On Sequential Random Distortion Testing of Non-Stationary Processes

— Source link ☑

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

Institutions: Syracuse University, Indian Institute of Science

Related papers:

- Truncated Sequential Non-Parametric Hypothesis Testing Based on Random Distortion Testing

- Sequential Tests of Statistical Hypotheses

- Sequential detection with limited memory

- Sequential hypothesis testing with off-line randomized sensor selection strategy

- Sequential Analysis in High Dimensional Multiple Testing and Sparse Recovery

Sequential Random Distortion Testing of Non-Stationary Processes

Prashant Khanduri, Dominique Pastor, Vinod Sharma, Pramod Varshney

Sequential Random Distortion Testing of Non-Stationary Processes

Prashant Khanduri, *Student Member, IEEE* Dominique Pastor, *Member, IEEE*, Vinod Sharma, *Senior Member, IEEE* and Pramod K. Varshney, *Life Fellow, IEEE*

Abstract—In this work, we propose a non-parametric sequential hypothesis test based on random distortion testing (RDT). RDT addresses the problem of testing whether or not a random signal, Ξ, observed in independent and identically distributed (i.i.d) additive noise deviates by more than a specified tolerance, τ, from a fixed model, ξ_0. The test is non-parametric in the sense that the underlying signal distributions under each hypothesis are assumed to be unknown. The need to control the probabilities of false alarm (PFA) and missed detection (PMD), while reducing the number of samples required to make a decision, leads to a novel sequential algorithm, *Seq*RDT. We show that under mild assumptions on the signal, *Seq*RDT follows the properties desired by a sequential test. We introduce the concept of a buffer and derive bounds on PFA and PMD, from which we choose the buffer size. Simulations show that *Seq*RDT leads to faster decision-making on an average compared to its fixed-sample-size (FSS) counterpart, *Block*RDT. These simulations also show that the proposed algorithm is robust to model mismatches compared to the sequential probability ratio test (SPRT).

Index Terms—Sequential testing, non-parametric testing, robust hypothesis testing, sequential probability ratio test (SPRT)

I. INTRODUCTION

Over the past few years, with the availability of large amounts of data, the need to devise novel but simple inference techniques has surged [2], [3]. Moreover, it is not always possible to fit the data to predefined models or distribution. Therefore, it is important to devise methods that are simple to implement, model-independent and robust [4], [5]. In this work, we propose one such algorithm, *Seq*RDT, for binary hypothesis testing.

In standard binary hypothesis testing problems, fixed-sample-size (FSS) tests, based on a fixed number of observations, decide between two possible statistical hypotheses, the so-called null ($\mathcal{H}_0$) and alternative ($\mathcal{H}_1$) hypotheses [6]. The decision is generally made under the Bayesian, minimax or Neyman-Pearson frameworks. However, many decision problems are inherently sequential in nature, where observations are collected sequentially [7]–[9]. In his seminal

works [10], [11], Wald developed a likelihood ratio based sequential approach namely the sequential probability ratio test (SPRT). SPRT is optimal in the sense that it is the fastest on an average compared to all other sequential and FSS tests that guarantee the same probabilities of false alarm (PFA) and missed detection (PMD) [10], [11]. However, this optimality is lost if there is a mismatch between the assumed and the true signal models or the hypotheses to be tested are composite [7], [11]–[14]. SPRT and other composite hypothesis testing approaches are extensions of likelihood theory in that they assume 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 [7], [15], [16]. However, in practice, prior knowledge or good models for the distributions under each hypothesis are often not available. This is all the more detrimental as likelihood ratio tests are not robust to uncertainty or model mismatch. Moreover, many approaches in sequential testing make stationarity or independent and identically distributed (i.i.d.) assumptions on the observed process under each hypothesis [7], [17]. Such assumptions are questionable in practice. SPRT and standard sequential testing solutions in the literature aimed at relaxing stationarity or i.i.d assumptions are still based on likelihood ratio tests [7] and, as such, may suffer from the lack of robustness of these tests. Therefore, it is important to devise robust tests that assume little knowledge of the underlying signal distributions.

To overcome these limitations, in this work, we propose one such algorithm, *Seq*RDT, which is based on an alternative sequential binary hypothesis testing problem formulation. We assume that the received observation is a real random process Y parameterized by ξ and the hypothesis testing problem considered is $\mathcal{H}_0 : \xi = \xi_0$ vs $\mathcal{H}_1 : \xi \neq \xi_0$. We argue that, in many cases, testing the signal exactly for ξ_0 might be too strict due to measurement errors, environmental fluctuations other than noise and other factors as mentioned in [18]. It is reasonable to allow for some distortions around ξ_0 and test for the parameter in the neighborhood of ξ_0. Therefore, we model the hypothesis test as Y being a noisy version of a random process Ξ, which is a distorted version of ξ_0 with unknown distribution. The hypothesis testing problem is then stated as

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

with $\tau \in [0, \infty)$ representing the distortion. The design of the hypotheses in (1) implies that we test for the distorted signal in the small neighborhood of ξ_0. This problem is referred to as random distortion testing (RDT) and was first proposed in [18].

The authors in [18] assumed that the signal, Ξ, is embedded in Gaussian noise and proposed the optimal single sample test for the problem. The RDT problem was extended for hypothesis testing with multiple samples in [19], in the form of *Block*RDT, where the authors replaced the signal Ξ in (1) by its empirical mean. However, the need for faster decision-making and the inherent sequential nature of many problems lead us to define a novel RDT based framework for sequential testing, which yields a new sequential algorithm, *Seq*RDT. In this work, we show that with a few key assumptions, *Seq*RDT guarantees pre-specified PFA and PMD. Importantly, we show that *Seq*RDT has an almost surely finite stopping time. Moreover, we introduce the notion of a buffer that makes it possible to control the probabilities of error. The theoretical analysis, the control over the error probabilities and the notion of buffer were not present in the past works [20]. A preliminary version without any analysis was presented in [1] and a truncated version of *Seq*RDT, T-*Seq*RDT, has been developed in [21], [22], where the sequential test is forced to stop if a decision is not made until a certain time. Beyond biomedical signal processing applications [20], the proposed theoretical framework can be of interest in many real world applications, including communications, radar, fault-detection and structural health monitoring. Below, we summarize the major contributions of the work:

- We propose a new RDT based framework for non-parametric sequential hypothesis testing and introduce a new sequential algorithm, *Seq*RDT, to solve the binary hypothesis testing problem.
- We analyze the properties of the thresholds, asymptotic error probabilities and introduce the notion of a buffer which helps in controlling the error probabilities of *Seq*RDT.
- Without any prior knowledge of the signal distribution, *Seq*RDT is shown to guarantee pre-specified PFA and PMD, whereas, in contrast, the likelihood ratio based tests need precise knowledge of the signal distributions under each hypothesis.
- Via numerical simulations, the proposed approach is shown to be faster on an average compared to its FSS counterpart, *Block*RDT. More importantly, *Seq*RDT is shown to be robust to mismatches compared to the likelihood ratio based tests.

Next, before discussing notations, in Section II we state the problem and discuss the assumptions. In Section III we discuss the test statistic. We introduce *Seq*RDT in Section IV and analyze PFA and PMD in Section V. In Section VI, the assumptions made in the previous sections are relaxed. The performance of *Seq*RDT is discussed in Section VII. Finally, Section VIII concludes the paper.

Notation: $\mathbb{N}$ is the set of natural numbers and $\mathbb{R}$ that of real numbers. Given $N \in \mathbb{N}$, $\mathbb{R}^N$ is the vector space of all N-dimensional row vectors with real components. The components of $x \in \mathbb{R}^N$ are denoted by $x_1, x_2, \ldots, x_N$ and we write $x = (x_1, x_2, \ldots, x_N)$.

All the random variables are defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote the set of all real random variables

defined on $(\Omega, \mathcal{F})$ by $\mathcal{M}(\Omega, \mathbb{R})$. Given $U \in \mathcal{M}(\Omega, \mathbb{R})$, $\mathbb{P}_U$ is the probability distribution of U: $\mathbb{P}_U(B) = \mathbb{P}\big[U \in B\big]$ with $\big[U \in B\big] = \{\omega \in \Omega : U(\omega) \in B\}$ when B is a Borel set of $\mathbb{R}$. A domain $\mathcal{D}$ of U is any Borel set B of $\mathbb{R}$ such that $\mathbb{P}_U(B) = 1$. Given $\zeta \in [0, \infty)$, $Z \sim \mathcal{U}(\zeta)$ means that Z is uniformly distributed in $[-\zeta, +\zeta]$. Given $\xi \in \mathbb{R}$ and $\sigma \in [0, \infty)$, $Z \sim \mathcal{N}(\xi, \sigma^2)$ means that Z is Gaussian distributed with mean ξ and variance σ^2. The generalized Marcum function [23] with order $1/2$ is denoted by $Q_{1/2}$. For any $Z \sim \mathcal{N}(\xi, 1)$, we have [24, Eq. (19)] and Remark V.3]:

$$\mathbb{P}\big[\,|Z| > \eta\,\big] = Q_{1/2}(|\xi|, \eta). \tag{2}$$

It follows that, for any $(a, b) \in [0, \infty) \times [0, \infty)$,

$$Q_{1/2}(a, b) = 1 - \Phi(b - a) + \Phi(-b - a) \tag{3}$$

where Φ is the cumulative distribution function (cdf) of Z. $Q_{1/2}$ increases with its first argument and decreases with its second [23]. Given $\gamma \in (0, 1)$ and $\rho \in [0, \infty)$, $\lambda_\gamma(\rho)$ is defined as the unique solution in x to $Q_{1/2}(\rho, x) = \gamma$ [18, Lemma 2, statement (ii)], so that:

$$Q_{1/2}(\rho, \lambda_\gamma(\rho)) = \gamma. \tag{4}$$

The set of all sequences defined on $\mathbb{N}$ (resp. $[\![1, N]\!] = \{1, 2, \ldots, N\}$) and valued in $\mathcal{M}(\Omega, \mathbb{R})$ is denoted by $\mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}$ (resp. $\mathcal{M}(\Omega, \mathbb{R})^{[\![1,N]\!]}$). Given U in $\mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}$ (resp. $U \in \mathcal{M}(\Omega, \mathbb{R})^{[\![1,N]\!]}$), the realization of U at $n \in \mathbb{N}$ (resp. $n \in [\![1, N]\!]$) is called a sample of U and denoted by U_n. Each U_n is an element of $\mathcal{M}(\Omega, \mathbb{R})$. Given $N \in \mathbb{N}$, $\langle U \rangle_N = (1/N) \sum_{n=1}^{N} U_n$ is the sample mean of U over the N samples $U_1, \ldots, U_N$. We say that $U, V \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}$ (resp. $\mathcal{M}(\Omega, \mathbb{R})^{[\![1,N]\!]}$) are independent if any two finite subsequences $(U_{i_1}, U_{i_2}, \ldots, U_{i_k})$ and $(V_{j_1}, V_{j_2}, \ldots, V_{j_r})$ are independent for $i_1 < \ldots < i_k$ and $j_1 < \ldots < j_r$. We denote by $\mathbb{1}_A(x)$ the indicator function of the set A.

II. PROBLEM STATEMENT AND ASSUMPTIONS

We model the random mixture of a distorted signal of interest and possible interferences by some $\Xi \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}$. We suppose that Ξ is observed in additive and independent noise X with i.i.d samples and unknown cdf $\mathbb{F}$. The observation process is Y with $Y_n = \Xi_n + X_n$ for any $n \in \mathbb{N}$ and we write $Y = \Xi + X$. We assume that Ξ under the null hypothesis (H_0) is generated from an underlying joint distribution $\mathcal{P}_{\xi_0}$. Under the alternative hypothesis (H_1), Ξ is supposed to be generated by any underlying arbitrary joint distribution other than $\mathcal{P}_{\xi_0}$, i.e., $\Xi \not\sim \mathcal{P}_{\xi_0}$. No assumption is made on the stationarity or the distribution of Ξ. The samples Ξ_n are thus not necessarily i.i.d. We summarize the problem as:

$$\begin{cases} \underline{\text{Observation}} : Y = \Xi + X \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}} \\ \text{with} \begin{cases} \Xi, X \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}, \\ X_1, X_2, \ldots \overset{\text{i.i.d}}{\sim} \mathbb{F}, \; \mathbb{F} \text{ unknown.} \end{cases} \\ \begin{cases} H_0 : & \Xi = (\Xi_n)_{n \in \mathbb{N}} \sim \mathcal{P}_{\xi_0}, \\ H_1 : & \Xi = (\Xi_n)_{n \in \mathbb{N}} \not\sim \mathcal{P}_{\xi_0} \end{cases} \end{cases} \tag{5}$$

The problem in its current form (5) is difficult to tackle since we do not assume any knowledge of the signal distributions under the two hypotheses. Therefore, we associate with

each hypothesis a non-parametric distance related criterion independent of these possible distributions as motivated in Section I. This criterion then serves as a surrogate to the actual hypotheses to be tested. This idea was first proposed in the form of RDT in [18] and was later extended for multiple samples by *Block*RDT [19] as discussed earlier in the Introduction. We now present the model more specifically.

In our formulation, Ξ models distortion other than noise around a fixed model ξ_0. For instance, such distortion can result from interferences as encountered in radar, sonar and telecommunication systems. We, however, expect that, for N sufficiently large, the empirical mean $\langle \Xi \rangle_N$ remains close to ξ_0 under H_0 and drifts significantly away from ξ_0 under H_1. We can then formalize the foregoing by quantifying the possible drift between $\langle \Xi \rangle_N$ and ξ_0 via $|\langle \Xi \rangle_N - \xi_0|$, introducing a tolerance $\tau \in [0, \infty)$ on this distance to specify the maximal acceptable drift and assuming the existence of some $N_0 \in \mathbb{N}$ such that, either $|\langle \Xi \rangle_N - \xi_0| \leqslant \tau$ for any $N \geqslant N_0$ or $|\langle \Xi \rangle_N - \xi_0| > \tau$ for any $N \geqslant N_0$. The former (resp. latter) hypothesis above becomes the null (resp. alternative) hypothesis $\mathcal{H}_0$ (resp. $\mathcal{H}_1$) in the surrogate model for (5). Given the observation Y, the *Seq*RDT problem addressed in this paper is then the sequential testing of $\mathcal{H}_0$ against $\mathcal{H}_1$, written as:

$$
\begin{cases}
\underline{\text{Observation}} : Y = \Xi + X \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}} \\
\text{with} \begin{cases} \Xi, X \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}, \\ X_1, X_2, \ldots \overset{\text{i.i.d}}{\sim} \mathbb{F}, \mathbb{F} \text{ unknown.} \\ \mathbb{E}\left[X_1 \right] = 0, \mathrm{Var}\left(X_1 \right) = 1 \text{ and } \mathbb{E}\left[X_1^3 \right] < \infty \\ \Xi \text{ and } X \text{ are independent.} \end{cases} \\
\exists N_0 \in \mathbb{N}, \begin{cases} \mathcal{H}_0 \; \forall N \geqslant N_0, 0 \leqslant |\langle \Xi \rangle_N - \xi_0| \leqslant \tau \text{ (a-s)} \\ \mathcal{H}_1 \; \forall N \geqslant N_0, \tau < |\langle \Xi \rangle_N - \xi_0| \leqslant \tau_H \text{ (a-s)} \end{cases}
\end{cases} \quad (6)
$$

where, throughout the paper and without recalling it, $\tau \in [0, \infty)$. The above testing model encompasses the *Block*RDT model [19] for a fixed sample size, N, and $X \sim \mathcal{N}(0, 1)$. Here, N_0 and the tolerances τ and τ_H are known *a priori*, based on some prior knowledge about the signal. The idea is that, although the exact distributions in play are unknown, the user's experience on some representative data may be sufficient to roughly bound the behavior of the empirical mean with respect to ξ_0. Discussion of procedures suitable for extracting such knowledge is beyond the scope of this work because they depend on the targeted application. The generic question is how far we can get if prior knowledge on the process is limited to such bounds.

Remark 1: The above problem (6) tests whether the deviation of the signal mean $\langle \Xi \rangle_N$ around ξ_0 is below (or above) a specified tolerance τ for the null hypothesis (or the alternate hypothesis) to be true. As indicated above, this non-parametric criterion then serves as a surrogate to the complete knowledge of the signal distributions and thus avoids their prior knowledge. Likelihood ratio based tests are, therefore, not feasible for the above problem. ∎

Example 1 (Change-in-mean testing): The *Seq*RDT problem (6) embraces the testing of a change in the mean of a Gaussian process [7] when, given two known real values ξ_0 and ξ_1, $\Xi_n = \xi_0$ for all $n \in \mathbb{N}$ under $\mathcal{H}_0$, $\Xi_n = \xi_1$ for all $n \in \mathbb{N}$ under $\mathcal{H}_1$, $\tau = 0$, $N_0 = 1$ and $X \sim \mathcal{N}(0, 1)$ in (6). ∎

Example 2 (Bounded regime testing): Given $\xi \in \mathbb{R}$ and $h \in [0, \infty)$, say that Ξ follows the (bounded) regime (ξ, h) and write $\Xi \sim (\xi, h)$ if, for any $N \in \mathbb{N}$, $|\langle \Xi \rangle_N - \xi| \leqslant h$. A sufficient condition for $\Xi \sim (\xi, h)$ is that $|\Xi_n - \xi| \leqslant h$ (a-s) for any $n \in \mathbb{N}$. Suppose that Ξ satisfies either $\mathcal{H}_0^* : \Xi \sim (\xi_0, h_0)$, where the regime (ξ_0, h_0) is given, or $\mathcal{H}_1^* : \Xi \sim (\xi_1, h_1)$, where (ξ_1, h_1) is any possibly unknown regime other than (ξ_0, h_0). Say that the regimes (ξ_0, h_0) and (ξ_1, h_1) are separate if $|\xi_1 - \xi_0| \geqslant h_0 + h_1$, which amounts to assuming that $(\xi_0 - h_0, \xi_0 + h_0) \cap (\xi_1 - h_1, \xi_1 + h_1) = \emptyset$. When (ξ_0, h_0) and (ξ_1, h_1) are separate, testing $\mathcal{H}_0^*$ against $\mathcal{H}_1^*$ is the particular problem (6) with $h_0 \leqslant \tau < |\xi_1 - \xi_0| - h_1$, $\tau_H \geqslant |\xi_1 - \xi_0| + h_1$ and $N_0 = 1$. ∎

Remark 2: Since our results hold for any random signal sequence, Ξ, they also hold for any deterministic sequence. The two statements are actually equivalent, as can be seen by simple conditioning on Ξ. The reader can thus think of Ξ as being deterministic and arbitrary. Clearly, no independence assumptions on Ξ are required. The assumptions that really matter are those on the noise sequence, X. Indeed, since the main quantity appearing in $\mathcal{H}_0$ and $\mathcal{H}_1$ is the empirical average of the distorted signal Ξ, which we observe up to some i.i.d. noise, we can reduce the problem to a Gaussian problem through the Berry-Esseen theorem.

A. Assumptions

To solve problem (6) sequentially, we introduce two additional assumptions. The first one is asymptotic and can be regarded as a weak notion of ergodicity. The second one concerns the case of finite sample sizes. Both assumptions are used below to state different results. Their use depends on the available amount of prior information on the process.

Assumption 2.1 ((a-s) convergence of $\langle \Xi \rangle_N$):

$$
\exists \langle \Xi \rangle_\infty \in \mathcal{M}(\Omega, \mathbb{R}), \lim_{N \to \infty} \langle \Xi \rangle_N = \langle \Xi \rangle_\infty \text{ (a-s)}. \quad (7)
$$

for which exist $\tau^- \in [0, \tau)$ and $\tau^+ \in (\tau, \infty)$ such that:

$$
\begin{cases}
\text{Under } \mathcal{H}_0 : |\langle \Xi \rangle_\infty - \xi_0| \leqslant \tau^- \text{ (a-s)}, \\
\text{Under } \mathcal{H}_1 : |\langle \Xi \rangle_\infty - \xi_0| \geqslant \tau^+ \text{ (a-s)}.
\end{cases}
$$

Remark 3: Assumption 2.1 is automatically satisfied if Ξ is stationary and ergodic [25, Ch. 4, Sec. 24]. In this case, there exists $\xi \in \mathbb{R}$ such that $\mathbb{E}\left[\Xi_n \right] = \xi$ for every $n \in \mathbb{N}$, so that Assumption 2.1 holds with $\langle \Xi \rangle_\infty = \xi$ and $\xi \in \{\xi_0, \xi_1\}$. ∎

Basically, Assumption 2.1 will prove helpful to characterize the relevance of the sequential procedure introduced in Section IV (see Theorem 4.2) in asymptotic conditions. The next assumption is aimed at establishing additional results in non-asymptotic situations.

Assumption 2.2 (Bounded behavior of $|\langle \Xi \rangle_N - \xi_0|$):
There exist $\tau^- \in [0, \tau)$ and $\tau^+ \in (\tau, \infty)$ such that:

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

Remark 4: At this stage, it is crucial to emphasize the significance of Assumption 2.1 and Assumption 2.2, as well as the differences between them with respect to the two hypotheses in (6). Assumption 2.1 is asymptotic in nature, whereas

Assumption 2.2 is not. Moreover, Assumption 2.2 does not require the existence of an (a-s) convergence of the empirical mean in contrast to Assumption 2.1. The two assumptions will be helpful to better control the behavior, specifically, PFA and PMD of the sequential test. This better control will actually be rendered possible via the strict inequalities between τ^- and τ, on the one hand, and between τ^+ and τ, on the other hand. By so proceeding, $|\langle \Xi \rangle_N - \xi_0|$ is kept away from τ, under $\mathcal{H}_0$ and $\mathcal{H}_1$. The decision will then turn out to be all the more reliable as τ^- and τ^+ drift away from τ, which can be seen as the frontier between the two hypotheses in (6). Note also that, if Assumption 2.2 and (7) hold true together, $|\langle \Xi \rangle_\infty - \xi_0|$ remains bounded away from τ and all the properties stated under Assumptions 2.1 and 2.2 add to guarantee an even better grip on the behavior of the sequential testing. ∎

Remark 5: It must be noted that the Assumption 2.2 was not required in both RDT [18] and its FSS version, *Block*RDT [19]. The tests proposed in these works were designed to guarantee PFA below a pre-specified level, while guaranteeing a minimal PMD for FSS tests, without any control over this probability. As already emphasized in the previous remark, Assumption 2.2 gives the algorithm designer control over both PFA and PMD for FSS tests, as well as for the sequential testing framework proposed in this work. ∎

We now give two simple examples to illustrate Assumption 2.1 and Assumption 2.2.

Example 3: (i) Consider a random variable U with unknown distribution. We assume, under $\mathcal{H}_0$: $|U| \leqslant \tau^-$ and under $\mathcal{H}_1$: $|U| \geqslant \tau^+$. Suppose further that $\Xi_n = U + \Delta_n$, for $n \in \mathbb{N}$ when the Δ_ns are i.i.d with zero mean and unknown distribution. This translates to Assumption 2.1 with $\xi_0 = 0$ with the application of the strong law of large numbers.
(ii) Consider a random variable $U_n \sim \mathcal{N}(\xi_i, 1)$ for $n \in \mathbb{N}$ and $i \in \{0, 1\}$. We assume, under $\mathcal{H}_0$: $\xi = \xi_0$ and under $\mathcal{H}_1$: $\xi \neq \xi_0$. Without loss of generality we assume $\xi_0 = 0$ and if we have $|\xi_1| \geqslant \tau^+ > \tau^-$. If $\Xi_n = \frac{1}{n}\sum_{i=1}^{n} U_i$, Assumption 2.1 is verified. Moreover, the above process is a non-stationary Markov process under both hypotheses and a variant of the random walk. Similarly, many examples of practical interest can be framed in terms of Assumption 2.2 also, with little prior knowledge of the underlying signals. This will become clearer in the experimental section. ∎

Next we discuss the structure of the tests used to design *Seq*RDT and their asymptotic properties.

III. TEST STATISTIC

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

$$\mathfrak{T}_{N,\gamma}(x) = \begin{cases} 0 \text{ if } |\langle x \rangle_N - \xi_0| \leqslant \lambda_\gamma(\tau\sqrt{N})/\sqrt{N} \\ 1 \text{ otherwise} \end{cases}. \quad (8)$$

Proposition 3.1 below describes the asymptotic behavior of such tests under Assumption 2.1. These tests play a crucial role in the design of *Seq*RDT for the problem stated in (6).

Throughout, $\delta_N = c\,\mathbb{E}\left[\,|X_1 - \mathbb{E}[X_1]|^3\,\right] \mathrm{Var}(X_1)^{-3/2} N^{-1/2}$ and $(2\pi)^{-1/2} \leqslant c < 0.8$ as specified in Lemma A.1.

Proposition 3.1: Given $\gamma \in (0, 1)$ and $\tau \geqslant 0$, $\mathfrak{T}_{N,\gamma}$ satisfies the following asymptotic behavior for testing $\mathcal{H}_0$ against $\mathcal{H}_1$

in (6):
(i) we have

under $\mathcal{H}_0$: $\quad \lim_{N\to\infty} \mathbb{P}[\mathfrak{T}_{N,\gamma}(Y) = 1] \leqslant \gamma,$ (9)

under $\mathcal{H}_1$: $\quad \lim_{N\to\infty} \mathbb{P}[\mathfrak{T}_{N,\gamma}(Y) = 0] \leqslant 1 - \gamma.$ (10)

(ii) under Assumption 2.1, we have,

$$\lim_{N\to\infty} \mathbb{P}[\mathfrak{T}_{N,\gamma}(Y) = 1] = \begin{cases} 0 & \text{under } \mathcal{H}_0 \\ 1 & \text{under } \mathcal{H}_1 \end{cases}. \quad (11)$$

Proof: Proof of statement (i): From (8) and Lemma A.1,

$$\mathbb{P}[\mathfrak{T}_{N,\gamma}(Y) = 1]$$
$$\leqslant \mathbb{E}\left[Q_{1/2}\left(\sqrt{N}|\langle \Xi \rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\right)\right] + 2\delta_N. \quad (12)$$

Therefore, under $\mathcal{H}_0$ and for any $N \geqslant N_0$, we have:

$$\mathbb{P}[\mathfrak{T}_{N,\gamma}(Y) = 1] \leqslant Q_{1/2}\left(\tau\sqrt{N}, \lambda_\gamma(\tau\sqrt{N})\right) + 2\delta_N.$$

According to (4), the upper-bound in the second inequality above equals $\gamma + 2\delta_N$. As N grows to ∞, δ_N vanishes and (9) follows.

We prove (10) similarly. We begin by combining (8) and Lemma A.1 to get

$$\mathbb{P}[\mathfrak{T}_{N,\gamma}(Y) = 0]$$
$$\leqslant 1 - \mathbb{E}\left[Q_{1/2}\left(\sqrt{N}|\langle \Xi \rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\right)\right] + 2\delta_N. \quad (13)$$

It then suffices to particularize (13) under $\mathcal{H}_1$ for $N \geqslant N_0$ and make N grow to ∞ to complete the proof.

Proof of statement (ii): Under $\mathcal{H}_0$ and Assumption 2.1, $\lim_N |\langle \Xi \rangle_N - \xi_0| = |\langle \Xi \rangle_\infty - \xi_0| \leqslant \tau^- < \tau$ (a-s). It then follows from Lemma A.2 that:

$$\lim_N Q_{1/2}\left(\sqrt{N}|\langle \Xi \rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\right) = 0 \quad \text{(a-s)}.$$

We then derive from (12) and the Lebesgue dominated convergence theorem that, under $\mathcal{H}_0$:

$$\lim_N \mathbb{P}[\mathfrak{T}_{N,\gamma}(Y) = 1] = 0.$$

Similarly, under $\mathcal{H}_1$ and Assumption 2.1 $\lim_N |\langle \Xi \rangle_N - \xi_0| = |\langle \Xi \rangle_\infty - \xi_0| \geqslant \tau^+ > \tau$ (a-s). Lemma A.2 induce that

$$\lim_N Q_{1/2}\left(\sqrt{N}|\langle \Xi \rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\right) = 1 \quad \text{(a-s)}.$$

By injecting this equality into (13) and using the Lebesgue dominated convergence theorem again, we obtain:

$$\lim_{N\to\infty} \mathbb{P}[\mathfrak{T}_{N,\gamma}(Y) = 1] = 1.$$

under $\mathcal{H}_1$, which concludes the proof. ∎

Proposition 3.1 enhances the interest of the tests defined in (8). Proposition 3.1 (i) implies that with the use of only one threshold, $\lambda_\gamma(\tau\sqrt{N})/\sqrt{N}$, PFA is guaranteed to stay below γ, but PMD is only guaranteed to be below $1 - \gamma$. Therefore, with only one threshold, we can design a test which simply controls PFA, without any control over PMD. In contrast, with the use of two thresholds along with Assumption 2.2,

the designer can control both PFA and PMD of the algorithm. Moreover, one of these thresholds should be small enough to diminish PFA. In contrast, the other one should be sufficiently high so as to make PMD small. Such a strategy naturally leads to a sequential approach. Also, Proposition 3.1 (ii) highlights the importance of Assumption 2.1 in achieving arbitrarily low PFA and PMDs for large but fixed sample sizes. However, we need to control the number of samples, which again highlights the need for a sequential approach. As a matter of fact, with the thresholds designed according to (8), we can design a sequential test capable of reducing the decision-making time for the testing problem defined in (6), while guaranteeing certain performance levels. This sequential approach is described below.

IV. ALGORITHM: *Seq*RDT

Section III motivates a sequential approach involving two thresholds designed using (8). One of these thresholds must guarantee an upper-bounded PFA, while the other aims at upper-bounding PMD. Given any natural number $M \geqslant N_0 - 1$, the sequential procedure *Seq*RDT suggested by Proposition 3.1 for testing $\mathcal{H}_0$ against $\mathcal{H}_1$ in (6) is specified by defining the stopping time:

$$T = \min \big\{ N \in \mathbb{N} : \mathcal{D}_M(N) \neq \infty \big\}, \qquad (14)$$

with:
$$\begin{cases} \mathcal{D}_M(1) = \mathcal{D}_M(2) = \ldots = \mathcal{D}_M(M) = \infty, \\[2pt] \text{for } N > M, \mathcal{D}_M(N) = \begin{cases} 0 & \text{if} \quad |\langle Y \rangle_N - \xi_0| \leqslant \lambda_L(N), \\ \infty & \text{if} \quad \lambda_L(N) < |\langle Y \rangle_N - \xi_0| \leqslant \lambda_H(N), \\ 1 & \text{if} \quad |\langle Y \rangle_N - \xi_0| > \lambda_H(N), \end{cases} \end{cases}$$
$$\qquad (15)$$

with $\lambda_L(N) = \lambda_\gamma(\tau\sqrt{N})/\sqrt{N}$, $\lambda_H(N) = \lambda_{\gamma'}(\tau\sqrt{N})/\sqrt{N}$, $\tau \in (0, \infty)$ and $\gamma, \gamma' \in (0, 1)$ must be such that $\gamma' < \gamma$, which implies $\lambda_L(N) < \lambda_H(N)$. Here, $\mathcal{D}_M(N)$ represents the decision variable with $\mathcal{D}_M(N) = 0$ equivalent to saying that $\mathcal{H}_0$ is decided, $\mathcal{D}_M(N) = 1$ equivalent to saying that $\mathcal{H}_1$ is decided and $\mathcal{D}_M(N) = \infty$ equivalent to saying that no decision is made at the Nth sample and that the algorithm will update the statistic and repeat the test with the $(N + 1)^{\text{th}}$ sample. Note that M is the number of samples *Seq*RDT waits for before starting the test. We refer to this M as the buffer size. An appropriate M can be chosen based on some elementary knowledge of the signal. This will be made clearer in the coming section and Section VII.

The choice for γ and γ' can be figured out as follows. The PFA of *Seq*RDT is:

$$\mathbb{P}_{\text{FA}}(\mathcal{D}_M) \stackrel{\text{def}}{=} \mathbb{P}\big[\mathcal{D}_M(T) = 1\big] \quad \text{under } \mathcal{H}_0. \qquad (16)$$

In the same way, the PMD is:

$$\mathbb{P}_{\text{MD}}(\mathcal{D}_M) \stackrel{\text{def}}{=} \mathbb{P}\big[\mathcal{D}_M(T) = 0\big] \quad \text{under } \mathcal{H}_1. \qquad (17)$$

Since the goal of the sequential algorithm is to guarantee $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ below certain pre-specified levels α and β, respectively, Proposition 3.1 leads us to choose $\gamma = 1 - \beta$ and $\gamma' = \alpha$ with $\alpha, \beta \in (0, 1/2)$. This assumption is required to ensure $\lambda_L(N) < \lambda_H(N)$ (refer to Proposition 4.1 (i)). Moreover, typical values of α and β are of the order of

10^{-1} to 10^{-4}, so the assumption is not particularly restrictive. Henceforth, we always assume $\alpha, \beta \in (0, 1/2)$ and set:

$$\lambda_L(N) = \lambda_{1-\beta}(\tau\sqrt{N})/\sqrt{N}, \, \lambda_H(N) = \lambda_\alpha(\tau\sqrt{N})/\sqrt{N}. \quad (18)$$

Next, we analyze the properties of the proposed thresholds.

A. Threshold properties

Proposition 4.1 below validates that such thresholds are appropriate for *Seq*RDT under both asymptotic and non-asymptotic conditions.

Proposition 4.1: We have:
(i) $\forall N \in \mathbb{N}$, $\lambda_L(N) < \lambda_H(N)$,
(ii) $\lambda_H(N)$ is decreasing in $N \in \mathbb{N}$ and lower bounded by τ,
(iii) for N large enough, the threshold $\lambda_L(N)$ is increasing in N and upper bounded by τ,
(iv) both thresholds approach τ as N increases:

$$\lim_{N\to\infty} \lambda_H(N) = \lim_{N\to\infty} \lambda_L(N) = \tau.$$

Proof: Apply Lemmas B.4, B.5 and B.6. ∎

Proposition 4.1 (i) & (ii) imply that, as $N \to \infty$ and thus, in vanishing noise, the test will resemble a non-sequential test as both thresholds become equal to τ.

The question addressed now is then *"Can this choice of thresholds give some performance guarantees?"*.

Before stating several theorems to answer this question, we establish the following straightforward inequalities, which will prove useful at several places in the sequel. With the same notation as above, for any given $\varepsilon \in \{0, 1\}$, we have:

$$\mathbb{P}\big[\mathcal{D}_M(T) = \varepsilon\big] = \mathbb{P}\big(\big[\mathcal{D}_M(T) = \varepsilon\big] \cap \big[T \geq M + 2\big]\big) \\ + \mathbb{P}\big[\mathcal{D}_M(M + 1) = \varepsilon\big]. \qquad (19)$$

Now using $[\mathcal{D}_M(T) = \epsilon] \subset [\mathcal{D}_{M+1}(M + 1) \neq 1 - \epsilon]$

$$\mathbb{P}\big[\mathcal{D}_M(T) = \varepsilon\big] \leqslant 1 - \mathbb{P}\big[\mathcal{D}_M(M + 1) = 1 - \varepsilon\big]. \qquad (20)$$

Theorem 4.2 (Asymptotics: T, $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$): If $\alpha, \beta \in (0, 1/2)$ and Assumption 2.1 holds true, then:
(i) $\mathbb{P}\big[T < \infty\big] = 1$ under both $\mathcal{H}_0$ and $\mathcal{H}_1$;
(ii) $\lim_{M\to\infty} \mathbb{P}_{\text{FA}}(\mathcal{D}_M) = \lim_{M\to\infty} \mathbb{P}_{\text{MD}}(\mathcal{D}_M) = 0$.

Proof: Proof of statement (i): We have $[T = \infty]$ if and only if $\mathcal{D}_M(N) = \infty$ for each $N > M$. Therefore, $\mathbb{P}[T = \infty] \leqslant \mathbb{P}[\mathcal{D}_M(N) = \infty]$ for any $N \geqslant M + 1$. Since

$$\big[\mathcal{D}_M(N) = \infty\big] = \big[\lambda_L(N) < |\langle \Xi \rangle_N + \langle X \rangle_N - \xi_0| \leqslant \lambda_H(N)\big],$$

we have:

$$\mathbb{P}\big[\mathcal{D}_M(N) = \infty\big] \\ = \mathbb{P}\Big[\mathcal{T}_{\lambda_{1-\beta}(\tau\sqrt{N})/\sqrt{N}}(Y) = 1\Big] - \mathbb{P}\Big[\mathcal{T}_{\lambda_\alpha(\tau\sqrt{N})/\sqrt{N}}(Y) = 1\Big].$$

According to Proposition 3.1(ii), $\lim_{N\to\infty} \mathbb{P}\big[\mathcal{D}_M(N) = \infty\big] = 0$. Hence the result.

Proof of statement (ii): The PFA is $\mathbb{P}_{\text{FA}}(\mathcal{D}_M) = \mathbb{P}\big[\mathcal{D}_M(T) = 1\big]$ under $\mathcal{H}_0$. According to (20), $\mathbb{P}_{\text{FA}}(\mathcal{D}_M) \leqslant 1 - \mathbb{P}\big[\mathcal{D}_M(M + 1) = 0\big]$. The right hand side (rhs) in this equality can be rewritten $\mathbb{P}\big[|\langle Y \rangle_{M+1} - \xi_0| > \lambda_L(M + 1)\big]$. It follows from (18) and Lemma A.1 that, regardless of Assumption 2.1:

$$\mathbb{P}_{\text{FA}}(\mathcal{D}_M) \leqslant \mathbb{E}\left[Q_{1/2}\left(\sqrt{M+1}|\langle\Xi\rangle_{M+1} - \xi_0|, \lambda_{1-\beta}(\tau\sqrt{M+1}) \right) \right]$$
$$+ \ 2\delta_{M+1}. \qquad (21)$$

We then derive from Assumption 2.1 & Lemma A.2 that, under $\mathcal{H}_0$,

$$\lim_{M \to \infty} Q_{1/2}\left(\sqrt{M+1}|\langle\Xi\rangle_{M+1} - \xi_0|, \lambda_{1-\beta}(\tau\sqrt{M+1}) \right) = 0 \quad \text{(a-s)}.$$

The Lebesgue dominated convergence theorem then implies that $\lim_{M \to \infty} \mathbb{P}_{\text{FA}}(\mathcal{D}_M) = 0$.

Similarly, we derive from (20), (18) and Lemma A.1 that, regardless of Assumption 2.1:

$$\mathbb{P}_{\text{MD}}(\mathcal{D}_M) \leqslant$$
$$1 - \mathbb{E}\left[Q_{1/2}\left(\sqrt{M+1}|\langle\Xi\rangle_{M+1} - \xi_0|, \lambda_{\alpha}(\tau\sqrt{M+1}) \right) \right] + 2\delta_{M+1}.$$
$$(22)$$

It then suffices to apply Assumption 2.1, Lemma A.2 and the Lebesgue dominated convergence theorem to obtain the second equality in (ii). ∎

The above theorem implies that under Assumption 2.1, i.e., if the empirical mean of the signal centered around ξ_0 converges away from τ, the sequential test (15) takes a decision in finite time with probability one. The theorem also implies that PFA and PMD diminish with the increasing buffer size, M. Next, we give some performance guarantees for the non-asymptotic regime. In this regard, the next theorem shows that without any assumption on the signal model, the bounds on PFA and PMD can be loose. Therefore, we make use of Assumption 2.2 to derive tighter bounds on PFA and PMD and use these bounds to choose an appropriate buffer size, M, such that PFA and PMD can be controlled. We derive two bounds in the next section.

V. ANALYSIS: PFA AND PMD

In this section, we derive bounds on PFA and PMD of *Seq*RDT in the next two theorems.

Theorem 5.1 (Non-Asymptotics: $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$): $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ are bounded as:

$$\begin{cases} Q_{1/2}\left(0, \lambda_\alpha(\tau\sqrt{M+1})\right) - 2\delta_{M+1} \\ \qquad\qquad \leqslant \ \mathbb{P}_{\text{FA}}(\mathcal{D}_M) \ \leqslant \ 1 - \beta + 2\delta_{M+1}, \\[1ex] 1 - Q_{1/2}\left(\tau_H\sqrt{M+1}, \lambda_{1-\beta}(\tau\sqrt{M+1})\right) - 2\delta_{M+1} \\ \qquad\qquad \leqslant \ \mathbb{P}_{\text{MD}}(\mathcal{D}_M) \ \leqslant \ 1 - \alpha + 2\delta_{M+1}. \end{cases}$$

Proof: Under $\mathcal{H}_0$, we derive from (15), (16), (18), (19) & Lemma A.1 that:

$$\mathbb{P}_{\text{FA}}(\mathcal{D}_M) \geqslant \mathbb{P}\left[\mathcal{D}_M(M+1) = 1\right] \geqslant$$
$$\mathbb{E}\left[Q_{1/2}\left(\sqrt{M+1}|\langle\Xi\rangle_{M+1} - \xi_0|, \lambda_{\alpha}(\tau\sqrt{M+1}) \right) \right] - 2\delta_{M+1}.$$
$$(23)$$

The bounds on $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ result from the inequalities satisfied by $|\langle\Xi\rangle_{M+1} - \xi_0|$ under $\mathcal{H}_0$ and (23), for the lower bound, and (21) along with (4), for the upper bound.

Similarly, for the probability of missed detection, under $\mathcal{H}_1$, (15), (17), (18), (19) & Lemma A.1 yield

$$\mathbb{P}_{\text{MD}}(\mathcal{D}_M) \geqslant \mathbb{P}\left[\mathcal{D}_M(M+1) = 0\right]$$
$$\geqslant 1 - \mathbb{E}\left[Q_{1/2}\left(\sqrt{M+1}|\langle\Xi\rangle_{M+1} - \xi_0|, \lambda_{1-\beta}(\tau\sqrt{M+1}) \right) \right]$$
$$- 2\delta_{M+1}. \quad (24)$$

We obtain the bounds on $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ from the inequalities satisfied by $|\langle\Xi\rangle_{M+1} - \xi_0|$ under $\mathcal{H}_1$ and (24), for the lower bound, and (22) along with (4), for the upper bound. ∎

The lower bounds for $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ derived in Theorem 5.1 always stay below levels α and β, respectively. Theorem 5.1 also states that, without any assumption, the upper bounds on $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ although bounded by unity, are loose. Hence, by assuming further knowledge of the signal through Assumption 2.2 and noise to be Gaussian distributed, which is of practical interest in many applications and which induces $\delta_k = 0$ for all $k \in \mathbb{N}$, we derive tighter upper bounds on $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ in the next theorem. These bounds will be used to choose appropriate buffer sizes for *Seq*RDT.

Theorem 5.2 (Non-Asymptotics: $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$): (i) Under Assumption 2.2 and Gaussian noise, $X \sim \mathcal{N}(0,1)$, $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ are bounded as:

$$\begin{cases} Q_{1/2}\left(0, \lambda_\alpha(\tau\sqrt{M+1})\right) \\ \qquad\qquad \leqslant \ \mathbb{P}_{\text{FA}}(\mathcal{D}_M) \ \leqslant \ \text{UB1}_{\text{FA}}(M) \bigwedge \text{UB2}_{\text{FA}}(M), \\[1ex] 1 - Q_{1/2}\left(\tau_H\sqrt{M+1}, \lambda_{1-\beta}(\tau\sqrt{M+1})\right) \\ \qquad\qquad \leqslant \ \mathbb{P}_{\text{MD}}(\mathcal{D}_M) \ \leqslant \ \text{UB1}_{\text{MD}}(M) \bigwedge \text{UB2}_{\text{MD}}(M). \end{cases} \quad (25)$$

where $a_1 \bigwedge a_2 = \min(a_1, a_2)$ for $a_1, a_2 \in \mathbb{R}$. $\text{UB1}_{\text{FA}}(M)$, $\text{UB2}_{\text{FA}}(M)$, $\text{UB1}_{\text{MD}}(M)$ and $\text{UB2}_{\text{MD}}(M)$ are finite and are given in (26) (27), (28) and (29), respectively.
(ii) $\text{UB1}_{\text{FA}}(M) \bigwedge \text{UB2}_{\text{FA}}(M)$ and $\text{UB1}_{\text{FA}}(M) \bigwedge \text{UB2}_{\text{FA}}(M)$ decrease with M.

Proof: Proof of statement (i): When Assumption 2.2 holds true, $0 \leqslant |\langle\Xi\rangle_{M+1} - \xi_0| \leqslant \tau^-$ under $\mathcal{H}_0$. Injecting these inequalities into (23) and (21) and using $\delta_k = 0$ for all $k \in \mathbb{N}$ for Gaussian noise yields the bounds:

$$Q_{1/2}\left(0, \lambda_\alpha(\tau\sqrt{M+1})\right) \leqslant \mathbb{P}_{\text{FA}}(\mathcal{D}_M) \leqslant \text{UB1}_{\text{FA}}(M).$$

We obtain $\text{UB2}_{\text{FA}}(M)$ by writing first:

$$[\mathcal{D}_M(T) = 1] = [\mathcal{D}_M(M+1) = 1]$$
$$\bigcup_{N=M+2}^{\infty} \left([\mathcal{D}_M(N) = 1] \cap [\mathcal{D}_M(K) = \infty, \forall K \in [\![M+1, N-1]\!]] \right).$$

Now using the union bound and from the Frechet inequality it follows that:

$$\mathbb{P}\left[\mathcal{D}_M(T) = 1\right] \leqslant \mathbb{P}\left[\mathcal{D}_M(M+1) = 1\right]$$
$$+ \sum_{N=M+2}^{\infty} \mathbb{P}\left[\mathcal{D}_M(N) = 1\right] \bigwedge \left(\bigwedge_{K=M+1}^{N-1} \mathbb{P}\left[\mathcal{D}_M(K) = \infty\right] \right).$$
$$(30)$$

For any $N \geqslant M+1$, Lemma A.1 and Assumption 2.2 along with $\delta_k = 0$ for all $k \in \mathbb{N}$ imply that under $\mathcal{H}_0$;

$$\mathbb{P}\left[\mathcal{D}_M(N) = 1\right] \leqslant Q_{1/2}\left(\tau^- \sqrt{N}, \lambda_\alpha(\tau\sqrt{N})\right). \qquad (31)$$

$$\text{UB1}_{\text{FA}}(M) = Q_{1/2}\left(\tau^-\sqrt{M+1}, \lambda_{1-\beta}(\tau\sqrt{M+1})\right), \tag{26}$$

$$\text{UB2}_{\text{FA}}(M) = Q_{1/2}\left(\tau^-\sqrt{M+1}, \lambda_{\alpha}(\tau\sqrt{M+1})\right)$$
$$+ \sum_{N=M+2}^{\infty}\left[\left(Q_{1/2}\left(\tau^-\sqrt{N}, \lambda_{\alpha}(\tau\sqrt{N})\right)\right)\bigwedge\left(\bigwedge_{K=M+1}^{N-1}\left(Q_{1/2}\left(\tau^-\sqrt{K}, \lambda_{1-\beta}(\tau\sqrt{K})\right) - Q_{1/2}\left(0, \lambda_{\alpha}(\tau\sqrt{K})\right)\right)\right)\right], \tag{27}$$

$$\text{UB1}_{\text{MD}}(M) = 1 - Q_{1/2}\left(\sqrt{M+1}\,\tau^+, \lambda_{\alpha}(\tau\sqrt{M+1})\right), \tag{28}$$

$$\text{UB2}_{\text{MD}}(M) = 1 - Q_{1/2}\left(\tau^+\sqrt{M+1}, \lambda_{1-\beta}(\tau\sqrt{M+1})\right)$$
$$+ \sum_{N=M+2}^{\infty}\left[\left(1 - Q_{1/2}\left(\tau^+\sqrt{N}, \lambda_{1-\beta}(\tau\sqrt{N})\right)\right)\bigwedge\left(\bigwedge_{K=M+1}^{N-1}\left(Q_{1/2}\left(\tau_H\sqrt{K}, \lambda_{1-\beta}(\tau\sqrt{K})\right) - Q_{1/2}\left(\tau^+\sqrt{K}, \lambda_{\alpha}(\tau\sqrt{K})\right)\right)\right)\right]. \tag{29}$$

For any $K \in [\![M+1, N-1]\!]$, we can write:

$$\mathbb{P}\left[\mathcal{D}_M(K) = \infty\right] = \mathbb{P}\left[\langle Y\rangle_K - \xi_0| \geqslant \lambda_L(K)\right]$$
$$- \mathbb{P}\left[\langle Y\rangle_K - \xi_0| > \lambda_H(K)\right]. \tag{32}$$

Again from Lemma A.1, Assumption 2.2 and $\delta_k = 0$ for all $k \in \mathbb{N}$ under $\mathcal{H}_0$ we have:

$$\mathbb{P}\left[\mathcal{D}_M(K) = \infty\right] \leqslant Q_{1/2}\left(\tau^-\sqrt{K}, \lambda_{1-\beta}(\tau\sqrt{K})\right)$$
$$- Q_{1/2}\left(0, \lambda_{\alpha}(\tau\sqrt{K})\right). \tag{33}$$

The bound $\text{UB2}_{\text{FA}}(M)$ follows by injecting (31) and (33) into (30). The bounds $\text{UB1}_{\text{MD}}(M)$ and $\text{UB2}_{\text{MD}}(M)$ on $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ follow similarly from (22), (24), (32) and Lemma A.1, under $\mathcal{H}_1$ and Assumption 2.2. Because of space limitations, the details are left to the reader. The proof of the convergence of $\text{UB1}_{\text{MD}}(M)$ and $\text{UB2}_{\text{MD}}(M)$ are given in Appendix C.

Proof of statement (ii): For any given $\rho \in [0, \infty)$ such that $\rho \neq \tau$ and for all $\gamma \in (0,1)$, it follows from Lemma B.3 that the map $N \in \mathbb{N} \mapsto Q_{1/2}\left(\rho\sqrt{N}, \lambda_{\gamma}(\tau\sqrt{N})\right)$ is decreasing if $\rho < \tau$ and increasing if $\rho > \tau$. Therefore, $\text{UB1}_{\text{FA}}(M)$ and $\text{UB1}_{\text{MD}}(M)$ also decrease with M. A careful inspection of $\text{UB2}_{\text{FA}}(M)$ and $\text{UB2}_{\text{MD}}(M)$ reveals that each term involved in these bounds is decreasing with M. Statement (ii) follows since the minimum of two decreasing terms is decreasing. ∎

Theorem 5.2 makes it possible to choose the least buffer size M guaranteeing specified values for the upper bounds $\text{UB1}_{\text{FA}}(M)\bigwedge\text{UB2}_{\text{FA}}(M)$ and $\text{UB1}_{\text{MD}}(M)\bigwedge\text{UB2}_{\text{MD}}(M)$. Therefore, with the choice of an appropriate buffer size M, we can expect to control $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ under desired levels. More precisely, if we want a test that guarantees $\mathbb{P}_{\text{FA}}(\mathcal{D}_M) \leqslant \alpha$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M) \leqslant \beta$ for specified $0 < \alpha < 1/2$ and $0 < \beta < 1/2$, we can choose an appropriate M as follows. First, choose M_1 such that $\text{UB1}_{\text{FA}}(M_1)\bigwedge\text{UB2}_{\text{FA}}(M_1) \leqslant \alpha$. Afterwards, choose M_2 such that $\text{UB1}_{\text{MD}}(M_2)\bigwedge\text{UB2}_{\text{MD}}(M_2) \leqslant \beta$. The buffer size can then be fixed to $M = \max(M_1, M_2)$. In Section VII, we will proceed in this manner to choose the buffer size. It is, however, important to emphasize that the upper bounds given in Theorem 5.2 could still be loose in some scenarios, as the terms $\text{UB2}_{\text{FA}}(M)$ and $\text{UB2}_{\text{MD}}(M)$ are derived from the intersection of multiple events. However, according to Theorem 5.2 (i) and (25) these bounds will always stay below $\text{UB1}_{\text{FA}}(M)$ and $\text{UB1}_{\text{MD}}(M)$ even if $\text{UB2}_{\text{FA}}(M)$ and

$\text{UB2}_{\text{MD}}(M)$ are loose. Moreover, $\text{UB1}_{\text{FA}}(M)$ and $\text{UB1}_{\text{MD}}(M)$ are always guaranteed to stay below one.

VI. AN EXTENSION

Suppose that, instead of (6) where the inequalities are assumed to be satisfied in (a-s) sense, we have $\mathcal{H}_0^* : \mathbb{P}\left[\forall N \geqslant N_0, |\langle\Xi\rangle_N - \xi_0| \leqslant \tau\right] \geqslant 1 - \varepsilon$ vs $\mathcal{H}_1^* : \mathbb{P}\left[\forall N \geqslant N_0, |\langle\Xi\rangle_N - \xi_1| \geqslant \tau\right] \geqslant 1 - \varepsilon$ with a small positive constant $\varepsilon \leqslant \min(\alpha, \beta)$. Under the assumptions of Theorem 5.2, *Seq*RDT can still be used as follows to test $\mathcal{H}_0^*$ against $\mathcal{H}_1^*$ with guaranteed bounds on PFA and PMD. Indeed, given $\alpha, \beta \in (0, 1/2)$, choose M so that $\text{UB1}_{\text{FA}}(M)\bigwedge\text{UB2}_{\text{FA}}(M) \leqslant \alpha$ and $\text{UB1}_{\text{MD}}(M)\bigwedge\text{UB2}_{\text{MD}}(M) \leqslant \beta$ in (25). Under $\mathcal{H}_0^*$, the PFA, $\mathbb{P}_{\text{FA}}^*$, of *Seq*RDT satisfies:

$$\mathbb{P}_{\text{FA}}^* = \mathbb{P}\left[\mathcal{D}_M(T) = 1\right] \leqslant \mathbb{P}(\Omega_0^c) + \mathbb{P}\left[\mathcal{D}_M(T) = 1|\Omega_0\right]\mathbb{P}(\Omega_0)$$
$$\leqslant \varepsilon + \mathbb{P}\left[\mathcal{D}_M(T) = 1|\Omega_0\right]\mathbb{P}(\Omega_0),$$

with $\Omega_0 = \left[\forall N \geqslant N_0, |\langle\Xi\rangle_N - \xi_0| \leqslant \tau\right]$. Consider the probability space $\left(\Omega_0, \mathcal{F}_{\Omega_0}, \mathbb{P}(\bullet|\Omega_0)\right)$, where $\mathcal{F}_{\Omega_0}$ is the trace σ-algebra of $\mathcal{F}$ on Ω_0 and $\mathbb{P}(\bullet|\Omega_0)$ is the conditional probability that assigns to each $A \in \mathcal{F}_{\Omega_0}$ the probability $\mathbb{P}(A|\Omega_0)$. According to Theorem 5.2, $\mathbb{P}\left[\mathcal{D}_M(T) = 1|\Omega_0\right] \leqslant \alpha - \varepsilon$ and thus $\mathbb{P}_{\text{FA}}^* \leqslant \alpha$. Similarly, we have $\mathbb{P}_{\text{MD}}^* \leqslant \beta$. This easy extension is very useful in practice. For example, Ξ can be any i.i.d (or not) random variable sequence with unbounded support under $\mathcal{H}_0$ and $\mathcal{H}_1$ and such that $\langle\Xi\rangle_N \to \xi_0$ under $\mathcal{H}_0$ and $\langle\Xi\rangle_N \to \xi_1$ under $\mathcal{H}_1$, with $\xi_0 \neq \xi_1$.

VII. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we perform some simulations to highlight the advantages of *Seq*RDT compared to *Block*RDT [19] and SPRT [10], [11]. We first present the detection problem considered to carry out the experiments, then we outline each algorithm and finally, carry out the comparison of the presented algorithms for different parameter values.

A. The Gaussian mean detection problem

In the classical Gaussian mean detection problem [7], the observation is $Y_n = \Xi_n + X_n$, with $X_n \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$ for $n \in \mathbb{N}$, $\Xi_n = \xi_0$ under $\mathcal{H}_0$ and $\Xi_n \neq \xi_0$ under $\mathcal{H}_1$, where ξ_0 is a deterministic constant. This problem can be formulated in the *Seq*RDT framework (6) with $\tau = 0$ and $N_0 = 1$. But, in many applications, there might be a mismatch between the model
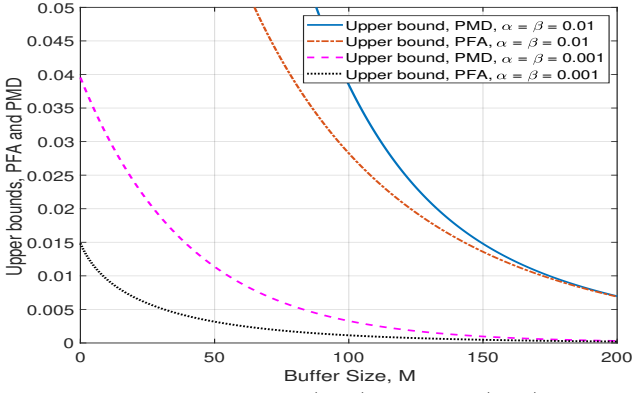
Fig. 1: Upper bound on $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ vs M for N curtailed to 500 in Theorem 5.2.

and the actual signal observed in practice. In reality, under either hypothesis, the actual signal would not be a constant assumed by the model, but rather a perturbed version of this constant. Such unavoidable perturbation may be hard to model in a parametric setup and likelihood ratio based tests can fail to guarantee reliable performance in the presence of model mismatches [18]. However, the *Block*RDT setup [19] and the *Seq*RDT setup as given in (6) and (15) are not limited by these drawbacks. Therefore, instead of dealing with the somewhat unrealistic model 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 $n \in \mathbb{N}$. Here, Δ_n models possible additive distortion with unknown distribution. In what follows, we compare different algorithms for mean testing of the distorted signal Ξ. Specifically, we consider two cases: first the bounded regime (bounded distortion) case introduced in Example 2 and then, the unbounded regime case (the distortion is not bounded). In the second case, we show that, even if the assumptions of *Seq*RDT are not strictly satisfied, *Seq*RDT will still be able to provide sufficient performance guarantees.

B. Bounded Distortion

We consider the case where there exist positive real values h_0, h_1 and H such that $2h_0 < h_1 < H$ and $\forall n \in \mathbb{N}$, $|\Delta_n| \leqslant h_0$ and $h_1 \leqslant |\xi_1 - \xi_0| \leqslant H$. We thus have $|\langle \Delta \rangle_N| \leqslant h_0$ and $h_0 < h_1 - h_0 < |\Delta_n + \xi_1 - \xi_0| \leqslant h_0 + H$, which imply:

$$\begin{cases} \mathcal{H}_0 : \forall N \geqslant 1, 0 \leqslant |\langle \Xi \rangle_N - \xi_0| \leqslant h_0, \text{ (a-s)} \\ \mathcal{H}_1 : \forall N \geqslant 1, h_0 < h_1 - h_0 \leqslant |\langle \Xi \rangle_N - \xi_0| \leqslant h_0 + H \text{ (a-s)}. \end{cases} \tag{34}$$

This makes it possible to test the distorted signal using *Seq*RDT with tolerances $\tau^- = h_0$, $\tau^+ = h_1 - h_0$, $\tau \in (\tau^-, \tau^+)$, $\tau_H \in [h_0 + H, \infty)$ and $N_0 = 1$. This is all *Seq*RDT needs to know to conduct the test, irrespective of the distortion distribution (please see Assumption 2.2).

For experimental illustration, we first choose $\Delta_n \sim \mathcal{U}(h_0)$, $(n \in \mathbb{N})$, with $h_1 \in (2h_0, |\xi_1 - \xi_0| - h_0]$ and $H \in [|\xi_1 - \xi_0| + h_0, \infty)$. This distortion distribution has to be fully known for SPRT, whereas *Seq*RDT and *Block*RDT are unaware of it and need to know only a few parameters. We now discuss the algorithms in the considered bounded regime.

$\alpha = \beta = 0.01$					
SNR $= \|\xi_1 - \xi_0\|$		0.4	0.5	0.6	0.8
*Block*RDT, using τ	$N_{\text{B-RDT}}$	2165	542	241	87
*Block*RDT, using τ^-	$N_{\text{B-RDT}}$	542	241	136	61
*Seq*RDT, Bounded Regime	$\mathbb{E}[T]$	167.28	142.83	137.31	136.78
*Seq*RDT, Unbounded Regime	$\mathbb{E}[T]$	171.34	145.83	141.12	140.26

TABLE I: *Seq*RDT vs *Block*RDT.

SeqRDT: By referring to (34) and Example 2, one potential choice for the tolerances in the *Seq*RDT framework can be $\tau^- = h_0$, $\tau^+ = 3h_0$, $\tau = 2h_0$ and $\tau_H = |\xi_1 - \xi_0| + h_0$. With this choice, $\mathcal{H}_0$ and $\mathcal{H}_1$ specify bounded and separate regimes with $N_0 = 1$. The choice of these values is not rigid. These values were chosen for ease of presentation and illustration purposes, but infinitely many others could have been considered to yet guarantee bounded and separate regimes. Note also that we need not know the distortion distribution to design the tolerances. These tolerances can be learned over time or can be available *a priori* to the user, as discussed earlier in Section II.

BlockRDT: In the *Block*RDT framework, problem (6) is tested for a fixed number of samples $N \geqslant N_0$. Specifically, we have N observation samples as $Y = \Xi + X \in \mathcal{M}(\Omega, \mathbb{R})^{[\![1,N]\!]}$ with $X_N \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$. A solution to this problem is proposed in [19] & [26]. 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

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

and $\mathcal{T}$ is said to have level $\gamma \in (0, 1)$ if $\alpha_{\mathcal{T}} \leqslant \gamma$. No Uniformly Most Powerful (UMP) test with level γ exists for *Block*RDT. By UMP test with level γ, we mean an N-dimensional test $\mathcal{T}^*$ such that $\alpha_{\mathcal{T}^*} \leqslant \gamma$ and $\mathbb{P}\left[\mathcal{T}^*(Y) = 1 \,\big|\, |\langle \Xi \rangle_N - \xi_0| > \tau\right] \geqslant \mathbb{P}\left[\mathcal{T}(Y) = 1 \,\big|\, |\langle \Xi \rangle_N - \xi_0| > \tau\right]$ for any N-dimensional test $\mathcal{T}$ and any $\Xi \in \mathcal{M}(\Omega, \mathbb{R})^{[\![1,N]\!]}$. We thus exhibit the subclass of *Block*RDT-coherent tests [26], among which a "best" test exists. An N-dimensional test $\mathcal{T}$ is *Block*RDT-coherent if:
[Invariance in mean] For all $y, y' \in \mathbb{R}^N$, if $\langle y \rangle_N = \langle y' \rangle_N$, then $\mathcal{T}(y) = \mathcal{T}(y')$
[Constant conditional power] For all $\Xi \in \mathcal{M}(\Omega, \mathbb{R})^{[\![1,N]\!]}$ independent of X, there exists a domain $\mathcal{D}$ of $|\langle \Xi \rangle_N - \xi_0|$ such that, for any $\rho \in \mathcal{D} \cap (0, \infty)$, $\mathbb{P}\left[\mathcal{T}(Y) = 1 \,\big|\, |\langle \Xi \rangle_N - \xi_0| = \rho\right]$ is independent of $\mathbb{P}_{|\langle \Xi \rangle_N - \xi_0|}$.

The rationale behind **[Invariance in mean]** is straightforward. **[Constant conditional power]** encapsulates the idea that $\mathcal{T}$ should not yield different results for different distributions of $|\langle \Xi \rangle_N - \xi_0|$, conditionally to $|\langle \Xi \rangle_N - \xi_0| = \rho$. The notion of domain is required for mathematical correctness.

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

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

$\alpha = \beta = 0.01$					
SNR $= \lvert \xi_1 - \xi_0 \rvert$		0.4	0.5	0.6	0.8
*Seq*RDT	$\mathbb{E}[T]$	167.28	142.83	137.31	136.78
	$\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$	2.1×10^{-4}	2.3×10^{-4}	3.1×10^{-4}	1.8×10^{-4}
	$\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$	2.4×10^{-4}	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$
SPRT	$\mathbb{E}[T_{\mathrm{SPRT}}]$	59.78	38.91	27.48	15.94
	$P_{\mathrm{FA}}^{\mathrm{SPRT}}$	7.8×10^{-3}	7.7×10^{-3}	6.7×10^{-3}	5.8×10^{-3}
	$P_{\mathrm{MD}}^{\mathrm{SPRT}}$	7.5×10^{-3}	7.5×10^{-3}	7.4×10^{-3}	6.7×10^{-3}
SPRT-MM	$\mathbb{E}[T_{\mathrm{SPRT\text{-}MM}}]$	59.62	38.74	27.37	15.89
	$P_{\mathrm{FA}}^{\mathrm{SPRT\text{-}MM}}$	1.2×10^{-2}	8.4×10^{-3}	7.9×10^{-3}	6.9×10^{-3}
	$P_{\mathrm{MD}}^{\mathrm{SPRT\text{-}MM}}$	1.1×10^{-2}	8.0×10^{-3}	7.5×10^{-3}	7.4×10^{-3}
$\alpha = \beta = 0.05$					
SNR $= \lvert \xi_1 - \xi_0 \rvert$		0.4	0.5	0.6	0.8
*Seq*RDT	$\mathbb{E}[T]$	123.13	112.77	111.63	110.54
	$\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$	9.1×10^{-4}	1.7×10^{-4}	7.2×10^{-4}	7.4×10^{-4}
	$\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$	3.5×10^{-4}	6×10^{-5}	$<10^{-5}$	$<10^{-5}$
SPRT	$\mathbb{E}[T_{\mathrm{SPRT}}]$	36.94	24.32	17.29	10.26
	$P_{\mathrm{FA}}^{\mathrm{SPRT}}$	3.9×10^{-2}	3.8×10^{-2}	3.6×10^{-2}	3.1×10^{-2}
	$P_{\mathrm{MD}}^{\mathrm{SPRT}}$	4.0×10^{-2}	3.8×10^{-2}	3.5×10^{-2}	3.3×10^{-2}
SPRT-MM	$\mathbb{E}[T_{\mathrm{SPRT\text{-}MM}}]$	36.74	24.23	17.24	10.22
	$P_{\mathrm{FA}}^{\mathrm{SPRT\text{-}MM}}$	5.8×10^{-2}	4.0×10^{-2}	3.8×10^{-2}	3.6×10^{-2}
	$P_{\mathrm{MD}}^{\mathrm{SPRT\text{-}MM}}$	5.2×10^{-2}	4.9×10^{-2}	3.8×10^{-2}	3.4×10^{-2}

TABLE II: *Seq*RDT vs SPRT for bounded regime.

$\alpha = \beta = 0.01$					
SNR $= \lvert \xi_1 - \xi_0 \rvert$		0.4	0.5	0.6	0.8
*Seq*RDT	$\mathbb{E}[T]$	171.34	145.83	141.12	140.26
	$\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$	1.5×10^{-4}	2×10^{-4}	2×10^{-4}	2.1×10^{-4}
	$\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$	4.4×10^{-4}	1.3×10^{-4}	4×10^{-5}	$<10^{-5}$
SPRT	$\mathbb{E}[T_{\mathrm{SPRT}}]$	68.74	44.74	31.52	18.27
	$P_{\mathrm{FA}}^{\mathrm{SPRT}}$	7.5×10^{-3}	7.6×10^{-3}	6.6×10^{-3}	6.6×10^{-3}
	$P_{\mathrm{MD}}^{\mathrm{SPRT}}$	7.5×10^{-3}	8.2×10^{-3}	7.4×10^{-3}	6.3×10^{-4}
SPRT-MM	$\mathbb{E}[T_{\mathrm{SPRT\text{-}MM}}]$	58.98	38.40	27.18	15.84
	$P_{\mathrm{FA}}^{\mathrm{SPRT\text{-}MM}}$	1.5×10^{-2}	1.46×10^{-2}	1.3×10^{-2}	1.2×10^{-2}
	$P_{\mathrm{MD}}^{\mathrm{SPRT\text{-}MM}}$	1.5×10^{-2}	1.4×10^{-2}	1.4×10^{-2}	1.2×10^{-2}
$\alpha = \beta = 0.001$					
SNR $= \lvert \xi_1 - \xi_0 \rvert$		4	5	6	8
*Seq*RDT	$\mathbb{E}[T]$	252.55	198.51	185.17	181.80
	$\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$	3×10^{-5}	1×10^{-5}	4×10^{-5}	2×10^{-5}
	$\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$	1.1×10^{-4}	3×10^{-5}	1×10^{-5}	$<10^{-5}$
SPRT	$\mathbb{E}[T_{\mathrm{SPRT}}]$	103.34	66.80	46.78	26.80
	$P_{\mathrm{FA}}^{\mathrm{SPRT}}$	9.6×10^{-4}	5.9×10^{-4}	5.3×10^{-4}	6.6×10^{-4}
	$P_{\mathrm{MD}}^{\mathrm{SPRT}}$	8.8×10^{-4}	9.2×10^{-4}	7.0×10^{-4}	5.7×10^{-4}
SPRT-MM	$\mathbb{E}[T_{\mathrm{SPRT\text{-}MM}}]$	89.24	57.82	40.56	23.32
	$P_{\mathrm{FA}}^{\mathrm{SPRT\text{-}MM}}$	2.2×10^{-3}	1.9×10^{-3}	1.8×10^{-3}	1.6×10^{-3}
	$P_{\mathrm{MD}}^{\mathrm{SPRT\text{-}MM}}$	2.3×10^{-3}	2.0×10^{-3}	1.9×10^{-3}	1.8×10^{-4}

TABLE III: *Seq*RDT vs SPRT for unbounded regime.

According to [19] & [26], the N-dimensional test $\mathcal{T}_{N,\gamma}^* : \mathbb{R}^N \to \{0,1\}$ defined for every $x \in \mathbb{R}^N$ by:

$$\mathcal{T}_{N,\gamma}^*(x) = \begin{cases} 0 \text{ if } \ \lvert \langle x \rangle_N - \xi_0 \rvert \leqslant \lambda_\gamma(\tau\sqrt{N})/\sqrt{N} \\ 1 \text{ otherwise} \end{cases} , \quad (35)$$

is maximal in $\mathcal{K}_\gamma$: for any $\mathcal{T} \in \mathcal{K}_\gamma$, $\mathcal{T} \preceq \mathcal{T}_{N,\gamma}^*$. Let $\mathbb{P}_{\mathrm{FA}}^{\mathrm{B\text{-}RDT}}(N,\gamma)$ and $\mathbb{P}_{\mathrm{MD}}^{\mathrm{B\text{-}RDT}}(N,\gamma)$ be PFA and PMD for *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 in (6).

Proposition 7.1: $\mathbb{P}_{\mathrm{FA}}^{\mathrm{B\text{-}RDT}}(N,\gamma) \leqslant \gamma$ & $\mathbb{P}_{\mathrm{MD}}^{\mathrm{B\text{-}RDT}}(N,\gamma) \leqslant 1-\gamma$. Under Assumption 2.2, $\mathcal{H}_0$ and $\mathcal{H}_1$ in (6) are satisfied for τ^- instead of τ. Thereby, $\lambda_\gamma(\tau\sqrt{N})/\sqrt{N}$ can be replaced by $\lambda_\gamma(\tau^-\sqrt{N})/\sqrt{N}$ in (35). The resulting better performance guaranteed by *Block*RDT is given by the next proposition.

Proposition 7.2: Under Assumption 2.2, $\mathbb{P}_{\mathrm{FA}}^{\mathrm{B\text{-}RDT}}(N,\gamma) \leqslant \gamma$ & $\mathbb{P}_{\mathrm{MD}}^{\mathrm{B\text{-}RDT}}(N,\gamma) \leqslant 1 - Q_{1/2}\left(\tau^+\sqrt{N}, \lambda_\gamma(\tau^-\sqrt{N})\right)$.

Given γ, the upper bound on $\mathbb{P}_{\mathrm{MD}}^{\mathrm{B\text{-}RDT}}(N,\gamma)$ is decreasing with N, therefore N can be chosen large enough to guarantee a specified $\mathbb{P}_{\mathrm{MD}}^{\mathrm{B\text{-}RDT}}(N,\gamma)$. *Block*RDT applies to the detection of a change in the mean of Ξ with bounded distortions by choosing $\tau^- = h_0$ and $\lvert \Delta_n \rvert \leqslant \tau^-$. Given $\alpha, \beta \in (0,1)$, we can compute the block size $N_{\mathrm{B\text{-}RDT}}$ required to drive the upper bound on $\mathbb{P}_{\mathrm{MD}}^{\mathrm{B\text{-}RDT}}(N,\alpha)$ below β. The decision is made by $\mathcal{T}_{N,\gamma}^*$ with $N = N_{\mathrm{B\text{-}RDT}}$ and $\gamma = \alpha$ in (35).

Sequential Probability Ratio Test (SPRT): Suppose that the probability density function f_i of the observation is known under $\mathcal{H}_i$ for $i = 0,1$. Given $\alpha, \beta \in (0,1)$, and with initialization $N = 1$, SPRT has the following form:

$$\begin{cases} \text{If } \Lambda_N \leqslant \lambda_L^{\mathrm{SPRT}}, \text{ decide } \mathcal{H}_0 \text{ and stops;} \\ \text{If } \Lambda_N \geqslant \lambda_H^{\mathrm{SPRT}}, \text{ decide } \mathcal{H}_1 \text{ and stops;} \\ \text{If } \lambda_L^{\mathrm{SPRT}} < \Lambda_N < \lambda_H^{\mathrm{SPRT}}, \text{compute } \Lambda_{N+1} \text{ and repeat;} \end{cases} , \quad (36)$$

with $\Lambda_N = \prod_{n=1}^{N} \frac{f_1(Y_i)}{f_0(Y_i)}$, $\lambda_L^{\mathrm{SPRT}} = \frac{\beta}{1-\alpha}$ and $\lambda_H^{\mathrm{SPRT}} = \frac{1-\beta}{\alpha}$. The PFA, $\mathbb{P}_{\mathrm{FA}}^{\mathrm{SPRT}}$ and the PMD, $\mathbb{P}_{\mathrm{MD}}^{\mathrm{SPRT}}$ of SPRT are guaranteed to stay

below α and β, respectively [10], [11]. In our case, $\Lambda_N =$

$$\prod_{n=1}^{N} \frac{\int_{-h_0}^{h_0} \exp\left(-(Y_n - (\xi_1 + \Delta))^2/2\right) \mathrm{d}\Delta}{\int_{-h_0}^{h_0} \exp\left(-(Y_n - (\xi_0 + \Delta))^2/2\right) \mathrm{d}\Delta}.$$ SPRT requires

the knowledge of distortion distribution. But in many practical scenarios, it is unaware of the distortion. In this case, we represent the algorithm by SPRT-MM and denote by $T_{\mathrm{SPRT\text{-}MM}}$, $\mathbb{P}_{\mathrm{FA}}^{\mathrm{SPRT\text{-}MM}}$ and $\mathbb{P}_{\mathrm{MD}}^{\mathrm{SPRT\text{-}MM}}$, its stopping time, PFA and PMD, respectively. SPRT-MM has same thresholds as SPRT and log-likelihood ratio is: $\log \Lambda_N = N \frac{\xi_0^2 - \xi_1^2}{2} + (\xi_1 - \xi_0) \sum_{n=1}^{N} Y_N$.

C. Unbounded Distortion

Because Assumption 2.2 might not be satisfied for all types of distortion, we relax Assumption 2.2 by considering unbounded distortions $\Delta_n \sim \mathcal{N}(0, \sigma^2)$ for $n \in \mathbb{N}$. The notation below remains the same as in the bounded regime case.

*Seq*RDT: For $\tau^- = h_0 = \sigma/4$, $\tau^+ = 3h_0$, $\tau = 2h_0$ and $\tau_H = \lvert \xi_1 - \xi_0 \rvert + h_0$, $\mathbb{P}[\lvert\langle\Delta\rangle_N\rvert \leqslant \tau] \geqslant 0.9545$, $\mathbb{P}[\lvert\langle\Delta\rangle_N\rvert \leqslant \tau^-] \geqslant 0.6827$, $\mathbb{P}[\lvert\langle\Delta\rangle_N + \xi_1 - \xi_0\rvert > \tau] \geqslant 0.9772$ and $\mathbb{P}[\lvert\langle\Delta\rangle_N + \xi_1 - \xi_0\rvert > \tau^+] \geqslant 0.8413$ for all $N \geqslant N_0$ with $N_0 = 16$ and $\lvert \xi_1 - \xi_0 \rvert \geqslant 4h_0$. Although the (a-s) convergence in (34) (Assumption 2.2) is not satisfied, simulations show that *Seq*RDT is still able to guarantee performance.

*Block*RDT: For $\tau^- = h_0$, the probabilities stay the same as derived above for *Seq*RDT.

SPRT: Because SPRT is aware of the distortion distribution $\Delta_n \sim \mathcal{N}(0, h_0^2)$ $(n \in \mathbb{N})$, the log-likelihood ratio in (36) becomes $\Lambda_N = N \frac{\xi_0^2 - \xi_1^2}{2(1+h_0^2)} + \frac{\xi_1 - \xi_0}{1+h_0^2} \sum_{n=1}^{N} Y_n$. Unaware of the distortion distribution, SPRT-MM remains the same as specified for the bounded regime.

D. Comparison: SeqRDT, BlockRDT and SPRT

For both the bounded distortion case described in Section VII-B and the unbounded distortion case described in Section

VII-C, we define $|\xi_1 - \xi_0|$ as the Signal-to-Noise Ratio (SNR) and we choose $h_0 = 0.1$ in our simulations. First, to perform *Seq*RDT, we must first choose an appropriate buffer size $M \geqslant N_0 - 1$. We fix M such that $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ are below specified levels α and β, respectively, by making use of the upper bounds derived in Theorem 5.2 (i). Figure 1 plots the upper bounds on $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ given by Theorem 5.2 (i), when M increases. Note that the terms $\text{UB2}_{\text{FA}}(M)$ and $\text{UB2}_{\text{MD}}(M)$ in Theorem 5.2 are infinite sums over $N = M + 2$ to ∞. For simulation purposes, we curtail the sum to $N = 500$ terms, which approximates the bounds sufficiently well. From Figure 1 we choose $M = 90$.

First, in Table I we analyze the average number of samples taken by *Seq*RDT compared to its FSS counterpart *Block*RDT for level $\alpha = \beta = 0.01$ and for two cases one when *Block*RDT utilizes τ^- and second when it uses τ to design the threshold. We compare $N_{\text{B-RDT}}$ needed for *Block*RDT to guarantee that $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}(N_{\text{B-RDT}}, \beta)$ stays below β against T of *Seq*RDT designed to guarantee that $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ stay below α and β, respectively. For *Block*RDT, we choose $\tau^+ = |\xi_1 - \xi_0| - h_0$ to compute $N_{\text{B-RDT}}$. This is the best τ^+ we can choose to have the smallest possible $N_{\text{B-RDT}}$. Note from Table I that *Seq*RDT is faster compared to *Block*RDT for both bounded and unbounded distortions, even when *Block*RDT has knowledge of τ^-, specially at low SNRs.

In Tables II and III, we compare the average stopping times and PFA and PMD of *Seq*RDT against SPRT and SPRT-MM, for different SNRs and for levels $\alpha = \beta = 0.01$ and $\alpha = \beta = 0.05$ for bounded distortions and for $\alpha = \beta = 0.01$ and $\alpha = \beta = 0.001$ for unbounded distortions. We average the stopping times, PFA and PMD over 10^5 Monte Carlo iterations. The stopping time is taken to be the average of the stopping times under the two hypotheses. As expected, SPRT is optimal if the distortion and noise distributions are completely known. On the other hand, when SPRT is unaware of the distortions, we notice that for SPRT-MM, PFA and PMD cannot be guaranteed to stay below the levels α and β, for unbounded distortions, especially at low SNRs. SPRT is sensitive to model mismatches, whereas *Seq*RDT is robust and requires knowledge of a few parameters only. The results of Table III imply that Assumption 2.2 need not be satisfied in a strict (a-s) sense. If the assumption is satisfied with sufficiently high probability, *Seq*RDT still provides sufficient performance guarantees. This means that *Seq*RDT is also robust to mismatches with respect to the choice of parameter values and can be expected to perform well even if the tolerances are not known precisely or are merely known approximately.

VIII. CONCLUSION

In this work, we proposed a new algorithm *Seq*RDT for non-parametric sequential hypothesis testing. We studied the properties of the thresholds and showed that the chosen thresholds are appropriate for conducting the sequential test. We introduced the concept of a buffer; the size of this buffer helps control PFA and PMD of the proposed test. Simulations showed that the *Seq*RDT approach leads to faster decision-making on an average compared to its FSS counterpart,

*Block*RDT [19]. Moreover, we showed that *Seq*RDT is robust to model mismatches compared to SPRT [10], [11]. Extension of *Seq*RDT to multi-dimensional observations may be addressed in future work.

APPENDIX A
USEFUL LEMMAS

Lemma A.1: For any $N \in \mathbb{N}$ and any $\eta \geqslant 0$, we have:

$$\mathbb{E}\left[Q_{1/2}\left(\sqrt{N} \, |\langle \Xi \rangle_N - \xi_0|, \eta\sqrt{N} \right) \right] - 2\delta_N$$
$$\leqslant \mathbb{P}\left[|\langle \Xi \rangle_N + \langle X \rangle_N - \xi_0| > \eta \right]$$
$$\leqslant \mathbb{E}\left[Q_{1/2}\left(\sqrt{N}|\langle \Xi \rangle_N - \xi_0|, \eta\sqrt{N} \right) \right] + 2\delta_N,$$

with $\delta_N = c\,\mathbb{E}\left[|X_1 - \mathbb{E}[X_1]|^3 \right] \text{Var}(X_1)^{-3/2} N^{-1/2}$ and $(2\pi)^{-1/2} \leqslant c < 0.8$.

Proof: Using Berry-Esseen inequality [27, Chap. 3]:

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}\left[\sqrt{N}\langle X \rangle_N \leqslant x \right] - \Phi(x) \right| \leqslant \delta_N. \qquad (37)$$

Since $\langle \Xi \rangle_N$ and $\langle X \rangle_N$ are independent, we have:

$$\mathbb{P}[|\langle \Xi \rangle_N + \langle X \rangle_N - \xi_0| \leqslant \eta]$$
$$= \int_{\mathbb{R}} \mathbb{P}\left[|x + \langle X \rangle_N| \leqslant \eta \right] \mathbb{P}_{\langle \Xi \rangle_N - \xi_0}(\mathrm{d}x).$$

It then follows from (37) that, for all $x \in \mathbb{R}$:

$$\psi(x\sqrt{N}, \eta\sqrt{N}) - 2\delta_N \leqslant \mathbb{P}\left[|x + \langle X \rangle_N| \leqslant \eta \right]$$
$$\leqslant \psi(x\sqrt{N}, \eta\sqrt{N}) + 2\delta_N, \quad (38)$$

with $\psi(a, b) = \Phi(b - a) - \Phi(-b - a)$ for any $(a, b) \in \mathbb{R} \times \mathbb{R}$. We derive from the foregoing that:

$$\mathbb{E}\left[\psi\left(\sqrt{N}\left(\langle X \rangle_N - \xi_0 \right), \eta\sqrt{N} \right) \right] - 2\delta_N$$
$$\leqslant \mathbb{P}[|\langle \Xi \rangle_N + \langle X \rangle_N - \xi_0| \leqslant \eta]$$
$$\leqslant \mathbb{E}\left[\psi\left(\sqrt{N}\left(\langle X \rangle_N - \xi_0 \right), \eta\sqrt{N} \right) \right] + 2\delta_N,$$

and conclude since $\psi(a, b) = \psi(|a|, b) = 1 - Q_{1/2}(a, b)$. ∎

Lemma A.2: If Ξ satisfies Assumption 2.1, then, for any $\gamma \in (0, 1)$, we have:

$$\lim_{N \to \infty} Q_{1/2}\left(\sqrt{N} \, |\langle \Xi \rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N}) \right) = \begin{cases} 0 \text{ under } \mathcal{H}_0 \\ 1 \text{ under } \mathcal{H}_1 \end{cases}.$$

Proof: Under Assumption 2.1, $\lim_{N \to \infty} |\langle \Xi \rangle_N - \xi_0| = |\langle \Xi \rangle_\infty - \xi_0|$ (a-s). Therefore, there exists $\Omega' \in \mathcal{F}$ such that $\mathbb{P}(\Omega') = 1$ and for all $\omega \in \Omega'$, $\lim_{N \to \infty} |\langle \Xi \rangle_N(\omega) - \xi_0| = |\langle \Xi \rangle_\infty(\omega) - \xi_0|$. It follows that, for any $\varepsilon > 0$ and any $\omega \in \Omega'$, there exists $N_0(\varepsilon, \omega) \in \mathbb{N}$ such that, for any $N \geqslant N_0(\varepsilon, \omega)$,

$$|\langle \Xi \rangle_\infty(\omega) - \xi_0| - \varepsilon \leqslant |\langle \Xi \rangle_N(\omega) - \xi_0| \leqslant |\langle \Xi \rangle_\infty(\omega) - \xi_0| + \varepsilon.$$

Because $Q_{1/2}(\bullet, \lambda_\gamma(\tau\sqrt{N}))$ increases, we can also write that, for all $\varepsilon > 0$ and all $\omega \in \Omega'$, there exists $N_0(\varepsilon, \omega) \in \mathbb{N}$ such that, for any $N \geqslant N_0(\varepsilon, \omega)$:

$$Q_{1/2}\Big(\sqrt{N}\big(|\langle\Xi\rangle_\infty(\omega)-\xi_0|-\varepsilon\big),\lambda_\gamma(\tau\sqrt{N})\Big)$$
$$\leqslant Q_{1/2}\Big(\sqrt{N}\,|\langle\Xi\rangle_N(\omega)-\xi_0|,\lambda_\gamma(\tau\sqrt{N})\Big)$$
$$\leqslant Q_{1/2}\Big(\sqrt{N}\,\big(|\langle\Xi\rangle_\infty(\omega)-\xi_0|+\varepsilon\big),\lambda_\gamma(\tau\sqrt{N})\Big). \quad (39)$$

Under $\mathcal{H}_0$, (7) implies the existence of $\Omega_0\in\mathcal{F}$ such that, for any $\omega\in\Omega_0$, $|\langle\Xi\rangle_\infty(\omega)-\xi_0|\leqslant\tau^-$. Choose $\varepsilon>0$ such that $\tau^-+\varepsilon<\tau$. For any $\omega\in\Omega_0$, $|\langle\Xi\rangle_\infty(\omega)-\xi_0|+\varepsilon<\tau$. Therefore, for any $\omega\in\Omega_0$, it follows from Lemma B.2 that:

$$\lim_{N\to\infty}Q_{1/2}\Big(\sqrt{N}\,\big(|\langle\Xi\rangle_\infty(\omega)-\xi_0|+\varepsilon\big),\lambda_\gamma(\tau\sqrt{N})\Big)=0. \quad (40)$$

Since $\mathbb{P}(\Omega'\cap\Omega_0)=1$, Eqs. (39) and (40) imply that:

$$\lim_{N\to\infty}Q_{1/2}\Big(\sqrt{N}\,|\langle\Xi\rangle_N-\xi_0|,\lambda_\gamma(\tau\sqrt{N})\Big)=0 \quad \text{(a-s)}.$$

Under $\mathcal{H}_1$, the proof is similar. From (7), there exists $\Omega_1\in\mathcal{F}$ such that: $\forall\omega\in\Omega_1$, $|\langle\Xi\rangle_\infty(\omega)-\xi_0|>\tau^+$. For $\varepsilon>0$ such that $\tau^+-\varepsilon>\tau$, Lemma B.2 induces that, for any $\omega\in\Omega_1$:

$$\lim_{N\to\infty}Q_{1/2}\Big(\sqrt{N}\,\big(|\langle\Xi\rangle_\infty(\omega)-\xi_0|-\varepsilon\big),\lambda_\gamma(\tau\sqrt{N})\Big)=1. \quad (41)$$

Since $\mathbb{P}(\Omega'\cap\Omega_1)=1$, it follows from (41) and (39) that:

$$\lim_{N\to\infty}Q_{1/2}\Big(\sqrt{N}\,|\langle\Xi\rangle_N-\xi_0|,\lambda_\gamma(\tau\sqrt{N})\Big)=1 \quad \text{(a-s)},$$

which concludes the proof. ∎

APPENDIX B
PROPERTIES OF THE THRESHOLDS

Lemma B.1: For any $\gamma\in(0,1)$:

(i) $\lim_{\rho\to\infty}\big(\lambda_\gamma(\rho)-\rho\big)=\Phi^{-1}(1-\gamma)$;

(ii) $\lim_{\rho\to\infty}\lambda_\gamma(\rho)/\rho=1$;

Proof: We only prove (i) since it directly implies (ii). Pose $g(\rho)=\lambda_\gamma(\rho)-\rho$ and $\theta=\Phi^{-1}(1-\gamma)$. Since $\Phi(x)+\Phi(-x)=1$, (3) and the definition of $\lambda_\gamma(\tau)$ induce that:

$$\Phi(g(\rho))+\Phi(g(\rho)+2\rho)=1+\Phi(\theta). \quad (42)$$

To prove that $g(\rho)$ tends to θ when $\rho\to\infty$, we proceed by contradiction. If $g(\rho)$ does not tend to θ when $\rho\to\infty$, there exists some positive real number ε such that, for all $n\in\mathbb{N}$, there exists some real number $\rho_n>n$ such that either $g(\rho_n)>\theta+\varepsilon$ or $g(\rho_n)<\theta-\varepsilon$. Basically, $\lim_{n\to\infty}\rho_n=\infty$. Consider any $\eta\in(0,\Phi(\theta)-\Phi(\theta-\varepsilon))$. Since $\lim_{n\to\infty}\Phi(2\rho_n+\theta+\varepsilon)=1$, there exists $N_0\in\mathbb{N}$ such that, for all $n\geqslant N_0$:

$$\Phi(2\rho_n+\theta+\varepsilon)>1-\eta. \quad (43)$$

Similarly, since $\lim_{n\to\infty}\Phi(2\rho_n+\theta-\varepsilon)=1$, there exists $N_1\in\mathbb{N}$ such that, for all $n\geqslant N_1$:

$$\Phi(2\rho_n+\theta-\varepsilon)<1+\eta. \quad (44)$$

Let n be any integer above $\max(N_0,N_1)$. If $g(\rho_n)<\theta-\varepsilon$, we then have $\Phi(g(\rho_n))<\Phi(\theta-\varepsilon)$ and $\Phi(2\rho_n+g(\rho_n))<\Phi(2\rho_n+\theta-\varepsilon)$. Eqs. (42) and (44) then imply that:

$$1+\Phi(\theta)<\Phi(\theta-\varepsilon)+\Phi(2\rho_n+\theta-\varepsilon)<\Phi(\theta-\varepsilon)+1+\eta,$$

which is impossible because of our choice for η. Therefore, we cannot have $g(\rho_n)<\theta-\varepsilon$. We cannot have $g(\rho_n)>\theta+\varepsilon$ either because, via (42) and (43), this inequality implies:

$$1+\Phi(\theta)>\Phi(\theta+\varepsilon)+\Phi(2\rho_n+\theta+\varepsilon)>\Phi(\theta+\varepsilon)+1-\eta, \quad (45)$$

which is contradictory to our choice for η. ∎

Lemma B.2 (Asymptotic behavior of $Q_{1/2}$ in vanishing noise): Consider $\tau\in]0,\infty)$ and $\rho\in(0,\infty)$ such that $\rho\neq\tau$.

$$\forall\gamma\in(0,1),\ \lim_{\sigma\to0}Q_{1/2}\big(\rho/\sigma,\lambda_\gamma(\tau/\sigma)\big)=\mathbb{1}_{(\tau,\infty)}(\rho).$$

Proof: Let $(\sigma_n)_{n\in\mathbb{N}}$ be a sequence of positive real values such that $\lim_{n\to\infty}\sigma_n=0$ and set $\rho_n=\tau/\sigma_n$ for each $n\in\mathbb{N}$. According to (2), $Q_{1/2}\big((\rho/\tau)\rho_n,\lambda_\gamma(\rho_n)\big)=\mathbb{P}\big[\,|(\rho/\tau)+X/\rho_n|>\lambda_\gamma(\rho_n)/\rho_n\big]$ for any $X\sim\mathcal{N}(0,1)$. It follows from Lemma B.1 (ii) that $|(\rho/\tau)+(X/\rho_n)|-\lambda_\gamma(\rho_n)/\rho_n=(\rho/\tau)-1$ (a-s). Therefore, the cdf of $|(\rho/\tau)+(X/\rho_n)|-\lambda_\gamma(\rho_n)/\rho_n$ converges weakly to $\mathbb{1}_{[(\rho/\tau)-1,\infty)}$. Since $\rho\neq\tau$, this weak convergence implies that $\lim_{n\to\infty}\mathbb{P}\big[\,|(\rho/\tau)+X/\rho_n|>\lambda_\gamma(\rho_n)/\rho_n\big]=\mathbb{1}_{(\tau,\infty)}(\rho)$. Thence the result since $(\sigma_n)_{n\in\mathbb{N}}$ is arbitrary. ∎

Lemma B.3 (Non-Asymptotic behavior of $Q_{1/2}$): Given $\tau\in[0,\infty)$, $\rho\in(0,\infty)$ and $\gamma\in(0,1)$, the map:

$$\sigma\in[0,\infty)\mapsto Q_{1/2}\big(\rho\sigma,\lambda_\gamma(\tau\sigma)\big)\ \text{is}\ \begin{cases}\text{constant }\gamma\text{ if }\rho=\tau\\\text{decreasing if }\rho<\tau\\\text{increasing if }\rho>\tau\end{cases}.$$

Proof: Given ρ and τ, we want to study the behavior of

$$\mathcal{Q}(\sigma)=Q_{1/2}\big(\rho\sigma,\lambda_\gamma(\tau\sigma)\big)=1-\Phi(r_-(\sigma))+\Phi(-r_+(\sigma)), \quad (46)$$

with $r_+(\sigma)=\lambda_\gamma(\tau\sigma)+\rho\sigma$ and $r_-(\sigma)=\lambda_\gamma(\tau\sigma)-\rho\sigma$. For $\rho=\tau$, it follows from (4) that $\mathcal{Q}$ is constant equal to γ. We thus have $1-\Phi(\lambda_\gamma(\tau\sigma)-\tau\sigma)+\Phi(-\lambda_\gamma(\tau\sigma)-\tau\sigma)=\gamma$. After differentiating the two members of this equality above and after some routine algebra, we obtain:

$$\lambda_\gamma'(\tau\sigma)=(1-e^{-2\tau\sigma\lambda_\gamma(\tau\sigma)})/(1+e^{-2\tau\sigma\lambda_\gamma(\tau\sigma)}), \quad (47)$$

where λ_γ' is the first derivative of λ_γ. We now differentiate $\mathcal{Q}$ defined by (46). Some easy computation yields:

$$\mathcal{Q}'(\sigma)=\frac{1}{\sqrt{2\pi}}\left(e^{-r_-^2(\sigma)/2}-e^{-r_+^2(\sigma)/2}\right)$$
$$\left(\rho-\tau\lambda_\gamma'(\tau\sigma)\frac{1+e^{-2\rho\sigma\lambda_\gamma(\tau\sigma)}}{1-e^{-2\rho\sigma\lambda_\gamma(\tau\sigma)}}\right).$$

By injecting (47) into the equality above, we obtain:

$$\mathcal{Q}'(\sigma)=\frac{\tau}{\sqrt{2\pi}}\left(e^{-r_-^2(\sigma)/2}-e^{-r_+^2(\sigma)/2}\right)\left(\frac{\rho}{\tau}-\frac{\Delta_{-1}(\rho,\tau)}{\Delta_{+1}(\rho,\tau)}\right). \quad (48)$$

with $\Delta_\varepsilon(\rho,\tau)=(1+\varepsilon e^{-2\tau\sigma\lambda_\gamma(\tau\sigma)})/(1+\varepsilon e^{-2\rho\sigma\lambda_\gamma(\tau\sigma)})$ and $\varepsilon\in\{-1,+1\}$. For all $\sigma\geqslant0$, the sign of $\mathcal{Q}'$ is therefore that of $(\rho/\tau)-(\Delta_{-1}(\rho,\tau)/\Delta_{+1}(\rho,\tau))$ We verify easily that:

$$\begin{cases}\rho<\tau\Leftrightarrow\Delta_-(\rho,\tau)>1\Leftrightarrow\Delta_+(\rho,\tau)<1\\\rho=\tau\Leftrightarrow\Delta_-(\rho,\tau)=\Delta_+(\rho,\tau))=1\end{cases}$$

Therefore, if $\rho<\tau$, $\rho/\tau<1<\Delta_-(\rho,\tau)/\Delta_+(\rho,\tau)$, which implies that $\mathcal{Q}'(\sigma)<0$ and, thus, that $\mathcal{Q}$ decreases. On the other hand, if $\rho>\tau$, we have $\rho/\tau>1>\Delta_-(\rho,\tau)/\Delta_+(\rho,\tau)$, so that $\mathcal{Q}$ increases in this case. ∎

Lemma B.4: Given $\rho \in (0, \infty)$, the map $\gamma \in (0, 1) \mapsto \lambda_\gamma(\rho)$ is decreasing.

Proof: A straightforward consequence of (4) and the decreasing nature of $Q_{1/2}$ with its second argument. ∎

Lemma B.5:

(P1) For any $\tau \in [0, \infty)$ and any $\eta \in (\tau, \infty)$, the map $\sigma \in (0, \infty) \mapsto Q_{1/2}(\tau/\sigma, \eta/\sigma)$ is increasing.

(P2) The map $\rho \in (0, \infty) \mapsto Q_{1/2}(\rho, \rho)$ is decreasing, lower-bounded by $1/2$ and $\lim_{\rho \to \infty} Q_{1/2}(\rho, \rho) = 1/2$.

(P3) For any $\gamma \in (0, 1/2)$, the map $\rho \in (0, \infty) \mapsto \lambda_\gamma(\rho)/\rho$ is decreasing and lower bounded by 1.

Proof: Proof of (P1): Using (3), define $\mathcal{Q}(\sigma)$ as:

$$\mathcal{Q}(\sigma) = Q_{1/2}(\tau/\sigma, \eta/\sigma) = 1 - \Phi(\eta/\sigma - \tau/\sigma) + \Phi(-\eta/\sigma - \tau/\sigma).$$

We now differentiate $\mathcal{Q}$ and some easy computation yields:

$$\mathcal{Q}'(\sigma) = (1/\sqrt{2\pi}\sigma^2)e^{-\frac{(\eta-\tau)^2}{2\sigma^2}}\left[(\eta - \tau) + (\eta + \tau)e^{-\frac{2\eta\tau}{\sigma^2}}\right].$$

Thence the result, since $\eta \in [\tau, \infty)$ implies that $\mathcal{Q}'(\sigma) > 0$.

Proof of (P2): The map $\rho \in (0, \infty) \mapsto Q_{1/2}(\rho, \rho)$ is decreasing as a consequence of (P1). Given $\rho \in (0, \infty)$, $Q_{1/2}(\rho, \rho) = 1/2 + \Phi(-2\rho)$ from (3). Hence the result.

Proof of statement (P3): Let ρ and ρ' be two positive real numbers such that $0 < \rho < \rho'$. According to (4), we have:

$$Q_{1/2}(\rho, \lambda_\gamma(\rho)) = Q_{1/2}(\rho', \lambda_\gamma(\rho')) = \gamma. \qquad (49)$$

Since $\gamma < 1/2$, it follows from (P2) and (49) that $Q_{1/2}(\rho, \rho) > 1/2 > Q_{1/2}(\rho, \lambda_\gamma(\rho))$. The decreasing behavior of $Q_{1/2}$ with its second argument implies that $\lambda_\gamma(\rho) > \rho$, so that $\lambda_\gamma(\rho)/\rho$ is lower bounded by 1. We then derive from (P1) that $x \in (0, \infty) \mapsto Q_{1/2}(\rho/x, \lambda_\gamma(\rho)/x)$ is an increasing map. Since $\rho/\rho' < 1$, we have $Q_{1/2}(\rho, \lambda_\gamma(\rho)) > Q_{1/2}(\rho', \rho'\lambda_\gamma(\rho)/\rho)$. This inequality and (49) induce that $Q_{1/2}(\rho', \lambda_\gamma(\rho')) > Q_{1/2}(\rho', \rho'\lambda_\gamma(\rho)/\rho)$. The decreasing nature of $Q_{1/2}(\rho', \bullet)$ then implies that $\lambda_\gamma(\rho') < \rho'\lambda_\gamma(\rho)/\rho$. Thereby, $\rho \in (0, \infty) \mapsto \lambda_\gamma(\rho)/\rho$ is decreasing in ρ. ∎

Lemma B.6: For $\gamma \in (1/2, 1)$ and ρ large enough, the map $\rho \in (0, \infty) \mapsto \lambda_\gamma(\rho)/\rho$ is increasing and upper bounded by 1.

Proof: According to statement (i) of Lemma B.1, $\lambda_\gamma(\rho) - \rho = \Phi^{-1}(1 - \gamma) + \kappa(\rho)$ where κ is such that $\lim_{\rho \to \infty} \kappa(\rho) = 0$. Since $\gamma > 1/2$, $\Phi^{-1}(1 - \gamma) < 0$. Given η such that $0 < \eta < -\Phi^{-1}(1-\gamma)$, there exists ρ_0 such that, for all $\rho \geqslant \rho_0$, $\kappa(\rho) \leqslant \eta$. Therefore, for all $\rho \geqslant \rho_0$, $\lambda_\gamma(\rho) - \rho \leqslant \Phi^{-1}(1 - \gamma) + \eta < 0$. We have hence proved that $\lambda_\gamma(\rho) < \rho$ for ρ large enough.

Pose $f(\rho) = \lambda_\gamma(\rho)/\rho$ so that $\Phi(\rho(f(\rho)-1)) - \Phi(-\rho(f(\rho)+1)) = 1 - \gamma$. By differentiation of this equality with respect to ρ and since f is differentiable via the implicit function theorem, we find that $f'(\rho)$ has the same sign as $\Upsilon(\rho) = \left(1 - f(\rho)\right)\left(e^{2\rho\lambda_\gamma(\rho)} + \frac{\lambda_\gamma(\rho)+\rho}{\lambda_\gamma(\rho)-\rho}\right)$. For ρ large enough, $f(\rho) < 1$ by the first part of the proof and Lemma B.1 implies that $\lim_{\rho \to \infty} \Upsilon(\rho) = \infty$. Therefore, $\Upsilon(\rho) > 0$ for ρ large enough and the proof is complete. ∎

APPENDIX C
CONVERGENCE OF THE UPPER BOUNDS IN THEOREM 5.2

Let us consider $\text{UB2}_{\text{FA}}(M)$ given in (27), we have:

$$\text{UB2}_{\text{FA}}(M) \leqslant \sum_{N=1}^{\infty} Q_{1/2}\left(\tau^-\sqrt{N}, \lambda_\alpha(\tau\sqrt{N})\right).$$

Each term of the above summation is positive and $t > 0 \mapsto Q_{1/2}\left(\tau^-\sqrt{t}, \lambda_\alpha(\tau\sqrt{t})\right)$ is decreasing since $\tau^- < \tau$, as a consequence of Lemma B.3. Therefore, to show that the above series converges, it suffices to show that

$$I_{\text{FA}}(\zeta) = \int_\zeta^\infty Q_{1/2}\left(\tau^-\sqrt{t}, \lambda_\alpha(\tau\sqrt{t})\right) \mathrm{d}t, \qquad (50)$$

is finite for some $\zeta > 0$, with $\alpha \in (0, 1/2)$ and $\tau^- < \tau$. From the definition of $Q_{1/2}(a, b)$ given in (3), we have:

$$Q_{1/2}(a, b) = \Phi^c(b - a) + \Phi^c(b + a), \qquad (51)$$

where $\Phi^c(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{\frac{-z^2}{2}}\mathrm{d}z$, also known as the standard Q function, is the complementary cdf of a standard normal distributed random variable. Using (51) in (50) and the fact that we have $\Phi^c(\lambda_\alpha(\tau\sqrt{t}) + \tau^-\sqrt{t}) \leqslant \Phi^c(\lambda_\alpha(\tau\sqrt{t}) - \tau^-\sqrt{t})$ for any t, we get

$$I_{\text{FA}}(\zeta) \leqslant 2\int_\zeta^\infty \Phi^c(\lambda_\alpha(\tau\sqrt{t}) - \tau^-\sqrt{t})\mathrm{d}t \overset{(a)}{\leqslant} 2\int_\zeta^\infty e^{\frac{-(\lambda_\alpha(\tau\sqrt{t})-\tau^-\sqrt{t})^2}{2}}\mathrm{d}t,$$

where (a) results from $\Phi^c(x) \leqslant e^{-x^2/2}$ for $x > 0$ as, for $\alpha < 1/2$, $\lambda_\alpha(\tau\sqrt{t}) - \tau^-\sqrt{t} > 0$ according to Lemma B.5 (P3). Therefore, to show that $I_{\text{FA}}(\zeta) < \infty$, it suffices to show

$$J_{\text{FA}}(\zeta) \overset{\text{def}}{=} \int_\zeta^\infty e^{\frac{-u(t)^2}{2}}\mathrm{d}t < \infty, \qquad (52)$$

with $u(t) \overset{\text{def}}{=} \lambda_\alpha(\tau\sqrt{t}) - \tau^-\sqrt{t}$.

To prove (52), we need the following properties.

(C1) For any $t > 0$, $u(t) \geqslant (\tau - \tau^-)\sqrt{t}$.

(C2) For any $t > 0$, $u'(t) = \frac{1}{2\sqrt{t}}\left[\tau\frac{1-G_\alpha(t)}{1+G_\alpha(t)} - \tau^-\right]$ where:

$$\forall\gamma \in (0, 1), G_\gamma(t) \overset{\text{def}}{=} e^{-2\tau\sqrt{t}\lambda_\alpha(\tau\sqrt{t})}.$$

(C3) There exists $A > 0$ such that, for any $t > A$, $u'(t) > 0$.

(C4) Given any $a \in (0, \tau - \tau^-)$, there exists $B > 0$ such that, for any $x > B$, $\tau\frac{1-G_\alpha(u^{-1}(x))}{1+G_\alpha(u^{-1}(x))} - \tau^- > a$.

Proof: **(C1)** Lemma B.5 (P3) implies $\lambda_\alpha(\tau\sqrt{t}) \geqslant \tau\sqrt{t}$. Hence **(C1)**

(C2) This property follows from (47) in Lemma B.3.

(C3) Lemma B.5 (P3) implies $G_\alpha(t) \leqslant e^{-2\tau^2 t}$. Thus, $\lim_{t\to\infty}\left(\tau\frac{1-G_\alpha(t)}{1+G_\alpha(t)} - \tau^-\right) = \tau - \tau^- > 0$. Whence the result.

(C4) From **(C3)** $u(t)$ increases for $t > A > 0$, which implies that $u^{-1}(x)$ increases for $x > u(A)$. It follows that $\lim_{x\to\infty}\left(\tau\frac{1-G_\alpha(u^{-1}(x))}{1+G_\alpha(u^{-1}(x))} - \tau^-\right) = \tau - \tau^-$. Thence **(C4)**. ∎

Choose any $a \in (0, \tau - \tau^-)$ and $\zeta \geqslant \max(A, u^{-1}(B))$ where A and B are given by **(C3)** and **(C4)**, respectively. By the Jacobi's transformation formula (see [28, Theorem 12.6], among others) and the fact that $u'(u^{-1}(x)) > 0$ for any $x > u(\zeta)$ since $\zeta > A$, we have:

$$J_{\text{FA}}(\zeta) = \int_{u(\zeta)}^\infty \frac{1}{u'(u^{-1}(x))}e^{-x^2/2}\mathrm{d}x. \qquad (53)$$

By **(C2)**, $\forall x \geqslant u(\zeta)$ we have:

$$u'(u^{-1}(x)) = \frac{1}{2\sqrt{u^{-1}(x)}}\left[\tau \frac{1 - G_\alpha(u^{-1}(x))}{1 + G_\alpha(u^{-1}(x))} - \tau^-\right]. \quad (54)$$

On the other hand, **(C1)** implies $\forall x \geqslant u(\zeta)$ we have:

$$\frac{1}{\sqrt{u^{-1}(x)}} \geqslant \frac{\tau - \tau^-}{x}. \quad (55)$$

Finally, **(C4)** induces that $\forall x \geqslant u(\zeta)$:

$$\tau \frac{1 - G_\alpha(u^{-1}(x))}{1 + G_\alpha(u^{-1}(x))} - \tau^- > a. \quad (56)$$

By injecting (55) and (56) into (54), we obtain $u'(u^{-1}(x)) \geqslant a(\tau - \tau^-)/2x$ for all $x \geqslant u(\zeta)$. Therefore, it results from (53) that $J_{\text{FA}}(\zeta) \leqslant \frac{2}{a(\tau - \tau^-)}\int_0^\infty xe^{-x^2/2}\mathrm{d}x < \infty$ and that (52) holds. The convergence of $\widehat{\text{UB2}}_{\text{MD}}(M)$ in (29) is proved similarly via Lemma B.6 and (47) in Lemma B.3.

REFERENCES

[1] 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.

[2] V. Chandrasekaran and M. I. Jordan, "Computational and statistical tradeoffs via convex relaxation," *Proceedings of the National Academy of Sciences*, vol. 110, no. 13, pp. E1181–E1190, 2013.

[3] P. Bühlmann and N. Meinshausen, "Magging: Maximin aggregation for inhomogeneous large-scale data," *Proceedings of the IEEE*, vol. 104, no. 1, pp. 126–135, Jan 2016.

[4] J.D. Gibson and J.L. Melsa, *Introduction to Nonparametric Detection with Applications*, Mathematics in science and engineering. Academic Press, 1975.

[5] P. Bühlmann and S. Geer, "Statistics for big data: A perspective," *Statistics & Probability Letters*, vol. 136, pp. 37 – 41, 2018.

[6] E. L. Lehmann and J. P. Romano, *Testing statistical hypotheses*, Springer, 3rd edition, 2005.

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

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

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

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

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

[12] B. K. Ghosh and P. K. Sen, *Handbook of Sequential Analysis*, Statistics: A Series of Textbooks and Monographs. Taylor & Francis, 1991.

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

[14] Tze Leung Lai, "Sequential analysis: Some classical problems and new challenges," *Statistica Sinica*, vol. 11, no. 2, pp. 303–351, 2001.

[15] Tze Leung Lai, "Likelihood ratio identities and their applications to sequential analysis," *Sequential Analysis*, vol. 23, no. 4, pp. 467–497, 2004.

[16] Xiaoou Li, Jingchen Liu, and Zhiliang Ying, "Generalized sequential probability ratio test for separate families of hypotheses," *Sequential Analysis*, vol. 33, no. 4, pp. 539–563, 2014.

[17] Y. Xin, H. Zhang, and L. Lai, "A low-complexity sequential spectrum sensing algorithm for cognitive radio," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 3, pp. 387–399, March 2014.

[18] 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.

[19] 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.

[20] Q.-T. Nguyen, D. Pastor, and E. L'Her, "Automatic detection of autopeep during controlled mechanical ventilation," *BioMedical Engineering OnLine*, vol. 11, no. 1, pp. 32, 2012.

[21] P. Khanduri, D. Pastor, V. Sharma, and P. K. Varshney, "Truncated sequential non-parametric hypothesis testing based on random distortion testing," *IEEE Transactions on Signal Processing*, vol. 67, no. 15, pp. 4027–4042, Aug 2019.

[22] P. Khanduri, D. Pastor, V. Sharma, and P. K. Varshney, "On random distortion testing based sequential non-parametric hypothesis testing," in *2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct 2018.

[23] 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.

[24] D. Pastor, R. Gay, and A. Gronenboom, "A sharp upper bound for the probability of error of likelihood ratio test for detecting signals in white gaussian noise," *IEEE Transactions on Information Theory*, vol. 48, no. 1, pp. 228 – 238, Jan. 2002.

[25] P. Billingsley, *Probability and Measure, 3rd edition*, Wiley, 1995.

[26] D. Pastor and F.-X. Socheleau, "Random distortion testing with linear measurements," *Signal Processing*, vol. 145, pp. 116 – 126, 2018.

[27] A. N. Shiryaev, *Probability*, Springer-Verlag, 1995.

[28] J. Jacod and Philip Protter, *Probability Essentials, 2nd edition*, Springer, 2004.