December 2019

SEQUENTIAL METHODS FOR NON-PARAMETRIC HYPOTHESIS TESTING

Prashant Khanduri
Syracuse University

ABSTRACT

In today's world, many applications are characterized by the availability of large amounts of complex-structured data. It is not always possible to fit the data to predefined models or distributions. Model dependent signal processing approaches are often susceptible to mismatches between the data and the assumed model. In cases where the data does not conform to the assumed model, providing sufficient performance guarantees becomes a challenging task. Therefore, it is important to devise methods that are model-independent, robust, provide sufficient performance guarantees for the task at hand and, at the same time, are simple to implement. The goal of this dissertation is to develop such algorithms for two-sided sequential binary hypothesis testing.

In this dissertation, we propose two algorithms for sequential non-parametric hypothesis testing. The proposed algorithms are based on the random distortion testing (RDT) framework. The RDT framework addresses the problem of testing whether a random signal, Ξ, observed in additive noise deviates by more than a specified tolerance, τ, from a fixed model, ξ_0. The data-based approach is non-parametric in the sense that the underlying signal distributions under each hypothesis are assumed to be unknown. Importantly, we show that the proposed algorithms are not only robust but also provide performance guarantees in the non-asymptotic regimes in contrast to the popular non-parametric likelihood ratio based approaches which provide only asymptotic performance guarantees.

In the first part of the dissertation, we develop a sequential algorithm *Seq*RDT. We first introduce a few mild assumptions required to control the error probabilities of the algorithm. We then analyze the asymptotic properties of the algorithm along with the behavior of the thresholds. Finally, we derive the upper bounds on the probabilities of false alarm (PFA) and missed detection (PMD) and demonstrate how to choose the algorithm parameters such that PFA and PMD can be guaranteed to stay below pre-specified levels. Specifically, we present two ways to design the algorithm: We first introduce the notion of a buffer and show that with the help of a few mild

assumptions we can choose an appropriate buffer size such that PFA and PMD can be controlled. Later, we eliminate the buffer by introducing additional parameters and show that with the choice of appropriate parameters we can still control the probabilities of error of the algorithm.

In the second part of the dissertation, we propose a truncated (finite horizon) algorithm, T-*Seq*RDT, for the two-sided binary hypothesis testing problem. We first present the optimal fixed-sample-size (FSS) test for the hypothesis testing problem and present a few important preliminary results required to design the truncated algorithm. Similar, to the non-truncated case we first analyze the properties of the thresholds and then derive the upper bounds on PFA and PMD. We then choose the thresholds such that the proposed algorithm not only guarantees the error probabilities to be below pre-specified levels but at the same time makes a decision faster on average compared to its optimal FSS counterpart. We show that the truncated algorithm requires fewer assumptions on the signal model compared to the non-truncated case. We also derive bounds on the average stopping times of the algorithm. Importantly, we study the trade-off between the stopping time and the error probabilities of the algorithm and propose a method to choose the algorithm parameters. Finally, via numerical simulations, we compare the performance of T-*Seq*RDT and *Seq*RDT to sequential probability ratio test (SPRT) and composite sequential probability ratio tests. We also show the robustness of the proposed approaches compared to the standard likelihood ratio based approaches.

To my family: Dadi, Nani,

Mumma, Papa, Didi, and Iti

SEQUENTIAL METHODS FOR NON-PARAMETRIC HYPOTHESIS TESTING

By

Prashant Khanduri
B.E., Kumaun University, India, 2009
M.E., Punjab Engineering College, India, 2011

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Electrical and Computer Engineering

Syracuse University
December 2019

Acknowledgments

"Give me six hours to chop down a tree and I will spend the first four sharpening the axe."

attributed to Abraham Lincoln

I would like to take this opportunity to thank my advisor Prof. Pramod K. Varshney for his invaluable guidance along with the continued intellectual and motivational support throughout my doctoral studies. Without his patience and consistent encouragement, this dissertation would not have been possible. I would like to express my gratitude to Prof. Dominique Pastor, Prof. Vinod Sharma, and Prof. Lakshmi N. Theagarajan, for mentoring me during the course of my studies. Also, thank you, Prof. Pastor, for hosting me in IMT Atlantique for the summer visit as well as for all the illuminating technical discussions we had. I got to learn a lot from those discussions. I would also like to thank my defense committee members Prof. Pinyuen Chen, Prof. Biao Chen, Prof. Mustafa Gursoy, Prof. Dominique Pastor, and Prof. Reza Zafarani for their valuable suggestions.

I would like to extend sincere thanks to my current and past "Sensor Fusion Lab" members and friends, including Aditya, Arun, Bao, Bhavya, Hao, Nianxia, Pranay, Qunwei, Raghed, Sai, Sandeep, Sid, Sijia, Shan, and Swatantra, for all the helpful technical and philosophical discussions we had during the course of this dissertation. Times spent with Aditya, Bhavya, Gogi, Manish, Pranay, Sai, and Swatantra, will forever be cherished. I would like to give special thanks to Aunty for the motherly love and making my stay in Syracuse feel like home.

I am grateful to Prof. Chandra R. Murthy for teaching me to take baby steps in research. Thanks to all the members of the SPC lab, IISc, who helped me in one way or other when I started this journey before coming to Syracuse. A heartfelt thanks to all my friends outside of my academic circle including my high-school and college friends for their unwavering support throughout this journey.

Last but not least, I will forever be indebted to my family for their sustained backing and patience during my graduate studies. I am grateful to my parents, sister, and brother-in-law, for their unconditional love, selfless care, and consistent support throughout this journey, it is needless to say that this dissertation would not have been possible without their blessings. Finally, I would like to thank my cheerful wife, Iti, whose companionship has made this journey worthwhile. Thank You for being there with me in good as well as in bad times.

Table of Contents

LIST OF TABLES

List of Figures

Chapter 1
Introduction

Past few decades have seen a proliferation of the sensor networks. It was predicted in 2017 that there will be more than 4 devices for every human on earth by the end of 2020, which might very well turn out to be an underestimate [17]. We are surrounded by sensors now more than ever. From every handheld device we use, to the watches we wear, to the cars we drive, to our homes and office buildings, all are equipped with multitude of sensors which we rely upon to make everyday decisions. These decisions are often about some phenomenon of interest and the devices have to make these decisions based on some noisy observations of the phenomenon. In this dissertation, we focus on one such decision making problem referred to as binary hypothesis testing. We define a hypothesis as [5]:

Definition 1.1. *A hypothesis is a statement about a population parameter.*

The population parameter here refers to the underlying phenomenon of interest. And the goal of a binary hypothesis test is to decide, based on the noisy observations, which of the two complementary hypotheses is true [5]:

Definition 1.2. *The two complementary hypotheses in a binary hypothesis test are called the null and the alternate hypotheses. They are denoted by* $\mathcal{H}_0$ *and* $\mathcal{H}_1$*, respectively.*

Hypothesis testing is one of the fundamental problems in the area of statistical signal processing. A majority of sensors we are surrounded with are performing some hypothesis testing on a daily basis. Some examples include, face or fingerprint based identification in mobile devices, localization and navigation sensors in self-driving vehicles, the sensors deployed in surveillance systems, fire alarms, radars detecting a target, and many more. With this increased number of sensors, the amount and the diversity of the available data has also grown exponentially. Therefore, it is not always possible to fit the data to predefined models. In cases where the data does not conform to the assumed model, providing sufficient performance guarantees for hypothesis testing becomes a challenging task. Therefore, it is important to devise methods that are model-independent, robust, provide sufficient performance guarantees for the task at hand and, at the same time, are simple to implement.

For the examples discussed above, in the context of face or fingerprint based identification in mobile devices, the goal of the designed algorithm would be to allow access of the device to the registered user and prevent an adversary from accessing the device. In this case, null hypothesis will refer to the signal corresponding to the registered user and the alternate to the signal corresponding to an adversary trying to get into the system. Note here that the algorithm designer has no knowledge of the alternate hypothesis other than the fact that it is different from the null hypothesis. These kind of problems, when we have some prior knowledge about the null hypothesis but have no knowledge of the alternate hypothesis, fall under the category of two-sided hypothesis testing. Moreover, some hypothesis testing problems can be sequential in nature, i.e., a decision about the hypotheses has to be made in an online fashion. For example, a radar detection system must decide in an online fashion, the presence or absence of a target by collecting observations sequentially. The goal of this dissertation is to develop algorithms for such two-sided sequential binary hypothesis testing that are model-independent, robust, provide sufficient performance guarantees for the task at hand and, at the same time, are simple to implement.

Next, we motivate the problem considered in this dissertation, list the major contributions of the dissertation, and finally, discuss the organization of the dissertation.

1.1 Motivation

Standard binary hypothesis testing problems [19], based on a fixed number of samples, test the null ($\mathcal{H}_0$) versus the alternate ($\mathcal{H}_1$) hypotheses, i.e.,

$$\begin{cases} \underline{\text{Observations}} : Y_1, Y_2, \ldots, Y_N \sim \mathcal{P}_\xi \\ \text{with} \begin{cases} \mathcal{H}_0 : \xi = \xi_0, \\ \mathcal{H}_1 : \xi = \xi_1 \end{cases} \end{cases}$$

where, $Y_1, Y_2, \ldots, Y_N$ represent N random observations generated from a probability distribution $\mathcal{P}_\xi$ and the goal is to make a decision based on the observations whether $\xi = \xi_0$ (hypothesis $\mathcal{H}_0$) or $\xi = \xi_1$ (hypothesis $\mathcal{H}_1$) is true. The decision is usually made through the Bayesian, minimax or Neyman-Pearson frameworks. Such tests are referred to as fixed-sample-size (FSS) tests. However, many decision making problems are inherently sequential in nature, i.e, observations are collected sequentially and are processed one after the other [4, 8, 41, 42].

$$\begin{cases} \underline{\text{Observations}} : Y_1, Y_2, \ldots \sim \mathcal{P}_\xi \\ \text{with} \begin{cases} \mathcal{H}_0 : \xi = \xi_0, \\ \mathcal{H}_1 : \xi = \xi_1 \end{cases} \end{cases}$$

where $Y_1, Y_2, \ldots$ refer to a sequence of observations generated from an underlying distribution $\mathcal{P}_\xi$. In contrast to the FSS tests, for a sequential test, the stopping time of the algorithm is random which is generally a function of the observations collected until that point. In his seminal works [39, 40], Wald proposed his celebrated sequential procedure, namely, the sequential probability ratio test (SPRT) for testing two simple hypotheses. These hypotheses are termed as simple as the values of the parameters ξ_0 and ξ_1 under both hypotheses are assumed to be precisely known [27]. For such problems, SPRT is optimal in the sense that it makes a decision faster on average, compared to all the procedures including FSS tests that guarantee the same probabilities of false alarm (PFA) and missed detection (PMD). However, in many practical scenarios the precise values of the parameters might not be available. In such cases, composite hypothesis testing models provide a popular approach to model the hypothesis testing problem [5, 27]. In this dissertation, we work with one such popular composite hypothesis testing model termed as two-sided testing [18]

$$\begin{cases} \underline{\text{Observations}} : Y_1, Y_2, \dots, \sim \mathcal{P}_\xi \\ \text{with} \begin{cases} \mathcal{H}_0 : \xi = \xi_0, \\ \mathcal{H}_1 : \xi \neq \xi_0 \end{cases} \end{cases}$$

Note that the parameter of interest is assumed to be precisely known, i.e., $\xi = \xi_0$, under the null hypothesis and there is no assumption on the parameter of interest under the alternate hypothesis, i.e., $\xi \neq \xi_0$. It is important to note that the optimality of SPRT (or SPRT based procedures) is lost when there is a mismatch between the assumed and the true signal models or the hypotheses to be tested are composite [9, 11, 16, 36, 40].

For composite binary hypothesis testing problems, variants of SPRT have been developed. Of particular interest are invariant SPRT (ISPRT), weighted SPRT (WSPRT) and generalized SPRT (GSPRT) [36]. ISPRT relies on the principle of invariance [19, 29] to reduce the composite hypothesis to a simple one, which then makes it possible to apply Wald's SPRT [40]. However, this simplification imposes strong restrictions on the hypotheses to be tested [13, 36]. On the other hand, WSPRT assigns a suitable weight function to the unknown parameters [36], although it is not always possible to upper bound the probabilities of error and find an appropriate weight function, even in asymptotic regimes. In contrast, GSPRT approximates the likelihood ratio by replacing the unknown parameters in the likelihood by their maximum likelihood (ML) estimates [20, 35, 36]. Various versions of GSPRT have been proposed in the literature with different thresholds [12, 14, 15] and most of the literature is focused on the design of one-sided tests for testing single parameter families of distributions. Moreover, it is important to note that most of the algorithms discussed above are developed for exponential families of distributions and guarantees are asymptotic, which do not upper bound the probabilities of error [12, 14, 15, 36]. Importantly, GSPRT based approaches have heavy computational complexity even for simplest of models and, therefore, are difficult to implement online [7]. The goal of this dissertation is to design sequential non-parametric binary hypothesis testing algorithms with the following properties:

- The underlying signal distribution under each hypothesis is assumed to be unknown, and importantly, the algorithms do not rely on independence (or i.i.d) assumptions on the observations either. This makes the algorithm robust to mismatches in the distributions of the signals, compared to likelihood ratio based approaches.

- The upper bounds on PFA and PMD are guaranteed to stay below pre-specified levels even in non-asymptotic regimes, which is naturally of practical interest. Moreover, the proposed algorithms are faster on an average compared to the FSS algorithm.

- The algorithms are simple in structure with low computational complexity and, therefore, are easy to implement online.

It must be noted that non-parametric sequential hypothesis testing approaches have been considered in the past, with limited to no success, as guaranteeing both PFA and PMD below certain pre-specified levels may not be feasible for such non-parametric sequential testing problems as shown in the works [9, 29, 30]. The approaches proposed in [9] are based on approximating the likelihood ratio by employing estimates of the unknown parameters to be tested. These approaches impose restrictive assumptions on these estimates to guarantee robustness and asymptotic optimality when there is a mismatch between the assumed and the true distribution. This is of limited use in practical problems, which are non-asymptotic in nature.

Moreover, SPRT and other composite hypothesis testing approaches discussed above are extensions of likelihood theory in that they assume precise knowledge of the distributions of the observations under each hypothesis to compute the likelihood ratio, perhaps up to a vector parameter in case of nuisance parameters [14, 20, 36]. 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 i.i.d. assumptions on the observed process under each hypothesis [36, 43]. Such assumptions are questionable in practice and emphasize the need for devising testing approaches that assume little knowledge of the underlying signals to be tested. To overcome these limitations, in this dissertation, we propose two such algorithms, T-*Seq*RDT and *Seq*RDT, which are based on an alternative binary hypothesis testing formulation. Importantly, we show that the proposed algorithms fulfill all the properties desired by a non-parametric sequential algorithm as listed above.

1.1.1 Main Idea

To begin with, let us assume that Y is a one-dimensional observation, with probability distribution parameterized by ξ. As discussed above, consider a two-sided hypothesis testing problem as

$$\begin{aligned} \mathcal{H}_0 &: \xi = \xi_0 \\ \mathcal{H}_1 &: \xi \neq \xi_0 \end{aligned}$$

In practice, testing the signal for a precise value of ξ_0 might be too stringent due to measurement errors, environmental fluctuations other than noise and other factors [24]. Therefore, it is reasonable to allow for some fluctuations around ξ_0 and design the null hypothesis $\mathcal{H}_0$ to test for the signal in the neighborhood of ξ_0. In this respect, we assume that Y is a corrupted observation of the signal to be tested, Ξ, and that Ξ is a random distorted version of ξ_0 with unknown distribution. The hypothesis testing problem then becomes:

$$\begin{aligned} \mathcal{H}_0 &: |\Xi - \xi_0| \leqslant \tau \\ \mathcal{H}_1 &: |\Xi - \xi_0| > \tau \end{aligned} \tag{1.1}$$

where $\tau \in [0, \infty)$ represents the distortion. Problem (1.1) was first considered in the form of random distortion testing (RDT) in [24], where the signal of interest, Ξ, with an unknown distribution, was embedded in i.i.d. Gaussian noise. The authors showed that the optimal tests (under certain criteria) were simple in design and, at the same time, independent of the signal distributions, thereby did not need the computation of the likelihood ratios in contrast to the SPRT based approaches which rely on approximating the likelihood ratios of the observations under the two hypotheses. The authors extended the RDT formulation to FSS tests, *Block*RDT [25] where the authors generalized the RDT formulation by replacing the signal Ξ, in (1.1), by its empirical mean over time. Although the detection performance improved with the number of samples, the designer had control only over PFA and no control over PMD. In this dissertation, we show that (in Chapter 4) with an additional assumption on the underlying hypothesis (1.1), the FSS test, *Block*RDT, can be designed to achieve desired PFA and PMD. However, the FSS test might need a very large number of samples to achieve the desired performance. Therefore, the need for faster decision making as well as the inherent sequential nature of many decision problems lead us to define a novel RDT based framework for sequential testing. For the proposed formulation we first propose a non-truncated (infinite horizon) sequential algorithm, *Seq*RDT and in the second part of the dissertation, we then develop a truncated (finite horizon) sequential algorithm, T-*Seq*RDT.

Below we list the main contributions and the organization of this dissertation.

1.2 Major Contributions

The goal of this dissertation is to develop sequential algorithms for non-parametric hypothesis testing. Specifically, we want the proposed algorithms to be simple in design but at the same time guarantee performance in the non-asymptotic regimes unlike the traditional composite (or non-parametric) likelihood ratio based schemes which generally only guarantee asymptotic performance. To this end, in this dissertation and as motivated earlier we use RDT based approaches to develop novel sequential algorithms which do not rely on the knowledge of the precise distributions of the underlying signals, and thereby, by design do not require the computations or even approximation of the likelihood ratios. Below we list the major contributions of the dissertation.

- We propose a novel RDT based framework for non-parametric two-sided sequential hypothesis testing and introduce two sequential algorithms to solve the two-sided binary hypothesis testing problem.

- We first motivate the structure of the tests and the thresholds used to design the sequential tests. We then propose a non-truncated algorithm, *Seq*RDT, and analyze its asymptotic performance. We analyze the properties of the thresholds and introduce the notion of a buffer which helps in controlling PFA and PMD of the algorithm. Next, we derive bounds on PFA and PMD and show that *Seq*RDT can be designed to achieve arbitrarily low PFA and PMD. Finally, we introduce additional parameters in the algorithm which we show can be chosen carefully to eliminate the buffer for *Seq*RDT.

- We introduce a truncated algorithm, T-*Seq*RDT. We design the truncation window for the algorithm using the optimal FSS test which is discussed before introducing T-*Seq*RDT along with a few important preliminary results necessary to design T-*Seq*RDT. We first analyze the properties of the proposed thresholds and then derive bounds on PFA and PMD. Importantly, we show that the designed thresholds can guarantee pre-specified PFA and PMD. Moreover, we analyze the average stopping time of T-*Seq*RDT and provide insights into the trade-off between the average stopping time and the error probabilities of T-*Seq*RDT.

- For both the algorithms *Seq*RDT and T-*Seq*RDT, we propose methods to choose the model parameters efficiently. Finally, we extend the proposed framework for testing of distorted signals and show that the proposed algorithms are not only efficient for testing of distorted signals but also are faster on average compared to the optimal FSS test. Moreover, we show the generalization of the proposed approach for different types of underlying signal (distortion) distributions. We show that the proposed algorithms are robust to mismatches compared to the likelihood ratio based approaches like SPRT, GSPRT and WSPRT.

We believe that the proposed non-parametric hypothesis testing approaches can be an alternative for the two-sided composite likelihood ratio based approaches especially when the knowledge of the underlying signal distributions are not precisely known.

Below we discuss the organization of the dissertation.

1.3 Organization of the Dissertation

The dissertation is organized into six chapters. In Chapter 2, we first introduce the problem and then discuss an important application where the proposed formulation is being applied. We make a few remarks about the proposed problem formulation and then in the latter part of the chapter, we discuss the past literature aimed at solving standard composite and non-parametric binary hypothesis testing problems.

In Chapter 3, we propose the non-truncated sequential algorithm, *Seq*RDT, to solve the hypothesis testing problem introduced in Chapter 2. We motivate the algorithm by analyzing asymptotic properties of the test statistic along with the thresholds. We then propose the algorithm and analyze the threshold properties before providing the performance guarantees for the algorithm. Importantly, in the design of the algorithm we introduce the concept of a buffer which helps in controlling PFA and PMD of *Seq*RDT. Later in the chapter, we introduce an additional parameter to avoid the need of the buffer and present the approach to design the algorithms both with and without the buffer.

In Chapter 4, we propose the truncated sequential algorithm, T-*Seq*RDT, to solve the hypothesis testing problem introduced in Chapter 2. We first introduce the optimal FSS test and provide a few preliminary results which we use to design T-*Seq*RDT. Specifically, we use the FSS test to design the truncation window (truncation time) of the algorithm, i.e., the time when we decide to stop the algorithm and make a decision if the algorithm has not reached a decision until that time instant. We then discuss the properties of the thresholds and the truncation window of the algorithm. After discussing these properties, we derive bounds on PFA and PMD of T-*Seq*RDT and show that the designed thresholds can guarantee pre-specified PFA and PMD. We also analyze the average stopping time of T-*Seq*RDT and provide insights into the trade-off between its average stopping time and the error probabilities. Finally, we propose an approach to choose the algorithm parameters which minimizes the upper bounds on the average stopping time while guaranteeing PFA and PMD to be below pre-specified levels.

In Chapter 5, we present the proposed framework for testing of distorted signals. Specifically, we test the signals for Gaussian, heavy-tailed and deterministic distortions and compare the proposed algorithms to SPRT, GSPRT and WSPRT. Moreover, we show that the proposed sequential algorithms are faster on average compared to the optimal FSS test. The simulations suggest that the proposed approaches are robust to mismatches compared to the standard likelihood approaches. Finally, in Chapter 6 we conclude the dissertation with some possible future directions we intend to pursue.

Before proceeding further we first discuss the notations along with a useful lemma we will use in the rest of the dissertation.

1.4 Notations

All the random variables are defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by $\mathcal{M}(\Omega, \mathbb{R})$ the set of all real random variables defined on $(\Omega, \mathcal{F})$. Given $U \in \mathcal{M}(\Omega, \mathbb{R})$: $\mathbb{P}_U(B) = \mathbb{P}\big[U \in B\big]$ with $\big[U \in B\big] = \{\omega \in \Omega : U(\omega) \in B\}$ when B is a Borel set of $\mathbb{R}$. A domain $\mathcal{B}$ of U is any Borel set B of $\mathbb{R}$ such that $\mathbb{P}_U(B) = 1$.

Given $\xi \in \mathbb{R}$ and $\sigma \in [0, \infty)$, $Z \sim \mathcal{N}(\xi, \sigma^2)$ implies Z is Gaussian distributed with mean ξ and variance σ^2. The Generalized Marcum Function [32] with order $1/2$ is denoted by [23, Eq. (19) and Remark V.3],

$$Q_{\frac{1}{2}}(|\xi|, \eta) = \mathbb{P}\big[\,|Z| > \eta\,\big], \tag{1.2}$$

for $Z \sim \mathcal{N}(\xi, 1)$. For any $(a, b) \in [0, \infty) \times [0, \infty)$, we have

$$Q_{\frac{1}{2}}(a, b) = 1 - \Phi(b - a) + \Phi(-b - a) \tag{1.3}$$

where Φ is the cumulative distribution function (cdf) of a standard normal Gaussian random variable, i.e., Gaussian random variable with zero mean and unit variance. Below we present a Lemma from [32, Theorem 1] to show the behavior of $Q_{\frac{1}{2}}$ with its two arguments:

Lemma 1.1 (Behavior of the Marcum function)**.** *Whatever its order, the Generalized Marcum function — and thus $Q_{\frac{1}{2}}$ — increases with its first argument and decreases with its second.*

Given $\gamma \in (0, 1)$ and $\rho \in [0, \infty)$, $\lambda_\gamma(\rho)$ is defined as the unique solution in x to $Q_{\frac{1}{2}}(\rho, x) = \gamma$ [24, Lemma 2, Statement (i)], so we have:

$$Q_{\frac{1}{2}}(\rho, \lambda_\gamma(\rho)) = \gamma. \tag{1.4}$$

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 $N \in \mathbb{N}$, the sample mean of U over the N samples $U_1, \ldots, U_N$ is denoted as:

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

The minimum of two real numbers a_1 and a_2 is denoted by $a_1 \bigwedge a_2$ and $\bigwedge_{i=1}^{n} a_i$ denotes the minimum of n real numbers $a_1, a_2, \ldots, a_n$.

1.5 Bibliographic Note

Most of the research work appearing in this dissertation has already been published at various venues and has appeared in the publications listed below.

Work Included in the Dissertation

Journal Papers:

- P. Khanduri, D. Pastor, V. Sharma, and P. K. Varshney, “Sequential Random Distortion Testing of Non-Stationary Processes", *IEEE Trans. Signal Process.*, vol. 67, no. 21, pp. 5450 - 5462, 2019.

- P. Khanduri, D. Pastor, V. Sharma, and P. K. Varshney, “Truncated Sequential Non-Parametric Hypothesis Testing Based on Random Distortion Testing", *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 4027 - 4042, 2019.

Conference Papers:

- P. Khanduri, D. Pastor, V. Sharma, and P. K. Varshney, “On Random Distortion Testing Based Sequential Non-parametric Hypothesis Testing", *Allerton*, 2018.

- P. Khanduri, D. Pastor, V. Sharma, and P. K. Varshney, “On Sequential Random Distortion Testing of Non-Stationary Processes", *ICASSP*, 2018.

Work not Included in the Dissertation

Journal Papers:

- S. Zhang, P. Khanduri, and P. K. Varshney, “Distributed Sequential Detection: Dependent Observations and Imperfect Communication", *submitted to IEEE Trans. Signal Process.*, 2019.

- P. Khanduri, L. N. Theagarajan, and P. K. Varshney, “Online Design of Optimal Precoders for High Dimensional Signal Detection", *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 4122 - 4135, 2019.

- P. Khanduri, B. Kailkhura, J. J. Thiagarajan, and P. K. Varshney, “Universal Collaboration Strategies for Signal Detection: A Sparse Learning Approach", *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1484 - 1488, 2016.

Under Preparation:

- P. Khanduri, S. Bulusu, P. Sharma, S. Kafle, and P. K Varshney, “Byzantine SVRG with Distributed Batch Gradient Computations".
- P. Khanduri, D. Pastor, V. Sharma, and P. K. Varshney, “Testing Mahalonobis Distances: Non-Parameteric Sequential Hypothesis Testing Framework".

Conference Papers:

- S. Zhang, P. Khanduri, and P. K. Varshney, “Distributed Sequential Hypothesis Testing with Dependent Sensor Observations", *Asilomar*, 2019.
- P. Khanduri, L. N. Theagarajan, and P. K. Varshney, “Online Linear Compression with Side Information for Distributed Detection of High Dimensional Signals", *SPAWC*, 2019.
- P. Khanduri, L. N. Theagarajan, and P. K. Varshney, “Online Design of Precoders for High Dimensional Signal Detection in Wireless Sensor Networks", *FUSION*, 2018.
- K. R. Varshney, P. Khanduri, P. Sharma, S. Zhang, and P. K. Varshney, “Why Interpretability in Machine Learning? An Answer Using Distributed Detection and Data Fusion Theory", *WHI, ICML* 2018.
- P. Khanduri, A. Vempaty, and P. K. Varshney, “A Unified Diversity Measure for Distributed Inference", *ICASSP*, 2017.
- P. Khanduri, V. Sharma, and P. K. Varshney, “Detection Diversity of Spatio-Temporal Data using Pitman’s Efficiency for low SNR Regimes", *IEEE GlobalSIP*, 2016.
- P. Khanduri, B. N. Bharath, and C. R. Murthy, “Coverage Analysis and Training Optimization for Uplink Cellular Networks with Practical Channel Estimation", *IEEE Globecom*, 2014.

Chapter 2
Background

In this chapter, we discuss the hypothesis testing problem to be solved. We first discuss an important real life example where the proposed frameworks are applicable. We make a few important remarks and finally, discuss the past literature where composite or non-parametric sequential hypothesis testing problems have been considered along with some popular sequential algorithms for binary composite hypothesis testing.

2.1 A Novel Sequential Testing Framework

Let $\Xi = (\Xi_n)_{n\in\mathbb{N}}$ be an element of $\mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}$. This discrete-time random process models the random mixture of a distorted signal of interest and possible interference. Standard two-sided composite hypothesis testing approaches test for

$$\mathcal{H}_0 : \xi = \xi_0$$

$$\mathcal{H}_1 : \xi \neq \xi_0$$

where ξ is the parameter of interest. As discussed earlier in Section 1.1.1 of Chapter 1, we assume the signal to be, Ξ, a distorted version of ξ. This implies that the binary testing problem can be framed in the standard hypothesis testing framework as: assume that the random process under null hypothesis ($\mathcal{H}_0$) is generated from a family of underlying joint distribution $\mathcal{P}_{\xi=\xi_0}$, i.e, $\Xi = (\Xi_n)_{n\in\mathbb{N}} \sim \mathcal{P}_{\xi=\xi_0}$, and under alternate hypothesis ($\mathcal{H}_1$), from a disjoint family of distributions, $\mathcal{P}_{\xi\neq\xi_0}$, i.e., $\Xi = (\Xi_n)_{n\in\mathbb{N}} \sim \mathcal{P}_{\xi\neq\xi_0}$. 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. We summarize the problem as:

$$\begin{cases} \underline{\text{Observation}} : Y = \Xi + X \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}} \\ \text{with} \begin{cases} \Xi, X \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}, \\ X_1, X_2, \ldots \overset{\text{i.i.d}}{\sim} \mathbb{F},\ \mathbb{F} \text{ unknown.} \end{cases} \\ \begin{cases} \mathcal{H}_0 : \Xi = (\Xi_n)_{n\in\mathbb{N}} \sim \mathcal{P}_{\xi=\xi_0}, \\ \mathcal{H}_1 : \Xi = (\Xi_n)_{n\in\mathbb{N}} \sim \mathcal{P}_{\xi\neq\xi_0} \end{cases} \end{cases}$$

This problem is difficult to tackle as very little or no knowledge of the underlying signal distributions is assumed under both hypotheses; thereby, likelihood ratio based tests (SPRT or GSPRT) are not suitable for such problems. As an alternative to likelihood ratio based approaches, we propose tests based on RDT [24], where we associate a non-parametric distance related criterion with each hypothesis which is independent of the distributions of the actual hypotheses. This non-parametric criterion serves as a surrogate to the actual hypotheses to be tested. Next, we present the model in more detail.

We assume that Ξ is observed in additive and independent Gaussian noise $X = (X_n)_{n\in\mathbb{N}}$. The observation process is $Y = (Y_n)_{n\in\mathbb{N}}$ such that $Y_n = \Xi_n + X_n$ for all $n \in \mathbb{N}$, and we write $Y = \Xi + X$. In our formulation, Ξ models the distortion around a fixed known and deterministic model ξ_0. We, however, expect that, for N large enough, 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 then say that this problem is the testing of the null hypothesis — a random event, actually — $\mathcal{H}_0 : |\langle\, \Xi\rangle_N - \xi_0| \leqslant \tau$ against the alternate hypothesis (event) $\mathcal{H}_1 : |\langle\, \Xi\rangle_N - \xi_0| > \tau$, on the basis of observation Y. The hypothesis testing problem is, therefore, given as:

$$\begin{cases} \underline{\text{Observation}} : Y = \Xi + X \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}} \\ \text{with} \begin{cases} \Xi = (\Xi_n)_{n\in\mathbb{N}} \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}, \\ X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathcal{N}(0,1), \\ \Xi \ \text{ and } \ X \ \text{ are independent.} \end{cases} \\ \exists N_0 \in \mathbb{N}, \begin{cases} \mathcal{H}_0 : \forall N \geqslant N_0,\ 0 \leqslant |\langle\, \Xi \rangle_N - \xi_0| \leqslant \tau \text{ (a-s)} \\ \mathcal{H}_1 : \forall N \geqslant N_0,\ \tau < |\langle\, \Xi \rangle_N - \xi_0| \leqslant \tau_H \text{ (a-s)} \end{cases} \end{cases} \tag{2.1}$$

where, $\tau \in [0, \infty)$ is the tolerance and $\tau < \tau_H < \infty$. Note that for the above hypothesis testing model when testing with a fixed number of samples N (a block of N samples), we refer to the designed algorithm as *Block*RDT [25]. We will discuss *Block*RDT in detail in Chapter 4 which we then use to develop the truncated sequential algorithm. Here, N_0 and the tolerances τ and τ_H are known *a priori* based on some prior knowledge (or experience) about the signal[1]. The algorithms based on formulation (2.1) have already been used for biomedical signal processing applications, specifically for the detection of Auto-positive end expiratory pressure (Auto-PEEP) [22] which we discuss below. Moreover, for illustration purposes, below we give a few simple examples where formulation (2.1) can be easily used. Before that we make a few useful remarks about the model discussed above:

Remark 2.1. *Note that the above problem* (2.1) *tests whether the deviation of the signal mean* $\langle\, \Xi \rangle_N$ *around a fixed model* 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.* □

Remark 2.2. *Note that the RDT framework of* (1.1) *is the same as that given in* (2.1) *for* $N = 1$.

[1]This knowledge can follow from machine learning training procedures or be based on some statistical knowledge of the signal. Discussion of these procedures is beyond the scope of this dissertation.

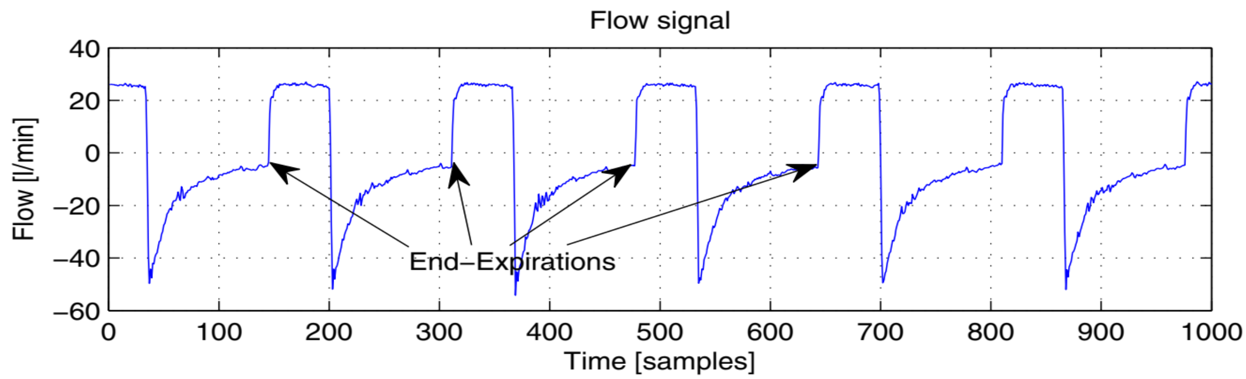

Fig. 2.1: Example 1: AutoPEEP Detection, An example of flow signal.

The formulation in (2.1) *generalizes the RDT framework of* (1.1) *for testing with multiple samples, i.e., for the FSS test, BlockRDT, and sequential hypothesis testing approaches. An alternative testing problem would be to use* $\langle\, |\Xi_n - \xi_0|\rangle_N$ *instead of* $|\langle\, \Xi\rangle_N - \xi_0|$ *in* (2.1)*. This would allow a larger class of distortions. However, designing such a test would require stronger assumptions of* $|\Xi_n - \xi_0| \leqslant \tau$ *under* $\mathcal{H}_0$ *and* $|\Xi_n - \xi_0| > \tau$ *under* $\mathcal{H}_1$ *for all* $n \in \mathbb{N}$ *in comparison to the condition of* (2.1)*, where introducing* N_0 *in* (2.1) *gives the designer the flexibility to design the testing problem for models when the condition* $|\langle\, \Xi\rangle_N - \xi_0| \leqslant \tau$ *(resp.* $|\langle\, \Xi\rangle_N - \xi_0| > \tau$*) might not hold true for smaller values of* $N \in \mathbb{N}$ *under* $\mathcal{H}_0$ *(resp.* $\mathcal{H}_1$*).* □

Example 1 (Automatic detection of AutoPEEP)**.** AutoPEEP is a common ventilatory abnormality that is usually observed in patients with severe asthma or chronic obstructive pulmonary disease [22]. The presence of AutoPEEP indicates an insufficient expiratory time and it is measured by a pressure transducer of a mechanical ventilator [37, 38]. It can be visually observed and detected through the flow signal as depicted in Figure 2.1. AutoPEEP is present if the flow signal at the end of each expiration as indicated in Figure 2.1 does not return to zero. The detection of AutoPEEP usually requires an expert at the patient's bedside [3]. To eliminate or reduce this human intervention an RDT based formulation was presented in [22] where a detector for automatic detection of AutoPEEP was developed. In the design, the authors accounted for various factors other

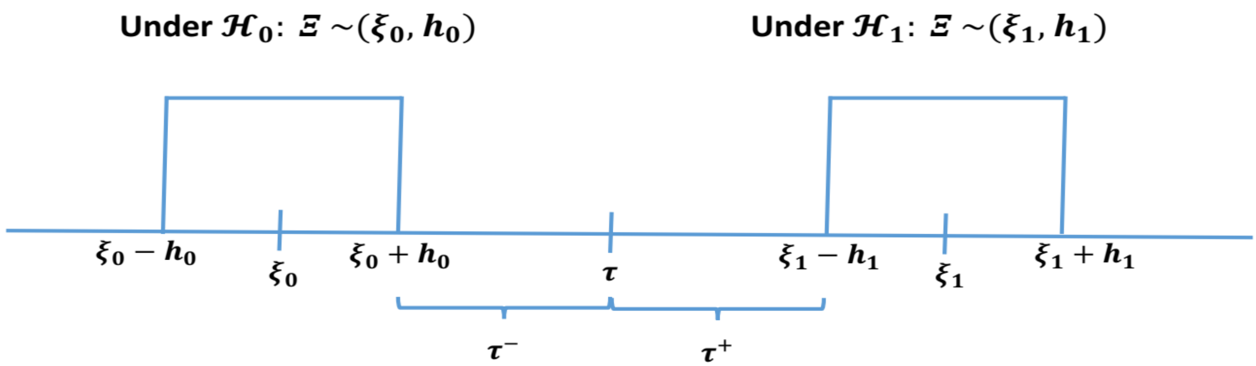

Fig. 2.2: Example 3: Bounded Regime Testing.

than noise like the mechanical vibration of the air tube, the patient movement, the electro-magnetic interference, by introducing a tolerance, τ, in the hypothesis test as discussed above in (2.1).

Example 2 (Gaussian-mean testing)**.** The sequential testing problem (2.1) embraces the testing of the mean of a Gaussian process [36] when, given two known real values ξ_0 and ξ_1 with $\xi_0 \neq \xi_1$, where we have

$$\text{Under } \mathcal{H}_0 : \Xi_n = \xi_0 \quad \text{for all} \quad n \in \mathbb{N}$$
$$\text{Under } \mathcal{H}_1 : \Xi_n = \xi_1 \quad \text{for all} \quad n \in \mathbb{N}$$

Note that here we have $\tau = 0$, $N_0 = 1$ and where the observations Y are corrupted version of the signal Ξ embedded in $X \sim \mathcal{N}(0, 1)$ in (2.1). □

Example 3 (Bounded regime testing)**.** Given $\xi \in \mathbb{R}$ and $h \in [0, \infty)$, we 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 $n \in \mathbb{N}$. From Figure 2.2, 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). We 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 $(\xi_0 - h_0, \xi_0 + h_0) \cap$

$(\xi_1 - h_1, \xi_1 + h_1) = \emptyset$, please refer to Figure 2.2. When (ξ_0, h_0) and (ξ_1, h_1) are separate, testing $\mathcal{H}_0$ against $\mathcal{H}_1$ is the particular problem (2.1) with $h_0 \leqslant \tau < |\xi_1 - \xi_0| - h_1$, $\tau_H \geqslant |\xi_1 - \xi_0| + h_1$ and $N_0 = 1$. □

2.1.1 Sequential Tests: Definitions

We first define a sequential test and the class of algorithms we are interested in. In later chapters, we develop two algorithms in the class $\mathcal{C}(\alpha, \beta)$ as defined in this section.

Following the standard terminology [36] with a slight change of notation, we define a sequential test for the binary hypothesis testing problem (2.1) as a pair $(T, \mathcal{D})$, where T is the stopping time and $\mathcal{D}$ is a decision rule taking values in $\{0, 1, \infty\}$ such that, for each $1 \leq N \leq T$:

$$\mathcal{D}(N) = \begin{cases} 0 & \mathcal{H}_0 \text{ is accepted} \\ 1 & \mathcal{H}_1 \text{ is accepted} \\ \infty & \text{repeat the test with } N+1 \text{ samples.} \end{cases} \tag{2.2}$$

Further, the stopping time T for the non-truncated (infinite horizon) test is defined as:

$$T = \min\{N \in \mathbb{N} : \mathcal{D}(N) \neq \infty\}. \tag{2.3}$$

It must also be noted that the stopping time T is a random variable and is a function of the random observations. From the definition, we notice that the non-truncated test can potentially run forever. Similarly, for the truncated (finite horizon) test we define the stopping time T as:

$$T = \inf\{N \in \mathbb{N} : N \leqslant N_0 + W^* - 1, \mathcal{D}(N) \neq \infty\}. \tag{2.4}$$

where the condition $N \leqslant N_0 + W^* - 1$ guarantees that $T \leq N_0 + W^*$ and we refer to W^* as the truncation window of the algorithm. Note that $W^* = \infty$ for non-truncated sequential procedures. It is also worth noticing that FSS tests are particular cases of tests $(T, \mathcal{D})$, with stopping time being a deterministic constant $T = N$ and $\mathcal{D}$ valued in $\{0, 1\}$.

Now to define the class of tests which are of interest to us, we define $\mathscr{C}(\alpha, \beta)$ as: Given two specified levels α and β in $(0, 1/2)$, we define the class of tests:

$$\mathscr{C}(\alpha, \beta) = \{(T, \mathcal{D}) : \sup_{\Omega_0} \mathbb{P}_{\text{FA}}(\mathcal{D}) \leq \alpha, \sup_{\Omega_1} \mathbb{P}_{\text{MD}}(\mathcal{D}) \leq \beta\} \tag{2.5}$$

with

$$\Omega_0 = \{\Xi \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}} : \forall N \geqslant N_0, |\langle \Xi \rangle_N - \xi_0| \leqslant \tau \ \text{(a-s)}\},$$

and

$$\Omega_1 = \{\Xi \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}} : \forall N \geqslant N_0, |\langle \Xi \rangle_N - \xi_0| > \tau \ \text{(a-s)}\}$$

and where

$$\mathbb{P}_{\text{FA}}(\mathcal{D}) \stackrel{\text{def}}{=} \mathbb{P}\left[\mathcal{D}(T) = 1\right], \quad \text{under } \mathcal{H}_0, \tag{2.6}$$

is the PFA and

$$\mathbb{P}_{\text{MD}}(\mathcal{D}) \stackrel{\text{def}}{=} \mathbb{P}\left[\mathcal{D}(T) = 0\right], \quad \text{under } \mathcal{H}_1. \tag{2.7}$$

is the PMD.

We are interested in the class of tests $\mathscr{C}(\alpha, \beta)$ which implies that a given test $(T, \mathcal{D})$ belongs to $\mathscr{C}(\alpha, \beta)$ if it can guarantee the PFA and PMD to stay below pre-specified levels α and β, respectively. Throughout this work, the levels α and β are chosen in the interval $(0, 1/2)$. The goal of this work is to first design a non-truncated and then a truncated sequential algorithm belonging to $\mathscr{C}(\alpha, \beta)$ which solves Problem 2.1. Moreover, we desire that the proposed algorithms make a decision faster on average compared to the optimal FSS test, *Block*RDT discussed in Chapter 4.

Before proceeding further, next we discuss in detail the popular methods addressed in the literature on sequential methods that deals with composite hypothesis testing.

2.2 Literature Review: Composite Hypothesis Testing

In this section, we discuss composite as well as the non-parametric approaches for sequential binary hypothesis testing and discuss how generalized SPRT (GSPRT) based approaches are not a good choice for the two-sided hypothesis testing models discussed in this work. Below, we discuss the approaches for the two-sided hypothesis testing problem considered in the work. Specifically, we consider the two-sided composite hypothesis testing problem of the mean of a Gaussian process with unknown mean under $\mathcal{H}_1$, i.e.,

$$\begin{cases} \underline{\text{Observation}} : Y = \xi + X \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}} \\ \text{with} \quad \begin{cases} X \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}, \\ X_1, X_2, \ldots \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1) \end{cases} \\ \begin{cases} \mathcal{H}_0 : \xi = \xi_0, \\ \mathcal{H}_1 : \xi \neq \xi_0. \end{cases} \end{cases}$$

Note that this is a particular case of the general problem considered above (cf Eq. (2.1) earlier), where the signal is a randomly distorted version of ξ. Now we discuss the methods to tackle composite hypothesis testing problems and the assumptions each method needs to impose on the signal model to perform the two-sided hypothesis tests as mentioned above. To the best of our knowledge, there are three ways to tackle composite or non-parametric hypothesis testing problems. Below we briefly discuss each of the methods:

1. **Principle of invariance:** For some hypothesis testing problems, one may use the principle of invariance [18, 29] to reduce the composite hypothesis to a simple one, which makes it possible to apply Wald's SPRT [40]. This type of test is referred to as an invariant SPRT (ISRPT). This reduction is useful but, in practice, it can be applied in only a handful of cases as it imposes strong restrictions on the hypothesis to be tested [13, Sec 2]. Please look at examples in [36, Chapter 3, Sec 6]. On the other hand, the hypothesis tests considered in this

dissertation are fairly general where the composite hypothesis cannot be reduced to simple hypothesis with invariant statistics.

2. **Composite hypothesis tests:** Composite hypothesis testing procedures are most popular in the literature and enjoy various asymptotic optimality properties for sequential testing of composite hypotheses [36, Chapter 5]. Composite sequential testing can be carried out in three frameworks namely, the generalized sequential likelihood ratio tests (GSLRT) or generalized sequential probability ratio tests (GSPRT), the minimax tests and weighted sequential probability ratio tests (WSPRT). Next we discuss these approaches one-by-one:

   a) **Generalized sequential probability ratio tests (GSPRT):** GSPRT or GSLRT compares the generalized likelihood ratio with predetermined thresholds. In its most general form, it is represented as

$$\begin{cases} \text{If } \hat{\Lambda}_N \leq \lambda_L, \text{ decide } \mathcal{H}_0 \text{ and stop;} \\ \text{If } \hat{\Lambda}_N \geq \lambda_H, \text{ decide } \mathcal{H}_1 \text{ and stop;} \\ \text{If } \lambda_L \leq \hat{\Lambda}_N \leq \lambda_H, \text{ compute } \hat{\Lambda}_{N+1} \text{ and repeat;} \end{cases}$$

   where $\hat{\Lambda}_N$ is the generalized likelihood ratio of the observations, i.e., the unknown parameter ξ under $\mathcal{H}_1$ denoted by ξ_1 is replaced with its maximum-likelihood (ML) estimate [20,35,36]. The thresholds λ_L and λ_H are chosen as

$$\lambda_L = \frac{\beta}{1-\alpha} \text{ and } \lambda_H = \frac{1-\beta}{\alpha}.$$

   which are the same as the SPRT thresholds. Various versions of GSPRT have been proposed in the literature with similar test statistics but different thresholds and testing rules [12, 14, 15]. Below, we list the drawbacks of the above approaches compared to the RDT based approaches proposed in this dissertation:

   - The literature on GSRPT is largely focused on the design of one-sided tests for testing single parameter hypotheses [12, 14, 15, 36]. In this sense, this dissertation takes a step towards advancing the state-of-the-art and addresses a two-sided hypothesis testing problem, where the signal distribution is unknown, but the signal is embedded in Gaussian noise.

- Most of the literature on sequential composite hypothesis testing provides guarantees when the observations follow *"exponential parameter families"* of distributions. Moreover, the guarantees are usually asymptotic and are provided in terms of minimizing the Bayesian cost [12, 14, 15, 36]. Usually, the asymptotic order of PFA and PMD are derived without any upper bound on the error probabilities, as it is difficult to bound the error probabilities. In contrast, for the algorithms proposed in this dissertation, we do not claim optimality. However, we guarantee upper bounds on the PFA and PMD, even in non-asymptotic regimes which are naturally of more practical interest compared to the asymptotic regimes.
- To derive asymptotic results, most of the literature assumes independence of observations over time, whereas our approach makes no such assumption on the signal model and, at the same time, guarantees performance in the non-asymptotic regimes as well.
- Most importantly, GSRPT needs the knowledge of the distributions of the observations whereas, in contrast, the algorithms proposed in this work do not rely on the knowledge of the signal distributions.
- Moreover, GSRPT based approaches have heavy computational complexity even in simplest of cases and for simplest of models. Therefore, it is difficult to implement them online [7], whereas the algorithms proposed in this dissertation are not only simple in structure but also have very low computational complexity.

Most of the same issues that exist with GSPRT also exist with minimax formulations [36, Chapter 5, Sec 3]. Here, we do not discuss them in detail to avoid duplication.

c) **Weighted sequential probability ratio test (WSPRT):** WSPRTs can be thought of as

Bayesian equivalent of FSS tests, in the sense that some suitable weight functions are assigned to the unknown parameters [36, Chapter 5] and the likelihood ratio is averaged over these weight functions. However, it is not always possible to upper bound the error probabilities and find an appropriate weight function even with asymptotic analysis. On the other hand, as discussed earlier, the algorithms proposed in this dissertation provide upper bounds on error probabilities even in non-asymptotic regimes, even without the knowledge of signal distributions or any independence assumptions. For simulation purposes in Chapter 5, we consider a problem where WSPRT proposed by Wald [36, 40] for Gaussian mean testing with unknown mean and variance can be applied. However, it must be noted that the test proposed by Wald can only be applied when the received observations are Gaussian distributed [40, Chapter 4], whereas for the models considered in this work the observations are modeled as *" unknown distributed signal + Gaussian noise"*.

3. **Non-parametric approaches:** In [9, 29–31], the authors mention that guaranteeing both PFA and PMD to stay below certain pre-specified levels for non-parametric sequential hypothesis testing approaches may not be feasible. Also, the authors propose some non-parametric approaches for sequential binary hypothesis testing. The proposed approaches are based on the likelihood ratio tests (GSPRT) in the sense that they approximate the likelihood statistic with estimates of the unknown parameters to be tested. With appropriate assumptions on the estimates, which can be restrictive in many practical scenarios, the authors show the asymptotic optimality of the operating characteristic and the average sample number of the sequential tests. However, in the simulations in Chapter 5 we show that this asymptotic optimality is of limited practical use as the non-asymptotic performance of the algorithm is not satisfactory even for simple models.

The above discussion implies that our model is more general than the above mentioned approaches as we do not need to know the signal distribution or even if it is deterministic or not. More importantly, the proposed tests with little knowledge (cf Assumption 2.1) of the signal distributions are able to provide performance guarantees. Also, many of the approaches either make i.i.d. (or independence) assumptions even for asymptotic analysis. In contrast, our proposed tests do not require the signals to be iid across time and at the same time the associated analyses are non-asymptotic in nature. Furthermore, the proposed tests are simple in structure as well as in computational complexity compared to the above tests where the test statistic as well as the thresholds might not even be available in closed form in many cases.

2.3 Summary

In this chapter, we presented the sequential framework we consider in this dissertation. We then discussed a few examples along with a few remarks detailing the scenarios where the proposed algorithms are applicable. Then we defined a sequential test along with the stopping time for both non-truncated and truncated formulations. We also defined the class of tests, $\mathscr{C}(\alpha, \beta)$, which are of interest to us. Finally, we discussed some key works on the composite hypothesis testing problem and highlighted the differences of each of the algorithms with the approach proposed in this dissertation.

CHAPTER 3
SEQUENTIAL RANDOM DISTORTION TESTING

3.1 Introduction

In this chapter, we propose an algorithm to solve the sequential hypothesis testing problem presented in Chapter 2. We propose a novel sequential algorithm, *Seq*RDT. We first introduce a few assumptions required to design the sequential algorithm. Then we motivate the algorithm design by analyzing the properties of the proposed test statistic. Next, we propose the algorithm and analyze its asymptotic properties. We introduce the notion of buffer which is then used to control the probabilities of error of the algorithm. We then derive upper bounds on PFA and PMD of the algorithm and give an approach to choose an appropriate buffer size. Importantly, we show that without any prior knowledge of the signal distribution, *Seq*RDT guarantees pre-specified values of PFA and PMD, whereas, in contrast, the likelihood ratio based tests need precise knowledge of the signal distributions under each hypothesis. We also introduce an additional parameter in the algorithm and present another algorithm which eliminates the need of the buffer. Finally, we present the complete algorithm with all the steps and conclude the chapter.

3.1.1 Problem Statement

In this section, we introduce some assumptions required to control the PFA and PMD corresponding to the proposed algorithm. Specifically, we first introduce an assumption which is required to analyze the asymptotic properties of the algorithm. Then we introduce another assumption, which makes it possible to control the probabilities of error of the proposed algorithm in non-asymptotic regimes. Before introducing the assumptions we state the problem again:

$$\begin{cases} \underline{\text{Observation}} : Y = \Xi + X \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}} \\ \text{with} \begin{cases} \Xi = (\Xi_n)_{n \in \mathbb{N}} \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}, \\ X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathcal{N}(0,1), \\ \Xi \text{ and } X \text{ are independent.} \end{cases} \\ \exists N_0 \in \mathbb{N}, \begin{cases} \mathcal{H}_0 : \forall N \geqslant N_0,\ 0 \leqslant |\langle\, \Xi \rangle_N - \xi_0| \leqslant \tau \text{ (a-s)} \\ \mathcal{H}_1 : \forall N \geqslant N_0,\ \tau < |\langle\, \Xi \rangle_N - \xi_0| \leqslant \tau_H \text{ (a-s)} \end{cases} \end{cases} \tag{3.1}$$

where, $\tau \in [0, \infty)$ is the tolerance and $\tau < \tau_H < \infty$. Now the goal is to solve the above problem in a sequential manner and propose an algorithm such that the algorithm belongs to the class $\mathscr{C}(\alpha, \beta)$ as defined in Chapter 2. For this purpose, we introduce the following assumptions.

3.1.2 Assumptions

To solve problem (3.1) sequentially, we introduce two assumptions. The first assumption 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 regarding the process.

Assumption 1 ((a-s) convergence of $\langle\, \Xi \rangle_N$)**.** *There exist $\tau^- \in [0, \tau)$ and $\tau^+ \in (\tau, \infty)$, such that:*

$$\begin{cases} \textit{Under } \mathcal{H}_0 : \limsup_{N\to\infty} |\langle\, \Xi \rangle_N - \xi_0| \leqslant \tau^- \textit{ (a-s)}, \\ \textit{Under } \mathcal{H}_1 : \liminf_{N\to\infty} |\langle\, \Xi \rangle_N - \xi_0| \geqslant \tau^+ \textit{ (a-s)}. \end{cases}$$

Remark 3.1. *The above Assumption **[(a-s) convergence of** $\langle\, \Xi \rangle_N$**]** is automatically satisfied if the signal, Ξ, is stationary and ergodic [2, Ch. 4, Sec. 24]. In this case, there exists $\xi \in \mathbb{R}$ such that $\mathbb{E}\,[\, \Xi_n \,] = \xi$ for every $n \in \mathbb{N}$, so that **[(a-s) convergence of** $\langle\, \Xi \rangle_N$**]** holds with $\langle \Xi \rangle_\infty = \xi$ and $\xi \in \{\xi_0, \xi_1\}$ with $\xi = \xi_0$ and $\xi = \xi_1$ under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively.* □

Basically, Assumption **[(a-s) convergence of** $\langle\, \Xi \rangle_N$**]** will prove helpful to characterize the relevance of the sequential procedure introduced later in the chapter in the asymptotic regime. The next assumption is aimed at establishing additional results in non-asymptotic situations.

Assumption 2 (Bounded behavior of $|\langle\, \Xi \rangle_N - \xi_0|$)**.** *There exist $\tau^- \in [0, \tau)$ and $\tau^+ \in (\tau, \infty)$ such that:*

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

Now in the following remark we discuss the implications of the two assumptions discussed above.

Remark 3.2. *At this stage, it is crucial to emphasize the significance of Assumption **[(a-s) convergence of** $\langle\, \Xi \rangle_N$**]** and Assumption **[Bounded behavior of** $|\langle\, \Xi \rangle_N - \xi_0|$**]**, as well as the differences between them with respect to the two hypotheses in* (3.1).

*As can be seen, the Assumption **[(a-s) convergence of** $\langle\, \Xi \rangle_N$**]** addresses the asymptotic regime, whereas the Assumption **[Bounded behavior of** $|\langle\, \Xi \rangle_N - \xi_0|$**]** does not. The two assumptions will be helpful to better control the performance of the test, specifically, PFA and PMD of the sequential test proposed later in the chapter. 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 both $\mathcal{H}_0$ 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 indifference zone between the two hypotheses in* (3.1). *Note also that, if the Assumption **[Bounded behavior of** $|\langle\, \Xi\rangle_N - \xi_0|$**]** holds true this implies that the Assumption **[(a-s) convergence of** $\langle\, \Xi\rangle_N$**]** will also hold true.* □

Next, we contrast the need for the two assumptions above in comparison to the RDT proposed in [24] and its FSS version, *Block*RDT proposed in [25].

Remark 3.3. *It must be noted that the Assumption **[Bounded behavior of** $|\langle\, \Xi\rangle_N - \xi_0|$**]** was not required in both RDT [24] and its FSS version, BlockRDT [25]. Motivated by the Neyman-Pearson framework [21, 27], 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, the Assumption **[Bounded behavior of** $|\langle\, \Xi\rangle_N - \xi_0|$**]** 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 the two assumptions **[(a-s) convergence of** $\langle\, \Xi\rangle_N$**]** and **[Bounded behavior of** $|\langle\, \Xi\rangle_N - \xi_0|$**]**.

Example 4. (i) Consider a random variable U with unknown distribution. We assume the two hypotheses as:

$$\text{Under } \mathcal{H}_0 : |U| \leqslant \tau^-$$

$$\text{Under } \mathcal{H}_1 : |U| \geqslant \tau^+$$

Suppose further that $\Xi_n = U + \Delta_n$, for $n \in \mathbb{N}$ where the Δ_ns are i.i.d with zero mean and unknown distribution. Now, with the application of strong law of large numbers [6], we know that $\langle\, \Xi\rangle_N$ will converge almost surely to U. This implies that Assumption **[(a-s) convergence of** $\langle\, \Xi\rangle_N$**]** will be satisfied.

(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,

$$\text{Under } \mathcal{H}_0 : \xi = \xi_0 = 0$$
$$\text{Under } \mathcal{H}_1 : \xi = \xi_1 \neq 0$$

with $|\xi_1| \geqslant \tau^+ > \tau^-$. If $\Xi_n = \frac{1}{n}\sum_{i=1}^{n} U_i$, Assumption **[(a-s) convergence of $\langle\, \Xi\rangle_N$]** is verified. Moreover, the above process is a non-stationary Markov process under both hypotheses and a variant of the random walk. Similarly, for many problems of practical interest as discussed earlier in Chapter 2, we can show that Assumption **[Bounded behavior of $|\langle\, \Xi\rangle_N - \xi_0|$]** can be verified with little prior knowledge of the underlying signals. This will become clearer in the Chapter 5 where we analyze the performance of the algorithms. □

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

3.2 Test statistic

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 :

$$\mathcal{T}_{N,\gamma}(x) = \begin{cases} 0 & \text{if} \quad |\langle\, x\rangle_N - \xi_0| \leqslant \lambda_\gamma(\tau\sqrt{N})/\sqrt{N} \\ 1 & \text{otherwise} \end{cases} . \tag{3.2}$$

Proposition 3.1 below describes the asymptotic behavior of such tests under Assumption **[(a-s) convergence of $\langle\, \Xi\rangle_N$]**. These tests play a crucial role in the design of *Seq*RDT for the problem stated in (2.1).

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

(i) we have

$$\textit{under } \mathcal{H}_0: \quad \mathbb{P}\left[\,\mathcal{T}_{N,\gamma}(Y) = 1\,\right] \leqslant \gamma, \tag{3.3}$$

$$\textit{under } \mathcal{H}_1: \quad \mathbb{P}\left[\,\mathcal{T}_{N,\gamma}(Y) = 0\,\right] \leqslant 1-\gamma. \tag{3.4}$$

(ii) *under Assumption **[(a-s) convergence of** $\langle\,\Xi\rangle_N$**]**, we have,*

$$\lim_{N\to\infty} \mathbb{P}\left[\,\mathcal{T}_{N,\gamma}(Y) = 1\,\right] = \begin{cases} 0 & \textit{under } \mathcal{H}_0 \\ 1 & \textit{under } \mathcal{H}_1 \end{cases}. \tag{3.5}$$

PROOF:

Proof of statement (i): From (3.2) and Lemma A.1,

$$\mathbb{P}\left[\,\mathcal{T}_{N,\gamma}(Y) = 1\,\right] \leqslant \mathbb{E}\left[Q_{\frac{1}{2}}\left(\sqrt{N}|\langle\,\Xi\rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\right)\right]. \tag{3.6}$$

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

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

According to (1.4), the upper-bound in the second inequality above equals γ and the inequality follows.

We prove (3.4) similarly. We begin by combining (3.2) and Lemma A.1 to get

$$\mathbb{P}\left[\,\mathcal{T}_{N,\gamma}(Y) = 0\,\right] \leqslant 1 - \mathbb{E}\left[Q_{\frac{1}{2}}\left(\sqrt{N}|\langle\,\Xi\rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\right)\right]. \tag{3.7}$$

It then suffices to use the hypothesis under $\mathcal{H}_1$ for $N \geqslant N_0$ in (3.7) above to get the final result.

Proof of statement (ii): Under $\mathcal{H}_0$ and Assumption **[(a-s) convergence of** $\langle\,\Xi\rangle_N$**]**, we have

$$\limsup_{N\to\infty} |\langle\,\Xi\rangle_N - \xi_0| \leqslant \tau^- < \tau \quad \text{(a-s)}.$$

It then follows from Lemma A.2 that:

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

We then derive from (3.6) and the Lebesgue dominated convergence theorem [6] that, under $\mathcal{H}_0$:

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

Similarly, under $\mathcal{H}_1$ and the Assumption **[(a-s) convergence of $\langle\,\Xi\rangle_N$]** we have

$$\liminf_{N\to\infty} |\langle\,\Xi\rangle_N - \xi_0| \geqslant \tau^+ > \tau \text{ (a-s)}.$$

Then from Lemma A.2 we have that

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

By injecting this equality into (3.7) and using the Lebesgue dominated convergence theorem [6] again, we obtain:

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

under $\mathcal{H}_1$.

This concludes the proof. ∎

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 the Assumption **[Bounded behavior of $|\langle\,\Xi\rangle_N - \xi_0|$]**, the designer can control both PFA and PMD of the algorithm as we demonstrate in the later part of this chapter. Moreover, intuition suggests that one of these thresholds should be small enough to reduce 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 the Assumption **[(a-s) convergence of $\langle\,\Xi\rangle_N$]** 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 (3.2), we can design a sequential test capable of reducing the decision-making time for the testing problem defined in (3.1), while guaranteeing certain performance levels. This sequential approach yields the algorithm *Seq*RDT described below.

3.3 The Non-Truncated Algorithm: *Seq*RDT

Section 3.2 above motivates a sequential approach involving two thresholds designed using (3.2). One of these thresholds must guarantee a PFA that is upper bounded from above, 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 (3.1) is specified by defining the stopping time:

$$T = \min\left\{N \in \mathbb{N} : \mathcal{D}_M(N) \neq \infty\right\}, \tag{3.8}$$

and the decision rule as:

$$\text{with: } \begin{cases} \mathcal{D}_M(1) = \mathcal{D}_M(2) = \ldots = \mathcal{D}_M(M) = \infty, \\ \text{for } N > M, \mathcal{D}_M(N) = \begin{cases} 0 & \text{if } \ |\langle\,Y\rangle_N - \xi_0| \leqslant \lambda_L(N), \\ \infty & \text{if } \ \lambda_L(N) < |\langle\,Y\rangle_N - \xi_0| \leqslant \lambda_H(N), \\ 1 & \text{if } \ |\langle\,Y\rangle_N - \xi_0| > \lambda_H(N), \end{cases} \end{cases} \tag{3.9}$$

with

$$\lambda_L(N) = \frac{\lambda_\gamma(\tau\sqrt{N})}{\sqrt{N}} \ \text{ and } \ \lambda_H(N) = \frac{\lambda_{\gamma'}(\tau\sqrt{N})}{\sqrt{N}},$$

$\tau \in (0, \infty)$ and $\gamma, \gamma' \in (0, 1)$ must be such that $\gamma' < \gamma$, which implies $\lambda_L(N, w_L) < \lambda_H(N, w_H)$. Here, $\mathcal{D}_M(N)$ represents the decision variable as defined in Chapter 2. Specifically, we have

$\mathcal{D}_M(N) = 0$ which is equivalent to saying that $\mathcal{H}_0$ is decided, $\mathcal{D}_M(N) = 1$ is equivalent to saying that $\mathcal{H}_1$ is decided and $\mathcal{D}_M(N) = \infty$ is equivalent to saying that no decision is made at the N^{th} 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 in Chapter 5.

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

$$\mathbb{P}_{\text{FA}}(\mathcal{D}_M) \stackrel{\text{def}}{=} \mathbb{P}\left[\,\mathcal{D}_M(T) = 1\,\right] \quad \text{under } \mathcal{H}_0. \tag{3.10}$$

In the same way, the PMD is:

$$\mathbb{P}_{\text{MD}}(\mathcal{D}_M) \stackrel{\text{def}}{=} \mathbb{P}\left[\,\mathcal{D}_M(T) = 0\,\right] \quad \text{under } \mathcal{H}_1. \tag{3.11}$$

Since the goal of the sequential algorithm is to guarantee $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ to be 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 3.2 (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 the lower and the upper thresholds, respectively as:

$$\lambda_L(N) = \frac{\lambda_{1-\beta}(\tau\sqrt{N})}{\sqrt{N}} \;\text{ and }\; \lambda_H(N) = \frac{\lambda_{\alpha}(\tau\sqrt{N})}{\sqrt{N}}. \tag{3.12}$$

Next, we analyze the properties of the proposed thresholds.

3.3.1 Properties of the Thresholds

Proposition 3.2 below validates that the thresholds proposed above in (3.12) are appropriate for *Seq*RDT under both asymptotic and non-asymptotic regimes.

Proposition 3.2. *We have:*

(i) *For all $N \in \mathbb{N}$, we have*

$$\lambda_L(N) < \lambda_H(N).$$

(ii) *The threshold $\lambda_H(N)$ 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:

The proof of statement (i) is given in Lemma A.6 of the Appendix. The proofs of statement (ii) and (iii) are provided in Lemma A.7 and Lemma A.8, respectively, given in the Appendix. The proof of statement (iv) is given in Lemma A.3 (ii) of the Appendix. ∎

Proposition 3.2 (i) and (ii) imply that, as $N \to \infty$, the test will reduce to a non-sequential test as both thresholds become equal to τ. In Figures 3.1 and 3.2, we plot the two thresholds $\lambda_H(N)$ and $\lambda_L(N)$ defined in (3.12) for different parameter values. We notice that the threshold behavior corroborates the result of Proposition 3.2 (ii), (iii) and (iv).

The question addressed now is then *"Can this choice of thresholds give some performance guarantees, i.e, can we control $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ of SeqRDT such that SeqRDT belongs to $\mathscr{C}(\alpha, \beta)$?"*.

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:

$$\mathbb{P}\big[\mathcal{D}_M(T) = \varepsilon\big] = \mathbb{P}\left(\big[\mathcal{D}_M(T) = \varepsilon\big] \cap \big[T \geq M+2\big]\right) + \mathbb{P}\big[\mathcal{D}_M(M+1) = \varepsilon\big]. \tag{3.13}$$

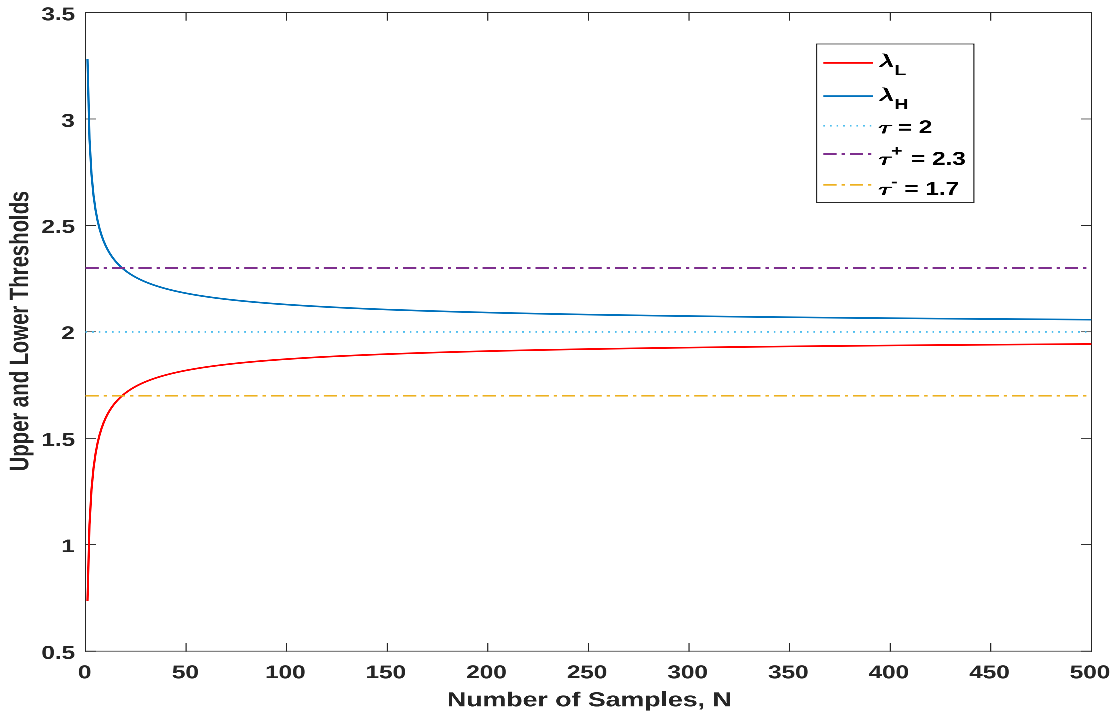

Fig. 3.1: $\lambda_L(N)$ and $\lambda_H(N)$ vs N for $\alpha = \beta = 0.1$ and $\tau = 2$.

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

$$\mathbb{P}\big[\mathcal{D}_M(T) = \varepsilon\big] \leqslant 1 - \mathbb{P}\big[\mathcal{D}_M(M+1) = 1-\varepsilon\big]. \tag{3.14}$$

Next, we discuss the asymptotic properties of the proposed algorithm.

3.3.2 Asymptotic Analysis of *Seq*RDT

Note that the algorithm *Seq*RDT (with stopping time defined in (3.8), the decision rule defined in (3.9) and the thresholds as defined in (3.12)) can potentially run forever. Therefore, it is important to guarantee that the stopping time of the algorithm stays finite with probability one. The next Theorem states this result and also studies the behavior of PFA and PMD with the buffer size, M.

Theorem 3.1 (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 **[(a-s) convergence of** $\langle \Xi \rangle_N$**]** holds true, then:*

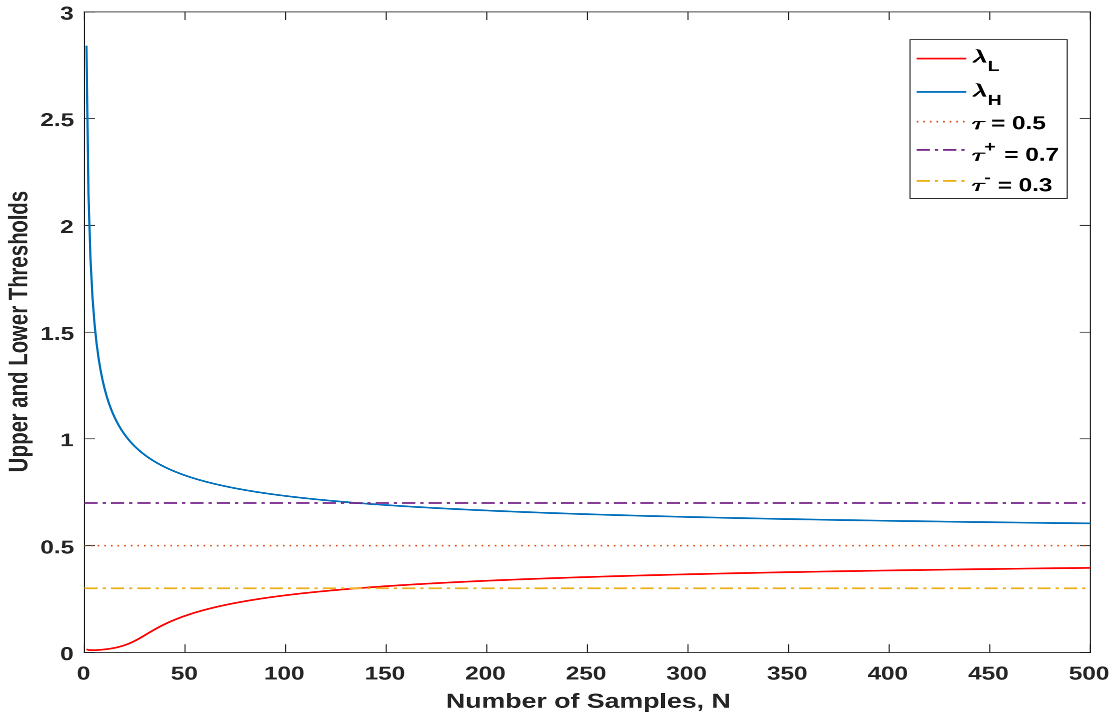

Fig. 3.2: $\lambda_L(N)$ and $\lambda_H(N)$ vs N for $\alpha = \beta = 0.01$ and $\tau = 0.5$.

(i) *We have*

$$\mathbb{P}\,[\,T < \infty\,] = 1 \textit{ under both } \mathcal{H}_0 \textit{ and } \mathcal{H}_1;$$

(ii) *We also have*

$$\lim_{M\to\infty} \mathbb{P}_{\text{FA}}(\mathcal{D}_M) = \lim_{M\to\infty} \mathbb{P}_{\text{MD}}(\mathcal{D}_M) = 0.$$

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 we have

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

This implies that we further have:

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

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

Proof of statement (ii): The PFA is

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

Using (3.14), we have

$$\mathbb{P}_{\text{FA}}(\mathcal{D}_M)\leqslant 1-\mathbb{P}\big[\mathcal{D}_M(M+1)=0\big].$$

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 (3.12) and Lemma A.1

$$\mathbb{P}_{\text{FA}}(\mathcal{D}_M)\leqslant\mathbb{E}\left[\,Q_{\frac{1}{2}}\left(\sqrt{M+1}|\langle\Xi\rangle_{M+1}-\xi_0|,\lambda_{1-\beta}(\tau\sqrt{M+1})\right)\,\right].\tag{3.15}$$

We then derive from Assumption **[(a-s) convergence of $\langle\,\Xi\rangle_N$]** and Lemma A.2 that, under $\mathcal{H}_0$,

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

The Lebesgue dominated convergence theorem [6] then implies that $\lim_{M\to\infty}\mathbb{P}_{\text{FA}}(\mathcal{D}_M)=0$.

Similarly, we derive from (3.14), (3.12) and Lemma A.1 that, regardless of **[(a-s) convergence**

of $\langle\,\Xi\rangle_N$]:

$$\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M) \leqslant 1 - \mathbb{E}\left[Q_{\frac{1}{2}}\left(\sqrt{M+1}|\langle\Xi\rangle_{M+1} - \xi_0|, \lambda_\alpha(\tau\sqrt{M+1})\right)\right]. \tag{3.16}$$

It then suffices to apply Assumption **[(a-s) convergence of $\langle\,\Xi\rangle_N$]**, Lemma A.2 and the Lebesgue dominated convergence theorem [6] to obtain the second equality in (ii).

Hence the proof. ■

The above theorem implies that under Assumption **[(a-s) convergence of $\langle\,\Xi\rangle_N$]**, i.e., if the empirical mean of the signal centered around ξ_0 converges away from τ, the sequential test (3.9) 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 the Assumption **[Bounded behavior of $|\langle\,\Xi\rangle_N - \xi_0|$]** 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 and the algorithm *Seq*RDT belongs to $\mathscr{C}(\alpha,\beta)$. We derive two bounds in the next section.

3.3.3 Non-Asymptotic Analysis of *Seq*RDT

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

Theorem 3.2 (Non-Asymptotics: $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$)**.** *$\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ are bounded as:*

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

PROOF: Under $\mathcal{H}_0$, we derive from (3.9), (3.10), (3.13), (3.12), and Lemma A.1 that:

$$\begin{aligned}\mathbb{P}_{\text{FA}}(\mathcal{D}_M) &\geqslant \mathbb{P}\left[\mathcal{D}_M(M+1)=1\right] \\ &\geqslant \mathbb{E}\left[Q_{\frac{1}{2}}\left(\sqrt{M+1}|\langle\Xi\rangle_{M+1}-\xi_0|, \lambda_\alpha(\tau\sqrt{M+1})\right)\right] \end{aligned} \tag{3.17}$$

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 (3.17), for the lower bound, and (3.15) along with (1.4) given in Chapter 1, for the upper bound.

Similarly, for the probability of missed detection, under $\mathcal{H}_1$, (3.9), (3.11), (3.12), (3.13) and Lemma A.1 yield

$$\begin{aligned}\mathbb{P}_{\text{MD}}(\mathcal{D}_M) &\geqslant \mathbb{P}\left[\mathcal{D}_M(M+1)=0\right] \\ &\geqslant 1-\mathbb{E}\left[Q_{\frac{1}{2}}\left(\sqrt{M+1}|\langle\Xi\rangle_{M+1}-\xi_0|, \lambda_{1-\beta}(\tau\sqrt{M+1})\right)\right]. \end{aligned} \tag{3.18}$$

We obtain the bounds on $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ from the inequalities satisfied by $|\langle\Xi\rangle_{M+1}-\xi_0|$ under $\mathcal{H}_1$ and (3.18), for the lower bound, and (3.16) along with (1.4) given in Chapter 1, 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 3.2 always stay below levels α and β, respectively. Theorem 3.2 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 **[Bounded behavior of $|\langle\,\Xi\rangle_N-\xi_0|$]** 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 and will help in guaranteeing that *Seq*RDT belongs to the class $\mathscr{C}(\alpha,\beta)$.

Theorem 3.3 (Non-Asymptotics: $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$)**.** *(i) Under Assumption **[Bounded be-***

$$\text{UB1}_{\text{FA}}(M) = Q_{\frac{1}{2}}\left(\tau^- \sqrt{M+1}, \lambda_{1-\beta}(\tau\sqrt{M+1})\right), \tag{3.20}$$

$$\text{UB2}_{\text{FA}}(M) = Q_{\frac{1}{2}}\left(\tau^- \sqrt{M+1}, \lambda_{\alpha}(\tau\sqrt{M+1})\right) + \sum_{N=M+2}^{\infty}\Bigg[\left(Q_{\frac{1}{2}}\left(\tau^-\sqrt{N}, \lambda_\alpha(\tau\sqrt{N})\right)\right) \bigwedge\left(\bigwedge_{K=M+1}^{N-1}\left(Q_{\frac{1}{2}}\left(\tau^-\sqrt{K}, \lambda_{1-\beta}(\tau\sqrt{K})\right) - Q_{\frac{1}{2}}\left(0, \lambda_\alpha(\tau\sqrt{K})\right)\right)\right)\Bigg], \tag{3.21}$$

$$\text{UB1}_{\text{MD}}(M) = 1 - Q_{\frac{1}{2}}\left(\sqrt{M+1}\,\tau^+, \lambda_{\alpha}(\tau\sqrt{M+1})\right), \tag{3.22}$$

$$\text{UB2}_{\text{MD}}(M) = 1 - Q_{\frac{1}{2}}\left(\tau^+ \sqrt{M+1}, \lambda_{1-\beta}(\tau\sqrt{M+1})\right) + \sum_{N=M+2}^{\infty}\Bigg[\left(1 - Q_{\frac{1}{2}}\left(\tau^+\sqrt{N}, \lambda_{1-\beta}(\tau\sqrt{N})\right)\right) \bigwedge\left(\bigwedge_{K=M+1}^{N-1}\left(Q_{\frac{1}{2}}\left(\tau_H\sqrt{K}, \lambda_{1-\beta}(\tau\sqrt{K})\right) - Q_{\frac{1}{2}}\left(\tau^+\sqrt{K}, \lambda_\alpha(\tau\sqrt{K})\right)\right)\right)\Bigg]. \tag{3.23}$$

***havior of** $|\langle\,\Xi\rangle_N - \xi_0|$**]**, $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ for SeqRDT are bounded as:*

$$\begin{cases} Q_{\frac{1}{2}}\left(0, \lambda_\alpha(\tau\sqrt{M+1})\right) \leqslant \mathbb{P}_{\text{FA}}(\mathcal{D}_M) \;\leqslant\; \mathit{UB1}_{\mathit{FA}}(M) \bigwedge \mathit{UB2}_{\mathit{FA}}(M), \\ 1 - Q_{\frac{1}{2}}\left(\tau_H\sqrt{M+1}, \lambda_{1-\beta}(\tau\sqrt{M+1})\right) \;\leqslant\; \mathbb{P}_{\text{MD}}(\mathcal{D}_M) \leqslant\; \mathit{UB1}_{\mathit{MD}}(M) \bigwedge \mathit{UB2}_{\mathit{MD}}(M). \end{cases} \tag{3.19}$$

where $a_1 \bigwedge a_2 = \min(a_1, a_2)$ for $a_1, a_2 \in \mathbb{R}$. UB1$_{FA}(M)$, UB2$_{FA}(M)$, UB1$_{MD}(M)$ and UB2$_{MD}(M)$ are finite and are given in (3.20) (3.21), (3.22) *and* (3.23)*, respectively.*
(ii) Moreover we have that UB1$_{FA}(M) \bigwedge$ UB2$_{FA}(M)$ and UB1$_{FA}(M) \bigwedge$ UB2$_{FA}(M)$ decrease with M.

PROOF:

Proof of statement (i): When Assumption **[Bounded behavior of** $|\langle\,\Xi\rangle_N - \xi_0|$**]** holds true, we have $0 \leqslant |\langle\Xi\rangle_{M+1} - \xi_0| \leqslant \tau^-$ under $\mathcal{H}_0$. Injecting these inequalities into (3.17) and (3.15) yields the bounds:

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

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

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

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

$$\begin{aligned}\mathbb{P}\left[\mathcal{D}_M(T)=1\right] &\leqslant \mathbb{P}\left[\mathcal{D}_M(M+1)=1\right] \\ &\qquad + \sum_{N=M+2}^{\infty} \mathbb{P}\left[\mathcal{D}_M(N)=1\right] \bigwedge \left(\bigwedge_{K=M+1}^{N-1} \mathbb{P}\left[\mathcal{D}_M(K)=\infty\right]\right). \qquad (3.24)\end{aligned}$$

For any $N \geqslant M+1$, Lemma A.1 and Assumption **[Bounded behavior of** $|\langle\, \Xi \rangle_N - \xi_0|$**]** imply that under $\mathcal{H}_0$, we have:

$$\mathbb{P}\left[\mathcal{D}_M(N)=1\right] \leqslant Q_{\frac{1}{2}}\left(\tau^-\sqrt{N}, \lambda_\alpha(\tau\sqrt{N})\right). \qquad (3.25)$$

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

$$\mathbb{P}\left[\mathcal{D}_M(K)=\infty\right] = \mathbb{P}\left[\langle Y \rangle_K - \xi_0| \geqslant \lambda_L(K)\right] - \mathbb{P}\left[\langle Y \rangle_K - \xi_0| > \lambda_H(K)\right]. \qquad (3.26)$$

Again from Lemma A.1, Assumption **[Bounded behavior of** $|\langle\, \Xi \rangle_N - \xi_0|$**]** under $\mathcal{H}_0$ we have:

$$\mathbb{P}\left[\mathcal{D}_M(K)=\infty\right] \leqslant Q_{\frac{1}{2}}\left(\tau^-\sqrt{K}, \lambda_{1-\beta}(\tau\sqrt{K})\right) - Q_{\frac{1}{2}}\left(0, \lambda_\alpha(\tau\sqrt{K})\right). \qquad (3.27)$$

The bound $\mathrm{UB2}_{\mathrm{FA}}(M)$ follows by injecting (3.25) and (3.27) into (3.24).

The bounds $\mathrm{UB1}_{\mathrm{MD}}(M)$ and $\mathrm{UB2}_{\mathrm{MD}}(M)$ on $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_{N_0})$ follow similarly from (3.16), (3.18), (3.26) and Lemma A.1, and using Assumption **[Bounded behavior of** $|\langle\, \Xi \rangle_N - \xi_0|$**]** under $\mathcal{H}_1$. The proof of the convergence of $\text{UB1}_{\text{MD}}(M)$ and $\text{UB2}_{\text{MD}}(M)$ are given in Appendix A.9.

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 A.5 that the mapping

$$N \mapsto Q_{\frac{1}{2}}\left(\rho\sqrt{N}, \lambda_\gamma(\tau\sqrt{N})\right)$$

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

Hence the proof. ■

3.3.4 Parameter Selection

Note that, Theorem 3.3 above makes it possible to choose the least buffer size M that guarantees 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 and ensure that *Seq*RDT belongs to the class $\mathscr{C}(\alpha, \beta)$. 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 we have

$$\text{UB1}_{\text{FA}}(M_1) \bigwedge \text{UB2}_{\text{FA}}(M_1) \leqslant \alpha.$$

Afterwards, choose M_2 such that we ahve

$$\text{UB1}_{\text{MD}}(M_2) \bigwedge \text{UB2}_{\text{MD}}(M_2) \leqslant \beta.$$

The buffer size can then be fixed to $M = \max(M_1, M_2)$. In Chapter 5, we will proceed in this man-

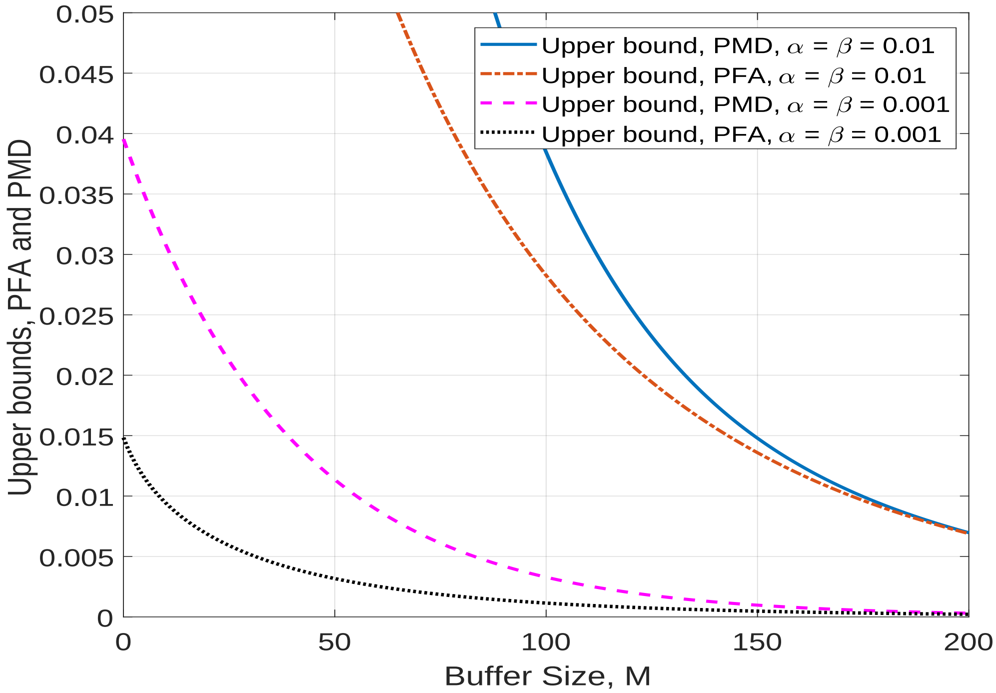

Fig. 3.3: Upper bound on $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ vs M (please see Theorem 3.3)

ner to choose the buffer size. It is, however, important to emphasize that the upper bounds given in Theorem 3.3 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 3.3 (i) and (3.19) 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. Importantly, we show in Chapter 5, that the tightness of these bounds depends on the underlying signal distributions and these bounds can in fact be tight for some signal distributions. Moreover, we show in Chapter 5 that the proposed algorithm, *Seq*RDT makes a decision faster on an average compared to the optimal FSS test discussed in the next chapter.

Now we know that with the test as defined in (3.9) with stopping time defined in (3.8) and with the thresholds as designed in (3.12), we can ensure that *Seq*RDT belongs to the class $\mathscr{C}(\alpha, \beta)$. This was ensured with the help of a buffer, M, as designed above. Next, we present an alternative design to *Seq*RDT where we eliminate the need for this buffer.

3.4 Alternate Design: Eliminating the Buffer

In this section, we eliminate the need for the buffer, M, by introducing an additional parameter in the design of the algorithm. Specifically, we update the thresholds given in (3.12) by introducing this additional parameter. Before introducing this parameter, let us have a look at the behavior of the upper bounds derived in Theorem 3.3 with increasing buffer size, M.

In Figure 3.3, we plot the upper upper bounds derived in Theorem 3.3 with the buffer size, M, for two pairs of α, β values, specifically, we have $\alpha = \beta = 0.01$ and $\alpha = \beta = 0.001$. The plot confirms the observation of Theorem 3.3 (ii), which implies that we can choose the buffer size, M, large enough such that *Seq*RDT belongs to $\mathscr{C}(\alpha, \beta)$. Importantly, we make another observation from Figure 3.3, which is, to achieve the same level of PFA and PMD performance the algorithm with smaller α and β values requires smaller buffer size, M. This follows from the Figure 3.3 above which shows that to achieve PFA $= 0.01$ and PMD $= 0.01$ the algorithm with thresholds designed using $\alpha = \beta = 0.001$ requires a smaller buffer size, M (approximately $M = 50$), compared to the algorithm with thresholds designed using $\alpha = \beta = 0.01$ (approximately $M = 170$).

Based on the above observation we introduce an additional scaling parameter w_{N_0} and design the thresholds given in (3.12) with α replaced by α/w_{N_0} and β replaced by β/w_{N_0} with $w_{N_0} \geqslant 1$. This choice of thresholds will reduce the buffer size, M, as can be seen in the Figure 3.3 above, which also suggests that we can choose the parameter w_{N_0} large enough such that the buffer size, M, becomes equal to N_0 (please see (3.1)).

3.4.1 Designing the Thresholds

Based on the discussion above we update the thresholds given in (3.12) with α replaced by α/w_{N_0} and β replaced by β/w_{N_0}, therefore we have the new thresholds as:

$$\lambda_L(N, w_{N_0}) = \frac{\lambda_{1-\frac{\beta}{w_{N_0}}}(\tau\sqrt{N})}{\sqrt{N}} \text{ and } \lambda_H(N, w_{N_0}) = \frac{\lambda_{\frac{\alpha}{w_{N_0}}}(\tau\sqrt{N})}{\sqrt{N}}. \tag{3.28}$$

Note that we only update the thresholds, the rest of the test including the stopping time (3.8) and the decision rule (3.9) stay the same. Next, we analyze the properties of the thresholds proposed above in (3.28).

Proposition 3.3. *For* $w_{N_0} \geqslant 1$*:*

(i) *We have*

$$\lambda_L(N, w_{N_0}) < \lambda_H(N, w_{N_0}),$$

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

(ii) *The threshold* $\lambda_H(N, w_{N_0})$ *is decreasing in* $N \in \mathbb{N}$ *and lower bounded by* τ.

(iii) *For* N *large enough, the threshold* $\lambda_L(N, w_{N_0})$ *is increasing in* N *and upper bounded by* τ.

(iv) *Both the thresholds approach* τ *as* N *increases:*

$$\lim_{N\to\infty} \lambda_H(N, w_{N_0}) = \lim_{N\to\infty} \lambda_L(N, w_{N_0}) = \tau.$$

PROOF: Since $\alpha, \beta \in (0, 1/2)$ and w_{N_0} is greater or equal to 1, we have

$$0 < \frac{\alpha}{w_{N_0}} < \frac{1}{2} < 1 - \frac{\beta}{w_{N_0}}.$$

Thus the proof of statement *(i)* is given in Lemma A.6 of the Appendix. Proof of Statements *(ii)* and *(iii)* are given in Lemmas A.7 and A.8, respectively, given in the Appendix. The proof of statement *(iv)* is given in Lemma A.3 of the Appendix. ■

Similar to the case of the thresholds $\lambda_H(N)$ and $\lambda_L(N)$ as defined in (3.12), Proposition 3.3 above ensures $\lambda_L(N, w_{N_0}) < \lambda_H(N, w_{N_0})$, which is made possible by the assumption that $w_{N_0} \geqslant 1$. This makes the thresholds defined in (3.28) a valid choice for designing a sequential algorithm. Moreover, both the thresholds tend to τ as N increases, which intuitively implies that the chance of making a decision should be higher for larger N.

Now, we analyze the behavior of the thresholds with respect to parameters w_H and w_L, when N is fixed.

Proposition 3.4. *For $w_{N_0} \geqslant 1$ we have*

(i) $\lambda_H(N, w_{N_0})$ *increases when* w_{N_0} *increases,*

(ii) $\lambda_L(N, w_{N_0})$ *decreases when* w_{N_0} *increases,*

(iii) *We have*

$$\lim_{w_{N_0}\to\infty} \lambda_H(N, w_{N_0}) = \infty \;\; \textit{and} \;\; \lim_{w_{N_0}\to\infty} \lambda_L(N, w_{N_0}) = 0.$$

PROOF:

The proof of statements *(i)* and *(ii)* are given in Lemma A.6 of the appendix. Statement *(iii)* follows from (1.4) in Chapter 1 and the fact that the Marcum function (1.2) (in Chapter 1) is a complementary cdf. ■

In this section, we analyzed the properties of the thresholds proposed in (3.28). Next, we derive the upper bounds on PFA and PMD of the algorithm and show that we can choose appropriate parameters such that the algorithm belongs to $\mathscr{C}(\alpha, \beta)$ for arbitrary pre-specified α and β.

3.4.2 Analysis of *Seq*RDT

In this section, we bound PFA and PMD of *Seq*RDT for the thresholds as defined in (3.28)

Theorem 3.4 (Non-Asymptotics: $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$)**.** *(i) Under Assumption* ***[Bounded behavior of** $|\langle\, \Xi \rangle_N - \xi_0|$**]** *and for the thresholds designed according to* (3.28)*,* $\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$ *and* $\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$ *for SeqRDT are bounded as:*

$$\begin{cases} Q_{\frac{1}{2}}\left(0, \lambda_{\frac{\alpha}{w_{N_0}}}(\tau\sqrt{M+1})\right) \leqslant \mathbb{P}_{\text{FA}}(\mathcal{D}_M) \\ \qquad\qquad\qquad\qquad \leqslant \textit{UB1}_{\textit{FA}}(M, w_{N_0}) \bigwedge \textit{UB2}_{\textit{FA}}(M, w_{N_0}), \\ 1 - Q_{\frac{1}{2}}\left(\tau_H\sqrt{M+1}, \lambda_{1-\frac{\beta}{w_{N_0}}}(\tau\sqrt{M+1})\right) \leqslant \mathbb{P}_{\text{MD}}(\mathcal{D}_M) \\ \qquad\qquad\qquad\qquad \leqslant \textit{UB1}_{\textit{MD}}(M, w_{N_0}) \bigwedge \textit{UB2}_{\textit{MD}}(M, w_{N_0}). \end{cases} \tag{3.29}$$

where $a_1 \bigwedge a_2 = \min(a_1, a_2)$ *for* $a_1, a_2 \in \mathbb{R}$*.* *UB1*$_{\textit{FA}}(M, w_{N_0})$*,* *UB2*$_{\textit{FA}}(M, w_{N_0})$*,* *UB1*$_{\textit{MD}}(M, w_{N_0})$

$$\text{UB1}_{\text{FA}}(M, w_{N_0}) = Q_{\frac{1}{2}}\left(\tau^-\sqrt{M+1}, \lambda_{1-\frac{\beta}{w_{N_0}}}(\tau\sqrt{M+1})\right), \tag{3.30}$$

$$\text{UB2}_{\text{FA}}(M, w_{N_0}) = Q_{\frac{1}{2}}\left(\tau^-\sqrt{M+1}, \lambda_{\frac{\alpha}{w_{N_0}}}(\tau\sqrt{M+1})\right) + \sum_{N=M+2}^{\infty}\Bigg[\left(Q_{\frac{1}{2}}\left(\tau^-\sqrt{N}, \lambda_{\frac{\alpha}{w_{N_0}}}(\tau\sqrt{N})\right)\right) \bigwedge\left(\bigwedge_{K=M+1}^{N-1}\left(Q_{\frac{1}{2}}\left(\tau^-\sqrt{K}, \lambda_{1-\frac{\beta}{w_{N_0}}}(\tau\sqrt{K})\right) - Q_{\frac{1}{2}}\left(0, \lambda_{\frac{\alpha}{w_{N_0}}}(\tau\sqrt{K})\right)\right)\right)\Bigg], \tag{3.31}$$

$$\text{UB1}_{\text{MD}}(M, w_{N_0}) = 1 - Q_{\frac{1}{2}}\left(\tau^+\sqrt{M+1}, \lambda_{\frac{\alpha}{w_{N_0}}}(\tau\sqrt{M+1})\right), \tag{3.32}$$

$$\text{UB2}_{\text{MD}}(M, w_{N_0}) = 1 - Q_{\frac{1}{2}}\left(\tau^+\sqrt{M+1}, \lambda_{1-\frac{\beta}{w_{N_0}}}(\tau\sqrt{M+1})\right) + \sum_{N=M+2}^{\infty}\Bigg[\left(1 - Q_{\frac{1}{2}}\left(\tau^+\sqrt{N}, \lambda_{1-\frac{\beta}{w_{N_0}}}(\tau\sqrt{N})\right)\right) \bigwedge\left(\bigwedge_{K=M+1}^{N-1}\left(Q_{\frac{1}{2}}\left(\tau_H\sqrt{K}, \lambda_{1-\frac{\beta}{w_{N_0}}}(\tau\sqrt{K})\right) - Q_{\frac{1}{2}}\left(\tau^+\sqrt{K}, \lambda_{\frac{\alpha}{w_{N_0}}}(\tau\sqrt{K})\right)\right)\right)\Bigg]. \tag{3.33}$$

and UB2$_{MD}(M, w_{N_0})$ *are finite and are given in* (3.30) (3.31), (3.32) *and* (3.33), *respectively.*

(ii) We have that UB1$_{FA}(M, w_{N_0}) \bigwedge$ *UB2*$_{FA}(M, w_{N_0})$ *and UB1*$_{FA}(M, w_{N_0}) \bigwedge$ *UB2*$_{FA}(M, w_{N_0})$ *decrease with* M.

(iii) Moreover, we have that UB1$_{FA}(M, w_{N_0}) \bigwedge$ *UB2*$_{FA}(M, w_{N_0})$ *and UB1*$_{FA}(M, w_{N_0}) \bigwedge$ *UB2*$_{FA}(M, w_{N_0})$ *decrease with* w_{N_0} *for* w_{N_0} *large enough.*

PROOF:

The proof of statement (i) and (ii) follows from the same series of arguments as the proof of Theorem 3.3

Proof of statement (iii): Proof of statement (iii) follows from a simple inspection of the individual terms of the upper bounds. Let us first consider the upper bound on PFA,

$$\text{UB1}_{\text{FA}}(M, w_{N_0}) \bigwedge \text{UB2}_{\text{FA}}(M, w_{N_0}).$$

Note that as w_{N_0} increases, the term $\text{UB1}_{\text{FA}}(M, w_{N_0})$ approaches 1 which follows from the definition of the threshold given in (1.4) from Chapter 1. Now consider the second term of the bound $\mathrm{UB2}_{\mathrm{FA}}(M, w_{N_0})$. Let us consider the first and the second terms of $\mathrm{UB2}_{\mathrm{FA}}(M, w_{N_0})$, we have the mapping

$$w_{N_0} \mapsto Q_{\frac{1}{2}}\left(\tau^- \sqrt{N}, \lambda_{\frac{\alpha}{w_{N_0}}}(\tau\sqrt{N})\right),$$

decreasing with w_{N_0}. Similarly, we have the mapping

$$w_{N_0} \mapsto Q_{\frac{1}{2}}\left(\tau^- \sqrt{N}, \lambda_{1-\frac{\beta}{w_{N_0}}}(\tau\sqrt{N})\right) - Q_{\frac{1}{2}}\left(0, \lambda_{\frac{\alpha}{w_{N_0}}}(\tau\sqrt{N})\right),$$

increasing with w_{N_0}. This implies that in the upper bound $\mathrm{UB2}_{\mathrm{FA}}(M, w_{N_0})$ as w_{N_0} increase the terms of type $Q_{\frac{1}{2}}\left(\tau^- \sqrt{N}, \lambda_{\frac{\alpha}{w_{N_0}}}(\tau\sqrt{N})\right)$ will dominate and will make the bound smaller and smaller as w_{N_0} increases further. This further implies that the term $\mathrm{UB2}_{\mathrm{FA}}(M, w_{N_0})$ will dominate in the bound $\mathrm{UB1}_{\mathrm{FA}}(M, w_{N_0}) \bigwedge \mathrm{UB2}_{\mathrm{FA}}(M, w_{N_0})$ and will capture the behavior of the upper bound at large values of w_{N_0}.

The result for $\mathrm{UB1}_{\mathrm{MD}}(M, w_{N_0}) \bigwedge \mathrm{UB2}_{\mathrm{MD}}(M, w_{N_0})$ follows from a similar argument.

Hence we have the proof. ■

The above Theorem 3.4 implies that the upper bounds on PFA and PMD of *Seq*RDT are a function of the buffer size, M, and the parameter w_{N_0}. Now, the goal is to choose these parameters such that *Seq*RDT belongs to the class $\mathscr{C}(\alpha, \beta)$ and at the same time makes a decision faster on average compared to the optimal FSS test.

3.4.3 Parameter Selection

In this section, we utilize the statements of Theorem 3.4 (ii) and (iii) to design an algorithm to choose the parameter, w_{N_0}, such that the buffer is eliminated, i.e., it becomes equal to N_0 (please see (3.1)).

Note that, from Theorem 3.4 (ii) and (iii) we see that if we increase the parameter, w_{N_0}, then in order to maintain

$$\mathrm{UB1}_{\mathrm{FA}}(M, w_{N_0}) \bigwedge \mathrm{UB2}_{\mathrm{FA}}(M, w_{N_0}) \approx \alpha,$$

i.e., to keep the upper bound tight, we will need to reduce the buffer size, M. This implies that we can potentially eliminate the buffer by choosing a sufficiently large value of the parameter, w_{N_0}. Note that this behavior was also captured in Figure 3.3. In Algorithm 1, we list the steps of *Seq*RDT. Note that we can choose to run Algorithm 1 with a buffer (Option I) or without a buffer (Option II). The thresholds in the algorithm are chosen according to the Option chosen.

Algorithm 1: *Seq*RDT

Initialize Given N_0, τ, τ^-, τ^+, α and β.

1. Parameter Selection
   Option I: Choose M as given in Section 3.3.4 or alternately choose
   Option II: w_{N_0} as given in Section 3.4.3

2. Compute Thresholds
   Option I: $\lambda_H(N)$ and $\lambda_L(N)$ using (3.12) or alternatively choose
   Option II: $\lambda_H(N, w_{N_0})$ and $\lambda_L(N, w_{N_0})$ (3.28)

While $\lambda_L(N) < |\langle Y \rangle_N - \xi_0| \leqslant \lambda_H(N)$
    $N = N + 1$
End
If $|\langle Y \rangle_N - \xi_0| \leqslant \lambda_L(N)$
    Accept $\mathcal{H}_0$
else if $|\langle Y \rangle_N - \xi_0| > \lambda_H(N)$
    Reject $\mathcal{H}_0$
End If

Finally, we discuss an important extension of the hypothesis testing frameworks presented in this dissertation. Note that for the hypothesis testing problem addressed in this chapter, with the hypotheses as given in (3.1), the inequalities are assumed to be satisfied in (a.s.) sense. However, in some cases the inequalities might not be satisfied in (a.s.) sense but might rather be satisfied in a weaker sense. The next section discusses such a scenario.

3.5 An Extension

Suppose that, instead of (3.1) where the inequalities are assumed to be satisfied in (a-s) sense, we have:

$$\text{Under } \mathcal{H}_0^* : \mathbb{P}\big[\text{for all } N \geqslant N_0, |\langle\,\Xi\rangle_N - \xi_0| \leqslant \tau\big] \geqslant 1-\varepsilon,$$
$$\text{Under } \mathcal{H}_1^* : \mathbb{P}\big[\text{for all } N \geqslant N_0, |\langle\,\Xi\rangle_N - \xi_1| \geqslant \tau\big] \geqslant 1-\varepsilon.$$

with a small positive constant $\varepsilon \leqslant \min(\alpha, \beta)$. Under the assumptions of Theorem 3.3, *Seq*RDT can still be used as follows to test $\mathcal{H}_0^*$ against $\mathcal{H}_1^*$ with guaranteed bounds on PFA and PMD, and thereby, can be guaranteed to belong to the class $\mathscr{C}(\alpha, \beta)$. Let us consider Option I in Algorithm 1 given above.

Indeed, given $\alpha, \beta \in (0,1)$, 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 (3.19). Under $\mathcal{H}_0^*$, the PFA, $\mathbb{P}_{\text{FA}}^*$, of *Seq*RDT satisfies:

$$\begin{aligned} \mathbb{P}_{\text{FA}}^* &= \mathbb{P}\,[\,\mathcal{D}_M(T) = 1\,] && (3.34)\\ &\leqslant \mathbb{P}(\Omega_0^c) + \mathbb{P}\big[\mathcal{D}_M(T) = 1\big|\Omega_0\big]\mathbb{P}(\Omega_0)\\ &\leqslant \varepsilon + \mathbb{P}\big[\mathcal{D}_M(T) = 1\big|\Omega_0\big]\mathbb{P}(\Omega_0), && (3.35) \end{aligned}$$

with $\Omega_0 = \big[\text{for all } N \geqslant N_0, |\langle\,\Xi\rangle_N - \xi_0| \leqslant \tau\big]$. Consider the probability space $\big(\Omega_0, \mathcal{F}_{\Omega_0}, \mathbb{P}\,(\bullet|\Omega_0)\,\big)$, 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 3.3, $\mathbb{P}\big[\mathcal{D}_M(T) = 1\big|\Omega_0\big] \leqslant \alpha - \varepsilon$ and thus $\mathbb{P}_{\text{FA}}^* \leqslant \alpha$. Similarly, we have $\mathbb{P}_{\text{MD}}^* \leqslant \beta$.

This extension can be useful in practice. For example, consider the case when the signal, $\Xi = (\Xi_n)_{n\in\mathbb{N}}$, comes from a Gaussian distribution, i.e., $\Xi_n \sim \mathcal{N}(0, \sigma^2)$. In this case, the probabilities given in (3.1) will not be satisfied in (a-s) sense for a finite N_0, but will be satisfied in a weaker sense as given in (3.34) above. Similarly, the above formulation is applicable for the cases when the signal, Ξ, is 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 $\xi_0 \neq \xi_1$.

However, in Chapter 5, we show that the proposed algorithms work with assumptions even weaker than as given in (3.34) above.

3.6 Summary

In this chapter, we proposed an algorithm to solve the sequential hypothesis testing problem proposed in Chapter 2 and stated in (3.1). We first introduced a few key assumptions required to design the sequential algorithm and help ensure that the proposed algorithms belong to the class $\mathscr{C}(\alpha, \beta)$. Then we motivated the algorithm design by analyzing the properties of the proposed test statistic. We then proposed the algorithm and analyzed its asymptotic properties. We introduced the notion of a buffer which is then used to control PFA and PMD of the proposed algorithm. We then derived the upper bounds on PFA and PMD of the algorithm and provided a method to choose an appropriate buffer size. We showed that we can choose a buffer size which ensures that the proposed algorithm, *Seq*RDT, belongs to the class $\mathscr{C}(\alpha, \beta)$. Importantly, we showed that without any prior knowledge of the underlying signal distributions, *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. In the later part of the chapter, we introduced an additional parameter in the algorithm (with updated thresholds) and presented another design of the algorithm which eliminated the need of the buffer. Finally, we presented the steps of the algorithm for both the designs with and without buffer size and showed a simple extension of the proposed frameworks when the proposed hypotheses are not true in (a.s.) sense but are rather true in a weaker sense.

Chapter 4
Truncated Sequential Random Distortion Testing

4.1 Introduction

As discussed earlier in Chapter 1, in his seminal works [39, 40], Wald proposed his celebrated sequential procedure, namely, SPRT for testing two simple hypotheses. SPRT is optimal in the sense that it makes a decision faster on average, compared to all the procedures including FSS tests achieving the same PFA and PMD. However, this optimality is lost in some cases when there is a mismatch between the assumed and true models for the underlying hypotheses to be tested [1, 33, 34], i.e, SPRT can have larger stopping times on average compared to the FSS tests that achieve the same error probabilities. To avoid these scenarios, a truncated version of SPRT was proposed in [33], where the truncation time was chosen based on the FSS test. However, the error probabilities achieved by truncated SPRT are usually higher than those achieved by non-truncated SPRT. In addition, larger truncation times are needed to guarantee error probabilities below predefined levels. The purpose of this chapter is to propose a truncated sequential algorithm for non-parametric hypothesis testing framework introduced in Chapter 2. In this chapter, we introduce a truncated version of the algorithm, *Seq*RDT, proposed in Chapter 3. Similar to truncated

SPRT, we also design the truncation window of the algorithm using the optimal FSS test.

In this chapter, we first state the problem and discuss the assumption required to control PFA and PMD of the proposed truncated algorithm. We then introduce the optimal FSS test and discuss some important preliminary results which play an important role in the design of the truncation time of the truncated sequential algorithm. We then extend the algorithm, *Seq*RDT, proposed in the preceding chapter and introduce a new truncated sequential algorithm, T-*Seq*RDT, to solve the binary hypothesis testing problem introduced in Chapter 2. Similar to *Seq*RDT, we first analyze the properties of the proposed thresholds. We derive bounds on PFA and PMD of T-*Seq*RDT and show that we can choose the parameters of the thresholds along with the truncation window to ensure that T-*Seq*RDT belongs to class $\mathscr{C}(\alpha, \beta)$. In contrast to *Seq*RDT, for T-*Seq*RDT we analyze the bounds on the average stopping time of T-*Seq*RDT and provide insights into the trade-off between the average stopping time and the error probabilities of the algorithm. Finally, we give an approach to choose the parameters of the thresholds and the truncation window size of the algorithm.

Next, we state the problem along with a crucial assumption which helps not only in controlling the PFA and PMD of the proposed algorithm but also helps in designing the truncation window of the truncated algorithm, T-*Seq*RDT.

4.2 Problem Statement

In this section, we introduce an important assumption required to control the PFA and PMD of the proposed algorithm. Also, as stated earlier in Chapter 3, the FSS tests only guarantee the PFA to be below a pre-specified level and cannot control the associated PMD similar to the Neyman-Pearson frameworks [21,27]. The assumption introduced in this section makes it possible to design FSS tests such that PMD can also be guaranteed to stay below pre-specified level. Importantly, the assumption introduced in this section also helps in designing the truncation window of the proposed algorithm. Before introducing the assumptions, we state the problem again:

$$\begin{cases} \underline{\text{Observation}} : Y = \Xi + X \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}} \\ \text{with} \begin{cases} \Xi = (\Xi_n)_{n\in\mathbb{N}} \in \mathcal{M}(\Omega, \mathbb{R})^{\mathbb{N}}, \\ X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathcal{N}(0,1), \\ \Xi \text{ and } X \text{ are independent.} \end{cases} \\ \exists N_0 \in \mathbb{N}, \begin{cases} \mathcal{H}_0 : \forall N \geqslant N_0,\ 0 \leqslant |\langle\, \Xi\rangle_N - \xi_0| \leqslant \tau \text{ (a-s)} \\ \mathcal{H}_1 : \forall N \geqslant N_0,\ \tau < |\langle\, \Xi\rangle_N - \xi_0| \leqslant \tau_H \text{ (a-s)} \end{cases} \end{cases} \tag{4.1}$$

where, $\tau \in [0, \infty)$ is the tolerance and $\tau < \tau_H < \infty$. Now, to exhibit elements of $\mathscr{C}(\alpha, \beta)$, we will make use of the following assumption.

Assumption 3 (Behavior of $|\langle\, \Xi\rangle_N - \xi_0|$ under $\mathcal{H}_1$)**.** *There exists $\tau^+ \in (\tau, \infty)$ such that:*

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

The Assumption **[Behavior of $|\langle\, \Xi\rangle_N - \xi_0|$ under $\mathcal{H}_1$]** states that under the alternate hypothesis, $\mathcal{H}_1$, the empirical mean of the signal centered around the model, ξ_0, is bounded away from τ. This assumption is similar in nature to that of the indifference zone assumed in [12,36]. Here the region (τ, τ^+) represents the indifference zone.

Remark 4.1. *SeqRDT proposed in Chapter 3 imposed stricter conditions on the signal compared to Assumption **[Behavior of** $|\langle\, \Xi\rangle_N - \xi_0|$ **under** $\mathcal{H}_1$**]**. Beyond Assumption **[Behavior of** $|\langle\, \Xi\rangle_N - \xi_0|$ **under** $\mathcal{H}_1$**]**, it was assumed in the SeqRDT framework that*

$$\textit{Under } \mathcal{H}_0, \textit{ for all } N \geqslant N_0, \textit{we have } |\langle\, \Xi\rangle_N - \xi_0| \leqslant \tau^- \textit{ (a-s)},$$

with $\tau^- \in [0, \tau)$. Therefore, SeqRDT required more parameters than T-SeqRDT. In addition, performance bounds were guaranteed by SeqRDT via the use of a buffer or via introducing a *parameter which was then used to eliminate the buffer. The buffer size as well as the parameter was selected using τ^- and τ^+ along with τ and τ_H defined in* (4.1). *In contrast, T-SeqRDT proposed in this chapter does not need to know τ^- or even τ_H. It requires the knowledge of τ and τ^+ only to guarantee performance, i.e., to ensure that the algorithm belongs to $\mathscr{C}(\alpha, \beta)$ with pre-defined levels α and β.* □

Next, we first define a FSS test *Block*RDT and show that with the use of Assumption **[Behavior of $|\langle\, \Xi \rangle_N - \xi_0|$ under $\mathcal{H}_1$]**, *Block*RDT can be designed so as to belong to $\mathscr{C}(\alpha, \beta)$. Then, by using *Block*RDT, we define the truncated sequential test, T-*Seq*RDT, that also belongs to $\mathscr{C}(\alpha, \beta)$ as well but at the same time makes a decision faster on average compared to *Block*RDT.

4.3 Optimal Fixed Sample Size (FSS) Test: *Block*RDT

In this section, we discuss the FSS testing framework to solve the binary hypothesis testing problem defined in (4.1) for a fixed number of samples $N \geqslant N_0$. Specifically, suppose that we have only N samples from our observation Y so that $Y = \Xi + X \in \mathcal{M}(\Omega, \mathbb{R})^{[\![1,N]\!]}$ in (4.1):

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

where, $\tau \in [0, \infty)$ is the tolerance and $\tau < \tau_H < \infty$. To solve this hypothesis testing problem, the authors in [25, 26] consider all the FSS tests $\mathcal{D}_{N_0}(N) = \mathcal{T}(Y)$, where $\mathcal{T}$ is any (measurable) mapping $\mathcal{T} : \mathbb{R}^N \to \{0, 1\}$. All such mappings $\mathcal{T}$ are hereafter called N-dimensional tests. In the

*Block*RDT framework [25, 26], we define the size of a given N-dimensional test $\mathcal{T}$ 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 $\mathcal{T}$ 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 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

$$\mathbb{P}\big[\mathcal{T}^*(Y) = 1 \,||\langle \Xi \rangle_N - \xi_0| > \tau\big] \geqslant \mathbb{P}\big[\mathcal{T}(Y) = 1 \,||\langle \Xi \rangle_N - \xi_0| > \tau\big]$$

for any N-dimensional test $\mathcal{T}$ and any $\Xi \in \mathcal{M}(\Omega,\mathbb{R})^{[\![1,N]\!]}$. We thus define the subclass of *Block*RDT-coherent tests [26], among which a "best" test exists. We say that an N-dimensional test $\mathcal{T}$ is *Block*RDT-coherent if:

[Invariance in mean] Given $y, y' \in \mathbb{R}^N$, if:

$$\langle y \rangle_N = \langle y' \rangle_N, \text{ then } \mathcal{T}(y) = \mathcal{T}(y').$$

[Constant conditional power] For all $\Xi \in \mathcal{M}(\Omega,\mathbb{R})^{[\![1,N]\!]}$ independent of X, there exists a Borel set $\mathcal{B}$ such that $|\langle \Xi \rangle_N - \xi_0| \in \mathcal{B}$ (a-s) and, for any $\rho \in \mathcal{B} \cap (0,\infty)$, $\mathbb{P}\big[\mathcal{T}(Y) = 1 \,|\, |\langle \Xi \rangle_N - \xi_0| = \rho\big]$ is independent of the distribution of $|\langle \Xi \rangle_N - \xi_0|$.

The rationale behind **[Invariance in mean]** is straightforward and implies that two different observation processes with the same empirical mean must yield the same decision for $\mathcal{T}$.

[Constant conditional power] means that $\mathcal{T}$ should not yield different results for different distributions of $|\langle \Xi \rangle_N - \xi_0|$, conditioned on $|\langle \Xi \rangle_N - \xi_0| = \rho$.

Let the class of all *Block*RDT-coherent tests with level γ be denoted by $\mathcal{K}_\gamma$. This class can be partially pre-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 $\mathcal{B}$ and

(ii) For all $\rho \in \mathcal{B} \cap (\tau, \infty)$,

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

According to [25, 26], the N-dimensional test defined for every $x \in \mathbb{R}^N$ by:

$$\mathcal{T}^*_{N,\gamma}(x) = \begin{cases} 0 & \text{if} \quad |\langle x \rangle_N - \xi_0| \leqslant \lambda_\gamma(\tau\sqrt{N})/\sqrt{N} \\ 1 & \text{otherwise.} \end{cases} \tag{4.2}$$

where $\lambda_\gamma(\tau\sqrt{N})/\sqrt{N}$ is defined using (1.4), is maximal in $\mathcal{K}_\gamma$: for any $\mathcal{T} \in \mathcal{K}_\gamma$, $\mathcal{T} \preceq \mathcal{T}_{N,\gamma}$. Let the PFA and PMD of $\mathcal{T}^*_{N,\gamma}$ for *Block*RDT be denoted by $\mathbb{P}^{\text{B-RDT}}_{\text{FA}}(N,\gamma)$ and $\mathbb{P}^{\text{B-RDT}}_{\text{MD}}(N,\gamma)$, respectively. We have the following proposition [25].

Proposition 4.1. *For any $\gamma \in (0,1)$ and $\tau \geqslant 0$, we have:*

$$\begin{cases} Q_{\frac{1}{2}}\left(0, \lambda_\gamma(\tau\sqrt{N})\right) & \leqslant \mathbb{P}^{\textit{B-RDT}}_{\textit{FA}}(N,\gamma) \leqslant \gamma \\ 1 - Q_{\frac{1}{2}}\left(\tau_H\sqrt{N}, \lambda_\gamma(\tau\sqrt{N})\right) & \leqslant \mathbb{P}^{\textit{B-RDT}}_{\textit{MD}}(N,\gamma) \leqslant 1-\gamma \end{cases}$$

According to the above proposition, although being optimal for *Block*RDT, $\mathcal{T}^*_{N,\gamma}$ controls $\mathbb{P}^{\text{B-RDT}}_{\text{FA}}(N,\gamma)$ but has no control over $\mathbb{P}^{\text{B-RDT}}_{\text{MD}}(N,\gamma)$.

This implies that, without further assumption and for any $\gamma \in (0,1)$, *Block*RDT cannot belong to the class $\mathscr{C}(\alpha,\beta)$ (with $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ replaced by $\mathbb{P}^{\text{B-RDT}}_{\text{FA}}(N,\gamma)$ and $\mathbb{P}^{\text{B-RDT}}_{\text{MD}}(N,\gamma)$ in (2.5), respectively) when $\alpha, \beta \in (0, 1/2)$. However, with Assumption **[Behavior of $|\langle \Xi \rangle_N - \xi_0|$ under $\mathcal{H}_1$]**, the next result implies that we can control $\mathbb{P}^{\text{B-RDT}}_{\text{FA}}(N,\gamma)$ such that *Block*RDT is in $\mathscr{C}(\alpha,\beta)$.

Proposition 4.2. *For any $\gamma \in (0,1)$, $\mathbb{P}^{\textit{B-RDT}}_{\textit{FA}}(N,\gamma)$ and $\mathbb{P}^{\textit{B-RDT}}_{\textit{MD}}(N,\gamma)$ are bounded under Assump-*

tion ***[Behavior of*** $|\langle\, \Xi\rangle_N - \xi_0|$ ***under*** $\mathcal{H}_1$***]*** *as:*

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

and the upper bound on $\mathbb{P}_{\text{\textit{MD}}}^{\text{\textit{B-RDT}}}(N,\gamma)$ *decreases to* 0 *with* N.

PROOF: The bound follows from Lemma A.1 and the application of Assumption **[Behavior of** $|\langle\, \Xi\rangle_N - \xi_0|$ **under** $\mathcal{H}_1$**]**. The upper-bound on $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}(N,\gamma)$ decreases with N as a consequence of Lemma A.5. ■

Proposition 4.2 implies that for $\gamma = \alpha$ and a sufficiently large N such that the bound on $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}(N,\gamma)$ is below β, *Block*RDT is in $\mathscr{C}(\alpha,\beta)$. Since this N might be very large in practice, we introduce a novel truncated sequential algorithm, T-*Seq*RDT, to control the number of samples and make a decision faster on average compared to the optimal FSS test, *Block*RDT.

4.4 The Truncated Algorithm: T-*Seq*RDT

In this section, we propose T-*Seq*RDT. In T-*Seq*RDT, if no decision has been reached until a specified time, the decision will be forced using *Block*RDT [25], since Proposition 4.2 guarantees that we can attain arbitrarily small PMD for a bounded PFA under Assumption **[Behavior of** $|\langle\, \Xi\rangle_N - \xi_0|$ **under** $\mathcal{H}_1$**]**.

Below we state the stopping time and the decision rule for T-*Seq*RDT. The stopping time is defined as:

$$T = \inf\{N \in \mathbb{N} : N \leqslant N_0 + W^* - 1, \mathcal{D}(N) \neq \infty\}. \tag{4.3}$$

and the decision variable $\mathcal{D}_{N_0}(N)$ for T-*Seq*RDT is defined as:

$$\begin{cases} \mathcal{D}_{N_0}(1) = \mathcal{D}_{N_0}(2) = \ldots = \mathcal{D}_{N_0}(N_0 - 1) = \infty, \\ \text{for } N_0 \leqslant N < N_0 + W^*, \\ \mathcal{D}_{N_0}(N) = \begin{cases} 0 \ \text{ if } |\langle Y \rangle_N - \xi_0| \leqslant \lambda_L(N) \\ 1 \ \text{ if } |\langle Y \rangle_N - \xi_0| > \lambda_H(N) \\ \infty \text{ if } \lambda_L(N) < |\langle Y \rangle_N - \xi_0| \leqslant \lambda_H(N) \end{cases} \\ \text{for } N = N_0 + W^*, \\ \mathcal{D}_{N_0}(N) = \begin{cases} 0 \text{ if } |\langle Y \rangle_N - \xi_0| \leqslant \lambda_{\text{B-RDT}}(N) \\ 1 \text{ if } |\langle Y \rangle_N - \xi_0| > \lambda_{\text{B-RDT}}(N) \end{cases} \end{cases} \tag{4.4}$$

with decisions taken according to (2.2) given in Chapter 2. At time instant $N = N_0 + W^*$, with $W^* \in \mathbb{N}$, the decision is made using *Block*RDT, if a decision has not been made until then. Recall that W^* is defined as the truncation window. The three thresholds $\lambda_L(N)$, $\lambda_H(N)$ and $\lambda_{\text{B-RDT}}(N)$ must be designed jointly so as to guarantee that T-*Seq*RDT is in $\mathscr{C}(\alpha, \beta)$. In any case, $\lambda_H(N)$ and $\lambda_L(N)$ must be such that $\lambda_L(N) < \lambda_H(N)$. Moreover, we want a decision faster compared to *Block*RDT, the optimal FSS counterpart of T-*Seq*RDT. The thresholds are chosen with respect to these constraints. We define the PFA and PMD of T-*Seq*RDT as:

$$\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0}) \stackrel{\text{def}}{=} \mathbb{P}\left[\mathcal{D}_{N_0}(T) = 1\right], \quad \text{under } \mathcal{H}_0, \tag{4.5}$$

is the PFA and

$$\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0}) \stackrel{\text{def}}{=} \mathbb{P}\left[\mathcal{D}_{N_0}(T) = 0\right], \quad \text{under } \mathcal{H}_1. \tag{4.6}$$

is the PMD.

Earlier in Chapter 3, we proposed a non-truncated sequential algorithm, *Seq*RDT, to solve the binary hypothesis testing problem (4.1). Moreover, *Seq*RDT was shown to belong to class $\mathscr{C}(\alpha, \beta)$ either with the help of a buffer or with the help of a parameter which helped in eliminating the buffer. The upper and lower thresholds respectively for *Seq*RDT when designed with a buffer were defined as:

$$\lambda_{\alpha}(\tau\sqrt{N})/\sqrt{N} \;\text{ and }\; \lambda_{1-\beta}(\tau\sqrt{N})/\sqrt{N}, \tag{4.7}$$

and for the design of *Seq*RDT without the buffer were defined as:

$$\lambda_{\frac{\alpha}{w_{N_0}}}(\tau\sqrt{N})/\sqrt{N} \;\text{ and }\; \lambda_{1-\frac{\beta}{w_{N_0}}}(\tau\sqrt{N})/\sqrt{N}, \tag{4.8}$$

Note that with both the designs we were able to control both PFA and PMD of *Seq*RDT. These thresholds were designed using τ^- and τ^+ along with τ and τ_H defined in (4.1), where the meaning of τ^- is recalled in Remark 4.1 above.

4.4.1 Designing the Thresholds and Their Properties

*Seq*RDT proposed in Chapter 3 was designed to belong to class $\mathscr{C}(\alpha, \beta)$ via the thresholds (4.7) or (4.8). T-*Seq*RDT by design eliminates the need for the buffer required in the design of *Seq*RDT, while being in $\mathscr{C}(\alpha, \beta)$. In view of the similarity between the T-*Seq*RDT statistic in (4.4) to that of *Block*RDT in (4.2), we define the thresholds similar in structure to those of *Block*RDT. The thresholds $\lambda_H(N)$, $\lambda_L(N)$ and $\lambda_{\text{B-RDT}}(N)$ for T-*Seq*RDT are designed as:

$$\begin{aligned}
\lambda_H(N) &= \lambda_H(N, w_H) = \lambda_{\alpha/w_H}(\tau\sqrt{N})/\sqrt{N} \\
\lambda_L(N) &= \lambda_L(N, w_L) = \lambda_{1-\beta/w_L}(\tau\sqrt{N})/\sqrt{N} \\
\lambda_{\text{B-RDT}}(N) &= \lambda_{\text{B-RDT}}(N, w_{BH}) = \lambda_{\alpha/w_{BH}}(\tau\sqrt{N})/\sqrt{N},
\end{aligned} \tag{4.9}$$

where the parameters w_H, w_L and w_{BH} give the algorithm designer control over these thresholds and are equal to or greater than 1. This constraint is necessary to ensure that T-*Seq*RDT is a valid sequential test, by guaranteeing that $\lambda_L(N, w_L) < \lambda_H(N, w_H)$, as shown in Proposition 4.3 below. In addition, the parameters w_H, w_L and w_{BH} must appropriately be chosen so as to guarantee that

T-*Seq*RDT belongs to $\mathscr{C}(\alpha,\beta)$. To this end, we study the properties of the thresholds (4.9) and establish that they satisfy suitable properties for T-*Seq*RDT.

Proposition 4.3. *For $w_L \geqslant 1$ and $1 \leqslant w_{BH} \leqslant w_H$ given:*

(i) *We have*

$$\lambda_L(N, w_L) < \lambda_{\textit{B-RDT}}(N, w_{BH}) \leqslant \lambda_H(N, w_H),$$

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

(ii) *The thresholds $\lambda_H(N, w_H)$ and $\lambda_{\textit{B-RDT}}(N, w_{BH})$ are decreasing in $N \in \mathbb{N}$ and lower bounded by τ.*

(iii) *For N large enough, the threshold $\lambda_L(N, w_L)$ is increasing in N and upper bounded by τ.*

(iv) *All the thresholds approach τ as N increases:*

$$\lim_{N\to\infty} \lambda_H(N, w_H) = \lim_{N\to\infty} \lambda_{\textit{B-RDT}}(N, w_{BH}) = \lim_{N\to\infty} \lambda_L(N, w_L) = \tau.$$

PROOF: Since $\alpha, \beta \in (0, 1/2)$, we have

$$0 < \frac{\alpha}{w_H} \leqslant \frac{\alpha}{w_{BH}} < \frac{1}{2} < 1 - \frac{\beta}{w_L}.$$

Thus the proof of *(i)* follows from Lemma A.6. Statements *(ii)* and *(iii)* follow from Lemmas A.7 and A.8, respectively. The proof of *(iv)* derives from Lemma A.3. ■

As discussed earlier, Proposition 4.3 ensures $\lambda_L(N, w_L) < \lambda_H(N, w_H)$, which is made possible by the assumption that w_H, w_L and $w_{BH} \geqslant 1$. Moreover, all the thresholds tend to τ as N increases, which intuitively implies that the chance of making a decision should be higher for larger N.

Now, we analyze the behavior of the thresholds with respect to parameters w_H, w_L and w_{BH}, when N is fixed.

Proposition 4.4. *We have*

(i) *$\lambda_H(N, w_H)$ increases when w_H increases,*

(ii) $\lambda_L(N, w_L)$ *decreases when* w_L *increases,*

(iii) $\lambda_{\textit{B-RDT}}(N, w_{BH})$ *increases when* w_{BH} *increases,*

(iv) *We have*

$$\lim_{w_H \to \infty} \lambda_H(N, w_H) = \infty \text{ and } \lim_{w_L \to \infty} \lambda_L(N, w_L) = 0.$$

PROOF: The proof of *(i)*, *(ii)* and *(iii)* follows from Lemma A.6. Statement *(iv)* follows from (1.4) and the fact that the Marcum function (1.2) is a complementary cdf. ■

According to Proposition 4.4, $\lambda_H(N, w_H)$ and $\lambda_L(N, w_L)$ grow further away as w_H and w_L increase. Therefore, thresholds designed with higher values of w_H and w_L should provide better $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ performance compared to thresholds tuned with lower w_H and w_L values, but at the expense of longer stopping times. For *Seq*RDT proposed in Chapter 3, the error probabilities were controlled via the buffer or an additional parameter, w_{N_0}, and no control over the stopping time was provided. For T-*Seq*RDT, the control over the error probabilities is achieved by choosing the parameters w_H, w_L and w_{BH} so as to move the thresholds away from or closer to each other. This gives the designer control over the average stopping time as well. This will be discussed in more detail later.

4.4.2 Designing the Truncation Window

The goal of T-*Seq*RDT is to make a decision faster on average compared to its FSS counterpart, *Block*RDT, while providing sufficient performance guarantees. Thus, it makes sense to base the choice of the truncation window W^* on *Block*RDT as follows. For the threshold $\lambda_{\text{B-RDT}}(N, w_{BH})$ given in (4.9), Proposition 4.2 implies that $\mathbb{P}_{\text{FA}}^{\text{B-RDT}}(N, \alpha/w_{BH})$ is always upper bounded by α/w_{BH} and hence by α as $w_{BH} \geq 1$. Moreover, the upper bound on $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}(N, \alpha/w_{BH})$ is a decreasing function of N. We thus propose to choose $W^* = W^*(w_{BH}, w_{BL})$ as:

$$\begin{aligned} W^* &= W^*(w_{BH}, w_{BL}) \\ &= \min\left\{ W \in \mathbb{N} : 1 - Q_{\frac{1}{2}}\left(\tau^+\sqrt{N_0 + W}, \lambda_{\frac{\alpha}{w_{BH}}}(\tau\sqrt{N_0 + W})\right) \leqslant \frac{\beta}{w_{BL}} \right\} \end{aligned} \quad (4.10)$$

with $w_{BL} \geq 1$.

Remark 4.2. *Note from Proposition 4.2 that BlockRDT with $\lambda_{\text{B-RDT}}(N, w_{BH})$ given in* (4.9) *for the number of samples $N = N_0 + W^*$ is in $\mathscr{C}(\alpha/w_{BH}, \beta/w_{BL})$.*

The parameters w_H and w_L defined earlier are used to control the upper and the lower thresholds, respectively, via (4.9). On the other hand, the parameters w_{BH} and w_{BL} control the truncation window, $W^*(w_{BH}, w_{BL})$, defined in (4.10) and the assumption $w_{BL} \geq 1$ is required to make sure that the PMD of T-*Seq*RDT stays below β (see Theorem 4.2 below). All the thresholds along with the truncation window, which are thus controlled by w_H, w_L, w_{BH} and w_{BL}, govern the performance of T-*Seq*RDT. Therefore, we next analyse the behavior of $W^*(w_{BH}, w_{BL})$ with w_{BH} and w_{BL} so that w_{BL}, w_H, w_L and w_{BH} can be fixed to guarantee that T-*Seq*RDT belongs to $\mathscr{C}(\alpha, \beta)$.

Proposition 4.5. *We have*

(i) *For fixed w_{BL}, $W^*(\bullet, w_{BL})$ does not decrease;*

(ii) *For fixed w_{BH}, $W^*(w_{BH}, \bullet)$ does not decrease;*

PROOF: For any $w_{BH} \geqslant 1$ and any $W \in \mathbb{N}$, set:

$$\text{UB}_{\text{MD}}^{\text{B-RDT}}(w_{BH}, W) = 1 - Q_{\frac{1}{2}}\left(\tau^+\sqrt{N_0 + W}, \lambda_{\frac{\alpha}{w_{BH}}}(\tau\sqrt{N_0 + W})\right)$$

For any $w_{BL} \geqslant 1$,

$$W^*(w_{BH}, w_{BL}) = \min \mathcal{A}(w_{BH}, w_{BL}) \tag{4.11}$$

with:

$$\mathcal{A}(w_{BH}, w_{BL}) = \left\{ W \in \mathbb{N} : \text{UB}_{\text{MD}}^{\text{B-RDT}}(w_{BH}, W) \leqslant \frac{\beta}{w_{BL}} \right\} \tag{4.12}$$

Proof of (i): Consider $w_{BH} \leqslant w'_{BH}$. According to Lemmas 1.1 and A.6, we have:

$$\text{UB}_{\text{MD}}^{\text{B-RDT}}(w_{BH}, W) \leqslant \text{UB}_{\text{MD}}^{\text{B-RDT}}(w'_{BH}, W) \tag{4.13}$$

Therefore, from (4.11) and (4.12), we have $\mathcal{A}(w'_{BH}, w_{BL}) \subseteq \mathcal{A}(w_{BH}, w_{BL})$ and thus

$$W^*(w'_{BH}, w_{BL}) \geqslant W^*(w_{BH}, w_{BL}).$$

Proof of (ii): Fix w_{BH}. If $w_{BL} \leqslant w'_{BL}$, then $\frac{\beta}{w'_{BL}} \leqslant \frac{\beta}{w_{BL}}$. This implies that

$$\mathcal{A}(w_{BH}, w'_{BL}) \subseteq \mathcal{A}(w_{BH}, w_{BL}).$$

Hence the result. ■

Proposition 4.5 tells us that the smaller the required PFA and PMD for truncation by *Block*RDT, the larger the truncation window for T-*Seq*RDT, which is natural. This will lead to the trade-off pinpointed in the next section between this truncation window and the error probabilities of T-*Seq*RDT. In addition, the choice of the truncation window using *Block*RDT will allow for easier comparison between T-*Seq*RDT and *Block*RDT.

Remark 4.3. *Assumption **[Behavior of** $|\langle\, \Xi \rangle_N - \xi_0|$ **under** $\mathcal{H}_1$**]** is instrumental in choosing an appropriate truncation window W^* for T-SeqRDT (see Proposition 4.2 and* (4.10)*). But, if W^* is known* a priori*, i.e., it is available via some preliminary training procedure or prior experience, Assumption **[Behavior of** $|\langle\, \Xi \rangle_N - \xi_0|$ **under** $\mathcal{H}_1$**]** is not needed, while the algorithm will still achieve the same performance.* □

Our next goal is to choose the appropriate thresholds (4.9) and window size (4.10), such that T-*Seq*RDT is in $\mathscr{C}(\alpha, \beta)$. We proceed by noticing that (4.9) and Proposition 4.5 show that this question is equivalent to choosing appropriate values of w_H, w_L, w_{BH} and w_{BL}.

4.5 Analysis of T-*Seq*RDT

In this section, we calculate bounds on the PFA and PMD of T-*Seq*RDT. These bounds are used to derive values for w_H, w_L, w_{BH} and w_{BL} that guarantee the required performance. Then, we study the average stopping time. Finally, we discuss the relationship between the error probabilities and the average stopping time.

4.5.1 False Alarm and Missed Detection Probabilities

Since closed form expressions for $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ cannot be derived, we instead calculate upper and lower bounds on these error probabilities, for the thresholds (4.9). These bounds provide useful insights into the behavior of T-*Seq*RDT. We begin with lower bounds.

Theorem 4.1 (Lower-bounds on $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$)**.**

$$\begin{cases} \mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0}) \geqslant Q_{\frac{1}{2}}\left(0, \lambda_{\alpha/w_H}(\tau\sqrt{N_0})\right), \\ \mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0}) \geqslant 1 - Q_{\frac{1}{2}}\left(\tau_H\sqrt{N_0}, \lambda_{1-\beta/w_L}(\tau\sqrt{N_0})\right). \end{cases}$$

PROOF: Since $\left[\mathcal{D}_{N_0}(T) = 1\right] \supseteq \left[\mathcal{D}_{N_0}(N_0) = 1\right]$, (2.6) implies that, under $\mathcal{H}_0$:

$$\begin{aligned} \mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0}) &\geqslant \mathbb{P}\left[\mathcal{D}_{N_0}(N_0) = 1\right] \\ &\stackrel{(a)}{=} \mathbb{E}\left[Q_{\frac{1}{2}}\left(\sqrt{N_0}|\langle\Xi\rangle_{N_0} - \xi_0|, \sqrt{N_0}\lambda_H(N_0, w_H)\right)\right] \\ &\stackrel{(b)}{\geqslant} Q_{\frac{1}{2}}\left(0, \lambda_{\alpha/w_H}(\tau\sqrt{N_0})\right) \end{aligned}$$

where (a) follows from Lemma A.1, (b) from (4.9), Lemma 1.1 and the fact that under $\mathcal{H}_0$, $0 \leqslant |\langle\Xi\rangle_N - \xi_0| \leqslant \tau$ (a-s). Similarly, consider the event $\left[\mathcal{D}_{N_0}(T) = 0\right]$ and follow the same procedure as above to get the lower bound for $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$. ■

Although the lower bounds do not play any role in designing the thresholds, note that they decrease with N_0 and approach 0 as $N_0 \to \infty$, which follows from Lemma A.5 and A.4.

$$\text{UB}_{\text{FA}} = \frac{\alpha}{w_H} + \left[\sum_{N=N_0+1}^{N_0+W^*-1} \frac{\alpha}{w_H} \bigwedge \left(\bigwedge_{K=N_0}^{N-1} \left(\left(1 - \frac{\beta}{w_L}\right) - Q_{\frac{1}{2}}\left(0, \lambda_{\frac{\alpha}{w_H}}(\tau\sqrt{K})\right) \right) \right) \right]$$
$$+ \frac{\alpha}{w_{BH}} \bigwedge \left(\bigwedge_{K=N_0}^{N_0+W^*-1} \left(\left(1 - \frac{\beta}{w_L}\right) - Q_{\frac{1}{2}}\left(0, \lambda_{\frac{\alpha}{w_H}}(\tau\sqrt{K})\right) \right) \right), \quad (4.14)$$

$$\text{UB}_{\text{MD}} = \frac{\beta}{w_L} + \left[\sum_{N=N_0+1}^{N_0+W^*-1} \frac{\beta}{w_L} \bigwedge \left(\bigwedge_{K=N_0}^{N-1} \left(Q_{\frac{1}{2}}\left(\tau_H\sqrt{K}, \lambda_{1-\frac{\beta}{w_L}}(\tau\sqrt{K})\right) - \frac{\alpha}{w_H} \right) \right) \right]$$
$$+ \frac{\beta}{w_{BL}} \bigwedge \left(\bigwedge_{K=N_0}^{N_0+W^*-1} \left(Q_{\frac{1}{2}}\left(\tau_H\sqrt{K}, \lambda_{1-\frac{\beta}{w_L}}(\tau\sqrt{K})\right) - \frac{\alpha}{w_H} \right) \right). \quad (4.15)$$

Theorem 4.2 (Upper-bounds on $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$)**.**

$$\begin{cases} \mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0}) \leqslant \mathit{UB}_{FA} \leqslant \left(\frac{W^*}{w_H} + \frac{1}{w_{BH}}\right)\alpha, \\ \mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0}) \leqslant \mathit{UB}_{MD} \leqslant \left(\frac{W^*}{w_L} + \frac{1}{w_{BL}}\right)\beta, \end{cases}$$

where UB_{FA} and UB_{MD} are given in (4.14) *and* (4.15)*, respectively, and $W^* = W^*(w_{BH}, w_{BL})$.*

PROOF: We have

$$\left[\mathcal{D}_{N_0}(T) = 1\right] = \left[\mathcal{D}_{N_0}(N_0) = 1\right]$$
$$\bigcup_{N=N_0+1}^{N_0+W^*} \left(\left[\mathcal{D}_{N_0}(N) = 1\right] \cap \left[\mathcal{D}_{N_0}(K) = \infty, \forall K \text{ s.t. } N_0 \leqslant K \leqslant N-1\right] \right).$$

Since these events are disjoint, we have

$$\begin{aligned}\mathbb{P}\big[\mathcal{D}_{N_0}(T)=1\big] &= \mathbb{P}\big[\mathcal{D}_{N_0}(N_0)=1\big] + \sum_{N=N_0+1}^{N_0+W^*} \mathbb{P}\Big(\big[\mathcal{D}_{N_0}(N)=1\big] \\ &\qquad \cap \big[\mathcal{D}_{N_0}(K)=\infty, \forall K \text{ s.t. } N_0 \leqslant K \leqslant N-1\big]\Big) \\ &\overset{(a)}{\leqslant} \mathbb{P}\big[\mathcal{D}_{N_0}(N_0)=1\big] + \sum_{N=N_0+1}^{N_0+W^*} \mathbb{P}\big[\mathcal{D}_{N_0}(N)=1\big] \\ &\qquad \bigwedge \Big(\bigwedge_{K=N_0}^{N-1} \mathbb{P}\big[\mathcal{D}_{N_0}(K)=\infty\big]\Big), \end{aligned} \tag{4.16}$$

where (a) follows from the Frechet inequality. We bound each individual probability on the right hand side (rhs) of (4.16) under $\mathcal{H}_0$. First, for all $N_0 \leqslant N \leqslant N_0 + W^* - 1$, we have:

$$\begin{aligned}\mathbb{P}\big[\mathcal{D}_{N_0}(N)=1\big] &\overset{(a)}{=} \mathbb{E}\left[Q_{\frac{1}{2}}\left(\sqrt{N}|\langle\Xi\rangle_N - \xi_0|, \sqrt{N}\lambda_H(N, w_H)\right)\right] \\ &\overset{(b)}{\leqslant} Q_{\frac{1}{2}}\left(\tau\sqrt{N}, \lambda_{\alpha/w_H}(\tau\sqrt{N})\right) \qquad (4.17) \\ &\overset{(c)}{=} \alpha/w_H, \qquad (4.18)\end{aligned}$$

where (a) follows from Lemma A.1; (b) results from (4.9), the fact that under $\mathcal{H}_0 : 0 \leqslant |\langle\Xi\rangle_N - \xi_0| \leqslant \tau$ and Lemma 1.1; (c) comes from (1.4).

Second, for $N = N_0 + W^*$, we have under $\mathcal{H}_0$:

$$\begin{aligned}\mathbb{P}\big[\mathcal{D}_{N_0}(N)=1\big] &= \mathbb{P}_{\text{FA}}^{\text{B-RDT}} \\ &\overset{(a)}{\leqslant} \alpha/w_{BH},\end{aligned} \tag{4.19}$$

where (a) follows from Proposition 4.2 and (4.9).

Now, for all $N_0 \leqslant K \leqslant N_0 + W^* - 1$, we have:

$$\begin{aligned}
\mathbb{P}\big[\mathcal{D}_{N_0}(K) = \infty\big] &= \mathbb{P}\big[|\langle Y\rangle_K - \xi_0| \geqslant \lambda_L(K, w_L)\big] - \mathbb{P}\big[|\langle Y\rangle_K - \xi_0| > \lambda_H(K, w_H)\big] \\
&\overset{(a)}{=} \mathbb{E}\left[Q_{\frac{1}{2}}\left(\sqrt{K}|\langle \Xi\rangle_K - \xi_0|, \sqrt{K}\lambda_L(K, w_L)\right)\right] \\
&\qquad\qquad - \mathbb{E}\left[Q_{\frac{1}{2}}\left(\sqrt{K}|\langle \Xi\rangle_K - \xi_0|, \sqrt{K}\lambda_H(K, w_H)\right)\right] \\
&\overset{(b)}{\leqslant} Q_{\frac{1}{2}}\left(\tau\sqrt{K}, \lambda_{1-\beta/w_L}(\tau\sqrt{K})\right) - Q_{\frac{1}{2}}\left(0, \lambda_{\alpha/w_H}(\tau\sqrt{K})\right) \\
&\overset{(c)}{=} 1 - \beta/w_L - Q_{\frac{1}{2}}\left(0, \lambda_{\alpha/w_H}(\tau\sqrt{K})\right),
\end{aligned} \tag{4.20}$$

where: (a) follows from Lemma A.1, (b) from the monotonicity of the Marcum function, (4.9) and the fact that under $\mathcal{H}_0 : 0 \leqslant |\langle \Xi\rangle_N - \xi_0| \leqslant \tau$, and (c) from (1.4).

The upper bounds on $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ follow by substituting (4.18), (4.19) and (4.20) into (4.16) and using that $a_1 \wedge a_2 \leqslant a_1$. The upper bounds for $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ result from a similar procedure and the definition of W^* via (4.10). ■

This theorem justifies the definition of the thresholds in (4.9). It is clear that $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ of T-*Seq*RDT can be controlled such that T-*Seq*RDT is in $\mathscr{C}(\alpha, \beta)$ by choosing appropriate parameters w_H, w_L, w_{BH} and w_{BL}, which are independent of the signal model. Moreover, to do so, all these parameters have to be greater than or equal to one. Hereafter, we work with the looser upper bounds stated in Theorem 4.2. They are simpler to analyze as they depend on fewer parameters than UB_{FA} and UB_{MD} and give useful insights into the behavior of T-*Seq*RDT.

We use the threshold $\lambda_{\text{B-RDT}}(N_0 + W^*, w_{BH})$ with $W^* = W^*(w_{BH}, w_{BL})$ to stop T-*Seq*RDT if a decision has not been taken until $N_0 + W^*$. As pointed out in Remark 4.2, the PFA (resp. PMD) of the corresponding *Block*RDT is upper-bounded by α/w_{BH} (resp. β/w_{BL}). Therefore, from Theorem 4.2, we see that T-*Seq*RDT may lose some detection performance compared to *Block*RDT. However, it follows from this same theorem and Subsection 4.4.2 that the upper-bounds on the false alarm and missed detection probabilities are of the same order for T-*Seq*RDT and *Block*RDT. For example, if $w_{BH} = w_{BL} = 1$ and $w_H = w_L = W^*$, T-*Seq*RDT is in $\mathscr{C}(2\alpha, 2\beta)$ whereas *Block*RDT is in $\mathscr{C}(\alpha, \beta)$. We can thus increase w_H, w_L, w_{BH} and w_{BL} such that T-

*Seq*RDT is in $\mathscr{C}(\alpha,\beta)$. Though this comes at the cost of increasing the average stopping-time, this is the same behavior as observed for SPRT and discussed in [33]. We show in the next section that this average stopping time remains always less than $N_0 + W^*$.

4.5.2 Stopping time of T-*Seq*RDT

Similar to $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$, a closed form for the average stopping time of T-*Seq*RDT is not derivable. We, however, get an insight into the stopping behavior of T-*Seq*RDT by bounding its average stopping time.

Theorem 4.3 (Bounds on the average stopping time)**.** *With* $W^* = W^*(w_{BH}, w_{BL})$*:*

(i) *We have*

$$\begin{cases} \textit{Under } \mathcal{H}_0 : \mathbb{E}[T] \leqslant \textit{UB}_{T_{\mathcal{H}_0}} \\ \textit{Under } \mathcal{H}_1 : \mathbb{E}[T] \leqslant \textit{UB}_{T_{\mathcal{H}_1}}, \end{cases}$$

where:

$$\textit{UB}_{T_{\mathcal{H}_0}} = N_0 + W^* - \beta W^*/w_L - \sum_{N=N_0}^{N_0+W^*-1} Q_{\frac{1}{2}}\left(0, \lambda_{\alpha/w_H}(\tau\sqrt{N})\right),$$

$$\textit{UB}_{T_{\mathcal{H}_1}} = N_0 + W^* - \alpha W^*/w_H - \sum_{N=N_0}^{N_0+W^*-1} \left[1 - Q_{\frac{1}{2}}\left(\tau_H\sqrt{N}, \lambda_{1-\beta/w_L}(\tau\sqrt{N})\right)\right].$$

(ii) $\mathbb{E}[T] < N_0 + W^*$.

PROOF:

Proof of statement (i): Since the random variable T is discrete and valued in $\{N_0, N_0+1, \cdots, N_0+W^*\}$ and

$$\mathbb{E}[T] = \sum_{N=0}^{\infty} \mathbb{P}[T > N] = N_0 + \sum_{N=N_0}^{N_0+W^*-1} \mathbb{P}[T > N].$$

By definition of T (4.3), $\big[T > N\big] \subset \big[\mathcal{D}_{N_0}(N) = \infty\big]$ for any $N \in \{N_0, N_0+1, \cdots, N_0+W^*\}$. Hence, the following inequality:

$$\mathbb{E}[T] \leqslant N_0 + \sum_{N=N_0}^{N_0+W^*-1} \mathbb{P}\big[\mathcal{D}_{N_0}(N) = \infty\big]. \tag{4.21}$$

According to Lemma A.1, we can write:

$$\begin{aligned}\mathbb{P}\big[\mathcal{D}_{N_0}(N) = \infty\big] = \mathbb{E}\left[Q_{\frac{1}{2}}\left(\sqrt{N}|\langle\,\Xi\rangle_N - \xi_0|, \sqrt{N}\lambda_L(N, w_L)\right)\right] \\ - \mathbb{E}\left[Q_{\frac{1}{2}}\left(\sqrt{N}|\langle\,\Xi\rangle_N - \xi_0|, \sqrt{N}\lambda_H(N, w_H)\right)\right].\end{aligned} \tag{4.22}$$

Under $\mathcal{H}_0$, $0 \leqslant |\langle\,\Xi\rangle_N - \xi_0| \leqslant \tau$ (a-s) for all $N \geqslant N_0$ and thus:

$$\begin{aligned}\mathbb{P}\big[\mathcal{D}_{N_0}(N) = \infty\big] &\overset{(a)}{\leqslant} Q_{\frac{1}{2}}\left(\sqrt{N}\tau, \lambda_{1-\beta/w_L}(\tau\sqrt{N})\right) - Q_{\frac{1}{2}}\left(0, \lambda_{\alpha/w_H}(\tau\sqrt{N})\right) \\ &\overset{(b)}{=} 1 - \beta/w_L - Q_{\frac{1}{2}}\left(0, \lambda_{\alpha/w_H}(\tau\sqrt{N})\right),\end{aligned}$$

where (a) results from the monotonicity of $Q_{\frac{1}{2}}$ and (b) from (1.4). The bound on $\mathbb{E}[T]$ under $\mathcal{H}_0$ follows by substituting the inequality above into (4.21). Following a similar procedure to bound (4.22) under $\mathcal{H}_1$, we can obtain the bound under $\mathcal{H}_1$.

Proof of (ii): The result follows from the bound

$$\mathbb{P}[\mathcal{D}_{N_0}(N) = \infty] < 1 \;\text{ for all }\; N \in \{N_0, \cdots, N_0 + W - 1\}.$$

Hence the proof. ■

Theorem 4.3 states that the average stopping time of T-*Seq*RDT is strictly less than the *Block*RDT block size N_0+W^*. Therefore, on the one hand, Theorem 4.2 suggests that T-*Seq*RDT will lose detection performance compared to *Block*RDT; but on the other hand, Theorem 4.3 shows that T-*Seq*RDT is faster on average than *Block*RDT. Moreover, the bounds derived in the two theorems depend on the choice of parameters w_H, w_L, w_{BH} and w_{BL}. As stated earlier, these parameters are used to select the three thresholds and the truncation window required for T-*Seq*RDT. Next, we study the behavior of the error probabilities and the stopping time with these parameters.

4.5.3 Trade-off: Error probabilities vs Stopping time

In this subsection, we study how increasing/decreasing $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ affect the average stopping time of T-*Seq*RDT. Since $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$, $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ and $\mathbb{E}[T]$ are not available in a closed form, we hereafter study the behavior of the upper bounds for $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ with respect to the upper bounds given for $\mathbb{E}[T]$.

Proposition 4.6 (Behavior with w_H and w_L)**.** *Given w_{BL} and w_{BH}, we have:*

(i) *As w_H and w_L tend to ∞, T-SeqRDT approaches BlockRDT in the sense that*

$$\lim_{w_L, w_H \to \infty} \mathbb{E}[T] = N_0 + W^*;$$

(ii) *As w_H and w_L increase, the upper bounds on $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ decrease while the upper bounds on $\mathbb{E}[T]$ increase under each hypothesis.*

PROOF:

Proof of (i): Using the definition of the expectation, we have

$$\begin{aligned}
\mathbb{E}[T] &= \sum_{N=N_0}^{N_0+W^*} N\mathbb{P}[T = N] \\
&\geqslant (N_0 + W^*)\mathbb{P}[T = N_0 + W^*] \\
&= (N_0 + W^*)\mathbb{P}\left[\bigcap_{N=N_0}^{N_0+W^*-1} [\mathcal{D}_{N_0}(N) = \infty]\right] \\
&\overset{(b)}{\geqslant} (N_0 + W^*)\left(1 - \sum_{N=N_0}^{N_0+W^*-1} \mathbb{P}\big[\mathcal{D}_{N_0}(N) \neq \infty\big]\right),
\end{aligned}$$

where (b) follows from the Boole inequality. Moreover, it follows from Proposition 3.4 (iv) that

$$\mathbb{P}\big[\mathcal{D}_{N_0}(N)=\infty\big]=\mathbb{P}\big[|\langle Y\rangle_N-\xi_0|\leqslant\lambda_H(N,w_H)\big]-\mathbb{P}\big[|\langle Y\rangle_N-\xi_0|\leqslant\lambda_L(N,w_L)\big],$$

tends to 1 when both w_H and w_L grow to ∞. Therefore, the result follows.

Proof of (ii): From Theorem 4.2, the looser upper bound on $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ is an inverse function of w_H, whereas the looser upper bound on $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$ is an inverse function of w_L. Hence, the first part of the statement.

Now, let us look at the upper bounds on the average stopping time from Proposition 4.3. Let us first look at the upper bound under $\mathcal{H}_0$. The first term N_0+W^* is independent of w_L, the second term $\beta W^*/w_L$ decreases when w_L increases and the third term $\sum_{N=N_0}^{N_0+W^*-1} Q_{\frac{1}{2}}\left(0,\lambda_{\alpha/w_H}(\tau\sqrt{N})\right)$ decreases with increasing w_H, as a consequence of Lemmas 1.1 and A.6. This implies that the upper bound $\mathbb{E}[T]$ under $\mathcal{H}_0$ will increase with increasing w_H and w_L. Similar reasoning follows for the upper bound under $\mathcal{H}_1$. ■

Proposition 4.6 plays an important role in helping us design T-*Seq*RDT. Proposition 4.6 *(i)* states that, as w_L and w_H increase, the stopping time of T-*Seq*RDT approaches the number of samples required by *Block*RDT to belong to $\mathscr{C}(\alpha/w_{BH},\beta/w_{BL})$ (see (4.10), Remark 4.2 and discussion afterwards). Moreover, from Theorem 4.2 notice that for all $\epsilon>0$, there exist w_L and w_H greater than or equal to one such that T-*Seq*RDT belongs to the class $\mathscr{C}(\alpha/w_{BH}+\epsilon,\beta/w_{BL}+\epsilon)$. This can be achieved by increasing w_H and w_L, which is equivalent to moving the thresholds $\lambda_H(N,w_H)$ and $\lambda_L(N,w_L)$ away from each other (see Proposition 4.4), hence, increasing the average stopping time of T-*Seq*RDT (Proposition 4.6*(ii)*). This implies that we can choose larger parameter values w_H and w_L, which moves the thresholds $\lambda_H(N,w_H)$ and $\lambda_L(N,w_L)$ away from each other in order to reduce $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$. At the same time, this choice of parameters (or thresholds) will increase $\mathbb{E}[T]$ of T-*Seq*RDT. Next, we analyze the behavior of $\mathbb{E}[T]$ with increasing w_{BH} and w_{BL}.

Proposition 4.7 (Behavior with w_{BH} and w_{BL})**.** *For fixed w_L and w_H, as w_{BH} and w_{BL} increase,* *the upper bounds on $\mathbb{E}[T]$ increase under each hypothesis.*

PROOF: According to Proposition 4.3 (i), we have:

$$\mathbb{E}[T] \leqslant N_0 + \sum_{N=N_0}^{N_0+W^*-1} \left[1 - \left(\beta/w_L + Q_{\frac{1}{2}}\left(0, \lambda_{\alpha/w_H}(\tau\sqrt{N})\right)\right)\right]. \tag{4.23}$$

We have $\beta/w_L < 1/2$ since $\beta < 1/2$ and $w_L \geqslant 1$. Similarly, since $\alpha < 1/2$ and $w_H \geqslant 1$, Lemma 1.1, [24, Lemma 2(ii)] and (1.4) imply that

$$\begin{aligned} Q_{\frac{1}{2}}\left(0, \lambda_{\alpha/w_H}(\tau\sqrt{N})\right) &\leqslant \frac{\alpha}{w_H} \\ &< 1/2. \end{aligned}$$

Therefore, the second term on the rhs of (4.23) is a sum of W^* positive terms. From Proposition 4.5, we know that W^* increases with increasing w_{BH} and w_{BL}. Hence the result under $\mathcal{H}_0$.

The proof under $\mathcal{H}_1$ follows similarly. ■

From Proposition 4.5, we know that increasing w_{BH} and w_{BL} will also increase the window size $W^*(w_{BH}, w_{BL})$. Proposition 4.7 above suggests that choosing a larger w_{BH} and w_{BL}, and hence a larger $W^*(w_{BH}, w_{BL})$, while keeping the parameters w_L and w_H fixed, will increase the upper bounds on the average stopping time. But, from Theorem 4.2, we see that, to guarantee that T-*Seq*RDT belongs to $\mathscr{C}(\alpha, \beta)$, w_L and w_H cannot stay fixed and must satisfy

$$w_L \geqslant \frac{w_{BL} W^*(w_{BH}, w_{BL})}{w_{BL} - 1},$$

and

$$w_H \geqslant \frac{w_{BH} W^*(w_{BH}, w_{BL})}{w_{BH} - 1}.$$

Clearly, varying w_H and w_L along with w_{BH} and w_{BL} will also have an impact on the average stopping time. It is not easy to characterize the average stopping time behavior of T-*Seq*RDT when w_{BH} and w_{BL} increase while maintaining PFA and PMD below levels α and β, respectively. However, Theorem 4.2 suggests that when w_{BH} and w_{BL} are chosen such that $w_{BH} > 1$ and $w_{BL} > 1$, any α and β can be achieved with:

$$w_L = \frac{w_{BL} W^*(w_{BH}, w_{BL})}{w_{BL} - 1} \tag{4.24}$$

and

$$w_H = \frac{w_{BH} W^*(w_{BH}, w_{BL})}{w_{BH} - 1}. \tag{4.25}$$

Now the question that arises is: how should we choose w_H, w_L, w_{BH} and w_{BH} such that $\mathbb{E}[T]$ is minimized and at the same time T-*Seq*RDT is in $\mathscr{C}(\alpha, \beta)$? The next subsection addresses this question.

4.5.4 Parameter Selection

We need to choose appropriate thresholds (4.9) and the window $W^*(w_{BH}, w_{BL})$ (4.10) such that T-*Seq*RDT belongs to $\mathscr{C}(\alpha, \beta)$, and at the same time minimizes the average stopping time. The parameters w_H, w_L and w_{BH} fully determine the thresholds (4.9), whereas w_{BH} and w_{BL} are required to design $W^*(w_{BH}, w_{BL})$. The choice of the appropriate thresholds and window thus boils down to selecting suitable values of parameters w_H, w_L, w_{BH} and w_{BL}. Using (4.24) and (4.25), we propose to choose the parameters such that the maximum of the two upper bounds on the stopping time derived in Theorem 4.3 are minimized, i.e.,

$$(w_{BH}^*, w_{BL}^*, w_H^*, w_L^*) = \underset{w_{BH}, w_{BL}, w_H, w_L}{\arg\min} \max(\text{UB}_{T_{\mathcal{H}_0}}, \text{UB}_{T_{\mathcal{H}_1}})$$
$$\text{s.t. (4.24) and (4.25) hold}, w_{BH} > 1, w_{BL} > 1, w_H \geq 1, w_L \geq 1. \tag{4.26}$$

If $w_{BH} = w_{BL}$, which implies that $w_L = w_H$, (4.26) becomes:

$$(w_{BH}^*, w_H^*) = \underset{w_{BH}, w_H}{\arg\min}\ \max(\mathrm{UB}_{T_{\mathcal{H}_0}}, \mathrm{UB}_{T_{\mathcal{H}_1}})$$
$$\text{s.t. (4.24) and (4.25) hold}, w_{BH} > 1, w_H \geq 1. \tag{4.27}$$

The above problem can be further simplified to one-dimensional search via Proposition 4.6*(ii)*, which tells us that for fixed w_{BH} (and w_{BL}), smaller w_H (and w_L) implies smaller bounds on $\mathbb{E}[T]$. Therefore, we can choose w_{BH} (and w_{BL}), hence $W^*(w_{BH}, w_{BL})$ so as to minimize w_H (and w_L) given by (4.24) and (4.25) as:

$$w_{BH}^* = \underset{w_{BH}}{\arg\min}\ \frac{w_{BH} W^*(w_{BH}, w_{BL})}{w_{BH} - 1} \text{ s.t. } w_{BH} > 1. \tag{4.28}$$

Thereby, the upper bounds derived in Theorem 4.2 are maintained equal to α and β and we expect

Algorithm 2: T-*Seq*RDT

Initialize Given N_0, τ, τ^+, α and β.

1. Choose w_{BL} and w_{BH}, thereby $W^* = W^*(w_{BH}, w_{BL})$, $w_L = \frac{w_{BL} W^*}{(w_{BL}-1)}$ and $w_H = \frac{w_{BH} W^*}{(w_{BH}-1)}$ using either of (4.26), (4.27) or (4.28).

2. Compute $\lambda_H(N, w_H)$, $\lambda_L(N, w_L)$ and $\lambda_{\text{B-RDT}}(N, w_{BH})$ from (4.9)

While $\lambda_L(N, w_L) < |\langle Y \rangle_N - \xi_0| \leqslant \lambda_H(N, w_H)$ and $N_0 \leqslant N < N_0 + W^*$
  $N = N + 1$
End
If $|\langle Y \rangle_N - \xi_0| \leqslant \lambda_L(N, w_L)$ and $N < N_0 + W^*$
  Accept $\mathcal{H}_0$
  else if $|\langle Y \rangle_N - \xi_0| > \lambda_H(N, w_H)$ and $N < N_0 + W^*$
  Reject $\mathcal{H}_0$
  else if $|\langle Y \rangle_N - \xi_0| \leqslant \lambda_{\text{B-RDT}}(N, w_{BH})$ and $N = N_0 + W^*$
  Accept $\mathcal{H}_0$
  else if $|\langle Y \rangle_N - \xi_0| > \lambda_{\text{B-RDT}}(N, w_{BH})$ and $N = N_0 + W^*$
  Reject $\mathcal{H}_0$
End If

to minimize the stopping time of T-*Seq*RDT. In the next Chapter, we experimentally show the effect of w_{BH} and w_{BL} on w_L, w_H and $\mathbb{E}[T]$, and point out that the parameters can be chosen over a wide range without significantly impacting $\mathbb{E}[T]$. Since we have a method to choose appropriate values for w_H, w_L, w_{BH} and w_{BL} from which one derives $W^* = W^*(w_{BH}, w_{BL})$, we can calculate the thresholds according to (4.9) and, then, perform T-*Seq*RDT. Algorithm 2 lists the steps of T-*Seq*RDT.

4.6 Summary

In this chapter, we extended the non-truncated algorithm, *Seq*RDT, proposed in the preceding chapter and introduced a novel truncated sequential algorithm, T-*Seq*RDT, to solve the binary hypothesis testing problem introduced in Chapter 2. In doing so, we first stated Assumption **[Behavior of $|\langle\, \Xi \rangle_N - \xi_0|$ under $\mathcal{H}_1$]**, which is required to control PFA and PMD of T-*Seq*RDT. We then introduced the optimal FSS test, *Block*RDT, and showed in Proposition 4.2 that Assumption **[Behavior of $|\langle\, \Xi \rangle_N - \xi_0|$ under $\mathcal{H}_1$]** also helps in controlling PFA and PMD of *Block*RDT. Similar to *Seq*RDT, we first analyzed the properties of the proposed thresholds and showed that the designed thresholds are appropriate for sequential testing. We derived bounds on PFA and PMD of T-*Seq*RDT and showed that we can choose the parameters of the thresholds along with the truncation window to ensure that T-*Seq*RDT belongs to the class $\mathscr{C}(\alpha, \beta)$. In contrast to *Seq*RDT, for T-*Seq*RDT we also analyzed the bounds on the average stopping time of T-*Seq*RDT and provided insights into the trade-off between the average stopping time and the error probabilities of the algorithm. Finally, we provided an approach to choose the parameters of the thresholds and the window size of the algorithm efficiently. One critical feature of the proposed algorithm, T-*Seq*RDT, is that it gives the algorithm designer freedom to choose these parameters, thus making it possible to test signals with arbitrarily low SNRs. Also, the algorithm is robust to mismatches in signal distributions as it does not rely on the underlying signal distributions.

CHAPTER 5
EXPERIMENTAL RESULTS FOR SEQUENTIAL RANDOM DISTORTION TESTING

5.1 Introduction

In this chapter, we perform some simulations to highlight the advantages of the proposed algorithms, *Seq*RDT and T-*Seq*RDT, proposed in Chapters 3 and 4, respectively, compared to the optimal FSS algorithm, *Block*RDT, discussed in Chapter 4 and SPRT as proposed in [39, 40]. Importantly, we compare the proposed algorithms to two popular composite hypothesis tests, GSPRT [36] and WSPRT as defined in [40]. Moreover, we compare the algorithms for different types of signal models and show that the proposed approaches are robust to mismatches compared to the likelihood ratio based approaches.

In the following, we first present the signal model and discuss the experimental setup. We show how popular mean-testing algorithms can be framed in the *Seq*RDT and T-*Seq*RDT frameworks proposed in this dissertation. Then we present three different signal models and compare the performance of the proposed algorithms with likelihood ratio based approaches. We demonstrate how to choose the parameters for the proposed algorithms using the techniques presented in Chapters 3 and 4. Finally, we conclude the section with some numerical results and analysis of the robustness of the proposed frameworks through numerical simulations. Next, we discuss the experimental setup for the numerical simulations.

5.2 Experimental Results and Discussion

In this section, we first present the signal model assumed for simulation purposes. Then we discuss how to frame a standard mean-testing algorithm in the sequential testing framework proposed in this dissertation. In the later part of the section, we discuss three different practical signal models we have chosen to conduct the experiments. Next, we discuss the signal model.

5.2.1 Detection with signal distortions

We address the problem of testing the mean of a signal. Let us first consider the case when:

$$Y_n = \Xi_n + X_n, \text{ for } n \in \mathbb{N}$$

with the signal, $\Xi = (\Xi)_{n\in\mathbb{N}}$, we have

$$\Xi_n = \xi_0 \text{ under } \mathcal{H}_0 \text{ vs } \Xi_n \neq \xi_0 \text{ under } \mathcal{H}_1.$$

Here, ξ_0 is a deterministic constant and the noise is Gaussian, i.e., $X_n \sim \mathcal{N}(0,1)$ for all $n \in \mathbb{N}$. This model can be formulated in the sequential framework presented in this work (please see (2.1)) with $\tau = 0$ and $N_0 = 1$. This is the classical Gaussian mean-testing problem.

However, in many practical systems, there might be a mismatch between the model and the actual signal. In many practical applications, the underlying signal, Ξ_n, will not be a constant ξ_0 under the null hypothesis, $\mathcal{H}_0$, but a perturbed version of this value. These unavoidable perturbations are difficult to model in a parametric setup. Therefore, likelihood ratio based tests fail to guarantee reliable performance [9, 24, 36]. However, the T-*Seq*RDT and the *Seq*RDT setups proposed in Chapters 3 and 4, respectively, are not limited by these drawbacks. Therefore, instead of dealing with a perfect model as described above, we consider the case when the actual signal, Ξ_n, is a distorted version of ξ_i for $i \in \{0, 1\}$:

$$\Xi_n = \xi_i + \Delta_n \text{ under } \mathcal{H}_i \text{ for } i \in \{0,1\} \text{ and } n \in \mathbb{N}.$$

Here, the Δ_ns model possible perturbations with unknown distribution. We thus want to experimentally assess different algorithms for testing $\Xi = (\Xi_n)_{n\in\mathbb{N}}$ when we observe $Y = (Y_n)_{n\in\mathbb{N}}$ (please see (2.1), (3.1) or (4.1)). We focus on algorithms in class $\mathscr{C}(\alpha, \beta)$. If the distributions in play are perfectly known, SPRT is optimal in the sense that it makes a faster decision on average, compared to all other algorithms in class $\mathscr{C}(\alpha, \beta)$. Otherwise, if the distributions are not completely known and only partial knowledge of the distortions is assumed, the above hypothesis testing problem can easily be formulated in the framework of (2.1). Then the problem can be solved by *Block*RDT, *Seq*RDT or T-*Seq*RDT. In this respect, we hereafter benchmark T-*Seq*RDT against WSPRT, GSPRT, SPRT, *Block*RDT and *Seq*RDT under experimental settings described below.

5.2.2 Experimental setup

We first list the parameters required to design each algorithm. *Block*RDT only requires τ, but guarantees $\mathbb{P}_{\text{FA}}^{\text{B-RDT}}$ only, with no control over $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}$. With additional knowledge of τ^+, *Block*RDT can control both $\mathbb{P}_{\text{FA}}^{\text{B-RDT}}$ and $\mathbb{P}_{\text{MD}}^{\text{B-RDT}}$ as illustrated in Proposition 4.2. Likewise, T-*Seq*RDT also requires τ and τ^+, whereas *Seq*RDT requires τ^-, τ, τ^+ and τ_H. On the other hand, SPRT requires complete knowledge of the signal distributions under each hypothesis. Similarly, WSPRT and GSPRT also require precise knowledge of the signal distributions at least up to an unknown (possibly vector) parameter. Note that *Block*RDT is a FSS algorithm whereas the rest of the algorithms are sequential and belong to class $\mathscr{C}(\alpha, \beta)$. For the experimental setup, let us assume τ^- to be some positive real value. We consider ξ_1 and ξ_0 such that $|\xi_1 - \xi_0| \geqslant 4\tau^-$. We set $\tau^+ = |\xi_1 - \xi_0| - \tau^-$

and $\tau_H \in [|\xi_1 - \xi_0| + \tau^-, \infty)$. Suppose that the empirical mean of the distortion $\Delta = (\Delta_n)_{n\in\mathbb{N}}$ exhibits the following bounded behavior: there exists some $N_0 \in \mathbb{N}$ such that, for all $N \geqslant N_0$, $0 \leqslant |\langle \Delta \rangle_N| \leqslant \tau^-$ and $\tau^+ \leqslant |\langle \Delta \rangle_N + \xi_1 - \xi_0| \leqslant \tau_H$. The first inequality captures the signal behavior under $\mathcal{H}_0$, whereas the second inequality captures the signal behavior under $\mathcal{H}_1$. The problem of testing the mean of $\Xi = (\Xi_n)_{n\in\mathbb{N}}$ can be rewritten as:

$$\begin{cases} \text{under } \mathcal{H}_0 : \forall N \geqslant N_0, 0 \leqslant |\langle \Xi \rangle_N - \xi_0| \leqslant \tau^- < \tau \ \text{ (a-s)}, \\ \text{under } \mathcal{H}_1 : \forall N \geqslant N_0, \tau < \tau^+ \leqslant |\langle \Xi \rangle_N - \xi_0| \leqslant \tau_H \ \text{ (a-s)}. \end{cases} \tag{5.1}$$

We can choose $\tau \in (\tau^-, \tau^+)$. For simulation purposes, we set $\tau = 2\tau^-$. Note that (5.1) is a special case of the hypothesis testing problem (2.1) and can thus be tested using the *Block*RDT, *Seq*RDT and T-*Seq*RDT frameworks. None of these algorithms need the complete knowledge of the distortion (or signal) distributions under either hypothesis, unlike SPRT, WSPRT and GSPRT which require the precise knowledge of these distributions under both hypotheses at least up to parametric uncertainty. We consider three different types of distortions, two when (5.1) is only required to be satisfied with high probability and the third when it is satisfied with probability 1 (in (a-s) sense).

Case 1: Gaussian distortion: We assume $\Delta_n \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$ for $n \in \mathbb{N}$. For simulation purposes, we choose $\tau^- = \sigma/4$. For this distortion type, the inequalities in (5.1) will only be satisfied with high-probability. Below, we list the probabilities corresponding to the Gaussian distortion. We have $\mathbb{P}[|\langle \Delta \rangle_N| \leqslant \tau] \geqslant 0.9545, \mathbb{P}[|\langle \Delta \rangle_N + \xi_1 - \xi_0| > \tau] \geqslant 0.9772, \mathbb{P}[|\langle \Delta \rangle_N| \leqslant \tau^-] \geqslant 0.6827$ and $\mathbb{P}[|\langle \Delta \rangle_N + \xi_1 - \xi_0| > \tau^+] \geqslant 0.8413$ for all $N \geqslant N_0$ with $N_0 = 16$ and $|\xi_1 - \xi_0| \geqslant 2\tau$. Note that these probabilities increase with N.

Case 2: Heavy-Tailed distortion: We model Δ_n as an $\bar{\alpha}$-stable random variable denoted as $\Delta_n \overset{\text{iid}}{\sim} \mathcal{S}(\bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{\delta})$ for $n \in \mathbb{N}$ [28]. The parameters $\bar{\alpha} \in (0, 2], \bar{\beta} \in [-1, 1], \bar{\gamma} > 0$ and $\bar{\delta} \in (-\infty, \infty)$ are the tail-index, location, dispersion and skewness parameters, respectively. In general, an $\bar{\alpha}$-stable distribution does not admit a closed-form probability density function, except in a few special cases like the Cauchy ($\bar{\alpha} = 1, \bar{\beta} = 0$) and Gaussian ($\bar{\alpha} = 2$) distributions. Moreover, for the Cauchy distribution and for $\bar{\alpha} \in (0, 1]$, none of the moments of the $\bar{\alpha}$-stable distribution exist. For $\bar{\alpha} \in (1, 2)$ the distribution is sometimes referred to as the Pareto-Lévy distribution and for this class of distributions, all higher moments beyond the mean do not exist. For simulation purposes, we consider the following two types of heavy-tailed distortions:

Case 2(i) [Pareto-Lévy distortion $(\bar{\alpha} \in (1, 2))$*]:* We assume the distortion to be Pareto-Lévy distributed with $\Delta_n \overset{\text{iid}}{\sim} \mathcal{S}(1.5, 0, \tau^-, 0)$ for $n \in \mathbb{N}$. We thus have: $\mathbb{P}[|\langle \Delta \rangle_N| \leqslant \tau] \geqslant 0.9885, \mathbb{P}[|\langle \Delta \rangle_N + \xi_1 - \xi_0| > \tau] \geqslant 0.9953, \mathbb{P}[|\langle \Delta \rangle_N| \leqslant \tau^-] \geqslant 0.9646, \mathbb{P}[|\langle \Delta \rangle_N + \xi_1 - \xi_0| > \tau^+] \geqslant 0.9832$ for all $N \geqslant N_0$ with $N_0 = 30$ and $|\xi_1 - \xi_0| \geqslant 2\tau$. Again, note that these probabilities increase with N.

Case 2(ii) [Cauchy distortion $(\bar{\alpha} = 1)$*]:* Note that, unlike in the cases involving Gaussian and Pareto-Lévy distortions, the empirical mean of i.i.d Cauchy distributed random variables is again Cauchy distributed [28] and none of the moments exist for the Cauchy distribution, thereby, none of the moments exist for the empirical mean as well. Therefore, the empirical mean of a Cauchy distorted signal does not converge in the neighborhood of ξ_0 and ξ_1 under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively, in contrast to the Gaussian and Pareto-Lévy distortions as discussed above. Below, we show that, although the Cauchy distortion does not exhibit the desired convergence properties, the proposed algorithms guarantee performance if (5.1) holds with sufficiently high probabilities. To experimentally show this, we assume the distortion to be Cauchy distributed as $\Delta_n \overset{\text{iid}}{\sim} \mathcal{S}(1, 0, \tau^-/10, 0)$ for $n \in \mathbb{N}$ with the associated probabilities given as: $\mathbb{P}[|\langle \Delta \rangle_N| \leqslant \tau] = 0.9682$ and $\mathbb{P}[|\langle \Delta \rangle_N| \leqslant \tau^-] = 0.9365$ for all $N \in \mathbb{N}$. Also, $\mathbb{P}[|\langle \Delta \rangle_N + \xi_1 - \xi_0| > \tau] \geqslant 0.9894$ and $\mathbb{P}[|\langle \Delta \rangle_N + \xi_1 - \xi_0| > \tau^+] \geqslant 0.9728$ for all $N \in \mathbb{N}$ and $|\xi_1 - \xi_0| \geqslant 2\tau$. Note that, unlike the *Cases 1* and *2(i)* these probabilities do not increase with N as the distribution of the empirical mean of a Cauchy distribution remains the same. As a consequence, the probabilities stay the same for all $N \in \mathbb{N}$. To ensure that (5.1) is satisfied with high probability, we need the dispersion parameter, $\bar{\gamma}$, to be small enough. Later in the chapter, we show how the above probabilities, PFA and PMD vary with $\bar{\gamma}$ for T-*Seq*RDT.

Case 3 [Deterministic unknown distortion]: The distortion is assumed to be unknown deterministic with $|\Delta_n| \leqslant \tau^-$ for all $n \in \mathbb{N}$. For simulation purposes, we choose $\Delta_n = \tau^-$. With this choice, the inequalities in (5.1) are satisfied with probability 1. However, not all types of distortions satisfy (5.1) with probability 1 as shown in *Cases 1* and *2*. In the next section, we discuss different algorithms.

5.3 Algorithms: Likelihood Ratio Based Approaches

In this section, we discuss the algorithms we use to solve the above mean testing problem. Certainly, using (5.1), we can cast the problem in the *Block*RDT, *Seq*RDT and T-*Seq*RDT frameworks presented in Chapters 3 and 4. Next, we discuss likelihood ratio based parametric and semi-parametric approaches that we use for comparison purposes.

5.3.1 Sequential probability ratio test (SPRT)

For SPRT, we assume that the probability density function, f_i, of the observations is known under $\mathcal{H}_i$ for $i = 0, 1$. For $\alpha, \beta \in (0, 1/2)$, and with initialization $\Lambda_N = 1$, SPRT with stopping time and decision pair $(T_{\text{SPRT}}, \mathcal{D})$ is defined as:

$$T_{\text{SPRT}} = \inf\{N \geqslant 0 : \Lambda_N \notin (\lambda_L^{\text{SPRT}}, \lambda_H^{\text{SPRT}})\}$$

$$\mathcal{D}(N) = \begin{cases} 1 \ \text{ if } \Lambda_N \geqslant \lambda_H^{\text{SPRT}} \\ 0 \ \text{ if } \Lambda_N \leqslant \lambda_L^{\text{SPRT}} \\ \infty \ \text{ if } \lambda_L^{\text{SPRT}} < \Lambda_N < \lambda_H^{\text{SPRT}} \end{cases}$$

where

$$\Lambda_N = \sum_{n=1}^{N} \frac{f_1(Y_i)}{f_0(Y_i)}$$

is the likelihood ratio based on the observations,

$$\lambda_L^{\text{SPRT}} = \frac{\beta}{1-\alpha} \text{ and } \lambda_H^{\text{SPRT}} = \frac{1-\beta}{\alpha}$$

are the lower and upper thresholds, respectively. We denote the stopping time, PFA and PMD of SPRT as T_{SPRT}, $\mathbb{P}_{\text{FA}}^{\text{SPRT}}$ and $\mathbb{P}_{\text{MD}}^{\text{SPRT}}$, respectively. For the model described above, SPRT for detecting the mean with unknown distortions we have

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

5.3.2 Composite hypothesis test, GSPRT

Note that as discussed in Chapters 1 and 2, standard GSPRT does not work for the two-sided hypothesis testing problem. However, a simple GSPRT can be designed for the case of Gaussian distortions when the means under $\mathcal{H}_0$ and $\mathcal{H}_1$ are known but the variances are unknown. Specifically, the algorithm is aware that the distortion is zero mean Gaussian distributed, but is unaware of its variance [9]. The generalized log likelihood ratio for such a test is given as:

$$\log\hat{\Lambda}_N = \frac{\xi_1-\xi_0}{s_n^2}\sum_{n=1}^{N}\left(Y_n - \frac{1}{2}(\xi_o+\xi_1)\right),$$

with

$$s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \langle Y\rangle_i),$$

for $N \geq 2$. GSPRT uses the same thresholds as SPRT [9]. We denote the stopping time, PFA and PMD of GSPRT as $\mathbb{E}\left[T_{\text{GSPRT}}\right]$, $\mathbb{P}_{\text{FA}}^{\text{GSPRT}}$ and $\mathbb{P}_{\text{MD}}^{\text{GSPRT}}$, respectively.

5.3.3 Composite hypothesis test, WSPRT

The idea of WSPRT is somewhat similar to the sequential testing frameworks introduced in this work. However, there are a few key differences as discussed below. WSPRT considers the problem of testing [40, Chapter 4]:

$$\mathcal{H}_0 : |\xi - \xi_0| < \tau \text{ vs } \mathcal{H}_1 : |\xi - \xi_0| > \tau.$$

In *Seq*RDT and T-*Seq*RDT, the signal Ξ is assumed to be a corrupted version of ξ and the distribution of Ξ is assumed to be unknown. In contrast, for WSPRT, ξ is deterministic and the test can only handle the case when the observations, Y_is, are Gaussian distributed [36,40]. We denote PFA, PMD and the stopping time of WSPRT as $\mathbb{P}_{\text{FA}}^{\text{WSPRT}}$, $\mathbb{P}_{\text{MD}}^{\text{WSPRT}}$ and T_{WSPRT}, respectively. WSPRT uses the same thresholds as SPRT. However, the likelihood ratio for WSPRT is given as:

$$\widehat{\Lambda}_N = \prod_{n=1}^{N} \frac{e^{-(Y_n - \xi_0 + \tau)^2/2} + e^{-(Y_n - \xi_0 - \tau)^2/2}}{2e^{-(Y_n - \xi_0)^2/2}}.$$

Note that in the above the distribution under the alternate hypothesis, $\mathcal{H}_1$, is replaced by a weighted average of two distributions and thus the name WSPRT.

5.3.4 Comparison

We define $|\xi_1 - \xi_0|$ as the SNR and for simulation purposes, we assume $\tau^- = 0.1$. For T-*Seq*RDT, the thresholds as given in (4.9):

$$\begin{aligned}
\lambda_H(N) &= \lambda_H(N, w_H) = \lambda_{\alpha/w_H}(\tau\sqrt{N})/\sqrt{N} \\
\lambda_L(N) &= \lambda_L(N, w_L) = \lambda_{1-\beta/w_L}(\tau\sqrt{N})/\sqrt{N} \\
\lambda_{\text{B-RDT}}(N) &= \lambda_{\text{B-RDT}}(N, w_{BH}) = \lambda_{\alpha/w_{BH}}(\tau\sqrt{N})/\sqrt{N},
\end{aligned}$$

and the truncation window $W^*(w_{BH}, w_{BL})$ (4.10) defined in Chapter 4 and is as given below:

$$\begin{aligned}
W^* &= W^*(w_{BH}, w_{BL}) \\
&= \min\left\{ W \in \mathbb{N} : 1 - Q_{\frac{1}{2}}\left(\tau^+\sqrt{N_0 + W}, \lambda_{\frac{\alpha}{w_{BH}}}(\tau\sqrt{N_0 + W})\right) \leqslant \frac{\beta}{w_{BL}} \right\}
\end{aligned}$$

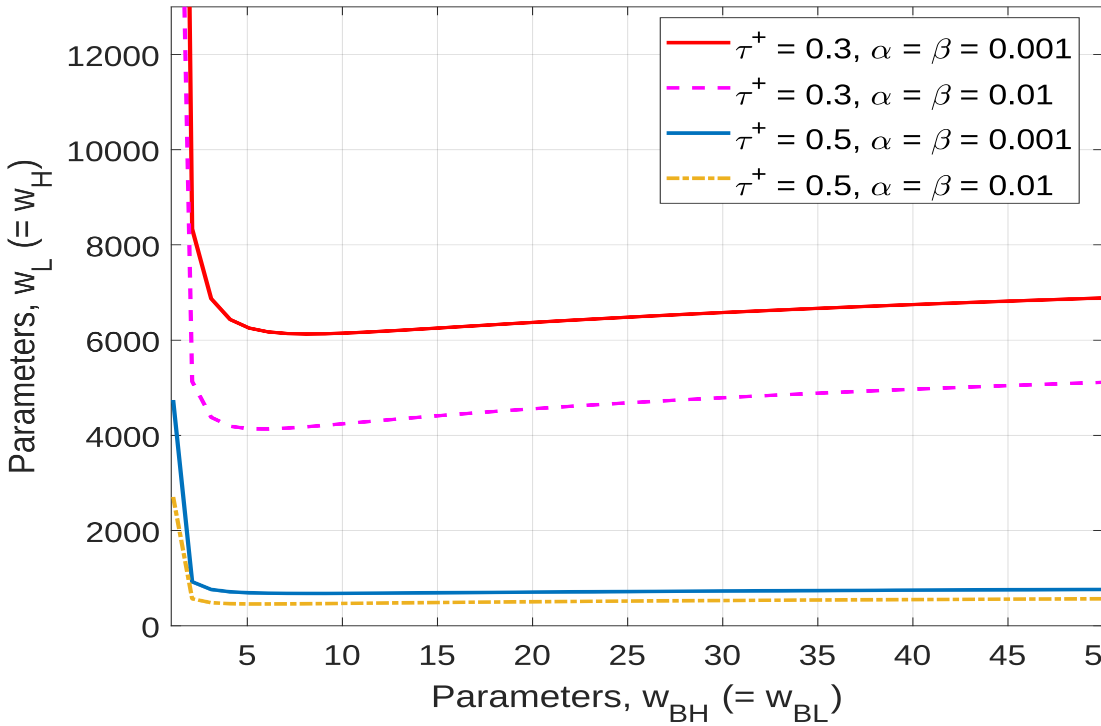

Fig. 5.1: $w_H = w_L$ vs $w_{BH} = w_{BL}$ such that UB_{FA} and UB_{MD} in Theorem 4.2 stay equal to α and β, respectively.

are selected via parameters w_H, w_L, w_{BH} and w_{BL}, using Algorithm 2 given in Chapter 4. For the simulations, we assume $w_{BL} = w_{BH}$, which implies that $w_H = w_L$ and make use of the method proposed in (4.28) to choose the parameters. We denote the PFA and PMD by $\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$ and $\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$, respectively, and the stopping time of T-*Seq*RDT bt T.

In Figures 5.1 and 5.2, we plot w_H against w_{BH} and the stopping time of T-*Seq*RDT, $\mathbb{E}[T]$, respectively, for Gaussian distortion as discussed in *Case 1* above and for different α, β and τ^+. We notice that w_H (and w_L) as plotted in Figure 5.1 varies in a similar manner to the average stopping time of T-*Seq*RDT, $\mathbb{E}[T]$, as plotted in Figure 5.2. This implies, by choosing a value of w_{BH} (and w_{BL}) which minimizes w_H (and w_L), we expect to minimize $\mathbb{E}[T]$ as well. This behavior was also suggested by Proposition 4.6*(ii)*. Moreover, we see that we can choose w_{BH} and w_{BL} (hence w_H and w_L) over a wide range without affecting $\mathbb{E}[T]$ considerably. For simulation purposes, we select $w_{BH} = w_{BL} = 2$, which is then used to choose an appropriate W^* via (4.10). We then set $w_H = w_L = 2W^*$. Albeit not unique, these choices guarantee that T-*Seq*RDT is in

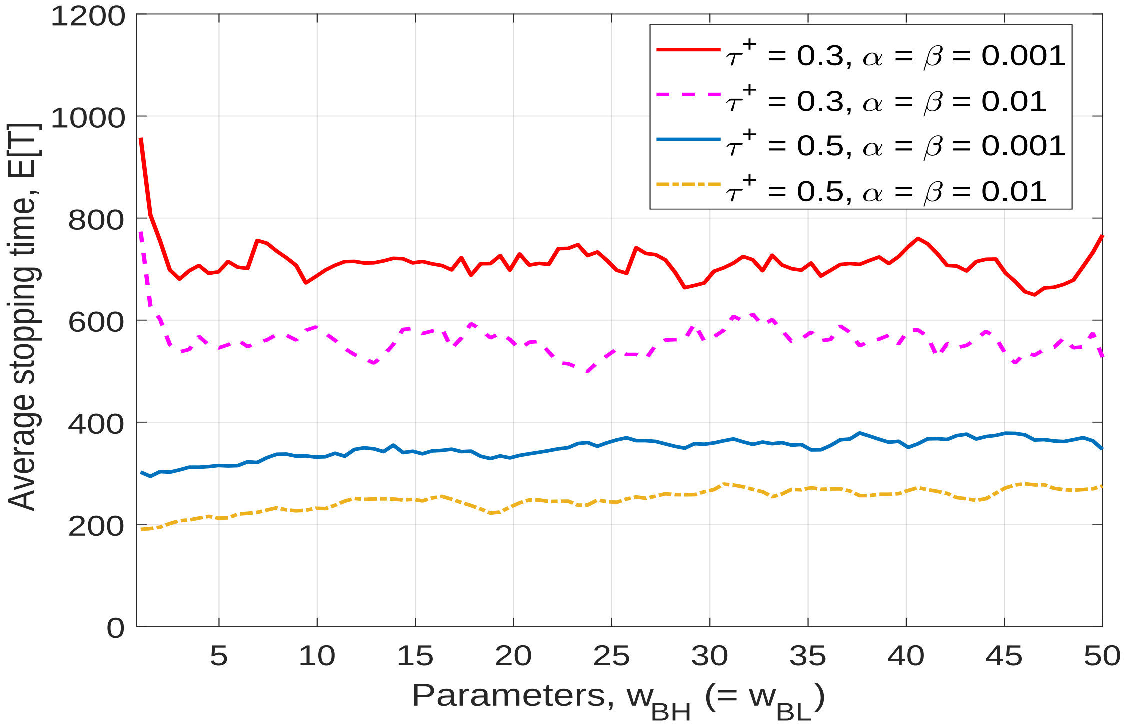

Fig. 5.2: $\mathbb{E}[T]$ vs $w_{BH} = w_{BL}$ such that $\mathrm{UB}_{\mathrm{FA}}$ and $\mathrm{UB}_{\mathrm{MD}}$ in Theorem 4.2 stay equal to α and β, respectively.

$\mathscr{C}(\alpha, \beta)$ by Theorem 4.2.

In contrast to T-*Seq*RDT, *Seq*RDT requires additional knowledge of τ^- and τ_H to design the test: τ along with levels α and β is used to design the thresholds; τ^+, τ^- and τ_H are used to choose an appropriate buffer size, M, or the parameter, w_{N_0}, which can be chosen to eliminate the buffer, M. The algorithm is explained in Algorithm 1 in Chapter 3. For simulation purposes, we work with Option I where instead of w_{N_0} we choose a buffer, M, to control PFA and PMD and ensure that the algorithm belongs to $\mathscr{C}(\alpha, \beta)$. The thresholds for *Seq*RDT are chosen as (3.12), and are as listed below:

$$\frac{\lambda_\alpha(\tau\sqrt{N})}{\sqrt{N}} \quad \text{and} \quad \frac{\lambda_{1-\beta}(\tau\sqrt{N})}{\sqrt{N}}$$

We denote by $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_M)$ and $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_M)$ the PFA and PMD, respectively, and by T_{SeqRDT} the stopping time of *Seq*RDT.

Case 1: For Gaussian distortion, we compare in Table 5.1 the average stopping times of T-*Seq*RDT and *Seq*RDT to the block-size of *Block*RDT, for different SNR values, $\alpha = \beta = 0.01$ and

$\alpha = \beta = 0.01$					
SNR $= \lvert\xi_1 - \xi_0\rvert$		0.4	0.5	0.6	0.8
*Block*RDT	$N_{\text{B-RDT}}$	2165	542	241	87
*Seq*RDT	$\mathbb{E}\left[T_{SeqRDT}\right]$	171.34	145.83	141.12	140.26
T-*Seq*RDT	$\mathbb{E}\left[T\right]$	567.73	349.42	192.82	73.92
$\alpha = \beta = 0.001$					
*Block*RDT	$N_{\text{B-RDT}}$	3820	955	425	153
*Seq*RDT	$\mathbb{E}\left[T_{SeqRDT}\right]$	252.55	198.51	185.17	181.80
T-*Seq*RDT	$\mathbb{E}\left[T\right]$	720.08	481.75	298.65	114.88

Table 5.1: Comparison of T-*Seq*RDT, *Seq*RDT and *Block*RDT for Gaussian distortion.

$\alpha = \beta = 0.001$. For *Seq*RDT, the buffer size $M = 90$ is selected. From Table 5.1, *Seq*RDT is the fastest on average, especially at low SNR values, but needs the most amount of information (all of τ^-, τ, τ^+ and τ_H to design M) about the signal. *Block*RDT is the slowest and requires the same information (τ and τ^+ only) as T-*Seq*RDT. However, T-*Seq*RDT is considerably faster on average. Moreover, at moderate to high SNRs, T-*Seq*RDT is the fastest among the three algorithms and considerably outperforms *Seq*RDT as well. It must be noted that the stopping time of *Seq*RDT is limited by the need and the choice of the buffer size, which makes *Seq*RDT relatively slower compared to T-*Seq*RDT, especially at higher SNRs. Importantly, it must be noted that since T-*Seq*RDT is a truncated algorithm, its stopping times will never be higher than that of the the FSS test, *Block*RDT, while achieving the probabilities of errors of the same order (please see Chapter 4 for more detailed discussion).

In Table 5.2, we compare the average stopping times, PFAs and PMDs of T-*Seq*RDT, *Seq*RDT, SPRT, WSPRT and GSPRT. From Table 5.2, we notice that, because of the distortion, WSPRT and SPRT do not belong to $\mathscr{C}(\alpha, \beta)$ as $\mathbb{P}_{\text{FA}}^{\text{SPRT}}$, $\mathbb{P}_{\text{MD}}^{\text{SPRT}}$ and $\mathbb{P}_{\text{FA}}^{\text{WSPRT}}$ are above the pre-specified levels α and β. Moreover, GSPRT, even with prior knowledge of the distortion, does not belong to $\mathscr{C}(\alpha, \beta)$, as the $\mathbb{P}_{\text{FA}}^{\text{GSPRT}}$ and $\mathbb{P}_{\text{MD}}^{\text{GSPRT}}$ are orders of magnitude higher compared to α and β, respectively. This implies that even though GSPRT is asymptotically optimal, it is of little practical significance as it does not guarantee performance in the non-asymptotic regimes. In contrast, both *Seq*RDT and T-*Seq*RDT, with only limited knowledge about the signal under each hypothesis, are in $\mathscr{C}(\alpha, \beta)$. Importantly, by design, T-*Seq*RDT eliminates the need for buffer M, whereas *Seq*RDT does need

$\alpha = \beta = 0.01$					
SNR = $\lvert \xi_1 - \xi_0 \rvert$		0.4	0.5	0.6	0.8
T-*Seq*RDT	$\mathbb{E}[T]$	567.73	349.42	192.82	73.92
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$	0.0004	0.0001	0.0001	0.0002
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$	0.0001	0.0003	0.0002	0.0003
*Seq*RDT	$\mathbb{E}[T_{SeqRDT}]$	171.34	145.83	141.12	140.26
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$	0.00015	0.0002	0.0002	0.00021
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$	0.00044	0.00013	4×10^{-5}	$< 10^{-5}$
SPRT	$\mathbb{E}[T_{\text{SPRT}}]$	58.98	38.40	27.18	15.84
	$\mathbb{P}_{\text{FA}}^{\text{SPRT}}$	0.0150	0.0146	0.0131	0.0114
	$\mathbb{P}_{\text{MD}}^{\text{SPRT}}$	0.0149	0.0143	0.0142	0.0124
WSPRT	$\mathbb{E}[T_{\text{WSPRT}}]$	209.77	198.55	191.85	184.07
	$\mathbb{P}_{\text{FA}}^{\text{WSPRT}}$	0.0192	0.0185	0.0187	0.0182
	$\mathbb{P}_{\text{MD}}^{\text{WSPRT}}$	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$
GSPRT	$\mathbb{E}[T_{\text{GSPRT}}]$	41.47	26.01	18.07	10.68
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$	0.1307	0.1339	0.1377	0.1292
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$	0.1430	0.1481	0.1433	0.1325
$\alpha = \beta = 0.001$					
SNR = $\lvert \xi_1 - \xi_0 \rvert$		0.4	0.5	0.6	0.8
T-*Seq*RDT	$\mathbb{E}[T]$	720.08	481.75	298.65	114.88
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$	3×10^{-5}	2×10^{-5}	5×10^{-5}	1×10^{-5}
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$	1×10^{-5}	3×10^{-5}	2×10^{-5}	$< 10^{-5}$
*Seq*RDT	$\mathbb{E}[T_{SeqRDT}]$	252.55	198.51	185.17	181.80
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$	3×10^{-5}	1×10^{-5}	4×10^{-5}	2×10^{-5}
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$	0.00011	3×10^{-5}	1×10^{-5}	$< 10^{-5}$
SPRT	$\mathbb{E}[T_{\text{SPRT}}]$	89.24	57.82	40.56	23.32
	$\mathbb{P}_{\text{FA}}^{\text{SPRT}}$	0.0022	0.0019	0.0018	0.0016
	$\mathbb{P}_{\text{MD}}^{\text{SPRT}}$	0.0023	0.0020	0.0019	0.00018
WSPRT	$\mathbb{E}[T_{\text{WSPRT}}]$	304.37	288.09	278.55	267.46
	$\mathbb{P}_{\text{FA}}^{\text{WSPRT}}$	0.0033	0.0033	0.0030	0.0036
	$\mathbb{P}_{\text{MD}}^{\text{WSPRT}}$	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$
GSPRT	$\mathbb{E}[T_{\text{GSPRT}}]$	68.75	42.39	28.76	16.05
	$\mathbb{P}_{\text{FA}}^{\text{GSPRT}}$	0.0937	0.0970	0.0981	0.0994
	$\mathbb{P}_{\text{MD}}^{\text{GSPRT}}$	0.1021	0.1059	0.1047	0.1003

Table 5.2: Comparison of T-*Seq*RDT, *Seq*RDT, SPRT and WSPRT for Gaussian distortion. Here, PFA $< 10^{-5}$ and PMD $< 10^{-5}$ indicate that probabilities of errors are at most of the order of 10^{-5}.

such a buffer to guarantee the pre-specified levels α and β. Moreover, the bounds on PFA and PMD are loose for Gaussian distortion. Therefore, we next consider different types of distortions

$\alpha = \beta = 0.01$					
SNR $= \lvert\xi_1 - \xi_0\rvert$		0.4	0.5	0.6	0.8
T-*Seq*RDT	$\mathbb{E}[T]$	576.14	350.71	192.296	75.17
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$	0.0097	0.0095	0.0079	0.0075
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$	0.0003	0.0002	0.0004	0.0013
*Seq*RDT	$\mathbb{E}[T_{SeqRDT}]$	171.71	146.85	142.97	141.81
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$	0.0088	0.0085	0.0086	0.0089
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$	0.0044	0.0013	0.0004	0.0006
SPRT	$\mathbb{E}[T_{\text{SPRT}}]$	58.56	37.83	27.10	15.82
	$\mathbb{P}_{\text{FA}}^{\text{SPRT}}$	0.0211	0.0204	0.0167	0.0132
	$\mathbb{P}_{\text{MD}}^{\text{SPRT}}$	0.0200	0.0199	0.0165	0.0148
WSPRT	$\mathbb{E}[T_{\text{WSPRT}}]$	207.63	195.07	189.20	182.07
	$\mathbb{P}_{\text{FA}}^{\text{WSPRT}}$	0.0454	0.0495	0.0517	0.0505
	$\mathbb{P}_{\text{MD}}^{\text{WSPRT}}$	0.0001	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$

Table 5.3: Comparison of T-*Seq*RDT, *Seq*RDT, SPRT and WSPRT for Pareto-Lévy Distortion.

$\alpha = \beta = 0.05$					
SNR $= \lvert\xi_1 - \xi_0\rvert$		0.4	0.5	0.6	0.8
T-*Seq*RDT	$\mathbb{E}[T]$	453.05	229.33	117.30	44.59
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$	0.0471	0.0424	0.0357	0.0271
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$	0.0047	0.0040	0.0052	0.0069
*Seq*RDT	$\mathbb{E}[T_{SeqRDT}]$	124.73	115.10	113.75	111.23
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$	0.0303	0.0299	0.0283	0.0286
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$	0.0087	0.0037	0.0025	0.0013
SPRT	$\mathbb{E}[T_{\text{SPRT}}]$	35.51	23.50	17.03	10.07
	$\mathbb{P}_{\text{FA}}^{\text{SPRT}}$	0.0570	0.0491	0.0481	0.0408
	$\mathbb{P}_{\text{MD}}^{\text{SPRT}}$	0.0612	0.0531	0.0510	0.0419
WSPRT	$\mathbb{E}[T_{\text{WSPRT}}]$	132.75	124.89	119.00	114.63
	$\mathbb{P}_{\text{FA}}^{\text{WSPRT}}$	0.1198	0.1162	0.1210	0.1115
	$\mathbb{P}_{\text{MD}}^{\text{WSPRT}}$	0.0006	0.0002	$< 10^{-4}$	$< 10^{-4}$

Table 5.4: Comparison of T-*Seq*RDT, *Seq*RDT, SPRT and WSPRT for Cauchy distortion.

to see if the bounds are tight for some other scenarios.

Case 2: For heavy-tailed distortions, we again compare T-*Seq*RDT, RDT, SPRT and WSPRT. For *Cases 2(i)* and *2(ii)*, via simulation, we obtain PFA, PMD and average stopping times for

$\alpha = \beta = 0.01$ and $\alpha = \beta = 0.05$ to obtain Tables 5.3 and 5.4, respectively. The average stopping time of T-*Seq*RDT and *Seq*RDT are similar to those obtained in the Gaussian distortion case. However, the bounds on PFA and PMD are tight, as a consequence of the heavy-tailed distribution of the distortion. Moreover, similar to *Case 1*, SPRT and WSPRT do not belong to $\mathscr{C}(\alpha, \beta)$ for both *Cases 2(i)* and *2(ii)*.

Case 3: Finally, we consider the unknown deterministic distortion case. In this case, (5.1) is satisfied with probability 1, unlike in *Cases 1* and *2*. We choose $\Delta_n = \tau^-$, and via simulation obtain PFA, PMD and average stopping times for $\alpha = \beta = 0.05$ and $\alpha = \beta = 0.01$. From Table 5.5, T-*Seq*RDT and *Seq*RDT belong to $\mathscr{C}(\alpha, \beta)$, whereas SPRT and WSPRT fail to. Also, note that, similar to *Case 2*, the bounds on PFA are tight.

In conclusion, the above discussion suggests that likelihood ratio based approaches are sensitive even to small mismatches between the assumed and the true models. On the other hand, GSPRT based approaches which account for the hypotheses being composite, although being asymptotically optimal, do not guarantee performance in non-asymptotic regimes. This implies that GSPRT based approaches are of little practical interest in the practical non-asymptotic scenarios. Moreover, as discussed in Chapter 2, GSPRT based approaches are computationally heavy and cannot be easily implemented online [7]. In contrast, the algorithms proposed in this dissertation are robust to mismatches, are capable of providing sufficient performance guarantees even in non-asymptotic regimes and at the same time are simple in design and are easy to implement online.

5.4 A note of caution

The above simulation results show that the proposed algorithms *Seq*RDT and T-*Seq*RDT are robust to mismatches and can guarantee performance even in the cases when (5.1) is not always satisfied with probability 1, as shown in *Cases 1* and *2* above. Now, the question that arises is: *"When* (5.1) *is not satisfied in (a-s) sense, how high do the probabilities of events in* (5.1) *need to be so*

$\alpha = \beta = 0.01$					
SNR $= \lvert\xi_1 - \xi_0\rvert$		0.4	0.5	0.6	0.8
T-*Seq*RDT	$\mathbb{E}\,[\,T\,]$	387.86	177.24	106.44	48.42
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$	0.0059	0.0062	0.0063	0.0063
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
*Seq*RDT	$\mathbb{E}\,[\,T_{SeqRDT}\,]$	220.10	210.61	206.25	202.16
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$	0.0050	0.0065	0.0061	0.0058
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$	9×10^{-4}	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
SPRT	$\mathbb{E}\,[\,T_{\text{SPRT}}\,]$	70.77	43.40	29.75	16.73
	$\mathbb{P}_{\text{FA}}^{\text{SPRT}}$	0.0802	0.0537	0.0333	0.0222
	$\mathbb{P}_{\text{MD}}^{\text{SPRT}}$	0.0006	0.0012	0.0012	0.0014
WSPRT	$\mathbb{E}\,[\,T_{\text{WSPRT}}\,]$	317.05	310.01	305.50	303.89
	$\mathbb{P}_{\text{FA}}^{\text{WSPRT}}$	0.4566	0.4439	0.4538	0.4569
	$\mathbb{P}_{\text{MD}}^{\text{WSPRT}}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
$\alpha = \beta = 0.05$					
SNR $= \lvert\xi_1 - \xi_0\rvert$		0.4	0.5	0.6	0.8
T-*Seq*RDT	$\mathbb{E}\,[\,T\,]$	232.18	109.74	63.85	29.34
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_{N_0})$	0.0244	0.0259	0.0232	0.0275
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_{N_0})$	0.0002	0.0004	0.0005	0.0010
*Seq*RDT	$\mathbb{E}\,[\,T_{SeqRDT}\,]$	120.23	105.07	101.30	98.78
	$\mathbb{P}_{\text{FA}}(\mathcal{D}_M)$	0.0441	0.0442	0.0440	0.0466
	$\mathbb{P}_{\text{MD}}(\mathcal{D}_M)$	0.0052	6×10^{-4}	2×10^{-4}	$< 10^{-4}$
SPRT	$\mathbb{E}\,[\,T_{\text{SPRT}}\,]$	39.56	25.38	17.87	10.50
	$\mathbb{P}_{\text{FA}}^{\text{SPRT}}$	0.1653	0.1265	0.1029	0.0694
	$\mathbb{P}_{\text{MD}}^{\text{SPRT}}$	0.0070	0.0100	0.0098	0.0113
WSPRT	$\mathbb{E}\,[\,T_{\text{WSPRT}}\,]$	160.72	156.14	152.06	149.89
	$\mathbb{P}_{\text{FA}}^{\text{WSPRT}}$	0.4299	0.4272	0.4253	0.4202
	$\mathbb{P}_{\text{MD}}^{\text{WSPRT}}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$

Table 5.5: Comparison of T-*Seq*RDT, *Seq*RDT, SPRT and WSPRT for deterministic distortion.

that the proposed algorithms belong to $\mathscr{C}(\alpha, \beta)$?" The simulation results of *Case 1* suggest that Gaussian distortions allow for large mismatches, i.e., *Seq*RDT and T-*Seq*RDT work even when the probabilities of events in (5.1) are not very high. On the other hand, *Case 2*, involving heavy-tailed distortions, requires these probabilities to be high, i.e., a relatively smaller mismatch. In the following, we focus our attention on T-*Seq*RDT for two different cases and show via numerical

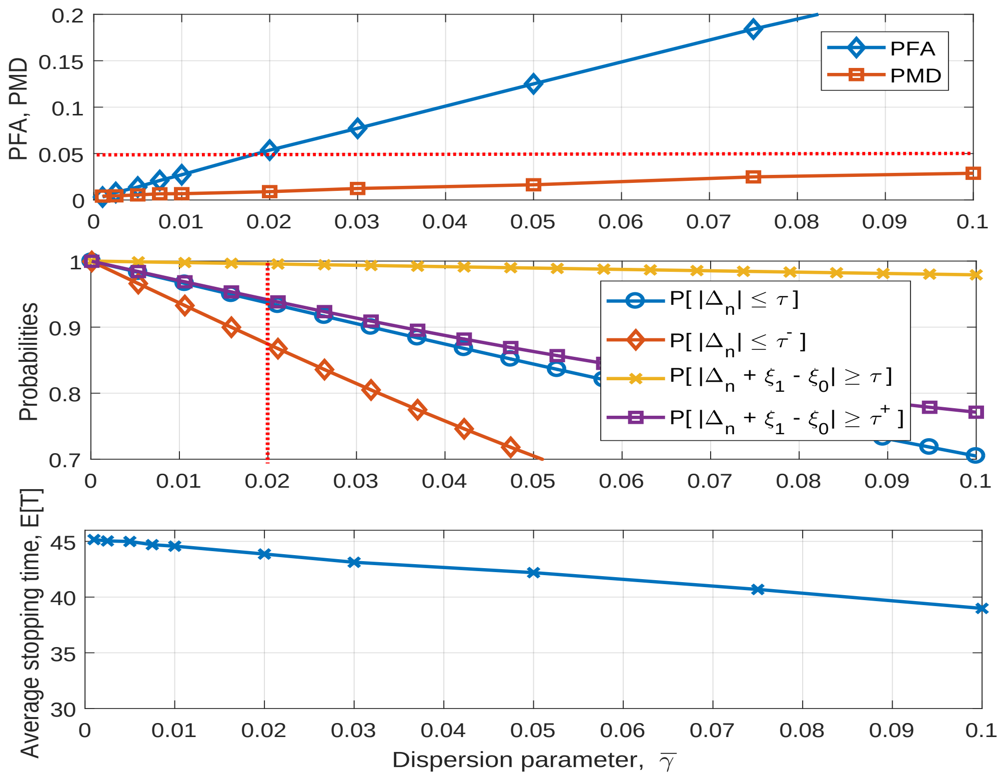

Fig. 5.3: $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_{N_0})$, $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_{N_0})$, Probabilities in (5.1) and $\mathbb{E}[T]$ against $\bar{\gamma}$ for T-*Seq*RDT.

experiments how high the probabilities of events under (5.1) need to be for T-*Seq*RDT to belong to $\mathscr{C}(\alpha, \beta)$.

Case A: We mentioned earlier in *Case 2(ii)* that, with Cauchy distortions, we needed the dispersion parameter $\bar{\gamma}$ to be small. In Figure 5.3, we show how PFA and PMD of T-*Seq*RDT and the probabilities associated with (5.1), when they are not satisfied in (a-s) sense, vary with increasing $\bar{\gamma}$, for $\alpha = \beta = 0.05$ and SNR $= 0.8$. Notice that there exists a threshold $\bar{\gamma} = 0.02$ above which the PFA of T-*Seq*RDT exceeds the pre-specified level, $\alpha = 0.05$. This implies that for $\bar{\gamma} > 0.02$, T-*Seq*RDT does not belong to $\mathscr{C}(\alpha, \beta)$. Moreover, note that from the middle plot in Figure 5.3 where we plot the probabilities: $\mathbb{P}[|\langle \Delta \rangle_N| \leqslant \tau]$, $\mathbb{P}[|\langle \Delta \rangle_N + \xi_1 - \xi_0| > \tau]$ and $\mathbb{P}[|\langle \Delta \rangle_N + \xi_1 - \xi_0| > \tau^+]$[1] (please see *Case 2(ii)* in Section 5.2.2 above), we notice that we need these probabilities to be as high as 95% for T-*Seq*RDT to belong to $\mathscr{C}(\alpha, \beta)$.

[1]For T-*Seq*RDT, there is no constraint on the probability $\mathbb{P}[|\langle \Delta \rangle_N| \leqslant \tau^-]$, as T-*Seq*RDT does not rely on the knowledge of τ^-.

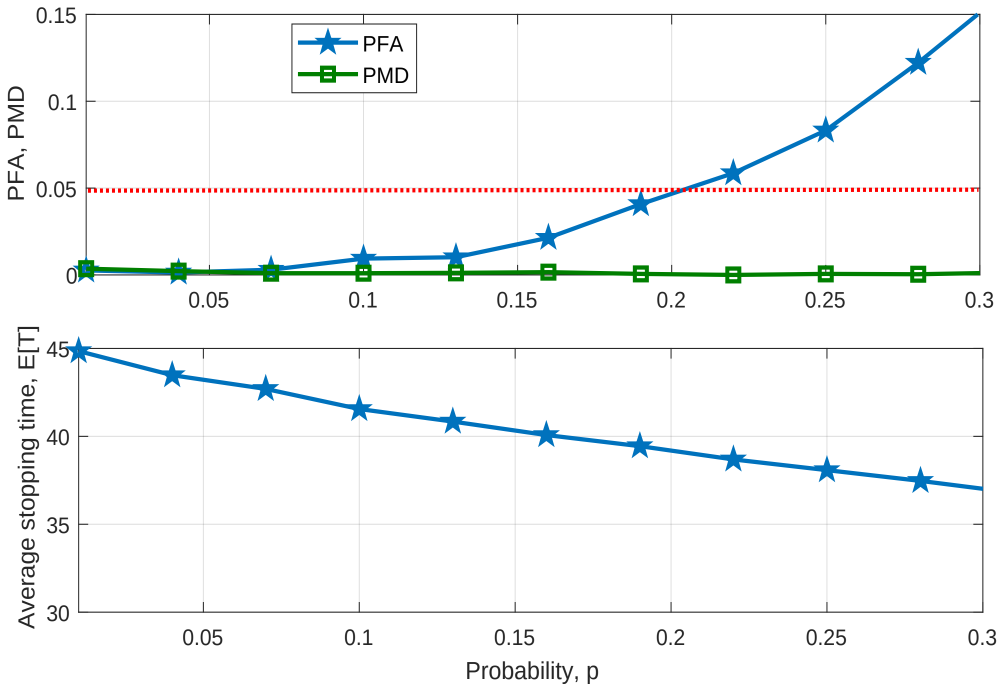

Fig. 5.4: $\mathbb{P}_{\mathrm{FA}}(\mathcal{D}_{N_0})$, $\mathbb{P}_{\mathrm{MD}}(\mathcal{D}_{N_0})$ and $\mathbb{E}[T]$ against probability of impulse, p for T-*Seq*RDT.

Case B: We perform further simulations for a simple model of impulsive distortion. We assume that Δ_n is Bernoulli distributed as

$$\Delta_n \sim \begin{cases} 10\tau^- & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$$

This implies that $\mathbb{P}\big[|\langle\,\Delta\rangle_N| \leqslant \tau\big] \geqslant 1-p$ and $\mathbb{P}\big[|\langle\,\Delta_n - \xi_0 + \xi_1\rangle_N| > \tau\big] = 1$ for all $N \geqslant N_0$ with $N_0 = 1$. In Figure 5.4, we show PFA and PMD of T-*Seq*RDT for $\alpha = \beta = 0.05$ and SNR $= 0.8$. As expected, as the mismatch grows, PFA grows and crosses the level α if p increases beyond 20%. The probabilities in the above cases depend on a multitude of parameters like, SNR, tolerances, levels α and β and, most importantly, on the underlying signal distribution, as shown in all the above cases. The above discussion shows the flexibility as well as the robustness of T-*Seq*RDT.

5.5 Summary

In this chapter, we highlighted the advantages of the algorithms, *Seq*RDT and T-*Seq*RDT, proposed in this dissertation compared to the optimal FSS algorithm, *Block*RDT, and SPRT as proposed in [39, 40]. We also compared the proposed algorithms to two popular composite hypothesis tests, GSPRT [36] and WSPRT as defined in [40]. We compared the algorithms for different types of signal models and showed that the proposed algorithms are not only robust to mismatches but also are capable of guaranteeing PFA and PMD performance, unlike GSPRT and WSPRT which fail to guarantee performance in non-asymptotic scenarios.

CHAPTER 6

CONCLUSION AND FUTURE DIRECTIONS

6.1 Summary

In this dissertation, we developed sequential algorithms for two-sided non-parametric hypothesis testing. We proposed a novel RDT based framework for sequential hypothesis testing and introduced two sequential algorithms to solve the binary hypothesis testing problem. We first proposed a non-truncated algorithm, *Seq*RDT, and analyzed its asymptotic performance. We then analyzed the properties of the thresholds and introduced the notion of a buffer which helped in controlling PFA and PMD of the algorithm. We finally, derived bounds on PFA and PMD and showed that *Seq*RDT can be designed to achieve arbitrarily low PFA and PMD. Finally, we introduced an additional parameter in the algorithm which we showed can be chosen such that the need for the buffer is eliminated.

We then introduced a truncated version of *Seq*RDT algorithm, T-*Seq*RDT. We designed the truncation window for T-*Seq*RDT using the optimal FSS test, *Block*RDT. We first analyzed the properties of the proposed thresholds and then derived bounds on PFA and PMD. Importantly, we showed that the designed thresholds guarantee pre-specified PFA and PMD. Moreover, we analyzed the average stopping time of T-*Seq*RDT and provided insights into the trade-off between the average stopping time and the error probabilities of T-*Seq*RDT. For both the algorithms

*Seq*RDT and T-*Seq*RDT, we proposed methods to choose the model parameters efficiently. Finally, we extended the proposed framework for testing of distorted signals and showed that the proposed algorithms are not only efficient for testing of distorted signals but also are faster compared to the optimal FSS test. We showed that the proposed algorithms are robust to mismatches compared to the likelihood ratio based approaches like SPRT, GSPRT and WSPRT.

Importantly, the proposed algorithms are simple in design and at the same time guarantee performance in the non-asymptotic regimes unlike the traditional composite (or non-parametric) likelihood ratio based schemes which generally only guarantee asymptotic performance. We used RDT based approaches to develop novel sequential algorithms which do not rely on the knowledge of the precise distributions of the underlying signals, and thereby, by design do not require the computations or even approximation of the likelihood ratios.

Next, we discuss some promising future directions of the work proposed in this dissertation.

6.2 Future Directions

In this section, we discuss some promising future directions of the work we presented in this dissertation that we intend to pursue.

6.2.1 Multi-Dimensional Signals

In this dissertation, we developed sequential algorithms for the case when the underlying signal as well as the observations were one-dimensional. However, in some hypothesis testing problems, the signal of interest as well as the observations can be multi-dimensional. We are currently in the process of extending the algorithms and frameworks presented in this work for the case of multi-dimensional signals and observations. In the multidimensional case, the testing problem can be stated as:

$$\exists N_0 \in \mathbb{N} \text{ such that } \forall N \geq N_0 \text{ we have } \begin{cases} \text{Under } \mathcal{H}_0 : \|\langle\, \Xi\rangle_N - \xi_0\|_\Sigma \leq \tau \;\; \text{(a.s)} \\ \\ \text{Under } \mathcal{H}_1 : \|\langle\, \Xi\rangle_N - \xi_0\|_\Sigma > \tau \;\; \text{(a.s).} \end{cases} \tag{6.1}$$

where, $\Xi = (\Xi_n)_{n\in\mathbb{N}}$ is assumed to lie in a d-dimensional space and $\|\cdot\|_\Sigma$ represents the Mahalanobis distance and is defined as: $\|\mathrm{x}\|_\Sigma = \sqrt{\mathrm{x}^T\Sigma^{-1}\mathrm{x}}$. Note, that the positive definite matrix, Σ, determines the ellipse the underlying signal mean, $\langle\, \Xi\rangle_N$, is assumed to converge to according to (6.1). Developing algorithms for the above case is a challenging task and will be addressed in detail in future.

6.2.2 Distributed Implementations

The algorithms proposed in the current work are presented for a centralized framework. However, in the current landscape of wireless sensor technologies, where a large number of sensor networks are used to solve inference problems. A natural and important extension of the work would be to develop the algorithms proposed in this dissertation for a distributed framework. Specifically, where the information is collected via multiple sensors distributed spatially in a region of interest. Moreover, since the algorithms proposed in this dissertation do not rely on the underlying signal distributions, their extension to distributed frameworks might help with multi-modal signal processing applications.

6.2.3 Optimality of the Proposed Tests

In this dissertation, we developed sequential algorithms for the non-parametric sequential testing framework proposed in Chapter 2. Specifically, we designed algorithms which belonged to the class $\mathscr{C}(\alpha,\beta)$, for pre-sepcified levels α and β of PFA and PMD, respectively. Importantly, we showed that the proposed algorithms does not only belong to class $\mathscr{C}(\alpha,\beta)$ but are also faster on average compared to the optimal FSS test. However, in all of the above analysis, we did not address one important question that is: *“Do the sequential tests proposed in this dissertation provide any optimality guarantees for the sequential testing framework proposed in Chapter 2?"* In future work, we want to address some questions regarding the optimality of the proposed tests.

APPENDIX A

APPENDIX: PROOFS OF VARIOUS RESULTS

A.1 Proof of Lemma A1

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

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

PROOF:

By property of a conditional and taking the independence of $\langle\, \Xi \rangle_N$ and $\langle\, X \rangle_N$ into account, we have:

$$\mathbb{P}\left[\, |\langle\, \Xi \rangle_N + \langle\, X \rangle_N - \xi_0| \leqslant \eta \,\right] = \int_0^\infty \mathbb{P}\left[\, |\rho + \langle\, X \rangle_N| \leqslant \eta \right] \mathbb{P}\, |\langle\, \Xi \rangle_N - \xi_0|^{-1}(\mathrm{d}\rho)$$

It follows from $X \sim \mathcal{N}(0, 1)$ that, for all $\rho \in [0, \infty)$:

$$\mathbb{P}\left[\, |\rho + \langle\, X \rangle_N| \leqslant \eta \,\right] = \Phi\left(\sqrt{N}(\eta - \rho)\right) - \Phi\left(-\sqrt{N}(\eta + \rho)\right)$$

The foregoing and (1.3) imply the result through the equality:

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

Hence the result. ■

A.2 Proof of Lemma A2

Lemma A.2. *If the signal, Ξ, satisfies Assumption **[(a-s) convergence of $\langle \Xi \rangle_N$]** given in Chapter 3, then for $\gamma \in (0,1)$ we have*

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

PROOF:

From Assumption **[(a-s) convergence of $\langle \Xi \rangle_N$]** given in Chapter 3 we have under $\mathcal{H}_0$:

$$\limsup_{N\to\infty} |\langle \Xi \rangle_N - \xi_0| \leq \tau^- \quad \text{(a.s.)}$$

This implies that

$$\mathbb{P}\{\omega \in \Omega : \limsup_{N\to\infty} |\langle \Xi \rangle_N - \xi_0| \leq \tau^-\} = 1.$$

which further implies that there exist $\Omega' \subset \Omega$ and $\Omega' \in \mathcal{F}$ such that for all $\omega \in \Omega'$ there exist a $N_0(\omega)$ such that $\mathbb{P}(\Omega') = 1$ and we have

$$|\langle \Xi \rangle_N - \xi_0| \leq \tau^- \text{ for all } N \geq N_0(\omega).$$

Therefore, for $N \geq N_0(\omega)$ from the fact that $Q_{\frac{1}{2}}(\bullet, \lambda_\gamma(\tau\sqrt{N}))$ is increasing in the first argument, we have

$$Q_{\frac{1}{2}}\left(\sqrt{N}|\langle \Xi \rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\right) \leq Q_{\frac{1}{2}}\left(\tau^-\sqrt{N}, \lambda_\gamma(\tau\sqrt{N})\right).$$

Now taking limits on both sides as $N \to \infty$ under $\mathcal{H}_0$, from Lemma A.4 we get

$$\lim_{N\to\infty} Q_{\frac{1}{2}}\left(\tau^-\sqrt{N}, \lambda_\gamma(\tau\sqrt{N})\right) = 0.$$

which implies the result.

Similarly, under $\mathcal{H}_1$ we have from Assumption **[(a-s) convergence of $\langle\,\Xi\rangle_N$]** given in Chapter 3 that:

$$\liminf_{N\to\infty} |\langle\,\Xi\rangle_N - \xi_0| \geq \tau^+ \quad \text{(a.s.)}$$

This implies that there exist $\Omega'' \subset \Omega$ and $\Omega'' \in \mathcal{F}$ such that for all $\omega \in \Omega''$ there exist a $N_0(\omega)$ such that $\mathbb{P}(\Omega'') = 1$ and we have

$$|\langle\,\Xi\rangle_N - \xi_0| \geq \tau^+ \;\text{ for all }\; N \geq N_0(\omega).$$

Therefore, for $N \geq N_0(\omega)$ and the fact that $Q_{\frac{1}{2}}(\bullet, \lambda_\gamma(\tau\sqrt{N}))$ is increasing in the first argument, we have

$$Q_{\frac{1}{2}}\left(\sqrt{N}|\langle\,\Xi\rangle_N - \xi_0|, \lambda_\gamma(\tau\sqrt{N})\right) \geq Q_{\frac{1}{2}}\left(\tau^+\sqrt{N}, \lambda_\gamma(\tau\sqrt{N})\right).$$

Now taking limits on both sides as $N \to \infty$ under $\mathcal{H}_1$, we get from Lemma A.4 that

$$\lim_{N\to\infty} Q_{\frac{1}{2}}\left(\tau^+\sqrt{N}, \lambda_\gamma(\tau\sqrt{N})\right) = 1.$$

which implies the result. Hence the proof ■

A.3 Proof of Lemma A3

Lemma A.3. *For any* $\gamma \in (0,1)$*:*

(i) We have

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

(ii) And

$$\lim_{\rho \to \infty} \frac{\lambda_\gamma(\rho)}{\rho} = 1.$$

PROOF:

We prove (i) only since it straightforwardly implies (ii). Pose $g_\gamma(\rho) = \lambda_\gamma(\rho) - \rho$ and $\theta = \Phi^{-1}(1-\gamma)$. Since $\Phi(x) + \Phi(-x) = 1$, (1.3) and the definition of $\lambda_\gamma(\tau)$ induce that:

$$\Phi(g_\gamma(\rho)) + \Phi(g_\gamma(\rho) + 2\rho) = 1 + \Phi(\theta). \tag{A.1}$$

To prove that $g_\gamma(\rho)$ tends to θ when $\rho \to \infty$, we proceed by contradiction. If $g_\gamma(\rho)$ 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 $g_\gamma(\rho_n) > \theta + \varepsilon$ or $g_\gamma(\rho_n) < \theta - \varepsilon$. 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 $n \geqslant N_0$:

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

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

$$\Phi(2\rho_n + \theta - \varepsilon) < 1 + \eta. \tag{A.3}$$

Let n be any integer above $\max(N_0, N_1)$. If $g_\gamma(\rho_n) < \theta - \varepsilon$, we then have $\Phi(g_\gamma(\rho_n)) < \Phi(\theta - \varepsilon)$

and $\Phi(2\rho_n + g_\gamma(\rho_n)) < \Phi(2\rho_n + \theta - \varepsilon)$. Eqs. (A.1) and (A.3) 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 $g_\gamma(\rho_n) < \theta - \varepsilon$. We cannot have $g_\gamma(\rho_n) > \theta + \varepsilon$ either because, via (A.1) and (A.2), this inequality implies:

$$1 + \Phi(\theta) > \Phi(\theta + \varepsilon) + \Phi(2\rho_n + \theta + \varepsilon) > \Phi(\theta + \varepsilon) + 1 - \eta, \tag{A.4}$$

which is contradictory to our choice for η. ■

A.4 Proof of Lemma A4

Lemma A.4 (Behavior of $Q_{\frac{1}{2}}$ in vanishing noise)**.** *Consider* $\tau \in [0, \infty)$ *and* $\rho \in (0, \infty)$ *such that* $\rho \neq \tau$.

$$\textit{for all } \gamma \in (0,1), \lim_{\sigma \to 0} Q_{\frac{1}{2}}\left(\frac{\rho}{\sigma}, \lambda_\gamma(\tau/\sigma)\right) = \mathbb{1}_{(\tau,\infty)}(\rho).$$

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 (1.2),

$$Q_{\frac{1}{2}}\big(\frac{\rho}{\tau}\rho_n, \lambda_\gamma(\rho_n)\big) = \mathbb{P}\left[|\frac{\rho}{\tau} + \frac{X}{\rho_n}| > \frac{\lambda_\gamma(\rho_n)}{\rho_n}\right]$$

for any $X \sim \mathcal{N}(0,1)$. It follows from Lemma A.3 (ii) that

$$|\frac{\rho}{\tau} + \frac{X}{\rho_n}| - \frac{\lambda_\gamma(\rho_n)}{\rho_n} = \frac{\rho}{\tau} - 1 \text{ a-s.}$$

Therefore, the cdf of $\big|(\rho/\tau) + (X/\rho_n)\big| - \lambda_\gamma(\rho_n)/\tau_n$ converges weakly to $\mathbb{1}_{[(\rho/\tau)-1,\infty)}$. Since $\rho \neq \tau$, this weak convergence implies that

$$\lim_{n\to\infty} \mathbb{P}\left[|\frac{\rho}{\tau} + \frac{X}{\rho_n}| > \frac{\lambda_\gamma(\rho_n)}{\rho_n}\right] = \mathbb{1}_{(\tau,\infty)}(\rho).$$

Hence the result since $(\sigma_n)_{n\in\mathbb{N}}$ is arbitrary. ■

A.5 Proof of Lemma A5

Lemma A.5 (Non-Asymptotic behavior of $Q_{\frac{1}{2}}$)**.** *Consider $\tau \in [0, \infty)$, $\rho \in (0, \infty)$ and $\gamma \in (0, 1)$, the map:*

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

PROOF:

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

$$\mathcal{Q}(\sigma) = Q_{\frac{1}{2}}\big(\rho\sigma, \lambda_\gamma(\tau\sigma)\big) = 1 - \Phi(r_-(\sigma)) + \Phi(-r_+(\sigma)) \tag{A.5}$$

with $r_+ = \lambda_\gamma(\tau\sigma) + \rho\sigma$ and $r_- = \lambda_\gamma(\tau\sigma) - \rho\sigma$. For $\rho = \tau$, it follows from (1.4) that $\mathcal{Q}$ is constant equal to γ. We thus have $1 - \Phi(\lambda_\gamma(\tau\sigma) - \rho\sigma) + \Phi(-\lambda_\gamma(\tau\sigma) - \rho\sigma) = \gamma$. After differentiating the two members of the equality above and after some routine algebra, we obtain:

$$\lambda'_\gamma(\tau\sigma) = \frac{1 - e^{-2\tau\sigma\lambda_\gamma(\tau\sigma)}}{1 + e^{-2\tau\sigma\lambda_\gamma(\tau\sigma)}} \tag{A.6}$$

where λ'_γ is the first derivative of λ_γ. We now differentiate $\mathcal{Q}$ defined by (A.5). 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 (A.6) into the equality above, we obtain:

$$\mathcal{Q}'(\sigma) = \frac{\tau}{\sqrt{2\pi}}\left(e^{-r_-^2(\sigma)/2} - e^{-r_+^2(\sigma)/2}\right)\left(\frac{\rho}{\tau} - \frac{\Delta_-(\rho,\tau)}{\Delta_+(\rho,\tau)}\right) \tag{A.7}$$

with $\Delta_\varepsilon(\rho,\tau) = \frac{1+\varepsilon e^{-2\tau\sigma\lambda_\gamma(\tau\sigma)}}{1+\varepsilon e^{-2\rho\sigma\lambda_\gamma(\tau\sigma)}}$ 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 $\rho < \tau$,

$$\frac{\rho}{\tau} < 1 < \frac{\Delta_-(\rho,\tau)}{\Delta_+(\rho,\tau)},$$

which implies that $\mathcal{Q}'(\sigma) \leqslant 0$ and, thus, that $\mathcal{Q}$ is decreasing. On the other hand, if $\rho > \tau$, we have

$$\frac{\rho}{\tau} > 1 > \frac{\Delta_-(\rho,\tau)}{\Delta_+(\rho,\tau)},$$

so that $\mathcal{Q}$ is increasing in this case. ∎

A.6 Proof of Lemma A6

Lemma A.6. *Given $\rho \in (0, \infty)$, the map $\gamma \in (0, 1) \mapsto \lambda_\gamma(\rho)$ is decreasing.*

PROOF:

The Lemma follows from the definition of $\lambda_\gamma(\rho)$ given in (1.4) (please see Chapter 1) and the decreasing nature of $Q_{\frac{1}{2}}$ with its second argument given in Lemma 1.1 (please see Chapter 1) . ■

A.7 Proof of Lemma A7

Lemma A.7.

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

(P2) *The map* $\rho \in (0, \infty) \mapsto Q_{\frac{1}{2}}\big(\rho, \rho\big)$ *is decreasing, lower-bounded by* $1/2$ *and*

$$\lim_{\rho \to \infty} Q_{\frac{1}{2}}\big(\rho, \rho\big) = \frac{1}{2}.$$

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

PROOF:

Proof of statement (P1)*:* Using (1.3), define $\mathcal{Q}(\sigma)$ as:

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

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

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

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

Proof of statement (P2)*:* The map $\rho \in (0, \infty) \mapsto Q_{\frac{1}{2}}\big(\rho, \rho\big)$ is decreasing as a consequence of (P1). Given $\rho \in (0, \infty)$,

$$Q_{\frac{1}{2}}(\rho, \rho) = \frac{1}{2} + \Phi(-2\rho)$$

from (1.3). Hence the result.

Proof of statement (P3)*:* Let ρ and ρ' be two positive real numbers such that $0 < \rho < \rho'$. According to (1.4), we have:

$$Q_{\frac{1}{2}}\big(\rho, \lambda_\gamma(\rho)\big) = Q_{\frac{1}{2}}\big(\rho', \lambda_\gamma(\rho')\big) = \gamma. \tag{A.8}$$

Since $\gamma < 1/2$ so that $1/2 < 1 - \gamma$, it follows from (P2) and (A.8) that:

$$Q_{\frac{1}{2}}(\rho, \rho) > 1/2 > Q_{\frac{1}{2}}(\rho, \lambda_\gamma(\rho)).$$

The decreasing behavior of $Q_{\frac{1}{2}}$ 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 $x \in (0, \infty) \mapsto Q_{\frac{1}{2}}\big(\rho/x, \lambda_\gamma(\rho)/x\big)$ is an increasing map. Since $\rho/\rho' < 1$, we thus have

$$Q_{\frac{1}{2}}\big(\rho, \lambda_\gamma(\rho)\big) > Q_{\frac{1}{2}}\left(\rho', \rho' \frac{\lambda_\gamma(\rho)}{\rho}\right).$$

This inequality and (A.8) induce that

$$Q_{\frac{1}{2}}\big(\rho', \lambda_\gamma(\rho')\big) > Q_{\frac{1}{2}}\left(\rho', \rho' \frac{\lambda_\gamma(\rho)}{\rho}\right).$$

The decreasing nature of $Q_{\frac{1}{2}}(\rho', \cdot)$ then implies that

$$\lambda_\gamma(\rho') < \rho' \frac{\lambda_\gamma(\rho)}{\rho}.$$

Thereby, $\rho \in (0, \infty) \mapsto \lambda_\gamma(\rho)/\rho$ is decreasing in ρ. Since the map $\rho \in (0, \infty) \mapsto \lambda_\gamma(\rho)/\rho$ is decreasing and lower bounded by 1, this map has a limit $\ell \geqslant 1$ when ρ tends to ∞. The result then follows as a consequence of Lemma A.3 (ii). ■

A.8 Proof of Lemma A8

Lemma A.8. *For $\gamma \in (1/2, 1)$ and ρ large enough, the map $\rho \in (0, \infty) \mapsto \lambda_\gamma(\rho)/\rho$ is increasing, upper bounded by* 1.

PROOF:

According to statement (i) of Lemma A.3,

$$\lambda_\gamma(\rho) - \rho = \Phi^{-1}(1-\gamma) + \kappa(\rho),$$

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.

With $h_\gamma(\rho) = \lambda_\gamma(\rho)/\rho$,

$$\Phi(\rho(h_\gamma(\rho) - 1)) - \Phi(-\rho(h_\gamma(\rho) + 1)) = 1 - \gamma.$$

By differentiation of this equality with respect to ρ and since h_γ is differentiable via the implicit function theorem, we find that $h'_\gamma(\rho)$ has the same sign as

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

For ρ large enough, $h_\gamma(\rho) < 1$ by the first part of the proof and Lemma A.3 implies that

$$\lim_{\rho\to\infty} \Upsilon(\rho) = \infty.$$

Therefore, $\Upsilon(\rho) > 0$ for ρ large enough and the proof is complete. ■

A.9 Convergence of the upper bounds in Theorem 3.3

PROOF:

Let us consider $\text{UB2}_{\text{FA}}(M)$ given in (3.21), we have:

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

Each term of the above summation is positive and $t > 0 \mapsto Q_{\frac{1}{2}}\left(\tau^{-}\sqrt{t}, \lambda_{\alpha}(\tau\sqrt{t})\right)$ is decreasing since $\tau^{-} < \tau$, as a consequence of Lemma A.5. Therefore, to show that the above series converges, it suffices to show that

$$I_{\text{FA}}(\zeta) = \int_{\zeta}^{\infty} Q_{\frac{1}{2}}\left(\tau^{-}\sqrt{t}, \lambda_{\alpha}(\tau\sqrt{t})\right) \mathrm{d}t, \tag{A.9}$$

is finite for some $\zeta > 0$, with $\alpha \in (0, 1/2)$ and $\tau^{-} < \tau$. From the definition of $Q_{\frac{1}{2}}(a, b)$ given in (1.3), we have:

$$Q_{\frac{1}{2}}(a, b) = \Phi^{c}(b - a) + \Phi^{c}(b + a), \tag{A.10}$$

where $\Phi^{c}(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \mathrm{e}^{\frac{-z^2}{2}} \mathrm{d}z$, also known as the standard Q function, is the complementary cdf of a standard normal distributed random variable. Using (A.10) in (A.9) 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

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

where (a) results from $\Phi^{c}(x) \leqslant \mathrm{e}^{-x^2/2}$ for $x > 0$ as, for $\alpha < 1/2$, $\lambda_{\alpha}(\tau\sqrt{t}) - \tau^{-}\sqrt{t} > 0$ according to Lemma A.7 (P3). Therefore, to show that $I_{\text{FA}}(\zeta) < \infty$, it suffices to show

$$J_{\text{FA}}(\zeta) \stackrel{\text{def}}{=} \int_{\zeta}^{\infty} \mathrm{e}^{\frac{-u(t)^2}{2}} \mathrm{d}t < \infty, \tag{A.11}$$

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

To prove (A.11), 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:

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

(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: [Statements (C1), (C2), (C3), and (C4)]

(C1) Lemma A.7 (P3) implies

$$\lambda_\alpha(\tau\sqrt{t}) \geqslant \tau\sqrt{t}.$$

Hence **(C1)** holds.

(C2) This property follows from (A.6) in Lemma A.5.

(C3) Lemma A.7 (P3) implies $G_\alpha(t) \leqslant \mathrm{e}^{-2\tau^2 t}$. Thus,

$$\lim_{t\to\infty}\left(\tau\frac{1-G_\alpha(t)}{1+G_\alpha(t)}-\tau^-\right)=\tau-\tau^->0.$$

Hence 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

$$\lim_{x\to\infty}\left(\tau\frac{1-G_\alpha(u^{-1}(x))}{1+G_\alpha(u^{-1}(x))}-\tau^-\right)=\tau-\tau^-.$$

Which implies **(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 [10, Theorem 12.6], among others) and the fact that $u'(u^{-1}(x)) > 0$ for any $x > u(\zeta)$ since $\zeta > A$, we have:

$$J_{\mathrm{FA}}(\zeta)=\int_{u(\zeta)}^{\infty}\frac{1}{u'(u^{-1}(x))}e^{-x^2/2}\mathrm{d}x. \tag{A.12}$$

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

$$u'(u^{-1}(x))=\frac{1}{2\sqrt{u^{-1}(x)}}\left[\tau\frac{1-G_\alpha(u^{-1}(x))}{1+G_\alpha(u^{-1}(x))}-\tau^-\right]. \tag{A.13}$$

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

$$\frac{1}{\sqrt{u^{-1}(x)}}\geqslant\frac{\tau-\tau^-}{x}. \tag{A.14}$$

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

$$\tau\frac{1-G_\alpha(u^{-1}(x))}{1+G_\alpha(u^{-1}(x))}-\tau^->a. \tag{A.15}$$

By injecting (A.14) and (A.15) into (A.13), we obtain $u'(u^{-1}(x)) \geqslant a(\tau-\tau^-)/2x$ for all $x \geqslant$

$u(\zeta)$. Therefore, it results from (A.12) that

$$J_{\text{FA}}(\zeta) \leqslant \frac{2}{a(\tau - \tau^{-})} \int_{0}^{\infty} x e^{-x^2/2} \mathrm{d}x < \infty$$

and that (A.11) holds.

The convergence of $\text{UB2}_{\text{MD}}(M)$ in (3.23) is proved similarly via Lemma A.8 and (A.6) in Lemma A.5. ■

REFERENCES

[1] R. Bechhofer, "A note on the limiting relative efficiency of the wald sequential probability ratio test," *Journal of the American Statistical Association*, vol. 55, no. 292, pp. 660–663, 1960. [Online]. Available: http://www.jstor.org/stable/2281589 53

[2] P. Billingsley, *Probability and Measure, 3rd edition*. Wiley, 1995. 28

[3] L. Blanch, F. Bernabé, and U. Lucangelo, "Measurement of air trapping, intrinsic positive end-expiratory pressure, and dynamic hyperinflation in mechanically ventilated patients," *Respiratory Care*, vol. 50, no. 1, pp. 110–124, 2005. [Online]. Available: http://rc.rcjournal.com/content/50/1/110 17

[4] E. Brodsky and B. Darkhovsky, *Nonparametric Methods in Change Point Problems*, ser. Mathematics and Its Applications. Springer Netherlands, 1993. 3

[5] G. Casella and R. Berger, *Statistical Inference*. Duxbury Resource Center, June 2001. 1, 4

[6] R. Durrett, *Probability: Theory and Examples*, 4th ed. New York, NY, USA: Cambridge University Press, 2010. 29, 32, 38, 39

[7] G. Fellouris and A. G. Tartakovsky, "Almost optimal sequential tests of discrete composite hypotheses," *Statistica Sinica*, vol. 23, no. 4, pp. 1717–1741, 2013. 5, 23, 91

[8] T. S. Ferguson, "Who solved the secretary problem?" *Statist. Sci.*, vol. 4, no. 3, pp. 282–289, 08 1989. [Online]. Available: https://doi.org/10.1214/ss/1177012493 3

[9] B. K. Ghosh and P. K. Sen, *Handbook of Sequential Analysis*, ser. Statistics: A Series of Textbooks and Monographs. Taylor & Francis, 1991. 4, 5, 24, 80, 84

[10] J. Jacod and P. Protter, *Probability Essentials, 2nd edition*. Springer, 2004. 115

[11] T. L. Lai, "Asymptotic optimality of invariant sequential probability ratio tests," *Ann. Statist.*, vol. 9, no. 2, pp. 318–333, 03 1981. 4

[12] ——, "Nearly optimal sequential tests of composite hypotheses," *Ann. Statist.*, vol. 16, no. 2, pp. 856–886, 06 1988. 4, 5, 22, 23, 55

[13] ——, "Sequential analysis: Some classical problems and new challenges," *Statistica Sinica*, vol. 11, no. 2, pp. 303–351, 2001. 4, 21

[14] ——, "Likelihood ratio identities and their applications to sequential analysis," *Sequential Analysis*, vol. 23, no. 4, pp. 467–497, 2004. 4, 5, 22, 23

[15] T. L. Lai and L. Zhang, "A modification of schwarz's sequential likelihood ratio tests in multivariate sequential analysis," *Sequential Analysis*, vol. 13, no. 2, pp. 79–96, 1994. 4, 5, 22, 23

[16] T. L. Lai, "Sequential analysis: Some classical problems and new challenges," *Statistica Sinica*, vol. 11, no. 2, pp. 303–351, 2001. 4

[17] K. Lant. (2017) By 2020, there will be 4 devices for every human on earth. [Online]. Available: https://futurism.com/by-2020-there-will-be-4-devices-for-every-human-on-earth 1

[18] E. L. Lehmann and J. P. Romano, *Testing statistical hypotheses*, 3rd ed. Springer, 2005. 4, 21

[19] ——, *Testing Statistical Hypotheses, 3rd edition*. Springer, 2005. 3, 4

[20] X. Li, J. Liu, and Z. Ying, "Generalized sequential probability ratio test for separate families of hypotheses," *Sequential Analysis*, vol. 33, no. 4, pp. 539–563, 2014. 4, 5, 22

[21] J. Neyman and E. S. Pearson, “On the problem of the most efficient tests of statistical hypotheses,” *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 231, pp. 289–337, 1933. [Online]. Available: http://www.jstor.org/stable/91247 29, 54

[22] Q.-T. Nguyen, D. Pastor, and E. L’Her, “Automatic detection of autopeep during controlled mechanical ventilation,” *BioMedical Engineering OnLine*, vol. 11, no. 1, p. 32, 2012. 16, 17

[23] 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. 10

[24] 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. 6, 7, 11, 15, 29, 74, 80

[25] ——, “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. 7, 16, 29, 56, 57, 58, 59

[26] D. Pastor and F.-X. Socheleau, “Random distortion testing with linear measurements,” *Signal Processing*, vol. 145, pp. 116 – 126, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168417304127 56, 57, 58

[27] H. V. Poor, *An introduction to signal detection and estimation*, 2nd ed. Springer-Verlag, 1994. 3, 4, 29, 54

[28] G. Samorodnitsky and M. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, ser. Stochastic Modeling Series. Taylor & Francis, 1994. [Online]. Available: https://books.google.com/books?id=wTTUfYwjksAC 81, 82

[29] P. K. Sen, *Sequential Nonparametrics: Invariance Principles and Statistical Inference*, ser. Wiley Series in Probability and Statistics - Applied Probability and Statistics Section. Wiley, 1981. 4, 5, 21, 24

[30] P. K. Sen, C. B. of the Mathematical Sciences, S. for Industrial, and A. Mathematics, *Theory and Applications of Sequential Nonparametrics*, ser. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1985. 5, 24

[31] P. K. Sen, "Nonparametric methods in sequential analysis," 1989. 24

[32] 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, 11

[33] S. Tantaratana and J. B. Thomas, "Truncated sequential probability ratio test," *Information Sciences*, vol. 13, no. 3, pp. 283 – 300, 1977. 53, 70

[34] ——, "Relative efficiency of the sequential probability ratio test in signal detection," *IEEE Transactions on Information Theory*, vol. 24, no. 1, pp. 22–31, January 1978. 53

[35] A. G. Tartakovsky, "Nearly optimal sequential tests of composite hypotheses revisited," *Proceedings of the Steklov Institute of Mathematics*, vol. 287, no. 1, pp. 268–288, Dec 2014. 4, 22

[36] A. G. Tartakovsky, M. Basseville, and I. Nikiforov, *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. Chapman & Hall/CRC, 2014. 4, 5, 6, 18, 19, 21, 22, 23, 24, 55, 78, 80, 85, 95

[37] A. W. Thille, P. Rodriguez, B. Cabello, F. Lellouche, and L. Brochard, "Patient-ventilator asynchrony during assisted mechanical ventilation," *Intensive Care Medicine*, vol. 32, no. 10, pp. 1515–1522, Oct 2006. [Online]. Available: https://doi.org/10.1007/s00134-006-0301-8 17

[38] L. Vignaux, F. Vargas, J. Roeseler, D. Tassaux, A. W. Thille, M. P. Kossowsky, L. Brochard, and P. Jolliet, “Patient–ventilator asynchrony during non-invasive ventilation for acute respiratory failure: a multicenter study,” *Intensive Care Medicine*, vol. 35, no. 5, p. 840, Jan 2009. [Online]. Available: https://doi.org/10.1007/s00134-009-1416-5 17

[39] A. Wald, “Sequential tests of statistical hypotheses,” *The Annals of Mathematical Statistics*, vol. 16, no. 2, pp. 117–186, 06 1945. 3, 53, 78, 95

[40] ——, *Sequential Analysis*. John Wiley and Sons, New York, 1948. 3, 4, 21, 24, 53, 78, 85, 95

[41] D. J. Wheeler and D. Chambers, *Understanding Statistical Process Control*. SPC Press, 2010. [Online]. Available: https://books.google.com/books?id=8JlHSAAACAAJ 3

[42] J. Whitehead, *The Design and Analysis of Sequential Clinical Trials*, ser. Statistics in Practice. Wiley, 1997. [Online]. Available: https://books.google.com/books?id=y0s1ACQTUskC 3

[43] 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. 6

VITA

NAME OF AUTHOR: Prashant Khanduri

MAJOR: Electrical and Computer Engineering

EDUCATION:

M.E. 2011 PEC University of Technology, Chandigarh, India

B.E. 2009 Kumaun University, Nainital, India

AWARDS AND HONORS:

Best student paper award, IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Cannes, France 2019

Student travel award, IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)

NSF travel grant, GlobalSIP, 2016