Randomizing a Sliding M-of-N Detector to Control False Alarm Rate

D. A. Abraham

Technical Report
APL-UW TR 2504
August 2025

Applied Physics Laboratory
1013 NE 40th Street

University of Washington
Seattle, Washington 98105-6698

Contract N00024-21-D-6400 under task orders N00024-22-F-8714 and N00024-25-F-8703

Acknowledgements

This report was sponsored by the Office of Naval Research, Code 32, Undersea Signal Processing, through Naval Sea Systems Command contract N00024-21-D-6400 under task orders N00024-22-F-8714 and N00024-25-F-8703.

Abstract

A simple and robust sequential detector commonly used in remote sensing applications declares a signal is present the first time M successes are observed in any N consecutive measurements. In scenarios where N is fixed by stationarity restrictions, the sliding M-of-N detector is designed by varying M. This provides coarse control of the false alarm rate (FAR), which decreases as M is increased from one to N. In this report, the value of M is randomized to allow precise control of the FAR, which can reduce the average delay before detection (i.e., latency) compared to using the smallest fixed value of M that meets or exceeds the FAR specification. The cost of using a randomized sliding M-of-N detector is an increase in the standard deviation of the number of measurements required to make a decision relative to its mean. Approximations to the detection performance measures for standard sliding M-of-N detectors are reviewed and employed to design and analyze the randomized sliding M-of-N detector. Precise control of the false alarm performance is then exploited to compare approaches for controlling FAR in a two-stage detection algorithm when the first-stage background is dominated by false-alarm-inducing clutter and the second stage employs a sliding M-of-N detector. Throttling the first stage to maintain a constant single-measurement probability of false alarm was seen to have a minor advantage in detection latency at very high signal-to-noise power ratio (SNR), compared with passing the clutter-induced false alarms to a randomized sliding M-of-N detector in the second stage. At moderate SNR with heavy clutter or at low SNR, however, throttling reduces the single-measurement probability of detection to the point where there is a significant increase in latency relative to using the randomized sliding M-of-N detector to control FAR. This analysis supports the commonly encountered engineering design approach where the first-stage single-measurement detector is run "hot" and the second-stage multiple-measurement detector cleans up the excessive false alarms, while providing a means for precise control of the FAR and adding the nuance of the high-SNR result.

Contents

1 Introduction

A common technique for detecting a short-duration signal or the onset of a longer duration one within a sequence of binary measurements is to require any M of each N consecutive observations to individually present detections [1, Sect. 10.4.2]. These sliding M-of-N detectors are straightforward to implement and tune. For example, increasing M enforces a stricter requirement for detection that increases the average time between false alarms (F [units: samples]), at the expense of an increase in the average delay to detection (D [units: samples]). Although increasing N elicits a minor relaxation of the detection conditions relative to decreasing M, it is often determined by other considerations, such as stationarity or contiguity of the underlying signal. Given N, guidance for choosing M can be found in [2] (also see [1, Sect. 10.3.1.1]) when the binary data are produced by thresholding measurements made in active sensing for various signal models in Gaussian noise and in [3] for a clutter-dominated background.

A limitation of the sliding M-of-N detector is that only a finite set of operating points (i.e., pairs of (F, D)) can be achieved by varying M when N is fixed. In this report, it is shown that a randomization of M when implementing the test provides full control of the false alarm rate (FAR), which is simply $1/F$. Randomized fixed-sample-size tests, which arise when the decision statistic has discrete components [4, Sect. 3.2], are often avoided owing to the detection decision potentially being relegated to a coin flip. Rather than a randomly missed detection, however, the cost of randomizing a sliding M-of-N detector is a change in latency and an increase in the standard deviation of the number of observations required to make a decision relative to its mean.

This report is structured as follows. Background on sliding M-of-N detectors is presented in Sect. 2, including approximations to the detection performance metrics. The randomized sliding M-of-N detector is described and analyzed in Sect. 3. In Sect. 4, examples are presented for optimizing the design when the Bernoulli probabilities describing the binary data are fixed and for controlling FAR when heavy-tailed noise (i.e., clutter) in the underlying sensing process of a two-stage detector causes an increase in false alarms in the binary data.

2 Background on sliding M-of-N detectors

A sequential detection algorithm declares a signal is present when a sequence of observations satisfies a criteria designed to quickly detect signals while having a large average time between false alarms.[1] The detection performance metrics (F, D) can be obtained from the probability distribution of the stopping time (K [units: samples]), which is defined as the number of observations required by the sequential detector to declare a signal is present. The average delay before detection is $D = E_1[K]$ [units: samples] and the average time between false alarms is $F = E_0[K]$ [units: samples], with the subscript on the expectation operator defining the hypothesis under which it is taken (H_0 for the noise-only hypothesis and H_1 for the signal-present hypothesis). Although the stopping time and the (F, D) performance metrics are all in units of samples, it is straightforward to convert them to temporal units when the time between observations is fixed.

Suppose the data presented to the algorithm $(B_1, B_2, \ldots)$ are modeled as an independent Bernoulli random process (i.e., $B_k = 0$ or 1 and B_k is independent of B_l for all $k \neq l$). Each observation is a Bernoulli random variable that is defined by its success probability: $\Pr_0\{B_k = 1\} = p_0$

[1] Under the assumption that the sequential detector starts anew each time it ends, the average time between false alarms is the average stopping time of the test under the noise-only hypothesis.

under H_0 and $\Pr_1\{B_k = 1\} = p_1$ under H_1. In the following analysis, a Bernoulli success probability $\Pr\{B_k = 1\} = p$ will be used when the results are not specific to one of the detection hypotheses.

To introduce the concept of representing the stopping time as a random variable, consider the example used in [1, Sect. 10.2], which declares a signal present the first time an individual measurement exhibits success (i.e., the first k for which $B_k = 1$). In the context of a sliding M-of-N detector, this is achieved by setting $M = 1$, irrespective of N. Although this is not a very interesting scenario, it is simple and instructive to evaluate. If this sequential detector declares a signal is present on the kth sample, then the first $k-1$ observations must be '0's and the kth a '1'. The probability of this event occurring is

$$f_K[k] = \Pr\{K = k\} = p(1-p)^{k-1} \;\text{ for }\; k \geq 1. \tag{1}$$

This describes the probability mass function[2] (PMF), $f_K[k]$, of the stopping time. A cumulative sum of the PMF yields the cumulative distribution function (CDF) of the stopping time,

$$F_K(k) = \Pr\{K \leq k\} = 1 - (1-p)^k \;\text{ for }\; k \geq 1, \tag{2}$$

which is the probability the signal is detected no later than the kth measurement.

An important performance metric is the average number of samples required to make a decision,

$$\begin{aligned} E[K] &= \sum_{k=1}^{\infty} k f_K[k] && (3) \\ &= \sum_{k=1}^{\infty} [1 - F_K(k-1)]\,. && (4) \end{aligned}$$

In the $M = 1$ example, it is straightforward to show using (1)–(4) that the average stopping time is $E[K] = 1/p$. This could also have been obtained by noting that K for this case is a geometric random variable that starts at one rather than zero (i.e., $K - 1$ follows the geometric distribution described in [5, Ch. 23] with success probability p). For this detector, the average time between false alarms is $F = 1/p_0$ and the average delay to detection is $D = 1/p_1$. As might be expected, good performance requires p_0 to be near zero and p_1 to be near one.

A more realistic sliding M-of-N detector incorporates more than one measurement in the decision. Suppose M measurements in a row are required to declare a signal is present (i.e., a sliding M-of-M detector). The stopping-time CDF can be obtained as a straightforward summation from [6] or [7, pg. 46, eq. 4.7],

$$F_K(k) = \sum_{j=1}^{J} (-1)^{j+1} \left[p + \frac{(k - jM + 1)}{j}(1-p) \right] \binom{k - jM}{j-1} p^{jM} (1-p)^{j-1} \;\text{ for }\; k \geq M, \tag{5}$$

where $\binom{n}{k}$ is the binomial coefficient, $J = \lfloor (k+1)/(M+1) \rfloor$, and $\lfloor \cdot \rfloor$ represents the floor function. Although this approach utilizes more measurements in the detection decision, it is important to note that $F_K(k) = 0$ for $k < M$. That is, a signal cannot be detected until at least M observations have been made. Although increasing M delays detection when a signal is present, it has the advantage of producing an exponential increase in the average time between false alarms.

[2] A probability mass function (PMF) is the equivalent to a probability density function (PDF) for discrete random variables.

Outside of these two simple cases, the general sliding M-of-N detector is more difficult to evaluate. As described in [8] (or [9, Sect. 3.1]), it can be represented as a finite-state Markov chain. This yields an exact analysis of the performance, but is not straightforward and is only feasible (computationally) for somewhat small values of N. An alternative approach that balances both accuracy and ease of implementation can be found in the field of *scan statistics* [7] through use of Naus's approximations for Bernoulli processes (see [10] or [7, Sect. 4.2]). After modifying the approximations to allow the signal-present decision to be made using fewer than N observations, they are applied in Sect. 2.2 to obtain the CDF of the stopping time and in Sect. 2.3 for the average delay before detection. A simpler approach found in [9, Sect. 3.3.1] is presented in Sect. 2.1 to approximate the average time between false alarms.

2.1 Approximating the average time between false alarms

Ideally, the Bernoulli success probability under the noise-only hypothesis (p_0) is small relative to when a signal is present. As described in [9, Sect. 3.3.1]), the noise-only scenario in a sliding M-of-N detector can be treated like a sequence of Bernoulli trials if $Np_0 \ll 1$. The event defining success in these Bernoulli trials is the observation of a sequence of the underlying binary data consisting of a miss, followed by $M-1$ successes in the next $N-1$ trials, followed by a success. The average time between false alarms in the sliding M-of-N detector is then approximated by

$$F = E_0[K] \approx \frac{\Gamma(M)\Gamma(N-M+1)}{\Gamma(N)p_0^M(1-p_0)^{N-M+1}} \tag{6}$$

from [9, eq. 56] (also found in [3, eq. 7]).

The p_0^M term in the denominator of (6) illustrates why M has a larger impact on the average time between false alarms than N. The approximation also suggests that any value of F might be achieved if M were not restricted to be an integer. This is what motivates the randomized sliding M-of-N detector described in Sect. 3.

2.2 Approximating the cumulative distribution function (CDF) of the stopping time

The CDF of the stopping time of the sliding M-of-N detector is useful in evaluating detection performance metrics obtained under the signal-present hypothesis, including the average delay before detection and the probability of detecting a signal within a given number of observations. As previously noted, Naus's approximation [10] (also found in [7, pg. 45, eq. 4.3]) has an appropriate level of accuracy and a straightforward evaluation. However, it was developed under the assumption that the test begins with a set of N observations and then observes new samples until a detection decision is made. In applications where rapid detection is imperative, the sliding M-of-N detector is typically allowed to make a detection decision starting with M observations. The straightforward modification of Naus's approximation to represent this implementation of the sliding M-of-N detector is described in this section.

Naus's approximation to the distribution of the stopping time of the sliding M-of-N detector essentially describes it as having a geometric tail. That is, the PMF is proportional to q_{g}^k as the stopping time k increases, where $p_{\mathrm{g}} = 1 - q_{\mathrm{g}}$ is the success probability of the geometric distribution. Although Naus sets p_{g} using an asymptotic (large k) approximation, it can be quite accurate for moderate values of k that do not significantly exceed N. For $k \leq N$, the exact probabilities can be obtained using the binomial distribution, where stopping on the kth sample or earlier requires observing at least M successes in the first k samples. Combining these two parts of the distribution produces

$$F_K(k) \approx \begin{cases} 0 & \text{for } k < M \\ 1 - F_\mathrm{b}(M-1; k, p) & \text{for } M \leq k \leq N \\ 1 - c_\mathrm{g}\, q_\mathrm{g}^{k-N} & \text{for } k > N \end{cases} \tag{7}$$

for the probability the sliding M-of-N detector stops on or before the kth measurement, where $F_\mathrm{b}(x; N, p)$ and $f_\mathrm{b}(x; N, p)$ (which is used below) are, respectively, the CDF and PMF of a binomial distribution representing N trials with success probability p and argument $x \in \{0, 1, \ldots, N\}$. The probability and scale parameters of Naus's approximation to the geometric tail of the distribution are

$$q_\mathrm{g} = \left(\frac{Q_3'}{Q_2'}\right)^{1/N} \quad \text{and} \tag{8}$$

$$c_\mathrm{g} = \frac{(Q_2')^2}{Q_3'} = F_\mathrm{b}(M-1; N, p) = \Pr\{K > N\}\,, \tag{9}$$

where Q_2' and Q_3' are defined in [10], [7, Sect. 4.2] and also here in App. A.

An approximation to the PMF can be obtained by the difference $f_K[k] = F_K(k) - F_K(k-1)$, which results in

$$f_K[k] \approx \begin{cases} 0 & \text{for } k < M \\ p f_\mathrm{b}(M-1; k-1, p) & \text{for } M \leq k \leq N \\ c_\mathrm{g}(1-q_\mathrm{g}) q_\mathrm{g}^{k-N-1} & \text{for } k > N. \end{cases} \tag{10}$$

For $k \in [M, N]$, this represents observing $M-1$ successes in the first $k-1$ samples followed by a success on the kth measurement. The PMF for $k > N$ can be seen to be that for a geometric random variable, given it exceeds N, multiplied by c_g. As seen in (9), c_g is the probability the stopping time of the sliding M-of-N detector exceeds N, which occurs when there are fewer than M successes in the first N observations.

2.3 Approximating the average delay before detection

Using the PMF from (10) in (3), the average stopping time of a sliding M-of-N detector that is allowed to declare a detection after M observations can be approximated by

$$E[K] \approx c_\mathrm{g}\left(N + \frac{1}{1-q_\mathrm{g}}\right) + p \sum_{k=1}^{N} k f_\mathrm{b}(M-1; k-1, p). \tag{11}$$

The summation term in (11) represents the contribution to the mean for stopping times $k \leq N$. The first term is seen to be N plus the mean of a geometric distribution (with success probability $p_\mathrm{g} = 1 - q_\mathrm{g}$), multiplied by the probability (c_g) that a sliding M-of-N detector requires more than N samples.

This approximation to the average stopping time is useful when evaluating the average delay before detection, which is obtained by setting p to p_1. Although it can also represent the average time between false alarms, the approximation shown in Sect. 2.1 is simpler to evaluate and less prone to computational errors when p_0 is small (e.g., $p_0 \ll 10^{-2}$).

2.4 Approximating the standard deviation of the stopping time

Although the average stopping time is the primary measure of performance in sequential detectors, the variance or standard deviation is useful in understanding how much the stopping time can vary from one instantiation of the test to the next. The standard deviation is most easily obtained here by its definition that combines the first two moments of the stopping time,

$$\mathrm{Std}\{K\} = \sqrt{E[K^2] - (E[K])^2} \quad \text{[units: samples]}. \tag{12}$$

Similar to how the first moment in (11) was obtained from the PMF of K in (10), the second moment can be shown to be approximated by

$$E[K^2] \approx c_{\mathrm{g}}\left(N^2 + \frac{2N}{1-q_{\mathrm{g}}} + \frac{1+q_{\mathrm{g}}}{(1-q_{\mathrm{g}})^2}\right) + p\sum_{k=1}^{N} k^2 f_{\mathrm{b}}(M-1; k-1, p) \quad \text{[units: samples}^2\text{]}. \tag{13}$$

Combining (13) with (11) in (12) then produces the standard deviation.

An approximation to the second moment based on the small-p_0 approximation used to obtain (6) leads to

$$E_0[K^2] \approx F^2 + (F-1)^2 \approx 2F^2, \tag{14}$$

which implies the standard deviation of the stopping time under the noise-only hypothesis (H_0) is $\mathrm{Std}\{K\} \approx F$. For small success probabilities, the standard deviation of a geometric random variable is approximately its mean, which is analogous to the exponential distribution for continuous random variables.

3 Randomizing M in a sliding M-of-N detector

Consider an example where stationarity conditions require $N = 5$, where the Bernoulli probability under the noise-only hypothesis is $p_0 = 10^{-3}$, and when an average time between false alarms of $F = 10^{10}$ samples is desired. Solving (6) for M given these values of N and F results in $M \approx 3.6$. However, the five integer-valued choices available for M lead to average times between false alarms of:

M	F
1	1.0×10^3
2	2.5×10^5
3	1.7×10^8
4	2.5×10^{11}
5	1.0×10^{15}

As might be expected, setting $M = 3$ does not achieve the desired false alarm performance. Although using $M = 4$ surpasses the false alarm specification, it will have a larger average delay before detecting the signal compared to a similarly efficacious test that precisely meets F.

In fixed-sample-size tests, the Neyman–Pearson lemma (e.g., see [4, Sect. 3.2]) dictates employing randomized tests to achieve a specific probability of false alarm (P_f) under certain circumstances. This requires that the decision statistic has at least one discrete value, which causes a discontinuity in its exceedance distribution function (EDF), which is simply one minus the CDF. If the desired value of P_f occurs at a discontinuity in the EDF, then a simple threshold test cannot meet the false alarm specification. When the discrete value is observed, a random detection decision is made with a probability that achieves the desired P_f. Another way of viewing the randomized test is to assume the threshold is adjusted (randomly) when the decision statistic equals the discrete value. Extending this interpretation to a fixed M-of-N test equates to a randomization of M.

Randomization is rare in practice, either because adjusting the threshold (and therefore P_f) is often acceptable or owing to an aversion to injecting randomness into the final detection decision. In the sliding M-of-N detector, the first option yields the coarse operating points seen in the above example when N is fixed. However, the detriment of a randomized final detection decision is avoided by the sequential nature of the test. Randomizing M in a sliding M-of-N detector can be viewed as a soft decision that incurs a delay in detection rather than a hard decision that the signal is present or not.

In the above example, suppose the test is initially run with $M = M_0 = 3$, which does not achieve the desired average time between false alarms. When 3 successes are first observed over 5 consecutive samples, suppose the detection is accepted 96.1% of the time. In the remaining 3.9% of the cases, the test is continued with a requirement for $M = M_1 = 4$ successes in any 5 consecutive samples. The average time between false alarms of this randomized test is a similar mixture of the values of F seen above,

$$F = 96.1\% \cdot \left(1.7 \times 10^8\right) + 3.9\% \cdot \left(2.5 \times 10^{11}\right) = 10^{10}, \tag{15}$$

which meets the specification (within rounding). The average delay to detection is also a mixture, so the randomized test will perform better than the sliding 4-of-5 detector, which is the best sliding M-of-5 detector that meets or exceeds the false alarm specification.

The parameters defining the randomized sliding M-of-N detector are M_0, ϵ, and N, where ϵ is the probability of continuing the test after first observing M_0-of-N successes and assuming that the continued test requires $M_1 = M_0 + 1$ successes. The notation "$M_0[+\epsilon]$-of-N" is adopted to define the test. For example, the above test would be denoted a 3[+0.039]-of-5 randomized sliding M-of-N detector.

3.1 Implementation

The steps required to implement the randomized sliding M-of-N detector described above are

1. Apply a sliding M_0-of-N detector to the binary data sequence
2. When M_0 successes are observed in N consecutive samples, pause testing and draw a uniformly distributed random variable U on $[0, 1]$

3. If $U > \epsilon$ stop the test and declare a detection
4. If $U \leq \epsilon$, continue testing until $M_1 = M_0 + 1$ successes are observed in N consecutive samples, at which point the test is stopped and a detection is declared

Alternatively, the detector can be implemented by randomizing M at the beginning of the test. This requires drawing a uniformly distributed random variable U on $[0, 1]$ and setting

$$M = M_0 + \mathcal{I}\{U \leq \epsilon\}, \tag{16}$$

where $\mathcal{I}\{\cdot\}$ is the indicator function returning one when the argument is true and is otherwise zero. A standard sliding M-of-N detector is then implemented with the value of M from (16). In both implementations, an independent draw of U must be performed for each instantiation of the detector.

From these descriptions, it is clear that the randomized sliding M-of-N detector is a standard sliding M_0-of-N detector if $\epsilon = 0$ and a standard sliding $M_0 + 1$-of-N detector if $\epsilon = 1$. It is also evident that if an application only requires a single instantiation of the randomized sliding M-of-N detector, the performance will be that for a standard sliding M-of-N detector with M in (16) after U is drawn. This likely limits the practical utility of the detector to scenarios where many instantiations of the test are implemented, with the desired average time between false alarms achieved over the ensemble.

3.2 Design and analysis

The CDF of the stopping time of a randomized sliding M-of-N detector can be obtained by conditioning on the observed value of M, using the standard CDF approximation from Sect. 2.2, and then removing the conditioning. When $M = M_0$ with probability $1 - \epsilon$ and $M_0 + 1$ with probability ϵ, the CDF is the mixture,

$$F_K(k) = (1 - \epsilon)F_K(k; M_0, N, p) + \epsilon F_K(k; M_0 + 1, N, p), \tag{17}$$

where $F_K(k; M, N, p)$ is the stopping-time CDF for the standard sliding M-of-N detector from (7), with explicit representation of the parameters M, N, and p. This implies that the PMF and any metric formed through an expectation are similarly weighted mixtures. For example, the average time between false alarms of this randomized sliding M-of-N detector is

$$F = (1 - \epsilon)F(p_0, M_0, N) + \epsilon F(p_0, M_0 + 1, N), \tag{18}$$

where $F(p, M, N)$ can be approximated by (6), with explicit arguments. The average delay to detection is approximated by a similar mixture using (11) evaluated with $p = p_1$.

The probability of extending the test in the randomized sliding M-of-N detector is obtained by solving (18) for ϵ, which produces

$$\epsilon = \frac{\bar{F} - F(p_0, M_0, N)}{F(p_0, M_0 + 1, N) - F(p_0, M_0, N)}, \tag{19}$$

where $\bar{F}$ is the desired average time between false alarms. This requires that M_0 be chosen so that

$$F(p_0, M_0, N) \leq \bar{F} \leq F(p_0, M_0 + 1, N). \tag{20}$$

As noted in [9, Sect. 3.4.1], a lower bound on M in a sliding M-of-N detector can be obtained from (6) as

$$M \geq \frac{\log_{10} \bar{F}}{-\log_{10} p_0}. \tag{21}$$

In the introductory example, this yields $M \geq 10/3$, which illustrates that a search for M_0 satisfying (20) should be initialized at the largest integer below the lower bound in (21).

Compared to using a standard sliding M-of-N detector with $M = M_0 + 1$, which surpasses the false alarm specification, the randomized version achieves it precisely. As previously noted, using $M = M_0$ with probability $1-\epsilon$ and $M = M_0+1$ with probability ϵ reduces the average delay before detection relative to the standard sliding M-of-N detector using $M = M_0 + 1$. This is illustrated in the receiver operating characteristic (ROC) curve shown in Fig. 1, which displays $\log_{10} F$ as a function of D. Ideal performance is in the upper left corner where D is minimized and F is maximized. The logarithmic scaling of the figure axes, which is required to display the large range of values, skews the depiction of the ROC curve connecting each standard sliding M-of-N detector (the black dots). If (F_0, D_0) and (F_1, D_1) are the operating points for the standard sliding M-of-N detectors using $M = M_0$ and $M = M_0 + 1$, respectively, then the ROC curve between these two points are linear functions between D and F,

$$F = F_0 + \epsilon(F_1 - F_0) = F_0 + \frac{(D - D_0)(F_1 - F_0)}{D_1 - D_0} \tag{22}$$

and

$$D = D_0 + \epsilon(D_1 - D_0) = D_0 + \frac{(F - F_0)(D_1 - D_0)}{F_1 - F_0}. \tag{23}$$

Thus, the randomized sliding M-of-N detector provides control of the FAR as a linear interpolation between the average stopping times under the noise-only hypothesis of the adjacent standard sliding M-of-N detectors. The average delay before detection is subject to a similarly weighted mixture, which represents a degradation compared to the standard sliding M_0-of-N detector (that does not meet the FAR specification) and an improvement over the standard sliding $M_0 + 1$-of-N detector (that surpasses the FAR specification).

A consequence of implementing the sliding M-of-N detector with different values of M is an increase in the standard deviation of the stopping time when taken relative to its mean. This is illustrated in Fig. 2 for the $3[+\epsilon]$-of-5 detector with $p_0 = 10^{-3}$ and various values of p_1. The effect can be significant under the noise-only hypothesis, owing to the large difference between F for $M_0 = 3$ and 4 (from the table on pg. 5, $F = 1.7 \times 10^8$ for $M = 3$ and 2.5×10^{11} for $M = 4$). However, the largest increase in relative spreading occurs for a very small value of ϵ. The value of $\epsilon = 0.039$ used in the introductory example, which leads to an average time between false alarms of 10^{10} samples, exhibits a seven-fold increase, which is significantly below the peak factor, which is over 27. The increase in the relative spread of detection latency in the example illustrates peaks on the order of a 50% increase that occur at larger values of ϵ.

The standard deviation of the stopping time of the randomized sliding M-of-N detector was obtained by evaluating the first two moments of the stopping time as mixtures (e.g., as done in (18) for F) and using them in (12). The increase in relative spreading is a direct result of using different values of M, which implies that the randomization should be limited to two adjacent values of M.

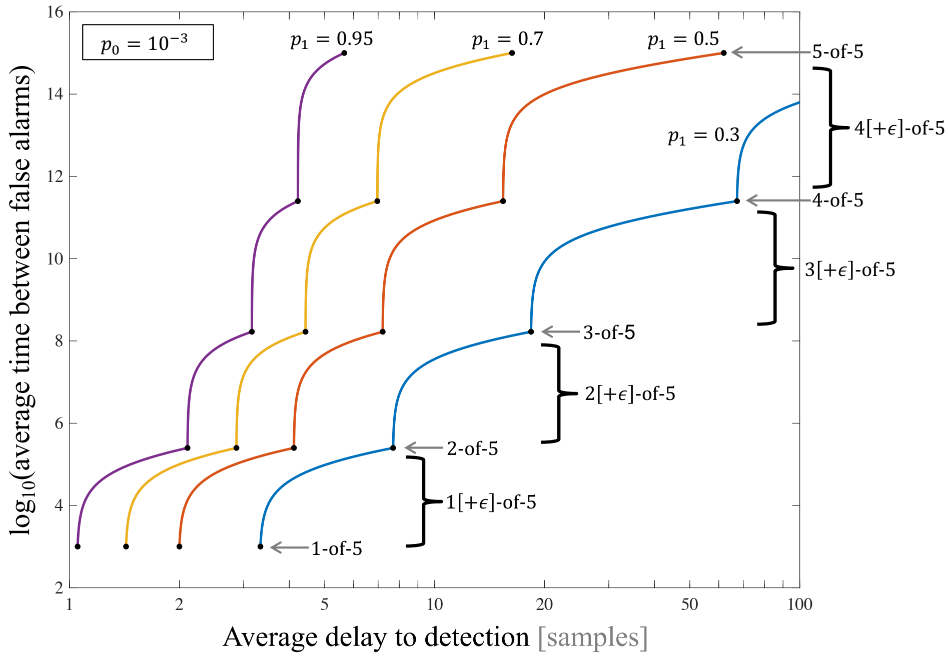

Figure 1: Example receiver operating characteristic curve of a randomized sliding M-of-N detector. Ideal performance is in the upper left corner.

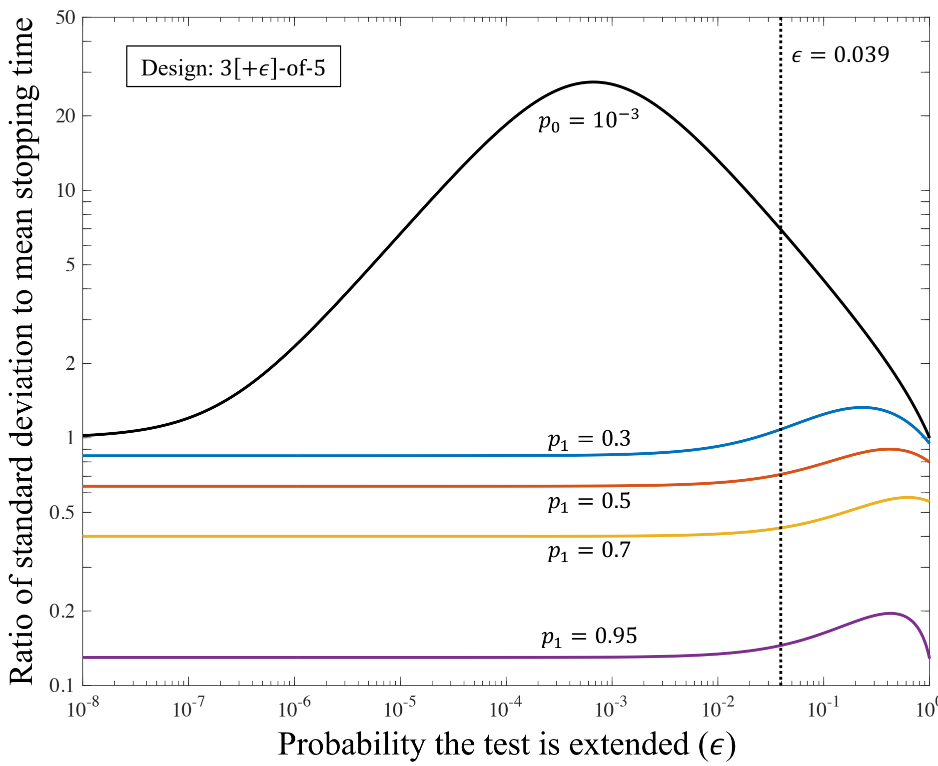

Figure 2: Ratio of the standard deviation of the stopping time to its mean for an example randomized sliding M-of-N detector (3[+ϵ]-of-5) as a function of the probability the test is extended (ϵ).

For example, one might meet the false alarm specification with a shorter average delay to detection than that shown in Fig. 1 by mixing M_0 with M_0+2 rather than M_0+1. However, this would lead to heavier tails in the stopping-time distribution and exacerbate the relative spreading. Although the increase in spread relative to the mean is undesirable, it is a reasonable cost for being able to precisely control the average time between false alarms.

4 Design examples

The basic design of a randomized sliding M-of-N detector entails finding M_0 and ϵ when given N and p_0, as was described in Sect. 3. MATLAB® code implementing the basic design process can be found in App. B.2. The examples considered in this section relax the design constraints on N and p_0. As shown in Sect. 4.1, increasing N can improve performance for weak signals. The example considered in Sect. 4.2 examines different approaches for controlling the false alarm rate when the severity of the background clutter worsens in the single-measurement detection process that forms the binary data in a two-stage detector.

4.1 Optimizing the design for fixed Bernoulli success probabilities

When N and p_0 are fixed, a randomized sliding M-of-N detector is designed by choosing M_0 and ϵ to meet a FAR specification. By allowing N to vary, the detector can also be optimized for a given Bernoulli success probability when a signal is present (p_1).

With other parameters fixed, increasing N reduces the average time between false alarms. When M_0 is fixed, this can be countered by increasing ϵ. At some point, however, increasing N requires an increase in M_0 to meet the FAR specification. Consider the introductory example where $F = 10^{10}$ and $p_0 = 10^{-3}$, which leads to a minimum value of $M_0 = 3$ from (21). As shown in Fig. 3, the probability the test continues (ϵ) increases with N for a fixed value of M_0 in order to meet the FAR specification. In this example, using $M_0 = 3$ sufficed while $N \leq 10$. However, M_0 needed to be increased to 4 for $N \in [11, 42]$ and to 5 for $N \in [43, 108]$.

Although any of these designs (or appropriate combinations for larger values of N) can be used to achieve the FAR specification, they do not all perform the same when a signal is present. The average delay before detection is presented in Fig. 4 for values of p_1 ranging from 0.1 to 0.95. Here it can be seen that the optimal value of N is inversely related to the strength of the signal. In particular, better detection performance can be achieved for weak signals by increasing M_0 above the minimum value required to meet the specification and using an appropriately larger value of N. A MATLAB® subroutine optimizing a randomized sliding M-of-N detector for a given pair of Bernoulli success probabilities can be found in App. B.2.

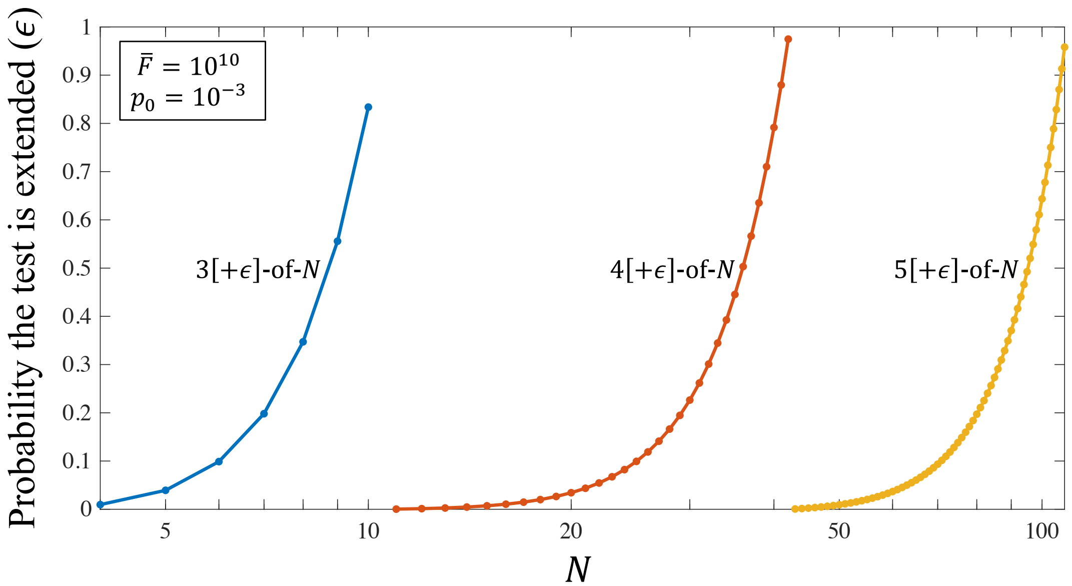

Figure 3: The probability (ϵ) that a randomized sliding M-of-N detector is extended as a function of N when designed to achieve an average time between false alarms of $\bar{F} = 10^{10}$ samples with $p_0 = 10^{-3}$.

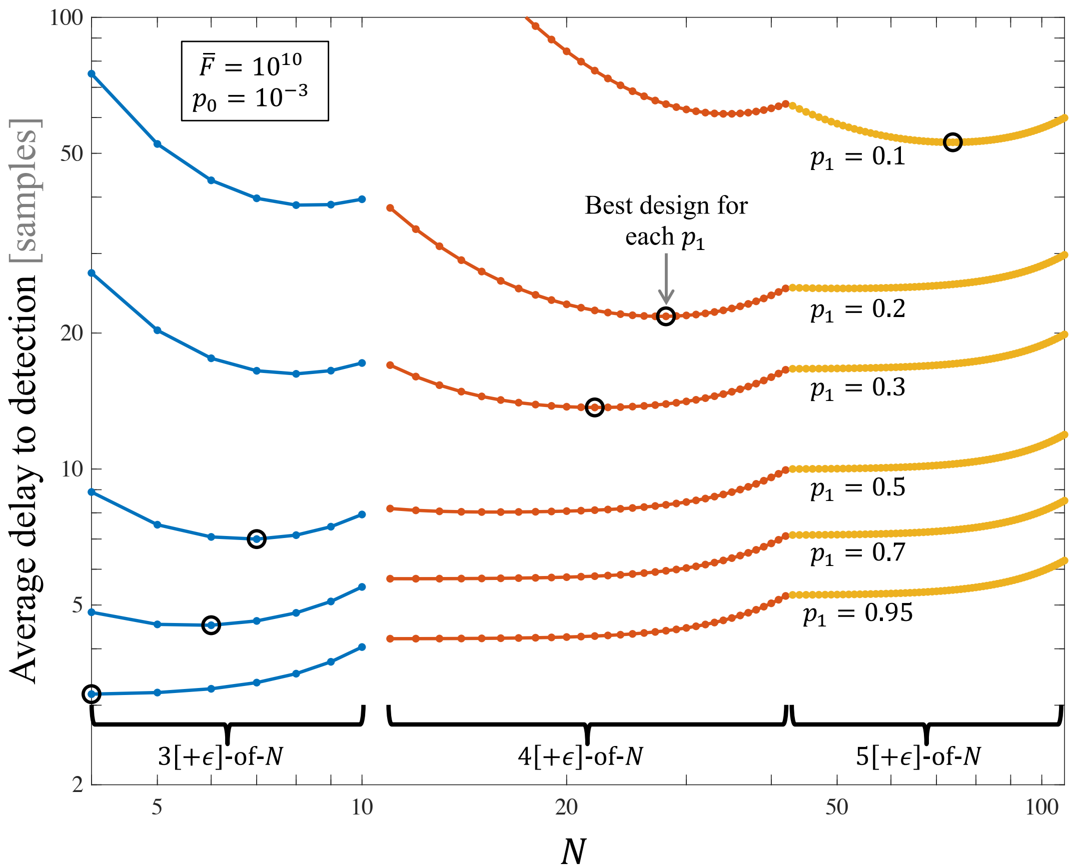

Figure 4: Average delay to detection as a function of N for different values of p_1 in the randomized sliding M-of-N detector designed for $\bar{F} = 10^{10}$ samples with $p_0 = 10^{-3}$.

4.2 Controlling false alarm rate in background clutter

It is common to apply a sliding M-of-N detector to the binary sequence created by threshold testing a consecutive set of measurements. For example, this occurs in radar and active sonar applications where target tracks are initialized with a sliding M-of-N detector [11, Sect. 3.6.2] applied to the binary detection decisions obtained over a sequence of pings. In applications where many such streams are used to initialize tracks (e.g., when searching a broad scene for targets), the FAR after tracking is inversely proportional to the average time between false alarms in a single sliding M-of-N track detector. Using the approximation in (6), this implies FAR is proportional to and dominated by p_0^M, where p_0 is the probability of false alarm when threshold testing a single measurement. When the sliding M-of-N detector is operated in different environments, the statistics of the background clutter can change. An increase in clutter severity typically causes an increase in the single-measurement probability of false alarm (i.e., p_0) and therefore the FAR after tracking. The FAR can be controlled by adjusting the decision threshold in the single-measurement detector to maintain a constant probability of false alarm or by altering the track-level detection algorithm. The randomized sliding M-of-N detector described in Sect. 3 provides the flexibility required to precisely meet a FAR specification in the latter approach, which permits a fair comparison of the two approaches to controlling FAR.

In the example considered here, the single-measurement decision statistics under the noise-only hypothesis are represented by an instantaneous intensity following a generalized Pareto distribution (GPD) [12]. Although this is representative of thresholding a normalized matched-filter intensity in an active sensing system, it omits many other components of the target-tracking application. Using [12, eq. 3], the single-measurement probability of false alarm is

$$p_0 = \left(1 + \frac{\gamma}{\lambda} h\right)^{-\gamma^{-1}}, \tag{24}$$

where $\gamma \in [0, 1)$ and $\lambda > 0$ are, respectively, the GPD shape and scale parameters and h [unitless] is the intensity decision threshold after perfect normalization. Under perfect normalization, which produces a unit-valued average intensity, the GPD scale parameter is $\lambda = 1 - \gamma$. The scintillation index (SI) of the GPD intensity is

$$\mathrm{SI} = \frac{1}{1 - 2\gamma}, \tag{25}$$

from [12, eq. 13], which requires $\gamma \in [0, 0.5)$ in order for the intensity variance (and therefore SI) to be finite. A GPD shape parameter of $\gamma = 0$ represents the benign background of (bandpass) Gaussian-distributed noise, which produces an exponentially distributed intensity with SI $= 1$. Increasing γ above zero exacerbates the severity of the background clutter, with the probability of false alarm in (24) typically increasing. This effect is clearly seen in the scintillation index: $\gamma = 0.25$ yields SI $= 2$ and $\gamma = 0.45$ produces an extremely large value of SI $= 10$.

In the first approach to controlling FAR, the single-measurement probability of false alarm (p_0) is held constant as γ increases by appropriately increasing the detector decision threshold (h). The constant p_0 implies the FAR after tracking does not change as the background clutter becomes more severe. However, the increase in h causes a decrease in the single-measurement probability of detection (p_1). The signal is modeled as having (bandpass) Gaussian fluctuations with the single-measurement probability of detection (p_1) approximated by that for a benign noise background, as described in [12, Sect. 4.1.3]. If s [unitless] is the linear-quantity SNR after coherent detection processing, then the single-measurement probability of detection is

$$p_1 \approx e^{-h/(1+s)}. \tag{26}$$

A reduction in p_1 leads to an increase in the average delay to detection (i.e., latency) in the sliding M-of-N algorithm.

In the second approach to controlling FAR, the single-measurement detector decision threshold (h) is held constant and a randomized sliding M-of-N detector is employed to meet the false alarm specification as the clutter severity worsens. This also increases detection latency relative to that observed in a benign background. Comparing these two approaches is straightforward here because they can be designed to have the same FAR. When the second approach employs a standard sliding M-of-N detector, the comparison can only be done at the FAR achieved by that detector as M and N are varied.

An example analysis of the average delay to detection as a function of the scintillation index of the GPD background clutter is presented in Fig. 5. The single-measurement probability of false alarm was assumed to be $p_0 = 10^{-3}$ in the benign background (i.e., when $\gamma = 0$). An average time between false alarms of $\bar{F} \approx 1.7 \times 10^8$ samples was set based on that achieved by a sliding 3-of-5 detector in the benign background. The solid lines in Fig. 5 represent the latency at three different SNRs for the first approach, where the increase in FAR is handled by increasing the single-measurement detector decision threshold, so the sliding 3-of-5 track-level detector can still be employed. The dots and circles represent the performance of a randomized sliding M-of-N detector with, respectively, N fixed at 5 (i.e., $M_0[+\epsilon]$-of-5) and when N is allowed to vary (i.e., $M_0[+\epsilon]$-of-N). The optimized values of M_0 and N for the latter design are shown on the figure (gray text presented as M_0/N). The values of M_0 for the former design (where $N = 5$) are equal to 3 for SI < 1.5 and otherwise equal to 4. Each detector represented on the figure achieves the desired average time between false alarms.

The 30 dB SNR case (gold markings) illustrates that controlling FAR in the single-measurement detector is generally the best approach at very high SNR. For example, the sliding 3-of-5 detector under the first approach has a maximum latency < 3.1 when the SNR is 30 dB (the minimum latency as SNR grows to infinity is 3 measurements). Each randomized test has a latency similarly close to its minimum, which is simply the average value of M in the test,

$$\lim_{\mathrm{SNR}\to\infty} D = (1-\epsilon)M_0 + \epsilon(M_0+1). \tag{27}$$

In the benign background ($\gamma = 0$), the 3-of-5 sliding M-of-N detector can be improved by reducing M and N to a 2-of-3 randomized sliding M-of-N detector. The minimum latency for this detector is between 2 and 3, so it will always be better than the sliding 3-of-5 detector as SNR tends to infinity. In most of the high-SNR scenarios, however, the cost of controlling FAR in the sliding M-of-N algorithm is greater than the minor reduction in p_1 incurred by increasing the decision threshold in the detector (e.g., $p_1 > 0.97$ for all of the 30 dB SNR cases, even after raising h to control p_0).

At low SNR or at moderate SNR when the scintillation index is large, controlling FAR by modifying a randomized sliding M-of-N detector can have significantly lower latency than adapting the single-measurement detector decision threshold. The 15 dB SNR case (red markings) exhibits improvement except for SI between 1.4 and slightly over 2, which is where ϵ increases toward one and then M_0 increases from 3 to 4. As the clutter severity worsens above this point, however, the randomized sliding M-of-N detector with $M_0 = 4$ is able to maintain the FAR specification

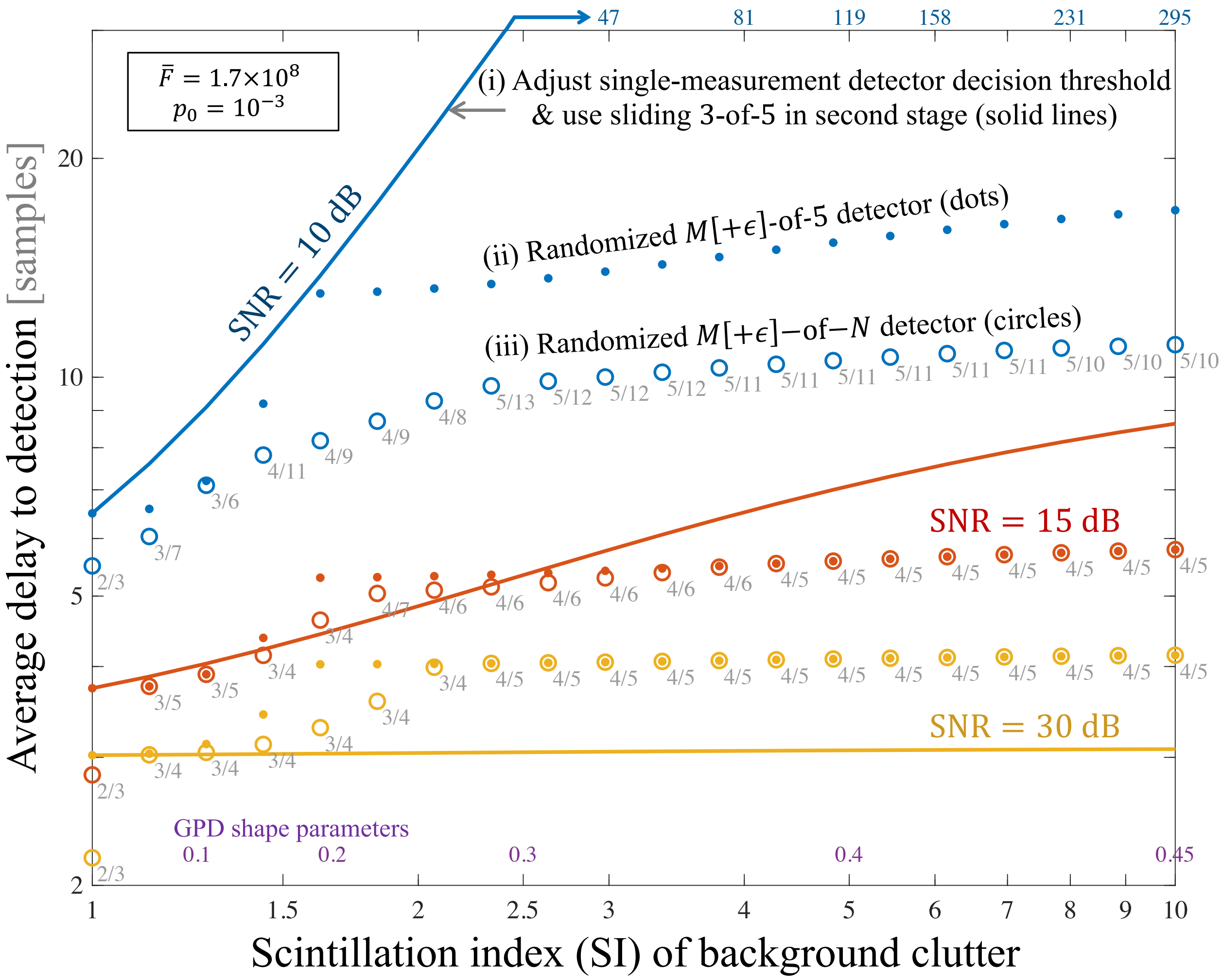

Figure 5: Average delay to detection as a function of the scintillation index of background clutter for three different approaches to controlling false alarm rate: (i) increasing the underlying single-measurement detector decision threshold to maintain its probability of false alarm (p_0) and using a sliding 3-of-5 detector (solid lines); (ii) using a randomized sliding M-of-N detector with $N = 5$ (dots); and (iii) using a randomized sliding M-of-N detector optimized over M and N (circles).

with only a small increase in latency. A similar improvement in robustness to increasing clutter severity can be seen in the 10 dB SNR case (blue markings) when M_0 is increased. In this low-SNR example, the detection performance when controlling FAR in the single-measurement detector (solid blue line) degrades rapidly, with the latency increasing to more than 17 and 26 times greater than that for the two randomized sliding M-of-N detectors when SI = 10. Similar to the results seen in Fig. 4, allowing N to vary can significantly reduce latency at low SNR (e.g., the blue circles in Fig. 5 are more than 30% lower than the blue dots at the highest SIs).

A common engineering design approach in multiple-stage detection algorithms states that it is better to operate the first-stage detector with a high probability of false alarm (i.e., let it run "hot") and then clean up excessive false alarms in a second-stage, multiple-measurement detector (here the sliding M-of-N detector) than to throttle the false alarms in the first stage by increasing the decision threshold. The heuristic argument for the approach is that multiple-measurement detection is difficult if the target is not detected often enough at the single-measurement stage. The low-SNR results seen in Fig. 5 directly support this design paradigm. The high-SNR result suggesting an increase in the first-stage decision threshold provides indirect support in that it only applies while the single-measurement probability of detection remains high enough to have only a minor impact on latency.

5 Conclusions

A technique for controlling the FAR of a sliding M-of-N detector was developed and analyzed in this report. An appropriate randomization of M allows tuning the detector to achieve operating points that are linear combinations of the standard sliding M-of-N detectors. The tools required to design and evaluate the detector were presented, along with techniques for their optimization. Although the randomized sliding M-of-N detectors may only have practical utility in applications where many instantiations are employed, they are useful in comparing different techniques for controlling FAR in multistage detection algorithms where the sliding M-of-N detector operates on binary data produced by a single-measurement-thresholding process. In low-SNR cases, it was observed that controlling FAR (as background clutter becomes more severe) with a randomized sliding M-of-N detector could significantly reduce detection latency compared to increasing the single-measurement decision threshold and retaining the standard sliding M-of-N detector.

References

[1] D. A. Abraham, *Underwater Acoustic Signal Processing: Modeling, Detection, and Estimation.* Springer, 2019.

[2] D. A. Shnidman, "Binary integration for Swerling target fluctuations," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 3, pp. 1043–1053, 1998.

[3] D. A. Abraham, "Optimization of M-of-N detectors in heavy-tailed noise," in *Proceedings of 2018 IEEE/MTS Oceans Conference*, Charleston, SC, 2018.

[4] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, 3rd ed. New York: Springer, 2005.

[5] C. Forbes, M. Evans, N. Hastings, and B. Peacock, *Statistical Distributions*, 4th ed. Hoboken, New Jersey: Wiley, 2011.

[6] G. Bateman, "On the power function of the longest run as a test for randomness in a sequence of alternatives," *Biometrika*, vol. 35, no. 1-2, pp. 97–112, May 1948.

[7] J. Glaz, J. Naus, and S. Wallenstein, *Scan Statistics.* New York: Springer-Verlag, 2001.

[8] P. Williams, "Evaluating the state probabilities of m out of n sliding window detectors," Maritime Operations Division Aeronautical and Maritime Research Laboratory, Tech. Rpt. DSTO-TN-0132, February 1998.

[9] D. A. Abraham, "Analysis and design of sliding m-of-n detectors," CausaSci LLC, Tech. Rpt. 2011-01, 2011. [Online]. Available: http://www.dtic.mil/get-tr-doc/pdf?AD=ADA565377

[10] J. I. Naus, "Approximations for distributions of scan statistics," *Journal of the American Statistical Association*, vol. 77, no. 377, pp. 177–183, 1982.

[11] A. Farina and F. A. Studer, *Radar Data Processing: Volume 1—Introduction and Tracking.* John Wiley & Sons, 1985.

[12] D. A. Abraham, "Parameter estimation and performance modeling in generalized-Pareto-distributed clutter," University of Washington, Applied Physics Laboratory, 1013 NE 40th Street, Seattle, WA 98105, Tech. Rpt. TR 2401, January 2024. [Online]. Available: http://apl.uw.edu/tr2401

A Terms from Naus's approximations

Definitions of Naus's Q_2' and Q_3' terms from [10] or [7, Sect. 4.2] are found below. The functions $f_\mathrm{b}(M;N,p)$ and $F_\mathrm{b}(M;N,p)$ are, respectively, the probability mass function (PMF) and cumulative distribution function (CDF) of a binomial (N,p) random variable evaluated at argument M.

The following terms are used in Sect. 2.2 to approximate the PMF and CDF of the stopping time of a sliding M-of-N detector and in Sect. 2.3 for the average stopping time.

$$Q_2' = F_\mathrm{b}^2(M-1;N,p) - (M-1)f_\mathrm{b}(M;N,p)F_\mathrm{b}(M-2;N,p) + Npf_\mathrm{b}(M;N,p)F_\mathrm{b}(M-3;N-1,p) \tag{A1}$$

$$Q_3' = F_\mathrm{b}^3(M-1;N,p) - A_1 + A_2 + A_3 - A_4 \tag{A2}$$

$$A_1 = 2f_\mathrm{b}(M;N,p)F_\mathrm{b}(M-1;N,p)[(M-1)F_\mathrm{b}(M-2;N,p) - NpF_\mathrm{b}(M-3;N-1,p)] \tag{A3}$$

$$A_2 = 0.5f_\mathrm{b}^2(M;N,p)\Big[(M-1)(M-2)F_\mathrm{b}(M-3;N,p) - 2(M-2)NpF_\mathrm{b}(M-4;N-1,p) + N(N-1)p^2F_\mathrm{b}(M-5;N-2,p)\Big] \tag{A4}$$

$$A_3 = \sum_{k=1}^{M-1} f_\mathrm{b}(2M-k;N,p)F_\mathrm{b}^2(k-1;N,p) \tag{A5}$$

$$A_4 = \sum_{k=2}^{M-1} f_\mathrm{b}(2M-k;N,p)f_\mathrm{b}(k;N,p)[(k-1)F_\mathrm{b}(k-2;N,p) - NpF_\mathrm{b}(k-3;N-1,p)] \tag{A6}$$

B MATLAB® code for randomized sliding M-of-N detectors

B.1 Analysis

Example evaluation

```
% File name:  test_rmofn.m
% Example analysis of a randomized sliding M-of-N design
N=5;        % Fixed size of sliding M-of-N test
p0=1e-3;    % Bernoulli success probability under H0
p1=0.5;     % Bernoulli success probability under H1
log10F=10;  % log10(average time between false alarms)

[M0,e]=rmofn_design(p0,N,log10F)
% Results:  M0=3, e=0.0393
D0=mofn_asn_naus_mod(p1,[M0 M0+1],N)*[1-e;e]
% Results:  D0=7.5285
[Mopt,Nopt,eopt,Dopt]=rmofn_optimize(p0,p1,log10F)
% Results:  Mopt=3, Nopt=7, eopt=0.1981, Dopt=6.9969
```

Average stopping time: asymptotic approximation for small p_0

```
function F=mofn_asn_asy(p0,M,N,qLog10)
% F=mofn_asn_naus(p0,M,N,qLog10)
%     Average sample number for a sliding M-of-N process using a small-p0 approximation
%     useful for evaluating the average time between false alarms.
% Parameters:  [mix of scalar and common-size matrix/vector]
%   p0 = Bernoulli probability of success (p0<<1)
%   M = number of successes required for stopping
%   N = window over which successes are counted
%   qLog10 = 1 to return log10(F)
%
if nargin<4, qLog10=0; end;
F=-M.*log(p0)-(N-M+1).*log(1-p0)+gammaln(M)+gammaln(N-M+1)-gammaln(N);
if qLog10,
  F=F*log10(exp(1));
else
  F=exp(F);
end;
```

Average stopping time: modified Naus approximation

```
function T=mofn_asn_naus_mod(p,M,N)
% T=mofn_asn_naus_mod(p,M,N)
%     Average sample number for a sliding M-of-N process using Naus'
%     approximation adjusted to allow stopping after M observations.
% Parameters:  [mix of scalar and common-size matrix/vector]
%   p = Bernoulli probability of success
%   M = number of successes required for stopping
%   N = window over which successes are counted
%
[p,M,N,T]=input_par_std(p,M,N,nan); Ni=numel(p);
for i=1:Ni,
  if M(i)>1,
    T(i)=mofn_stats_scan_fcn(M(i),N(i),p(i));
  else
    T(i)=1/p(i);
  end;
end;
%-----------------
function [Tout,cg,qg]=mofn_stats_scan_fcn(m,n,p)
fb=binopdf(m,n,p); Fbm1=binocdf(m-1,n,p);
Fbm2=binocdf(m-2,n,p); Fbm3=binocdf(m-3,n,p);
Fbm3n1=binocdf(m-3,n-1,p); Fbm4n1=binocdf(m-4,n-1,p);
Fbm5n2=binocdf(m-5,n-2,p);
Q2=Fbm1.^2-(m-1).*fb.*Fbm2+n.*p.*fb.*Fbm3n1;
A1=2*fb.*Fbm1.*((m-1).*Fbm2-n.*p.*Fbm3n1);
A2=0.5*(fb.^2).*((m-1).*(m-2).*Fbm3-2*(m-2).*n.*p.*Fbm4n1+n.*(n-1).*(p.^2).*Fbm5n2);
r=1:(m-1);
A3=sum(binopdf(2*m-r,n,p).*(binocdf(r-1,n,p).^2));
r=2:(m-1);
A4=sum(binopdf(2*m-r,n,p).*binopdf(r,n,p).*((r-1).*binocdf(r-2,n,p)-n.*p.*binocdf(r-3,n-1,p)));
Q3=Fbm1.^3-A1+A2+A3-A4;
qg=(Q3/Q2)^(1/n); cg=Fbm1; j=(1:n)';
Tout=cg*(n+1/(1-qg))+p*sum(j.*binopdf(m-1,j-1,p));
```

Support function

```
function varargout=input_par_std(varargin)
% [x1,x2,x3,...] = input_par_std(x1,x2,x3,...)
%   Finds the maximum dimension of the input variables (up to 2-D) and
%   fills scalar and vector inputs to be that size 2-D array in the output.
%
nrc=[cellfun('size',varargin,1)' cellfun('size',varargin,2)'];
varargout=cell(1,nargout);
for i=1:nargout, varargout{i}=repmat(varargin{i},max(nrc)-nrc(i,:)+1); end
```

B.2 Design and optimization

Basic design

```
function [M0,e]=rmofn_design(p0,N,log10F)
% [M0,e]=rmofn_design(p0,N,log10F)
%    Designs a randomized sliding M-of-N detector to achieve a desired average
%    time between false alarms
% Parameters:
%    p0 = Bernoulli success probability when only noise is present (e.g., Pf)
%    N = size of test
%    log10F = log10(F) where F = average time between false alarms [units: samples]
%
logFfun=@(M,N,p0) log10(exp(1))*(-M.*log(p0)-(N-M+1).*log(1-p0)+gammaln(M)+gammaln(N-M+1)-gammaln(N));
M0=floor(log10F/-log10(p0)); % Initialize M0 at lower bound
while logFfun(M0+1,N,p0)<log10F, M0=M0+1; end; % Increase to bound F
% Find probability detection is delayed (i.e., prob. M=M0+1)
logF01=logFfun([M0 M0+1],N,p0);
e=(10^(log10F-logF01(1))-1)./((10^diff(logF01))-1);
```

Bounds on N given M_0

```
function [N0,N1]=mofn_bound_N(M0,p0,log10F)
% [N0,N1]=mofn_bound_N(M0,p0,log10F)
%    Bounds on N in a randomized sliding M-of-N detector for achieving a
%    false alarm rate (FAR) specification given M0 and p0
% Input parameters: [all are scalar]
%    M0 = largest value of M in a sliding M-of-N detector not achieving FAR specification
%    p0 = Bernoulli probability under noise-only hypotheses
%    log10F = log10(average number of samples between false alarms) = log10(1/FAR)
% Output paremeters:
%    [N0,N1] = min and max values of N achieving or exceeding FAR specification
%
N0=M0; % Increase N0 until below the FAR spec for this value of M
while mofn_asn_asy(p0,M0,N0,1)>log10F,
  N0=N0+1;
end;
% Now increase N0 using M0+1 until no longer meet spec
N1=N0;
while mofn_asn_asy(p0,M0+1,N1+1,1)>log10F,
  N1=N1+1;
end;
```

Optimization over M_0 and N

```
function [Mopt,Nopt,eopt,Dopt]=rmofn_optimize(p0,p1,log10F)
% [Mopt,Nopt,eopt,Dopt]=rmofn_optimize(p0,p1,log10F)
%    Optimal design of a randomized sliding M-of-N detector achieving a specified
%    average time between false alarms for fixed Bernoulli probabilities
% Input parameters: [all are scalar]
%    (p0,p1) = Bernoulli probabilities under noise-only and signal-present hypotheses
%    log10F = log10(average number of samples between false alarms) = log10(1/FAR)
% Output paremeters:
%    (Mopt,Nopt,eopt) = design parameters of the optimized randomized sliding M-of-N detector
%    Dopt = average number of samples before detection
%

% Find the smallest M that will satisfy the FAR spec
M0=floor(log10F/-log10(p0));
while mofn_asn_asy(p0,M0+1,M0+1,1)<log10F, M0=M0+1; end;
% Increase M until find global optimum
Qstop=0; Dopt=inf;
while ~Qstop,
  % Optimize test for this M
  [N0,N1]=mofn_bound_N(M0,p0,log10F);
  N0=max(N0,M0+1);
  % Loop through these until find minimum over N
  Dtmp=inf;
  for n=N0:N1,
    % Tune randomized sliding M-of-N test to achieve FAR specification
    log10F01=mofn_asn_asy(p0,M0+[0 1],n,1);
    e=(10^(log10F-log10F01(1))-1)./((10^diff(log10F01))-1);
    Dmn=mofn_asn_naus_mod(p1,M0+[0 1],n)*[1-e;e];
    if Dmn>Dtmp, % Stop searching for this M0 if D is now increasing
      break;
    else  % Otherwise update and continue
      Dtmp=Dmn; etmp=e; Ntmp=n;
    end;
  end;
  % Test for minimum over M0
  if Dtmp>Dopt, % Stop if D for this M0 is larger
    Qstop=1;
  else  % Otherwise set these to the minimum and continue
    Mopt=M0; Nopt=Ntmp; eopt=etmp; Dopt=Dtmp;
    M0=M0+1; % Increment M0 until global minimum is found
  end;
end;
```