HAL archives-ouvertes.fr

Sequential Random Distortion Testing of Non-Stationary Processe

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

To cite this version:

Prashant Khanduri, Dominique Pastor, Vinod Sharma, Pramod K. Varshney. Sequential Random Distortion Testing of Non-Stationary Processe. [Research Report] IMT-Atlantique; Syracuse University, Syracuse, NY, USA; Indian Institute of Science, Bangalore, India. 2018. hal-01904985v2

HAL Id: hal-01904985

https://hal-imt-atlantique.archives-ouvertes.fr/hal-01904985v2

Submitted on 28 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

IMT Atlantique
Dépt. Signal & Communications
Technopôle de Brest-Iroise - CS 83818
29238 Brest Cedex 3
Téléphone: +33 (0)2 29 00 13 04
Télécopie: +33 (0)2 29 00 10 12
URL: **www.imt-atlantique.fr**

Prashant Khanduri, *Student Member, IEEE*, Syracuse University, NY 13244, USA
Dominique Pastor, *Member, IEEE*, IMT Atlantique, Lab-STICC, UBL, 29238 Brest, France
Vinod Sharma, *Senior Member, IEEE*, Indian Institute of Science, Bangalore, 560012, India
Pramod K. Varshney, *Life Fellow, IEEE*, Syracuse University, NY 13244, USA

Sequential Random Distortion Testing of Non-Stationary Processes

Collection des rapports de recherche d'IMT Atlantique

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

Date d'édition : February 11, 2019
Version du document : 1.2

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Summary

Abstract: Random distortion testing (RDT) [1] addresses the problem of testing whether or not a random signal, Ξ, deviates by more than a specified tolerance, τ, from a fixed value, ξ_0. The test is non-parametric in the sense that the distribution of the signal under each hypothesis is assumed to be unknown. The signal is observed in independent and identically distributed (i.i.d) additive noise. The need to control the probabilities of false alarm and missed detection while reducing the number of samples required to make a decision leads to the *Seq*RDT approach, which generalizes [2, 3]. We show that under mild assumptions on the signal, *Seq*RDT follows the properties desired by a sequential test for practical applications. We introduce the concept of buffer size for sequential testing and derive bounds on the probabilities of false alarm and missed detection, from which we choose the buffer size. Simulations show that *Seq*RDT leads to faster decision making compared to its fixed sample counterpart *Block*RDT [4]. It is robust to model mismatches compared to the Sequential Probability Ratio Test (SPRT) [5], when the actual signal is a distorted version of the assumed signal, especially at low Signal-to-Noise Ratios (SNRs).

Keywords: Sequential testing, non-parametric testing, robust hypothesis testing, random distortion testing (RDT), sequential probability ratio test (SPRT)

1. Introduction

In standard binary hypothesis testing problems, on the basis of a fixed number of observations, a decision is made between two possible statistical hypotheses, the so-called null ($\mathcal{H}_0$) and alternative ($\mathcal{H}_1$) hypotheses. The decision is generally made under the Bayesian, minimax or Neyman-Pearson frameworks. In his seminal works [5, 6], Wald moved from standard likelihood theory with fixed sample size to sequential procedures, where observations are collected and processed one after another, until a decision can be made with specified confidence. Basically, at any stage of a sequential procedure, the same decision rule is applied. This rule has three possible outcomes, instead of two: it may either 1) accept $\mathcal{H}_0$ and stop the testing, or 2) accept $\mathcal{H}_1$ and stop the testing or 3) make no decision and acquire a new observation. These three steps are repeated sequentially until a decision is reached, in which case the testing stops. In sequential testing, the sample size and the time instant when the decision is made are random. The issue is then to devise a decision rule that optimizes a certain criterion *"to achieve a tradeoff between the average observation time and the quality of the decision. ...It has been shown that the sequential procedure performs significantly better than the classical Neyman-Pearson test in the case of two simple hypotheses."* [7]. We recall that simple hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ correspond to two possible distributions for the observations. For details on Wald's approach, the reader can refer to [7].

Standard sequential testing is an extension of likelihood theory in that it assumes prior knowledge regarding the distributions of the observations under each hypothesis to derive the likelihood ratio, perhaps up to a vector parameter in case of nuisance parameters. This procedure has the following limitations. In practice, prior knowledge or good models for the distributions under each hypothesis are usually not available This is all the more detrimental when likelihood ratio tests are not robust to uncertainty or model mismatch. Moreover, many approaches in sequential testing make stationarity or independent and identically distributed (iid) assumptions on the observed process under each hypothesis [7]. Such assumptions are questionable in practice. For instance, in many practical applications such as radar, sonar and communication systems, signals of interest, distorted by the environment, are acquired in noise and are cluttered by interfering echoes. The observed random process resulting from this mixture — not necessarily additive — of signal, distortions and interferences, can thus hardly be modeled as a stationary random process with known distribution. Solutions proposed in the literature and aimed at relaxing stationarity or iid assumptions are still based on likelihood ratio tests [7] and, as such, may suffer from the lack of robustness of these tests.

To overcome these limitations, the observation process is hereafter modeled as the sum of a non-stationary signal with unknown distribution and independent noise. We introduce a theoretical framework suited for statistical signal processing applications such as those considered in [2, 3, 8], where the issue is to sequentially test the empirical mean of a non-stationary random signal. The signal has non-i.i.d samples and unknown sample distributions in additive and independent Gaussian noise. In contrast to the preliminary approach [2, 3], the theory presented below introduces a sequential procedure *Seq*RDT that guarantees an almost surely finite stopping time and error probabilities that can be rendered arbitrarily small. We introduce the notion of buffer that makes it possible to control the probabilities of false alarm and missed detection. Both the notion of buffer and the control over the probabilities of false alarm and missed detection were not present in the past works [2, 3]. We exhibit nested models and assumptions that help predict the behavior of *Seq*RDT without prior knowledge of the signal distribution and with no stationarity or iid assumption. Moreover, we compare the number of samples taken by the proposed sequential algorithm to its fixed sample counterpart *Block*RDT and show that it is faster compared to *Block*RDT, specially in the low SNR regimes. Finally, we show the robustness of the proposed sequential algorithm compared to conventional SPRT.

We explain the notation in Section 2 and state the problem in Section 3. We introduce *Seq*RDT in Section 4. In Section 5, the assumptions made in the previous sections are relaxed. The performance of *Seq*RDT is discussed in Section 6. Section 7 concludes the paper.

2. 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 $(\Omega, \mathcal{F}, \mathbb{P})$. The set of all real random variables defined on $(\Omega, \mathcal{F})$ is denoted by $\mathcal{M}(\Omega, \mathbb{R})$. Given $U \in \mathcal{M}(\Omega, \mathbb{R})$, $\mathbb{P}_U$ is the probability distribution of U: for any Borel set B of $\mathbb{R}$, $\mathbb{P}_U(B) = \mathbb{P}\big[U \in B\big]$. A domain $\mathcal{D}$ of a U is any Borel set 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 [9] with order 1/2 is denoted by $Q_{1/2}$. For any $Z \sim \mathcal{N}(\xi, 1)$, we have [10, Eq. (19) and Remark V.3]:

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

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

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

where Φ is the cumulative distribution function (cdf) of Z. $Q_{1/2}$ increases with its first argument and decreases with its second [9]. Given $\gamma \in (0, 1)$ and $\rho \in [0, \infty)$, $\lambda_\gamma(\rho)$ is defined as the unique solution in x to $Q_{1/2}(\rho, x) = \gamma$ [1, Lemma 2, statement (ii)], so that:

$$Q_{1/2}(\rho, \lambda_\gamma(\rho)) = \gamma. \tag{3}$$

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 $N \in \mathbb{N}$, the sample mean of U over the N samples $U_1, \ldots, U_N$ is $\langle U \rangle_N = \frac{1}{N} \sum_{n=1}^{N} U_n$. We say that $U, V \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}$ (resp. $\mathcal{M}(\Omega, \mathbb{R})^{[\![1,N]\!]}$) are independent if U_n and V_n are independent for each $n \in \mathbb{N}$ (resp. $n \in [\![1, N]\!]$).

3. Problem statement

Let $\Xi = (\Xi_n)_{n \in \mathbb{N}}$ be an element of $\mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}$. This random process models the random mixture of a distorted signal of interest and possible interferences. We suppose that Ξ is observed in additive and independent noise $X = (X_n)_{n \in \mathbb{N}}$, whose samples are iid with unknown cdf $\mathbb{F}$. The observation process is then $Y = (Y_n)_{n\in\mathbb{N}}$ such that $Y_n = \Xi_n + X_n$ for any $n \in \mathbb{N}$ and we write $Y = \Xi + X$. We assume that the random process under the null hypothesis (H_0) is generated form an underlying joint distribution $\mathcal{P}_0$, i.e, $\Xi = (\Xi_n)_{n\in\mathbb{N}} \sim \mathcal{P}_0$. Under alternative hypothesis (H_1), Ξ is supposed to be generated by any underlying arbitrary joint distribution other than $\mathcal{P}_0$, i.e., $\Xi = (\Xi_n)_{n\in\mathbb{N}} \nsim \mathcal{P}_0$. No assumption is made on the stationarity or the distribution of $\Xi = (\Xi_n)_{n\in\mathbb{N}}$. In this respect, the samples Ξ_n are not necessarily i.i.d. The foregoing can be summarized by writing:

$$\begin{cases} \underline{\text{Observation}} : Y = \Xi + X \in \mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}} \\ \text{with} \begin{cases} \Xi \in \mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}}, \\ X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathbb{F},\ \mathbb{F} \text{ unknown.} \end{cases} \\ \begin{cases} H_0 : & \Xi = (\Xi_n)_{n\in\mathbb{N}} \sim \mathcal{P}_0, \\ H_1 : & \Xi = (\Xi_n)_{n\in\mathbb{N}} \nsim \mathcal{P}_0 \end{cases} \end{cases} \tag{4}$$

The problem in its current form (4) 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 which independent of the distributions of these distributions. This non-parametric criterion serves as a surrogate to the actual hypotheses to be tested. This idea was first proposed in the form of Random Distortion Testing (RDT) [1] and was later extended to multiple samples by *Block*RDT [4]. We now present our 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 $\mathcal{H}_0$ and drifts significantly away from ξ_0 under $\mathcal{H}_1$. 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 our null (resp. alternative) hypothesis $\mathcal{H}_0$ (resp. $\mathcal{H}_1$) in our surrogate model to (4). Given the observation Y, the *Seq*RDT problem addressed in this paper is then the sequential testing of $\mathcal{H}_0$ against $\mathcal{H}_1$. Its formulation can be written as:

$$\begin{cases} \underline{\text{Observation}} : Y = \Xi + X \in \mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}} \\ \text{with} \begin{cases} \Xi = (\Xi_n)_{n\in\mathbb{N}} \in \mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}}, \\ X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathbb{F},\ \mathbb{F} \text{ unknown.} \\ \mathbb{E}\,[\,X_1\,] = 0, \mathrm{Var}\,(X_1) = 1 \text{ and } \mathbb{E}\left[\,X_1^3\,\right] < \infty \\ \Xi \text{ and } X \text{ are independent.} \end{cases} \\ \exists N_0 \in \mathbb{N}, \begin{cases} \mathcal{H}_0\, \forall N \geqslant N_0, 0 \leqslant |\langle\,\Xi\rangle_N - \xi_0| \leqslant \tau \text{ (a-s)} \\ \mathcal{H}_1\, \forall N \geqslant N_0, \tau < |\langle\,\Xi\rangle_N - \xi_0| \leqslant \tau_H \text{ (a-s)} \end{cases} \end{cases} \tag{5}$$

where, throughout the paper and without recalling it, $\tau \in [0,\infty)$. In the proposed formulation, the assumption on the noise third moment is technical and made clear below. The above testing model encompasses the *Block*RDT model [4] for a fixed sample size N. 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* (5) *tests whether the deviation of the signal mean* $\langle\,\Xi\rangle_N$ *around* ξ_0 *is below (or above) a specified tolerance level* τ *for the null hypothesis (or the alternate hypothesis) to be true. As* *indicated above, this non-parametric criterion serves as a surrogate to the complete knowledge of the distributions and thus avoids prior knowledge of them. Likelihood ratio based tests are, therefore, not feasible for the above problem. Moreover, the absence of assumption on the signal stationarity makes the problem even more challenging.* ■

Example 1 (Change-in-mean testing) *A sufficient condition for Ξ to follow $\mathcal{H}_0$ (resp. $\mathcal{H}_1$) is that $|\Xi_n - \xi_0| \leqslant \tau$ (a-s) (resp. $|\Xi_n - \xi_0| > \tau$ (a-s)) for any $n \in \mathbb{N}$. The* Seq*RDT problem* (5) *thus embraces the testing of a change in the mean of a Gaussian process [7, 11] 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 $\mathcal{H}_1$, $\tau = 0$ and $N_0 = 1$ in* (5). ■

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 (5) with $h_0 \leqslant \tau < |\xi_1 - \xi_0| - h_1$, $\tau_H \geqslant |\xi_1 - \xi_0| + h_1$ and $N_0 = 1$.* ■

To solve the *Seq*RDT problem (5), we introduce two 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 respective use depend on the available amount of prior information on the process.

Assumption 3.1 ((a-s) convergence of $\langle \Xi \rangle_N$)

$$\exists \langle \Xi \rangle_\infty \in \mathcal{M}(\Omega, \mathbb{R}), \lim_{N \to \infty} \langle \Xi \rangle_N = \langle \Xi \rangle_\infty \text{ (a-s)} \tag{6}$$

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

Basically, Assumption 3.1 will prove helpful to characterize the relevance of the sequential procedure introduced below, especially in asymptotic conditions. The next assumption is aimed at establishing additional results in non-asymptotic situations.

Assumption 3.2 (Bounded behavior of $|\langle \Xi \rangle_N - \xi_0|$) *There exist $\tau^- \in [0, \tau)$ and $\tau^+ \in (\tau, \infty)$ such that:*

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

Remark 3 *At this stage, it is crucial to emphasize the significance of Assumption 3.1 and Assumption 3.2, as well as the differences between them with respect to the two hypotheses in* (5). *To begin with, Assumption 3.1 is asymptotic in nature whereas Assumption 3.2 is not. Moreover, Assumption 3.2 does not require the existence of an (a-s) convergence of the empirical mean in contrast to Assumption 3.1. The two assumptions will be helpful to better control the behavior, specifically, the probability of false alarm and missed detection 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* (5). *Note also that, if Assumption 3.2 and* (6) *hold true together, $|\langle \Xi \rangle_\infty - \xi_0|$ remains bounded away from τ and all the properties stated under Assumptions 3.1 and 3.2 add to guarantee an even better grip on the behavior of the sequential testing.* ■

Remark 4 *We could have introduced the tolerances τ^- and τ^+ under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively, in place of τ in* (5) *as in Assumption 3.2. However, a single tolerance $\tau \in (\tau^-, \tau^+)$ in* (5) *keeps the model general and, more importantly, consistent with the similar models introduced in [1] and [4]. No assumption such as Assumption 3.2 was actually required in both [1] and [4]. The tests proposed in these works were designed to guarantee the probability of false alarm below a pre-specified level, while guaranteeing a maximal detection probability for fixed sample tests, without any control over this probability. As already emphasized in the previous remark, Assumption 3.2 will allow for a control of both the probability of false alarm and the probability of missed detection for fixed sample tests, as well as for the sequential testing framework proposed in this work.* ■

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

$$\mathcal{T}_{N,\gamma}(x) = \begin{cases} 0 \text{ if } \quad |\langle x \rangle_N - \xi_0| \leqslant \lambda_\gamma(\tau\sqrt{N})/\sqrt{N} \\ 1 \text{ otherwise.} \end{cases} \tag{7}$$

Proposition 3.3 below describes the asymptotic behavior of such tests under Assumption 3.1. These tests play a crucial role in the design of *Seq*RDT for the problem stated in (5).

Throughout, we use δ_N as specified in Lemma A.1.

Proposition 3.3 *Given $\gamma \in (0,1)$ and $\tau \geqslant 0$, $\mathcal{T}_{N,\gamma}$ satisfies the following asymptotic behavior for testing $\mathcal{H}_0$ against $\mathcal{H}_1$ in (5):*

(i) *we have*

$$\textit{under } \mathcal{H}_0: \quad Q_{1/2}\left(0, \lambda_\alpha(\tau\sqrt{N})\right) \leqslant \lim_{N\to\infty} \mathbb{P}\left[\,\mathcal{T}_{N,\gamma}(Y) = 1\,\right] \leqslant \gamma \tag{8}$$

$$\textit{under } \mathcal{H}_1: \quad 1 - Q_{1/2}\left(\tau_H\sqrt{N}, \lambda_\alpha(\tau\sqrt{N})\right) \leqslant \lim_{N\to\infty} \mathbb{P}\left[\,\mathcal{T}_{N,\gamma}(Y) = 0\,\right] \leqslant 1-\gamma \tag{9}$$

(ii) *under Assumption 3.1, we have,*

$$\lim_N \mathbb{P}\left[\,\mathcal{T}_{N,\gamma}(Y) = 1\,\right] = \begin{cases} 0 & \textit{under } \mathcal{H}_0 \\ 1 & \textit{under } \mathcal{H}_1 \end{cases} \tag{10}$$

Proof:

Proof of statement (i): According to (7) and Lemma A.1,

$$\begin{aligned} &\mathbb{E}\left[\,Q_{1/2}\left(\sqrt{N}\,|\langle \Xi \rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\right)\,\right] - 2\delta_N \\ &\quad \leqslant \mathbb{P}\left[\,\mathcal{T}_{N,\gamma}(Y) = 1\,\right] \\ &\qquad \leqslant \mathbb{E}\left[\,Q_{1/2}\left(\sqrt{N}|\langle \Xi \rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\right)\,\right] + 2\delta_N \end{aligned} \tag{11}$$

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

$$Q_{1/2}\left(0, \lambda_\gamma(\tau\sqrt{N})\right) - 2\delta_N \leqslant \mathbb{P}\left[\,\mathcal{T}_{N,\gamma}(Y) = 1\,\right] \leqslant Q_{1/2}\left(\tau\sqrt{N}, \lambda_\gamma(\tau\sqrt{N})\right) + 2\delta_N$$

According to (3), the upper-bound in the second inequality above equals $\gamma + 2\delta_N$. As N grows to ∞, δ_N vanishes and (8) follows.

We prove (9) similarly. We begin by combining (7) and Lemma A.1 to get

$$\begin{aligned} &1 - \mathbb{E}\left[\,Q_{1/2}\left(\sqrt{N}\,|\langle \Xi \rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\right)\,\right] - 2\delta_N \\ &\quad \leqslant \mathbb{P}\left[\,\mathcal{T}_{N,\gamma}(Y) = 0\,\right] \\ &\quad \leqslant 1 - \mathbb{E}\left[\,Q_{1/2}\left(\sqrt{N}|\langle \Xi \rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\right)\,\right] + 2\delta_N \end{aligned} \tag{12}$$

It then suffices to particularize (12) 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 3.1, $\lim_N |\langle \Xi \rangle_N - \xi_0| = |\langle \Xi \rangle_\infty - \xi_0| \leqslant \tau^- < \tau$ (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 (11) and the Lebesgue dominated convergence theorem that, under $\mathcal{H}_0$:

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

Similarly, under $\mathcal{H}_1$ and Assumption 3.1 $\lim_N |\langle \Xi \rangle_N - \xi_0| = |\langle \Xi \rangle_\infty - \xi_0| \geqslant \tau^+ > \tau$ (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 (12) and using Lebesgue's dominated convergence theorem again, we obtain:

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

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

The asymptotic result of Proposition 3.3 enhances the interest of the tests defined in (7). In addition, Proposition 3.3 (i) suggests the use of two thresholds as the false alarm and missed detection probabilities cannot both be controlled through one single threshold designed for a fixed γ. One of these thresholds should be small enough to diminish the probability of false alarm. In contrast, the other one should be sufficiently high so as to make the probability of missed detection small. Such a strategy naturally leads to a sequential approach. Proposition 3.3 (ii) also highlights the importance of Assumption 3.1 in achieving arbitrarily low false alarm and missed detection probabilities for large sample sizes. However, we need to control the number of samples, which pleads again for a sequential approach. As a matter of fact, with the thresholds chosen according to (7), we can design a sequential test capable of reducing the decision making time for the mean testing problem defined in (5), while guaranteeing certain performance levels. This sequential approach is the *Seq*RDT described below.

4. *Seq*RDT

Section 3 above has highlighted the motivation for sequential approach involving two different tests of the form (7). One of these tests must guarantee an upper-bounded false alarm probability, while the other aims at upper-bounding the missed detection probability. Given any natural number $M \geqslant N_0 - 1$, Proposition 3.3 then suggests specifying a sequential test for $\mathcal{H}_0$ against $\mathcal{H}_1$ by defining the stopping time:

$$T = \min\left\{N \in \mathbb{N} : \mathcal{D}_M(N) \neq \infty\right\}, \tag{13}$$

$$\text{with: } \begin{cases} \mathcal{D}_M(1) = \mathcal{D}_M(2) = \ldots = \mathcal{D}_M(M) = \infty, \\ \text{for } N > M, \mathcal{D}_M(N) = \begin{cases} 0 & \text{if } \ |\langle Y \rangle_N - \xi_0| \leqslant \lambda_L(N), \\ \infty & \text{if } \ \lambda_L(N) < |\langle Y \rangle_N - \xi_0| \leqslant \lambda_H(N), \\ 1 & \text{if } \ |\langle Y \rangle_N - \xi_0| > \lambda_H(N). \end{cases} \end{cases} \tag{14}$$

where $\lambda_L(N) = \lambda_\gamma(\tau\sqrt{N})/\sqrt{N}$ and $\lambda_H(N) = \lambda_{\gamma'}(\tau\sqrt{N})/\sqrt{N}$, with $\tau \in (0, \infty)$ and $\gamma, \gamma' \in (0, 1)$ must be such that $\lambda_L(N) \leqslant \lambda_H(N)$. 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 Nth sample and that the algorithm will update the statistic and repeat the test with $N + 1$ samples.

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 and noise. This will be made clearer in the coming section and Section 6.

The choice for γ and γ' can then be figured out as follows. The false alarm probability of the proposed sequential testing is:

$$\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M) \stackrel{\text{def}}{=} \mathbb{P}\,[\,\mathcal{D}_M(T) = 1\,] \quad \text{under } \mathcal{H}_0 \tag{15}$$

In the same way, the missed detection probability is:

$$\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M) \stackrel{\text{def}}{=} \mathbb{P}\,[\,\mathcal{D}_M(T) = 0\,] \quad \text{under } \mathcal{H}_1. \tag{16}$$

Since the goal of the sequential algorithm is to guarantee $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ below certain pre-specified levels α and β, respectively, Proposition 3.3 leads us to choose $\gamma = 1 - \beta$ and $\gamma' = \alpha$ with $\alpha, \beta \in (0, 1/2)$. This is a technical assumption required to ensure $\lambda_L(N) \leqslant \lambda_H(N)$ (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:

$$\lambda_L(N) = \lambda_{1-\beta}(\tau\sqrt{N})/\sqrt{N} \ \text{ and } \ \lambda_H(N) = \lambda_\alpha(\tau\sqrt{N})/\sqrt{N} \tag{17}$$

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

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

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

Proof: The proof of the above proposition follows from Lemmas B.4, B.5 and B.6 in Appendix B. ■

Proposition 4.1 (i) & (ii) imply that, asymptotically 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:

$$\mathbb{P}\big[\mathcal{D}_M(T) = \varepsilon\big] = \mathbb{P}\left(\big[\mathcal{D}_M(T) = \varepsilon\big] \cap \big[T \geq M + 2\big]\right) \ + \ \mathbb{P}\big[\mathcal{D}_M(M+1) = \varepsilon\big] \tag{18}$$

Now, using

$$\big[\mathcal{D}_M(T) = \varepsilon\big] \cap \big[T \geq M + 2\big] \subset \big[\mathcal{D}_M(M+1) = \infty\big], \tag{19}$$

we derive from (18) that:

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

Now, let us consider the first term on the rhs of (20). We can write that

$$\mathbb{P}\big[\mathcal{D}_M(M+1) = \infty\big] = 1 - \mathbb{P}\big[\mathcal{D}_M(M+1) = 1\big] - \mathbb{P}\big[\mathcal{D}_M(M+1) = 0\big]$$

By injecting this equality into (20), we get:

$$\mathbb{P}\big[\mathcal{D}_M(T) = \varepsilon\big] \leqslant 1 - \mathbb{P}\big[\mathcal{D}_M(M+1) = 1 - \varepsilon\big] \tag{21}$$

Theorem 4.2 (Asymptotics: T, $\mathbb{P}_{\mathbf{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathbf{MD}}(\mathcal{D}_M)$) *If $\alpha, \beta \in (0, 1/2)$ and Assumption 3.1 holds true, then:*
(i) $\mathbb{P}\,[\, T = \infty \,] = 0$ *under* $\mathcal{H}_0$ *and* $\mathcal{H}_1$,
(ii) $\lim_{M\to\infty} \mathbb{P}_{FA}(\mathcal{D}_M) = \lim_{M\to\infty} \mathbb{P}_{MD}(\mathcal{D}_M) = 0$

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

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

It then follows from Proposition 3.3 (ii) that $\lim_{N\to\infty} \mathbb{P}\,[\, \mathcal{D}_M(N) = \infty \,] = 0$. Thence the result.

Proof of statement (ii): The probability of false alarm is $\mathbb{P}_{\text{FA}}(\mathcal{D}_M) = \mathbb{P}\big[\mathcal{D}_M(T) = 1\big]$ under $\mathcal{H}_0$. According to (21), $\mathbb{P}_{\text{FA}}(\mathcal{D}_M) \leqslant 1 - \mathbb{P}\big[\mathcal{D}_M(M+1) = 0\big]$. The 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 (17) and Lemma A.1 that, regardless of Assumption 3.1:

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

We then derive from Assumption 3.1 & 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 $\lim_{M\to\infty} \mathbb{P}_{\text{FA}}(\mathcal{D}_M) = 0$.

The probability of missed detection $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ is $\mathbb{P}\big[\mathcal{D}_M(T) = 0\big]$ under $\mathcal{H}_1$. As above, we derive from (21), (17) and Lemma A.1 that, regardless of Assumption 3.1:

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

It then suffices to apply Assumption 3.1, Lemma A.2 and the Lebesgue dominated convergence theorem to obtain the second equality in (ii). ■

The above theorem implies that if the two hypotheses in (5) converge asymptotically away from τ, the sequential test (14) takes a decision in finite time with probability one. Moreover, the theorem also implies that $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ approach zero if the buffer size is large enough. However, in practice all the tests will be non-asymptotic in nature. Therefore, we must give some performance guarantees for the non-asymptotic regimes. In this regard, the next theorem shows that $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ are bounded away from one. Moreover, we need to fix the buffer size M before conducting the sequential test. For this purpose we make use of Assumption 3.2 to derive upper and lower bounds on $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$. We then use these bounds to choose the buffer size M. We derive two bounds in the coming theorems.

Theorem 4.3 (Non-Asymptotics: $\mathbb{P}_{\mathbf{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathbf{MD}}(\mathcal{D}_M)$) *$\mathbb{P}_{FA}(\mathcal{D}_M)$ and $\mathbb{P}_{MD}(\mathcal{D}_M)$ are bounded as:*

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

PROOF: Under $\mathcal{H}_0$, we derive from (14), (15), (17), (18) & Lemma A.1 that:

$$\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M) \geqslant \mathbb{P}\left[\mathcal{D}_M(M+1)=1\right] \geqslant \mathbb{E}\left[Q_{1/2}\left(\sqrt{M+1}|\langle\Xi\rangle_{M+1}-\xi_0|, \lambda_\alpha(\tau\sqrt{M+1})\right)\right] - 2\delta_{M+1} \qquad (24)$$

The bounds on $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ then results from the inequalities satisfied by $|\langle\Xi\rangle_{M+1}-\xi_0|$ under $\mathcal{H}_0$ and (24) for the lower bound and (22) along with the definition given in (3) for the upper bound.

Similarly, for the probability of missed detection, under $\mathcal{H}_1$, (14), (16), (17), (18) & Lemma A.1 yield

$$\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M) \geqslant \mathbb{P}\left[\mathcal{D}_M(M+1)=0\right] \geqslant 1-\mathbb{E}\left[Q_{1/2}\left(\sqrt{M+1}|\langle\Xi\rangle_{M+1}-\xi_0|, \lambda_{1-\beta}(\tau\sqrt{M+1})\right)\right] - 2\delta_{M+1}. \qquad (25)$$

We obtain the bounds on $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ from the inequalities satisfied by $|\langle\Xi\rangle_{M+1}-\xi_0|$ under $\mathcal{H}_1$ and (25) for the lower bound and (23) along with the definition given in (3) for the upper bound. ■

This theorem states that, without any assumptions, the upper bounds on $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ tend to be bounded away from unity. In particular, when noise is Gaussian distributed — which is of special interest as many systems of practical interest are modeled using this assumption — or M grows to ∞, $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ are upper-bounded by $1-\beta$ and $1-\alpha$, respectively. Moreover, the lower bounds for $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ always stay below levels α and β, respectively. The simulations later suggest that the upper bounds given in Theorem 4.3 are actually loose, i.e., $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ will be considerably smaller than these upper bounds. At this stage, Assumption 3.2 is then useful for deriving tighter bounds for $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$. These bounds are given in the next theorem and will be employed to choose appropriate buffer sizes for *Seq*RDT.

Theorem 4.4 (Non-Asymptotics: $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$) (i) *Under Assumption 3.1, $\mathbb{P}_{FA}(\mathcal{D}_M)$ and $\mathbb{P}_{MD}(\mathcal{D}_M)$ are bounded as:*

$$\begin{cases} Q_{1/2}\left(0, \lambda_\alpha(\tau\sqrt{M+1})\right) - 2\delta_{M+1} \\ \qquad \leqslant \mathbb{P}_{FA}(\mathcal{D}_M) \leqslant \mathit{UB1}_{FA}(M) \bigwedge \mathit{UB2}_{FA}(M) \\ 1 - Q_{1/2}\left(\tau_H\sqrt{M+1}, \lambda_{1-\beta}(\tau\sqrt{M+1})\right) - 2\delta_{M+1} \\ \qquad \leqslant \mathbb{P}_{MD}(\mathcal{D}_M) \leqslant \mathit{UB1}_{MD}(M) \bigwedge \mathit{UB2}_{MD}(M) \end{cases} \qquad (26)$$

where $a_1 \bigwedge a_2 = \min(a_1, a_2)$ for $a_1, a_2 \in \mathbb{R}$ and $\mathit{UB1}_{FA}(M)$, $\mathit{UB2}_{FA}(M)$, $\mathit{UB1}_{MD}(M)$ and $\mathit{UB2}_{MD}(M)$ are given at the top of the next page;

(ii) *$\mathit{UB1}_{FA}(M) \bigwedge \mathit{UB2}_{FA}(M)$ and $\mathit{UB1}_{FA}(M) \bigwedge \mathit{UB2}_{FA}(M)$ do not increase with M.*

PROOF:
Proof of statement (i): When Assumption 3.2 holds true, $0 \leqslant |\langle\Xi\rangle_{M+1}-\xi_0| \leqslant \tau^-$ under $\mathcal{H}_0$. Injecting these inequalities into (24) and (22)yield the bounds:

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

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

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

Since all the events involved in the equality above are disjoint, it follows from the Frechet inequality that:

$$\mathbb{P}\left[\mathcal{D}_M(T)=1\right] \leqslant \mathbb{P}\left[\mathcal{D}_M(M+1)=1\right] + \sum_{N=M+2}^{\infty} \mathbb{P}\left[\mathcal{D}_M(N)=1\right] \bigwedge \left(\bigwedge_{K=M+1}^{N-1} \mathbb{P}\left[\mathcal{D}_M(K)=\infty\right]\right) \qquad (27)$$

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

$$\mathrm{UB2}_{\mathrm{FA}}(M) = Q_{1/2}\left(\tau^{-}\sqrt{M+1}, \lambda_{\alpha}(\tau\sqrt{M+1})\right) + 2\delta_{M+1}$$
$$+ \sum_{N=M+2}^{\infty}\left[\left(Q_{1/2}\left(\tau^{-}\sqrt{N}, \lambda_{\alpha}(\tau\sqrt{N})\right) + 2\delta_N\right) \bigwedge \left(\bigwedge_{K=M+1}^{N-1}\left(Q_{1/2}\left(\tau^{-}\sqrt{K}, \lambda_{1-\beta}(\tau\sqrt{K})\right) - Q_{1/2}\left(0, \lambda_{\alpha}(\tau\sqrt{K})\right) + 4\delta_K\right)\right)\right]$$

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

$$\mathrm{UB2}_{\mathrm{MD}}(M) = 1 - Q_{1/2}\left(\tau^{+}\sqrt{M+1}, \lambda_{1-\beta}(\tau\sqrt{M+1})\right) + 2\delta_{M+1}$$
$$+ \sum_{N=M+2}^{\infty}\left[\left(1 - Q_{1/2}\left(\tau^{+}\sqrt{N}, \lambda_{1-\beta}(\tau\sqrt{N})\right) + 2\delta_N\right) \bigwedge \left(\bigwedge_{K=M+1}^{N-1}\left(Q_{1/2}\left(\tau_H\sqrt{K}, \lambda_{1-\beta}(\tau\sqrt{K})\right) - Q_{1/2}\left(\tau^{+}\sqrt{K}, \lambda_{\alpha}(\tau\sqrt{K})\right) + 4\delta_K\right)\right)\right]$$

For any $N \geqslant M + 1$, Lemma A.1 and Assumption 3.2 imply that under $\mathcal{H}_0$;

$$\mathbb{P}\left[\mathcal{D}_M(N) = 1\right] \leqslant Q_{1/2}\left(\tau^{-}\sqrt{N}, \lambda_{\alpha}(\tau\sqrt{N})\right) + 2\delta_N \tag{28}$$

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

$$\mathbb{P}\left[\mathcal{D}_M(K) = \infty\right] = \mathbb{P}\left[\langle Y\rangle_K - \xi_0| \geqslant \lambda_L(K)\right] - \mathbb{P}\left[\langle Y\rangle_K - \xi_0| > \lambda_H(K)\right] \tag{29}$$

It follows again from Lemma A.1 and Assumption 3.2 that, under $\mathcal{H}_0$:

$$\mathbb{P}\left[\mathcal{D}_M(K) = \infty\right] \leqslant Q_{1/2}\left(\tau^{-}\sqrt{K}, \lambda_{1-\beta}(\tau\sqrt{K})\right) - Q_{1/2}\left(0, \lambda_{\alpha}(\tau\sqrt{K})\right) + 4\delta_K \tag{30}$$

The desired bound $\mathrm{UB2}_{\mathrm{FA}}(M)$ follows by injecting (28) and (30) into (27). The bounds $\mathrm{UB2}_{\mathrm{MD}}(M)$ and $\mathrm{UB2}_{\mathrm{MD}}(M)$ on $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ follow similarly from (23), (25), (29) and Lemma A.1, under $\mathcal{H}_1$ and Assumption 3.2. Because of space limitations, the details are left to the user.

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 of Appendix B that the map $N \in \mathbb{N} \mapsto Q_{1/2}\left(\rho\sqrt{N}, \lambda_{\gamma}(\tau\sqrt{N})\right)$ is non-increasing if $\rho < \tau$ and non-decreasing if $\rho > \tau$. In addition, $(\delta_N)_{N\in\mathbb{N}}$ is a non-increasing sequence. Therefore, $\mathrm{UB1}_{\mathrm{FA}}(M)$ and $\mathrm{UB1}_{\mathrm{MD}}(M)$ do no increase with M. Careful inspection of $\mathrm{UB2}_{\mathrm{FA}}(M)$ and $\mathrm{UB2}_{\mathrm{MD}}(M)$ reveals that each term involved in these bounds is non-increasing with M. Statement (ii) follows since the minimum of two non-increasing terms is also non-increasing. ■

Theorem 4.4 makes it possible to choose the least buffer size M guaranteeing specified values for the upper bounds $\mathrm{UB1}_{\mathrm{FA}}(M) \bigwedge \mathrm{UB2}_{\mathrm{FA}}(M)$ and $\mathrm{UB1}_{\mathrm{MD}}(M) \bigwedge \mathrm{UB2}_{\mathrm{MD}}(M)$. Therefore, with the choice of an appropriate buffer size M, we can expect to control $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ under desired levels. More precisely, if we want a test that guarantees $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M) \leqslant \alpha$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M) \leqslant \beta$ for specified $0 < \alpha < 1/2$ and $0 < \beta < 1/2$, we can choose an appropriate M as follows, provided that the first three moments of X_1 are known. First, choose M_1 such that $\mathrm{UB1}_{\mathrm{FA}}(M_1) \bigwedge \mathrm{UB2}_{\mathrm{FA}}(M_1) \leqslant \alpha$. Afterwards, choose M_2 such that $\mathrm{UB1}_{\mathrm{MD}}(M_2) \bigwedge \mathrm{UB2}_{\mathrm{MD}}(M_2) \leqslant \beta$. The buffer size can then be fixed to $M = \max(M_1, M_2)$. In Section 6, we will proceed in this manner to choose the buffer size. It is, however, important to emphasize that the bounds given in Theorem 4.4 are loose, so that *Seq*RDT is expected to work for considerably smaller buffer sizes, as shown in the simulation section.

5. An extension

Suppose that, instead of Eq. (5), we have either $\mathcal{H}_0^* : \mathbb{P}\left[\forall N \geqslant N_0, |\langle\,\Xi\rangle_N - \xi_0| \leqslant \tau\right] \geqslant 1 - \varepsilon$ or $\mathcal{H}_1^* : \mathbb{P}\left[\forall N \geqslant N_0, |\langle\,\Xi\rangle_N - \xi_1| \geqslant \tau\right] \geqslant 1 - \varepsilon$ with a small positive constant $\varepsilon \leqslant \min(\alpha, \beta)$. Under the

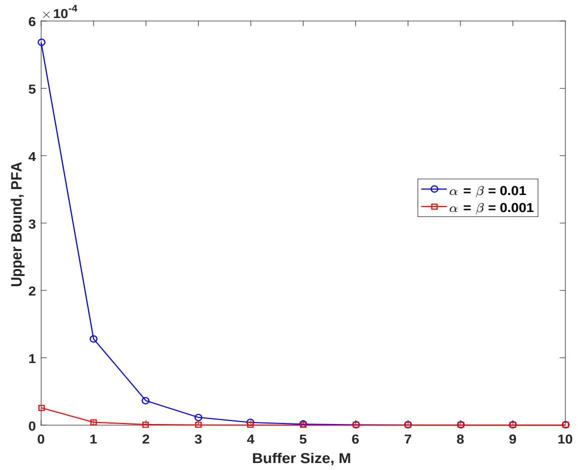

Figure 1 – Upper bound on $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ vs M for N curtailed to 500 in Theorem 4.4 with $h_0 = 1$.

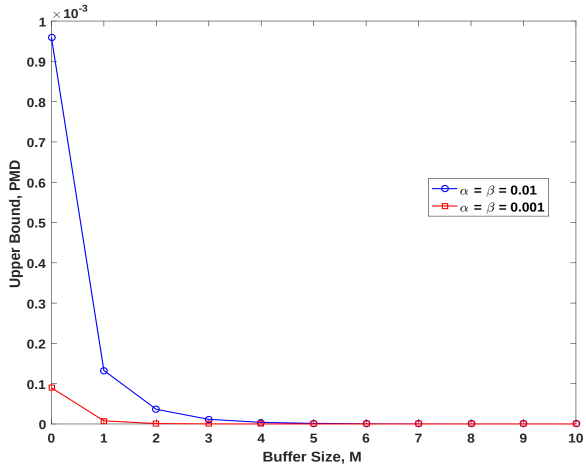

Figure 2 – Upper bound on $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ vs M for N curtailed to 500 in Theorem 4.4 with $h_0 = 1$.

assumptions of Theorem 4.4, *Seq*RDT can still be used as follows to test $\mathcal{H}_0^*$ against $\mathcal{H}_1^*$ with guaranteed bounds on the false alarm and missed detection probabilities. Indeed, given $\alpha, \beta \in (0,1)$, choose M so that $\text{UB1}_{\text{FA}}(M) \wedge \text{UB2}_{\text{FA}}(M) \leqslant \alpha$ and $\text{UB1}_{\text{MD}}(M) \wedge \text{UB2}_{\text{MD}}(M) \leqslant \beta$ in (26). Under $\mathcal{H}_0^*$, the false alarm probability $\mathbb{P}_{\text{FA}}^*$ of *Seq*RDT satisfies:

6. Experimental Results and Discussion

In this section, we perform some simulations to highlight the advantages of *Seq*RDT compared to *Block*RDT [4] and SPRT [5, 6]. 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.

6.1. The Gaussian mean change detection problem with model mismatch

In the classical Gaussian mean detection problem [7, 11], the observation is $Y_n = \Xi_n + X_n$ $(n \in \mathbb{N})$, with $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$, $\Xi_n = \xi_0$ under $\mathcal{H}_0$ and $\Xi_n = \xi_1$ under $\mathcal{H}_1$, where ξ_0 and ξ_1 are deterministic constants. This problem can be formulated in the *Seq*RDT framework (5) with $\tau = 0$ and $N_0 = 1$. But, in many applications, there might be a mismatch between the model 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 be 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 [1]. However, the *Block*RDT setup [4] and the *Seq*RDT setup as given in (5) and (14) 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 around ξ_i. 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.

6.2. Bounded Distortion

We consider the case where there exist known 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:

$$\begin{cases} \text{Under } \mathcal{H}_0 : \forall N \geqslant 1, 0 \leqslant |\langle\,\Xi\rangle_N - \xi_0| \leqslant h_0, \\ \text{Under } \mathcal{H}_1 : \forall N \geqslant 1, h_0 < h_1 - h_0 \leqslant |\langle\,\Xi\rangle_N - \xi_0| \leqslant h_0 + H. \end{cases} \tag{31}$$

This makes it possible to test the distorted signal using *Seq*RDT with the 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. Note that Assumption 3.2 is satisfied so that we can benefit from the most elaborate results established above.

For experimental illustration, we first choose $\Delta_n \sim \mathcal{U}(h_0)$, $\forall n \in \mathbb{N}$, with $h_1 \in (h_0, |\xi_1 - \xi_0| - h_0]$ 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 a few parameters only. We now discuss the algorithms and their behavior in the considered bounded regime.

SeqRDT

By referring to (31) and Example 2, one potential choice of 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 3. To perform *Seq*RDT, we must first choose an appropriate buffer size $M \geqslant N_0 - 1$. We fix M such that $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ are below specified levels α and β, respectively, by making use of the upper bounds derived in Theorem 4.4 (i). Henceforth, we choose $h_0 = 1$, $\alpha = \beta = 10^{-2}$ or $\alpha = \beta = 10^{-3}$ in our simulations. Figures 1 and 2 display the upper bounds on $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ given by Theorem 4.4 (i), when M increases. Note that the terms $\mathrm{UB2}_{\mathrm{FA}}(M)$ and $\mathrm{UB2}_{\mathrm{MD}}(M)$ in Theorem 4.4 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. As stated by Theorem 4.2 (ii), these bounds decrease. Figures 1 and 2 suggest that we can choose $M = N_0 - 1 = 0$. We have all the required information to perform detection of a change in the mean of Y_n by *Seq*RDT (5). We again emphasize that the bounds on $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ depend on the distribution of the signal through τ, τ^-, τ^+ and τ_H only. So, to choose an appropriate buffer size M, *Seq*RDT only needs to know these parameters rather than the complete distribution of the signal under both hypotheses, which is relevant in practice.

*Block*RDT

In the *Block*RDT framework, the problem of testing the mean of a random process observed over a block of N samples can be summarized as:

$$\begin{cases} \underline{\text{Observation:}}\ Y = \Xi + X \in \mathcal{M}(\Omega, \mathbb{R})^{[\![1,N]\!]} \\ \text{with} \begin{cases} \Xi \in \mathcal{M}(\Omega, \mathbb{R})^{[\![1,N]\!]}, X_1, X_2, \ldots, X_N \overset{\text{iid}}{\sim} \mathcal{N}(0,1), \\ \Xi \text{ and } X \text{ independent}, \end{cases} \\ \mathcal{H}_0 :\ |\langle\,\Xi\rangle_N - \xi_0| \leqslant \tau, \\ \mathcal{H}_1 :\ |\langle\,\Xi\rangle_N - \xi_0| > \tau. \end{cases} \tag{32}$$

The solution to this problem is proposed in [4] & [12]. 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 any such test $\mathcal{T}$ is defined as

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

and the test 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 $\mathbb{P}\left[\mathcal{T}^*(Y) = 1 \,|\, |\langle\, \Xi \rangle_N - \xi_0| > \tau\right] \geqslant \mathbb{P}\left[\mathcal{T}(Y) = 1 \,|\, |\langle\, \Xi \rangle_N - \xi_0| > \tau\right]$ 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, among which a "best" test exists. Say that 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 $\rho \in \mathcal{D} \cap (0, \infty)$, $\mathbb{P}\left[\mathcal{T}(Y) = 1 \,|\, |\langle\, \Xi \rangle_N - \xi_0| = \rho\right]$ 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 N-dimensional tests with level γ that satisfy (i) and (ii). 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 $\rho \in \mathcal{D} \cap (\tau, \infty)$,

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

According to [4] & [12], the N-dimensional test $\mathcal{T}^*_{N,\gamma} : \mathbb{R}^N \to \{0, 1\}$ defined for every $x \in \mathbb{R}^N$ by:

$$\mathcal{T}^*_{N,\gamma}(x) = \begin{cases} 0 \text{ if } & |\langle\, x \rangle_N - \xi_0| \leqslant \lambda_\gamma(\tau\sqrt{N})/\sqrt{N} \\ 1 \text{ otherwise.} & \end{cases} \tag{33}$$

is maximal in $\mathcal{K}_\gamma$: for any $\mathcal{T} \in \mathcal{K}_\gamma$, $\mathcal{T} \preceq \mathcal{T}^*_{N,\gamma}$. Let $\mathbb{P}^{\text{B-RDT}}_{\text{FA}}(N, \gamma)$ and $\mathbb{P}^{\text{B-RDT}}_{\text{MD}}(N, \gamma)$ be the probabilities of false alarm and missed detection 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 (32).

Proposition 6.1 $\mathbb{P}^{\text{B-RDT}}_{\text{FA}}(N, \gamma) \leqslant \gamma$ & $\mathbb{P}^{\text{B-RDT}}_{\text{MD}}(N, \gamma) \leqslant 1 - \gamma$.

Proof: Apply $\mathcal{H}_0$ and $\mathcal{H}_1$ specified in (32) to (11) and (12), with $\delta_N = 0$ since noise is Gaussian. ■

Under Assumption 3.2, $\mathcal{H}_0$ and $\mathcal{H}_1$ in (32) 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 (33). The resulting better performance guaranteed by *Block*RDT is given by the next proposition.

Proposition 6.2 *Under Assumption 3.2,* $\mathbb{P}^{\text{B-RDT}}_{\text{FA}}(N, \gamma) \leqslant \gamma$ & $\mathbb{P}^{\text{B-RDT}}_{\text{MD}}(N, \gamma) \leqslant 1 - Q_{1/2}\left(\tau^+\sqrt{N}, \lambda_\gamma(\tau^-\sqrt{N})\right)$.

Proof: Apply Assumption 3.2 to (11) and (12) under $\mathcal{H}_0$ and $\mathcal{H}_1$ specified by (32). ■

Given γ, the upper bound on $\mathbb{P}^{\text{B-RDT}}_{\text{MD}}(N, \gamma)$ is non-increasing with N, which can thus be chosen large enough to guarantee a specified missed detection probability. *Block*RDT applies to the detection of a change in the mean of Ξ with bounded distortions by choosing $\tau^- = h_0$ in (32) and $|\Delta_n| \leqslant \tau^-$. Given $\alpha, \beta \in (0, 1)$, we can compute the block size $N_{\text{B-RDT}}$ required to drive the upper bound on $\mathbb{P}^{\text{B-RDT}}_{\text{MD}}(N, \alpha)$ below β. The decision is made by $\mathcal{T}^*_{N,\gamma}$ with $N = N_{\text{B-RDT}}$ and $\gamma = \alpha$ in (33)

Sequential Probability Ratio Test (SPRT)

Suppose that the density 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:

$$\begin{cases} \text{If } \Lambda_N \leqslant \lambda^{\text{SPRT}}_L, \text{ decide } \mathcal{H}_0 \text{ and stops} \\ \text{If } \Lambda_N \geqslant \lambda^{\text{SPRT}}_H, \text{ decide } \mathcal{H}_1 \text{ and stops} \\ \text{If } \lambda^{\text{SPRT}}_L < \Lambda_N < \lambda^{\text{SPRT}}_H : \text{compute } \Lambda_{N+1} \text{ and repeat} \end{cases} \tag{34}$$

with $\Lambda_N = \prod_{n=1}^{N} \frac{f_1(Y_i)}{f_0(Y_i)}$, $\lambda_L^{\text{SPRT}} = \log \frac{\beta}{1-\alpha}$ and $\lambda_H^{\text{SPRT}} = \log \frac{1-\beta}{\alpha}$. The probability of false alarm $\mathbb{P}_{\text{FA}}^{\text{SPRT}}$ and the probability of missed detection $\mathbb{P}_{\text{MD}}^{\text{SPRT}}$ of SPRT are guaranteed to stay below α and β, respectively [5, 6].
In our case,

$$\Lambda_N = \sum_{n=1}^{N} \log \frac{\int_{-h_0}^{h_0} \exp\left(-\frac{(Y_n-(\xi_1+\Delta))^2}{2}\right) \mathrm{d}h}{\int_{-h_0}^{h_0} \exp\left(-\frac{(Y_n-(\xi_0+\Delta))^2}{2}\right) \mathrm{d}h}.$$

SPRT is possible when the distortion distribution is completely known. But in many practical scenarios, it is unaware of the distortion. In this case, we represent the algorithm by SPRT-MM and denote by $T_{\text{SPRT-MM}}$, $\mathbb{P}_{\text{FA}}^{\text{SPRT-MM}}$ and $\mathbb{P}_{\text{MD}}^{\text{SPRT-MM}}$, its stopping time, probability of false alarm and probability of missed detection, respectively. SPRT-MM has same thresholds as SPRT but log-likelihood updated as

$$\Lambda_N = N\frac{\xi_0^2 - \xi_1^2}{2} + (\xi_1 - \xi_0)\sum_{n=1}^{N} Y_N. \tag{35}$$

6.3. Unbounded Distortion

Because Assumption 3.2 might not be satisfied for all types of distortion, we relax Assumption 3.2 by considering unbounded distortions $\Delta_n \sim \mathcal{N}(0, h_0^2)$ for all $n \in \mathbb{N}$. The notation below remains the same as in the bounded regime case. The robustness of *Seq*RDT is however shown with respect to this experimental setting.

SeqRDT

Since $\tau^- = h_0, \tau^+ = 3h_0, \tau = 2h_0$ and $\tau_H = |\xi_1 - \xi_0| + h_0$, $\mathbb{P}[|\langle\Delta\rangle_\lambda| \leqslant \tau] = 0.9545$, $\mathbb{P}[|\langle\Delta\rangle_\lambda| \leqslant \tau^-] = 0.6827$, $\mathbb{P}[|\langle\Delta\rangle_\lambda + \xi_1 - \xi_0| > \tau] \approx 0.9772$ and $\mathbb{P}[|\langle\Delta\rangle_\lambda + \xi_1 - \xi_0| > \tau^+] \approx 0.8413$ for $|\xi_1 - \xi_0| = 4h_0$. According to Figures 1 and 2 again, we choose $M = N_0 - 1 = 0$ for $\alpha = \beta = 0.01$ and $\alpha = \beta = 0.001$. Although the (a-s) convergence in Assumption 3.2 is not satisfied with $N_0 = 1$, simulations show that this buffer size does not impact the results significantly.

*Block*RDT

For $\tau^- = h_0$ as in the bounded regime case, $\mathbb{P}[|\Delta_n| \leqslant \tau^-] = 0.6827$ for all $n \in \mathbb{N}$ so that $|\Delta_n| \leqslant \tau^-$ does not hold. However, as shown by the experimental results below, this does not really impact the performance of *Block*RDT for the detection of the distorted signal.

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 (34) 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 above with log-likelihood ratio (35).

6.4. Comparison: *Seq*RDT, *Block*RDT and SPRT

For both the bounded distortion case described in Section 6.2 and the unbounded distortion case described in Section 6.3, we define $|\xi_1 - \xi_0|$ as the Signal-to-Noise Ratio (SNR). First, in Table 1 we analyze the average number of samples taken by *Seq*RDT compared to its fixed sample size counterpart *Block*RDT for level $\alpha = \beta = 0.01$ and $\alpha = \beta = 0.001$. We choose the distortion parameter to be $h_0 = 1$ for both bounded and unbounded distortion cases. 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 chose $\tau^+ = |\xi_1 - \xi_0| - h_0$ for computing $N_{\text{B-RDT}}$. This is the best τ^+ we can choose to have the minimum possible $N_{\text{B-RDT}}$. Note from Table 1 that *Seq*RDT is faster compared to *Block*RDT for both bounded and unbounded distortions even when the best possible τ^+ is known, specially at low SNRs.

$\alpha = \beta = 0.01$					
SNR = $\lvert\xi_1 - \xi_0\rvert$		4	5	6	8
*Block*RDT	Number of samples, $N_{\text{B-RDT}}$	6	3	2	1
*Seq*RDT, $M = 0$, Bounded Regime	Average stopping time, $\mathbb{E}[T]$	2.58	2.18	2.05	2.01
*Seq*RDT, $M = 0$, Unbounded Regime	Average stopping time, $\mathbb{E}[T]$	2.79	2.37	2.23	2.17
$\alpha = \beta = 0.001$					
*Block*RDT	Number of samples, $N_{\text{B-RDT}}$	10	5	3	2
*Seq*RDT, $M = 0$, Bounded Regime	Average stopping time, $\mathbb{E}[T]$	3.74	3.05	2.81	2.70
*Seq*RDT, $M = 0$, Unbounded Regime	Average stopping time, $\mathbb{E}[T]$	3.98	3.28	3.04	2.90

Table 1 – *Seq*RDT vs *Block*RDT.

In Tables 2 and 3, we compare the average stopping times and probabilities of false alarm and missed detection of *Seq*RDT against SPRT and SPRT-MM, for different SNRs and for levels $\alpha = \beta = 0.01$ and $\alpha = \beta = 0.001$. We average the stopping times, the probabilities of false alarm and probabilities of missed detection over 10^6 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. Otherwise, the SPRT-MM probabilities of false alarm and missed detection are higher and do not 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 only a few parameters. The results of Table 3 implies that Assumption 3.2 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 known approximately.

7. Conclusion and Perspectives

In this work, we proposed a new framework *Seq*RDT for non-parametric sequential mean testing. Under mild assumptions on the signal, *Seq*RDT satisfies the properties desired for a sequential test. 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 buffer size; this buffer size helps in controlling $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ of the proposed test. Simulations showed that the *Seq*RDT approach leads to faster decision making compared to its fixed sample counterpart *Block*RDT [4]. Moreover, we showed that *Seq*RDT is robust to model mismatches compared to conventional sequential testing procedures [5, 6].

Many mathematical results established above generalize to any dimension for the observation, signal and noise. Extension of *SeqRDT* to multi-dimensional observations may thus be addressed in future work. Future directions could also concern truncated versions of *Seq*RDT, where the algorithm stops if a decision is not made within a specified time interval. Such an algorithm can be expected to require fewer assumptions on the signal compared to *Seq*RDT.

$\alpha = \beta = 0.01$					
SNR $= \lvert\xi_1 - \xi_0\rvert$		4	5	6	8
*Seq*RDT, $M = 0$	Average stopping time, $\mathbb{E}[T]$	2.58	2.18	2.05	2.01
	$\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$	1.23×10^{-4}	1.23×10^{-4}	1.18×10^{-4}	1.23×10^{-4}
	$\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$	1.27×10^{-4}	6×10^{-6}	0	0
SPRT	Average stopping time, $\mathbb{E}[T_{\mathrm{SPRT}}]$	1.40	1.14	1.04	1
	$P_{\mathrm{FA}}^{\mathrm{SPRT}}$	1.3×10^{-3}	6.34×10^{-4}	2.25×10^{-4}	1.3×10^{-5}
	$P_{\mathrm{MD}}^{\mathrm{SPRT}}$	1.3×10^{-3}	6.1×10^{-4}	1.89×10^{-4}	1.4×10^{-5}
SPRT-MM	Average stopping time, $\mathbb{E}[T_{\mathrm{SPRT\text{-}MM}}]$	1.27	1.09	1.03	1
	$P_{\mathrm{FA}}^{\mathrm{SPRT\text{-}MM}}$	3.4×10^{-3}	1.4×10^{-3}	4.33×10^{-4}	2.2×10^{-5}
	$P_{\mathrm{MD}}^{\mathrm{SPRT\text{-}MM}}$	3.4×10^{-3}	1.4×10^{-3}	3.91×10^{-4}	2.2×10^{-5}
$\alpha = \beta = 0.001$					
SNR $= \lvert\xi_1 - \xi_0\rvert$		4	5	6	8
*Seq*RDT, $M = 0$	Average stopping time, $\mathbb{E}[T]$	3.74	3.05	2.81	2.70
	$\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$	4×10^{-6}	3×10^{-6}	6×10^{-6}	3×10^{-6}
	$\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$	3×10^{-5}	0	0	0
SPRT	Average stopping time, $\mathbb{E}[T_{\mathrm{SPRT}}]$	1.74	1.29	1.10	1
	$P_{\mathrm{FA}}^{\mathrm{SPRT}}$	1.67×10^{-4}	9×10^{-5}	3.8×10^{-5}	1×10^{-6}
	$P_{\mathrm{MD}}^{\mathrm{SPRT}}$	1.78×10^{-4}	7.7×10^{-5}	3.8×10^{-5}	3×10^{-6}
SPRT-MM	Average stopping time, $\mathbb{E}[T_{\mathrm{SPRT\text{-}MM}}]$	1.50	1.18	1.06	1
	$P_{\mathrm{FA}}^{\mathrm{SPRT\text{-}MM}}$	6.89×10^{-4}	3.54×10^{-4}	1.16×10^{-4}	7×10^{-6}
	$P_{\mathrm{MD}}^{\mathrm{SPRT\text{-}MM}}$	6.88×10^{-4}	3.07×10^{-4}	1.04×10^{-4}	9×10^{-6}

Table 2 – *Seq*RDT vs SPRT for bounded regime.

Appendices

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

Proof: The Berry-Essen inequality [13, p. 374, Chap. 3] implies:

$$\sup_{x\in\mathbb{R}} \left| \mathbb{P}\left[\sqrt{N}\langle\,X\rangle_N \leqslant x\right] - \Phi(x)\right| \leqslant \delta_N \tag{36}$$

$\alpha = \beta = 0.01$					
SNR $= \lvert\xi_1 - \xi_0\rvert$		4	5	6	8
*Seq*RDT, $M = 0$	Average stopping time, $\mathbb{E}[T]$	2.79	2.37	2.23	2.17
	$\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$	2.4×10^{-3}	2.5×10^{-3}	2.4×10^{-3}	2.4×10^{-3}
	$\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$	1.1×10^{-3}	1.32×10^{-4}	2×10^{-6}	0
SPRT	Average stopping time, $\mathbb{E}[T_{\mathrm{SPRT}}]$	1.84	1.38	1.16	1.02
	$P_{\mathrm{FA}}^{\mathrm{SPRT}}$	2.2×10^{-3}	1.3×10^{-3}	7.35×10^{-4}	1.43×10^{-4}
	$P_{\mathrm{MD}}^{\mathrm{SPRT}}$	2.1×10^{-3}	1.3×10^{-3}	7.49×10^{-4}	1.26×10^{-4}
SPRT-MM	Average stopping time, $\mathbb{E}[T_{\mathrm{SPRT\text{-}MM}}]$	1.32	1.14	1.06	1.01
	$P_{\mathrm{FA}}^{\mathrm{SPRT\text{-}MM}}$	1.68×10^{-2}	8.8×10^{-3}	4.2×10^{-3}	5.94×10^{-4}
	$P_{\mathrm{MD}}^{\mathrm{SPRT\text{-}MM}}$	1.66×10^{-2}	8.6×10^{-3}	4×10^{-3}	6.56×10^{-4}
$\alpha = \beta = 0.001$					
SNR $= \lvert\xi_1 - \xi_0\rvert$		4	5	6	8
*Seq*RDT, $M = 0$	Average stopping time, $\mathbb{E}[T]$	3.98	3.28	3.04	2.90
	$\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$	3.33×10^{-4}	3.47×10^{-4}	3.19×10^{-4}	3.17×10^{-4}
	$\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$	1.24×10^{-4}	5×10^{-6}	5×10^{-6}	0
SPRT	Average stopping time, $\mathbb{E}[T_{\mathrm{SPRT}}]$	2.44	1.73	1.34	1.05
	$P_{\mathrm{FA}}^{\mathrm{SPRT}}$	2.08×10^{-4}	1.59×10^{-4}	1.03×10^{-4}	2.85×10^{-5}
	$P_{\mathrm{MD}}^{\mathrm{SPRT}}$	2.09×10^{-4}	1.54×10^{-4}	1.03×10^{-4}	2.57×10^{-5}
SPRT-MM	Average stopping time, $\mathbb{E}[T_{\mathrm{SPRT\text{-}MM}}]$	1.57	1.24	1.10	1.01
	$P_{\mathrm{FA}}^{\mathrm{SPRT\text{-}MM}}$	6.2×10^{-3}	3.5×10^{-3}	1.8×10^{-3}	2.88×10^{-4}
	$P_{\mathrm{MD}}^{\mathrm{SPRT\text{-}MM}}$	6.2×10^{-3}	3.6×10^{-3}	1.8×10^{-3}	3.05×10^{-4}

Table 3 – *Seq*RDT vs SPRT for unbounded regime.

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

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

It then follows from (36) that, for all $x \in \mathbb{R}$:

$$\psi(x\sqrt{N}, \eta\sqrt{N}) - 2\delta_N \leqslant \mathbb{P}\big[|\rho + \langle X \rangle_N| \leqslant \eta\big] \leqslant \psi(x\sqrt{N}, \eta\sqrt{N}) + 2\delta_N \tag{37}$$

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:

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

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

Lemma A.2 *If Ξ satisfies Assumption 3.1, then, for any $\gamma \in (0,1)$, we have:*

$$\lim_{N\to\infty} Q_{1/2}\Big(\sqrt{N}\,|\langle\,\Xi\rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\Big) = \mathbb{1}_{[\tau^+,\infty)}(|\langle\Xi\rangle_\infty - \xi_0|) \text{ (a-s).}$$

Proof: Under Assumption 3.1, $\lim_{N\to\infty} |\langle\,\Xi\rangle_N - \xi_0| = |\langle\Xi\rangle_\infty - \xi_0|$ (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 $N \geqslant N_0(\varepsilon,\omega)$:

$$\begin{aligned} &Q_{1/2}\Big(\sqrt{N}(|\langle\Xi\rangle_\infty(\omega) - \xi_0| - \varepsilon), \lambda_\gamma(\tau\sqrt{N})\Big) \\ &\quad\leqslant Q_{1/2}\Big(\sqrt{N}\,|\langle\,\Xi\rangle_N(\omega) - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\Big) \\ &\quad\quad\leqslant Q_{1/2}\Big(\sqrt{N}\,(|\langle\Xi\rangle_\infty(\omega) - \xi_0| + \varepsilon), \lambda_\gamma(\tau\sqrt{N})\Big) \end{aligned} \tag{38}$$

Under $\mathcal{H}_0$, (6) 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:

$$\lim_{N\to\infty} Q_{1/2}\left(\sqrt{N}\,(|\langle\Xi\rangle_\infty(\omega) - \xi_0| + \varepsilon), \lambda_\gamma(\tau\sqrt{N})\right) = 0. \tag{39}$$

Since $\mathbb{P}(\Omega' \cap \Omega_0) = 1$, Eqs. (38) and (39) imply that:

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

Under $\mathcal{H}_1$, the proof is similar. From (6), 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 $\omega \in \Omega_1$:

$$\lim_{N\to\infty} Q_{1/2}\left(\sqrt{N}\,(|\langle\Xi\rangle_\infty(\omega) - \xi_0| - \varepsilon), \lambda_\gamma(\tau\sqrt{N})\right) = 1. \tag{40}$$

Since $\mathbb{P}(\Omega' \cap \Omega_1) = 1$, it follows from (40) and (38) that:

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

which concludes the proof. ■

B. Asymptotic 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 prove (i) only since it straightforwardly implies (ii). Pose $g(\rho) = \lambda_\gamma(\rho) - \rho$ and $\theta = \Phi^{-1}(1-\gamma)$. Since $\Phi(x) + \Phi(-x) = 1$, (2) and the definition of $\lambda_\gamma(\tau)$ induce that:

$$\Phi(g(\rho)) + \Phi(g(\rho) + 2\rho) = 1 + \Phi(\theta). \tag{41}$$

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, $\lim\limits_{n\to\infty} \rho_n = \infty$. Consider any $\eta \in (0, \Phi(\theta) - \Phi(\theta - \varepsilon))$. Since $\lim\limits_{n\to\infty} \Phi(2\rho_n + \theta + \varepsilon) = 1$, there exists $N_0 \in \mathbb{N}$ such that, for all $n \geqslant N_0$:

$$\Phi(2\rho_n + \theta + \varepsilon) > 1 - \eta. \tag{42}$$

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

$$\Phi(2\rho_n + \theta - \varepsilon) < 1 + \eta. \tag{43}$$

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. (41) and (43) 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 (41) and (42), this inequality leads to:

$$1 + \Phi(\theta) > \Phi(\theta + \varepsilon) + \Phi(2\rho_n + \theta + \varepsilon) > \Phi(\theta + \varepsilon) + 1 - \eta, \tag{44}$$

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 $\rho \neq \tau$.*

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

Proof: Let $(\sigma_n)_{n\in\mathbb{N}}$ be a sequence of positive real values such that $\lim\limits_{n\to\infty} \sigma_n = 0$ and set $\rho_n = \tau/\sigma_n$ for each $n \in \mathbb{N}$. According to (1), $Q_{1/2}\big((\rho/\tau)\rho_n, \lambda_\gamma(\rho_n)\big) = \mathbb{P}\left[\left|(\rho/\tau) + X/\rho_n\right| > \lambda_\gamma(\rho_n)/\rho_n\right]$ 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 $\left|(\rho/\tau) + (X/\rho_n)\right| - \lambda_\gamma(\rho_n)/\tau_n$ converges weakly to $\mathbb{1}_{[(\rho/\tau)-1,\infty)}$. Since $\rho \neq \tau$, this weak convergence implies that $\lim\limits_{n\to\infty} \mathbb{P}\left[\left|(\rho/\tau) + X/\rho_n\right| > \lambda_\gamma(\rho_n)/\rho_n\right] = \mathbb{1}_{(\tau,\infty)}(\rho)$. 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:*

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

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

$$\mathcal{Q}(\sigma) = Q_{1/2}\big(\rho\sigma, \lambda_\gamma(\tau\sigma)\big) = 1 - \Phi(r_-(\sigma)) + \Phi(-r_+(\sigma)) \tag{45}$$

with $r_+(\sigma) = \lambda_\gamma(\tau\sigma) + \rho\sigma$ and $r_-(\sigma) = \lambda_\gamma(\tau\sigma) - \rho\sigma$. For $\rho = \tau$, it follows from (3) 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 the equality above and after some routine algebra, we obtain:

$$\lambda'_\gamma(\tau\sigma) = (1 - e^{-2\tau\sigma\lambda_\gamma(\tau\sigma)})/(1 + e^{-2\tau\sigma\lambda_\gamma(\tau\sigma)}) \tag{46}$$

where λ'_γ is the first derivative of λ_γ. We now differentiate $\mathcal{Q}$ defined by (45). 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 (46) into the equality above, we obtain:

$$\mathcal{Q}'(\sigma) = \frac{\tau}{\sqrt{2\pi}}\left(e^{-r_-^2(\sigma)/2} - e^{-r_+^2(\sigma)/2}\right)\left(\frac{\rho}{\tau} - \frac{\Delta_{-1}(\rho,\tau)}{\Delta_{+1}(\rho,\tau)}\right) \tag{47}$$

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 $\sigma \geqslant 0$, 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 $\mathcal{Q}'(\sigma) \leqslant 0$ and, thus, that $\mathcal{Q}$ is non-increasing. On the other hand, if $\rho > \tau$, we have $\rho/\tau > 1 > \Delta_-(\rho,\tau)/\Delta_+(\rho,\tau)$, so that $\mathcal{Q}$ is non-decreasing 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 (3) 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 non-decreasing.*
(P2) *The map $\rho \in (0,\infty) \mapsto Q_{1/2}(\rho,\rho)$ is non-increasing, lower-bounded by $1/2$ and $\lim_{\rho\to\infty} Q_{1/2}(\rho,\rho) = 1/2$.*
(P3) *For any $\gamma \leqslant 1/2$, the map $\rho \in (0,\infty) \mapsto \lambda_\gamma(\rho)/\rho$ is non-increasing and lower bounded by* 1.

Proof: Throughout the proof, $W \sim \mathcal{N}(0,1)$. For any pair $(\tau,\eta) \in (0,\infty)\times(0,\infty)$ and any $\sigma \in (0,\infty)$, (1) implies that:

$$Q_{1/2}(\tau/\sigma, \eta/\sigma) = \mathbb{P}\left[\, |\tau + \sigma W| > \eta \,\right]. \tag{48}$$

We then define $f(\sigma,x) = |\tau+\sigma x|^2 - \eta^2 = x^2\sigma^2 + 2\tau x\sigma + \tau^2 - \eta^2$, for every $\sigma \in (0,\infty)$ and every $x \in \mathbb{R}$. We can thus write that:

$$Q_{1/2}(\tau/\sigma, \eta/\sigma) = \mathbb{P}\left[\, f(\sigma, W) > 0 \,\right]. \tag{49}$$

Proof of statement (P1)*:* Given any $x \in \mathbb{R}$, the reduced discriminant of the quadratic polynomial $f(\bullet,x)$ is $\Delta'(x) = x^2\eta^2$. Therefore, for $x \neq 0$, $\Delta'(x) \geqslant 0$ and $f(\bullet,x)$ has two roots, $\sigma_0(x) = -(\tau/x) - (\eta/|x|)$ and $\sigma_1(x) = -(\tau/x) + (\eta/|x|)$, possibly equal. Since we assume $\eta \geqslant \tau$, $\sigma_0(x) \leqslant 0 \leqslant \sigma_1(x)$. Therefore, $\left[f(\sigma,W) > 0\right] \cap \left[W \neq 0\right] = \left[\sigma > \sigma_1(W)\right]$. Because $W \neq 0$ (a-s), we derive from the foregoing and (49) that

$$Q_{1/2}(\tau/\sigma, \eta/\sigma) = \mathbb{P}\left[\, \sigma_1(W) < \sigma \,\right]. \tag{50}$$

If $0 < \sigma < \sigma'$, $\left[\sigma_1(W) < \sigma\right] \subset \left[\sigma_1(W) < \sigma'\right]$. It follows from (50) that $Q_{1/2}(\tau/\sigma, \eta/\sigma) \leqslant Q_{1/2}(\tau/\sigma', \eta/\sigma')$. Thence the result.
Proof of statement (P2)*:* The map $\rho \in (0,\infty) \mapsto Q_{1/2}(\rho,\rho)$ is non-increasing as a consequence of (P1). Given $\rho \in (0,\infty)$, it follows from (48) that

$$Q_{1/2}(\rho,\rho) = \mathbb{P}\left[\ |1 + W/\rho| > 1\ \right]. \tag{51}$$

Since $\left[|1+W/\rho| > 1\right] \supset \left[W > 0\right]$ and $\mathbb{P}\left[\, W > 0 \,\right] = 1/2$, (51) induces that $Q_{1/2}(\rho,\rho) \geqslant 1/2$. When ρ tends to ∞, $(1/\rho)W^2 + 2W$ converges (a-s) to $2W$. Therefore, for any sequence $(\rho_n)_{n\in\mathbb{N}}$ of positive real

values such that $\lim_{N\to\infty} \rho_n = \infty$, $\mathbb{F}_{(1/\rho_n)W^2+2W} \Rightarrow \mathbb{F}_{2W}$. Because $\mathbb{F}_{2W}$ is continuous everywhere, it follows from (51) that $\lim_{n\to\infty} Q_{1/2}(\rho_n, \rho_n) = 1 - \lim_{n\to\infty} \mathbb{F}_{(1/\rho_n)W^2+2W}(0) = 1 - \mathbb{F}_{2W}(0) = 1/2$.
Proof of statement (P3)*:* Let ρ and ρ' be two positive real numbers such that $0 < \rho \leqslant \rho'$. According to (3), we have:

$$Q_{1/2}\big(\rho, \lambda_\gamma(\rho)\big) = Q_{1/2}\big(\rho', \lambda_\gamma(\rho')\big) = \gamma. \tag{52}$$

Since $\gamma \leqslant 1/2$, it follows from (P2) and (52) that $Q_{1/2}(\rho, \rho) \geqslant 1/2 \geqslant Q_{1/2}(\rho, \lambda_\gamma(\rho))$. The non-increasing behavior of $Q_{1/2}$ with its second argument implies that $\lambda_\gamma(\rho) \geqslant \rho$, so that $\lambda_\gamma(\rho)/\rho$ is lower bounded by 1. We then derive from (P1) that $x \in (0, \infty) \mapsto Q_{1/2}\big(\rho/x, \lambda_\gamma(\rho)/x\big)$ is a non-decreasing map. Since $\rho/\rho' \leqslant 1$, we have $Q_{1/2}\big(\rho, \lambda_\gamma(\rho)\big) \geqslant Q_{1/2}\big(\rho', \rho'\lambda_\gamma(\rho)/\rho\big)$. This inequality and (52) induce that $Q_{1/2}\big(\rho', \lambda_\gamma(\rho')\big) \geqslant Q_{1/2}\big(\rho', \rho'\lambda_\gamma(\rho)/\rho\big)$. The non-increasing nature of $Q_{1/2}(\rho', \bullet)$ then implies that $\lambda_\gamma(\rho') \leqslant \rho'\lambda_\gamma(\rho)/\rho$. Thereby, $\rho \in (0, \infty) \mapsto \lambda_\gamma(\rho)/\rho$ does not increase. ■

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 non-decreasing 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 $\Upsilon(\rho) = \big(1 - f(\rho)\big)\left(e^{2\rho\lambda_\gamma(\rho)} + \frac{\lambda_\gamma(\rho)+\rho}{\lambda_\gamma(\rho)-\rho}\right)$. 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. ■

References

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

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

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

[4] 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.

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

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

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

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

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

[10] 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.

[11] M. Basseville and I.V. Nikiforov, *Detection of Abrupt Changes - Theory and Application*, Prentice-Hall, Inc., 1993.

[12] Dominique Pastor and Francois-Xavier Socheleau, "Random distortion testing with linear measurements," *Signal Processing*, vol. 145, pp. 116 – 126, 2018.

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

OUR WORLDWIDE PARTNERS UNIVERSITIES - DOUBLE DEGREE AGREEMENTS

3 CAMPUS, 1 SITE

IMT Atlantique Bretagne–Pays de la Loire – **http://www.imt-atlantique.fr/**

Campus de Brest
Technopôle Brest-Iroise
CS 83818
29238 Brest Cedex 3
France
T +33 (0)2 29 00 11 11
F +33 (0)2 29 00 10 00

Campus de Nantes
4, rue Alfred Kastler
CS 20722
44307 Nantes Cedex 3
France
T +33 (0)2 51 85 81 00
F +33 (0)2 99 12 70 08

Campus de Rennes
2, rue de la Châtaigneraie
CS 17607
35576 Cesson Sévigné Cedex
France
T +33 (0)2 99 12 70 00
F +33 (0)2 51 85 81 99

Site de Toulouse
10, avenue Édouard Belin
BP 44004
31028 Toulouse Cedex 04
France
T +33 (0)5 61 33 83 65

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

© IMT Atlantique, 2018
Imprimé à IMT Atlantique
Dépôt légal : March 2018
ISSN : en cours