ROBUST DETECTION OF NONSTATIONARY RANDOM SIGNALS BELONGING TO *P*-POINT UNCERTAINTY CLASSES*

Gerald Matz and Alfred Raidl

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Wien, Austria
phone: +43 1 58801 38916, fax: +43 1 58801 38999, email: g.matz@ieee.org
web: www.nt.tuwien.ac.at/dspgroup/gmatz.html

ABSTRACT

We present two robust detectors for nonstationary random signals that belong to *p*-point uncertainty classes, one being based on an estimator-correlator approach, the other using the deflection criterion. Apart of stable performance, these robust detectors have the advantage of requiring only reduced prior knowledge. Using local cosine bases, we provide an intuitive and highly efficient time-frequency implementation of these robust detectors along with an extension that permits signal-adaptive operation. Simulation results illustrate the robustness of the proposed detectors.

1. INTRODUCTION

Problem Formulation. We consider the detection a nonstationary random signal[1] $s(t)$ that is corrupted by additive nonstationary Gaussian noise $n(t)$. Signal and noise are assumed uncorrelated. This problem corresponds to the hypothesis test $\mathcal{H}_0: x(t) = n(t)$ versus $\mathcal{H}_1: x(t) = s(t) + n(t)$, with $x(t)$ being the observed signal.

If $s(t)$ is also Gaussian, the *likelihood ratio* test statistic (which is optimal in the Neyman-Pearson and Bayesian sense) equals [1]

$$T_{\rm LR}(x) = \langle \mathbf{H}_{\rm LR} x, x \rangle, \qquad \text{with } \mathbf{H}_{\rm LR} = \mathbf{R}_n^{-1}\mathbf{R}_s(\mathbf{R}_s+\mathbf{R}_n)^{-1}. \quad (1)$$

$\mathbf{R}_s$ and $\mathbf{R}_n$ denote the correlation operators[2] of signal and noise.

If the distribution of $s(t)$ is unknown, the likelihood ratio cannot be determined. Here, the *deflection* d^2 of a test statistic $T(x)$ is a reasonable alternative performance criterion. It is defined as [1]

$$d^2 \triangleq \frac{\big(\mathbb{E}_1\{T(x)\} - \mathbb{E}_0\{T(x)\}\big)^2}{\mathrm{var}_0\{T(x)\}},$$

where $\mathbb{E}_i\{\cdot\}$ and $\mathrm{var}_i\{\cdot\}$ respectively denote expectation and variance conditioned on $\mathcal{H}_i$. For real-valued quadratic test statistics $T(x) = \langle \mathbf{H}x, x\rangle$ induced by a linear self-adjoint [2] operator (linear time-varying filter) $\mathbf{H}$, the deflection can be shown to equal [3]

$$d^2(\mathbf{H};\mathbf{R}_s,\mathbf{R}_n) = \frac{\mathrm{tr}^2\{\mathbf{H}\mathbf{R}_s\}}{\mathrm{tr}\{(\mathbf{H}\mathbf{R}_n)^2\}}. \quad (2)$$

This expression is maximized by the detection filter [3]

$$\mathbf{H}_{\rm D} \triangleq \arg\max_{\mathbf{H}} d^2(\mathbf{H};\mathbf{R}_s,\mathbf{R}_n) = \mathbf{R}_n^{-1}\mathbf{R}_s\mathbf{R}_n^{-1}, \quad (3)$$

with $d^2_{\max}(\mathbf{R}_s,\mathbf{R}_n) \triangleq d^2(\mathbf{H}_{\rm D};\mathbf{R}_s,\mathbf{R}_n) = \mathrm{tr}\{(\mathbf{R}_s\mathbf{R}_n^{-1})^2\}$. Note that at low SNR, $\mathbf{H}_{\rm D} \approx \mathbf{H}_{\rm LR}$. The final decisions are obtained by comparing the test statistics $T_{\rm LR}(x)$ and $T_{\rm D}(x) = \langle \mathbf{H}_{\rm D}x, x\rangle$ to a threshold.

In practice, the correlations $\mathbf{R}_s$ and $\mathbf{R}_n$ required for the calculation of $\mathbf{H}_{\rm D}$ and $\mathbf{H}_{\rm LR}$ are rarely exactly known (usually, they have to be estimated and/or modeled using physical reasoning). Unfortunately, deviations of the assumed (nominal) correlations from the true correlations can result in a dramatic performance loss.

p-point Uncertainty Classes. Instead of requiring that the correlations $\mathbf{R}_s$, $\mathbf{R}_n$ be exactly known, a reduced-detail prior knowledge can be modeled using *p-point uncertainty classes* [4, 5]. Consider a partition of $L_2(\mathbb{R})$ into mutually orthogonal subspaces $\mathcal{X}_i$, $i = 1,\ldots,N$, i.e., $\bigoplus_{i=1}^N \mathcal{X}_i = L_2(\mathbb{R})$ and $\mathcal{X}_i \perp \mathcal{X}_j$ for $i \neq j$. Then, the signal and noise uncertainty classes are given by

$$\begin{aligned} \mathcal{S} &= \big\{\mathbf{R}_s:\ \mathrm{tr}\{\mathbf{P}_i\mathbf{R}_s\} = s_i, \quad i = 1,\ldots,N\big\}, \\ \mathcal{N} &= \big\{\mathbf{R}_n:\ \mathrm{tr}\{\mathbf{P}_i\mathbf{R}_n\} = n_i, \quad i = 1,\ldots,N\big\}, \end{aligned} \quad (4)$$

where $\mathbf{P}_i$ denotes the orthogonal projection operator on $\mathcal{X}_i$ and s_i, n_i are fixed positive numbers assumed known. Since $\mathrm{tr}\{\mathbf{P}_i\mathbf{R}_x\} = \mathbb{E}\{\|\mathbf{P}_i x\|^2\}$, these classes contain all signal and noise processes with average subspace energies s_i and n_i, respectively, within $\mathcal{X}_i$.

Contributions. In this paper, we develop robust detectors for nonstationary processes that belong to p-point uncertainty classes:

- we propose a robust estimator-correlator detector that replaces the Wiener filter with a minimax robust Wiener filter (Section 2);
- we derive a maximin robust detector based on the deflection criterion (Section 3);
- we provide an efficient time-frequency (TF) implementation of both robust detectors (Section 4);
- we extend the robust detectors to allow for signal-adaptive operation without requiring any prior knowledge (Section 4).

Simulation results illustrate the performance of our detectors (Section 5). The proposed robust detection schemes have the advantage of i) having (almost) constant performance within the uncertainty classes; ii) requiring little prior knowledge that can be estimated reliably; and iii) allowing for highly efficient (signal-adaptive) online implementation. Our results are related to previous work on minimax robust detection and estimation in [4, 6–8].

2. ROBUST ESTIMATOR-CORRELATOR DETECTOR

The formulation of a detector that has robust (stable) performance within the p-point uncertainty classes $\mathcal{S}$, $\mathcal{N}$ in (4) is difficult in general since the usual performance criteria (detection and false alarm probability) depend highly nonlinearly on the correlations $\mathbf{R}_s$ and $\mathbf{R}_n$. We thus first consider an *ad hoc* approach that is motivated by the estimator-correlator interpretation of $T_{\rm LR}(x)$ in (1) and by *minimax robust nonstationary Wiener filters* [4].

Assume that signal and noise belong to the uncertainty classes $\mathcal{S}$, $\mathcal{N}$ in (4), with known subspace energies s_i, n_i. The minimax robust Wiener filter optimizes the worst-case MSE within $\mathcal{S}$, $\mathcal{N}$,

$$\mathbf{H}_{\rm W}^{\rm R} \triangleq \arg\min_{\mathbf{H}} \max_{\substack{\mathbf{R}_s\in\mathcal{S} \\ \mathbf{R}_n\in\mathcal{N}}} \mathbb{E}\{\|s - \mathbf{H}x\|^2\}. \quad (5)$$

The solution for the robust Wiener filter was obtained in [4] as

$$\mathbf{H}_{\rm W}^{\rm R} = \sum_{i=1}^N \frac{s_i}{s_i+n_i}\mathbf{P}_i. \quad (6)$$

It has the advantage of achieving the same MSE $\mathbb{E}\{\|s - \mathbf{H}_{\rm W}^{\rm R}x\|^2\} = \sum_{i=1}^N \frac{s_i n_i}{s_i+n_i}$ for all processes within the uncertainty classes $\mathcal{S}$, $\mathcal{N}$.

*Funding by FWF grant P15156.

[1] All random processes are assumed to be zero-mean.

[2] The correlation operator $\mathbf{R}_x$ of a random process $x(t)$ is the linear operator whose kernel equals the correlation function $r_x(t,t') = \mathbb{E}\{x(t)x^*(t')\}$.

To obtain a robust detector, we start by rewriting the likelihood ratio detector (1) in its estimator-correlator form, $T_{\rm LR}(x) = \langle \mathbf{R}_n^{-1}\hat{s}, x\rangle$. Here, $\hat{s}(t) = (\mathbf{H}_{\rm W}x)(t)$ is the MMSE estimate of $s(t)$ obtained with the Wiener filter $\mathbf{H}_{\rm W} = \mathbf{R}_s(\mathbf{R}_s+\mathbf{R}_n)^{-1}$ [1]. Thus, $T_{\rm LR}(x)$ is obtained by correlating the observation $x(t)$ with the signal estimate $\hat{s}(t)$ (with $\mathbf{R}_n^{-1}$ as weight). This estimator-correlator structure suggests to replace the Wiener filter $\mathbf{H}_{\rm W}$ with the robust Wiener filter $\mathbf{H}_{\rm W}^{\rm R}$ in (6) to "robustify" $T_{\rm LR}(x)$. Unfortunately, the test statistic additionally involves the unknown noise correlation. However, (6) suggests that robustness is achieved by replacing the operator algebra involving $\mathbf{R}_s$, $\mathbf{R}_n$ with the scalar algebra involving s_i, n_i. Motivated by this observation, we propose the test statistic

$$T_{\rm LR}^{\rm R}(x) = \langle \mathbf{H}_{\rm LR}^{\rm R}x, x\rangle, \qquad \text{with } \mathbf{H}_{\rm LR}^{\rm R} \triangleq \sum_{i=1}^{N} \frac{s_i}{n_i(s_i+n_i)}\,\mathbf{P}_i\,.$$

We will refer to $T_{\rm LR}^{\rm R}(x)$ as *robust estimator-correlator detector*. It can alternatively be written as

$$T_{\rm LR}^{\rm R}(x) = \sum_{i=1}^{N} \frac{s_i}{n_i(s_i+n_i)}\,x_i\,, \qquad \text{with } x_i \triangleq \langle \mathbf{P}_i x, x\rangle = \|\mathbf{P}_i x\|^2. \quad (7)$$

Simulation results indicated that $T_{\rm LR}^{\rm R}(x)$ is indeed more robust than $T_{\rm LR}(x)$ (cf. Section 5). A property supporting this claim is

$$\mathbb{E}_1\{T_{\rm LR}^{\rm R}(x)\} - \mathbb{E}_0\{T_{\rm LR}^{\rm R}(x)\} = \sum_{i=1}^{N} \frac{s_i^2}{n_i(s_i+n_i)}\,, \qquad (8)$$

valid for $\mathbf{R}_s \in \mathcal{S}$, $\mathbf{R}_n \in \mathcal{N}$. In the sense of (8), $T_{\rm LR}^{\rm R}(x)$ achieves constant separation of the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$.

3. MAXIMIN ROBUST DEFLECTION DETECTOR

We next consider an alternative to the robust estimator-correlator detector that is based on the deflection criterion in (2) [6, 7]. In a sense, our results are nonstationary extensions of [8]. For technical reasons, we introduce the modified p-point noise uncertainty class

$$\mathcal{N}' = \{\mathbf{R}_n\colon \operatorname{tr}\{\mathbf{P}_i\mathbf{R}_n\mathbf{P}_j\mathbf{R}_n\} = n_i^2\delta_{ij}\,, \quad i,j = 1,\dots,N\}, \qquad (9)$$

with prescribed numbers n_i^2. Note that $\operatorname{tr}\{\mathbf{P}_i\mathbf{R}_n\mathbf{P}_j\mathbf{R}_n\} = \|\mathbf{P}_i\mathbf{R}_n\mathbf{P}_j\|^2$ and $\mathbf{P}_i\mathbf{R}_n\mathbf{P}_j = \mathbf{R}_{\mathbf{P}_i n,\mathbf{P}_j n}$ is the cross-correlation operator of the noise subspace components $(\mathbf{P}_i n)(t)$ and $(\mathbf{P}_j n)(t)$. Thus, (9) means that different noise subspace components are uncorrelated and the energy (squared Hilbert-Schmidt norm [2]) of the correlation operator $\mathbf{R}_{\mathbf{P}_i n} = \mathbf{P}_i\mathbf{R}_n\mathbf{P}_i$ of the noise subspace component $(\mathbf{P}_i n)(t)$ equals n_i^2.

Analogous to the minimax robust Wiener filter (5), the *maximin robust deflection detector* is defined as $T_{\rm D}^{\rm R}(x) = \langle \mathbf{H}_{\rm D}^{\rm R}x, x\rangle$ with [6, 7]

$$\mathbf{H}_{\rm D}^{\rm R} \triangleq \arg\max_{\mathbf{H}} \min_{\substack{\mathbf{R}_s\in\mathcal{S}\\ \mathbf{R}_n\in\mathcal{N}'}} d^2(\mathbf{H};\mathbf{R}_s,\mathbf{R}_n)\,. \qquad (10)$$

Thus, $\mathbf{H}_{\rm D}^{\rm R}$ maximizes the worst-case deflection within $\mathcal{S}$, $\mathcal{N}'$. The usual approach of finding *least favorable* correlations that solve the dual problem (i.e., in our case minimize $d^2_{\max}(\mathbf{R}_s,\mathbf{R}_n)$) and designing the optimal filter for the resulting least favorable priors [4, 6] is not feasable since $d^2(\mathbf{H};\mathbf{R}_s,\mathbf{R}_n)$ is no convex cost function [6, 7].

Alternatively, one can try to find a detection filter $\mathbf{H}_{\rm D}^{\rm R}$ and correlations $\mathbf{R}_s^{\rm R}$, $\mathbf{R}_n^{\rm R}$ that solve the saddle point problem

$$d^2(\mathbf{H};\mathbf{R}_s^{\rm R},\mathbf{R}_n^{\rm R}) \le d^2(\mathbf{H}_{\rm D}^{\rm R};\mathbf{R}_s^{\rm R},\mathbf{R}_n^{\rm R}) \le d^2(\mathbf{H}_{\rm D}^{\rm R};\mathbf{R}_s,\mathbf{R}_n), \qquad (11)$$

where the inequalities are to hold for all self-adjoint $\mathbf{H}$ and all $\mathbf{R}_s \in \mathcal{S}$, $\mathbf{R}_n \in \mathcal{N}'$. The left-hand inequality is satisfied by chosing $\mathbf{H}_{\rm D}^{\rm R} = (\mathbf{R}_n^{\rm R})^{-1}\mathbf{R}_s^{\rm R}(\mathbf{R}_n^{\rm R})^{-1}$ (cf. (3)). The right-hand inequality implies that when using $\mathbf{H}_{\rm D}^{\rm R}$, worst performance is attained for $\mathbf{R}_s^{\rm R}$, $\mathbf{R}_n^{\rm R}$, i.e., $\mathbf{H}_{\rm D}^{\rm R}$ performs better for all correlations $\mathbf{R}_s \in \mathcal{S}$, $\mathbf{R}_n \in \mathcal{N}'$ than for those it was designed with. Eq. (11) thus implies that $\mathbf{H}_{\rm D}^{\rm R}$ is the minimax robust detection filter in (10) [6, 7]. We caution the reader, however, that a robust detection filter obtained according to (10) does not necessarily satisfy the saddle point condition (11).

It remains to find correlations $\mathbf{R}_s^{\rm R}$, $\mathbf{R}_n^{\rm R}$ such that the right-hand inequality in (11) is satisfied with $\mathbf{H}_{\rm D}^{\rm R} = (\mathbf{R}_n^{\rm R})^{-1}\mathbf{R}_s^{\rm R}(\mathbf{R}_n^{\rm R})^{-1}$. This is done in the appendix, leading to the following result:

Theorem 1. *The robust detection filter defined by* (10) *is given by*

$$\mathbf{H}_{\rm D}^{\rm R} = \sum_{i=1}^{N} \frac{s_i}{n_i^2}\,\mathbf{P}_i.$$

For all $\mathbf{R}_s \in \mathcal{S}$, $\mathbf{R}_n \in \mathcal{N}'$, *it achieves the same deflection,*

$$d^2(\mathbf{H}_{\rm D}^{R};\mathbf{R}_s,\mathbf{R}_n) = \sum_{i=1}^{N} \frac{s_i^2}{n_i^2}.$$

With x_i as in (7), the corresponding robust detection statistic is

$$T_{\rm D}^{\rm R}(x) = \sum_{i=1}^{N} \frac{s_i}{n_i^2}\,x_i\,. \qquad (12)$$

Discussion. The test statistics of the robust estimator-correlator detector (7) and the robust deflection detector (12) have a similar structure: they are weighted averages of the observation's subspace energies x_i. Furthermore, both detection filters $\mathbf{H}_{\rm D}^{\rm R}$ and $\mathbf{H}_{\rm LR}^{\rm R}$ can be written as $\mathbf{H} = \sum_{i=1}^{N} h_i\,\mathbf{P}_i$ (leading to the same efficient implementation in Section (4)) with $h_i^{\rm D} = \frac{s_i}{n_i^2}$ and $h_i^{\rm LR} = \frac{s_i}{n_i(s_i+n_i)}$. In accordance with the prior knowledge specified by the uncertainty classes, both detectors treat signal components within the same subspace $\mathcal{X}_i$ alike and ignore correlations of signal components within different subspaces. Finally, it is seen that in the robust detection filters, the operator algebra involving $\mathbf{R}_s$, $\mathbf{R}_n$ is replaced by a simple scalar algebra involving the subspace energies s_i, n_i. This is advantageous with regard to the computational complexity and numerical stability of practical implementations.

4. TF IMPLEMENTATION OF ROBUST DETECTORS

We next present a physically intuitive choice of the subspace partition underlying the uncertainty classes $\mathcal{S}$, $\mathcal{N}$, and $\mathcal{N}'$. This subspace partition is based on *local cosine bases* (LCB) [9, 10].

LCB Subspaces. We choose[3] $\mathcal{X}_{k,l} = \operatorname{span}\{u_{k,l}^{(1)}(t),\dots,u_{k,l}^{(M)}(t)\}$, $k \in \mathbb{Z}$, $l \in \mathbb{N}_0$, with the LCB functions

$$u_{k,l}^{(m)}(t) \triangleq w_k(t)\sqrt{\frac{2}{T_k}}\cos\left(\frac{2(lM+m)-1}{2T_k}\pi(t-t_k)\right).$$

Here, the t_k define a partition of the time axis into disjoint intervals $[t_k, t_{k+1}]$ of duration $T_k = t_{k+1}-t_k$. Furthermore, $w_k(t)$ is a window associated to the kth interval $[t_k, t_{k+1}]$ (for details see [9]). Thus, any partition $\{t_k\}_{k\in\mathbb{Z}}$ of the time axis corresponds to a partition $\{\mathcal{X}_{k,l}\}_{k\in\mathbb{Z},l\in\mathbb{N}_0}$ into M-dimensional subspaces. Since the subspace $\mathcal{X}_{k,l}$ can be shown [10] to be supported within the TF region $[t_k, t_k+T_k]\times[lMF_k,(l+1)MF_k]$ of area M, the subspace partition $\{\mathcal{X}_{k,l}\}_{k\in\mathbb{Z},l\in\mathbb{N}_0}$ corresponds to a rectangular tiling of the TF plane.

Efficient Implementation. With[4] $\mathbf{P}_{k,l} = \sum_{m=1}^{M} u_{k,l}^{(m)} \otimes u_{k,l}^{(m)*}$ denoting the orthogonal projection operators on $\mathcal{X}_{k,l}$ and $s_{k,l}$, $n_{k,l}$ being specified by the uncertainty classes, the robust detectors can directly be implemented. However, the LCB subspaces allow for a much more efficient online implementation. In particular, the LCB coefficients $\langle x, u_{k,l}^{(m)}\rangle$ entering the subspace energies according to

$$x_{k,l} = \langle \mathbf{P}_{k,l}x, x\rangle = \sum_{m=1}^{M} |\langle x, u_{k,l}^{(m)}\rangle|^2 \qquad (13)$$

[3] In the following, the index i is replaced by the double index k,l.

[4] The rank one operator $x\otimes y^*$ has the outer product $x(t)y^*(t')$ as kernel.

can be computed efficiently using fast DCT based algorithms [9].

Let us define the partial test statistic at time $t_{k'}$ as $T_{k'}(x) = \sum_{k=-\infty}^{k'} \sum_{l=1}^{L} h_{k,l}\, x_{k,l}$ (with $h_{k,l}^{\mathrm{D}} = \frac{s_{k,l}}{n_{k,l}^2}$ and $h_{k,l}^{\mathrm{LR}} = \frac{s_{k,l}}{n_{k,l}(s_{k,l}+n_{k,l})}$). For each block (interval) $[t_k, t_{k+1}]$, the computation of the robust test statistic can be summarized as follows:

1. compute the LCB coefficients $\langle x, u_{k,l}^{(m)} \rangle$, $l \in \mathbb{N}_0$, $m = 1, \ldots, M$;
2. determine the subspace energies $x_{k,l}$, $l \in \mathbb{N}_0$, according to (13);
3. compute the weighted sum $\Delta T_k = \sum_{l=0}^{\infty} h_{k,l}\, x_{k,l}$;
4. update the test statistic as $T_k(x) = T_{k-1}(x) + \Delta T_k$.

This procedure is repeated over the time interval of interest. Note that only the current data block needs to be stored and processed by this algorithm. In a discrete-time setting (i.e., after appropriate sampling) the computational complexity of the update of the test statistic in the kth interval can be shown to be $O(L_k \log L_k)$ with L_k denoting the number of signal samples in the kth block (interval).

Signal-Adaptive Implementation. We finally consider a signal-adaptive implementation of our robust detectors that incorporates the estimation of the prior knowledge (i.e., of $s_{k,l}$, $n_{k,l}$). For simplicity, we restrict to the case of stationary white noise, $\mathbf{R}_n = N_0\, \mathbf{I}$, with known N_0. More general situations can be treated analogous to [10]. Our goal is now to augment the previous LCB based implementation by a reliable estimation of the subspace energies $s_{k,l}$ and by that the estimation of the detection filter coefficients $h_{k,l}$.

Let us assume that $\mathcal{H}_1$ is in force, i.e., $x(t) = s(t) + n(t)$. In that case, $\mathbb{E}\{x_{k,l}\} = s_{k,l} + n_{k,l}$ with $n_{k,l} = N_0 M$. This suggests $x_{k,l} - N_0 M$ as en estimate of $s_{k,l}$. Since $s_{k,l} \geq 0$, a better estimate is obviously given by

$$\hat{s}_{k,l} = \max\{0, x_{k,l} - N_0 M\}. \qquad (14)$$

Similar to [10], it can be shown that the variance of this estimate typically decreases inversely proportional with the subspace dimension M. Thus, the estimation of these subspace energies typically is statistically much more stable than, e.g., the estimation of the correlation operator $\mathbf{R}_s$ required for the nominally optimum detectors.

If $\mathcal{H}_0$ is in force, i.e., $x(t) = n(t)$, (14) will lead to a systematic under-estimation of $s_{k,l}$. However, this will result in a test statistic that is even smaller than with the ideal $s_{k,l}$, and thus will typically not affect performance adversely (i.e., the false alarm probability will not increase noticeably). Using the estimate (14) to calculate the filter coefficients $h_{k,l}$ results in

$$\hat{h}_{k,l}^{\mathrm{D}} = \frac{\max\{0, x_{k,l} - N_0 M\}}{N_0^2 M} = \max\left\{0, \frac{x_{k,l}}{N_0^2 M} - \frac{1}{N_0}\right\}$$

$$\hat{h}_{k,l}^{\mathrm{LR}} = \frac{\max\{0, x_{k,l} - N_0 M\}}{N_0 M\, x_{k,l}} = \max\left\{0, \frac{1}{N_0 M} - \frac{1}{x_{k,l}}\right\}$$

With these filter coefficients, the robust test statistics equal

$$T_{\mathrm{D}}^{\mathrm{R}}(x) = \sum_{k=-\infty}^{\infty} \sum_{l=0}^{\infty} f_{\mathrm{D}}(x_{k,l}), \qquad T_{\mathrm{LR}}^{\mathrm{R}}(x) = \sum_{k=-\infty}^{\infty} \sum_{l=0}^{\infty} f_{\mathrm{LR}}(x_{k,l}),$$

where $f_{\mathrm{D}}(x_{k,l}) = h_{k,l}^{\mathrm{D}} x_{k,l} = \frac{1}{N_0} \max\left\{0, \frac{x_{k,l}^2}{N_0 M} - x_{k,l}\right\}$ and $f_{\mathrm{LR}}(x_{k,l}) = h_{k,l}^{\mathrm{LR}} x_{k,l} = \max\left\{0, \frac{x_{k,l}}{N_0 M} - 1\right\}$ are soft-thresholding functions with $f_{\cdot}(x_{k,l}) = 0$ for $x_{k,l} < N_0 M$. Thus, both detectors discard signal components with corresponding subspace energy $x_{k,l}$ below the average (i.e., expected) noise level of $N_0 M$. With regard to implementation, the update of the test statistic in step 3 is now computed as $\Delta T_k = \sum_{l=0}^{\infty} f_{\cdot}(x_{k,l})$ with essentially the same costs.

5. SIMULATION RESULTS

Experiment 1. To illustrate the performance of the proposed robust detectors, we prescribed nominal nonstationary second-order statistics $\mathbf{R}_s^0$, $\mathbf{R}_n^0$ and calculated the corresponding optimal detection filter $\mathbf{H}_{\mathrm{D}}$. The uncertainty classes $\mathcal{S}$, $\mathcal{N}'$ were constructed by using orthogonal projection operators $\mathbf{P}_i$ that correspond to regular rectangular partitions of the TF plane and choosing $s_i = \mathrm{tr}\{\mathbf{P}_i \mathbf{R}_s^0\}$, $n_i^2 = \|\mathbf{P}_i \mathbf{R}_n^0 \mathbf{P}_i\|^2$ such that $\mathbf{R}_s^0 \in \mathcal{S}$ and $\mathbf{R}_n^0 \in \mathcal{N}'$. We then computed the robust detection filter $\mathbf{H}_{\mathrm{D}}^{\mathrm{R}}$ as well as the "saddle point" correlations $\mathbf{R}_s^{\mathrm{R}}$, $\mathbf{R}_n^{\mathrm{R}}$.

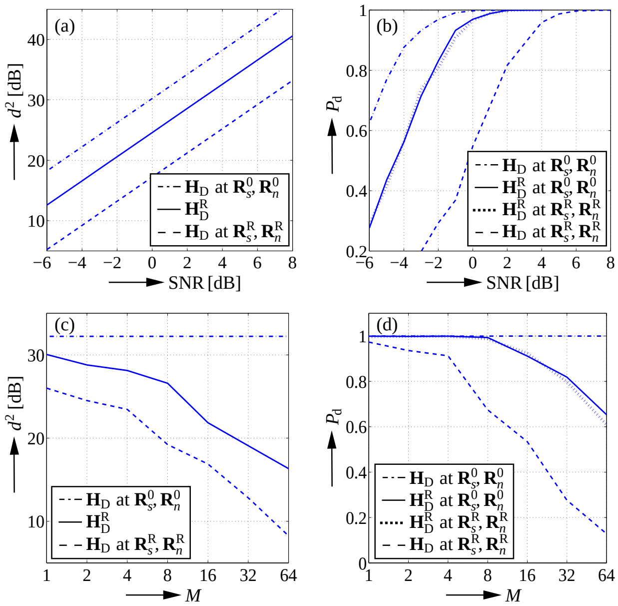

Figure 1: *(a) Deflection and (b) detection probability of nominal and robust detectors versus SNR; (c) deflection and (d) detection probability of nominal and robust detectors versus subspace dimension M. In (b) and (d), the false alarm probability was set to 0.01.*

Performance of $\mathbf{H}_{\mathrm{D}}$ and $\mathbf{H}_{\mathrm{D}}^{\mathrm{R}}$ was assessed by evaluating (2). leading to the deflection versus SNR curve in Fig. 1(a) (here, the dimension of all subspaces $\mathcal{X}_i$ was $M = 8$, SNR was varied simply by scaling $\mathbf{R}_s^0$). Obviously, the best deflection is achieved by $\mathbf{H}_{\mathrm{D}}$ if the actual correlations equal the nominal correlations. However, for the saddle point correlations, the deflection of $\mathbf{H}_{\mathrm{D}}$ is dramatically reduced (the worst-case deflection of $\mathbf{H}_{\mathrm{D}}$ within $\mathcal{S}$, $\mathcal{N}'$ might be even smaller). In contrast, the deflection achieved by the $\mathbf{H}_{\mathrm{D}}^{\mathrm{R}}$ is independent of the actual correlations and lies between $d^2(\mathbf{H}_{\mathrm{D}}; \mathbf{R}_s^0, \mathbf{R}_n^0)$ and $d^2(\mathbf{H}_{\mathrm{D}}; \mathbf{R}_s^{\mathrm{R}}, \mathbf{R}_n^{\mathrm{R}})$. This advantage of the robust detector $\mathbf{H}_{\mathrm{D}}^{\mathrm{R}}$ is even more pronounced in Fig. 1(b) that shows corresponding power curves, i.e., probability of detection P_{d} versus SNR at fixed false alarm rate of 0.01 (these curves were obtained by Monte-Carlo simulations involving 5000 normal distributed realizations). At SNR$= -2$ dB the detection probability of $\mathbf{H}_{\mathrm{D}}$ may drop from 0.97 to 0.29 while $\mathbf{H}_{\mathrm{D}}^{\mathrm{R}}$ maintains a detection probability of 0.8 (note that the detection probability of the robust detector is *not* constant within the uncertainty classes). We conclude that $\mathbf{H}_{\mathrm{D}}^{\mathrm{R}}$ achieves a significant gain over $\mathbf{H}_{\mathrm{D}}$ in case of adverse operating conditions while loosing little at nominal operating conditions.

Figs. 1(c) and (d) show the deflection and detection probability of $\mathbf{H}_{\mathrm{D}}$ and $\mathbf{H}_{\mathrm{D}}^{\mathrm{R}}$ versus the subspace dimension M for SNR=1 dB. It is seen that the performance variation of the optimum filter $\mathbf{H}_{\mathrm{D}}$ increases with increasing M. While the robust filter performs much better than $\mathbf{H}_{\mathrm{D}}$ at $\mathbf{R}_s^{\mathrm{R}}$, $\mathbf{R}_n^{\mathrm{R}}$, the gap to $\mathbf{H}_{\mathrm{D}}$ at nominal operating conditions also grows noticeably with increasing M.

We note that qualitatively similar results (not shown due to lack of space) were obtained for the estimator-correlator detectors $\mathbf{H}_{\mathrm{LR}}$ and $\mathbf{H}_{\mathrm{LR}}^{\mathrm{R}}$. However, $\mathbf{H}_{\mathrm{LR}}^{\mathrm{R}}$ always performed worse than $\mathbf{H}_{\mathrm{D}}^{\mathrm{R}}$.

Experiment 2. We next analyze the performance of the signal-adaptive versions of our robust detectors using LCB subspaces. The signal process was modeled as $s(t) = s_0(t - t_0)\, e^{j2\pi f_0 t}$. Here, $s_0(t)$

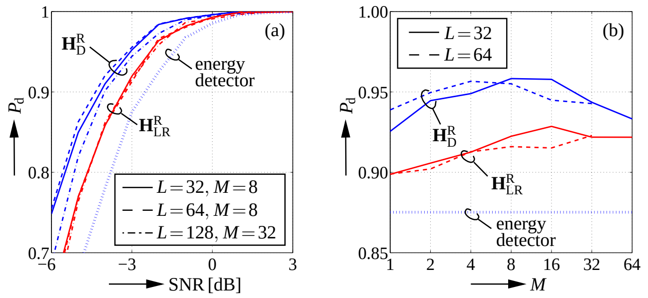

Figure 2: *Detection probability of $\mathbf{H}_\mathrm{D}^\mathrm{R}$, $\mathbf{H}_\mathrm{LR}^\mathrm{R}$, and energy detector versus (a) SNR and (b) subspace dimension M (the false alarm probability was 0.01).*

is a Gaussian random process consisting of an initial short broadband component followed by a longer narrowband component and t_0 and f_0 are unknown time and frequency offsets. The noise was Gaussian, stationary, and white. SNR was varied between -6 dB and 3 dB. The only prior knowledge available is the noise power N_0, usually necessitating the use of a simple energy detector.

We simulated our signal-adaptive robust detectors with equal block lengths (duration of LCB basis functions) of $L_k = L = 32$, 64, and 128, and subspace dimensions M between 2 and L. The power curves obtained using Monte Carlo simulations involving 4096 realizations are shown in Fig. 2(a) for various block lengths and subspace dimensions. It is seen that $\mathbf{H}_\mathrm{D}^\mathrm{R}$ outperforms $\mathbf{H}_\mathrm{LR}^\mathrm{R}$ and both robust detectors outperform the energy detector (by as much as ≈ 2 dB in the case of $\mathbf{H}_\mathrm{D}^\mathrm{R}$). Furthermore, the curves suggest that the best performance is achieved for $L = 64$ and $M = 8$. Indeed, as can be seen from Fig. 2(b), which shows detection probability P_d versus M, the optimum subspace dimension in case of $L = 64$ is $M = 8$ for $\mathbf{H}_\mathrm{D}^\mathrm{R}$ and $M = 16$ for $\mathbf{H}_\mathrm{LR}^\mathrm{R}$. The existence of an optimum subspace dimension is due to the fact that detectors with small M suffer from large estimation variance in the $\hat{s}_{k,l}$ whereas large M results in a significant loss of TF resolution.

6. CONCLUSIONS

We considered robust detection for signal and noise processes belonging to so-called p-point uncertainty classes. We proposed a robust estimator-correlator detector based on an *ad hoc* approach and we derived a maximin robust detector based on the deflection criterion. For uncertainty classes defined in terms of local cosine bases, both robust detectors allow for a computationally very efficient implementation. This implementation can be augmented to allow for signal-adaptive operation without requiring any prior knowledge. Theoretical and numerical results illustrated the robustness of our detector under a variety of operating conditions.

APPENDIX: PROOF OF THEOREM 1

We first exhibit (non-unique) correlations $\mathbf{R}_s^\mathrm{R}$, $\mathbf{R}_n^\mathrm{R}$ and a corresponding detection filter $\mathbf{H}_\mathrm{D}^\mathrm{R}$ that satisfy the saddle-point inequalities (11). Consider correlations $\mathbf{R}_s^\mathrm{R}, \mathbf{R}_n^\mathrm{R}$ constructed according to

$$\mathbf{R}_s^\mathrm{R} \triangleq \sum_{i=1}^{N} \mathbf{R}_{s,i}, \qquad \mathbf{R}_n^\mathrm{R} \triangleq \sum_{i=1}^{N} \mathbf{R}_{n,i}$$

where $\mathbf{R}_{s,i}$, $\mathbf{R}_{n,i}$ are positive semi-definite operators with range $\mathcal{X}_i$ that satisfy $n_i^2 \mathbf{R}_{s,i} = s_i \mathbf{R}_{n,i}^2$ or, equivalently,

$$n_i \mathbf{R}_{s,i}^{1/2} = \sqrt{s_i}\, \mathbf{R}_{n,i}\,. \tag{15}$$

$\mathbf{R}_{s,i}$ and $\mathbf{R}_{n,i}$ can always be normalized such that $\mathbf{R}_s^\mathrm{R} \in \mathcal{S}$, $\mathbf{R}_n^\mathrm{R} \in \mathcal{N}'$.

Starting from (3), the deflection-optimal detection filter $\mathbf{H}_\mathrm{D}^\mathrm{R}$ for $\mathbf{R}_s^\mathrm{R}$, $\mathbf{R}_n^\mathrm{R}$ can be developed as (superscript # denotes pseudo-inverse)

$$\mathbf{H}_\mathrm{D}^\mathrm{R} = \big(\mathbf{R}_n^\mathrm{R}\big)^{-1}\mathbf{R}_s^\mathrm{R}\big(\mathbf{R}_n^\mathrm{R}\big)^{-1} = \left[\sum_{i=1}^{N}\mathbf{R}_{n,i}\right]^{-1}\sum_{j=1}^{N}\mathbf{R}_{s,j}\left[\sum_{k=1}^{N}\mathbf{R}_{n,k}\right]^{-1}$$

$$= \sum_{i=1}^{N}\mathbf{R}_{n,i}^{\#}\sum_{j=1}^{N}\mathbf{R}_{s,j}\sum_{k=1}^{N}\mathbf{R}_{n,k}^{\#} = \sum_{i=1}^{N}\mathbf{R}_{n,i}^{\#}\mathbf{R}_{s,i}\mathbf{R}_{n,i}^{\#}\,,$$

where we used the fact that by construction $\mathbf{R}_{n,i}^{\#}\mathbf{R}_{s,j}\mathbf{R}_{n,k}^{\#} = \mathbf{0}$ unless $i = j = k$. Now inserting (15) yields

$$\mathbf{H}_\mathrm{D}^\mathrm{R} = \sum_{i=1}^{N}\left(\frac{n_i}{\sqrt{s_i}}\mathbf{R}_{s,i}^{1/2}\right)^{\#}\mathbf{R}_{s,i}\left(\frac{n_i}{\sqrt{s_i}}\mathbf{R}_{s,i}^{1/2}\right)^{\#}$$

$$= \sum_{i=1}^{N}\frac{s_i}{n_i^2}\mathbf{R}_{s,i}^{\#/2}\mathbf{R}_{s,i}\mathbf{R}_{s,i}^{\#/2} = \sum_{i=1}^{N}\frac{s_i}{n_i^2}\mathbf{P}_i\,.$$

Note that by definition $\mathbf{H}_\mathrm{D}^\mathrm{R}$ and $\mathbf{R}_s^\mathrm{R}$, $\mathbf{R}_n^\mathrm{R}$ satisfy the left-hand inequality in (11). To prove the right-hand side inequality of (11), we next consider the deflection achieved with $\mathbf{H}_\mathrm{D}^\mathrm{R}$. For all $\mathbf{R}_s \in \mathcal{S}$, the numerator of $d^2(\mathbf{H}_\mathrm{D}^\mathrm{R};\mathbf{R}_s,\mathbf{R}_n)$ equals

$$\mathrm{tr}^2\{\mathbf{H}_\mathrm{D}^\mathrm{R}\mathbf{R}_s\} = \mathrm{tr}^2\left\{\sum_{i=1}^{N}\frac{s_i}{n_i^2}\mathbf{P}_i\mathbf{R}_s\right\} = \left(\sum_{i=1}^{N}\frac{s_i}{n_i^2}\mathrm{tr}\{\mathbf{P}_i\mathbf{R}_s\}\right)^2$$

$$= \left(\sum_{i=1}^{N}\frac{s_i}{n_i^2}\mathrm{tr}\{\mathbf{P}_i\mathbf{R}_s\}\right)^2 = \left(\sum_{i=1}^{N}\frac{s_i^2}{n_i^2}\right)^2. \tag{16}$$

For all $\mathbf{R}_n \in \mathcal{N}'$, the denominator of $d^2(\mathbf{H}_\mathrm{D}^\mathrm{R};\mathbf{R}_s,\mathbf{R}_n)$ equals

$$\mathrm{tr}\{(\mathbf{H}_\mathrm{D}^\mathrm{R}\mathbf{R}_n)^2\} = \mathrm{tr}\left\{\sum_{i=1}^{N}\frac{s_i}{n_i^2}\mathbf{P}_i\mathbf{R}_n\sum_{j=1}^{N}\frac{s_j}{n_j^2}\mathbf{P}_j\mathbf{R}_n\right\}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{s_i s_j}{n_i^2 n_j^2}\mathrm{tr}\{\mathbf{P}_i\mathbf{R}_n\mathbf{P}_j\mathbf{R}_n\} = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{s_i s_j}{n_i^2 n_j^2}n_i^2\delta_{i,j} = \sum_{i=1}^{N}\frac{s_i^2}{n_i^2}. \tag{17}$$

By combining (16) and (17), it follows that for all $\mathbf{R}_s \in \mathcal{S}$, $\mathbf{R}_n \in \mathcal{N}'$, the deflection equals $d^2(\mathbf{H}_\mathrm{D}^\mathrm{R};\mathbf{R}_s,\mathbf{R}_s) = \sum_{i=1}^{N}\frac{s_i^2}{n_i^2}$. Thus, the right-hand inequality of (11) is satisfied with equality. Together with the fact that (11) implies (10), this establishes Theorem 1.

REFERENCES

[1] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer, 1988.

[2] A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science*. New York: Springer, 2nd ed., 1982.

[3] C. R. Baker, "Optimum quadratic detection of a random vector in Gaussian noise," *IEEE Trans. Comm. Technol.*, vol. 14, pp. 802–805, Dec. 1966.

[4] G. Matz and F. Hlawatsch, "Minimax robust nonstationary signal estimation based on a p-point uncertainty model," *J. Franklin Inst.*, vol. 337, pp. 403–419, July 2000.

[5] L. J. Cimini and S. A. Kassam, "Robust and quantized Wiener filters for p-point spectral classes," in *Proc. Conf. Information Sciences and Systems*, (Princeton, NJ), pp. 314–318, March 1980.

[6] S. Verdú and H. V. Poor, "On minimax robustness: A general approach and applications," *IEEE Trans. Inf. Theory*, vol. 30, pp. 328–340, March 1984.

[7] R. J. Barton and H. V. Poor, "An RKHS approach to robust L_2 estimation and signal detection," *IEEE Trans. Inf. Theory*, vol. 36, pp. 485–501, May 1990.

[8] E. K. Al-Hussaini and S. A. Kassam, "Robust Eckart filters for time-delay estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 32, pp. 1052–1063, Oct. 1984.

[9] S. G. Mallat, *A Wavelet Tour of Signal Processing*. San Diego: Academic Press, 1998.

[10] G. Matz, F. Hlawatsch, and A. Raidl, "Signal-adaptive robust time-varying Wiener filters: Best subspace selection and statistical analysis," in *Proc. IEEE ICASSP-2001*, (Salt Lake City, UT), pp. 3945–3948, May 2001.