ON SEQUENTIAL RANDOM DISTORTION TESTING OF NON-STATIONARY PROCESSES

Prashant Khanduri[†] *Dominique Pastor*[‡] *Vinod Sharma*[*] *Pramod K. Varshney*[†]

[†] Electrical Engineering and Computer Science department, Syracuse University, Syracuse, NY, USA
[‡] IMT Atlantique, Lab-STICC, Univ. Bretagne Loire, Brest, France
[*] Electrical Communication Engineering department, IISc, Bangalore, India

ABSTRACT

Random distortion testing (RDT) addresses the problem of testing whether or not a random signal, Ξ, deviates by more than a specified tolerance, τ, from a fixed value, ξ_0 [1]. The test is non-parametric in the sense that the distribution of the signal under each hypothesis is assumed to be unknown. The signal is observed in independent and identically distributed (i.i.d) additive noise. The need to control the probabilities of false alarm and missed detection while reducing the number of samples required to make a decision leads to the *Seq*RDT approach. We show that under mild assumptions on the signal, *Seq*RDT will follow the properties desired by a sequential test. Simulations show that the *Seq*RDT approach leads to faster decision making compared to its fixed sample counterpart Block-RDT [2] and is robust to model mismatches compared to the Sequential Probability Ratio Test (SPRT) [3] when the actual signal is a distorted version of the assumed signal especially at low Signal-to-Noise Ratios (SNRs).

Index Terms— Sequential tests, non-parametric tests, random distortion testing, non-stationary signals

1. INTRODUCTION

In standard binary hypothesis testing problems, on the basis of a fixed number of observations, a decision is made between two possible statistical hypothesis, the so-called null ($\mathcal{H}_0$) and alternative ($\mathcal{H}_1$) hypotheses. The decision is generally made under the Bayesian, minimax or Neyman-Pearson frameworks. In his seminal works [3, 4], Wald moved from standard likelihood theory with a fixed sample size to sequential procedures where observations are collected and processed one after another, until a decision can be made with specified confidence. Basically, at any stage of a sequential procedure, the same decision rule is applied. This rule has three possible outcomes, instead of two: it may either 1) accept $\mathcal{H}_0$ and stop the testing, or 2) accept $\mathcal{H}_1$ and stop the testing or 3) make no decision and acquire a new observation. These three steps are repeated sequentially until a decision is reached, in which case the testing stops. In sequential testing, the sample size and the time instant when the decision is made are random. The issue is then to devise a decision rule that optimizes a certain criterion *"to achieve a tradeoff between the average observation time and the quality of the decision. ...It has been shown that the sequential procedure performs significantly better than the classical Neyman-Pearson test in the case of two simple hypotheses."* [5]. We recall that simple hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ correspond to two possible distributions for the observations. For details on Wald's approach the reader can refer to [5].

This work was supported by ARO grant W911NF-14-1-0339.

Standard sequential testing is an extension of likelihood theory in that it assumes prior knowledge regarding the distributions of the observations under each hypothesis to derive the likelihood ratio, perhaps up to a vector parameter in case of nuisance parameters. This procedure has the following limitations. In practice, prior knowledge or good models for the distributions under each hypothesis are usually not available. This is all the more detrimental when likelihood ratio tests are not robust to uncertainty or model mismatch. Moreover, many approaches in sequential testing make stationarity or independent and identically distributed (iid) assumptions on the observed process under each hypothesis [5]. Such assumptions are questionable in practice. In addition, proposed solutions that relax stationarity or iid assumptions are still based on likelihood ratio tests and suffer from the same drawbacks.

In many practical applications such as radar, sonar and communication systems signals of interest, distorted by the environment, are acquired in noise and are cluttered by interfering echoes. The observed random process resulting from this mixture — not necessarily additive — of signal, distortions and interferences, can hardly be modeled as a stationary random process with known distribution. Therefore, the observation process here is modeled as the sum of a non-stationary signal with unknown distribution and independent noise. We introduce a theoretical framework suited for statistical signal processing applications such as those considered in [6–8], where the issue is to sequentially test the empirical mean of a non-stationary random signal that has non-iid samples and unknown sample distributions in additive and independent Gaussian noise. In contrast to the preliminary approach in [6, 7], the theory presented below introduces a sequential procedure that guarantees an almost surely finite stopping time and error probabilities that can be rendered arbitrarily small. In particular, the analysis conducted in the paper exhibits nested models and assumptions that help predict the behavior of sequential testing without prior knowledge of the distribution of the signal and without any stationarity or iid assumption. Because of space limitations, detailed proofs of the results stated below are postponed to the longer version of the paper.

2. NOTATIONS

All the random variables encountered below are defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. As usual, for any given $\xi \in \mathbb{R}$ and any $\sigma \in [0, \infty)$, $Z \sim \mathcal{N}(\xi, \sigma^2)$ means that Z is Gaussian distributed with mean ξ and variance σ^2. In what follows, $Q_{1/2}$ denotes the Generalized Marcum Function [9] with order 1/2. Basically, we have $\mathbb{P}\big[|Z| > \eta\big] = Q_{1/2}(|\xi|, \eta)$ for any given $Z \sim \mathcal{N}(\xi, 1)$. Given $\gamma \in (0, 1)$ and $\rho \in [0, \infty)$, $\lambda_\gamma(\rho)$ is defined as the unique solution

in x to $Q_{1/2}(\rho,x)=\gamma$ [1, Lemma 2, statement (ii)], so that:

$$Q_{1/2}(\rho,\lambda_\gamma(\rho))=\gamma. \tag{1}$$

The set of all real random variables defined on $(\Omega,\mathcal{F})$ is denoted by $\mathcal{M}(\Omega,\mathbb{R})$. Accordingly, $\mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}}$ (resp. $\mathcal{M}(\Omega,\mathbb{R})^{[\![1,N]\!]}$) is the set of all sequences or random processes defined on $\mathbb{N}$ (resp. $[\![1,N]\!]=\{1,2,\dots,N\}$). Given U in $\mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}}$, the realization by U of $n\in\mathbb{N}$ (resp. $n\in[\![1,N]\!]$) is called a sample of U and denoted by U_n. Each U_n is an element of $\mathcal{M}(\Omega,\mathbb{R})$. The empirical mean of $U\in\mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}}$ is the real random variable defined by:

$$\langle U\rangle_N=\frac{1}{N}\sum_{n=1}^{N}U_n.$$

Two elements U, V of $\mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}}$ are said to be independent if U_n and V_n are independent for each $n\in\mathbb{N}$.

3. MATHEMATICAL FORMULATION AND ANALYSIS

Let $\Xi=(\Xi_n)_{n\in\mathbb{N}}$ be an element of $\mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}}$. This random process models the random mixture of a distorted signal of interest and possible interferences. In accordance with the introduction, no assumption is made on the stationarity or the distribution of $\Xi=(\Xi_n)_{n\in\mathbb{N}}$. In this respect, the samples Ξ_n are not necessarily iid. We suppose that Ξ is observed in additive and independent noise $X=(X_n)_{n\in\mathbb{N}}$. The observation process is then $Y=(Y_n)_{n\in\mathbb{N}}$ such that $Y_n=\Xi_n+X_n$ for any $n\in\mathbb{N}$ and we write $Y=\Xi+X$. In our formulation, Ξ models distortion around a fixed model ξ_0, in the absence of any distortion, $\Xi=\xi_0$, but due to distortions this will not be true. We expect the empirical mean $\langle\Xi\rangle_N$ to remain close to ξ_0 under $\mathcal{H}_0$ and drift significantly away from ξ_0 under $\mathcal{H}_1$ for all $N\geqslant N_0$. The role of τ is to distinguish small distortions from large ones. The *Seq*RDT formulation is then given as

$$\begin{cases}\underline{\text{Observation}}: Y=\Xi+X\in\mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}}\\ \text{with}\begin{cases}\Xi\in\mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}},\\ X_1,X_2,\dots\overset{\text{iid}}{\sim}\mathbb{F},\\ \mathbb{E}[X_1]=0,\mathrm{Var}(X_1)=1 \text{ and } \mathbb{E}\left[X_1^3\right]<\infty\\ \Xi \text{ and } X \text{ are independent.}\end{cases}\\ \begin{cases}\mathcal{H}_0: & \Xi=(\Xi_n)_{n\in\mathbb{N}},\ \forall N\geqslant N_0,\ |\langle\Xi\rangle_N-\xi_0|\leqslant\tau \text{ (a-s)}\\ \mathcal{H}_1: & \Xi=(\Xi_n)_{n\in\mathbb{N}},\ \forall N\geqslant N_0,\ |\langle\Xi\rangle_N-\xi_0|>\tau \text{ (a-s)}\end{cases}\end{cases} \tag{2}$$

where, $\tau\in[0,\infty)$ is the tolerance and $\mathbb{F}$ is an unknown cdf. Note, that the above testing model is the same as the Block-RDT model [2] if the sample size, N, is fixed. Here, N_0 is known *a priori* based on some prior knowledge about the signal.

To solve the mean testing problem (2), we introduce the following assumptions, which can be regarded as weak notions of ergodicity.

Assumption 3.1 (Asymptotic Convergence of $|\langle\Xi\rangle_N-\xi_0|$) *We assume that*

$$\exists\langle\Xi\rangle_\infty\in\mathcal{M}(\Omega,\mathbb{R}),\lim_N\langle\Xi\rangle_N=\langle\Xi\rangle_\infty \text{ (a-s).}$$

we further assume that as N grows:

$$\begin{aligned}&\textit{Under}\,\mathcal{H}_0: && \lim_N|\langle\Xi\rangle_N-\xi_0|=|\langle\Xi\rangle_\infty-\xi_0|<\tau \text{ (a-s),}\\ &\textit{Under}\,\mathcal{H}_1: && \lim_N|\langle\Xi\rangle_N-\xi_0|=|\langle\Xi\rangle_\infty-\xi_0|>\tau \text{ (a-s).}\end{aligned}$$

Remark 1 *The first assumption in Assumption 3.1 is automatically satisfied if Ξ is stationary and ergodic. Indeed, in this case, there exists $\xi\in\mathbb{R}$ such that $\mathbb{E}[\Xi_n]=\xi$ for every $n\in\mathbb{N}$, so that the assumption holds true with $\langle\Xi\rangle_\infty=\xi$.* ■

Given $\gamma\in(0,1)$ and $\tau\geqslant 0$, define $\mathcal{T}_{\lambda_\gamma(\tau\sqrt{N})/\sqrt{N}}:\mathbb{R}^{\mathbb{N}}\to\{0,1\}$ for any sequence $x=(x_n)_{n\in\mathbb{N}}\in\mathbb{R}^{\mathbb{N}}$ by:

$$\mathcal{T}_{\lambda_\gamma(\tau\sqrt{N})/\sqrt{N}}(x)=\begin{cases}0 & \text{if } |\langle x\rangle_N-\xi_0|\leqslant\lambda_\gamma(\tau\sqrt{N})/\sqrt{N}\\ 1 & \text{otherwise.}\end{cases} \tag{3}$$

Next, in Proposition 3.2, we analyze the behavior of the type of tests given by (3) as the number of samples N increases.

Proposition 3.2 *For any $\gamma\in(0,1)$ and $\tau\geqslant 0$, $\mathcal{T}_{\lambda_\gamma(\tau\sqrt{N})/\sqrt{N}}$ satisfies the following asymptotic behavior for testing $\mathcal{H}_0$ vs $\mathcal{H}_1$ in (2):*

(i) *we have*

$$\begin{aligned}&\limsup_N\mathbb{P}\left[\mathcal{T}_{\lambda_\gamma(\tau\sqrt{N})/\sqrt{N}}(Y)=1\right]\leqslant\gamma && \textit{under}\,\mathcal{H}_0\\ &\limsup_N\mathbb{P}\left[\mathcal{T}_{\lambda_\gamma(\tau\sqrt{N})/\sqrt{N}}(Y)=0\right]\leqslant 1-\gamma && \textit{under}\,\mathcal{H}_1\end{aligned} \tag{4}$$

(ii) *under Assumption 3.1 we have,*

$$\lim_N\mathbb{P}\left[\mathcal{T}_{\lambda_\gamma(\tau\sqrt{N})/\sqrt{N}}(Y)=1\right]=\begin{cases}0 & \textit{under}\,\mathcal{H}_0\\ 1 & \textit{under}\,\mathcal{H}_1\end{cases} \tag{5}$$

The asymptotic result of Proposition 3.2 enhances the interest of the tests of type as defined in (3). In the above, Proposition 3.2 (i) suggests the use of two thresholds as the probabilities of false alarm and missed detection both cannot be controlled with the use of a single threshold designed for a fixed γ. One level must be small to diminish the probability of false alarm and the second level must be close to 1 so as to make the probability of missed detection small. Such a strategy will naturally lead to a sequential approach. Proposition 3.2 (ii) highlights the importance of Assumption 3.1 in achieving arbitrarily low probabilities of false alarm and missed detection for a large sample size. But we need to control the number of samples, so to this end we resort to a sequential approach. We show that with the thresholds chosen according to (3), one can design a sequential test which can reduce the decision making time while guaranteeing certain levels of performance.

Next, we propose the *Seq*RDT approach for the mean testing problem defined in (2).

4. SEQUENTIAL TEST

Given any natural number $M\geqslant N_0-1$, we can completely specify a sequential test for $\mathcal{H}_0$ against $\mathcal{H}_1$ by defining the stopping time:

$$T=\min\{N\in\mathbb{N}:\mathcal{D}_M(N)\neq\infty\}, \tag{6}$$

$$\text{with:}\begin{cases}\mathcal{D}_M(1)=\mathcal{D}_M(2)=\dots=\mathcal{D}_M(M)=\infty,\\ \text{for } N>M,\ \mathcal{D}_M(N)=\begin{cases}0 & \text{if } |\langle Y\rangle_N-\xi_0|\leqslant\lambda_L(N),\\ \infty & \text{if } \lambda_L(N)<|\langle Y\rangle_N-\xi_0|\leqslant\lambda_H(N),\\ 1 & \text{if } |\langle Y\rangle_N-\xi_0|>\lambda_H(N).\end{cases}\end{cases} \tag{7}$$

where the two thresholds $\lambda_L(N)$ and $\lambda_H(N)$ must be such that $\lambda_L(N)\leqslant\lambda_H(N)$. Note that M is the number of samples *Seq*RDT waits for before starting the test. M can be chosen based on some elementary knowledge of the signal and noise.

We now define the False Alarm Probability of the proposed sequential test as:

$$\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M) \overset{\text{def}}{=} \mathbb{P}\left[\mathcal{D}_M(T)=1\right] \quad \text{under } \mathcal{H}_0$$

In the same way, the Missed Detection Probability is defined as:

$$\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M) \overset{\text{def}}{=} \mathbb{P}\left[\mathcal{D}_M(T)=0\right] \quad \text{under } \mathcal{H}_1$$

The goal of any sequential algorithm is to design the thresholds so as to guarantee that $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ stay below certain pre-specified levels α and β, respectively.

Next, in Proposition 4.1 we show that the thresholds $\lambda_H(N)$ and $\lambda_L(N)$ designed according to (3) for levels α and β, respectively are indeed appropriate for *Seq*RDT.

Proposition 4.1 *For $\alpha,\beta \in (0,1/2)$, $\tau \in (0,\infty)$ and the thresholds $\lambda_H(N) = \lambda_\alpha(\tau\sqrt{N})/\sqrt{N}$ and $\lambda_L(N) = \lambda_{1-\beta}(\tau\sqrt{N})/\sqrt{N}$ we have,*

$$\lambda_L(N) \leqslant \lambda_H(N),$$

for all $N \in \mathbb{N}$.

We thus know that the thresholds $\lambda_H(N)$ and $\lambda_L(N)$ satisfy the criterion $\lambda_L(N) \leqslant \lambda_H(N)$. The question that arises is then *"Can this choice of thresholds give some performance guarantees?"*. To answer this, we next present the main Theorem of the paper.

Theorem 4.2 *Given $\alpha,\beta \in \left(0,\frac{1}{2}\right)$, set $\lambda_L(N) = \lambda_{1-\beta}(\tau\sqrt{N})/\sqrt{N}$ and $\lambda_H(N) = \lambda_\alpha(\tau\sqrt{N})/\sqrt{N}$. If Assumption 3.1 is satisfied, then:*

(i) $\mathbb{P}\left[T=\infty\right] = 0$ *under $\mathcal{H}_0$ and $\mathcal{H}_1$,*

(ii) $\lim\limits_{M\to\infty} \mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M) = \lim\limits_{M\to\infty} \mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M) = 0$

The above theorem shows that the thresholds $\lambda_H(N)$ and $\lambda_L(N)$ will guarantee that the stopping time T is finite with probability 1. Moreover, these thresholds guarantee that *Seq*RDT can achieve arbitrarily low $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ provided one has the freedom to choose sufficiently large M. Therefore, all the above results indicate that one can use the test defined in (7) for testing the mean of a non-stationary process with the choice of thresholds as given in Proposition 4.1 and Theorem 4.2.

Next, we perform some simulations to analyze the performance of the *Seq*RDT.

5. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we perform some simulations to highlight the advantages of *Seq*RDT compared to Block-RDT and SPRT as proposed in [2] and [3, 4], respectively. We first present the detection problem, then we outline each algorithm and finally, carry out the comparison of the presented algorithms.

5.1. Detection with signal distortions

We consider the detection of change in mean in Gaussian noise with some model mismatch. Let us first consider the model $Y_n = \Xi_n + X_n$, for $n \in \mathbb{N}$, with the signal $\Xi_n = \xi_0$ under $\mathcal{H}_0$ and $\Xi_n = \xi_1$ under $\mathcal{H}_1$, with ξ_0 and ξ_1 as deterministic constants. The noise is assumed to be Gaussian, i.e., $X_n \sim \mathcal{N}(0,1)$ for any $n \in \mathbb{N}$. This model can be formulated in the *Seq*RDT framework as defined in (2) with $\tau = 0$ and $N_0 = 1$ as

$$\begin{cases} \underline{\text{Observation}}: Y = \xi_i + X \in \mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}}, \text{ for } i = \{0,1\} \\ \text{with} \quad X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathcal{N}(0,1). \\ \begin{cases} \mathcal{H}_0: & \Xi = (\xi_0)_{n\in\mathbb{N}},\ \forall N \geqslant 1,\ |\langle \Xi\rangle_N - \xi_0| \leqslant 0 \text{ (a-s)} \\ \mathcal{H}_1: & \Xi = (\xi_1)_{n\in\mathbb{N}},\ \forall N \geqslant 1,\ |\langle \Xi\rangle_N - \xi_0| > 0 \text{ (a-s)}, \end{cases} \end{cases} \tag{8}$$

which is the classical Gaussian mean detection problem. But, in many practical systems there might unfortunately be a mismatch between the model and the actual signal observed in practice. In reality, the actual signal would not be a constant ξ_0 or ξ_1, under either hypothesis, the signal will be a perturbed version of the actual signal and these perturbations are hard to model in a parametric setup. Therefore, likelihood ratio based tests will fail to guarantee reliable performance if there are model mismatches. However, the Block-RDT setup as proposed in [2] and the *Seq*RDT setup as proposed in (2) and (7) are not limited by these drawbacks.

Similar to [2], instead of dealing with the perfect and somewhat unrealistic model as described above, we consider the case when the signal is $\Xi_n = \xi_i + \Delta_n$ under $\mathcal{H}_i$ for $i \in \{0,1\}$ and for all $n \in \mathbb{N}$, with $|\xi_1 - \xi_0| > \tau$. Here Δ_ns model the possible additive distortions in the above deterministic model with unknown distribution. Let us assume for all $N \geqslant N_0$ we have some positive value τ such that $\mathbb{P}[|\langle\Delta\rangle_N| \leqslant \tau] = 1$ and $\mathbb{P}[|\langle\Delta\rangle_N + \xi_1 - \xi_0| > \tau] = 1$ are satisfied. Next, we show that even if these probabilities are not strictly satisfied the algorithm will still be able to provide sufficient performance guarantees.

5.2. *Seq*RDT, Block-RDT and SPRT

For illustration, let us consider the distortions $\Delta_n \sim \mathcal{N}(0,\sigma_\Delta^2)$ for all $n \in \mathbb{N}$. Consider the *Seq*RDT framework. For the choice of tolerance $\tau = 2\sigma_\Delta$, we have $\mathbb{P}[|\langle\Delta\rangle_N| \leqslant \tau] \geqslant 0.9545$ for all $N \geqslant 1$, and equality for $N = 1$, i.e., $\mathbb{P}[|\langle\Delta\rangle_1| \leqslant \tau] = 0.9545$. Moreover, we have $\mathbb{P}[|\langle\Delta\rangle_N + \xi_1 - \xi_0| > \tau] > 0.5$ for all $N \geqslant 1$. This probability will be increasing in $|\xi_1 - \xi_0|$. For example, if $|\xi_1 - \xi_0| = 4\sigma_\Delta$, we have $\mathbb{P}[|\langle\Delta\rangle_N + \xi_1 - \xi_0| > \tau] \approx 0.9772$. Here, we choose the buffer size $M = N_0 - 1 = 0$. Although we do not have $\mathbb{P}[|\langle\Delta\rangle_N| \leqslant \tau] = 1$ and $\mathbb{P}[|\langle\Delta\rangle_N + \xi_1 - \xi_0| > \tau] = 1$ satisfied strictly for $N_0 = 1$, we show via simulations that this does not impact the results significantly. Therefore, the hypothesis test for this system with model mismatch using (2) and (7) can be written as

$$\begin{cases} \underline{\text{Observation}}: Y = \xi_i + \Delta + X \in \mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}}, \text{ for } i = \{0,1\} \\ \text{with} \quad \begin{cases} \Delta_1, \Delta_2, \ldots \overset{\text{iid}}{\sim} \mathcal{N}(0,\sigma_\Delta^2) \\ X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathcal{N}(0,1). \end{cases} \\ \begin{cases} \mathcal{H}_0: & \Xi = (\xi_0 + \Delta)_{n\in\mathbb{N}},\ \forall N \geqslant 1,\ |\langle \Xi\rangle_N - \xi_0| \leqslant 2\sigma_\Delta \text{ (a-s)} \\ \mathcal{H}_1: & \Xi = (\xi_1 + \Delta)_{n\in\mathbb{N}},\ \forall N \geqslant 1,\ |\langle \Xi\rangle_N - \xi_0| > 2\sigma_\Delta \text{ (a-s)}, \end{cases} \end{cases}$$

Now, consider the detection of the same distorted signal with the Block-RDT approach [2]. In this framework, the problem of testing the mean of a random process observed over N samples can be summarized as:

$$\begin{cases} \underline{\text{Observation:}}\ Y = \Xi + X \in \mathcal{M}(\Omega,\mathbb{R})^{[\![1,N]\!]} \\ \text{with} \begin{cases} \Xi \in \mathcal{M}(\Omega,\mathbb{R})^{[\![1,N]\!]}, X_1, X_2, \ldots, X_N \overset{\text{iid}}{\sim} \mathcal{N}(0,1), \\ \Xi_n \text{ and } X_n \text{ independent for each } n \in [\![1,N]\!], \end{cases} \\ \mathcal{H}_0: |\langle \Xi\rangle_N - \xi_0| \leqslant \tau, \\ \mathcal{H}_1: |\langle \Xi\rangle_N - \xi_0| > \tau. \end{cases} \tag{9}$$

Before proceeding further let us first analyze the invariance properties of the above problem for the Block-RDT framework. The

$\alpha = \beta = 0.001$					
SDR $\left(=\frac{\text{SNR}}{\tau}\right)$ (dB)		6.02	7.96	9.54	12.00
*Seq*RDT	Stopping time, T	3.98	3.28	3.04	2.90
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$	3.33×10^{-4}	3.47×10^{-4}	3.19×10^{-4}	3.17×10^{-4}
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$	1.24×10^{-4}	5×10^{-6}	5×10^{-6}	0
Block-RDT	Number of samples, $N_{\text{B-RDT}}$	14	7	4	2
	$P_{\text{FA}}^{\text{B-RDT}}$	0	0	0	2.2×10^{-5}
	$P_{\text{MD}}^{\text{B-RDT}}$	9.45×10^{-4}	3.12×10^{-4}	2.44×10^{-4}	6.80×10^{-5}
SPRT	Stopping time, T_{SPRT}	2.44	1.73	1.34	1.05
	$P_{\text{FA}}^{\text{SPRT}}$	2.08×10^{-4}	1.59×10^{-4}	1.03×10^{-4}	2.85×10^{-5}
	$P_{\text{MD}}^{\text{SPRT}}$	2.09×10^{-4}	1.54×10^{-4}	1.03×10^{-4}	2.57×10^{-5}
SPRT-MM	Stopping time, $T_{\text{SPRT-MM}}$	1.57	1.24	1.10	1.01
	$P_{\text{FA}}^{\text{SPRT-MM}}$	6.2×10^{-3}	3.5×10^{-3}	1.8×10^{-3}	2.88×10^{-4}
	$P_{\text{MD}}^{\text{SPRT-MM}}$	6.2×10^{-3}	3.6×10^{-3}	1.8×10^{-3}	3.05×10^{-4}

Table 1: *Seq*RDT versus Block-RDT and SPRT for buffer size $M = 0$.

set $\mathcal{G} = \{e, s\}$, where $e(x) = x$ is the identity in $\mathbb{R}$ and $s(x) = -x + 2\xi_0$, is a group. It is even the smallest group for which the sequel holds true. Consider the group action [10, Definition 6.3, p. 186] π that associates each $g \in \mathcal{G}$ to the map $\pi_g : \mathbb{R}^{[\![1,N]\!]} \to \mathbb{R}^{[\![1,N]\!]}$ by: $\pi_g(x) = (g(x_1), g(x_2), \ldots, g(x_N))$. It can be easily shown that $\pi_g(Y)$ satisfies the same model as Y, i.e., $\pi_g(Y) = \pi_g(\Xi) + X'$, where, $\pi_g(\Xi)$ is independent of X' and $X' = (X'_1, X'_2, \ldots, X'_N) \sim \mathcal{N}(0, I_N)$. Moreover, we have $|\langle \pi_g(\Xi) \rangle_N - \xi_0| = |\langle \Xi \rangle_N - \xi_0|$. Therefore, the above change-in-mean detection problem (9) remains unchanged by substituting $\pi_g(\Xi)$ for Ξ and X' for X. In [2], the authors seek π-invariant tests, that is, N dimensional tests $\mathcal{T} : \mathbb{R}^{[\![1,N]\!]} \to \{0, 1\}$ that are invariant under the action of $\pi : \mathcal{T}(\pi_g(x)) = \mathcal{T}(x)$ for any $g \in \mathcal{G}$ and any $x \in \mathbb{R}^N$. Because signal averaging reduces noise variance, we naturally consider π-invariant integrator tests, that is, tests $\mathcal{T} : \mathbb{R}^{[\![1,N]\!]} \to \{0, 1\}$ that are not merely π-invariant but for which exists $\bar{\mathcal{T}} : \mathbb{R} \to \{0, 1\}$, such that $\forall\ x \in \mathbb{R}^{[\![1,N]\!]}$, $\mathcal{T}(x) = \bar{\mathcal{T}}(\langle x \rangle_N)$. The test $\bar{\mathcal{T}}$ is called the reduced form of $\mathcal{T}$ and turns to be $\mathcal{G}$-invariant: $\bar{\mathcal{T}}(g(x)) = \bar{\mathcal{T}}(x)$ for any $x \in \mathbb{R}$ and any $g \in \mathcal{G}$. Tests of type $\mathcal{T}_{\lambda_\gamma(\tau\sqrt{N})/\sqrt{N}}(Y)$, as defined in (3) with $\gamma \in (0, 1)$, are then UMP among all π-invariant integrator tests with level γ for the above change-in-mean detection problem [2].

Now, consider the Block-RDT approach for the detection of the same distorted signal model as considered for *Seq*RDT and for the same tolerance, $\tau = 2\sigma_\Delta$ we have $\mathbb{P}[|\Delta_n| \leqslant \tau] = 0.9545$ for all $n \in \mathbb{N}$. Here we do not have $|\Delta_n| \leqslant \tau$, however, we show via simulations that this does not impact the results below. The goal of Block-RDT is to design an α level test such that the probability of missed detection $P_{\text{MD}}^{\text{B-RDT}}$ stays below level β. The threshold is chosen to be $\lambda_\alpha(\sqrt{N}\tau)/\sqrt{N}$ from (3) and [2]. We denote the probability of false alarm as $P_{\text{FA}}^{\text{B-RDT}}$ and the number of samples required to achieve $P_{\text{MD}}^{\text{B-RDT}}$ by $N_{\text{B-RDT}}$. Now, let us consider SPRT for the cases when the distortion is known and when the it is unknown.

SPRT computes the likelihood ratio (Λ_N) of the observations under the two hypotheses as a function of the number of observations. The aim of SPRT is to decide which hypothesis is true as soon as possible. For this purpose, two thresholds $\lambda_H^{\text{SPRT}} > \lambda_L^{\text{SPRT}}$ are chosen as $\lambda_H^{\text{SPRT}} = (1 - \beta)/\alpha$ and $\lambda_L^{\text{SPRT}} = \beta/(1 - \alpha)$ such that the probability of false alarm $P_{\text{FA}}^{\text{SPRT}}$ and the probability of missed detection $P_{\text{FA}}^{\text{SPRT}}$ stay below α and β, respectively. SPRT stops once $\Lambda_N \geqslant \lambda_H^{\text{SPRT}}$ and decides in favor of $\mathcal{H}_1$ or once $\Lambda_N \leqslant \lambda_L^{\text{SPRT}}$ and decides in favor of $\mathcal{H}_0$. Otherwise SPRT updates the likelihood ratio for $N + 1$th observation and repeats the procedure. Let T_{SPRT} denote the stopping time for SPRT [3, 4]. We represent the algorithm by SPRT when the distortion is completely known and by SPRT-MM when the distortion is unknown. We denote by $T_{\text{SPRT-MM}}$, $P_{\text{FA}}^{\text{SPRT-MM}}$ and $P_{\text{MD}}^{\text{SPRT-MM}}$ the stopping time, the probability of false alarm and the probability of missed detection, respectively, for SPRT-MM. Knowing the distortion completely is infeasible in most practical situations, therefore, SPRT-MM will capture the performance loss when there are model mismatches. Next, we compare *Seq*RDT to SPRT and SPRT-MM.

5.3. Comparison

We define $|\xi_1 - \xi_0|$ as the Signal-to-Noise Ratio (SNR) and $\frac{|\xi_1 - \xi_0|}{\tau}$ as Signal-to maximum-Distortion Ratio (SDR) [1, 2]. We analyze the average number of samples taken by *Seq*RDT compared to its fixed sample size counterpart Block-RDT, SPRT and SPRT-MM. We choose the distortion variance to be $\sigma_\Delta = 1$. We average the stopping times and count the probability of false alarm and missed detection over 10^6 Monte carlo iterations for *Seq*RDT, SPRT and SPRT-MM. Whereas, for Block-RDT the P_{FA}^{B-RDT} and P_{MD}^{B-RDT} can be derived in closed form for the given signal model. In Table 1, we compare the number of samples taken by *Seq*RDT versus Block-RDT, SPRT and SPRT-MM for different SDR (SNR) values and for levels $\alpha = \beta = 0.001$. We also compare the probability of false alarm and missed detection achieved by each of the tests. Notice that *Seq*RDT is faster compared to the Block-RDT especially at low SNRs. SPRT is optimal when distortion and noise distributions are completely known but note that for SPRT-MM the probability of false alarm and missed detection do not stay below the levels α and β, respectively at low SNRs.

6. CONCLUSION AND FUTURE WORK

In this work, we proposed a sequential algorithm *Seq*RDT for detecting the change-in-mean of a non-stationary random process. The performance of the algorithm was analyzed and compared to Block-RDT and SPRT for a simple Gaussian change-in-mean detection problem. It was shown that the proposed algorithm makes a decision faster on average compared to its equivalent fixed sample test Block-RDT and is robust to model mismatches compared to SPRT, especially at low SNRs. Future directions include a detailed study of the threshold behavior. Also, the bounds on the probabilities of errors need to be derived in the future.

7. REFERENCES

[1] D. Pastor and Q.-T. Nguyen, "Random Distortion Testing and Optimality of Thresholding Tests," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 4161 – 4171, Aug. 2013.

[2] D. Pastor and Q. T. Nguyen, "Robust statistical process control in Block-RDT framework," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2015, pp. 3896–3900.

[3] A. Wald, "Sequential tests of statistical hypotheses," *The Annals of Mathematical Statistics*, vol. 16, no. 2, pp. 117–186, 06 1945.

[4] A. Wald, *Sequential Analysis*, John Wiley and Sons, New York, 1948.

[5] A. Tartakovsky, M. Basseville, and I.V. Nikiforov, *Sequential Analysis: Hypothesis Testing and Changepoint Detection*, Chapman & Hall/CRC, 2014.

[6] Quang-Thang Nguyen, *Contributions to statistical signal processing with applications in biomedical engineering*, Phd thesis, Telecom Bretagne, 2012.

[7] Quang-Thang Nguyen, Dominique Pastor, and Erwan L'Her, "Automatic detection of autopeep during controlled mechanical ventilation," *BioMedical Engineering OnLine*, vol. 11, no. 1, pp. 32, 2012.

[8] K. R. Sahasranand and V. Sharma, "Distributed nonparametric sequential spectrum sensing under electromagnetic interference," in *ICC 2015, London, UK*, June 8-12, 2015.

[9] Y. Sun, A. Baricz, and S. Zhou, "On the Monotonicity, Log-Concavity, and Tight Bounds of the Generalized Marcum and Nuttall Q-Functions," *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1166 – 1186, Mar. 2010.

[10] M. L. Eaton, *Multivariate statistics. A vector space approach*, Wiley, 1983.