[8] S. Ventosa, C. Simon, M. Schimmel, J. J. Danobeitia, and A. Manuel, "The S-transform from a wavelet point of view," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2771–2780, Jul. 2008.
[9] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, A. Jeffrey and D. Zwillinger, Eds. Orlando, FL: Harcourt Brace Jovanovich, 2000.
[10] C. R. Pinnegar, M. W. Wong, and H. Zhu, "Integral representations of the TT-transform," *Applicable Analysis*, vol. 85, no. 8, pp. 933–940, Aug. 2006.
[11] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1972.

Matched Subspace Detection With Hypothesis Dependent Noise Power

Franois Vincent, Olivier Besson, and Cédric Richard

***Abstract*—We consider the problem of detecting a subspace signal in white Gaussian noise when the noise power may be different under the null hypothesis—where it is assumed to be known—and the alternative hypothesis. This situation occurs when the presence of the signal of interest (SOI) triggers an increase in the noise power. Accordingly, it may be relevant in the case of a mismatch between the actual SOI subspace and its presumed value, resulting in a modelling error. We derive the generalized likelihood ratio test (GLRT) for the problem at hand and contrast it with the GLRT which assumes known and equal noise power under the two hypotheses. A performance analysis is carried out and the distributions of the two test statistics are derived. From this analysis, we discuss the differences between the two detectors and provide explanations for the improved performance of the new detector. Numerical simulations attest to the validity of the analysis.**

***Index Terms*—Generalized likelihood ratio test (GLRT), robustness, signal detection.**

I. Introduction

Detecting a partly known signal in additive noise is a widespread task in many signal processing applications. It is the main goal of radar or sonar systems and can be encountered in most communication or seismic schemes as well as in pattern recognition, to cite a few [1], [2]. Optimum detectors, that maximize the probability of detection (P_d) for a given probability of false alarm (P_{fa}) have been developed for a wide class of signal and noise modeling. This kind of detector is designed for a specific signal waveform and a given noise probability density function (pdf). Unfortunately, it turns out that in many cases optimum detectors can suffer a drastic degradation in performance for small deviations from the nominal assumptions. Such deviations can occur because of signal distortion, scattering, imprecise calibration, jitter or errors in sensor localization for instance. Since the real signal waveform and the noise pdf are rarely exactly known in practice, one usually needs to develop robust detectors. Many studies have been carried out on robust detectors, with a view to maintain acceptable performances for a large class of deviations [3]–[8]. In this paper, we consider the problem of detecting a deterministic signal belonging to a known subspace in white Gaussian additive noise [see (1) below]. This modeling is widely used each time one has to detect a partly known narrowband signal such as in radar, sonar or communication systems [9]. In the literature two different cases have been studied depending if the noise power is supposed to be known or not. In many applications, one can precisely evaluate the noise power under the null hypothesis when secondary data are available (this is classically the case in radar systems) because this power is directly the data power. However, it is difficult to verify that this noise power remains the same under H_1 because its estimation is linked to the assumed signal model. In addition, in many cases, one can suppose that this power is modified if the signal is present. This variation could be due, e.g., to the *receiver electronics*. For instance, when automatic gain control (AGC) is used, the noise factor depends on the signal amplitude. This effect is a well-known problem for digital cameras where signal-dependent noise is always present [10]. One may have the same problem in magnetic recordings [11]. More generally, any nonlinearity in the electronics can modify the noise power by creating products between the noise part and the signal part. Moreover, quantification can be another source of noise power modification. Quantification noise could be considered as additive uniform white noise whose power depends on the gauge and the number of quantification bit used. Then, if the gauge changes when the signal is present, the noise power will change too. In this case, one can observe a noise power reduction. Accordingly, a *non-complete knowledge* of the signal to be detected can lead to noise power variations. One can encounter this problem in acoustic recognition or in automotive engine knock detection for instance [12]. Knocking is an undesired auto-ignition occurring in the cylinder chamber that limits the efficiency of modern engines and has to be controlled. The generated shockwave stimulates characteristic oscillations analyzed through a vibration sensor. However, this shockwave can also create other non modeled noises due to other mechanic vibrations and subsequently increase the noise power under H_1 [13]. More generally, every *signal modeling error* will be added to the non modeled data part and will thus increase the noise power under H_1 [14], [15]. In this paper, we propose to study the robustness to noise power variation between the two hypotheses by introducing and analyzing a new robust detector. In contrast to most robust detectors introduced for signal modeling error, we do not need any hypothesis about the noise power variation.

II. Generalized Likelihood Ratio Test

The decision problem to be studied in this paper can be described as follows. We are given N samples from a complex scalar time series which are gathered in the N-dimensional measurement vector $\boldsymbol{x} = [x(0) \quad x(1) \quad \ldots \quad x(N-1)]^T$. The problem is to decide between the null hypothesis (H_0) and the alternative one (H_1):

$$\begin{cases} H_0, & \boldsymbol{x} = \boldsymbol{n}_0 \\ H_1, & \boldsymbol{x} = \boldsymbol{s} + \boldsymbol{n}_1 \end{cases} \tag{1}$$

where $\boldsymbol{s} = \boldsymbol{A}\boldsymbol{a}$ is the deterministic signal of interest which belongs to a known subspace $\langle \boldsymbol{A} \rangle$ of size R, and the complex amplitude vector $\boldsymbol{a}$ is unknown. $\boldsymbol{n}_0$ (respectively, $\boldsymbol{n}_1$) stands for the noise vector, and is supposed to be zero-mean Gaussian distributed with known (respectively, unknown) covariance matrix $\sigma_0^2 \boldsymbol{I}$ (respectively, $\sigma_1^2 \boldsymbol{I}$). It is usually assumed that $\sigma_0^2 = \sigma_1^2$, and the latter may be known or not. In the case where $\sigma_0^2 = \sigma_1^2$ is known, the GLRT is the so-called matched subspace detector (MSD) [1] and consists in comparing the test statistic

$$T_{\mathrm{kn}} = \frac{\boldsymbol{x}^H \boldsymbol{P}_A \boldsymbol{x}}{N \sigma_0^2} \tag{2}$$

Manuscript received January 16, 2008; revised July 9, 2008. First published August 19, 2008; current version published October 15, 2008. The associate editor coordinating the review of this paper and approving it for publication was Dr. Chong-Meng Samson See.

F. Vincent and O. Besson are with the ISAE, Department of Electronics, Optronics and Signal, University of Toulouse, 31055 Toulouse, France (e-mail: vincent@isae.fr; besson@isae.fr).
C. Richard is with the Charles Delaunay Institute, FRE CNRS 2848, BP 2060, 10010 Troyes, France (e-mail: cedric.richard@utt.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.2008.929667

to a threshold, where $\boldsymbol{P}_A = \boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H$ is the orthogonal projection on the signal subspace. In the sequel we will use the short hand notation $\boldsymbol{x}_A = \boldsymbol{P}_A\boldsymbol{x}$. When σ_0^2 is unknown, the GLRT becomes the CFAR MSD, whose test statistic is [1]

$$T_{\mathrm{un}} = \frac{\boldsymbol{x}^H\boldsymbol{P}_A\boldsymbol{x}}{\boldsymbol{x}^H\boldsymbol{P}_\perp\boldsymbol{x}} \tag{3}$$

where $\boldsymbol{P}_\perp = \boldsymbol{I} - \boldsymbol{P}_A$. In both cases, T_{kn} and T_{un} can be viewed as signal-to-noise ratio (SNR) estimates.

As discussed in the introduction, in many applications, the noise power under H_0 can be very accurately estimated, and hence it is natural to assume that σ_0^2 is known. In contrast, σ_1^2 may be different from σ_0^2 and is unknown. Under the stated assumptions the pdf of $\boldsymbol{x}$ under H_1 is given by [1], [16]

$$p(\boldsymbol{x};\boldsymbol{a},\sigma_1^2|H_1) = \pi^{-N}\sigma_1^{-2N}\exp\left\{-\frac{\|\boldsymbol{x}-\boldsymbol{A}\boldsymbol{a}\|^2}{\sigma_1^2}\right\}. \tag{4}$$

It is well known that, for a given $\boldsymbol{a}$,

$$\max_{\sigma_1^2} p(\boldsymbol{x};\boldsymbol{a},\sigma_1^2|H_1) = (e\pi)^{-N}\left[N^{-1}\|\boldsymbol{x}-\boldsymbol{A}\boldsymbol{a}\|^2\right]^{-N} \triangleq g(\boldsymbol{a}) \tag{5}$$

and is achieved when $\sigma_1^2 = N^{-1}\|\boldsymbol{x}-\boldsymbol{A}\boldsymbol{a}\|^2$. Next, the maximum of $g(\boldsymbol{a})$ with respect to $\boldsymbol{a}$ is obtained for $\hat{\boldsymbol{a}} = \left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H\boldsymbol{x}$, resulting in

$$\begin{aligned}\max_{\boldsymbol{a}} g(\boldsymbol{a}) &= (e\pi)^{-N}\left[N^{-1}\|\boldsymbol{x}-\boldsymbol{A}\hat{\boldsymbol{a}}\|^2\right]^{-N}\\ &= \left[e\pi N^{-1}\|\boldsymbol{x}_\perp\|^2\right]^{-N}\end{aligned} \tag{6}$$

where $\boldsymbol{x}_\perp = \boldsymbol{P}_\perp\boldsymbol{x}$. Let $\hat{\sigma}_1^2 = N^{-1}\|\boldsymbol{x}-\boldsymbol{A}\hat{\boldsymbol{a}}\|^2 = N^{-1}\|\boldsymbol{x}_\perp\|^2$ denote the maximum-likelihood estimate (MLE) of σ_1^2. Then, the GLR is given by

$$\begin{aligned}\frac{p(\boldsymbol{x};\hat{\boldsymbol{a}},\hat{\sigma}_1^2|H_1)}{p(\boldsymbol{x}|H_0)} &= \frac{\left[e\pi N^{-1}\|\boldsymbol{x}_\perp\|^2\right]^{-N}}{\pi^{-N}\sigma_0^{-2N}\exp\left\{-\frac{\|\boldsymbol{x}\|^2}{\sigma_0^2}\right\}}\\ &= \left[e\frac{\|\boldsymbol{x}_\perp\|^2}{N\sigma_0^2}\exp\left\{-\frac{\|\boldsymbol{x}\|^2}{N\sigma_0^2}\right\}\right]^{-N}.\end{aligned} \tag{7}$$

Taking the logarithm of the Nth root of the GLR, it ensues that the GLRT amounts to comparing

$$T_r = \frac{\|\boldsymbol{x}\|^2}{N\sigma_0^2} - \ln\frac{\|\boldsymbol{x}_\perp\|^2}{N\sigma_0^2} - 1 \tag{8}$$

to an appropriate threshold. Noticing that T_{kn} can be rewritten as

$$T_{\mathrm{kn}} = \frac{\|\boldsymbol{x}\|^2}{N\sigma_0^2} - \frac{\|\boldsymbol{x}_\perp\|^2}{N\sigma_0^2} \tag{9}$$

we observe that T_r and T_{kn} differ in the way they remove the power in the subspace orthogonal to $\langle\boldsymbol{A}\rangle$ from the total power. In the sequel, we will analyze this modification and explain why it leads to improved performance when the ratio $b \triangleq \left(\sigma_1^2/\sigma_0^2\right)$ is not equal to one. As a final remark, we also introduce the clairvoyant detector which knows σ_1^2, in addition to σ_0^2. This detector is obviously not realizable but will serve as an upper limit in our comparison. Under the stated assumptions, it is straightforward to show that the test statistic of the clairvoyant detector is given by

$$T_c = \frac{\|\boldsymbol{x}\|^2}{N\sigma_0^2} - \frac{\|\boldsymbol{x}_\perp\|^2}{N\sigma_1^2} - \ln\frac{\sigma_1^2}{\sigma_0^2}. \tag{10}$$

We can obviously notice that $T_c = T_{\mathrm{kn}}$ if $b = 1$.

III. Performance Analysis and Discussion

In this section, we derive the (possibly asymptotic) distributions of the three test statistics T_{kn}, T_r, and T_c under both hypotheses with a view 1) to predict their performance in terms of probability of detection and 2) to comprehend the differences between them in order to explain why and how the new detector may improve the conventional one. In the sequel, $\mathcal{N}\left(\mu,\sigma^2\right)$ denotes the Gaussian distribution with mean μ and variance σ^2 and $\chi_n^2(\lambda)$ denotes the chi-square distribution with n degrees of freedom and noncentrality parameter λ. The central chi-square distribution will be denoted as $\chi_n^2(0)$. When a random variable ζ can be written as the product of a scalar α and a chi-square distributed random variable, we will use the notation $\zeta \sim \alpha\chi_n^2(\lambda)$. Accordingly, the notation $\zeta \sim \alpha_1\chi_{n_1}^2(\lambda_1) + \alpha_2\chi_{n_2}^2(\lambda_2)$ means that ζ is distributed as the sum of (possibly scaled) independent chi-square distributed random variables with n_1, n_2 degrees of freedom and noncentrality parameters λ_1, λ_2, respectively.

A. Distribution of T_{kn}

We begin by analyzing T_{kn}. Since $\boldsymbol{x}$ is drawn from a complex multivariate Gaussian distribution, with zero mean and covariance matrix $\sigma_0^2\boldsymbol{I}$ or $\sigma_1^2\boldsymbol{I}$, it follows immediately that [1]

$$T_{\mathrm{kn}} = \frac{\|\boldsymbol{x}_A\|^2}{N\sigma_0^2} \sim \begin{cases}\frac{1}{2N}\chi_{2R}^2(0) & \text{under } H_0\\ \frac{b}{2N}\chi_{2R}^2\left(2\frac{\boldsymbol{s}^H\boldsymbol{s}}{\sigma_1^2}\right) & \text{under } H_1.\end{cases} \tag{11}$$

B. Distribution of T_c

Let us recall that

$$\begin{aligned}T_c &= \frac{\|\boldsymbol{x}\|^2}{N\sigma_0^2} - \frac{\|\boldsymbol{x}_\perp\|^2}{N\sigma_1^2} - \ln\frac{\sigma_1^2}{\sigma_0^2}\\ &= \frac{\|\boldsymbol{x}_A\|^2}{N\sigma_0^2} + \frac{b-1}{b}\frac{\|\boldsymbol{x}_\perp\|^2}{N\sigma_0^2} - \ln b.\end{aligned} \tag{12}$$

Since $\boldsymbol{x}_A$ and $\boldsymbol{x}_\perp$ are independent and Gaussian distributed, T_c is distributed as

$$T_c \sim \begin{cases}\frac{1}{2N}\chi_{2R}^2(0) + \frac{b-1}{2bN}\chi_{2(N-R)}^2(0) - \ln b & \text{under } H_0\\ \frac{b}{2N}\chi_{2R}^2\left(\frac{2\boldsymbol{s}^H\boldsymbol{s}}{\sigma_1^2}\right) + \frac{b-1}{2N}\chi_{2(N-R)}^2(0) - \ln b & \text{under } H_1.\end{cases} \tag{13}$$

C. Asymptotic Distribution of T_r

In order to obtain the distribution of T_r, we first relate it to T_{kn}. Towards this end, note that

$$T_r - T_{\mathrm{kn}} = \frac{\|\boldsymbol{x}_\perp\|^2}{N\sigma_0^2} - \ln\frac{\|\boldsymbol{x}_\perp\|^2}{N\sigma_0^2} - 1 = f(\gamma) \tag{14}$$

where $\gamma = \|\boldsymbol{x}_\perp\|^2/\left(N\sigma_0^2\right)$ and $f(\gamma) = \gamma - \ln\gamma - 1$. We can notice that f is a positive function with $f(1) = 0$, and that γ is an estimate of the noise power normalized by the presumed one. From the Gaussian distribution of $\boldsymbol{x}$, we infer that

$$\gamma = \frac{\|\boldsymbol{x}_\perp\|^2}{N\sigma_0^2} \sim \begin{cases}\frac{1}{2N}\chi_{2(N-R)}^2(0) & \text{under } H_0\\ \frac{b}{2N}\chi_{2(N-R)}^2(0) & \text{under } H_1.\end{cases} \tag{15}$$

However, since $f(\gamma)$ is nonlinear and its inverse cannot be obtained in closed form, deriving the exact distribution of $T_r - T_{\mathrm{kn}}$ is intractable. In order to come up with manageable expressions, we thus investigate an asymptotic approach, assuming that the number of samples N is large. We first approximate the distribution of γ, show that its PDF is concentrated around 1, and then use a Taylor series expansion of $f(\cdot)$ around 1. As n grows large, it is well known that the chi-square distribution $\chi_n^2(0)$ converges to a Gaussian distribution with mean n and variance $2n$. It follows that

$$\gamma \overset{a}{\sim} \begin{cases} \mathcal{N}\left(1-r, \frac{1-r}{N}\right) & \text{under } H_0 \\ \mathcal{N}\left(b(1-r), \frac{b^2(1-r)}{N}\right) & \text{under } H_1 \end{cases} \tag{16}$$

where $r = R/N$ and $\overset{a}{\sim}$ means asymptotically distributed. Next we assume that $b(1-r)$ is close to 1. Note that this is a mild assumption as, in the applications we consider, the number of samples can be quite large (e.g., a few hundreds) while the dimension of the signal subspace is typically small (inferior to 10). In addition, the noise power increase $b \simeq 1+\delta$ is close to 1, and hence $b(1-r)$ is also close to 1. It follows that the asymptotic PDF of γ will be highly concentrated around 1. Now, in the vicinity of $\gamma = 1$, $f(\gamma) \simeq (1/2)(\gamma-1)^2$. Using the latter approximation along with (16), one obtains that

$$f(\gamma) \overset{a}{\sim} \begin{cases} \frac{1-r}{2N}\chi_1^2\left(\frac{Nr^2}{1-r}\right) & \text{under } H_0 \\ \frac{b^2(1-r)}{2N}\chi_1^2\left(\frac{N[b(1-r)-1]^2}{b^2(1-r)}\right) & \text{under } H_1. \end{cases} \tag{17}$$

We now observe that T_{kn} depends on $\|\boldsymbol{x}_A\|^2$ while $f(\gamma)$ depends on $\|\boldsymbol{x}_\perp\|^2$ only. Since $\|\boldsymbol{x}_A\|^2$ and $\|\boldsymbol{x}_\perp\|^2$ are independent, T_{kn} and $f(\gamma)$ are also independent. Therefore, the asymptotic distribution of T_r is given by

$$T_r \overset{a}{\sim} \begin{cases} \frac{1-r}{2N}\chi_1^2\left(\frac{Nr^2}{1-r}\right) + \frac{1}{2N}\chi_{2R}^2(0) & \text{under } H_0 \\ \frac{b^2(1-r)}{2N}\chi_1^2\left(\frac{N[b(1-r)-1]^2}{b^2(1-r)}\right) + \frac{b}{2N}\chi_{2R}^2\left(2\frac{\boldsymbol{s}^H\boldsymbol{s}}{\sigma_1^2}\right) & \text{under } H_1. \end{cases} \tag{18}$$

The exact pdf and the cumulative distribution function (cdf) of a linear combination of central or non-central chi-square distributed random variables can be found, e.g., in [17], where they are expressed in terms of an infinite series of Laguerre polynomials. Although the expressions in [17] are exact, they are not easily manageable when it comes to derive the receiver operating characteristics. In order to come up with exploitable expressions, we examine a further approximation to (18). Indeed, under H_0,

$$\begin{aligned} T_r &\overset{a}{\sim} \frac{1}{2N}\left[(1-r)\chi_1^2\left(\frac{Nr^2}{1-r}\right) + \chi_{2R}^2(0)\right] \\ &\simeq \frac{1}{2N}\left[\chi_1^2(0) + \chi_{2R}^2(0)\right] = \frac{1}{2N}\chi_{2R+1}^2(0). \end{aligned} \tag{19}$$

Accordingly, under H_1,

$$\begin{aligned} T_r &\overset{a}{\sim} \frac{b}{2N}\left[b(1-r)\chi_1^2\left(\frac{N[b(1-r)-1]^2}{b^2(1-r)}\right) + \chi_{2R}^2\left(2\frac{\boldsymbol{s}^H\boldsymbol{s}}{\sigma_1^2}\right)\right] \\ &\simeq \frac{b}{2N}\chi_{2R+1}^2\left(\frac{N[b(1-r)-1]^2}{b^2(1-r)} + 2\frac{\boldsymbol{s}^H\boldsymbol{s}}{\sigma_1^2}\right). \end{aligned} \tag{20}$$

Let $\lambda_1 = N[b(1-r)-1]^2/(b^2(1-r))$ and $\lambda_s = 2\left(\boldsymbol{s}^H\boldsymbol{s}/\sigma_1^2\right) = 2N \times \text{SNR} = 2\text{SNR}_p$ where SNR (respectively, SNR_p) denotes the input (respectively, output) signal-to-noise ratio. Then

$$T_r \overset{a}{\sim} \begin{cases} \frac{1}{2N}\chi_{2R+1}^2(0) & \text{under } H_0 \\ \frac{b}{2N}\chi_{2R+1}^2(\lambda_1+\lambda_s) & \text{under } H_1. \end{cases} \tag{21}$$

The above expression holds for large N and $b(1-r) \simeq 1$. We verified that the pdf and cdf of (21) are very close to those of (18). Furthermore, through extensive Monte Carlo simulations, we checked that the pdf of (21) matches very accurately the exact pdf of T_r, obtained from a large number of independent test statistics T_r drawn from (8).

D. Receivers Operating Characteristics

The distributions derived above enable one to obtain the receivers operating characteristics (ROC), that is the probability of detection P_d as a function of the probability of false alarm P_{fa}. Let

$$Q_{\chi^2}(y; n, \lambda) = \int_y^\infty p_{\chi^2(n,\lambda)}(x)dx$$

represent the complementary cumulative distribution function (ccdf) of a noncentral χ^2 random variable with n degrees of freedom and noncentrality parameter λ. Also, let $Q_{\chi^2}^{-1}(p; n, \lambda)$ denote its inverse function. Then, using the fact that T_r is asymptotically chi-square distributed [see (21)], one can write

$$P_{fa}(T_r) = Q_{\chi^2}(2N\eta_r; 2R+1, 0) \tag{22}$$

$$P_d(T_r) = Q_{\chi^2}\left(\frac{2N}{b}\eta_r; 2R+1, \lambda_1+\lambda_s\right). \tag{23}$$

It follows that

$$\begin{aligned} 2N\eta_r &= Q_{\chi^2}^{-1}(P_{fa}(T_r); 2R+1, 0) \\ &= bQ_{\chi^2}^{-1}(P_d(T_r); 2R+1, \lambda_1+\lambda_s) \end{aligned} \tag{24}$$

and hence the ROC of the new detector is characterized by one of the following equations:

$$P_d(T_r) = Q_{\chi^2}\left(b^{-1}Q_{\chi^2}^{-1}(P_{fa}(T_r); 2R+1, 0);\ 2R+1, \lambda_1+\lambda_s\right) \tag{25a}$$

$$P_{fa}(T_r) = Q_{\chi^2}\left(bQ_{\chi^2}^{-1}(P_d(T_r); 2R+1, \lambda_1+\lambda_s);\ 2R+1, 0\right). \tag{25b}$$

Using similar arguments for the conventional detector, one obtains

$$P_d(T_{\text{kn}}) = Q_{\chi^2}\left(b^{-1}Q_{\chi^2}^{-1}(P_{fa}(T_{\text{kn}}); 2R, 0);\ 2R+1, \lambda_s\right) \tag{26a}$$

$$P_{fa}(T_{\text{kn}}) = Q_{\chi^2}\left(bQ_{\chi^2}^{-1}(P_d(T_{\text{kn}}); 2R, \lambda_s);\ 2R+1, 0\right). \tag{26b}$$

E. Discussion

The previous sections provide a theoretical performance analysis of the new detectors. In order to comprehend how and why the new detector may outperform the conventional one, we now provide an intuitive and qualitative analysis of the differences between T_r and T_{kn} in presence of noise power variation using Fig. 1. The classical test, T_{kn} projects the data onto the signal subspace to estimate the signal power since, under H_1

$$\mathcal{E}\left\{\frac{\|\boldsymbol{x}_A\|^2}{N}\right\} = P_s + \sigma_1^2 r \simeq P_s \text{ if } r \ll 1. \tag{27}$$

It compares this estimated power to a threshold depending on the presumed noise level σ_0^2. In case of noise power variation, T_{kn} is marginally modified if the size of the signal subspace is small compared to N ($r \ll 1$). In contrast, the new robust test, T_r, can be written as

$$\begin{aligned} T_r &= T_{\text{kn}} + f(\gamma) \simeq T_{\text{kn}} + \frac{1}{2}(\gamma-1)^2 \\ &= T_{\text{kn}} + \frac{1}{2}\left(\frac{\frac{\|x_\perp\|^2}{N} - \sigma_0^2}{\sigma_0^2}\right)^2 \\ &\simeq T_{\text{kn}} + \frac{1}{2}\left(\frac{\hat{\sigma}_1^2 - \sigma_0^2}{\sigma_0^2}\right)^2 \end{aligned} \tag{28}$$

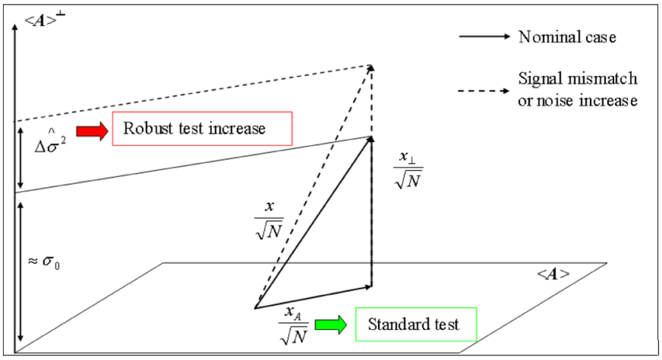

Fig. 1. Geometric interpretation of the detectors.

and $\mathcal{E}\left\{\hat{\sigma}_1^2|H_k\right\} = \sigma_k^2(1-r) \simeq \sigma_k^2$. Then it is the sum of two independent positive random variables. The first one is the classical test, depending only on the projection of the data on the signal subspace; the second depends on the projection on the noise subspace. This corrective term estimates the noise power and calculates the difference with the presumed one $\left(\sigma_0^2\right)$. Then T_r modifies $T_{\rm kn}$ by adding a corrective factor proportional to the square of this noise power variation. This correction increases the probability of detection.

In Fig. 1, the continuous lines correspond to the case where no power variation is present and the dashed ones correspond to the opposite case. In the first case, $\mathcal{E}\left\{\hat{\sigma}_1^2\right\} \simeq \sigma_0^2$ and $T_r \simeq T_{\rm kn}$. In the second case, if $r \ll 1$, the projection on the signal subspace remains almost the same, and the projection on the noise subspace increases by the noise power variation $\Delta\sigma^2$. In this case, $T_{\rm kn}$ remains approximately the same and T_r shifts depending of the estimated noise power variation (noted $\widehat{\Delta\sigma^2}$ in Fig. 1). Under H_0, we know that $\sigma^2 = \sigma_0^2$, so $T_r \simeq T_{\rm kn}$. Under H_1, if $b \neq 1$, T_r increases compared to $T_{\rm kn}$ in order to improve the probability of detection. One can notice that the only information used by T_r to modify $T_{\rm kn}$ is the power in the noise subspace. One can then expect a good behavior of T_r each time the estimated noise power is different from the expected one. This will be the case if signal mismatches are present. The projection onto the signal subspace will decrease and the projection onto the noise subspace will increase. T_r could recover a part of the energy having moved from one subspace to the other and try to maintain the test performance. This behavior will be illustrated by examples in Section IV. The previous geometric interpretation enables one to understand the differences between T_r and $T_{\rm kn}$. We now investigate how it results in terms of the ROCs. In order to assess the gain provided by T_r, one could evaluate the difference between $P_d(T_r)$ and $P_d(T_{\rm kn})$ for a given P_{fa}. Conversely, one could compute the P_{fa} ratio for a given P_d, i.e.,

$$G_{P_{fa}} = 10\log_{10}\frac{Q_{\chi^2}\left(bQ_{\chi^2}^{-1}\left(P_d;2R,\lambda_s\right);2R,0\right)}{Q_{\chi^2}\left(bQ_{\chi^2}^{-1}\left(P_d;2R+1,\lambda_1+\lambda_s\right);2R+1,0\right)}. \tag{29}$$

Although the expression in (29) is closed form, it does not lead to a simple and clear understanding of the respective influences of b, N, and SNR. In order to gain insights we make some approximations, with a view to obtain simple expressions that could state the conditions under which the new detector is more performant than the conventional one. First note that the P_{fa} ratio between two detectors is usually maximum around $P_d \simeq 0.5$. Hence, we will calculate $G_{P_{fa}}$ between the two detectors for this particular value of P_d, which gives a good information about the performance gain. A first difficulty for computing (29) stems from the evaluation of the inverse CCDF. Note that, for any distribution, $P_d = 0.5$ is obtained when the threshold is set to the median of the distribution. Furthermore, under H_1 the noncentrality parameter for the distributions of T_r and $T_{\rm kn}$ are $\lambda_1 + \lambda_s$ and λ_s, respectively, with $\lambda_s = 2N\mathrm{SNR} = \mathrm{SNR}_p$ equal to twice the output SNR. Therefore, assuming high (output) signal to noise ratio, the noncentrality parameters will be large, and thus the chi-square distributions can be fairly well approximated by Gaussian distributions with respective means $2R+1+\lambda_1+\lambda_s$ and $2R+\lambda_s$. Since the median of a Gaussian distribution coincides with its mean, we have approximately

$$Q_{\chi^2}^{-1}\left(0.5;2R,\lambda_s\right) = \frac{2N}{b}\eta_{\rm kn} \simeq 2R+\lambda_s$$
$$Q_{\chi^2}^{-1}\left(0.5;2R+1,\lambda_1+\lambda_s\right) = \frac{2N}{b}\eta_r \simeq 2R+1+\lambda_1+\lambda_s.$$

Therefore, $G_{P_{fa}}$ for $P_d = 0.5$ is approximately

$$G_{P_{fa}} \simeq 10\log_{10}\frac{Q_{\chi^2}\left(b\left(2R+\lambda_s\right);2R,0\right)}{Q_{\chi^2}\left(b\left(2R+1+\lambda_1+\lambda_s\right);2R+1,0\right)}. \tag{30}$$

Obtaining simple expressions for the CCDF $Q_{\chi^2}(y;n,0)$ is not easy, except in some particular cases, namely small n. In the sequel, we thus consider the case $R=1$, i.e., the SOI belongs to a one-dimensional subspace. Then we have [2]

$$Q_{\chi^2}(y;3,0) = 2Q_G\left(\sqrt{y}\right) + \sqrt{\frac{2}{\pi}}y^{1/2}e^{-y/2}$$
$$\simeq \sqrt{\frac{2}{\pi}}y^{-1/2}\left[1+y\right]e^{-y/2}$$

where $Q_G(y)$ stands for the ccdf of a zero-mean, unit variance Gaussian distributed random variable, and the approximation holds for large y. Therefore, for $R=1$ and high SNR,

$$G_{P_{fa}} \simeq \frac{5b}{\ln 10}\left[1+\frac{N(b-1)^2}{b^2}\right] + 5\log_{10}\frac{\pi}{2} - 5\log_{10}\left[b\left(3+\frac{N(b-1)^2}{b^2}+2\mathrm{SNR}_p\right)\right]. \tag{31}$$

Considering this simpler expression of the P_{fa} ratio, we can first evaluate the performance loss when there is no power variation $(b=1)$:

$$G_{pfa}|_{b=1} \simeq \frac{5}{\ln 10} + 5\log_{10}\frac{\pi}{2} - 5\log_{10}\left[3+2\mathrm{SNR}_p\right]. \tag{32}$$

One can notice that, for large N, this loss only depends on the post processing signal to noise ratio. Moreover from expression (31), one can derive two approximated expressions of the noise power variation, b_{0+} and b_{0-}, needed to have the same performance for T_r and $T_{\rm kn}$. That is to say, if $b > b_{0+}$ or $b < b_{0-}$, one should use T_r instead of the classical test $T_{\rm kn}$ with

$$b_{0+} \simeq 1 + \left[N^{-1}\left(\ln\left[\frac{2}{\pi}(3+2\mathrm{SNR}_p)\right]-1\right)\right]^{1/2} \tag{33a}$$
$$b_{0-} \simeq 1 - \left[N^{-1}\left(\ln\left[\frac{2}{\pi}(3+2\mathrm{SNR}_p)\right]-1\right)\right]^{1/2}. \tag{33b}$$

We can remark that in contrast to (32), these threshold values for b depend on N.

IV. Numerical Illustrations

In this section, we compare the performances of T_r with those of $T_{\rm kn}$, $T_{\rm un}$, and T_c for different noise power variations. Then, we show that our new detector can also be used to mitigate signal modelling errors through a convincing example. We also compare it to the robust matched detector introduced in [14]. In all simulations, we consider the case of a single complex tone $\left(x(k) = e^{2i\pi fk} + n(k) \text{ for } k=1,\ldots,N\right)$ with $f = 0.15$. First we take $N = 1024$ points, and fix the noise power under H_0 $\left(\sigma_0^2\right)$ so that $\mathrm{SNR}_0 = 10\log\left(P_s/\sigma_0^2\right) = -23$ dB. Figs. 2–4 show the receiver

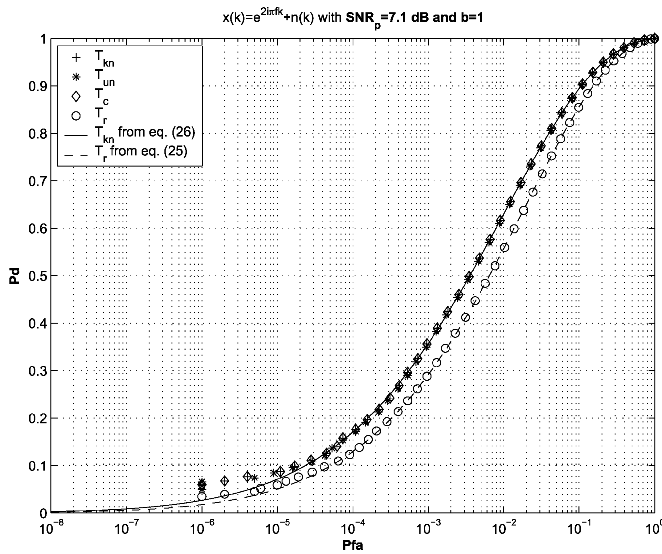

Fig. 2. Detectors comparison for $b = 1$.

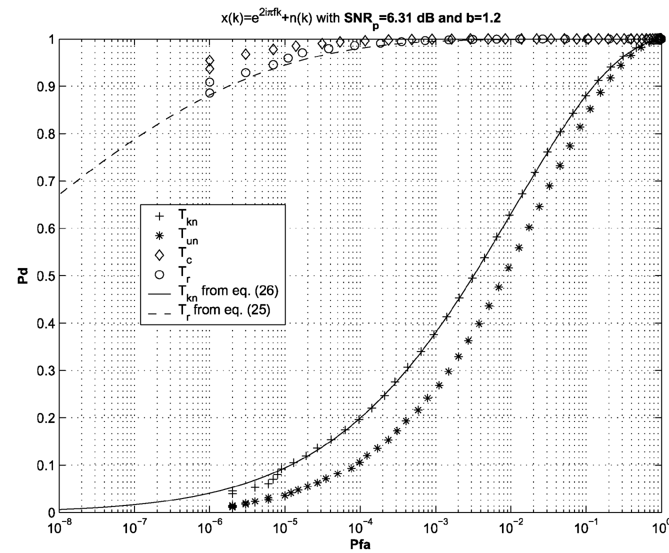

Fig. 4. Detectors comparison for $b = 1.2$.

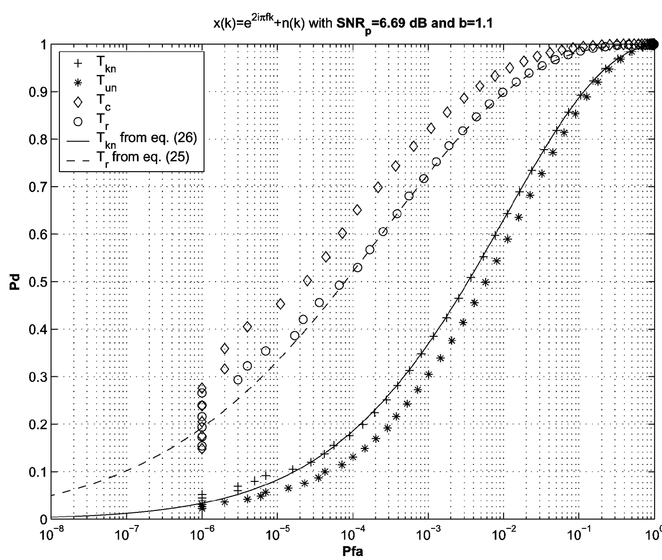

Fig. 3. Detectors comparison for $b = 1.1$.

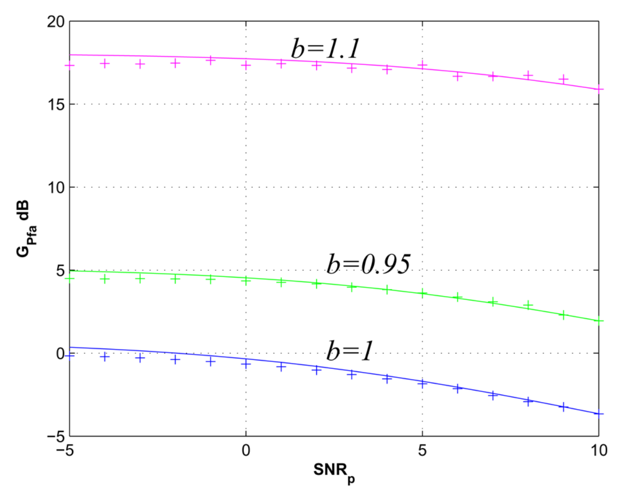

Fig. 5. Performance gain/loss for different values of b. Solid lines refer to (31), markers to the real G_{Pfa}.

operating characteristics resulting from 10^6 Monte Carlo trials for the four detectors and for three different values of b ($b = 1$, $b = 1.1$, $b = 1.2$). At the same time, we have plotted the approximated analytical performance given by (25) and (26) for T_r and T_{kn} (dashed line for T_r and continuous for T_{kn}). We can first notice that the closed-form expressions (25) and (26) give a very precise approximation of the real tests performances. This validates the different assumptions made in Section II. Then, we can see that T_r is slightly poorer than the other detectors in the case where there is no noise power variation ($b = 1$, Fig. 2) as we had seen in (32). This is somehow the price to be paid when using a more robust detector. The maximum loss is about $\Delta Pd = 0.06$ or $G_{Pfa} = 2.5$ dB in this case and will be studied more precisely in Fig. 5. This loss is compensated as soon as b increases or decreases (a variation of 4% is actually sufficient) and the gain provided by T_r could reach some false alarm decades in case of 10% increase (Figs. 3 and 4).

The P_{fa} loss when one uses T_r instead of T_{kn} is studied in Fig. 5 when there is no noise power variation -$b = 1$- and when there is power variation, $b = 0.95$ and $b = 1.1$. The lines represent this loss calculated with (31). These approximated values are also compared with the loss evaluated by Monte Carlo simulations (represented with markers). First, observe that equation (31) gives a good approximation of the loss for all values of SNR_p and b. Moreover, for $b = 1$, this loss is only a few decibels even for very good SNR corresponding to only some hundredth on the Pd loss. In contrast, for $b = 0.95$ or $b = 1.1$, one can observe gains in terms of P_{fa} on the order of 5 to 15 dB, which is considerable for a 10% only noise power variation.

The threshold values of b needed to have a gain using T_r is plotted on Fig. 6. One can see that for classical SNR_p, only a few percent of noise power increase is necessary for T_r to be better than T_{kn}. We have also plotted the real P_{fa} gain estimated by Monte Carlo simulations to verify that this is close to unity. Indeed, the difference in P_{fa} between the two detectors is within 1 dB for all N and SNR_p, which proves that the simple expressions for b_{0+} and b_{0-} are quite accurate, and thus assesses the validity of the closed-form expression for this threshold.

As we have seen in Fig. 1, T_r differs from the conventional detector T_{kn} in adding a corrective term based on the energy of the projection on the noise subspace. Then, one can expect a good behavior of T_r each time this projection moves away from its expected value. Hence, T_r could also be robust to a very large class of signal model mismatches.

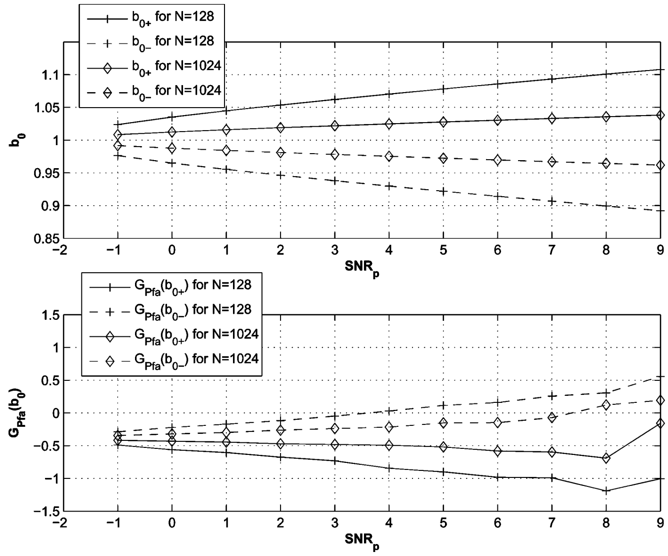

Fig. 6. Threshold on the noise power variation to have a benefit.

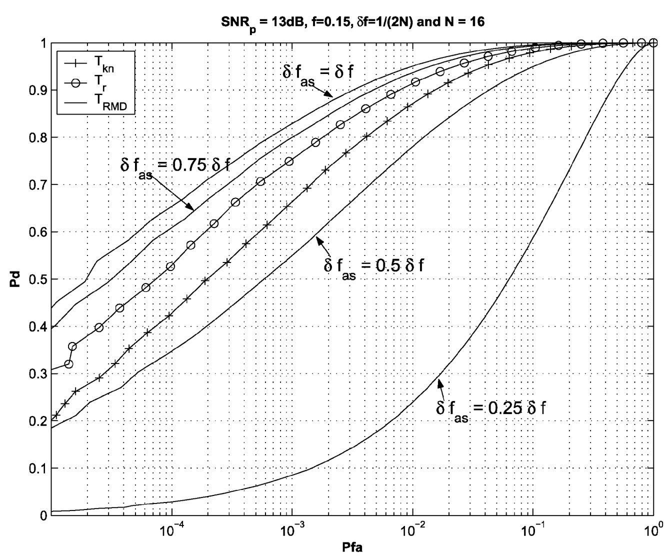

Fig. 7. Robustness against frequency mismatch.

To illustrate this behavior, we still consider the case of the detection of a complex tone when the frequency is not perfectly known. We consider $N = 16$, $\mathrm{SNR}_p = 13$ dB and an actual frequency of $f = 0.15$ as the assumed one is $f + \delta f = 0.15 + (1/2N)$. We can notice that this frequency mismatch corresponds to the maximum intrinsic error made by a classical discrete Fourier algorithm. We compare the performances of T_r with T_{kn} and with the robust matched detector (RMD), T_{RMD} introduced in [14]. This last detector has to know the *a priori* signal mismatch power coefficient noted α in [14]. Then, we have plotted different results of the RMD depending on the assumed frequency mismatch, δf_{as} (from the real δf to $(\delta f/4)$) on Fig. 7. We can first notice that T_r gives better results than the classical T_{kn} even if it has not been designed for such a model. Then, as stated in [14], the RMD maintains a good performance as δf_{as} is roughly well estimated and outperforms T_r in this case. One can notice, however, that the model at hand is closer to the one used to design the RMD than our T_r. In the case where the frequency mismatch is strongly underestimated, one should prefer T_r than T_{RMD}.

V. Conclusion

In this paper, we have extended the formulation of the matched subspace detectors (with noise power known or not) to the case where the noise power is only known under the null hypothesis. We derived the GLRT for the problem at hand and carried out a performance analysis of this new detector, in order to compare it to the classical MSD. From this analysis, we first draw an explanation of the performance improvement for large N. Then, we analyzed more precisely the case of a one-dimensional subspace SOI to determine the conditions under which this new detector should be used. Numerical simulations attest to the validity of the theoretical analysis and show that this new detector could outperform the classical one. Moreover, we compare it to the RMD and show it performs well even in cases quite far from the model used to design it. Compared to other robust schemes, the new detector does not need any assumption about the *a priori* mismatch amplitude.

References

[1] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.

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

[3] P. J. Huber, "A robust version of the probability ratio test," *Ann. Math. Stat.*, vol. 36, no. 4, pp. 1753–1758, Dec. 1965.

[4] S. Kassam and H. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, no. 3, pp. 433–481, Mar. 1985.

[5] F. Hampel, E. Ronchetti, P. Rousseeuw, and W. Stahel, *Robust Statistics—The Approach Based on Influence Functions*. New York: Wiley, 1986.

[6] I. Song and S. A. Kassam, "Locally optimum detection of signals in a generalized observation model: The known signal case," *IEEE Trans. Inf. Theory*, vol. 36, no. 3, pp. 502–515, May 1990.

[7] H. Poor, "Robustness in signal detection," in *Communications and Networks: A Survey of Recent Advances*, I. Blake and H. Poor, Eds. New York: Springer-Verlag, 1986, ch. 6.

[8] G. V. Moustakides, "Robust detection of signals: A large deviations approach," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 6, pp. 822–825, Nov. 1985.

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

[10] H. F. MacLean, "CCD noise removal in digital images," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2676–2685, Oct. 2003.

[11] L. C. Barbosa, "A model for magnetic recording channels with signal dependent noise," *IEEE Trans. Magn.*, vol. 31, no. 2, pp. 1062–1064, Mar. 1995.

[12] N. Harle and J. Bohme, "Detection of knocking for spark ignition engines based on structural vibrations," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Dallas, Tx, Apr. 1987, vol. 12, pp. 1744–1747.

[13] M. Zadnik, F. Vincent, R. Vingerhoeds, and F. Galtier, "Si engine knock detection method robust to resonance frequency changes," presented at the 8th Int. Conf. Engines For Automobile, Naples, Italy, Sep. 16–20, 2007.

[14] J.-J. Fuchs, "A robust matched detector," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5133–5142, Nov. 2007.

[15] J.-J. Fuchs, "Matched detector and estimator with signature uncertainty," presented at the 41st Asilomar Conf., Pacific Grove, CA, Nov. 4–7, 2007.

[16] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[17] A. Castano-Martínez and F. López-Blázquez, "Distribution of a sum of weighted noncentral chi-square variables," *Test*, vol. 14, no. 2, pp. 397–415, 2005.