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, \( Z \), deviates by more than a specified tolerance, \( \tau \), from a fixed value, \( \zeta_0 \) [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 SeqRDT approach. We show that under mild assumptions on the signal, SeqRDT will follow the properties desired by a sequential test. Simulations show that the SeqRDT 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 (\( H_0 \)) and alternative (\( H_1 \)) hypotheses. The decision is generally made under the Bayesian, minimax or Neyman-Pearson frameworks. In 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 \( H_0 \) and stop the testing, or 2) accept \( H_1 \) 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 \( H_0 \) and \( H_1 \) correspond to two possible distributions for the observations. For details on Wald's approach the reader can refer to [5].

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 \( \zeta \in \mathbb{R} \) and any \( \sigma \in (0, \infty) \), \( Z \sim \mathcal{N}(\zeta, \sigma^2) \) means that \( Z \) is Gaussian distributed with mean \( \zeta \) and variance \( \sigma^2 \). In what follows, \( Q_{1/2} \) denotes the Generalized Marcum Function [9] with order 1/2. Basically, we have \( \mathbb{P}[|Z| > \eta] = Q_{1/2}(\eta, \eta) \) for any given \( Z \sim \mathcal{N}(\zeta, 1) \). Given \( \gamma \in (0, 1) \) and \( \rho \in [0, \infty) \), \( \lambda_{\gamma}(\rho) \) is defined as the unique solution

This work was supported by ARO grant W911NF-14-1-0339.
in $x$ to $Q_{1/2}(\rho, x) = \gamma [1, \text{Lemma 2, statement (ii)}]$, so that:

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

(1)

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

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

Two elements $U, V$ of $M(\Omega, \mathcal{F})^{N}$ 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 $M(\Omega, \mathcal{F})^{N}$. 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 $\Xi_n$ are not necessarily iid.

Proposition 3.2

Remark 1

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

Next, in Proposition 3.2, we study 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 $\tau \geq 0$, $T_{A_{\gamma}(\sqrt{\tau N})}$ satisfies the following asymptotic behavior for testing $H_0$ vs $H_1$ in (2):

(i) we have

$$\lim_{N \to \infty} P \left[ \frac{T_{A_{\gamma}(\sqrt{\tau N})}}{\sqrt{\tau N}} = 1 \right] = \frac{\gamma}{\gamma} \quad \text{under } H_0$$

(ii) under Assumption 3.1 we have,

$$\lim_{N \to \infty} P \left[ \frac{T_{A_{\gamma}(\sqrt{\tau N})}}{\sqrt{\tau N}} = 1 \right] = \frac{0}{1} \quad \text{under } H_1$$

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 $\gamma$. 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 SeqRDT approach for the mean testing problem defined in (2).

4. SEQUENTIAL TEST

Given any natural number $N \geq N_0$, we can completely specify a sequential test for $H_0$ against $H_1$ by defining the stopping time:

$$T = \min \{ N \in \mathbb{N} : D_M(N) \neq \infty \},$$

(6)

with:

$$D_M(1) = D_M(2) = \ldots = D_M(M) = \infty$$

for $N > M$, $D_M(N) = \begin{cases} 0 & \text{if } |(Y)_N - \xi_0| < \lambda_L(N), \\ \infty & \text{if } \lambda_L(N) < |(Y)_N - \xi_0| < \lambda_H(N), \\ 1 & \text{if } |(Y)_N - \xi_0| > \lambda_H(N). \end{cases}$

(7)

where the two thresholds $\lambda_L(N)$ and $\lambda_H(N)$ must be such that $\lambda_L(N) < \lambda_H(N)$. Note that $M$ is the number of samples SeqRDT 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:

\[
P_{FA}(D_M) \triangleq \mathbb{P}[D_M(T) = 1] \quad \text{under } \mathcal{H}_0
\]

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

\[
P_{MD}(D_M) \triangleq \mathbb{P}[D_M(T) = 0] \quad \text{under } \mathcal{H}_1
\]

The goal of any sequential algorithm is to design the thresholds so as to guarantee that \(P_{FA}(D_M)\) and \(P_{MD}(D_M)\) stay below certain pre-specified levels \(\alpha\) and \(\beta\), respectively.

Next, in Proposition 4.1 we show that the thresholds \(\lambda_H(N)\) and \(\lambda_L(N)\) designed according to (3) for levels \(\alpha\) and \(\beta\), respectively are indeed appropriate for SeqRDT.

**Proposition 4.1** For \(a, \beta \in (0, 1/2)\), \(\tau \in (0, \infty)\) and the thresholds

\[
\lambda_H(N) = \lambda_a(\tau \sqrt{N})/\sqrt{N} \quad \text{and} \quad \lambda_L(N) = \lambda_{1-a}(\tau \sqrt{N})/\sqrt{N}
\]

we have,

\[
\lambda_L(N) \leq \lambda_H(N),
\]

for all \(N \in \mathbb{N}\).

We thus know that the thresholds \(\lambda_H(N)\) and \(\lambda_L(N)\) satisfy the criterion \(\lambda_L(N) \leq \lambda_H(N)\). 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 \(a, \beta \in (0, 1/2)\), set \(\lambda_L(N) = \lambda_{1-c}(\tau \sqrt{N})/\sqrt{N}\) and

\[
\lambda_H(N) = \lambda_a(\tau \sqrt{N})/\sqrt{N}
\]

If Assumption 3.1 is satisfied, then:

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

(ii) \(\lim_{M \to \infty} P_{FA}(D_M) = \lim_{M \to \infty} P_{MD}(D_M) = 0\)

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 SeqRDT can achieve arbitrarily low \(P_{FA}(D_M)\) and \(P_{MD}(D_M)\) 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 SeqRDT.

**5. Experimental Results and Discussion**

In this section, we perform some simulations to highlight the advantages of SeqRDT 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 \(\xi_0\) and \(\xi_1\) 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 SeqRDT framework as defined in (2) with \(\tau = 0\) and \(N_0 = 1\) as

\[
\text{Observation: } Y = \xi_i + X \sim \mathcal{N}(\Omega, \Omega)^N, \text{ for } i = \{0, 1\}
\]

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 \(\xi_0\) or \(\xi_1\), 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 SeqRDT 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 \(\Xi_n = \xi_1 + \Delta_n\) under \(\mathcal{H}_1\) for \(i \in \{0, 1\}\) and for all \(n \in \mathbb{N}\), with \(|\xi_1 - \xi_0| > \tau\). Here \(\Delta_n\) model the possible additive distortions in the above deterministic model with unknown distribution. Let us assume that \(N_0 \geq \tau\) we have some positive value \(\tau\) such that \(P[|\Delta_N| < \tau] = 1\) and \(P[|\Delta_N| + \xi_1 - \xi_0| > \tau] = 1\) 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. SeqRDT, Block-RDT and SPRT**

For illustration, let us consider the distortions \(\Delta_n \sim \mathcal{N}(0, \sigma^2)\) for all \(n \in \mathbb{N}\). Consider the SeqRDT framework. For the choice of tolerance \(\tau = 2\sigma\Delta\), we have \(P[|\Delta_N| < \tau] = 0.9545\) for all \(N \geq 1\), and equality for \(N = 1\), i.e., \(P[|\Delta_1| < \tau] = 0.9545\). Moreover, we have \(P[|\Delta_N| + \xi_1 - \xi_0| > \tau] = 0.5\) for all \(N \geq 1\). This probability will be increasing in \(|\xi_1 - \xi_0|\). For example, if \(|\xi_1 - \xi_0| = 4\sigma\Delta\), we have \(P[|\Delta_N| + \xi_1 - \xi_0| > \tau] = 0.9772\). Here, we choose the buffer size \(M = N_0 - 1 = 0\). Although we do not have \(P[|\Delta_N| < \tau] = 1\) and \(P[|\Delta_N| + \xi_1 - \xi_0| > \tau] = 1\) 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

\[
\text{Observation: } Y = \xi_i + \Delta + X \sim \mathcal{N}(\Omega, \Omega)^N, \text{ for } i = \{0, 1\}
\]

with \(X_1, X_2, \ldots \sim \mathcal{N}(0, 1)\) and \(\Delta_n \sim \mathcal{N}(0, \sigma^2)\), respec-

\[
\text{Observation: } Y = \Xi + X \sim \mathcal{N}(\Omega, \Omega)^N, \text{ with } \Xi \sim \mathcal{N}(0, \sigma^2)
\]

Before proceeding further let us first analyze the invariance properties of the above problem for the Block-RDT framework. The
The table below shows the stopping times, the probability of false alarm and missed detection for different SDR values and for Block-RDT, SPRT and SPRT-MM. We denote the probability of false alarm as $P_{FA}$ and the probability of missed detection as $P_{MD}$. The thresholds are chosen as $\lambda_{FA}$ and $\lambda_{MD}$ that satisfy the constraints in (3) and [2]. We compare the number of samples taken by SeqRDT versus Block-RDT and SPRT-MM for different SDR values and for the same tolerance, $\tau = 2\sigma_{d}$ for all $n \in \mathbb{N}$. Here we do not have $|\Delta_n| \leq \tau$, however, we show via simulations that this does not impact the results below. The goal of Block-RDT is to design an $\alpha$ level test such that the probability of missed detection $P_{MD}$ stays below level $\beta$. The threshold is chosen to be $\lambda_{MD}(\sqrt{\tau}N)/\sqrt{\tau} N$ from (3) and [2]. We denote the probability of false alarm as $P_{FA}$ and the number of samples required to achieve $P_{FA}$ for different SDR values and for the conditions when the threshold is known and when the threshold is unknown.

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

<table>
<thead>
<tr>
<th>SDR $= \frac{SNR}{T}$ (dB)</th>
<th>6.02</th>
<th>7.96</th>
<th>9.54</th>
<th>12.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>SeqRDT Stopping time, $T$</td>
<td>3.98</td>
<td>3.28</td>
<td>3.04</td>
<td>2.90</td>
</tr>
<tr>
<td>$P_{FA}(\tau_{MD})$</td>
<td>$3.33 \times 10^{-4}$</td>
<td>$3.47 \times 10^{-4}$</td>
<td>$3.19 \times 10^{-4}$</td>
<td>$3.17 \times 10^{-4}$</td>
</tr>
<tr>
<td>$P_{MD}(\tau_{MD})$</td>
<td>$1.24 \times 10^{-4}$</td>
<td>$5 \times 10^{-6}$</td>
<td>$5 \times 10^{-6}$</td>
<td>0</td>
</tr>
<tr>
<td>Block-RDT Number of samples, $N_{B-RDT}$</td>
<td>14</td>
<td>7</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>$P_{FA}^{B-RDT}$</td>
<td>$9.45 \times 10^{-4}$</td>
<td>$3.12 \times 10^{-4}$</td>
<td>$2.44 \times 10^{-4}$</td>
<td>$2.2 \times 10^{-5}$</td>
</tr>
<tr>
<td>$P_{MD}^{B-RDT}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6.80 $\times 10^{-5}$</td>
</tr>
<tr>
<td>SPRT Stopping time, $T_{SPRT}$</td>
<td>2.44</td>
<td>1.73</td>
<td>1.34</td>
<td>1.05</td>
</tr>
<tr>
<td>$P_{FA}^{SPRT}$</td>
<td>$2.08 \times 10^{-4}$</td>
<td>$1.59 \times 10^{-4}$</td>
<td>$1.03 \times 10^{-4}$</td>
<td>$2.85 \times 10^{-5}$</td>
</tr>
<tr>
<td>$P_{MD}^{SPRT}$</td>
<td>$2.09 \times 10^{-4}$</td>
<td>$1.54 \times 10^{-4}$</td>
<td>$1.03 \times 10^{-4}$</td>
<td>$2.57 \times 10^{-5}$</td>
</tr>
<tr>
<td>SPRT-MM Stopping time, $T_{SPRT-MM}$</td>
<td>1.57</td>
<td>1.24</td>
<td>1.10</td>
<td>1.01</td>
</tr>
<tr>
<td>$P_{FA}^{SPRT-MM}$</td>
<td>$6.2 \times 10^{-3}$</td>
<td>$3.5 \times 10^{-3}$</td>
<td>$1.8 \times 10^{-3}$</td>
<td>$2.88 \times 10^{-4}$</td>
</tr>
<tr>
<td>$P_{MD}^{SPRT-MM}$</td>
<td>$6.2 \times 10^{-3}$</td>
<td>$3.6 \times 10^{-3}$</td>
<td>$1.8 \times 10^{-3}$</td>
<td>$3.05 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

5.3. Comparison

We define $|\xi_1 - \xi_0|$ as the Signal-to-Noise Ratio (SNR) and $|\xi_1 - \xi_0|$ as Signal-to-maximum-Distortion Ratio (SDR) [1, 2]. We analyze the average number of samples taken by SeqRDT compared to its fixed sample size counterpart Block-RDT, SPRT and SPRT-MM. We choose the threshold variance to be $\sigma_{d} = 1$. We average the stopping times and count the probability of false alarm and missed detection over $10^4$ Monte Carlo iterations for SeqRDT, 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 SeqRDT versus Block-RDT and SPRT-MM for different SDR values and for the cases when the threshold is known and when the threshold is unknown.

6. CONCLUSION AND FUTURE WORK

In this work, we proposed a sequential algorithm SeqRDT 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


