Detection and localization of breaks in railroad tracks using time domain reflectometry

by

Nathaniel P. Diehl

B.S., Kansas State University, 2024

A THESIS

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Mike Wiegers Department of Electrical and Computer Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2025

Approved by:

Major Professor
Balasubramaniam Natarajan

Copyright

Abstract

A new technique is proposed for detecting and locating breaks in railroad tracks using time domain reflectometry that provides real time measurement of the track condition while requiring minimal additions to existing railroad hardware. Time domain reflectometry is a technique that measures reflection delay of electrical pulses to determine the location of impedance changes. The feasibility of the technique is evaluated using an experimental setup with a real-world railroad track as well as a digital twin model. The experimental platform includes equipment designed to transmit and detect signals along the rail, which are used to measure the reflection delay of signals on a real-world railroad track. The simulation done with a digital twin observes the change in reflection delay caused by a break in the rail. Monte Carlo simulation is used to augment the observations from the digital twin by adding random noise. These noisy observations are used to assess the effectiveness of the break detection and break location estimation methods. Results demonstrate that time domain reflectometry is a viable approach for detecting and locating breaks in railroad tracks. Future work could analyze this technique under a nonconstant propagation velocity and on a wider range of defect types.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

TDR	Time Domain Reflectometry
SNR	Signal-to-Noise Ratio
PCB	Printed Circuit Board
ROC	Receiver Operating Characteristic
DAC	Digital to Analog Converter
ADC	Analog to Digital Converter
RFSoC	Radio Frequency System-on-Chip
QQ	Quantile-Quantile
CRLB	Cramer-Rao Lower Bound
i.i.d.	Independently and Identically Distributed

Chapter 1

Introduction

1.1 Motivation

Railroads form a critical component of modern transportation infrastructure. In 2024 alone, railroads transported 1.53 billion tons of cargo valued at $584 billion across the United States [1]. The reliability of the railroads depends heavily on the condition of the track, which must withstand constant mechanical stress from heavy loads and continuous operation. Maintaining reliable railroad tracks is essential for maintaining safe and efficient transportation.

Track usage inevitably diminishes the condition of the rails resulting in the development of defects. Rail defects can arise from a variety of sources, including material fatigue, temperature fluctuations, improper installation, and repeated stress from heavy loads [2]. Over time, small imperfections such as cracks, corrosion, or misalignments can grow into significant defects. The growth of these defects can be accelerated by environmental conditions such as moisture and seasonal temperature variation.

Subtle rail defects cause additional wear and tear on the train cars and other railroad property any time a train passes over them. This additional wear causes maintenance and repairs to become more frequent and extensive. When rail defects are severe enough, they can cause train derailments [3] which can be incredibly destructive. Derailment causes economic damage to the railroad in the form of lost cargo and damage to railroad property like the train engine, cars, and track [4]. Derailments put the health and lives of railroad operators and bystanders at risk. When the cargo of the trains is hazardous, derailments can pollute the area and put nearby communities at risk [5]. Just such an accident happened on February 3, 2023 in East Palestine, Ohio when the derailment of a Norfolk Southern train cause hazardous materials to ignite resulting in the evacuation of nearly 2000 residents.

Train derailments in the United States are currently decreasing [6], but any improvement in rail safety is in the interest of both the railroad and the public. To mitigate these risks, railroads are highly motivated to identify when a break has occurred and get the defect repaired as soon as possible. Railroads rely on detection methods to identify when and where a defect occurs, so repair teams can be sent to perform the required maintenance.

1.2 Related Work

The oldest method of defect detection was human inspection, where a trained inspector walked the length of the rail and physically examined the tracks for visible signs of damage. This is a labor intensive, time inefficient, and error prone task. To improve this process, several automated methods have been developed.

The first class of methods are those that use inspection vehicles [7, 8] where a robot equipped with a sensor drives along the rail and scans for defects. Vehicles have been equipped with a large variety of sensors, including ultrasonic, electromagnetic, image processing, and thermal imaging, to detect a range of different defects. These sensing methods are able to give an impressive amount of detail. Providing estimations of the defect's properties like size, shape and orientation that can be used to assess the severity of the defect. Some methods are even able to detect internal flaws that would be invisible to surface scanning methods. Additionally inspection vehicles are able to give an accurate location of the defect using a mounted GPS.

All inspection vehicle methods share similar inherent downsides. While faster than a human inspector, it still takes time for the vehicles to be set up and test a track section. The result is routine testing where each section of track is only inspected once every few months. During the test itself the inspection vehicle must occupy the rail which means no trains can use it while the test is being run. Also the moving components of the mechanical device are prone to failure.

While unable to give the same amount of details as the inspection vehicle, there are various methods that have been designed by permanently placing sensors on the rails themselves. These methods have eliminated the need for human operators to configure the device before each test. It also allows the operators to run inspections remotely. Several different strategies of detection have been used for these permanent sensor methods. The vibration method [9] monitors how the track vibrations induced by a passing train change when a defect is present. The acoustic emission method [10, 11] monitors transient elastic waves released when a defect occurs. Both of these methods remove the need to have testing rotations, instead they provide a new measurement upon an instigating event.

These permanent sensor methods solve most of the problems of inspection vehicles, but due to large amounts of acoustic noise introduced in rail road environments they operate with low signal to noise ratios. As a result, these methods rely on complex noise suppression methods to extract useful information from the data. These permanent sensor methods also require an upfront cost from the railroads since the new sensing equipment will need to be purchased and installed on each section of rail.

Another permanent sensor method uses DC voltages to detect breaks in rails. This method sets a voltage at one end of the rail and then measures the voltage at the other end. It there is no break in the rail, the rail acts as a conductor and the voltage at both ends should be approximately the same. If there is a break in the rail, the rail acts as an open circuit, which means the ends will have different voltages. This method is able to provide real time measurements of the status of the rail, but it is only able to detect complete breaks in the rails and is unable to provide any information about the location of the break. This thesis's proposed method of defect detection, which will be described in Section 1.3, can be thought of as an improvement to this DC method.

1.3 Break Detection with TDR

Transmission lines provide a medium for electromagnetic waves to travel. Unlike in traditional circuit theory where all voltage changes are assumed to travel instantaneously, when working with transmission lines, a change in voltage at one point will create a wave that propagates through the transmission line at a fixed velocity [12].

A transmission line (shown in Figure 1.1) has two defining properties, the propagation time it takes a wave to travel from one end of the line to the other, and the characteristic impedance. When the wave reaches the end of the transmission line it will encounter a load element or termination. If the impedance of the load element is the same as that of the transmission line, then the signal will be absorbed by the element. In this case the load is said to be matched [13].

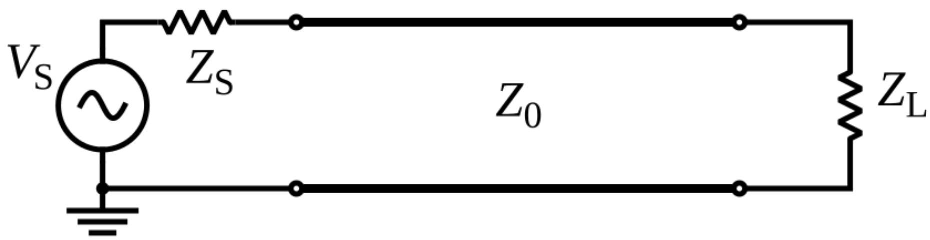

Figure 1.1: *Circuit Diagram of a Transmission Line*
V_S: source voltage, Z_S: source impedance, Z_0: characteristic impedance, Z_L: load impedance

If the impedance at the end of the rail is different from that of the transmission line, there is an impedance mismatch. This mismatch causes part of the wave to be reflected back towards the start of the line. The size of the reflected signal equals the size of the original wave multiplied by a quantity called the reflection coefficient Γ [13] corresponding to,

$$\Gamma = \frac{Z_{new} - Z_{old}}{Z_{new} + Z_{old}} \tag{1.1}$$

here Z_{new} is the load impedance and Z_{old} is the impedance at the start of the transmission line, but the equation is the same for any change in impedance [14].

When the load impedance is much smaller than the transmission line impedance, the termination acts like a short circuit and reflects an inverted version of the wave. If the load impedance is larger, it acts as an open circuit and reflects a large positive version of the wave.

The propagation velocity v_p of the wave is a fraction of the speed of light. When the permeability and permittivity of the transmission line's dielectric are equivalent to air, the propagation velocity is the speed of light. When the permeability and permittivity are different the propagation velocity will be slower. This relationship can be represented by the velocity factor $VF \in [0, 1]$ where $VF = 1$ means the propagation velocity equals the speed of light. The velocity factor VF can be determined from the dielectric properties of the material [15], i.e.

$$VF = \frac{1}{\sqrt{\mu_r \varepsilon_r}} \approx \frac{1}{\sqrt{\varepsilon_r}} \tag{1.2}$$

where μ_r is the relative permeability and ε_r is the relative permittivity.

If the velocity factor is known, then the time between when a signal is transmitted and when the reflection is received can be converted into a measurement of the distance to that change in impedance. A signal would be transmitted, travel down the line, reflect off the impedance change, travel that distance a second time, and be received by the transmitter. This means the transmit to receive time (the reflection period T) is the time it takes the signal to travel twice the distance to the impedance mismatch, i.e.

$$T = \frac{2l}{v_p} = \frac{2l}{VF \cdot c} \tag{1.3}$$

where v_p is the propagation velocity, VF is the velocity factor, c is the speed of light, and l is the length of the transmission line or equivalently the distance to the impedance mismatch. This means that the distance to a point of impedance mismatch corresponds to,

$$l = \frac{T \cdot v_p}{2} = \frac{T \cdot VF \cdot c}{2} \tag{1.4}$$

The process of converting the reflection period to distance is called time domain reflectometry. Time domain reflectometry has been used for fault detection in a variety of applications such as automotive [16], avionics [17], fiberoptic [18], and even in complex networks with multiple branches [19]. This thesis proposes a new application of time domain reflectometry to detect breaks in railroad tracks.

If the railroad track were treated as a transmission line as shown in Figure 1.2, then the rail would be the wire that the signal travels on. If the signal reaches the end of the rail, location l_r in Figure 1.2, it will encounter an impedance mismatch and reflect back to the start of the rail. If there were a break in the rail, the signal would encounter the break before the end of the rail and the impedance mismatch from the break would cause the signal to reflect sooner. The time that earlier reflection arrives T_b can be used to calculate where the break is located on the rail l_b.

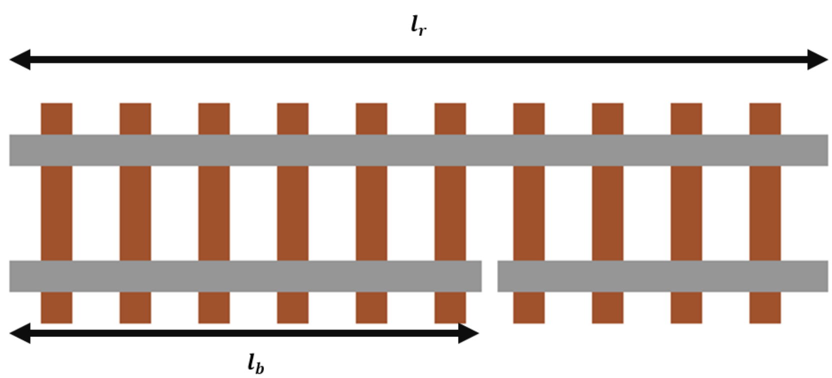

Figure 1.2: *Illustration of a Track with a Break*

1.4 Contributions

This thesis proposes a new method for non-intrusively detecting and locating breaks in railroad tracks using time domain reflectometry. This method will be able to

- provide real-time measurements of the condition of the track without the need to wait for an instigating event

- utilize existing railroad hardware with only minimal additional measurement equipment

The first major contribution of this thesis is described in Chapter 2, which aims to test the feasibility of using time domain reflectometry on railroad tracks. To accomplish this real-world experiments were performed on a section of track provided by BNSF Railway. These experiments measured the signal behavior as it travels on a non-defected railroad track. Specifically, the contributions of Chapter 2 include:

- Detecting the end of the rail using reflectometry
- Conducting an experiment to estimate the bandwidth of the rail
- Estimating the propagation velocity of a signal as it travels on the rail
- Determining the signal suited for use in the time domain reflectometry method.

The work described in Chapter 3 aims to analyze the proposed method on rails with breaks. The results from Chapter 2 where used to inform the creation of a digital twin [20] of a railroad track in the software COMSOL Multiphysics [21]. The digital twin could be programed with defects at arbitrary locations. When a simulated input signal is launched on the rail, COMSOL simulates the reflections a real-world rail would produce. Iterating through multiple defect locations, a dataset is created that could be used in the analysis of the detection and estimation aspects of the proposed method.

Chapter 3 continues by deriving the theoretical performance of the detection and estimation methods. The theoretical performance is compared with the results of running the detection and estimation method implementations on data from the digital twin. Specifically, the contributions of Chapter 3 include:

- Designing a digital twin informed by the real-world rail
- Proposing a rail break detection method
- Proposing a rail break location estimation method

- Theoretical and simulation based analysis of the implemented detection and estimation methods

Chapter 2

Field Testing

In this chapter, the feasibility of the TDR method is evaluated on a test track. Section 2.1 details the testing environment and equipment set. Sections 2.2 describe the system response of the measurement equipment and rail. Sections 2.3 discusses the signal design and the limiting factors that constrain the choice of signal. The impact of noise on the observations and the ensemble averaging process are described in Section 2.4. The chapter concludes with the experimental results in Section 2.5 showing a clear pulse where a reflection occurs from the end of the rail.

2.1 Test Set Up

BNSF Railway provided Kansas State with a section of railroad tracks to use in testing (see Figure 2.1). The tracks are composed of two $480\,in$ long steel rails placed $56.375\,in$ apart. Under the rails are 24 wooden support beams spaced $9\,in$ in apart. A ballast of crushed gravel, limestone, basalt and granite is poured to fill the gaps between the wooden support beams. The tracks were set up in a field behind Kansas State's Civil Infrastructure Systems Laboratory (CISL).

The measurement equipment is located inside the CISL building to shield them from outdoor elements. A $59.74\,m$ cable travels from the output of the measurement equipment,

Figure 2.1: *Photo of Track*

underneath the facilities garage door, across the field to a connection point on the rail. The connection point is 21 *in* off from the start of the rail which makes the distance from the end of the cable to the end of the rail 459 *in*. A photo of the cable connection point is shown in Figure 2.2.

Figure 2.2: *Photo of Cable Connection Point on Track*

Signal generation and measurement is done using the Real Digital RFSoC 4x2 board [22]. The RFSoC 4x2 is a powerful platform that can be used for a large variety of RF applications, but for this project it is essentially being used as a function generator and oscilloscope. The RFSoC is designed to support the PYNQ framework, which allows the board to be controlled by python software running from Jupyter Notebooks. This framework makes developing and running tests easy to implement. A photo of the Measurement Equipment is shown in Figure 2.3.

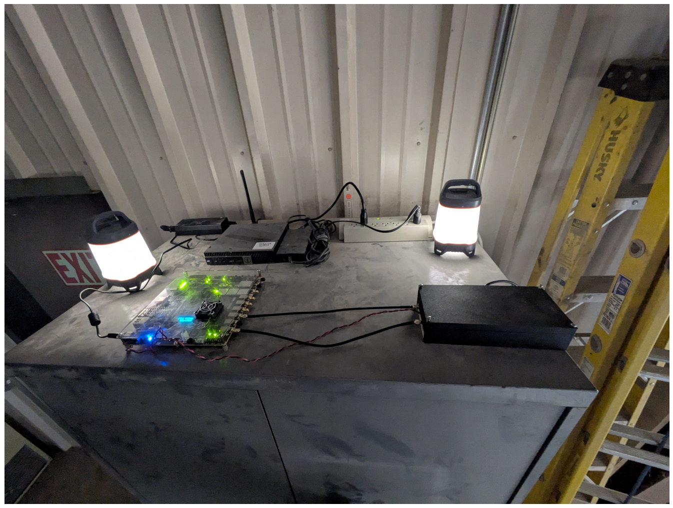

Figure 2.3: *Photo of Measurement Equipment*

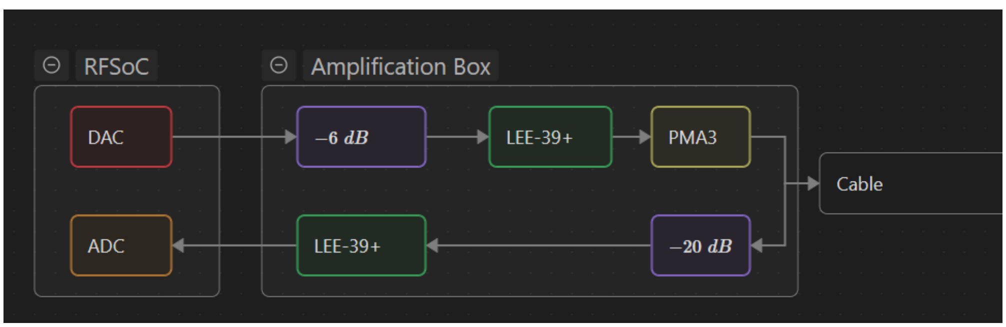

Figure 2.4: *Amplification Flowchart*

In our configuration, the DAC and ADC are running at 4 giga samples per second. This gives the measurements a generous $2GHz$ of bandwidth. The biggest limitation the measurement equipment has is the amplitude of the DAC and ADC. Both are designed to operate with voltages between $\pm 1V$. Given the amount of signal energy expected to be lost to the environment, such small signals would not work for this application.

To resolve the signal amplitude problem, an amplifier system was designed to boost the signal to usable levels. To maintain the large bandwidth of the RFSoC, the decision was made to use high frequency amplifiers. Specifically the Lee-39+ [23] and PMA3-43-1w+ [24]

were chosen. PCBs were designed for each amplifier, and they were arranged as seen in Figure 2.4.

Name	Bandwidth
PMA3-43-1w+	10 MHz to 4 GHz
Lee-39+	8 GHz

Table 2.1: *Table of Amplifier Part Bandwidths*

The amplifiers have broad bandwidths as seen in table 2.1. Both parts ensured that the bandwidth of the signal would not be limited on the high end. The PMA3 however was not a baseband amplifier, being unable to generate frequencies below 10 MHz. This design decision committed us to using a modulated input signal to avoid unwanted distortion. This will be discussed more in the section on signal design. An illustration of the entire field testing set up is shown in Figure 2.5.

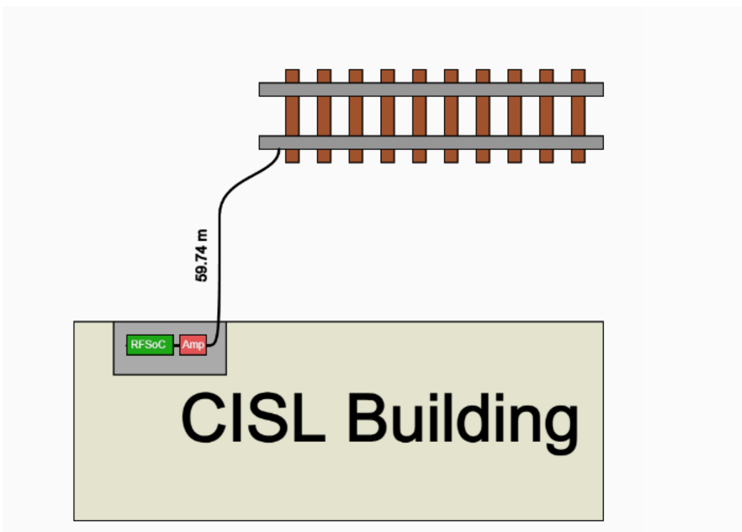

Figure 2.5: *Illustration of Field Testing Set Up*

2.2 System Response

When a signal travels through the system, there are three notable components that appear in the observations. An ideal illustration of these components is shown in Figure 2.6. The first component is the initial signal created by the DAC. This component has the largest amplitude and is relatively undistorted. It is used to define the start time of $t = 0$.

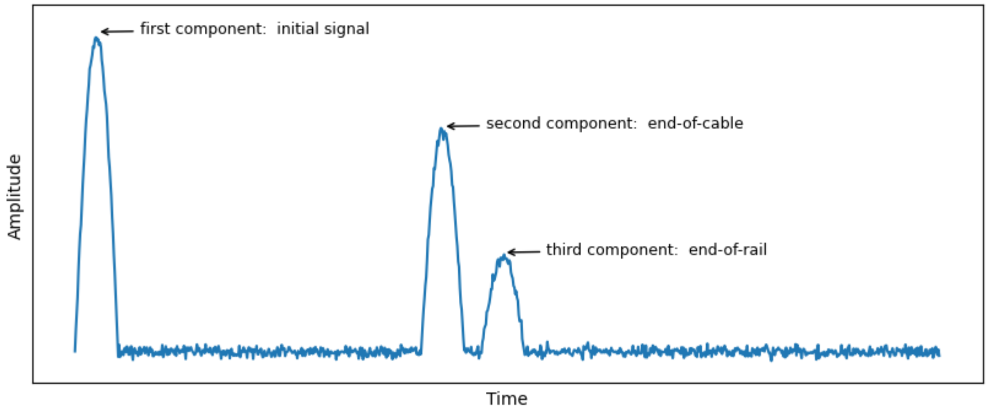

Figure 2.6: *Illustration of System Response*

The second component is the reflection from the end of the cable. This reflection is caused by the impedance change from the cable to rail. It is the second largest component in the observations. Because the impedance of a coax cable and that of an uninsulated steel rail are very different, a large amount of the signal energy gets reflected here and will never enter the rail.

The third component is the reflection of the end of the rail. This reflection is much smaller than the reflection from the end of the cable, but the most significant for detecting a break. If this pulse occurs earlier then the end of the rail, that indicates the presents of a break, and the exact location of the pulse corresponds to the break's location in the rail.

The are other reflections that occur in the response from marginal impedance mismatches and reflections of other pulses. These other reflections are much smaller or occur well after the notable pulses. Knowing the locations of these small pulses doesn't provide any particularly helpful information, so these pulses will be ignored or treated as noise.

2.3 Signal Design

When the transmitted signal returns to the receiver, it is passed into a matched filter to distinguish the signal from random noise. The matched filter correlates the received observations $y(t)$ with a copy of the transmitted signal $s(t)$ to generate a correlation function $R_{YS}(\tau)$ by shifting the delay on the transmitted signal and integrating over time.

$$R_{YS}(\tau) = \int y(t)s(t+\tau)dt \tag{2.1}$$

Points in time where the observations look like the transmitted signal will have high correlation and points in time where the observations look like random noise will have almost no correlation. When the received signal in the observations is exactly lined up with the transmitted signal there will be a peak in the correlation function. When the received signal is only partially lined up with the transmitted signal there will still be correlation, but it won't be as strong as the peak. These points of partial correlation are called sidelobes. The effectiveness of the matched filter is determined by how well it can separate the signal from the noise and distinguish the correlation peaks from the sidelobes. A visual depiction of this relationship is shown in the Figure 2.7, where $s(t)$ is a waveform corresponding to the 5-bit barker code [25].

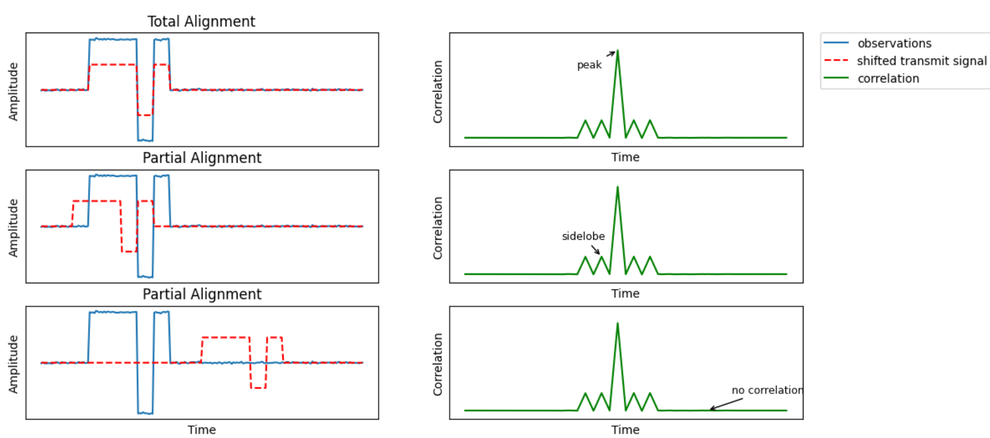

Figure 2.7: *Illustration of Correlation Alignment*

In general the effectiveness of the matched filter detection of a reflected pulse is controlled by the choice of the transmitted signal. Barker codes were selected as the transmission signal. The autocorrelation functions of barker codes have high peak-to-sidelobe ratios, provide low sidelobe energy, and distribute the sidelobe energy uniformly [25]. These features make any barker code a good candidate for a correlator detector (i.e. matched filter) but choosing the best waveform for this application requires real measurements from the rail and design choices based on the hardware capabilities.

The choice of input signal used in the field testing was determined by several limiting factors that affect the length of the signal in time as well as the minimum and maximum frequencies of the signal's power spectrum. An optimum signal was selected based on these limiting factors to provide the best measurement of the end of the rail.

The first restriction on the input signal is the length of signal that can be used. Due to the reflections off the end-of-cable and off the end-of-rail being so close together, there is a limit to how long the signal period can be before the reflection from the end-of-cable overlaps with the reflection from the end-of-rail. Overlapping signals would cause distortion that would prevent the correlation step from giving accurate locations for when the reflections occur.

The next limiting factor is the symbol period of the signal. Each bit in the barker code corresponds to a symbol in the signal. The period of each symbol is equal to the signal period divided by the number of bits in the signal $T_{symbol} = \frac{T_{signal}}{N_{bits}}$. This symbol period will be the largest factor in the bandwidth of the signal. The bandwidth between the first nulls of the signal is where most of the signal energy is and is equal to twice the reciprocal of the symbol period $B = \frac{2}{T_{symbol}}$. This means the smaller the symbol is the more signal bandwidth will be needed. If the signal period is fixed, then the symbol bandwidth can be controlled by choosing the number of bits used in the signal.

The first restriction on bandwidth is the smallest frequency that can be used as an input. As discussed in the Measurement Equipment section, the amplification hardware is ineffective at low frequencies, being unable to generate signals with low frequency components without causing distortion. This distortion is easy to remove with a low pass filter, but such a filter further distorts any signal components in its stopband. Thus the chosen input signal should have minimal signal energy below $10\,MHz$.

The second bandwidth restriction is the largest frequency that the input signal can contain. An experiment was run on the rail to determine which frequencies would reach the end of the rail without being attenuated. Multiple input signals were sent, each with the same bandwidth but having a different center frequency. The plot in Figure 2.8 shows the reflection correlation for each input signal in a 2D plot. On the x-axis is time. On the y-axis is the center frequency of the input signal. The color bar on the right shows the correlation magnitude in dBs. Visually inspecting this plot a red region of large correlation can be seen between 600 ns and 650 ns. This first region corresponds to the end of the cable. After that region a smaller light green region can be seen just before 700 ns. This second region corresponds to the end of the rail. Unlike the cable region whose reflection is seen for all the input signals, the rails reflection only appears for inputs with center frequencies smaller than 210 MHz. The last input that detected the rail had a maximum frequency of 410 MHz. Thus the chosen input signal should have minimal signal energy above 410 MHz

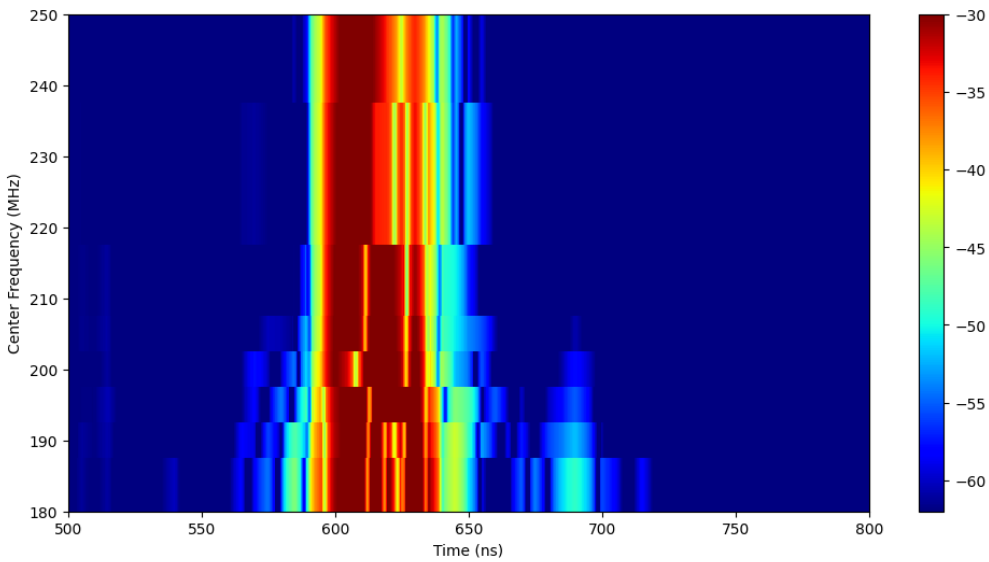

Figure 2.8: *Results from Rail Bandwidth Test*

Based on the upper bound on the signal bandwidth from the rail's attenuation $\max\{f\} = 410\,MHz$ and the lower bound from the distortion region of the amplifier $\min\{f\} = 10\,MHz$, the usable bandwidth of the input signal is $B = 400\,MHz$. Based on this bandwidth, the length of each symbol in the barker code must be larger than $T_{symbol} = 5\,ns$. To ensure all the sidelobes from the end-of-cable are finished well before the start of the end-of-rail reflection, the total length of this signal has been chosen to be 25 ns. The signal must now be modulated such that most of the signal energy is in the feasible region between 10 MHz and 410 MHz as shown in Figure 2.9.

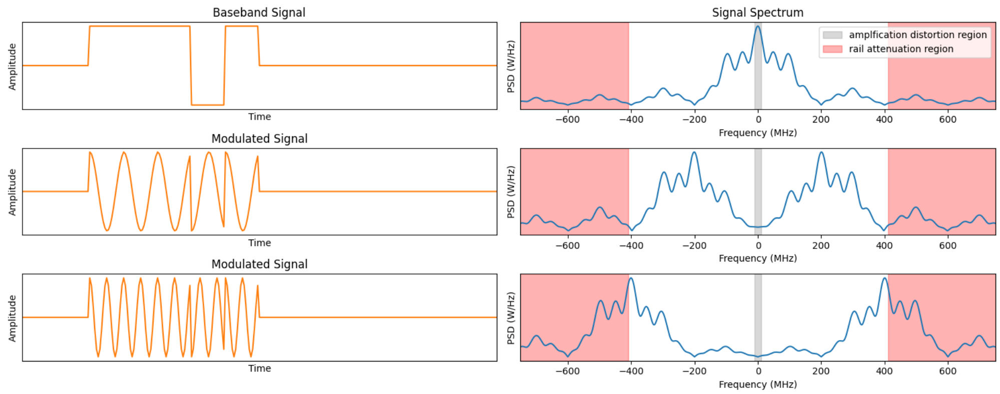

Figure 2.9: *Illustration of Input Signal Bandwidth*

2.4 Noise Effect

The reflection measurements from the rail contain the deterministic behavior of the signal along with the random electromagnetic noise from the environment and electronics. Results from testing show that the noise is consistently white and gaussian, but the amount of noise varies from day to day. The variability in noise power isn't surprising since the noise is likely thermal noise that is proportional to the rail temperature.

The size of reflected signal from the end of the rail is small due to the amount of signal attenuation and the amount of signal energy lost in the transition from cable to rail. As a result of this, when the amount of noise is high there is a noise floor that makes it challenging to determine where that reflection is. An ideal illustration of this issue is shown in Figure 2.10.

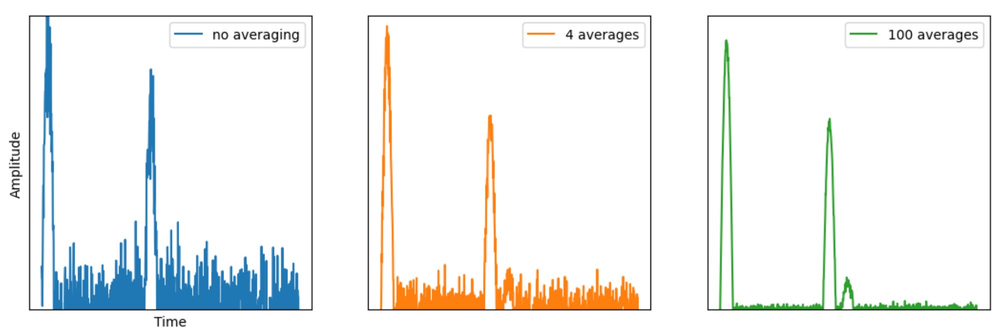

Figure 2.10: *Illustrated Effect of Averaging on Noise*

While the presence of this noise might be unavoidable, the signal can be made clearer by increasing the signal to noise ratio (SNR). The amplifier system describes in Section 2.1 increases the about of signal energy sent into the system and the matched filter described in Section 2.3 uses correlation to distinguish the signal from the random noise. In addition to these previously discussed methods, the signal to noise ratio is also improved by reducing the amount of noise energy in the measurement with an ensemble average. Ensemble averaging runs multiple tests in quick succession then averages the results together. The process reduces the noise in all frequency bands while leaving the signal unchanged [26].

If the noise power in each individual test is σ^2, then the noise power in the average of n tests would be σ^2/n [27]. This well-known relationship says that the noise power can be reduced, but there is a diminishing return for each additional tested average. The number of tests can be chosen to ensure the reflection is always sufficiently larger than the noise in the measurement.

2.5 Experimental Results

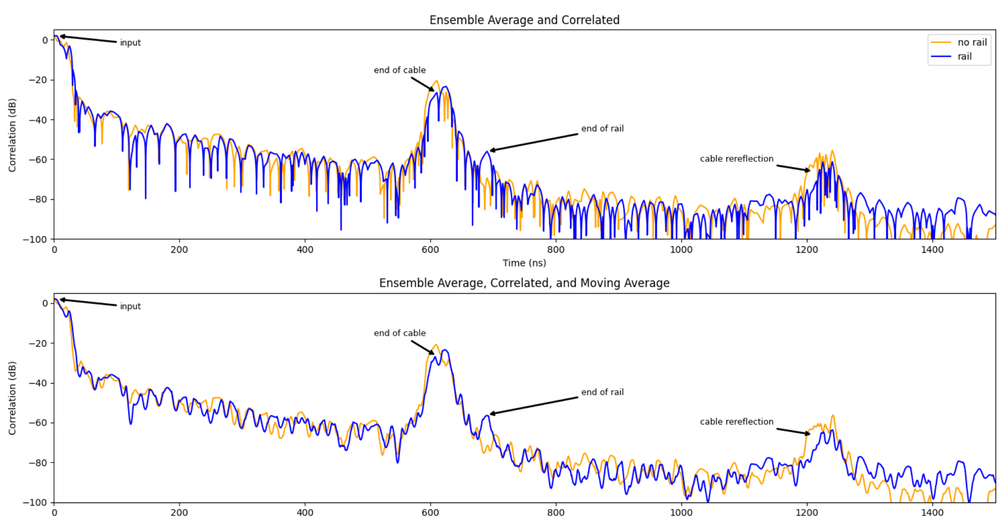

Figure 2.11: *Plot of Experimental Results*

In Figure 2.11 are plots of the correlation of the track's response to the input signal. In the top plot is the data from the measurement equipment ensemble averaged together and correlated with the transmitted signal. In the bottom plot the data is still ensemble averaged and correlated with the transmitted signal like the top plot, but it is then smoothed with a moving average filter. The moving average filter reduces the oscillation caused by the transition from positive to negative correlation and makes it easier to see the underlying behavior. Both plots include two tests. In orange there is a test where the measurement equipment is connected to the coax cable, but the cable remains disconnected from the rail. In blue there is a test where the coax cable is connected to the rail. The reason for both test is to distinguish between reflection pulses from the cable and reflection pulses from the rail.

In both tests a large pulse can be seen at time $t = 0$ where the signal is launched from the DAC. An approximately $-20\,dB$ reflection of the input signal occurs at $609.75\,ns$. This reflection corresponds to the end of the coax cable, which is why $609.75\,ns$ later another even smaller reflect occurs which is the re-reflection of the end-of-cable. Where the tests differ is at $t = 689.997\,ns$ where the rail test has an $-56\,dB$ reflection that does not appear in the no rail test. This correlation pulse is interpreted as the reflection from the end of the rail. Given this interpretation, the signal would have traveled $22.032m$ in $80.247\,ns$. Which means the signal's propagation velocity was $2.74 \cdot 10^8 \frac{m}{s}$ and the corresponding velocity factor 0.9158. This result is reasonable since the rail is essentially an unshielded cable surrounded by mostly air. If the rail was an ideal cable surrounded by nothing but air the velocity factor would be 1. Having a value that is close to 1 matches the environment the rail is in.

2.6 Summary

This chapter described the experimental setup and testing of the TDR method. The results showed that the TDR method is able to measure the reflection from the end of a rail without breaks. This means TDR is a viable means for rail inspection. The next chapter will investigate using the TDR method when there are breaks present in the rail.

Chapter 3

Digital Twin Design and Analyses

This chapter analyzes the TDR method performance when detecting breaks in a rail. The analyses are based on results from a rail digital twin informed by measurements from the Chapter 2. Section 3.1 introduces the digital twin and describes the design. Section 3.2 discusses the break detection and break location estimation methods. The section derives probability distributions and theoretical performance for the detection and estimation methods. Section 3.3 compares the performance of the detection and estimation methods on the digital twin data with the theoretical performance to evaluate the performance of the TDR method.

3.1 Digital Twin Setup

Testing on the real-world rails provides an understanding of how a signal travels on a non-defected rail. Measuring a reflection of a break would require a rail with a break in it. Measuring a reflection from a break in a different location would require yet another rail. Using only measurements from real rails to build a dataset of break reflections would require far more rails than were available.

Instead, a digital twin of the rail was created using COMSOL Multiphysics. This digital twin was informed by measurements from the actual rail section. It can simulate the reflection that would occur if a test had been run on a real section of track. The digital twin can be programmed with a break at arbitrary locations, and its output would simulate what the reflection would be if the real rail had a break at that location. By thoroughly looping through break locations, a dataset could be generated that can be used to analyze detection and estimation methods. Therefore, a digital twin model is essential to study the feasibility of the proposed TDR method.

The geometry of the digital twin is the shape and location of all the parts of the model. Figure 3.1 shows a screen shot of the geometry of the digital twin. The digital twin is made out of five types of parts:

- The $480\,in$ long steel rail.
- The oak wood support beams that are placed directly below the rail and are spaced every $9\,in$.
- The ballast, which is a material made from various relatively uniform crushed rocks fills the space between the rails.
- The soil which is the foundation that the rest of the track rests on.
- A block of air that is defined to surround the rest of the model.

The digital twin uses the COMSOL Transient Electromagnetic Waves physics interface to simulate how the signal propagates through the model. To configure the physics interface, two perfect electric conductors are defined, one at the start of the rail and one on a small patch of soil directly below the start of the rail. A lumped port is defined between the two electric conductors. The lumped port sets a voltage difference between the rail and the ground. Once a voltage difference appears, the physics interface will simulate the resulting electromagnetic waves as they propagate down the rail in time. The voltage set by the lumped port is controlled by a mathematical function. This allows the simulator to send arbitrary waveforms just like the DAC would be able to in the real world. The final part of

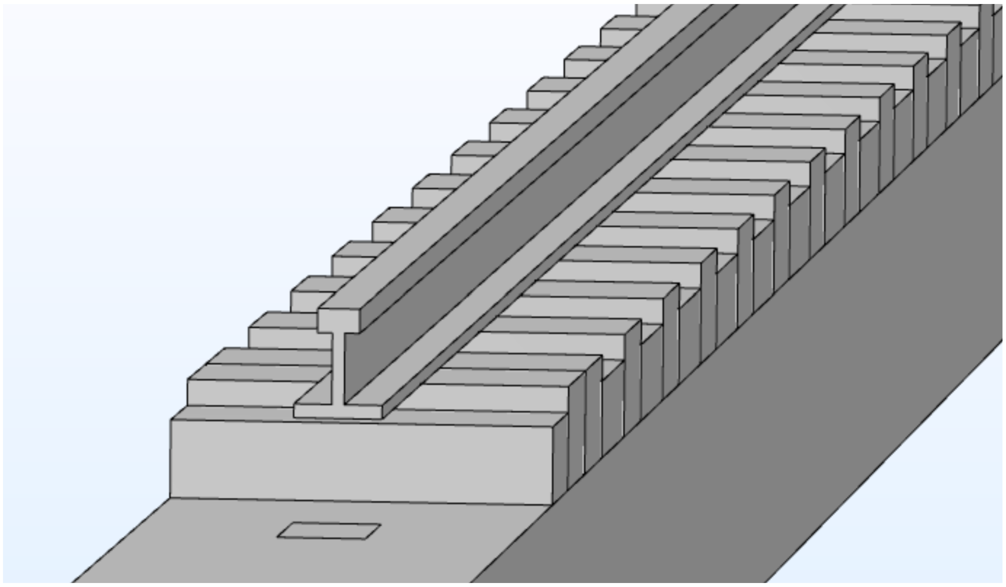

Figure 3.1: *Digital Twin Screenshot*

the physics interface is the borders of the block of air which are also defined to be perfect conductors to ensure the model acts as a closed and isolated system.

The physics interface uses the material properties of the digital twin parts to determine how a wave would behave at that point. Each of the materials in the digital twin has its own dielectric properties: electric conductivity, electric permittivity, and magnetic permeability. The material properties affect the propagation speed of the wave and which frequencies in the wave get attenuated. This in turn affects the size, shape, and delay of the returning reflections. In order to match the behavior of the model elements with the real-world behavior, reasonable values had to be selected for each of the materials. These values were selected from a combination of literature on similar materials under varying moisture contents, measurements collected from our real-world rail, and measurements provided by industry professionals.

For several of the materials, the dielectric properties are affected by environmental factors like the amount of moisture in the material [28, 29]. Materials that absorb water increase in conductivity and permittivity as the moisture content increases. The detection and estimation can be done regardless of whether the materials are damp or dry, but both require that the propagation speed be known in advance. Since the amount of moisture in the materials when damp varies on factors like the frequency and accumulation of rain, the design decision is made to use the material properties when the materials are dry. There is still some variability since the exact line between what is considered dry and damp can be hard to define, but the decisions made should reflect a common single state of the rail.

The effect of magnetic permeability of all of these materials is negligible. All of these materials have a relative permeability $\mu_r \approx 1$ which remains unaffected by environmental factors. This means the material properties of interest are the electric conductivity and electric permittivity. For the surrounding air and steal rail, the conductive and permittivity are known and remain almost constant for changing environmental factors. The soil is assumed to be loamy and dry. Based on these assumptions conductivity and permittivity are determined from preexisting literature [28]. The wood beams are known to be made of oak. The beams are oriented such that the electromagnetic waves will be perpendicular to the rings of the wood. Equations exist for the permittivity and resistivity per the wood moisture content [29]. The resistivity can be algebraically converted to conductivity.

The dielectric properties of the ballast are harder to calculate. Partly because it is a mixture of an unknown ratio of several materials, each with their own properties, but also because the asymmetric grains cause air gabs. This complexity prevents the dielectric parameters from being determined directly from the materials involved. Measurements provided by railroad professionals state that properly maintained ballast has a Resistance Per Length of $30\frac{\Omega}{ft}$. The measurement can be algebraically converted into a measurement of conductivity using the equation $\sigma = \rho^{-1} = \frac{1}{A\frac{R}{L}}$. The value of the permittivity of the ballast is the most speculative of the material properties. Since the relative permeability $\mu_r \approx 1$ the equation for the velocity factor can be simplified to,

$$VF = \frac{1}{\sqrt{\mu_r \varepsilon_r}} \approx \frac{1}{\sqrt{\varepsilon_r}} \tag{3.1}$$

using the velocity factor determined in Chapter 2 the permittivity can be estimated to be $\varepsilon_r \approx VF^{-2}$.

Table 3.1 summarized the material dielectric properties chosen for the simulation.

Material Property	Symbol	Dry Value
Air conductivity	σ_{air}	0
Steel conductivity	σ_{steel}	4.032×10^6
Soil conductivity	σ_{soil}	10^{-4}
Wood conductivity	σ_{wood}	1.278×10^{-4}
Ballast conductivity	σ_{ballast}	4×10^{-3}
Air permittivity	ε_{air}	1.0
Steel permittivity	$\varepsilon_{\text{steel}}$	1.0
Ballast permittivity	$\varepsilon_{\text{ballast}}$	1.19
Soil permittivity	$\varepsilon_{\text{soil}}$	4.0
Wood permittivity	$\varepsilon_{\text{wood}}$	1.6

Table 3.1: *Electrical Conductivity and Permittivity of Materials (Dry Values)*

Once the model was working several simplifying assumptions were made to reduce the runtime of the model. The first simplification was to only include one rail in the digital twin. The real track used in testing has two steel rails placed 56.375 *in* apart, but due to the inverse square law, any effect the second rail has on the propagation of the electromagnetic waves is negligible. Removing the second rail and shrinking the model around the first rail significantly reduced the number of COMSOL mesh elements and improved the model runtime. The second simplification was ignoring the cross-sectional shape of the rail; treating it like a simple rectangular beam of steel. The cross-section of the rail has no effect on the timing of the reflections. By simplifying the shape of the beam, COMSOL was able to use a simpler surface when generating the COMSOL mesh which reduced the number of mesh elements.

With this complete digital twin model, breaks can be nondestructively placed at arbitrary positions at the rail. By programmatically looping through multiple break locations, the physics simulation can create a dataset of break reflections that can be used in analysis.

The output of the digital twin is the deterministic behavior of the rail reflections. For each break location in the digital twin dataset, there is one unique function of observations that describes the rail reflections at each point in time. Figure 3.2 shows plots of observations from two break locations: one at 50% of the rail and one at 70%. The top plot displays the raw amplitude of the observed reflections plotted against time. The bottom plot displays a plot of the observations correlated with the transmitted signal. In the correlation plot, the shift in the reflection location can be seen by visual inspection.

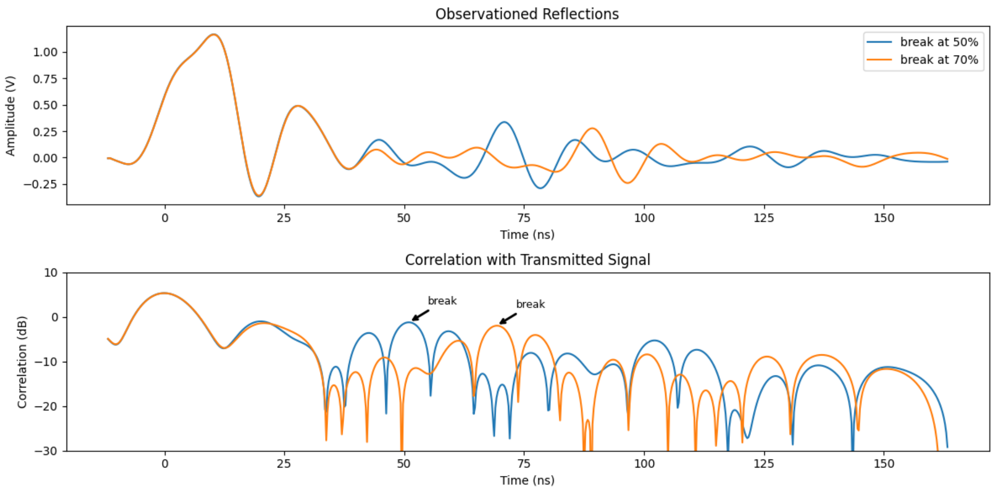

Figure 3.2: *Data from Digital Twin*

When doing testing in the real world, this deterministic behavior is only one component of the observations. The observations will also contain random noise that will obscure what is happening. To evaluate the detection and estimation of breaks the randomness must be included in the analysis.

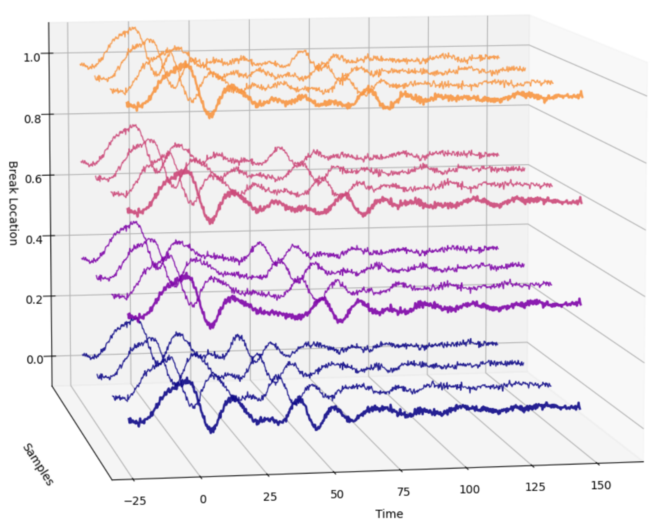

Figure 3.3: *Illustration of the Monte Carlo simulation*

A Monte Carlo simulation is used to extend the digital twin dataset. The Monte Carlo simulation converts the deterministic functions of observations to a collection of sample functions from the underlying random process [30]. It does this by generating multiple sample functions for each deterministic function. Each with independent white gaussian noise added to each element in the observations as illustrated in Figure 3.3. When enough sample functions are generated, the dataset approximates the random process. This extended dataset is used as the input to the detection and estimation implementations.

3.2 Detection and Estimation Analysis

This section derives the theoretical performance of the detection and estimation methods. The detection Subsection 3.2.1 begins by defining the null and alternative hypotheses for the observed reflections. It then derives probability distributions of the observations under the null and alternative hypothesis. These distributions are used to derive the detection statistic and the associated false alarm and detection rates. The estimation Subsection 3.2.2 begins with the definition of maximum likelihood estimation and converts it to an equivalent estimation problem of estimating the time delay of a matched filter. The subsection then derives the optimum bias and variance of a matched filter. The results from these subsections are compared to the performance of the detection and estimation implementations on the digital twin data in Section 3.3.

3.2.1 Detection

The first part of the detection analysis is defining what is being detected. In this problem the observation can be classified into one of two cases. The first case is the null hypothesis $\mathcal{H}_0$ when there is no break in the rail. The observations in this case should just be noise from the environment. The second case is the alternative hypothesis $\mathcal{H}_1$ when there is a break in the rail. In this case the observations should contain a time shifted version of the transmitted signal $s(t)$ at some point before the reflection from the end of the track $t < T_r$. This shifted version of the transmitted signal will also be impacted by the additive noise from the environment along with whatever distortion the environment makes to the signal as it travels down the rail.

$$y(t) = \begin{cases} n(t), & \mathcal{H}_0 \quad \text{(No reflection)} \\ h(t) * s(t-k_b) + n(t), & \mathcal{H}_1 \quad \text{(Reflection present)} \end{cases} \tag{3.2}$$

here $y(t)$ is the observation at time t, $s(t)$ is the transmitted signal, $n(t) \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ is assumed to be i.i.d. additive white noise encountered by the signal, σ^2 is the variance of the noise, and $h(t)$ is the impulse response of the rail. The corresponding discrete time expressions can equivalently be written in vector from as,

$$\mathbf{y}_i = \begin{cases} \mathbf{n}_i, & \mathcal{H}_0 \quad \text{(No reflection)} \\ \mathbf{h} * \mathbf{s}_{i-k_b} + \mathbf{n}_i, & \mathcal{H}_1 \quad \text{(Reflection present)} \end{cases} \tag{3.3}$$

where y_i is the i^{th} element of an observation vector $\mathbf{y}$, s_i is the i^{th} element of the transmitted signal s, k_b is the discrete reflection delay of the signal. Since the noise of each element n_i is i.i.d. gaussian noise, the observation vector is a gaussian random vector. Based on this definition of the observation vector, probability distributions can be derived for each hypothesis.

$$p(\mathbf{y}|\mathcal{H}_0) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left(-\frac{\mathbf{y}^T\mathbf{y}}{2\sigma^2}\right) \tag{3.4}$$

$$p(\mathbf{y}|\mathcal{H}_1) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left(-\frac{(\mathbf{y}-\mathbf{s})^T(\mathbf{y}-\mathbf{s})}{2\sigma^2}\right) \tag{3.5}$$

where N is the number of observations in vector $\mathbf{y}$.

The observation vector probability distributions can be combined to make the Likelihood Ratio,

$$L(\mathbf{y}) = \frac{p(\mathbf{y}|\mathcal{H}_1)}{p(\mathbf{y}|\mathcal{H}_0)} = \exp\left(\frac{1}{\sigma^2}\mathbf{s}^T\mathbf{y} - \frac{1}{2\sigma^2}\mathbf{s}^T\mathbf{s}\right) \tag{3.6}$$

This likelihood ratio will be used as a test statistic [31] to summarize the evidence in the observation vector into a single number that can be compared to a threshold to classify the rail as either having or not having a break. This paper uses Neyman-Pearson Criterion to define a decision threshold η for a fixed false alarm rate $P_{FA} = \alpha$.

$$L(\mathbf{y}) \gtrless \eta \tag{3.7}$$

$$\eta = \arg_{\eta} \{P_{FA} = \alpha\} = \arg_{\eta} \left\{ \int_{L(\mathbf{y})>\eta} p(\mathbf{y}|H_0) d\mathbf{y} = \alpha \right\} \tag{3.8}$$

By substituting the equation for likelihood ratio this decision rule can be rearranged as,

$$\exp\left(\frac{1}{\sigma^2}\mathbf{s}^T\mathbf{y} - \frac{1}{2\sigma^2}\mathbf{s}^T\mathbf{s} \right) \gtrless \eta \tag{3.9}$$

$$\frac{1}{\sigma^2}\mathbf{s}^T\mathbf{y} - \frac{1}{2\sigma^2}\mathbf{s}^T\mathbf{s} \gtrless ln(\eta) \tag{3.10}$$

$$\mathbf{s}^T\mathbf{y} - \frac{1}{2}\mathbf{s}^T\mathbf{s} \gtrless \sigma^2 ln(\eta) \tag{3.11}$$

$$\mathbf{s}^T\mathbf{y} \gtrless \sigma^2 ln(\eta) + \frac{1}{2}\mathbf{s}^T\mathbf{s} \tag{3.12}$$

$$z \gtrless \eta' \tag{3.13}$$

where z is the dot product of the observation vector $\mathbf{y}$ with the transmitted signal $\mathbf{s}^T$ and η' is a new threshold value. The new parameter z would have the following hypotheses,

$$\mathbf{z} = \mathbf{s}^T\mathbf{y} = \begin{cases} \mathbf{s}^T\mathbf{n}, & \mathcal{H}_0 \quad \text{(No reflection)} \\ \mathbf{h} * \mathbf{s}^T\mathbf{s} + \mathbf{s}^T\mathbf{n}, & \mathcal{H}_1 \quad \text{(Reflection present)} \end{cases} \tag{3.14}$$

$$= \begin{cases} n_s, & \mathcal{H}_0 \quad \text{(No reflection)} \\ \mathbf{h} * \mathbf{s}^T\mathbf{s} + n_s, & \mathcal{H}_1 \quad \text{(Reflection present)} \end{cases} \tag{3.15}$$

where $n_s \sim \mathcal{N}(0, \mathbf{s}^T\mathbf{s}\sigma^2)$ is a single-dimensional quantity of additive gaussian noise whose variance is $\mathbf{s}^T\mathbf{s}$ times the variance of noise added to each point of the signal. Which means the parameter z is a random variable whose mean and variance under the hypotheses are,

$$= \begin{cases} \mathcal{H}_0 : z \sim \mathcal{N}(0, \mathbf{s}^T\mathbf{s}\sigma^2), \\ \mathcal{H}_1 : z \sim \mathcal{N}(\mathbf{h} * \mathbf{y}^T\mathbf{s}, \mathbf{s}^T\mathbf{s}\sigma^2) \end{cases} \tag{3.16}$$

Using the gaussian distribution of the parameter z, the false alarm rate and detection rate can be derived for a fixed threshold η',

$$P_{FA} = P(z > \eta'|H_0) = \int_{\eta'}^{\infty} p(z|H_0)dy = Q\left(\frac{\eta'}{\sqrt{\mathbf{s}^T\mathbf{s}\sigma^2}}\right) \tag{3.17}$$

$$P_D = P(z > \eta'|H_1) = \int_{\eta'}^{\infty} p(z|H_1)dy = Q\left(\frac{\eta' - \mathbf{s}^T\mathbf{s}}{\sqrt{\mathbf{s}^T\mathbf{s}\sigma^2}}\right) \tag{3.18}$$

where Q is the cumulative distribution function of the standard normal distribution.

Sweeping through threshold values and plotting the resulting detection rates against the false alarm rate generates a theoretical ROC curve which shows the performance of the detector. If the area under the ROC curve is 1 then the detector is perfect. If the area under the curve is 0.5 then the detection is equivalent to random guessing.

The theoretical ROC curves will be treated as the optimum performance of a detector implementation. The analysis will compare the results of the detector on the generated dataset with these theoretical ROC curves to quantify the effectiveness of the implementation.

3.2.2 Parameter Estimation

Once a break reflection has been determined to be present in the observations, the next task determines where the break in the rail is located. This starts by doing parameter estimation to determine the reflection period of the break k_b. Once the reflection period is estimated, time domain reflectometry can be used to convert the measurement of time to a measurement of location.

This estimator uses the maximum likelihood criterion which chooses the value of k_b that has the maximum likelihood of generating the observations $\mathbf{y}$ and uses that value as the estimate $\hat{k}_b$.

$$\hat{k}_b = \arg\max_{k_b} \{p(\mathbf{y}|k_b)\} \tag{3.19}$$

$$\mathbf{y} = \mathbf{s} + \mathbf{n} \tag{3.20}$$

Since the noise $\mathbf{n}$ in the observations is a gaussian random vector, maximizing the likelihood is equivalent to minimizing the least square error of the estimate [32].

$$\hat{k}_b = \arg\min_{k_b} \left\{ \sum (\mathbf{y}_i - \mathbf{s}_{i-k_b})^2 \right\} \tag{3.21}$$

$$= \arg\min_{k_b} \left\{ \sum \mathbf{y}_i^2 - 2 \sum \mathbf{y}_i \mathbf{s}_{i-k_b} + \sum \mathbf{s}_{i-k_b}^2 t \right\} \tag{3.22}$$

The first and third terms in this equation are constant for all values of k_b, so this equation is equivalent to,

$$= \arg\min_{k_b} \left\{ -2 \sum \mathbf{y}_i \mathbf{s}_{i-k_b} \right\} \tag{3.23}$$

$$= \arg\max_{k_b} \left\{ \sum \mathbf{y}_i \mathbf{s}_{i-k_b} \right\} \tag{3.24}$$

In this final form, the definition of the maximum likelihood estimator is equivalent to finding the time delay of a matched filter. A matched filter finds the cross-correlation between the observations $\mathbf{y}$ and the transmitted signal $\mathbf{s}$ in order to extract the signal in the observations from the observation noise.

The quality of an estimator can be quantified by treating the estimator as a random variable and calculating its bias and variance. The bias of the estimator is defined by the difference between the mean of the estimator and the true value of the parameter being estimated. The estimator would be considered unbiased if the difference between the mean and the true value is zero.

$$\text{Bias} = E\left[\hat{k}_b\right] - \theta \tag{3.25}$$

where θ is the parameter being estimated. In this case the parameter being estimated is the reflection period k_b.

The variance of the estimator is defined by the amount of uncertainty in the estimate. Since the observations used for estimation are noisy, the value predicted by the estimate won't be the same each time. The smaller the variance the better the estimator. The theoretical optimum variance for an estimator is the Cramér-Rao lower bound which is defined by,

$$\text{Var}\left[\hat{k}_b\right] \geq CRLB = \frac{1}{I(\theta)} \tag{3.26}$$

where $I(\theta)$ is the Fisher information. The Fisher information measures how much information a random variable carries about an unknown parameter [32]. It is defined as,

$$I(\theta) = E\left[\left(\frac{\partial}{\partial\theta}\log p(\mathbf{y}|\theta)\right)^2\right] \tag{3.27}$$

Deriving the equation for the Fisher information begins by defining a new parameter x_i which integrates the observations,

$$\mathbf{x}_t = \int_0^t \mathbf{y}_u du \tag{3.28}$$

The null and alternative hypotheses of this new parameter are,

$$\mathbf{x}_t = \begin{cases} W_t, & \mathcal{H}_0 \quad \text{(No reflection)} \\ \int_0^t \mathbf{s}_u(\theta)du + W_t, & \mathcal{H}_1 \quad \text{(Reflection present)} \end{cases} \tag{3.29}$$

where $\mathbf{s}_u(\theta)$ is the transmitted signal time shifted by θ and W_i is a Winer process whose increments are the noise in the observation $n(t)$. The likelihood ration of the new parameter [33] is,

$$L(x_t) = \frac{p(x|\mathcal{H}_1,\theta)}{p(x|\mathcal{H}_0,\theta)} = exp\left\{\frac{1}{N_0/2}\int_0^T \mathbf{s}_t(\theta)dx_t - \frac{1}{N_0}\int_0^T \mathbf{s}_t^2(\theta)dt\right\} \tag{3.30}$$

where $N_0/2$ is the magnitude of the power spectral density of the noise and T is the period being integrated over. Assuming $p(x|\mathcal{H}_1) >> p(x|\mathcal{H}_0)$ [34] the likelihood of $\mathbf{x}$ when $\mathcal{H}_1$ is true is approximately equal to the likelihood ratio,

$$p(x|\mathcal{H}_1,\theta) \approx exp\left\{\frac{1}{N_0/2}\int_0^T \mathbf{s}_t(\theta)dx_t - \frac{1}{N_0}\int_0^T \mathbf{s}_t^2(\theta)dt\right\} \tag{3.31}$$

Estimation will be performed when the reflected signal has been detected in the observations so $p(x|\theta) = p(x|\mathcal{H}_1,\theta)$. Substituting this likelihood into the equation for Fisher information gives,

$$I(\theta) = E\left[\left(\frac{\partial}{\partial\theta}\left\{exp\left\{\frac{1}{N_0/2}\int_0^T \mathbf{s}_t(\theta)dx_t - \frac{1}{N_0}\int_0^T \mathbf{s}_t^2(\theta)dt\right\}\right\}\right)^2\right] \tag{3.32}$$

$$= E\left[\left(\frac{\partial}{\partial\theta}\frac{1}{N_0/2}\int_0^T \mathbf{s}_t(\theta)dx_t - \frac{\partial}{\partial\theta}\frac{1}{N_0}\int_0^T \mathbf{s}_t^2(\theta)dt\right)^2\right] \tag{3.33}$$

$$= E\left[\left(\frac{2}{N_0}\left(\frac{\partial}{\partial\theta}\int_0^T \mathbf{s}_t(\theta)dx_t - \frac{\partial}{\partial\theta}\frac{1}{2}\int_0^T \mathbf{s}_t^2(\theta)dt\right)\right)^2\right] \tag{3.34}$$

$$= E\left[\frac{4}{N_0^2}\left(\int_0^T \mathbf{s}_t'(\theta)dx_t - \int_0^T \mathbf{s}_t(\theta)\mathbf{s}_t'(\theta)dt\right)^2\right] \tag{3.35}$$

where $\mathbf{s}_t'(\theta) = \frac{\partial \mathbf{s}_t(\theta)}{\partial\theta}$,

$$= \frac{4}{N_0^2}E\left[\left(\int_0^T \mathbf{s}_t'(\theta)(dx_t - \mathbf{s}_t(\theta)dt)\right)^2\right] \tag{3.36}$$

using equation 3.29,

$$= \frac{4}{N_0^2}E\left[\left(\int_0^T \mathbf{s}_t'(\theta)dW_t\right)^2\right] \tag{3.37}$$

$$= \frac{4}{N_0^2}E\left[\left(\int_0^T \mathbf{s}_t'(\theta)dW_t\right)^2\right] \tag{3.38}$$

$$= \frac{2}{N_0}\left(\int_0^T \left[\mathbf{s}_t'(\theta)\right]^2 dt\right) \tag{3.39}$$

Since θ is a time shift the equation can be written in the following form,

$$= \frac{2}{N_0}\left(\int_0^T \left[\mathbf{s}'(t-\theta)\right]^2 dt\right) \tag{3.40}$$

letting $T \to \infty$,

$$= \frac{2}{N_0}\left(\int_0^\infty \left[\mathbf{s}'(t)\right]^2 dt\right) \tag{3.41}$$

$$= \frac{2}{N_0}\left(\int_{-\infty}^\infty \omega^2 |\mathbf{S}(\omega)|^2 d\omega\right) \tag{3.42}$$

where $\mathbf{S}(\omega)$ is the Fourier transform of the signal, ω is the angular frequency. This makes the Cramér-Rao lower bound for a matched filter,

$$\mathrm{Var}\left[\hat{k}_b\right] \geq CRLB = \left[\frac{2}{N_0}\left(\int_{-\infty}^\infty \omega^2 |\mathbf{S}(\omega)|^2 d\omega\right)\right]^{-1} \tag{3.43}$$

which can be rewritten as [34],

$$\mathrm{Var}\left[\hat{k}_b\right] \geq CRLB = \left[\frac{2E_0}{N_0}\beta^2\right]^{-1} \tag{3.44}$$

$$\beta^2 = \frac{\int \omega^2 |S(\omega)|^2 d\omega}{\int |S(\omega)|^2 d\omega} \tag{3.45}$$

where $\frac{2E_0}{N_0}$ is the signal to noise ratio. In descrete form, β^2 can be written as,

$$\beta^2 = \frac{\sum \omega_i^2 |S(\omega_i)|^2}{\sum |S(\omega_i)|^2} \tag{3.46}$$

Using a numeric approximation of β^2, the CRLB is calculated for multiple signal to noise ratios. This CRLB are treated at the optimum performance of the estimator implementation. The analysis will compare variance of the estimator on data with a set SNR with the CRLB as that SNR quantify the effectiveness of the implementation.

3.3 Detection and Estimation Performance

For the detection analysis, sample functions from the dataset are passed into a likelihood ratio and a decision rule method. The outputs were then classified as a valid detection (True Positive) or a false alarm (False Positive). Figure 3.4 is a plot of ROC Curves at different signal to noise ratios. The dotted black lines are the theoretical ROC Curves for a given noise ratio. A visual inspection will show that the detection simulation is able to very closely match the theoretical limit. The largest signal to noise ratio in this plot is only 1.5, but even at this rather low ratio the True Positive to False Positive rate is almost perfect.

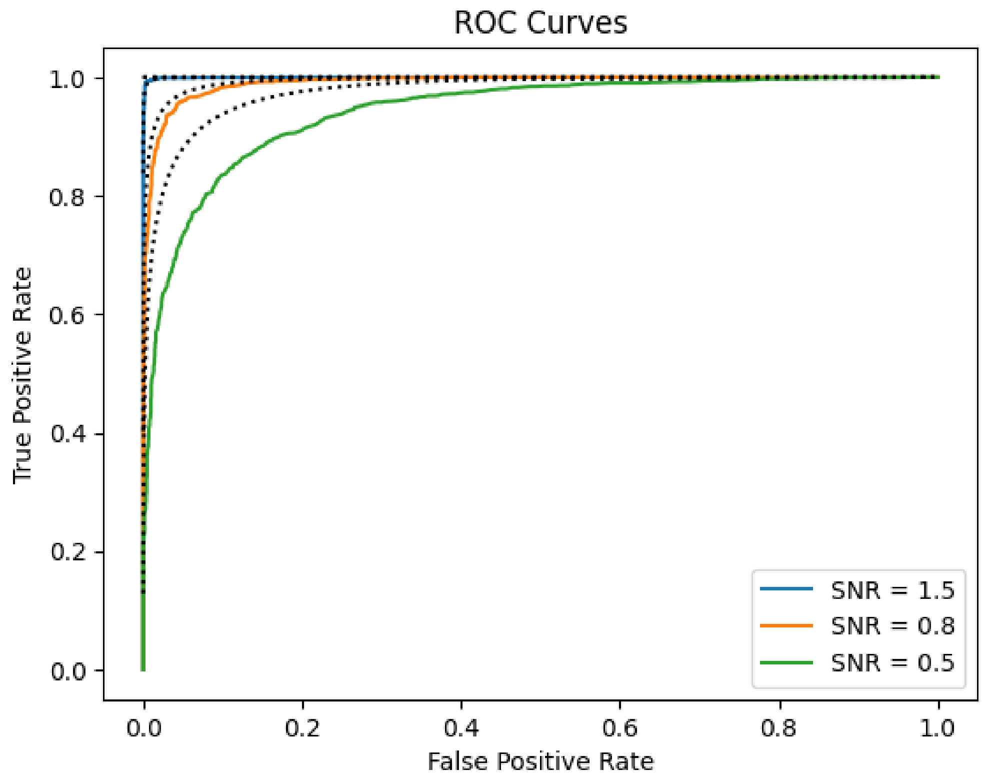

Figure 3.4: *ROC Curves*

The estimation method takes in sample functions from the dataset and provides an estimate for what time the reflection pulse arrives. The time estimate is then converted to the physical distance to the rail to provide a more intuitive metric of the estimator's effectiveness. The total length of the rail is $11.52\,m$.

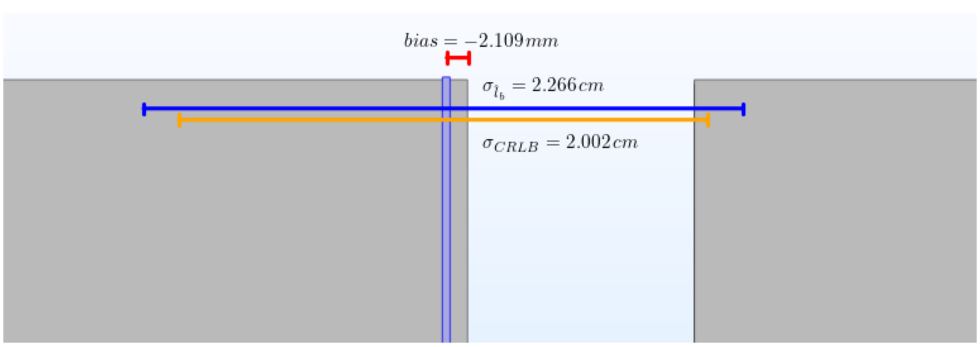

Figure 3.5: *Estimator Bias, Standard Deviation, and Cramér-Rao Lower Bound*

When the sample functions from the random process have a signal to noise ratio of 15, the estimator has a slight bias of $-2.109\ mm$ and a standard deviation of $2.266\,cm$. Numerically calculating the Cramér-Rao lower bound gives a minimum standard deviation of $2.002\,cm$. The simulation estimator's variance is slightly higher than the theoretical minimum variance. The ratio between the Cramér-Rao lower bound and the true estimator variance is the efficiency of the estimator $E = \frac{CRLB}{\mathrm{Var}(\hat{k}_b)} = 0.78$. These estimator bias, standard deviation, and CRLB on standard deviation are plotted on a screen shot of the digital twin to provide a since of scale in Figure 3.5.

The probability distribution of the simulation estimator is approximately normal. The QQ plot in Figure 3.6 shows how close the estimator distribution is to a true normal distribution. The first two quantiles on both sides of the mean are quite close to those of a normal distribution. The tails of the estimator distribution are thinner than those of a normal distribution, meaning it is not a perfect fit.

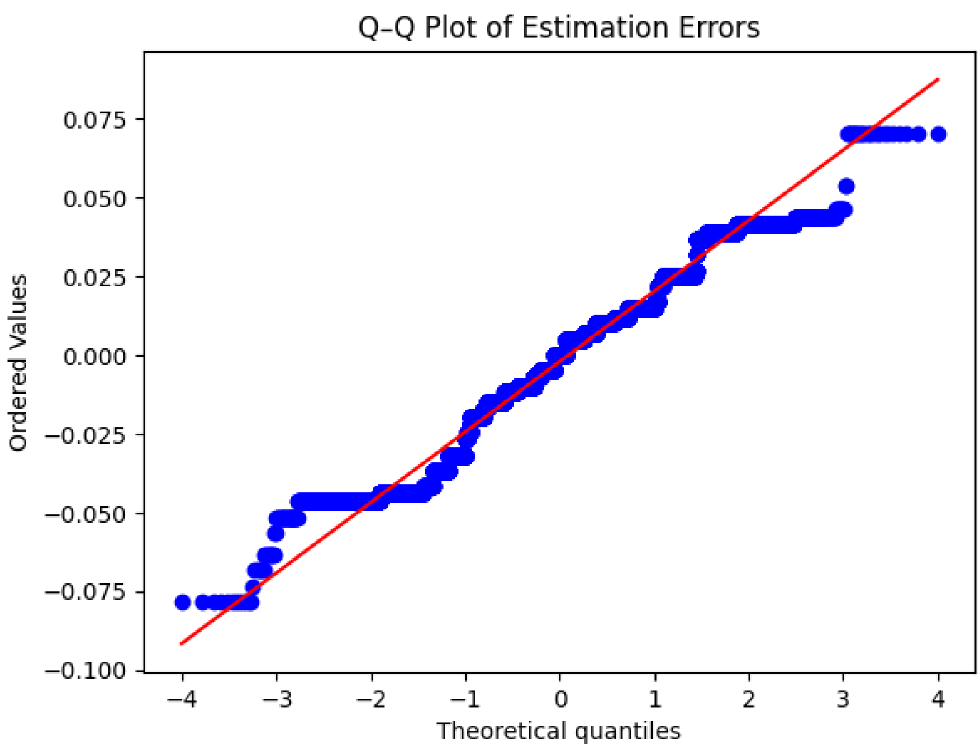

Figure 3.6: *QQ Plot*

Changing the signal to noise ratio used in the simulation results in a change in the variance of the estimator. The equation for the Cramér-Rao lower bound predicts the estimator variance for each signal to noise ratio. In the plots in Figure 3.7 the variance predicted by the Cramér-Rao lower bound is plotted in orange. In blue is the true variance measured from the results of the implementation. The ratio of the Cramér-Rao lower bound and the true estimator variance is plotted in green in the right plot.

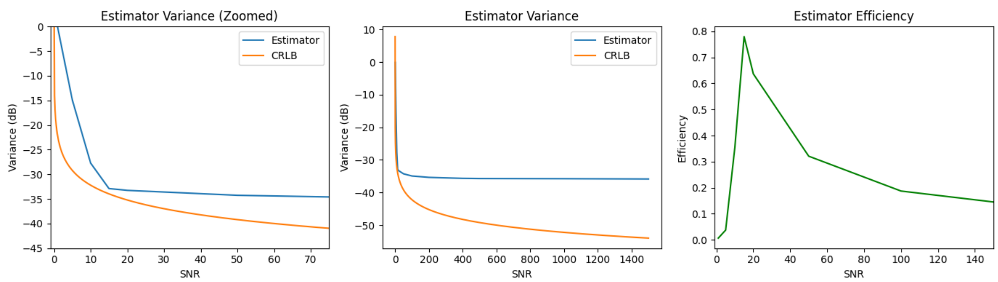

Figure 3.7: *CRLB Plot*

At low signal to noise ratios, the estimator variance is much larger than the Cramér-Rao lower bound resulting in very poor estimator efficiency. As the signal to noise ratio gets larger the variance of the estimator gets much closer to that of the Cramér-Rao lower bound. The estimator's efficiency is at its greatest around a signal to noise ratio of 15. After that the estimator's variance seems to hit an asymptotic limit around $-36\,dB$ as seen in the plot on the right. This limit seems to be caused by limiting factors in the digital twin and thus remains relatively unaffected by increasing the signal to noise ratio used by the Monte Carlo simulation.

3.4 Summary

This chapter introduces a digital twin of a railroad track and shows how the TDR method can be used to detect and estimate the location of breaks in the track. The results show that the TDR method is able to detect breaks in the track and estimate the location of the break with a high degree of accuracy.

Chapter 4

Conclusions and Future Work

The work discussed throughout this thesis is summarized here. Chapter 1 introduced the need for defect detection and provided an overview of several of the current methods. A proposed new method of defect detection was introduced that uses time domain reflectometry that would have several advantages over current methods.

The work in Chapter 2 demonstrated the feasibility of using time domain reflectometry on railroad tracks. The results showed that the reflections from the end-of-rail were large enough to detect. The behavior of the real-world track was used to inform the creation of a digital twin in Chapter 3.

Chapter 3 analyzed the performance of the proposed detection and estimation methods. This chapter provided the methodology for the design of a digital twin of the tracks, and the deterministic outputs of the digital twin simulated the reflections from breaks in the rail. After augmenting the deterministic behavior of the rail with additive gaussian noise, the detection and estimation methods were stochastically analyzed against their theoretical optimum performance and determined to be effective implementations.

During the research for this proposed method, several lessons were learned concerning the methodology:

- The importance of a good electrical connection at the start of the rail. Establishing a good electrical connection between the end of the cable and the start of the rail proved harder to do in practice then expected. There were times when a physical connection was made that seemed sturdy, but when testing was done the observations looked the same as when the measurement equipment was disconnected from the rail. It is important when setting up the field-testing equipment to measure the connection point with an ohm meter to ensure there is a good electrical connection before collecting measurements.

- The importance of using a long enough cable between the measurement equipment and the rail. The large reflection from the end-of-cable will continue to re-reflect off the measurement equipment and the end-of-cable causing periodic reflections to appear in the observations. Early field testing used a short coax cable that was about $1\,m$ in length. This experimental set up caused the end-of-cable reflections to be close together. Since they are much larger than the end-of-rail reflection, the end-of-cable re-reflections obscured the end-of-rail reflection. Switching to a coax cable that was much longer than the rail, spaced these re-reflections further apart so the end-of-rail reflection could now be seen clearly.

- The importance of testing in all weather conditions. Early experimentation suffered from a sampling bias since measurements were mostly collected on sunny and dry days when the field conditions were most pleasant. When later testing was done on a rainy day, the importance of more thorough testing became apparent. This ceased to be an issue once testing became automated and periodic.

The work described throughout this thesis provides a foundation for the proposed method. Continued development of this method provides several opportunities for future research. Some specific examples include:

- The analysis throughout this thesis has treated the propagation velocity as a constant. It is known from the real-world experimentation that the propagation velocity changes as the weather conditions of the track change. The extent of this change, its underlying probability distribution, and which weather conditions cause the change remain unknown. Further research involving long-term monitoring of the rail reflections could record the reflection delay/propagation velocity every few minutes. This database could then be correlated with weather features like humidity, temperature, or rain fall.

- The Monte Carlo simulation done in the analysis was focused on the random noise introduced by the environment. Additional analysis could be done by treating the dielectric properties in the digital twin as random variables rather than constants. Monte Carlo simulation could then be done on the dielectric properties before running the simulation to stochastically analyze how the observations are effected by the track's dielectric properties. This combined with the long-term monitoring could generate more realistic estimates for the track's dielectric properties.

- The design of the digital twin was informed by measurements from a real-world rail without defects. The break reflections the digital twin simulated were not confirmed against break reflections from a real-world rail since a rail with a break was not available for review. Future research and experimentation with a real-world rail having a break could confirm the findings of this thesis.

- The work in this thesis has been exclusively focused on detecting defects that cause a complete break in the rail. In theory Time Domain Reflectometry could also detect the impedance change caused by more subtle defects. Future research could analyze the proposed method's performance when detecting a wider selection of defects.

- The work in this thesis has focused on detecting breaks in sections of continuous rail. Future research could analyze the proposed method's performance when detecting breaks in turnouts and crossings [35] which have asymmetrical structure and act as electric deadzones since no existing hardware is installed for collecting measurements.

Bibliography

[1] Transportation Statistics Annual Report 2024. [Online]. Available: https://rosap.ntl.bts.gov

[2] D. F. Cannon, K.-O. Edel, S. L. Grassie, and K. Sawley, "Rail defects: An overview," vol. 26, no. 10, pp. 865–886. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1460-2695.2003.00693.x

[3] "(PDF) Probabilistic Risk Analysis of Broken Rail-Caused Train Derailments," in *ResearchGate*. [Online]. Available: https://www.researchgate.net/publication/343492378_Probabilistic_Risk_Analysis_of_Broken_Rail-Caused_Train_Derailments

[4] V. Hromádka, J. Korytárová, E. Vítková, H. Seelmann, and T. Funk, "Economic Impact of Occurrences on Railways," vol. 181, pp. 76–83. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050921001447

[5] S. Welch, S. Youn, A. Nichols, S. Na, and R. Shen, "Transportation of hazardous material via railroad: Incident investigation and a case study of derailment in 2023," vol. 43, no. 3, pp. 570–578. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/prs.12598

[6] X. Liu, M. Rapik Saat, and C. P. L. Barkan, "Freight-train derailment rates for railroad safety and risk analysis," vol. 98, pp. 1–9. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001457516303426

[7] R. Mordia and A. K. Verma, "Nondestructive Testing Methods for Rail Defects Detection." [Online]. Available: https://www.sciencedirect.com/science/article/pii/S294986782500008X

[8] Y. Zhao, Z. Liu, D. Yi, X. Yu, X. Sha, L. Li, H. Sun, Z. Zhan, and W. J. Li, "A Review on Rail Defect Detection Systems Based on Wireless Sensors," vol. 22, no. 17, p. 6409. [Online]. Available: https://www.mdpi.com/1424-8220/22/17/6409

[9] B. Li, X. Chen, Z. Wang, and S. Tan, "Vibration Signal Analysis For Rail Flaw Detection," in *2019 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, pp. 830–835. [Online]. Available: https://ieeexplore.ieee.org/document/9213353

[10] X. Zhang, Y. Cui, Y. Wang, M. Sun, and H. Hu, "An improved AE detection method of rail defect based on multi-level ANC with VSS-LMS," vol. 99, pp. 420–433. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327017303412

[11] X. Zhang, K. Wang, Y. Wang, Y. Shen, and H. Hu, "Rail crack detection using acoustic emission technique by joint optimization noise clustering and time window feature detection," vol. 160, p. 107141. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0003682X19303548

[12] A. K. Verma, *Introduction To Modern Planar Transmission Lines: Physical, Analytical, and Circuit Models Approach.* John Wiley & Sons.

[13] M. De Piante and A. M. Tonello, "On Impedance Matching in a Power-Line-Communication System," vol. 63, no. 7, pp. 653–657. [Online]. Available: https://ieeexplore.ieee.org/document/7407357

[14] "Transmission Lines," in *Fundamentals of Engineering Electromagnetics*, R. Bansal, Ed. CRC/Taylor & Francis, pp. 185–225.

[15] F. T. Ulaby and U. Ravaioli, "Transmission Lines," in *Fundamentals of Applied Electromagnetics*, 7th ed., pp. 70–144. [Online]. Available: https://www.abebooks.com/9780132139311/Fundamentals-Applied-Electromagnetics-Ulaby-Fawwaz-0132139316/plp

[16] W. Winter and M. Herbrig, "Time Domain Measurements in automotive applications," in *2009 IEEE International Symposium on Electromagnetic Compatibility*, pp. 109–115. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/5284604

[17] P. Smith, C. Furse, and J. Gunther, "Analysis of spread spectrum time domain reflectometry for wire fault location," vol. 5, no. 6, pp. 1469–1478. [Online]. Available: https://ieeexplore.ieee.org/document/1532290

[18] D. Uttam and B. Culshaw, "Precision time domain reflectometry in optical fiber systems using a frequency modulated continuous wave ranging technique," vol. 3, no. 5, pp. 971–977. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/1074315

[19] M. K. Smail, L. Pichon, M. Olivas, F. Auzanneau, and M. Lambert, "Detection of Defects in Wiring Networks Using Time Domain Reflectometry," vol. 46, no. 8, pp. 2998–3001. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/5512859

[20] N. Crespi, A. T. Drobot, and R. Minerva, "The Digital Twin: What and Why?" in *The Digital Twin*, N. Crespi, A. T. Drobot, and R. Minerva, Eds. Springer International Publishing, pp. 3–20. [Online]. Available: https://doi.org/10.1007/978-3-031-21343-4_1

[21] B. Ali Rachedi, A. Babouri, and F. Berrouk, "A study of electromagnetic field generated by high voltage lines using COMSOL MULTIPHYSICS," in *2014 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM)*, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7076989/references

[22] L. H. Crockett, D. Northcote, and R. W. Stewart, *Software Defined Radio with Zynq UltraScale+ RFSoC*. Strathclyde Academic Media. [Online]. Available: www.RFSoCbook.com

[23] LEE-39+. DigiKey Electronics. [Online]. Available: https://www.digikey.com/en/products/detail/mini-circuits/LEE-39/13927293

[24] PMA3-43-1W+. DigiKey Electronics. [Online]. Available: https://www.digikey.com/en/products/detail/mini-circuits/PMA3-43-1W/27669457

[25] A. Hossain, S. Islam, and S. Ali. Performance Analysis of Barker Code Based on Their Correlation Property in Multiuser Environment. [Online]. Available: https://papers.ssrn.com/abstract=3812910

[26] D. A. Skoog, F. J. Holler, and S. R. Crouch, "Signals and Noise," in *Principles of Instrumental Analysis*. Cengage Learning, pp. 98–110.

[27] J. A. Gubner, "Statistcs," in *Probability and Random Processes for Electrical and Computer Engineers*. Cambridge University Press, pp. 240–272.

[28] N. J. Cassidy, "Chapter 2 - Electrical and Magnetic Properties of Rocks, Soils and Fluids," in *Ground Penetrating Radar Theory and Applications*, H. M. Jol, Ed. Elsevier, pp. 41–72. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780444533487000028

[29] A. Fediuk, D. Wilken, T. Wunderlich, and W. Rabbel, "Physical Parameters and Contrasts of Wooden Objects in Lacustrine Environment: Ground Penetrating Radar and Geoelectrics," vol. 10, no. 4, p. 146. [Online]. Available: https://www.mdpi.com/2076-3263/10/4/146

[30] K. S. Shanmugan and A. M. Breipohl, "Random Processes and Sequences," in *Random Signals: Detection, Estimation and Data Analysis*, 1st ed. Wiley, January 16, 1991, pp. 110–203.

[31] H. V. Poor, "Signal Detection in Discrete Time," in *An Introduction to Signal Detection and Estimation*, H. V. Poor, Ed. Springer, pp. 45–140. [Online]. Available: https://doi.org/10.1007/978-1-4757-2341-0_3

[32] ——, "Elements of Parameter Estimation," in *An Introduction to Signal Detection and Estimation*, H. V. Poor, Ed. Springer, pp. 141–204. [Online]. Available: https://doi.org/10.1007/978-1-4757-2341-0_4

[33] ——, "Signal Detection in Continuous Time," in *An Introduction to Signal Detection*

and Estimation, H. V. Poor, Ed. Springer, pp. 263–326. [Online]. Available: https://doi.org/10.1007/978-1-4757-2341-0_6

[34] ——, "Signal Estimation in Continuous Time," in *An Introduction to Signal Detection and Estimation*, H. V. Poor, Ed. Springer, pp. 327–385. [Online]. Available: https://doi.org/10.1007/978-1-4757-2341-0_7

[35] M. Z. Hamarat, M. Silvast, and S. Kaewunruen, "Railway turnouts and inspection technologies," in *Rail Infrastructure Resilience*, ser. Woodhead Publishing Series in Civil and Structural Engineering, R. Calçada and S. Kaewunruen, Eds. Woodhead Publishing, pp. 319–340. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128210420000162