AN ULTRA-WIDEBAND SPATIAL FILTER FOR
TIME-OF-ARRIVAL LOCALIZATION IN TUNNELS

by

Natalie A. Jones

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Graduate Department of Electrical and Computer Engineering
University of Toronto

© Copyright 2012 by Natalie A. Jones

Abstract

An Ultra-wideband Spatial Filter for
Time-of-Arrival Localization in Tunnels

Natalie A. Jones

Master of Applied Science

Graduate Department of Electrical and Computer Engineering

University of Toronto

2012

An ultra-wideband (UWB) spatial filter is proposed to mitigate multipath effects in a one-way, one-dimensional time-of-arrival (TOA) localization system for use inside a tunnel. The spatial filter is a weighted array of judiciously placed antennas and it exploits the fact that electromagnetic waves propagate as modes in a tunnel by selectively extracting these mode(s). The design of several spatial filters is presented alongside vigorous analyses to characterize the localization performance afforded by them in a noisy environment. The filters are evaluated using data from an analytical equation waveguide model, a ray tracer model and measurements. These spatial filters deliver accurate localization estimates across distance and well-designed filters can operate at higher SNRs and further distances than single sensors. Insights into successful spatial filter design are provided and this spatial filtering technique has created a new branch of multipath-aware localization systems.

Acknowledgements

Foremost, I would like to sincerely thank my supervisor Professor Sean Hum. He has provided insightful guidance, energetic helpfulness and immeasurable support throughout the entire duration of this research. With his help, I have expanded my knowledge and developed skills that will be instrumental in my future endeavours. It has been a valuable experience working with him.

I would like to thank my committee members, Professors Costas Sarris, Ravi Adve and Olivier Trescases, for their time and helpful suggestions. I also thank Thales for funding this research.

A special thanks to Professor Costas Sarris and Neeraj Sood for providing, and customizing, the ray tracing software used in this work. It has been a great collaboration. I would also like to thank Tse Chan, Tony Liang, Krishna Kishor, Neeraj Sood and Alex Wong, who helped me with my measurement campaign in one way or another. I am also grateful to all of my fellow students in the Electromagnetics Group who gave me insight, helped me troubleshoot, challenged me to think from new perspectives and supported me in many others ways.

I would like to thank my parents, Allan and Lisa, and my siblings, Teresa, Kristin, Dexter and Alina, for their support, encouragement, advice and inspiration throughout my education. I am also grateful to my extended family and friends for their support and encouragement throughout the years. Finally, I profoundly thank my partner, Erik, for his patience, support and encouragement over the last few years and, especially, throughout the duration of my research.

Natalie Jones

University of Toronto, 2012

Contents

1	Introduction	1
1.1	Ultra-wideband Communications	2
1.2	Ultra-wideband Localization	4
1.2.1	Localization Techniques	4
1.2.2	UWB Localization in a Tunnel	8
1.3	Research Goals and Outline	10
2	Modelling Propagation in a Tunnel	12
2.1	Analytical Equation Waveguide Model	13
2.1.1	Derivation of the Electromagnetic Fields in a Hollow Rectangular Dielectric Waveguide	14
2.1.2	Derivation of Source Expansion Coefficients	19
2.1.3	Summary of Waveguide Model	20
2.2	Ray Tracer Model	21
2.3	Relationship between the Models	23
3	Spatial Filter Concept and Its Use in Tunnel Localization Systems	27
3.1	Concept and Design of a Spatial Filter	28
3.2	A Spatial Filter as Part of an UWB Localization System	30
3.3	Evaluation Methods for a Spatial Filter	34
3.3.1	A Receive SNR Definition for RMSE Evaluation in Varying Noise Levels	37
3.3.2	A Transmit SNR Definition for RMSE Evaluation with Varying Transmit Power Levels	38
3.3.3	Evaluating Spatial Filter Distance Range with a Realistic Transmit SNR	39

4	Simulation and Measurement Results	41
4.1	Environment Description	41
4.2	Simulation Settings and Potential Spatial Filter Configurations	44
4.3	Measurement Campaign	50
4.4	Comparison of Measured and Simulated Electric Fields	51
4.5	Results	58
4.5.1	Performance of an Ideal Spatial Filter	59
4.5.2	Performance of Practical Spatial Filters using a Varying Receive SNR	65
4.5.3	Performance of Practical Spatial Filters using a Varying Transmit SNR	66
4.5.4	Performance of Practical Spatial Filters across Distance using a Constant Transmit SNR	71
4.5.5	Results Summary	73
5	Conclusion	75
5.1	Contributions	78
5.2	Future Work	78
A	Derivation of Fields in a Hollow Rectangular Dielectric Waveguide	80
A.1	Field Formulations	80
A.1.1	Inside Field Components	81
A.1.2	Region a Field Components	82
A.1.3	Region b Field Components	83
A.2	Imposing Boundary Conditions	84
A.3	Eigenvalue Solutions	86
A.3.1	Solving Boundary Condition Matrices	86
A.3.2	Obtaining the Eigenvalue Expressions	89
A.4	Solutions of Fields using Eigenvectors	92
B	Derivation of Weighting Coefficients due to a y-polarized Current Source	96
B.1	E_y in a Tunnel due to a y -polarized Current Source	96
B.2	E_y in a Tunnel due to a y -polarized Infinitesimal Dipole	99
	Bibliography	101

List of Tables

2.1	Transverse propagation constants.	17
4.1	Hallway properties.	43
4.2	Sample modes passed with minimal or no attenuation by various spatial filters.	49
A.1	Propagation constants.	91

List of Figures

1.1	FCC UWB EIRP emissions mask for indoor applications.	4
1.2	Position estimation in TOA systems.	5
1.3	Two-dimensional trilateration.	6
1.4	Position estimation in TDOA systems.	7
1.5	Position estimation in AOA systems.	7
1.6	Concept of a Rake receiver.	9
2.1	Tunnel geometry.	14
2.2	Tunnel geometry adopted in this work.	18
2.3	Example of ray paths using an image-based ray tracer.	22
2.4	Comparison of simulation electric fields (e_y) for the $(-0.55, -0.65, z)$ m point across distance.	24
2.5	Mode propagation angles at a receiver.	25
3.1	Sample spatial filter in tunnel cross-section shown with the electric field at one frequency.	29
3.2	Proposed one-way TOA UWB localization system.	31
3.3	Receive SNR.	37
3.4	Transmit SNR.	38
3.5	UWB mask-complying EIRP.	40
4.1	Floor plan of measurement hallway in the Galbraith Building, including transmitter and receiver locations.	43
4.2	Mode coefficients in transmitting plane.	46
4.3	Sampled normalized outputs for the idealized spatial filter.	46
4.4	Mode profile on x -axes for several modes.	48
4.5	Location of thirteen sensors and combinations of them used in various spatial filters.	48
4.6	Alignment of $n = 2$ mode extrema with $n = 1$ and $n = 3$ modes.	49

4.7	Measurement set-up.	50
4.8	Comparison of measurement and simulation electric fields (e_y) for the $x = 0$ m point across distance.	53
4.9	Comparison of measurement and simulation electric fields (e_y) for the $x = 0$ m point across distance (zoomed view).	54
4.10	Comparison of measurement and simulation electric fields (e_y) for the $x = 0.55$ m point across distance.	55
4.11	Comparison of measurement and simulation electric fields (e_y) for the $x = 0.55$ m point across distance (zoomed view).	56
4.12	Comparison of measurement data and ray tracer electric fields (e_y) from different input files at $(0, -0.65, 19, 51)$ m.	57
4.13	RMSE vs. receive SNR plots for various ideal filters.	60
4.14	Matched filter output for the (9,2) ideal filter at 4.88 m.	61
4.15	RMSE vs. transmit SNR plots for various ideal filters.	62
4.16	Filter output for various ideal filters.	63
4.17	Mode attenuation in 2.4×2.6 m tunnel.	65
4.18	RMSE vs. receive SNR plots for sample spatial filters.	66
4.19	RMSE vs. transmit SNR plots for various spatial filters using analytical equation data.	68
4.20	RMSE vs. transmit SNR plots for (17,2)-type spatial filters using analytical equation and ray tracer data.	69
4.21	RMSE vs. transmit SNR plots for (9,2)-type spatial filters using analytical equation and ray tracer data.	69
4.22	RMSE vs. transmit SNR plots for (8,2)-type spatial filters using analytical equation and ray tracer data.	69
4.23	RMSE vs. transmit SNR plots for spatial filters using measurement data.	70
4.24	RMSE vs. transmit SNR plot at 25 m for a (1,1) spatial filter designed for a center transmitter.	70
4.25	RMSE vs. distance for various spatial filters.	72
4.26	Energy decay in waveguides.	73

List of Abbreviations

AWGN:	Additive White Gaussian Noise
BAVA:	Balanced Antipodal Vivaldi Antenna
BN:	Base Node
CRB:	Cramér-Rao Bound
DS-CDMA:	Direct Sequence Code Division Multiple Access
EIRP:	Effective Isotropic Radiated Power
ESD:	Energy Spectral Density
FCC:	Federal Communications Commission
GNSS:	Global Navigation Satellite System
GPS:	Global Positioning System
IR:	Impulse Radio
LNA:	Low Noise Amplifier
MSE:	Mean Square Error
MIMO:	Multiple-Input Multiple-Output
OFDM:	Orthogonal Frequency Division Multiplexing
PEC:	Perfectly Electrically Conducting
PNA:	Precision Network Analyzer
PRF:	Pulse Repetition Frequency
PSD:	Power Spectral Density
RSS:	Received Signal Strength
RMSE:	Root Mean Square Error
Rx:	Receiver
SNR:	Signal-to-Noise Ratio
TN:	Target Node

TDOA:	Time Difference of Arrival
TOA:	Time of Arrival
Tx:	Transmitter
UWB:	Ultra-Wideband
WLAN:	Wireless Local Area Network

List of Symbols

a	x -dimension of tunnel
α_{mn}	Mode attenuation constant
B	Bandwidth
b	y -dimension of tunnel
β_{mn}	Mode phase constant
C_{mn}	Mode weighting coefficient
c	Speed of light
χ^2	Second moment of the spectrum $P(f)$
D	Antenna directivity
$\delta(\cdot)$	Dirac delta function
$\vec{E}, E_x, E_y, E_z$	Electric field intensity
E_{mn}^{eig}	y -oriented electric field eigenmode
E_p	Energy of one pulse
$E[\cdot]$	Expected value
$ESD(f)$	Energy spectral density
$e_y(t)$	Time-domain, y -oriented electric field intensity
ϵ	Permittivity
η	Free-space wave impedance
$\mathfrak{F}(\cdot)$	Fourier transform
f	Frequency
G	Antenna gain
γ	Threshold level in threshold detector
$\vec{H}, H_x, H_y, H_z$	Magnetic field intensity
I_0	Antenna excitation current

K	Number of receivers in a spatial filter
k	Wavenumber
k_B	Boltzmann's constant
l	Infinitesimal dipole's physical length
$\vec{l}_{eff}$	Vector effective length of an antenna
λ	Wavelength
m	Mode index in the x -dimension
μ	Permeability
N, N_f, N_x, N_y	Number of points for various quantities
NF	Noise figure
N_0	One-sided PSD of AWGN
n	Mode index in the y -dimension
P_{Tx}	Transmit power
P_{noise}	Noise power
$P(f)$	Frequency spectrum of $p(t)$
$p(t)$	Transmitted pulse shape
ϕ_x, ϕ_y	Appropriate angles for even and odd modes
R_{rad}	Radiation resistance
ρ	Approximated correlation coefficient
$S_x(f)$	PSD of $x(t)$
SNR_{Tx}	Transmit signal-to-noise ratio
SNR_{Rx}	Receive signal-to-noise ratio
σ	Conductivity
T	Temperature
T_p	Transmitted pulse width
TOA_{est}	Estimated time-of-arrival
$TOA_{theoretical}$	Theoretical time-of-arrival
t	Time
θ_i	Angle of incidence
θ_{mn}	Mode arrival angle
$V_{oc}(f)$	Open-circuit antenna voltage
$v_k(t)$	Voltage at k^{th} receiver in a spatial filter

$v_{tot}(t)$	Voltage output by a spatial filter
$v_{g,mn}$	Mode group velocity
U, U_{decomp}	Energy
W_k	Weight of k^{th} receiver in a spatial filter
ω	Angular frequency
$x(t)$	Train of $p(t)$ pulses
x_k	x -dimension of the k^{th} receiver in a spatial filter
x_0	x -dimension of the transmitting antenna location
y_k	y -dimension of the k^{th} receiver in a spatial filter
y_0	y -dimension of the transmitting antenna location
$y(t)$	Matched filter output in the TOA localization system

Chapter 1

Introduction

As wireless localization systems are increasingly important in daily lives, it is vital that localization systems continue to strive to deliver high performance and accuracy in all environments. Localization systems have been used for decades and the most ubiquitous of these are outdoor radar technologies, for military and aircraft applications, and Global Navigation Satellite Systems (GNSS), such as the Global Positioning System (GPS). However, these types of localization systems cannot be exploited to localize objects in indoor or enclosed environments. Therefore, alternative localization strategies must be implemented in specific environments. One enclosed environment, in which few localization strategies have been proposed, is a tunnel. Precise tunnel localization is paramount in several situations, such as object and vehicle tracking in tunnel sections of mines, vehicular monitoring in automobile tunnels, train positioning in train tunnels and object or pedestrian tracking in large corridors.

In underground mines, wireless sensor networks and radio-frequency tags have been proposed to locate objects [1], [2], [3], [4]. In train tunnels, train locations can be monitored using technologies such as transponders, radio-frequency tags and infrared sensors [5]. However, some of these mine and train tunnel systems may not continuously track an object, have limited ranging abilities (considering distance and accuracy), require complicated positioning algorithms and require a large investment for system installation. Thus, the motivation for this research is to provide a simple tunnel localization system that has the ability to operate over long distances while delivering accurate localization, and potentially provide real-time information.

Ultra-wideband communications is a forefront candidate for localization applications because short ultra-wideband pulses enable centimeter-level accuracy in these systems. UWB technology has been recognized as a promising and suitable technology for tunnel positioning because of potential accuracy,

ranging limits and low system costs [5]. However, to implement an UWB localization system in a tunnel multipath propagation must be considered because there are many significant multipath arrivals and some will interfere with one another.

In this research an UWB localization system is proposed for a tunnel environment. The novelty of the proposed localization system is that it entails the use of a spatial filter, which is essentially a weighted receiving array of antennas. The spatial filter is the focus of this work.

1.1 Ultra-wideband Communications

An ultra-wideband signal is defined as having an absolute bandwidth of at least 500 MHz or having a fractional bandwidth of at least 20%, where the fractional bandwidth is given by

$$B_{frac} = 2 \frac{f_h - f_l}{f_h + f_l} = \frac{f_h - f_l}{f_{avg}}, \quad (1.1)$$

and f_h , f_l and f_{avg} are the upper, lower and average frequencies for the -10 dB emission points of the effective isotropic radiated power (EIRP), respectively [6].

There are several properties inherent to UWB signals [7], [8]. First, through Fourier analysis a wide-band signal corresponds to a short time-domain pulse. This has several advantages. From a capacity point of view, a signal is essentially compressed and more pulses can be transmitted in a given amount of time. Thus, UWB signals can support high data rate communications. From a resolution point of view, fine UWB pulses reflect off of objects, in imaging and localization applications, and can be appropriately detected to achieve centimeter-level precision. From a channel propagation point of view, UWB multipath components are typically resolvable, enabling simple time-gating receivers to be used to remove multipath, or most sophisticated Rake receivers to extract energy from the multipath components.

To illustrate the potential of UWB signals further, consider Shannon's single-channel capacity equation where a channel's capacity, C , can be related to its bandwidth, B , and SNR; that is,

$$C = B \log_2(1 + \text{SNR}) \text{ bits/s}, \quad (1.2)$$

where additive white Gaussian noise (AWGN) is assumed [9]. From this equation, which is an upper limit, it can be seen that UWB systems are extremely attractive for high data rate applications. Additionally, it is evident that in order to maintain the same channel capacity as a narrowband system with B_{NB} and SNR_{NB} , an UWB system that has B_{UWB} will have a lower SNR_{UWB} . Thus, UWB signals can

be transmitted at lower power levels than narrowband systems while preserving channel capacity. By operating at low power levels, and in the situation that carrier-less transmission schemes are adopted, UWB systems can be implemented with low-cost transmitters and receivers.

The concept of UWB signals and systems is not new. For example, UWB radar systems have used short high power pulses since the 1960s [7]. However, interest in ultra-wideband has rapidly grown since 2002 when the Federal Communications Commission (FCC) authorized unlicensed use of spectrum for UWB signals providing that power emission conditions are met. The FCC also outlined categories of UWB systems (for example, ground penetrating radar, medical imaging, indoor and outdoor applications, vehicular radar, etc.) and provided specific power constraints and implementation conditions for each type. In the case of tunnel localization, indoor system constraints are assumed. Power constraints on UWB signals are necessary to ensure that legacy systems' functionality are not compromised¹. To legacy systems, UWB signals appear as low-level white noise.

UWB signal emissions must not exceed an effective isotropic radiated power mask defined by the FCC. This mask is shown for indoor communications in Figure 1.1. The maximum EIRP defined by the mask is -41.3 dBm, which is measured over a 1 MHz resolution bandwidth in the range of 3.1 to 10.6 GHz. EIRP is a far-field quantity and is defined as the product of the power supplied to the antenna and antenna gain in any direction. The mask shown in Figure 1.1 is an average emissions mask, and there are also peak power emission limits for UWB signals. Peak power limits are 0 dBm over a 50 MHz bandwidth in the 3.1 to 10.6 GHz bandwidth. Most UWB emissions are average limited if they have a pulse repetition frequency (PRF) above 1 MHz [6].

There are numerous UWB applications. The main categories of systems have been briefly mentioned but the breadth of these categories has not been fully elucidated. High data rate applications use signalling schemes such as Direct Sequence Code Division Multiple Access (DS-CDMA) and Orthogonal Frequency Division Multiplexing (OFDM) for short-range communications between consumer electronics, industrial equipment, etc. [7]. However, the potential for these systems has not been realized for a number of reasons and interest for UWB high data rate applications has been shifted to around 60 GHz where higher power levels can be used. Nevertheless, non-communication applications of UWB have since attracted considerable interest. UWB imaging applications range from cancer screening, non-destructive testing and more. In these applications, as well as for localization and sensing applications, low data rates and impulse radio (IR) signalling schemes are commonly used. In IR UWB signalling short UWB pulses are transmitted with or without a modulation scheme. Examples of sensing applications are body area networks, search and rescue disaster aids, inventory control trackers and smart home applications [8].

¹Examples of legacy systems are GPS, wireless local area networks (WLAN), cellular telephone systems, etc.

Figure 1.1: FCC UWB EIRP emissions mask for indoor applications.

Localization applications will be reviewed in the following section.

1.2 Ultra-wideband Localization

It has been established that UWB is extremely well suited for positioning applications and there have been many UWB localization systems proposed in recent years. Prior to exploring possible localization system implementations in a tunnel, types of localization are reviewed in Section 1.2.1. In Section 1.2.2 current and potential UWB localization systems are presented.

1.2.1 Localization Techniques

A localization or positioning system is defined as a system that obtains location information about an object by using a limited amount of prior knowledge (for example, reference sensor positions) and by using and processing measurements. Types of measurements may be time of arrival, time difference of arrival (TDOA), angle of arrival (AOA) and received signal strength (RSS). In the following types of measurements and processing methods that positioning systems use are reviewed.

Time of arrival localization systems may make one-way or two-way time-of-flight measurements. Figures 1.2(a) and 1.2(b) show the bases for one-way and two-way TOA systems, respectively. In one-way TOA systems a transmitting node² (Tx Node) emits a pulse at time t_1 and that pulse is detected

²A node is a term used to describe a sensor or antenna (transmitting or receiving) in a localization system.

at a receiving node (Rx Node) at a later time, t_2 . The resultant time-of-flight time, $t_f = t_2 - t_1$, is used to calculate the distance between the nodes: $d = t_f c$, where c is the speed of light. In two-way TOA systems the round-trip time of a pulse from one node (Tx/Rx Node) to another (Target Node) and back is used to calculate the distance between the nodes. Two-way TOA systems do not rely on synchronized clocks at each node, which alleviates the implementation challenge facing one-way TOA systems. Additionally, the target node may be passive and reflect the transmitted pulse back, or it may be a transponder and actively emit a pulse back.

(a) One-way TOA localization.

(b) Two-way TOA localization.

Figure 1.2: Position estimation in TOA systems.

In a three-dimensional TOA system trilateration, using three or four known transmitting nodes, can be used to find the exact position of an unknown target node. An example of two-dimensional trilateration is shown in Figure 1.3 where three known nodes, labelled BN for base node, and one target node (TN) is shown. Assuming a one-way TOA system, the radial distance between BN 1 and TN can be calculated using $d_1 = t_{f,1}c$. With only this measurement TN may be located anywhere on the circumference of a circle with the radius d_1 . By using the time-of-flight measurement from BN 2 to TN, $d_2 = t_{f,2}c$ and the TN must be located at one of the two intersections of the circles with radii d_1 and d_2 . By using a third measurement from BN 3 $d_3 = t_{f,3}c$ is determined and the TN must be located at the intersection of all three circles, denoted by the star in Figure 1.3.

Trilateration can be expanded to three-dimensions where the intersection of two spheres forms a circle, and the intersection of three spheres is two points. Only one point is valid, as the other may be outside the region of interest; thus, the exact point of the target node can be easily determined.

Figure 1.3: Two-dimensional trilateration.

Alternatively, a fourth base node can be used to find one intersection point and its use also eliminates the problem of unsynchronized clocks between the base nodes and target node in one-way systems [10]. GPS uses trilateration where base nodes are orbiting satellites with known positions and highly accurate atomic clocks.

Time difference of arrival localization is the second type of system considered. In TDOA systems synchronized receiving nodes, with knowledge of each others' position, each receive a signal from a target transmitting node and then time differences are used to calculate the target node's location. The time difference between signals at two receiving nodes can be used to define a branch of a hyperbola for the possible position of the transmitting node³ [11]. The foci of the hyperbola are at the two receiving nodes. Figure 1.4 shows the resultant hyperbola for a time difference between BN 1 and BN 2. By using a second TDOA measurement, with BN 1 and BN 3, a second hyperbola can be formed. The transmitting target node is located at the intersection of the two hyperbolas, which is shown by a star in Figure 1.4. This scheme can be expanded to three-dimensions where hyperboloids are formed.

Angle of arrival localization systems use an antenna array that receives a signal from an unknown transmitting node and differences in the signal arriving at array elements are used to calculate the transmitter's position [8], [11]. More specifically, the angle of the incoming plane wave, α , is determined because arrivals at each antenna have an offset of $d/c \sin \alpha$ seconds from the neighbouring antenna, assuming the geometry in Figure 1.5(a). In narrowband systems the phase differences between antenna

³Two-dimensions are assumed.

Figure 1.4: Position estimation in TDOA systems.

signals can be used to determine the AOA. In UWB systems time delayed versions of received signals should be considered [8].

To find the transmitting node's location in an AOA system triangulation can be used. For example, the AOA of two receiving nodes can be found and then intersecting lines reveal the transmitter's location, as Figure 1.5(b) shows.

Figure 1.5: Position estimation in AOA systems.

In received signal strength localization systems the received signal strength of a signal detected at a known receiving node is used to predict the unknown transmitting node's location [8], [12]. By using path loss models a given signal strength is associated with a specific distance. The accuracy of an RSS system depends on the path loss model and that the assumption that signal strength decreases as mode separation distance increases is true. Trilateration can be used with RSS measurements to find a target

node's exact location. Additionally, RSS localization is more common in narrowband systems where TOA techniques are difficult to implement.

There are also other types of localization systems and position estimation algorithms. For example, only geometrical positioning estimation algorithms, such as trilateration, were presented. There are many statistical and mapping techniques as well [8], [11]. The former efficiently solves position-related equations with or without noise, and the latter uses a database which consists of previously estimated signal parameters at known positions.

In the case of UWB localizing in a tunnel time-based ranging (TOA or TDOA) is most suitable as it exploits the potential of very short UWB pulses. RSS systems rely on ideal propagation environments and are not robust enough for imperfect situations. Multipath in tunnel environments will seriously affect AOA system accuracy and the receiver complexity is high. Between TOA and TDOA, TOA is preferred in a tunnel because multipath may interfere with TDOA measurements, unless receivers are placed very close together.

In TOA systems there are many ways to extract the TOA of a received pulse. First, a matched filter or correlator is typically used, and then the output of it is processed to find a maximum peak or threshold mark at which the TOA is recorded. There are many other search algorithms for finding the precise TOA [8], [12]. In TOA localization systems there are several potential sources of error. For example, error sources may be multipath, multiple user interference, the inability to fully digitally resolve UWB pulses and clock inaccuracies.

1.2.2 UWB Localization in a Tunnel

Possible implementations for TOA tunnel localization systems are now explored. In the tunnel localization system it is of primary interest to localize along the tunnel length; that is, only one-dimensional localization is required. The simplest implementation of this system would consist of a transmitting and receiving antenna spaced a distance apart. However questions are, where should the antennas be placed and how will the system handle multipath? Multipath arrivals are significant and vary over separation distances in tunnels, and they can corrupt TOA algorithms. Therefore, a two antenna TOA localization system that does not consider multipath arrivals is suboptimal.

There have been many recently proposed and tested time-based UWB localization systems [13], [14], [15], [16]. These systems perform two- or three-dimensional localization in indoor environments where the area or volume considered is limited. In some situations, optimal receiver positions were determined [13], [14]. These systems are, however, not designed for use in a tunnel, where propagation is very

different and the placement of transmitters and receivers is constrained.

Another potential solution for tunnel localization could use a Rake receiver [17], [18], [19]. Rake receivers have been proposed to mitigate multipath effects by combining multipath components to create a signal with a larger SNR at the receiver output, in comparison to a single receiver. Figure 1.6 shows a Rake receiver with N taps that can add N multipath components. The taps or fingers in a Rake receiver each have their own delay and weight that are designated to extract a specific multipath component. The total number of taps depends on the design used, as there are many types of Rake receivers [17]. There are, however, a couple substantial challenges in using Rake receivers. First, the Rake receiver must be trained in order to arrive at suitable tap delays and weights. Second, assuming that the Rake is digitally realized, very high sampling rates must be used so that tap delays are appropriate. For example, if the entire UWB range is used in signal transmission, a minimum sampling rate of 20 Gsamples/s would be required and to resolve the closely spaced tunnel multipaths an even higher sampling rate would be required for accurate resolution. As a result, Rake receiver implementations are very complex and a simpler receiver implementation is an objective in this research.

Figure 1.6: Concept of a Rake receiver.

UWB communication and localization in tunnels has been considered [20], [21], [22]. In one case, multiple access performance and time reversal was considered, where time reversal pre-filtering was used to help mitigate interference and multipath effects [20]. However, estimating channels is difficult in time reversal solutions. In another case, a TDOA/RSS localization system was used to track an object in a mine tunnel. However, several base nodes were required to track the moving mobile node and the mobile node was only 2 to 8 m from a base node, and the channel model considered was not tailored to the tunnel environment. In the final case considered, the performance of an UWB multiple-input multiple-output (MIMO) system operating in a tunnel section of a mine was investigated. However, the

focus of the study was on MIMO system characteristics.

In this research it is proposed to implement a single-user, one-dimensional localization system that accounts for tunnel propagation phenomenon. In this localization system a single transmitting node is proposed to be placed along a tunnel wall, such that vehicular traffic or tunnel activities can move unconstrained in the tunnel. A receiving node, that consists of a filtering array of antennas or spatial filter, is proposed to be placed in the tunnel's cross-section at a variable position. The receiving node may move longitudinally in the tunnel. The filter shall mitigate multipath effects and capitalize on the fact that electromagnetic waves propagate as modes in a tunnel. In contrast to Rake receiver, time-reversal receiver and other receiver implementations, this receiving node uses comparatively simple processing of the received signals to perform TOA calculations.

1.3 Research Goals and Outline

The motivation for this research is to improve the accuracy of one-way TOA-based UWB localization techniques in tunnels using spatial filters. As such, filtering techniques shall be investigated and the goals of this thesis are as follows.

1. To design a filter(s) to improve the performance of one-way TOA systems in a tunnel. Straight, rectangular tunnels with discontinuities need only be considered for the proof-of-concept filter(s). In current literature no such systems have been proposed and, thus, this is a novel system.
2. To investigate the performance of the filter compared to systems not equipped with a filter, in order to gauge the relative performance improvements afforded by the filter. Moreover, filter performance shall be verified with simulations and measurements, and compared to theoretical bounds.
3. To explore different filter configurations to lead to practically realizable filters. Practical filters must utilize only a low number of receiving antennas for implementation.

The organization of this thesis is as follows. Chapter 2 provides an overview of wave propagation in a tunnel. Two propagation models are introduced: an analytical equation-based waveguide model in which electromagnetic waves are modelled as modes propagating in a waveguide, and an image-based ray tracer model in which electromagnetic waves are modelled as rays. The waveguide model is used as the premises on which the spatial filter is designed. Chapter 3 presents the spatial filter concept, its design and evaluation metrics for it. Chapter 4 provides sample spatial filter designs, measurement details and results showing data from the waveguide model, ray tracer model and measurements. A thorough

discussion of results is also provided in the latter half of Chapter 4. Finally, Chapter 5 concludes this work and proposes future extensions of this work.

Chapter 2

Modelling Propagation in a Tunnel

The knowledge of how electromagnetic waves propagate in a tunnel is essential in designing an UWB localization system to operate within it. Wave propagation in a tunnel is unlike wave propagation in a terrestrial environment, where empirical models are typically utilized in radio system design. In the case of terrestrial UWB communications, models in the 802.15.3 standard that are based primarily on the Saleh-Valenzuela model are commonly used [23].

Electromagnetic wave propagation in a tunnel was first studied in the 1970s [24], [25], [26]. An equivalent study on infrared or optical electromagnetic waves propagating in a hollow rectangular dielectric waveguide was also performed in this same time period [27]. Irregardless of the structure, in both cases the wavelength of operation is much smaller than the structure's cross-sectional dimensions, and as such the structures act as waveguides in which modes naturally propagate. Concurrently, it was also proposed to use cables, pipes, etc. to help guide low-frequency modes in a tunnel [28]. Leaky coaxial cables, or feeders, were also proposed for tunnel communications and the use of leaky feeders has extended up to 2 GHz [26], [29]. However, in both of these cases extensive infrastructure must be in place and the systems are not designed for localization. The concept of modelling a tunnel as a waveguide and launching naturally propagating modes in it with a transmitting antenna is more appropriate for tunnel localization, especially for UWB signals.

Modes propagate efficiently in tunnels if several conditions are met, such as if the wavelength of operation is much smaller than the tunnel's dimensions. The modes in this oversized, dielectric waveguide are lossy and attenuate as they propagate through the tunnel, unlike the modes in a perfectly conducting rectangular waveguide. Furthermore, the modes are hybrid modes but can be cast into a simplified form that is used in propagation models. Waveguide theory has been confirmed with measurements in several

studies [24], [29], [30], [31].

Modelling electromagnetic waves as modes in a waveguide is one deterministic model available to study tunnel propagation. Alternative models include full-wave solvers, ray tracer models and hybrid models. In full-wave solvers, such as those used in the Finite-Difference Time-Domain method, the propagation environment is described and electromagnetic fields are found in the entire domain, which may not be feasible as memory requirements are very high for large tunnel dimensions and wavelengths at the centimeter scale. In ray tracer models, optical rays that take different paths between the transmitter and receiver are used to characterize radiation [32], [33]. Hybrid models may also be used and they combine ray theory and waveguide theory [31], [32].

In this research the waveguide model is used foremost to study propagation in a rectangular tunnel and, as it will be seen, waveguide theory will be used to formulate the basis of the spatial filter. A ray tracer model will be used to verify the waveguide theory propagation and vice versa. Additionally, both models will be compared to measurements taken in a hallway in Chapter 4. In the following, waveguide theory is introduced and necessary derivations are provided, and then further details on ray tracer modelling are provided.

2.1 Analytical Equation Waveguide Model

An analytical equation, based on waveguide theory, to find the dominant electric field component at any point in the tunnel is now presented. By solving the analytical equation at field points within the tunnel, electromagnetic wave propagation in the tunnel can be characterized.

The electromagnetic field components, including attenuation and phase constants, for modes within a rectangular tunnel were first presented in studies of coal mine tunnels [24], and later presented more formally for a generic hollow rectangular dielectric waveguide [27]. In Section 2.1.1 the derivation of these components is provided. In these studies weighting coefficients for each mode were not provided. However, in a recent study weighting coefficients for modes were provided considering a transmitting antenna source expansion [34]. Mode weighting coefficients can also be found using other field expansions [31]. The weighting coefficient of a given mode depends on the transmitter and its location in the tunnel's cross-section. In this research it is of interest to weight modes based on a realistic UWB antenna excitation, which other weighting coefficient derivations did not provide. In order to do this, Green's function is solved subject to tunnel conditions in Section 2.1.2.

2.1.1 Derivation of the Electromagnetic Fields in a Hollow Rectangular Dielectric Waveguide

The derivation of the approximate characteristic modes, or eigenmodes, in a tunnel or hollow rectangular dielectric waveguide is summarized in the following and detailed in Appendix A. This derivation is provided because the only other rigorous analysis found contained numerous errors and a brief summary of steps taken [27]. The resultant fields agree with fields used in [24], [30], [31], [34], [35].

In this derivation the geometry and region definition is in accordance to that used in [27]; however, later equations are transformed to correspond to a slightly different geometry. The tunnel is defined to have a cross-section of $2a \times 2b$ where the center of the waveguide is at the origin of the $x - y$ plane. The geometry of the guide is shown in Figure 2.1, in which complex permittivities are labeled in the regions they are present. The permittivity of free-space, ϵ_0 , is present in the central core of the waveguide and the permeability of free-space, μ_0 , is assumed in all regions. The complex relative permittivities are $\epsilon_a = \epsilon_{a,r}\epsilon_0 + \frac{\sigma_a}{j2\pi f}$ and $\epsilon_b = \epsilon_{b,r}\epsilon_0 + \frac{\sigma_b}{j2\pi f\epsilon_0}$, where σ is the corresponding conductivity and f is frequency.

Figure 2.1: Tunnel geometry.

An exact analytical solution is not possible for this geometry due to boundary conditions, but an approximate solution can be developed if the following assumptions are used:

1. The dimensions of the guide are much larger than the wavelength of operation and the mode order is not too high; that is

$$\left(\frac{n\lambda}{4b}\right) \ll 1 \text{ and } \left(\frac{m\lambda}{4a}\right) \ll 1, \quad (2.1)$$

where m, n are the mode indices for the x and y components, respectively and λ is the wavelength

in free-space.

2. The dielectric constants satisfy

$$\text{for } x\text{-polarized modes: } \begin{cases} \sqrt{|\bar{\epsilon}_b - 1|} \gg \frac{n\lambda}{4b} \\ \frac{\sqrt{|\bar{\epsilon}_a - 1|}}{\bar{\epsilon}_a} \gg \frac{m\lambda}{4a} \end{cases} \quad (2.2)$$

and

$$\text{for } y\text{-polarized modes: } \begin{cases} \sqrt{|\bar{\epsilon}_a - 1|} \gg \frac{m\lambda}{4a} \\ \frac{\sqrt{|\bar{\epsilon}_b - 1|}}{\bar{\epsilon}_b} \gg \frac{n\lambda}{4b}, \end{cases} \quad (2.3)$$

where $\bar{\epsilon}_b = \epsilon_b/\epsilon_0$ and $\bar{\epsilon}_a = \epsilon_a/\epsilon_0$.

3. The boundary conditions are matched only along the four sides and not in the corners of the waveguide; that is, the boundary conditions are effectively decoupled. Additionally, the two side walls have the same dielectric properties, and the ceiling and floor have the same dielectric properties.

In order to find the fields in the waveguide the assumed form for the field components in each region in Figure 2.1 must be written. In the interior region of the waveguide a sinusoidal variation is assumed for the transverse dimensions, and in exterior regions a decaying exponential behaviour away from the waveguide is assumed. By using Helmholtz's equations and the source-free curl equations longitudinal fields can be assumed and then transverse fields can be found from them. As a result, in the inner region of the waveguide

$$E_z^i = \mathcal{E}_z^i \cos(k_x^i x + \phi_x) \cos(k_y^i y + \phi_y) e^{-jk_z z}, \quad (2.4)$$

$$H_z^i = \mathcal{H}_z^i \sin(k_x^i x + \phi_x) \sin(k_y^i y + \phi_y) e^{-jk_z z}, \quad (2.5)$$

$$E_x^i = \frac{j\omega\mu_0}{k_0^2 - k_z^2} \left[\frac{k_z k_x^i}{\omega\mu_0} \mathcal{E}_z^i - k_y^i \mathcal{H}_z^i \right] \sin(k_x^i x + \phi_x) \cos(k_y^i y + \phi_y) e^{-jk_z z}, \quad (2.6)$$

$$E_y^i = \frac{j\omega\mu_0}{k_0^2 - k_z^2} \left[\frac{k_z k_y^i}{\omega\mu_0} \mathcal{E}_z^i + k_x^i \mathcal{H}_z^i \right] \cos(k_x^i x + \phi_x) \sin(k_y^i y + \phi_y) e^{-jk_z z}, \quad (2.7)$$

$$H_x^i = -\frac{j\omega\epsilon_0}{k_0^2 - k_z^2} \left[k_y^i \mathcal{E}_z^i + \frac{k_z k_x^i}{\omega\epsilon} \mathcal{H}_z^i \right] \cos(k_x^i x + \phi_x) \sin(k_y^i y + \phi_y) e^{-jk_z z}, \quad (2.8)$$

$$H_y^i = \frac{j\omega\epsilon_0}{k_0^2 - k_z^2} \left[k_x^i \mathcal{E}_z^i - \frac{k_z k_y^i}{\omega\epsilon} \mathcal{H}_z^i \right] \sin(k_x^i x + \phi_x) \cos(k_y^i y + \phi_y) e^{-jk_z z}, \quad (2.9)$$

where $k_0 = \sqrt{(k_x^i)^2 + (k_y^i)^2 + (k_z)^2} = \sqrt{\omega^2 \mu_0 \epsilon_0}$ is the wavenumber in the inner region, k_z is the wavenumber in the z direction, k_x^i is the interior wavenumber in the x direction, k_y^i is the interior wavenumber in the y direction, $\mathcal{E}_z^i$ is the interior electric field amplitude in the z direction and $\mathcal{H}_z^i$ is the

internal magnetic field amplitude in the z direction. The notations ϕ_x and ϕ_y are used to account for even and odd symmetry cases. That is, if $\phi_x, \phi_y = 0$ there is even symmetry, and if $\phi_x, \phi_y = \pi/2$ there is odd symmetry. Field components in regions a and b can be found in Appendix A.1, alongside further details for the field formulation.

Boundary conditions must now be enforced at the interface of the tunnel's interior and walls. All media are dielectric media and the permeability of free-space is assumed in all regions. It is necessary to find expressions for k_x^i and k_y^i by solving boundary conditions, but there are also many other quantities in field equations and many are coupled to k_x^i or k_y^i . The approach taken to find k_x^i and k_y^i is formulating matrix equations and then finding simple expressions in which k_x^i and k_y^i are contained and can be solved for.

To find the matrix equations boundary conditions are met at the $y = \pm b$ and $x = \pm a$ boundaries. For example at $y = \pm b$ tangential electric field intensity components, tangential magnetic field intensity components and normal magnetic field density components are equal. By using these three conditions the following matrix relation can be written

$$\begin{bmatrix} \frac{\epsilon_0 k_y^i}{\Delta k_i^2} - \frac{\epsilon_b \alpha_y^b}{\Delta k_b^2} \cot(k_y^i b + \phi_y) & \frac{k_z k_x^i}{\omega} \left[\frac{1}{\Delta k_i^2} - \frac{1}{\Delta k_b^2} \right] \\ \frac{k_x^i \epsilon_b}{\Delta k_b^2} - \frac{k_x^i \epsilon}{\Delta k_i^2} & \frac{k_z}{\omega} \left[\frac{k_y^i}{\Delta k_i^2} + \frac{\alpha_y^b}{\Delta k_b^2} \tan(k_y^i b + \phi_y) \right] \end{bmatrix} \times \begin{bmatrix} \mathcal{E}_z^i \\ \mathcal{H}_z^i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad (2.10)$$

where $\Delta k_i^2 = k_0^2 - k_z^2$, $\Delta k_b^2 = k_b^2 - k_z^2$, k_b is the wavenumber in region b, α_y^b is the region b wavenumber (attenuation) in the y direction and ω is the angular frequency. Similarly, a matrix relation can be written by matching fields at the $x = \pm a$ boundaries and it is

$$\begin{bmatrix} -\frac{\epsilon_a k_y^i}{\Delta k_a^2} + \frac{\epsilon_o k_y^i}{\Delta k_i^2} & \frac{k_z}{\omega} \left[\frac{\alpha_x^a}{\Delta k_a^2} \tan(k_x^i a + \phi_x) + \frac{k_x^i}{\Delta k_i^2} \right] \\ \frac{\alpha_x^a \epsilon_a}{\Delta k_a^2} \cot(k_x^i a + \phi_x) - \frac{k_x^i \epsilon_0}{\Delta k_i^2} & \frac{k_z k_y^i}{\omega} \left[-\frac{1}{\Delta k_a^2} + \frac{1}{\Delta k_i^2} \right] \end{bmatrix} \times \begin{bmatrix} \mathcal{E}_z^i \\ \mathcal{H}_z^i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad (2.11)$$

where $\Delta k_a^2 = k_a^2 - k_z^2$, k_a is the wavenumber in region a and α_x^a is the region a wavenumber (attenuation) in the x direction. More details on the matrices' formulation are provided in Appendix A.2.

To find simple expressions in which k_x^i and k_y^i are contained and can be easily solved for, the determinants of the matrices are used. By setting determinants to zero unique solutions can be found for the matrices. The determinants can be simplified using expressions relating different wavenumbers and by solving the determinants it is found that

$$\tan(k_y^i b + \phi_y) = \begin{cases} \alpha_y^b / k_y^i \\ -k_y^i \epsilon_b / \alpha_y^b \end{cases}, \quad (2.12)$$

from Equation (2.10), and

$$\tan(k_x^i a + \phi_x) = \begin{cases} \alpha_x^a / k_x^i \\ -k_x^i \bar{\epsilon}_a / \alpha_x^a \end{cases}, \quad (2.13)$$

from Equation (2.11). Details for these solutions are provided in Appendix A.3.

To find k_x^i and k_y^i using Equations (2.12) and (2.13) assumptions that $|\bar{\epsilon}_a \frac{k_x^i}{\alpha_x^a}| \ll 1$ or $|\frac{\alpha_x^a}{k_x^i}| \gg 1$ and $|\bar{\epsilon}_b \frac{k_y^i}{\alpha_y^b}| \ll 1$ or $|\frac{\alpha_y^b}{k_y^i}| \gg 1$, respectively, need to be made. Then it must be assumed that the term in the tangent expression is approximately $\pi/2$ or π , or a multiple thereof. By using these assumptions k_x^i and k_y^i can be found, which is provided in Appendix A.3, and they are listed for all possible cases in Table 2.1.

		For x-polarized modes		For y-polarized modes	
k_x^i	$\phi_x = 0$	$\frac{m\pi}{2a}$	$1 + j \frac{\bar{\epsilon}_a}{k_0 a \sqrt{\bar{\epsilon}_a - 1}}$	$\frac{m\pi + \pi/2}{2a}$	$1 + \frac{j}{k_0 a \sqrt{\bar{\epsilon}_a - 1}}$
	$\phi_x = \pi/2$	$\frac{m\pi + \pi/2}{2a}$	$1 + j \frac{\bar{\epsilon}_a}{k_0 a \sqrt{\bar{\epsilon}_a - 1}}$	$\frac{m\pi}{2a}$	$1 + \frac{j}{k_0 a \sqrt{\bar{\epsilon}_a - 1}}$
k_y^i	$\phi_y = 0$	$\frac{n\pi + \pi/2}{2b}$	$1 + \frac{j}{k_0 b \sqrt{\bar{\epsilon}_b - 1}}$	$\frac{n\pi}{2b}$	$1 + j \frac{\bar{\epsilon}_b}{k_0 b \sqrt{\bar{\epsilon}_b - 1}}$
	$\phi_y = \pi/2$	$\frac{n\pi}{2b}$	$1 + \frac{j}{k_0 b \sqrt{\bar{\epsilon}_b - 1}}$	$\frac{n\pi + \pi/2}{2b}$	$1 + j \frac{\bar{\epsilon}_b}{k_0 b \sqrt{\bar{\epsilon}_b - 1}}$

Table 2.1: Transverse propagation constants.

With the knowledge of k_x^i and k_y^i , the wavenumber in the z dimension, k_z , can be found using $k_z = \sqrt{k_0^2 - (k_x^i)^2 - (k_y^i)^2}$. From k_z the phase constant and attenuation constant can be found. This solution is provided in Appendix A.4 alongside final field components. In summary, the attenuation constants for the x -polarized and y -polarized modes are

$$\alpha_{mn} = \frac{1}{a} \left(\frac{m\pi}{2ak_0} \right)^2 \text{Re} \left\{ \frac{1}{\sqrt{\bar{\epsilon}_a - 1}} \right\} + \frac{1}{b} \left(\frac{n\pi}{2bk_0} \right)^2 \text{Re} \left\{ \frac{\bar{\epsilon}_b}{\sqrt{\bar{\epsilon}_b - 1}} \right\}, \quad (2.14)$$

$$\alpha_{mn} = \frac{1}{a} \left(\frac{m\pi}{2ak_0} \right)^2 \text{Re} \left\{ \frac{\bar{\epsilon}_a}{\sqrt{\bar{\epsilon}_a - 1}} \right\} + \frac{1}{b} \left(\frac{n\pi}{2bk_0} \right)^2 \text{Re} \left\{ \frac{1}{\sqrt{\bar{\epsilon}_b - 1}} \right\}, \quad (2.15)$$

respectively. The phase constant for both types of modes is

$$\beta_{mn} = \sqrt{k_0^2 - \left(\frac{m\pi}{2a} \right)^2 - \left(\frac{n\pi}{2b} \right)^2}. \quad (2.16)$$

The field components in the interior region are solved for by substituting k_x^i and k_y^i into Equations (2.4) to (2.9) and, thus, all field components are present. However, many simplifications can be made by using the assumption that terms with (λ/a) and (λ/b) are negligible. The result of this is a very simple form for the field components and only two terms remain for each polarization. For x -polarized

modes the non-zero fields components are

$$E_{mn}^x(x, y) = \sin\left(\frac{m\pi}{2a}x + \phi_x\right) \sin\left(\frac{n\pi}{2b}y + \phi_y\right), \quad (2.17)$$

$$H_{mn}^y(x, y) = \sqrt{\frac{\epsilon_0}{\mu_0}} E_{mn}^x, \quad (2.18)$$

where $m, n > 0$ and ϕ_x and ϕ_y are as previously defined.

Correspondingly, for y-polarized modes the non-zero field components are

$$E_{mn}^y(x, y) = \sin\left(\frac{m\pi}{2a}x + \phi_x\right) \sin\left(\frac{n\pi}{2b}y + \phi_y\right), \quad (2.19)$$

$$H_{mn}^x(x, y) = \sqrt{\frac{\epsilon_0}{\mu_0}} E_{mn}^y. \quad (2.20)$$

Figure 2.2: Tunnel geometry adopted in this work.

In this analysis the geometry in Figure 2.2 is used, where $\epsilon_w = \epsilon_{w,r} + \frac{\sigma_a}{j2\pi f\epsilon_0}$ is the normalized permittivity. Also, a y -polarized source and receiver is assumed and, thus, only the dominant E_y component of the electric field is of interest. For the E_y component in this tunnel,

$$E_y(x, y, z) = \sum_m \sum_n C_{mn} E_{mn}^{eig}(x, y) e^{-(\alpha_{mn} + j\beta_{mn})z}, \quad (2.21)$$

where

$$E_{mn}^{eig}(x, y) = \sin\left(\frac{m\pi}{a}x + \phi_x\right) \sin\left(\frac{n\pi}{b}y + \phi_y\right), \quad (2.22)$$

$$\alpha_{mn} = \frac{2}{a} \left(\frac{m\pi}{ak_0}\right)^2 \text{Re}\left\{\frac{1}{\sqrt{\epsilon_w - 1}}\right\} + \frac{2}{b} \left(\frac{n\pi}{bk_0}\right)^2 \text{Re}\left\{\frac{\epsilon_w}{\sqrt{\epsilon_w - 1}}\right\}, \quad (2.23)$$

$$\beta_{mn} = \sqrt{k_0^2 - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2}. \quad (2.24)$$

The electric field is composed of many modes each with their own attenuation constant, propagation constant and weighting coefficient, C_{mn} , that is provided in the following section. The propagation constant in Equation (2.24) is actually identical to that for a perfectly conducting rectangular waveguide. Additionally, as for the case for all single-conductor waveguides, signals disperse as they propagate in them. The group velocity of a mode can be predicted using

$$v_{g,mn} = \frac{c^2}{v_p} = c^2 \frac{\beta}{\omega} = c \sqrt{1 - \left(\frac{m\lambda}{2a}\right)^2 - \left(\frac{n\lambda}{2b}\right)^2}. \quad (2.25)$$

The group velocity of a tunnel is frequency dependent, and as a result an UWB pulse will disperse as it travels through it.

2.1.2 Derivation of Source Expansion Coefficients

Assuming that a single transmitter excites modes in a tunnel, the weighting coefficient of a given mode depends on the type of transmitting antenna used and its placement in the tunnel. The weighting coefficient in a tunnel has previously been derived using geometric optics [34], and by projecting a field onto a reference plane and correlating it with orthogonal modes [31], [32]. Here, the weighting coefficient is found by finding the resultant E_y field in a tunnel due to a y -polarized electric current source. Green's functions are used for current sources in a waveguide [36] and the derivation is provided in Appendix B.1. The resulting E_y component is

$$E_y = \frac{2\omega\mu_0}{k_0^2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{1}{ab\beta_{mn}} \left(k_0^2 - \left(\frac{n\pi}{b}\right)^2\right) \sin\left(\frac{m\pi x}{a} + \phi_x\right) \sin\left(\frac{n\pi y}{b} + \phi_y\right) \cdot \int \int \int J_y(x', y', z') \sin\left(\frac{m\pi x'}{a} + \phi_x\right) \sin\left(\frac{n\pi y'}{b} + \phi_y\right) e^{-j\beta_{mn}(z-z')} dV', \quad (2.26)$$

where J_y is the electric current source and the primed coordinates denote the source coordinates.

By assuming a specific antenna type Equation (2.26) can be simplified further and C_{mn} can be identified by comparing to Equation (2.21). In this research an infinitesimal dipole is considered because J_y is easily integrable. Furthermore, the ray tracer model uses an infinitesimal dipole by default; thus, the models can be compared easily. In Appendix B.2 the weighting coefficient is derived assuming an

infinitesimal dipole and the result is

$$C_{mn} = \frac{2I_0\omega\mu_0}{k_0^2} \frac{2b}{n\pi} \frac{1}{ab\beta_{mn}} \left(k_0^2 - \left(\frac{n\pi}{b} \right)^2 \right) \sin \left(\frac{n\pi l}{2b} \right) \sin \left(\frac{m\pi x_0}{a} + \phi_x \right) \sin \left(\frac{n\pi y_0}{b} + \phi_y \right), \quad (2.27)$$

where I_0 is the amplitude of the dipole's current, $(x_0, y_0, 0)$ is the dipole's location and l is the dipole's length.

This weighting coefficient strongly depends on β_{mn} , which monotonically decreases as mode order increases. For most tunnel sizes of interest (e.g. 5×5 m) β_{mn} will reach zero for very high order modes (e.g. $m \approx 100$ for $n = 1$ at 3.1 GHz). At this point high order modes must be considered negligible or else they will be disproportionately weighted. The weighting coefficient also shows more dependency on the n index which means that low order n indices are strong.

2.1.3 Summary of Waveguide Model

An analytical equation waveguide model has been provided in Equations (2.21) to (2.24) and (2.27). This model is valid for a straight, rectangular tunnel at UWB frequencies, providing that tunnel dimensions are in the meter range. The permittivity and conductivity of the walls should also be large; however, there is a greater dependence on tunnel dimensions for the model. This waveguide model assumes that there are an infinite number of modes propagating because there is no cut-off frequency for hollow dielectric waveguide modes. That is, β_{mn} can become imaginary and the corresponding mode will decay very fast because the attenuation constant, α_{mn} , is effectively enhanced. A cut-off must, however, be adopted to perform simulations with a reasonable run-time. Additionally, high order modes can be cut-off because they attenuate very quickly and do not abide the first and second derivation assumptions used and, thus, may not exist in reality. In Section 4.2 mode cut-offs used in this research are presented.

The waveguide model of propagation in a tunnel has been compared to measurements in previous studies [30], [34] and will be compared to measurements in this research. In previous studies propagating signal strength was studied versus tunnel distance and two distinct distance regions were identified. In the first zone, called the near zone, many modes propagate and the signal strength varies rapidly versus distance. In the second zone, the far zone, only a couple of the lowest order modes propagate and the signal strength fluctuates smoothly versus distance. The distance at which the zone transition occurs depends on the tunnel size and frequency. For increasing frequency and tunnel size, the further the transition distance is at. In this study, UWB frequencies are used and meter-scale tunnel dimensions are considered; thus, this transition distance is expected to be around 500 m or more [30], [34]. That is, many modes are strongly present for distances less than approximately 500 m and this is the primary

region in which the proposed localization system must operate.

This analytical equation waveguide model has several limitations. The first is the tunnel shape. Though modes for circular tunnels can be derived [30], propagation in other tunnel geometries may not be predictable by using rectangular or circular waveguide modes. The second limitation is with modelling discontinuities, tunnel roughness and objects in the tunnel. The waveguide model cannot do this. Extra factors to accommodate for objects in a tunnel have been introduced [31], [37], but have not been used in this research.

As mentioned previously, the scope of this research is limited to straight, rectangular tunnels and, therefore, the limitations with the waveguide model are avoided. Despite this an alternative tunnel model is used, a ray tracer, to verify that mode theory is in fact valid.

2.2 Ray Tracer Model

Ray tracer models are versatile modelling techniques because propagation in any environment can be studied after the environment is described by planar facets and antenna parameters are set. For example in a tunnel, discontinuities and tunnel branches can be described and various transmitting antennas can be used. In this research an image-based ray tracer model developed at the University of Toronto was used [32].

Ray tracing techniques are based on the geometric optics approximation where a propagating electromagnetic wave is a ray that traverses a path from a transmitting antenna to a receiving antenna. The path may have any number of reflections off of surfaces. The geometric optics approximation of wave propagation is valid when the dimensions of the objects in the environment are much larger than the wavelength of the frequency of operation. This approximation is valid for UWB signals propagating in tunnels with cross-sections on the order of meters. Image-based ray tracers use image theory to generate image sources that correspond to ray paths that have one or more reflections. Details of the image-based ray tracing algorithm used can be found in [32]. A ray must reflect off of a surface at a point that corresponds to the intersection of the surface and a line connecting the image source or actual source, and the receiver or previous reflection point. Thus, the angle at the receiver between a normal plane (to the tunnel cross-section at the receiver) and the incoming ray depends on the path taken, as Figure 2.3 shows for two different examples.

In using the ray tracing software at the University of Toronto an input file must be specified. The input file must consist of a description of planar surfaces representing the environment, the transmitter and receiver positions, the frequency of operation and the maximum number of reflections allowed for a

Figure 2.3: Example of ray paths using an image-based ray tracer.

given ray.

In the ray tracer version used, the expression for the far-field electric field generated by the direct ray path is found as [38]

$$\vec{E}_{dir} = ZI_0 \frac{e^{-jkr}}{r} \vec{E}_\theta(\theta, \phi) = \sqrt{\frac{\eta P_T}{2\pi}} e^{j\psi} \frac{e^{-jkr}}{r} \vec{E}_\theta(\theta, \phi), \quad (2.28)$$

where Z is the equivalent impedance for the transmitting antenna, I_0 is the current supplied to the transmitting antenna, $\vec{E}_\theta(\theta, \phi)$ is the normalized radiation pattern, P_T is the power radiated from the transmitting antenna, ψ is the phase of the product of ZI_0 for the antenna and (r, θ, ϕ) are the coordinates for the receiving point considered. Paths that reflect off of at least one surface have the same form as Equation (2.28), but are adjusted with appropriate Fresnel reflection coefficients. The overall electric field calculated at the receiving antenna's position comprises of many rays each with their own electric field unique to the path they took. Thus, a ray tracing model is a multipath model of propagation.

In order to compare Equation (2.21), which describes a current to electric field transfer function considering a infinitesimal dipole, to the resultant electric field using the ray tracer model, Equation (2.28) can be modified. For an infinitesimal dipole oriented on the z -axis the far-field electric field is [38], [39]

$$\vec{E} = j\eta \frac{l}{2\lambda} I_0 \frac{e^{-jkr}}{r} \sin \theta \hat{\mathbf{a}}_\theta, \quad (2.29)$$

where l is the dipole's length, I_0 is the dipole's current and η is the free-space wave impedance.

In comparing this to Equation (2.28) the impedance is $Z = j\eta l/(2\lambda)$ and the radiation pattern is $\sin\theta\hat{\mathbf{a}}_\theta$. In the ray tracing software used $j\eta/(2\lambda)$ is considered; however, the effect of Il must be accounted for in post-processing in order to accurately model an infinitesimal dipole and to compare field amplitudes to the analytical equation model.

2.3 Relationship between the Models

In the previous sections it was shown that the analytical equation model and ray tracer model are both appropriate models for studying tunnel propagation and both can calculate the resultant electric field from an infinitesimal dipole excitation. In Figure 2.4 the y -component of the time domain electric field is compared at one transverse point at several longitudinal z distances in a 2.4×2.6 m tunnel that is described in detail in Chapter 4. The notation ‘AE’ and ‘RT’ is used in Figure 2.4 to denote if data from ray tracing simulations or analytical equation simulations, respectively, is used. It is observed that the predicted fields agree in terms of multipath arrival times and only the amplitude of multipath components arriving after the main arrival disagree. Further comparisons of fields from the two models can be found in Chapter 4.

The relation between a propagating mode and a ray path is not one-to-one. In the ray tracer model each k^{th} multipath component traverses a path length of d_k between the transmitter and receiver. The corresponding time-of-arrival for each path is $\tau_k = d_k/c$. Additionally, the angle between the normal of the tunnel’s aperture and the arriving path at the receiver, θ_k , depends on the reflections taken, as shown in Figure 2.3.

In the analytical equation model propagating modes may be thought of as sets of multipath components and their approximate time of arrival is $\tau_{mn} = d/v_{g,mn}$. Now consider that the walls of the tunnel are perfectly electrically conducting (PEC). A propagating mode at one frequency can be interpreted as the superposition of four plane waves bouncing obliquely between waveguide walls. The waves propagate at a certain angle, which can be compared to arrival angles predicted with ray tracing. Take for example the $m = 1$ mode in two dimensions where the dominant E_y component can be written as [40]

$$E_y(m = 1) = \cos\left(\frac{\pi x}{a}\right)e^{-j\beta_1 z} = \frac{1}{2}\left(e^{j\frac{\pi}{a}x} + e^{-j\frac{\pi}{a}x}\right)e^{-j\beta_1 z} = \frac{1}{2}\left(e^{-j(\beta_1 z - \frac{\pi}{a}x)} + e^{-j(\beta_1 z + \frac{\pi}{a}x)}\right), \quad (2.30)$$

which is two plane waves obliquely propagating between the top and bottom surfaces of the waveguide. The first wave travels in the $+z$ and $-x$ directions and the second waves travels in the $+z$ and $+x$ directions, and both have phase constants of β_1 and π/a , respectively. This can be written for any mode

(a) $z = 4.88$ m (b) $z = 9.75$ m (c) $z = 14.63$ m (d) $z = 19.51$ m (e) $z = 24.38$ mFigure 2.4: Comparison of simulation electric fields (e_y) for the $(-0.55, -0.65, z)$ m point across distance.

and be expanded to three dimensions which reveals that four waves are propagating. Equation (2.30) can then be compared to the case of parallel polarization oblique incidence on a PEC surface and it is found that

$$k_0 \sin \theta_i = \beta_1, \quad (2.31)$$

$$k_0 \cos \theta_i = \frac{\pi}{a}, \quad (2.32)$$

where θ_i is the angle of incidence to the wave measured from the normal to the wall. These relations show that the angle of incidence can be expressed in terms of mode properties. Additionally, this relation holds for the case of perpendicular polarization and can be expanded to the three-dimensional case. As a result the overall angle of arrival for a given mode, measured from the normal of the tunnel's aperture is

$$\theta_{mn} = \cos^{-1} \left(\frac{\beta_{mn}}{k_0} \right) = \cos^{-1} \left(\frac{\sqrt{k_0^2 - \left(\frac{m\pi}{a} \right)^2 - \left(\frac{n\pi}{b} \right)^2}}{k_0} \right) = \cos^{-1} \left(\sqrt{1 - \left(\frac{m\lambda}{2a} \right)^2 - \left(\frac{n\lambda}{2b} \right)^2} \right), \quad (2.33)$$

where Figure 2.5 represents the angle in this scenario.

Figure 2.5: Mode propagation angles at a receiver.

This relationship approximately holds in a tunnel with dielectric walls for modes with small grazing angles. That is, the Fresnel transmission coefficients at a wall are considered negligible while the Fresnel reflection coefficients are very large. A phase shift will be produced upon reflection at a wall, that will not necessarily be 180° like that for reflection at a PEC interface.

Equation (2.33) provides an angle for a given mode at a given frequency; thus, many modes may have the same θ_{mn} for different frequencies. On the other hand, a ray with a given number of reflections has a given angle for all frequencies and that angle changes depending on the distance between the

transmitter and receiver. For example, if the $(1, 1)$ mode is propagating it will have θ_{mn} angles between approximately 0.23° and 0.81° in the UWB frequency range in a 5×5 m tunnel. In the ray tracer model no paths will have this angle unless the longitudinal distance of the tunnel is very large, for example 500 m. Even if the mode's arriving angle and the ray's arriving angle match for one scenario, it will not hold when with distance variations and it will only hold at one frequency. A ray can thus be interpreted to be a composition of modes at many frequencies. Alternatively, a mode can be interpreted as many different rays. Despite this complex relationship both models predict fields that generally agree, which will be shown in Chapter 4. Additionally, since these models are intrinsically different both are used to verify spatial filter results to prove its validity.

Chapter 3

Spatial Filter Concept and Its Use in Tunnel Localization Systems

In Chapter 2 two different modelling techniques were presented: an analytical equation model, based on the fact that modes exist in waveguide-like tunnels, and a ray tracer model, based on the geometric optics approximation of electromagnetic wave propagation. These models were shown to be inherently different. For example, in the analytical equation model multipath at a single receiving point may be considered as the arrival of many superposed modes. On the other hand, in the ray tracer model multipath arrivals are viewed as different rays that take unique transmission paths to the receiver. For a TOA localization system it is crucial that the first arriving multipath is resolved and isolated in order to provide accurate localization estimates. If a multipath corrupted signal passes through a matched filter the extracted TOA will not be accurate. Constructive interference of later multipath components may result in erroneous TOA estimates because the first multipath component is overwhelmed. This is because the matched filter output will be asymmetrical and the peak or high threshold levels, at which the TOA could be extracted, would not correspond with the first multipath component's arrival. Moreover, since the localization system must move in the tunnel the accuracy of estimates will change as the electromagnetic wave distribution fluctuates.

Evidently, implementing a TOA localization system using a multipath corrupted signal would result in poor performance. However, in a tunnel ridden with multipath, isolating and extracting the first arriving multipath, or any multipath component, is difficult. In an ideal TOA localization system only one multipath would exist. With this one multipath the signal would pass through the matched filter and the output would have a clear peak at which the TOA could be, theoretically, perfectly extracted.

In a ray tracer context, one ray would have to be filtered out of a received signal. To do this there must be knowledge of, perhaps, which angle the direct ray arrives from and which angles indirect rays arrive from; that way arrivals from certain angles could be disregarded. However, rays' angles change with distance and the once reflecting paths (which have the largest effect on the first multipath component) may arrive at angles less than a few degrees apart from the direct path. Hence, implementing such an adaptable, angle-sensitive filter would be challenging.

In a waveguide theory context, one mode could be filtered out of a received signal. An UWB mode has a clear, distinct pulse shape, like one multipath component, but it disperses as it propagates. By receiving one mode the TOA could be accurately extracted. To do this other modes must be filtered out and the clearest way to do this is by spatially filtering the incidence electromagnetic field such that only one mode passes, on which the TOA calculation would be based. This solution is conceptually simple and more robust than filtering signals from specific arrival angles.

In this chapter the design of such a mode-extracting filter is presented. The filter is a spatial filter as it is a weighted array of receiving antennas. Following this the use of the spatial filter in a time-of-arrival localization system is outlined and methods for evaluating a given spatial filter's performance are detailed.

3.1 Concept and Design of a Spatial Filter

A spatial filter is a weighted array of receivers designed to selectively receive a specific mode in a tunnel and it is shown conceptually in the tunnel's cross-sectional plane in Figure 3.1 and systematically in the dashed box in Figure 3.2. As Figure 3.1 shows the incidence electric field combined with the spatial filter is equivalent to receiving one mode, when the resultant time-domain signal is viewed. The spatial filter is a two-dimensional array of antennas and the weights for each antenna are not shown in Figure 3.1. The mode to be passed by the filter is chosen based on the strength of the modes launched by a transmitter and whether the degree of dispersion in the selected mode is tolerable for accurate TOA estimates. The effect of dispersion is significant at far distances because an UWB mode's corresponding pulse width increases with time. Recall that there will be many propagating modes in the distance range considered for the spatial filter's operation, which is from $z = 0$ m until the received signals are overcome by receiver noise.

An arbitrary mode can be perfectly resolved by using a very dense grid of receiving antennas placed across the tunnel's cross-section. However, even fifty antennas may be too many for a practically implementable spatial filter. Thus, the following three steps are proposed to design a spatial filter with

Figure 3.1: Sample spatial filter in tunnel cross-section shown with the electric field at one frequency.

a limited number of receiving antennas, providing a feasible transmitter position has been identified.

1. **Choose the Mode to be Extracted** (m_f, n_f): Analyze mode coefficients excited by the infinitesimal dipole by evaluating Equation (2.27). A mode coefficient is effectively the strength of a given mode at the transmitting plane in the tunnel. The mode to be extracted by the filter, denoted by (m_f, n_f) , should have a high C_{mn} coefficient and as low a mode order as possible. The former condition is to ensure the mode is easily detectable. The latter condition is to ensure that the mode is easily detectable in noise and to ensure the mode does not disperse and attenuate significantly as it propagates, which means it correlates with a non-dispersed pulse well.
2. **Place K Receiving Antennas in the Sampling Plane**: Using a limited number of K receiving antennas, choose their locations such that they coincide with or near extrema of the chosen mode. The signal from each antenna is weighted, with W_k , to correspond to the filtered mode (m_f, n_f) ; that is,

$$W_k = \sin\left(\frac{m_f \pi}{a} x_k + \phi_x\right) \sin\left(\frac{n_f \pi}{b} y_k + \phi_y\right), \quad (3.1)$$

where (x_k, y_k) is the location of the k^{th} antenna. The weights of the antennas correlate with the orthogonal tunnel modes and, thus, only the desired mode and spatially aliased modes pass. The effect of aliasing is discussed more later.

3. **Evaluate Performance**: Spatial filter performance can be evaluated in many ways to ensure design criteria are met. For example, one can study the performance of the spatial filter in additive white Gaussian noise (AWGN) and compare to theoretical performance bounds, a single sensor's performance and an ideal spatial filter's performance.

All practically-sized spatial filters will undersample a tunnel's aperture because many modes propagate in meter-sized tunnels at UWB frequencies and a high number of antennas are necessary to satisfy

Nyquist conditions. If a mode is given by $\cos(\frac{m\pi}{a}x)$ then $f_x = \frac{m\pi}{a}(\frac{1}{2\pi}) = \frac{m}{2a}$ and $f_{Nyquist,x} = 2f_x = \frac{m}{a}$. This means that to satisfy Nyquist conditions there must be sample points every $\frac{a}{m}$ meters and every $\frac{b}{n}$ meters in the x and y directions, respectively. Correspondingly, this means there must be a sample point at every extrema in a given mode to resolve it adequately. For example, to satisfy Nyquist requirements one sample point is necessary to sample the $m = 1$ mode, two samples points are necessary to sample the $m = 2$ mode, three sample points are necessary to sample the $m = 3$ mode and so on.

In overmoded tunnels the number of sample points necessary to resolve modes accurately depends on the highest order mode that is to be resolved or is to be extracted. For example, in a 5×5 m tunnel that has an offset dipole as the source, modes up to about $m = 50$, $n = 50$ will be present at close distances. Thus, 50×50 antennas must be used to resolve any given mode perfectly in this environment. Fewer antennas, such as 30×10 , can be used if modes with $m > 30$ and $n > 10$ have much less energy and attenuate quickly. The only trade-off here is that some very high order modes may be aliased, but they will have a minute effect on TOA localization as they arrive later, and these modes will only be aliased for short distances. By accepting the fact that mode aliasing is inevitable, a more realistic spatial filter with 15 or less antennas can be used.

In the second design step it was instructed to place antennas at or near extrema of the mode to be extracted by the filter. This is done so that the targeted mode passes through the spatial filter with maximum energy. Other modes may have nulls, extrema or values in between at these antenna positions, and as a result the majority of other modes will experience attenuation as they pass through the filter. There will be some modes that pass through the filter without experiencing much attenuation and their effect on the spatial filter's performance should be studied. If these passed modes have a significant energy in the tunnel they will affect filter performance, if not, their effect will be minimal.

In Chapter 4 several sample spatial filter configurations are presented and for each configuration a list of modes that are minimally attenuated by the spatial filter is shown. These modes are found by plotting filter weight positions alongside the given mode and observing if the filter weights are also near extrema of that mode. The effect of aliasing specific modes is then analyzed in performance results for these configurations.

3.2 A Spatial Filter as Part of an UWB Localization System

The UWB localization system proposed is simple, as the focus of this research is the spatial filter itself. Figure 3.2 shows the proposed system architecture. A sole transmitting antenna emits an UWB pulse and the transmitter is to be placed near a tunnel wall, which allows tunnel activities to coexist with

the system's infrastructure. If it is feasible the transmitter can also be placed anywhere in the tunnel's cross-section and it is assumed that the transmitter does not do any signal filtering. The spatial filter, consisting of a weighted array of antennas, receives the UWB signal. The receiving antennas are assumed to be on a transverse cross-sectional plane in the tunnel, as Figure 3.1 shows, and they are assumed to be mounted on the object that is to be localized. The output of the spatial filter, which is a single mode or superposition of a few modes, passes through a matched filter that is matched to the transmitter's pulse shape. The output signal is then thresholded and time of arrival calculations are made based off of the instant at which the threshold is passed. It is assumed that the transmitter and receiver are synchronized in time and, thus, this is a one-way ranging system that calculates the longitudinal tunnel distance between the transmitting and receiving planes. Digital sampling may be present at the matched filter's output. The effects of digital sampling are not considered in this analysis in order to show the potential accuracy of the system. It is emphasized that the receiver consists of multiple antennas but only one matched filter is required and signal processing must be completed on only one signal. Specific system components are now described in more detail.

Figure 3.2: Proposed one-way TOA UWB localization system.

A matched filter is used in this system because in a channel with AWGN the matched filter maximizes the signal to root mean square noise amplitude ratio [9]. The matched filter uses a time inverted template of the transmit signal pulse, $p(t)$. That is,

$$h(t) = p(T_o - t), \quad (3.2)$$

is the response of the matched filter where T_o is an arbitrary time delay. The signal output by the

matched filter is

$$y(t) = v_{tot}(t) * h(t)dt = \int_0^t v_{tot}(\tau)h(t - \tau)d\tau. \quad (3.3)$$

In the case of the localization system proposed the input of the matched filter, $v_{tot}(t)$, does not usually match the transmit waveform $p(t)$, due to multiple modes passing, or in the case of a perfect spatial filter, dispersion. The result is that the matched filter's output is delayed and dispersed in time, or it not a symmetrical autocorrelation function. However, by using the matched filter, the TOA system with the spatial filter can tolerate more AWGN than without it. Because the output of the matched filter may not be symmetrical peak detection algorithms for determining the signal's TOA may lead to erroneous results. Thus, a threshold detector is used to determine the TOA of a given signal.

In this research the threshold detector operates by normalizing the matched filter output and the threshold is triggered when the absolute signal passes a given level, denoted by γ . The time at which the threshold is first exceeded is the estimated TOA, or TOA_{est} and it corresponds to

$$TOA_{est} = \arg \min_t ||y_n(t)| - \gamma > 0|, \quad (3.4)$$

where $y_n(t)$ is the matched filter output normalized to an equivalent clean matched filter output and $0 < \gamma < 1$. There are more accurate ways to set the threshold as well [12].

This is a one-way TOA ranging system that must have a time synchronized transmitter and receiver in order to calculate $TOA_{est} = t_2 - t_1$. Synchronization error can corrupt one-way TOA systems as both sides of the link can have clock drift and clock offset [12]. Clock error and jitter has not been considered in the scope of this research.

In this localization system a periodic pulse train is assumed to be transmitted with a low pulse repetition frequency (PRF of f_0). This is an impulse radar transmission approach. Modulation schemes would be typically adopted in such a transmission scheme, but to keep the approach general no modulation scheme is used. That is, a train of unmodulated pulses is transmitted. For signal-to-noise ratio calculations shown later a PRF of 20 MHz is assumed, but lower rates can be considered for this system. The pulse shape used is a fifth order Gaussian derivative as it fits the UWB spectral mask [41]. In the time domain the pulse is

$$p(t) = A \left(-\frac{t^5}{\sqrt{2\pi}\sigma^{11}} + \frac{10t^3}{\sqrt{2\pi}\sigma^9} - \frac{15t}{\sqrt{2\pi}\sigma^7} \right) e^{-t^2/(2\sigma^2)}, \quad (3.5)$$

and in the frequency domain it is

$$P(f) = A(j2\pi f)^5 e^{-\frac{(2\pi f\sigma)^2}{2}}, \quad (3.6)$$

where $\sigma = 51$ ps and the pulse width, T_p , is approximately $10\sigma = 0.51$ ns, and it contains more than 99.99% of the pulse's energy [41]. A is an arbitrary amplitude.

A pulse train may be written as

$$x(t) = \sum_{n=-\infty}^{\infty} p(t) * \delta(t - nT_0) = \sum_{n=-\infty}^{\infty} p(t - nT_0), \quad (3.7)$$

where the pulse period is $T_0 = 1/f_0$. The power spectral density (PSD) of this train of unmodulated pulses needs to be found because later it will be used to ensure that the transmitted signal is compliant with the UWB spectral mask. To derive the PSD, first consider the complex exponential Fourier series for a train of pulses

$$x(t) = \sum_{n=-\infty}^{\infty} x_n e^{jn2\pi f_0 t}, \quad (3.8)$$

which holds for any periodic signal. In Equation (3.8) $x(t)$ is the train of pulses and the coefficients, x_n , are

$$x_n = \frac{1}{T_0} \int_{T_0} x(t) e^{-jn2\pi f_0 t} dt. \quad (3.9)$$

From this representation the power spectrum of the pulse train can be easily obtained by plotting the coefficients squared, $|x_n|^2$, at nf_0 . Correspondingly, the power spectral density, $S_x(f)$, shows the same information on a density scale (in [W/Hz], considering a 1 Ω resistor). That is, the PSD consists of delta functions weighted by $|x_n|^2$. An important feature of the pulse train's frequency domain representation is that since it is a periodic signal the components in the frequency spectrum are discrete impulses spaced at multiples of f_0 .

Different types of pulse trains can be described by their Fourier series and their Fourier spectra can be plotted using coefficients. However, a final equation to describe $X(f)$ or $S_x(f)$ must be formulated after finding the x_n coefficients. The Fourier transform, $\mathfrak{F}()$, of Equation (3.8) is

$$\begin{aligned} X(f) &= \int_{-\infty}^{\infty} x(t) e^{-j2\pi f t} dt \\ &= \int_{-\infty}^{\infty} \left(\sum_{n=-\infty}^{\infty} x_n e^{jn2\pi f_0 t} \right) e^{-j2\pi f t} dt \\ &= \sum_{n=-\infty}^{\infty} x_n \left(\int_{-\infty}^{\infty} e^{jn2\pi f_0 t} e^{-j2\pi f t} dt \right) \\ &= \sum_{n=-\infty}^{\infty} x_n \left(\int_{-\infty}^{\infty} e^{j2\pi t(nf_0 - f)} dt \right) \\ &= \sum_{n=-\infty}^{\infty} x_n \delta(f - nf_0). \end{aligned} \quad (3.10)$$

To find x_n either the previous approach can be taken, or consider that for a train of pulses (Equation (3.7)) the Fourier transform is

$$\begin{aligned} X(f) &= P(f) \mathfrak{F} \left(\sum_{n=-\infty}^{\infty} \delta(t - nT_0) \right) \\ &= P(f) \left(\frac{1}{T_0} \sum_{n=-\infty}^{\infty} \delta(f - n/T_0) \right) \\ &= \frac{1}{T_0} \sum_{n=-\infty}^{\infty} P(n/T_0) \delta(f - n/T_0), \end{aligned} \quad (3.11)$$

where the Fourier transform of a train of pulses is taken from [42]. Comparing Equation (3.11) to Equation (3.10) it is seen that $x_n = \frac{1}{T_0} P(n/T_0)$.

Now, the weights of the power spectrum, $|x_n|^2$, are in a clear form for a pulse train and can be plotted at f_0 multiples. From this the PSD of the pulse train can be written as

$$S_x(f) = \sum_{n=-\infty}^{\infty} |x_n|^2 \delta(f - n/T_0) = \frac{1}{T_0^2} \sum_{n=-\infty}^{\infty} |P(n/T_0)|^2 \delta(f - n/T_0). \quad (3.12)$$

This PSD expression can also be verified using another approach where the autocorrelation function of an impulse train is found, $x_{imp}(t)$. Then $S_x(f) = S_{x_{imp}}(f) |P(f)|^2$ can be calculated where it is considered that the pulse train of fifth order Gaussian derivatives is the output of a system that has a transfer function of $P(f)$ and an input of a train of delta functions.

3.3 Evaluation Methods for a Spatial Filter

The design of a spatial filter and its use in an UWB tunnel localization system has been described. It is now necessary to establish evaluation methods that can quantify the performance of a spatial filter in order to see how accurate its ranging abilities are, how it stands up to varying AWGN levels, how far of a distance it can be used for while considering realistic power and noise levels, and how it compares to other spatial filters and benchmarks. Furthermore, equivalent spatial filters are to be evaluated and compared using the analytical equation model, ray tracer model and measurements.

The root mean square error (RMSE) is used to show the ranging accuracy of a spatial filter for different levels of AWGN [8]. It is defined as the square-root of the mean square error (MSE) of an estimate and is given by

$$RMSE = \sqrt{MSE} = \sqrt{E[(TOA_{theoretical} - TOA_{est})^2]}, \quad (3.13)$$

where $E[\cdot]$ is the expected value of the error. The RMSE is effectively the standard deviation of the estimated TOA from the theoretical TOA. The RMSE has the same units as the TOA. In practice the MSE can be calculated across many equivalent measurements, also known as Monte Carlo trials. The result is

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (TOA_{theoretical} - TOA_{est,i})^2}. \quad (3.14)$$

where N is the number of Monte Carlo trials and $TOA_{est,i}$ is the estimated TOA for the i^{th} trial. The estimated TOA is output from the threshold detector after the spatial filter output passes through the matched filter. The theoretical TOA is based off of the time it takes a pulse to propagate directly from the transmitting to receiving planes in the tunnel.

There are adjustments that must be made to $TOA_{true} = d/c$, where d is the distance between the transmitting and receiving planes. One alternative is adding two time-adjusting factors to TOA_{true} . The first factor is half of the pulse width, $T_p/2$, which is added because the matched filter output is delayed from the true theoretical TOA by this amount. The second factor considers that threshold detection is used and the TOA is not at the peak of the matched filter output; thus, $TOA_{theoretical} = TOA_{true} + T_p/2 - \Delta t_{threshold}$ is more accurate. With these adjustments and if non-dispersed pulses are received, theoretical performance bounds can be met.

Another alternative time adjustment that can be made when calculating the theoretical TOA accommodates for the time a specific mode arrives, as each mode travels at its own group velocity and not at the speed of light. In order to consider this time delay the TOA of a perfectly resolved mode, at a specific threshold and distance, is used as the $TOA_{theoretical}$ for a spatial filter that is designed to extract the same mode. As such, $TOA_{mn^{th}mode} \approx d/v_{g,mn}$, but $TOA_{theoretical}$ must be found through simulations because $v_{g,mn}$ is frequency dependent and the time at which the pulse crosses a threshold does not correspond to one frequency, or $v_{g,mn}$, across distance. This consideration is feasible because it is known beforehand which mode is to be extracted and this delay can be used in a realistic system. One thing to note is that the difference between TOA_{true} and $TOA_{theoretical}$ is not a constant for all distances. This fact is considered when processing results. This adjustment is especially important for spatial filters that extract modes higher than the first few propagating modes.

The performance limits of TOA estimation in an AWGN channel is bounded by the Cramér-Rao bound (CRB) for large SNRs [12]. The MSE of a TOA estimate is bounded by the following CRB when multipath and dispersion effects are not considered and a matched filter is used:

$$CRB = \frac{N_0/2}{(2\pi)^2 E_p \chi^2} = \frac{1}{8\pi^2 \chi^2 SNR}, \quad (3.15)$$

where $SNR = E_p/N_0$, N_0 is the one-sided PSD of the AWGN, and χ^2 is the second moment for the spectrum $P(f)$ of $p(t)$, the transmitted pulse

$$\chi^2 = \frac{\int_{-\infty}^{\infty} f^2 |P(f)|^2 df}{\int_{-\infty}^{\infty} |P(f)|^2 df}. \quad (3.16)$$

It can be shown that for an n^{th} order derivative of the Gaussian pulse that [12]

$$\chi^n = \sqrt{\frac{2n+1}{8\pi^2\sigma^2}}, \quad (3.17)$$

where σ is as previously defined. The CRB is approached for non-dispersed received pulses and if the delay between the correlation peak and threshold mark is considered in TOA calculations, as mentioned before. In the case of spatial filter outputs, the received pulse will be dispersed and may contain multipath arrivals; thus, the CRB will not be followed. However, the comparison of the spatial filter's performance to a theoretical measure is still useful to see if values are reasonable. Additionally, the CRB will be approximately followed if the ideal filter output and the second $TOA_{theoretical}$ is used, which will be shown in Chapter 4.

Other theoretical bounds can also be used to evaluate performance. It is known that the CRB is not followed for low and mid-range SNRs, and other theoretical bounds can alternatively be used to model these regions more accurately. An example is the Ziv-Zakai bound that accounts for these different regions and can account for other parameters [12].

It has been established that the RMSE can be used to gauge localization performance and that by varying SNR a spatial filter's robustness can be characterized and compared to the theoretical CRB. The SNR definition used to compare to the CRB is a receive SNR definition, which is introduced in Section 3.3.1. This receive SNR is, however, not appropriate for comparing different types of spatial filters since in practical systems the receiver noise is fixed and the transmit power is varied. Therefore, a transmit SNR definition for evaluating RMSE is introduced in Section 3.3.2. By using this transmit SNR definition with realistic noise and power levels the distance range of a given spatial filter can be predicted, and this is described in Section 3.3.3. These three types of methods will be used to present spatial filter results in Chapter 4.

3.3.1 A Receive SNR Definition for RMSE Evaluation in Varying Noise Levels

The receive signal to noise ratio is defined as the ratio of the energy in the received signal, corresponding to one transmitted pulse, at the output of the spatial filter to the noise variance of AWGN; that is

$$SNR_{Rx} = \frac{E_p}{N_0}, \quad (3.18)$$

where $E_p = \int_0^t |v_{tot}(t)|^2 dt$ and Figure 3.3 shows the scenario considered. The amount of noise added to $v_{tot}(t)$ depends on a given SNR. That is, $N_0 = E_p/SNR_{Rx}$ is calculated and randomly distributed AWGN with standard deviation of $\sigma_n = \sqrt{N_0 B}$ is added to the signal where $B = 7.5$ GHz, the bandwidth of the UWB signal.

Figure 3.3: Receive SNR.

By using receive SNR the performance of a given spatial filter can be compared to the CRB. Furthermore, in situations where signals have different energies there is a normalization using this SNR definition so that these signals can be compared on a similar scale.¹

However, when using receive SNR to compare different spatial filters it must be considered that E_p may not be the same for these spatial filters and, thus, different levels of N_0 will be used for equivalent SNR_{Rx} values.

¹This is the case when measurement data and simulation data is compared in Chapter 4; the correct amplitude of measurement data was not deembedded.

3.3.2 A Transmit SNR Definition for RMSE Evaluation with Varying Transmit Power Levels

The transmit signal to noise ratio is defined as the ratio of the power of the radiated signal to the noise power at each antenna; that is

$$SNR_{Tx} = \frac{P_{Tx}}{P_{noise}}, \quad (3.19)$$

where Figure 3.4 shows the scenario considered. The transmit power, P_{Tx} , can be varied to achieve different SNR_{Tx} values and P_{noise} is the receiver noise. In this definition the amplitude of the received signal at a given antenna and the noise power are realistically modelled. The value of SNR_{Tx} itself is not significant; it is only a quantity used to illustrate the relative performance between various spatial filters. The voltage detected by an antenna is the open circuit voltage where $V_{oc}(f) = \vec{E}(f) \cdot \vec{\ell}_{eff}(f)$

Figure 3.4: Transmit SNR.

and $\vec{\ell}_{eff}(f)$ is the vector effective length of the antenna. The voltage developed across a load resistor, assuming conjugate matching, is $V_{oc}(f)/2 = V_k(f)$. The voltage from each antenna, $v_k(t)$, is given by

$$v_k(t) = \Im\{E_y(k, f)l(f)\} \frac{1}{2}, \quad (3.20)$$

where $E_y(k, f)$ is the y -oriented electric field at the k^{th} antenna, and it is assumed that a y -oriented infinitesimal dipole is used and the vector effective length of it is equal to its physical length, l .

The noise at the receiver is assumed to be thermal noise, which is a white process and has a one-sided power spectrum equal to N_0 where $N_0 = k_B T$ [42]. Boltzmann's constant is k_B and is 1.38×10^{-23} J/K. The system temperature is T , in Kelvin, where $T = (F - 1)T_0 + T_a$ and $F = 10^{NF/10}$ where NF is the noise figure of the receiver in decibels, T_0 is the room temperature and T_a is the antenna temperature. In this research $NF = 2$ dB, $T_0 = 290$ K and $T_a = 290$ K are used, and as a result $T \approx 460$ K. The overall noise power is given by $P_{noise} = N_0 B = k_B T B$ where $B = 7.5$ GHz. The noise added to $v_k(t)$

has standard deviation of $\sigma_n = \sqrt{N_0 B}$.

3.3.3 Evaluating Spatial Filter Distance Range with a Realistic Transmit SNR

When using the transmit SNR definition the radiated power can be set such that the UWB emissions mask is obeyed. Therefore, by using a conforming radiated power and the noise power described the performance of a spatial filter can be characterized across the tunnel's distance. Recall that the average-limited UWB spectral mask was shown in Figure 1.1. The average-limited constraints are applicable in this situation as it is assumed that the PRF is 20 MHz, which is greater than 1 MHz where approximately peak emission levels need to be considered. The UWB spectral mask must be satisfied in the far-field of an antenna. It is thus required that the output power satisfies

$$\begin{aligned} 10 \log_{10} \left(\int_{f_c - 0.5 \times 10^6}^{f_c + 0.5 \times 10^6} G S_x(f) df \right) + 30 &\leq M(f_c) \quad [\text{dBm}] \\ 10 \log_{10} \left(\int_{f_c - 0.5 \times 10^6}^{f_c + 0.5 \times 10^6} \frac{1}{T_0^2} G |P(nf_0)|^2 \delta(f - nf_0) df \right) + 30 &\leq M(f_c) \quad [\text{dBm}] \\ 10 \log_{10} \left(\frac{1}{T_0^2} G |P(nf_0)|^2 \right) + 30 &\leq M(f_c) \quad [\text{dBm}], \end{aligned} \quad (3.21)$$

where G is the gain of the antenna, $S_x(f)$ is the PSD of the unmodulated pulse train shown in Equation (3.12) where $|P(nf_0)|^2$ corresponds to the pulse shape emitted by the infinitesimal dipole, and $M(f_c)$ is the UWB spectral mask over a 1 MHz bandwidth. For example, $M(f_c)$ is -41.3 dBm from 3.1 to 10.6 GHz [43], [44], [45]. Note that other impulse radar transmission schemes can also be used in Equation (3.12) by modifying $S_x(f)$ appropriately.

The gain and pulse shape considered in Equation (3.21) corresponds to an infinitesimal dipole radiating in free-space. It is assumed that the gain is equal to the antenna's directivity and is constant with frequency at $G = D = 3/2$. For a dipole radiating in free-space with perfect efficiency the pulse shape can be related to the far-field radiated power [39], and is

$$|P(f)|^2 = \frac{\alpha}{2} |I_0(f)|^2 R_{rad}(f), \quad (3.22)$$

where α is an arbitrary factor introduced to make the EIRP comply with the UWB emissions mask, I_0 is the current input to the dipole and the radiation resistance is

$$R_{rad}(f) = 80\pi^2 \left(\frac{l}{\lambda} \right)^2. \quad (3.23)$$

The wavelength of operation, λ , and dipole length is considered in the calculation of the radiation resistance.

The scale α can be found by first using an unscaled $|Pnf_0|^2$ or EIRP and the UWB emissions mask. Providing that an unscaled EIRP is given by $EIRP'(f)$, the new mask-satisfying EIRP can be found using

$$EIRP(f) = \frac{\max(\text{UWB mask in [W/Hz]})}{\max(EIRP'(f) \text{ in [W/Hz]})} EIRP'(f). \quad (3.24)$$

As a result²,

$$\alpha = \frac{\max(\text{UWB mask in [W/Hz]})}{\max(EIRP'(f) \text{ in [W/Hz]})}. \quad (3.25)$$

By using this α value and the realistic noise power presented the range of a spatial filter can be predicted. Additionally, transmit power in the SNR_{Tx} definition can be found by summing the power contributions over 1 MHz bandwidths over the UWB frequency range. Power contributions in 1 MHz bandwidths can be found by integrating the PSD, $S_x(f)$, over 1 MHz in a manner similar to that shown in Equation (3.21). Figure 3.5 shows an appropriately scaled EIRP for an unmodulated train of fifth order Gaussian derivative pulses that has a PRF of 20 MHz.

Figure 3.5: UWB mask-complying EIRP.

²Note that this assumes that $EIRP(f)$ is smooth and has a fifth order Gaussian derivative shape. If it is not smooth the maximum EIRP may not be the limiting value; EIRP values less than 3.1 GHz or greater than 10.6 GHz may be the limiting values.

Chapter 4

Simulation and Measurement Results

Spatial filter design steps and evaluation methods presented in Chapter 3 will now be used to prove that the spatial filter can mitigate multipath and lead to more accurate TOA estimates, to quantify the performance improvements afforded by spatial filters and to explore different configurations and trade-offs between them. First, in this chapter example spatial filters are designed, simulated and tested for a tunnel-mimicking environment: a hallway. In Section 4.1 a description of the hallway is presented. In Section 4.2 potential spatial filters using a limited number of receiving antennas are designed using the design steps outlined in Section 3.1 and electric fields from the analytical equation model, which are used to visualize mode properties in the hallway. In Section 4.3 details of the measurement set-up used to take frequency domain measurements in the hallway are presented.

Secondly, in this chapter's latter half, measurement data is compared to simulated data both from the ray tracer and analytical equation model, and then spatial filter results are thoroughly compared. In Section 4.4 electric fields from each of the three techniques (measurements and two models) are compared. In Section 4.5 spatial filter results are presented for ideal spatial filters, and then for the designed and measured spatial filters.

4.1 Environment Description

The spatial filter concept relies on the assumption that the propagation environment is a tunnel that obeys the constraints provided in Section 2.1.1. In this research, a tunnel was not available and is

not necessary to prove the concept of the spatial filter. Therefore, prior to launching a measurement campaign inside a tunnel a substitute environment was used: a hallway that has cinder block walls, minimal doorways and no obstructions in it. The hallway used is the west hallway on the third floor of the Galbraith building at the University of Toronto. Figure 4.1 shows a blueprint of the hallway, which also includes measurement locations and the measurement origin, and properties of the hallway are listed in Table 4.1. Hallway materials have been estimated in Table 4.1.

Hallways in the Galbraith building have been used in a previous narrowband measurement campaign [46], and the estimated effective relative permittivity and conductivity of the wall materials was 9.0 and 0.05 S/m, respectively. These estimates produced waveguide-based simulation results that agreed with measurements taken at 900 MHz [46], which is slightly lower than the UWB frequency range considered here.

Several other sources in literature have stated electrical properties of concrete at UWB frequencies. For example, in one study the relative permittivity and conductivity of concrete without rebar was stated as 7.63 to 9.54 and 0.0352 to 0.6028 S/m in the frequency range of 500 MHz to 6 GHz [47]. In another study concrete slabs were measured having an average permittivity of 8.16 and conductivity of 0.749 S/m between 3 and 5 GHz [48]. Concrete slabs were also measured within the range of 500 MHz to 2.5 GHz in another study and the relative permittivity was reported as 9.5 to 9.625 and the conductivity as 0.007 to 0.146 S/m [49]. In this analysis, all stated values are considered and the electrical properties are adjusted to account for a larger frequency range and the fact that cinder blocks are present and not solid concrete. Thus, electrical properties are set to the conservative value of 7.5 and 0.05 S/m, for the relative permittivity and conductivity, respectively. In the situation that the relative permittivity and conductivity values are higher the hallway will in fact guide waves better.

It was also verified that electric fields calculated using the analytical waveguide model and ray tracer model do not change significantly for minor permittivity or conductivity variations; that is, for permittivities ranging between 7 and 10 and conductivities up to 0.5 S/m. The electrical properties of the floor and ceiling were assumed to be the same as the walls. This assumption is acceptable because the floor is predicted to be concrete covered with vinyl composition tile and the dropped ceiling most likely has concrete behind it. Additionally, the effect of the floor and ceiling is not as significant on overall fields in comparison to the walls, which have a pronounced effect on the fields due to the vertically-polarized antenna used. Note that the height of the ceiling without the dropped panels is unknown; however, similar to the case of electrical properties sensitivity, a slightly erred ceiling height will not invalidate simulation results.

Figure 4.1: Floor plan of measurement hallway in the Galbraith Building, including transmitter and receiver locations.

Property	Description or Measurement
Height	2.6 m
Width	2.4 m
Length	38.3 m
Wall materials	cinder blocks reinforced with rebar and covered with ceramic tile
Floor material	vinyl composition tile covering concrete
Ceiling material	dropped ceiling backed by concrete
Estimated relative permittivity	7.5
Estimated conductivity	0.05 S/m

Table 4.1: Hallway properties.

4.2 Simulation Settings and Potential Spatial Filter Configurations

Using the hallway properties in Table 4.1 electric fields were generated across several hallway cross-sections at numerous distances. The analytical equation model was used to generate these fields so that various spatial filter configurations could be investigated, which is conducted in the following. Several spatial filter configurations are proposed so that their performance can be compared to one another. The potential spatial filters, which must be practically implementable, are then verified with measurement and ray tracer simulation data, and are characterized further.

In the analytical equation model hallway parameters were set to those listed in Table 4.1 and a grid of 49×53 points (0.05 m spacing) was generated at 4.8768, 9.7536, 14.6304, 19.5072 and 24.3840 m or 16, 32, 48, 64 and 80 feet.¹ Recall in Chapter 2 it was mentioned that the number of modes used in simulation had to be limited. The number of modes used to calculate the electric field, E_y , at each of the points was determined by the following steps.

1. The maximum mode indices, m and n , were specified and then the equation of an ellipse revised the maximum n calculated for a specified m up to m_{max} ; that is,

$$1 = \left(\frac{m}{m_{max}}\right)^2 + \left(\frac{n}{n_{max}}\right)^2 \Rightarrow n_{max,mod} = \sqrt{1 - \left(\frac{m}{m_{max}}\right)^2} n_{max}. \quad (4.1)$$

2. The condition $(\frac{m\lambda}{2a}) \ll 1$ and $(\frac{n\lambda}{2b}) \ll 1$ was enforced. The maximum limit for this ratio was 0.5, which is high but at lower limits the low frequencies' contribution to significant modes is eliminated.
3. If a mode at a certain distance attenuated to 0.5% of its original strength at the transmitting plane it was not considered.
4. Considering C_{mn} (see Equation (2.27)), recall that when β_{mn} approaches zero C_{mn} becomes impractically large. To eliminate this problem, when $|\beta_{mn}|$ was less than 15 rad/m the corresponding C_{mn} coefficient was nulled. Note that the C_{mn} coefficients affected by this were of a very high order.

These limitations were made in the order listed to have accurate fields as well as optimal run-times. In most situations the first and second condition limited the modes used and the third condition did not

¹The distances were chosen using imperial units because the hallway tiles are one foot long and, while considering measurement procedures, it would be easier to align the receiver with lines on the floor.

need to be enforced. The fourth condition was always used and it did not affect dominant low-order modes.

The transmitter location used in all simulations and measurements is $(1.15, -0.774, 0)$ m. This location was chosen as it is an appropriately scaled transmitter location from $(2.4, 1.5, 0)$ m in a 5×5 m tunnel, which was identified as a feasible transmitter location in a tunnel because the transmitter is on the side of the tunnel, which may have vehicular traffic or other activities in it. Moreover, this transmitter location excites many modes with a significant strength and by using it it is shown that spatial filter combinations can work in very non-ideal situations.²

To identify potential spatial filter designs the spatial filter design steps outlined in Section 3.1 are used. However, before proceeding to the steps the following design considerations shall be made: (1) the number of receivers should be as low as possible to keep the filter practically implementable; and (2) the receivers must be located across a practically-sized sampling aperture.

The first design step is to choose the mode that is to be extracted by the spatial filter. Corresponding to this, mode coefficients, or intensities, in the transmitting plane are shown in Figure 4.2. Modes centered around $m = 25$ and $n = 2, 3, 7, 8, 12, 13, \dots$ have the maximum intensity, which is directly attributed to the chosen transmitter location. That is, since the transmitter location is very close to the $x = -1.2$ m vertical wall modes centered around $m = 25$ have the maximum intensity. The choice of which mode to extract with a spatial filter is not trivial in this situation because the pulse shape of the mode at a given distance, and to a lesser degree, the mode attenuation, must be considered. Attenuation is a usually a secondary consideration because for large tunnels (e.g. 5×5 m) the attenuation constants of practically-observable modes do not vary from each other. For this smaller tunnel attenuation has a larger effect. Considering the n index, choosing $n = 2$ is optimal because it is the lowest order n index with a high intensity. Considering the m index, the choice is more involved because of the aforementioned reasons.

The normalized pulse shape, which is also the ideal spatial filter output, of several modes at a propagation distance of $z = 4.88$ m and $z = 24.38$ m are shown in Figures 4.3(a) and 4.3(b), respectively. As mode order increases dispersion affects a mode's pulse shape more, as Figure 4.3 shows, especially at further distances like $z = 24.38$ m. In Figure 4.3 the $(1, 1)$ pulse is the least dispersed, but this mode has the least energy and smallest unnormalized amplitude of the three. Therefore, higher order modes must be used despite their dispersed pulse shapes. However, if a pulse is too dispersed the spatial filter's performance may also be compromised; a balance between having a high mode energy and minimal

²An ideal transmitter position would be in the center of the tunnel, where it would excite the $(1, 1)$ mode with significant energy.

Figure 4.2: Mode coefficients in transmitting plane.

Figure 4.3: Sampled normalized outputs for the idealized spatial filter.

dispersion must be made in choosing a mode to extract. Moreover, in this small tunnel the $(25, 2)$ mode decays markedly faster than the $(17, 2)$ mode. Also, note that having a high mode energy does not mean that mode has the highest amplitude when analyzing pulse shapes, due to attenuation. The $(17, 2)$ pulse in Figure 4.3(a) actually has the highest unnormalized amplitude while the $(25, 2)$ pulse has the highest energy.

It is now evident that the choice of which mode to extract is not trivial. To complicate the situation further, when using any practical number of antennas in the spatial filter, spatial aliasing will occur and multiple modes will pass through the filter. This issue was highlighted in Chapter 3.

At this point in the design, several receiving sensor positions must be set. Recall that the receiving antennas are to be located at positions that coincide with or near a given mode's extrema. Additionally,

it is desired to compare spatial filter configurations that extract low-, mid and high-order modes (e.g. $(9, 2)$, $(17, 2)$ and $(25, 2)$) in order to see which mode is better to extract in the hallway.

To extract $n = 2$ type modes there are 2 extrema where receivers can be located: at $y = -0.65$ m and $y = 0.65$ m. Also, recall that the transmitter is located at $y = -0.774$ m. Due to the transmitter location, signals at $y = -0.65$ m are generally stronger than those at $y = 0.65$ m, and it can be shown that a spatial filter with antennas located only along $y = -0.65$ m is adequate as the filter output does not change significantly when antennas at $y = 0.65$ m are added. Thus, it is concluded to locate receiving antennas only on the $y = -0.65$ m axis to reduce the total number of antennas used and measurement complexity.

Regarding the choice of m for candidate filters, a low-order mode is chosen first: the $(9, 2)$ mode. This mode is chosen because it is low-order but still has appreciable strength in the excitation plane. The $(9, 2)$ mode has 9 extrema across the cross-section of the tunnel; however, only the 5 center-most extrema will be chosen as receiver locations so that the receivers are located across a small sampling aperture. The top plot in Figure 4.4 shows the $(9, 2)$ mode and dashed lines show receiver positions for the 5 center extrema. The receiver positions are $x = 0, \pm 0.3, \pm 0.55$ m. The 5 receiver positions also alias the $(25, 2)$ mode quite strongly, which is evident in the second plot in Figure 4.4 where the dashed receiver position lines are extended into the second plot and they are close to the $(25, 2)$ extrema, and the extrema have the same orientation as the $(9, 2)$ mode. This spatial filter will thus extract the $(9, 2)$ and $(25, 2)$ mode, and the effect of doing this will be analyzed later.

A mid-order mode is now chosen: one that shares extrema near $x = 0, \pm 0.3, \pm 0.55$ m so that the amount of measurement points needed are minimized. The $(17, 2)$ mode approximately shares these extrema as the bottom plot in Figure 4.4 shows. Additionally on this plot, there are vertical lines indicating $x = \pm 0.15, \pm 0.45, \pm 0.7$ and ± 1 m, which approximately correspond to more $(17, 2)$ extrema. By placing sensors at these locations alternative configurations can also be tested. For example, an $(8, 2)$ filter that also extracts the $(24, 2)$ mode has extrema at $x = \pm 0.15$ and ± 0.45 m. The points $x = \pm 0.7$ and ± 1 m are also chosen sensor locations so that some non-central locations are used and all of the modes are non-zero at these locations.

In summary, thirteen points are chosen to conduct measurements at: $x = 0, \pm 0.15, \pm 0.3, \pm 0.45, \pm 0.55, \pm 0.7$ and ± 1 m and $y = -0.65$ m. By using these thirteen points the following filter configurations will be tested:

- A $(9, 2)/(25, 2)$ spatial filter with sensors at $x = 0, \pm 0.3, \pm 0.55$ m or $x = 0, \pm 0.3, \pm 0.55, \pm 0.7, \pm 1$ m.

Figure 4.4: Mode profile on x -axes for several modes.

- A $(17, 2)$ spatial filter with sensors at $x = 0, \pm 0.15, \pm 0.3$ or $x = 0, \pm 0.15, \pm 0.3, \pm 0.45, \pm 0.55, \pm 0.7, \pm 1$ m.
- A $(8, 2)/(24, 2)$ spatial filter with sensors at $x = \pm 0.15, \pm 0.45$ m or $x = \pm 0.15, \pm 0.45, \pm 0.7, \pm 1$ m.

Figure 4.5 shows the locations of the sensors for each filter combination listed (excluding decimated filters).

Figure 4.5: Location of thirteen sensors and combinations of them used in various spatial filters.

Table 4.2 contains significant modes passed, or aliased, by each of the spatial filter configurations. Modes are listed in the approximate order of increasing attenuation by the spatial filter; that is, the first mode(s) listed is not or is hardly attenuated by the spatial filter (because the filter is designed to extract this mode), while following modes also pass through the spatial filter, but experience some attenuation as they do. These aliased modes have been found in a manner similar to what was illustrated in Figure 4.4, where receiver positions were marked and aligned with other modes to see if an extrema or a large value existed in that mode at those locations (or at most of the locations). Take for example the $(9,1)$ and $(9,3)$ modes that are aliased with the $(9,2)$ spatial filter having 5 or 9 sensors. The $(9,1)$ and $(9,3)$

modes are minimally attenuated by the spatial filter because there are sensors at the extrema for the x mode index ($m = 9$ in all filters). The y -position of these sensors is at the extrema for $n = 2$ at $y = -0.65$ m, where the $n = 1$ and $n = 3$ modes also have a high value, as the vertical line in Figure 4.6 shows.

Figure 4.6: Alignment of $n = 2$ mode extrema with $n = 1$ and $n = 3$ modes.

The fields in the hallway must also be considered when evaluating the effect of aliasing. Taking the (9,2) spatial filter as an example again, the (9,1) mode will be aliased but its effect on the filter output will be minimal because the (9,1) mode has a weak presence in the hallway, or a low C_{mn} value. Meanwhile, the effect of aliasing the (9,3) mode will be larger because the (9,3) mode has a greater presence.

Configuration	x sensors	Sample modes passed by the filter
(9,2) 9	0, ± 0.3 , ± 0.55 , ± 0.7 , ± 1 m	(9,2), (25,2), (9,3), (25,3), (9,1), (25,1), (23,2), (19,2), (11,2), (17,2)
(9,2) 5	0, ± 0.3 , ± 0.55 m	(9,2), (25,2), (27,2), (9,3), (25,3), (9,1), (25,1), (7,2), (23,2), (11,2), (1,2), (17,2)
(8,2) 8	± 0.15 , ± 0.45 , ± 0.7 , ± 1 m	(8,2), (26,2), (24,2), (8,3), (26,3), (24,3), (8,1), (26,1), (24,1), (10,2), (18,2)
(8,2) 4	± 0.15 , ± 0.45 m	(8,2), (24,2), (26,2), (8,3), (24,3), (26,3), (8,1), (24,1), (26,1), (10,2)
(17,2) 13	0, ± 0.15 , ± 0.3 , ± 0.45 , ± 0.55 , ± 0.7 , ± 1 m	(17,2), (17,1), (17,3), (45,2), (27,2), (15,2), (25,2), (1,2)
(17,2) 5	0, ± 0.15 , ± 0.3 m	(17,2), (17,1), (17,3), (15,2), (45,2), (13,2), (19,2), (1,2)

Table 4.2: Sample modes passed with minimal or no attenuation by various spatial filters.

In the following, by using these sensor locations these six spatial filter configurations are measured and characterized. Additionally, insights into what produces a good or poor performing spatial filter are

revealed.

4.3 Measurement Campaign

Frequency domain measurements were made at the thirteen points previously listed at five distances in the described Galbraith building hallway. Figure 4.7 shows the measurement set-up. An N5244A Precision Network Analyzer (PNA) was used to measure the channel's transfer function. The transmitting and receiving antennas used were UWB Balanced Antipodal Vivaldi Antennas (BAVAS) [50]. Additionally, an ultra-wideband low noise amplifier (LNA) was used to amplify received signals. The transmit power was set to 12 dBm and the frequency range used extended from 2.2 to 13.4 GHz to match simulation parameters. An S_{21} response calibration was performed on the PNA without the LNA. The response of the LNA was removed in post-processing from the measurement data.

Figure 4.7: Measurement set-up.

In measurements the transmitter remained at the same location while the receiver was manually moved to the thirteen points that were all at the same height, at a given distance. Figure 4.1 shows the approximate transmitter and receiver locations. The transmitter and receiver positions never came within 6.5 m of the end walls in the hallway which reduced the effect of end-wall reflections on measurements. Additionally, doors were closed in the hallway and all objects, such as garbage bins, were removed to make the hallway as waveguide-mimicking as possible. For the longest two distances the PNA was in the measurement path and it was located alongside a doorway to minimize its effect. In Section 4.4 electric fields are compared to simulated electric fields and it is seen that the effect of the PNA is not significant.

Small positioning errors may have been introduced during measurements, but they are estimated

to be only 2 to 3 cm at maximum for the following reasons: on the transmitting side the antenna was accurately placed and then it remained in the same position throughout all measurements; and on the receiving side the antenna height stayed the same (as it sat on a box and foam), and tile lines were used to align the antenna longitudinally and self-measured markers were used for alignment across the width of the hallway.

In order to extract the received electric field from gathered S_{21} data for each of the measurement points the effect of the LNA was removed by dividing out the LNA's S_{21} response. Then the data was multiplied with the pulse shape used in simulations and its inverse fast Fourier transform was taken.

4.4 Comparison of Measured and Simulated Electric Fields

Prior to introducing spatial filter results using data from the analytical equation model, ray tracer model and measurements, a comparison of electric fields from each of the techniques is provided. In the following the vertical component, e_y , of the time-domain electric field from each of the techniques is quantitatively compared at a sample of the thirteen measurement points across the five distances. Additionally, discrepancies between the data are discussed.

Before comparing fields, details regarding the ray tracing simulations are discussed. To run the ray tracing software an input file describing the hallway was created. In it six surfaces were defined (for the ceiling, floor and four walls) with dimensions and electrical properties reflecting those in Table 4.1. The transmitter location was defined to be that used in measurements, at $(1.15, -0.774, 0)$ m, which is shown in Figure 4.1. A frequency sweep was performed in each ray tracer simulation for a given receiver point. Simulations were performed for each of the thirteen receiver points at five distances. More accurate ray tracer simulations, that included doorways and the PNA's location, were also performed. Further details for running ray tracer simulations can be found in [32].

Figures 4.8 to 4.11 show the y -oriented time domain electric field from each of the three data techniques at each distance. Fields from each technique are labelled 'Meas.' if data from measurements is used, 'RT' if data from ray tracing simulations is used and 'AE' if data from analytical equation simulations is used. This notation is used in all plots showing data from more than one technique. Figure 4.8 shows e_y for the $(0, -0.65, z)$ m point and Figure 4.9 shows the same data on a 5 ns wide time scale. Similarly, Figure 4.10 shows e_y for the $(0.55, -0.65, z)$ m point and Figure 4.11 shows the same data on a 5 ns wide time scale. Quantitatively, all fields agree as the main groupings of pulses³ are together for each data source, which was also the case for Figure 2.4 that showed only analytical equation data

³The term pulse is used to refer to the arrival of a multipath.

and ray tracer data. However, pulse amplitudes are not in agreement while the shape of each pulse is generally matched, despite the fact that in simulations an infinitesimal dipole, and not a BAVA, was assumed. Considering the distances $z = 4.88$ m and $z = 9.75$ m in each figure, the first two groups of pulses are aligned for each data type and their shapes are also similar. For the responses shown at $z = 14.63$ m, $z = 19.51$ m and $z = 24.38$ m this trend also exists; however, at these distances the measurement data appears to have more arrivals between main pulse arrivals, as its response is not as smooth as the other two techniques. This discrepancy may be attributed to additional reflecting surfaces in the hallway, for example from doorways, so a few experiments with the ray-tracing software that were conducted are now shown.

Figure 4.12 shows e_y for the measurement point $(0, -0.65, 19.51)$ m using two different data sets from the ray tracer and measurement data. The first ray tracer curve was generated using the previously used, basic input file with six surfaces. The second ray tracer curve was generated using an input file that modelled doorways and the PNA equipment rack. The differences between the ray tracer fields are minimal, but there is one reflection that has moved (because the PNA is present in this situation). However, for many other measurement points at this distance there are no visible changes in multipath arrivals, considering time-domain profiles. In this case the second ray tracer curve matches the measurement data, as expected, but it does not match all of the reflections in the measurement data even though all doorways have been modelled. Hence, either the roughness of the hallway, measurement non-idealities or other unaccounted for details are the cause of the simulation and measurement disagreement. In the scope of this research it is not necessary to match measurement results exactly; rather, it is of interest to see if the spatial filter design, which is based on mode theory, can operate in a realistic environment which may not act like an ideal hollow rectangular dielectric waveguide.

Also note that in Figures 4.8 to 4.11 the analytical equation data does not agree with the ray tracer and measurement data for late-arriving multipath arrivals, such as those at about 50 ns and 53 ns at $z = 14.63$ m, 66 ns and 68 ns at 19.51 m and 84 ns at $z = 24.38$ m. This trend is also observable in Figure 2.4 between the analytical equation and ray tracer data for another measurement point. Through additional simulations it can be seen that the direct line-of-sight multipath component from the ray tracer aligns perfectly with the analytical equation model, whereas reflecting components have a smaller amplitude. This observation may suggest that high order modes do not propagate according to the analytical equation model or at all, which is feasible because the model's dimension assumptions are not satisfied for those modes in this narrow tunnel. Modes may attenuate faster than what the waveguide model predicts; however, the ray tracer simulation may also be underestimating reflections as measurement amplitudes sometimes lie between the two models.

Figure 4.8: Comparison of measurement and simulation electric fields (e_y) for the $x = 0$ m point across distance.

Figure 4.9: Comparison of measurement and simulation electric fields (e_y) for the $x = 0$ m point across distance (zoomed view).

Figure 4.10: Comparison of measurement and simulation electric fields (e_y) for the $x = 0.55$ m point across distance.

Figure 4.11: Comparison of measurement and simulation electric fields (e_y) for the $x = 0.55$ m point across distance (zoomed view).

Figure 4.12: Comparison of measurement data and ray tracer electric fields (e_y) from different input files at $(0, -0.65, 19, 51)$ m.

In the case of measurements the moderately-directional BAVA antennas enhance multipath arrivals received from small incidence angles and, effectively, the arrivals from larger incidence angles are received with a lower gain. Correspondingly, the effect of high order modes on the received electric field is smaller. The propagation of the low order modes, which correspond to the first arrivals, aligns very well across the two models and measurement data. Note that the sample spatial filters, that extract up to the (17,2) mode, primarily depend on the first multipath component and rarely depend on later arriving multipaths unless they are spaced very closely to the first arrival or the propagation distance is large. Thus, in terms of showing the spatial filter's functionality all three techniques are suitable.

In Figures 4.8 to 4.11, as well as for all other measurement points, the measurement data had to be time-shifted to align with the analytical equation data and ray tracer data. This shift had to be performed because the time delay in the BAVA antennas was not measured and the delay from changes in cable connections between calibration and measurement were not accounted for. Regarding the antennas' time delay it is estimated to be .886 ns for the two antennas combined. The time delay can be coarsely estimated using the antenna's length and its dielectric substrate's relative permittivity and it is

$$t_g = \frac{l}{c} \sqrt{\epsilon_r} = \frac{0.084}{c} \sqrt{2.5} = 0.443 \text{ ns.} \quad (4.2)$$

This value is an estimate and the actual time delay introduced by the antenna is angle and frequency dependent, which can only be obtained from an anechoic chamber antenna measurement. Connector delays are estimated to be approximately 30 ps. Additionally, small positioning misalignments could

also introduce a delay and because of the non-automized measurement procedure, misalignment errors within ± 1 cm are very reasonable, but devastating in terms of aligning data sets. Accounting for time delay and measurement misalignment in measurement data is possible but difficult; therefore, here each set of measurement data was aligned with simulation data. Across all measurement points the average time shift was 0.7539 ns, which is close to the estimated delay from the antennas alone, as expected because the antenna delay is the primary cause for the offset.

In Figures 4.8 to 4.12 normalized electric fields are plotted. Recall that in Figure 2.4 analytical equation and ray tracer data was compared without normalization and the amplitudes were in agreement. The analytical equation model and ray tracer model can accurately predict field amplitudes due to a radiating infinitesimal dipole, but assumptions have to be made to generate field amplitudes from a BAVA antenna. Additionally, the correct amplitude of e_y cannot be deembedded from measurement data without the antenna's transfer function, which is angle and frequency dependent, and the antenna's frequency dependent input impedance. These quantities were not available and thus unnormalized measurement data cannot be accurately obtained and compared with unnormalized simulation data.

4.5 Results

In this section performance results for several spatial filters are presented and discussed. As Figure 3.2 showed, the signal output from a given spatial filter is presented by appropriately weighting signals from various receiving antennas and then combining the signals.

Recall that the localization accuracy, robustness to noise, and maximum resolvable distance for a given spatial filter is to be determined in order to evaluate the overall ranging ability of that spatial filter. Furthermore, optimal spatial filter configurations are to be determined. The metric used to demonstrate the ranging abilities of a spatial filter is RMSE, which is averaged across many Monte Carlo trials.

The RMSE of a given spatial filter will be shown across transmit and receive SNRs in order to show how well the filter stands up to varying AWGN levels, and how it compares to other spatial filters and benchmarks. Receive SNR was defined in Section 3.3.1 and it will be used to compare spatial filter performance to the CRB, and to compare spatial filter performance using data from the three techniques. Transmit SNR was defined in Section 3.3.2 and it will be used to compare spatial filter performance for varying filter configurations and for a fair comparison to a single sensor. Additionally, transmit SNR can be used to compare filters using data from the ray tracer model and analytical equation model. Filters using data from measurements can also be compared, but they are on an arbitrary scale and can only be compared relative to each other and qualitatively with the analytical equation and ray tracer model

filters. In order to estimate the maximum distance resolvable with the spatial filter a realistic UWB transmit SNR, discussed in Section 3.3, is used while the transmitter to receiver distance is varied.

In the following the performance of ideal spatial filters is first presented in Section 4.5.1 in order to introduce performance trends. A sweep across SNR is used for the ideal filters and a comparison to the CRB and to a single median-performing sensor is completed. In Section 4.5.2 the performance of the six practically-sized proposed spatial filters is evaluated using the RMSE for various receive SNRs. In Section 4.5.3 the performance of the six spatial filters is investigated further, but in terms of varying transmit SNR levels. Finally, in Section 4.5.4 the localization accuracy across distance and maximum ranging distance for the six spatial filters is analyzed.

4.5.1 Performance of an Ideal Spatial Filter

An ideal spatial filter is one that perfectly resolves a given mode using an infinitely dense grid of receiving antennas. Practically, a finite number of closely-spaced sensors can be used to implement the ideal filter without significant aliasing effects creeping in. The performance of several ideal filters is analyzed here to introduce features of RMSE plots and the performance potential of spatial filters.

For the 2.4×2.6 m tunnel, ideal filters using 49×53 antennas (with a uniform spacing of 0.05 m) are simulated using data from the analytical equation model. In the 2.4×2.6 m tunnel there are hundreds of modes; consequently, 49×53 antennas are not quite sufficient to prevent aliasing. However, modes up to $m = 25$ and $n = 2$, which is the maximum mode order considered here, are resolved as they are adequately sampled. Thus, 49×53 antennas are used in a nearly-ideal filter.

In Figure 4.13 the RMSE of ideal spatial filters extracting the (1,1), (9,2) or (17,2) modes are compared to the CRB for varying receive SNR levels. There are two curves for each mode, for example (1,1) a and (1,1) b. The “a” curves use the first theoretical TOA adjustment, presented in Section 3.3, which is $TOA_{theoretical} = TOA_{true} + T_p/2 - \Delta t_{threshold}$ where $\Delta t_{threshold}$ is a factor that is considered for thresholding a fifth order Gaussian derivative. An error floor is reached in all “a” curves because this estimated theoretical TOA will never precisely match a dispersing mode’s TOA. The “b” curves use the second theoretical TOA adjustment, or group velocity adjustment, which considers the thresholded time of arrival that a specific mode has at that distance. With this adjustment, the “b” RMSE curves become more accurate for increasing SNR levels.

In Figure 4.13 that the “b” curves achieve a lower error at mid- to high-SNRs and come very close to the CRB. Since the matched filter uses a fifth order Gaussian derivative template none of the dispersed arriving modes match with it and the CRB is not closely approached. In Figure 4.13(b) the (1,1) b curve

Figure 4.13: RMSE vs. receive SNR plots for various ideal filters.

actually exceeds the CRB because of dispersion and the resolution of the signal is not fine enough for very low error levels. Note that sometimes even when the group velocity adjustment is made, it appears like an error floor is reached, such is the case for the (9,2) b curve in Figure 4.13(c). If the SNR scale was extended it would be seen that this error floor does not extend, rather, the error eventually decreases. Figure 4.14 shows the oscillating ideal spatial filter output for the (9,2) ideal filter at 4.88 m, which is one case where an apparent error floor is reached. Since the threshold is 0.6 the first oscillation passing this may be missed in small amounts of noise. Thus, the apparent error floor behaviour is attributed to the threshold misidentifying the arrival mark on the matched filter output, which can occur even in the smallest amount of noise. RMSE is extremely sensitive to outlying estimates and, therefore, the apparent error floor lasts across a wide range of SNRs.⁴

As mentioned, error floors are reached for the “a” type curves in Figure 4.13, and the error floor’s

⁴The jagged behaviour of RMSE curves in the transition area from a high RMSE to a low RMSE is also due to the RMSE’s sensitivity to outlying estimates. This highlights the importance of performing many Monte Carlo trials, but limits must be used for reasonable simulation times.

Figure 4.14: Matched filter output for the (9,2) ideal filter at 4.88 m.

value decreases for decreasing mode error. This feature is directly attributable to the fact that as mode order increases, mode dispersion increases and, thus, the difference between the mode's TOA and the true TOA increases. For long distances, e.g. 100 m, this error can become very large for high-order modes, e.g. 100 cm or 1% relative error. If this amount of error is tolerable, or a very low order mode is used where the error is small, then this TOA adjustment is acceptable. However, in order to compare the potential overall error of spatial filters that extract different modes it is fairer to consider group velocity adjustments. Furthermore, in system design the knowledge of which mode is to be extracted is known, and thus, group velocity adjustments are feasible. As such, in all further analysis the group velocity adjustment is used; that is, the “b” type curves will be used from now on.

Now the ideal spatial filters will be compared to one another using RMSE across transmit SNR plots, which are shown in Figure 4.15. A single median-performing sensor is also plotted in Figure 4.15, which is selected considering all of the 49×53 sensors used and the three distances used.⁵ At all distances in Figure 4.15 three of the ideal filters perform better than the median-performing single sensor as they operate at lower SNR levels and have more accurate ranging estimates. For the (9,2), (17,2) and (25,2) ideal filters an SNR gain of 7 – 15 dB is achieved across the three distances. Additionally, an error of approximately 2 cm or less (or 0.4% relative error at a maximum) is achieved with the spatial filters. These ideal filters have an SNR performance gain over the single sensor because many antennas are used versus one, and because the filter's outputs match the template signal better in the matched filter, compared to the multipath-corrupted single sensor. SNR performance gains are always achieved in ideal

⁵The median-performing sensor is defined as one that transitions to a low error floor at an SNR that is a median distance away from the low and high SNR transition area limits, considering all sensors. The same sample sensor is used at all distances, so it must approximately have a median-performance at all of these distances.

spatial filters if the extracted mode has significant energy in the transmitting plane. If the number of sensors is doubled in this ideal filter there will be an approximate 3 dB shift left in the error curves in Figure 4.15. In the case of non-ideal or practically-sized filters, performance gains over single sensors cannot be guaranteed because the number of sensors used is not much greater. Additionally, in this case an increase in the number of sensors does not necessarily translate to a performance gain increase because mode undersampling is still present.

Figure 4.15: RMSE vs. transmit SNR plots for various ideal filters.

The best ideal spatial filter for this tunnel scenario is now investigated. At $z = 4.88$ m the (17,2) filter has the most SNR gain, at $z = 14.63$ m the (17,2) filter has the most SNR gain and at $z = 24.38$ m the (9,2) filter has the most gain. Additionally, the (1,1) filter performs the worst at all distances, as it shows no SNR gain. The performance of these ideal spatial filters in noise depends on how much energy and dispersion a mode has at a specific distance. Figure 4.16 helps to illustrate this by showing spatial filter outputs. Outputs with high amplitudes correspond to spatial filters that perform well at

low transmit power levels; for example, at 24.38 m the (9,2) filter has the highest amplitude and also performs the best. On the other hand, the (1,1) mode does not perform well at low transmit power levels as it contains very little energy and is not plotted in Figure 4.16 because it is hardly visible. Additionally, at 4.88 m the (25,2) filter outputs, and at 14.63 m and 24.38 m the (17,2) and (25,2) filter outputs are very dispersed which means that a peak can be easily misidentified in the presence of noise. This was observed in evaluating receive SNR plots where dispersion introduced temporary noise floors and it is seen in transmit SNR plots as well where dispersion reduces the potential SNR gain of a spatial filter.

Figure 4.16: Filter output for various ideal filters.

The change in which filter performs the best at a given distance can also be predicted by looking at the predicted attenuation of a given mode or the energy decomposition of a given mode at a specific distance, as Figure 4.17 shows. The best performing filter will extract a mode with substantial energy at a given distance. Figures 4.17 shows the simulated energy attenuation of three modes and the energy

decomposition of total energy across the entire tunnel aperture into those three modes. The predicted energy attenuation is obtained by

$$U(z, m, n) = \int_{f_{min}}^{f_{max}} ESD(f)(e^{-\alpha_{m,n}z})^2 df, \quad (4.3)$$

where $\alpha_{m,n}$ was provided in Equation (2.23), $ESD(f)$ is an energy spectral density (ESD) that has the shape of the Fourier transform of a fifth order Gaussian derivative, and f_{min} and f_{max} are the -10 dB attenuation frequencies for $ESD(f)$. The energy decomposition of aperture fields into respective modes is obtained by numerically correlating the overall field at a given distance with each mode. That is,

$$U_{decomp}(z, m, n) \approx \frac{\sum_f^{N_f} \rho(z, m, n, f) df}{N_f}, \quad (4.4)$$

where N_f is the number of frequency samples used and

$$\rho(z, m, n, k) = \sum_x^{N_x} \sum_y^{N_y} E_y(x, y, z, f) E_{mn}^{eig}(x, y) dx dy. \quad (4.5)$$

Equation (4.5) is an approximated correlation coefficient for a given mode where $E_y(x, y, z, f)$ is the overmoded electric field across the tunnel's cross-section at a given distance, $E_{mn}^{eig}(x, y)$ is the eigenfunction for a mode (see Equation (2.22)), and N_x and N_y are the number of x and y positions considered, respectively.

In Figure 4.17 the predicted energy attenuation for a given mode is scaled such that at $z = 4.88$ m it is the same value as the energy decomposition. The curves decay at approximately the same rate. Small amounts of aliasing may be contributing to the mode decomposition energy calculation. The main feature of Figure 4.17 is that the higher the mode order, the higher the attenuation rate; thus, at far distances spatial filters based on lower order modes (like (9, 2)) will perform better. Note that in larger tunnels the attenuation rates for these three modes are very close together and the cross-over points do not occur until longer distances.

Returning to Figure 4.15, the (9,2) ideal spatial filter performs the best at $z = 24.38$ m, although the energy cross-over point of the (9,2) mode having maximum energy does not occur until about $z = 100$ m. Therefore, it is primarily dispersion that influences the performance of the (17,2) and (25,2) ideal spatial filters at these close distances. It can be seen that it is very important to consider mode initial energy, attenuation and dispersion in spatial filter design.

Figure 4.17: Mode attenuation in 2.4×2.6 m tunnel.

4.5.2 Performance of Practical Spatial Filters using a Varying Receive SNR

In Figure 4.18 the RMSE performance of two practically-sized spatial filters is presented, alongside the CRB, across varying receive SNR levels for all five measured distances. Data from the analytical equation model, ray tracer model and measurements is plotted in Figure 4.18. For each spatial filter 10,000 Monte Carlo trials were performed.

First note that in comparison to Figure 4.13 these spatial filters do not come as close to the CRB. An error floor is usually reached because the filter's output does not necessarily match the mode's pulse shape precisely, on which the TOA values are based. Error floors of approximately 10 cm or less (or 2% or less of relative error at $z = 4.88$ m) are achieved across all distances. Secondly, the shapes of the RMSE performance curves for each filter configuration are similar across the three techniques. For example, for the (17,2) 13 receiver spatial filter at $z = 4.88$ m the AE, RT and Meas. curves all have a similar shape. At other distances there is, at times, horizontal shifts between the AE, RT and Meas. curves for the same filter. These horizontal shifts are because the one filter output (e.g. AE) may have more energy than another (e.g. RT), and thus, more noise will be added to the former output for a given receive SNR. Therefore, the horizontal shifts can be disregarded and only the shapes of the curves from the different techniques should be evaluated. In general, curve shapes agree and the resultant error floors are similar, especially for the AE and RT data.

Figure 4.18: RMSE vs. receive SNR plots for sample spatial filters.

4.5.3 Performance of Practical Spatial Filters using a Varying Transmit SNR

All six proposed spatial filters presented in Section 4.2 are now evaluated and compared in terms of their RMSE performance across varying transmit SNRs. In Figure 4.19 all six spatial filters' results are plotted using data from the analytical equation model, in order to see the relative performance between them. In Figures 4.20 to 4.22 spatial filter results are plotted at select distances using data from the analytical equation model and ray tracer model to allow an easy performance comparison. In

Figure 4.23 spatial filters using data from measurements are shown. Recall that the correct amplitude of the measurement data was not deembedded and, thus, it has to be arbitrarily scaled so that a receiver noise is proportional. As a result, the transmit SNR scale for measurement data is arbitrary. Finally, Figure 4.24 offers a performance comparison to another spatial filtering situation where a center transmitter is used and the lowest order mode is extracted. Five-thousand Monte Carlo trials were used in all RMSE calculations.

In Figures 4.19 to 4.23 single median-performing sensors are also plotted and they correspond to the same sensor for each data type. This sensor was chosen as an analytical equation data sensor whose error had a median performance in comparison to the other twelve sensors, across all five distances considered. The median sensor used is that at $(0, -0.65, z)$ m. This sensor performs better at some distances compared to others, relative to the spatial filters. In Figure 4.19 some spatial filters, like the $(17, 2)$ -type filters, typically outperform the single sensor by a large margin. For example $(17, 2)$ -type filters have an SNR gain of usually 1 to 4 dB over the single sensors. One advantage of any spatial filter over a single sensor is that the spatial filter's performance is more dependable across distance. A single sensor's performance highly depends on distance because the tunnel is highly overmoded and the electric field as a function of longitudinal distance fluctuates rapidly.

Considering the spatial filters plotted alongside one another in Figure 4.19, the question of why a given filter works better than another does not have a simple answer. In Section 4.5.1 it was observed that ideal filters that extracted modes with high initial energy, and reasonable attenuation and dispersion in the range of interest performed the best. This trend is also partially observed here as the $(17, 2)$ -type filters perform well, but that may not be the only reason they perform well. Additionally, the $(9, 2)$ -type filters perform quite poorly in contrast to the ideal filter case.

Recall that in Table 4.2 modes admitted, or minimally attenuated, by each of the six spatial filters were listed. For the $(8, 2)$ - and $(9, 2)$ -type spatial filters many modes admitted lay around $m = 7$ to 10 or $m = 23$ to 27 and $n = 1$ to 3, whereas for the $(17, 2)$ -type spatial filters modes admitted mainly ranged from $m = 13$ to 19 and $n = 1$ to 3. Due to these distributions the $(17, 2)$ filters can gather much more energy as the modes in the range $m = 15$ to 19 and $n = 2$ to 3 have more cumulative and focused energy than the other spatial filter types considered.

Regarding the number of sensors used in the spatial filter, using 13 versus 5 sensors in the $(17, 2)$ -type spatial filters or 9 versus 5 sensors in the $(9, 2)$ -type spatial filters and so on has no guaranteed advantage and sometimes results in poorer performance. This is because all of these spatial filters are undersampling the modes in the tunnel. A guarantee in performance improvement can only be made if the number of sensors is increased in a situation where the tunnel aperture fields are already adequately

Figure 4.19: RMSE vs. transmit SNR plots for various spatial filters using analytical equation data.

sampled in space.

Transmit SNR RMSE results are now compared using ray tracer data as well. Recall that the electric fields from the analytical equation model and ray tracer model model an infinitesimal dipole and have similar amplitudes. Thus, AE and RT spatial filter results can be fairly compared using a transmit SNR definition and this is done in Figures 4.20 to 4.22 for select distances.

In Figure 4.20 RMSE for the (17,2)-type spatial filters is presented for the analytical equation and ray

Figure 4.20: RMSE vs. transmit SNR plots for (17,2)-type spatial filters using analytical equation and ray tracer data.

Figure 4.21: RMSE vs. transmit SNR plots for (9,2)-type spatial filters using analytical equation and ray tracer data.

Figure 4.22: RMSE vs. transmit SNR plots for (8,2)-type spatial filters using analytical equation and ray tracer data.

Figure 4.23: RMSE vs. transmit SNR plots for spatial filters using measurement data.

Figure 4.24: RMSE vs. transmit SNR plot at 25 m for a (1,1) spatial filter designed for a center transmitter.

tracer techniques, alongside their appropriate one sensor samples. The spatial filters have a performance gain versus the one sensor of 1 to 4 dB considering analytical equation predictions and a performance loss considering ray tracer predictions. The disagreeing ray tracer results are attributed to the fact that late-arriving multipaths have a low magnitude according to the ray tracer model and these multipaths may contribute to the (17,2) mode. In Figures 4.21 and 4.22 the RMSE for the (9,2)-type and (8,2)-type spatial filters, respectively, is presented in a similar format. The analytical equation and ray tracer models predict more of a similar performance for these filters, especially for the (9,2) 5 receiver and (8,2)-type spatial filters. The analytical equation model's predicted performance for the (9,2)- and (8,2)-type filters has a performance loss over the single sensor, but the ray tracer model's predicted performance for these filters has a performance gain at $z = 19.51$ m of approximately 2 to 4 dB.

Regarding the measurement data in Figures 4.23 it is noted that the one sensor's performance is poorer than more spatial filters than in previous cases. At $z = 9.75$ m two filters have at least a 2.5 dB performance gain over the single sensor and at $z = 19.51$ m five filters have performance gains of 2 to

5 dB over the single sensor.

The six designed spatial filters are suitable in the situation where the transmitter is very close to the tunnel wall. In order to show that a spatial filter can also improve the TOA localization performance in the situation where the transmitter is placed in the center of the tunnel's aperture Figure 4.24 is shown. Figure 4.24 shows the performance of a (1,1) 5 receiver spatial filter in the hallway where a center transmitter is used and its performance is compared to the performance of each of the 5 individual receivers in it. Receiver positions are at $x = 0, \pm 0.15, \pm 0.3$ m and $y = 0$ m. In Figure 4.24 it is seen that the spatial filter's performance is better than the individual receivers as it performs at lower transmit power levels and reaches a lower error. Thus, a spatial filter can also improve localization performance in this centrally-located transmitter case, where less modes are propagating.

4.5.4 Performance of Practical Spatial Filters across Distance using a Constant Transmit SNR

The performance of the sample six spatial filters in various noise levels has so far been analyzed for five distances, because at these five distances measurements were taken. Now, by using a reasonable receiver temperature of 460 K, which was calculated in Section 3.3.2, each spatial filter's and the single median-performing sensor's performance is analyzed across distance. The pulse repetition rate of the unmodulated pulse train is set to be 20 MHz and Equation (3.21) is satisfied by finding an appropriate scaling value. With regard to the transmit SNR plots shown in Section 4.5.3, all spatial filters using analytical equation data are in the low RMSE error region for this SNR which is 56 dB.

Figure 4.25 shows the RMSE averaged over 5000 trials versus distance for the six spatial filters, three ideal spatial filters and median-performing sensor used previously. Data from the analytical equation model is used. In Figure 4.25 it is seen that usually a spatial filter's error is better overall than the single sensor across distance, because the single sensor's error is higher. For example, the spatial filter's error is usually less than 10 cm (2% to 0.1% relative error at the minimum and maximum distances, respectively), while the single sensor's error is usually greater than 10 cm. This enforces the fact that a spatial filter delivers more precise TOA ranging. Furthermore, all the spatial filters can be used for a slightly longer distance than the single sensor, whose error spikes at approximately 62 m.

In Figure 4.25 a constant offset was used across distance to accommodate for the difference between the true TOA and the mode's TOA. The offset chosen corresponds to that at $z = 25$ m. If variable offsets are used across distance the spatial filters' error will be less across distance for all spatial filters, especially the (17,2)-type spatial filters.

Recall that it was mentioned that modes in larger tunnels attenuate slower. Therefore, in these large tunnels spatial filters can be used for greater longitudinal distances. The range for a 5×5 m tunnel can be predicted using Figure 4.26. In Figure 4.26 the energy across a tunnel's aperture is calculated at various distances in the 2.4×2.6 m hallway and a 5×5 m tunnel. The energy is calculated using

$$U_{aperture}(z) = \sum_f^{N_f} \sum_x^{N_x} \sum_y^{N_y} \frac{1}{2\eta} |E_y(f, x, y, z)|^2 dx dy df, \quad (4.6)$$

where quantities are as previously defined. The energy attenuation in a tunnel (approximately 12 to 15 dB/decade for the smaller tunnel) is less than that in free-space, which is 20 dB/decade. At approximately 70 – 75 m the error for practically-sized spatial filters in the 2.4×2.6 m tunnel becomes very large and the corresponding energy level is 61.5 to 62 J. By intersecting this energy value range with the 5×5 m curve a distance range of 175 – 200 m can be predicted for spatial filters operating in it.

Figure 4.25: RMSE vs. distance for various spatial filters.

Figure 4.26: Energy decay in waveguides.

4.5.5 Results Summary

It has been shown that a spatial filter, that extracts a mode propagating in a tunnel, can be used for TOA localization. Ideal spatial filters were presented, in which many sensors were used. These ideal filters were compared to the CRB using receive SNR and it was shown that the ideal filters came very close to the CRB, especially for low order modes. Following this the ideal filters were compared to one another and to a single sensor using transmit SNR, and it was noted that the best performing ideal spatial filters used a mode that had a strong initial energy in the tunnel, attenuated slowly and, most importantly, did not disperse considerably in the range of distances considered.

Practically-sized spatial filters, using 4 to 13 sensors, were then analyzed using data from an analytical equation model, a ray tracer model and measurements. The three data techniques were compared for select filters and compared to the CRB using receive SNR. Generally, the data agreed which verified the models and measurements. To gauge the performance of the spatial filters further, the six spatial filters were compared to one another and to a single sensor using transmit SNR. Considering the analytical equation model's data, it was observed that the (17,2)-type spatial filters usually offered 1 to 4 dB in performance gain over the single sensors. The other spatial filters always reached a lower error floor than the single sensor. Considering the ray tracer model's data, spatial filter results agreed with the analytical equation model for most filter types, albeit discrepancies for the (17,2)-type filters attributed to the fact that the ray tracer predicts low amplitudes for mid-order modes. Performance gains were achieved at longer distances considering ray tracer model predictions. Considering measurement data, the sample single sensor usually performed poorer than spatial filters, and the spatial filters' performance was in

accordance with the models. For measurement data more spatial filter performance gains over the single sensor were observed. For example, five of the spatial filters had a performance gain of 2 to 5 dB over the single sensor at $z = 19.51$ m.

At different tunnel distances the six spatial filters performed better than the single sensor as a low error was achieved and many spatial filters could extend the useful range of the localization system. The predicted range in the hallway was estimated to be 70 – 75 m, while complying to the FCC UWB average emissions mask. In a larger 5×5 m tunnel the spatial filters' ranges was predicted to be 175 – 200 m.

Chapter 5

Conclusion

UWB signals are extremely well suited to localization applications as short UWB pulses enable centimeter level accuracy. Due to multipath propagation in a tunnel the potential localization accuracy available by UWB signals can not be realized in simple one-transmitter, one-receiver TOA localization systems. Thus, a filtering technique has been proposed to improve the performance of one-way TOA systems in a tunnel.

Prior to proposing the spatial filter, propagation in a tunnel was studied so that propagation behaviour could be considered in the filter's design. An analytical equation waveguide model was described and mode weighting coefficients were appropriately derived for a vertically-polarized infinitesimal dipole in the hollow rectangular dielectric waveguide by using a source expansion. This approach for deriving mode coefficients has not previously been used for tunnel modes. A ray tracer model was also used in this research and it was compared to the analytical equation model to show that rays and modes are inherently different and not equivalent, although both models rely on the assumption that the wavelength of operation is much smaller than the tunnel's dimensions. A comparison of time-domain electric fields predicted by each of the models was presented and it was seen that the models agree for the initial part of the response, but differ for later multipath arrivals as once or more reflected rays have much smaller amplitudes than late arriving modes. However, it is the initial part of the response that is crucial as it corresponds to low- to mid-order modes.

By exploiting the fact that electromagnetic waves propagate as modes in a tunnel a spatial filtering technique, that uses a weighted array of judiciously placed antennas, was proposed and it effectively filters the incidence UWB electric field and allows only one mode, ideally, to pass. As each mode has a unique UWB pulse that propagates, the filtered signal essentially contains only one multipath component

and, thus, very accurate or perfect TOA estimates can be made. This spatial filtering technique is very simple as there is no training required to set filter parameters. However, practically not only one mode will be extracted by the filter due to practical limits on the number of antennas that can be used in the filter's implementation. As a result, spatial undersampling will occur in the filter and multiple modes will be passed by the filter. Design steps for a spatial filter with a limited number of antennas were presented in Chapter 3. In the one-way TOA system proposed a single transmitting antenna mounted on the side of the tunnel sends a pulse train of mask-compliant UWB pulses and the spatial filter would receive, weight and combine the signals and the resultant signal would proceed to pass through a matched filter and threshold detector for TOA determination.

Several sample spatial filters were designed for a hallway, that was perhaps smaller than most tunnels but was easily accessible for measurements. The sample spatial filters were evaluated using data from three techniques: measurements, the analytical equation model and the ray tracer model. The RMSE of the TOA estimate from a spatial filter's output was evaluated various ways to gauge a given spatial filter's performance. The RMSE was evaluated for various receive SNRs to compare to theoretical bounds and the RMSE was evaluated for various transmit SNRs to compare various spatial filters and to compare to a single median-performing sensor. Moreover, it was ensured that UWB mask-compliant signals were used to evaluate RMSE versus distance for spatial filters.

Sample ideal spatial filters were first evaluated. Each of the ideal spatial filters provided precise localization accuracy and approached the theoretical bound established for noise-limited TOA localization systems. However, depending on the mode that was extracted with the spatial filter, the performance varied. Best performing ideal spatial filters extract a mode with sufficient energy and reasonable dispersion in the distance range considered. All ideal spatial filters considered in this work that extracted a mode with reasonable strength had a 7 to 15 dB SNR advantage over the single sensor and had an error of 2 cm at most compared to approximately 15 to 40 cm with the single sensor.

Six practical spatial filters were then evaluated using the three techniques. The six filters considered used a combination of up to 13 antenna positions and extracted different modes so that performance trends could be seen. The spatial filters did not typically approach the CRB, due to significant aliasing of multiple modes by the sparse array, but still delivered approximately 10 cm of error or less across distance (or a maximum of 2% relative error at the closest distance considered), which is more than accurate for most tunnel localization applications. Spatial filter results from the three techniques were in general agreement as the RMSE profile versus varying SNR levels was approximately the same shape for a given filter.

Considering data from the analytical equation model at various distances, the spatial filters designed

to extract the (17,2) mode primarily performed the best in comparison to a single sensor in varying transmit SNR levels as they generally offered 1 to 4 dB of performance gain. The spatial filters designed to extract the (8,2) or (9,2) modes did not perform well at low transmitter power levels, but still provided a very good error floor. The (8,2)- and (9,2)-type spatial filters performed poorer because these filters aliased modes around (25,2) and, thus, the spatial filter output did not have a heavy energy concentration in one mode range, like the (17,2)-type filters that primarily aliased modes around $m = 13$ to 19 or at much higher indices where the modes decayed quickly and did not affect filter performance. In summary, the effect of mode aliasing must be considered in spatial filter design and it is more advantageous to alias neighbouring modes than modes spaced further apart, considering that all modes have a significant presence in the tunnel.

However, when spatial filter results were analyzed considering data from measurements and the ray tracer model the (8,2)- and (9,2)-type filters performed better, especially at the longer distance considered. At least a 2 dB performance gain was afforded by most of these spatial filters at 19.51 m. The (17,2)-type filters did not perform as well at this distance, for these two techniques, perhaps because mid- to high-order modes decay quicker than the analytical equation model predicts and, in the case of measurements, a moderately-directional antenna was used where this mode was not fully received. This observation highlights the importance of performing rigorous model validations; the limits on mode theory applicability need to be defined.

With respect to this research's goals identified in Section 1.3 all goals have been achieved. A spatial filter has been presented that clearly improves the performance of one-way TOA systems in a tunnel, which was the first research goal, by offering SNR gain and reduced error levels if the spatial filter is well designed. Furthermore, the distance range for a well-designed spatial filter is better than that for a single sensor. By way of meeting the first research goal, the second research goal of investigating the performance of the filter compared to other systems not equipped with a filter has also been achieved.

Finally the last goal of this research, to explore different filter configurations, has also been achieved. The best performing spatial filter for a given tunnel scenario is not well defined, but many insights have been gained in this research. The extracted mode should have reasonable energy and dispersion, and aliased modes should be neighbouring modes or contain minimal energy. Furthermore, reconfigurable spatial filters may be appropriate for spatial filters operating over large distance ranges, as high order modes may be more appropriate to extract at close distances and low order modes are more appropriate to extract at further distances. Reconfigurable spatial filters may use a limited number of receivers and the weights of a given receiver will vary with distance to deliver optimal performance. Ideal spatial filters are practically unreasonable; however, even if 25 sensors were used across the x -axis and one sensor was

used across the y -axis for the offset-transmitter scenario considered a very well-performing spatial filter could be realized because for modes with $m > 25$ initial energy decreases and their amplitudes rapidly attenuate across distance. Regarding the number of receivers used in the spatial filter, design constraints must be considered and an increase in performance does not always correlate with an increase in the number of receivers used. The modes extracted or aliased by the filter also highly influences the number of receivers used. In the design example presented in this research, 4 to 13 receivers were shown to be adequate, which is practically implementable. Using more receivers does, however, increase system robustness in situations of receiver blockage and other non-tunnel like propagation behaviour. In the situation that different transmitter locations are used, such as a centrally located transmitter, spatial filter design can be simplified considerably as a low order mode can be extracted and the performance of such a filter will still be better than single sensors in this scenario. In conclusion, when carefully designed the spatial filter is very valuable for TOA localization in a tunnel as it can deliver accurate TOA estimates, can provide better ranging than a single sensor and considers multipath propagation in a tunnel, which current systems lack.

5.1 Contributions

In this research the following article was presented at the International Conference on Ultra-Wideband in September 2012:

N. Jones and S. Hum, "An Ultra-wideband Spatial Filter for Tunnel Environments," in *IEEE International Conference on Ultra-Wideband*, Sep. 2012, pp. 491-495.

5.2 Future Work

The spatial filter was presented as part of a proof-of-concept TOA localization system. Therefore, there are many potential research avenues that can be taken with the tunnel propagation models, spatial filter and localization system presented in this research.

The analytical equation model presented was for mode propagation in straight, rectangular tunnels with no discontinuities. Realistically, a tunnel like this is hard to find. In mining tunnels, walls may be uneven and the geometry of the tunnel may be varying. In train tunnels, rails, platforms and branches are present. In vehicular tunnels, there are vehicles and the shape is rarely rectangular. In pedestrian tunnels or hallways, there are people, objects and doorways. The effect of different tunnel geometries, discontinuities and objects on mode propagation needs to be studied to evaluate if mode theory can

still be applied. The framework of the analytical equation model presented can be used to study modes in tunnels with different cross-sectional shapes. For example, modes in different cross-sections can be predicted using numerical software packages like Comsol, and these modes can be approximated by basis functions in the waveguide model. Various antenna types or multiple antennas may also be integrated into the analytical equation model by deriving or approximating appropriate mode weighting coefficients. The waveguide model presented cannot, however, be used to study discontinuities and objects, unless additional factors are introduced. The ray tracer model in conjunction with measurements is more appropriate for studying these effects. These effects need to be characterized as no tunnels are ideal and the spatial filtering concept relies on mode propagation in ideal tunnels.

In the TOA localization system proposed many assumptions were made, such as that one-way measurements are accurate because the transmitter and receiver are time synchronous, clock error and jitter are negligible, the threshold was adaptable, the TOA calculation considered timing offsets and the output was fully digitally resolved. First, the effect of clock error and jitter on a one-way TOA system is typically significant and needs to be studied. Furthermore, the implementation of two-way TOA or TDOA localization could be investigated. Threshold algorithms and automatic gain control based thresholds could also be proposed so that spatial filter outputs are accurately and optimally thresholded across distance. The ability to change timing offsets versus distance could also be investigated by using a system that monitors range in real-time to adjust the offset appropriately. In such a system, reconfigurable spatial filter weights could also be used to, for example, extract high order modes at close distances and low order modes at further distances. Finally, the effect of digital sampling of UWB spatial filter outputs must be studied to see how accurate TOA estimates can realistically be.

With respect to the antennas used in the TOA localization system, their design could be optimized. For example, in the spatial filter high order modes have large incidence angles when they are received, especially at low frequencies; thus, they can be filtered out by using a directional antenna. Transmitting antenna(s) and their configurations can also be proposed to launch specific modes to alleviate the filtering burden on the receiving side of the system.

Appendix A

Derivation of Fields in a Hollow Rectangular Dielectric Waveguide

In the following the derivation of the approximate electromagnetic fields inside a hollow rectangular dielectric waveguide is shown. This derivation follows and expands steps outlined in the most complete derivation found [27]. Numerous errors were found in the original derivation and some nontrivial steps were not provided. This derivation aims at providing a detailed explanation of steps taken and shows many intermediates steps.

This Appendix is to be referred to alongside Section 2.1.1. In Section 2.1.1 the waveguide dimensions and regions are shown, and assumptions are listed. An exact analytical solution is not possible for the geometry at hand, and the fields can be derived only when the assumptions in Section 2.1.1 are used.

A.1 Field Formulations

Fields are written from Helmholtz's equation and are then matched using the boundary conditions on the sides of the waveguide. The homogeneous vector Helmholtz equation for the electric field is

$$\nabla^2 \vec{E} + k^2 \vec{E} = 0, \quad (\text{A.1})$$

where $k = \omega \sqrt{\mu_0 \epsilon}$ is the wavenumber, $\vec{E}(x, y, z; t) = [e_{xy}^-(x, y) + e_z(x, y)\hat{z}]e^{j\omega t - \gamma z}$, and $\gamma = \alpha + j\beta$ is the propagation constant. Equation A.1 can be rearranged as

$$\begin{aligned} \nabla^2 \vec{E} + k^2 \vec{E} &= (\nabla_{xy}^2 + \nabla_z^2) \vec{E} + k^2 \vec{E} = (\nabla_{xy}^2 + \frac{\partial^2}{\partial z^2}) \vec{E} + k^2 \vec{E} \\ &= \nabla_{xy}^2 \vec{E} + (-\gamma)^2 \vec{E} + k^2 \vec{E} = \nabla_{xy}^2 \vec{E} + (k^2 + \gamma^2) \vec{E} = 0, \end{aligned} \quad (\text{A.2})$$

Similarly the same steps can be taken with the magnetic field which also results in $\nabla_{xy}^2 \vec{H} + (k^2 + \gamma^2) \vec{H} = 0$. Equation A.2 and the equivalent equation for the magnetic field are both second order partial differential equations. They can be decomposed into components (x, y, z) such that there are six

equations. These equations are not all independent. From the source-free curl equations

$$\nabla \times \vec{E} = -j\omega\mu_0\vec{H} \implies \begin{cases} \frac{\partial E_z}{\partial y} + \gamma E_y = -j\omega\mu_0 H_x \\ -\gamma E_x - \frac{\partial E_z}{\partial x} = -j\omega\mu_0 H_y \\ \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -j\omega\mu_0 H_z \end{cases} \quad (\text{A.3})$$

$$\nabla \times \vec{H} = j\omega\epsilon\vec{E} \implies \begin{cases} \frac{\partial H_z}{\partial y} + \gamma H_y = j\omega\epsilon E_x \\ -\gamma H_x - \frac{\partial H_z}{\partial x} = j\omega\epsilon E_y \\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = j\omega\epsilon E_z \end{cases} \quad (\text{A.4})$$

These equations can be written in terms in E_z and H_z which results in the following

$$E_x = -\frac{1}{h^2} \left(\gamma \frac{\partial E_z}{\partial x} + j\omega\mu_0 \frac{\partial H_z}{\partial y} \right), \quad (\text{A.5})$$

$$E_y = -\frac{1}{h^2} \left(\gamma \frac{\partial E_z}{\partial y} - j\omega\mu_0 \frac{\partial H_z}{\partial x} \right), \quad (\text{A.6})$$

$$H_x = -\frac{1}{h^2} \left(\gamma \frac{\partial H_z}{\partial x} - j\omega\epsilon \frac{\partial E_z}{\partial y} \right), \quad (\text{A.7})$$

$$H_y = -\frac{1}{h^2} \left(\gamma \frac{\partial H_z}{\partial y} + j\omega\epsilon \frac{\partial E_z}{\partial x} \right). \quad (\text{A.8})$$

In these equations $h^2 = \gamma^2 + k^2$ is used.

Now the fields for each region in Figure 2.1 are written: the inner region, region a and region b. Solutions of E_z and H_z from Helmholtz's equations are provided and then other field components are found using Equations (A.5) to (A.8).

A.1.1 Inside Field Components

In this region the transverse field components have a sinusoidal, standing wave behavior and the longitudinal component is a travelling wave. The notations ϕ_x and ϕ_y are used to account for even and odd symmetry cases (that is, if $\phi_x, \phi_y = 0$ there is even symmetry, if $\phi_x, \phi_y = \pi/2$ there is odd symmetry). The longitudinal components are

$$E_z^i = \mathcal{E}_z^i \cos(k_x^i x + \phi_x) \cos(k_y^i y + \phi_y) e^{-jk_z z}, \quad (\text{A.9})$$

$$H_z^i = \mathcal{H}_z^i \sin(k_x^i x + \phi_x) \sin(k_y^i y + \phi_y) e^{-jk_z z}, \quad (\text{A.10})$$

where

- $\gamma = jk = \sqrt{(jk_x^i)^2 + (jk_y^i)^2 + (jk_z)^2}$ ¹, thus, $k = \sqrt{(k_x^i)^2 + (k_y^i)^2 + (k_z)^2} = \sqrt{\omega^2 \mu_0 \epsilon_0}$ = wavenumber in inner region,
- k_z = wavenumber in z direction,
- k_x^i = internal wavenumber in x direction,
- k_y^i = internal wavenumber in y direction,

¹This definition is used to be in accordance with notation in [27].

- ϵ_0 = internal complex dielectric constant (free-space),
- $\mathcal{E}_z^i$ = internal electric field amplitude in z direction and
- $\mathcal{H}_z^i$ = internal magnetic field amplitude in z direction.

By using Equations (A.5) to (A.8) the remaining field components can be found and are

$$\begin{aligned} E_x^i &= -\frac{1}{h^2} \left[-\gamma k_x^i \mathcal{E}_z^i + j\omega\mu_0 k_y^i \mathcal{H}_z^i \right] \sin(k_x^i x + \phi_x) \cos(k_y^i y + \phi_y) e^{-jk_z z} \\ &= \frac{j\omega\mu_0}{k^2 - k_z^2} \left[\frac{k_z k_x^i}{\omega\mu_0} \mathcal{E}_z^i - k_y^i \mathcal{H}_z^i \right] \sin(k_x^i x + \phi_x) \cos(k_y^i y + \phi_y) e^{-jk_z z}, \end{aligned} \quad (\text{A.11})$$

$$E_y^i = -\frac{1}{h^2} \left[-\gamma k_y^i \mathcal{E}_z^i - j\omega\mu_0 k_x^i \mathcal{H}_z^i \right] \cos(k_x^i x + \phi_x) \sin(k_y^i y + \phi_y) e^{-jk_z z} \quad (\text{A.12})$$

$$= \frac{j\omega\mu_0}{k^2 - k_z^2} \left[\frac{k_z k_y^i}{\omega\mu_0} \mathcal{E}_z^i + k_x^i \mathcal{H}_z^i \right] \cos(k_x^i x + \phi_x) \sin(k_y^i y + \phi_y) e^{-jk_z z}, \quad (\text{A.13})$$

$$\begin{aligned} H_x^i &= -\frac{1}{h^2} \left[\gamma k_x^i \mathcal{H}_z^i + j\omega\epsilon k_y^i \mathcal{E}_z^i \right] \cos(k_x^i x + \phi_x) \sin(k_y^i y + \phi_y) e^{-jk_z z} \\ &= -\frac{j\omega\epsilon_0}{k^2 - k_z^2} \left[k_y^i \mathcal{E}_z^i + \frac{k_z k_x^i}{\omega\epsilon} \mathcal{H}_z^i \right] \cos(k_x^i x + \phi_x) \sin(k_y^i y + \phi_y) e^{-jk_z z}, \end{aligned} \quad (\text{A.14})$$

$$\begin{aligned} H_y^i &= -\frac{1}{h^2} \left[\gamma k_y^i \mathcal{H}_z^i - j\omega\epsilon k_x^i \mathcal{E}_z^i \right] \sin(k_x^i x + \phi_x) \cos(k_y^i y + \phi_y) e^{-jk_z z} \\ &= \frac{j\omega\epsilon_0}{k^2 - k_z^2} \left[k_x^i \mathcal{E}_z^i - \frac{k_z k_y^i}{\omega\epsilon} \mathcal{H}_z^i \right] \sin(k_x^i x + \phi_x) \cos(k_y^i y + \phi_y) e^{-jk_z z}. \end{aligned} \quad (\text{A.15})$$

Note that in these equations $\gamma \approx jk_z$ and $h^2 = (jk_z)^2 + k^2 = k^2 - k_z^2$ because there is propagation in the $+z$ direction.

A.1.2 Region a Field Components

In this region the field components have a sinusoidal distribution in the y direction and a decaying exponential distribution in the x direction. Definitions presented for the inner field region hold for this region as well, and additional definitions in this region are

- $\gamma_a = jk_a = \sqrt{(\alpha_x^a)^2 + (jk_y^i)^2 + (jk_z)^2}$, thus, $k_a = \sqrt{-(\alpha_x^a)^2 + (k_y^i)^2 + (k_z)^2} = \sqrt{\omega^2 \mu_0 \epsilon_a} =$ wavenumber in region a ,
- α_x^a = region a wavenumber (attenuation) in the x direction,
- ϵ_a = region a complex dielectric constant,
- $\mathcal{E}_z^a$ = region a electric field amplitude in z direction and
- $\mathcal{H}_z^a$ = region a magnetic field amplitude in z direction.

Region a field components are (assuming that $x \geq a$ and $\alpha_x^a \geq 0$)

$$E_z^a = \mathcal{E}_z^a \cos(k_y^i y + \phi_y) e^{-\alpha_x^a x} e^{-jk_z z}, \quad (\text{A.16})$$

$$H_z^a = \mathcal{H}_z^a \sin(k_y^i y + \phi_y) e^{-\alpha_x^a x} e^{-jk_z z}, \quad (\text{A.17})$$

$$\begin{aligned} E_x^a &= -\frac{1}{h^2} \left[-\gamma \alpha_x^a \mathcal{E}_z^a + j\omega\mu_0 k_y^i \mathcal{H}_z^a \right] \cos(k_y^i y + \phi_y) e^{-\alpha_x^a x} e^{-jk_z z} \\ &= \frac{j\omega\mu_0}{k_a^2 - k_z^2} \left[\frac{\alpha_x^a k_z}{\omega\mu_0} \mathcal{E}_z^a - k_y^i \mathcal{H}_z^a \right] \cos(k_y^i y + \phi_y) e^{-\alpha_x^a x} e^{-jk_z z}, \end{aligned} \quad (\text{A.18})$$

$$\begin{aligned} E_y^a &= -\frac{1}{h^2} \left[-\gamma k_y^i \mathcal{E}_z^a + j\omega\mu_0 \alpha_x^a \mathcal{H}_z^a \right] \sin(k_y^i y + \phi_y) e^{-\alpha_x^a x} e^{-jk_z z} \\ &= \frac{j\omega\mu_0}{k_a^2 - k_z^2} \left[\frac{k_y^i k_z}{\omega\mu_0} \mathcal{E}_z^a - \alpha_x^a \mathcal{H}_z^a \right] \sin(k_y^i y + \phi_y) e^{-\alpha_x^a x} e^{-jk_z z}, \end{aligned} \quad (\text{A.19})$$

$$\begin{aligned} H_x^a &= -\frac{1}{h^2} \left[-\gamma \alpha_x^a \mathcal{H}_z^a + j\omega\epsilon_a k_y^i \mathcal{E}_z^a \right] \sin(k_y^i y + \phi_y) e^{-\alpha_x^a x} e^{-jk_z z} \\ &= \frac{j\omega\epsilon_a}{k_a^2 - k_z^2} \left[-k_y^i \mathcal{E}_z^a + \frac{k_z \alpha_x^a}{\omega\epsilon_a} \mathcal{H}_z^a \right] \sin(k_y^i y + \phi_y) e^{-\alpha_x^a x} e^{-jk_z z}, \end{aligned} \quad (\text{A.20})$$

$$\begin{aligned} H_y^a &= -\frac{1}{h^2} \left[\gamma k_y^i \mathcal{H}_z^a - j\omega\epsilon_a \alpha_x^a \mathcal{E}_z^a \right] \cos(k_y^i y + \phi_y) e^{-\alpha_x^a x} e^{-jk_z z} \\ &= \frac{j\omega\epsilon_a}{k_a^2 - k_z^2} \left[\alpha_x^a \mathcal{E}_z^a - \frac{k_z k_y^i}{\omega\epsilon_a} \mathcal{H}_z^a \right] \cos(k_y^i y + \phi_y) e^{-\alpha_x^a x} e^{-jk_z z}. \end{aligned} \quad (\text{A.21})$$

A.1.3 Region b Field Components

In this region the field components have a sinusoidal distribution in the x direction and a decaying exponential distribution in the y direction. Additional definitions in this region are

- $\gamma_b = jk_b = \sqrt{(jk_x^i)^2 + (\alpha_y^b)^2 + (jk_z)^2}$, thus, $k_b = \sqrt{(k_x^i)^2 - (\alpha_y^b)^2 + (k_z)^2} = \sqrt{\omega^2 \mu_0 \epsilon_b}$ = wavenumber in region b ,
- α_y^b = region b wavenumber (attenuation) in the y direction,
- ϵ_b = region b complex dielectric constant,
- $\mathcal{E}_z^b$ = region b electric field amplitude in z direction and
- $\mathcal{H}_z^b$ = region b magnetic field amplitude in z direction.

Region b field components are (assuming that $y \geq b$ and $\alpha_y^b \geq 0$):

$$E_z^b = \mathcal{E}_z^b \cos(k_x^i x + \phi_x) e^{-\alpha_y^b y} e^{-jk_z z}, \quad (\text{A.22})$$

$$H_z^a = \mathcal{H}_z^b \sin(k_x^i x + \phi_x) e^{-\alpha_y^b y} e^{-jk_z z}, \quad (\text{A.23})$$

$$\begin{aligned} E_x^b &= -\frac{1}{h^2} \left[-\gamma k_x^i \mathcal{E}_z^b - j\omega\mu_0 \alpha_y^b \mathcal{H}_z^b \right] \sin(k_x^i x + \phi_x) e^{-\alpha_y^b y} e^{-jk_z z} \\ &= \frac{j\omega\mu_0}{k_b^2 - k_z^2} \left[\frac{k_x^i k_z}{\omega\mu_0} \mathcal{E}_z^b + \alpha_y^b \mathcal{H}_z^b \right] \sin(k_x^i x + \phi_x) e^{-\alpha_y^b y} e^{-jk_z z}, \end{aligned} \quad (\text{A.24})$$

$$\begin{aligned} E_y^b &= -\frac{1}{h^2} \left[-\gamma \alpha_y^b \mathcal{E}_z^b - j\omega\mu_0 k_x^i \mathcal{H}_z^b \right] \cos(k_x^i x + \phi_x) e^{-\alpha_y^b y} e^{-jk_z z} \\ &= \frac{j\omega\mu_0}{k_b^2 - k_z^2} \left[\frac{\alpha_y^b k_z}{\omega\mu_0} \mathcal{E}_z^b + k_x^i \mathcal{H}_z^b \right] \cos(k_x^i x + \phi_x) e^{-\alpha_y^b y} e^{-jk_z z}, \end{aligned} \quad (\text{A.25})$$

$$\begin{aligned} H_x^b &= -\frac{1}{h^2} \left[\gamma k_x^i \mathcal{H}_z^b + j\omega\epsilon_b \alpha_y^b \mathcal{E}_z^b \right] \cos(k_x^i x + \phi_x) e^{-\alpha_y^b y} e^{-jk_z z} \\ &= -\frac{j\omega\epsilon_b}{k_b^2 - k_z^2} \left[\alpha_y^b \mathcal{E}_z^b + \frac{k_x^i k_z}{\omega\epsilon_b} \mathcal{H}_z^b \right] \cos(k_x^i x + \phi_x) e^{-\alpha_y^b y} e^{-jk_z z}, \end{aligned} \quad (\text{A.26})$$

$$\begin{aligned} H_y^b &= -\frac{1}{h^2} \left[-\gamma \alpha_y^b \mathcal{H}_z^b - j\omega\epsilon_b k_x^i \mathcal{E}_z^b \right] \sin(k_x^i x + \phi_x) e^{-\alpha_y^b y} e^{-jk_z z} \\ &= \frac{j\omega\epsilon_b}{k_b^2 - k_z^2} \left[k_x^i \mathcal{E}_z^b + \frac{k_z \alpha_y^b}{\omega\epsilon_b} \mathcal{H}_z^b \right] \sin(k_x^i x + \phi_x) e^{-\alpha_y^b y} e^{-jk_z z}. \end{aligned} \quad (\text{A.27})$$

A.2 Imposing Boundary Conditions

At the interface of two dielectric media the following applies: $E_{t,1} = E_{t,2}$, $D_{n,1} = D_{n,2}$, $H_{t,1} = H_{t,2}$ and $B_{n,1} = B_{n,2}$. The t and n notations denote tangential and normal components, respectively.

The first step of solving the wavenumber components in the inner region, k_x^i and k_y^i , is matching the fields across the boundaries. In [27] fields are matched along the direction of propagation, z , using $E_{t,1} = E_{t,2}$ and $H_{t,1} = H_{t,2}$ at the $y = \pm b$ and $x = \pm a$ boundaries.

Thus, at $y = \pm b$

$$\begin{aligned} E_z^i = E_z^b &\implies \mathcal{E}_z^i \cos(k_x^i x + \phi_x) \cos(k_y^i b + \phi_y) e^{-jk_z z} = \mathcal{E}_z^b \cos(k_x^i x + \phi_x) e^{-\alpha_y^b b} e^{-jk_z z} \\ &\implies \mathcal{E}_z^b = \mathcal{E}_z^i \cos(k_y^i b + \phi_y) e^{\alpha_y^b b}, \end{aligned} \quad (\text{A.28})$$

$$\begin{aligned} H_z^i = H_z^b &\implies \mathcal{H}_z^i \sin(k_x^i x + \phi_x) \sin(k_y^i b + \phi_y) e^{-jk_z z} = \mathcal{H}_z^b \sin(k_x^i x + \phi_x) e^{-\alpha_y^b b} e^{-jk_z z} \\ &\implies \mathcal{H}_z^b = \mathcal{H}_z^i \sin(k_y^i b + \phi_y) e^{\alpha_y^b b}. \end{aligned} \quad (\text{A.29})$$

Now, by matching the y components of the magnetic field ($H_{n,1} = H_{n,2} \Rightarrow H_{y,i} = H_{y,b}$) and using

Equations (A.28) and (A.29) to replace $\mathcal{E}_z^b$ and $\mathcal{H}_z^b$ the following can be written:

$$\begin{aligned} \frac{j\omega\epsilon_0}{k^2 - k_z^2} \left[k_x^i \mathcal{E}_z^i - \frac{k_z k_y^i}{\omega\epsilon} \mathcal{H}_z^i \right] \sin(k_x^i x + \phi_x) \cos(k_y^i b + \phi_y) e^{-jk_z z} = \\ \frac{j\omega\epsilon_b}{k_b^2 - k_z^2} \left[k_x^i \mathcal{E}_z^i \cos(k_y^i b + \phi_y) + \frac{k_z \alpha_y^b}{\omega\epsilon_b} \mathcal{H}_z^i \sin(k_y^i b + \phi_y) \right] e^{\alpha_y^b b} \sin(k_x^i x + \phi_x) e^{-\alpha_y^b b} e^{-jk_z z} \end{aligned} \quad (\text{A.30})$$

Letting $\Delta k_i^2 = k^2 - k_z^2$ and $\Delta k_b^2 = k_b^2 - k_z^2$, and dividing by $j\omega$

$$\mathcal{E}_z^i \left[\frac{\epsilon_0}{\Delta k_i^2} k_x^i \cos(k_y^i b + \phi_y) - \frac{\epsilon_b}{\Delta k_b^2} k_x^i \cos(k_y^i b + \phi_y) \right] + \mathcal{H}_z^i \left[-\frac{k_z k_y^i}{\omega \Delta k_i^2} \cos(k_y^i b + \phi_y) - \frac{k_z \alpha_y^b}{\omega \Delta k_b^2} \sin(k_y^i b + \phi_y) \right] = 0. \quad (\text{A.31})$$

Then by dividing by $\cos(k_y^i b + \phi_y)$ and rearranging Equation (A.31)

$$\mathcal{E}_z^i \left[\frac{k_x^i \epsilon_b}{\Delta k_b^2} - \frac{k_x^i \epsilon}{\Delta k_i^2} \right] + \mathcal{H}_z^i \frac{k_z}{\omega} \left[\frac{k_y^i}{\Delta k_i^2} + \frac{\alpha_y^b}{\Delta k_b^2} \tan(k_y^i b + \phi_y) \right] = 0. \quad (\text{A.32})$$

A similar approach will now be used in matching the x components of the magnetic field ($H_{t,1} = H_{t,2} \Rightarrow H_{x,i} = H_{x,b}$), which results in

$$\begin{aligned} -\frac{j\omega\epsilon_0}{k^2 - k_z^2} \left[k_y^i \mathcal{E}_z^i + \frac{k_z k_x^i}{\omega\epsilon_0} \mathcal{H}_z^i \right] \sin(k_y^i b + \phi_y) = -\frac{j\omega\epsilon_b}{k_b^2 - k_z^2} \left[k_y^i \mathcal{E}_z^i \cos(k_y^i b + \phi_y) + \frac{k_z k_x^i}{\omega\epsilon_b} \mathcal{H}_z^i \sin(k_y^i b + \phi_y) \right] \\ \Rightarrow \mathcal{E}_z^i \left[\frac{\epsilon_0 k_y^i}{\Delta k_i^2} \sin(k_y^i b + \phi_y) - \frac{\epsilon_b \alpha_y^b}{\Delta k_b^2} \cos(k_y^i b + \phi_y) \right] + \mathcal{H}_z^i \frac{k_z}{\omega} \left[\frac{k_x^i}{\Delta k_i^2} \sin(k_y^i b + \phi_y) - \frac{k_x^i}{\Delta k_b^2} \sin(k_y^i b + \phi_y) \right] = 0 \\ \Rightarrow \mathcal{E}_z^i \left[\frac{\epsilon_0 k_y^i}{\Delta k_i^2} - \frac{\epsilon_b \alpha_y^b}{\Delta k_b^2} \cot(k_y^i b + \phi_y) \right] + \mathcal{H}_z^i \frac{k_z k_x^i}{\omega} \left[\frac{1}{\Delta k_i^2} - \frac{1}{\Delta k_b^2} \right] = 0. \end{aligned} \quad (\text{A.33})$$

Equations (A.32) and (A.33) can be cast into a matrix formulation as

$$\begin{bmatrix} \frac{\epsilon_0 k_y^i}{\Delta k_i^2} - \frac{\epsilon_b \alpha_y^b}{\Delta k_b^2} \cot(k_y^i b + \phi_y) & \frac{k_z k_x^i}{\omega} \left[\frac{1}{\Delta k_i^2} - \frac{1}{\Delta k_b^2} \right] \\ \frac{k_x^i \epsilon_b}{\Delta k_b^2} - \frac{k_x^i \epsilon}{\Delta k_i^2} & \frac{k_z}{\omega} \left[\frac{k_y^i}{\Delta k_i^2} + \frac{\alpha_y^b}{\Delta k_b^2} \tan(k_y^i b + \phi_y) \right] \end{bmatrix} \times \begin{bmatrix} \mathcal{E}_z^i \\ \mathcal{H}_z^i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (\text{A.34})$$

A matrix with the same form can be found by applying the boundary conditions at $x = \pm a$. First expressions for E_z^a and H_z^a are arranged to be

$$\begin{aligned} E_z^i = E_z^a \Rightarrow \mathcal{E}_z^i \cos(k_x^i a + \phi_x) \cos(k_y^i y + \phi_y) e^{-jk_z z} = \mathcal{E}_z^a \cos(k_y^i y + \phi_y) e^{-\alpha_x^a a} e^{-jk_z z} \\ \Rightarrow \mathcal{E}_z^a = \mathcal{E}_z^i \cos(k_x^i a + \phi_x) e^{\alpha_x^a a}, \end{aligned} \quad (\text{A.35})$$

$$\begin{aligned} H_z^i = H_z^a \Rightarrow \mathcal{H}_z^i \sin(k_x^i a + \phi_x) \sin(k_y^i y + \phi_y) e^{-jk_z z} = \mathcal{H}_z^a \sin(k_y^i y + \phi_y) e^{-\alpha_x^a a} e^{-jk_z z} \\ \Rightarrow \mathcal{H}_z^a = \mathcal{H}_z^i \sin(k_x^i a + \phi_x) e^{\alpha_x^a a}. \end{aligned} \quad (\text{A.36})$$

Then by matching the y components of the magnetic field ($H_{t,1} = H_{t,2} \Rightarrow H_{y,i} = H_{y,a}$)

$$\begin{aligned}
\frac{j\omega\epsilon_0}{\Delta k_i^2} \left[k_x^i \mathcal{E}_z^i - \frac{k_z k_y^i}{\omega\epsilon_0} \mathcal{H}_z^i \right] \sin(k_x^i a + \phi_x) &= \frac{j\omega\epsilon_a}{\Delta k_a^2} \left[\alpha_x^a \mathcal{E}_z^i \cos(k_x^i a + \phi_x) - \frac{k_z k_y^i}{\omega\epsilon_a} \mathcal{H}_z^i \sin(k_x^i a + \phi_x) \right] \\
\Rightarrow \mathcal{E}_z^i \left[\frac{k_x^i \epsilon_0}{\Delta k_i^2} \sin(k_x^i a + \phi_x) - \frac{\alpha_x^a \epsilon_a}{\Delta k_a^2} \cos(k_x^i a + \phi_x) \right] &+ \mathcal{H}_z^i \frac{k_z}{\omega} \left[-\frac{k_y^i}{\Delta k_i^2} \sin(k_x^i a + \phi_x) + \frac{k_y^i}{\Delta k_a^2} \sin(k_x^i a + \phi_x) \right] = 0 \\
\Rightarrow \mathcal{E}_z^i \left[-\frac{k_x^i \epsilon_0}{\Delta k_i^2} + \frac{\alpha_x^a \epsilon_a}{\Delta k_a^2} \cot(k_x^i a + \phi_x) \right] &+ \mathcal{H}_z^i \frac{k_z k_y^i}{\omega} \left[-\frac{1}{\Delta k_a^2} + \frac{1}{\Delta k_i^2} \right] = 0.
\end{aligned} \tag{A.37}$$

Then by matching the x components of the magnetic field ($H_{n,1} = H_{n,2} \Rightarrow H_{x,i} = H_{x,a}$)

$$\begin{aligned}
-\frac{j\omega\epsilon_0}{\Delta k_i^2} \left[k_y^i \mathcal{E}_z^i + \frac{k_z k_x^i}{\omega\epsilon_0} \mathcal{H}_z^i \right] \cos(k_x^i a + \phi_x) &= \frac{j\omega\epsilon_a}{\Delta k_a^2} \left[-k_y^i \mathcal{E}_z^i \cos(k_x^i a + \phi_x) + \frac{k_z \alpha_x^a}{\omega\epsilon_a} \mathcal{H}_z^i \sin(k_x^i a + \phi_x) \right] \\
\Rightarrow \mathcal{E}_z^i \left[-\frac{k_y^i \epsilon_0}{\Delta k_i^2} + \frac{k_y^i \epsilon_a}{\Delta k_a^2} \right] \cos(k_x^i a + \phi_x) &+ \mathcal{H}_z^i \frac{k_z}{\omega} \left[-\frac{k_x^i}{\Delta k_i^2} \cos(k_x^i a + \phi_x) - \frac{\alpha_x^a}{\Delta k_a^2} \sin(k_x^i a + \phi_x) \right] = 0 \\
\Rightarrow \mathcal{E}_z^i \left[\frac{k_y^i \epsilon_0}{\Delta k_i^2} - \frac{k_y^i \epsilon_a}{\Delta k_a^2} \right] &+ \mathcal{H}_z^i \frac{k_z}{\omega} \left[\frac{k_x^i}{\Delta k_i^2} + \frac{\alpha_x^a}{\Delta k_a^2} \tan(k_x^i a + \phi_x) \right] = 0.
\end{aligned} \tag{A.38}$$

Similar to before, Equations (A.37) and (A.38) can be cast into a matrix form as

$$\begin{bmatrix} -\frac{\epsilon_a k_y^i}{\Delta k_a^2} + \frac{\epsilon_0 k_y^i}{\Delta k_i^2} & \frac{k_z}{\omega} \left[\frac{\alpha_x^a}{\Delta k_a^2} \tan(k_x^i a + \phi_x) + \frac{k_x^i}{\Delta k_i^2} \right] \\ \frac{\alpha_x^a \epsilon_a}{\Delta k_a^2} \cot(k_x^i a + \phi_x) - \frac{k_x^i \epsilon_0}{\Delta k_i^2} & \frac{k_z k_y^i}{\omega} \left[-\frac{1}{\Delta k_a^2} + \frac{1}{\Delta k_i^2} \right] \end{bmatrix} \times \begin{bmatrix} \mathcal{E}_z^i \\ \mathcal{H}_z^i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{A.39}$$

The matrices in (A.34) and (A.39) form the eigenvalue relations. The solution of these matrices leads to the eigenvalue solution, which is shown in the following section.¹

A.3 Eigenvalue Solutions

A.3.1 Solving Boundary Condition Matrices

To find the solutions of the wavenumber in the inner region of the waveguide k , k_x^i and k_y^i are isolated in the eigenvalue relations shown the matrices (A.34) and (A.39). By using the fact that $k = \sqrt{\omega^2 \mu_0 \epsilon}$, k_z can be found using $k_z = \sqrt{k^2 - (k_x^i)^2 - (k_y^i)^2}$. The expression for k is then substituted back into the field components in the inner region to find the field distribution of the hollow dielectric rectangular waveguide.

Recall that in matrix relations the system has a unique solution if and only if the determinant of the matrix is nonzero; that is: $ad - bc \neq 0$ or $ad = bc$ where the matrix is defined as

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{A.40}$$

¹The matrix equations in (A.34) and (A.39) exactly match those in [27], except normalized permittivities are used.

The determinant of matrix (A.34) is

$$\begin{aligned} \frac{k_z k_y^{i,2} \epsilon_0}{\omega \Delta k_i^4} + \frac{k_z k_y^i \alpha_y^b \epsilon_0}{\omega \Delta k_i^2 \Delta k_b^2} \tan(k_y^i b + \phi_y) - \frac{k_z \alpha_y^{b,2} \epsilon_b}{\omega \Delta k_b^4} - \frac{k_z \alpha_y^b k_y^i \epsilon_b}{\omega \Delta k_i^2 \Delta k_b^2} \cot(k_y^i b + \phi_y) = \\ \frac{k_z k_x^{i,2} \epsilon_b}{\omega \Delta k_i^2 \Delta k_b^2} - \frac{k_z k_x^{i,2} \epsilon_0}{\omega \Delta k_i^4} - \frac{k_z k_x^{i,2} \epsilon_b}{\omega \Delta k_b^4} + \frac{k_z k_x^{i,2} \epsilon_0}{\omega \Delta k_i^2 \Delta k_b^2}. \end{aligned} \quad (\text{A.41})$$

Now the substitution $X = \tan(k_y^i b + \phi_y)$ will be used and $X\omega/k_z$ will be multiplied through Equation (A.41). Rearranging such that like terms are together

$$\begin{aligned} X^2 \left[\frac{k_y^i \alpha_y^b \epsilon_0}{\Delta k_b^2 \Delta k_i^2} \right] + X \left[\frac{\epsilon_0}{\Delta k_i^4} (k_y^{i,2} + k_x^{i,2}) - \frac{k_x^{i,2}}{\Delta k_i^2 \Delta k_b^2} (\epsilon_b + \epsilon_0) + \frac{\epsilon_b}{\Delta k_b^4} (k_x^{i,2} - \alpha_y^{b,2}) \right] - \frac{\alpha_y^b k_y^i \epsilon_b}{\Delta k_i^2 \Delta k_b^2} = 0 \\ \implies X^2 [k_y^i \alpha_y^b] + X \left[\frac{\Delta k_b^2}{\Delta k_i^2} (k_y^{i,2} + k_x^{i,2}) - k_x^{i,2} (\bar{\epsilon}_b + 1) + \frac{\Delta k_i^2}{\Delta k_b^2} \bar{\epsilon}_b (k_x^{i,2} - \alpha_y^{b,2}) \right] - \alpha_y^b k_y^i \bar{\epsilon}_b = 0. \end{aligned} \quad (\text{A.42})$$

In the last line of Equation (A.42) $\bar{\epsilon}_b = \epsilon_b/\epsilon_0$ and $\Delta k_b^2 \Delta k_i^2$ was multiplied through. Equation (A.42) is a quadratic equation and will be solved for X . Prior to that the ‘ X ’ terms must be simplified. To do this all $k_x^{i,2}$, Δk_i^2 and Δk_b^2 occurrences will be substituted with

$$\begin{aligned} \Delta k_i^2 &= k^2 - k_z^2, \\ k_x^{i,2} &= k^2 - k_y^{i,2} - k_z^2, \\ \Delta k_b^2 &= k_b^2 - k_z^2 = (k_x^{i,2} - \alpha_y^{b,2} + k_z^2) - k_z^2 = k^2 - k_y^{i,2} - k_z^2 - \alpha_y^{b,2}. \end{aligned} \quad (\text{A.43})$$

Using the substitutions in the ‘ X ’ term

$$\begin{aligned} \frac{\Delta k_b^2}{\Delta k_i^2} (k_y^{i,2} + k_x^{i,2}) - k_x^{i,2} (\bar{\epsilon}_b + 1) + \frac{\Delta k_i^2}{\Delta k_b^2} \bar{\epsilon}_b (k_x^{i,2} - \alpha_y^{b,2}) \\ = \frac{k^2 - \alpha_y^{b,2} - k_y^{i,2} - k_z^2}{k^2 - k_z^2} (k_y^{i,2} + k^2 - k_y^{i,2} - k_z^2) - (k^2 - k_y^{i,2} - k_z^2) (\bar{\epsilon}_b + 1) + \\ \frac{k^2 - k_z^2}{k^2 - \alpha_y^{b,2} - k_y^{i,2} - k_z^2} (k^2 - k_y^{i,2} - k_y^{b,2} - k_z^2) \\ = k^2 - \alpha_y^{b,2} - k_y^{i,2} - k_z^2 - k^2 + k_y^{i,2} + k_z^2 + \bar{\epsilon}_b (k_y^{i,2} + k_z^2 - k^2 + k^2 - k_z^2) \\ = -\alpha_y^{b,2} + \bar{\epsilon}_b k_y^{i,2}. \end{aligned} \quad (\text{A.44})$$

Solving for $X = \tan(k_y^i b + \phi_y)$ using the quadratic equation results in

$$\begin{aligned} \tan(k_y^i b + \phi_y) &= \frac{-(-\alpha_y^{b,2} + \bar{\epsilon}_b k_y^{i,2}) \pm \sqrt{\alpha_y^{b,4} - 2\bar{\epsilon}_b \alpha_y^{b,2} k_y^{i,2} + \bar{\epsilon}_b^2 k_y^{i,4} + 4\bar{\epsilon}_b \alpha_y^{b,2} k_y^{i,2}}}{2k_y^b k_y^i} \\ &= \frac{\alpha_y^{b,2} - \bar{\epsilon}_b k_y^{i,2} \pm (\alpha_y^{b,2} + \bar{\epsilon}_b k_y^{i,2})}{2\alpha_y^b k_y^i} \\ &= \begin{cases} \alpha_y^b / k_y^i \\ -k_y^i \bar{\epsilon}_b / \alpha_y^b \end{cases}. \end{aligned} \quad (\text{A.45})$$

In the following section $\tan(k_y^i b + \phi_y)$ will be simplified to obtain an expression for k_y^i . First an

expression for $\tan(k_x^i a + \phi_a)$ will be obtained following an analysis very similar to that just developed. The determinant of matrix (A.39) is

$$\begin{aligned} & \frac{k_z k_y^{i,2} \epsilon_a}{\omega \Delta k_a^4} - \frac{k_z k_y^{i,2} \epsilon_a}{\omega \Delta k_i^2 \Delta k_a^2} - \frac{k_z k_y^{i,2} \epsilon_0}{\omega \Delta k_i^2 \Delta k_a^2} + \frac{k_z k_y^{i,2} \epsilon_0}{\omega \Delta k_i^4} = \\ & \frac{k_z \alpha_x^{a,2} \epsilon_a}{\omega \Delta k_a^4} - \frac{k_z \alpha_x^a k_x^i \epsilon_0}{\omega \Delta k_i^4 \Delta k_a^4} \tan(k_x^i a + \phi_a) + \frac{k_z k_x^i \alpha_x^a \epsilon_a}{\omega \Delta k_i^2 \Delta k_a^2} \cot(k_x^i a + \phi_a) - \frac{k_z k_x^{i,2} \epsilon_0}{\omega \Delta k_i^4}. \end{aligned} \quad (\text{A.46})$$

$X = \tan(k_x^i a + \phi_x)$ will be used and $X\omega/k_z$ will be multiplied through Equation (A.46). Rearranging such that like terms are together

$$\begin{aligned} & X^2 \left[\frac{k_x^i \alpha_x^a \epsilon_0}{\Delta k_a^2 \Delta k_i^2} \right] + X \left[\frac{\epsilon_0}{\Delta k_i^4} (k_y^{i,2} + k_x^{i,2}) - \frac{k_y^{i,2}}{\Delta k_i^2 \Delta k_a^2} (\epsilon_a + \epsilon_0) + \frac{\epsilon_a}{\Delta k_a^4} (k_y^{i,2} - \alpha_x^{a,2}) \right] - \frac{\alpha_x^a k_x^i \epsilon_a}{\Delta k_i^2 \Delta k_a^2} = 0 \\ \implies & X^2 [k_x^i \alpha_x^a] + X \left[\frac{\Delta k_a^2}{\Delta k_i^2} (k_y^{i,2} + k_x^{i,2}) - k_y^{i,2} (\bar{\epsilon}_a + 1) + \frac{\Delta k_i^2}{\Delta k_a^2} \bar{\epsilon}_a (k_y^{i,2} - \alpha_x^{a,2}) \right] - k_x^i k_x^b \bar{\epsilon}_a = 0. \end{aligned} \quad (\text{A.47})$$

In the last line of Equation (A.47) $\bar{\epsilon}_a = \epsilon_a/\epsilon_0$ and $\Delta k_a^2 \Delta k_i^2$ was multiplied through. To simplify the ‘X’ terms all $k_y^{i,2}$, Δk_i^2 and Δk_a^2 occurrences will be substituted with

$$\begin{aligned} \Delta k_i^2 &= k^2 - k_z^2, \\ k_y^{i,2} &= k^2 - k_x^{i,2} - k_z^2, \\ \Delta k_a^2 &= k_a^2 - k_z^2 = (-\alpha_x^{a,2} + k_y^{i,2} + k_z^2) - k_z^2 = -\alpha_x^{a,2} + k^2 - k_x^{i,2} - k_z^2. \end{aligned} \quad (\text{A.48})$$

Using the substitutions in the ‘X’ term

$$\begin{aligned} & \frac{\Delta k_a^2}{\Delta k_i^2} (k_y^{i,2} + k_x^{i,2}) - k_y^{i,2} (\bar{\epsilon}_a + 1) + \frac{\Delta k_i^2}{\Delta k_a^2} \bar{\epsilon}_a (k_y^{i,2} - \alpha_x^{a,2}) \\ &= \frac{k^2 - \alpha_x^{a,2} - k_x^{i,2} - k_z^2}{k^2 - k_z^2} (k^2 - k_x^{i,2} - k_z^2 + k_x^{i,2}) - (k^2 - k_x^{i,2} - k_z^2) (\bar{\epsilon}_a + 1) + \\ & \quad \frac{k^2 - k_z^2}{k^2 - \alpha_x^{a,2} - k_x^{i,2} - k_z^2} (k^2 - k_x^{i,2} - \alpha_x^{a,2} - k_z^2) \\ &= k^2 - \alpha_x^{a,2} - k_x^{i,2} - k_z^2 - k^2 + k_x^{i,2} + k_z^2 + \bar{\epsilon}_a (-k^2 + k_x^{i,2} + k_z^2 + k^2 - k_z^2) \\ &= -\alpha_x^{a,2} + \bar{\epsilon}_a k_x^{i,2}. \end{aligned} \quad (\text{A.49})$$

Solving for $X = \tan(k_x^i a + \phi_a)$

$$\begin{aligned} \tan(k_x^i a + \phi_x) &= \frac{-(-\alpha_x^{a,2} + \bar{\epsilon}_a k_x^{i,2}) \pm \sqrt{\alpha_x^{a,4} - 2\bar{\epsilon}_a \alpha_x^{a,2} k_x^{i,2} + \bar{\epsilon}_a^2 k_x^{i,4} + 4\bar{\epsilon}_a \alpha_x^{a,2} k_x^{i,2}}}{2\alpha_x^a k_x^i} \\ &= \frac{\alpha_x^{a,2} - \bar{\epsilon}_a k_x^{i,2} \pm (\alpha_x^{a,2} + \bar{\epsilon}_a k_x^{i,2})}{2\alpha_x^a k_x^i} \\ &= \begin{cases} \alpha_x^a / k_x^i \\ -k_x^i \bar{\epsilon}_a / \alpha_x^a \end{cases}. \end{aligned} \quad (\text{A.50})$$

Now expressions for $\tan(k_x^i a + \phi_a)$ and $\tan(k_y^i b + \phi_b)$ are known and k_x^i and k_y^i can be found.

A.3.2 Obtaining the Eigenvalue Expressions

To obtain expressions for k_x^i or k_y^i using Equations (A.45) or (A.50) one of the respective assumptions must be made

$$\begin{aligned} \text{for } k_x^i : |\bar{\epsilon}_a \frac{k_x^i}{\alpha_x^a}| &\ll 1 \text{ or } |\frac{\alpha_x^a}{k_x^i}| \gg 1, \\ \text{for } k_y^i : |\bar{\epsilon}_b \frac{k_y^i}{\alpha_y^b}| &\ll 1 \text{ or } |\frac{\alpha_y^b}{k_y^i}| \gg 1. \end{aligned} \quad (\text{A.51})$$

Essentially these assumptions must be consistent with the second assumptions in Section 2.1.1. Before solving Equations (A.45) and (A.50) they will be modified. For example, α_x^a will be eliminated by using

$$\begin{aligned} \text{Since } k_a^2 &= k^2 \bar{\epsilon}_a = -\alpha_x^{a,2} + k_y^{i,2} + k_z^2 \implies k_y^{i,2} = k^2 \bar{\epsilon}_a + \alpha_x^{a,2} - k_z^2, \\ \text{then } k_x^{i,2} &= k^2 - k_y^{i,2} - k_z^2 = k^2(1 - \bar{\epsilon}_a) - \alpha_x^{a,2}, \\ \text{and thus } -\alpha_x^{a,2} &= k_x^{i,2} + k^2(\bar{\epsilon}_a - 1). \end{aligned} \quad (\text{A.52})$$

Similarly, $-\alpha_y^{b,2}$ can be expressed as $-\alpha_y^{b,2} = k_y^{i,2} + k^2(\bar{\epsilon}_b - 1)$. Reformulating Equation (A.45) as

$$\begin{aligned} \tan(k_y^i b + \phi_y) &= \begin{cases} \frac{\alpha_y^b}{k_y^i} = \frac{j\sqrt{k_y^{i,2} + k^2(\bar{\epsilon}_b - 1)}}{k_y^i} = j\sqrt{1 + k^2/k_y^{i,2}(\bar{\epsilon}_b - 1)} \\ -\frac{k_y^i \bar{\epsilon}_b}{\alpha_y^b} = -\frac{k_y^i \bar{\epsilon}_b}{j\sqrt{k_y^{i,2} + k^2(\bar{\epsilon}_b - 1)}} = -\frac{\bar{\epsilon}_b}{j\sqrt{1 + k^2/k_y^{i,2}(\bar{\epsilon}_b - 1)}} \end{cases} \\ &\approx \begin{cases} j\frac{k\sqrt{\bar{\epsilon}_b - 1}}{k_y^i} \gg 1 \\ j\frac{\bar{\epsilon}_b k_y^i}{k\sqrt{\bar{\epsilon}_b - 1}} \ll 1 \end{cases}. \end{aligned} \quad (\text{A.53})$$

Reformulating Equation (A.54) as

$$\begin{aligned} \tan(k_x^i a + \phi_x) &= \begin{cases} \frac{\alpha_x^a}{k_x^i} = \frac{j\sqrt{k_x^{i,2} + k^2(\bar{\epsilon}_a - 1)}}{k_x^i} = j\sqrt{1 + k^2/k_x^{i,2}(\bar{\epsilon}_a - 1)} \\ -\frac{k_x^i \bar{\epsilon}_a}{\alpha_x^a} = -\frac{k_x^i \bar{\epsilon}_a}{j\sqrt{k_x^{i,2} + k^2(\bar{\epsilon}_a - 1)}} = -\frac{\bar{\epsilon}_a}{j\sqrt{1 + k^2/k_x^{i,2}(\bar{\epsilon}_a - 1)}} \end{cases} \\ &\approx \begin{cases} j\frac{k\sqrt{\bar{\epsilon}_a - 1}}{k_x^i} \gg 1 \\ j\frac{\bar{\epsilon}_a k_x^i}{k\sqrt{\bar{\epsilon}_a - 1}} \ll 1 \end{cases}. \end{aligned} \quad (\text{A.54})$$

The results of Equations (A.53) and (A.54) are consistent with assumptions 1 and 2 in Section 2.1.1. Equations (A.53) and (A.54) are transcendental equations and to solve them the substitutions

$$\begin{aligned} k_y^i b + \phi_y &= l\pi/2 + \delta_b, \quad \delta_b \ll 1, \\ k_x^i a + \phi_x &= l\pi/2 + \delta_a, \quad \delta_a \ll 1, \end{aligned} \quad (\text{A.55})$$

are used, where l is an integer. Now, if $k_y^i = (l\pi/2 + \delta_b)/b$ (since $\phi_y = \pi/2$ or 0), then

$$\tan\left(\left(\frac{l\pi/2 + \delta_b}{b}\right)b + \phi_y\right) \rightarrow \begin{cases} \infty \text{ as } \delta_b \rightarrow 0, \text{ if } l \text{ odd and } \phi_y = 0; \text{ or } l \text{ even and } \phi_y = \pi/2 \\ 0 \text{ as } \delta_b \rightarrow 0, \text{ if } l \text{ odd and } \phi_y = \pi/2; \text{ or } l \text{ even and } \phi_y = 0 \end{cases}. \quad (\text{A.56})$$

Solving for k_y^i in Equation (A.53) using this results in the following cases.

Case I: $\tan(k_y^i b + \phi_y) \rightarrow 0$ which corresponds to $\tan(k_y^i b + \phi_y) \approx j\frac{\bar{\epsilon}_b k_y^i}{k\sqrt{\bar{\epsilon}_b - 1}} \ll 1$. This case applies when

l is odd and $\phi_y = 0$, or when l is even and $\phi_y = \pi/2$. Thus

$$\tan(k_y^i b + \phi_y) = \tan(l\pi/2 + \delta_b + \phi_y) \approx \delta_b, \quad (\text{A.57})$$

which applies because δ_b is very small. Additionally

$$l\pi/2 + \phi_y = m\pi, \quad (\text{A.58})$$

where $m = 0, 1, 2, 3$. Thus k_y^i can be expressed as

$$k_y^i b + \phi_y = m\pi + \delta_b \Rightarrow k_y^i = \frac{m\pi + \delta_b - \phi_y}{b}. \quad (\text{A.59})$$

Note that $+\phi_y$ can be equivalently used on the left hand side of Equation (A.59) and it will be in following equations. An expression for δ_b is now arranged using

$$\tan(k_y^i b + \phi_y) = j \frac{\bar{\epsilon}_b k_y^i}{k \sqrt{\bar{\epsilon}_b} - 1} \approx \delta_b. \quad (\text{A.60})$$

By substituting Equation (A.59) into (A.60) to eliminate k_y^i δ_b becomes

$$\begin{aligned} \delta_b &= j \frac{\bar{\epsilon}_b}{k \sqrt{\bar{\epsilon}_b} - 1} \left(\frac{m\pi + \delta_b + \phi_y}{b} \right) \\ \Rightarrow \delta_b &= j \frac{\bar{\epsilon}_b}{k \sqrt{\bar{\epsilon}_b} - 1} \left(\frac{m\pi + \phi_y}{b} \right) \left(\frac{kb \sqrt{\bar{\epsilon}_b} - 1}{kb \sqrt{\bar{\epsilon}_b} - 1 - j \bar{\epsilon}_b} \right) \\ &= j \bar{\epsilon}_b (m\pi + \phi_y) \left(\frac{1}{kb \sqrt{\bar{\epsilon}_b} - 1 - j \bar{\epsilon}_b} \right). \end{aligned} \quad (\text{A.61})$$

Finally, using Equations (A.59) and (A.61) k_y^i can be written as

$$k_y^i = \frac{m\pi + \phi_y}{b} + j \frac{\bar{\epsilon}_b (m\pi + \phi_y)}{b} \left(\frac{1}{kb \sqrt{\bar{\epsilon}_b} - 1 - j \bar{\epsilon}_b} \right) \approx \frac{m\pi + \phi_y}{b} \left[1 + j \frac{\bar{\epsilon}_b}{kb \sqrt{\bar{\epsilon}_b} - 1} \right]. \quad (\text{A.62})$$

Note that the $j \bar{\epsilon}_b$ was dropped in the denominator of the last term because $kb \sqrt{\bar{\epsilon}_b} - 1 \gg \bar{\epsilon}_b$, which is in accordance with y -polarized modes in the second assumption of Section 2.1.1. Thus

$$k_y^i = \begin{cases} \frac{m\pi}{b} \left[1 + j \frac{\bar{\epsilon}_b}{kb \sqrt{\bar{\epsilon}_b} - 1} \right], & \text{for } \phi_y = 0 \\ \frac{m\pi + \pi/2}{b} \left[1 + j \frac{\bar{\epsilon}_b}{kb \sqrt{\bar{\epsilon}_b} - 1} \right], & \text{for } \phi_y = \pi/2 \end{cases}. \quad (\text{A.63})$$

Case II: $\tan(k_y^i b + \phi_y) \rightarrow \infty$ which corresponds to $\tan(k_y^i b + \phi_y) \approx j \frac{k \sqrt{\bar{\epsilon}_b} - 1}{k_y^i} \gg 1$. This case applies when l is odd and $\phi_y = \pi/2$, or when l is even and $\phi_y = 0$. Thus

$$\tan(k_y^i b + \phi_y) = \tan(l\pi/2 + \delta_b + \phi_y) \approx \frac{1}{\delta_b}, \quad (\text{A.64})$$

which applies because δ_b is very small, resulting in $1/\delta_b$ very large. Therefore

$$l\pi/2 + \phi_y = (2m + 1)\pi/2, \quad (\text{A.65})$$

where $m = 0, 1, 2, 3$. Thus k_y^i can be expressed as²:

$$k_y^i b + \phi_y = (2m + 1)\pi/2 - \delta_b \Rightarrow k_y^i = \frac{(2m + 1)\pi/2 - \delta_b + \phi_y}{b}. \quad (\text{A.66})$$

An expression for δ_b is now arranged using:

$$\tan(k_y^i b + \phi_y) = \frac{jk\sqrt{\epsilon_b - 1}}{k_y^i} \approx \frac{1}{\delta_b}. \quad (\text{A.67})$$

By substituting Equation (A.66) into (A.67) to eliminate k_y^i $1/\delta_b$ becomes

$$\begin{aligned} \frac{1}{\delta_b} &= j \frac{kb\sqrt{\epsilon_b - 1}}{(2m + 1)\pi/2 - \delta_b + \phi_y} \\ \Rightarrow \delta_b &= \frac{(2m + 1)\pi/2 + \phi_y}{kb\sqrt{\epsilon_b - 1}} \frac{kb\sqrt{\epsilon_b - 1}}{jkb\sqrt{\epsilon_b - 1} + 1} \\ &= \frac{(2m + 1)\pi/2 + \phi_y}{jkb\sqrt{\epsilon_b - 1} + 1}. \end{aligned} \quad (\text{A.68})$$

Finally, using Equations (A.65) and (A.67) k_y^i can be written as

$$k_y^i = \frac{(2m + 1)\pi/2 + \phi_y}{b} - \frac{(2m + 1)\pi/2 + \phi_y}{b} \left(\frac{1}{jkb\sqrt{\epsilon_b - 1} + 1} \right) \approx \frac{(2m + 1)\pi/2 + \phi_y}{b} \left[1 + \frac{j}{kb\sqrt{\epsilon_b - 1}} \right]. \quad (\text{A.69})$$

The extra 1 was dropped in the denominator of the last term because $kb\sqrt{\epsilon_b - 1} \gg 1$, which is in accordance with x -polarized modes in the second assumption of Section 2.1.1. Thus

$$k_y^i = \begin{cases} \frac{m\pi + \pi/2}{b} \left[1 + \frac{j}{kb\sqrt{\epsilon_b - 1}} \right], & \text{for } \phi_y = 0 \\ \frac{m\pi}{b} \left[1 + \frac{j}{kb\sqrt{\epsilon_b - 1}} \right], & \text{for } \phi_y = \pi/2 \end{cases}. \quad (\text{A.70})$$

This same analysis can be applied to find k_x^i and is not shown. To summarize the results they are listed in Table A.1. The results in Table A.1 match those given in [27]. However, an additional factor of 1/2 must be multiplied with all terms in order to find k_z , which is shown next and the resultant k_z expression also agrees with that in [27]. Moreover, in the only other source that provides k_x^i and k_y^i the 1/2 factor is present [24]. Thus, the additional factor of 1/2 is assumed correct and is used in following equations.

		For x-polarized modes	For y-polarized modes
k_x^i	$\phi_x = 0$	$\frac{m\pi}{a} \left[1 + j \frac{\epsilon_a}{ka\sqrt{\epsilon_a - 1}} \right]$	$\frac{m\pi + \pi/2}{a} \left[1 + \frac{j}{ka\sqrt{\epsilon_a - 1}} \right]$
	$\phi_x = \pi/2$	$\frac{m\pi + \pi/2}{a} \left[1 + j \frac{\epsilon_a}{ka\sqrt{\epsilon_a - 1}} \right]$	$\frac{m\pi}{a} \left[1 + \frac{j}{ka\sqrt{\epsilon_a - 1}} \right]$
k_y^i	$\phi_y = 0$	$\frac{n\pi + \pi/2}{b} \left[1 + \frac{j}{kb\sqrt{\epsilon_b - 1}} \right]$	$\frac{n\pi}{b} \left[1 + j \frac{\epsilon_b}{kb\sqrt{\epsilon_b - 1}} \right]$
	$\phi_y = \pi/2$	$\frac{n\pi}{b} \left[1 + \frac{j}{kb\sqrt{\epsilon_b - 1}} \right]$	$\frac{n\pi + \pi/2}{b} \left[1 + j \frac{\epsilon_b}{kb\sqrt{\epsilon_b - 1}} \right]$

Table A.1: Propagation constants.

²Note that $-\delta_b$ or δ_b can be used. Here $-\delta_b$ is used so results agree with those in [27].

A.4 Solutions of Fields using Eigenvectors

In this section the fields and propagation constants in the inner region are solved for the x -polarized and y -polarized modes.

In Section A.1.1 the fields for the inner region were provided. Before solving the fields using k_x^i and k_y^i the z -direction propagation constant, k_z is found. The $1/2$ factor mentioned previously is used and the substitution $k = 2\pi/\lambda$ is made in wavenumbers in Table A.1.

The propagation constant is solved for x -polarized modes with $\phi_x = 0$ and $\phi_y = 0$ and is

$$\begin{aligned}
k_z^2 &= k^2 - k_x^{i,2} - k_y^{i,2} \\
&= \left(\frac{2\pi}{\lambda}\right)^2 - \left(\frac{m\pi}{2a}\right)^2 \left(1 + j \frac{\bar{\epsilon}_a}{\sqrt{\bar{\epsilon}_a - 1}} \frac{\lambda}{2\pi a}\right)^2 - \left(\frac{n\pi}{2b}\right)^2 \left(1 + j \frac{1}{\sqrt{\bar{\epsilon}_b - 1}} \frac{\lambda}{2\pi b}\right)^2 \\
&= \left(\frac{2\pi}{\lambda}\right)^2 - \left(\frac{m\pi}{2a}\right)^2 \left(1 + 2j \frac{\bar{\epsilon}_a}{\sqrt{\bar{\epsilon}_a - 1}} \left(\frac{\lambda}{2\pi a}\right) - \frac{\bar{\epsilon}_a^2}{\bar{\epsilon}_a - 1} \left(\frac{\lambda}{2\pi a}\right)^2\right) - \left(\frac{n\pi}{2b}\right)^2 \left(1 + 2j \frac{1}{\sqrt{\bar{\epsilon}_b - 1}} \left(\frac{\lambda}{2\pi b}\right) - \frac{1}{\bar{\epsilon}_b - 1} \left(\frac{\lambda}{2\pi b}\right)^2\right) \\
&= \left(\frac{2\pi}{\lambda}\right)^2 \left(1 - \left(\frac{m\lambda}{4a}\right)^2 - \left(\frac{n\lambda}{4b}\right)^2\right) + \left(\frac{2\pi}{\lambda}\right)^2 \left(\left(\frac{m\lambda}{4a}\right)^2 \frac{\bar{\epsilon}_a^2}{\bar{\epsilon}_a - 1} \left(\frac{\lambda}{2\pi a}\right)^2 + \left(\frac{n\lambda}{4b}\right)^2 \frac{1}{\bar{\epsilon}_b - 1} \left(\frac{\lambda}{2\pi b}\right)^2\right) \dots \\
&\quad - j \left(\frac{2\pi}{\lambda}\right)^2 \left(\left(\frac{m\lambda}{4a}\right)^2 \frac{\bar{\epsilon}_a}{\sqrt{\bar{\epsilon}_a - 1}} \frac{\lambda}{2\pi a} + \left(\frac{n\lambda}{4b}\right)^2 \frac{1}{\sqrt{\bar{\epsilon}_b - 1}} \frac{\lambda}{2\pi b}\right). \tag{A.71}
\end{aligned}$$

Note that a very similar equation can be written for the other cases.

Recall that the notation $\gamma = jk_z = \alpha + j\beta$ was adopted. It is desired to find $\alpha + j\beta$, the attenuation and phase constant, from k_z^2 in Equation (A.71). To take the square root of a complex number (e.g. k_z^2) the following can be used, where $z = a + jb$, $r = \sqrt{a^2 + b^2}$ and $\cos \theta = a/r$,

$$\begin{aligned}
\sqrt{z} = \sqrt{a + jb} &= \sqrt{r} \left(\cos \frac{\theta}{2} + j \sin \frac{\theta}{2} \right) \\
&= \sqrt{r} \sqrt{\frac{1}{2}(1 + \cos \theta)} + j \sqrt{r} \sqrt{\frac{1}{2}(1 - \cos \theta)} \\
&= \sqrt{\frac{r}{2}} \sqrt{1 + \frac{a}{r}} + j \sqrt{\frac{r}{2}} \sqrt{1 - \frac{a}{r}} \\
&= \frac{1}{\sqrt{2}} \sqrt{r + a} + j \frac{1}{\sqrt{2}} \sqrt{r - a} \\
&= \frac{1}{\sqrt{2}} \sqrt{\sqrt{a^2 + b^2} + a} + \frac{j}{\sqrt{2}} \sqrt{\sqrt{a^2 + b^2} - a} \\
&= \frac{1}{\sqrt{2}} \sqrt{a \sqrt{1 + b^2/a^2} + a} + \frac{j}{\sqrt{2}} \sqrt{a \sqrt{1 + b^2/a^2} - a}, \quad \text{as } a \gg b, \text{ use } \sqrt{1 + x} = 1 + 1/2x \\
&= \frac{1}{\sqrt{2}} \sqrt{a(1 + \frac{b^2}{2a^2}) + a} + \frac{j}{\sqrt{2}} \sqrt{a(1 + \frac{b^2}{2a^2}) - a} \\
&\approx \frac{1}{\sqrt{2}} \sqrt{2a} + \frac{j}{\sqrt{2}} \sqrt{\frac{b^2}{2a}} \\
&\approx \sqrt{a} + \frac{jb}{2\sqrt{a}} \\
&= p + jq \tag{A.72}
\end{aligned}$$

If $b < 0$, which is the case here, then the sign of the second term must be switched. Since $\sqrt{z} = \sqrt{k_z^2} = k_z = -j\gamma$ then $(p + jq) = -j(\alpha + j\beta) = -\alpha j + \beta$; that is, $\beta = p$ and $\alpha = -q$. First β is found and to

find it the second real term in Equation (A.71) is ignored since it is much smaller than the first term and it is

$$\begin{aligned}
\beta &= p = \sqrt{a} \\
&= \sqrt{\left(\frac{2\pi}{\lambda}\right)^2 \left(1 - \left(\frac{m\lambda}{4a}\right)^2 - \left(\frac{n\lambda}{4b}\right)^2\right)} \\
&\approx \left(\frac{2\pi}{\lambda}\right) \left(1 - \frac{1}{2} \left(\frac{m\lambda}{4a}\right)^2 - \frac{1}{2} \left(\frac{n\lambda}{4b}\right)^2\right).
\end{aligned} \tag{A.73}$$

Now α is found:

$$\begin{aligned}
\alpha &= -q = -\left(-\frac{b}{2\sqrt{a}}\right) = \frac{b}{2\beta} \\
&= \frac{1}{2\beta} \left(\frac{2\pi}{\lambda}\right)^2 2 \left(\left(\frac{m\lambda}{4a}\right)^2 \frac{\bar{\epsilon}_a}{\sqrt{\bar{\epsilon}_a - 1}} \frac{\lambda}{2\pi a} + \left(\frac{n\lambda}{4b}\right)^2 \frac{1}{\sqrt{\bar{\epsilon}_b - 1}} \frac{\lambda}{2\pi b}\right) \\
&\approx \left(\frac{2\pi}{\lambda}\right) \left[\left(\frac{m\lambda}{4a}\right)^2 \frac{\bar{\epsilon}_a}{\sqrt{\bar{\epsilon}_a - 1}} \frac{\lambda}{2\pi a} + \left(\frac{n\lambda}{4b}\right)^2 \frac{1}{\sqrt{\bar{\epsilon}_b - 1}} \frac{\lambda}{2\pi b}\right] \\
&\approx \frac{1}{a} \left(\frac{m\lambda}{4a}\right)^2 \frac{\bar{\epsilon}_a}{\sqrt{\bar{\epsilon}_a - 1}} \frac{1}{b} \left(\frac{n\lambda}{4b}\right)^2 \frac{1}{\sqrt{\bar{\epsilon}_b - 1}},
\end{aligned} \tag{A.74}$$

where β was approximated with the first term. The derived α and β agree with values in [27], [30], [31], [34] and [35]. Also, since $\bar{\epsilon}_a$ and $\bar{\epsilon}_b$ may have imaginary terms the attenuation constant is revised to be

$$\alpha_{mn} = -\frac{1}{a} \left(\frac{m\pi}{4ak_0}\right)^2 \text{Re}\left\{\frac{\epsilon_a}{\sqrt{\epsilon_a - 1}}\right\} - \frac{1}{b} \left(\frac{n\pi}{4bk_0}\right)^2 \text{Re}\left\{\frac{1}{\sqrt{\epsilon_b - 1}}\right\}, \tag{A.75}$$

which is appropriate for x -polarized modes. The notation α_{mn} is used to denote the attenuation constant so that it is evident that it is for a specific mode. For y -polarized modes, the equivalent α_{mn} can be found and it is

$$\alpha_{mn} = -\frac{1}{a} \left(\frac{m\pi}{4ak_0}\right)^2 \text{Re}\left\{\frac{1}{\sqrt{\epsilon_a - 1}}\right\} - \frac{1}{b} \left(\frac{n\pi}{4bk_0}\right)^2 \text{Re}\left\{\frac{\epsilon_b}{\sqrt{\epsilon_b - 1}}\right\}. \tag{A.76}$$

The phase constant is the same as before and will be denoted by β_{mn} .

To get the field components in the inner region of the waveguide expressions for k_x^i and k_y^i must be substituted into Equations (A.9) to (A.15). There are many approximations used to find the final expressions for the field components (for example, λ/a terms with a power greater than 1 are ignored), and they are not derived here because an even more simplified form of the field equations is used in this research. The final approximated field components are stated as they are in [27]. For x -polarized or EH_{mn}^x modes

$$E_z^i = -j \left(\frac{m\lambda}{4a}\right) \cos\left(\frac{m\pi}{2a}x + \phi_x\right) \cos\left(\frac{n\pi}{2b}y + \phi_y\right) e^{-j(\alpha_{mn} + j\beta_{mn})z}, \tag{A.77}$$

$$H_z^i = j \left(\frac{\epsilon_0}{\mu_0}\right)^{1/2} \left(\frac{n\lambda}{4b}\right) \sin\left(\frac{m\pi}{2a}x + \phi_x\right) \sin\left(\frac{n\pi}{2b}y + \phi_y\right) e^{-j(\alpha_{mn} + j\beta_{mn})z}, \tag{A.78}$$

$$E_x^i = \left[\sin\left(\frac{m\pi}{2a}x + \phi_x\right) + \sin\left[\frac{j\bar{\epsilon}_a}{\sqrt{\bar{\epsilon}_a} - 1ka}\left(\frac{m\pi}{2a}x\right)\right] \cos\left(\frac{m\pi}{2a}x + \phi_x\right) \right] \times \left[\cos\left(\frac{n\pi}{2b}y + \phi_y\right) - \sin\left[j\frac{1}{\sqrt{\bar{\epsilon}_b} - 1} \frac{1}{kb}\left(\frac{n\pi}{2b}y\right)\right] \sin\left(\frac{n\pi}{2b}y + \phi_y\right) \right], \quad (\text{A.79})$$

$$H_y^i = \left(\frac{\epsilon_0}{\mu_0}\right)^2 E_x^i, \quad (\text{A.80})$$

$$E_y^i = 0, \quad (\text{A.81})$$

$$H_x^i = 0. \quad (\text{A.82})$$

For y -polarized or EH_{mn}^x modes

$$E_z^i = -j\left(\frac{m\lambda}{4a} \cos\left(\frac{m\pi}{2a}x + \phi_x\right) \cos\left(\frac{n\pi}{2b}y + \phi_y\right) e^{-j(\alpha_{mn} + j\beta_{mn})z}\right), \quad (\text{A.83})$$

$$H_z^i = j\left(\frac{\epsilon_0}{\mu_0}\right)^{1/2} \left(\frac{n\lambda}{4b} \sin\left(\frac{m\pi}{2a}x + \phi_x\right) \sin\left(\frac{n\pi}{2b}y + \phi_y\right) e^{-j(\alpha_{mn} + j\beta_{mn})z}\right), \quad (\text{A.84})$$

$$E_x^i = \left[\sin\left(\frac{m\pi}{2a}x + \phi_x\right) + \sin\left[\frac{j\bar{\epsilon}_a}{\sqrt{\bar{\epsilon}_a} - 1ka}\left(\frac{m\pi}{2a}x\right)\right] \cos\left(\frac{m\pi}{2a}x + \phi_x\right) \right] \times \left[\cos\left(\frac{n\pi}{2b}y + \phi_y\right) - \sin\left[j\frac{1}{\sqrt{\bar{\epsilon}_b} - 1} \frac{1}{kb}\left(\frac{n\pi}{2b}y\right)\right] \sin\left(\frac{n\pi}{2b}y + \phi_y\right) \right], \quad (\text{A.85})$$

$$H_y^i = \left(\frac{\epsilon_0}{\mu_0}\right)^2 E_x^i, \quad (\text{A.86})$$

$$E_y^i = 0, \quad (\text{A.87})$$

$$H_x^i = 0. \quad (\text{A.88})$$

The modes of a hollow rectangular dielectric waveguide are hybrid modes and all components of the electric and magnetic fields exist. The field equations can be simplified further by assuming that all terms containing λ/a and λ/b are small. The result is a very simplified form and this form is what is used in the analytical equation waveguide model. For x -polarized modes non-zero field components are

$$E_x^i(x, y) = \left[\frac{\sin\left(\frac{m\pi x}{2a}\right)}{\cos\left(\frac{m\pi x}{2a}\right)} \right] \cdot \left[\frac{\sin\left(\frac{n\pi y}{2b}\right)}{\cos\left(\frac{n\pi y}{2b}\right)} \right], \quad (\text{A.89})$$

$$H_y^i(x, y) = \left(\frac{\epsilon_0}{\mu_0}\right)^2 E_x^i. \quad (\text{A.90})$$

The first line applies when m is even and n is even. The second line applies when m is odd and n is odd. Therefore, appropriate quantities are chosen for each mode type. In [27] no z -dependence is explicitly written for the transverse components, but these components do propagate longitudinally in the waveguide with the appropriately derived attenuation and phase constants. For y -polarized modes the result is in the same form as Equations A.89 and A.90 except that E_y^i replaces E_x^i and H_x^i replaces H_y^i .

Appendix B

Derivation of Weighting Coefficients due to a y -polarized Current Source

This first part of this derivation derives the resultant y -oriented fields in a hollow rectangular dielectric waveguide due to a y -oriented electric current source. A similar, more general analysis has been completed for a rectangular waveguide and it is the basis for this derivation [36]. The second part of the derivation uses knowledge of a specific current source, an infinitesimal dipole, to reduce the expression of the resultant electric field.

B.1 E_y in a Tunnel due to a y -polarized Current Source

In this derivation it is assumed that the hollow rectangular dielectric waveguide or tunnel has dimensions $a \times b$ in the xy -plane, where the origin is situated in the center of the waveguide's cross section and the waveguide extends in the z -direction. This convention is in accordance with Figure 2.2, which is followed in this work.

The basis functions in a hollow rectangular dielectric waveguide are

$$E_{m,n}^{eig}(x, y) = \sin\left(\frac{m\pi}{a}x + \phi_x\right) \sin\left(\frac{n\pi}{b}y + \phi_y\right), \quad (\text{B.1})$$

where $\phi_x = 0$ if m is even, $\phi_x = \pi/2$ if m is odd, $\phi_y = 0$ if n is even and $\phi_y = \pi/2$ if n is odd.

Consider an electric current source, $\vec{J}$, in the waveguide oriented in the y -direction such that $\vec{J} = J_y \hat{y}$. The electromagnetic field inside the waveguide due to $J_y \hat{y}$ is determined from

$$\vec{H}(x, y, z) = \frac{1}{\mu_0} \nabla \times \vec{A} \quad (\text{B.2})$$

and

$$\vec{E}(x, y, z) = -\frac{j\omega}{k_0^2} \left(k_0^2 \vec{A} + \nabla(\nabla \cdot \vec{A}) \right), \quad (\text{B.3})$$

where $e^{j\omega t}$ is suppressed and $\vec{A}$ is the magnetic vector potential that satisfies the inhomogeneous wave equation

$$\nabla^2 \vec{A}(x, y, z) + k_0^2 \vec{A}(x, y, z) = -\mu_0 \vec{J}(x', y', z'). \quad (\text{B.4})$$

A superposition of weighted Green's functions will now be used to represent $\vec{A}$. A Green's function represents the response of a physical system in space due to a point exciting source [51]. If $\vec{G}$ is the dyadic Green's function for the tunnel waveguide for a unit impulse current source, $I(x', y', z')$ inside the waveguide, $\vec{A}$ can be written as

$$\vec{A}(x, y, z) = \int \int \int_{source} \vec{G}(x, y, z, x', y', z') \cdot \vec{J}(x', y', z') dx' dy' dz'. \quad (\text{B.5})$$

It is insufficient to determine Green's function uniquely and, thus, in this problem the waveguide boundary conditions will be used to solve it. Equation (B.5) can be written in component form as

$$\nabla^2 G_{yy}(\cdot) + k_0^2 G_{yy}(\cdot) = \mu_0 \delta(x - x') \delta(y - y') \delta(z - z'). \quad (\text{B.6})$$

Only the y -component is shown since only the y -directed electric field excited by the y -oriented current source is of interest. The $(\cdot)$ notation is used to represent (x, y, z, x', y', z') . The solution of Equation (B.6) may be assumed in the following form where g_{yy} is unknown

$$\begin{aligned} G_{yy}(\cdot) &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} g_{yy}(x', y', z', z) E_{m,n}^{eig}(x, y) \\ &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} g_{yy}(x', y', z', z) \sin\left(\frac{m\pi}{a}x + \phi_x\right) \sin\left(\frac{n\pi}{b}y + \phi_y\right). \end{aligned} \quad (\text{B.7})$$

This form can be assumed because any function can be expressed by a series of eigenfunctions, and these eigenfunctions satisfy the boundary conditions in the tunnel. By substituting Equation (B.7) into Equation (B.6), the right-hand-side is

$$\begin{aligned} \nabla^2 G_{yy}(\cdot) + k_0^2 G_{yy}(\cdot) &= \frac{d^2}{dx^2} G_{yy}(\cdot) + \frac{d^2}{dy^2} G_{yy}(\cdot) + \frac{d^2}{dz^2} G_{yy}(\cdot) + k_0^2 G_{yy}(\cdot) \\ &= \left(-g_{yy}(\cdot) \left(\frac{m\pi}{a}\right)^2 - g_{yy}(\cdot) \left(\frac{n\pi}{b}\right)^2 + \frac{d^2}{dz^2} g_{yy}(\cdot) + k_0^2 g_{yy}(\cdot) \right) \sin\left(\frac{m\pi}{a}x + \phi_x\right) \sin\left(\frac{n\pi}{b}y + \phi_y\right) \\ &= \left(\frac{d^2}{dz^2} g_{yy}(\cdot) + \beta_{mn}^2 g_{yy}(\cdot) \right) \sin\left(\frac{m\pi}{a}x + \phi_x\right) \sin\left(\frac{n\pi}{b}y + \phi_y\right), \end{aligned} \quad (\text{B.8})$$

where the substitution for the propagation constant β_{mn} has been made ($k_0^2 - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2 = \beta_{mn}^2$). Thus,

$$\left(\frac{d^2}{dz^2} g_{yy}(\cdot) + \beta_{mn}^2 g_{yy}(\cdot) \right) \sin\left(\frac{m\pi}{a}x + \phi_x\right) \sin\left(\frac{n\pi}{b}y + \phi_y\right) = -\mu_0 \delta(x - x') \delta(y - y') \delta(z - z'). \quad (\text{B.9})$$

Equation (B.9) is now multiplied by $\sin\left(\frac{m'\pi}{a}x + \phi_x\right) \sin\left(\frac{n'\pi}{b}y + \phi_y\right)$ and integrated over the waveguide dimensions, a and b .¹ The left- and right-hand-sides of integration are shown separately. The left-hand-

¹Note that modes are orthogonal so only when $m = m'$ or $n = n'$ will the integration over the aperture be nonzero.

side is

$$\begin{aligned}
& \int_{-a/2}^{a/2} \int_{-b/2}^{b/2} \left(\frac{d^2}{dz^2} g_{yy}(\cdot) + \beta_{mn}^2 g_{yy}(\cdot) \right) \sin\left(\frac{m\pi}{a}x + \phi_x\right) \sin\left(\frac{n\pi}{b}y + \phi_y\right) \sin\left(\frac{m'\pi}{a}x + \phi_x\right) \sin\left(\frac{n'\pi}{b}y + \phi_y\right) dx dy \\
&= \left(\frac{d^2}{dz^2} g_{yy}(\cdot) + \beta_{mn}^2 g_{yy}(\cdot) \right) \frac{1}{4} \int_{-a/2}^{a/2} \int_{-b/2}^{b/2} \left(1 - \cos\frac{2m\pi}{a}x \right) \left(1 - \cos\frac{2n\pi}{b}y \right) dx dy \\
&= \left(\frac{d^2}{dz^2} g_{yy}(\cdot) + \beta_{mn}^2 g_{yy}(\cdot) \right) \frac{1}{4} \left(1 - \frac{a}{2m\pi} \sin\left(\frac{2m\pi}{a}x\right) - \frac{b}{2n\pi} \sin\left(\frac{2n\pi}{b}y\right) + \dots \right. \\
&\quad \left. \frac{ab}{4mn\pi^2} \sin\left(\frac{2m\pi}{a}x\right) \sin\left(\frac{2n\pi}{b}y\right) \right) \Big|_{-a/2}^{a/2} \Big|_{-b/2}^{b/2} \\
&= \left(\frac{d^2}{dz^2} g_{yy}(\cdot) + \beta_{mn}^2 g_{yy}(\cdot) \right) \frac{ab}{4}, \tag{B.10}
\end{aligned}$$

where it was assumed that $\phi_x = 0$ and $\phi_y = 0$ and the result is identical for any other combinations of ϕ_x and ϕ_y . The right-hand-side is

$$\begin{aligned}
& -\mu_0 \int_{-a/2}^{a/2} \int_{-b/2}^{b/2} \delta(x - x') \sin\left(\frac{m'\pi}{a}x + \phi_x\right) \delta(y - y') \sin\left(\frac{n'\pi}{b}y + \phi_y\right) \delta(z - z') dx dy \\
&= -\mu_0 \sin\left(\frac{m\pi}{a}x' + \phi_x\right) \sin\left(\frac{n\pi}{b}y' + \phi_y\right) \delta(z - z'), \tag{B.11}
\end{aligned}$$

where the sampling property of the delta function was used.² Combining Equations (B.10) and (B.11) results in

$$\left(\frac{d^2}{dz^2} g_{yy}(\cdot) + \beta_{mn}^2 g_{yy}(\cdot) \right) = -\mu_0 \frac{4}{ab} \sin\left(\frac{m\pi}{a}x' + \phi'_x\right) \sin\left(\frac{n\pi}{b}y' + \phi'_y\right) \delta(z - z'). \tag{B.12}$$

The inhomogeneous differential equation in Equation (B.12) is solved using the Fourier transform method (in order to treat $\delta(z - z')$ on the right-hand-side). The details of this method are not shown here but can be found in [36] and [51]. The resultant expression for $g_{yy}(\cdot)$ is

$$g_{yy}(\cdot) = \frac{-j}{2\beta_{mn}} \frac{\mu_0 4}{ab} \sin\left(\frac{m\pi}{a}x' + \phi'_x\right) \sin\left(\frac{n\pi}{b}y' + \phi'_y\right) e^{-j\beta_{mn}(z-z')}, \tag{B.13}$$

where it is assumed that propagation is in the $+z$ direction of the waveguide.

By substituting Equation (B.13) back into Equation (B.7) the resultant dyadic Green's function is

$$G_{yy}(\cdot) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{-j}{2\beta_{mn}} \frac{\mu_0 4}{ab} \sin\left(\frac{m\pi}{a}x' + \phi'_x\right) \sin\left(\frac{n\pi}{b}y' + \phi'_y\right) \sin\left(\frac{m\pi}{a}x + \phi_x\right) \sin\left(\frac{n\pi}{b}y + \phi_y\right) e^{-j\beta_{mn}(z-z')}. \tag{B.14}$$

$\vec{A}_y$ can now be obtained using Equation (B.5) and is

$$\begin{aligned}
\vec{A}(x, y, z) &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{-j}{2\beta_{mn}} \frac{\mu_0 4}{ab} \sin\left(\frac{m\pi}{a}x + \phi_x\right) \sin\left(\frac{n\pi}{b}y + \phi_y\right) \\
&\quad \int \int \int_{source} \vec{J}(x', y', z') \left(\frac{m\pi}{a}x' + \phi'_x\right) \sin\left(\frac{n\pi}{b}y' + \phi'_y\right) e^{-j\beta_{mn}(z-z')} dx' dy' dz'. \tag{B.15}
\end{aligned}$$

²That is, $\int_{-\infty}^{\infty} \phi(t) \delta(t - t_o) dt = \phi(t_o)$.

Finally, E_y can be found using Equation (B.3) where, in this situation $\nabla(\nabla \cdot A_y) = \frac{d^2}{dy^2} A_y$, and it is

$$\begin{aligned}
E_y(A_y) &= \frac{-j\omega}{k_0^2} \left[k_0^2 A_y + \nabla(\nabla \cdot A_y) \right] \\
&= \frac{-j\omega}{k_0^2} \left[k_0^2 A_y - \left(\frac{n\pi}{b} \right)^2 A_y \right] \\
&= \frac{-j\omega}{k_0^2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{-j\mu_0}{2\beta_{mn}} \frac{4}{ab} \left(k_0^2 - \left(\frac{n\pi}{b} \right)^2 \right) \sin \left(\frac{m\pi x}{a} + \phi_x \right) \sin \left(\frac{n\pi y}{b} + \phi_y \right) \\
&\quad \int \int \int J_y(x', y', z') \sin \left(\frac{m\pi x'}{a} + \phi_x \right) \sin \left(\frac{n\pi y'}{b} + \phi_y \right) e^{-j\beta_{mn}(z-z')} dV' \\
&= \frac{2\omega\mu_0}{k_0^2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{1}{ab\beta_{mn}} \left(k_0^2 - \left(\frac{n\pi}{b} \right)^2 \right) \sin \left(\frac{m\pi x}{a} + \phi_x \right) \sin \left(\frac{n\pi y}{b} + \phi_y \right) \\
&\quad \int \int \int J_y(x', y', z') \sin \left(\frac{m\pi x'}{a} + \phi_x \right) \sin \left(\frac{n\pi y'}{b} + \phi_y \right) e^{-j\beta_{mn}(z-z')} dV', \tag{B.16}
\end{aligned}$$

which is the resultant y -component of the electric field due to a y -directed current source.

B.2 E_y in a Tunnel due to a y -polarized Infinitesimal Dipole

Now, Equation (B.16) will be reduced using an infinitesimal dipole as the source. The infinitesimal dipole has a finite length, l , and the current density reduces to

$$J_y(x', y', z') = \begin{cases} I_0 \delta(x' - x_0) \delta(z'), & \text{for } y_0 - l/2 \leq y' \leq y_0 + l/2 \\ 0, & \text{otherwise} \end{cases}, \tag{B.17}$$

where $(x_0, y_0, 0)$ is the center of the dipole's location and I_0 is the current on the dipole.

Solving for the non-zero current in the waveguide (assuming $\phi_y = 0$) results in

$$\begin{aligned}
&\int \int \int J_y(x', y', z') \sin \left(\frac{m\pi x'}{a} + \phi_x \right) \sin \left(\frac{n\pi y'}{b} + \phi_y \right) e^{-j\beta_{mn}(z-z')} dV' \\
&= \int \int \int I_0 \delta(x' - x_0) \delta(z') \sin \left(\frac{m\pi x'}{a} + \phi_x \right) \sin \left(\frac{n\pi y'}{b} + \phi_y \right) e^{-j\beta_{mn}(z-z')} dx' dy' dz' \\
&= I_0 \int_{y'=y_0-l/2}^{y'=y_0+l/2} \sin \left(\frac{m\pi x_0}{a} + \phi_x \right) \sin \left(\frac{n\pi y'}{b} \right) e^{-j\beta_{mn}z} dy' \\
&= I_0 \sin \left(\frac{m\pi x_0}{a} + \phi_x \right) \frac{b}{n\pi} \left[-\cos \left(\frac{n\pi(y_0 + l/2)}{b} \right) + \cos \left(\frac{n\pi(y_0 - l/2)}{b} \right) \right] e^{-j\beta_{mn}z} \\
&= I_0 \sin \left(\frac{m\pi x_0}{a} + \phi_x \right) \frac{b}{n\pi} \left[-\cos \left(\frac{n\pi y_0}{b} \right) \cos \left(\frac{n\pi l}{2b} \right) + 2 \sin \left(\frac{n\pi y_0}{b} \right) \sin \left(\frac{n\pi l}{2b} \right) + \cos \left(\frac{n\pi y_0}{b} \right) \cos \left(\frac{n\pi l}{2b} \right) \right] e^{-j\beta_{mn}z} \\
&= I_0 \sin \left(\frac{m\pi x_0}{a} + \phi_x \right) \frac{2b}{n\pi} \sin \left(\frac{n\pi y_0}{b} \right) \sin \left(\frac{n\pi l}{2b} \right) e^{-j\beta_{mn}z}. \tag{B.18}
\end{aligned}$$

The same results are obtained if $\phi_y = \pi/2$ in the above except, as expected, $\sin \left(\frac{n\pi y_0}{b} \right)$ changes to

$\cos\left(\frac{n\pi y_0}{b}\right)$. By substituting the simplification in Equation (B.18) back into Equation (B.16)

$$\begin{aligned}
E_y(A_y) &= \frac{2I_0 \omega \mu_0}{k_0^2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{1}{ab\beta_{mn}} \left(k_0^2 - \left(\frac{n\pi}{b}\right)^2\right) \sin\left(\frac{m\pi x}{a} + \phi_x\right) \sin\left(\frac{n\pi y}{b} + \phi_y\right) \\
&\quad \frac{2b}{n\pi} \sin\left(\frac{m\pi x_0}{a} + \phi_x\right) \sin\left(\frac{n\pi y_0}{b} + \phi_y\right) \sin\left(\frac{n\pi l}{2b}\right) e^{-j\beta_{mn}z} \\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \left[\frac{2I_0 \omega \mu_0}{k_0^2} \frac{2b}{n\pi} \frac{1}{ab\beta_{mn}} \left(k_0^2 - \left(\frac{n\pi}{b}\right)^2\right) \sin\left(\frac{n\pi l}{2b}\right) \sin\left(\frac{m\pi x_0}{a} + \phi_x\right) \sin\left(\frac{n\pi y_0}{b} + \phi_y\right) \right] \\
&\quad \sin\left(\frac{m\pi x}{a} + \phi_x\right) \sin\left(\frac{n\pi y}{b} + \phi_y\right) e^{-(\alpha_{mn} + j\beta_{mn})z} \\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} C_{mn} \sin\left(\frac{m\pi x}{a} + \phi_x\right) \sin\left(\frac{n\pi y}{b} + \phi_y\right) e^{-(\alpha_{mn} + j\beta_{mn})z}, \tag{B.19}
\end{aligned}$$

where C_{mn} can be identified as

$$\begin{aligned}
C_{mn} &= \frac{2I_0 \omega \mu_0}{k_0^2} \frac{2b}{n\pi} \frac{1}{ab\beta_{mn}} \left(k_0^2 - \left(\frac{n\pi}{b}\right)^2\right) \sin\left(\frac{n\pi l}{2b}\right) \sin\left(\frac{m\pi x_0}{a} + \phi_x\right) \sin\left(\frac{n\pi y_0}{b} + \phi_y\right) \\
&\approx \frac{480I_0 k}{\beta_{mn} ab} \frac{b}{n} \sin\left(\frac{n\pi l}{2b}\right) \sin\left(\frac{m\pi x_0}{a} + \phi_x\right) \sin\left(\frac{n\pi y_0}{b} + \phi_y\right). \tag{B.20}
\end{aligned}$$

Bibliography

- [1] X. Wang, X. Zhao, Z. Liang, and M. Tan, “Deploying a wireless sensor network on the coal mines,” in *IEEE International Conference on Networking, Sensing and Control*, Apr. 2007, pp. 324–328.
- [2] D. Wang, H. Feng, T. Xing, and J. Sun, “Optimized anchor nodes placement for underground mine localization system based on zigbee technology,” in *International Conference on Mechatronic Science, Electric Engineering and Computer*, Aug. 2011, pp. 478–481.
- [3] L. Zhang, X. Li, L. Chen, S. Yu, and N. Xiao, “Localization system of underground mine trackless facilities based on wireless sensor networks,” in *IEEE International Conference on Mechatronics and Automation*, Aug. 2008, pp. 347–351.
- [4] S. Rusu, M. Hayes, and J. Marshall, “Localization in large-scale underground environments with RFID,” in *Canadian Conference on Electrical and Computer Engineering*, May 2011, pp. 1140–1143.
- [5] L. Hui, Y. Lei, and W. Yuanfei, “UWB, multi-sensors and Wifi-mesh based precision positioning for urban rail traffic,” in *Ubiquitous Positioning Indoor Navigation and Location Based Service*, Oct. 2010, pp. 1–8.
- [6] Federal Communications Commission, “First report and order 02-48,” 2002.
- [7] B. Allen *et al.*, *Ultra-wideband Antennas and Propagation*, B. Allen, M. Dohler, E. Okon, W. Malik, A. Brown, and D. Edwards, Eds. John Wiley & Sons Ltd., 2007.
- [8] S. Gezici and H. Poor, “Position estimation via ultra-wide-band signals,” *Proceedings of the IEEE*, vol. 97, no. 2, pp. 386–403, Feb. 2009.
- [9] B. Lathi, *Modern Digital and Analog Communication Systems*, 3rd ed. Oxford University Press, 1998.
- [10] P. Groves, *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*. Artech House, 2008. [Online]. Available: <http://common.books24x7.com.myaccess.library.utoronto.ca/toc.aspx?bookid=27195>
- [11] G. Mao and B. Fidan, Eds., *Localization Algorithms and Strategies for Wireless Sensor Networks: Monitoring and Surveillance Techniques for Target Tracking*, 2009. [Online]. Available: IGI Global
- [12] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Win, “Ranging with ultrawide bandwidth signals in multipath environments,” *Proceedings of the IEEE*, vol. 97, no. 2, pp. 404–426, Feb. 2009.

- [13] L. Zwirello, M. Janson, C. Ascher, U. Schwesinger, G. Trommer, and T. Zwick, "Localization in industrial halls via ultra-wideband signals," in *Workshop on Positioning Navigation and Communication*, Mar. 2010, pp. 144–149.
- [14] R. Ye and H. Liu, "Ultra-wideband localization with collocated receivers," in *IEEE International Conference on Ultra-Wideband*, Sep. 2011, pp. 155–159.
- [15] A. Bahr, A. Feldman, J. Colli-Vignarelli, S. Robert, C. Dehollain, and A. Martinoli, "Modeling and benchmarking ultra-wideband localization for mobile robots," in *IEEE International Conference on Ultra-Wideband*, Sep. 2012, pp. 443–447.
- [16] O. Cetin, H. Nazli, R. Gurcan, H. Ozturk, H. Guneren, Y. Yelkovan, M. Cayir, H. Celebi, and H. Partal, "An experimental study of high precision TOA based UWB positioning systems," in *IEEE International Conference on Ultra-Wideband*, Sep. 2012, pp. 357–361.
- [17] W. Malik, C. Stevens, and D. Edwards, "Multipath effects in ultrawideband rake reception," *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 2, pp. 507–514, Feb. 2008.
- [18] J. Ono, T. Tsunoda, K. Ohno, and M. Itami, "A study on RAKE reception using multicarrier template waveform for UWB-IR system," in *IEEE International Conference on Ultra-Wideband*, Sep. 2012, pp. 212–216.
- [19] Z. Xiao, D. Jin, and N. Ge, "Fractionally spaced RAKE receiver in realistic digital SC-UWB systems," in *IEEE International Conference on Ultra-Wideband*, Sep. 2012, pp. 347–351.
- [20] H. Saghir, M. Heddebaut, F. Elbahhar, J. Rouvaen, and A. Rivenq, "Multi-user, time-reversal UWB communication for railway systems," in *IEEE Vehicular Technology Conference*, Oct. 2007, pp. 1416–1420.
- [21] D. Zhu and K. Yi, "EKF localization based on TDOA/RSS in underground mines using UWB ranging," in *IEEE International Conference on Signal Processing, Communications and Computing*, Sep. 2011, pp. 1–4.
- [22] I. Mabrouk, L. Talbi, and M. Nedil, "Improvement of wireless channel performance using MIMO-UWB system in underground mine gallery," in *IEEE International Conference on Ultra-Wideband*, 2012, pp. 155–159.
- [23] A. Molisch, "Ultra-wide-band propagation channels," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 353–371, Feb. 2009.
- [24] A. Emslie, R. Lagace, and P. Strong, "Theory of the propagation of UHF radio waves in coal mine tunnels," *IEEE Transactions on Antennas and Propagation*, vol. 23, no. 2, pp. 192–205, Mar. 1975.
- [25] S. Mahmoud and J. Wait, "Geometrical optical approach for electromagnetic wave propagation in rectangular mine tunnels," *Radio Science*, vol. 9, no. 12, pp. 1147–1158, 1974.
- [26] P. Delogne, "Basic mechanisms of tunnel propagation," *Radio Science*, vol. 11, no. 4, pp. 295–303, 1976.
- [27] K. D. Laakmann and W. H. Steier, "Waveguides: characteristic modes of hollow rectangular dielectric waveguides," *Applied Optics*, vol. 15, no. 5, pp. 1334–1340, May 1976.

- [28] P. Delogne, "EM propagation in tunnels," *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 3, pp. 401–406, Mar. 1991.
- [29] M. Lienard and P. Degauque, "Natural wave propagation in mine environments," *IEEE Transactions on Antennas and Propagation*, vol. 48, no. 9, pp. 1326–1339, Sep. 2000.
- [30] D. Dudley, M. Lienard, S. Mahmoud, and P. Degauque, "Wireless propagation in tunnels," *IEEE Antennas and Propagation Magazine*, vol. 49, no. 2, pp. 11–26, Apr. 2007.
- [31] F. Fuschini and G. Falciasecca, "A mixed rays-modes approach to the propagation in real road and railway tunnels," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 2, pp. 1095–1105, Feb. 2012.
- [32] N. Sood, "Realistic assessment of novel wireless systems with ray-tracing based techniques," Master's thesis, University of Toronto, 2012.
- [33] D. Didascalou, T. Schafer, F. Weinmann, and W. Wiesbeck, "Ray-density normalization for ray-optical wave propagation modeling in arbitrarily shaped tunnels," *IEEE Transactions on Antennas and Propagation*, vol. 48, no. 9, pp. 1316–1325, Sep. 2000.
- [34] Z. Sun and I. Akyildiz, "Channel modeling and analysis for wireless networks in underground mines and road tunnels," *IEEE Transactions on Communications*, vol. 58, no. 6, pp. 1758–1768, Jun. 2010.
- [35] S. Mahmoud, "Modal propagation of high frequency electromagnetic waves in straight and curved tunnels within the earth," *Journal of Electromagnetic Waves and Applications*, vol. 19, no. 12, pp. 1611–1627, 2005.
- [36] M. Deshpande, "Analysis of discontinuities in a rectangular waveguide using dyadic Green's function approach in conjunction with method of moments," *NASA Report 201692*, April 1997.
- [37] Z. Sun and I. Akyildiz, "A mode-based approach for channel modeling in underground tunnels under the impact of vehicular traffic flow," *IEEE Transactions on Communications*, vol. 10, no. 10, pp. 3222–3231, Oct. 2011.
- [38] H. Bertoni, *Radio Propagation for Modern Wireless Systems*. Prentice Hall PTR, 2000.
- [39] C. A. Balanis, *Antenna Theory*, 3rd ed. John Wiley & Sons, Inc., 2005.
- [40] D. Cheng, *Field and Wave Electromagnetics*, 2nd ed. Addison-Wesley Publishing Company, Inc., 1989.
- [41] H. Sheng, P. Orlik, A. Haimovich, L. Cimini, and J. Zhang, "On the spectral and power requirements for ultra-wideband transmission," in *IEEE International Conference on Communications*, vol. 1, May 2003, pp. 738–742.
- [42] J. G. Proakis and M. Salehi, *Fundamentals of Communication Systems*. Pearson Prentice Hall, 2005.

- [43] A. Ott, C. Eisner, and T. Eibert, “Enhanced investigations on effective isotropic radiated power emissions of impulse radio devices,” in *European Conference on Antennas and Propagation*, Mar. 2012, pp. 1589–1593.
- [44] R. A. Scholtz, D. M. Pozar, and W. Namgoong, “Ultra-wideband radio,” *European Association for Signal Processing Journal on Applied Signal Processing*, vol. 2005, pp. 252–272, Jan. 2005.
- [45] M. Schmidt, H. Jakel, and F. Jondral, “Spectral properties of ultra wideband signals with general time-hopping codes,” in *International Conference on Computer as a Tool*, vol. 1, Nov. 2005, pp. 124–127.
- [46] J. Leung, “Hybrid waveguide theory-based modeling of indoor wireless propagation,” Master’s thesis, University of Toronto, 2009.
- [47] C. Grosvenor, R. Johnk, J. Baker-Jarvis, M. Janezic, and B. Riddle, “Time-domain free-field measurements of the relative permittivity of building materials,” *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 7, pp. 2275–2282, Jul. 2009.
- [48] Y. Shen, C. Law, and W. Dou, “Ultra-wideband measurement of the dielectric constant and loss tangent of concrete slabs,” in *China-Japan Joint Microwave Conference*, Sep. 2008, pp. 537–540.
- [49] J. Zhang and Y. Huang, “Extraction of dielectric properties of building materials from free-space time-domain measurement,” in *High Frequency Postgraduate Student Colloquium*, 1999, pp. 127–132.
- [50] L. Liang, “Evaluation of UWB beamformers in a wireless channel and potential microwave implementations,” Master’s thesis, University of Toronto, 2010.
- [51] A. Ishimaru, *Electromagnetic wave propagation, radiation, and scattering*. Prentice Hall, 1991.