

## REFERENCES

- [1] G. C. Carter, Ed., *Coherence and Time Delay Estimation: An Applied Tutorial for Research, Development, Test, and Evaluation Engineers*. New York: IEEE, 1993, pp. 134–137.
- [2] B. Emile and P. Comon, "Estimation of time delays between unknown colored signals," *Signal Process.*, vol. 68, no. 1, pp. 93–100, Aug. 1998.
- [3] M. J. Hinich and G. R. Wilson, "Time delay estimation using the cross bispectrum," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. 40, pp. 106–113, Jan. 1992.
- [4] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*. New York: Springer, 1988.
- [5] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 320–327, Aug. 1976.
- [6] H. Messer and Y. Anu, "A novel algorithm for time delay estimation with improved performance for non-Gaussian signals," in *Proc. ATHOS Workshop Syst. Identification High Order Stat.*, Sophia-Antipolis, France, Sept. 20–21, 1993.
- [7] C. L. Nikias and R. Pan, "Time delay estimation in unknown Gaussian spatially correlated noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1706–1714, Nov. 1988.
- [8] P. M. Schultheiss, H. Messer, and G. Shor, "Maximum likelihood time delay estimation in non-Gaussian noise," *IEEE Trans. Signal Processing*, vol. 45, pp. 2571–2575, Oct. 1997.
- [9] J. K. Tugnait, "On time delay estimation with unknown spatially correlated Gaussian noise using fourth-order cumulants and cross cumulants," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 39, pp. 1258–1267, June 1991.

## Quadratic Tests for Detection of Abrupt Changes in Multivariate Signals

Igor V. Nikiforov

**Abstract**—This correspondence considers the problem of detecting abrupt changes in the mean of a multivariate Gaussian random signal. A fixed sample size  $\chi^2$ -test is compared against the optimum sequential tests ( $\chi^2$ -CUSUM and  $\chi^2$ -GLR).

### I. INTRODUCTION

The change detection problem considered in this correspondence arises from practical tasks when it is necessary to detect quickly any significant change at an unknown time  $\nu$  in the mean vector  $\theta$  of an independent Gaussian multivariate ( $r > 1$ ) sequence

$$\mathcal{L}(Y_t) = \begin{cases} \mathcal{N}(\theta_0, \Sigma), & \text{if } t < \nu \\ \mathcal{N}(\theta_1, \Sigma), & \text{if } t \geq \nu \end{cases} \quad (1)$$

taken from some process. Two main ways of extending the scalar detection scheme are known in the literature [2]: *linear* and *quadratic*. The *linear* multivariate scheme, when it is necessary to specify both a nominal value  $\theta_0$  (hypothesis  $\mathcal{H}_0$ ) and an alternative  $\theta_1$  (hypothesis  $\mathcal{H}_1$ ), is an elementary extension of the scalar one [2]. In this correspondence, we will discuss the *quadratic* multivariate

Manuscript received October 8, 1997; revised February 11, 1999. The associate editor coordinating the review of this paper and approving it for publication was Dr. Shubha Kadambe.

The author is with the Département Genie des Systèmes d'Information et de Décision, Laboratoire de Modélisation et Sécurité des Systèmes, Université De Technologie de Troyes, Troyes Cedex, France.

Publisher Item Identifier S 1053-587X(99)06740-9.

scheme when it is sufficient to specify a nominal value  $\theta_0$  and an ellipsoid around  $\theta_0$  in the following manner:

$$\mathcal{H}_0 : \{\theta_0\} \text{ vs } \mathcal{H}_1 : \{\theta_1 \in \Theta_1 : (\theta - \theta_0)^T \Sigma^{-1} (\theta - \theta_0) = b^2\} \quad (2)$$

where  $\theta_0$ ,  $\Sigma$ , and the signal-to-noise ratio (SNR)  $b$  are known. We assume the non-Bayesian (min-max) approach; hence, the change time  $\nu$  is an unknown but *nonrandom* integer value. Let  $F_\nu$  ( $\nu \geq 1$ ) be the distribution of the observations  $Y_1, \dots, Y_{\nu-1}, Y_\nu, \dots, Y_t$  when  $Y_\nu$  is the first observation distributed according to the law  $\mathcal{N}(\theta_1, \Sigma)$ ,  $\theta_1 \in \Theta_1$ . The associated probability and the expectation are denoted by  $\mathbf{P}_\nu$  and  $E_\nu$ , respectively. The notation  $F_\infty$  corresponds to the case when all observations are distributed according to  $\mathcal{N}(\theta_0, \Sigma)$ . Hence,  $\mathbf{P}_\infty(\cdot) = \mathbf{P}_{\theta_0}(\cdot)$ , and  $E_\infty(\cdot) = E_{\theta_0}(\cdot)$ . The statistical performance of non-Bayesian algorithms is measured with the aid of a criterion proposed by Lorden [7]. The stopping (alarm) time of a change detection algorithm is denoted by  $N$ . We require that the 'worst case' mean detection delay

$$\bar{\tau}^* = \sup_{\nu \geq 1} \text{esssup } E_\nu(N - \nu + 1 \mid N \geq \nu, Y_1, \dots, Y_{\nu-1}) \quad (3)$$

should be as small as possible for a given mean time before a false alarm

$$\bar{T} = E_\infty(N). \quad (4)$$

Lorden has proved that the cumulative sum (CUSUM) scheme minimizes  $\bar{\tau}^*$  (3) in the class  $\mathcal{K}_\gamma = \{N : E_\infty(N) \geq \gamma\}$  and that the infimum of  $\bar{\tau}^*$  as  $N$  ranges over this class is

$$n(\gamma) = \inf_{N \in \mathcal{K}_\gamma} \bar{\tau}^* \sim \frac{\log \gamma}{\rho(\theta_1, \theta_0)} \quad \text{as } \gamma \rightarrow \infty \quad (5)$$

where  $\rho(\theta_1, \theta_0) = E_{\theta_1}(\log \frac{\varphi_{\theta_1}(Y_t)}{\varphi_{\theta_0}(Y_t)})$  is the Kullback–Leibler information, and  $\varphi_{\theta, \Sigma}(Y) = (2\pi)^{-\frac{r}{2}} (\det \Sigma)^{-\frac{1}{2}} \exp\{-\frac{1}{2}(Y - \theta)^T \Sigma^{-1} (Y - \theta)\}$  is the probability density of the multivariate Gaussian distribution  $\mathcal{N}(\theta, \Sigma)$ . Moustakides [8] and Ritov [11] investigated *nonasymptotic* aspects of optimality for the CUSUM scheme. Finally, recent optimality results are summarized in the exhaustive survey of Lai [6].

The goal of the correspondence is to compare optimal sequential and nonsequential, or fixed sample size (FSS), detection strategies for the model (1), (2) by using the criterion (3), (4). The starting point of this direction is the classical work of Wald [13]. It is known that the sequential strategy is optimal for detection of abrupt changes, but a practical motivation to use the nonsequential one is the fact that this strategy uses observations 'block-by-block' that seriously simplify the transmission and processing of the input information. The complexity of a detector is proportional to the mean number of the likelihood ratio (LR) computations at time  $t$ . In the case considered, the sequential detector leads to the number of the LR computations, which grows to infinity with  $t$ . Unlike the sequential detector, the FSS rule involves only one LR computation at every stage  $t$ . The results of the comparison between optimal sequential and FSS strategies for scalar signals can be found in [12] for the Bayesian approach and in [10] for the min-max approach. It is shown in [10] that the scalar CUSUM is *asymptotically twice as good as* the nonsequential (FSS) optimal competitor. We will show that this fact can be extended to the model (1), (2). This means that in some practical detection problems, it is reasonable to use a *very simple*  $\chi^2$ -FSS test rather than the optimal  $\chi^2$  sequential tests with a *high* computational cost.

## II. SEQUENTIAL STRATEGY

To solve the problem of sequential detection for unknown  $\theta_1$ , Wald [13] proposed two possible solutions. The first one is to replace the LR by the *weighted* LR  $\hat{\Lambda}_k^t = \int \cdots \int_{\Theta_1} \prod_{i=k}^t \frac{\varphi_{\theta_1, \Sigma}(Y_i)}{\varphi_{\theta_0, \Sigma}(Y_i)} f(\theta_1) dS$ , where  $dS$  is the surface element of the ellipsoid  $\Theta_1$ , and the *weighting function*  $f(\theta_1)$  is concentrated on the surface of  $\Theta_1$ . The second solution consists of maximizing the LR  $\prod_{i=k}^t \frac{\varphi_{\theta_1, \Sigma}(Y_i)}{\varphi_{\theta_0, \Sigma}(Y_i)}$  with respect to  $\theta_1 \in \Theta_1$ :  $\hat{\Lambda}_k^t = \sup_{\theta_1 \in \Theta_1} \prod_{i=k}^t \frac{\varphi_{\theta_1, \Sigma}(Y_i)}{\varphi_{\theta_0, \Sigma}(Y_i)}$ , which results in the generalized likelihood ratio (GLR) test. Therefore, the  $\chi^2$ -CUSUM is based on the *weighted* LR  $\hat{\Lambda}_k^t$ , and the  $\chi^2$  GLR is based on the GLR  $\hat{\Lambda}_k^t$ . The detailed proofs of the  $\chi^2$ -CUSUM and GLR tests are given in [2, ch. 4, 7]. The stopping time of the  $\chi^2$ -CUSUM test is expressed in the following form:

$$\tilde{N} = \inf \left\{ t \geq 1 : \max_{1 \leq k \leq t} \tilde{S}_k^t \geq h \right\}$$

$$\tilde{S}_k^t = -\frac{(t-k+1)b^2}{2} + \log G\left(\frac{r}{2}, \frac{b^2(\chi_k^t)^2}{4}\right) \quad (6)$$

$$(\chi_k^t)^2 = (\tilde{S}_k^t)^T \Sigma^{-1} \tilde{S}_k^t$$

where  $\tilde{S}_k^t = \sum_{i=k}^t (Y_i - \theta_0)$ ,  $G(d, z) = 1 + \frac{z}{d} + \cdots + \frac{z^d}{d(d+1)\cdots(d+n-1)n!} + \cdots$  is the generalized hypergeometric function, and  $h > 0$  is a threshold. Let us consider the model given by (1) and (2). The  $\chi^2$ -CUSUM test is asymptotically optimal in the sense of the min-max criterion (3), (4). The result is stated by the following theorem (see the proofs in [2] and [9]).

*Theorem 1:* The ‘worst case’ mean detection delay for the  $\chi^2$ -CUSUM test (6) is given by the following asymptotic equation  $\bar{\tau}^*(\bar{T}) \sim n(\bar{T}) \sim \frac{\log \bar{T}}{\rho(\theta_1, \theta_0)} = \frac{2 \log \bar{T}}{b^2}$  as  $\bar{T} \rightarrow \infty$ , where  $\rho(\theta_1, \theta_0) = \frac{b^2}{2}$ .

The stopping time of the GLR test is

$$\hat{N} = \inf \left\{ t \geq 1 : \max_{1 \leq k \leq t} \hat{S}_k^t \geq h \right\}$$

$$\hat{S}_k^t = -\frac{(t-k+1)b^2}{2} + b|\chi_k^t| \quad (7)$$

where  $\chi_k^t$  is defined in (6), and  $h > 0$  is a threshold. Let us show now that the  $\chi^2$ -GLR test is asymptotically (as  $\bar{T} \rightarrow \infty$ ) equivalent to the above  $\chi^2$ -CUSUM test. It follows from the properties of the generalized hypergeometric function  $G(d, z)$  [1, Ch. 13] that the function  $F : y \mapsto \log G(d, \frac{y^2}{4})$  is convex and its derivative exists and is non-negative when  $y \geq 0$  and  $d \geq 1$ . It then can be shown that the inequalities

$$\alpha(x)y + \beta(x) \leq \log G\left(d, \frac{y^2}{4}\right) \leq y \quad (8)$$

where  $\alpha(x) = \frac{d}{dy}(\log G(d, \frac{y^2}{4}))|_{y=x}$  and  $\beta(x) = -\alpha(x)x + \log G(d, \frac{x^2}{4})$ , hold for any  $x, y \geq 0$ . The first inequality has a simple geometric interpretation. The linear function  $L : y \mapsto \alpha(x)y + \beta(x)$  is tangent to the graph of  $F(y) = \log G(d, \frac{y^2}{4})$  at the point  $(x, F(x))$ . In this inequality,  $x$  plays the role of a tuning parameter. Taking into account the equation  $G(d, z) = e^{\sqrt{z}} M(d - \frac{1}{2}, 2d - 1; 4\sqrt{z})$ , where  $M(a, c, z)$  is a confluent hypergeometric function [1], we get the following asymptotic equations for  $\alpha, \beta$ :  $\alpha(x) = 1 - O(\frac{1}{2x})$  and  $\beta(x) \sim -\frac{r-1}{2} \log 2x$  as  $x \rightarrow \infty$ . From [7], it follows that the ‘worst case’ mean detection delay  $\bar{\tau}^*$  for the stopping time  $\hat{N}$  is less than or equal to  $E_\nu\{\inf(t \geq \nu : \hat{S}_\nu^t \geq h)\}$ . From (8), it follows that  $E_\nu\{\inf(t \geq \nu : \hat{S}_\nu^t \geq h)\} \leq E_\nu\{\inf(t \geq \nu : \tilde{S}_\nu^t \geq h)\}$ . According to [2] and [9],  $E_\nu\{\inf(t \geq \nu : \tilde{S}_\nu^t \geq h)\} \sim \frac{2h}{b^2}$  as  $h \rightarrow \infty$ . Combining these results, we get the asymptotic relation

$\bar{\tau}^* \leq \frac{2h}{b^2}(1 + o(1))$  as  $h \rightarrow \infty$  for the  $\chi^2$ -GLR test. Let us define the modified SNR  $\hat{b}(x) = \frac{b}{\alpha(x)}$  and the following stopping time  $\hat{T}_b = \inf\{t \geq 1 : \hat{S}_1^t(\hat{b}) \geq h\}$  ( $\hat{T}_b = \infty$  if no such  $t$  exists), where  $\hat{S}_1^t(\hat{b}) = -\frac{tb^2}{2} + \hat{b}|\chi_1^t|$ . Taking into account the left-hand side of (8), we get  $\mathbf{P}_{\theta_0}(\hat{T}_b < \infty) \leq \mathbf{P}_{\theta_0}(\bar{T} < \infty)$ , where  $\bar{T} = \inf\{t \geq 1 : \tilde{S}_1^t \geq \alpha^2 h + \beta\}$ . From [2], it results that  $E_\infty(\hat{N}_b) \geq \mathbf{P}_{\theta_0}(\hat{T}_b < \infty)^{-1}$ , where  $\hat{N}_b$  is the stopping time of the  $\chi^2$ -GLR test when the SNR equals  $\hat{b}$ . Hence,  $E_\infty(\hat{N}_b) \geq e^{\alpha^2 h + \beta}$ . Let us choose the parameter  $x \sim h$  as  $h \rightarrow \infty$ . Since  $\alpha(h) \rightarrow 1$  and  $\beta(h) \sim -\frac{r-1}{2} \log 2h$  as  $h \rightarrow \infty$ , we get  $\log E_\infty(\hat{N}_b) \geq (h - \frac{r-1}{2} \log 2h)(1 + o(1))$  as  $h \rightarrow \infty$ . Taking into account that  $\hat{b}(h) = \frac{b}{\alpha(h)} \rightarrow b$  as  $h \rightarrow \infty$ , we get  $\bar{\tau}^*(\bar{T}) \leq \frac{2 \log \bar{T}}{b^2}(1 + o(1))$  as  $\bar{T} \rightarrow \infty$ . This yields the following conclusion.

*Corollary 1:* The asymptotic relation  $\bar{\tau}^*(\bar{T}) \sim \frac{2 \log \bar{T}}{b^2}$  as  $\bar{T} \rightarrow \infty$  is also valid for the  $\chi^2$ -GLR test.

## III. FSS AGAINST SEQUENTIAL

The FSS strategy is based on the following rule: Samples of fixed size  $m \geq 1$  are taken, and at the end of each sample, a decision function is computed to test between the hypotheses  $\mathcal{H}_0$  and  $\mathcal{H}_1$  (2). Sampling is stopped after the first sample  $j$  of observations for which the decision  $d_j$  is taken in favor of  $\mathcal{H}_1$ . The solution to the optimal hypotheses testing problem is given by the  $\chi^2$ -test [2, ch. 4 and 5]. Therefore, the stopping time  $\bar{N}$  of the *repeated*  $\chi^2$ -test is

$$\bar{N} = \inf_{j \geq 1} \{mj : d_j = 1\}; \quad d_j = \begin{cases} 1, & \text{if } |\chi_{(j-1)m+1}^j| \geq mh \\ 0, & \text{if } |\chi_{(j-1)m+1}^j| < mh \end{cases} \quad (9)$$

where  $h > 0$  is a threshold, and  $(\chi_k^t)^2$  is given by (6). The  $\chi^2$ -FSS test has two tuning parameters:  $m$  and  $h$ . Hence, the statistical properties of this test depend on  $m$  and  $h$ . Let us compute the ‘worst case’ mean detection delay  $\bar{\tau}^*$  (3) for the  $\chi^2$ -FSS test as a function of these parameters.

*Lemma 1:* The ‘worst case’ mean detection delay for the  $\chi^2$ -FSS test (9) is given by the following function of  $m$  and  $h$ :

$$\bar{\tau}^*(m, h) = \max \left\{ \frac{m}{1-\beta}, \max_{1 \leq l \leq m} \left[ m-l+1 + \frac{m}{1-\beta} \mathbf{P} \left( \chi_{r,0}^2 < \frac{m^2 h^2}{m-l+1} \right) \right] \right\} \quad (10)$$

where  $\beta = \mathbf{P}(\chi_{r,mb^2}^2 < mh^2)$ , and  $\chi_{r,\lambda}^2$  is distributed according to a noncentral  $\chi^2$  law, with  $r$  degrees of freedom and noncentrality parameter  $\lambda$ .

*Proof of Lemma 1:* See Appendix A. It follows from [2] that the mean time before a false alarm for the  $\chi^2$ -FSS test is

$$E_\infty(\bar{N}) = m[1 - \mathbf{P}(\chi_{r,0}^2 < mh^2)]^{-1}. \quad (11)$$

To compare the  $\chi^2$ -FSS test against the optimum sequential candidates, we have to find the best possible values of the tuning parameters  $m$  and  $h$  for the  $\chi^2$ -FSS test. As it follows from (10) and (11), the optimal choice of the tuning parameters  $m$  and  $h$  reduces to the constrained minimization problem in (12), shown at the bottom of the next page. Therefore, the optimal values of the tuning parameters  $m^*$  and  $h^*$  are functions of  $\bar{T}$ . In the rest of this correspondence, we will consider the minimum ‘worst case’ mean detection delay  $\bar{\tau}^*(\bar{T}) = \bar{\tau}^*(m^*(\bar{T}), h^*(\bar{T}))$  as a function of  $\bar{T}$  for the  $\chi^2$ -FSS test.

An *asymptotic upper bound* for the minimum ‘worst case’ mean detection delay  $\bar{\tau}^*$  is established in the following Theorem.

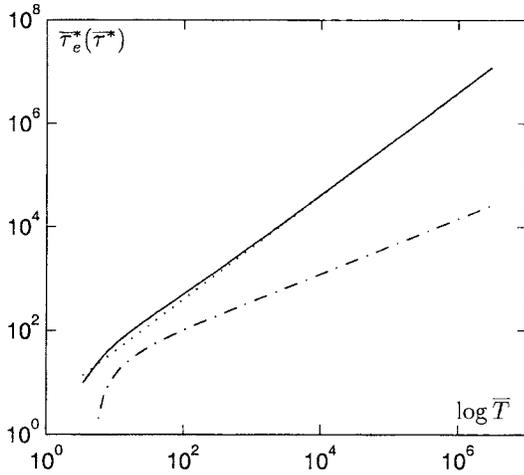


Fig. 1. “Exact” solution  $\bar{\tau}_e^*$  (solid line), the asymptotic upper bound  $\bar{\tau}^* = \frac{4 \log \bar{T}}{b^2}$  (dashed line), and the bias  $\bar{\tau}_e^* - \frac{4 \log \bar{T}}{b^2}$  (dashed-dot line) as functions of  $\log \bar{T}$ .

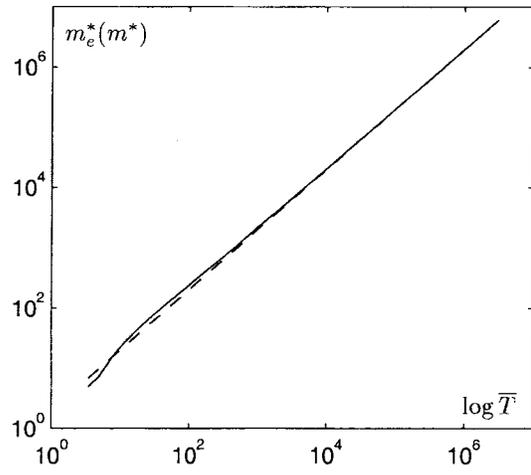


Fig. 2. “Exact” solution  $m_e^*$  (solid line) and the asymptotic formula  $m^* = \frac{2 \log \bar{T}}{b^2}$  (dashed line) as functions of  $\log \bar{T}$ .

**Theorem 2:** The minimum “worst case” mean detection delay for the  $\chi^2$ -FSS test (9) is given by the following asymptotic inequality:

$$\bar{\tau}^*(\bar{T}) \leq \frac{4 \log \bar{T}}{b^2} (1 + o(1)) \quad \text{as } \bar{T} \rightarrow \infty. \quad (13)$$

*Proof of Theorem 2:* See Appendix B.

**Corollary 2:** Let us consider model (1), (2) and compare the optimal quadratic sequential and FSS detection procedures. Asymptotically, as  $\bar{T}$  goes to infinity, the properties of these procedures are given by the following relations:

$$\bar{\tau}^*(\bar{T}) \sim \begin{cases} \frac{2 \log \bar{T}}{b^2}, & \text{for the } \chi^2\text{-CUSUM (GLR) test} \\ \leq \frac{4 \log \bar{T}}{b^2}, & \text{for the } \chi^2\text{-FSS test} \end{cases} \quad \text{as } \bar{T} \rightarrow \infty.$$

**Remark 1:** First, the asymptotic optimal choice of the tuning parameters  $m, h$  is

$$\begin{cases} m^* \sim \frac{2 \log \bar{T}}{b^2} \\ h^* \sim b \{1 + 2 \sqrt{[\log(\sqrt{2} \log \bar{T} / 2) - \log 2 \sqrt{2\pi}] / \log \bar{T}}\}^{-1/2} \end{cases} \quad \text{as } \bar{T} \rightarrow \infty. \quad (14)$$

Second, let us compare the asymptotic upper bound  $\bar{\tau}^* = \frac{4 \log \bar{T}}{b^2}$  and (14) with the “exact” (nonasymptotic!) optimal solution for the  $\chi^2$ -FSS when  $r = 10$  and  $b = 1$ . The “exact” (with the subscript “e”) and asymptotic values of  $\bar{\tau}^*, m^*$  and  $h^*$  as functions of  $\log \bar{T}$  are presented in Figs. 1–3. For the “exact” solution, we deduce the optimal tuning parameters  $m_e^*$  and  $h_e^*$  by numeric constrained minimization of the objective function  $\bar{\tau}^*(m, h)$  (12) for a given  $\bar{T}$ . It is worth noting that a very slow convergence of the “exact” functions  $\bar{\tau}_e^*, m_e^*,$  and  $h_e^*$  to the asymptotic ones as  $\bar{T} \rightarrow \infty$ . The explanation

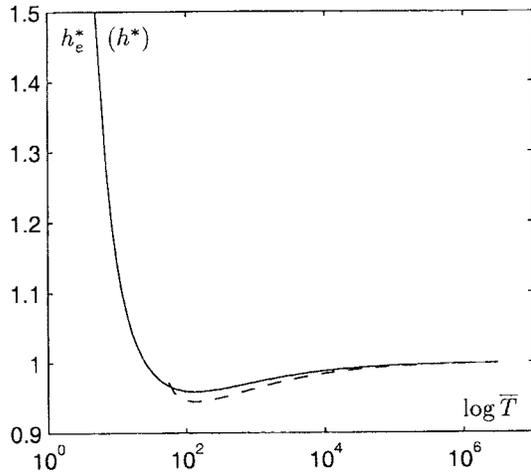


Fig. 3. “Exact” solution  $h_e^*$  (solid line) and the asymptotic formula  $h^* = h^*(\log \bar{T})$  [see (14)] (dashed line) as a function of  $\log \bar{T}$ .

of this fact lies in the feature of the function  $\hat{f}(h, \bar{T})$  (see the proof of Theorem 2). The rapidity of the convergence is defined by the term  $1/\sqrt{2 \log \bar{T}}$ . It follows from Theorem 2 that  $\bar{\tau}_e^*(\bar{T}) - \frac{4 \log \bar{T}}{b^2} \leq o(\frac{4 \log \bar{T}}{b^2})$ . Fig. 1 shows that the bias  $\bar{\tau}_e^*(\bar{T}) - \frac{4 \log \bar{T}}{b^2}$  (dash-dot line) is increasing when  $\bar{T} \rightarrow \infty$ , but it is  $o(\frac{4 \log \bar{T}}{b^2})$  as  $\bar{T} \rightarrow \infty$ . Fig. 4 shows that the ratio  $\bar{\tau}_e^*/(\frac{4 \log \bar{T}}{b^2})$  tends to one as  $\bar{T} \rightarrow \infty$ .<sup>1</sup> This completely confirms the results of Theorem 2.

APPENDIX A  
PROOF OF LEMMA 1

We assume without any loss of generality that  $\Sigma = I$  in model (1), (2). Because the observations  $Y_1, Y_2, \dots, Y_t$  are independant,

<sup>1</sup>To explain this fact, let us recall the definition of the  $\sim$  and  $(1 + o(1))$ -notations. We consider two functions:  $f, g : \mathbb{R}_+ \rightarrow \mathbb{R}$ ,  $\{f(x) \sim g(x) \text{ as } x \rightarrow \infty\} \Leftrightarrow \{f(x) = g(x)(1 + o(1)) \text{ as } x \rightarrow \infty\} \Leftrightarrow \{f(x) = \lambda(x)g(x), \text{ where } \lambda(x) \rightarrow 1 \text{ as } x \rightarrow \infty\}$ .

$$\begin{cases} \text{minimize} & \bar{\tau}^*(m, h) = \max \left\{ \frac{m}{1 - \beta}, \max_{1 \leq l \leq m} \left[ m - l + 1 + \frac{m}{1 - \beta} \mathbf{P} \left( \chi_{r,0}^2 < \frac{m^2 h^2}{m - l + 1} \right) \right] \right\} \\ \text{subject to:} & \bar{T} - m [1 - \mathbf{P}(\chi_{r,0}^2 < m h^2)]^{-1} = 0, \quad m \geq 1 \quad \text{and} \quad h > 0. \end{cases} \quad (12)$$

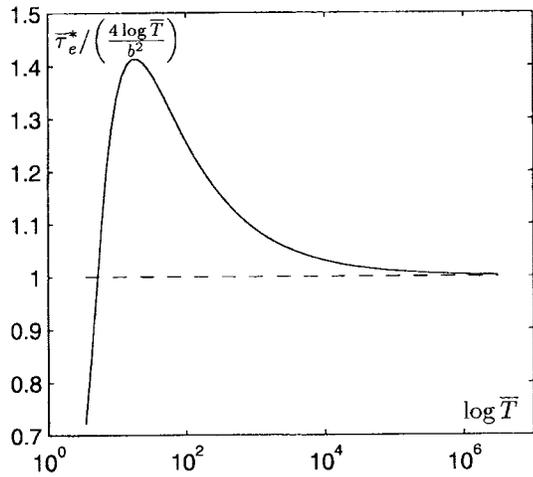


Fig. 4. Ratio  $\bar{\tau}_e^*/(\frac{4 \log \bar{T}}{b^2})$  (solid line) as a function of  $\log \bar{T}$ . The horizontal dashed straight line shows the limit value  $\bar{\tau}_e^*/(\frac{4 \log \bar{T}}{b^2}) \rightarrow 1$  as  $\log \bar{T} \rightarrow \infty$ .

the ‘worst case’ mean detection delay for the  $\chi^2$ -FSS (3) can be rewritten as

$$\begin{aligned} \bar{\tau}^* &= \sup_{\nu \geq 1} \text{esssup } E_\nu(\bar{N} - \nu + 1 \mid \bar{N} \geq \nu, Y_1^{\nu-1}) \\ &= \max_{1 \leq l \leq m} \text{esssup } E_l(\bar{N} - l + 1 \mid Y_1^{l-1}) \end{aligned} \quad (15)$$

where  $Y_1^k = (Y_1, \dots, Y_k)$ , and  $Y_1^0 = \emptyset$ . If  $l = 1$ , then

$$\text{esssup } E_1(\bar{N} \mid Y_1^0) = m(1 - \beta)^{-1} \quad (16)$$

where  $\beta$  is the probability of miss detection. If  $1 < l \leq m$ , then

$$\begin{aligned} \text{esssup } E_l(\bar{N} - l + 1 \mid Y_1^{l-1}) \\ = m - l + 1 + m(1 - \beta)^{-1} \text{esssup } \mathbf{P}_l(d_1 = 0 \mid Y_1^{l-1}). \end{aligned} \quad (17)$$

Let us consider the conditional probability  $\mathbf{P}_l(d_1 = 0 \mid Y_1^{l-1})$ . It is easy to see that the vector  $\bar{S}_1^m = \sum_{i=1}^{l-1} (Y_i - \theta_0) + \sum_{i=l}^m (Y_i - \theta_0)$  given  $X = \sum_{i=1}^{l-1} (Y_i - \theta_0)$  follows a multivariate normal distribution with (conditional) expectation  $E_{\bar{S}}(X) = (m - l + 1)(\theta_1 - \theta_0) + X$  and variance-covariance matrix  $\Sigma_{\bar{S}} = (m - l + 1)I$ . Hence, we get  $\mathbf{P}_l(d_1 = 0 \mid Y_1^{l-1}) = \mathbf{P}_l(\|\bar{S}_1^m\|_2 < mh \mid X) = \mathbf{P}(\chi_{r,\lambda}^2 < \frac{m^2 h^2}{m-l+1} \mid X)$ , where  $\lambda = (m - l + 1)^{-1} \|(m - l + 1)(\theta_1 - \theta_0) + X\|_2^2$ . The distribution function of a noncentral  $\chi^2$  law is given by [5, ch. 28]

$$\mathbf{P}(\chi_{r,\lambda}^2 < y) = e^{-\frac{\lambda}{2}} \sum_{i=0}^{\infty} \frac{\lambda^i}{2^i i!} \mathbf{P}(\chi_{r+2i,0}^2 < y) \quad (18)$$

where  $0! = 1$ , and  $\mathbf{P}(\chi_{r,0}^2 < y) = \frac{1}{2^{\frac{r}{2}} \Gamma(\frac{r}{2})} \int_0^y x^{\frac{r}{2}-1} e^{-\frac{x}{2}} dx$ . It results from (18) that  $\mathbf{P}(\chi_{r,\lambda}^2 < y) < \mathbf{P}(\chi_{r,0}^2 < y)$  for any  $y, \lambda > 0$ . Therefore, we get

$$\begin{aligned} \text{esssup } \mathbf{P}_l(d_1 = 0 \mid Y_1^{l-1}) \\ = \mathbf{P}_l[d_1 = 0 \mid X = -(m + l - 1)(\theta_1 - \theta_0)] \\ = \mathbf{P}[\chi_{r,0}^2 < m^2 h^2 (m - l + 1)^{-1}]. \end{aligned}$$

Combining this with (15)–(17) yields (10). The proof of Lemma 1 is complete.

#### APPENDIX B PROOF OF THEOREM 2

We apply (10), (11) (where  $E_\infty(\bar{N}) = \bar{T}$ ), and (12) to compute  $m^*$  and  $h^*$ . Equation (10) implies

$$\begin{aligned} m[\mathbf{P}(\chi_{r,mb^2}^2 \geq mh^2)]^{-1} \\ \leq \bar{\tau}^*(m, h) \leq m - 1 + m[\mathbf{P}(\chi_{r,mb^2}^2 \geq mh^2)]^{-1}. \end{aligned} \quad (19)$$

We consider the asymptotic case  $\bar{T} \rightarrow \infty$ . The idea of the proof consists of showing the fact that  $m^*(\bar{T}) \rightarrow \infty$ , and  $h^*(\bar{T}) \sim b$  as  $\bar{T} \rightarrow \infty$ . Let us denote  $x = mh^2$ . First, we will show that the optimal value is  $x^*(\bar{T}) \rightarrow \infty$  as  $\bar{T} \rightarrow \infty$ . From (11) and (19), it results that

$$\begin{aligned} \bar{\tau}^* &\geq \bar{T} \mathbf{P}(\chi_{r,0}^2 \geq mh^2) [\mathbf{P}(\chi_{r,mb^2}^2 \geq mh^2)]^{-1} \\ &\geq \bar{T} \mathbf{P}(\chi_{r,0}^2 \geq mh^2) = \bar{T} \mathbf{P}(\chi_{r,0}^2 \geq x). \end{aligned} \quad (20)$$

It follows from (5) that the asymptotically best result is  $\bar{\tau}^* = O(\log \bar{T})$  as  $\bar{T} \rightarrow \infty$ ; hence, only the case  $x^*(\bar{T}) \rightarrow \infty$  can lead to this solution. Because of this fact, we will use the following tail probability of a central  $\chi^2$  law [3]  $\mathbf{P}(\chi_{r,0}^2 \geq x) \sim \Gamma(\frac{r}{2})^{-1} (\frac{x}{2})^{\frac{r}{2}-1} e^{-\frac{x}{2}}$  as  $x \rightarrow \infty$  to compute the mean time before a false alarm in the remainder of this proof. It follows from (11) that  $\log \bar{T} = \log m - \log \mathbf{P}(\chi_{r,0}^2 \geq mh^2)$ ; hence

$$\begin{aligned} \log \bar{T} &= \log m + \log \Gamma\left(\frac{r}{2}\right) + \frac{x}{2} - \left(\frac{r}{2} - 1\right) \log \frac{x}{2} + o(1) \\ &\text{as } x = mh^2 \rightarrow \infty. \end{aligned} \quad (21)$$

Second, let us show that  $m^*(\bar{T}) \rightarrow \infty$  as  $\bar{T} \rightarrow \infty$ . Let us suppose that  $m^* = \text{const}$ . Because  $x^*(\bar{T}) \rightarrow \infty$  as  $\bar{T} \rightarrow \infty$ , we use the tail probability of a noncentral  $\chi^2$  law with  $\lambda = mb^2$  [3]  $\mathbf{P}(\chi_{r,\lambda}^2 \geq x) \sim e^{-\frac{\lambda}{2}} \Gamma(\frac{r}{2}) G(\frac{r}{2}, \frac{\lambda x}{4}) \mathbf{P}(\chi_{r,0}^2 \geq x)$  as  $x \rightarrow \infty$ . It follows from (21) that  $h(\bar{T}) \sim \sqrt{\frac{2 \log \bar{T}}{m}}$  as  $\bar{T} \rightarrow \infty$ . Combining this with the asymptotic expansion [1, ch. 13]  $G(\frac{r}{2}, y) \sim \frac{\Gamma(r-1)}{\Gamma(\frac{r-1}{2})} e^{2\sqrt{y}} (4\sqrt{y})^{-\frac{r-1}{2}}$  as  $y \rightarrow \infty$  yields  $\bar{\tau}^* \geq O(\bar{T}(\log \bar{T})^{\frac{r-1}{4}} e^{-\sqrt{\log \bar{T}}})$  as  $\bar{T} \rightarrow \infty$ . Therefore, the conditions  $x^*(\bar{T}) \rightarrow \infty$  and  $m^*(\bar{T}) \rightarrow \infty$  as  $\bar{T} \rightarrow \infty$  are *necessary* to obtain an optimal solution.

Now, let us show that  $h^*(\bar{T}) \sim \mathcal{A}$ , where  $\mathcal{A}$  is a positive constant, as  $\bar{T} \rightarrow \infty$ , and then find the optimal value  $h^*(\bar{T})$ . It results from (19) and (21) with  $x = mh^2 \rightarrow \infty$  and  $m \rightarrow \infty$  as  $\bar{T} \rightarrow \infty$  that  $\bar{\tau}^* \geq \frac{2(\log \bar{T} - \log \Gamma(\frac{r}{2}) - o(1))}{[h^2 + o(1)] \mathbf{P}(\chi_{r,mb^2}^2 \geq mh^2)}$  as  $\bar{T} \rightarrow \infty$ . Thus, it is quite obvious that if  $h^*(\bar{T}) \rightarrow 0$  as  $\bar{T} \rightarrow \infty$ , then the denominator in the right side of the last equation tends to 0 when  $\bar{T} \rightarrow \infty$ . Now, let  $h^*(\bar{T}) \rightarrow \infty$  as  $\bar{T} \rightarrow \infty$ . By using Chebyshev’s inequality for  $\mathbf{P}(\chi_{r,mb^2}^2 \geq mh^2)$ , we again have the denominator that tends to 0, and therefore, the optimal parameter  $h^*(\bar{T})$  must be  $\sim \mathcal{A}$  as  $\bar{T} \rightarrow \infty$ . It then follows that  $m^*(\bar{T}) \sim \frac{2 \log \bar{T}}{h^2}$  as  $\bar{T} \rightarrow \infty$ . Let  $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_z^\infty \exp\{-\frac{t^2}{2}\} dt$ . Applying a normal approximation to the  $\chi_{r,\lambda}^2$  distribution (see details in [5, ch. 28]) with  $\lambda = mb^2$ , we get  $\mathbf{P}(\chi_{r,mb^2}^2 \geq mh^2) = \Phi(\frac{mh^2 - mb^2 - r}{\sqrt{4mb^2 + 2r}}) + O(\frac{1}{\sqrt{mb^2}})$  as  $\lambda = mb^2 \rightarrow \infty$ , uniformly in  $x = mh^2$ . Combining the last result with (19) and  $m^* \sim \frac{2 \log \bar{T}}{h^2}$ , we obtain  $\bar{\tau}^*(h, \bar{T}) \leq \frac{2f(h, \bar{T}) \log \bar{T}}{b^2} (1 + o(1))$  as  $\bar{T} \rightarrow \infty$  with  $f(h, \bar{T}) = \{\frac{b^2}{h^2} + \frac{b^2}{h^2} [\Phi(\frac{\sqrt{2 \log \bar{T}}}{2} (\frac{h}{b} - \frac{b}{h})) + O(\frac{h}{b \sqrt{2 \log \bar{T}}})]^{-1}\}$ . Because  $h^*(\bar{T}) \sim \mathcal{A}$ ,  $f(h^*, \bar{T}) \sim \hat{f}(h^*, \bar{T}) = \{\frac{b^2}{h^{*2}} + \frac{b^2}{h^{*2}} [\Phi(\frac{\sqrt{2 \log \bar{T}}}{2} (\frac{h^*}{b} - \frac{b}{h^*}))]^{-1}\}$  as  $\bar{T} \rightarrow \infty$ . Let us denote  $g(\bar{T}) = \min_{h>0} \hat{f}(h, \bar{T})$  and  $h^* = \arg \min_{h>0} \hat{f}(h, \bar{T})$ . It can be shown that  $\lim_{\bar{T} \rightarrow \infty} g(\bar{T}) = 2$  and  $h^* \sim b \{1 + 2\sqrt{[\log(\sqrt{2 \log \bar{T}}/2) - \log 2\sqrt{2\pi}]/\log \bar{T}}\}^{-1/2}$  as  $\bar{T} \rightarrow \infty$ . This implies (13), and the theorem is proved. Let us add the following remark about a relation between  $\bar{\tau}^*(\bar{T})$  and the asymptotic lower bound  $n(\bar{T}) \sim \frac{2 \log \bar{T}}{b^2}$ . Applying the same approach to the left side inequality of (19), we get an obvious result  $\frac{2 \log \bar{T}}{b^2} (1 + o(1)) \leq \bar{\tau}^*(\bar{T}) \leq \frac{4 \log \bar{T}}{b^2} (1 + o(1))$  as  $\bar{T} \rightarrow \infty$ .

#### ACKNOWLEDGMENT

The author is grateful to Reviewer 1 for his very helpful comments that improved the paper.

## REFERENCES

- [1] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, U.S. Dept. Commerce, Nat. Bureau Stand., Applied Math. Series 55, 1964.
- [2] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes. Theory and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [3] R. Beran, "Tail probabilities of noncentral quadratic forms," *Ann. Stat.*, vol. 3, no. 4, pp. 969–974, 1975.
- [4] T. S. Ferguson, *Mathematical Statistics. A Decision Theoretic Approach*. New York: Academic, 1967.
- [5] N. L. Johnson and S. Kotz, *Continuous Univariate Distributions—2*. New York: Wiley, 1970.
- [6] T. L. Lai, "Sequential changepoint detection in quality control and dynamical systems," *J. R. Stat. Soc. B*, vol. 57, no. 4, pp. 613–658, 1995.
- [7] G. Lorden, "Procedures for reacting to a change in distribution," *Ann. Math. Stat.*, vol. 42, pp. 1897–1908, 1971.
- [8] G. Moustakides, "Optimal procedures for detecting changes in distributions," *Ann. Stat.*, vol. 14, pp. 1379–1387, 1986.
- [9] I. V. Nikiforov, "On first order optimality of a change detection algorithm in a vector case," *Automat. Remote Contr.*, vol. 55, no. 1, pp. 87–105, 1994.
- [10] ———, "Two strategies in the problem of change detection and isolation," *IEEE Trans. Inform. Theory*, vol. 43, pp. 770–776, Mar. 1997.
- [11] Y. Ritov, "Decision theoretic optimality of the CUSUM procedure," *Ann. Stat.*, vol. 18, pp. 1464–1469, 1990.
- [12] A. N. Shiryaev, "The problem of the most rapid detection of a disturbance in a stationary regime," *Sov. Math. Dokl.*, no. 2, pp. 795–799, 1961.
- [13] A. Wald, *Sequential Analysis*. New York: Wiley, 1947.

## The CFAR Adaptive Subspace Detector is a Scale-Invariant GLRT

Shawn Kraut and Louis L. Scharf

**Abstract**—The constant false alarm rate (CFAR) matched subspace detector (CFAR MSD) is the uniformly most-powerful-invariant test and the generalized likelihood ratio test (GLRT) for detecting a target signal in noise whose covariance structure is known but whose level is unknown. Recently, the CFAR adaptive subspace detector (CFAR ASD), or adaptive coherence estimator (ACE), was proposed for detecting a target signal in noise whose covariance structure and level are both unknown and whose covariance structure is estimated with a sample covariance matrix based on training data. We show here that the CFAR ASD is GLRT when the test measurement is not constrained to have the same noise level as the training data. As a consequence, this GLRT is invariant to a more general scaling condition on the test and training data than the well-known GLRT of Kelly.

**Index Terms**—Adaptive arrays, matched filters, maximum likelihood detection, multidimensional signal detection, radar detection.

### I. INTRODUCTION

Recently, we have suggested the constant false alarm rate (CFAR) adaptive subspace detector (CFAR ASD) [3] for detecting a target signal  $\psi$  in a complex multivariate measurement  $\underline{y}$  whose distribution is complex normal  $\underline{y} \sim CN[\mu e^{j\alpha} \psi, \sigma^2 \mathbf{R}]$ . The signal scaling  $\mu$  determines the null hypothesis  $H_0: \mu = 0$  and alternate hypothesis  $H_1: \mu > 0$ . We factor out a noise scaling  $\sigma^2$  from the noise covariance structure  $\mathbf{R}$ : a step to be clarified in the subsequent discussion.

When the noise covariance structure and scaling  $\mathbf{R}$  and  $\sigma^2$  are both known, the appropriate noncoherent detection statistic is the matched filter magnitude-squared or the matched subspace detector (MSD). This uses the inner product of the whitened measurement  $\underline{z} = \mathbf{R}^{-(1/2)} \underline{y}$  with the whitened signal template  $\underline{\phi} = \mathbf{R}^{-(1/2)} \psi$

$$\chi^2 = \frac{|\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{y}|^2}{\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{\psi} \sigma^2} = \frac{\underline{z}^\dagger \mathbf{P}_\phi \underline{z}}{\sigma^2} \geq \eta \quad (1)$$

where  $\mathbf{P}_\phi = \underline{\phi}(\underline{\phi}^\dagger \underline{\phi})^{-1} \underline{\phi}^\dagger$  is the projection onto  $\underline{\phi}$ . This statistic is complex chi-squared (or gamma) distributed; the MSD compares it with the threshold  $\eta$  to decide on hypothesis  $H_0$  or  $H_1$ .

When the covariance matrix  $\mathbf{R}$  is known but the scaling  $\sigma^2$  is unknown, the MSD may be normalized by the magnitude squared of the measurement weighted by  $\mathbf{R}^{-1}$ . This measures the direction-cosine squared of the angle that  $\underline{z}$  makes with  $\underline{\phi}$ :

$$\cos^2 = \frac{|\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{y}|^2}{(\underline{\psi}^\dagger \mathbf{R}^{-1} \underline{\psi})(\underline{y}^\dagger \mathbf{R}^{-1} \underline{y})} \geq \eta. \quad (2)$$

This statistic has a "beta" density under  $H_0$ ; under  $H_1$ , it is most clearly described as a monotone function of a statistic with a scaled noncentral "F" distribution

$$\cos^2 = \frac{F}{F + 1}; \quad F = \frac{\underline{z}^\dagger \mathbf{P}_\phi \underline{z}}{\underline{z}^\dagger \mathbf{P}_\phi^\perp \underline{z}} \quad (3)$$

Manuscript received August 4, 1998; revised November 25, 1998. This work was supported by the Office of Naval Research under Contract N00014-89-J-1070 and by the National Science Foundation under Contract MIP-9529050. The associate editor coordinating the review of this paper and approving it for publication was Dr. Lal C. Godara.

The authors are with the Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309-0425 USA.

Publisher Item Identifier S 1053-587X(99)06753-7.