Quickest Detection of a Random Signal in Background Noise Using a Sensor Array

Taragay Oskiper

Zargis Medical Corp., Princeton, NJ 08540, USA
Email: toskiper@zargis.com

H. Vincent Poor

Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA
Email: poor@princeton.edu

The problem of detecting the onset of a signal impinging at an unknown angle on a sensor array is considered. An algorithm based on parallel CUSUM tests matched to each of a set of discrete beamforming angles is proposed. Analytical approximations are developed for the mean time between false alarms, and for the detection delay of this algorithm. Simulations are included to verify the results of this analysis.

Keywords and phrases: sensor array, quickest detection, angle-of-arrival estimation.

1. INTRODUCTION

In this paper, we consider the problem of detecting, as soon as possible, a target that appears abruptly and at an unknown angle in a sensor array. This is a problem that arises in a number of applications including radar, sonar, and communications. For a fixed angle of incidence and known signal and noise distributions, this is a classical problem in statistical change detection, and can be solved, for example, by the Page's CUSUM algorithm. However, here we consider the situation in which the angle of incidence and the signal and noise statistics are unknown. In this case, alternatives to the classic CUSUM must be considered, and a number of such methods have been developed for such problems [1, 2, 3, 4, 5].

Here, we use an approach motivated by Nikiforov [4] in which we discretize the set of incidence angles and run parallel change-detection algorithms, each one matched to a beamformer pointed at a particular angle. The presence of a signal is announced the first time the test statistic associated with any of these parallel algorithms crosses a threshold. The angle of incidence is then estimated as the pointing angle corresponding to the first test to detect. This test can be analyzed by adapting the methodology of Lorden [6], and we do so by deriving expressions for the mean time between false alarms and the asymptotic mean detection delay for our test. We include a number of simulation results to verify these expressions and to illustrate further properties of the proposed

algorithm, including the effects on the performance of increasing the number of array elements.

This paper is organized as follows. In Section 2, we describe a model for the problem of interest, including relevant performance criteria. In Section 3, we review briefly the action and properties of the classic Page's CUSUM test to provide a framework for our algorithm. Section 4 develops our parallel beamformer-based CUSUM algorithm, while Section 5 contains an analysis of the algorithm under the assumption of Gaussian noise. We also measure the performance of the proposed method against the optimal algorithm that has perfect knowledge of the signal and noise distributions together with the direction of arrival, for the case when both the signal and noise are Gaussian distributed. Section 6 discusses simulation results that illustrate the algorithm's properties. And, finally, Section 7 contains some concluding remarks.

2. STATEMENT OF THE PROBLEM

We assume a uniform linear sensor array with L elements and consider the following signal model:

$$\mathbf{y}_i = \begin{cases} \mathbf{n}_i, & i = 1, 2, \dots, \nu - 1, \\ \mathbf{a}(\phi)X_i + \mathbf{n}_i, & i = \nu, \nu + 1, \dots, \end{cases} \quad (1)$$

where ν is an unknown change point after which $\{X_i\}$, an independent identically distributed (i.i.d.) narrowband
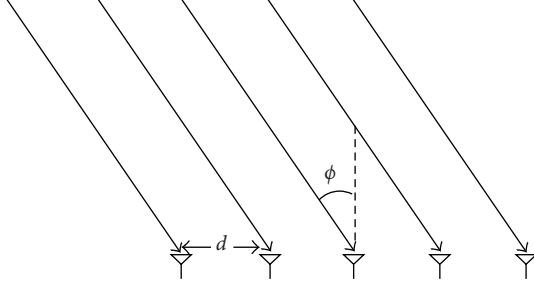
FIGURE 1: Linear sensor array with 5 elements with a narrowband source in the far-field impinging on the array from a direction ϕ.

complex-valued random signal source, is incident on the array (see Figure 1) at an unknown angle $\phi \in [-\pi/2, \pi/2)$, and $\mathbf{a}(\phi)$ is the $L \times 1$ array response vector (also called the steering vector) associated with it. The array response has the following form:

$$\mathbf{a}(\phi) = \begin{bmatrix} 1 & e^{-j2\pi(d/\lambda)\sin(\phi)} & \cdots & e^{-j(L-1)2\pi(d/\lambda)\sin(\phi)} \end{bmatrix}^T, \quad (2)$$

where λ is the wavelength and d is the sensor spacing, typically chosen as half the wavelength. Finally, $\{\mathbf{n}_i\}$ is the ambient noise independent of the source signal, and white in both space and time, with covariance matrix $\sigma_n^2 \mathbf{I}_L$.

According to the above model, before an unknown time instant ν (the change point), there is only noise in the system, and after the change point, a random signal appears at an angle ϕ in addition to the noise. We wish to detect the appearance of this random signal as soon as possible and also to estimate the angle of arrival of this source. In particular, we would like to design a detection algorithm that does not rely on the knowledge of the distribution of the random process $\{X_i\}$ and the noise process $\{\mathbf{n}_i\}$, except that the noise variance is assumed to be known. In the following, we pose this problem formally as a quickest change detection problem [1, 6], and define the criteria involved in designing an algorithm for this purpose.

Let $\mathbb{P}_{(\nu)}$ denote the distribution of the sequence of observations $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{\nu-1}, \mathbf{y}_\nu, \mathbf{y}_{\nu+1}, \ldots$, where ν is the change point, and let $\mathbb{E}_{(\nu)}$ denote expectation under $\mathbb{P}_{(\nu)}$. We assume that, under $\mathbb{P}_{(\nu)}$, the random variables $\{\mathbf{y}_i\}$ are independent with a marginal probability density function (pdf) p_0 for $i < \nu$ and $p_1 \neq p_0$ for $i \geq \nu$. Let $\mathbb{P}_0$ correspond to the case where $\nu = \infty$, that is, $\{\mathbf{y}_i\} \sim p_0$ for all $i \geq 1$ (the no-change situation), and let expectation under $\mathbb{P}_0$ be denoted by $\mathbb{E}_0$. Also, we use $\mathbb{P}_1$ and $\mathbb{E}_1$ instead of $\mathbb{P}_{(1)}$ and $\mathbb{E}_{(1)}$ for the case when $\nu = 1$, that is, $\{\mathbf{y}_i\} \sim p_1$ for all $i \geq 1$. The goal is to minimize, over all possible stopping times N, the worst-case mean delay for detection,

$$\bar{\tau}(N) = \sup_{\nu \geq 1} \text{ess sup } \mathbb{E}_{(\nu)}\{N - \nu + 1 \mid N \geq \nu, \mathbf{y}_1, \ldots, \mathbf{y}_{\nu-1}\}, \quad (3)$$

such that the mean time before a false alarm satisfies

$$\mathbb{E}_0\{N\} \geq \gamma \quad (4)$$

for a given $\gamma > 0$. So, the idea is to detect the presence of a change as soon as possible while keeping the false alarm rate below a desired level.

3. PAGE'S CUSUM TEST

Before moving on to the more complex case of composite hypotheses, in the following, we review the basic case in which both the pre- and postchange hypotheses are simple. Later on, we will discretize the parameter space of the angle of arrival, reduce the composite alternative to a set of simple hypotheses in that parameter, and apply parallel simple-change detection tests.

When the likelihood ratio of the observations under the two hypotheses can be written explicitly, the following algorithm, called Page's CUSUM test [7], is known to be optimal in the minimax sense described above [6, 8]. Page's test declares the detection of a change point the first time the CUSUM statistic

$$\max_{1 \leq n \leq i} G_n^i, \quad (5)$$

or the equivalent and computationally efficient recursive form

$$Q_i = (Q_{i-1} + g(\mathbf{y}_i))^+ \quad (Q_0 = 0) \quad (6)$$

exceeds a threshold $h > 0$, where $G_n^i = \sum_{m=n}^{i} g(\mathbf{y}_m)$ and $g(\mathbf{y}_i) = \log p_1(\mathbf{y}_i)/p_0(\mathbf{y}_i)$ is the log-likelihood function (or score function) which should satisfy $0 < \rho \triangleq \mathbb{E}_1\{g(\mathbf{y}_i)\} < \infty$. Recall that $\rho = \mathbb{E}_1\{g(\mathbf{y}_i)\}$ is the Kullback-Leibler distance between the two densities and is always positive, while $\mathbb{E}_0\{g(\mathbf{y}_i)\}$ is always negative.

Thus, the stopping time N of the CUSUM algorithm is given by

$$N = \inf\left\{i \geq 1 : \max_{1 \leq n \leq i} G_n^i \geq h\right\} \quad (7)$$

or, equivalently,

$$N = \inf\{i \geq 1 : Q_i \geq h\}, \quad (8)$$

where the equivalence is true under the condition that we use $h > 0$. For this algorithm, we have the following well-known result of Lorden [6]:

$$\mathbb{E}_0\{N\} \geq e^h, \quad \bar{\tau}(N) = \mathbb{E}_1\{N\} \sim \frac{h}{\rho} \quad \text{as } h \longrightarrow \infty. \quad (9)$$

Note that for the CUSUM stopping rule, the worst-case mean detection delay corresponds to the change point $\nu = 1$, since this is when the CUSUM statistic $Q_i \geq 0$, for all i, is at its minimum ($Q_0 = 0$) and hence is the farthest from the threshold $h > 0$.

The idea behind the CUSUM algorithm is that it stops at the first time instant i such that for some $n \leq i$, the log-likelihood ratio test to decide between the hypotheses $H_0[i] : \mathbf{y}_n, \ldots, \mathbf{y}_i \sim$ i.i.d. p_0 and $H_{(n)}[i] : \mathbf{y}_n, \ldots, \mathbf{y}_i \sim$ i.i.d. p_1 exceeds a certain threshold. The basic operating principle of

the recursive form (6) is that, before the change, $\mathbb{E}_0\{g(\mathbf{y}_i)\} < 0$, so that Q_i remains close to zero; whereas after the change, Q_i starts drifting upward with a positive mean $\rho = \mathbb{E}_1\{g(\mathbf{y}_i)\}$ until it ultimately crosses the threshold h.

In general, when the likelihood ratio is not known explicitly, which we assume to be so in our situation, the score function $g(\cdot)$ can be replaced by any other function with negative mean before the change, and positive mean after the change; that is, satisfying the conditions $\mathbb{E}_0\{g(\mathbf{y}_i)\} < 0$ and $\mathbb{E}_1\{g(\mathbf{y}_i)\} > 0$. In this case, the stopping time is no longer guaranteed to be optimal but still a very good candidate once an appropriate function is chosen [1, 9, 10] satisfies

$$\mathbb{E}_0\{N\} \geq e^{s_0 h}, \quad \bar{\tau}(N) = \mathbb{E}_1\{N\} \sim \frac{h}{\mathbb{E}_1\{g(\mathbf{y}_i)\}} \quad \text{as } h \longrightarrow \infty, \tag{10}$$

where $s_0 > 0$ is the (nontrivial) root of the equation $\mathbb{E}_0\{e^{sg(\mathbf{y}_i)}\} = 1$, which exists and is unique under certain regularity conditions.

4. PARALLEL BEAMFORMER-BASED CUSUM ALGORITHM

Since the angle of arrival ϕ after the change point is unknown, we have a composite alternative hypothesis in that parameter. Motivated by Nikiforov's approach [4] for detecting a change under such a condition, we propose a simple scheme in which we run simultaneously K parallel CUSUM algorithms, each using a conventional beamformer. Since we assume no knowledge of the probability distributions of the target signal and noise, this suboptimal method acts as an energy detection scheme with each CUSUM "tuned" for the detection of signal energy from a particular direction of arrival.

The array weight vector, $\mathbf{w}(\theta)$, for the conventional beamformer [11] (also called the fixed-phased array beamformer) is given by $\mathbf{w}(\theta) = \mathbf{a}(\theta)/L$, where $\mathbf{a}(\theta)$, defined in (2), is the array response or the steering vector associated with a source incident at angle θ. Hence, the output of the conventional beamformer is unity in the look direction $\mathbf{w}(\theta)^H \mathbf{a}(\theta) = (1/L)\mathbf{a}(\theta)^H \mathbf{a}(\theta) = 1$, where $\mathbf{w}(\theta)^H$ is the conjugate transpose of $\mathbf{w}(\theta)$. Note that the beamformer response is maximal in the look direction θ, that is,

$$\frac{1}{L}\mathbf{a}(\theta)^H \mathbf{a}(\phi) \leq 1 \quad \forall \theta, \phi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right), \tag{11}$$

with equality if and only if $\theta = \phi$. In general, given an array weight vector $\mathbf{w}(\theta)$, the function

$$z(\theta, \phi) = \left|\mathbf{w}(\theta)^H \mathbf{a}(\phi)\right| \tag{12}$$

for fixed θ, and as a function of ϕ, is called the *beampattern* corresponding to the beamformer pointing in the direction θ, and it is the collection of that beamformer's responses as the angle of incidence varies over ϕ. On the other hand, for fixed ϕ, $z(\theta, \phi)$, as a function of θ, is called the *steered response* corresponding to the angle of incidence ϕ, and it is the collection of beamformer's responses as the look direction varies over θ.

To devise a test for (1), we discretize the parameter space $[-\pi/2, \pi/2]$ into K angles $\{\theta_1, \theta_2, \ldots, \theta_K\}$ such that $-\pi/2 \leq \theta_1 < \cdots < \theta_K < \pi/2$. We can then design a CUSUM test for the detection of a target incident from each of such angles, operate them in parallel to provide the coverage for the whole space, and then combine them into a single change-detection algorithm. For $d = \lambda/2$, the fixed-phased array beamformer pointing in the direction θ_k is defined as

$$\mathbf{w}_k = \frac{1}{L}\begin{bmatrix} 1 & e^{-j\pi \sin(\theta_k)} & \cdots & e^{-j\pi(L-1)\sin(\theta_k)} \end{bmatrix}^T,$$
$$\text{for } \theta_k \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right), \ k = 1, \ldots, K. \tag{13}$$

The stopping time, N, of our parallel beamformer-based CUSUM test is then given by

$$N = \inf\{i \geq 1 : \bar{Q}_i \geq h\}, \tag{14}$$

with $\bar{Q}_i$ defined as follows:

$$\bar{Q}_i \triangleq \max_{1 \leq k \leq K} Q_i^k, \tag{15}$$

and, for each $1 \leq k \leq K$, the CUSUM statistic Q_i^k is given by

$$Q_i^k = \left(Q_{i-1}^k + g_i(k)\right)^+ \quad \text{with } Q_0^k = 0, \tag{16}$$

where we use the following equation:

$$g_i(k) = \left|\mathbf{w}_k^H \mathbf{y}_i\right|^2 - \frac{\sigma_n^2}{L} - c. \tag{17}$$

Defining the function $g_i(k)$ in the above fashion makes each CUSUM test act as an energy accumulator "tuned" to the direction θ_k, which will make the alarm go off in the presence of a target once it starts drifting upward collecting signal energy coming from the look direction. In order to see this, first, we examine the behavior under the prechange (noise only) case. Notice that, since $\mathbf{y}_i = \mathbf{n}_i$ for $i < \nu$, we have $\mathbb{E}_{(\nu)}\{\mathbf{y}_i \mathbf{y}_i^H\} = \sigma_n^2 \mathbf{I}_L$, and using the fact that $\mathbf{w}_k^H \mathbf{w}_k = 1/L$, we obtain $\mathbb{E}_{(\nu)}\{g_i(k)\} = \mathbf{w}_k^H \mathbb{E}_{(\nu)}\{\mathbf{y}_i \mathbf{y}_i^H\}\mathbf{w}_k - \sigma_n^2/L - c = -c$. Here, the bias term c must be chosen to satisfy the condition $c > 0$, so that the expected value of $g_i(k)$ is negative when there is no target present, that is, $\mathbb{E}_{(\nu)}\{g_i(k)\} < 0$ for $i < \nu$. This negative drift is needed to keep the CUSUM statistic Q_i^k, $1 \leq k \leq K$, (16) close to zero before the target presence. Next, looking at the postchange behavior (signal plus noise), it is easy to see that, since $\mathbf{y}_i = \mathbf{a}(\phi)X_i + \mathbf{n}_i$ for $i \geq \nu$, and $\{X_i\}$ and $\{\mathbf{n}_i\}$ are independent, we have $\mathbb{E}_{(\nu)}\{g_i(k)|\phi\} = |\mathbf{w}_k^H \mathbf{a}(\phi)|^2 \sigma_s^2 - c$, where $\sigma_s^2 = \mathbb{E}\{|X_i|^2\}$ is the target signal energy and $|\mathbf{w}_k^H \mathbf{a}(\phi)|$ is the kth beamformer's response for the signal coming from direction ϕ. Now, we can see that the bias term $c > 0$ plays an important role in the postchange situation and should be chosen based on the designed set of beamformers and the minimum signal to-noise-ratio (SNR) requirements so that it satisfies the condition $c < \max_{1 \leq k \leq K} |\mathbf{w}_k^H \mathbf{a}(\phi)|^2 \sigma_s^2$, for all ϕ. Then, for any angle of arrival ϕ, we have $\max_{1 \leq k \leq K} \mathbb{E}_{(\nu)}\{g_i(k)|\phi\} > 0$ for $i \geq \nu$, which will guarantee that at least one of the CUSUM statistics (16) will start drifting upward with a positive mean that
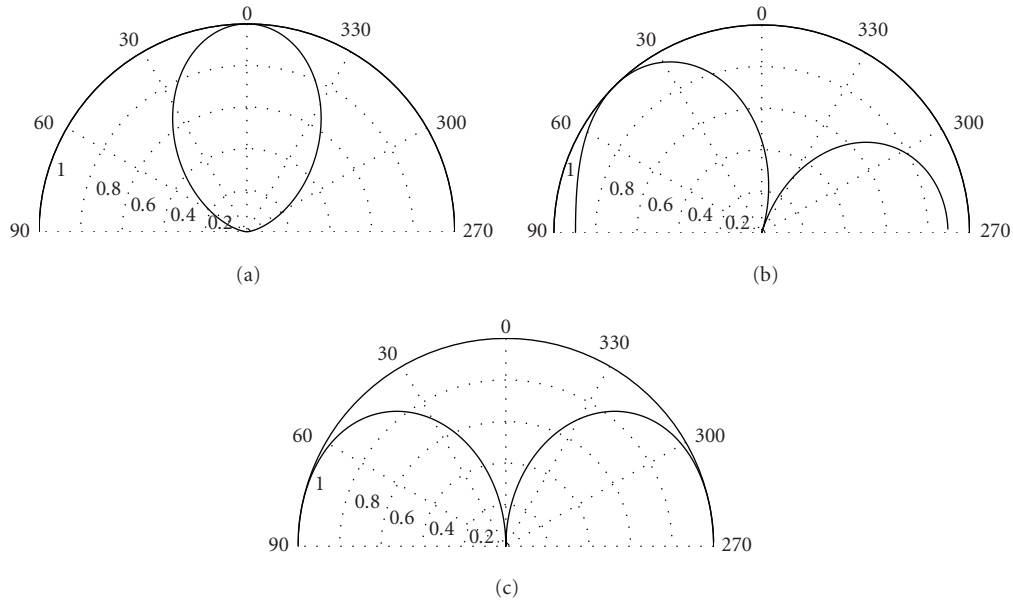
(a)

(b)

(c)

FIGURE 2: Beampattern samples for a two-sensor array ($L = 2$) for look directions (a) $\theta = 0°$, (b) $\theta = 45°$, and (c) $\theta = 90°$.
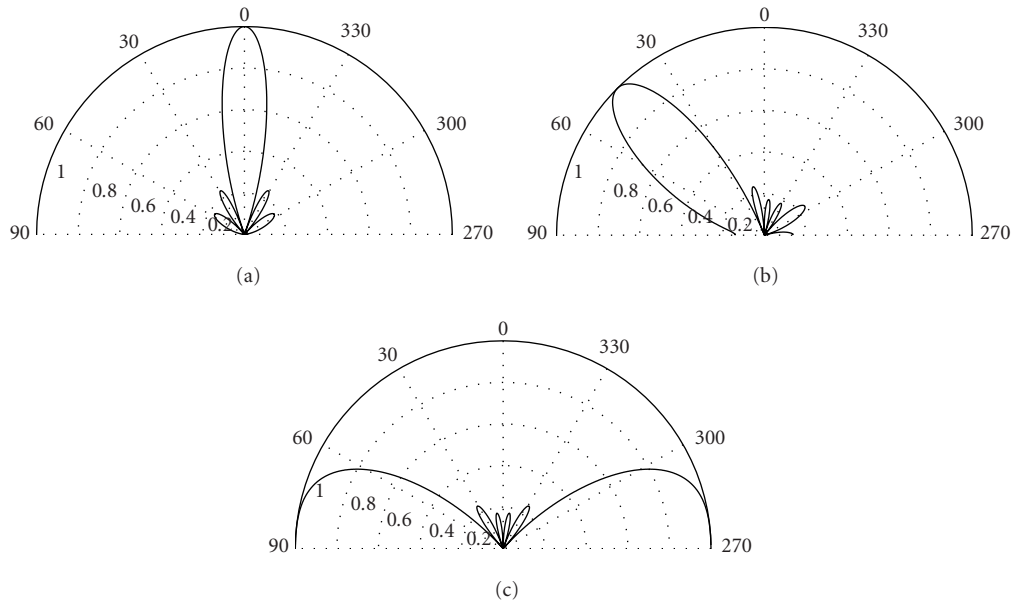

(a)

(b)

(c)

FIGURE 3: Beampattern samples for a six-sensor array ($L = 6$) for look directions (a) $\theta = 0°$, (b) $\theta = 45°$, and (c) $\theta = 90°$.

is proportional to the target signal energy σ_s^2, and the beamformer response for that direction $|\mathbf{w}_k^H \mathbf{a}(\phi)|^2$, enabling the detection of the target as soon as it exceeds a certain threshold value h. In particular, let the angle of arrival of the incoming signal ϕ^* be such that for some ℓ, $\theta_\ell = \phi^*$, namely, we assume that the incoming angle of arrival matches exactly one of the look directions in our set of beamformers. Then, the ℓth CUSUM test based on that beamformer is expected to be responsible for the detection of the target since that beamformer will have the unity (and maximal) response com-

pared to the others, that is, $|\mathbf{w}_\ell^H \mathbf{a}(\phi^*)| = 1$, for $\phi^* = \theta_\ell$. Now, if K, the number of beamformers, is chosen large enough, then we can cover the whole interval $[-\pi/2, \pi/2]$, and we do not need to restrict ϕ^* to belong to the set $\{\theta_1, \ldots, \theta_K\}$, that is, ϕ^* can be any real number in the interval. In this situation, the response of the beamformer whose look direction is the closest to ϕ^* will be approximately unity again. This is of course true if, for a given number of sensors L, the number of beamformers K is chosen such that the main lobe width of the beampatterns (see Figures 2a, 2b, 2c, 3a, 3b, and 3c) is

not very small compared to the difference between the beamformer look directions. Namely, for large L, we have better resolution and, accordingly, K must also be chosen proportionately large so that we have close-to-unit response for all angles of arrival ϕ. Thus, each beamformer will have an interval of responsibility for detection, and the union of these intervals will cover the whole region.

We can now proceed to find the mean time to a false alarm for the parallel CUSUM algorithm whose stopping time is given in (14). First, we define the following K stopping times corresponding to each CUSUM rule (16):

$$N^k \triangleq \inf \{i \geq 1 : Q_i^k \geq h\}, \qquad (18)$$

and the final stopping time, which is equivalent to (14), is then given by

$$N = \min_{1 \leq k \leq K} N^k. \qquad (19)$$

Now, using the equivalent representation (7), the stopping time N^k (18) can be expressed as

$$N^k = \min \{T^k(n) : n = 1, 2, \dots\}, \qquad (20)$$

where

$$T^k(n) = \inf \{i \geq n : G_n^i(k) \geq h\} \qquad (21)$$

with $G_n^i(k) = \sum_{m=n}^{i} g_m(k)$. Hence, the stopping time (19) is given by

$$\begin{aligned} N &= \min_{1 \leq k \leq K} \min_{n \geq 1} T^k(n) \\ &= \min_{n \geq 1} \min_{1 \leq k \leq K} T^k(n). \end{aligned} \qquad (22)$$

For any finite n, $T^k(n)$ is the stopping time of a sequential hypothesis test acting on observations $\mathbf{y}_n, \mathbf{y}_{n+1}, \dots$ for which we have [1, 10]

$$\mathbb{P}_0\big(T^k(n) < \infty\big) \leq e^{-s_0(k)h} \qquad (23)$$

which implies

$$\mathbb{P}_0\left(\min_{1 \leq k \leq K} T^k(n) < \infty\right) \leq \sum_{k=1}^{K} e^{-s_0(k)h}, \qquad (24)$$

where $s_0(k)$ is the nonzero root of the following equation:

$$\mathbb{E}_0\{e^{s g_i(k)}\} = 1. \qquad (25)$$

Now, using Lorden's theorem [6] which states that, for a stopping time $\widetilde{T}$ with respect to a sequence of random variables $\mathbf{y}_i$, $i = 1, 2, \dots$, such that $\mathbb{P}_0(\widetilde{T} < \infty) \leq \alpha$, the extended stopping time $\widetilde{N} \triangleq \min\{\widetilde{T}(n) \mid n = 1, 2, \dots\}$, where $\widetilde{T}(n)$ is

obtained by applying $\widetilde{T}$ to $\mathbf{y}_n, \mathbf{y}_{n+1}, \dots$, satisfies $\mathbb{E}_0\{\widetilde{N}\} \geq 1/\alpha$, we get from (20)–(24),

$$\mathbb{E}_0\{N^k\} \geq e^{s_0(k)h}, \qquad (26)$$

and most importantly,

$$\mathbb{E}_0\{N\} \geq \frac{1}{\sum_{k=1}^{K} e^{-s_0(k)h}}. \qquad (27)$$

Next, based on the definition (19), the asymptotic detection delay can be derived using the upper bound

$$\mathbb{E}_1\{N\} \leq \min_{1 \leq k \leq K} \mathbb{E}_1\{N^k\} \qquad (28)$$

together with (10), which leads to

$$\mathbb{E}_1\{N\} \sim \frac{h}{\max_{1 \leq k \leq K} \mathbb{E}_1\{g_i(k)\}} \quad \text{as } h \longrightarrow \infty. \qquad (29)$$

Finally, at the alarm time N, as a byproduct of this algorithm, we can obtain an estimate of the angle of arrival, which follows from the idea that the CUSUM rule corresponding to the beamformer with the largest response (the one whose look direction is the closest to the angle of incidence) will have the sharpest increase and reach the threshold more quickly. Hence the estimate $\hat{\phi}$ is obtained via

$$\hat{\phi} = \theta_{k^*}, \quad k^* = \arg\max_{k} Q_N^k. \qquad (30)$$

In the next section, we analyze the properties of our algorithm in the case where the noise in our model (1) is i.i.d. zero-mean Gaussian, $\mathbf{n}_i \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I}_L)$.

5. ANALYSIS UNDER GAUSSIAN NOISE

We first recall the parallel beamformer-based CUSUM test, presented in (14) with $\bar{Q}_i$ and Q_i^k defined as before, and

$$g_i(k) = \frac{L |\mathbf{w}_k^H \mathbf{y}_i|^2}{\sigma_n^2} - 1 - c, \qquad (31)$$

where it is easy to see that the threshold h and bias c values above are related to the ones in (14), (15), (16), and (17), by a scale factor σ_n^2/L. From now on, we will use the above equivalent form of the CUSUM test, which makes the following analysis and interpretation of the results simpler.

Now, we look at the mean time to a false alarm for each CUSUM test in the parallel algorithm. Under $\mathcal{H}_0$, that is, $\mathbf{y}_i = \mathbf{n}_i$ for $i \geq 1$, we have

$$g_i(k) = \frac{L |\mathbf{w}_k^H \mathbf{n}_i|^2}{\sigma_n^2} - 1 - c, \qquad (32)$$

and the solution to (25) is the nonzero root of the equation

$$\mathbb{E}_0\left\{e^{s(L|\mathbf{w}_k^H\mathbf{n}_i|^2/\sigma_n^2)}\right\}e^{-s(1+c)} = 1. \qquad (33)$$

If we also assume that the real and imaginary components of the zero-mean complex Gaussian noise vector $\mathbf{n}_i = \Re\mathbf{n}_i + j\Im\mathbf{n}_i$ are independent and have equal covariance, that is, $\mathbb{E}\{\Re\mathbf{n}_i\Re\mathbf{n}_i^T\} = \mathbb{E}\{\Im\mathbf{n}_i\Im\mathbf{n}_i^T\} = (\sigma_n^2/2)\mathbf{I}_L$, and $\mathbb{E}\{\Re\mathbf{n}_i\Im\mathbf{n}_i^T\} = 0$, then $Y = L|\mathbf{w}_k^H\mathbf{n}_i|^2/\sigma_n^2$ is an exponential random variable with mean $\mathbb{E}\{Y\} = (L/\sigma_n^2)\mathbf{w}_k^H\mathbb{E}\{\mathbf{n}_i\mathbf{n}_i^H\}\mathbf{w}_k = 1$, where we used the fact that $\mathbf{w}_k^H\mathbf{w}_k = 1/L$. In order to see this, we write $Y = Y_R^2 + Y_I^2$ and obtain

$$\begin{aligned} Y_R &= \frac{\sqrt{L}}{\sigma_n}\left(\Re\mathbf{w}_k^T\Re\mathbf{n}_i + \Im\mathbf{w}_k^T\Im\mathbf{n}_i\right), \\ Y_I &= \frac{\sqrt{L}}{\sigma_n}\left(\Re\mathbf{w}_k^T\Im\mathbf{n}_i - \Im\mathbf{w}_k^T\Re\mathbf{n}_i\right). \end{aligned} \qquad (34)$$

Under the assumptions made above, Y_R and Y_I are zero-mean jointly Gaussian random variables with zero covariance, and hence are independent, that is,

$$\mathrm{Cov}\,(Y_R, Y_I) = \mathbb{E}\{Y_R Y_I\} = -\frac{L}{2}\Re\mathbf{w}_k^T\Im\mathbf{w}_k + \frac{L}{2}\Im\mathbf{w}_k^T\Re\mathbf{w}_k = 0. \qquad (35)$$

Thus, Y is a chi-square random variable with two degrees of freedom, which is equivalent to an exponential random variable. The moment generating function of Y is then given by

$$\mathbb{E}\{e^{sY}\} = \frac{1}{1-s} \qquad (36)$$

and the solution to (25) is the nonzero root of the equation

$$e^{-s(1+c)} = 1 - s \qquad (37)$$

which can be solved numerically. The solution is a function of c and does not depend on k, so we will denote the root as $s_0(c)$ and from (26) and (27), we obtain

$$\mathbb{E}_0\{N^k\} \geq e^{s_0(c)h}, \qquad \mathbb{E}_0\{N\} \geq \frac{e^{s_0(c)h}}{K}. \qquad (38)$$

In Figure 4, we plot the nonzero root s_0 of the above equation as a function of the bias term $c > 0$. Notice that, s_0 approaches 1 rapidly as the bias c increases from value 0.

Next, we will look at the mean detection delay performance. Under the assumption that the incoming signal always coincides with one of the look directions, that is, $\phi = \theta_\ell, 1 \leq \ell \leq K$, the stopping time corresponding to the beamformer that is perfectly tuned to the incoming signal (the one with unit response) will be N^ℓ. Under $\mathcal{H}_1$, that is, when the target is present, we have $\mathbf{y}_i = \mathbf{a}(\theta_\ell)X_i + \mathbf{n}_i$ for $i \geq 1$. Then using the independence of $\{X_i\}$ and $\{\mathbf{n}_i\}$, the fact that $\mathbf{w}_\ell^H\mathbf{a}(\theta_\ell) = 1$, and $\mathbf{w}_\ell^H\mathbf{w}_\ell = 1/L$, we obtain

$$\begin{aligned} \mathbb{E}_1&\{g_i(\ell) \mid \phi = \theta_\ell\} \\ &= \mathbb{E}_1\left\{L\frac{|\mathbf{w}_\ell^H(\mathbf{a}(\theta_\ell)X_i + \mathbf{n}(i))|^2}{\sigma_n^2}\right\} - 1 - c \\ &= L\frac{\sigma_s^2}{\sigma_n^2} - c, \end{aligned} \qquad (39)$$
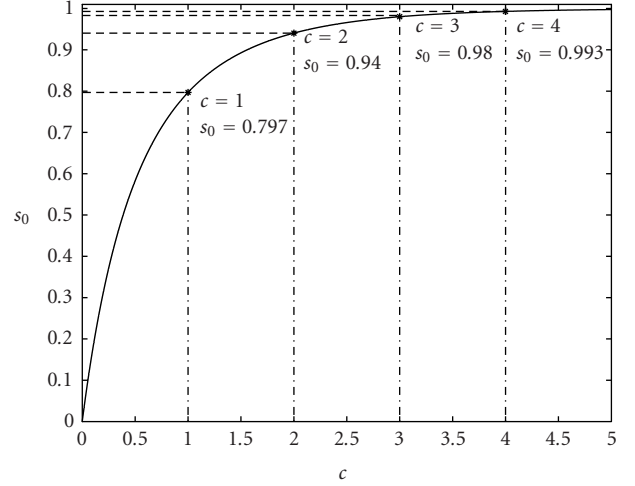

FIGURE 4: Exponent s_0 versus bias c (cf. (37)).

where $\sigma_s^2 = \mathbb{E}\{|X_i|^2\}$ and c is chosen so that it satisfies $0 < c < L\sigma_s^2/\sigma_n^2$. Since the ℓth beamformer's look direction is the closest to the target, clearly we have $\max_{1 \leq k \leq K} \mathbb{E}_1\{g_i(k) \mid \phi = \theta_\ell\} = L\sigma_s^2/\sigma_n^2 - c$, and therefore from (29), we obtain $\mathbb{E}_1\{N \mid \phi = \theta_\ell\} \sim h/(L\sigma_s^2/\sigma_n^2 - c)$ as $h \to \infty$, which does not depend on ℓ, and hence

$$\mathbb{E}_1\{N\} \sim \frac{h}{L\sigma_s^2/\sigma_n^2 - c} \quad \text{as } h \longrightarrow \infty. \qquad (40)$$

Now, in the case when the incoming signal is not constrained to belong to the discrete set of look directions, we will get

$$\mathbb{E}_1\{N\} \sim \frac{h}{\beta L\sigma_s^2/\sigma_n^2 - c} \quad \text{as } h \longrightarrow \infty, \qquad (41)$$

where

$$\beta = \min_\phi \max_{1 \leq k \leq K} |\mathbf{w}_k^H\mathbf{a}(\phi)|^2 \qquad (42)$$

will be close to unity, according to the preceding discussion in Section 4, provided that a sufficient number of beamformers K is used according to the given array size L such that for every angle of arrival ϕ, the beamformer with the closest look direction has approximately unity response. Of course, this happens when the main lobes of our set of beamformers considerably overlap with each other.

Now, assuming h is large enough that the asymptotically linear relationship applies, we can make the following observations. Mean detection delay is inversely proportional to the number of antenna elements L and increases linearly with the threshold value h; whereas, the lower bound for mean time to a false alarm is independent of L and increases exponentially with h. So, for a given SNR value, keeping the threshold h and bias value c fixed, we can reduce the mean detection delay by increasing L while satisfying the same false alarm requirement, or by increasing L and keeping c fixed, we can achieve the same mean detection delay with a higher h, yielding an exponential reduction in the false alarm rate.

Next, for the case where both the signal and noise are Gaussian, we will compare the performance of the proposed method to that of the optimal detector that has perfect knowledge of the signal and noise distributions and the angle of incidence. Then, as described in Section 3, the optimal CUSUM test takes the following form:

$$N^* = \inf \{i \geq 1 : Q_i^* \geq h\} \qquad (43)$$

with $Q_i^* = (Q_{i-1}^* + g^*(\mathbf{y}_i))^+$ and $g^*(\mathbf{y}_i) = \log p_1(\mathbf{y}_i)/p_0(\mathbf{y}_i)$. So, we assume that under $\mathcal{H}_0$, we have $\mathbf{y}_i = \mathbf{n}_i$ and $\mathbf{y}_i \sim \mathcal{N}(0, \mathbf{\Sigma}_0)$, where $\mathbf{\Sigma}_0 = \sigma_n^2 \mathbf{I}_L$, and under $\mathcal{H}_1$, we have $\mathbf{y}_i = \mathbf{a}(\phi)X_i + \mathbf{n}_i$ and $\mathbf{y}_i \sim \mathcal{N}(0, \mathbf{\Sigma}_1(\phi))$, where $\mathbf{\Sigma}_1(\phi) = \sigma_s^2 \mathbf{a}(\phi)\mathbf{a}(\phi)^H + \sigma_n^2 \mathbf{I}_L$. Then

$$
\begin{aligned}
g^*(\mathbf{y}_i \,|\, \phi) &= \log \frac{p_1(\mathbf{y}_i \,|\, \phi)}{p_0(\mathbf{y}_0 \,|\, \phi)} \\
&= -\frac{1}{2} \log \frac{|\mathbf{\Sigma}_1(\phi)|}{|\mathbf{\Sigma}_0|} - \frac{1}{2} \mathbf{y}_i^H (\mathbf{\Sigma}_1^{-1}(\phi) - \mathbf{\Sigma}_0^{-1})\mathbf{y}_i.
\end{aligned}
\qquad (44)
$$

It is easy to see that the determinants $|\mathbf{\Sigma}_1(\phi)|$ and $|\mathbf{\Sigma}_0|$ are given by $|\mathbf{\Sigma}_0| = (\sigma_n^2)^L$ and $|\mathbf{\Sigma}_1(\phi)| = (\sigma_n^2)^{L-1}(L\sigma_s^2 + \sigma_n^2)$, where we used the fact that $\mathbf{a}(\phi)^H \mathbf{a}(\phi) = L$. Then

$$
\begin{aligned}
\rho^*(\phi) &= \mathbb{E}_1\{g(\mathbf{y}_i \,|\, \phi)\} \\
&= -\frac{1}{2} \log \frac{|\mathbf{\Sigma}_1(\phi)|}{|\mathbf{\Sigma}_0|} - \frac{1}{2} \operatorname{trace}\left((\mathbf{\Sigma}_1^{-1}(\phi) - \mathbf{\Sigma}_0^{-1})\mathbf{\Sigma}_1(\phi)\right) \\
&= -\frac{1}{2} \log \left(1 + L\frac{\sigma_s^2}{\sigma_n^2}\right) + \frac{1}{2}L\frac{\sigma_s^2}{\sigma_n^2},
\end{aligned}
\qquad (45)
$$

where we have used the fact that $\mathbb{E}_1\{\mathbf{y}_i\mathbf{y}_i^H \,|\, \phi\} = \mathbf{\Sigma}_1(\phi)$. Note that ρ is independent of direction of arrival ϕ. Hence, from (9), the mean detection delay for the optimal detector is

$$\mathbb{E}_1\{N^*\} \sim \frac{h}{1/2\,(L\sigma_s^2/\sigma_n^2) - 1/2\log(1 + L\sigma_s^2/\sigma_n^2)} \quad \text{as } h \longrightarrow \infty, \qquad (46)$$

and the lower bound on the mean time to a false alarm is given by

$$\mathbb{E}_0\{N^*\} \geq e^h. \qquad (47)$$

And for the proposed parallel CUSUM method, we have

$$\mathbb{E}_0\{N\} \geq \frac{e^{s_0(c)h}}{K}, \quad \mathbb{E}_1\{N\} \sim \frac{h}{\beta L\sigma_s^2/\sigma_n^2 - c} \quad \text{as } h \longrightarrow \infty. \qquad (48)$$

Now, we can investigate under what circumstances the proposed method's performance is close to that of the optimal one. First, we know from Figure 4 that, as the bias term c increases, the exponent $s_0(c)$ will get closer to 1, and hence the false alarm performance of the parallel CUSUM method
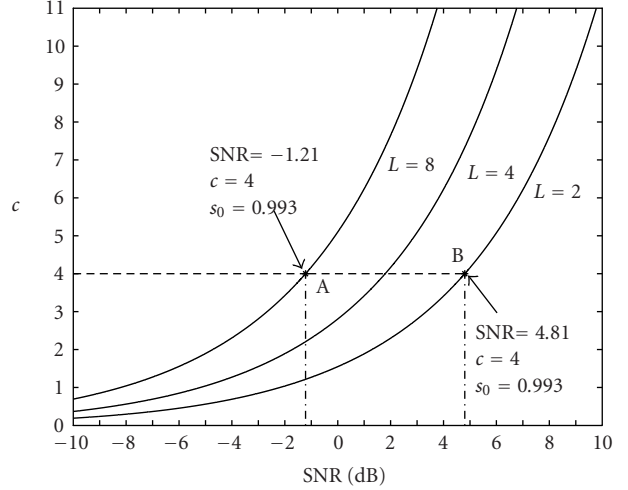

FIGURE 5: Bias values for optimal mean detection delay versus σ_s^2/σ_n^2 (SNR) (cf. (50)).

will get closer to the optimal algorithm, assuming the performance is measured with respect to the exponent term in the lower bound expression for the mean time to a false alarm. On the other hand, the detection delay performance is also related to the choice of c, and given the antenna array size L and desired target SNR to be detected, we can determine the bias value c in our proposed method that is needed in order to get the same asymptotic detection delay performance obtained from the optimal algorithm, assuming that the performance is measured with respect to the slope term in the asymptotic expression for the mean detection delay. So, from (41) and (46), with $\beta = 1$, we can write

$$\frac{L\sigma_s^2}{\sigma_n^2} - c = \frac{1}{2}\frac{L\sigma_s^2}{\sigma_n^2} - \frac{1}{2}\log\left(1 + \frac{L\sigma_s^2}{\sigma_n^2}\right) \qquad (49)$$

from which we get

$$c = \frac{1}{2}\frac{L\sigma_s^2}{\sigma_n^2} + \frac{1}{2}\log\left(1 + \frac{L\sigma_s^2}{\sigma_n^2}\right). \qquad (50)$$

In Figure 5, we plot the values of c as a function of SNR for different numbers of antenna elements L. Now, we can make the following interpretation. Given L and SNR values, we can find the value of the bias term c that is needed to give us mean detection delay performance equivalent to that of the optimal algorithm, and from Figure 4, we can find the value of the exponent s_0 corresponding to that bias value c, which in turn specifies the achievable performance on the false alarm rate given by (38). Or, given L and the requirement on the false alarm rate determined by the value of the exponent s_0, we can obtain from (37) $c = (1/s)\log(1/(1-s)) - 1$, the value of the bias term c that is needed to satisfy this requirement; and corresponding to that bias value, we can retrieve from Figure 5 the SNR value for which we can achieve mean detection delay performance equivalent to that of the optimal algorithm.

Given the number of sensor elements L, we can see that the parallel CUSUM method, with c chosen according to (50), can achieve optimal detection delay performance, while its false alarm performance gets closer to the optimal as the SNR increases. For instance, looking at point B in Figure 5, we see that with $L = 2$ and for values of SNR ≥ 4.81 dB, we can find a bias value $c \geq 4$, for which we get optimal detection delay with false alarm performance measured in terms of the exponent $s_0 \geq 0.993$, where $s_0 = 1$ is the optimal exponent. Also, looking at point A, we see that with a larger array size $L = 8$, we can use the same bias value $c \geq 4$, and achieve the optimal detection delay performance with the same false alarm performance ($s_0 \geq 0.993$) under lower SNR conditions (SNR ≥ -1.21 dB). So by increasing the number of antenna elements L, we can achieve the same performance levels under lower SNR conditions.

Next, we will present some simulation results in order to make the ideas developed in this paper more concrete.

6. SIMULATION RESULTS

For the simulations, we will take both the noise and the incoming signal to be i.i.d. zero-mean complex Gaussian with noise covariance $\sigma_n^2 \mathbf{I}$ and signal covariance $\sigma_s^2 \mathbf{I}$. We choose a value of SNR $= \sigma_s^2/\sigma_n^2 = 5$ dB, and for different array sizes L, the bias term is chosen according to (50). Overall, we perform 10 000 Monte Carlo simulations for each data point in detection delay measurements, and 1000 Monte Carlo simulations for each data point in false alarm measurements. The reason why we used fewer simulations for false alarm measurements is that even for threshold values h that are much smaller than the ones we used for detection delay measurements, the false alarm times were much longer which made each Monte Carlo simulation take that much longer to complete.

In Figures 2a through 3c, some sample beampatterns for a conventional beamformer with two and six sensors are plotted. We see that as the number of sensors increases, the width of the main lobe decreases and the resolution of our detector improves, which is expected to result in a better performance in terms of angle-of-arrival estimation. For simplicity, we divide the interval $[-\pi/2, \pi/2]$ equally into 180 points and employ 180 beamformers such that each beamformer $\mathbf{w}_k$ points in the direction $-\pi/2 + ((k-1)/180)\pi$, $k = 1, \ldots, 180$. We also note that even for 10 sensors ($L = 10$), where the main lobe width is very small, a separation of one degree provides the essential overlap of the main lobes of the collection of beampatterns to cover the whole region. That is to say, according to this configuration with $K = 180$, as L ranges from 2 to 10, we get $\beta = 0.9999$ through $\beta = 0.9969$, where we have defined β as in (42).

Now looking at Figures 2a through 3c, we notice the following. As the look direction moves away from $0°$ in either direction, another sidelobe starts to appear, the beampattern starts to have asymmetry, and the peak of the second lobe increases in the opposite direction. For $90°$, there is perfect symmetry and the beamformer aimed at $-90°$ will have unity response for a signal coming from $90°$. For small L, this phe-
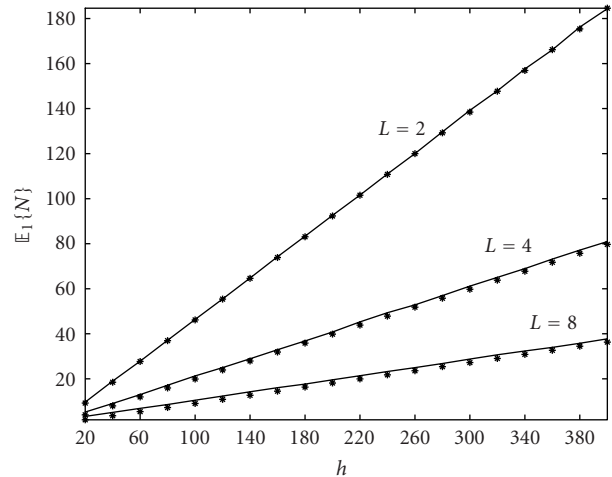

Figure 6: Mean detection delay versus h for $L = 2$, 4, and 8. Solid lines: simulation; asterisks: theoretical.

nomenon is more dramatic; when we compare Figures 2b and 3b, we see that the second lobe for $L = 2$ is much larger than it is for $L = 6$.

We estimate the angle of arrival (30) using the idea that the beamformer whose look direction corresponds to (or is the closest to) the angle of arrival will be the one responsible for the detection, since it will have the largest response among all beamformers. Ideally, we would like to have a collection of beampatterns in the shape of a daisy in which each petal corresponds to a specific beamformer so that we have the large responses from those beamformers that are looking in the direction of arrival, and not from those beamformers looking in the opposite direction. Now, for small L, and for a signal coming from, say $-70°$, the beamformer looking in the $+70°$ direction will have a significant response due to its large second lobe. Because of this, we may get large estimation errors coupled with a sign ambiguity. For this reason, in our simulations, we restrict the target's angle of arrival to the interval $[-50°, 50°]$ so that even for $L = 2$, the response of the second lobe remains at a comfortably low level and we can get reliable estimates.

In the following, we investigate the properties of the algorithm under the case of a minimal configuration with two sensors ($L = 2$), and we also vary the number of sensor elements to observe the performance improvement achieved by increasing the array size.

Figure 6 shows the mean detection delay for $L = 2$, 4, and 8 as threshold h varies from 20 to 400 in steps of 20. Solid line shows the simulation results and the asterisks show the locations obtained from the theoretical asymptotic expression (41). We observe that the mean detection delay follows the asymptotic expression very closely and increases linearly with h, meaning that the asymptotic result is fairly general (it holds for finite values of h), and it can be used as a design guideline. We also see that for fixed h, the mean detection delay decreases as the array size L increases, as expected. In particular, we can observe this in
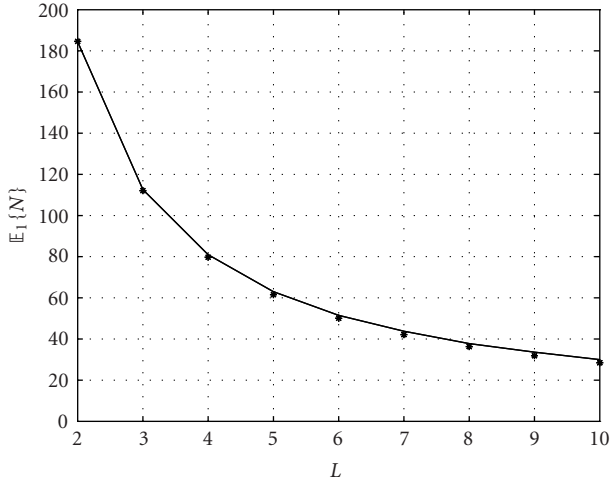
FIGURE 7: Mean detection delay versus L for $h = 400$. Solid line: simulation; asterisks: theoretical.
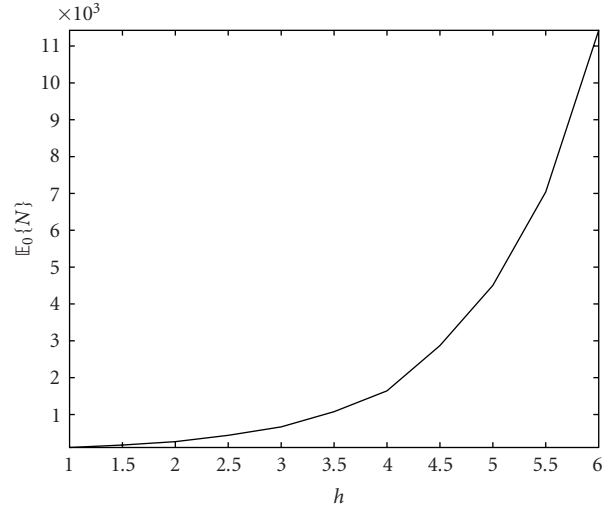

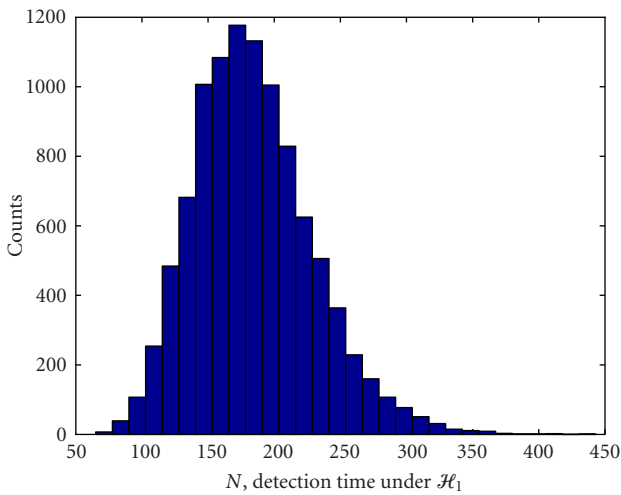FIGURE 9: Simulation results for mean time to a false alarm versus h for $L = 2$.

FIGURE 8: Histogram of detection times for $h = 400$ and $L = 2$.
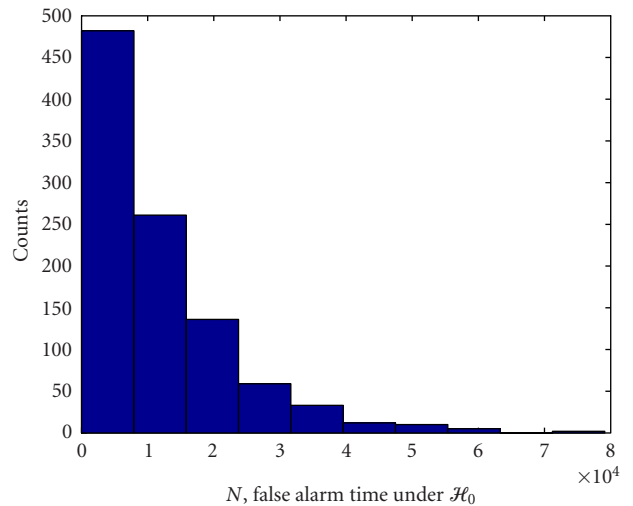

FIGURE 10: Histogram of false alarm times for $h = 6$ and $L = 2$.

Figure 7, where we plot the mean detection delay as a function of L, for threshold $h = 400$. In Figure 8, we show the histogram of detection delays for $h = 400$ and $L = 2$, based on which we can make the observation that it has a gamma-like density.

In Figure 9, we plot the mean time to a false alarm for $L = 2$, as h varies from 1 to 6 in steps of 0.5. We see that the mean time to a false alarm increases exponentially with h. For this value of L, we used $c = 4.16$ obtained from (50), and corresponding to that, we have $s_0(c) = 0.99$. Based on this, we can conclude that the theoretical lower bound (38) is rather loose since we get $e^6 \simeq 403$, which is much smaller than the simulation results. In Figure 10, we show the histogram of false alarm times for $h = 6$ and $L = 2$, based on which we

can make the observation that it has an exponential-like distribution.

In Figures 11 and 12, we look at the mean detection delay and the mean false-alarm time properties of the optimal detector we have defined in Section 5. Comparing Figure 6 with Figure 11, and Figure 9 with Figure 12, we see that the mean detection delays for the parallel CUSUM exactly match those of the optimal detector, as expected, whereas the mean time to a false alarm for the optimal detector is higher than that achieved by the parallel CUSUM method for the same threshold h. We note that the theoretical lower bound (47) is very loose for the optimal algorithm as well. We also compare, in Figure 13, the parallel CUSUM method to the optimal algorithm in terms of the standard deviation of the
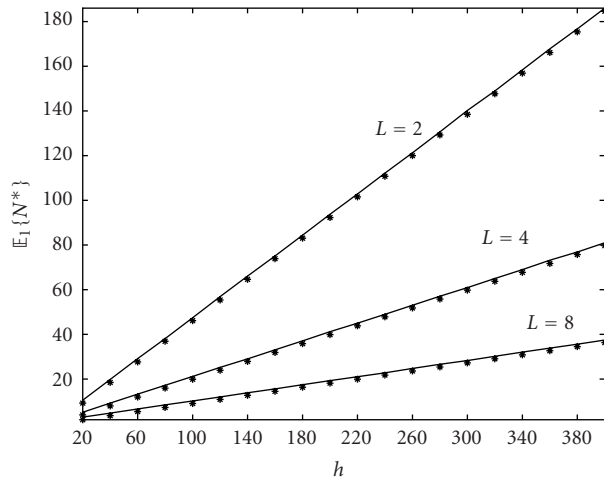
FIGURE 11: Mean detection delay versus h for $L = 2$, 4, and 8 with the optimal algorithm. Solid lines: simulation; asterisks: theoretical.
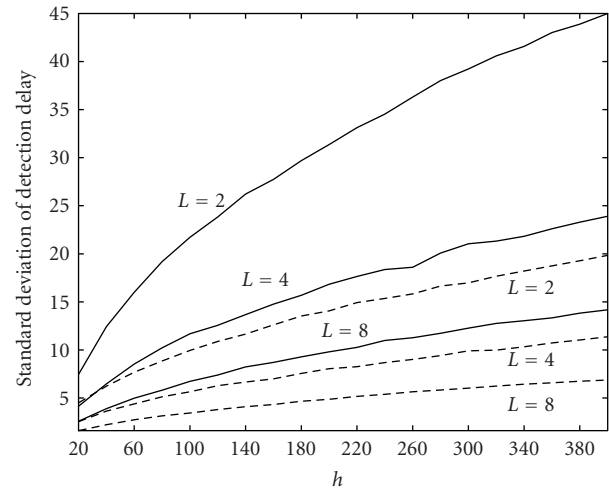

FIGURE 13: Simulation results for the standard deviation of the detection delay times versus h for $L = 2$, 4, and 8. Solid lines: parallel CUSUM method; dashed lines: optimal algorithm.
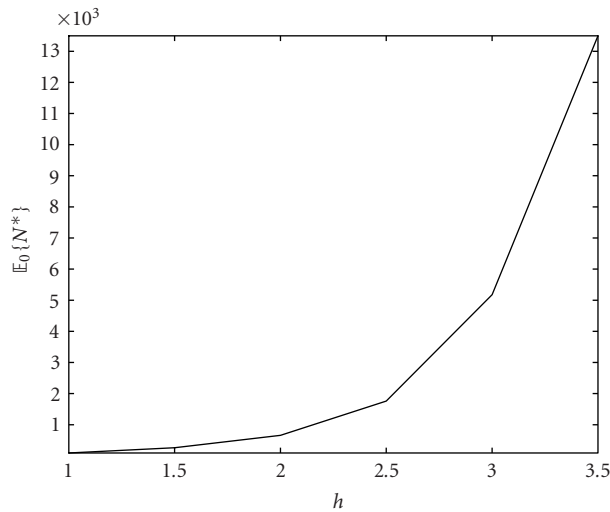

FIGURE 12: Simulation results for mean time to a false alarm versus h for $L = 2$, with the optimal algorithm.
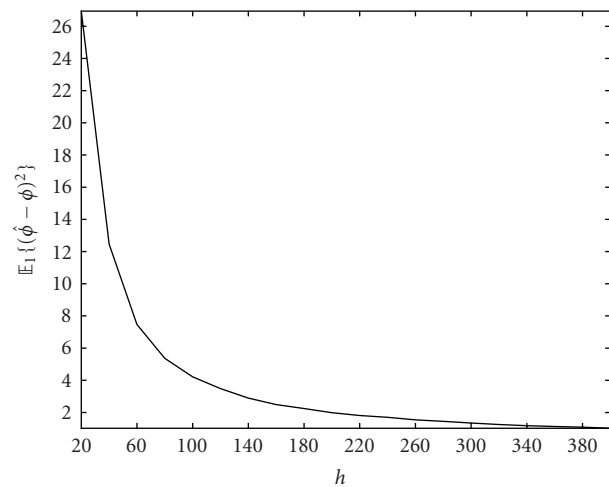

FIGURE 14: Mean-squared angle-of-arrival estimation error versus h for $L = 2$.

detection delay as a function of threshold h for different array sizes $L = 2$, 4, and 8. We see the standard deviation increases with h and the gap between the optimal and the proposed method is smaller for larger L.

In Figure 14, we consider the mean-squared angle-of-arrival estimation error as a function of threshold h, while keeping the number of sensors fixed ($L = 2$). As we increase h (causing more delay in detection), initially we see considerable improvement which diminishes slowly as we further increase the threshold size. In Figure 15, we again look at the mean-squared estimation error; this time with a larger array size ($L = 6$), and see that the estimation errors are significantly lower. In Figure 16, we vary the number of sensors in the array for a fixed detection threshold $h = 20$,

and observe that the mean-squared error exhibits a sharp decrease initially as L increases from 2, and then tapers off. Note that while comparing the estimation errors as a function of L for a fixed threshold h, we should keep in mind that we not only get a reduction in the mean-squared estimation error with a larger array size, but also we detect the target sooner as evident from Figures 6 and 7. Finally, Figure 17 shows the histogram of squared estimation error for $h = 100$ and $L = 6$.

7. CONCLUSION

We have examined the problem of detecting a target that appears abruptly in a noisy environment. For this purpose,
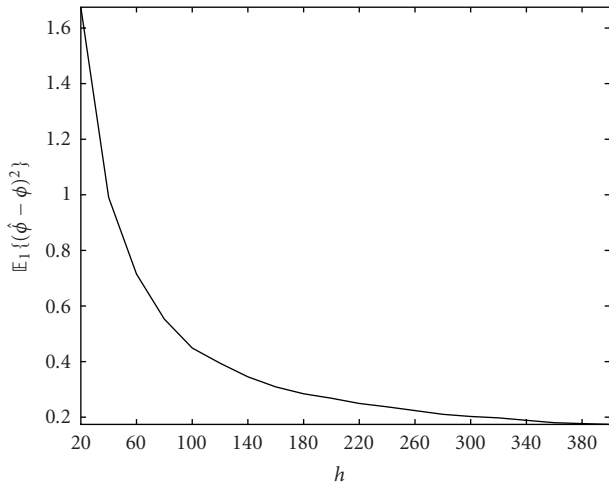
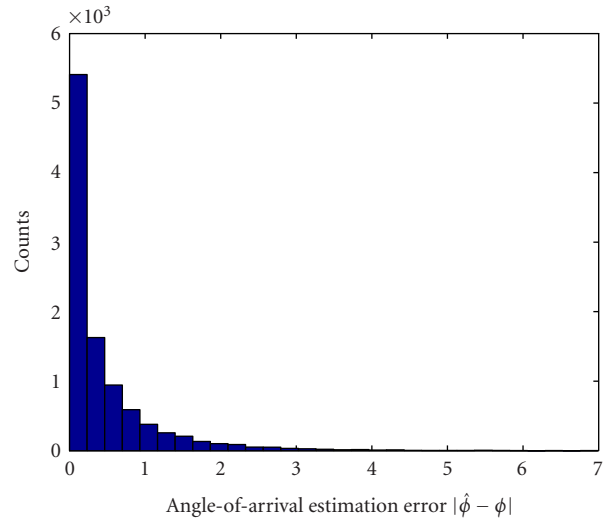FIGURE 15: Mean-squared angle-of-arrival estimation error versus h for $L = 6$.

FIGURE 16: Mean-squared angle-of-arrival estimation error versus L for $h = 20$.

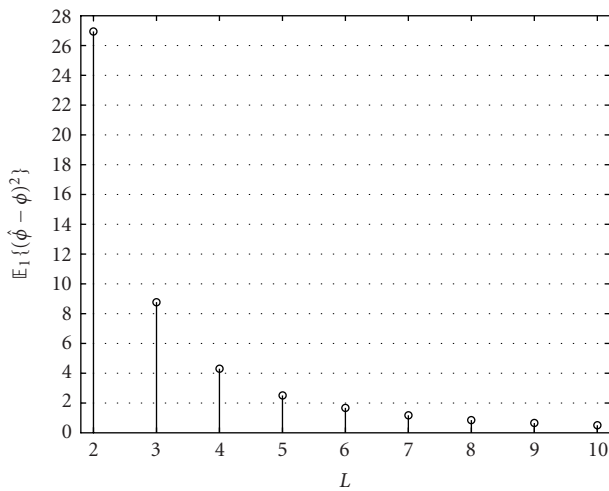FIGURE 17: Histogram of angle-of-arrival estimation error for $L = 6$ and $h = 100$.

we have applied a sensor array and devised a parallel CUSUM algorithm based on beamforming. The algorithm not only detects quickly the existence of the target, but it also provides an estimate on the target's angular direction. We have developed analytical bounds on the algorithm performance and verified these bounds through simulation, and also demonstrated the algorithm's effectiveness by varying different parameters in the system. We have also compared, under the Gaussian signal and noise case, the proposed algorithm's performance against the optimal algorithm that has perfect knowledge of the signal and noise distributions, together with the direction of arrival.

REFERENCES

[1] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*, Prentice Hall, Englewood Cliffs, NJ, USA, 1993.

[2] G. Lorden, "Nearly-optimal sequential tests for finitely many parameter values," *Annals of Statistics*, vol. 5, no. 1, pp. 1–21, 1977.

[3] I. V. Nikiforov, "A generalized change detection problem," *IEEE Transactions on Information Theory*, vol. 41, no. 1, pp. 171–187, 1995.

[4] I. V. Nikiforov, "A suboptimal quadratic change detection scheme," *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 2095–2107, 2000.

[5] T. Oskiper and H. V. Poor, "Online activity detection in a multiuser environment using the matrix CUSUM algorithm," *IEEE Transactions on Information Theory*, vol. 48, no. 2, pp. 477–493, 2002.

[6] G. Lorden, "Procedures for reacting to a change in distribution," *Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1897–1908, 1971.

[7] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 2, pp. 100–115, 1954.

[8] G. Moustakides, "Optimal procedures for detecting changes in distributions," *Annals of Statistics*, vol. 14, pp. 1379–1387, 1986.

[9] B. Broder, *Quickest detection procedures and transient signal detection*, Ph.D. dissertation, Department of Electrical Engineering, Princeton University, Princeton, NJ, USA, 1990.

[10] B. K. Ghosh, *Sequential Tests of Statistical Hypotheses*, Addison-Wesley, Reading, Mass, USA, 1970.

[11] L. C. Godara, "Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.

Taragay Oskiper was born in Ankara, Turkey, in 1974. He received the B.S. degree in electrical engineering from Bilkent University, Ankara, Turkey, in 1996, and the M.S. and Ph.D. degrees in electrical engineering from Princeton University, Princeton, NJ, in 1998 and 2001, respectively. His research interests are in the areas of statistical signal processing, theory of detection and estimation, and wireless communications. Since 2001, Dr. Oskiper's employment has been with Zargis Medical Corp., where he is developing advanced signal processing algorithms for computer-aided analysis of heart sounds.

H. Vincent Poor received the Ph.D. degree in electrical engineering and computer science from Princeton University in 1977. From 1977 till 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton University, where he is the George Van Ness Lothrop Professor in engineering. Dr. Poor's research interests are in the areas of wireless networks, advanced signal processing, and related fields. Among his publications in these areas is the recent book, *Wireless Communication Systems: Advanced Techniques for Signal Reception* (Prentice Hall, 2004). Dr. Poor is a Member of the US National Academy of Engineering, and is a Fellow of the IEEE, the Institute of Mathematical Statistics, and other organizations. In 1990, he served as a President of the IEEE Information Theory Society, and he is currently serving as the Editor-in-Chief of the IEEE Transactions on Information Theory. Among his recent honors are the Joint Paper Award of the IEEE Communications and Information Theory Societies (2001), the NSF Director's Award for Distinguished Teaching Scholars (2002), a Guggenheim Fellowship (2002–2003), and the IEEE EAB Major Educational Innovation Award (2004).