PERFORMANCE EVALUATION OF ADAPTIVE SUBSPACE DETECTORS, BASED ON STOCHASTIC REPRESENTATIONS

Shawn Kraut and Louis Scharf
Dept. of ECE, University of Colorado, Boulder
Campus Box 425
Boulder, C0 80309-0425 USA
Tel: (303) 492-2759; fax: (303) 492-2758
e-mail: kraut@colorado.edu, scharf@colorado.edu

ABSTRACT

In this paper we present a technique for evaluating the moments of "adaptive" detectors, where the noise covariance is estimated from training data, rather than assumed to be known *a priori*. It is based on the method of "stochastic representations" recently presented in [1]. These representations express adaptive detectors as simple functions of the same set of five statistically independent scalar random variables. They may be applied to a whole class of detectors, which includes the adaptive versions of the matched subspace detectors shown to be UMP-invariant and GLRT in [2] and [3], and to the adaptive GLRT detector of Kelly [4]. Using a stochastic representation, the moments of any member of this class may be evaluated without the need to derive its density or characteristic function. The first two moments give a convenient measure for how the SNR loss improves as the number of training vectors, M, increases. In this paper, the analysis is presented using the example of the adaptive version of the matched filter, illustrated in Figure 1.

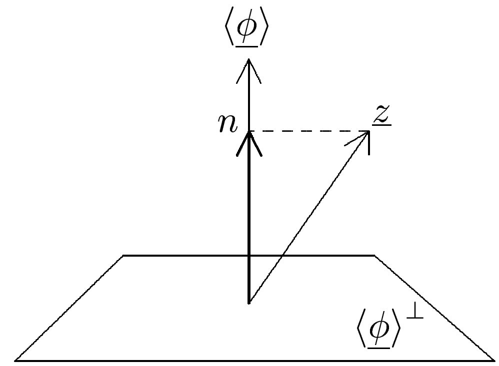

Figure 1: **The matched filter** statistic resolves the *projection* of $\underline{z}$ onto $\langle \underline{\phi} \rangle$

1 INTRODUCTION

This paper is concerned with the general problem of testing the presence of a signal in a multivariate measurement $\underline{y} \in \mathbb{C}^N$ that has a complex normal distribution, $\underline{y} \sim CN[\mu\underline{\psi}, \gamma^2\mathbf{R}]$. The question is whether or not the signal $\underline{\psi}$ is present in the data; this characterizes a hypothesis test on the scaling parameter μ: $H_0 : \mu = 0$, vs. $H_1 : \mu \neq 0$. When the structure of the noise covariance $\mathbf{R}$ is known, there is a class of four "matched subspace detectors" (MSDs) that can be applied, depending on whether the noise scaling γ and the signal phase are known or unknown. These detectors have been shown to be Uniformly Most Powerful Invariant (UMP Invariant) [2] and Generalized Likelihood Ratio Tests (GLRT) [3].

Here we consider the "adaptive" case in which the noise covariance $\mathbf{R}$ is unknown and is replaced by an estimate $\hat{\mathbf{R}} = \mathbf{S}$ based on training data vectors $\{\underline{x}_i\}$, which are independent of $\underline{y}$. In [1], we presented the method of "stochastic representations" for analyzing the statistical behavior of adaptive detectors constructed in this manner. This method applies to the adaptive versions of all four MSDs, as well as to Kelly's adaptive GLRT [4]. All of these detectors have a stochastic representation in terms of the same set of five scalar, independent random variables, which are either normal or chi-squared distributed. These representations distill the probabilistic expressions of the adaptive detectors down to their simplest form, in terms of the *same* scalar random variables. They may be used to directly evaluate the mean, variance, and higher moments of an adaptive detector, without requiring its density or characteristic function.

The ratio of the difference in means to the variance, or signal-to-noise ratio (SNR) is one simple scalar measure of how well the probability densities are separated under the two hypotheses, and thus of detection performance. In the following sections, we will demonstrate the evaluation of the SNR using the example of an adaptive matched filter, $\hat{n}$. This will yield a simple expression for how the SNR improves as M gets larger and the sample covariance estimate improves.

To see this behavior graphically, refer to Figure 2, which shows how the densities of $\hat{n}$ become better sep-

This work was supported by the Office of Naval Research under Contract No. N00014-89-J-1070, and by the National Science Foundation under Contract No. MIP-9529050.

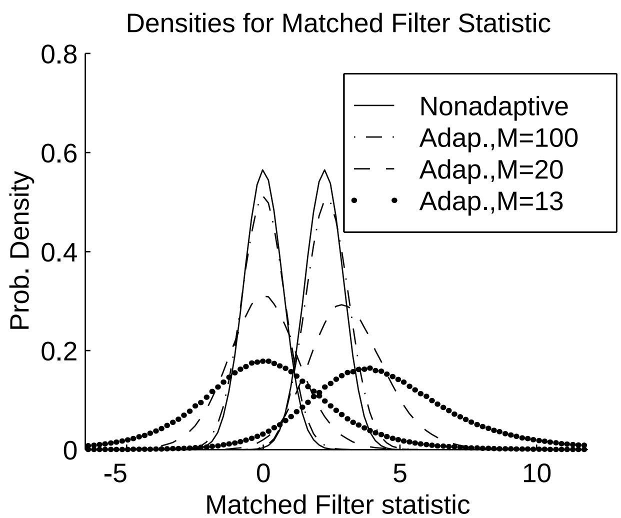

Figure 2: **Densities for the matched filter** n and the adaptive matched filter $\hat{n}$, for different numbers of training samples M; these were obtained from a Monte Carlo simulation using the canonical representation of $\hat{n}$. As M increases, the separation between the two hypotheses improves. (Here $N = 10$ and $|\mu^2| \underline{\psi}^\dagger \mathbf{R}^{-1} \underline{\psi} = 5$).

arated as the number of training vectors, M, increases. Figure 3 indicates the corresponding improvement in detection performance, as shown by the Receiver Operating Characteristics (ROC).

2 MATCHED SUBSPACE DETECTORS

The detector we will analyze in the following sections is the adaptive version of the *matched filter*, which resolves the *projection* of the measurement $\underline{y}$ onto the subspace of the signal (or steering vector) $\underline{\psi}$, in coordinates whitened by $\mathbf{R}^{-1/2}$, and compares it with a threshold η:

$$Re[n] \gtrless \eta; \quad n = \frac{\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{y}}{\gamma \sqrt{\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{\psi}}} \tag{1}$$

This is depicted geometrically in Figure 1, where the whitened measurement and signal are denoted by $\underline{z} = \mathbf{R}^{-\frac{1}{2}} \underline{y}$ and $\underline{\phi} = \mathbf{R}^{-\frac{1}{2}} \underline{\psi}$. When $\mathbf{R}$ is known, the matched filter is normally distributed: $n \sim CN[\frac{\mu}{\gamma}\sqrt{\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{\psi}}, 1]$.

When the noise covariance has an unknown scaling, γ, then the appropriate detector is the *CFAR matched filter* [2, 3]:

$$Re[cos] \gtrless \eta; \quad cos = \frac{\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{y}}{\sqrt{\underline{y}^\dagger \mathbf{R}^{-1} \underline{y}} \sqrt{\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{\psi}}} \tag{2}$$

This is simply the matched filter divided by the magnitude of the whitened measurement vector. This operation makes the detector invariant to arbitrary scaling of the measurement, and makes it CFAR with respect to the variance level γ. Referring again to Figure 1, the CFAR MSD measures the cosine of the *angle* that $\underline{z}$ makes with $\underline{\psi}$, rather than its projection onto $\langle \underline{\psi} \rangle$.

In the noncoherent case, when the phase of the signal $\underline{\psi}$ is unknown, these two detection statistics are modified by magnitude squaring them, yielding two more statistics, $\chi^2 = |n|^2$ and $\beta = |cos|^2$.

When the covariance $\mathbf{R}$ is unknown, all four detectors may be modified by simply replacing $\mathbf{R}$ with its estimate, $\mathbf{S}$. The resulting detectors, along with the stochastic representations that are explained in the next section, are summarized in Figure 4. For the coherent and noncoherent versions of the matched filter, this procedure is *ad hoc*, as the true GLRTs are given by the Kelly detector of [4], and its coherent version (neither are shown here, though the method of stochastic representations can be applied to them as well). For the coherent and noncoherent CFAR matched filters, this procedure is *not ad hoc*; it does produce the true GLRT, as we have recently shown in [5].

3 DERIVING STOCHASTIC REPRESENTATIONS

Now we present a method for succinctly characterizing the statistical behavior of adaptive detectors which employ the sample estimate of $\mathbf{R}$, namely $\hat{\mathbf{R}} = \mathbf{S} = \frac{1}{M}\sum_{i=1}^{M} \underline{x}_i \underline{x}_i^\dagger$, (We assume that the *training data* used to build $\hat{\mathbf{R}}$ is complex Gaussian distributed: $\underline{x}_i \sim CN[\underline{0}, \mathbf{R}]$.) An example is the adaptive matched filter statistic, $\hat{n}$, given by

$$\hat{n} = \frac{\underline{\psi}^\dagger \mathbf{S}^{-1} \underline{y}}{\gamma \sqrt{\underline{\psi}^\dagger \mathbf{S}^{-1} \underline{\psi}}} \tag{3}$$

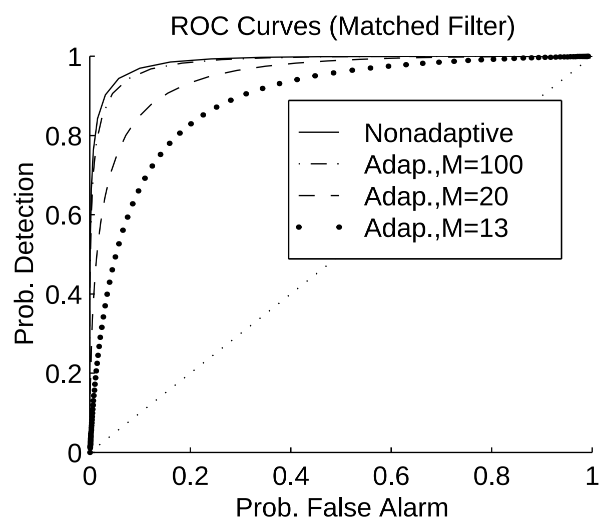

Figure 3: **Receiver operating characteristics** for the matched filter n and the adaptive matched filter $\hat{n}$. As M increases, the detection performance improves.

$\hat{n} = \frac{\underline{\psi}^\dagger \hat{\mathbf{S}}^{-1}\underline{y}}{\gamma\sqrt{\underline{\psi}^\dagger\hat{\mathbf{S}}^{-1}\underline{\psi}}}$ $= \frac{\sqrt{M}}{\sqrt{h_1}}\left[n+\sqrt{g}\frac{h_3}{\sqrt{h_2}}\right]$	$\widehat{cos} = \frac{\underline{\psi}^\dagger\hat{\mathbf{S}}^{-1}\underline{y}}{\sqrt{\underline{y}^\dagger\hat{\mathbf{S}}^{-1}\underline{y}}\sqrt{\underline{\psi}^\dagger\hat{\mathbf{S}}^{-1}\underline{\psi}}}$ $= \frac{\hat{t}}{\sqrt{\vert\hat{t}\vert^2+1}}$, where $\hat{t} = \frac{1}{\sqrt{h_1}}\left[\frac{n}{\sqrt{g}}\sqrt{h_2}+h_3\right]$
$\widehat{\chi^2} = \frac{\vert\underline{\psi}^\dagger\hat{\mathbf{S}}^{-1}\underline{y}\vert^2}{\gamma^2\underline{\psi}^\dagger\hat{\mathbf{S}}^{-1}\underline{\psi}}$ $= \frac{M}{h_1}\left\vert n+\sqrt{g}\frac{h_3}{\sqrt{h_2}}\right\vert^2$	$\hat{\beta} = \frac{\vert\underline{\psi}^\dagger\hat{\mathbf{S}}^{-1}\underline{y}\vert^2}{(\underline{y}^\dagger\hat{\mathbf{S}}^{-1}\underline{y})(\underline{\psi}^\dagger\hat{\mathbf{S}}^{-1}\underline{\psi})}$ $= \frac{\hat{F}}{\hat{F}+1}$, where $\hat{F} = \frac{1}{h_1}\left\vert\frac{n}{\sqrt{g}}\sqrt{h_2}+h_3\right\vert^2$
$n,\ g,\ h_1,\ h_2,\ h_3$ all $\perp\!\!\!\perp$ (independent)	
$n \sim CN[\frac{\mu}{\gamma}\sqrt{\underline{\psi}^\dagger\mathbf{R}^{-1}\underline{\psi}}, 1]$ $(H_0 : \mu = 0)$, $g \sim \gamma_{N-1}[0]$	
$h_1 \sim \gamma_{M-N+1}[0]$, $h_2 \sim \gamma_{M-N+2}[0]$, $h_3 \sim CN[0,1]$	

Figure 4: **The stochastic representations** of the four adaptive matched subspace detectors.

The derivation of the representation for $\hat{n}$ will only be briefly described here; the method is presented in complete detail in [1, 6]. The method may be outlined in four steps: (1) Apply the whitening transformation $\mathbf{R}^{-\frac{1}{2}}$ to the training and test data; this generates the transformed signal vector $\underline{\phi} = \mathbf{R}^{-\frac{1}{2}}\underline{\psi}$, and test vector $\underline{z} = \mathbf{R}^{-\frac{1}{2}}\underline{y}$. (2) Next apply a unitary transformation to rotate to a coordinate system in which the first two basis vectors are set in the direction of $\underline{\phi}$ and $\mathbf{P}^{\perp}_{\underline{\phi}}\underline{z}$. (3) Resolve the inverse of the sample covariance matrix $\mathbf{S}$ onto the 2×2 subspace $\langle\underline{\phi}, \mathbf{P}^{\perp}_{\underline{\phi}}\underline{z}\rangle$. (4) Perform a change of variables on the elements of the resulting 2×2 covariance matrix so that these variables are now *statistically independent*.

This results in the following stochastic representation for $\hat{n}$:

$$\hat{n} = \frac{\sqrt{M}}{\sqrt{h_1}}\left[n+\sqrt{g}\frac{h_3}{\sqrt{h_2}}\right] \tag{4}$$

This representation is in terms of five statistically independent variables. Two of these, n and g, depend solely on the measurement $\underline{y}$: n is the nonadaptive matched filter of Equation 1, and $\gamma^2 g = \underline{z}^\dagger(\mathbf{I}-\mathbf{P}_{\underline{\phi}})\underline{z}$ is an estimate of $(N-1)\gamma^2$. Then g has a gamma distribution (it is a chi-squared random variable scaled by $\frac{1}{2}$): $g \sim \gamma_{N-1}[0]$. The other three, h_1,h_2, and h_3, depend solely on the training data $\{\underline{x}_i\}$, and have gamma or normal distributions. The canonical representations for $\hat{n}$, as well as for the other three MSDs are summarized in Figure 4.

4 USING THE CANONICAL REPRESENTATIONS TO EVALUATE MOMENTS

The expressions in Figure 4 can be written in terms of sums and products of scalar random variables. For example, an adaptive matched filter $\hat{n}$ involves sums and products of n,$\sqrt{g}$,$\frac{1}{\sqrt{h_1}}$, $\frac{1}{\sqrt{h_2}}$, and h_3. For two independent random variables, a and b, the expectation of their sums and products is given by

$$\begin{aligned} E[a+b] &= E[a]+E[b] \\ E[a\cdot b] &= E[a]\cdot E[b]. \end{aligned} \tag{5}$$

Similarly, one can also obtain "propagation of variance" formulas for the variance of their sums and products (this is in analogy with the "propagation of error" formulas used to evaluate the uncertainty in an experimentally derived quantity, due to the uncertainties in measured quantities). Note however that the following expressions are *exact*:

$$\begin{aligned} var[a+b] &= var[a]+var[b] \\ var[a\cdot b] &= (E[a])^2 var[b] + (E[b])^2 var[a] \\ &\quad + var[a]var[b] \end{aligned} \tag{6}$$

Let us consider densities of the variables in the representation for $\hat{n}$: n and h_3 have complex normal distributions; $\sqrt{g}$ has a scaled Rayleigh distribution, and $\frac{1}{\sqrt{h_1}}$ and $\frac{1}{\sqrt{h_2}}$ are inverse scaled Rayleighs. The square root of a gamma with p degrees of freedom,$\sqrt{\gamma_p}$, has the following moments:

$$E[\sqrt{\gamma_p}] = \frac{\Gamma(p+\frac{1}{2})}{\Gamma(p)}, \qquad E[(\sqrt{\gamma_p})^2] = \frac{\Gamma(p+1)}{\Gamma(p)} \tag{7}$$

Similar expressions may be found for the first two moments of the inverse scaled Rayleigh, $\frac{1}{\sqrt{\gamma_q}}$.

$$E\left[1/\sqrt{\gamma_q}\right] = \frac{\Gamma(q-\frac{1}{2})}{\Gamma(q)}, \quad E\left[\left(1/\sqrt{\gamma_q}\right)^2\right] = \frac{\Gamma(q-1)}{\Gamma(q)} \tag{8}$$

By repeated application of Equations 5-6, we can express the mean and variance of $\hat{n}$ and the other adaptive MSDs detectors in terms of the moments of the gamma and Rayleigh densities. Under the signal-absent hypothesis H_0, $\hat{n}$ has zero mean. Under the signal-present hypothesis H_1, $\hat{n}$ has a mean given by

$$E_1(\hat{n}) = \sqrt{M}\frac{\Gamma(M-N+\frac{1}{2})}{\Gamma(M-N+1)}\bar{n} \approx \sqrt{\frac{M}{M-N-\frac{1}{4}}}\bar{n}, \tag{9}$$

where $\bar{n}$ is the mean of the nonadaptive matched filter under the signal-present hypothesis: $\bar{n} = E_1(n) = \frac{\mu}{\gamma}\sqrt{\underline{\psi}^\dagger\mathbf{R}^{-1}\underline{\psi}}$. (The approximation of the ratio of Gamma functions is obtained from the asymptotic approximation $\Gamma(z+\frac{1}{4}) \approx \sqrt{2\pi}z^{z-\frac{1}{4}}e^{-z}$.) The exact variance under H_0 is given by

$$\sigma_0^2(\hat{n}) = \frac{M^2}{(M-N)(M-N+1)}. \tag{10}$$

And the variance under H_1 is given by

$$\begin{aligned} \sigma_1^2(\hat{n}) &= \sigma_0^2(\hat{n})\left[1+\bar{n}^2\epsilon_n(M,N)\right], \\ \epsilon_n(M,N) &= \frac{M-N+1}{M}\left[1-\frac{(M-N)\Gamma(M-N+\frac{1}{2})^2}{\Gamma(M-N+1)^2}\right] \\ &\approx \frac{M-N+1}{4M(M-N+\frac{1}{4})} \approx \frac{1}{4M}. \end{aligned} \tag{11}$$

The variance under H_1 is larger, so a pessimistic estimate of the SNR is given by

$$\begin{aligned} SNR(\hat{n}) &= \frac{\left[E_1(\hat{n})-E_0(\hat{n})\right]^2}{\sigma_1^2(\hat{n})} = \frac{\left[E_1(\hat{n})\right]^2}{\sigma_0^2(\hat{n})[1+\bar{n}^2\epsilon_n(M,N)]} \\ &= \frac{(M-N)(M-N+1)\Gamma(M-N+\frac{1}{2})^2}{M\Gamma(M-N+1)^2}\cdot\frac{\bar{n}^2}{1+\bar{n}^2\epsilon_n(M,N)} \\ &\approx \frac{M-N}{M}\cdot\frac{\bar{n}^2}{1+\frac{\bar{n}^2}{4M}}. \end{aligned} \tag{12}$$

Normalizing this by the SNR of the nonadaptive matched filter, namely $SNR(n) = \bar{n}^2/1$, gives

$$\begin{aligned} \frac{SNR(\hat{n})}{SNR(n)} &\approx \frac{M-N}{M}\cdot\frac{1}{1+\frac{SNR(n)}{4M}} \\ &\approx \begin{cases} \frac{4(M-N)}{SNR(n)}, & 4M \ll SNR(n) \\ \frac{M-N}{M}, & 4M \gg SNR(n) \end{cases} \end{aligned} \tag{13}$$

For the performance of the adaptive detector to approach that of the non-adaptive detector, this ratio needs to be close to unity. This in turn gives a minimal requirement that $M \gg SNR(n)/4$ (incidently, this condition makes $\sigma_1^2(\hat{n}) \approx \sigma_0^2(\hat{n})$). Then the second approximation of Equation 13 can be used to determine how large M (number of training vectors) needs to be to only suffer $x\%$ of SNR loss:

$$\begin{aligned} \frac{M-N}{M} &= \frac{100-x}{100} \\ &\rightarrow\ M = \frac{N}{x}\cdot 100. \end{aligned} \tag{14}$$

To lose less than 5% of the SNR, this says M has to be larger than $20 \cdot N$ (where N is the dimension of the measurement and signal vectors), etc.

5 Conclusion

We have presented stochastic representations for four adaptive matched subspace detectors and shown how they may be applied to find the detector moments and evaluate the SNR as a figure-of-merit for detector performance. It should be noted that by integrating over the densities of the constituent random variables in the representations given in Figure 4, explicit analytic functions for the density and characteristic function of an adaptive detector can be obtained. However, these expressions involve infinite sums and/or indefinite integrals that need to be numerically approximated [4, 7, 8, 9]. By making use of the fact the the constituent random variables in a stochastic representation are statistically independent, we have shown how exact analytic expressions for the moments can be obtained, bypassing the complexity of the density and characteristic function. In this way we can characterize how the performance of the adaptive detectors approaches that of their non-adaptive counterparts, as the number of training samples in the estimated covariance matrix gets large.

References

[1] S. Kraut, L. T. McWhorter, and L. L. Scharf. A canonical representation for distributions of adaptive matched subspace detectors. In *Proceedings of the 31st Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 1997.

[2] L. L. Scharf. *Statistical Signal Processing*, chapter 4. Addison-Wesley, 1991.

[3] L. L. Scharf and B. Friedlander. Matched subspace detectors. *IEEE Trans. Signal Processing*, 42(8):2146–2157, 1994.

[4] E. J. Kelly. An adaptive detection algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, 22(1):115–127, 1986.

[5] S. Kraut and L. L. Scharf. The adaptive CFAR matched subspace detector is a scale-invariant GLRT. *IEEE Trans. Signal Processing*. submitted for review, April 1998.

[6] S. Kraut, L. T. McWhorter, and L. L. Scharf. Stochastic representations for the Kelly, AMF, and ACE detectors, and related quadratic forms. *IEEE Trans. Signal Processing*. submitted for review, April 1998.

[7] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. A. Nitzberg. A CFAR adaptive matched filter detector. *IEEE Transactions on Aerospace and Electronic Systems*, 28(1):208–216, 1992.

[8] L. T. McWhorter, L. L. Scharf, and L. J. Griffiths. Adaptive coherence estimation for radar signal processing. In *Proceedings of the 30th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 1996.

[9] Christ D. Richmond. Statistical performance analysis of the adaptive sidelobe blanker detection algorithm. In *Proceedings of the 31st Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 1997.