QUICKEST CHANGE DETECTION WITH APPLICATIONS TO LINE OUTAGE DETECTION

BY

GEORGIOS ROVATSOS

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Adviser:

Professor Venugopal V. Veeravalli

ABSTRACT

In this work, we focus on applications of quickest change detection (QCD) theory in the problem of line outage detection and identification. We start by discussing fundamental results of sequential hypothesis testing and QCD, and by proposing an algorithm for the QCD setting under transient dynamics. Following, we apply these results in the line outage detection problem. QCD algorithms are applied on measurements of voltage phase angles, which are collected using phasor measurement units (PMUs), sampling units that sample at an approximate rate of 30 samples per second and that are placed in the buses of the system. The goal is to detect a line outage as fast as possible, under false alarm constraints. First, we study the line outage setting where no transient dynamics are present. Then, we propose a QCD algorithm for the case where transient dynamics are present. Line outage identification schemes are also discussed.

To my family

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Professor Venugopal V. Veeravalli for his support throughout this research process. Without his counsel and guidance this work would not have been possible.

Lastly, I would like to express the thanks I owe to my family for their patience and support throughout my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The motivation behind this work stems from the increasing presence of phasor measurement units (PMUs) across the power grid. The introduction of this new measurement unit has led to significant advances in the stability monitoring and state estimation capabilities for the power system. As a result, the integration of the PMUs has paved the way for the use of real-time algorithms that can be exploited to detect line outages in an efficient and robust manner. Furthermore, the statistical behavior of the observed measurement process is another motivation. In particular, a line outage event leads to a change in the statistical behavior of the observed sequence of observations. The detection of changes of this nature is a problem that is studied thoroughly in the QCD literature.

1.1 Background

Timely detection of line outages in a power system is crucial for maintaining operational reliability. In this regard, many online decision-making tools rely on a system model that is obtained offline, which can be inaccurate due to bad historical or telemetry data. Such inaccuracies have been a contributing factor in many recent blackouts. For example, in the 2011 San Diego blackout, operators were unable to determine overloaded lines because the network model was not up to date [1]. This lack of situational awareness limited the ability of the operators to identify and prevent the next critical contingency, and led to a cascading failure. Similarly, during the 2003 US Northeastern blackout, operators failed to initiate the correct remedial schemes because they had an inaccurate model of the power system and could not identify the loss of key transmission elements [2]. These blackouts highlight the importance of developing online measurement-based techniques to

detect and identify system topological changes that arise from line outages. In this work, we tackle such topology change detection problems by utilizing measurements provided by PMUs.

1.2 Problem Statement

Our work extends the results of [3], where the authors developed a method for line outage detection and identification based on the theory of quickest change detection (QCD) [4], [5], [6]. In this method, the incremental changes in real power injections are modeled as independent zero-mean Gaussian random variables. Then, the probability distribution of such incremental changes is mapped to that of the incremental changes in voltage phase angles via a linear transformation obtained from the power flow balance equations. The PMUs provide a random sequence of voltage phase angle measurements in real-time; when a line outage occurs, the probability distribution of the incremental changes in the voltage phase angles changes abruptly. The objective is to detect a change in this probability distribution after the occurrence of a line outage as quickly as possible while maintaining a desired false alarm rate. In this work, we focus on the problem of line outage detection and identification in two settings. First, we assume that after the outage the system changes state instantaneously, and later we study a more realistic setting, where the system model is characterized by transient behavior.

1.3 Related Work

Early approaches for topological change detection include algorithms based on state estimation [7], [8], and rule-based algorithms that mimic system operator decisions [9]. More recent methods exploit the fast sampling of voltage magnitudes and phases provided by PMUs [10]–[12]. However, these schemes do not exploit the persistent nature of line outages and do not incorporate transient behavior. Only the most recent PMU measurement is used to determine if an outage has occurred. The authors of [13] proposed a method to detect line outages using statistical classifiers where a maximum likelihood estimation is performed on the PMU data. The authors also considered the

transient response of the system after a line outage by comparing synthesized data against actual data. However, their method requires knowledge of the exact time of the line outage before applying the algorithm, whereas our proposed methods do not have this restriction.

1.4 Contribution of Thesis

We first study the line outage problem when no transient dynamics are present, i.e., the shift from pre- to post- change distribution happens almost instantly. As in [3], the algorithm that is proposed as a solution is based on adapting the Generalized Cumulative Sum (G-CuSum) test from the QCD literature (see, e.g., [4], [6]) to the line outage detection problem. Our algorithm not only takes the persistent covariance change into consideration, but it also exploits past observations to detect the occurrence of an outage. In [3], the statistics for each individual line are compared to a common predetermined threshold, and an outage is declared if one of these statistics crosses the threshold. In this work, we present a method for setting a different threshold for each line outage statistic by taking the dissimilarity between the pre- and post-change distribution into consideration. This difference between pre- and post-change distributions is described by the Kullback-Leibler (KL) divergence, a metric that quantifies the distance between two distributions. In addition, we compare the performance of our test to that of the Shewhart test, the meanshift test, and the algorithm of [3], and observe notable improvements in terms of performance.

Next, we study the line outage problem under the presence of an arbitrary number of transient periods. We improve on the method proposed in [3] by considering the power system transient response immediately following the line outage. For example, after an outage, the transient behavior of the system is dominated by the inertial response from the generators. This is followed by the governor response and then the automatic generation control (AGC). We incorporate these dynamics into the power system model by relating incremental changes in active power demand to active power generation. We use this model to develop the Dynamic CuSum test (D-CuSum), which is used to capture the transient behavior in the non-composite QCD problem (see e.g., [4], [6]). Then, the Generalized Dynamic CuSum (G-D-CuSum) test

is derived by calculating a D-CuSum statistic for each possible line outage scenario; an outage is declared the first time any of the test statistics crosses a pre-specified threshold. The proposed test has better performance because it considers the transient behavior in addition to the persistent change in the distribution that results from the outage. Furthermore, we discuss line outage identification techniques that can be employed easily in practice.

The remainder of this thesis is organized as follows. In Chapter 2, we review the problem of binary sequential hypothesis testing. In Chapter 3, we study the problem of quickest change detection (QCD) and provide theoretical results that are going to be used throughout this work. In Chapter 4, we formulate the QCD problem under transient dynamics and propose the D-CuSum test. In Chapter, 5 we study the problem of line outage detection and identification when no transient dynamics are present, and propose a G-CuSum based algorithm that uses varying thresholds to achieve performance gains. In Chapter, 6 we introduce the G-D-CuSum test as a proposed algorithm for detecting outages under transient behavior. We also demonstrate methods for identifying outaged lines. In Chapter, 7 we provide simulation results on the IEEE 14-bus and 118-bus systems. Finally, concluding remarks are made in Chapter 8.

CHAPTER 2

BINARY SEQUENTIAL HYPOTHESIS TESTING

In this chapter, we provide a brief review of the problem of binary sequential hypothesis testing, introduced in [14] (also see [15]). In this setting, measurements following one of two candidate distributions are fed to a decision maker sequentially. The goal is to design stopping procedures that dictate when to stop sampling, and use the data available up to the time of stop to decide in favor of one of the two hypotheses. Note that, in contrast to the traditional detection theory techniques where the sample size is fixed beforehand, in sequential hypothesis testing the sample size is determined online, by managing a tradeoff between the number of samples and the desired level of accuracy.

2.1 Problem Statement

In the present setting, the measurements come from a stream of observations that is characterized by one of two potential statistical behaviors (in this work we will focus on the binary testing problem). As a result, we have two hypotheses on the sampled data:

$$H_0 : X_k \sim f_0 \text{ i.i.d.}$$
$$H_1 : X_k \sim f_1 \text{ i.i.d.}$$

The goal here is to design a sequential test (τ, δ), which is essentially a stopping time τ accompanied by a decision rule δ. The notion of stopping time is defined formally as follows:

Definition 1. A stopping time τ adapted on a random process $\{X_k\}_{k=1}^{\infty}$ is a random variable with the property that $\{\tau = k\} \subset \sigma(X_1, \ldots, X_k)$, where $\sigma(X_1, \ldots, X_k)$ the σ-algebra generated by $X_1, \ldots, X_k$. Intuitively, this means

that no observations after time instant k are necessary to decide whether to stop sampling at k.

The stopping time τ should be designed so that the observations $\{X_k\}_{k=1}^{\tau}$ are sufficient to make a correct decision with respect to given accuracy. Note that the present problem is characterized by a pair of conflicting goals: on one hand we need to make an accurate decision regarding the statistical behavior of the process, which may require a large number of samples, on the other hand, a larger sample size is more costly. Thus, we have to manage a tradeoff which will depend on our budget and the desired level of accuracy.

2.2 Bayesian and non-Bayesian Formulation

In [14], two formulations are proposed to capture the tradeoff between sample size and detection accuracy. For the first one, we follow a Bayesian approach, meaning we assume known priors π_0, π_1, on the hypotheses. The second is based on a Neyman-Pearson (NP) approach. The solution to both instances is the Sequential Probability Ratio Test (SPRT) proposed in [14]. We start by defining the Bayes risk for the first formulation:

Definition 2. For a sequential test (τ, δ), the Bayes risk is given by

$$r(\tau, \delta) = c[\pi_0 \mathbb{E}_0[\tau] + \pi_1 \mathbb{E}_1[\tau]] + \pi_0 P_F(\tau, \delta) + \pi_1 P_M(\tau, \delta), \tag{2.1}$$

where $\mathbb{E}_0$ is the expectation under distribution f_0, $\mathbb{E}_1$ is the expectation under distribution f_1, c is the cost that we suffer for each additional sample measured, P_F is the probability of false alarm, and P_M is the probability of misdetection. A Bayesian sequential test is the sequential test $(\tau, \delta)_B$ that minimizes the Bayes risk, i.e.

$$(\tau, \delta)_B = \arg\min_{(\tau,\delta)} r(\tau, \delta). \tag{2.2}$$

When no priors on the hypotheses are available, we can approach the sequential hypothesis testing problem through an NP approach. In particular,

such a formulation is given as follows:

$$\begin{aligned} \min \quad & \mathbb{E}_0[\tau] \text{ and } \mathbb{E}_1[\tau] \\ \text{subject to} \quad & P_F(\tau, \delta) \leq \alpha \\ & P_M(\tau, \delta) \leq \beta, \end{aligned} \tag{2.3}$$

where $\alpha, \beta \in (0, 1)$. In practical terms, the goal in this formulation is to find the sequential test that minimizes $\mathbb{E}_0[\tau]$ and $\mathbb{E}_1[\tau]$ among all tests that have sufficiently small P_F and P_M.

2.3 The Sequential Probability Ratio Test (SPRT)

Here, we present the Sequential Probability Ratio Test (SPRT), proposed in [14], which is the solution to the two tradeoff formulations presented in Sec. 2.2. To define the SPRT, we use the log-likelihood ratio of $X_1, \ldots, X_k$ as a test statistic. Define the test statistic at time k as:

$$S_k = \sum_{j=1}^{k} \log \frac{f_1(X_j)}{f_0(X_j)}. \tag{2.4}$$

The corresponding stopping and decision rule for the SPRT is given by:

$$\tau^{SPRT} = \inf\{k \geq 1 : S_k \notin (a, b)\} \tag{2.5}$$

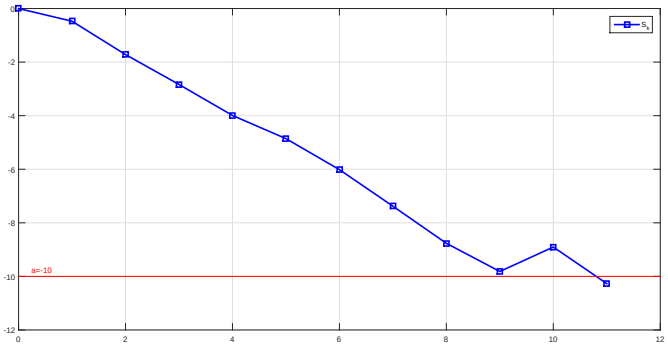
and

$$\delta^{SPRT} = \begin{cases} 1 & \text{if } S_{\tau^{SPRT}} > b \\ 0 & \text{if } S_{\tau^{SPRT}} < a, \end{cases} \tag{2.6}$$

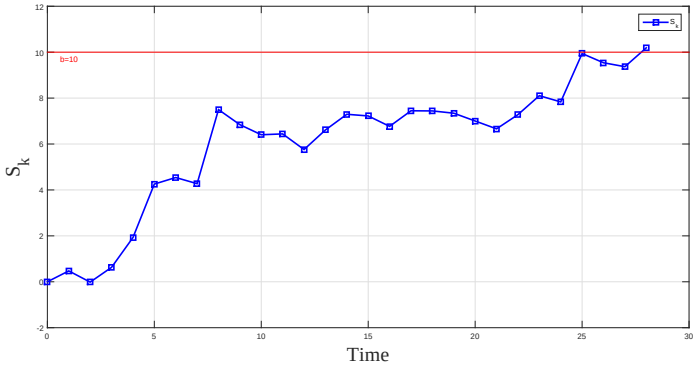
where $a < 0 < b$ are the test thresholds.

The SPRT test involves calculating a test statistic S_k at each time instant, with $S_0 \coloneqq 0$. The SPRT statistic is then compared to thresholds a and b. If $S_k > b$, we stop and decide in favor of H_1. If $S_k < a$, we stop and decide in favor of H_0. Otherwise, we continue sampling until we have accumulated enough data for an accurate decision.

In [14] it was shown that the SPRT is optimal for both the Bayesian and the non-Bayesian tradeoff formulations that were presented in Sec. 2.2. There, it is shown that the optimal test with respect to (2.2) is derived by selecting a



(a) When f_0 is the true distribution.



(b) When f_1 is the true distribution.

Figure 2.1: Typical evolution of the SPRT statistic with $f_0 = \mathcal{N}(0,1)$ and $f_1 = \mathcal{N}(0,\sqrt{2})$.

and b, so that the Bayes risk is minimized. For (2.3) it is shown that selecting thresholds that satisfy the inequality constraints with equality results in an optimal scheme.

In Fig. 2.1 we show the evolution of the SPRT statistic for two different hypotheses on the data, namely, H_0 corresponding to $f_0 = \mathcal{N}(0,1)$ and H_1 corresponding to $f_1 = \mathcal{N}(0,\sqrt{2})$. In particular, in Fig. 2.1(a) we see a sample path for the case that the data is generated by the distribution $f_0 = \mathcal{N}(0,1)$. Note how the test statistic decreases until threshold $a = -10$ is crossed and decision in favor of H_0 is taken. Similarly, when the data are generated by f_1, the test statistic grows until it crosses threshold $b = 10$ and we take a decision in favor of H_1.

For an intuitive interpretation of the detection algorithms discussed in this work, we will be using the Kullback-Leibler (KL) divergence, which is

an information theoretic measure of the discrepancy between two probability distributions.

Definition 3. The KL divergence between two probability density functions, f and g, is defined as:

$$D(f \parallel g) \coloneqq \int f(x) \log \frac{f(x)}{g(x)} dx \coloneqq \mathbb{E}_f \left[\log \frac{f(X)}{g(X)} \right]. \tag{2.7}$$

It is easy to show that $D(f \parallel g) \geq 0$, with equality if and only if $f = g$. For a detailed study of the KL divergence see, e.g., [16].

A simple justification as to why the SPRT is a suitable algorithm for the binary sequential hypothesis testing problem can be given by examining the expected value of the test statistic under both regimes. In particular, for hypotheses H_0 and H_1, respectively, we have that

$$\mathbb{E}_{f_0}[S_k] = \mathbb{E}_{f_0} \left[\sum_{j=1}^{k} \log \frac{f_1(X_j)}{f_0(X_j)} \right] = -\sum_{j=1}^{k} \mathbb{E}_{f_0} \left[\log \frac{f_0(X_j)}{f_1(X_j)} \right] = -kD(f_0 \parallel f_1) < 0$$

and

$$\mathbb{E}_{f_1}[S_k] = \mathbb{E}_{f_1} \left[\sum_{j=1}^{k} \log \frac{f_1(X_j)}{f_0(X_j)} \right] = \sum_{j=1}^{k} \mathbb{E}_{f_1} \left[\log \frac{f_1(X_j)}{f_0(X_j)} \right] = kD(f_1 \parallel f_0) > 0.$$

As a result, when H_0 is the true hypothesis, the test statistic will decrease with an expected drift of $-D(f_0 \parallel f_1)$, eventually passing a negative threshold of a. Similarly, when H_1 is the true hypothesis, the test statistic will grow with an expected drift of $D(f_1 \parallel f_0)$, eventually crossing a positive threshold of b. Note that since the average drifts are fixed for a specific pair of distributions, making b larger or a smaller will result in a larger number of sampled measurements on average, thus improving the decision accuracy. For a thorough analysis of the SPRT performance evaluation see [15].

CHAPTER 3

QUICKEST CHANGE DETECTION

In this chapter, we study the problem of quickest change detection (QCD). In QCD, the goal is to detect an abrupt change in the statistical behavior of a sequentially observed process. Detection techniques aim to minimize the delay under false alarm constraints, i.e., detect a change in the statistical behavior of the process fast enough while maintaining a sufficiently low occurrence rate of false alarm events. The theoretical results of this chapter are fundamental for the rest of the thesis. For a deeper analysis of QCD theory we refer the reader to [4]; also see [5] and [6].

3.1 Problem Statement

In Chapter 2, we studied the problem of binary sequential hypothesis testing, where a sequence of measurements is characterized by one of two statistical behaviors, and we aim to detect the true behavior by processing samples in a sequential manner. In QCD, we observe a random process that initially follows a distribution f_0 i.i.d. At some unknown time instant γ, the process switches to a distribution f_1 i.i.d. In summary, the statistical behavior of the process is as follows:

$$X_k \sim f_0, \text{for } k < \gamma$$
$$X_k \sim f_1, \text{for } k \geq \gamma.$$

This is also called the i.i.d. setting. For the traditional, non-composite instance of the problem, both f_0 and f_1 are known beforehand. In Fig. 3.1 we generate a sequence of samples that are characterized by a distribution shift of this form. In this work, we will focus on the minimax setting of the problem, where it is assumed that $\gamma \geq 1$ is unknown but not random.

The goal in QCD is to create a procedure that will be used to detect the

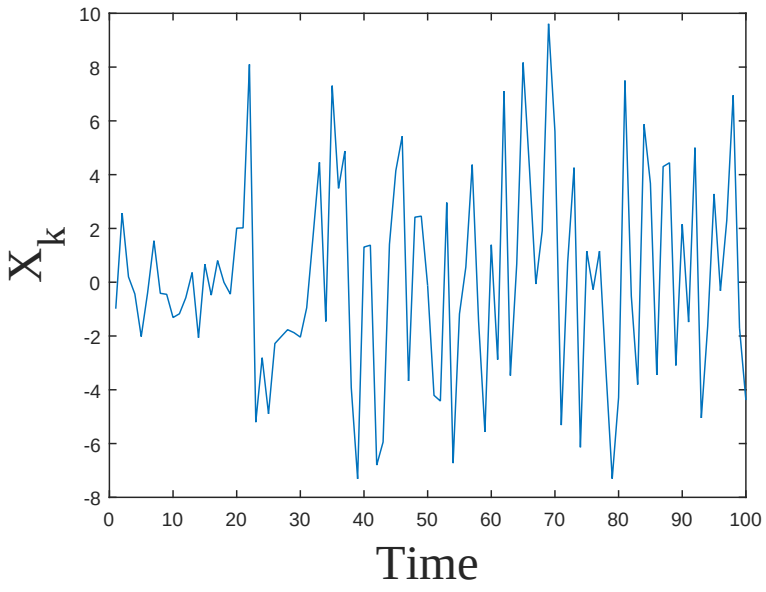


Figure 3.1: Example of sample path of a process under the QCD setting. At $\gamma = 20$ the distribution changes from $\mathcal{N}(0,1)$ to $\mathcal{N}(0,2)$.

abrupt statistical behavior change that occurs at γ. This procedure will have the structure of a stopping time τ adapted on $\{X_k\}$, with the understanding that instead of stopping sampling, we declare that a change has occurred.

3.1.1 Minimax QCD Tradeoff Formulations

The QCD problem is characterized by an underlying tradeoff: on one hand, we want to detect distribution changes as fast as possible, while on the other hand we want to avoid frequent false alarm events, i.e., avoid declaring a change has occurred, when it has not occurred yet. To this end, we present two popular delay–mean time to false alarm formulations that are used in the QCD literature, due to Lorden [17] and Pollak [18] respectively.

Before presenting these two tradeoff formulations we need to define two delay metrics that we will be using. The first delay metric, which was proposed by Lorden, is based on the expected value of $(\tau - \gamma)^+$ conditioned on the worst possible measurements before change. In particular, for a stopping time τ, define the first delay metric as:

$$\mathsf{WADD}(\tau) = \sup_{\gamma \geq 1} \operatorname{ess\,sup} \mathbb{E}_\gamma \left[(\tau - \gamma)^+ \middle| X_1, \ldots, X_{\gamma-1} \right], \tag{3.1}$$

where $\mathbb{E}_\gamma$ denotes the expected value when the underlying distribution is the one induced on the sequence of observations when an outage occurs at γ, and $(x)^+ \coloneqq \max\{x, 0\}$. Note that (3.1) involves taking the expected value of $(\tau - \gamma)^+$ after conditioning on a set of observations $\{X_1, \ldots, X_{\gamma-1}\}$. Using the ess sup can be seen as choosing the worst possible set of observations to condition on, i.e., the set of observations that maximize the expected delay for given γ. Finally, since the time instant during which the distribution change happens is unknown, we have to take the sup over all possible changepoints.

The second delay metric was proposed by Pollak, and is defined as follows:

$$\mathsf{CADD}(\tau) = \sup_{\gamma \geq 1} \mathbb{E}_\gamma \left[\tau - \gamma \middle| \tau \geq \gamma \right]. \tag{3.2}$$

It is easy to show that for any stopping rule τ we have that

$$\mathsf{CADD}(\tau) \leq \mathsf{WADD}(\tau),$$

i.e., the WADD metric is a more pessimistic way of measuring the delay.

It should be noted that although the delay metrics of (3.1) and (3.2) are generally difficult to compute, for the algorithms studied in this thesis, they can be easily computed by Monte Carlo simulations (see [6] for more details).

With the two defined delays in mind we move on to present the related tradeoff formulations. We start with Lorden's formulation:

Formulation 1.

$$\begin{aligned} \min_{\tau} \quad & \mathsf{WADD}(\tau) \\ \text{subject to} \quad & \mathbb{E}_\infty[\tau] \geq \beta. \end{aligned} \tag{3.3}$$

Lorden's formulation involves searching among the set of stopping times that satisfy $\mathbb{E}_\infty[\tau] \geq \beta$, where $\mathbb{E}_\infty$ the expectation under the measure that no distribution change occurs, to find the one that minimizes WADD. The inequality constraint is imposed to guarantee that false alarm events will be sufficiently rare.

Similarly, for Pollak's delay we have a respective tradeoff formulation:

Formulation 2.

$$\begin{aligned} \min_{\tau} \quad & \mathsf{CADD}(\tau) \\ \text{subject to} \quad & \mathbb{E}_\infty[\tau] \geq \beta. \end{aligned} \tag{3.4}$$

In the next section we present algorithms that are used in QCD theory.

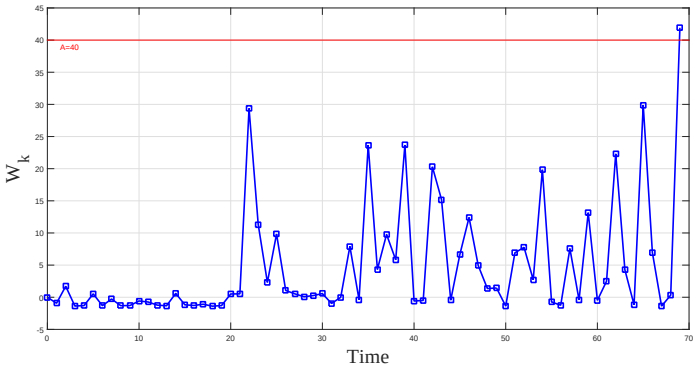


Figure 3.2: Shewhart test run for data of Figure 3.1.

3.2 QCD Algorithms

A simple QCD algorithm was proposed by Shewhart in [19]. In defining the Shewhart test, only the current observation is used. In particular, define the Shewhart test statistic by

$$W_k^{\mathrm{SH}} = \log \frac{f_1(X_k)}{f_0(X_k)}. \tag{3.5}$$

The Shewhart test is based on the fact that the expected value of the log-likelihood ratio after change is given by $D(f_1 \parallel f_0)$, which is a positive quantity; thus, detection can be achieved by using a positive threshold. In particular, the Shewhart stopping time is defined as

$$\tau^{\mathrm{SH}} = \inf\{k \geq 1 : W_k^{\mathrm{SH}} > A\}, \tag{3.6}$$

where $A > 0$ the threshold. In Fig. 3.2, we show the evolution of the Shewhart statistic for the samples shown in Fig. 3.1.

A detection scheme that enjoys optimality properties with respect to Pollak's and Lorden's formulation is the so-called Cumulative Sum (CuSum) algorithm, proposed by Page in [20]. The CuSum algorithm involves accumulating log-likelihood ratios between the pre- and post-change distribution (hence the name of the algorithm) to form the test statistic, and declaring a distribution change has occurred when said test statistic crosses a positive threshold. In particular, the CuSum statistic at time k is given by the

following recursion:

$$W_k^{\mathrm{C}} = \left(W_{k-1}^{\mathrm{C}} + \log \frac{f_1(X_k)}{f_0(X_k)} \right)^+, \tag{3.7}$$

with $W_0^{\mathrm{C}} = 0$. The corresponding CuSum stopping time is defined as

$$\tau^{\mathrm{C}} = \inf\{k \geq 1 : W_k^{\mathrm{C}} > A\}. \tag{3.8}$$

In Fig. 3.3, we show the evolution of the CuSum statistic for the samples shown in Fig. 3.1.

The notion of KL divergence can also be used here to provide an intuitive interpretation of the CuSum test. In particular, understanding the idea behind the algorithm boils down to understanding the behavior of the logarithmic term of (3.7) before and after the distribution change. Before the change occurs, we have that for the expected value of the logarithmic term of (3.7):

$$\mathbb{E}_{f_0}\left[\log \frac{f_1(X_k)}{f_0(X_k)} \right] = -D(f_0 \parallel f_1) < 0,$$

which will cause the test statistic to take non-negative values around zero. After the change occurs we have that

$$\mathbb{E}_{f_1}\left[\log \frac{f_1(X_k)}{f_0(X_k)} \right] = D(f_1 \parallel f_0) > 0.$$

Thus, the CuSum statistic will grow with a positive average drift of $D(f_0 \parallel f_1)$, eventually crossing a positive threshold and declaring a distribution change has occurred.

An algorithm that has a very strong connection to the CuSum algorithm is the so called Shiryaev-Roberts (SR) algorithm. The SR test statistic is given by the following recursion:

$$W_k^{\mathrm{SR}} = (1 + W_{k-1}^{\mathrm{SR}}) \log \frac{f_1(X_k)}{f_0(X_k)}, \tag{3.9}$$

with $W_0^{\mathrm{SR}} = 0$. The corresponding SR stopping time is defined as

$$\tau^{\mathrm{SR}} = \inf\{k \geq 1 : W_k^{\mathrm{C}} > A\}. \tag{3.10}$$

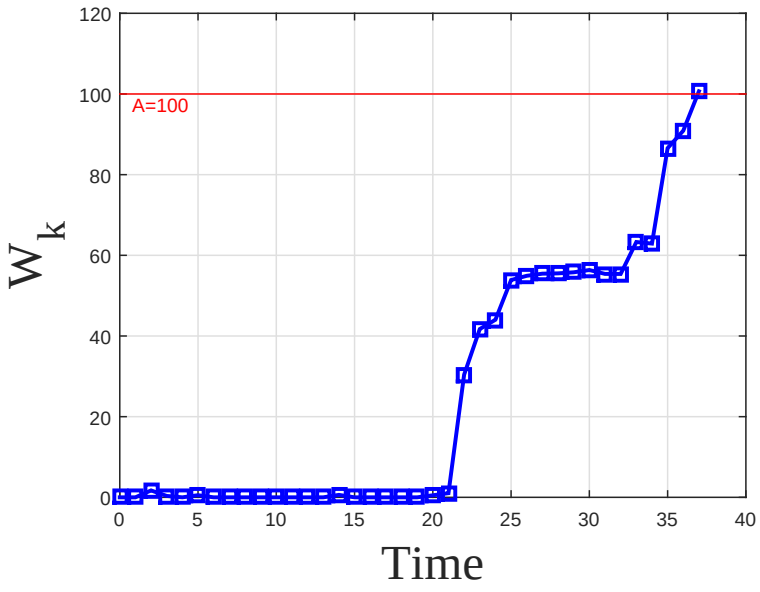


Figure 3.3: CuSum algorithm run for data of Figure 3.1.

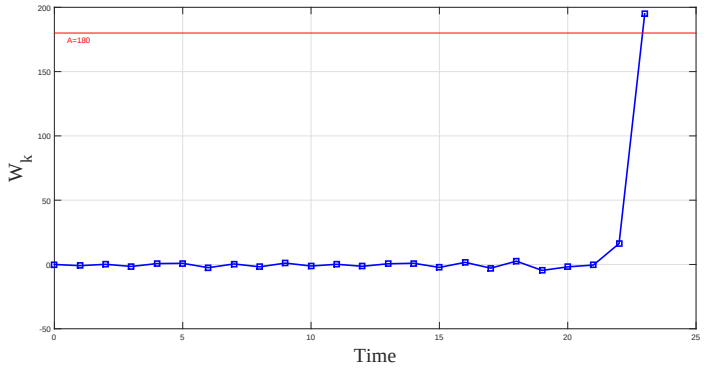


Figure 3.4: Shiryaev-Roberts algorithm run for data of Figure 3.1.

In Fig. 3.4 we show the evolution of the SR statistic for the samples shown in Fig. 3.1.

3.3 Optimality Properties of CuSum and Shiryaev-Roberts Algorithms

In Sec. 3.2, we presented algorithms that can be used to detect an abrupt change in the statistical behavior of a sequence of measurements under the usual i.i.d. setting. In the current section, we provide theoretical justification for the use of the CuSum and SR algorithms by reviewing properties regard-

ing the optimality of the algorithms for Lorden's and Pollak's formulations of the QCD problem.

For the i.i.d. setting, Lorden showed in [17] that the CuSum algorithm is asymptotically optimal with respect to Formulation 1 as $\gamma \to \infty$. In detail we have the following theorem:

Theorem 1. With a threshold choice of $A = \log \gamma$, the CuSum stopping rule (3.8) satisfies

$$\mathbb{E}_\infty[\tau^{\mathrm{C}}] \geq \gamma$$

and

$$\inf_{\tau:\mathbb{E}_\infty(\tau)\geq\gamma} \mathrm{WADD}(\tau) \sim \mathrm{WADD}(\tau^{\mathrm{C}}) \sim \frac{\log \gamma}{D(f_1 \parallel f_0)},$$

as $\gamma \to \infty$, where the $\sim$ notation is used to denote that the ratio of the quantities on the two sides of the $\sim$ approaches 1 in the limit as $\gamma \to \infty$.

A stronger result was proved in [21] and later in [22], where it was shown that the CuSum algorithm is exactly optimal with respect to Formulation 1.

For the same QCD setting, Pollak showed in [18] that the SR algorithm is asymptotically optimal with respect to Formulation 2 as $\gamma \to \infty$. In detail, we have the following theorem:

Theorem 2. With a threshold choice of $A = \gamma$, the SR stopping rule (3.10) satisfies

$$\mathbb{E}_\infty[\tau^{\mathrm{SR}}] \geq \gamma$$

and

$$\inf_{\tau:\mathbb{E}_\infty(\tau)\geq\gamma} \mathrm{CADD}(\tau) \simeq \mathrm{CADD}(\tau^{\mathrm{SR}}) \simeq \frac{\log \gamma}{D(f_1 \parallel f_0)},$$

as $\gamma \to \infty$, where the $\simeq$ notation is used to denote that the difference of the quantities on the two sides of the $\simeq$ approaches 0 in the limit as $\gamma \to \infty$.

Since the performance of both the CuSum and the SR algorithms are asymptotically equal, we have that Theorem 1 and Theorem 2 imply that both the algorithms are asymptotically optimal with respect to Formulation 1 and Formulation 2.

CHAPTER 4

QUICKEST CHANGE DETECTION UNDER TRANSIENT DYNAMICS

Up to now, we have studied the traditional QCD problem, where at some time instant γ the distribution of the observed process changes from an initial distribution to a final distribution. In this section, we further generalize the QCD problem by incorporating transient dynamics. As a result, the shift from the initial to the final distribution does not happen instantaneously, but after a series of cascading transient stages of finite duration, each one corresponding to a different probability distribution. We study the non-composite version of the transient QCD problem, where all distributions are known beforehand. We introduce the Dynamic CuSum (D-CuSum) algorithm as the proposed algorithm.

4.1 The Dynamic CuSum Algorithm

Assume a random process $\{X_k\}_{k=1}^{\infty}$ with the following statistical behavior:

$$X_k \sim \begin{cases} f_0, \text{ if } 1 \leq k < \gamma_0, \\ f^{(0)}, \text{ if } \gamma_0 \leq k < \gamma_1, \\ \vdots \\ f^{(i)}, \text{ if } \gamma_i \leq k < \gamma_{i+1}, \\ \vdots \\ f^{(T)}, \text{ if } \gamma_T \leq k, \end{cases} \tag{4.1}$$

where $\gamma_i \in \mathbb{N}$, $i = 0, \ldots, T$. Note that the case of $\gamma_{i+1} = \gamma_i + 1$ corresponds to a transient stage with a duration of one time instant. The goal is to design a stopping rule that will detect the change in the statistical behavior of the observed process that takes place at time instant γ_0.

A heuristic test solution can be derived by considering this problem as a dynamic composite hypothesis testing problem. Thus, at every time instant

k, choose between the following two hypotheses:

$$H_0^k : k < \gamma_0,$$
$$H_1^k : k \geq \gamma_0.$$

The nominal hypothesis H_0^k corresponds to the case that the time instant γ_0 has not been reached yet, while the alternative hypothesis H_1^k corresponds to the case that γ_0 has been reached. Each hypothesis induces a different set of distributions on the data $X_1, X_2, \ldots, X_k$. In particular, H_0^k is a single hypothesis under which the data follow distribution f_0 i.i.d. and H_1^k is a composite hypothesis, i.e., it induces one distribution belonging to a set of distributions. The distribution that is induced depends on the values of the γ's and k. To find the test statistic we first form the likelihood ratio of this hypothesis testing problem for an arbitrary choice of γ's:

$$\frac{\prod_{j=\gamma_0}^{\min\{\gamma_1-1,k\}} f^{(0)}(X_j) \cdots \prod_{j=\min\{\gamma_T-1,k\}+1}^{k} f^{(T)}(X_j)}{\prod_{j=\gamma_0}^{k} f_0(X_j)}.$$

Note that this likelihood ratio should be interpreted with the understanding that $\prod_{j=k+1}^{k} \frac{f^{(i)}(X_j)}{f_0(X_j)} := 1$ for $i = 0, \ldots, T$. This is a natural generalization of the maximum likelihood interpretation of the CuSum statistic [6]. The test statistic is derived by taking the maximum with respect to $\gamma_0, \ldots, \gamma_T$. An equivalent test statistic can be derived by maximizing the logarithm of the above quantity. As a result, we have that

$$W_k = \max_{\gamma_0 < \cdots < \gamma_T} \left\{ \sum_{j=\gamma_0}^{\min\{\gamma_1-1,k\}} \log \frac{f^{(0)}(X_j)}{f_0(X_j)} + \cdots + \sum_{j=\min\{\gamma_T-1,k\}+1}^{k} \log \frac{f^{(T)}(X_j)}{f_0(X_j)} \right\},$$

with the understanding that $\gamma_0 \leq k$ holds. This maximization is the reason the test is independent of the transient duration, as will be seen later. W_k can be written in the following way:

$$W_k = \max\{\Omega_k^{(0)}, \ldots, \Omega_k^{(i)}, \ldots, \Omega_k^{(T)}\}, \tag{4.2}$$

where

$$\Omega_k^{(i)} = \max_{\gamma_0<\gamma_1<\cdots<\gamma_i\leq k} \left\{ \sum_{j=\gamma_0}^{\gamma_1-1} \log \frac{f^{(0)}(X_j)}{f_0(X_j)} + \ldots + \sum_{j=\gamma_i}^{k} \log \frac{f^{(i)}(X_j)}{f_0(X_j)} \right\}, i = 0, \ldots, T, \tag{4.3}$$

by using the fact that $\sum_{j=k+1}^{k} \log \frac{f^{(i)}(X_k)}{f_0(X_k)} = 1$. We claim that the (4.3) can be written in a recursive manner as follows:

$$\Omega_k^{(i)} = \max\{\Omega^{(i)}[k-1], \Omega^{(i-1)}[k-1]\} + \log \frac{f^{(i)}(X_k)}{f_0(X_k)},$$

for $i = 0, \ldots, T$ and $\Omega_k^{(0)} \coloneqq 0$ for all $k \in \mathbb{Z}$. First, consider the case $i = 0$:

$$\begin{aligned}
\Omega_k^{(0)} &= \max_{\gamma_0\leq k} \left\{ \sum_{j=\gamma_0}^{k} \log \frac{f^{(0)}(X_j)}{f_0(X_j)} \right\} \\
&= \max_{\gamma_0\leq k} \left\{ \sum_{j=\gamma_0}^{k-1} \log \frac{f^{(0)}(X_j)}{f_0(X_j)} + \log \frac{f^{(0)}(X_k)}{f_0(X_k)} \right\} \\
&= \max_{\gamma_0\leq k} \left\{ \sum_{j=\gamma_0}^{k-1} \log \frac{f^{(0)}(X_j)}{f_0(X_j)} \right\} + \log \frac{f^{(0)}(X_k)}{f_0(X_k)} \\
&= \max \left\{ \max_{\gamma_0\leq k-1} \left[\sum_{j=\gamma_0}^{k-1} \log \frac{f^{(0)}(X_j)}{f_0(X_j)} \right], \sum_{j=k}^{k-1} \log \frac{f^{(0)}(X_j)}{f_0(X_j)} \right\} + \log \frac{f^{(0)}(X_k)}{f_0(X_k)} \\
&= \max\{\Omega^{(0)}[k-1], 0\} + \log \frac{f^{(0)}(X_k)}{f_0(X_k)}.
\end{aligned}$$

Since $\Omega_k^{(-1)} \coloneqq 0$, the argument we attempt to prove holds for the case of $i = 0$. Now for the case of an arbitrary i:

$$\Omega_k^{(i)} = \max_{\gamma_0<\gamma_1<\cdots<\gamma_i\leq k} \left\{ \sum_{j=\gamma_0}^{\gamma_1-1} \log \frac{f^{(0)}(X_j)}{f_0(X_j)} + \cdots + \sum_{j=\gamma_i}^{k} \log \frac{f^{(i)}(X_j)}{f_0(X_j)} \right\} =$$
$$\max_{\gamma_0<\gamma_1<\cdots<\gamma_i\leq k} \left\{ \sum_{j=\gamma_0}^{\gamma_1-1} \log \frac{f^{(0)}(X_j)}{f_0(X_j)} + \cdots + \sum_{j=\gamma_i}^{k-1} \log \frac{f^{(i)}(X_j)}{f_0(X_j)} \right\} +$$
$$\log \frac{f^{(i)}(X_k)}{f_0(X_k)}.$$

Consider the first term of this expression. We have that:

$$\max_{\gamma_0<\gamma_1<\cdots<\gamma_i\leq k} \left\{ \sum_{j=\gamma_0}^{\gamma_1-1} \log \frac{f^{(0)}(X_j)}{f_0(X_j)} + \cdots + \sum_{j=\gamma_i}^{k-1} \log \frac{f^{(i)}(X_j)}{f_0(X_j)} \right\}$$
$$= \max \left\{ \max_{\gamma_0<\gamma_1<\cdots<\gamma_i\leq k-1} \left[\sum_{j=\gamma_0}^{\gamma_1-1} \log \frac{f^{(0)}(X_j)}{f_0(X_j)} + \sum_{j=\gamma_i}^{k-1} \log \frac{f^{(i)}(X_j)}{f_0(X_j)} \right] \right.$$
$$\left. = \max_{\gamma_0<\gamma_1<\cdots<\gamma_{i-1}\leq k-1} \left[\sum_{j=\gamma_0}^{\gamma_1-1} \log \frac{f^{(0)}(X_j)}{f_0(X_j)} + \sum_{j=\gamma_{i-1}}^{k-1} \log \frac{f^{(i-1)}(X_j)}{f_0(X_j)} \right] \right\}$$
$$= \max\{\Omega^{(i)}[k-1], \Omega^{(i-1)}[k-1]\}.$$

Since the test statistic will be compared to a positive threshold, an equivalent test can be derived by not allowing the test statistic to take negative values. Thus, the final D-CuSum test statistic is defined as follows:

$$W_k = \max \left\{ \Omega_k^{(0)}, \ldots, \Omega_k^{(T)}, 0 \right\}, \tag{4.4}$$

where

$$\Omega_k^{(i)} = \max\{\Omega_{k-1}^{(i)}, \Omega_{k-1}^{(i-1)}\} + \log \frac{f^{(i)}(X_k)}{f_0(X_k)}, \tag{4.5}$$

for $i = 0, \ldots, T$, $\Omega_k^{(-1)} \coloneqq 0$ for all $k \in \mathbb{Z}$ and $\Omega_0^{(i)} \coloneqq 0$ for all i.

The corresponding stopping time is given by comparing W_k against a predetermined positive threshold:

$$\tau = \min\{k \geq 1 : W_k > A\}.$$

To demonstrate the performance of the algorithm we generate a process

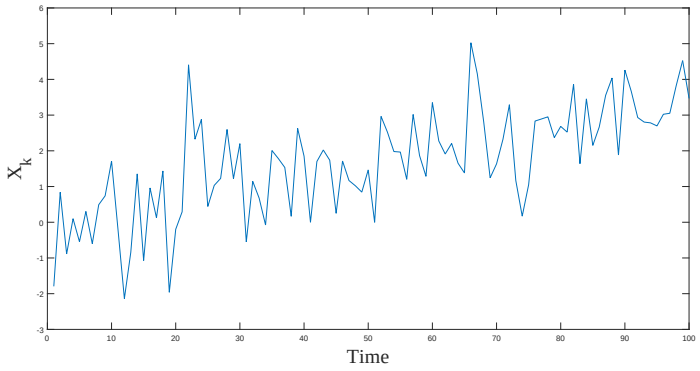


Figure 4.1: A typical realization of a sequence of i.i.d. Gaussian variables. The statistics of the process are characterized by (4.6).

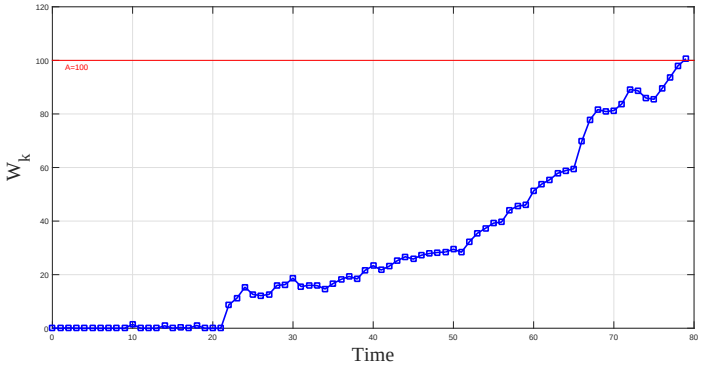


Figure 4.2: D-CuSum test run for data of Figure 4.1.

with the following statistical behavior:

$$X_k \sim \begin{cases} \mathcal{N}(0,1), \text{ if } 1 \leq k \leq 19, \\ \mathcal{N}(0,1.5), \text{ if } 20 \leq k \leq 39, \\ \mathcal{N}(0,2), \text{ if } 40 \leq k \leq 59, \\ \mathcal{N}(0,2.5), \text{ if } 60 \leq k \leq 79, \\ \mathcal{N}(0,3), \text{ if } 80 \leq k. \end{cases} \tag{4.6}$$

A realization of this process up to sample 100 is shown in Fig. 4.1. The evolution of the D-CuSum algorithm for this process realization is shown in Fig. 4.2. We see that the D-CuSum statistic starts to grow after $\gamma_0 = 20$, eventually crossing a threshold of $A = 100$.

CHAPTER 5

QCD ALGORITHMS FOR NON-TRANSIENT POWER SYSTEM LINE OUTAGE DETECTION

In this Chapter, we study the line outage detection problem in the case of no transient dynamics. We start by presenting the complete underlying power system model, which includes an arbitrary number of transient periods. Next, we propose a statistical algorithm for detecting line outages in a power system, for the special case of one post-change stage, and show that it has better performance than other schemes proposed in the literature. Our algorithm is based on the Generalized Cumulative Sum (G-CuSum) test from the quickest change detection (QCD) literature, a test which is formulated by using the CuSum statistic in a generalized manner, i.e., by calculating a test statistic for each post-change distribution, and comparing each statistic with a corresponding threshold. Different methods of selecting the test thresholds, including using the notion of KL divergence, are examined. Our algorithm exploits the statistical properties of the measured voltage phase angles before, during, and after a line outage, whereas other methods in the literature only utilize the change in statistics that occurs at the instant of outage. From now on, the time indexes for every process will be placed in braces and not as a subscript, so that reading is made easier.

5.1 Power System Model

Let $\mathcal{L} = \{1, \ldots, L\}$ denote the set of lines in a system with N buses. A transmission line can be denoted either by an integer ℓ, or by a couple of integers (m, n) denoting that this line connects bus m to bus n. At time t, let $V_i(t)$ and $\theta_i(t)$ denote the voltage magnitude and phase angle at bus i respectively, and let $P_i(t)$ and $Q_i(t)$ denote the net active and reactive power injection at bus i, respectively. Then, the quasi-steady-state behavior of the system can be described by the power flow equations (see e.g., [23]), which

for bus i can be compactly written as:

$$\begin{aligned} P_i(t) &= p_i(\theta_1(t), \ldots, \theta_N(t), V_1(t), \ldots, V_N(t)), \\ Q_i(t) &= q_i(\theta_1(t), \ldots, \theta_N(t), V_1(t), \ldots, V_N(t)), \end{aligned} \tag{5.1}$$

where the dependence on the system network parameters is implicitly captured by $p_i(\cdot)$ and $q_i(\cdot)$. The outage of line $\ell \in \mathcal{L}$ at time $t = t_f$ is assumed to be persistent (i.e., the line is not restored until it is detected to be outaged), with $\gamma_0 \Delta t \leq t_f < (\gamma_0 + 1)\Delta t$, where Δt is the time between successive PMU samples. In addition, assume that the loss of line ℓ does not cause islands to form in the post-event system (i.e., the underlying graph representing the internal power system remains connected).

5.1.1 Pre-outage Model

Let $P_i[k] \coloneqq P_i(k\Delta t)$ and $Q_i[k] \coloneqq Q_i(k\Delta t)$, $\Delta t > 0$, $k = 0, 1, 2, \ldots$, denote the k^{th} measurement sample of active and reactive power injections into bus i. Similarly, let $V_i[k]$ and $\theta_i[k]$, $k = 0, 1, 2, \ldots$, denote bus i's k^{th} voltage magnitude and angle measurement sample. Furthermore, define variations in voltage magnitudes and phase angles between consecutive sampling times $k\Delta t$ and $(k+1)\Delta t$ as $\Delta V_i[k] \coloneqq V_i[k+1] - V_i[k]$, and $\Delta\theta_i[k] \coloneqq \theta_i[k+1] - \theta_i[k]$, respectively. Similarly, variations in the active and reactive power injections at bus i between two consecutive sampling times are defined as $\Delta P_i[k] = P_i[k+1] - P_i[k]$ and $\Delta Q_i[k] = Q_i[k+1] - Q_i[k]$.

Proceeding in the same manner as in [3], we linearize the power flow equations of (5.1) around $(\theta_i[k], V_i[k], P_i[k], Q_i[k])$, $i = 1, \ldots, N$, and use the DC power flow assumptions (see e.g., [23]), namely, (i) flat voltage profile, (ii) negligible line resistances, and (iii) small phase angle differences, to decouple the real and reactive power flow equations. Then, after omitting the equation corresponding to the reference bus, the relationship between voltage phase angles and the variations in the real power injection can be expressed as:

$$\Delta P[k] \approx H_0 \Delta\theta[k], \tag{5.2}$$

where $\Delta P[k]$, $\Delta\theta[k] \in \mathbb{R}^{(N-1)}$ and $H_0 \in \mathbb{R}^{(N-1)\times(N-1)}$ is the imaginary part of the system admittance matrix with the row and column corresponding to

the reference bus removed.

In an actual power system, random fluctuations in the load drive the generator response. Therefore, in this thesis, we use the so-called governor power flow model (see e.g., [24]), which is more realistic than the conventional power flow model, where the slack bus picks up any changes in the load power demand. In the governor power flow model, at time instant k, the relation between changes in the load demand vector, $\Delta P^d[k] \in \mathbb{R}^{N_d}$, and changes in the power generation vector, $\Delta P^g[k] \in \mathbb{R}^{N_g}$, is described by

$$\Delta P^g[k] = B(t)\Delta P^d[k], \tag{5.3}$$

where $B(t)$ is a time dependent matrix of participation factors. We approximate $B(t)$ by quantizing it to take values B_i, $i = 0, 1, \ldots, T$, where i denotes the time period of interest. Let $B(t) = B_0$ and $M_0 \coloneqq H_0^{-1}$ during the pre-outage period. Then, we can substitute (5.3) into (5.2) to obtain a pre-outage relation between the changes in the voltage angles and the real power demand at the load buses as follows:

$$\begin{aligned}
\Delta\theta[k] &\approx M_0 \Delta P[k] \\
&= M_0 \begin{bmatrix} \Delta P^g[k] \\ \Delta P^d[k] \end{bmatrix} \\
&= [M_0^1 \; M_0^2] \begin{bmatrix} B_0 \Delta P^d[k] \\ \Delta P^d[k] \end{bmatrix} \\
&= (M_0^1 B_0 + M_0^2)\Delta P^d[k] \\
&= \tilde{M}_0 \Delta P^d[k],
\end{aligned} \tag{5.4}$$

where $\tilde{M}_0 = M_0^1 B_0 + M_0^2$.

5.1.2 Instantaneous Change During Outage

At the time of outage, $t = t_f$, there is an instantaneous change in the mean of the voltage phase angle measurements that affects only one incremental sample, namely, $\Delta\theta[\gamma_0] = \theta[\gamma_0 + 1] - \theta[\gamma_0]$. The measurement $\theta[\gamma_0]$ is taken immediately prior to the outage, whereas $\theta[\gamma_0 + 1]$ is the measurement taken immediately after the outage. Suppose the outaged line ℓ connects buses m

and n. Then, the effect of an outage in line ℓ can be modeled with a power injection of $P_\ell[\gamma_0]$ at bus m and $-P_\ell[\gamma_0]$ at bus n, where $P_\ell[\gamma_0]$ is the pre-outage line flow across line ℓ from m to n. Following a similar approach as that in [3], the relation between the incremental voltage phase angle at the instant of outage, $\Delta\theta[\gamma_0]$, and the variations in the real power flow can be expressed as:

$$\Delta\theta[\gamma_0] \approx M_0 \Delta P[\gamma_0] - P_\ell[\gamma_0 + 1] M_0 r_\ell, \tag{5.5}$$

where $r_\ell \in \mathbb{R}^{N-1}$ is a vector with the $(m-1)^{\text{th}}$ entry equal to 1, the $(n-1)^{\text{th}}$ entry equal to -1, and all other entries equal 0. Furthermore, by using the governor power flow model of (5.3) and substituting into (5.5), and simplifying, we obtain:

$$\Delta\theta[\gamma_0] \approx \tilde{M}_0 \Delta P^d[\gamma_0] - P_\ell[\gamma_0 + 1] M_0 r_\ell. \tag{5.6}$$

5.1.3 Post-Outage

Following a line outage, the power system undergoes a transient response governed by B_i, $i = 1, 2, \ldots, T-1$ until quasi-steady-state is reached, in which $B(t)$ settles to a constant B_T. For example, immediately after the outage occurs, the power system is dominated by the inertial response of the generators, which is then followed by the governor response. As a result of the line outage, the system topology changes, which manifests itself in the matrix H_0. This change in the matrix H_0 resulting from the outage can be expressed as the sum of the pre-outage matrix and a perturbation matrix, ΔH_ℓ, i.e., $H_\ell = H_0 + \Delta H_\ell$. Then, by letting $M_\ell \coloneqq H_\ell^{-1} = [M_\ell^1 \; M_\ell^2]$, and deriving in the same manner as the pre-outage model of (5.4), we obtain the post-outage relation between the changes in the voltage angles and the real power demand as:

$$\Delta\theta[k] \approx \tilde{M}_{\ell,i} \Delta P^d[k], \quad \gamma_{i-1} \leq k < \gamma_i, \tag{5.7}$$

where $\tilde{M}_{\ell,i} = M_\ell^1 B_i + M_\ell^2$, $i = 1, 2, \ldots, T$.

5.1.4 Measurement Model

Since the voltage phase angles, $\theta[k]$, are assumed to be measured by PMUs, we allow for the scenario where the angles are measured at only a subset of the load buses, and denote this reduced measurement set by $\hat{\theta}[k]$. Suppose that there are N_d load buses and we select $p \leq N_d$ locations to deploy the PMUs. Then, there are $\binom{N_d}{p}$ possible locations to place the PMUs. In this thesis, we assume that the PMU locations are fixed; in general, the problem of optimal PMU placement is NP-hard and its treatment is beyond the scope of this thesis.

Let

$$\tilde{M} = \begin{cases} \tilde{M}_0, \text{ if } 1 \leq k < \gamma_0, \\ \vdots \\ \tilde{M}_{\ell,T}, \text{ if } k \geq \gamma_T. \end{cases} \tag{5.8}$$

Then, the absence of a PMU at bus i corresponds to removing the i^{th} row of $\tilde{M}$. Thus, let $\hat{M} \in \mathbb{R}^{p \times N_d}$ be the matrix obtained by removing $N - p - 1$ rows from $\tilde{M}$. Therefore, we can relate $\hat{M}$ to $\tilde{M}$ in (5.8) as follows:

$$\hat{M} = C\tilde{M}, \tag{5.9}$$

where $C \in \mathbb{R}^{p \times (N-1)}$ is a matrix of 1's and 0's that appropriately selects the rows of $\tilde{M}$. Accordingly, the increments in the phase angle can be expressed as follows:

$$\Delta\hat{\theta}[k] \approx \hat{M}\Delta P^d[k]. \tag{5.10}$$

The small variations in the real power injections at the load buses, $\Delta P^d[k]$, can be attributed to random fluctuations in electricity consumption. In this regard, we may model the $\Delta P^d[k]$'s as independent and identically distributed (i.i.d.) random vectors. By the Central Limit Theorem [25], it can be argued that each $\Delta P^d[k]$ is a Gaussian vector, i.e., $\Delta P^d[k] \sim \mathcal{N}(0, \Lambda)$, where Λ is the covariance matrix. Note that the elements $\Delta P^d[k]$ are roughly independent. Since $\Delta\hat{\theta}[k]$ depends on $\Delta P^d[k]$ through the linear relationship given in (5.10), we have that:

$$\Delta\hat{\theta}[k] \sim \begin{cases} f_0 \coloneqq \mathcal{N}(0, \hat{M}_0 \Lambda \hat{M}_0^T), \text{ if } 1 \leq k < \gamma_0, \\ f_\ell^{(0)} \coloneqq \mathcal{N}(-P_\ell[\gamma+1]CM_0 r_\ell, \\ \qquad\qquad \hat{M}_0 \Lambda \hat{M}_0^T), \text{ if } k = \gamma_0, \\ \vdots \\ f_\ell^{(T)} \coloneqq \mathcal{N}(0, \hat{M}_{\ell,T} \Lambda \hat{M}_{\ell,T}^T), \text{ if } k \geq \gamma_T, \end{cases} \tag{5.11}$$

It is important to note that for $\mathcal{N}\left(0, \hat{M}\Lambda\hat{M}^T\right)$ to be a nondegenerate p.d.f., its covariance matrix, $\hat{M}\Lambda\hat{M}^T$, must be full rank. We enforce this by ensuring that the number of PMUs allocated, p, is less than or equal to the number of load buses, N_d.

5.1.5 Non-transient Statistical Model

In this chapter, we focus on the special case where $T = 1$ (also note that due to the meanshift $\gamma_1 = \gamma_0 + 1$), i.e., after an outage the distribution changes from pre-change to a post-change distribution after going through the meanshift phase. In particular, in this chapter we will develop line outage detection techniques for the case that $\{\Delta\hat{\theta}[k]\}_{k=1}^{\infty}$ is characterized by the following statistical behavior:

$$\Delta\hat{\theta}[k] \sim \begin{cases} f_0 \coloneqq \mathcal{N}(0, \hat{M}_0 \Lambda \hat{M}_0^T), \text{ if } 1 \leq k < \gamma_0, \\ f_\ell^{(0)} \coloneqq \mathcal{N}(-P_\ell[\gamma+1]CM_0 r_\ell, \\ \qquad\qquad \hat{M}_0 \Lambda \hat{M}_0^T), \text{ if } k = \gamma_0, \\ f_\ell^{(1)} \coloneqq \mathcal{N}(0, \hat{M}_{\ell,1} \Lambda \hat{M}_{\ell,1}^1), \text{ if } k > \gamma_0. \end{cases} \tag{5.12}$$

5.2 Line Outage Detection Using QCD

In the line outage detection problem setting of the present chapter, the outage induces a change in the statistical characteristics of the observed sequence $\{\Delta\hat{\theta}[k]\}_{k\geq 1}$ which is summarized by (5.12). The goal is to detect the outage in line ℓ as quickly as possible subject to false alarm constraints. It is quite apparent that the present statistical model is almost identical to the QCD setting studied in Chapter 3. Thus, we will use QCD-based stopping rules to detect the statistical behavior shift that occurs after the outage.

5.2.1 Problem Setup

The goal in line outage detection is to design stopping rules that will detect line outages as fast as possible under false alarm constraints. The false alarm constraint that we choose is based on the mean time to false alarm; thus, we would like $\mathbb{E}_\infty[\tau] \geq \beta$, where $\beta > 0$ is a pre-determined parameter, and $\mathbb{E}_\infty$ is the expectation under the probability measure where no outage has occurred.

In order to quantify the detection delay for line outages, we will be using Lorden's delay metric:

$$\mathsf{WADD}_\ell(\tau) = \sup_{\gamma_0 \geq 1} \operatorname{ess\,sup} \mathbb{E}_{\gamma_0,\ell}\left[(\tau - \gamma_0)^+ \middle| \Delta\hat{\theta}[1], \ldots, \Delta\hat{\theta}[\gamma_0 - 1]\right]. \tag{5.13}$$

The difference between the metric of (5.13) and (3.1) is that here we suffer a different delay depending on which line is outaged.

5.3 QCD-based Line Outage Detection Algorithms

With the statistical model for $\{\Delta\hat{\theta}[k]\}_{k\geq 1}$ in place, the problem of detecting a line outage was formulated as a problem of detecting a change in the probability distribution of the sequence of observations $\{\Delta\hat{\theta}[k]\}_{k\geq 1}$ as quickly as possible given false alarm constraints. As a result, we can use the QCD theory presented in this work to design algorithms for line outage detection.

5.3.1 The Generalized CuSum Algorithm

In our setting, the line in which the outage occurs is unknown, i.e., the post-change distribution induced on the observation sequence $\{\Delta\hat{\theta}[k]\}_{k\geq 1}$ is unknown. Since there are L lines, we have L different post-change scenarios, i.e., we have L possible post-change distributions. The problem of detecting an abrupt change in the distribution of a process, in which the post-change distribution belongs to a known set of distributions is called the composite QCD problem. Since we have a total of L post change scenarios, we use the generalized version of the CuSum algorithm, namely, the Generalized CuSum (G-CuSum) algorithm. This algorithm involves calculating a CuSum statistic for each possible post-change scenario. We present two versions of this algorithm: one proposed in [3], in which a common threshold is used for all CuSum statistics, and another in which the thresholds are selected based on each line's KL divergence.

We define the G-CuSum statistic corresponding to line ℓ outage recursively as:

$$W_\ell^{\mathrm{GC}}[k] = \max\left\{W_\ell^{\mathrm{GC}}[k-1] + \log\frac{f_\ell^{(1)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}, \log\frac{f_\ell^{(0)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}, 0\right\}, \tag{5.14}$$

with $W_\ell^{\mathrm{GC}}[0] = 0$ for all $\ell \in \mathcal{L}$. Note that an extra log-likelihood ratio is added inside the maximum. This term is used to capture the instantaneous meanshift that occurs at the time of the outage. The G-CuSum stopping time is defined as:

$$\tau^{\mathrm{GC}} = \min_{\ell\in\mathcal{L}}\left\{\inf\{k \geq 1 : W_\ell^{\mathrm{GC}}[k] > A_\ell^{\mathrm{GC}}\}\right\}. \tag{5.15}$$

We now present different ways of choosing the thresholds for the G-CuSum test. It can be shown (see, e.g., [26]) that by choosing

$$A_\ell^{\mathrm{GC}} = \log\beta - \log\xi_\ell, \tag{5.16}$$

with ξ_ℓ being a positive constant independent of β, the expected delay for each possible outage differs from the corresponding minimum delay among the class of stopping times $\mathcal{C}_\beta = \{\tau : \mathbb{E}_\infty(\tau) \geq \beta\}$, as $\beta \to \infty$, by a bounded constant.

A choice of thresholds for the G-CuSum algorithm is obtained by setting $\beta_\ell = \frac{1}{L}$ for all $\ell \in \mathcal{L}$. This way we get a common threshold, i.e., $A_\ell^{\mathrm{GC}} = A^{\mathrm{GC}} = \log(\beta L)$ for all $\ell \in \mathcal{L}$. It can be shown (see, e.g., [27]) that by choosing the thresholds this way, we can guarantee that $\mathbb{E}_\infty[\tau^{\mathrm{GC}}] \geq \beta$.

Using the results in [26], another choice of the thresholds could be based on a relative performance loss criterion, i.e.,

$$\beta_\ell = \frac{1}{D(f_\ell^{(1)} \parallel f_0)L(\zeta_\ell)^2}, \tag{5.17}$$

where

$$\zeta_\ell = \lim_{b\to\infty}\mathbb{E}_\ell^{(1)}\left[e^{\{-(S_\ell[\tau_\ell^{\mathrm{b}}]-b)\}}\right], \tag{5.18}$$

with

$$\tau_\ell^{\mathrm{b}} = \inf\{k \geq 1 : S_\ell[k] \geq b\}, \tag{5.19}$$

and

$$S_\ell[k] = \sum_{j=1}^{k}\log\frac{f_\ell^{(1)}(\Delta\hat{\theta}[j])}{f_0(\Delta\hat{\theta}[j])}. \tag{5.20}$$

Note that this choice of threshold depends on the asymptotic overshoot of an SPRT-based test. As we show later through case studies in Chapter 7, the threshold choice of (5.17)-(5.20) results in performance gains compared to choosing a

common threshold for all the lines.

5.4 Other Line Outage Detection Algorithms

In this section, we present some other change detection algorithms that can be shown to be equivalent to other techniques proposed in the literature. For example, the line outage detection algorithm proposed in [10] can be shown to be equivalent to a log-likelihood ratio test that only uses the most recent measurements.

5.4.1 Meanshift Test

The meanshift test is a "one-shot" detection scheme in that the algorithm uses only the most recent observation to decide whether a change in the mean has occurred and ignores all past observations. The meanshift statistic corresponding to line ℓ is defined as follows:

$$W_\ell^{\mathrm{MS}}[k] = \log \frac{f_\ell^{(0)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}. \tag{5.21}$$

The decision maker declares a change when one of the L statistics crosses a corresponding threshold, A_ℓ^{MS}. The stopping time for this algorithm is defined as:

$$\tau^{\mathrm{MS}} = \min_{\ell\in\mathcal{L}} \left\{ \inf\{k \geq 1 : W_\ell^{\mathrm{MS}}[k] > A_\ell^{\mathrm{MS}}\} \right\}. \tag{5.22}$$

The meanshift test ignores the persistent covariance change that occurs after the outage. In particular, note that the meanshift test is using the likelihood ratio between the distribution of the observations before and at the changepoint. More specifically, assuming that an outage occurs in line ℓ, the expected value of the statistic at the changepoint is given by

$$\mathbb{E}_\ell^{(0)}\left[\log \frac{f_\ell^{(0)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])} \right] = D(f_\ell^{(0)} \parallel f_0) > 0, \tag{5.23}$$

where $\mathbb{E}_\ell^{(0)}$ denotes the expectation under distribution $f_\ell^{(0)}$. On the other hand, after the changepoint ($k > \gamma_0$), the expected value of the statistic is given by

$$\mathbb{E}_\ell^{(1)}\left[\log \frac{f_\ell^{(0)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])} \right] = D(f_\ell^{(1)} \parallel f_0) - D(f_\ell^{(1)} \parallel f_\ell^{(0)}), \tag{5.24}$$

which could be either positive or negative.

5.4.2 Shewhart Test

Similar to the meanshift test, the Shewhart-based test presented here is also a "one-shot" detection scheme. This test attempts to detect a change on the observation sequence through the meanshift and the change in the covariance of the data. The Shewhart test statistic for line ℓ outage is defined as:

$$W_\ell^{\mathrm{SH}}[k] = \max\left\{\log\frac{f_\ell^{(0)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}, \log\frac{f_\ell^{(1)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}\right\}, \tag{5.25}$$

where the first log-likelihood ratio is used to detect the meanshift, while the second log-likelihood ratio is used to detect the persistent change in the covariance. The stopping time is:

$$\tau^{\mathrm{SH}} = \min_{\ell\in\mathcal{L}}\left\{\inf\{k \geq 1 : W_\ell^{\mathrm{SH}}[k] > A_\ell^{\mathrm{SH}}\}\right\}. \tag{5.26}$$

Since the Shewhart test exploits the covariance change in addition to the meanshift statistic, it should perform better than the meanshift test, at least as the meantime to false alarm goes to infinity, which is verified in the case studies.

CHAPTER 6

LINE OUTAGE DETECTION AND IDENTIFICATION UNDER TRANSIENT DYNAMICS

In this section, we propose a quickest change detection (QCD) algorithm to address the problem of detecting and identifying line outages in a power system when transient dynamics are present, i.e., when we are using the general power system model presented in Chapter 5. The proposed algorithm is applied to the measurements of voltage phase angles, which are collected using phasor measurement units (PMUs). This adaptive algorithm is based on the D-CuSum algorithm also proposed in this work.

6.1 QCD–Based Line Outage Detection Algorithms Under Transient Dynamics

In the present chapter, we assume that the underlying statistical model of the observed process is given by (5.11) for an arbitrary value of T, i.e., we assume an arbitrary number of transient periods post-change. In the setting described in Sec. 5.1, we assume that a sequence of observations $\{\Delta\hat{\theta}[k]\}_{k\geq 1}$ is measured by PMUs and passed sequentially to a decision maker. According to the statistical model in (5.11), before an outage has occurred, $\Delta\hat{\theta}[k] \sim f_0$. At an unknown time instant t_f, an outage occurs in line ℓ and the distribution of $\Delta\hat{\theta}[k]$ changes from f_0 to $f_\ell^{(0)}$. Then, the system undergoes a series of transient responses which corresponds to the distribution of $\Delta\hat{\theta}[k]$ evolving from $f_\ell^{(0)}$ to $f_\ell^{(T)}$. First, a meanshift takes place during the instant of change t_f, where the pdf is $f_\ell^{(0)}$. Then, the statistical behavior of the process is characterized by a series of changes only in the covariance matrix of the measurements.

As seen in Chapters 3 and 5, in classic QCD theory, the CuSum algorithm for the non-composite setting (with known pre- and post-outage distribution) and the Generalized CuSum (G-CuSum) algorithm for the composite setting (where pre-outage distribution is known and post-outage distribution belongs to a known set of distributions) are used to detect a persistent change in the distribution of a sequence. These tests have optimality properties with respect to popular delay-

false alarm formulations (see e.g., [18], [28], [21]). However, these algorithms are derived for statistical models that do not consider the transient behavior of the system following a line outage. Here, we modify the D-CuSum test proposed in Chapter 4 to derive a line outage detection algorithm that incorporates the transient phenomena.

To measure the delay of the line outage detection algorithms in this chapter, we will be using the delay metric of (5.13).

6.1.1 Generalized Dynamic CuSum Test

Since the statistical model used in this chapter includes an arbitrary number of transient periods with finite duration, each one corresponding to a respective transient distribution induced on the observations, it is clear that the proposed test needs to be designed to take this transient behavior into consideration. Toward this end, we introduce the Generalized Dynamic CuSum (G-D-CuSum) test. This test is derived by using the Dynamic CuSum (D-CuSum) algorithm proposed in Chapter 4 in a generalized test manner; i.e., we calculate a D-CuSum statistic for each possible line outage in parallel, and declare an outage when one of the statistics crosses a pre-determined positive threshold corresponding to the line.

By using the D-CuSum test statistic as a basis, we propose the G-D-CuSum test. The statistic for line ℓ is given as follows:

$$W_\ell^{\mathrm{GD}}[k] = \max\left\{\Omega_\ell^{(1)}[k], \ldots, \Omega_\ell^{(T)}[k], \log\frac{f_\ell^{(0)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}, 0\right\}, \tag{6.1}$$

where

$$\Omega_\ell^{(i)}[k] = \max\{\Omega_\ell^{(i)}[k-1], \Omega_\ell^{(i-1)}[k-1]\} + \log\frac{f_\ell^{(i)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}, \tag{6.2}$$

for $i \in \{1, \ldots, T\}$, $\Omega_\ell^{(0)}[k] \coloneqq 0$ for all $k \in \mathbb{Z}$ and $\Omega_\ell^{(i)}[0] \coloneqq 0$ for all $\ell \in \mathcal{L}$ and all i. The corresponding stopping rule is defined as

$$\tau^{\mathrm{GD}} = \min_{\ell\in\mathcal{L}}\left\{\inf\{k \geq 1 : W_\ell^{\mathrm{GD}}[k] > A_\ell\}\right\}. \tag{6.3}$$

Calculating the test statistic for line ℓ involves calculating the statistics $\Omega_\ell^{(1)}$, $\ldots$,$\Omega_\ell^{(T)}$. The final test statistic is given by taking the maximum of these terms together with the log-likelihood ratio between the distribution at the outage and the pre-outage distribution. Note that to renew each Ω statistic, the value of the

statistic in the previous time instant and the value of the statistic used to detect the previous distribution change are used. The basis of this algorithm is that each statistic is used to capture one of the transient distributions. As a result, at each different period that the process goes through, one of the Ω statistics will dominate the others, leading to the adaptive nature of the algorithm. The test statistics are designed to use prior information from other test statistics, exploiting the fact that distribution changes occur in a sequential manner. It is also important to note that the structure of the algorithm is not affected by the duration of any of the transient periods. This is because the test statistic is calculated by maximizing a log-likelihood ratio over all possible changepoint allocations.

6.1.2 Generalized CuSum Test

The Generalized CuSum (G-CuSum) based test we studied in Chapter 5 was proposed as a line outage detection scheme with the understanding that the transition between pre- and post-outage periods is not characterized by any transient behavior other than the meanshift that occurs at the instant of outage. The meanshift was captured by introducing an additional log-likelihood ratio term between the distribution at the time of change and the distribution before the change. The final test statistic takes the maximum of this log-likelihood ratio and the traditional G-CuSum test recursion.

Although the G-CuSum algorithm does not take any transient dynamics into consideration, it can still perform well when the transient distributions and the final post-change distribution are "similar", i.e., when the KL divergence between $f_\ell^{(i)}$, $i = 1, 2, \ldots, T-1$, and $f_\ell^{(T)}$ is small. As a result, it is useful to compare the performance of the G-CuSum test with the performance of the G-D-CuSum test that is proposed in the present chapter.

For the G-CuSum test in the present setting, we compute L CuSum statistics in parallel, one corresponding to each line outage scenario. The CuSum recursion for line ℓ is calculated by accumulating log-likelihood ratios between $f_\ell^{(T)}$ and f_0. In particular, define the G-CuSum statistic corresponding to line ℓ outage recursively as:

$$W_\ell^{\mathrm{GC}}[k] = \max\left\{W_\ell^{\mathrm{GC}}[k-1] + \log\frac{f_\ell^{(T)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}, \log\frac{f_\ell^{(0)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}, 0\right\}, \tag{6.4}$$

with $W_\ell^{\mathrm{GC}}[0] = 0$ for all $\ell \in \mathcal{L}$.

The goal is to declare an outage as soon as any line is outaged; thus, the algorithm declares a detection the first time any of the line statistics crosses its

corresponding threshold. Accordingly, the stopping time of the test is as follows:

$$\tau^{\mathrm{GC}} = \min_{\ell \in \mathcal{L}} \left\{ \inf\{k \geq 1 : W_\ell^{\mathrm{GC}}[k] > A_\ell\} \right\}, \tag{6.5}$$

with $A_\ell > 0$ being the threshold corresponding to line ℓ.

6.2 Other Algorithms for Line Outage Detection Under Transient Dynamics

In this section, we study two additional change detection algorithms that are of lower complexity than the G-D-CuSum test. First, we review the meanshift test presented in Chapter 5. Next, we present a modified version of the Shewhart test of Chapter 5 that accounts for transient dynamics. We also provide the main idea behind the two algorithms as well as an intuition as to why the Shewhart test might be superior to the meanshift test in the present setting.

6.2.1 Meanshift Test

For this algorithm, a single log-likelihood ratio between the distribution of the observations at the changepoint and before the changepoint is used to detect the outage. Thus, when using the meanshift test, the line outage detection problem is treated as a problem of detecting the meanshift that occurs at the changepoint. In particular, define the meanshift statistic corresponding to line ℓ as follows:

$$W_\ell^{\mathrm{MS}}[k] = \log \frac{f_\ell^{(0)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}. \tag{6.6}$$

Since an outage can occur at any line, a generalized test structure is used; i.e., the decision maker declares a change when one of the L statistics crosses the corresponding threshold, A_ℓ. Consequently, the stopping time for the meanshift test is defined as

$$\tau^{\mathrm{MS}} = \min_{\ell \in \mathcal{L}} \left\{ \inf\{k \geq 1 : W_\ell^{\mathrm{MS}}[k] > A_\ell\} \right\}. \tag{6.7}$$

A justification for this algorithm is given by examining the expected value of the test statistic for line ℓ at the changepoint assuming an outage has occurred in the same line. At the changepoint ($k = \gamma_0$) the expected value of the ℓ statistic is

given by

$$\mathbb{E}_{\ell}^{(0)}\left[\log\frac{f_{\ell}^{(0)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}\right] = D(f_{\ell}^{(0)} \parallel f_0) > 0. \tag{6.8}$$

Since the expected value of (6.8) is positive, it is likely to cross a positive predetermined threshold and an outage is declared.

It is expected that the meanshift test will perform worse than the G-D-CuSum test for the following reasons. First, the transient behavior is not incorporated in the definition of the scheme, in contrast to the proposed G-D-Cusum test. Furthermore, the meanshift test is designed without taking the persistency of the covariance shifts into account and without exploiting past observations. As a result, the log-likelihood ratio used in the test does not match the true distribution of the observations after the time of change. For example, during the first transient period ($\gamma_0 < k \leq \gamma_1$) the expected value of the test statistic is given by

$$\mathbb{E}_{\ell}^{(1)}\left[\log\frac{f_{\ell}^{(0)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}\right] = D(f_{\ell}^{(1)} \parallel f_0) - D(f_{\ell}^{(1)} \parallel f_{\ell}^{(0)}). \tag{6.9}$$

By following the same line of reasoning, it can be argued that for every time instant after the outage occurs, the expected value of the test statistic is given by a difference of two KL divergences. Consequently, (6.9) can even take negative values. In such cases, the algorithm suffers significant delay if a change is not declared at the changepoint, since the test statistic will be negative post-outage with high probability.

6.2.2 Shewhart Test

For the shewhart test under the classic QCD setting, where no transient dynamics are present, a log-likelihood ratio between the persistent post-outage distribution and the pre-outage distribution is used, as seen in Chapter 3. By modifying the structure of the test to account for the meanshift and transient phenomenon after an outage, we derive a test that is better in terms of performance compared to the meanshift test. This is done by introducing an additional log-likelihood ratio term for each transient response period and one for the meanshift. Similar to the meanshift test, the Shewhart test is also a "oneshot" detection scheme; thus, its performance is inferior to our proposed G-D-CuSum test. However, the introduction of the additional log-likelihood ratio terms allows the Shewhart test to have superior performance to that of the meanshift test.

Define the Shewhart test statistic for line ℓ outage as:

$$W_\ell^{\mathrm{SH}}[k] = \max_{i\in\{0,1,\ldots,T\}} \left\{ \log \frac{f_\ell^{(i)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])} \right\}. \tag{6.10}$$

The Shewhart test includes the meanshift log-likelihood ratio of (6.6) along with T additional terms that are associated with different transient periods. From (6.10), it is easy to see that the Shewhart test uses a matching log-likelihood ratio even after the outage. The Shewhart stopping time is defined as:

$$\tau^{\mathrm{SH}} = \min_{\ell\in\mathcal{L}} \left\{ \inf\{k \geq 1 : W_\ell^{\mathrm{SH}}[k] > A_\ell\} \right\}. \tag{6.11}$$

Assuming that an outage occurs in line ℓ, the expected value of the statistic at the changepoint ($k = \gamma_0$) satisfies

$$\begin{aligned} \mathbb{E}_\ell^{(0)}\left[W_\ell^{\mathrm{SH}}[k]\right] &\geq \mathbb{E}_\ell^{(0)}\left[\log \frac{f_\ell^{(0)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}\right] \\ &= D(f_\ell^{(0)} \parallel f_0) > 0. \end{aligned} \tag{6.12}$$

In addition, for some arbitrary transient period i, the expected value of the corresponding test statistic satisfies

$$\begin{aligned} \mathbb{E}_\ell^{(i)}\left[W_\ell^{\mathrm{SH}}[k]\right] &\geq \mathbb{E}_\ell^{(i)}\left[\log \frac{f_\ell^{(i)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}\right] \\ &= D(f_\ell^{(i)} \parallel f_0) > 0. \end{aligned} \tag{6.13}$$

6.3 Line Outage Identification

The detection algorithm proposed in Sec. 6.1.1 can also be used to identify the outaged line. One strategy would be to declare the outaged line as the one corresponding to the largest statistic; i.e., the line that is identified as outaged is given by:

$$\hat{l} = \arg\max_{\ell\in\mathcal{L}} W_\ell[\tau]. \tag{6.14}$$

A drawback to this method of line identification is that the statistics for other lines may also increase following a line outage. Due to the structure of a power system, certain line outages may cause multiple line statistics, in addition to the one corresponding to the true outaged line, to increase. Therefore, in order to reduce the probability of false isolation, a set of lines can be identified as potentially outaged. In this case, more than one line should be checked by the system operator

after an outage is declared.

To this end, we generalize the idea behind (6.14) to account for the case of multiple growing statistics. In particular, after an outage is declared, we create a ranked list containing r entries of line indices, one for each of the large line statistics at the stopping time, and the indices are ordered with respect to the values of these statistics. The idea is similar to list decoding in digital communications (see e.g. [29]). Define the ranked list for an outage in line ℓ as

$$\mathcal{R} = \{\ell_1, \ldots, \ell_r\}, \tag{6.15}$$

where r is the cardinality of the ranked list. Either the cardinality r can be fixed beforehand, or additional constraints can be added to make the size variable across different sample paths (e.g., by imposing an additional constraint that a statistic not only has to be among the largest, but also has to be comparable to the largest one for the corresponding line to belong to the ranked list).

To quantify the performance of our algorithm with respect to its ability to identify the outaged line accurately, we define the probability of false isolation (PFI). For the case of line ℓ outage, a false isolation event occurs when ℓ is not included in the ranked list $\mathcal{R}$. Define the PFI when line ℓ is outaged as:

$$\mathrm{PFI}_\ell(\tau) = \mathbb{P}\{\ell \notin \mathcal{R} | \text{line } \ell \text{ outage}\}. \tag{6.16}$$

The maximum length of the ranked list should be chosen to optimize the tradeoff between PFI and number of lines that need to be checked after an outage detection has occurred. In particular, larger ranked lists lead to lower PFI, but to a larger set of possibly outaged lines to check.

CHAPTER 7

CASE STUDIES

In this chapter, we provide numerical results for the line outage detection algorithms proposed in this work. In Sec. 7.1 we provide simulation results for detection schemes that were proposed in this work for the line outage problem without transients that was studied in Chapter 5. The algorithms are applied on the IEEE 14-bus test system, with the power model of Chapter 5. We show that the proposed G-CuSum based test offers superior performance to that of the other detection algorithms. In Sec. 7.2 we provide simulation results for the line outage detection problem under transient dynamics that was studied in Chapter 6. The detection algorithms that are presented in Chapter 6 are applied on the IEEE 118-bus test system, while assuming transient dynamics after the outage. We demonstrate the superiority of the G-D-CuSum test in terms of delay performance. Finally, we show how this test can be combined with the line outage identification techniques studied in Sec. 6.3 to achieve low probability of false isolation values when identifying outaged lines. For our simulations, we found that the error bounds for all the simulated values are within 5% of the means.

7.1 Simulation Results and Discussion for Chapter 5

In this section, we demonstrate the effectiveness of the proposed G-CuSum based line outage detection algorithm, which was used as a solution for the line outage problem presented in Chapter 5, where no transient behavior is present. We applied the G-CuSum algorithm to the IEEE 14-bus test system for an outage in the line connecting bus 2 and 5, with the thresholds chosen according to (5.16). The entries of $\Delta P[k]$ are sampled from a zero-mean Gaussian p.d.f. with covariance matrix, $\Lambda = \text{diag}(0.5)$. We simulate a line outage at $k = 10$ and the results are shown in Fig. 7.1. The $W^{\text{GC}}_{(2,5)}[k]$ statistic (blue) crosses the threshold of $A^{\text{GC}}_{(2,5)} = 100$ at $\tau^{\text{GC}} = 57$, resulting in a detection delay of 47 samples. As expected, all of the other CuSum statistics (red) either remain close to zero, or increase at a slower rate.

Next, we perform Monte Carlo simulations for the Shewhart, meanshift, and G-CuSum algorithms to obtain plots of average detection delay versus mean time to false alarm. The values for the average detection delay are obtained by simulating an outage in the line connecting buses 4 and 5 and running the corresponding detection algorithms for different thresholds until a detection of the outaged line is declared. For computing the mean time to false alarm, the detection algorithms are executed for the power system under normal operation until a false alarm occurs. Since false alarm events are in general rare, averaging many sample runs would incur significant computation time. In order to reduce the simulation time, importance sampling is used for the meanshift and Shewhart tests.

Figure 7.2 shows the average detection delay versus mean time to false alarm for all of the detection methods mentioned in this work. We note that the meanshift test performs better than the Shewhart test. It can be verified from QCD theory that the slope of delay versus log(mean time to false alarm) for the Shewhart and CuSum tests is given by $\frac{1}{D(f^{(1)}_{(m,n)} \parallel f_0)}$ for large mean time to false alarm [6].

From the plots, we conclude that for the same value of mean time to false alarm, both CuSum-based algorithms have a much lower average detection delay than the Shewhart and meanshift algorithms. In addition, the figure shows that when we use varied thresholds for the G-CuSum test (red lines) as opposed to a fixed threshold (green lines), even lower detection delay can be achieved for the same mean time to false alarm. This illustrates that our KL threshold G-CuSum based algorithm is an improvement over that of [3]. It should be noted that simulation results demonstrate that the detection delay scales exponentially with the selected thresholds for both the meanshift and Shewhart tests, and linearly for the CuSum-based tests.

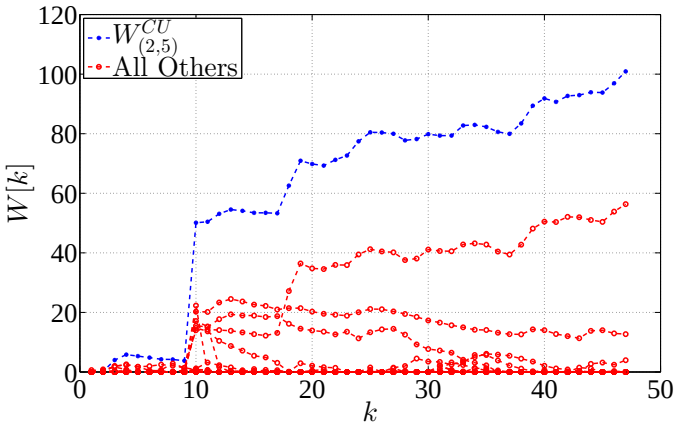


Figure 7.1: Example of a run of the G-CuSum for the 14-bus system.

7.2 Simulation Results and Discussion for Chapter 6

In this section, the algorithm proposed in (6.1)-(6.3) is applied to the IEEE 118-bus test system (for the model data, see [30]). In order to compute the transient dynamics following a line outage, we use the simulation tool Power System Toolbox (PST) [31]. For simplicity, we used the statistical model in (5.11) with $T = 2$; i.e., we assumed one transient period, with a duration of 100 samples, after the line outage occurs. Additional transient periods could easily be incorporated into the simulations. The power injection profiles at the load buses are assumed to be independent Gaussian random variables with variance of 0.03 and the PMU sampling rate is assumed to be 30 measurements per second.

7.2.1 Line Statistic Evolution

First, we simulate two different line outages, one in which the detection occurs during the transient period and one in which the detection occurs after the transient period; the results are shown in Fig. 7.3. Figure 7.3(a) shows some typical progressions of $W_{180}[k]$s for the various line outage detection schemes discussed earlier. The fault occurs at $k = 10$ and the threshold is $A_\ell = 120$. From the figure, we conclude that for this sample run, a line outage is declared after 50 samples when the G-D-CuSum stream crosses the threshold first. The other algorithms incur a much larger detection delay since they do not cross the threshold of $A_\ell = 120$. Figure 7.3(b) shows the typical progressions of $W_{32}[k]$s for an outage in line 32. For a threshold of $A_\ell = 125$, the G-D-CuSum detects a line outage 156 samples after the outage occurs. In this example, the detection occurs after the transient dynamics have subsided. From the plots, we conclude that even though detection takes place after the transient dynamics subside (at $k = 110$), the G-D-CuSum algorithm still has a smaller detection delay than the G-CuSum algorithm. Note that the G-CuSum statistic does not grow during the transient period, which

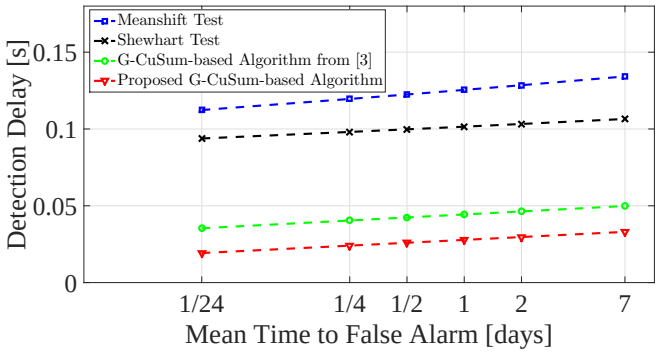
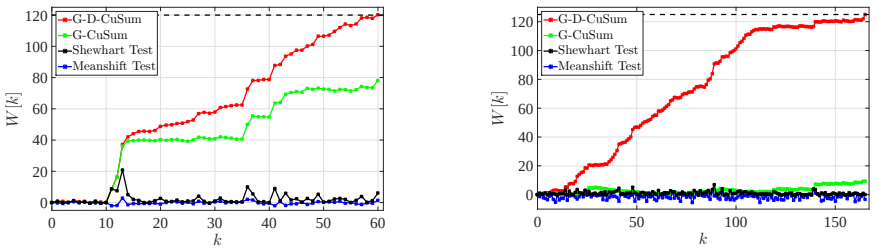
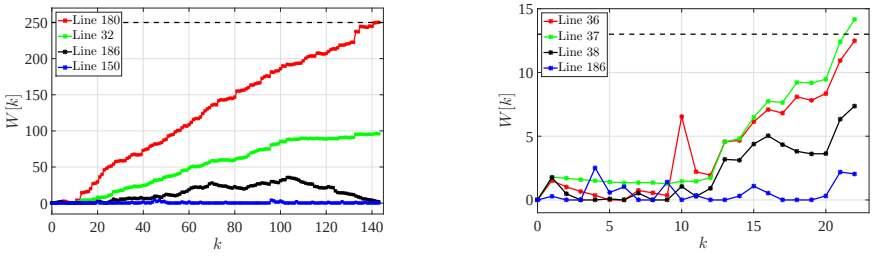


Figure 7.2: Detection delay vs. mean time to false alarm.



(a) Line 180 outage with A_l=120. (b) Line 32 outage with A_l=125.

Figure 7.3: Sample paths of different algorithms for IEEE 118-bus system.



(a) Line 180 outage with A_l=250. (b) Line 36 outage with A_l=13.

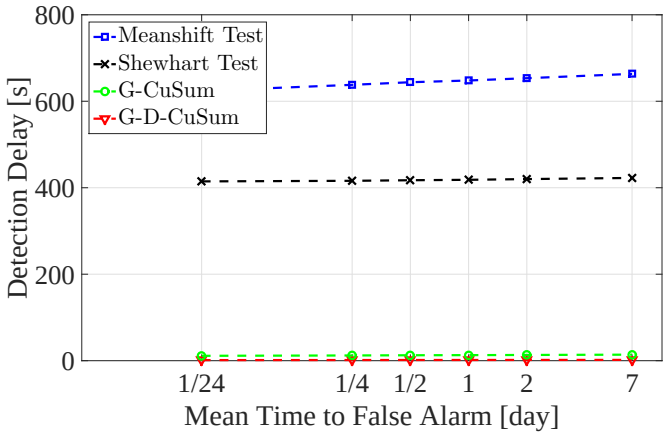
Figure 7.4: Sample paths of the G-D-CuSum algorithm for IEEE 118-bus system.

results in the G-CuSum having a large delay. Through close inspection, we also notice that the slopes of the G-D-CuSum and G-CuSum statistics are identical after the transient period is over, something that can be verified by the theory.

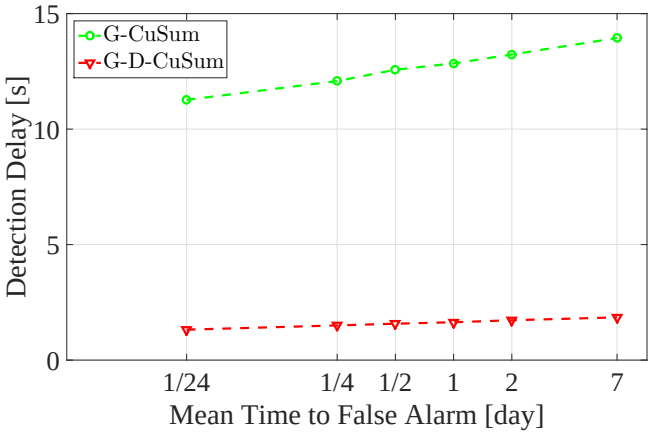
Next, we simulate two different line outages and demonstrate the evolution of the G-D-CuSum statistic for different lines in Fig. 7.4. For an outage in line 180, it is seen in Fig. 7.4(a) that $W_{32}[k]$ grows faster that other line statistics. An outage is declared after 135 samples, when $W_{32}[k]$ crosses a threshold of $A_\ell = 250$. In Fig. 7.4(b), we show an example of a misdetection event. In particular, it can be noted that, for an outage in line 36, other line statistics can sometimes cross the test threshold before $W_{36}[k]$. In particular, in Fig. 7.4(b) we see that $W_{37}[k]$ crosses a threshold of $A_\ell = 13$, thus a misdetection event occurs. This can be seen as a justification of the use of a ranked list to identify outaged lines.

7.2.2 Delay Performance

We performed Monte Carlo simulations for outages in lines 36, 180, and 104, and show detection delay versus mean time to false alarm results for all the detection schemes presented in this work. We compared the performance of our proposed



(a) Detection delay vs. mean time to false alarm for different algorithms.



(b) Detection delay vs. mean time to false alarm for G-D-CuSum and G-CuSum algorithms.

Figure 7.5: IEEE 118-bus Monte Carlo simulation results for an outage in line 36.

algorithm against the Meanshift test, Shewhart test, and the G-CuSum algorithm for an outage in line 36, which is the line outage case that corresponds to the largest delay. The results are shown in Fig. 7.5. From Fig. 7.5(a), we conclude that the G-D-CuSum algorithm achieves the lowest detection delays among all algorithms for a given mean time to false alarm. The performances of the Meanshift and Shewhart tests are considerably worse than those of the G-CuSum and G-D-CuSum tests. In Fig. 7.5(b) we demonstrate the performance gain achieved when using the G-D-CuSum test. For an outage in line 36, the G-D-CuSum test achieves more than an order of magnitude less delay for given false alarm rate. Next, we evaluate the performance of the proposed algorithm for different line outage cases. Among

all the lines of the system, detection delay for line 104 is the lowest for a fixed mean time to false alarm while line 36 has the worst detection delay. Line 180 was chosen as a representative line for intermediate delay values. The results are shown in Fig. 7.6.

7.2.3 Probability of False Isolation

Finally, the PFI versus mean time to false alarm is obtained for outages in lines 36, 104, and 180; the results are recorded in Tables 7.1, 7.2, and 7.3. The PFI was calculated by using the ranked list method discussed in Section 6.3 for a ranked list of fixed length 1, 3, and 5. In Table 7.1 we demonstrate the PFI results for a ranked lists of length 1. This is equal to identification using 6.14, i.e., identifying as outaged the line with the highest statistic at the stopping time. We note that, although some outages can be handled efficiently with this simple technique (e.g. outage in line 180 and 104), some line outages may lead to large PFI values, e.g. an outage in line 36. This happens due to the fact that many line statistics other than the one corresponding to the outaged line grow post-outage, as was discussed in Section 6.3. In Table 7.2 we demonstrate the PFI values for a ranked list of length 3. The PFI is significantly reduced for line 36. Finally, in Table 7.3 the PFI results for a ranked list of length 5 are shown. In this case, the PFI for line 36 is below 5%. Note that the PFI decreases as the mean time to false alarm increases. This is because larger mean time to false alarm corresponds to larger thresholds, which result in smaller PFI values.

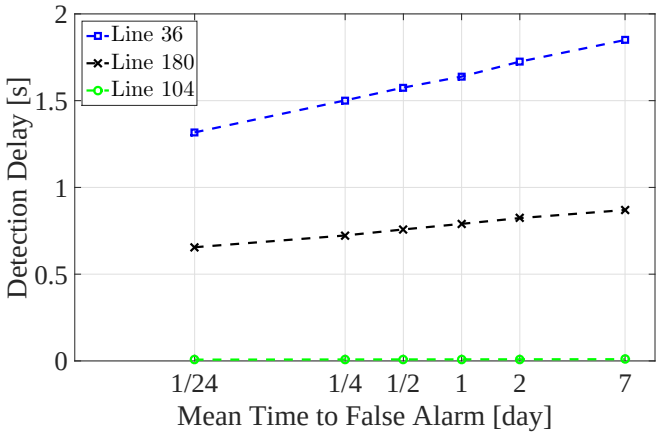


Figure 7.6: 118-bus: Detection delay vs. mean time to false alarm for different line outages.

Table 7.1: Probability of false isolation for IEEE 118-bus system simulated with a ranked list of length of 1.

$\mathbb{E}_\infty[\tau]$ [day]	1/24	1/4	1/2	1	2	7
Line 36	0.5999	0.5472	0.5266	0.5182	0.5102	0.4911
Line 180	0.0344	0.0237	0.0186	0.0160	0.0148	0.0091
Line 104	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$

Table 7.2: Probability of false isolation for IEEE 118-bus system simulated with a ranked list of length of 3.

$\mathbb{E}_\infty[\tau]$ [day]	1/24	1/4	1/2	1	2	7
Line 36	0.1477	0.1392	0.1371	0.1303	0.1251	0.1219
Line 180	0.0197	0.0127	0.0112	0.0097	0.0062	0.0049
Line 104	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$

Table 7.3: Probability of false isolation for IEEE 118-bus system simulated with a ranked list of length of 5.

$\mathbb{E}_\infty[\tau]$ [day]	1/24	1/4	1/2	1	2	7
Line 36	0.0295	0.0203	0.0158	0.0137	0.0108	0.0067
Line 180	0.0097	0.0074	0.0052	0.0041	0.0036	0.0027
Line 104	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$

CHAPTER 8

CONCLUSIONS

PMUs provide synchronized real-time measurements of voltage and current phasors across the power system. These measurements are used in a sequential manner to detect and identify line outage events.

In this work, we demonstrate the use of theoretical tools from QCD theory to tackle the line outage detection and identification problem in an efficient and robust manner. We first study the line outage problem without considering transient behavior. A G-CuSum based algorithm is proposed as a solution. Furthermore, different techniques for choosing the test thresholds are discussed. In particular, we exploit the notion of KL-divergence to choose the thresholds as a function of the difference between the pre- and post-change distributions. This choice leads to an algorithm that detects outages faster than other detection schemes proposed in the literature. Finally, we study the problem of line outage detection under transient dynamics. We introduce the G-D-CuSum algorithm, a scheme that is designed after taking the transient behavior into consideration. We use this test to introduce line outage identification techniques that can be easily implemented in practice. Through numerical results and simulations, we show that this algorithm is superior compared to other algorithms in the literature.

Future work in this area includes developing schemes to optimally place limited PMUs to maximize algorithmic performance in terms of detection delay and probability of false isolation, as well as implementing low complexity solutions for detecting multiple and cascading line outages. In addition, a more accurate transient model for power system line outages should be developed so that more appropriate detection schemes that are model specific could be used. Finally, distributed versions of the detection schemes presented in this thesis can be explored.

REFERENCES

[1] FERC and NERC, "Arizona - Southern California outages on September 8, 2011: Causes and recommendations," Apr. 2012. [Online]. Available: http://www.ferc.gov.

[2] U.S.-Canada Power System Outage Task Force, "Final report on the August 14th blackout in the United States and Canada: causes and recommendations," Apr. 2004. [Online]. Available: http://energy.gov.

[3] Y. C. Chen, T. Banerjee, A. D. Domínguez-García, and V. V. Veeravalli, "Quickest line outage detection and identification," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 749–758, 2016.

[4] H. V. Poor and O. Hadjiliadis, *Quickest Detection*. Cambridge University Press, 2009.

[5] A. G. Tartakovsky, I. V. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Change-Point Detection*, ser. Statistics. CRC Press, 2014.

[6] V. V. Veeravalli and T. Banerjee, *Quickest Change Detection*. Elsevier: E-reference Signal Processing, 2013.

[7] K. A. Clements and P. W. Davis, "Detection and identification of topology errors in electric power systems," *IEEE Transactions on Power Systems*, vol. 3, no. 4, pp. 1748–1753, 1988.

[8] F. F. Wu and W. E. Liu, "Detection of topology errors by state estimation," *IEEE Transactions on Power Systems*, vol. 4, no. 1, pp. 176–183, 1989.

[9] N. Singh and H. Glavitsch, "Detection and identification of topological errors in online power system analysis," *IEEE Transactions on Power Systems*, vol. 6, no. 1, pp. 324–331, 1991.

[10] J. E. Tate and T. J. Overbye, "Line outage detection using phasor angle measurements," *IEEE Transactions on Power Systems*, vol. 23, no. 4, pp. 1644–1652, 2008.

[11] H. Zhu and G. B. Giannakis, "Sparse overcomplete representations for efficient identification of power line outages," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2215–2224, 2012.

[12] G. Feng and A. Abur, "Identification of faults using sparse optimization," in *Proc. of Communication, Control, and Computing (Allerton Conference)*, Sept 2014, pp. 1040–1045.

[13] M. Garcia, T. Catanach, S. V. Wiel, R. Bent, and E. Lawrence, "Line outage localization using phasor measurement data in transient state," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 3019–3027, July 2016.

[14] A. Wald and J. Wolfowitz, "Optimum character of the sequential probability ratio test," *Ann. Math. Statist.*, vol. 19, no. 3, pp. 326–339, 09 1948. [Online]. Available: http://dx.doi.org/10.1214/aoms/1177730197

[15] A. Wald, *Sequential Analysis*. Dover Publications, Inc., 1973.

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.

[17] G. Lorden, "Procedures for reacting to a change in distribution," *Ann. Math. Statist.*, vol. 42, no. 6, pp. 1897–1908, 12 1971. [Online]. Available: http://dx.doi.org/10.1214/aoms/1177693055

[18] M. Pollak, "Optimal detection of a change in distribution," *Annals of Statistics*, vol. 13, no. 1, pp. 206–227, Mar. 1985.

[19] W. Shewhart, *Economic Control of Quality of Manufactured Product*. D. Van Nostrand Company, Inc., 1931.

[20] E. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1-2, pp. 100–115, 1954. [Online]. Available: http://biomet.oxfordjournals.org/content/41/1-2/100.short

[21] G. V. Moustakides, "Optimal stopping times for detecting changes in distributions," *Ann. Statist.*, vol. 14, no. 4, pp. 1379–1387, 12 1986. [Online]. Available: http://dx.doi.org/10.1214/aos/1176350164

[22] Y. Ritov, "Decision theoretic optimality of the cusum procedure," *Ann. Statist.*, vol. 18, no. 3, pp. 1464–1469, 09 1990. [Online]. Available: http://dx.doi.org/10.1214/aos/1176347761

[23] A. R. Bergen and V. Vittal, *Power Systems Analysis*. Prentice Hall, 2000.

[24] M. Lotfalian, R. Schlueter, D. Idizior, P. Rusche, S. Tedeschi, L. Shu, and A. Yazdankhah, "Inertial, governor, and AGC/economic dispatch load flow simulations of loss of generation contingencies," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-104, no. 11, pp. 3020 –3028, Nov. 1985.

[25] B. Hajek, *Random Processes for Engineers*. Cambridge University Press, 2015.

[26] G. Fellouris and G. Sokolov, "Second-order asymptotic optimality in multisensor sequential change detection," October 2014, http://arxiv.org/abs/1410.3815v2. [Online]. Available: http://arxiv.org/abs/1410.3815v2

[27] A. G. Tartakovsky and A. S. Polunchenko, "Quickest changepoint detection in distributed multisensor systems under unknown parameters," in *Proc. of IEEE International Conference on Information Fusion*, July 2008, pp. 1–8.

[28] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2917–2929, Nov. 1998.

[29] P. Elias, "Error-correcting codes for list decoding," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 5–12, Jan 1991.

[30] "Power system test case archive," Oct. 2012. [Online]. Available: http://www.ee.washington.edu/research/pstca.

[31] J. Chow and K. Cheung, "A toolbox for power system dynamics and control engineering education and research," *IEEE Transactions on Power Systems*, vol. 7, no. 4, pp. 1559–1564, Nov. 1992.