretical development made around the robust matched detector (RMD) are valid when applied to data simulated according to the model (1) as soon as (11) is satisfied. In the second we apply the RMD to data simulated according to models that are fully unrelated to (1), evaluate its performances, and compare them with those of other detectors when possible. This allows us to check the *true* robustness of the RMD.

A. Checking the Theoretical Properties

Unless stated otherwise in (1) we take $K = 12$, $\sigma^2 = 1$, $x = \sqrt{2}\mathbf{1}$, with $\mathbf{1}$ a column vector filled with ones, tune μ to attain the announced SNRs (13), and perform 10^5 simulations. Note that with these parameter settings, one has $DR = \alpha^2/2$.

1) *Receiver Operating Characteristics:* The receiver operating characteristics (ROCs) for the RMD are presented in Fig. 1. The threshold η is calculated using relation (16), and the probabilities of detection are compared in Fig. 1 to the

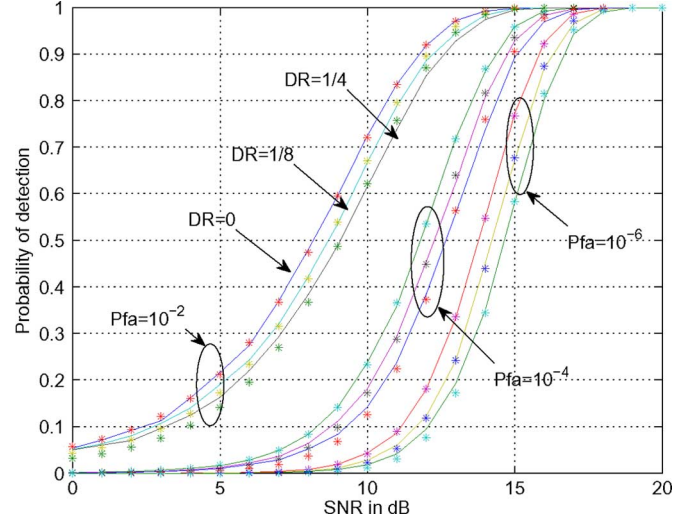


Fig. 1. ROC for proposed method, P_{FA} equal to 10^{-2} , 10^{-4} , and 10^{-6} and $DR = 1/80, 1/8, 1/4$. The solid line are the simulations results and the *'s the theoretical values.

theoretical P_D 's obtained by assuming $L((y, z))$ is a $\chi_1^2(\lambda^2)$ random variable.

For three different P_{FA} 's we represent the P_D as a function of the SNR defined in (13). In each case we further consider three different values of DR namely $DR = 0, 1/8, 1/4$ or equivalently α equal to 0, .5 and .7. There is a good agreement between the observed and theoretical values of P_D .

When $DR = 0$, the signature vector x is perfectly known; the RMD becomes then the standard matched detector (9). One sees in Fig. 1 that the P_D diminishes as DR increases. For a given P_{FA} additional robustness is somehow paid by a decrease in P_D .

To check that (16) is indeed a valid way to fix the threshold, we present in Table I the false alarm rates obtained using 10^8 simulations for six different P_{FA} 's and $DR = 1/8, 1/4$ and $1/2$. For $DR = 1/2$ some discrepancies appear, but when $DR \leq 1/4$ there is a good agreement between the expected and observed P_{FA} 's. It is thus fair to say that the tail of the probability of $L(y, z|H_0)$ is well approximated by the rule used in (16) as long as the condition (11) is satisfied.

2) *Behavior With Respect to False Values of the DR:* This is a pretty important issue for our approach since we need to fix *a priori* the value of the DR (2) by tuning α , which is the ratio between the standard deviations of the noises perturbing the signature x and the observations μx , (1). We have addressed this point in Section V-D.

Let us start with the case where one overestimates the DR. This happens for instance when one wrongly believes that the signature is imperfectly known. This is the situation we simulate.

The true simulated model is $y = \mu x + e$, $z = x$ ($DR = 0$) while we wrongly believe that it is $y = \mu x + e$, $z = x + \alpha n$ with $\alpha = .5$ ($DR = 1/8$). In Fig. 2, we compare for the same 3 values of P_{FA} and the same range of SNRs as in Fig. 1, the observed P_D for $DR = 1/8$, the optimal P_D , i.e., the best achievable results for the simulated model, and the expected P_D .

TABLE I
EXPECTED VERSUS ESTIMATED P_{FA} FOR THREE VALUES OF THE DR AND 10^8 REALIZATIONS

Expected P_{FA}	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
Obs. P_{FA} , DR = 1/8	$1.00 \cdot 10^{-1}$	$1.00 \cdot 10^{-2}$	$.99 \cdot 10^{-3}$	$.98 \cdot 10^{-4}$	$.95 \cdot 10^{-5}$	$.9 \cdot 10^{-6}$
Obs. P_{FA} , DR = 1/4	$1.01 \cdot 10^{-1}$	$1.0 \cdot 10^{-2}$	$.98 \cdot 10^{-3}$	$.94 \cdot 10^{-4}$	$.80 \cdot 10^{-5}$	$.9 \cdot 10^{-6}$
Obs. P_{FA} , DR = 1/2	$1.08 \cdot 10^{-1}$	$1.10 \cdot 10^{-2}$	$1.01 \cdot 10^{-3}$	$.85 \cdot 10^{-4}$	$.65 \cdot 10^{-5}$	$.6 \cdot 10^{-6}$

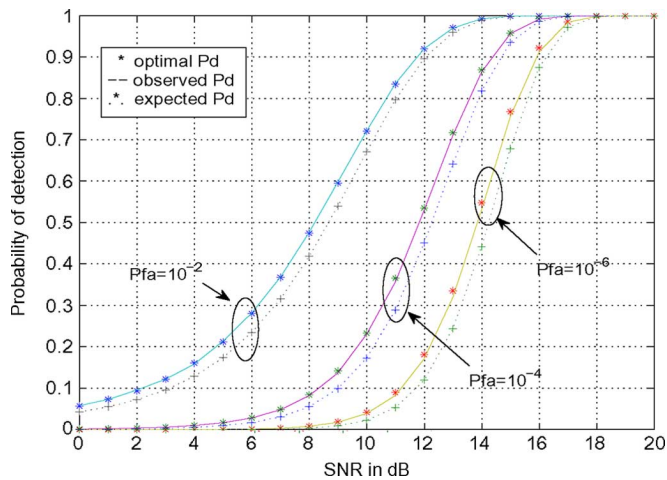


Fig. 2. Overestimating the uncertainty. The signature is perfectly known while we take DR = 1/8. Solid lines: the observed P_D 's; *: optimal P_D 's; ..+: the expected P_D 's.

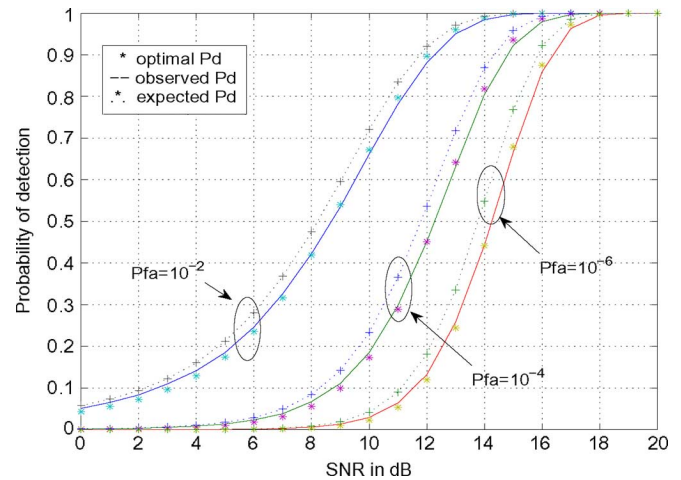


Fig. 3. Underestimating the uncertainty. The signature is imperfectly known (DR = 1/8) while we assume it is known. Solid lines: the observed P_D 's; *: the optimal P_D 's; ..+: the expected P_D 's.

The discrepancy between the observed and optimal P_D 's never exceeds 1.5%. At low SNR one tends to gain slightly in P_D while at high SNR one loses slightly. This trend can actually be deduced from (17). Note that it is quite possible to outperform locally the optimal detector, the one one would have designed if the true/simulated model had been known.

Of course, since our model is pessimistic, the observed and optimal probabilities of detection are both better than the expected ones. Note that the so-called optimal and expected P_D 's in Fig. 2 are obtained theoretically and are identical to the corresponding curves in Fig. 1.

As far as the P_{FA} 's are concerned, the simulations we performed show that the observed P_{FA} 's are strictly coherent with the expected ones. We do not report these results.

We now consider the case where one under-estimates the uncertainty on x . We simulate the extreme case where one believes the signature to be known exactly while it is indeed corrupted by noise.

The true simulated model is $y = \mu x + e, z = x + \gamma n$ with $\gamma = .5$ that corresponds to DR = 1/8, while we wrongly believe that it is $y = \mu x + e, z = x$ (DR = 0). In Fig. 3, we compare for the same 3 values of P_{FA} and the same range of SNRs, the observed P_D 's for the proposed detector to the P_D the optimal detector using the true model would attain, and to the expected P_D , the one one expects in the absence of uncertainty.

The difference between the optimal and the observed P_D 's is again quite small and never exceeds 2%. The overall picture is the same as above and can somehow be deduced from (18)

with $\alpha = 0$ and $\gamma = .5$: at low SNR the observed P_D is slightly better than the optimal and the opposite is true at high SNR.

With respect to the type of modeling errors we simulated in this section, it appears that the proposed RMD has the surprising property that, for moderate modeling errors, under H_0 , the observed (actual) P_{FA} 's have the expected values and, under H_1 , the observed P_D 's are quite close to the optimal ones, even though both can be quite greater or smaller than the expected ones depending upon whether the supposed model is pessimistic or optimistic.

B. Checking the Robustness Properties

We now apply the RMD to data that are not simulated according to (1) to test its *true* robustness. While the observation vector y is clearly identified, we will need to specify the value of the DR, or equivalently α , and define the vector that will play the role of z . We will replace the unknown x by z where needed.

While the vector z is taken equal to the expected signature, the tuning of α is less obvious. If, as in the three examples that follow, one has *a priori* a good idea of the expected discrepancies, one can evaluate the average angle between x and z and thus get a good idea of the aperture of the cone containing most x vectors. Since DR is roughly the square of the tangent of the half-angle of the aperture of this cone, this allows us to tune α which, is in any case, upper-bounded by the constraint (11). This is how we proceed below. The satisfaction of (11) guarantees that the P_{FA} 's are strictly equal to the expected ones and thus the detection performances that we observe are meaningful.

1) *A Ricean Channel*: To ease the comparison, we start from a scenario proposed in [11] and transpose it to our present context to end up comparing the RMD to the MSD [1], [13] described in Section IV. In [11], the data are complex, K is a number of sensors, the signature x represents steering vectors, the observations y are snapshots and a multisnapshot test is proposed that, in the single snapshot case, becomes the MSD.

The true steering vector x is assumed to lie in a known low dimensional subspace as $x = H\theta$ (5). This allows us to model or to approximate a Ricean channel [11], [18] for which one has

$$x = x_o + \frac{1}{\sqrt{q}} \sum_1^q g_k x(\phi_k)$$

where $x_o = x(\phi_o)$ corresponds to the line-of-sight component and the other vectors $x(\phi_k)$ represent neighboring components. There is thus a dominant mode x_o and additional smaller ones. With $H = [x_o \ H_2]$, the basic model $x = H\theta$ becomes $x = x_o + H_2\theta_2$. In [11], one scales H_2 to achieve a given uncertainty ratio (UR) defined as

$$\text{UR} = 10 \log_{10} \frac{\text{Tr}(H_2^T H_2)}{x_o^T x_o} \quad (19)$$

and θ_2 is drawn from $N(0, I_q)$. It is then natural to define the SNR as

$$\text{SNR} = 10 \log_{10} \frac{\mu^2 \{x_o^T x_o + \text{Tr}(H_2^T H_2)\}}{\sigma^2} \quad (20)$$

a function of the scalar μ . For a uniform array with half-wavelength spacings, the complex steering vector $x_c(\phi)$ is given by $[1 \ e^{i\omega} \ \dots \ e^{i(K-1)\omega}]^T$ with $\omega = \pi \sin \phi$. To stay with real vectors, we consider $x_r(\phi)$ the real part of $x_c(\phi)$ and define $x(\phi)$ as $x_r(\phi)$ (re)normalized to \sqrt{K} : $x(\phi) = \sqrt{K}(x_r(\phi)/\|x_r(\phi)\|)$.

The model we simulate below is

$$y = \mu x + e, \quad x = x_o + H_2 \theta_2$$

with $e \sim N(0, \sigma^2 I_K)$ and $\theta_2 \sim N(0, I_q)$

with $K = 12$, $\sigma = 1/2$, $x_o = x(0)$, $q = 2$, $H_2 = [x(5^\circ) \ x(10^\circ)]$. We scale H_2 to get $\text{UR} = -6$ dB, as in [11]. In the RMD, we take $z = x_o$ which yields $\alpha n = -H_2 \theta_2$ in (1). The obvious similarity between UR and DR in (2) yields a natural value for α : we take $\alpha = .7$ below.

We compare the probabilities of detection of both approaches as a function of the SNR (20) for different P_{FA} 's in Fig. 4. It appears that the probability of detections we obtain are slightly better than those obtained by the MSD. This is not expected but it is not impossible; indeed, even if the matched subspace detector is UMPU, it can under some circumstances be outperformed; this is the case for small UR.

As a matter of fact, the UR concept is of little interest for the MSD since its performance is independent of the UR; it is the same wherever x lies in the subspace spanned by x_o and H_2 . If one knows beforehand that its main contribution is in one direction and that the other directions are only marginally used, it may not be optimal to project on the whole potential subspace. This is what happens here and was also observed in [11]. For

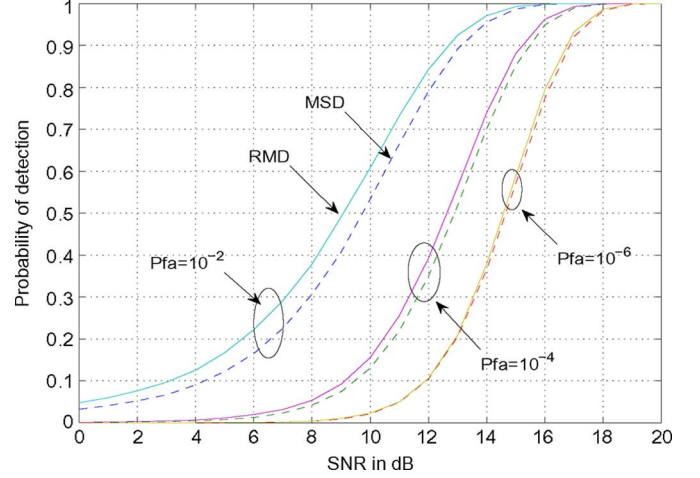


Fig. 4. Ricean channel: MSD (dashed lines) and RMD (solid lines).

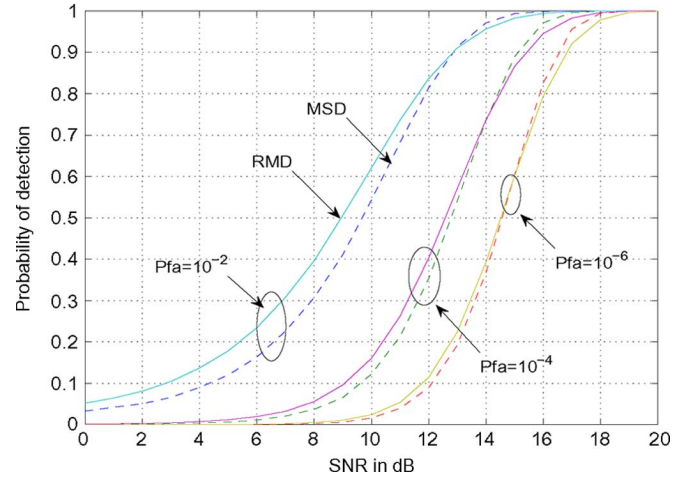


Fig. 5. Pointing error: MSD (dashed lines) and RMD (solid lines).

large values of UR, the MSD will have better performance than the RMD.

2) *A Pointing Error*: As a second example, we consider the case where x is a steering vector associated with a direction ϕ that is slightly different from the expected one ϕ_o . This is a scenario that is of interest in a radar context, for instance, where one will at best implement the test aiming to detect a target for directions belonging to a grid [16], [17]. It is well suited for the MSD since there is no privileged direction and the RMD should be able to handle it for small cells. We use the steering vectors and notations introduced above in Section VI-B-1). We take $K = 12$, $\sigma = 1/2$, the observation y satisfies $y = \mu x(\phi) + e$ with ϕ uniformly distributed between $\phi_{\min} = 0$ and $\phi_{\max} = 4^\circ$, $H = [x(\phi_{\max}/6) \ x(\phi_{\max}/2) \ x(5\phi_{\max}/6)]$ for the MSD and $z = x(\phi_{\max}/2)$ for RMD. For $\phi \in (0^\circ, 4^\circ)$, most $x(\phi)$ vectors are within a cone with half-angle $\theta \simeq 30^\circ$ around z ; this yields a potential value of $\alpha = 1$ using $\tan^2(\theta) \simeq \text{DR}$, but we take $\alpha = .7$ below to satisfy (11).

In Fig. 5, we compare for the same three values of P_{FA} and the SNR varying between 0 and 20 dB, the P_D of the RMD and of the MSD (6) which assumes that the signature lies in the range of H . The RMD has slightly better performance than the

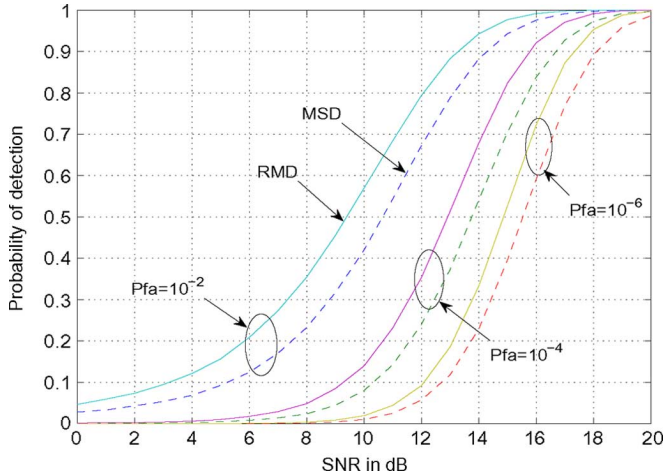


Fig. 6. A more realistic example: RMD (solid lines) versus MSD (dashed lines), the gain in performance is quite important.

MSD. We also took $H = [x(0) x(\phi_{\max}/2) x(\phi_{\max})]$ or H built using a Taylor series expansion around $\phi_{\max}/2$ with no better performance for the MSD. For larger ϕ_{\max} , the performance of the RMD degrades, those of the MSD remain remarkably steady and become better than those of the RMD.

3) *A Third Example:* We consider the case where one observes a sinusoid with jitter on the sampling period. The nominal sampling period is taken equal to 1 but the true sampling instants are

$$t_k = k + g\tau_k, \quad g \geq 0$$

with g fixed and τ_k independent random variables uniformly distributed in $[-.5, .5]$. This can, for instance, represent the effect of a fluctuating array in underwater acoustics [7]. The model is then $y = \mu x_p + e$ where $e \sim \mathcal{N}(0, \sigma^2 I)$, x_p (the subscript p for perturbed) is the unobserved signature with jittered sampling period that changes for each realization, and we take $z = x_r$ (the subscript r for regular), the expected signature with regular sampling period, which remains constant.

We take $K = 12$, $\sigma^2 = 1$, μ is tuned to achieve the expected SNR defined as $\mu^2(z^T z)/\sigma^2$ and the k th-component in z is equal to $A \cos(\omega k)$ with $A = 2$ and $\omega = 4$. The coefficient g in the jitter is taken equal to $1/2$. The discrepancies can thus be quite large. By simulations one can observe that the average angle between x_r and x_p is about 32° , the half-angle of the cone containing most x_p 's is about 40° corresponding to a DR and α larger than its upper-bound .7 given by (11).

In order to compare the RMD with the MSD, we propose to consider as above that the perturbed x_p vectors belong to a 3-D subspace. Since there is no natural or obvious way to build its basis, we estimate the covariance matrix of x_p and use the eigenvectors corresponding to the three dominant eigenvalues as basis vectors for the subspace [11], [16].

We performed 10^5 independent simulations, and in Fig. 6, we compare for the three values of P_{FA} and the SNR varying between 0 and 20 dB, the P_D of the RMD with $\alpha = 0.7$ to the P_D of the MSD (6). The RMD outperforms the MSD and the difference is quite important. This remains true for larger values of g and the difference actually increases mainly because, for

larger g , the dimension of the subspace necessary to represent the x_p 's has to be increased. Clearly, for this example, a cone is preferable to a finite-dimensional subspace.

In summary, one can say that the RMD has good performances even when the model differs from the one it has been built for, though a theoretical analysis or justification seems out of reach.

VII. CONCLUSION

We have proposed a new detection problem (1) and developed the associated approximate GLRT which we call the robust matched detector (RMD). It represents an extension of the standard matched detector. While in the latter the problem is to recognize the presence of a signal that lies in a known direction, we consider that the signal is imperfectly known, i.e., we model the signal to be detected as being known up to additive white Gaussian noise. Using first-order expansions, we have obtained approximate probability density functions for the GLR under both hypotheses and defined a way to compute the threshold that yields the desired probability of false alarm and to predict the associated probabilities of detection as a function of the SNR. We verify by simulations that the analysis we made is valid as soon as a condition that can be checked on the available data is satisfied.

Further simulations indicate that the proposed scheme is indeed efficient even when the true system cannot be modeled as (1) and that quite generally it applies when the signature is imperfectly known whatever the model behind this uncertainty may be. It is indeed this robustness of the RMD with respect to the model of the uncertainty that is particularly valuable and justifies its use.

As opposed to the well known matched subspace detector where one assumes the signal to lie in or close to a finite dimensional subspace, the model used for the RMD somehow assumes that it belongs to a cone with known axis whose aperture allows to tune the degree of uncertainty. The RMD works as long as the half-angle of the cone aperture is about 25 degrees (DR = 1/4). Such a cone is probably a better representation of the reality than a finite dimensional subspace in many practical situations and a rough value of its aperture is quite easy to learn.

APPENDIX

We are interested in $s_{\min}^2(\hat{C})$, the square of the minimal singular value of the K by 2 matrix

$$\hat{C} = \begin{bmatrix} z \\ \alpha y \end{bmatrix} = C + \Delta \quad \text{with } C = \begin{bmatrix} x \\ \alpha \mu x \end{bmatrix} \quad \text{and } \Delta = \begin{bmatrix} n \\ e \end{bmatrix}$$

where n and e are independent $\mathcal{N}(0, \sigma^2 I)$ random vectors. More precisely, we seek an expression of the first-order term of the expansion of $s_{\min}^2(\hat{C})$, we also denote \hat{s}_m^2 , with respect to σ^2 around $\sigma^2 = 0$.

We propose a quite ad hoc and simple analysis. One could also resort to more general results from (deterministic) matrix perturbation theory to establish the same result.

Since we are interested in \hat{s}_m^2 we introduce the order-2 matrix $\hat{C}^T \hat{C}$ and denote its two eigenvalues $\hat{\lambda}_M$ and $\hat{\lambda}_m = \hat{s}_m^2$. For $\sigma = 0$, the corresponding eigenvalues are λ_M and λ_m with

$\lambda_m = 0$ and $\lambda_M = (x^T x)((1/\alpha^2) + \mu^2) \gg \lambda_m$. For σ small enough, one then expects to have $\hat{\lambda}_M \gg \hat{\lambda}_m$ and therefore also

$$\hat{\lambda}_M \gg \hat{\lambda}_m \Rightarrow \hat{\lambda}_M^2 \gg \hat{\lambda}_M \hat{\lambda}_m \Rightarrow S^2 \gg P$$

where $S = \hat{\lambda}_M + \hat{\lambda}_m$ and $P = \hat{\lambda}_M \hat{\lambda}_m$ represent respectively the sum and the product of the two eigenvalues. Now $S^2 \gg P$ yields $\sqrt{S^2 - 4P} \simeq S - (2P/S)$ and thus $\hat{\lambda}_m = (1/2)(S - \sqrt{S^2 - 4P}) \simeq P/S$. From the expression of $\hat{C}^T \hat{C}$, it follows that $S = \hat{\lambda}_M + \hat{\lambda}_m = \text{Trace}(\hat{C}^T \hat{C}) = (z^T z/\alpha^2) + y^T y$ and $P = \hat{\lambda}_M \hat{\lambda}_m = \text{Det}(\hat{C}^T \hat{C}) = (z^T z/\alpha^2)(y^T y) - (z^T y/\alpha)^2$. Replacing now z by $x + \alpha n$ and y by $\mu x + e$, one gets the following approximations:

$$S = (x^T x) \left(\frac{1}{\alpha^2} + \mu^2 \right) + \text{higher order terms}$$

$$P = (x^T x) \left(\frac{e^T e}{\alpha^2} + \mu^2 n^T n \right) - \frac{(x^T e)^2}{\alpha^2} - \mu^2 (x^T n)^2 + \text{higher order terms}$$

and finally

$$\begin{aligned} \hat{s}_m^2 &\simeq \frac{P}{S} \\ &\simeq \frac{1}{1 + \alpha^2 \mu^2} \left(e^T \left(I - \frac{x x^T}{x^T x} \right) e \right. \\ &\quad \left. + \alpha^2 \mu^2 n^T \left(I - \frac{x x^T}{x^T x} \right) n \right). \end{aligned}$$

The expectation of this approximation of \hat{s}_m^2 is easily found to be equal to $(K - 1)\sigma^2$, a result that can also be deduced from matrix perturbation theory [26], [27].

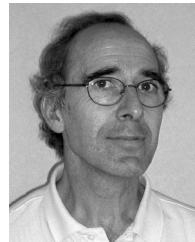
Note that we started assuming σ small but that more precisely what we used is $\hat{\lambda}_M \gg \hat{\lambda}_m$, a condition that can be easily checked in practice and which, with $\hat{\lambda}_M \simeq S$ and $\hat{\lambda}_m \simeq K\sigma^2$, is equivalent to

$$(x^T x) \left(\frac{1}{\alpha^2} + \mu^2 \right) \gg K\sigma^2.$$

REFERENCES

- [1] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2146–2157, Aug. 1994.
- [2] M. N. Desai and R. S. Mangoubi, "Robust Gaussian and nongaussian matched subspace detection," *IEEE Trans. Signal Process.*, vol. 51, no. 12, pp. 3115–3127, Dec. 2003.
- [3] H. V. Trees, *Optimum Array Processing*. New York: Wiley, 2002.
- [4] A. B. Gershman, "Robust adaptive beamforming in sensor arrays," *Int. J. Electron. Commun.*, vol. 56, no. 6, pp. 305–314, Dec. 1999.
- [5] O. Besson, A. A. Monakov, and C. Chalus, "Signal waveform estimation in the presence of uncertainties about the steering vector," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2432–2440, Sep. 2004.
- [6] O. Besson and F. Vincent, "Performance analysis of beamformers using generalized loading of the covariance matrix," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 452–459, Feb. 2005.

- [7] S. A. Vorobyov, A. B. Gershman, and Z. Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–323, Feb. 2003.
- [8] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.
- [9] M. Viberg and A. Swindlehurst, "A Bayesian approach to autocalibration for parametric array processing," *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3495–3507, Dec. 1994.
- [10] O. Besson, L. L. Scharf, and F. Vincent, "Matched direction detectors," in *Proc. Int. Conf. Acoustics, Speech Signal Process.*, 2005, vol. 4, pp. 977–980.
- [11] O. Besson, L. L. Scharf, and F. Vincent, "Matched direction detectors and estimators for array processing with subspace steering vector uncertainties," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4453–4463, Dec. 2005.
- [12] O. Besson and L. L. Scharf, "CFAR matched direction detector," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2840–2844, Jul. 2006.
- [13] L. L. Scharf, *Statistical Signal Process. Detection, Estimation and Time Series Analysis*. Boston, MA: Addison Wesley, 1991.
- [14] L. L. Scharf and M. McCloud, "Blind adaptation of zero forcing projections and oblique pseudo-inverses for subspace detection and estimation when interference dominates noise," *IEEE Trans. Signal Process.*, vol. 50, no. 12, pp. 2938–2946, Dec. 2002.
- [15] S. Bose and A. O. Steinhardt, "Adaptive array detection of uncertain rank one waveforms," *IEEE Trans. Signal Process.*, vol. 44, no. 11, pp. 2801–2809, Nov. 1996.
- [16] A. Zeira and B. Friedlander, "Robust subspace detectors," in *Proc. 31th Asilomar Conf. Signals, Syst. Comput.*, Nov. 1997, vol. 1, pp. 778–792.
- [17] R. S. Raghavan, N. Pulsone, and D. J. McLaughlin, "Performance of the GLRT for adaptive vector subspace detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 10, pp. 1473–1487, Oct. 1996.
- [18] G. Fuks, J. Goldberg, and H. Messer, "Bearing estimation in Ricean channel—Part 1: Inherent accuracy limitations," *IEEE Trans. Signal Process.*, vol. 49, no. 5, pp. 925–937, May 2001.
- [19] D. Astely and B. Ottersten, "The effects of local scattering on direction of arrival estimation with MUSIC," *IEEE Trans. Signal Process.*, vol. 47, no. 12, pp. 3220–3234, Dec. 1999.
- [20] 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.
- [21] A. Zeira and B. Friedlander, "Robust adaptive subspace detectors for space-time processing," in *Proc. Int. Conf. Acoustics, Speech Signal Process.*, 1998, vol. 4, pp. 1965–1968.
- [22] L. J. Gleser, "Estimation in a multivariate "errors in variables" regression model: Large sample results," *Annals Statist.*, vol. 9, no. 1, pp. 24–44, 1981.
- [23] G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," *SIAM J. Numer. Anal.*, vol. 17, no. 6, pp. 883–893, 1980.
- [24] B. Porat, *Digital Process. of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1994.
- [25] E. L. Lehmann, *Testing Statistical Hypotheses*. New York: Wiley, 1986.
- [26] G. W. Stewart and J. G. Sun, *Matrix Perturbation Theory*. New York: Academic, 1990, pp. 263–266.
- [27] J. J. Fuchs, "Structure and order estimation of multivariable stochastic processes," *IEEE Trans. Autom. Control*, vol. 35, no. 12, pp. 1338–1341, Dec. 1990.



Jean Jacques Fuchs was born in France in 1950. He graduated from the École Supérieure d'Électricité, Paris, France, in 1973 and received the M.S. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1974.

After a short period in industry with Thomson C.S.F., he joined the Institut de Recherche en Informatique et Systèmes Aléatoires (IRISA) in 1976. Since 1983, he has been a Professor at the Université de Rennes 1, Rennes, France. His research interests shifted from adaptive control and identification, in

which he obtained the Thèse d'Etat in 1982, towards signal processing. He is now involved in array processing and sparse representations.