Two Strategies in Transient Change Detection

Daniel Egea-Roca, Blaise Kévin Guépié, José A. López-Salcedo, *Senior Member, IEEE*, Gonzalo Seco-Granados, *Senior Member, IEEE*, and Igor V. Nikiforov

***Abstract*—The paper addresses the transient change detection (TCD) problem, assuming that the duration of change is finite. The TCD criterion minimizes the worst-case probability of missed detection among all tests with a prescribed worst-case probability of false alarm. We study the fixed sample size (FSS) test as a solution to the TCD problem. First, the operating characteristics of the FSS test have been established for arbitrary pre- and post-change distributions. Next, a numerical method of the sample (block) size optimization has been considered for three particular log-likelihood ratio distributions, i.e., Gaussian, χ^2 and exponential. Moreover, simple asymptotic equations for the optimal operating characteristics and block size have been proposed in the Gaussian case. Numerical results are provided to confirm the theoretical findings for the above-mentioned distributions. The accuracy and sharpness of the asymptotic analytical equation is analyzed in the Gaussian case. Finally, the FSS test is compared to the finite moving average (FMA) test obtained by optimizing the CUSUM-type test with respect to the TCD optimality criterion for the above-mentioned distributions. The application of the FSS and FMA tests to the radio-navigation integrity monitoring is also considered.**

***Index Terms*—Transient change detection, Neyman-Pearson test, fixed sample size test, finite moving average, hypothesis testing.**

I. Introduction

THE field of detection theory has usually been linked to the classical problem of hypothesis testing, which has been treated in so many textbooks since its inception [1], [2], and its has been applied to a wide range of applications [3]–[5]. On the other side, the problem of abrupt change detection in random signals has many important novel applications such as fault detection, on-line monitoring of safety-critical infrastructures, and segmentation of signals, just to mention a few [6]–[8]. This kind of detection lies outside the field of classical hypothesis testing and it uses the background of the sequential change detection (SCD), including *quickest* change detection (QCD) and more recently *transient* change detection (TCD).

Manuscript received June 23, 2021; revised December 10, 2021 and February 18, 2022; accepted February 21, 2022. Date of publication March 9, 2022; date of current version March 25, 2022. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Pia Addabbo. This work was supported in part by the Spanish Agency of Research under the research and development project under Grant PID2020-118984GB-I00, in part by the Catalan ICREA Academia Program, and in part by the University of Technology of Troyes, France. *(Corresponding author: Igor V. Nikiforov.)*

Daniel Egea-Roca, José A. López-Salcedo, and Gonzalo Seco-Granados are with the Department of Telecommunications and Systems Engineering, IEEC-CERES, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain (e-mail: daniel.egea@uab.cat; jose.salcedo@uab.cat; gonzalo.seco@uab.cat).

Blaise Kévin Guépié and Igor V. Nikiforov are with the Computer Science and Digital Society Lab, Université de Technologie de Troyes, 10004 Troyes, France (e-mail: blaise_kevin.guepie@utt.fr; igor.nikiforov@utt.fr).

Digital Object Identifier 10.1109/TSP.2022.3158008

In the QCD problem the post-change period is infinitely long. Let $\{y_n\}_{n\geq 1}$ be a sequence of independent random variables and let ν be the index of the first post-change observation:

$$y_n \sim \begin{cases} F_0 \text{ if } n < \nu, \\ F_1 \text{ if } n \geq \nu, \end{cases} \tag{1}$$

where F_0 is the pre-change cumulative distribution function (CDF) and F_1 is the post-change CDF. Let $\mathcal{P}^\nu$ be the joint distribution of the observations $y_1, y_2, \ldots, y_\nu, y_{\nu+1}, \cdots$ when $\nu < \infty$. Let $\mathcal{P}^\infty$ denote the same when $\nu = \infty$, i.e. there is no change and all the observations $y_1, y_2, \ldots$ are i.i.d. with CDF F_0. Let $\mathbb{E}^\nu$ (resp. $\mathbb{E}^\infty$) and $\mathbb{P}^\nu$ (resp. $\mathbb{P}^\infty$) be the expectation and probability with respect to (w.r.t.) the distribution $\mathcal{P}^\nu$ (resp. $\mathcal{P}^\infty$). Lorden [9] proposed an optimality criterion which involves the minimization of the worst-worst-case mean detection delay:

$$\overline{\mathbb{E}}(T) = \sup_{\nu\geq 1} \operatorname{esssup} \mathbb{E}^\nu \left[(T-\nu+1)^+ | y_1, \ldots, y_{\nu-1} \right] \tag{2}$$

among all stopping times T belonging to the class $C_\eta = \{T : \mathbb{E}^\infty(T) \geq \eta\}$, where $(x)^+ = \max(0, x)$ and $\eta > 0$ is a prescribed value of the average run length to a false alarm. He proved that the cumulative sum (CUSUM) test, previously introduced by Page [10], is asymptotically optimal w.r.t. criterion (2) when $\eta \to \infty$. A non asymptotic optimality of the CUSUM test has been established by Moustakides [11].

In contrast to the QCD, the TCD problem is motivated by some applications when the observed phenomena (say, an underwater acoustic signal [12], a navigation system fault [13] or a cyber-physical attack on the supervisory control and data acquisition (SCADA) system [14]) is of short and maybe unknown (and random) duration L:

$$y_n \sim \begin{cases} F_0 \text{ if } 1 \leq n < \nu \text{ or } n > \nu + L - 1 \\ F_1 \text{ if } \nu \leq n \leq \nu + L - 1 \end{cases}. \tag{3}$$

There are two main scenarios of TCD. The first scenario corresponds to the situation when the observed phenomena is of unknown and random duration L. The case of geometrically distributed $L \sim \text{Geom}(\psi)$ with a known parameter ψ and an optimal solution to the TCD problem in this case are considered in [15].

The second scenario arises when the observed phenomena (fault, attack or intrusion) leads to a serious degradation of the system security when the transient change is detected with the delay greater than a required time-to-alert L, i.e. $T - \nu + 1 > L$. For this reason the transient changes detected with the delay greater than L are assumed to be missed. Typical examples of safety and security critical systems include but are not limited to: navigation systems integrity monitoring [13], [16], [17], detection of cyber-physical attacks on SCADA systems [14], drinking water monitoring [18] or sodium-cooled fast reactors monitoring [19] and its heat exchanger leak detection [20]. A detailed numerical example of the GNSS integrity monitoring is discussed in Section VII-A.

The TCD problem has been considered in [12]–[15], [18], [21]–[28]. In particular, it was shown that the Shewhart test minimizes the probability of missed detection among all stopping times T belonging to the class C_η for the special case $L = 1$ in [24], [25]. A solution minimizing the probability of missed detection in a restricted class of truncated sequential probability ratio tests provided that the probability of false alarm is upper bounded have been established in [18], [26]. It was shown that the window-limited CUSUM test with an optimal variable threshold is reduced to the finite moving average (FMA) test. However, the issue of optimality or asymptotic optimality for the second scenario is still open.

If the pre- and post-change distributions F_0 and F_1 are known, the non-sequential or fixed sample size (FSS) statistical test is based on the *repeated* Neyman-Pearson (N-P) test. The conventional N-P test minimizes the probability of missed detection (PMD) provided that the probability of false alarm (PFA) is upper bounded [2] for a sample of a given size. In contrast to the sequential tests, the FSS test decides in favor of one of two alternative hypotheses by using a sample (block) of an optimal size. Hence, the FSS (*repeated* N-P test) necessitates the optimization of the block size. The FSS strategy is usually suboptimal w.r.t. the QCD criterion but it has some practical advantages. For instance, the FSS strategy is more simple, flexible and easily adaptable to the large-scale systems with time-variable structures than the sequential strategy. Typical examples of large-scale systems with time-variable structures are satellite navigation in (sub-) urban environment and network tomography for anomaly detection (see some details and references in [29]).

The global navigation satellite systems (GNSS) are extensively used in rail transport and in-car positioning and navigation in (sub-) urban environment. For some safety/security critical services, the GNSS autonomous integrity monitoring (fault detection and exclusion) is an important function. In contrast to the civil aviation application, some lines-of-sight to satellites are blocked by buildings, other urban constructions, trees, and interference. Hence, some channels disappear and reappear (after re-acquisition) in an unpredictable way. For the fault detection problem, this means that the dimensions of observations, nuisance and informative parameters, and the distribution of noise are time-variable in pre- and post-change modes. There are no results regarding the sequential approach applicable to such a situation in the literature. The FSS test is a suboptimal but practicable solution easily (re-) adjustable in each data block for such a system with a time-variable structure.

Furthermore, the FSS-type tests represent an alternative to the sequential tests when the sensors and computers are connected by a network or the data streams are necessarily preprocessed by blocks. In fact, the raw data is usually transmitted via network by blocks, because it is difficult to send each new observation one-by-one as required by the sequential approach. A similar example is the data preprocessing of subsequent signal blocks by the discrete Fourier transform (especially in the case of high sampling rate) in the spectral analysis. This makes the FSS-type tests interesting alternative to the sequential tests for these batch-type applications. Such a perspective also justifies the study of the FSS test, especially its block size optimization.

The history of comparisons between sequential and non-sequential, i.e., FSS strategies, in the theory of statistical hypotheses testing and signal detection is quite long, some results and references can be found in [6], [8], [30], [31]. The first comparison between optimal sequential and FSS strategies in QCD was performed by A.N. Shiryaev in [32]–[34] for the Bayesian approach and under assumption of a long stationary pre-change regime (when $\nu \to \infty$). Next, a method for selecting the asymptotically optimum sample size of the FSS test in the QCD problem has been considered in [35] by using the criterion, which is slightly different from (2). The comparison between optimal sequential and FSS strategies in QCD for the non-Bayesian approach by using Lorden's worst-worst-case criterion (2) has been considered in [36] and its generalization to the case of composite hypothesis and to the case of multiple hypotheses after change have been considered in [37] and [29], [36], [38], respectively. It was shown that the optimal minimax sequential change detection/isolation test is asymptotically twice as good as its FSS competitor.

There are no results in the literature on the comparison between the sequential and FSS strategies in the TCD problem. The current paper represents the first attempt in this direction.

The rest of the paper is organized as follows. In Section II we state the TCD problem which is treated in the paper, i.e., the optimality criterion and the original contributions. The proposed FSS test is defined in Section III. Its competitor, i.e., the FMA test is also briefly presented here. The operating characteristics of the FSS test are studied in Section IV for an arbitrary distribution. We consider three particular distributions in Section V. Subsection V-A is devoted to the Gaussian mean case. First, we propose upper bounds for the operating characteristics of the FSS test. Second, we consider the block size optimization. Two non-Gaussian log-likelihood ratio (LLR) settings, i.e., the detection of transient changes in the Gaussian variance and in the rate of the exponential distribution are considered in Subsection V-B and V-C. The accuracy and sharpness of the asymptotic solution of Gaussian mean case is examined in Section VI by its comparison to the numerical block size optimization. The numerical optimization for two non-Gaussian settings is also considered here. Section VII is dedicated to the comparison of the FSS test against the FMA test for three different LLR distributions (Gaussian, χ^2 and exponential) and to the application of these tests to the radio-navigation integrity monitoring. The discussion of the obtained results is given in Section VIII and Section IX concludes the paper.

II. Problem Statement and Contributions

Let $\{y_n\}_{n\geq 1}$ be a sequence of independent random variables defined by the generative model of transient changes (3). The goal of the current paper is to study the operating characteristics of the FSS test by using the TCD optimality criterion proposed in [18], [26]:

$$\inf_{T\in C_\alpha}\left\{\mathbb{P}_{\text{md}}(T)\stackrel{\text{def}}{=}\sup_{\nu\geq 1}\mathbb{P}^{\nu}(T-\nu+1>L|T\geq\nu)\right\} \quad (4)$$

among all stopping times $T\in C_\alpha$ satisfying

$$C_\alpha=\left\{T:\mathbb{P}_{\text{fa}}(T)\stackrel{\text{def}}{=}\sup_{\ell\geq 1}\mathbb{P}^{\infty}(\ell\leq T<\ell+m_\alpha)\leq\alpha\right\}, \quad (5)$$

where m_α is a given reference period. Next, we optimize the operating characteristics of the FSS test by choosing the optimal block size and, finally, we compare the FSS test against the FMA test.

The original contributions of this work are the following.

- Expressions for the worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}})$ (4) and PFA $\mathbb{P}_{\text{fa}}(T_{\text{FSS}})$ (5) of the FSS test are established for arbitrary pre-change and post-change distributions.
- The optimization of the FSS test w.r.t the TCD criterion (4) – (5) is performed in the Gaussian mean case. An asymptotically optimal block size is $N^*=\lceil L/2\rceil$. It minimizes an upper bound for the worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}})$ provided that the worst-case PFA $\mathbb{P}_{\text{fa}}(T_{\text{FSS}})$ in a given period m_α is upper bounded. The numerical optimization confirms the sharpness of the asymptotic upper bound for the worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}})$ for small values of the worst-case PFA $\mathbb{P}_{\text{fa}}(T_{\text{FSS}})$, typical in applications.
- The numerical optimization of the FSS test w.r.t the TCD criterion (4)–(5) is performed for two non-Gaussian LLR settings, i.e., the transient change detection in the variance of Gaussian distribution and in the rate of the exponential distribution.
- The comparison between the optimized FSS and FMA tests is performed for three different LLR distributions (Gaussian, χ^2 and exponential). This comparison is interesting because no optimal solution for the TCD problem is available, so it is worth considering the FSS as a possible candidate. The window-limited CUSUM test is also considered in the Gaussian mean case.
- Finally, the FSS and FMA tests are applied to the radio-navigation integrity monitoring and their statistical characteristics are compared in this case.

III. Transient Change Detection: FSS and FMA Tests

The goal of this section is to define the competitors, i.e. FSS and FMA tests for TCD.

A. FSS Test

The general FSS strategy can be described as follows: sample blocks with a fixed size N are collected, and at the end of each sample block a decision between $\mathcal{H}_0=\{y_n\sim F_0\}$ and $\mathcal{H}_1=\{y_n\sim F_1\}$ is taken. We stop sampling as soon as a decision is made in favour of $\mathcal{H}_1$. The solution to the optimal hypothesis testing problem is given by the fundamental Neyman-Pearson lemma [2], and then the stopping time of the FSS (or *repeated* N-P test) is given by

$$T_{\text{FSS}}(N,h)\stackrel{\text{def}}{=}N\cdot\inf\{\kappa\geq 1:\delta_\kappa(h)=1\} \quad (6)$$

with $\delta_\kappa(h)$ the decision at the κ-th block of N observations of the N-P test, defined as

$$\delta_\kappa(h)\stackrel{\text{def}}{=}\begin{cases}1 & \text{if } S_\kappa\geq h\\ 0 & \text{if } S_\kappa<h\end{cases}, \quad (7)$$

where h is a chosen threshold and S_κ is the LLR corresponding to the κ-th block of N observations

$$S_\kappa=\sum_{i=(\kappa-1)N+1}^{\kappa N}\lambda_i,\quad \lambda_i=\log\frac{f_1(y_i)}{f_0(y_i)}, \quad (8)$$

where f_j is the probability density function (PDF) of the CDF F_j, $j=0,1$ and λ_i is the LLR of observation y_i.

Remark 1: To simplify the notations, we consider the case of absolutely continuous distributions F_0 and F_1 (with densities). In the case of discrete or continuous-discrete distributions, a randomization on the boundary h is necessary (see details in [2]).

B. FMA Test

Initially, the window-limited CUSUM with a constant threshold has been considered by Lai in [39] and shown to be asymptotically optimal in the QCD problem for minimizing average detection delay for i.i.d. and non-i.i.d stochastic models. Next, in papers [18], [26], a window-limited CUSUM with a variable threshold

$$T_{\text{WL}}=\inf\left\{n\geq L:\max_{1\leq k\leq L}\left[\sum_{i=n-k+1}^{n}\lambda_i-a(k)\right]\geq 0\right\}, \quad (9)$$

where $a(k)$ is a variable threshold, has been proposed as a solution to the TCD problem. It was shown in [18], [26] that the optimization of the variable threshold $a(k)$ w.r.t. the criterion (4)–(5) reduces the window-limited CUSUM test T (9) to the FMA test given by the stopping time

$$T_{\text{FMA}}(a)=\inf\left\{n\geq L:\sum_{i=n-L+1}^{n}\lambda_i\geq a\right\}. \quad (10)$$

In conclusion, as it follows from [26], the FMA test is the best sequential competitor (benchmark for the TCD) of the FSS test for today's level of knowledge w.r.t. the criterion (4)–(5). In particular, the FMA test outperforms the conventional CUSUM (see [26]), window-limited CUSUM (see Section VII) and modified CUSUM tests (see [15]). As it follows from (6)–(8) and (10), the FSS and FMA tests are equivalent, i.e. $T_{\text{FSS}}=T_{\text{FMA}}$ when $N=L=1$ and $a=h$.

IV. Performance of FSS Test for an Arbitrary Distribution

The goal of this section is to examine the statistical properties of the FSS test w.r.t. the criterion (4)–(5). First, we provide the error probabilities of the N-P test w.r.t the TCD criterion. Next, the optimization of the FSS test is examined.

Let us consider a sample of size N, $y_1,\ldots,y_N\sim F_j$, $j=0,1$. Let $G_j=G_j(N,h)=\mathbb{P}_j(S\leq h)$ be the CDF of the LLR $S=\sum_{k=1}^{N}\lambda_k$, when $y_1,\ldots,y_N\sim F_j$, $j=0,1$. It follows from

(7)–(8), that the error probabilities are given by:

$$\alpha_0(N,h) \overset{\text{def}}{=} \mathbb{P}_0(S \geq h) = 1 - G_0(N,h),$$

$$\alpha_1(N,h) \overset{\text{def}}{=} \mathbb{P}_1(S < h) = G_1(N,h), \tag{11}$$

where $\mathbb{P}_j(\dots)$ stands for the probability of an event when the hypothesis $\mathcal{H}_j = \{y_1, \dots, y_N \sim F_j\}$ is true, $j = 0, 1$.

A. *Worst-Case PFA and PMD*

Let us consider the FSS stopping time defined in (6), where $\delta_\kappa(h)$ is the decision of the N-P test at the κ-th block of N observations. Thereby, the performance of the FSS stopping time in terms of the TCD criterion is given by the two following theorems for the worst-case probabilities of missed detection (4) and false alarm (5), respectively.

Theorem 1 (Worst-case PFA of the FSS test): Let T_{FSS} be the FSS stopping time (6) with the decision function $\delta_\kappa(h)$ (7)–(8), where the block size is $N \geq 1$ and the detection threshold is h. The worst-case PFA in a given reference period m_α boils down to

$$\mathbb{P}_{\text{fa}}(T_{\text{FSS}}; N, h) = 1 - [G_0(N,h)]^{\lceil \frac{m_\alpha}{N} \rceil}, \tag{12}$$

where $\lceil x \rceil = \min\{n \in \mathbb{Z} | n \geq x\}$ stands for the ceiling function.

Proof: The proof is given in Appendix A. ■

Theorem 2 (Worst-case PMD of the FSS test): Let T_{FSS} be the FSS stopping time (6) with the decision function $\delta_\kappa(h)$ (7) – (8), where the block size is $N \geq 1$ and the detection threshold is h. Let $L \geq 1$ be the required time-to-alert. Then, the worst-case PMD of the FSS test, $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N, h)$, is given by

$$\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N, h) = \begin{cases} \max\limits_{1 \leq \nu \leq N} A(\nu, N, h) & \text{if } 1 \leq N \leq L \\ 1 & \text{if } N > L \end{cases} \tag{13}$$

with

$$A(\nu, N, h) = \begin{cases} \mathbb{P}^\nu(S_1^\nu < h)[G_1(N,h)]^{\lfloor \frac{L}{N} \rfloor - 1} & \text{if } 1 \leq \nu \leq \nu^* \\ \mathbb{P}^\nu(S_1^\nu < h)[G_1(N,h)]^{\lfloor \frac{L}{N} \rfloor} & \text{if } \nu^* + 1 \leq \nu \leq N \end{cases}, \tag{14}$$

where $\lfloor x \rfloor = \max\{n \in \mathbb{Z} | n \leq x\}$ stands for the floor function, $\nu^* = (\lfloor \frac{L}{N} \rfloor + 1)N - L$ and S_κ^ν stands for the LLR of the transition block of observations with $\nu \in [(\kappa-1)N + 1, \kappa N]$. The first part of the transition block, i.e. $y_{(\kappa-1)N+1}, \dots, y_{\nu-1}$ belongs to the pre-change period and the second part, i.e. $y_\nu, \dots, y_{\kappa N}$ belongs to the post-change period and $\mathbb{P}^\nu(S_\kappa^\nu < h)$ is the PMD of the N-P test applied to this transition block.

Proof: The proof is given in Appendix B. The classical hypothesis testing is based on the assumption that all observations are identically distributed under each hypothesis. Because this assumption is not fulfilled for the above mentioned transition block, the classical hypothesis testing cannot be applied here. The main technical novelty in the proof is the calculation of $A(\nu, N, h)$ as a function of two terms. The term $G_1(N,h)$ is regular and can be calculated using the classical hypothesis testing. The other term, $\mathbb{P}^\nu(S_1^\nu < h)$, is irregular and the classical hypothesis testing cannot be applied to it. The transition block PMD $\mathbb{P}^\nu(S_1^\nu < h)$ and the function $A(\nu, N, h)$ are used in Corollary 2, Theorem 3, Corollary 4 and Corollary 6 for the calculation of the worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N, h)$. ■

Remark 2: It is worth noting that the worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N, h)$ of the FSS test is a periodic function of the change-point ν (the period is equal to N) due to the fact that the observations are i.i.d. under the pre-change regime. Without loss of generality and seeking simplicity, in the sequel we consider only the first block ($\kappa = 1$) of observations $\{y_n\}_{1 \leq n \leq N}$ and the change-point $\nu \in [1, N]$ for calculation of $\mathbb{P}^\nu(S_\kappa^\nu < h)$.

B. *Optimization of FSS Test*

The FSS test (6)–(7) has two tuning parameters: the block size N and the threshold h. Taking into account the criterion (4) – (5), the optimal choice of N and h is reduced to the following optimization problem:

$$\begin{cases} \operatorname*{arginf}\limits_{\{N,h\}} \{\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N, h)\} = \{N^*, h^*\} \\ \text{subject to } \mathbb{P}_{\text{fa}}(T_{\text{FSS}}; N, h) = \alpha \end{cases}, \tag{15}$$

where N^* and h^* are the optimal parameters of the FSS test. In general, this problem has to be numerically solved. The reason is that it is not trivial to obtain a closed-form expression for the statistical properties of the FSS test for the TCD criterion.

V. Performance of FSS Test for Three Particular Distributions

A. *Transient Change in the Mean*

The goal of this section is to examine the Gaussian mean case of the TCD problem and to apply the general results of Section IV to this particular case. With the previous results we realize how difficult is to obtain a closed-form expression for the worst-case PFA and PMD of the FSS test. Even if we know an expression for the worst-case PMD (13)–(14), we would still need numerical computation for the term $A(\nu, N, h)$ in (14). In this section we limit ourselves to the Gaussian mean case allowing us to obtain closed-form expressions for the worst-case PFA and PMD of the FSS test. Even though some additional upper bounding procedure for the worst-case PFA given by (13) – (14) will be necessary to optimize the FSS test. Fortunately, the negative impact of this upper bound is asymptotically negligible when $\alpha \to 0^+$. We will show these results below.

1) The Worst-Case Probabilities in the Gaussian Mean Case: Let us first re-write the generative TCD model (3) for the Gaussian mean case

$$y_n \sim \begin{cases} \mathcal{N}(0, \sigma^2) & \text{if } 1 \leq n < \nu \\ \mathcal{N}(\theta, \sigma^2) & \text{if } \nu \leq n \leq \nu + L - 1 \end{cases}, \tag{16}$$

where $\mathcal{N}(\mu, \sigma^2)$ stands for the Gaussian law with mean μ and variance σ^2. The PDF of $\mathcal{N}(\mu, \sigma^2)$ is given by $f(x; \mu, \sigma^2) = (1/\sqrt{2\pi\sigma^2}) \exp\{-(x-\mu)^2/2\sigma^2\}$. Because a transient change detected with the delay greater than L is assumed to be missed, we do not consider the generative TCD model for the period after $\nu + L - 1$. Let us also re-write the FSS test (7) for the

Gaussian mean case. After simple algebra, we obtain

$$\delta_\kappa(h) \overset{\text{def}}{=} \begin{cases} 1 \text{ if } S_\kappa \geq h \\ 0 \text{ if } S_\kappa < h \end{cases}, S_\kappa = \frac{\theta}{\sigma^2} \sum_{i=(\kappa-1)N+1}^{\kappa N} \left(y_i - \frac{\theta}{2} \right). \tag{17}$$

The case of zero-mean pre-change Gaussian law is assumed here and in the sequel without loss of generality.

Based on the above model, we provide the theoretical results needed to obtain closed-form expressions for the worst-case PMD and PFA of the FSS test and to optimize it w.r.t. the criterion (4)–(5).

Corollary 1 (of Theorem 1: Worst-case PFA in the Gaussian mean case): Let T_{FSS} be the FSS stopping time (6) with the block size $N \geq 1$ and detection threshold h, where the decision function is defined in (17). The worst-case PFA in a given reference period m_α is

$$\mathbb{P}_{\mathrm{fa}}(T_{\mathrm{FSS}}; N, h) = 1 - \left[\Phi \left(\frac{h + N\varrho/2}{\sqrt{\varrho N}} \right) \right]^{\left\lceil \frac{m_\alpha}{N} \right\rceil}, \tag{18}$$

where $\varrho = \theta^2/\sigma^2$ is the signal-to-noise ratio (SNR) and $\Phi(x) = \int_{-\infty}^{x} (1/\sqrt{2\pi}) \exp\{-u^2/2\} du$ denotes the CDF of the standard normal distribution.

Proof: It follows from (16) and (17) that the LLR S_κ obeys the Gaussian distribution $\mathcal{N}(-N\varrho/2, N\varrho)$ under the pre-change regime. It follows from (11) that

$$G_0(N, h) = \mathbb{P}_0(S_\kappa < h) = \Phi \left(\frac{h + N\varrho/2}{\sqrt{\varrho N}} \right). \tag{19}$$

Putting together (12) and (19), we prove (18). ■

Corollary 2 (of Theorem 2: Worst-case PMD in the Gaussian mean case): Let T_{FSS} be the FSS stopping time (6) with the block size N $(1 \leq N \leq L)$ and detection threshold h, where the decision function is defined in (17). Then the PMD γ_1 of the N-P test applied to the transition block is given by

$$\gamma_1(\nu, N, h) = \mathbb{P}^\nu(S_1^\nu < h) = \Phi \left(\frac{h + (\nu - N/2 - 1)\varrho}{\sqrt{\varrho N}} \right), \tag{20}$$

where $S_1^\nu = \frac{\theta}{\sigma^2} \sum_{i=1}^{N} (y_i - \frac{\theta}{2})$ is the LLR of the transition block with $1 \leq \nu \leq N$. The PMD γ_2 of the sequence of N-P tests applied to $\lfloor \frac{L}{N} \rfloor - 1$ blocks contained only the post-change observations is

$$\gamma_2(N, h) = [G_1(N, h)]^{\lfloor \frac{L}{N} \rfloor - 1} = \left[\Phi \left(\frac{h - N\varrho/2}{\sqrt{\varrho N}} \right) \right]^{\lfloor \frac{L}{N} \rfloor - 1}. \tag{21}$$

The worst-case PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, h)$ (13)–(14) of the FSS test is given by

$$\begin{aligned} \mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, h) &= \max_{1 \leq \nu \leq N} A(\nu, N, h) = A(\nu^*, N, h) \\ &= \gamma_1(\nu^*, N, h)\gamma_2(N, h). \end{aligned} \tag{22}$$

Proof: The proof is given in Appendix C. ■

Because (18) defines a bijective function, there exists a unique solution $h = h(\alpha)$ of the equation $\mathbb{P}_{\mathrm{fa}}(T_{\mathrm{FSS}}; N, h) = \alpha$. In the sequel we use α instead of h in the notations, wherever it is not confusing.

2) Asymptotic Optimization in the Gaussian Mean Case: Corollaries 1 and 2 from the previous section provide us with the elements needed to get the optimal block size N^* for the FSS test. Nevertheless, even in the Gaussian mean case, the problem (15) is analytically intractable due to the numerous combinations of L and N, the presence of the functions floor $\lfloor . \rfloor$, ceil $\lceil . \rceil$ and $\nu^* = (\lfloor \frac{L}{N} \rfloor + 1)N - L$. The main difficulty corresponds to the worst-case PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, h)$ given by (22) with the terms $\gamma_1(\nu, N, h)$ and $\gamma_2(N, h)$ defined in (20) and (21).

Therefore, the FSS test optimization will be done by using the criterion of optimality (4)–(5), the worst case PFA (18) and the upper bound $\overline{\gamma}(N, \alpha) \geq \mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, \alpha) = \gamma_1(\nu^*, N, \alpha)\gamma_2(N, \alpha)$ for the worst-case PMD (22). Nevertheless, even with such an upper bound for $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, \alpha)$, its minimization is not a trivial problem and it can be done only asymptotically when $\alpha \to 0^+$.

Theorem 3 (Optimization problem): Let T_{FSS} be the FSS stopping time (6) with the block size $N \geq 1$ and the detection threshold h, where the decision function is defined in (17). It is assumed that $T_{\mathrm{FSS}} \in \mathcal{C}_\alpha$. The optimal choice of the block size N is reduced to the following minimization problem

$$N^* = \underset{1 \leq N \leq L}{\operatorname{argmin}} \, \overline{\gamma}(N, \alpha) \text{ as } \alpha \to 0^+, \tag{23}$$

where $\overline{\gamma}(N, \alpha)$ is an upper bound for $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, \alpha)$ defined as follows

$$\overline{\gamma}(N, \alpha) = \begin{cases} \gamma_1(N, N, \alpha)\gamma_2(N, \alpha) & \text{if } 1 \leq N \leq \lfloor \frac{L}{2} \rfloor \\ \gamma_1(\nu^*, N, \alpha) & \text{if } \lfloor \frac{L}{2} \rfloor + 1 \leq N \leq L. \end{cases} \tag{24}$$

$$\gamma_1(\nu, N, \alpha) = \Phi \left[\Phi^{-1} \left((1 - \alpha)^{\frac{1}{\left\lceil \frac{m_\alpha}{N} \right\rceil}} \right) - \frac{(N - \nu + 1)\sqrt{\varrho}}{\sqrt{N}} \right] \tag{25}$$

$$\gamma_2(N, \alpha) = \left(\Phi \left[\Phi^{-1} \left((1 - \alpha)^{\frac{1}{\left\lceil \frac{m_\alpha}{N} \right\rceil}} \right) - \sqrt{\varrho N} \right] \right)^{\lfloor \frac{L}{N} \rfloor - 1}. \tag{26}$$

The asymptotically optimal parameters N^* and h^* of the FSS test (6), (17), which realize the optimum in (23) as $\alpha \to 0^+$, are given by

$$N^* = \lceil L/2 \rceil \tag{27}$$

$$h^* = \sqrt{\varrho \lceil L/2 \rceil} \Phi^{-1} \left((1 - \alpha)^{\frac{1}{\left\lceil \frac{m_\alpha}{\lceil L/2 \rceil} \right\rceil}} \right) - \frac{\varrho}{2} \lceil L/2 \rceil. \tag{28}$$

Proof: The detailed proof is given in Appendix D but a short sketch of the proof is as follows. We note the following technical novelty and subtlety of the worst-case PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, \alpha)$ minimization w.r.t. the criterion (4)–(5). Let $T_{\mathrm{FSS}} \in \mathcal{C}_\alpha$. The exact worst-case PMD $N \mapsto \mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, \alpha)$ as a function of the block size N for a given value of α is untractable due to a number of local minima and maxima. This fact motivated the authors to divide the interval of definition $[1, L]$ $(L \geq 2)$ of natural numbers N into two subintervals $[1, \lfloor L/2 \rfloor]$ and $[\lfloor L/2 \rfloor + 1, L]$ and to use the following upper bounds instead of exact function $N \mapsto \mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, \alpha)$. The PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, \alpha)$ is a product of the two terms: γ_1 and γ_2, see (22). If $N \in [1, \lfloor L/2 \rfloor]$, the term to minimize is $\gamma_2(N,\alpha)$ but the term $\gamma_1(\nu^*, N, \alpha)$ is just upper bounded by $\gamma_1(N,N,\alpha)$, which tends to 1 when $\alpha \to 0^+$. If $N \in [\lfloor L/2 \rfloor + 1, L]$, the term to minimize is $\gamma_1(\nu^*, N, \alpha)$ but $\gamma_2(N,\alpha) = 1$. Let $\overline{\gamma}_1$ (resp. $\overline{\gamma}_2$) be an upper bound for γ_1 (resp. γ_2). To find the optimal block size N^*, which minimizes the upper bound for $A(\nu^*, N, \alpha)$, we show that $N \mapsto \overline{\gamma}_2(N,\alpha)$ is a decreasing function in $[1, \lfloor L/2 \rfloor]$ and $N \mapsto \overline{\gamma}_1(\nu^*, N, \alpha)$ is an increasing function in $[\lfloor L/2 \rfloor + 1, L]$. Finally, to "glue" these two functions together, we consider two different cases: *i)* $L = 2n$, $n \in \mathbb{Z}^+$, is an even number; *ii)* $L = 2n+1$ is an odd number and we get the optimal block size $N^* = \lceil L/2 \rceil$. ∎

Let us examine now two TCD problems, where the LLR S_κ of the FSS test is non-Gaussian. The first case study is the detection of transient changes in the variance of Gaussian distribution and the second one is the detection of transient changes in the rate of the exponential distribution. The operating characteristics will be used for the numerical optimization of the block size N.

B. Transient Change in the Variance

Let us re-write the generative TCD model (3) for the Gaussian variance case

$$y_n \sim \begin{cases} \mathcal{N}(0,\sigma_0^2) \text{ if } 1 \le n < \nu \\ \mathcal{N}(0,\sigma_1^2) \text{ if } \nu \le n \le \nu + L - 1 \end{cases}, \tag{29}$$

where $\mathcal{N}(0,\sigma^2)$ stands for the zero-mean Gaussian law with variance σ^2. Let us also re-write the FSS test (7) for the Gaussian variance case. Without loss of generality, let us assume in the sequel that $\sigma_0 < \sigma_1$ [1]. After simple algebra, we obtain

$$\delta_\kappa(h) \overset{\text{def}}{=} \begin{cases} 1 & \text{if } S_\kappa \ge h \\ 0 & \text{if } S_\kappa < h \end{cases}, \; S_\kappa = \sum_{i=(\kappa-1)N+1}^{\kappa N} y_i^2. \tag{30}$$

Based on the above model, we provide the theoretical results needed to obtain the expressions for the worst-case PMD and PFA of the FSS test and to numerically optimize it w.r.t. the criterion (4)–(5).

Corollary 3 (of Theorem 1: Worst-case PFA in the Gaussian variance case): Let T_{FSS} be the FSS stopping time (6) with the block size $N \ge 1$ and detection threshold h, where the decision function is defined in (30). The worst-case PFA in a given reference period m_α is

$$\mathbb{P}_{\text{fa}}(T_{\text{FSS}}; N, h) = 1 - \left[F_{\chi^2}\left(h/\sigma_0^2; N\right)\right]^{\left\lceil \frac{m_\alpha}{N} \right\rceil}, \tag{31}$$

where $F_{\chi^2}(x;N) = \int_0^x \frac{e^{-u/2}u^{N/2-1}}{2^{N/2}\Gamma(N/2)} du$ denotes the CDF of the χ_N^2 distribution with N degrees of freedom and $x \mapsto \Gamma(x)$ is the gamma function.

Proof: By substituting the definition of CDF $G_0(N,h) = \mathbb{P}^\infty(S_\kappa < h) = F_{\chi^2}(h/\sigma_0^2; N)$ into (12), we immediately obtain (31). ∎

Corollary 4 (of Theorem 2: Worst-case PMD in the Gaussian variance case): Let T_{FSS} be the FSS stopping time (6) with the block size N ($1 \le N \le L$) and detection threshold h, where the decision function is defined in (30). Then the PMD γ_1^{Var} of the N-P test applied to the transition block is given by

$$\gamma_1^{\text{Var}}(\nu, N, h) = \mathbb{P}^\nu(S_1^\nu < h) = \mathbb{P}\left(\sum_{i=1}^{N} d_i(\nu)\xi_i^2 < h\right), \tag{32}$$

where the sum $S_1^\nu = \sum_{i=1}^N y_i^2$ corresponds to the transition block with $1 \le \nu \le N$, $\xi_i \sim \mathcal{N}(0,1)$ and

$$d_i(\nu) = \begin{cases} \sigma_0^2 \text{ if } i < \nu \\ \sigma_1^2 \text{ if } \nu \le i \le N \end{cases}. \tag{33}$$

The CDF of S_κ under the distribution $\mathcal{N}(0,\sigma_1^2)$ is given by

$$G_1^{\text{Var}}(N,h) = \mathbb{P}_1(S_\kappa < h) = F_{\chi^2}\left(h/\sigma_1^2; N\right). \tag{34}$$

The worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N, h)$ of the FSS test (6), (30) is given by (13)–(14) with $\gamma_1 = \gamma_1^{\text{Var}}(\nu, N, h)$ (32) and $G_1(N,h) = G_1^{\text{Var}}(N,h)$ (34).

Proof: The proof of (32)–(33) follows from the definitions of the transient change generative model (29) and the decision function (30). The proof of (34) follows immediately from the definition of the χ_N^2 distribution. ∎

Because (31) defines a bijective function, there exists a unique solution $h = h(\alpha)$ of the equation $\mathbb{P}_{\text{fa}}(T_{\text{FSS}}; N, h) = \alpha$. The main problem is to compute γ_1^{Var}. The exact CDF of the quadratic form in Gaussian variables $\sum_{i=1}^N d_i(\nu)\xi_i^2$ is available only for some special cases and its calculation is very difficult (see Chapter 4.6 in [40]). Instead, we use an approximation given in [40], [41]:

$$\mathbb{P}\left(\sum_{i=1}^{N} \widetilde{d}_i \xi_i^2 < x\right) = \min\left\{H_1(x), H_2(x)\right\}, \tag{35}$$

where $\xi_i \sim \mathcal{N}(0,1)$, $\widetilde{d}_i > 0$, $i = 1, \ldots, N$, $\sum_{i=1}^N \widetilde{d}_i = 1$,

$$H_1(x) = \sum_{i=1}^{N} \widetilde{d}_i \frac{\gamma\left(1/(2\widetilde{d}_i), x/(2\widetilde{d}_i)\right)}{\Gamma(1/(2\widetilde{d}_i))}, \tag{36}$$

the incomplete gamma function is $\gamma(p,x) = \int_0^x u^{p-1}e^{-u} du$ and

$$H_2(x) = F_{\chi^2}(x/\delta_d; N), \quad \delta_d = \left(\prod_{i=1}^{N} \widetilde{d}_i\right)^{1/N}. \tag{37}$$

Finally, we get

$$\gamma_1^{\text{Var}}(\nu, N, h) = \mathbb{P}\left(\sum_{i=1}^{N} \widetilde{d}_i(\nu)\xi_i^2 < \frac{h}{d(\nu)}\right), \tag{38}$$

where $\xi_i \sim \mathcal{N}(0,1)$, $\widetilde{d}_i = d_i(\nu)/d(\nu)$ stands for the normalized coefficients and $d(\nu) = \sum_{i=1}^N d_i(\nu)$ for $1 \le \nu \le N$.

C. Transient Change in Exponential Distribution

Let us re-write the generative TCD model (3) for the exponential distribution $\text{Exp}(\lambda)$

$$y_n \sim \begin{cases} \text{Exp}(\lambda_0) \text{ if } 1 \le n < \nu \\ \text{Exp}(\lambda_1) \text{ if } \nu \le n \le \nu + L - 1 \end{cases}, \tag{39}$$

[1] If $\sigma_0 > \sigma_1$ then the inequality in the definition of $\delta_\kappa(h)$ (30) must be reversed.

where λ stands for the rate of the exponential distribution. The PDF of $\mathrm{Exp}(\lambda)$ is given by

$$f(x;\lambda)=\begin{cases}\lambda e^{-\lambda x} & \text{if } x\geq 0\\ 0 & \text{if } x<0\end{cases}. \tag{40}$$

Let us also re-write the FSS test (7) for the exponential distribution $\mathrm{Exp}(\lambda)$. Without loss of generality, let us assume in the sequel that $\lambda_0>\lambda_1$ [2]. After simple algebra, we obtain

$$\delta_\kappa(h)\stackrel{\text{def}}{=}\begin{cases}1 & \text{if } S_\kappa\geq h\\ 0 & \text{if } S_\kappa<h\end{cases},\ S_\kappa=\sum_{i=(\kappa-1)N+1}^{\kappa N} y_i. \tag{41}$$

Based on the above model, we provide the theoretical results needed to obtain the expressions for the worst-case PMD and PFA of the FSS test and to numerically optimize it w.r.t. the criterion (4) – (5).

Corollary 5 (of Theorem 1: Worst-case PFA in the exponential case): Let T_{FSS} be the FSS stopping time (6) with the block size $N\geq 1$ and detection threshold h, where the decision function is defined in (41). The worst-case PFA in a given reference period m_α is

$$\mathbb{P}_{\mathrm{fa}}(T_{\mathrm{FSS}};N,h)=1-\left[F_{\chi^2}\left(2h\lambda_0;2N\right)\right]^{\left\lceil\frac{m_\alpha}{N}\right\rceil}. \tag{42}$$

Proof: It follows from the definition of χ^2_n distribution that the exponentially distributed random variable with $\lambda=1/2$ obeys the χ^2_2 distribution with 2 degrees of freedom, i.e., $\xi\sim\mathrm{Exp}(1/2)\sim\chi^2_2$. By substituting the definition of CDF $G_0(N,h)=\mathbb{P}_0(S_\kappa<h)=F_{\chi^2}(2h\lambda_0;2N)$ into (12), we immediately obtain (42).

As previously, the main problem is to compute the PMD γ_1^{Exp} of the N-P test applied to the transition block. The distribution of the sum of exponential variables $\sum_{i=1}^N y_i$, where $y_i\sim\mathrm{Exp}(\lambda_i)$, obeys the hypoexponential distribution (or the generalized Erlang distribution) [42]. The exact CDF of the hypoexponential distribution is available only for some special cases of $\lambda_1,\ldots,\lambda_N$ and its calculation is very difficult. Hence, based on the fact that $\xi\sim\mathrm{Exp}(1/2)\sim\chi^2_2$, we use again the approximation proposed in [41].

Corollary 6 (of Theorem 2: Worst-case PMD in the exponential case): Let T_{FSS} be the FSS stopping time (6) with the block size N ($1\leq N\leq L$) and detection threshold h, where the decision function is defined in (41). Then the PMD γ_1^{Exp} of the N-P test applied to the transition block is given by

$$\gamma_1^{\mathrm{Exp}}(\nu,N,h)=\mathbb{P}^\nu(S_1^\nu<h)=\mathbb{P}\left(\sum_{i=1}^N d_i(\nu)\zeta_i<h\right), \tag{43}$$

where the sum $S_1^\nu=\sum_{i=1}^N y_i$ corresponds to the transition block with $1\leq\nu\leq N$, $\zeta_i\sim\chi^2_2$ and

$$d_i(\nu)=\begin{cases}1/(2\lambda_0) & \text{if } i<\nu\\ 1/(2\lambda_1) & \text{if } \nu\leq i\leq N\end{cases}. \tag{44}$$

The CDF of S_κ under the distribution $\mathrm{Exp}(\lambda_1)$ is given by

$$G_1^{\mathrm{Exp}}(N,h)=\mathbb{P}_1(S_\kappa<h)=F_{\chi^2}\left(2h\lambda_1;2N\right). \tag{45}$$

[2]If $\lambda_0<\lambda_1$ then the inequality in the definition of $\delta_\kappa(h)$ (41) must be reversed.

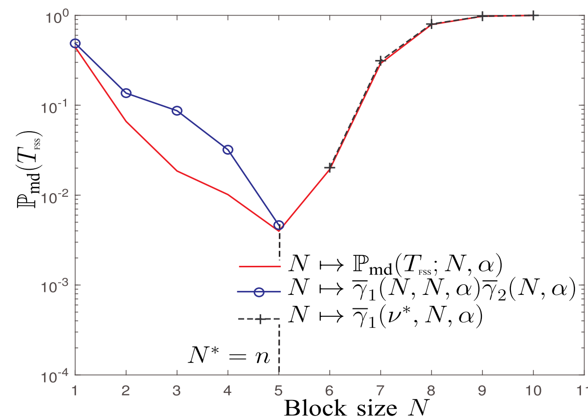

Fig. 1. The true worst-case PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}};N,\alpha)$ given by (22) and its upper bounds as functions of the block size N. The SNR is $\varrho=9$, the required time-to-alert is $L=10$, the reference period is $m_\alpha=20$, and the worst-case PFA is $\alpha=10^{-4}$.

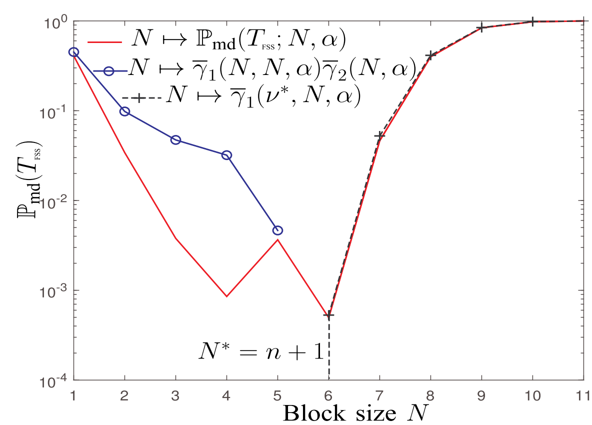

Fig. 2. The true worst-case PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}};N,\alpha)$ given by (22) and its upper bounds as functions of the block size N. The SNR is $\varrho=9$, the required time-to-alert is $L=11$, the reference period is $m_\alpha=20$, and the worst-case PFA is $\alpha=10^{-4}$.

The worst-case PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}};N,h)$ of the FSS test (6), (41) is given by (13)–(14) with $\gamma_1=\gamma_1^{\mathrm{Exp}}(\nu,N,h)$ (43) and $G_1(N,h)=G_1^{\mathrm{Exp}}(N,h)$ (45).

Proof: The proof of (43)–(44) follows from the definitions of the transient change generative model (39) and the decision function (41). The proof of (45) follows immediately from the definition of the χ^2_{2N} distribution. ∎

VI. Numerical Results

A. Gaussian Mean Case

This section is dedicated to the numerical analysis of the accuracy and sharpness of the asymptotic solution obtained in optimization Theorem 3. First of all, let us compare the true PMD $N\mapsto\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}};N,\alpha)$ given by (22) with its upper bounds $N\mapsto\overline{\gamma}_1(N,N,\alpha)\overline{\gamma}_2(N,\alpha)$ defined in the interval $[1,\lfloor L/2\rfloor]$ and $N\mapsto\overline{\gamma}_1(\nu^*,N,\alpha)$ defined in the interval $[\lfloor L/2\rfloor+1,L]$ for two different cases of $L=2n$ and $L=2n+1$. The results of this comparison are presented in Figs. 1 and 2 for the

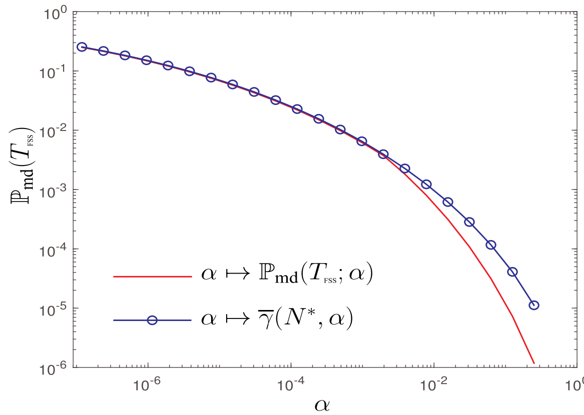

Fig. 3. The minimum worst-case PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}})$ given by (46) and its upper bound $\overline{\gamma}(N^*)$ (24) with the asymptotically optimal N^* as functions of the worst-case PFA α. The SNR is $\varrho = 4$, the required time-to-alert is $L = 20$, the reference period is $m_\alpha = 150$, and the worst-case PFA belongs to the interval $\alpha \in [10^{-7}, 0.25]$.

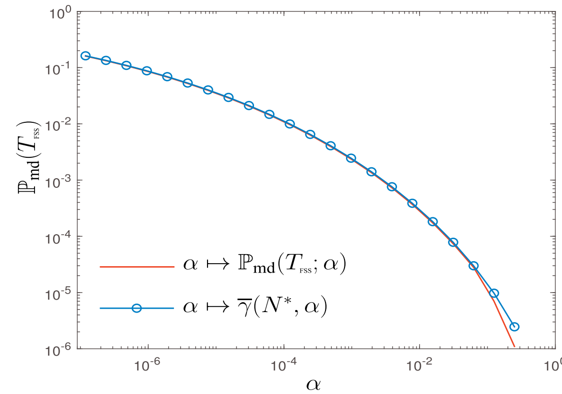

Fig. 4. The minimum worst-case PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}})$ given by (46) and its upper bound $\overline{\gamma}(N^*)$ (24) with the asymptotically optimal N^* as functions of the worst-case PFA α. The SNR is $\varrho = 4$, the required time-to-alert is $L = 21$, the reference period is $m_\alpha = 150$, and the worst-case PFA belongs to the interval $\alpha \in [10^{-7}, 0.25]$.

following parameters: the SNR $\varrho = 9$, $L = 10, 11$, $m_\alpha = 20$, and $\alpha = 10^{-4}$. Here and in the rest of the paper, we use the parameters typical for some real-time applications, such as the navigation systems integrity monitoring and the drinking water monitoring (see [13], [16]–[18], [43] and the example on GNSS integrity monitoring given in Section VII-A). First, these figures illustrate the main idea of Theorem 3: if the required time-to-alert is an even number, i.e. $L = 2n$, then the asymptotically optimal block size is $N^* = n$ but if $L = 2n + 1$ then $N^* = n + 1$. Second, as it follows from the definition of the function $N \mapsto \gamma_1(N, N, \alpha)$ in the interval $[1, \lfloor L/2 \rfloor]$, this upper bound cannot be very accurate for $N \mapsto \gamma_1(\nu^*, N, \alpha)$ due to the fact that $\nu^* = (\lfloor \frac{L}{N} \rfloor + 1)N - L$ is simply replaced by N. Hence, it is clear that for the non-asymptotic values of α, the upper bounding in the interval $[1, \lfloor L/2 \rfloor]$ is less accurate than in the interval $[\lfloor L/2 \rfloor + 1, L]$. Hence, a more rapid convergence of the upper bound to the true worst-case PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; \alpha)$ can be expected in the case of $L = 2n + 1$. This conclusion is confirmed by Figs. 3 and 4. Here, the non-asymptotic minimum worst-case PMD

$$\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; \alpha) = \min_{1 \le N \le L} \{\gamma_1(\nu^*, N, \alpha)\gamma_2(N, \alpha)\} \qquad (46)$$

obtained by numerical optimization of (22) and its upper bound $\overline{\gamma}(N^*, \alpha)$ in (24) are presented as functions of the worst-case PFA α for the following parameters typical for some real-time applications: SNR $\varrho = 4$, $L = 20, 21$, $m_\alpha = 150$, and $\alpha \in [10^{-7}, 0.25]$. It follows that the 5% accuracy is reached at $\alpha = 10^{-3}$ for $L = 20$ and the same accuracy is reached at $\alpha = 1.5 \cdot 10^{-2}$ for $L = 21$.

B. Two Non-Gaussian LLR Settings

Let us return to the examination of the TCD problem when the sum S_κ of the FSS test is non-Gaussian. The calculation of the worst case PMD and PFA of the FSS test for the detection of transient changes in the variance of Gaussian distribution has been considered in Section V-B and for the detection of transient changes in the rate of the exponential distribution in Section V-C. The goal of this section is to examine the numerical optimization of the block size N for non-Gaussian setting and extend the results obtained for the Gaussian mean case to some practical situations where observations are non-negative (e.g., daily number of infected people in the epidemics).

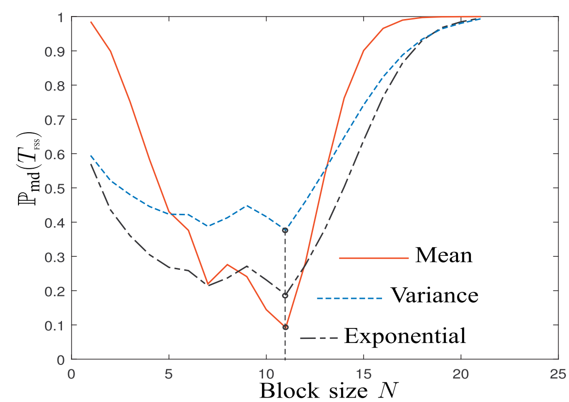

Fig. 5. The worst-case PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, \alpha)$ as a function of the block size N for the following three cases: – Gaussian mean; – Gaussian variance; – exponential distribution. The required time-to-alert is $L = 21$, the reference period is $m_\alpha = 20$, and the worst-case PFA is $\alpha = 10^{-5}$.

Let us compare the worst case PMD as a function of the block size $N \mapsto \mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, \alpha)$ for the Gaussian mean case (22) with the the same functions for the transient changes in the variance, calculated by (13)–(14) with $\gamma_1 = \gamma_1^{\mathrm{Var}}(\nu, N, h)$ (32) and $G_1(N, h) = G_1^{\mathrm{Var}}(N, h)$ (34), and for the transient changes in the rate of the exponential distribution, calculated by (13)–(14) with $\gamma_1 = \gamma_1^{\mathrm{Exp}}(\nu, N, h)$ (43) and $G_1(N, h) = G_1^{\mathrm{Exp}}(N, h)$ (45). The results of comparison are presented in Fig. 5 for the following parameters: $L = 21$, $m_\alpha = 20$, and $\alpha = 10^{-5}$. The models of transient changes are defined as follows:

- Gaussian mean: $\varrho = 3$;
- Gaussian variance: $\sigma_1^2/\sigma_0^2 = 5$;
- Exponential distribution: $\lambda_1/\lambda_0 = 0.25$.

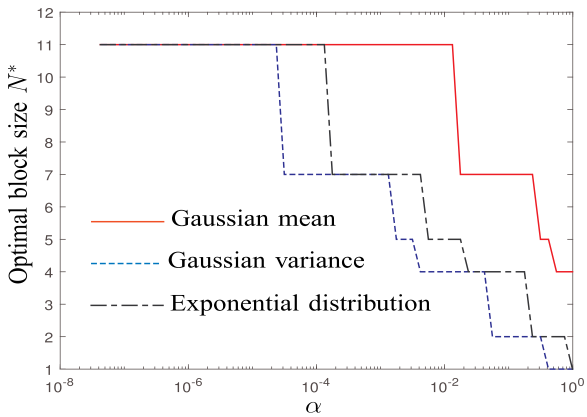

Fig. 6. The optimal block size N^* as a function of the worst-case PFA α for the following three cases: – Gaussian mean; – Gaussian variance; – exponential distribution. The required time-to-alert is $L = 21$ and the reference period is $m_\alpha = 20$.

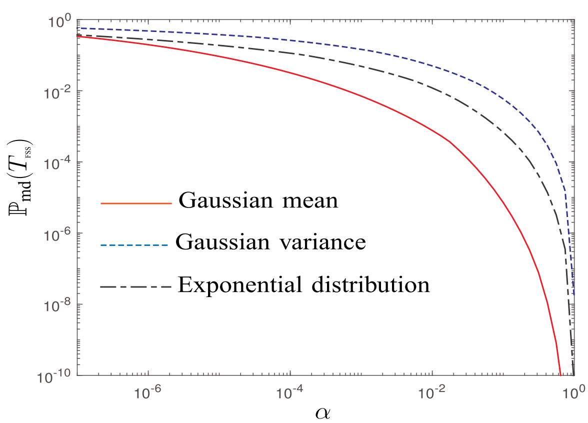

Fig. 7. The minimum worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N^*, \alpha)$ as a function of the worst-case PFA α for the following three cases: – Gaussian mean; – Gaussian variance; – exponential distribution. The required time-to-alert is $L = 21$ and the reference period is $m_\alpha = 20$.

Fig. 5 shows that the behaviour of the function $N \mapsto \mathbb{P}_{\text{md}}(T_{\text{FSS}}; N, \alpha)$ for two non-Gaussian LLR settings is similar to the Gaussian mean case. Less pronounced minima of the function $N \mapsto \mathbb{P}_{\text{md}}(T_{\text{FSS}}; N, \alpha)$ for the non-Gaussian LLR settings is explained by the reduced separability between the pre-change and post-change χ^2 and exponential distributions due to the fact that their support is limited to the real semi-axis $x \in [0, \infty[$.

As it follows from Theorem 3, the main source of the minimum worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N^*(\alpha), \alpha)$ overestimation by its upper bound is a non-asymptotic behaviour of the optimal block size $N^*(\alpha) < \lceil L/2 \rceil$ for significant values of the PFA α. To illustrate this situation, the optimal block sizes N^* as functions of the worst-case PFA α are shown in Fig. 6. Next, the minimum worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N^*(\alpha), \alpha)$ as functions of the worst-case PFA α are shown in 7. Both figures represent the three following settings: – Gaussian mean; – Gaussian variance; – exponential distribution. Due to the reduced separability between the pre-change and post-change χ^2 and exponential distributions, the convergence rate of the optimal block size N^* to its asymptotic value is reduced for the non-Gaussian LLR settings (see Fig. 6). For the three above-mentioned settings, the functions $\alpha \mapsto \mathbb{P}_{\text{md}}(T_{\text{FSS}}; N^*(\alpha), \alpha)$ are quite similar with correction for the separability between the pre-change and post-change distributions (see Fig. 7).

VII. Comparison Between the FSS, Window-Limited CUSUM and FMA Tests

A. Gaussian Mean Case

Let us compare the FSS, window-limited CUSUM and FMA tests in the Gaussian mean case. The upper bound for the worst-case PMD $\overline{\mathbb{P}_{\text{md}}}(T_{\text{FMA}}; \alpha)$ as a function of the worst-case PFA α is given as follows [18], [26]:

$$\overline{\mathbb{P}_{\text{md}}}(T_{\text{FMA}}; \alpha) = \Phi\left(\Phi^{-1}\left((1-\alpha)^{\frac{1}{m_\alpha}}\right) - \sqrt{\varrho L}\right). \quad (47)$$

The minimum worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; \alpha)$ of the FSS test is given by (46) with $\gamma_1(\nu^*, N, \alpha)$ (25) and $\gamma_2(N, \alpha)$ (26). Its upper bound $\overline{\gamma}(N^*, \alpha)$ is given by (24) with the asymptotically optimal block size $N^* = \lceil L/2 \rceil$.

We exclude from our comparison the case of $L = 1$ when the FSS, window-limited CUSUM and FMA tests coincide (see the analysis of this case in [24], [25]). In the case of $L = 1$, the upper bounds for the worst-case PMD become equal to the worst-case PMD for both tests $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; \alpha) = \overline{\gamma}(1, \alpha) = \mathbb{P}_{\text{md}}(T_{\text{FMA}}; \alpha) = \overline{\mathbb{P}_{\text{md}}}(T_{\text{FMA}}; \alpha)$.

Let us compare the equation for $\overline{\mathbb{P}_{\text{md}}}(T_{\text{FMA}}; \alpha)$ with the equations for the minimum worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N^*(\alpha), \alpha)$ and its upper bound $\overline{\gamma}(N^*, \alpha)$ for the FSS test. Their structures are very similar (compare (24) – (26) and (47)): in both cases the argument of the function $x \mapsto \Phi(x)$ is a sum of two terms. The first terms, i.e. $\Phi^{-1}((1-\alpha)^{\frac{1}{m_\alpha}})$ for the FMA test and $\Phi^{-1}((1-\alpha)^{\frac{1}{\lceil \frac{m_\alpha}{N^*} \rceil}})$ for the FSS test, are responsible for the relation between α and $\mathbb{P}_{\text{md}}(T)$. The second terms, i.e. $\sqrt{\varrho L}$ for the FMA test and $\sqrt{\varrho \lceil L/2 \rceil}$ or $\frac{(\lceil L/2 \rceil - \nu + 1)\sqrt{\varrho}}{\sqrt{\lceil L/2 \rceil}}$ for the FSS test, define the square root of information (in Kullback–Leibler sense) obtained during the time-to-alert L. It follows from Theorem 3 that $N^* \to \lceil L/2 \rceil$ when $\alpha \to 0^+$ and in any case, $N^* \geq 1$. Hence, the first term of the FSS test reduces the worst-case PMD not worse (even better) than the first term of the FMA test. On the other hand, the quantity of information extracted by the FMA test is asymptotically (when $\alpha \to 0^+$) twice as large as the quantity of information extracted by the FSS test. Hence, it can be expected that the FMA test should perform better for practically interesting situations when $\alpha \to 0^+$ and the second terms are dominant.

The first comparison between the FSS, FMA and window-limited CUSUM tests is presented in Fig. 8 for the following parameters: $\varrho = 0.25$, $L = 100$, $m_\alpha = 200$, and $\alpha \in [10^{-5}, 0.99]$. Because we are interested in non-asymptotic values of α, we use the minimum worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N^*(\alpha), \alpha)$ and its upper bound $\overline{\gamma}(N^*, \alpha)$ for the FSS tests (which is closed to the true PMD), the simulated worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FMA}}; \alpha)$ and its upper bound for the FMA test and the simulated PMD $\mathbb{P}_{\text{md}}(T_{\text{WL}}; \alpha)$ for the window-limited CUSUM

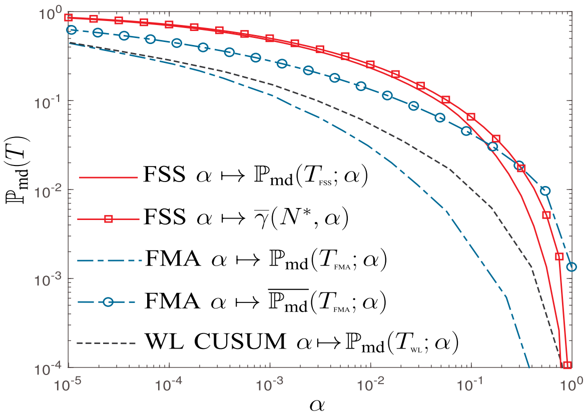

Fig. 8. The minimum worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}})$ of the FSS test, its upper bound $\overline{\gamma}(N^*)$, the simulated worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FMA}})$ of the FMA test, its upper bound given by (47) and the simulated worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{WL}})$ of the window-limited CUSUM test as functions of the worst-case PFA α. The SNR is $\varrho = 0.25$, the required time-to-alert is $L = 100$, the reference period is $m_\alpha = 200$, and the worst-case PFA belongs to the interval $\alpha \in [10^{-5}, 0.99]$.

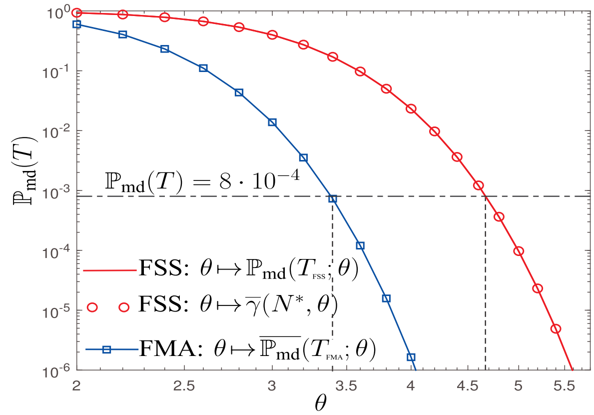

Fig. 9. The minimum worst-case PMD of the FSS and FMA tests as functions of the fault amplitude $\theta \in [2, 5.8]$.

test (9) with $a(k) = a, k = 1, \ldots, L$. To get the simulated PMD $\mathbb{P}_{\text{md}}(T_{\text{FMA}};\alpha)$ and $\mathbb{P}_{\text{md}}(T_{\text{WL}};\alpha)$, 10^6 Monte-Carlo runs have been performed. It follows that the FMA test outperforms the FSS and window-limited CUSUM tests.

The second comparison between the FSS and FMA tests is related to the important problem of radio-navigation integrity monitoring. Certainly, the considered model of signal is simplified but it is sufficient to show very promising results for the GNSS (GPS, Glonass, Galileo, Beidou,...) integrity monitoring.

Example [GNSS integrity monitoring] For safety-critical civil aircraft navigation modes (landing, takeoff, etc.), a major problem of existing navigation systems consists of their lack of integrity. The integrity monitoring concept requires a navigation system to detect the faults and remove them from the navigation solution before they sufficiently contaminate the output. Let us consider the receiver autonomous integrity monitoring (RAIM) of the GNSS channels. Some insight into the RAIM can be found in [13], [43]. The RAIM is a method of GNSS integrity monitoring that uses redundant measurements (satellite pseudoranges) at the user's satellite receiver. In the nominal situation the pseudorange measurements are affected by random noise obeying a zero-mean Gaussian distribution. Suddenly arriving additional biases (faults) in pseudoranges lead to a degradation of the position estimation, which is clearly undesirable. Due to the redundant pseudorange measurements, the residuals of the least squares estimator provide the user with information on the presence or absence of faults in measured pseudoranges. Hence, a simplified model of pre-change and post-change pseudorange residuals is reduced to (16), where ν is the fault onset time. The minimum operational performance standards (MOPS) for the GPS system [43] specifies, for the (non-) precision approach of typical duration m_α, the required time-to-alert L (both are measured in number of sampling periods), the worst-case PMD and the worst-case PFA during a period m_α. The signal sampling period is usually 1 sec. Let us compare the FMA and FSS tests for the following typical parameters: $L = 6$, $m_\alpha = 150$, $\alpha = 2 \cdot 10^{-5}$. The variance in the TCD model (16) is $\sigma^2 = 1$ to simplify notations. The worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}})$ and $\mathbb{P}_{\text{md}}(T_{\text{FMA}})$ as functions of the residual fault amplitude θ are shown in Fig. 9. It follows that the FMA test significantly outperforms the FSS test with the optimal block size $N^* = 3$ for all values of $\theta \in [2, 5.8]$. A typical MOPS in terms of the worst-case PMD is $\mathbb{P}_{\text{md}}(T) = 8 \cdot 10^{-4}$. Therefore, the FMA test respects this MOPS beginning from $\theta = 3.4$ but the FSS test only from $\theta = 4.7$. The minimum detectable fault amplitude (respecting the MOPS) defines the horizontal/vertical alarm limits. The reduction of these alarm limits improves the safety of aircraft and air traffic control.

B. Two Non-Gaussian LLR Settings Vs. The Gaussian Mean Case

Let us continue the comparison of the FSS and FMA tests for two non-Gaussian LLR settings (χ^2 and exponential) to verify how the operating characteristics of the FSS and FMA tests differ from those obtained in Section VII-A for the Gaussian mean case.

In the case of transient change in the variance, the upper bound for the worst-case PMD $\overline{\mathbb{P}_{\text{md}}}(T_{\text{FMA}};\alpha)$ as a function of the worst-case PFA α is given as follows [18], [26]:

$$\overline{\mathbb{P}_{\text{md}}}(T_{\text{FMA}};\alpha) = F_{\chi^2}\left(\frac{\sigma_0^2}{\sigma_1^2} F_{\chi^2}^{-1}\left((1-\alpha)^{\frac{1}{m_\alpha}}; L\right); L\right), \quad (48)$$

where $x \mapsto F_{\chi^2}(x; L)$ stands for the CDF of the χ^2 distribution with L degrees of freedom, and in the case of exponential distribution, the upper bound for the worst-case PMD $\overline{\mathbb{P}_{\text{md}}}(T_{\text{FMA}};\alpha)$ is

$$\overline{\mathbb{P}_{\text{md}}}(T_{\text{FMA}};\alpha) = F_{\chi^2}\left(\frac{\lambda_1}{\lambda_0} F_{\chi^2}^{-1}\left((1-\alpha)^{\frac{1}{m_\alpha}}; 2L\right); 2L\right). \quad (49)$$

As previously (see Subsection V-B and V-C), it is assumed that $\sigma_0 < \sigma_1$ and $\lambda_0 > \lambda_1$. We use the same set of parameters as in Section VI-B: $L = 21$ and $m_\alpha = 20$. First, we compare the FSS and FMA tests for the transient change detection in Gaussian mean as a reference case for the non-Gaussian settings. The

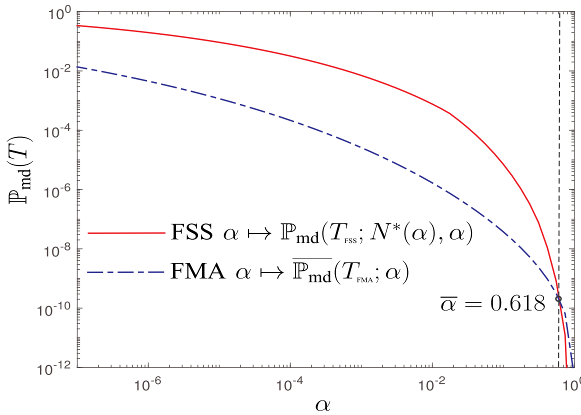

Fig. 10. The minimum worst-case PMD $\mathbb{P}_{\rm md}(T_{\rm FSS})$ of the FSS test and the upper bound for the PMD $\overline{\mathbb{P}_{\rm md}}(T_{\rm FMA})$ of the FMA test given by (47) as functions of the worst-case PFA α for the Gaussian mean case. The SNR is $\varrho = 3$, the required time-to-alert is $L = 21$, $m_\alpha = 20$, and the worst-case PFA belongs to the interval $\alpha \in [10^{-7}, 0.99]$.

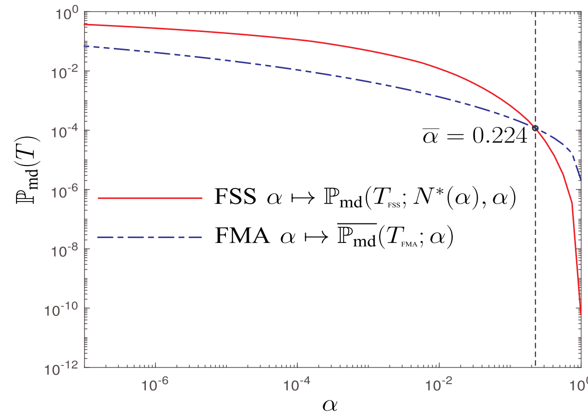

Fig. 12. The minimum worst-case PMD $\mathbb{P}_{\rm md}(T_{\rm FSS})$ of the FSS test and the upper bound for the PMD $\overline{\mathbb{P}_{\rm md}}(T_{\rm FMA})$ of the FMA test given by (49) as functions of the worst-case PFA α for the exponential case. The parameter ratio is $\lambda_1/\lambda_0 = 0.25$, the required time-to-alert is $L = 21$, $m_\alpha = 20$, and the worst-case PFA belongs to the interval $\alpha \in [10^{-7}, 0.99]$.

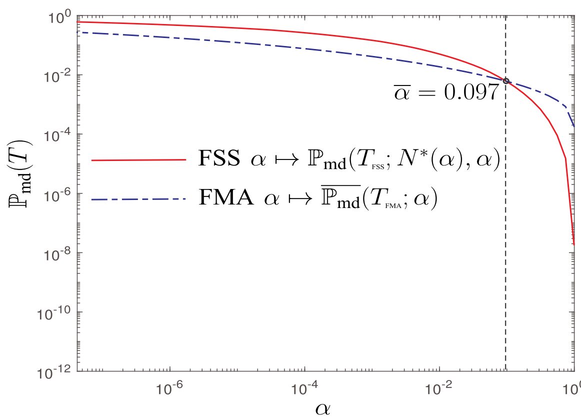

Fig. 11. The minimum worst-case PMD $\mathbb{P}_{\rm md}(T_{\rm FSS})$ of the FSS test and the upper bound for the PMD $\overline{\mathbb{P}_{\rm md}}(T_{\rm FMA})$ of the FMA test given by (48) as functions of the worst-case PFA α for the Gaussian variance case. The variances ratio is $\sigma_1^2/\sigma_0^2 = 5$, the required time-to-alert is $L = 21$, $m_\alpha = 20$, and the worst-case PFA belongs to the interval $\alpha \in [10^{-7}, 0.99]$.

results for the SNR $\varrho = 3$ are presented in Fig. 10. The comparison of FSS and FMA tests for the Gaussian variance with the ratio $\sigma_1^2/\sigma_0^2 = 5$ is shown in Fig. 11 and for the exponential case with the ratio $\lambda_1/\lambda_0 = 0.25$ is shown in Fig. 12. It follows from Figs. 10, 11 and 12 that the FMA test outperforms the FSS test for both the Gaussian LLR setting and the non-Gaussian LLR settings for all values of the PFA $\alpha \in [10^{-7}, \overline{\alpha}]$, where $\overline{\alpha}$ takes the values 0.618 for the Gaussian mean, 0.097 for the Gaussian variance and 0.224 for the exponential distribution. Hence, the relationship between the FSS and FMA tests is similar for three different settings with correction for the separability between the pre-change and post-change distributions, which only impacts the value of $\overline{\alpha}$.

VIII. Discussion

As it follows from the FSS stopping time definition (6) – (8), the FSS test is based on the repeated N-P tests applied to the sequence of blocks of a fixed size N. Hence, a key problem of the FSS test designer is to find an optimal block size N^* by using the transient change detection criterion (4) – (5) which minimizes the worst-case probability of missed detection, i.e., the situation when $T_{\rm FSS} - \nu + 1 > L$, provided that the worst-case probability of false alarm in a given period m_α is upper bounded. Knowing N^*, the optimal threshold h^* follows immediately from the definition of the class C_α (5).

To realize such a scheme, Theorems 1 and 2 provide the designer with the equations for the worst-case probabilities of missed detection $\mathbb{P}_{\rm md}(T_{\rm FSS})$ and false alarm $\mathbb{P}_{\rm fa}(T_{\rm FSS})$ as functions of the FSS tuning parameters N, h and the pre- and post-change distributions F_0 and F_1. Corollaries 1 – 6 concretize these general results in the special cases of the Gaussian, χ^2 and exponential LLR distributions. These three cases cover two different types of measured signals defined by their support. The support of the Gaussian distribution is the real axis $x \in]-\infty, \infty[$ but the supports of the χ^2 and exponential distributions are limited to the real semi-axis $x \in [0, \infty[$. Next, Theorem 3 establishes an asymptotically optimal solution for the Gaussian mean case. Here, the asymptotically optimal FSS tuning parameters N^*, h^* and the upper bound $\overline{\gamma}(N^*, \alpha)$ for the probability of missed detection $\mathbb{P}_{\rm md}(T_{\rm FSS}; \alpha)$ are represented as functions of the upper bound α for $\mathbb{P}_{\rm fa}(T_{\rm FSS})$, required time-to-alert L, reference period m_α and SNR ϱ. The comparison between the optimal probability of missed detection $\mathbb{P}_{\rm md}(T_{\rm FSS}; \alpha)$, obtained by the numerical FSS test optimization, and its asymptotically optimal upper bound $\overline{\gamma}(N^*, \alpha)$ shows that this upper bound is rather sharp in practically interesting case of small α. A numerical optimization of the FSS test has been considered for the χ^2 and exponential distributions.

The optimal FSS test has been compared against the FMA test for the three above-mentioned LLR distributions. The choice of the FMA test as a sequential competitor is justified for the following reasons. The general issue of optimality or asymptotic optimality in the TCD problem (4) – (5) is still open but the FMA test is obtained by optimizing the window-limited

CUSUM test w.r.t. criterion (4) – (5) in a restricted class of truncated sequential tests. These comparisons show that the relationship between the operating characteristics of the FSS and FMA tests in the case of non-Gaussian LLR settings does not differ substantially from the Gaussian mean case. Some slight difference is due to the different separability between the pre-change and post-change distributions in the cases of Gaussian, χ^2 and exponential settings. It is shown that the FSS test outperforms the FMA test for significant probability of false alarm, which has perhaps only theoretical rather than practical importance. For small probabilities of false alarm, which is typical in applications, the FMA test outperforms the FSS test for all settings.

It is worth noting that the relationship between the FSS test and the FMA test as sequential competitor in the TCD problem qualitatively corresponds to the relationship between the FSS and CUSUM tests in the QCD problem (see [29], [32]–[38], [44]). For both the QCD and the TCD problems the FSS test outperforms the FMA (resp. CUSUM) test for a high rate of false alarms, but for a lower rate of false alarms (which is typical in applications) the FMA (resp. CUSUM) test performs better than the FSS test. Theorem 3 shows that the asymptotically optimal block size N^* of the FSS test is equal to a half of the required time-to-alert L in the TCD problem. Accordingly, as follows from [29], [36], [38], the asymptotically optimal block size N^* of the FSS test is equal to a half of the worst-worst-case mean detection delay $\overline{\mathbb{E}}(T)$ (2) in the QCD problem.

Often the algorithm designer has to choose between the operating characteristic of the transient change detector and its complexity. Hence, another performance metric to consider in comparing the FSS and FMA tests is their respective complexity of implementation. Although the realization of the FSS and FMA tests involves the same number of LLR calculation per observation, the FSS test is simpler and easily adaptable for practical implementation due to its block-by-block method of data transmission, LLR calculation and decision-making. By using the obtained theoretical results, the designer can find a tradeoff between the loss of optimality and the computational burden of the detection scheme.

IX. Conclusion

In this paper, we have studied the statistical properties of the FSS test w.r.t. the transient change detection criterion which minimizes the worst-case probability of missed detection provided that the worst-case probability of false alarm in a given period is upper bounded. The operating characteristics of the FSS test are obtained in the general case of arbitrary pre- and post-change distributions. Next, these operating characteristics have been concretized for three particular LLR distributions: Gaussian, χ^2 and exponential. The optimization of the FSS tuning parameters has been considered for these distributions. An asymptotically optimal analytical solution has been obtained in the Gaussian case. Next, the FSS test has been compared against its sequential competitor, i.e., the FMA test. Finally, the application of both the FSS and the FMA tests to the radio-navigation integrity monitoring and their comparison demonstrate some promising results.

Appendix A
Proof of Theorem 1

To prove (12), let us first calculate the maximum number N_d of decisions taken by the N-P test (7) during the reference period m_α. It is obvious that $N_d = \left\lceil \frac{m_\alpha}{N} \right\rceil$. Due to the fact that the pre-change observations are i.i.d., the random number $M \in [0, N_d]$ of the events $\{d_n = 1\}$ during the reference period m_α follows the binomial distribution $B(N_d, p)$, where $p = 1 - G_0(N, h)$. Hence, the probability of the event $\{M > 0\}$ defines the worst-case PFA of the FSS test with the stopping time T_{FSS} (6)

$$\mathbb{P}_{\mathrm{fa}}(T_{\mathrm{FSS}}; N, h) = 1 - (1-p)^{N_d} = 1 - [G_0(N,h)]^{\left\lceil \frac{m_\alpha}{N} \right\rceil}. \tag{50}$$

Appendix B
Proof of Theorem 2

The proof of Theorem 2 is divided in two parts. In *Parts 1* the worst-case PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}})$ of the FSS test is defined as a product of N-P test error probabilities. *Parts 2* is devoted to the calculation of the number of factors in this product and the worst-case PMD $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}})$.

Part 1: Since the proof for $N > L$ is trivial, i.e. the worst-case PMD is $\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}) = 1$, we will prove the case when $1 \le N \le L$. Let us consider the following probability

$$\mathbb{P}^{\nu \ge 1}(T_{\mathrm{FSS}} > \nu + L) = [\mathbb{P}_0(S_\kappa < h)]^{\left\lfloor \frac{\nu-1}{N} \right\rfloor} \cdot \prod_{\kappa = \left\lfloor \frac{\nu-1}{N} \right\rfloor + 1}^{\left\lfloor \frac{\nu+L-1}{N} \right\rfloor} \mathbb{P}^{\nu \ge N}(S_\kappa < h), \tag{51}$$

where S_κ is the LLR of the N-P test in the κ-th block of observations. Next, we define the probability that no false alarms occur before the change-point

$$\mathbb{P}^{\nu \ge 1}(T_{\mathrm{FSS}} > \nu) = [\mathbb{P}_0(S_\kappa < h)]^{\left\lfloor \frac{\nu-1}{N} \right\rfloor}. \tag{52}$$

Hence, we get the following conditional PMD

$$\mathbb{P}^{\nu \ge 1}(T_{\mathrm{FSS}} > \nu + L | T_{\mathrm{FSS}} \ge \nu) = \prod_{\kappa = \left\lfloor \frac{\nu-1}{N} \right\rfloor + 1}^{\left\lfloor \frac{\nu+L-1}{N} \right\rfloor} \mathbb{P}^{\nu \ge N}(S_\kappa < h). \tag{53}$$

The pre-change observations are i.i.d. and the sequential blocks of observations are also i.i.d. Therefore, this conditional probability is a periodic function of the change-point ν (the period is equal to N). In order to simplify the notation, it is sufficient to consider that $\nu = \{1, 2, \ldots, N\}$:

$$\mathbb{P}_{\mathrm{md}}(T_{\mathrm{FSS}}; N, h) = \max_{\nu \in [1,N]} A(\nu, N, h) = \max_{\nu \in [1,N]} \prod_{\kappa=1}^{\left\lfloor \frac{\nu+L-1}{N} \right\rfloor} \mathbb{P}^{\nu}(S_\kappa < h). \tag{54}$$

Part 2: Let us calculate the number of factors in this product (54). Let us consider two real numbers $x, y \in \mathrm{IR}$. The following inequality takes place $\lfloor x \rfloor + \lfloor y \rfloor \le \lfloor x + y \rfloor \le \lfloor x \rfloor + \lfloor y \rfloor + 1$.

Hence

$$\left\lfloor \frac{\nu-1}{N} \right\rfloor + \left\lfloor \frac{L}{N} \right\rfloor \le \left\lfloor \frac{\nu+L-1}{N} \right\rfloor \le \left\lfloor \frac{\nu-1}{N} \right\rfloor + \left\lfloor \frac{L}{N} \right\rfloor + 1. \tag{55}$$

Since $\nu \in [1, N]$, we get $\lfloor \frac{\nu-1}{N} \rfloor = 0$. Therefore

$$\left\lfloor \frac{L}{N} \right\rfloor \le \left\lfloor \frac{\nu+L-1}{N} \right\rfloor \le \left\lfloor \frac{L}{N} \right\rfloor + 1. \tag{56}$$

It follows that the number of factors $\lfloor \frac{\nu+L-1}{N} \rfloor$ can be equal to $\lfloor \frac{L}{N} \rfloor$ or to $\lfloor \frac{L}{N} \rfloor + 1$. For given N and L, it depends on ν. The function $\nu \mapsto \lfloor \frac{\nu+L-1}{N} \rfloor$ is increasing, for this reason there exists a boundary value of change-point ν^* switching the number of factors from $\lfloor \frac{L}{N} \rfloor$ to $\lfloor \frac{L}{N} \rfloor + 1$. Let us define the following equation for the number of factors:

$$\left\lfloor \frac{\nu+L-1}{N} \right\rfloor = \begin{cases} \lfloor \frac{L}{N} \rfloor & \text{if } 1 \le \nu \le \nu^* \\ \lfloor \frac{L}{N} \rfloor + 1 & \text{if } \nu^*+1 \le \nu \le N \end{cases}. \tag{57}$$

Let us suppose that $\nu^* = (\lfloor \frac{L}{N} \rfloor + 1)N - L$. The validity of this definition of ν^* is established by simple direct calculation for end points of the intervals. For example, let $\nu = \nu^*$. It is easy to see that

$$\left\lfloor \frac{\nu^*+L-1}{N} \right\rfloor = \left\lfloor \frac{(\lfloor L/N \rfloor + 1)\, N - L + L - 1}{N} \right\rfloor = \left\lfloor \frac{L}{N} \right\rfloor. \tag{58}$$

Another end point is $\nu = \nu^* + 1$

$$\left\lfloor \frac{\nu^*+1+L-1}{N} \right\rfloor = \left\lfloor \frac{(\lfloor L/N \rfloor + 1)\, N - L + 1 + L - 1}{N} \right\rfloor = \left\lfloor \frac{L}{N} \right\rfloor + 1. \tag{59}$$

Putting together (11), (54) and (57), we get the product

$$A(\nu, N, h) = \begin{cases} \mathbb{P}^\nu(S_1^\nu < h)[G_1(N,h)]^{\lfloor \frac{L}{N} \rfloor - 1} & \text{if } 1 \le \nu \le \nu^* \\ \mathbb{P}^\nu(S_1^\nu < h)[G_1(N,h)]^{\lfloor \frac{L}{N} \rfloor} & \text{if } \nu^*+1 \le \nu \le N \end{cases}. \tag{60}$$

The worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N, h)$ (13) – (14) of the FSS test follows from (54) and (60) thus completing the proof of Theorem 2.

Appendix C
Proof of Corollary 2

First, we prove the PMD (20) of the transition block. Let us consider the observations $\{y_n\}_{1 \le n \le N}$ governed by the probability measure $\mathcal{P}^\nu$:

$$y_n \sim \begin{cases} \mathcal{N}\left(0, \sigma^2\right) & \text{if } 1 \le n \le \nu - 1 \\ \mathcal{N}\left(\theta, \sigma^2\right) & \text{if } \nu \le n \le N \end{cases}. \tag{61}$$

As it follows from (61), the first $\nu - 1$ pre-change observations obey $\mathcal{N}(0, \sigma^2)$ and the last $N - \nu + 1$ post-change observations obey $\mathcal{N}(\theta, \sigma^2)$. Hence, the LLR $S_1^\nu = \frac{\theta}{\sigma^2} \sum_{i=1}^N (y_i - \frac{\theta}{2})$ obeys the Gaussian distribution

$$S_1^\nu \sim \mathcal{N}(\mu_\nu, N\varrho), \tag{62}$$

where $\varrho = \theta^2/\sigma^2$ is the SNR and

$$\mu_\nu = -(\nu-1)\varrho/2 + (N-\nu+1)\varrho/2 = (N/2 - \nu + 1)\varrho.$$

Therefore, we prove (20)

$$\gamma_1(\nu, N, h) = \mathbb{P}^\nu(S_1^\nu < h) = \Phi\left(\frac{h - \mu_\nu}{\sqrt{\varrho N}}\right) = \Phi\left(\frac{h + (\nu - N/2 - 1)\varrho}{\sqrt{\varrho N}}\right). \tag{63}$$

Formula (21) corresponds to the block filled in only with the post-change observations $y_n \sim \mathcal{N}(\theta, \sigma^2)$ for $1 \le n \le N$. Hence, $G_1(N, h)$ in (21) follows immediately from (63) when $\nu = 1$, i.e., $G_1(N, h) = \gamma_1(1, N, h)$.

Finally, let us prove the last assertion of Corollary 2 given by (22). It follows from (13) – (14) that

$$\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N, h) = \max_{1 \le \nu \le N} A(\nu, N, h)$$
$$= \max \Big\{ \max_{1 \le \nu \le \nu^*} \{\gamma_1(\nu, N, h)\} \cdot \gamma_2(N, h),$$
$$\max_{\nu^*+1 \le \nu \le N} \{\gamma_1(\nu, N, h)\} \cdot \gamma_2(N, h) \cdot G_1(N, h) \Big\}. \tag{64}$$

The function $x \mapsto \Phi(x)$ is increasing and hence $\nu \mapsto \gamma_1(\nu, N, h)$ (20) is also an increasing function of ν and its minimum in the interval $[1, N]$ is reached at $\nu = 1$. Taking into account that $\max_{\nu^*+1 \le \nu \le N} \{\gamma_1(\nu, N, h)\} < 1$, we get

$$\max_{1 \le \nu \le \nu^*} \{\gamma_1(\nu, N, h)\} > \max_{\nu^*+1 \le \nu \le N} \{\gamma_1(\nu, N, h)\} \cdot G_1(N, h). \tag{65}$$

The worst-case PMD $\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N, h)$ (13) – (14) of the FSS test is given by

$$\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N, h) = \max_{1 \le \nu \le N} A(\nu, N, h)$$
$$= \max_{1 \le \nu \le \nu^*} \{\gamma_1(\nu, N, h)\}\, \gamma_2(N, h) = \gamma_1(\nu^*, N, h)\gamma_2(N, h) \tag{66}$$

thus completing the proof of Corollary 2.

Appendix D
Proof of Theorem 3

The proof of Theorem 3 is divided in four parts. *Parts 1–3* are devoted to the case of $L = 2n$ and *Part 4* adapts the previously obtained results to the case of $L = 2n + 1$.

Part 1: Let the required time-to-alert L be an even number, i.e. $L = 2n, n \in \mathbb{Z}^+$ and $T_{\text{FSS}} \in C_\alpha$. As it follows from Corollary 1, the worst-case PMD is given by

$$\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N) = \gamma_1(\nu^*, N)\gamma_2(N), \tag{67}$$

where

$$\gamma_1(\nu, N) = \Phi\left[\Phi^{-1}\left(\beta^{\frac{1}{\lceil \frac{m_\alpha}{N} \rceil}}\right) - \frac{(N-\nu+1)\sqrt{\varrho}}{\sqrt{N}}\right] \tag{68}$$

$$\gamma_2(N)=\left(\Phi\left[\Phi^{-1}\left(\beta^{\frac{1}{\lceil \frac{m_\alpha}{N}\rceil}}\right)-\sqrt{\varrho N}\right]\right)^{\lfloor \frac{L}{N}\rfloor-1},\beta=1-\alpha \tag{69}$$

for any fixed $\alpha : 0<\alpha<1$ (we omit α to simplify the notation).

Let us divide the interval $[1, 2n]$ of possible values of N into two disjoint subintervals: $[1, n]$ and $[n+1, 2n]$. First, we consider the interval $[1, n]$. We cannot establish a simple relation between ν^*, N and L. From the other hand, the function $x\mapsto\Phi(x)$ is increasing and hence $\gamma_1(\nu, N)$ (68) is also an increasing function of ν and its maximum in the interval $[1, N]$ is reached at $\nu = N$. Therefore,

$$\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N)\leq\gamma_1(N,N)\gamma_2(N)\leq\gamma_2(N) \tag{70}$$

over the interval $[1, n]$. Moreover, for any given constants N, m_α and ϱ, the following limit can be established for the first term

$$\begin{aligned}\lim_{\beta\to1^-}\gamma_1(N,N,\beta)&=\lim_{\beta\to1^-}\Phi\left[\Phi^{-1}\left(\beta^{\frac{1}{\lceil \frac{m_\alpha}{N}\rceil}}\right)-\frac{\sqrt{\varrho}}{\sqrt{N}}\right]\\&=1.\end{aligned} \tag{71}$$

Hence, an asymptotically sharp upper bound can be expected for $\mathbb{P}_{\text{md}}(T_{\text{FSS}})$ in the interval $[1, n]$ as $\beta\to1^-$.

Let us consider the interval $[n+1, 2n]$. As it follows from the definition (69), the second term is $\gamma_2(N) = 1$ when $N\in[n+1, 2n]$. Therefore,

$$\mathbb{P}_{\text{md}}(T_{\text{FSS}}; N)=\gamma_1(\nu^*,N) \tag{72}$$

over the interval $[n+1, 2n]$. Therefore, the first assertion (24) of Theorem 3 follows immediately from (70) and (72):

$$\overline{\gamma}(N)=\begin{cases}\gamma_1(N,N)\gamma_2(N) & \text{if } 1\leq N\leq n\\ \gamma_1(\nu^*,N) & \text{if } n+1\leq N\leq 2n\end{cases}. \tag{73}$$

Because $L = 2n$ is an even number, the block size is given by $N = n + k$, where $k\in\{1, 2, \ldots, n\}$. Next, we get $\nu^* = (\lfloor \frac{L}{N}\rfloor+1)N-L=2k$ and hence (68) can be re-written as a function of $k\in\{1,\ldots,n\}$

$$\gamma_1(k)=\Phi\left[\Phi^{-1}\left(\beta^{\frac{1}{\lceil \frac{m_\alpha}{n+k}\rceil}}\right)-\frac{(n-k+1)\sqrt{\varrho}}{\sqrt{n+k}}\right]. \tag{74}$$

Part 2: In the sequel an upper bound for γ_i is denoted by $\overline{\gamma}_i$, $i = 1, 2$. Let us show that the upper bound $N\mapsto\overline{\gamma}_2(N)$ is a decreasing function in the interval $[1, n]$, when $L = 2n$. It is easy to see that

$$\left\lfloor\frac{L}{N}\right\rfloor-1\geq\max\left\{\frac{L}{N}-2,1\right\}\text{ for }1\leq N\leq n \tag{75}$$

and

$$\left\lceil\frac{m_\alpha}{N}\right\rceil\leq\frac{m_\alpha}{N}+1 \tag{76}$$

Therefore, taking into account that the functions $x\mapsto\Phi(x)$ and $x\mapsto\Phi^{-1}(x)$ are increasing, $\beta = 1-\alpha < 1$ and $\Phi(x) < 1$, we get the following upper bound for $\gamma_2(N)$

$$\gamma_2(N)\leq\overline{\gamma}_2(N)$$

$$=\left\{\Phi\left[\Phi^{-1}\left(\beta^{\frac{1}{\frac{m_\alpha}{N}+1}}\right)-\sqrt{\varrho N}\right]\right\}^{\max\left\{\frac{L}{N}-2,1\right\}}. \tag{77}$$

Let $N = Lx$. The upper bound $\overline{\gamma}_2(N)$ can be re-written as a function of x:

$$\overline{\gamma}_2(x)=\left\{\Phi\left(\Phi^{-1}\left(\beta^{\frac{xL}{m_\alpha+xL}}\right)-\sqrt{\varrho Lx}\right)\right\}^{\max\left\{\frac{1}{x}-2,1\right\}} \tag{78}$$

for $x\in[1/L; 1/2]$. Our goal is to show that the function $x\mapsto\overline{\gamma}_2(x)$ is decreasing over the interval $[1/L; 1/2]$ beginning from a sufficiently small $\alpha_0 > 0$ (or beginning from $\beta_0 = 1-\alpha_0$ close to 1). Let us change the variable x to $y=\frac{xL}{m_\alpha+xL}$ for $y\in[\frac{1}{m_\alpha+1},\frac{L}{L+2m_\alpha}]$. This leads to the following differentiable piecewise-defined function of class $\mathbb{C}^1$

$$\overline{\gamma}_2(y)=\left\{\Phi\left[\Phi^{-1}\left(\beta^y\right)-\sqrt{\frac{\varrho m_\alpha y}{1-y}}\right]\right\}^{g(y)}, \tag{79}$$

where $g(y)=\max\{\frac{L}{ym_\alpha}-\frac{L}{m_\alpha}-2,1\}$. Let us re-write (79) as follows

$$\overline{\gamma}_2(y)=f(y)^{g(y)}, \tag{80}$$

where $f(y)=\Phi(\Phi^{-1}(\beta^y)-\sqrt{\frac{\varrho m_\alpha y}{1-y}})$. The derivative of $\overline{\gamma}_2(y)$ is

$$\frac{d\,\overline{\gamma}_2(y)}{dy}=f(y)^{g(y)}\left[g'_y(y)\log f(y)+\frac{g(y)f'_y(y)}{f(y)}\right]. \tag{81}$$

First, we compute the derivative over the interval $[\frac{1}{m_\alpha+1},\frac{L}{L+3m_\alpha}]$. The function g over this interval is given by $g(y)=\frac{L}{ym_\alpha}-\frac{L}{m_\alpha}-2$. The derivatives of the functions f and g are given as follows

$$\begin{aligned}f'_y(y)&=\varphi\left(\Phi^{-1}\left(\beta^y\right)-\sqrt{\frac{\varrho m_\alpha y}{1-y}}\right)\\&\cdot\left[\frac{\beta^y\log\beta}{\varphi\left[\Phi^{-1}\left(\beta^y\right)\right]}-\frac{\sqrt{\varrho m_\alpha}}{2\sqrt{y(1-y)^3}}\right],\end{aligned} \tag{82}$$

where $\varphi(x)=(1/\sqrt{2\pi})\exp\{-x^2/2\}$ denotes the PDF of the standard normal distribution,

$$g'_y(y)=-\frac{L}{m_\alpha y^2}. \tag{83}$$

By the definition of the CDF $\Phi(x)$, the first factor $f(y)^{g(y)}$ in (81) is positive. Hence, the sign of the derivative $\frac{d\overline{\gamma}_2(y)}{dy}$ is defined by two terms inside brackets in (81). We are interested in the asymptotic analysis when $\alpha\to0^+$ or $\beta=1-\alpha\to1^-$. Let us define the two terms inside brackets in (81) as functions of β and y:

$$h_1(\beta;y)=g'_y(y)\log f(\beta;y)\text{ and }h_2(\beta;y)=\frac{g(y)f'_y(\beta;y)}{f(\beta;y)}. \tag{84}$$

Because $\log\Phi(x) < 0\;\forall x\in\text{IR}$, the first term inside brackets is positive $h_1(\beta; y) = g'(y)\log f(\beta; y) > 0$. After simple algebra, we get

$$h_1(\beta;y)\sim-\frac{L}{m_\alpha y^2}\log\beta^y\quad\text{as}\quad\beta\to1^- \tag{85}$$

for all y in the closed interval $[\frac{1}{m_\alpha+1}, \frac{L}{L+3m_\alpha}]$. The second term h_2 becomes negative as $\beta \to 1^-$:

$$h_2(\beta; y) \sim -\frac{\sqrt{\varrho m_\alpha}}{2\sqrt{y(1-y)^3}} \cdot \varphi\left(\Phi^{-1}\left(\beta^y\right) - \sqrt{\frac{\varrho m_\alpha y}{1-y}}\right) \cdot \left(\frac{L}{y m_\alpha} - \frac{L}{m_\alpha} - 2\right) \quad (86)$$

for all y in the closed interval $[\frac{1}{m_\alpha+1}, \frac{L}{L+3m_\alpha}]$. Finally, let us establish that $h_1(\beta; y) = o(h_2(\beta; y))$ as $\beta \to 1^-$. By using the L'Hôpital's rule, we get

$$\lim_{\beta \to 1^-} \frac{h_1(\beta; y)}{h_2(\beta; y)} = \lim_{\beta \to 1^-} \frac{h'_{\beta,1}(\beta; y)}{h'_{\beta,2}(\beta; y)} = 0. \quad (87)$$

for all y in the closed interval $[\frac{1}{m_\alpha+1}, \frac{L}{L+3m_\alpha}]$.

Second, we compute the derivative over the interval $[\frac{L}{L+3m_\alpha}, \frac{L}{L+2m_\alpha}]$. The function $y \mapsto g(y)$ over this interval is equal to 1. Hence,

$$\frac{d\,\overline{\gamma}_2(y)}{dy} \sim -\frac{\sqrt{\varrho m_\alpha}}{2\sqrt{y(1-y)^3}} \varphi\left(\Phi^{-1}\left(\beta^y\right) - \sqrt{\frac{\varrho m_\alpha y}{1-y}}\right) \quad (88)$$

as $\beta \to 1^-$ for all y in the closed interval $[\frac{1}{m_\alpha+1}, \frac{L}{L+3m_\alpha}]$. Hence, $\exists \alpha_0 > 0$ such that for all $\alpha \leq \alpha_0$ the function $x \mapsto \overline{\gamma}_2(x)$ is decreasing over the interval $[1/L; 1/2]$.

Part 3: Let us show that the upper bound $N \mapsto \overline{\gamma}_1(\nu^*, N)$ is an increasing function in the interval $[n+1, 2n]$, or (it is equivalent to) that the function $k \mapsto \overline{\gamma}_1(k)$, where $\gamma_1(k)$ is given by (74), is an increasing function in the interval $[1, n]$. By using the inequality (76), we get an upper bound for $\gamma_1(x)$:

$$\gamma_1(x) < \overline{\gamma}_1(x) = \Phi\left[\Phi^{-1}\left(\beta^{\frac{n+x}{m_\alpha+n+x}}\right) - \frac{(n-x+1)\sqrt{\varrho}}{\sqrt{n+x}}\right], \quad (89)$$

where $x \in [1, n]$, $x \in \mathrm{I\!R}$. The derivative of $\overline{\gamma}_1(x)$ is

$$\frac{d\,\overline{\gamma}_1(x)}{dx} = \varphi\left(\Phi^{-1}\left(\beta^{\frac{n+x}{m_\alpha+n+x}}\right) - \frac{(n-x+1)\sqrt{\varrho}}{\sqrt{n+x}}\right) \cdot \left[\frac{m_\alpha \beta^{\frac{n+x}{m_\alpha+n+x}} \log\beta}{\varphi\left[\Phi^{-1}\left(\beta^{\frac{n+x}{m_\alpha+n+x}}\right)\right](m_\alpha+n+x)^2} + \frac{(3n+x+1)\sqrt{\varrho}}{2(n+x)^{3/2}}\right]. \quad (90)$$

By the definition of the PDF $\varphi(x)$, the first factor in (90) is positive. Hence, the sign of the derivative $\frac{d\overline{\gamma}_1(x)}{dx}$ is defined by two terms inside brackets in (90). We are interested in the asymptotic analysis when $\beta = 1 - \alpha \to 1^-$. The first term inside brackets in (90) is a function of β and x. Exactly in the same way as previously in *Part 2* (see (85)), it can be shown that this term tends to 0^- as $\beta \to 1^-$ for any $x \in [n+1, 2n]$. The second term is not a function of β and it is positive for any $x \in [1, n]$. Hence, $\exists \alpha_0 > 0$ such that for all $\alpha \leq \alpha_0$ the function $x \mapsto \overline{\gamma}_1(x)$ is increasing in the interval $[1, n]$.

Finally, to conclude with the case of $L = 2n$, it is necessary to "glue" both functions $N \mapsto \overline{\gamma}_1(N)$ and $N \mapsto \overline{\gamma}_2(N)$ and to show that

$$\overline{\gamma}_2(n) = \Phi\left[\Phi^{-1}\left(\beta^{\frac{n}{m_\alpha+n}}\right) - \sqrt{\varrho n}\right] \lesssim \overline{\gamma}_1(1) = \Phi\left[\Phi^{-1}\left(\beta^{\frac{n+1}{m_\alpha+n+1}}\right) - \frac{\sqrt{n}}{\sqrt{n+1}}\sqrt{\varrho n}\right] \quad (91)$$

as $\beta =\to 1^-$. To prove inequality (91), let us recall that $x \mapsto \Phi(x)$ is an increasing function. Due to the fact that $\beta^{\frac{n}{m_\alpha+n}} \sim \beta^{\frac{n+1}{m_\alpha+n+1}}$ as $\beta =\to 1^-$, the first terms inside brackets in (91) are equivalent. The second terms inside brackets in (91) are not functions of β, hence these terms define the relation between $\overline{\gamma}_2(n)$ and $\overline{\gamma}_1(1)$. The validity of inequality (91) follows from the inequality $\frac{\sqrt{n}}{\sqrt{n+1}} < 1$. Finally, it can be deduced from (91) that an upper bound for $\overline{\gamma}(N)$ (73), composed of two parts, asymptotically attains its minimum at the right end point $N = n$ of the interval $[1, n]$. This proves the asymptotically optimal solution given by equations (27) – (28) in the case of $L = 2n$.

Part 4: Let the required time-to-alert L be an odd number, i.e. $L = 2n+1$, $n \in \mathbb{Z}^+$ and $T_{\mathrm{FSS}} \in C_\alpha$. The proof is organized in the same way as for $L = 2n$ (see *Parts 1 – 3*), the only difference is how to "glue" the upper bounds $N \mapsto \overline{\gamma}_2(N)$ and $N \mapsto \overline{\gamma}_1(N)$ defined in the subintervals: $[1, n]$ and $[n+1, 2n+1]$, respectively. The bound $N \mapsto \overline{\gamma}_2(N)$ is defined as previously. Let us consider the bound $x \mapsto \overline{\gamma}_1(x)$ in the interval $[1, n+1]$. Now, $\overline{\gamma}_1(x)$ is slightly different from (89), $(n-x+1)$ in the second term between the brackets is replaced with $(n-x+2)$:

$$\overline{\gamma}_1(x) = \Phi\left[\Phi^{-1}\left(\beta^{\frac{n+x}{m_\alpha+n+x}}\right) - \frac{(n-x+2)\sqrt{\varrho}}{\sqrt{n+x}}\right], \quad (92)$$

where $x \in [1, n+1]$, $x \in \mathrm{I\!R}$. The proof of *Part 3* is completely applicable to the case of $L = 2n+1$. Hence, the function $x \mapsto \overline{\gamma}_1(x)$ is also increasing in the interval $[1, n+1]$. Nevertheless, in contrast to inequality (91), we get from (88) that

$$\overline{\gamma}_2(n) = \Phi\left[\Phi^{-1}\left(\beta^{\frac{n}{m_\alpha+n}}\right) - \sqrt{\varrho n}\right] \gtrsim \overline{\gamma}_1(1) = \Phi\left[\Phi^{-1}\left(\beta^{\frac{n+1}{m_\alpha+n+1}}\right) - \sqrt{\varrho(n+1)}\right] \quad (93)$$

as $\beta \to 1^-$. It can be deduced from (93) that the upper bound $\overline{\gamma}(N)$ (73), composed of two parts, asymptotically attains its minimum at the left end point $N = n+1$ of the interval $[n+1, 2n+1]$. This proves the asymptotically optimal solution given by equations (27)–(28) in the case of $L = 2n+1$.

Acknowledgment

The authors would like to thank the associate editor and anonymous reviewers for their comments which improved the paper.

References

[1] H. Cramér, *Mathematical Methods of Statistics*. Princeton, NJ, USA: Princeton Univ. Press, 1946.

[2] E. L. Lehmann, *Testing Statistical Hypotheses*. New York, NY, USA: Wiley, 1968.

[3] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

[4] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. Berlin, Germany: Springer, 1994.

[5] B. C. Levy, *Principles of Signal Detection and Parameter Estimation*. Berlin, Germany: Springer, 2008.

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

[7] H. V. Poor and O. Hadjiliadis, *Quickest Detection*. New York, NY, USA: Cambridge Univ. Press, 2009.

[8] A. G. Tartakovsky, I. V. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Changepoint Detection* (Monographs on Statistics and Applied Probability 136). Boca Raton, FL, USA: CRC Press, 2015.

[9] G. Lorden, "Procedures for reacting to a change in distribution," *Ann. Math. Statist.*, vol. 42, no. 6, pp. 1897–1908, 1971.

[10] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. 100–114, Jun. 1954.

[11] G. V. Moustakides, "Optimal stopping times for detecting changes in distributions," *Ann. Statist.*, vol. 14, no. 4, pp. 1379–1387, 1986.

[12] B. Broder and C. Schwartz, "Quickest detection procedures and transient signal detection," Office Naval Res., Arlington, VA, USA, Tech. Rep. 21, 1990.

[13] B. Bakhache and I. Nikiforov, "Reliable detection of faults in measurement systems," *Int. J. Adaptive Control Signal Process.*, vol. 14, no. 7, pp. 683–700, Nov. 2000.

[14] V. L. Do, L. Fillatre, I. Nikiforov, and P. Willett, "Security of SCADA systems against cyber-physical attacks," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 32, no. 5, pp. 28–45, May 2017.

[15] A. Tartakovsky, A. Berenkov, N. R. Kolessa, and I. Nikiforov, "Optimal sequential detection of signals with unknown appearance and disappearance points in time," *IEEE Trans. Signal Process.*, vol. 69, no. 4, pp. 2653–2662, 2021, doi: 10.1109/TSP.2021.3071016.

[16] D. Egea-Roca, G. Seco-Granados, J. A. Lopez-Salcedo, and H. V. Poor, "A finite moving average test for transient change detection in GNSS signal strength monitoring," in *Proc. IEEE Stat. Signal Process. Workshop*, 2016, pp. 1–5.

[17] D. Egea-Roca, G. Seco-Granados, and J. A. López-Salcedo, "Sequential change detection for next-generation RAIM algorithms," in *Proc. 30th Int. Tech. Meeting Satell. Division Inst. Navigation*, 2017, pp. 2417–2427.

[18] B. K. Guépié, L. Fillatre, and I. Nikiforov, "Sequential detection of transient changes," *Sequential Anal.*, vol. 31, no. 4, pp. 528–547, 2012.

[19] I. V. Nikiforov *et al.*, "Sequential detection of a total instantaneous blockage occurred in a single subassembly of a sodium-cooled fast reactor," *Nucl. Eng. Des.*, vol. 366, Sep. 2020, Art. no. 110733.

[20] B. K. Guépié, E. Grall-Maës, P. Beauseroy, I. Nikiforov, and F. Michel, "Reliable leak detection in a heat exchanger of a sodium-cooled fast reactor," *Ann. Nucl. Energy*, vol. 142, 2020, Art. no. 107357.

[21] Z. Wang and P. Willett, "A performance study of some transient detectors," *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2682–2685, Sep. 2000.

[22] Z. J. Wang and P. Willett, "A variable threshold page procedure for detection of transient signals," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4397–4402, Nov. 2005.

[23] Z. J. Wang and P. Willett, "Detecting transients of unknown length," in *Proc. IEEE Aerosp. Conf.*, 2005, pp. 2236–2247.

[24] M. Pollak and A. M. Krieger, "Shewhart revisited," *Sequential Anal.*, vol. 32, no. 2, pp. 230–242, Jun. 2013.

[25] G. V. Moustakides, "Multiple optimality properties of the Shewhart test," *Sequential Anal.*, vol. 33, pp. 318–344, 2014.

[26] B. K. Gu, L. Fillatre, and I. Nikiforov, "Detecting a suddenly arriving dynamic profile of finite duration," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 3039–3052, May 2017.

[27] D. Egea-Roca, J. A. López-Salcedo, G. Seco-Granados, and H. V. Poor, "Performance bounds for finite moving average tests in transient change detection," *IEEE Trans. Signal Process.*, vol. 66, no. 6, pp. 1594–1606, Jan. 2018.

[28] J. Noonan and A. Zhigljavsky, "Power of the MOSUM test for online detection of a transient change in mean," *Sequential Anal.*, vol. 39, no. 2, pp. 269–293, 2020.

[29] L. Fillatre and I. Nikiforov, "A fixed-size sample strategy for the sequential detection and isolation of non-orthogonal alternatives," *Sequential Anal.*, vol. 29, no. 2, pp. 176–192, Apr. 2010.

[30] A. Wald, *Sequential Analysis*. New York, NY, USA: Wiley, 1947.

[31] B. K. Ghosh, Ed., *Sequential Tests of Statistical Hypotheses*. Reading, MA, USA: Addison-Wesley, 1970.

[32] A. N. Shiryaev, "The problem of the most rapid detection of a disturbance in a stationary process," *Soviet Math. Dokl.*, vol. 2, pp. 795–799, 1961.

[33] A. Shiryaev, "On the detection of disorder in a manufacturing process," *Theory Probability Appl.*, vol. 8, no. 3, pp. 247–265, 1963.

[34] A. N. Shiryaev, "On optimum methods in quickest detection problems," *Theory Probability Appl.*, vol. 8, no. 1, pp. 22–46, 1963.

[35] L. Pelkowitz and S. C. Schwartz, "Asymptotically optimum sample size for quickest detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, no. 2, pp. 263–272, Mar. 1987.

[36] I. V. Nikiforov, "Two strategies in the problem of change detection and isolation," *IEEE Trans. Inf. Theory*, vol. 43, no. 2, pp. 770–776, Mar. 1997.

[37] I. Nikiforov, "Quadratic tests for detection of abrupt changes in multivariate signals," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2534–2538, Sep. 1999.

[38] T. L. Lai, "Sequential multiple hypothesis testing and efficient fault detection-isolation in stochastic systems," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 595–608, Mar. 2000.

[39] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2917–2929, Nov. 1998.

[40] A. Mathai and S. Provost, *Quadratic Forms in Random Variables: Theory and Applications*, Statistics: Textbooks and Monographs. New York, NY, USA: Marcel Dekker, 1992.

[41] S. Gabler and C. Wolff, "A quick and easy approximation to the distribution of a sum of weighted chi-square variables," *Statistische Hefte*, vol. 28, pp. 317–325, 1987.

[42] S. M. Ross, *Introduction to Probability Models*, 10th ed. San Diego, CA, USA: Academic, 2011.

[43] RTCA/DO-229D, *Minimum Operational Performance Standards for Global Positioning Systems/Wide Area Augmentation System Airborne Equipment*, RTCA, Inc., Washington, D.C., USA, 2006.

[44] A. G. Tartakovsky, "Efficiency of the generalized Neyman–Pearson test for detecting changes in a multichannel system," *Problems Inf. Transmiss.*, vol. 28, pp. 341–350, 1992.