# On Random Distortion testing based sequential Non parametric Hypothesis Testing

Prashant Khanduri, Dominique Pastor, Vinod Kumar Sharma, Pramod K. Varshney

# On Random Distortion Testing Based Sequential Non-Parametric Hypothesis Testing*

Prashant Khanduri[1], Dominique Pastor[2], Vinod Sharma[3] and Pramod K. Varshney[1]

***Abstract*— In this work, we propose a new method for sequential binary hypothesis testing. The approach is non-parametric in the sense that it does not assume any knowledge of signal distributions under each hypothesis. The proposed framework is based on Random distortion testing (RDT) which addresses the problem of testing whether or not a random signal, $\Xi$, deviates by more than a specified tolerance, $\tau$, from a fixed value, $\xi_0$. We first state the problem setup and then discuss earlier approaches to solve the problem. We then propose a new sequential algorithm, T-*Seq*RDT, which is shown to control the probabilities of error while reducing the number of samples required to make a decision compared to the fixed-sample-size version of RDT. Finally, via simulations we compare T-*Seq*RDT to other algorithms and show its robustness compared to standard likelihood ratio based approaches.**

## I. INTRODUCTION

Standard binary hypothesis testing problems, tests for the null ($\mathcal{H}_0$) against the alternate ($\mathcal{H}_1$) hypothesis based on a fixed number of samples [1]. However, many decision problems encountered in practical scenarios are inherently sequential [2]–[4]. In his seminal works [5], [6], Wald proposed a likelihood ratio based approach, Sequential Probability Ratio Test (SPRT). Optimality of SPRT was shown in the sense that it is faster on average compared to all the procedures with same error probabilities. Which implies that it also improved on decision making time compared to the standard fixed-sample-size tests. However, fixed-sample-size tests with same error probabilities can turn out to be faster than SPRT in some cases [7], [8]. A truncated version of SPRT was proposed in [9] to avoid such scenarios. The algorithm guaranteed the average stopping time to always stay below the fixed-sample-size tests at the expense of little increase in error probabilities.

All the approaches discussed above are based on likelihood ratio tests and computing the likelihood ratio requires complete knowledge of the distributions of observations under both hypotheses. In many practical scenarios, it might not be possible to model these distributions with precision. In such cases, guaranteeing performance becomes even more challenging [10]. Moreover, a majority of approaches for sequential testing make independent and identically distributed (i.i.d) assumptions on the observations under each hypothesis [11]. But, this assumption might not necessarily hold true in many practical systems like radar, sonar and communication systems, where the signals of interest are distorted by environment and observed in noise. Solutions in the literature aimed at relaxing i.i.d assumptions are still based on likelihood ratio tests [11]–[14] which may suffer from lack of robustness. This suggests the need for designing approaches that assume little knowledge of the underlying signals to be tested.

In this work, we propose an alternative formulation to binary hypothesis testing which is different from the likelihood ratio based approaches. The proposed approach does not assume knowledge of the signal distributions and also avoids the i.i.d. assumption. Specifically, we consider the hypothesis test $\mathcal{H}_0 : \xi = \xi_0$ versus $\mathcal{H}_1 : \xi \neq \xi_0$. We observe one dimensional $Y$, whose probability distribution is parameterized by $\xi$. Therefore, the motive is to figure out whether $Y$ is a corrupted version of some specified deterministic model, $\xi_0$. However, this $\xi_0$ in practical situations will itself be distorted because of measurement errors and environmental fluctuations other than noise [10]. Therefore, it is justifiable to allow the null hypothesis $\mathcal{H}_0$ to test for the signal in neighborhood of $\xi_0$. In this regard, we assume the signal to be tested, $\Xi$, as the distorted version of $\xi_0$. The hypothesis testing problem then becomes:

$$\mathcal{H}_0 : |\Xi - \xi_0| \leqslant \tau \quad \text{vs} \quad \mathcal{H}_1 : |\Xi - \xi_0| > \tau \tag{1}$$

where, $\tau \in [0, \infty)$ models the possible distortion. This problem was first considered in [10] where $\Xi$ with unknown distribution, was embedded in i.i.d. Gaussian noise. The authors showed that the optimal tests (under certain criteria) didn't need the computation of the likelihood ratios and are therefore independent of signal distributions. To improve the detection performance, the authors extended RDT to fixed-sample-size tests, *Block*RDT [15]. They generalized the RDT formulation by replacing $\Xi$ with its empirical mean over time. Later, in [16], [17] the authors further extended RDT framework for sequential testing where *Seq*RDT was proposed to guarantee both false alarm and missed detection probabilities below certain levels, while making a decision faster on average compared to *Block*RDT. In this work, we first briefly review *Block*RDT and *Seq*RDT and then propose the truncated version of *Seq*RDT: T-*Seq*RDT. Just like the truncated version of SPRT, we show that T-*Seq*RDT guarantees the average stopping time to always stay below the fixed-sample-size test, *Block*RDT, at the expense of little increase in error probabilities.

*The work of P. Khanduri and P. K. Varshney was supported by ARO Grant W911NF-14-1-0339. The work of D. Pastor was supported by Region Bretagne (France) and the European Regional Development Fund (ERDF).

[1]Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244, USA

[2] IMT Atlantique, Lab-STICC, UBL, Brest 29238, France

[3] Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012, India

In the following, we summarize the main contributions of the work:

- We review the RDT framework for fixed-sample-size tests, *Block*RDT and its sequential extension, *Seq*RDT.
- We propose the truncated version of *Seq*RDT: T-*Seq*RDT.
- We derive bounds on false alarm and missed detection probabilities for the proposed algorithm and show that they stay below pre-specified levels.
- We compare different algorithms via simulations and show the robustness of the proposed framework compared to likelihood ratio based approaches.

In Section II, we state the testing problem. In Section III, we review the fixed-sample-size test *Block*RDT to solve the testing problem. In Section IV, we review the untruncated sequential approach *Seq*RDT to solve the testing problem stated in Section II. Then in Section V, we introduce the truncated version of *Seq*RDT. In Section VI, we compare T-*Seq*RDT to other algorithms via simulations. Finally in Section VII, we conclude the paper.

Next we define the notations before stating the problem.

*Notation:* 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 $\xi$ and variance $\sigma^2$. In what follows, $Q_{1/2}$ denotes the Generalized Marcum Function [18] 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 $Q_{1/2}(\rho, x) = \gamma$ [10, Lemma 2, statement (ii)], so that:

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

We also recall that the Marcum function increases with its first parameter and decreases with the second one [18]. The set of all real random variables defined on $(\Omega, \mathcal{F})$ is denoted by $\mathcal{M}(\Omega, \mathbb{R})$. Accordingly, $\mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}$ is the set of all real valued sequences or random processes defined on $\mathbb{N}$ (resp. $[\![1, N]\!] = \{1, 2, \ldots, N\}$). Given $U$ in $\mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}$, a realization of $U$ for $n \in \mathbb{N}$ (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 $U \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}$ defined as: $\langle U \rangle_N = \frac{1}{N} \sum_{n=1}^{N} U_n$.

## II. Problem Statement

Let $\Xi = (\Xi_n)_{n \in \mathbb{N}}$ the signal to be tested 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. No assumption is made on the stationarity or the distribution of $\Xi = (\Xi_n)_{n \in \mathbb{N}}$. Therefore, the samples $\Xi_n$ are not necessarily i.i.d. No knowledge of underlying signal distributions makes the likelihood ratio based tests (SPRT) unsuitable for such problems. As a substitute to likelihood ratio based approaches, we propose RDT based tests [10] as follows. The observation process is $Y = (Y_n)_{n \in \mathbb{N}}$ with $Y_n = \Xi_n + X_n$ for all $n \in \mathbb{N}$ and we write $Y = \Xi + X$. Noise, $X = (X_n)_{n \in \mathbb{N}}$, is assumed to be independent and Gaussian distributed. In the absence of any distortions $\Xi = \xi_0$, but this will not be true in practical scenarios as discussed earlier. Therefore, we consider the signal, $\Xi$, as random which models distortion around a fixed model $\xi_0$. We expect the empirical mean $\langle \Xi \rangle_N$ to remain close to $\xi_0$ under $\mathcal{H}_0$ and drift away from $\xi_0$ under $\mathcal{H}_1$ for all $N \geqslant N_0$. Therefore, the problem is then of testing the null event $|\langle \Xi \rangle_N - \xi_0| \leqslant \tau$ against alternate event $|\langle \Xi \rangle_N - \xi_0| > \tau$ on the basis of observations $Y$. The hypothesis testing problem is then given as:

$$\begin{cases} \underline{\text{Observation}} : Y = \Xi + X \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}} \\ \text{with} \quad \begin{cases} \Xi = (\Xi_n)_{n \in \mathbb{N}} \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}, \\ X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathcal{N}(0, 1), \\ \Xi \text{ and } X \text{ are independent.} \end{cases} \\ \exists N_0 \in \mathbb{N}, \begin{cases} \mathcal{H}_0 : \forall N \geqslant N_0,\ |\langle \Xi \rangle_N - \xi_0| \leqslant \tau \text{ (a-s)} \\ \mathcal{H}_1 : \forall N \geqslant N_0,\ |\langle \Xi \rangle_N - \xi_0| > \tau \text{ (a-s)} \end{cases} \end{cases} \tag{3}$$

where, $\tau \in [0, \infty)$ is the tolerance. Here, $N_0$ and the tolerance $\tau$ are known *a priori* based on some prior knowledge about the signal[1]. Note, that instead of knowing the complete distribution of $\Xi$ under each hypothesis we only need to know a few parameters to perform the test. In earlier works, the authors have proposed two methods to solve problem (3). First is *Block*RDT proposed in [15], which solves this problem for a fixed-sample-size test, i.e., for a fixed $N$. And the second is *Seq*RDT proposed in [16], [17], which solves the problem using a sequential approach.

*Remark 1:* The RDT formulation as given in (1) performs the binary hypothesis test given in (3) for $N = 1$ and the *Block*RDT formulation generalizes this to any fixed $N > 1$. Note, that $N \geqslant N_0$ in (3) makes it possible to generalize the testing problem for sequential approaches.

Now, we define the following type of tests.

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

$$\mathcal{T}_{N,\gamma}(x) = \begin{cases} 0 & \text{if} \quad |\langle x \rangle_N - \xi_0| \leqslant \lambda_\gamma(\tau\sqrt{N})/\sqrt{N} \\ 1 & \text{otherwise.} \end{cases} \tag{4}$$

where the threshold $\lambda_\gamma(\tau\sqrt{N})/\sqrt{N}$ is defined by (2). This type of tests are of special interest to us because of their optimality properties described next. We first review *Block*RDT and *Seq*RDT before proposing the truncated version of *Seq*RDT: T-*Seq*RDT.

## III. *Block*RDT

In this section, we discuss *Block*RDT, we show how *Block*RDT naturally leads to the need to design sequential approaches to test (3). *Block*RDT framework tests the hypotheses defined in (3) for a fixed number of samples, i.e., $Y = \Xi + X \in \mathcal{M}(\Omega, \mathbb{R})^{[\![1, N]\!]}$. A solution to this problem is proposed in [15] & [19]. Next, we discuss the optimality

[1]This knowledge can be obtained by using some standard machine learning procedures or based on some statistical knowledge of the signal. Discussion of these procedures are beyond the scope of this work.

properties of the test defined in (4) for *Block*RDT. First, we define a test as, given $N \in \mathbb{N}$, any (measurable) map $\mathcal{T} : \mathbb{R}^N \to \{0, 1\}$ is called an $N$-dimensional test. The size of such a test $\mathcal{T}$ is defined as

$$\alpha_{\mathcal{T}} = \sup_{\Xi \in \mathcal{M}(\Omega,\mathbb{R})^{[\![1,N]\!]} : \mathbb{P}[|\langle \Xi \rangle_N - \xi_0| \leqslant \tau] \neq 0} \mathbb{P}\left[\mathcal{T}(Y) = 1 \,\middle|\, |\langle \Xi \rangle_N - \xi_0| \leqslant \tau\right]$$

where the test $\mathcal{T}$ is said to have level $\gamma \in (0, 1)$ if $\alpha_{\mathcal{T}} \leqslant \gamma$. Moreover, if we have an $N$-dimensional test $\mathcal{T}^*$ such that $\alpha_{\mathcal{T}^*} \leqslant \gamma$ and $\mathbb{P}\big[\mathcal{T}^*(Y) = 1 \,|\langle \Xi \rangle_N - \xi_0| > \tau\big] \geqslant \mathbb{P}\big[\mathcal{T}(Y) = 1 \,|\langle \Xi \rangle_N - \xi_0| > \tau\big]$ for any $N$-dimensional test $\mathcal{T}$, then $\mathcal{T}^*$ is said to be Uniformly Most Powerful (UMP). However, no UMP test with level $\gamma$ exists for *Block*RDT. We therefore show the subclass of *Block*RDT-coherent tests, among which a "best" test exists [19]. An $N$-dimensional test $\mathcal{T}$ is for *Block*RDT-coherent if it satisfies the following properties:

1) **[Invariance in mean]** For all $y, y' \in \mathbb{R}^N$, if $\langle y \rangle_N = \langle y' \rangle_N$, then we should have $\mathcal{T}(y) = \mathcal{T}(y')$.
2) **[Constant conditional power]** If for all $\Xi \in \mathcal{M}(\Omega, \mathbb{R})^{[\![1,N]\!]}$ such that $\Xi$ and $X$ are independent, there exists a domain $\mathcal{D}$ of $|\langle \Xi \rangle_N - \xi_0|$ such that, for any $\rho \in \mathcal{D} \cap (0, \infty)$, the decision $\mathbb{P}\big[\mathcal{T}(Y) = 1 \,|\, |\langle \Xi \rangle_N - \xi_0| = \rho\big]$ is independent of $\mathbb{P}\,|\langle \Xi \rangle_N - \xi_0|^{-1}$.

**[Invariance in mean]** implies that $\mathcal{T}$ should yield the same decision for two different observation processes $\big(y, y' \in \mathbb{R}^N\big)$ with same empirical mean $(\langle y \rangle_N = \langle y' \rangle_N)$. On the other hand, **[Constant conditional power]** implies that $\mathcal{T}$ should yield same decision conditioned on $|\langle \Xi \rangle_N - \xi_0| = \rho$, irrespective of the distribution of $|\langle \Xi \rangle_N - \xi_0|$.

Let us denote by $\mathcal{K}_\gamma$ the class of all *Block*RDT-coherent tests with level $\gamma$. This class of tests can be partially ordered as: given $\mathcal{T}$ and $\mathcal{T}'$ in the set $\mathcal{K}_\gamma$, we say $\mathcal{T} \preceq \mathcal{T}'$ if, for any $\Xi \in \mathcal{M}(\Omega, \mathbb{R})^{[\![1,N]\!]}$, $\mathcal{T}$ and $\mathcal{T}'$ satisfy **[Constant conditional power]** on the same domain $\mathcal{D}$ and for all $\rho \in \mathcal{D} \cap (\tau, \infty)$, we have

$$\mathbb{P}\big[\mathcal{T}(Y) = 1 \,|\, |\langle \Xi \rangle_N - \xi_0| = \rho\big] \leqslant \mathbb{P}\big[\mathcal{T}'(Y) = 1 \,|\, |\langle \Xi \rangle_N - \xi_0| = \rho\big].$$

The $N$-dimensional test defined for all $y \in \mathbb{R}^N$ given by (4) is maximal in $\mathcal{K}_\gamma$, i.e., for any $\mathcal{T} \in \mathcal{K}_\gamma$ we have $\mathcal{T} \preceq \mathcal{T}_{N,\gamma}$ [15], [19].

Let $\mathbb{P}_{\text{FA}}^{\text{B-RDT}}(N, \gamma)$ and $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}(N, \gamma)$ denote the false alarm and missed detection probabilities of *Block*RDT, respectively, when the testing on $Y$ is performed by $\mathcal{T}_{N,\gamma}$, so that $\mathcal{T}_{N,\gamma}(Y)$ is the accepted hypothesis. We have the following proposition.

*Proposition 3.1:* For any $\gamma \in (0, 1)$ and $\tau \geqslant 0$, we have $\mathbb{P}_{\text{FA}}^{\text{B-RDT}} \leqslant \gamma$ and $\mathbb{P}_{\text{MD}}^{\text{B-RDT}} \leqslant 1 - \gamma$.

Proposition 3.1 suggests that the tests defined in (4) although optimal for *Block*RDT, are unable to control both false alarm and missed detection probabilities with a single threshold designed for a fixed $\gamma$. It suggests the use of two thresholds to control both false alarm and missed detection probabilities, which naturally leads to a sequential approach. In sequential algorithms, a decision is not made until there is sufficient evidence to decide in favor of one hypothesis, thereby, making the stopping time a random variable. Moreover, sequential approaches save on the number of samples required to make a decision compared to fixed-sample-size tests.

Next, we discuss an untruncated sequential approach to solve (3).

## IV. *Seq*RDT

In this section, we briefly review *Seq*RDT proposed in [16], [17]. *Seq*RDT tests the hypothesis defined in (3) in a sequential manner. Let us assume, $\alpha, \beta \in (0, 1/2)$ are the prespecified levels for false alarm and missed detection probabilities, respectively. The goal is then to design *Seq*RDT such that the false alarm and missed detection probabilities stay below $\alpha$ and $\beta$, respectively. To do so, next assumption plays an important role in the design of *Seq*RDT.

*Assumption 4.1 (Bounded behavior of $|\langle \Xi \rangle_N - \xi_0|$):* There exist $\tau^- \in [0, \tau)$, $\tau^+ \in (\tau, \infty)$ and $\tau_H \in (\tau^+, \infty)$ such that:

$$\begin{cases} \text{Under } \mathcal{H}_0 : \forall N \geqslant N_0, 0 \leqslant |\langle \Xi \rangle_N - \xi_0| \leqslant \tau^- \text{ (a-s)}, \\ \text{Under } \mathcal{H}_1 : \forall N \geqslant N_0, \tau^+ \leqslant |\langle \Xi \rangle_N - \xi_0| \leqslant \tau_H \text{(a-s)}. \end{cases}$$

*Remark 2:* In [16], [17], *Seq*RDT used this assumption to design a buffer, $M$, which made it possible to control the false alarm and missed detection probabilities below $\alpha$ and $\beta$, respectively. The assumption implies that the empirical mean of the signal under the two hypotheses is bounded away from $\tau$.

Now, we define the test. Given any natural number $M \geqslant N_0 - 1$, *Seq*RDT for testing $\mathcal{H}_0$ against $\mathcal{H}_1$ in (3) is specified by defining the stopping time:

$$T = \min\big\{N \in \mathbb{N} : \mathcal{D}_M(N) \neq \infty\big\}, \tag{5}$$

$$\text{with:} \begin{cases} \mathcal{D}_M(1) = \mathcal{D}_M(2) = \ldots = \mathcal{D}_M(M) = \infty, \\ \text{for } N > M, \\ \mathcal{D}_M(N) = \begin{cases} 0 & \text{if } |\langle Y \rangle_N - \xi_0| \leqslant \lambda_L(N), \\ \infty & \text{if } \lambda_L(N) < |\langle Y \rangle_N - \xi_0| \leqslant \lambda_H(N), \\ 1 & \text{if } |\langle Y \rangle_N - \xi_0| > \lambda_H(N). \end{cases} \end{cases} \tag{6}$$

where $\lambda_L(N)$ and $\lambda_H(N)$ were designed from (2) as:

$$\lambda_L(N) = \lambda_{1-\beta}(\tau\sqrt{N})/\sqrt{N} \text{ and } \lambda_H(N) = \lambda_\alpha(\tau\sqrt{N})/\sqrt{N}$$

This choice of thresholds guaranteed $\lambda_L(N, w_L) \leqslant \lambda_H(N, w_H)$. Here, $\mathcal{D}_M(N)$ represents the decision variable. $\mathcal{D}_M(N) = 0$ implies $\mathcal{H}_0$ is decided, $\mathcal{D}_M(N) = 1$ implies $\mathcal{H}_1$ is decided and $\mathcal{D}_M(N) = \infty$ is equivalent to saying that no decision is made at the $N^{\text{th}}$ time sample and that the algorithm will update the statistic and repeat the test with $N + 1$ samples. Note that the buffer size, $M \geqslant N_0 - 1$, is the number of samples *Seq*RDT waits for before starting the test. This $M$ was chosen based on the knowledge of $\tau^-$, $\tau^+$ and $\tau_H$ given in Assumption 4.1, which subsequently

guaranteed that the false alarm and missed detection probabilities stay below levels $\alpha$ and $\beta$, respectively. Moreover, it was shown via simulations that *Seq*RDT makes a decision faster compared to *Block*RDT and is robust to mismatches unlike likelihood ratio based approaches.

Next we, propose T-*Seq*RDT, truncated version of *Seq*RDT, which relaxes Assumption 4.1 and at the same time eliminates the need for the buffer, $M$.

## V. T-*Seq*RDT

Assumption 4.1 played a crucial role in the design of *Seq*RDT. Next we relax this assumption and present a milder assumption required for designing T-*Seq*RDT.

*Assumption 5.1 (Behavior of $|\langle \Xi \rangle_N - \xi_0|$ under $\mathcal{H}_1$):* There exist $\tau^+ \in (\tau, \infty)$ such that:

$$\text{Under } \mathcal{H}_1: \ \forall N \geqslant N_0, \ |\langle \Xi \rangle_N - \xi_0| \geqslant \tau^+ \text{(a-s)}.$$

In contrast, *Seq*RDT [16], [17] required additional assumptions on the signal (see Assumption 4.1). Now, we show that unlike the likelihood based approaches, T-*Seq*RDT can be designed with only the knowledge of this $\tau^+$ and $\tau$ instead of knowing the complete distribution of the signal. The test is truncated at a specified time if a decision has not been reached yet. At truncation time the decision is forced using *Block*RDT (Section III). We define the stopping time $T$ of T-*Seq*RDT as:

$$T = \min\left\{N \leqslant N_0 + W - 1, N \in \mathbb{N} : \mathcal{D}_{N_0}(N) \neq \infty\right\},$$

$$\text{with: } \begin{cases} \mathcal{D}_{N_0}(1) = \mathcal{D}_{N_0}(2) = \ldots = \mathcal{D}_{N_0}(N_0 - 1) = \infty, \\ \text{for } N_0 \leqslant N < N_0 + W, \\ \mathcal{D}_{N_0}(N) = \begin{cases} 0 & \text{if } |\langle Y \rangle_N - \xi_0| \leqslant \lambda_L(N) \\ 1 & \text{if } |\langle Y \rangle_N - \xi_0| > \lambda_H(N) \\ \infty & \text{if } \lambda_L(N) < |\langle Y \rangle_N - \xi_0| \leqslant \lambda_H(N) \end{cases} \\ \text{for } N = N_0 + W, \\ \mathcal{D}_{N_0}(N) = \begin{cases} 0 \text{ if } |\langle Y \rangle_N - \xi_0| \leqslant \lambda_{\text{B-RDT}}(N) \\ 1 \text{ if } |\langle Y \rangle_N - \xi_0| > \lambda_{\text{B-RDT}}(N) \end{cases} \end{cases}$$

where $N = N_0 + W$ is the truncation time. With $W \in \mathbb{N}$ defined as the truncation window. $\mathcal{D}_{N_0}(N)$ represents the decision made at the $N^{\text{th}}$ time instant and if a decision is not made at time $N$ the algorithm will update the statistic and repeat the test with $N+1$ samples. The algorithm waits for $N_0$ samples before starting the test, note that this $N_0$ is different from the buffer, $M$ designed for *Seq*RDT. $N_0$ is given *a priori* as defined in (3) whereas the buffer $M$ was designed using Assumption 4.1. The thresholds $\lambda_L(N)$, $\lambda_H(N)$ and $\lambda_{\text{B-RDT}}(N)$ must be designed jointly so as to guarantee false alarm and missed detection probabilities to be bounded below fixed levels. Also, $\lambda_H(N)$ and $\lambda_L(N)$ must be such that $\lambda_L(N) \leqslant \lambda_H(N)$. For notational simplicity we use the same notations for the thresholds here although these thresholds are different than the ones defined in Section IV.

We now define the False Alarm Probability of T-*Seq*RDT as:

$$\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0}) \stackrel{\text{def}}{=} \mathbb{P}\left[\mathcal{D}_{N_0}(T) = 1\right], \quad \text{under } \mathcal{H}_0.$$

Similarly, the Missed Detection Probability is defined as:

$$\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0}) \stackrel{\text{def}}{=} \mathbb{P}\left[\mathcal{D}_{N_0}(T) = 0\right], \quad \text{under } \mathcal{H}_1.$$

The goal of any sequential algorithm is to design the thresholds so that $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ stay below certain prespecified levels $\alpha$ and $\beta$, respectively. Moreover, we want to make a decision faster compared to *Block*RDT, the fixed-sample-size counterpart of T-*Seq*RDT. We now define the thresholds for T-*Seq*RDT.

### A. Thresholds

Now we define the thresholds used for T-*Seq*RDT based on the tests defined in (4). The goal is to get rid of the buffer required by *Seq*RDT [16], [17], while maintaining a fixed level for false alarm and missed detection probabilities and at the same time making a decision faster compared to fixed-sample-size test, *Block*RDT. One way to get rid of this buffer is to increase the upper and decrease the lower threshold. This can be achieved by defining the thresholds as:

$$\begin{aligned} \lambda_H(N) &= \lambda_{\frac{\alpha}{2W}}(\tau\sqrt{N})/\sqrt{N} \\ \lambda_L(N) &= \lambda_{1-\frac{\beta}{2W}}(\tau\sqrt{N})/\sqrt{N} \\ \lambda_{\text{B-RDT}}(N) &= \lambda_{\frac{\alpha}{2}}(\tau\sqrt{N})/\sqrt{N} \end{aligned} \tag{7}$$

Note that the thresholds $\lambda_H(N)$ and $\lambda_L(N)$ depend on the window size, $W$ as defined earlier. Now the goal is to select an appropriate $W$ such that the $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ are bounded below $\alpha$ and $\beta$, respectively. But, before that we first make the case that these thresholds actually satisfy the properties desired for T-*Seq*RDT in the following proposition.

*Proposition 5.2:* We have, $\lambda_L(N) \leqslant \lambda_{\text{B-RDT}}(N) \leqslant \lambda_H(N)$, for all $N \in \mathbb{N}$.

Now we have stated that the thresholds defined in (7) follow the properties desired by T-*Seq*RDT. Next, we use Assumption 5.1 to design $W$ and show that the $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ are bounded below $\alpha$ and $\beta$, respectively.

### B. Truncation window

As discussed in the introduction section we base the choice of the truncation window $W$ on *Block*RDT. It follows from Proposition 3.1 that $\mathbb{P}_{\text{FA}}^{\text{B-RDT}}$ of *Block*RDT is always bounded below by $\alpha/2$ for threshold $\lambda_{\text{B-RDT}}(N)$ given in (7). However, without any additional knowledge one cannot control $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}$ of *Block*RDT. To control this $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}$ we need Assumption 5.1, which is then used to design $W$ for T-*Seq*RDT. The next proposition shows how $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}$ for *Block*RDT behaves under Assumption 5.1.

*Proposition 5.3:* For any $\gamma \in (0, 1)$ the tests of type $\mathcal{T}_{N,\gamma}(Y)$ for the hypotheses testing problem as defined in

(3), $\mathbb{P}_{\text{FA}}^{\text{B-RDT}}$ and $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}$ under Assumption 5.1 are bounded as

$$\begin{cases} \mathbb{P}_{\text{FA}}^{\text{B-RDT}} \leqslant \gamma \\ \mathbb{P}_{\text{MD}}^{\text{B-RDT}} \leqslant 1 - Q_{1/2}\left(\tau^+\sqrt{N}, \lambda_\gamma(\tau\sqrt{N})\right) \end{cases}$$

where the upper bound on $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}$ decreases to 0 with $N$.

*Remark 3:* Without Assumption 5.1 $\mathbb{P}_{\text{FA}}^{\text{B-RDT}}$ and $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}$ are bounded as given in Proposition 3.1.

Proposition 5.3 states that the upper bound on $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}$ is decreasing in $N$. We use this and propose to choose the truncation window such that the upper bound on $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}$ stays below $\beta/2$. Thus we define $W^*(\alpha, \beta)$ as the smallest $W$ such that,

$$1 - Q_{1/2}\left(\tau^+\sqrt{N_0+W}, \lambda_{\frac{\alpha}{2}}(\tau\sqrt{N_0+W})\right) \leqslant \frac{\beta}{2}, \quad (8)$$

Therefore, the thresholds $\lambda_H(N)$ and $\lambda_L(N, w_L)$ become: $\lambda_H(N) = \lambda_{\frac{\alpha}{2W^*}}(\tau\sqrt{N})/\sqrt{N}$ and $\lambda_L(N) = \lambda_{1-\frac{\beta}{2W^*}}(\tau\sqrt{N})/\sqrt{N}$.

Next, we analyze the $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ performance of T-*Seq*RDT designed with the thresholds as defined in (7) and the truncation window, $W = W^*(\alpha, \beta)$.

*Theorem 5.4 (Bounds on $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$):* We have: $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0}) \leqslant \alpha$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0}) \leqslant \beta$.

This theorem states that for the thresholds defined in (7) the testing problem (3) can be solved using T-*Seq*RDT. The algorithm will guarantee levels $\alpha$ and $\beta$ for false alarm and missed detection probabilities, respectively, irrespective of the underlying signal distributions. The error probabilities guaranteed by T-*Seq*RDT are a little higher than that of *Block*RDT ($\alpha/2$ and $\beta/2$), this loss in error performance is compensated by the faster decision making of T-*Seq*RDT as shown in the next section. Moreover, we know that the average stopping time of T-*Seq*RDT will always be smaller than or equal to that of *Block*RDT as T-*Seq*RDT will stop at the truncation time $N_0 + W^*$ if a decision is not made until that time. More importantly, this is achieved with T-*Seq*RDT without Assumption 4.1 and without the need to design a buffer, $M$, although T-*Seq*RDT required a milder Assumption 5.1. Next, we perform some simulations to analyze the performance of T-*Seq*RDT.

## VI. Experimental Results and Discussion

In this section, we compare T-*Seq*RDT to *Block*RDT, *Seq*RDT and SPRT via simulations. We first present the detection problem considered to perform the experiments. Then we carry out the comparison of the algorithms for different parameter values.

### A. Detection with signal distortions

We consider the case when $Y_n = \Xi_n + X_n$, for $n \in \mathbb{N}$, with $\Xi_n = \xi_0$ under $\mathcal{H}_0$ and $\Xi_n = \xi_1$ under $\mathcal{H}_1$. Here $\xi_0$ and $\xi_1$ are deterministic constants and $X_n \sim \mathcal{N}(0,1)$ for any $n \in \mathbb{N}$. If $\xi_0$ is known precisely, this model can be formulated in the framework of (3) with $\tau = 0$ and $N_0 = 1$. However, in many practical systems $\xi_0$ might not be known precisely and there might be a mismatch between the model and the signal observed in practice. In such cases, the signal, $\Xi_n$, will not be a constant $\xi_0$ or $\xi_1$ under each hypothesis, but a distorted version of these values. These distortions are difficult to model in a parametric setup, therefore, likelihood ratio based tests will fail to guarantee reliable performance [10]. However, the testing framework proposed in (3) for *Block*RDT, *Seq*RDT and T-*Seq*RDT is not limited by these drawbacks. Therefore, instead of dealing with a precise model, we consider the case when $\Xi_n = \xi_i + \Delta_n$ under $\mathcal{H}_i$ for $i \in \{0, 1\}$ and for all $n \in \mathbb{N}$. Here, $\Delta_n$s model the possible additive distortions. We thus want to experimentally compare different algorithms for testing $\Xi = (\Xi_n)_{n\in\mathbb{N}}$ when we observe $Y = (Y_n)_{n\in\mathbb{N}}$. We focus on algorithms guaranteeing false alarm and missed detection probabilities below specified levels $\alpha$ and $\beta$, respectively. For the sake of conciseness, we simply say that these algorithms guarantee $\alpha$ and $\beta$.

### B. Experimental setup

First, we list the parameters required to design each algorithm: *Block*RDT needs $\tau$, T-*Seq*RDT needs $\tau$ and $\tau^+$, *Seq*RDT needs $\tau^-$, $\tau$, $\tau^+$ and $\tau_H$ whereas, SPRT requires complete knowledge of the signal distributions under both hypotheses. For the experimental setup, we consider $\xi_1$ and $\xi_0$ such that $|\xi_1 - \xi_0| \geqslant 4\tau^-$, we set $\tau^+ = |\xi_1 - \xi_0| - \tau^-$ and $\tau_H \in [|\xi_1 - \xi_0| + \tau^-, \infty)$. Assume that the distortion $\Delta = (\Delta_n)_{n\in\mathbb{N}}$ satisfies: $|\Delta_n| \leqslant \tau^-$ and $\tau^+ \leqslant |\Delta_n + \xi_1 - \xi_0| \leqslant \tau_H$. The first inequality captures the behavior of the signal under $\mathcal{H}_0$ and the second inequality captures the behavior of the signal under $\mathcal{H}_1$. Using these, the problem of testing $\Xi$ can be written as:

$$\begin{cases} \text{under } \mathcal{H}_0: \ 0 \leqslant |\langle \Xi \rangle_N - \xi_0| \leqslant \tau^- < \tau, \ \forall N \geqslant 1, \\ \text{under } \mathcal{H}_1: \ \tau < \tau^+ \leqslant |\langle \Xi \rangle_N - \xi_0| \leqslant \tau_H, \ \forall N \geqslant 1. \end{cases} \quad (9)$$

with $\tau \in (\tau^-, \tau^+)$. For simulations, we set $\tau = 2\tau^-$. *Block*RDT, *Seq*RDT and T-*Seq*RDT can be designed using (9). However, all types of distortion may not satisfy (9) with probability 1, we therefore consider the case when $\Delta_1, \Delta_2, \cdots \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$. The algorithms *Block*RDT, *Seq*RDT and T-*Seq*RDT are still able to guarantee $\alpha$ and $\beta$ when (9) is satisfied only with high probability [15]–[17]. This indicates that the proposed algorithms are flexible as well as robust to model mismatches. Next, we discuss the algorithms for testing of the distorted signal as defined above.

Block*RDT:* For the rest of the section we choose $\tau^- = \sigma$. Since the distortion distribution is assumed to be Gaussian, we can compute the probabilities associated with (9). We have for tolerance $\tau$, $\mathbb{P}[|\langle \Delta \rangle_N| \leqslant \tau] \geqslant 0.9545$ and $\mathbb{P}[|\langle \Delta \rangle_N + \xi_1 - \xi_0| > \tau] \geqslant 0.9772$ for all $N \geqslant 1$ and $|\xi_1 - \xi_0| \geqslant 2\tau$, with equality for $N = 1$ and $|\xi_1 - \xi_0| = 2\tau$. Although the above probabilities are not equal to 1, *Block*RDT can be experimentally shown to guarantee $\alpha$ and $\beta$ [15]. From Proposition 3.1, we know that the threshold $\lambda_\alpha(\tau\sqrt{N})/\sqrt{N}$

| SNR $= \lvert\xi_1 - \xi_0\rvert$ | | 4 | 5 | 6 | 8 |
|---|---|---|---|---|---|
| *Block*RDT | Number of samples, $N_{\text{B-RDT}}$ | 39 | 10 | 5 | 2 |
| *Seq*RDT | Average stopping time, $\mathbb{E}[T_{SeqRDT}]$ | 3.98 | 3.28 | 3.04 | 2.90 |
| T-*Seq*RDT | Average stopping time, $\mathbb{E}[T]$ | 6.3318 | 4.4941 | 3.7770 | 3.1204 |
| SPRT | Average stopping time, $\mathbb{E}[T_{\text{SPRT}}]$ | 2.44 | 1.73 | 1.34 | 1.05 |
| SPRT-MM | Average stopping time, $\mathbb{E}[T_{\text{SPRT-MM}}]$ | 1.57 | 1.24 | 1.10 | 1.01 |
| | $\mathbb{P}_{\text{FA}}^{\text{SPRT-MM}}$ | $6.2 \times 10^{-3}$ | $3.5 \times 10^{-3}$ | $1.8 \times 10^{-3}$ | $2.88 \times 10^{-4}$ |
| | $\mathbb{P}_{\text{MD}}^{\text{SPRT-MM}}$ | $6.2 \times 10^{-3}$ | $3.6 \times 10^{-3}$ | $1.8 \times 10^{-3}$ | $3.05 \times 10^{-4}$ |

TABLE I

T-*Seq*RDT, *Seq*RDT, SPRT VS *Block*RDT FOR $\alpha = \beta = 0.001$.

will guarantee the false alarm probability to stay below $\alpha$. However, it does not guarantee the missed detection probability to stay below $\beta$. However, Proposition 5.3 shows that with the knowledge of $\tau^+$, we can control the missed detection probability by choosing an appropriate block-size such that the missed detection probability stays below pre-specified level $\beta$.

*SeqRDT:* For the same choice of tolerances $\tau$, $\tau^-$ and $\tau^+$, we have $\mathbb{P}[|\langle\Delta\rangle_N| \leqslant \tau^-] \geqslant 0.6827$ and $\mathbb{P}[|\langle\Delta\rangle_1 + \xi_1 - \xi_0| > \tau^+] \geqslant 0.8413$ for all $N \geqslant 1$ and $|\xi_1 - \xi_0| \geqslant 2\tau$, with equality when $N = 1$ and $|\xi_1 - \xi_0| = 2\tau$. Again, although these probabilities are not exactly one, *Seq*RDT, will perform the testing while guaranteeing $\alpha$ and $\beta$. For *Seq*RDT, we need to know all $\tau^-$, $\tau$, $\tau^+$ and $\tau_H$ where, $\tau$ is used to design the threshold, $\tau^+$, $\tau^-$ and $\tau_H$ are used to choose an appropriate buffer size, $M$, such that the algorithm guarantees $\alpha$ and $\beta$. As shown in [16], [17] for the above model, thresholds $\lambda_H(N) = \lambda_\alpha(\tau\sqrt{N})/\sqrt{N}$, $\lambda_L(N) = \lambda_{1-\beta}(\tau\sqrt{N})/\sqrt{N}$ and $M = 0$ will guarantee $\alpha$ and $\beta$ for the algorithm. We denote by $\mathbb{E}[T_{SeqRDT}]$ the average stopping time of *Seq*RDT.

*T-SeqRDT:* Similar to *Block*RDT, T-*Seq*RDT guarantees $\alpha$ and $\beta$ by the knowledge of only $\tau$ and $\tau^+$. We choose the same $\tau$ and $\tau^+$ as for *SeqRDT*. Hence, the probabilities $\mathbb{P}[|\langle\Delta\rangle_N| \leqslant \tau]$ and $\mathbb{P}[|\langle\Delta\rangle_1 + \xi_1 - \xi_0| > \tau^+]$ remain the same as for both *Block*RDT and *Seq*RDT. Therefore, the testing of the distorted signal as described above can be performed using T-*Seq*RDT by casting the problem in the framework as defined in (3). The truncation window, $W^*$, for T-*Seq*RDT is selected via (8). This $W^*$ along with $\alpha$ and $\beta$ is then used to design the thresholds as defined in (7). Theorem 5.4 then ensures that T-*Seq*RDT will guarantee $\alpha$ and $\beta$.

*Sequential Probability Ratio Test (SPRT):* For SPRT we assume that complete knowledge of probability density functions of the observations $f_i$ is available under $\mathcal{H}_i$ for $i = 0, 1$. For $\alpha, \beta \in (0, 1/2)$, and with initialization $N = 1$, SPRT is defined as:

$$\begin{cases} \text{If } \Lambda_N \leqslant \lambda_L^{\text{SPRT}}, \text{ decide } \mathcal{H}_0 \text{ and stop;} \\ \text{If } \Lambda_N \geqslant \lambda_H^{\text{SPRT}}, \text{ decide } \mathcal{H}_1 \text{ and stop;} \\ \text{If } \lambda_L^{\text{SPRT}} < \Lambda_N < \lambda_H^{\text{SPRT}}, \text{compute } \Lambda_{N+1} \text{ and repeat;} \end{cases} \tag{10}$$

where $\Lambda_N = \sum_{n=1}^{N} \log \frac{f_1(Y_i)}{f_0(Y_i)}$ is the log-likelihood ratio of the observations, $\lambda_L^{\text{SPRT}} = \log \frac{\beta}{1-\alpha}$ is the lower threshold and $\lambda_H^{\text{SPRT}} = \log \frac{1-\beta}{\alpha}$ is the upper threshold. The probability of false alarm $\mathbb{P}_{\text{FA}}^{\text{SPRT}}$ and the probability of missed detection $\mathbb{P}_{\text{MD}}^{\text{SPRT}}$ of SPRT are guaranteed to stay below $\alpha$ and $\beta$, respectively [5], [6], and the average stopping time is denoted by $\mathbb{E}[T_{\text{SPRT}}]$. In our case,

$$\Lambda_N = N\frac{\xi_0^2 - \xi_1^2}{2(1+\sigma^2)} + \frac{\xi_1 - \xi_0}{1+\sigma^2}\sum_{n=1}^{N} Y_n.$$

This likelihood ratio can be computed only when the distortion distribution is completely known. However, SPRT might be unaware of this distortion in many practical scenarios. We represent the algorithm by SPRT-MM in such cases and denote by $\mathbb{E}[T_{\text{SPRT-MM}}]$, $\mathbb{P}_{\text{FA}}^{\text{SPRT-MM}}$ and $\mathbb{P}_{\text{MD}}^{\text{SPRT-MM}}$, its average stopping time, false alarm and missed detection probabilities, resp. The thresholds for SPRT-MM are same as for SPRT, however, since it is unaware of the distortion, its log-likelihood is updated with $\sigma = 0$ in the above as

$$\Lambda_N = N\frac{\xi_0^2 - \xi_1^2}{2} + (\xi_1 - \xi_0)\sum_{n=1}^{N} Y_N. \tag{11}$$

### C. Comparison: T-SeqRDT, SeqRDT, BlockRDT and SPRT

We define $|\xi_1 - \xi_0|$ as the Signal-to-Noise Ratio (SNR). In Table I, we compare the average stopping times of T-*Seq*RDT, *Seq*RDT and SPRT to the block size of Block-RDT, for different SNR values, such that all the algorithms guarantee $\alpha = \beta = 0.001$. We choose $\tau^- = \sigma = 1$. Note from Table I that SPRT-MM is fastest but it is unable to guarantee $\alpha$ and $\beta$ as $\mathbb{P}_{\text{FA}}^{\text{SPRT-MM}}$ and $\mathbb{P}_{\text{MD}}^{\text{SPRT-MM}}$ does not stay below $\alpha$ and $\beta$, resp. This is a consequence of the model mismatch as SPRT-MM is unaware of the distortion distribution. SPRT is the fastest among the algorithms which guarantee $\alpha$ and $\beta$, but it comes at the cost of knowing precisely the complete distributions of the signal and the distortion. Among the algorithms *Block*RDT, *Seq*RDT and T-*Seq*RDT, *Seq*RDT is the fastest on average, especially at low SNR values, but needs most information (all $\tau^-$, $\tau$ and $\tau^+$) about the distorted signal. *Block*RDT is the slowest and to guarantee $\alpha$ and $\beta$, it requires the same information ($\tau$ and $\tau^+$) as T-*Seq*RDT. However, T-*Seq*RDT is considerably faster

on average in comparison and only marginally slower than *Seq*RDT, especially at low SNRs. Moreover, T-*Seq*RDT by design eliminates the need for buffer, $M$, whereas *Seq*RDT needs a buffer.

## VII. Conclusion and Perspectives

In this work, we proposed a novel framework for hypothesis testing. We can test arbitrary signals with unknown distributions, as long as the problem can be formulated in the form of (3) and $|\langle \Xi \rangle_N - \xi_0|$ is lower bounded away for $\tau$ (see Assumption 5.1). Instead of knowing the complete distribution of the signals, we need to know only a few parameters. We introduced new non-parametric algorithm, T-*Seq*RDT, for sequential testing. The algorithm is truncated version of *Seq*RDT [16], [17]. *Seq*RDT, although marginally faster than T-*Seq*RDT, needed more information about the signal. Moreover, *Seq*RDT guaranteed $\alpha$ and $\beta$ by use of a buffer, $M$, which waits for the observations to accumulate before starting the test. T-*Seq*RDT, by design eliminated the need for this buffer while guaranteeing $\alpha$ and $\beta$. This was made possible by theoretically choosing an appropriate truncation window and thresholds. Finally, simulations showed that T-*Seq*RDT, even with little knowledge of the signal, is able to provide sufficient performance guarantees while making a decision faster on average compared to *Block*RDT. Moreover, the algorithm was shown to be robust compared to SPRT. In conclusion, we believe that the proposed testing approach provides an appealing alternative to likelihood ratio based frameworks.

## References


[1] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses, 3rd edition*, Springer, 2005.

[2] D.J. Wheeler and D.S. Chambers, *Understanding Statistical Process Control*, SPC Press, 2010.

[3] J. Whitehead, *The Design and Analysis of Sequential Clinical Trials*, Statistics in Practice. Wiley, 1997.

[4] Thomas S. Ferguson, "Who solved the secretary problem?," *Statist. Sci.*, vol. 4, no. 3, pp. 282–289, 08 1989.

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

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

[7] Robert Bechhofer, "A note on the limiting relative efficiency of the wald sequential probability ratio test," *Journal of the American Statistical Association*, vol. 55, no. 292, pp. 660–663, 1960.

[8] S. Tantaratana and J. Thomas, "Relative efficiency of the sequential probability ratio test in signal detection," *IEEE Transactions on Information Theory*, vol. 24, no. 1, pp. 22–31, January 1978.

[9] Sawasd Tantaratana and John B. Thomas, "Truncated sequential probability ratio test," *Information Sciences*, vol. 13, no. 3, pp. 283 – 300, 1977.

[10] 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.

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

[12] V. P. Draglia, A. G. Tartakovsky, and V. V. Veeravalli, "Multihypothesis sequential probability ratio tests .I. asymptotic optimality," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2448–2461, Nov 1999.

[13] Tze Leung Lai, "Asymptotic optimality of invariant sequential probability ratio tests," *Ann. Statist.*, vol. 9, no. 2, pp. 318–333, 03 1981.

[14] Alexander Tartakovsky, "Asymptotically optimal sequential tests for nonhomogeneous processes," *Sequential Analysis*, vol. 17, no. 1, pp. 33–61, 1998.

[15] 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.

[16] P. Khanduri, D. Pastor, V. Sharma, and P.K. Varshney, "Sequential random distortion testing of non-stationary processes," *Manuscript submitted for publication*, 2018.

[17] P. Khanduri, D. Pastor, V. Sharma, and P.K. Varshney, "On sequential random distortion testing of non-stationary processes," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2018.

[18] 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.

[19] Dominique Pastor, "Random distortion testing with linear measurements," *Accepted for publication in Signal Processing*.