Sparse Directional Estimation of the Time-Varying Underwater Acoustic Channel

by

Graham McIntyre

Submitted in partial fulfillment of the requirements
for the degree of Master of Applied Science

at

Dalhousie University
Halifax, Nova Scotia
October 2015

Table of Contents

List of Figures

Abstract

This thesis focuses on the directional estimate of the underwater acoustic channel using modal beamforming with an orthogonal frequency division multiplexed (OFDM) communication system. An underwater channel model and simulator is first presented that is capable of implementing the time-scaling associated with the wideband time-varying acoustic channel. Modal beamforming is then introduced as a unique wideband beamforming method that makes use of circular arrays. The last Section explores the effect of a time-varying acoustic channel on an OFDM signal and produces a frequency domain channel that models the distortions that are introduced. This frequency domain model is then used to create a sparse approximation dictionary that, when coupled with the modal beamformer and a sparse approximation technique, is capable of producing accurate sparse directional channel estimates. These estimates include the complex amplitude, delay, time-scaling factor, and direction of arrival for each cluster arrival in the channel. This directional channel estimate is shown to greatly improve the performance of an OFDM receiver in a time-varying underwater acoustic channel when compared to classic estimation techniques.

Chapter 1

Introduction

1.1 Motivation for Studying Underwater Acoustic Communication

Underwater acoustic communication can trace its roots back to the early 1900's when the technology was explored as a method to inform ships of shallow water and the presence of rocks as an alternative to sirens and lights from lighthouses. While the original idea was simply to use a submerged bell as the sound source, a Canadian named Reginald Fessenden suggested using an electronic acoustic source and later demonstrated that the electronic source he created could also be used to transmit morse-encoded messages. By pulsing the underwater source Fessenden was able to communicate between two tug boats in the Boston harbour in 1913, giving birth to wireless underwater acoustic communication [1].

While Fessenden introduced the original idea, underwater acoustic communication was further developed and pioneered by the military mainly for use by navies as a method of communication with submarines. The use of sonar to track submarines also led to extensive studies into the propagation characteristics of acoustic waves underwater to determine how different environments would affect the performance of an acoustic system [1].

Recently there has been a shift back to the original roots of underwater communication for commercial purposes for a range of applications. Examples include underwater sensor networks for environmental monitoring and ocean sampling networks, as well as communication with underwater vehicles for exploration and other tasks such as

pipeline inspection and seismic monitoring of underwater oil and gas fields [2]. Examples of systems that require underwater communication links can also be found here in Nova Scotia.

Tracking of underwater mammals is one such example that is important to biologists to study the movement of the animals, as well as to the oil and gas industry in order to mitigate harm caused to animals during seismic testing. One method to detect mammals is to use their vocalisations and determine their position using multiple sensors distributed underwater. A second option, pursued by a local company, is to tag each animal and track the movement of the tag using reception devices that record the communication pings transmitted by the tags.

There are a number of other applications for underwater communication links that can be found in the local industry. Some examples are diver communication systems, acoustic releases, and monitoring and control of tidal turbines, to name a few. A commonality between all of the systems mentioned above is that while there are communication platforms currently available that can be used for their needs, they could all be improved by an underwater wireless link that functions more reliably at higher data rates.

For the marine mammal tracking problem in particular, a higher rate link could produce better estimates of location using the first method by allowing for more accurate time synchronisation between the distributed sensors. For the tagging solution, a wideband pulse could increase the probability of reception either over longer distances or when multiple tags are located in the same area. Benefits of higher bandwidth communication can be found for all of the examples above. The challenge with a high rate link, however, lies in the channel through which the data must be sent.

The underwater acoustic communication channel has a number of extra challenges when compared to the radio communication channel. The low speed of propagation of sound underwater when compared to the speed of light (for radio-frequency communication) leads to much longer propagation delays between transmission and reception when traveling over the same distance.. This low speed of propagation also means that even small transmitter and/or receiver velocities can lead to distortions in the transmitted signal. Over any substantial distance bandwidth is limited due to low carrier frequencies required to limit the frequency dependent loss experienced by a time-varying signal. Most underwater channels also have severe multipath resulting in echoes that can be perceived over an extended period of time after the first recep-

tion. And lastly, the varying sound speed profile in the ocean can lead to 'shadow zones' caused by refraction where no acoustic energy can be transmitted, thus leading to uncontrollable communication outages [3].

The wireless revolution that has occurred in the past twenty years in the radio world has yielded a number of new physical layer techniques for communication that can help to alleviate some of these problems. One of the main purposes of the underwater communications lab here at Dalhousie is to explore how these new techniques can be applied. However, some problems are largely unique to the underwater channel and require their own solutions.

Therefore, the motivation for this thesis is to apply modern communication techniques to underwater acoustic communication in order to improve both the data rate and link reliability of a wireless underwater acoustic link. More specifically, this thesis focuses on communication through time-varying underwater channels either due to intentional receiver/transmitter motion or due to movement of, or in the medium. The reason for this is that as the data rates increase with the low carrier frequencies used even minor variations of the channel over a symbol duration can cause drastic performance reduction. The techniques presented in this work will help to improve the performance of the applications described above, and may also facilitate the development of new ideas and products that can make use of a high data rate underwater wireless communication link.

1.2 Thesis Contributions

This thesis presents a wideband digital underwater communications system that can reliably transmit over medium to long ranges (> 500 m) in extreme channel conditions and in both shallow and deep water applications. While over short ranges there are a number of plausible solutions for wireless communication through water, such as optical or magneto inductive communications, over any significant distance greater than a few hundred meters the only viable solution remains acoustic waves.

In this thesis Orthogonal Frequency Division Multiplexing (OFDM) is used to modulate the acoustic wave that travels through the underwater channel. OFDM is a multi-carrier modulation technique that has a number of benefits over traditional single carrier techniques. These benefits will be discussed in Chapter 4. The layout of an underwater communication system making use of OFDM modulation is shown in

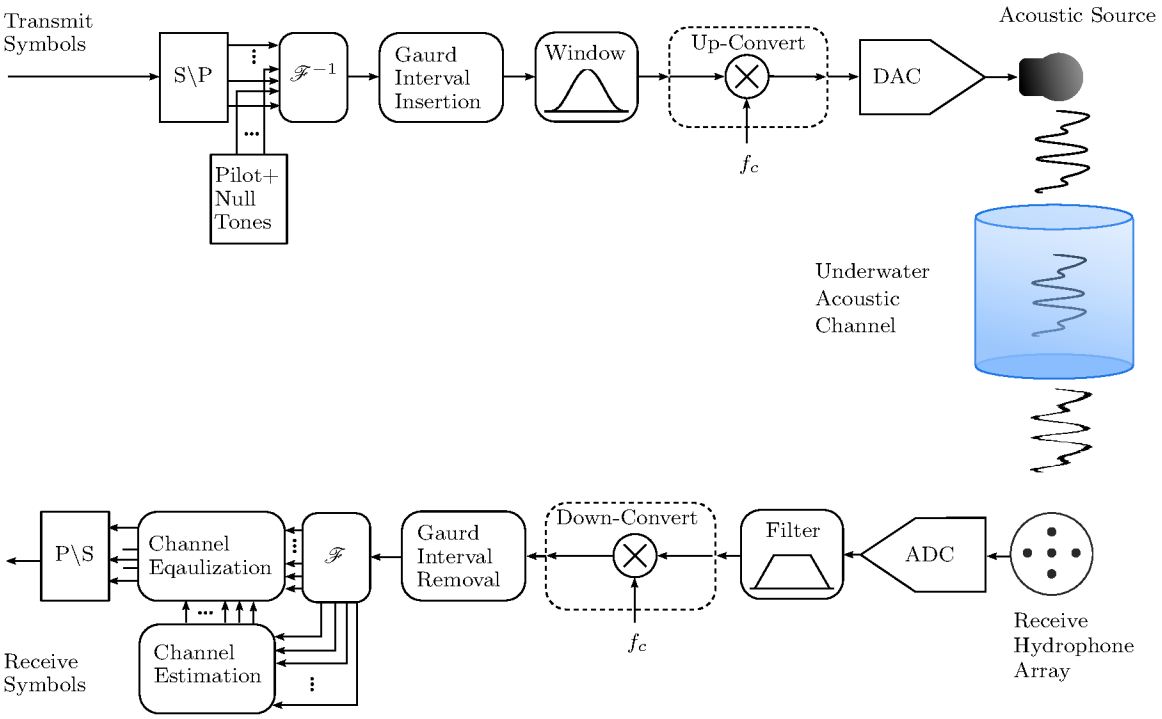


Figure 1.1: Block diagram of an underwater OFDM communication system with channel equalization.

Figure 1.1 [4].

This thesis focuses on three main components of the communication system in Figure 1.1 when the system parameters are chosen to transmit data at a high rate such that the transmitted signal is wideband in nature. The term *wideband* is discussed in detail in Chapter 2. These components are the channel model, the receive hydrophone array processing, and the channel estimation technique.

A channel model is required in order to be able to test communication systems in a simulated environment. An accurate model which properly simulates the effects of the channel that distort a signal as it travels from the transmitter to the receiver is required as it leads to more accurate analyses of the effect of changing other components. In Chapter 2 a stochastic model is presented that captures many of the predominant effects of the underwater acoustic channel. The main contribution of Chapter 2 is the introduction of the time-varying fractional delay line (TVFDL) as a method of implementing the stochastic channel. This is a new method to implement a time-varying channel that is able to simply and effectively implement all of the effects that the underwater channel has on a transmitted signal.

Chapter 2 also studies the underwater channel in detail and shows that it introduces

both temporal and spatial diversity to the signal, meaning the same information is received over different paths. In order to make use of spatial diversity, a method of capturing the angle-of-arrival of the received signal needs to be implemented. These methods are often referred to as beamforming techniques as beams are pointed in the direction of interest.

The second contribution of this thesis is the analysis, implementation and experimental measurement of a wideband modal beamformer, which is presented in Chapter 3. Modal beamforming is a non-conventional beamforming technique that makes use of a circular array of receive elements. The benefit of this technique is that once the wideband orthogonal modes of interest are produced, a wideband beamformer can be implemented that points in any direction with any pattern. This cannot be done with conventional beamforming techniques. While modal beamforming is a known technique for circular arrays, the implementation of a wideband modal beamformer using real weights in a digital environment has not been fully explored for the underwater environment.

The last contribution of this thesis is a wideband sparse directional channel estimation technique that makes use of the channel model presented in Chapter 2 and the modal beamformer presented in Chapter 3. This channel estimation technique uses the inherent sparsity of the received signal in both the temporal and spatial domain in order to more accurately estimate the channel parameters. This technique is unique in the fact that it estimates the directional channel parameters for a wideband signal using a circular array. Chapter 3 shows through simulation that the use of the directional sparse channel estimator greatly improves the performance of a underwater receiver in a time-varying environment.

1.3 Thesis Outline

The outline of the thesis is briefly reviewed here. Chapter 2 explores the physical properties of the channel itself and proposes a time-varying sparse multipath channel model to simulate it. Section 2.1 looks at the effect of the underwater acoustic channel on a propagating acoustic wave and highlights the important aspects of the underwater channel that need to be included in a channel model. Section 2.2 discusses how time-variance in the channel should be modeled for wideband signals and a mathematical model that uses time-scaling is proposed. The parameters of this equation are then

described by creating a stochasitc channel model. Finally, in Section 2.3 the time-varying fractional delay line (TVFDL) is presented as a method to implement the wideband channel model as well as any other model or operation that requires time-scaling, such as a resampling operation.

Chapter 3 focuses on the receive hydrophone array and more specifically on different methods of non-adaptive wideband beamforming. In Section 3.1, beamforming using a circular array is reviewed and the differences between narrowband and wideband beamforming are discussed. Section 3.2 then reviews classic wideband beamforming techniques that could be used with the circular array and discusses why these techniques are not appropriate. Section 3.3 introduces modal beamforming as a novel alternative to the classic technique and uses complex phase modes to introduce the techniques and present some of its drawbacks. Amplitude modes are then presented as the real alternative to phase modes that can be used to practically implement a modal beamformer and the implementation of a wideband amplitude mode beamformer is summarized. Section 3.4 then presents experimental results that demonstrate that modal beamforming can be used in an underwater environment.

Chapter 4 explores the estimation of a time-varying channel using sparse estimation techniques in order to improve equalization when the channel varies with time over an OFDM block. In Section 4.1 the effect of this time-variance on the received OFDM signal is first reviewed in detail and a frequency domain model is then produced to account for the effects of a time-varying channel on an OFDM signal. Using this frequency domain model along with the wideband modal beamformer from Chapter 3, an estimation problem is then developed that poses time-varying channel estimation as a sparse estimation problem that can be solved using sparse estimation techniques. In Section 4.3 simulations that implement this estimate are run and the estimate is then used to equalize an OFDM signal that has been passed through the time-varying UAC outlined in Chapter 1 and implemented using the TVFDL. These simulations show that a channel estimation technique that accounts for time-variance greatly improves the performance of an OFDM system in a time-varying channel and also demonstrates that adding directionality to the estimate improves the performance further.

Chapter 2

The Underwater Acoustic Channel

The underwater acoustic channel (UAC) is often considered as one of the most challenging communication channels. The difficulty associated with deploying a physical system also means that proper physical trials are rare, and so an accurate channel model is required in order to accurately assess the performance of a communication scheme in the lab. The purpose of this Chapter is to familiarize the reader with the underwater channel and introduce how the underwater channel can be modeled and implemented in a simulation environment. The simulated channel presented in this Chapter is used throughout the rest of the thesis for simulation.

Section 2.1 explores the physical properties of underwater acoustic wave propagation in the UAC and focuses on modeling multipath propagation and time-variance. Section 2.2 then presents the time-scale domain as an accurate domain in which to model the time-variance introduced to a wideband signal by the UAC, and a stochastic model is presented to describe the parameters associated with a simulated UAC. Section 2.3 concludes the Chapter by introducing the time-varying fractional delay line (TVFDL) as a practical method for implementing a time-varying UAC simulator using the stochastic channel description outlined in Section 2.2.

2.1 Underwater Acoustic Wave Propagation

Due to the rapid attenuation of electromagnetic (EM) waves in electrically conductive salt water and the inability to use optical communications in many underwater environments, acoustic waves are often considered the most well suited tool for communi-

cation over any substantial distance underwater. Unfortunately, there are a number of aspects of acoustic waves that make them undesirable for this application. First and foremost, acoustic waves propagate at speeds on the order of 5 magnitudes slower than EM waves. This means that the delay between transmission and reception of information is greatly increased and the ability to have full duplex communications in real-time is extremely difficult. Other undesirable properties of the UAC can be placed into two distinct groups, large-scale effects and small-scale effects. Large scale effects are dealt with in Section 2.1.1 and indicate that the effects are only noticeable over tens or hundreds of wavelengths. Small scale effects are covered in Section 2.1.2 and indicate the effects take place over much smaller distances, generally over a few wavelengths or less [5].

2.1.1 Large Scale Effects of the Underwater Acoustic Channel

The first large scale effect that needs to be considered is transmission loss, or the decrease in the intensity of the wave with respect to distance. Transmission loss is generally modeled by considering two effects, spreading and absorption [6]. Spreading occurs as the wave propagates and the area of the wave front increases. In order to maintain the transmitted energy the intensity of the wave must decrease as the wave propagates. The simplest case to consider is spherical spreading, where the wave is emitted from a point source and spreads through a homogeneous, infinite medium spreading along the surface of a sphere.

Unfortunately, the UAC channel cannot be modeled using spherical spreading exactly because the channel is bounded over large enough distances by the surface and the sea floor, and the ocean cannot be considered a homogeneous medium over large depth variations. A better approximation is to assume spherical spreading until the boundaries are met, and then cylindrical spreading, which takes into account refraction in the medium and the reflection of the wave front off of the upper and lower boundaries of the channel, thereafter [6].

The transmission loss can then be modeled as a piecewise functions where spherical spreading is assumed when the transmission range, r, is less than half of the depth of the water, r_0, and cylindrical spreading is assumed in excess of the spherical spreading when r is greater than r_0 [6].

$$\mathrm{TL}_{\mathrm{spreading}} = \begin{cases} 20\log r & r < r_0 \\ 20\log r_0 + 10\log(r - r_0) & r > r_0. \end{cases} \tag{2.1}$$

The second cause of transmission lose in the UAC is absorption, where the water absorbs part of the wave energy through both viscosity and chemical reactions. Specifically, the three main sources of absorption are shear viscosity, relaxation of magnesium sulphate ($MgSO_4$), and relaxation of boric acid ($B(OH)_3$) [7].

The absorption in sea water is frequency dependent and can be described using the Francois-Garrison model [6],

$$\alpha(f) = A_1 P_1 \frac{f_1 f^2}{f_1^2 + f^2} + A_2 P_2 \frac{f_2 f^2}{f_2^2 + f^2} + A_3 P_3 f^2. \tag{2.2}$$

In this Equation $\alpha(f)$ is a frequency dependent loss with units of dB/km and the three terms represent the loss due to $B(OH)_3$, $MgSO_4$ and shear viscosity respectively. The values for A_n, P_n and f_n are calculated differently for each term and depend on the salinity, temperature, depth and pH of the water. The equations for these values can be found in [6]. Figure 2.1 shows the effect that frequency has on the attenuation due to absorption and demonstrates that under 10 kHz the loss per kilometer remains fairly small, under 1 dB/km. However, at a frequency of 100 kHz the loss increases greatly to close to 40 dB/km. So for transmission over any distance low frequencies in the tens of kilohertz range should be used.

When absorption and spreading loss are considered together, then a simple and relatively accurate model for transmission loss is

$$TL = \begin{cases} 20\log r + \alpha(f)\frac{r}{1000}, & \text{when} r < r_0 \\ 20\log r_0 + 10\log(r - r_0) + \alpha(f)\frac{r}{1000}, & \text{when} r > r_0, \end{cases} \tag{2.3}$$

where r is in meters and needs to be converted to kilometers for the absorption term. Figure 2.2 is a plot of the total attenuation experienced by an acoustic wave in the UAC with respect to range for a number of frequencies and shows the fairly drastic difference in attenuation between a wave with a frequency of 10 and 50 kHz.

The UAC also presents a unique challenge in deep underwater environments due to refraction. The speed of propagation of an acoustic wave in water is predicted by the temperature, pressure and salinity of the water and can vary greatly over large depth

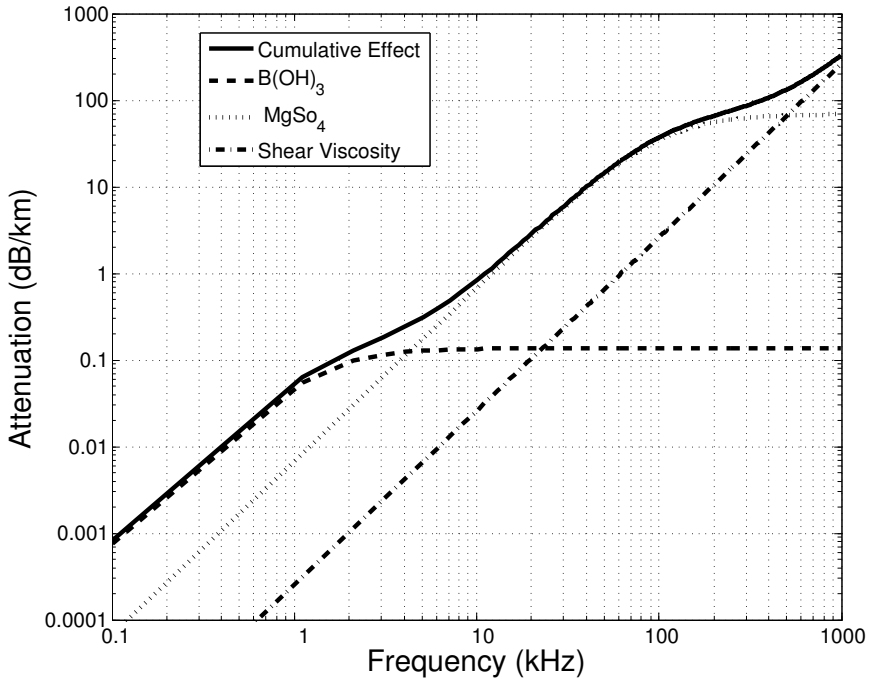


Figure 2.1: Attenuation in seawater from absorption in dB/km with respect to frequency. These curves were created using a temperature of 15° C (close to the average temperature of seawater), salinity of 35 p.s.u, a pH of 8 (median pH of seawater), and a depth of 0 meters. Varying the depth does not drastically change this curve.

variations. This leads to channels with velocity, or sound speed profiles (SSP) that vary with depth and cause refraction of the acoustic wave as it travels in any direction off the horizontal.

Figure 2.3 shows the effect that a depth-varying SSP can have on the overall coverage from an acoustic projector in a simple, deep underwater environment. This simulation was run using the Bellhop model with a 250 Hz wave in a deep water environment with a depth of 5000 meters. The source was placed at 1500 meters and the simulation was run using a constant SSP and a depth varying SSP. The plots in Figures 2.3a and 2.3c show these profiles and Figures 2.3b and 2.3d show the corresponding Bellhop simulations, respectively. The refraction caused by the the depth varying SSP leads to shadow zones [8] where, when compared to the model with the constant SSP, the intensity of the received wave is reduced or sometimes even vanishes.

The effect of large scale variations can therefore mainly be modeled as a range and frequency dependent loss term. In the simplest case with a constant SSP, the loss is simply due to transmission loss as shown in Eqn. (2.3). This is often applicable to shallow water deployments. For deep water deployments where the variable SSP

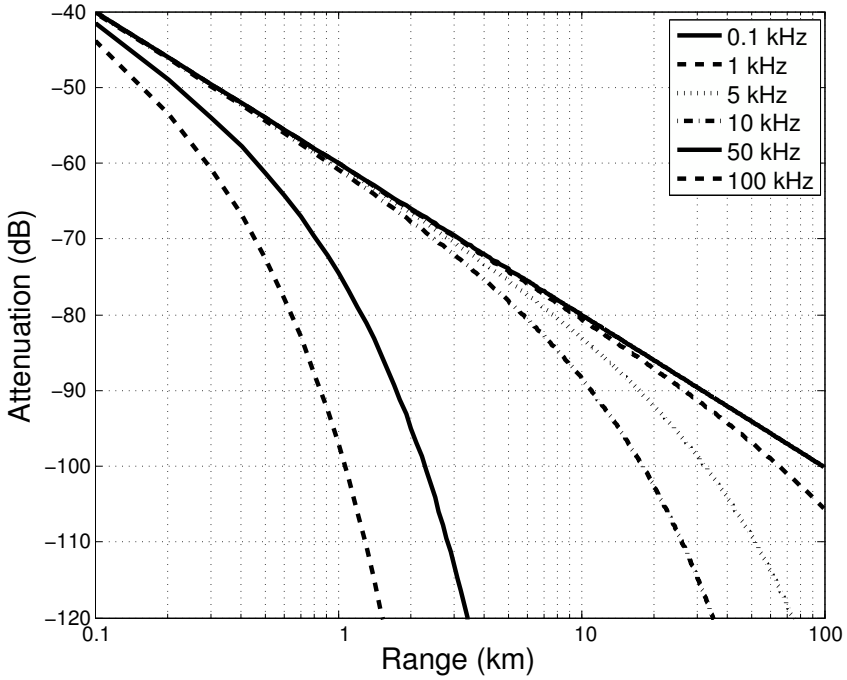
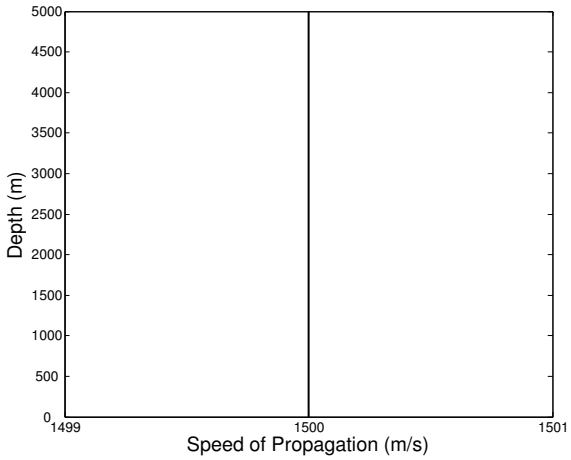


Figure 2.2: Total transmission loss in seawater from absorption and spreading for different frequencies. These curves were created using a temperature of 15° C, salinity of 35 p.s.u, depth of 0 meters, and pH of 8.

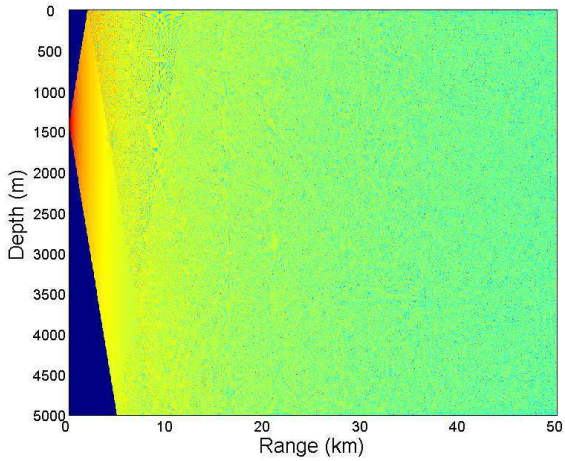
causes shadow zones, if the receiver is constantly moving it may be important to simulate the probability of an outage to model when a shadow zone is encountered. However, this is beyond the scope of this work as the focus is on the channel and its variations over small time scales, on the order of seconds, where it is assumed propagation paths are present. So, in this Chapter and the Chapters that follow the large scale effects will be modeled using a frequency dependent transmission loss filter, $H_{TL}(f)$, with a frequency response based off of Eqn. (2.3) for a desired range, r.

2.1.2 Large Scale Effects of the Underwater Acoustic Channel

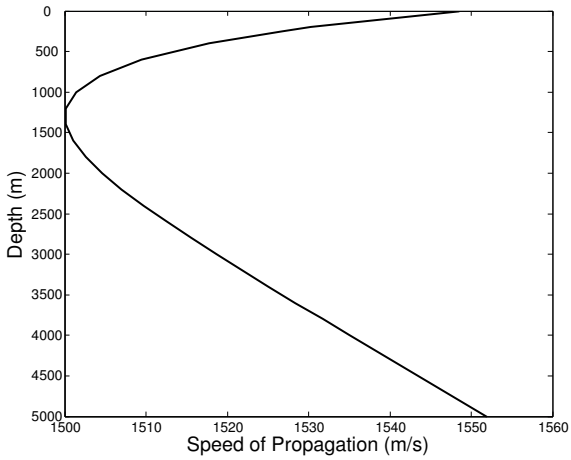
With large scale fading covered, the focus now switches to the small scale effects of the UAC on the transmitted signal. While refraction causes a large scale effect in the form of shadow zones, it is also one of the causes of a significant small scale effect, mulitpath. Multipath occurs when the transmitted signal is received from a number of different paths rather than just a single, direct path. Due to the fact that each non-direct path has a longer distance to travel than the direct path, these multipath components have a delay, or excess delay, with respect to the time of arrival of the



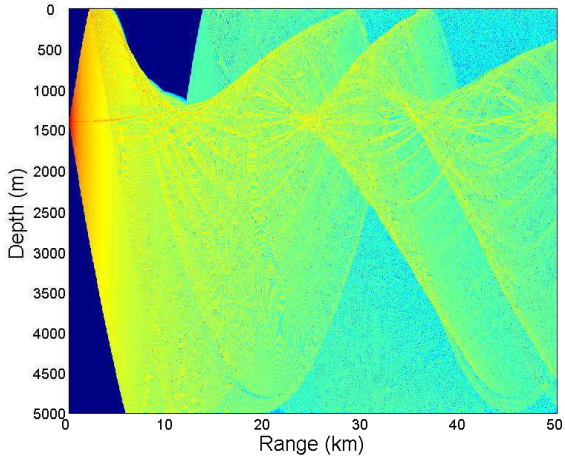
(a) Constant SSP.



(b) Bellhop model output for constant SSP.



(c) Munk SSP.



(d) Bellhop model output for Munk SSP.

Figure 2.3: Comparison of Bellhop outputs, (b) and (d), for a constant and depth-varying sound speed profile, (a) and (c), respectively. All parameters except for the SSP remain constant for both simulations.

direct path signal. A possible refracted path, RP_{t0}, that could cause multipath is shown in Figure 2.4 for time t_0.

The second source of multipath is reflection off the surface and bottom of the ocean. The underwater medium can therefore be thought of as a waveguide with the transmitted acoustic wave reaching the receiver through a number of reflected paths interacting with the water surface and bottom [8].

Similar to the refracted paths, the reflected paths travel along a longer path when compared to the direct path, as shown by the bottom and surface reflections in Figure 2.4, and so also have a delay when compared to the direct path arrival. An

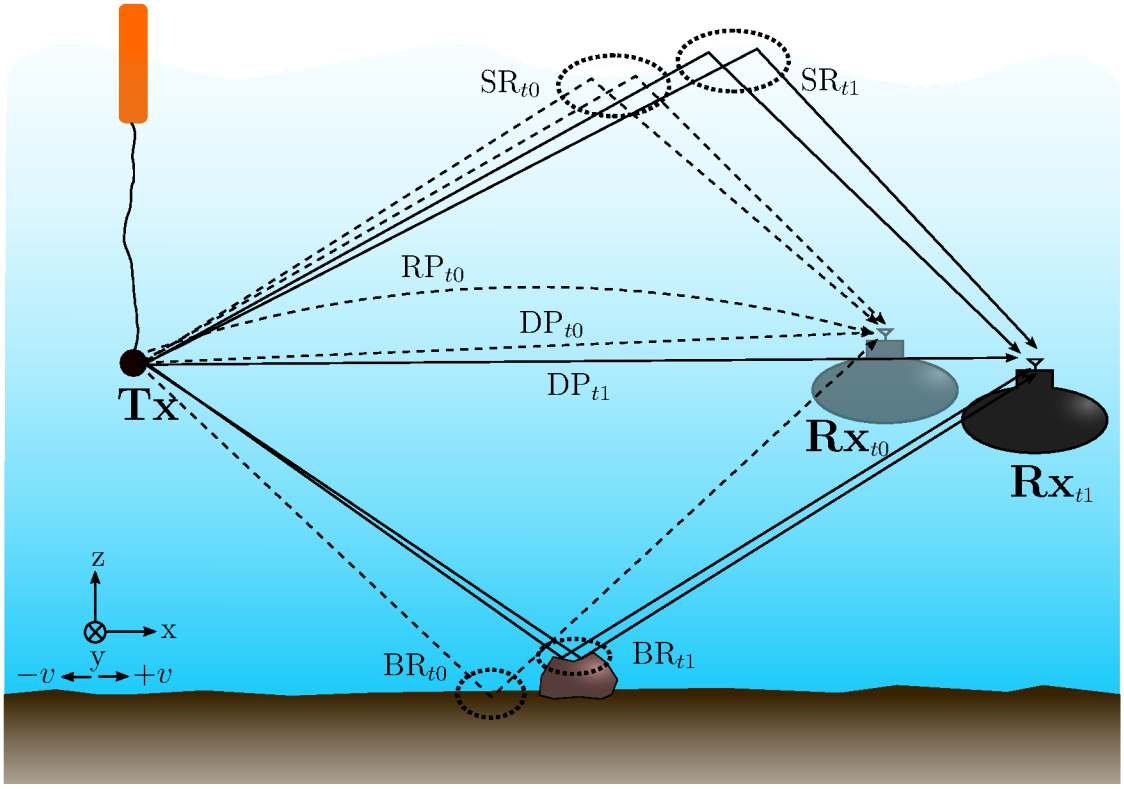


Figure 2.4: A simple underwater communication deployment with a stationary acoustic transmitter (Tx) and a moving receiver (Rx), shown at an initial time, t_0, and a final time, t_1. The dotted paths represent the dominant paths at t_0 and the solid lines represent the dominant paths at t_1. DP is the direct path, RP is a possible refracted path, SR is a surface reflection and BR is a bottom reflection.

equation that is often used to model the small scale effects of the UAC, or channel impulse response (CIR), is a sum of time delayed and amplitude scaled versions of the transmitted signal [9], [10], [11],

$$\tilde{h}(t,\tau) = \sum_{p=1}^{N_p} a_p \delta(\tau - \bar{\tau}_p(t)), \tag{2.4}$$

where N_p is the number of distinct paths received at the transmitter, a_p is the path-dependent gain, $\bar{\tau}_p(t)$ is the path-dependent delay at time t. $\tilde{h}$ is used to show that this is the passband channel measured at the carrier frequency. The relevance of this will be discussed in more detail in Chapter 4.

It is important to note here that due to the low speed of propagation in water, even minor variations in path length can lead to substantial time-delays between the received paths. Also, due to the sparsity of the scattering environment, the number

of paths included in the summation, N_p, is often low. Therefore the UAC for most deployments is often considered to be a sparse channel, where $\tilde{h}(t,\tau)$ only contains a small number of non-zero terms [9], [10].

This sparsity also manifests itself in the direction of arrival of the different paths. The small number of paths coupled with the large displacement between the two major reflectors, the waters surface and bottom, indicates there should be a good spread in the angle-of-arrival for each path. Therefore, the path arrivals should also be sparse with respect to angle-of-arrival [12], [13]. This angular sparsity can be seen in Figure 2.4 by comparing the angles-of-arrival of the three main paths. This sparsity in delay and direction-of-arrival will be exploited in the channel estimation scheme presented in Chapter 4.

The model in (2.4) is often referred to as a linear time-varying (LTV) channel model, where the dependence on time is motivated by two factors for underwater communication, motion of the transmitter and/or receiver and movement of the scatterers.

The first form of time-variance that will be discussed is movement of the scatters. For the UAC when the transmitter and receiver are stationary, the most prominent source of time-variance comes from movement of the waters surface, while the reflection from the bottom is generally considered to be non time-varying [8]. This movement presents a number of challenges.

First, the movement of surface waves causes the length of the surface reflected paths to vary with time. The rough surface also causes multiple localized scattering points, which means that instead of a single surface reflection, often there will be multiple reflections at very similar delays [8]. This is shown for both surface reflections at time t_0 and at time t_1 in Figure 2.4. Large waves that pass can also cause focusing of multiple beams for short range deployments ($r < 500$ m) that can produce large fluctuations in received power [14]. Finally, the rough water surface and the bubbles produced by crashing waves can also produce a loss that can be frequency dependent [14].

Therefore, in order to account for the effects of the time-varying water surface it's helpful to rewrite the LTV channel so that each multipath arrival is in fact made up of a cluster of scatterers as

$$\tilde{h}(t,\tau) = \sum_{c=1}^{N_c} \sum_{p=1}^{N_{p,c}} a_p \delta(\tau - \bar{\tau}_c(t) - \bar{\tau}_p(t)), \tag{2.5}$$

where N_c is the number of cluster arrivals, $N_{p,c}$ is the number of paths in a cluster, $\bar{\tau}_c(t)$ is the delay associated with each cluster arrival, and $\bar{\tau}_p(t)$ is the deviation from the cluster delay for each path arrival within the cluster. Within a cluster, $\bar{\tau}_p(t))$ can be positive or negative and is centered around 0. For the rest of this thesis, a cluster path, or cluster arrival is used to refer to the cluster of path arrivals with the delay $\bar{\tau}_c(t)$.

In general, when there is no transmitter or receiver motion then for the direct path or bottom reflection $N_{p,c} = 1$ and the cluster should consist of a single path. While for a surface reflection $N_{p,c} > 1$ and so the cluster consists of a number of different paths closely separated in delay. For the surface reflection, if $|\bar{\tau}_p(t)|$ is small for all paths in the cluster, then the paths all arrive at roughly the same time, $\bar{\tau}_c(t)$, but with different amplitudes and phases due to the different reflecting points. Their summation will therefore cause a time-varying amplitude for the c^{th} cluster. This will be discussed in more detail in the Section 2.2.

Above it was assumed that the bottom reflection is time-invariant when the transmitter and/or receiver are stationary. If receiver/transmitter motion is now considered, then it seems plausible that the bottom reflection and possibly the direct path might become time-variant as well. Specifically, it seems justified that due to the fact that the bottom reflected path is now moving as either the transmitter or receiver moves, that there will be some change in the bottom reflected path. This is shown in Figure 2.4 as the receiver moves from the position at t_0 to the position at t_1. While the bottom reflected cluster consists of a single path at t_0 that reflects off of the smooth sea floor, at t_1 the bottom reflection is now reflecting off of a rock and has two paths. This rock could potentially block the bottom reflection in between t_0 and t_1 and then change the gain as more paths are included in the reflection.

Although this is a contrived situation, it seem plausible that if the receiver and/or transmitter are moving that the bottom path will vary and that $N_{p,c} > 1$ will exist for the bottom-reflected clusters as well. Therefore, the bottom reflection can be treated in a similar way to the surface reflection. A second consequence of intentional transmitter/receiver motion is the rapid variation of the cluster delays, $\bar{\tau}_c(t)$, which causes the received signal to be dilated or compressed depending on the relative velocity between the transmitter and receiver.

With a general intuition into the effects that the time-varying underwater channel has on an acoustic signal, the goal now is to first determine how best to model the

time-varying underwater channel making use of Eqn. (2.5), and then how to select the parameters used to describe this model. This is the focus of Section 2.2.

2.2 A Stochastic Description of the Underwater Acoustic Channel

This Section focuses on creating a stochastic description of the UAC that can be used to simulate and model an actual underwater communication deployment. Section 2.2.1 looks at the best way to describe the channel and focuses on how the time-variance should be modeled, specifically looking into which domain should be used to describe the channel variations caused by the rapid movement of the cluster paths due to intentional transmitter/receiver motion. Section 2.2.2 then introduces a method for statistically describing the channel parameters of interest and briefly discusses the drawbacks of these descriptions.

2.2.1 The Wideband Channel Model

The UAC is often referred to as a doubly-selective, or doubly-spread channel, which indicates that the transmitted signal will have been distorted by spreading it in two dimensions to produce the received signal. Using Eqn.(2.5), the received passband signal, $\tilde{r}(t)$, produced by transmitting a passband signal, $\tilde{s}(t)$, through the passband UAC can be written as

$$\tilde{r}(t) = \tilde{s}(t) * \tilde{h}(t,\tau) = \sum_{c=1}^{N_c} \sum_{p=1}^{N_p} a_p \tilde{s}(t - \bar{\tau}_c(t) - \bar{\tau}_p(t)). \tag{2.6}$$

It is easy to see that the first dimension the signal has been spread over is the delay-domain, τ, due to the multipath arrivals of the original signal. This is referred to as delay-spread, or frequency-selective fading as multipath arrivals cause different frequencies to experience different gains. The second dimension over which the signal is spread is not as easy to identify and manifests itself in the variation of the channel with respect to time, t. The second spreading domain is then related to the time-variation of the multipath arrivals, $\bar{\tau}_c(t)$ and $\bar{\tau}_p(t)$.

There are two domains that can be used to quantify the spread caused by time-

variance, the Doppler domain and the time-scale domain. The Doppler domain assumes a narrowband signal and takes the Fourier transform of the channel, $\tilde{h}(t,\tau)$, with respect to time. Specifically, the delay-Doppler spreading function, $S_H(\tau,\nu)$, is defined as

$$S_H(\tau,\nu) = \int_{-\infty}^{\infty} \tilde{h}(t,\tau)e^{-j2\pi t\nu}d\nu. \tag{2.7}$$

Two important parameters of the delay-Doppler spreading function that give more insight into each cluster arrival are the Doppler-spread and the Doppler-shift. The Doppler-spread is the root-mean-square (RMS) width of the spreading function with respect to frequency, ν, for a specific tap and indicates the frequency shifts on the paths included in a cluster arrival. Due to the non-coherent combination of these paths to form a cluster arrival, their superposition causes the amplitude of the cluster arrival to vary, or fade with time. The rate of this amplitude fading depends on the Doppler-spread. If the spread is large then each path varies quickly causing the superposition of the paths to vary quickly as well [15]. Similarly, if the Doppler-spread is small, the fading occurs more slowly. So the Doppler-spread is caused by the superposition of the paths that form a cluster arrival and their respective time-varying path delays, $\bar{\tau}_p(t)$.

The Doppler-shift is a measure of the mean frequency shift, and therefore is associated with intentional transmitter/receiver velocity [15]. The Doppler-spread is usually small around the mean Doppler-shift due to the fact that the underwater channel is a sparse scattering environment with mainly stationary or slowly time-varying scatters. If the relative velocity between the transmitter and receiver is large, leading to rapidly changing cluster delays, $\bar{\tau}_c(t)$, the Doppler-shift will also be large. A fundamental question that needs to be answered then is when can a constant frequency shift, a Doppler-shift, be used to model the effects of velocity in a channel on a transmitted signal and when should a time-scale be used. A time-scale models the time-variance by compressing or dilating the signal in time, which more realistically models the effect of time variance for any frequency while a Doppler-shift only applies to a specific group of frequencies.

This question then is directly related to what constitutes a small group, or narrowband of frequencies and what constitutes a wideband system. This leads back to when should the Doppler-domain be used and when should the time-scale domain be used to model the channel. In order to answer this question it's more informative to separate

the causes of the Doppler-shift and Doppler-spread, which involves separating the effects of the path variance with respect to time, $\bar{\tau}_p(t)$, and the cluster variance with respect to time, $\bar{\tau}_c(t)$. This can be done by rewriting Eqn. (2.6) as

$$\tilde{r}(t) = \sum_{c=1}^{N_c} \tilde{s}(t - \bar{\tau}_c(t)) * \sum_{p=1}^{N_p} a_p \delta(t - \bar{\tau}_p(t)). \tag{2.8}$$

With this rearrangement and knowing $\bar{\tau}_p(t)$ is only a small variation around $\bar{\tau}_c(t)$, this question can be reposed as when can $\tilde{s}(t - \bar{\tau}_c(t))$ no longer be represented by a narrowband Doppler shift. The time-varying cluster delay, $\bar{\tau}_c(t)$, causes the time-variation of the signal and can be related to the change in the path length associated with a specific cluster, $d_c(t)$.

$$d_c(t) = d_{0,c} - v_c(\tau - \tau_{0,c}), \tag{2.9}$$

where $d_{0,c}$ is the initial path length and $\tau_{0,c}$ is the initial delay of the c^{th} cluster. Here it is assumed the change in path length can be represented as a first-order polynomial with a constant cluster velocity, v_c. The time variable multiplying the velocity is chosen as $(t - \tau_{0,c})$ rather than simply t to ensure that at time $\tau_{0,c}$ the path length $d_c(t) = d_{0,c}$ [16]. The velocity term includes both the transmitter and receiver velocity and can be expressed as

$$v_c = v_{Tx,c} - v_{Rx,c}, \tag{2.10}$$

where $v_{Tx,c}$ and $v_{Rx,c}$ indicate the transmitter and receiver velocity projected onto the cluster direction of departure from the transmitter and direction of arrival at the receiver, respectively. Figure 2.4 shows the sign convention for the velocities.

This path length can be converted into a time-varying delay by dividing by the speed of propagation of the wave in water, c_w,

$$\bar{\tau}_c(t) = \frac{d_c(t)}{c_w} = \tau_{0,c} - \frac{v_c}{c_w}(t - \tau_{0,c}). \tag{2.11}$$

So the $(t - \bar{\tau}_c(t))$ argument in Eqn. (2.8) can be written as

$$t - \bar{\tau}_c(t) = t - \tau_{0,c} + \frac{v_c}{c_w}(t - \tau_{0,c}) = (t - \tau_{0,c})(1 + \frac{v_c}{c_w}). \tag{2.12}$$

The received signal on the c^{th} cluster path is then

$$\tilde{s}(t - \bar{\tau}_c(t)) = \tilde{s}(\alpha(t - \tau_{0,c})) \tag{2.13}$$

where α_c represents the time-scaling factor for each cluster arrival

$$\alpha_c = (1 + \frac{v_c}{c_w}). \tag{2.14}$$

The time scaling factor $\alpha_c > 1$ when $v_c > 0$ and the path length is decreasing, causing compression of the signal with respect to time. When $\alpha_c < 1$, $v_c < 0$ and the path length is increasing, causing dilation of the signal with respect to time.

In order to derive a narrowband condition the received passband signal on the c^{th} cluster path in (2.13) can be converted to the frequency domain,

$$\tilde{R}_c(f) = \mathscr{F}\Big\{\tilde{s}(t - \bar{\tau}_c(t))\Big\} = \frac{1}{\alpha_c}\tilde{S}\left(\frac{f}{\alpha_c}\right)e^{-j2\pi f\tau_{0,c}}. \tag{2.15}$$

The frequency term can be rewritten as

$$\frac{f}{\alpha_c} = f\left(1 - \frac{v_c}{c_w}\right) = f - f\frac{v_c}{c_w}. \tag{2.16}$$

$\tilde{R}_c(f)$ represents the passband signal received on the c^{th} path in the frequency domain. If it is assumed the signal is bandlimited around a carrier frequency, f_c, with a bandwidth, B, then the frequency range of the signal is $[f_c - \frac{B}{2}, f_c + \frac{B}{2}]$. The first narrowband assumption that can be made is that as long as $B << f_c$ then $f \approx f_c$, and the scaled frequency becomes

$$\frac{f}{\alpha_c} \approx f - f_c\frac{v_c}{c_w}. \tag{2.17}$$

Therefore, as long as the fractional bandwidth is much less than 1, $\frac{B}{f_c} << 1$, then the scaling can be represented as a constant frequency shift of $f_c\frac{v_c}{c_w}$. The channel can also be assumed narrowband as long as the scaling factor is close to one, $\frac{v_c}{c_w} << 1$, then the assumption in Eqn. (2.17) could also be made for most systems even when the bandwidth is large relative to the center frequency [17].

In underwater communication systems with any intentional velocity these assumptions do not hold due to two main properties of the underwater acoustic channel. First,

because attenuation increases with frequency, for any signal that is transmitted over a long range the carrier frequency must be low, on the order of kHz's, as was shown in Section 2.1. So to transmit data at any substantial rate, $\frac{B}{f_c} < 1$, which violates the first narrowband condition. Often signals are referred to by their fractional bandwidth. For narrowband signals, $\frac{B}{f_c} < 0.01$, for wideband signals, $0.01 < \frac{B}{f_c} < 0.2$, and for ultrawideband signals, $\frac{B}{f_c} > 0.2$ [18]. Therefore, in underwater acoustic communication with any substantial data rate the signals are often wideband or ultrawideband in nature simply because the center frequencies used are so low.

Secondly, because the speed of propagation of acoustic waves in water is low, $c_w \approx 1500$ m/s, any velocity caused by deliberate transmitter and receiver movement will be substantial compared to c_w. So the second assumption is also not satisfied when the receiver and/or transmitter velocity is substantial rather than just slow drifting.

Because many underwater channels will be inherently wideband in nature, in this work the time-scale domain is considered as the best domain to use in order to quantify how the signal varies with time due to intentional velocity [15], [19], [11]. The channel in Eqn. (2.8) can then be written as a summation of time scaled and delayed cluster arrivals, each of which has been convolved with separate path arrivals that implement only small time-variations about the mean cluster delay. This can be expressed as

$$\tilde{r}(t) = \sum_{c=1}^{N_c} \sqrt{\alpha_c}\tilde{s}(\alpha_c(t - \tau_{0,c})) * \sum_{p=1}^{N_p} a_p \delta(t - \bar{\tau}_p(t)), \tag{2.18}$$

where the transmitted signal is multiplied by a normalization factor, $\sqrt{\alpha_c}$, to ensure the received and transmitted signal have the same energy [17].

In Eqn. (2.18) $\bar{\tau}_p(t)$ is assumed to only be a very small variation around $\bar{\tau}_c(t)$ and this indicates that the change of each path in the cluster is very close to the change of the main cluster arrival. So v_p, the residual difference in velocity for each path from the cluster velocity, should only be very small compared to v_c. This indicates that according to the second narrow band assumption that a Doppler shift is sufficient to represent the change of each path within the cluster. Therefore, each path can be represented using a frequency, or Doppler shift of the already time-scaled signal. This shift will have a magnitude of

$$f_p = f_c \frac{v_p}{c_w}. \tag{2.19}$$

The final passband received signal can be described as [20], [21], [22]

$$\tilde{r}(t) = \sum_{c=1}^{N_c} a_c(t)\tilde{s}(\alpha_c(t-\tau_c)), \tag{2.20}$$

where $\tau_{0,c}$ has been replaced here by τ_c for notational simplicity. The time variation of the amplitudes of each cluster path is captured by

$$a_c(t) = l_c(\tau_c)\sqrt{\alpha_c}\sum_{p=1}^{N_{p,c}} a_p e^{j(2\pi f_p t+\theta_p)}. \tag{2.21}$$

This time-fading term is the sum of Doppler shifted version of the already time-scaled signal on the c^{th} path, corresponding to the different path arrivals that make up a cluster arrival. Each path arrival has a residual Doppler frequency, f_p, a random phase, θ_p, due to specular scattering, and an amplitude, a_p, that could vary due to the strength of the reflection. The last term, $l_c(\tau_c)$, represents a constant amplitude scaling term that accounts for the power scaling due to the excess delay with respect to the first path arrival.

While this model holds for most wideband systems, if the signal is ultrawideband it may be more accurate to represent each path in a cluster as its own delay-scaled version of the transmit signal and spread the signal in delay closely around the cluster delay, but not at the exact same time [18]. However, for the rest of this work it is assumed the model in Eqn.(2.20) holds for the received signal.

2.2.2 Stochastic Channel Description

The wideband channel model from Section 2.2.1 provides a mathematical model to accurately represent the distortions introduced by a time-varying channel. In order for this model to apply to a UAC, the parameters in Eqn. (2.20) need to be chosen properly. The key parameters in this model being: 1) the delays associated with each cluster arrival, τ_c, 2) the number of cluster arrivals required for the model, N_c, 3) the time-scaling associated with each cluster, α_c, 4) the Doppler shifts caused by the path arrivals in each cluster, f_p and θ_p, and 5) the angle-of-arrival of each path, ϕ_p. The angle-of-arrival is important when multiple receive elements are used as this angle is used to describe how the signal is received in space. This will be discussed further in Chapter 3.

Stochastic models are often used to describe communication channels because while a deterministic model may capture the desired channel response for one specific deployment at different instances in time, it is not representative of the ensemble of channels from similar deployments [15]. Often, it is also very difficult to quantify what is causing specific phenomenon in the channel which makes reproducing the effects using a deterministic set of parameters more difficult as well. Therefore, in this work the channel model is produced using channel parameters that are described stochastically.

Here, simple stochastic models are presented for each parameter of interest and the combination of these models when applied using Eqn. (2.20) constitutes the final stochastic channel model. Some of these parameter models are justified through measurements, while for others simplifying assumptions are made. The main goal of this stochastic description of the channel is to reproduce a time-varying channel that *could* represent an underwater channel, but that does not perfectly recreate a channel for a given geometry or deployment.

2.2.2.1 Cluster Delay and Time-Scale

To describe the delay and time-scale parameters of the cluster arrivals a channel delay-scale, or wideband, spreading function is used, $F_H(\tau, \alpha)$. The received signal is then

$$\tilde{r}(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_H(\tau, \alpha) \sqrt{\alpha} \tilde{s}(\alpha(t - \tau)) d\tau d\alpha. \tag{2.22}$$

The double integration in (2.22) occurs for a single instant in time and so the time-varying gain term, $a_c(t)$, in Eqn. (2.20) remains constant and so it is excluded here for simplicity. Due to the nature of the channel, this function has finite support and is usually limited in delay between 0 and τ_{max} and in scale between $-\alpha_{max}$ and $+\alpha_{max}$ [15]. Also, because the channel is considered sparse with respect to the number of multipath arrivals, only a few delay-scale pairs will be non-zero in $F_H(\tau, \alpha)$ and can be represented by N_c delta functions at different locations in the support space for τ and α. This leads to the approximation in Eqn. (2.20) where the double integral over the delta functions is replaced by a single summation with N_c cluster arrivals each with a delay and time-scaling value associated with the delta functions.

In this work the delay-scale spreading function is represented using a wide sense sta-

tionary uncorrelated scattering (WSSUS) model. This model makes two assumptions. First, it is assumed that the statistics associated with each channel cluster, c, are wide-sense stationary. This implies that the mean delay and time-scale are constant with respect to time and the correlation function for a cluster, or tap in $F_H(\tau, \alpha)$, only depends on the time-difference and does not vary with time itself. The uncorrelated scattering assumption simply implies that all paths are uncorrelated, or are produced by independent scatterers [15].

With the WSSUS assumption, the delay-scale function is a 2D white process where each delay-scale pair is uncorrelated with every other delay-scale pair. Using this assumption, a statistical description of the the delay-scale function can be written as [15]

$$E\{F_H(\tau, \alpha), F_H(\tau', \alpha')\} = B_H(\tau, \alpha)\delta(\tau - \tau')\delta(\alpha - \alpha'). \tag{2.23}$$

The delay-scale scattering function, $B_H(\tau, \alpha)$, is a function that describes the average power of a scatterer with delay, τ, and scale, α. Therefore, the delay-scale scattering function is a statistical description of the likelihood of the channel containing a cluster arrival with a specific delay-scale pair. A useful property of the scattering function is that it is proportional to the probability density function (pdf), $p_H(\tau, \alpha)$, describing the distribution of the channel with respect to delay and time-scale [23].

A single realization of the channel can then be described by drawing a number of cluster arrivals from the pdf, $p_H(\tau, \alpha)$, describing the channel. Specifically, N_c cluster arrivals are drawn where the specific number is a deterministic value set to match the average, or expected number of paths encountered in a deployment.

Before a channel realization can be produced, the scattering function has to be created. In order to produce the scattering function it can be assumed that the delay and time-scaling are independent of each other and the function is separable in delay and time-scale [24]. This assumption is not based off of the physics of the deployment. Inherently the delay will be related to the time-scaling factor based on the geometry of the path. However, the goal of this model is to implement the general effect that the channel has on a signal and not to model a specific deployment. So the exact relationship between the two variables is not of interest and for this reason the two variables are assumed to be separable. Using this assumption, the wideband scattering

function can then be described as

$$B_H(\tau, \alpha) = b_H(\tau) b_H(\alpha), \tag{2.24}$$

and the task of creating the scattering function then becomes a task of describing $b_H(\tau)$, the delay power profile of the cluster arrivals, and $b_H(\alpha)$, the scale-spread.

A model that is commonly used for the delay power profile is a truncated negative exponential function [15], [25], [5], of the form

$$b_H(\tau) = \frac{1}{\tau_{rms}} e^{-\frac{\tau}{\tau_{rms}}}, \quad \tau \geq 0, \tag{2.25}$$

where τ_{rms} is the root mean square delay spread. In order to limit this function to a maximum delay, τ_{max}, the function can be reformulated as

$$b_H(\tau) = \begin{cases} \frac{b}{\tau_{rms}} e^{-\frac{\tau}{\tau_{rms}}}, & \tau \geq 0, \\ 0, & \tau > \tau_{max} \end{cases} \tag{2.26}$$

where

$$b = \frac{1}{1 - e^{-\frac{\tau_{max}}{\tau_{rms}}}}. \tag{2.27}$$

The factor b acts as a normalization term that ensures

$$\int_0^{\tau_{max}} p(\tau) = 1, \tag{2.28}$$

and so the power delay profile can be treated as a true pdf that integrates to 1 [25].

While the power delay profile follows a conventional form for selecting multipath components in a scattering environment, there is no such function for describing the statistics associated with the scale term, $b_H(\alpha)$. One approach would be to use a rough geometry-based model to determine the velocity as well as angle of arrival associated with each tap. The velocity for each path could then be used to determine the scale parameter using Eqn. (2.14).

Obtaining a deterministic model of the channel however requires a substantial amount of knowledge about the deployment and does not constitute a statistical description. Instead, a simple model can be produced where the possible cluster velocities can be represented as a uniform distribution from $v_c \in [v_{min}\ v_{max}]$. Where v_{min} and v_{max}

represent the minimum and maximum velocities expected, respectively. Using Eqn. (2.14), the distribution of the scale term can then be represented as

$$b_H(\alpha) = \begin{cases} \frac{1}{\alpha_{max}-\alpha_{min}}, & \alpha_{min} \leq \alpha_p \leq \alpha_{max} \quad \text{when} \quad v_{max} > 0, \\ \frac{1}{\alpha_{min}-\alpha_{max}}, & \alpha_{max} \leq \alpha_p \leq \alpha_{max} \quad \text{when} \quad v_{max} < 0, \\ 0, & \text{otherwise}, \end{cases} \tag{2.29}$$

where α_{min} and α_{max} correspond to the time scale terms calculated using v_{min} and v_{max}, respectively.

With definitions for $b_H(\alpha)$ and $b_H(\tau)$, the delay-scale scattering functions that describes the channel can be produced. This scattering function requires no normalization to be treated as the pdf describing the delay, time-scale distribution.

2.2.2.2 Path Doppler Frequencies and Phases

The channel scattering function provides a method for selecting cluster arrivals with delays and time-scaling factors. However it does not account for the fading behaviour that has been observed on the cluster arrivals with respect to time [26], represented in the model in (2.20) by $a_c(t)$. In Section 2.2.1 it was shown that while the time-variation of each cluster is best represented by a time-scaling factor rather than a Doppler shift, α_c, the variation of the path arrivals within each cluster is best represented as a sum of frequency shifted version of the time scaled signal. So the task of describing $a_c(t)$ is in fact a task of describing the Doppler profile for each cluster arrival and selecting $N_{p,c}$ frequencies from this profile that corresponds to the Doppler shifts on each path arrival within a cluster.

The choice of the Doppler profile is different for underwater acoustic communication than for radio communication. While Jake's Doppler profile is often used to model mobility in the channel in radio communication system when the receiver is surrounded by a ring of scatterers [15], [27], extensive tests carried out in [26] show that a better Doppler profile for underwater acoustics is a stretched exponential function of the form

$$S(f) = Ae^{-\left(\frac{|f|}{a}\right)^B}. \tag{2.30}$$

When describing a Doppler profile, a quantity that is often used is the coherence

time, T_c. This time indicates how long a cluster tap remains correlated with itself in time with respect to a coherence level between 0 and 1. This value therefore provides insight into the Doppler spread of the profile. With a coherence time specified, the Doppler profile itself can be fully described.

The full derivations for the parameters associated with the Doppler profile in (2.30) are shown in Appendix A but are summarized here. With a coherence time specified, the parameters describing the stretched exponential in Eqn. (2.30) can be chosen as $B = 1$,

$$a = \sqrt{\frac{1.5}{2\pi^2 T_c^2}}, \tag{2.31}$$

and

$$A = \frac{1}{2a}. \tag{2.32}$$

A is a normalizing factor that is chosen so that $S(f)$ can be treated as a pdf. a is the shape parameter that controls the Doppler spread of the profile and thus is related to the coherence time. Here, a is specified for a coherence level of 0.5, however this parameter can be recalculated for other coherence levels using Eqn. (A.6).

To account for the fading of the cluster taps with time, $a_c(t)$, a can be calculated using the desired coherence time, T_c, and A can be calculated to normalize the integral of $S(f)$ to 1. $N_{p,c}$ frequencies, f_p, corresponding to the residual velocities on each path within a cluster can then be drawn from Eqn. (2.30) and these frequencies can be used in Eqn. (2.21) to produce the fading on each cluster arrival.

To complete the model of the time-varying amplitude, Eqn. (2.21) requires two more inputs, a phase for each path, θ_p, and a constant scaling term to account for the power-scaling due to the delay of each cluster with respect to the direct path, $l_c(\tau_c)$. The phase term can simply be drawn from a uniform distribution for 0 to 2π for each path. The power-scaling term however is related to the delay, and is therefore related to the power-delay profile. The scaling term can then be calculated as

$$l_c(\tau_c) = e^{-\frac{\tau_c}{2\tau_{rms}}}. \tag{2.33}$$

2.2.2.3 Cluster Angle-of-Arrival

When the receiver has multiple receive elements, it is also useful to model the angle-of-arrival of each cluster at the receiver, ϕ_c. The last parameter that needs to be described then is the angle-of-arrival. Similar to the cluster velocity, the angle-of-arrival is best described using a geometry based solution. For simplification purposes, in this work it is assumed that the angle-of-arrival of each cluster is directly related to the delay of the cluster. So as the delay increases the angle of arrival increases up to a maximum angle of arrival, ϕ_{max}. The cluster angle-of-arrival can then be described as

$$\phi_c = (-1)^b \tau_c \frac{\phi_{max}}{\tau_{max}} + n_\phi, \tag{2.34}$$

where b is Bernoulli random variable, with $p = 0.5$, that randomly distributes the sign of the received cluster to simulate a signal coming from above and below the direct path, which is assumed to have a relative angle of arrival of 0. This simulates surface and bottom reflections. A small noise term, n_ϕ, is also inserted to add random variation.

2.2.2.4 Stochastic Model Summary

So, the stochastic model of the channel can be summarized as follows for a time-domain implementation:

1. Initialization

   - Specify model parameters
     (a) The range of the deployment, r.
     (b) Number of cluster arrivals, N_c.
     (c) Number of path arrivals for each cluster (excluding direct path), $N_{p,c}$.
     (d) The RMS and maximum delay of the channel, τ_{RMS} and τ_{max}.
     (e) The coherence time of the channel, T_c.
     (f) The maximum and minimum velocities associated with the cluster arrivals, v_{max} and v_{min}.
     (g) The maximum angle of arrival for the clusters, ϕ_{max}.

- Create the frequency dependent transmission loss filter, $H_{TL}(f)$, using r and Eqn. (2.3).
- Create wideband scattering function , $B_H(\tau, \alpha)$, using τ_{RMS} and τ_{max} and v_{max} and v_{min} with Eqns. (2.26) and (2.29). Treat as pdf and draw N_c delay-scale pairs for cluster arrivals.
- Create Doppler profile using T_c and Eqn. (2.30). Treat as pdf and draw $N_{p,c}$ path Doppler frequencies, f_p, for each cluster arrival.
- For each cluster arrival calculate the constant excess delay amplitude scaling coefficient, $l_c(\tau_c)$, using Eqn. (2.33).
- Calculate the angle-of-arrival for each cluster arrival using Eqn. (2.34).

2. Implementation

   (a) Filter the transmitted passband signal, $\tilde{s}(t)$, using $H_{TL}(f)$.

   (b) For each discrete time sample, n:

      i. Calculate the scaled and delayed sample of the signal, $\tilde{s}(\alpha_c(t - \tau_{0,c}))$.

      ii. Calculate the fading amplitude coefficient using Eqn. (2.21) and multiply time-scaled and delayed sample by fading coefficient.

      iii. Repeat steps i. and ii. for all clusters, following Eqn. (2.20). Sum to produce received signal at time instance n.

   (c) Repeat step (b) incrementing discrete time-sample until the end of $\tilde{s}(t)$ is reached.

This model does contain a number of simplifying assumptions. First, it was assumed that the cluster velocity is uncorrelated with the angle-of-arrival and delay associated with each cluster. In most circumstances, there will be a relationship between these parameters, and further studies will be required to demonstrate the interrelationship between them. Second, the Doppler profile used here may be a simplification of a true deployment with delay-spread and transmitter/receiver velocity. Lastly, while the WSSUS assumption may hold over short time periods, over long periods this assumption most likely won't hold in most deployments and so this model may be limited to short simulation durations.

So while the model presented above does have a few drawbacks, it still captures all of the desired properties of a wideband channel and reproduces the physical distortions of the received signal introduced by a time-varying UAC.

2.3 Channel Implementation

In order to simulate the UAC using the stochastic model presented above a simulation model needs to be created. Section 2.3.1 first reviews two common simulation methods used for both narrowband and wideband cases, and then discusses some of the drawbacks of these models. The time-varying fractional delay line (TVFDL) is then introduced in Section 2.3.2 as a method for overcoming these shortcomings. Section 2.3.3 concludes the Chapter with simulations demonstrating the ability of the TVFDL to implement the stochastic channel model presented in Section 2.2.2 as well as a channel with a varying time-scaling factors.

2.3.1 Current Simulation Models

When it is assumed the sampled channel parameters do not change with time (i.e. the the delay, scale and Doppler frequencies remain unchanged) then both the wideband and narrowband models can be implemented in similar ways. This involves discretizing the spreading function, and then implementing methods to either shift or scale the signal in the frequency domain. This method is most easily understood using the narrowband model where the time-variation is captured using a Doppler shift and a narrowband delay-Doppler spreading function.

For the narrowband case this method is called a basis expansion model (BEM), where the basis that is chosen represents how the channel varies with time. A basis that is often used with the delay-Doppler domain spreading function is the complex exponential, or Fourier basis due to the inherent relationship with the delay-Doppler domain, refer to Eqn. (2.7).

So, using a complex exponential basis function to model the channel time variation corresponds to taking the inverse Fourier transform of the narrowband delay-Doppler spreading function, $S_H(\tau, \nu)$. If the spreading function is discretized in both the delay and Doppler domain, then the discrete time-varying channel impulse response can be written as [15]

$$\tilde{h}[n, m] = \frac{1}{N} \sum_{d=0}^{N-1} S_H[m, d] e^{j2\pi \frac{nd}{N}}. \tag{2.35}$$

The time variance of the channel can then be captured using a multiplication by a

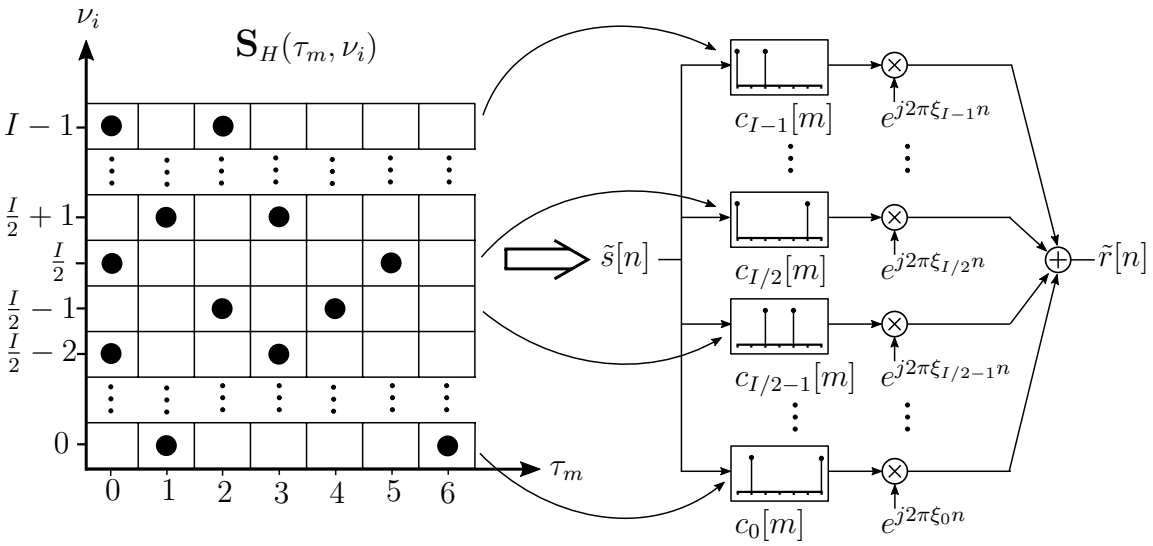


Figure 2.5: Discrete delay-Doppler spreading function with I complex exponential basis being converted to a basis expansion model with corresponding FIR filters and basis.

complex exponential basis, $e^{j2\pi\frac{nd}{N}}$, with respect to the discrete time variable, n, and the delay, or delays associated with each basis function is captured using an FIR filter, $c[m] = S_H[:, d]$. The number of basis functions used can be oversampled for better simulation results. If I basis functions are used, then the channel model using the complex exponential basis becomes [15]

$$\tilde{h}[n, m] = \sum_{i=0}^{I-1} c_i[m] e^{j2\pi\xi_i n}, \tag{2.36}$$

where ξ_i is the frequency associated with each basis. This model is summarized in Figure 2.5.

A similar approach is used for the wideband simulation model, however a Mellin transform is used with the wideband spreading function, $F_H(\tau, \alpha)$, instead of a Fourier transform as the Mellin transform can easily handle time-scaling [15]. Two main differences exist between the narrowband BEM and the wideband model. The first and most obvious difference is that the wideband model uses a time-scale instead of a exponential basis to account for time-variance. The second difference is how the delay and scales are sampled during the discretization process. While with the BEM the delay and Doppler parameters are uniformly sampled, in the wideband model the delay and scale parameters are geometrically sampled due to the use of the Mellin transform [15].

Aside from these two differences, the wideband channel implementation is very similar to the implementation in Figure 2.5 with $S_H(\tau, \nu)$ replaced with $F_H(\tau, \alpha)$ and with the frequency shifts replaced by time-scaling operations.

The similarities of the models means they share similar problems. The main drawback is that if the parameters of the model change slightly a new implementation is required. As an example, it the tap locations in Figure 2.5 change for multiple basis functions, then new filters are required for each of basis functions affected. Similarly, if the frequencies or time-scaling factors of the bases change, then the basis functions need to change as well.

While for the BEM when the basis function changes this simply involves changing the basis frequency, for the wideband model this involves resampling the signal, which is a much more computationally intense process. To show that resampling is in fact equivalent to time-scaling, assume a signal $x(t)$ is sampled satisfying the Nyquist criterion, then it can be represented using discrete samples as

$$x[n] = x_s(nT_s), \quad \text{with} \quad n = -\infty, \ldots, 0, \ldots, \infty. \tag{2.37}$$

If the continuous time signal is then scaled by a factor, α, then the sampled time domain signal becomes

$$x(\alpha t) = x_s(\alpha n T_s). \tag{2.38}$$

This time-scaling factor can be applied to the sampling period, T_s. So for a constant value of α the new sampling period becomes

$$T_{s,new} = \alpha T_s, \tag{2.39}$$

which leads to a new sampling frequency of

$$F_{s,new} = \frac{1}{\alpha} F_s. \tag{2.40}$$

A signal can therefore be time-scaled by resampling at $F_{s,new}$ and then by running the resampled data at the original sampling times, n, and at the original sampling frequency, F_s. If the time domain signal is compressed, $\alpha > 1$ and $F_{s,new} < F_s$. This leads to fewer samples, which when played at the original sampling rate leads to time-compression. Conversely, if $\alpha < 1$ then the signal is dilated and $F_{s,new} > F_s$, leading

to more samples and time-dilation. Therefore, resampling the signal at $F_{s,new}$ and playing it back at the original sampling rate F_s does produce the desired time-scaling.

A common method used for resampling is to convert, or approximate the scaling factor, α, by a rational number, M/N. Resampling is then implemented by upsampling, or interpolating by M and then decimating the upsampled signal by N. This, however, is a fairly computationally expensive process if M and N are large values [28], which is often the case if the time-scaling factor is close to 1. The time-varying fractional delay line presents a second method that can be used to interpret the time-scaling operation.

2.3.2 The Time-Varying Fractional Delay Line

A second method that can be used to implement time-scaling is found by reinterpreting Eqn. (2.38). The time scaled signal can be rewritten as

$$x(\alpha t) = x_s((1 - (1 - \alpha))nT_s) = x_s(nT_s - (1 - \alpha)nT_s), \tag{2.41}$$

which shows that time-scaling can also be implemented by sampling the original signal with a time-varying delay. So the time-scaled signal can be written as

$$x(\alpha t) = x_s(nT_s - \tau[n]), \tag{2.42}$$

where

$$\tau[n] = (1 - \alpha)nT_s. \tag{2.43}$$

The effect of varying the delay with respect to time is shown in Figure 2.6.

There are two main benefits of working with a time-varying delay over a resampling operation. First, almost any sampling rate can be implemented with the same computational overhead since the time-varying delay, $\tau[n]$, will simply be different for different resampling rates. Second, a time-varying delay allows for time-varying time-scaling factors, so now α can be a function of time which mimics changes in velocity without having to implement a new resampling operation for each new velocity encountered.

In order to implement a time-varying delay using discrete samples, it is easiest to break the operation up into the application of a fractional delay and then the time-variation

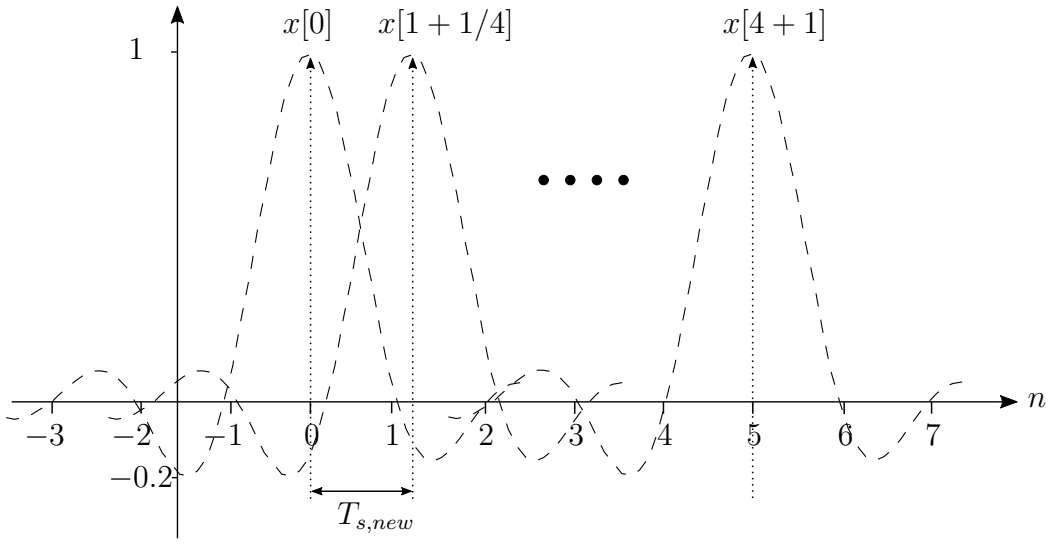


Figure 2.6: Example of a time-varying delay implementing a resampling operation. For $\alpha = 1.25$, the time-varying delay can be expressed as $\tau[n] = 0, -1/4, -1/2, -3/4, -1, ...$ for $n = 0, 1, 2, 3, 4....$ The time-varying delay changes the sampling frequency from F_s to $\frac{4}{5}F_s$, or changes the sampling period to $T_{s,new} = \frac{5}{4}T_s$.

of this fractional delay. A fractional delay is a delay that falls in between the sampling points of a sampled signal. A fractional delay can be realized by implementing a delayed ideal interpolation filter, or a sinc filter. The sampled, time-delayed signal then becomes [29]

$$x(t-\tau) = x_s(nT_s - \tau) = x[n - \tau_s] = \sum_{m=-\infty}^{\infty} \text{sinc}[m - \tau_s]x[n-m], \tag{2.44}$$

where τ_s is the normalized delay with respect to the sampling period, T_s. The normalized delay can be written as

$$\tau_s = \frac{\tau}{T_s} = D_{\tau_s} + d_{\tau_s}, \tag{2.45}$$

where D_{τ_s} is the integer part of the delay and d_{τ_s} is the fractional part.

The interpolation filter can be implemented using an FIR filter, however in order to practically implement this filter the sinc function needs to be truncated to reduce the order of the filter to a finite length. The order chosen will limit the accuracy of the truncation and the length becomes a trade off between precision and performance. If the length of the FIR filter is chosen as $2M + 1$, then higher values of M will lead to longer computational times, but more accurate results [28].

Due to this truncation, the sinc function will be subject to Gibbs phenomenon and therefore contain ripple in the passband and stopband. To reduce this effect a windowing function can be used on the time domain signal which puts more emphasis on the center of the time-domain FIR filter and less on the time-domain boundaries at $-M$ and $+M$. This reduces the ripple at the expense of reducing the passband width and increasing the transition region from the passband to the stopband [30].

By truncating and windowing the sinc filter as well as delaying it by the integer part of the delay, D_{τ_s}, the new fractionally delayed output signal can be expressed as

$$x[n-\tau_s] = \sum_{m=-M+D_{\tau_s}}^{M+D_{\tau_s}} w[m]\,\mathrm{sinc}[m-d_{\tau_s}]x[n-m], \tag{2.46}$$

where $w[m]$ is some desired windowing function. This filter can also be made causal by delaying the signal further. In the case that $D_{\tau_s} = 0$ the filter could be delayed by M samples and if $D_{\tau_s} > 0$ then the filter could be delayed by $M - D_{\tau_s}$.

The fractional delay filter can therefore be implemented using a tapped delay line, and if multiple fractional delays are required then the same tapped delay line can be used. A fractional delay line (FDL) implementing multiple fractional delays would correspond to a multipath channel with stationary non-integer cluster arrivals with respect to the sampling frequency. A simple 2-tap channel with fractional delays is shown in Figure 2.6 with cluster arrivals at $0.5T_s$ and $4.8T_s$.

If time-variance is required for the taps then the time-varying delays associated with the paths can be discretized and used in Eqn. (2.46). If this time-variance is caused by a constant velocity then $\tau_s[n]$ can be represented as shown in Eqn. (2.47). This can also be expressed in terms of a time-varying integer and fractional delay as

$$\tau_s[n] = \frac{(1-\alpha)nT_s}{T_s} = -\frac{v_c}{c_w}n = D_{\tau_s}[n] + d_{\tau_s}[n]. \tag{2.47}$$

This is a discrete delay that is represented by samples, and not continuous time. The integer delay is therefore represented by the integers and the fractional delay is always a real number between 0 and 1. The value of the fractional delay could be discritized to limit the number of fractional positions that need to be implemented, however for now it is assumed to be a real number.

If the velocity changes with time then this same technique can also be used with a

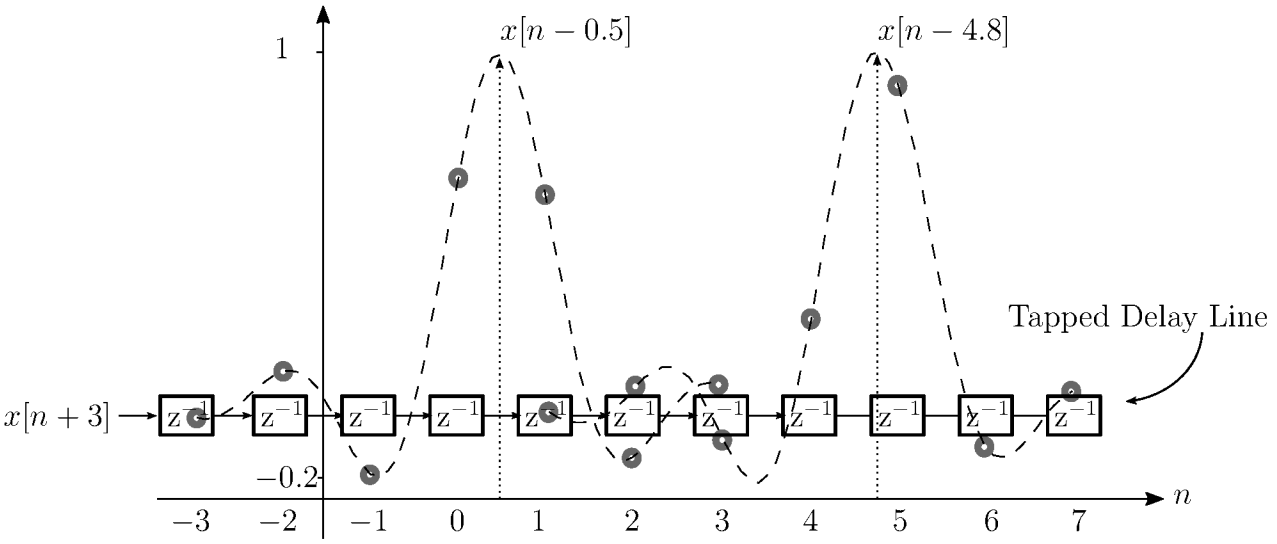


Figure 2.7: Example of a fractional delay line (FDL) implemented using a tapped delay line with two fractional delays at $\tau_s = 0.5$ and $\tau_s = 4.8$. Both taps use sinc filters with $M = 3$. The dashed line indicates the continuous sinc interpolation filter and the grey circles indicate the discrete, sampled points of the function. For the first tap, $D_{\tau_s} = 0$ and the sinc function samples from -3 to 3. For the second tap, $D_{\tau_s} = 4$ and the sinc function samples from 1 to 7. The input signal, $x[n]$, has been previously delayed by 3 samples to maintain causality. This figure shows the output for the fractional taps for a single instant in time, the next instant is calculated by advancing the delay line.

time-varying time-scaling term, $\alpha(t)$. The discrete time-varying delay then becomes

$$\tau_s[n] = \frac{(1-\alpha(nT_s))nT_s}{T_s} = -\frac{v_c(nT_s)}{c_w}n = D_{\tau_s}[n] + d_{\tau_s}[n] \tag{2.48}$$

where

$$\alpha(t) = 1 + \frac{v_c(t)}{c_w}. \tag{2.49}$$

Either of the two discrete time-varying delays above leads to the final equation for the time-varying fractional delay line (TVFDL)

$$x[n-\tau_s[n]] = \sum_{m=-M+D_{\tau_s}[n]}^{M+D_{\tau_s}[n]} w[m]\,\mathrm{sinc}[m-d_{\tau_s}[n]]x[n-m]. \tag{2.50}$$

This equation only applies for a single tap, or cluster arrival, at a single discrete time instance, n. This equation therefore needs to be repeated for each tap and for each

time instance corresponding to a new delay value. The benefit however remains that only a single delay line is required to model the time-scaling of all taps, for all times.

While this technique can be applied to resample the signal at any rate, care needs to be taken in choosing the initial sampling rate of the delay line, F_s, so that the sinc function used to resample the signal in Eqn. (2.50) will bandlimit properly for all desired resampling frequencies [28]. While the resampling rate can be changed by varying $\alpha(t)$, there should be some maximum velocity and therefore a maximum value of $\alpha(t)$, α_{max}. This leads to a minimum resampling rate of

$$F_{s,min} = \frac{F_s}{\alpha_{max}}. \tag{2.51}$$

Due to the fact that $F_{s,min} < F_s$, the highest frequency that can be included in the delay line will be dictated by the minimum sampling frequency. Or conversely, the sampling rate of the delay line will be dictated by assuring the maximum frequency included is less than $\frac{F_{s,min}}{2}$. So, if we assume a perfect brick-wall bandlimiting filter the sampling frequency can be set using

$$f_{max} \leq \frac{F_{s,1}}{2\alpha_{max}} \rightarrow F_{s,1} \geq 2f_{max}\alpha_{max}. \tag{2.52}$$

However, due to the fact that the bandlimiting, or interpolating filter, is truncated, this will not be a perfect brick-wall filter and so a second sampling frequency criterion can be set.

$$f_{max} \leq \frac{F_{s,2}}{2}\omega_n \rightarrow F_{s,2} \geq \frac{2f_{max}}{\omega_n}, \tag{2.53}$$

where ω_n is the normalized cutoff frequency of the truncated and windowed interpolating sinc function, with a value between 0 and 1.

So, the sampling frequency of the delay should be chosen as the maximum value of the two conditions above, or

$$F_s = \max\{F_{s,1}, F_{s,2}\}. \tag{2.54}$$

This condition means that if α_{max} is close to 1 and $F_{s,min} \approx F_s$, then F_s should be chosen so that f_{max} is in the passband of the truncated interpolating filter. If α_{max} is large, then F_s should be chosen so that f_{max} is always less than $\frac{F_{s,min}}{2}$.

Although the exact implementation will not be discussed here, a computationally efficient way to implement the delayed sinc functions is using a lookup table (LUT) [31], [28]. This is the reason why F_s is chosen to satisfy the conditions above, rather than just scaling the sinc interpolating function to bandlimit as required. Specifically, in [29], [31], and [28] it is suggested that the sinc function should be broadened when $F_{s,resample} < F_s$ to narrow the frequency range of the filter. However, if a variable sampling rate is required this means that time-scaled versions of the interpolating function would have to be created and stored for all $F_{s,resample} < F_s$. Therefore it's much easier in the variable sampling rate case to choose F_s according to (2.54) rather than to scale the sinc function.

2.4 Channel Simulation using Time-Varying Fractional Delay Line

In this Section, simulations of the TVFDL are presented. One of the appealing aspects of the TVFDL is that it can resample a signal both at a set resampling rate and also at a variable resampling rate. This is due to the fact that the TVFDL solely relies on the delays associated with a path and how they vary, and not specifically on a resampling rate. In order to show that the TVFDL is capable of implementing a wideband channel model three simulations were run.

The first simulation that was completed shows that the TVFDL is capable of resampling a signal. To do this, a reference signal was produced and then this signal was resampled two ways. The first resampling method was using a classic M/N interpolation/decimation scheme and the second way was using a TVFDL with an equivalent value of α. More specifically, α was chosen as

$$\alpha = \frac{N}{M}. \tag{2.55}$$

This value of α was then used with Eqn. (2.47) to create the required time varying delay to implement the desired resampling rate. The plot in Figure 2.8 shows a simple cosine wave with a frequency of 100 Hz sampled at $F_s = 1$ kHz, and resampled at a rate of $\frac{M}{N} = \frac{3}{2}$, or with $\alpha = \frac{2}{3}$.

Figure 2.8 shows that the TVFDL produces very similar results to those obtained using the conventional M/N method, which are both very similar to the desired resampled

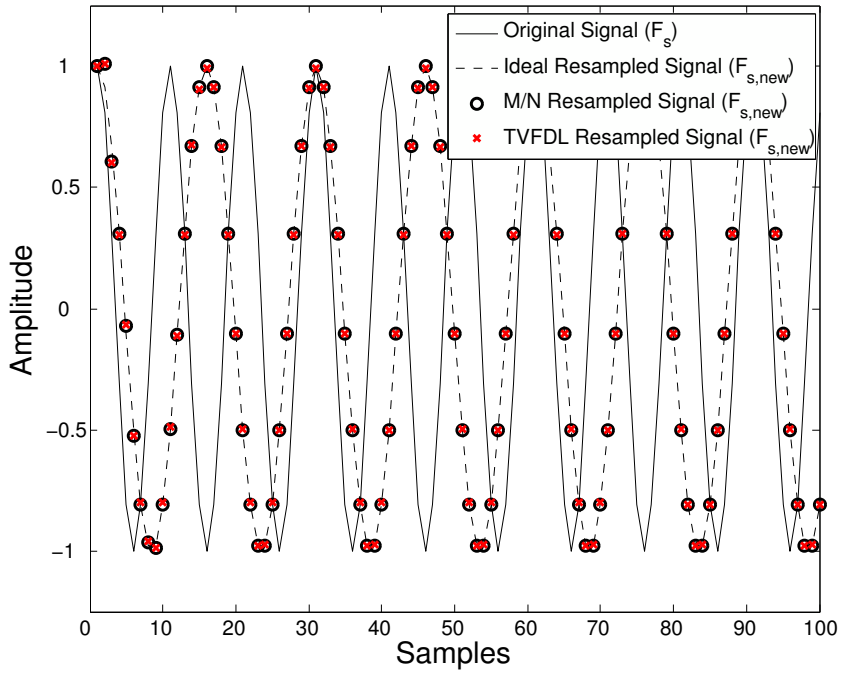


Figure 2.8: Plot of simple cosine wave resampled at $F_s = \frac{3}{2}F_s$ using both the classic M/N technique and a TVFDL, the ideal resampled response is also shown to compare with resampled values.

signal. The benefit of the the TVFDL however is that it can vary the resampling rate simply by varying α with respect to time, while changing the resampling rate of the M/N method involves a completely new implementation.

In order to determine the ability of the TVFDL to implement a simulated wide-band channel with time-variance, first a deterministic channel was created that had 4 known clusters with constant amplitudes, delays and time-scaling factors. Table 2.1 below shows these values for each cluster arrival.

Table 2.1: Cluster parameters for the the deterministic channel implemented using a TVFDL.

Cluster Arrival	**Amplitude**	**Delay (ms)**	**Time-Scaling Factor (α)**
1	1	0	1.0054
2	0.5	8	1.0063
3	0.3	15	1.0058
4	0.4	35	1.0067

In order to determine whether the TVFDL was implementing the channel properly, a channel sounder was implemented. A regular correlative channel sounder cannot be used for this channel as the time-scaled signals will not correlate well with the original

transmitted signal. Instead, the time-varying channel sounder presented in Appendix B was used. This channel sounder uses a bank of filters created by time-scaling the original signal, and then combines their outputs to produce the final CIR.

For this channel sounder the PN sequence chip rate, R_{chip}, was chosen to be 12 kHz, the length of the signal, T_{seq}, was chosen to be 0.1 seconds, and the carrier frequency was chosen to be 10 kHz. The final passband channel that was being estimated then was between 4 and 16 kHz. 50 PN sequences were concatenated together to form the final pass band transmit signal that had a length of 5 seconds. This signal was then used as the input to the deterministic TVFDL channel.

The output of the deterministic channel model was passed into the time-varying channel sounder to produce the output CIR's for each transmitted PN sequence. These PN sequences were then stacked on top of each other to produce a delay-time plot showing the evolution of the CIR with respect to time. The delay-time plot for the deterministic channel is shown in Figure 2.9. The red lines that overlay the CIR are the ideal delay paths produced using the channel parameters. The match between the channel sounder output and the ideal paths show that the TVFDL is capable of producing a channel with the desired time-varying delays.

While the path delays in Figure 2.9 do match the desired time-varying delays, the amplitudes seem to fluctuate slightly almost as if there is some fading rather than the desired constant amplitudes. The reason for this is due to the channel sounder itself and not the TVFDL. When the signal does not fall fully into one delay bin, it is spread out over neighboring ones, causing a split in energy. These splits then cause the amplitude to appear lower even though between the neighboring taps the energy should be the same. A proper averaging scheme should remove this phenomenon and show the proper tap amplitudes.

A second simulation was also run that implemented the stochastic model presented in Section 2.2.2 using the TVFDL. For this model, the following parameters were chosen: $N_c = 6$, $N_{p,c} = 10$, $\tau_{RMS} = 0.03$ s, $\tau_{max} = 0.06$ s, $T_c = 2$ s, $v_{max} = 10$ m/s, $v_{min} = 8$ m/s, and $\theta_{max} = \frac{\pi}{4}$ radians. The same correlative sounding signal used for the deterministic channel was used again with the stochastic model, and the time-delay plot produced using the time-varying channel sounder is shown in Figure 2.10.

While Figure 2.10 will also contain some of the minor amplitude fading due to the channel sounder, this plot clearly shows that the stochastic model implements time-

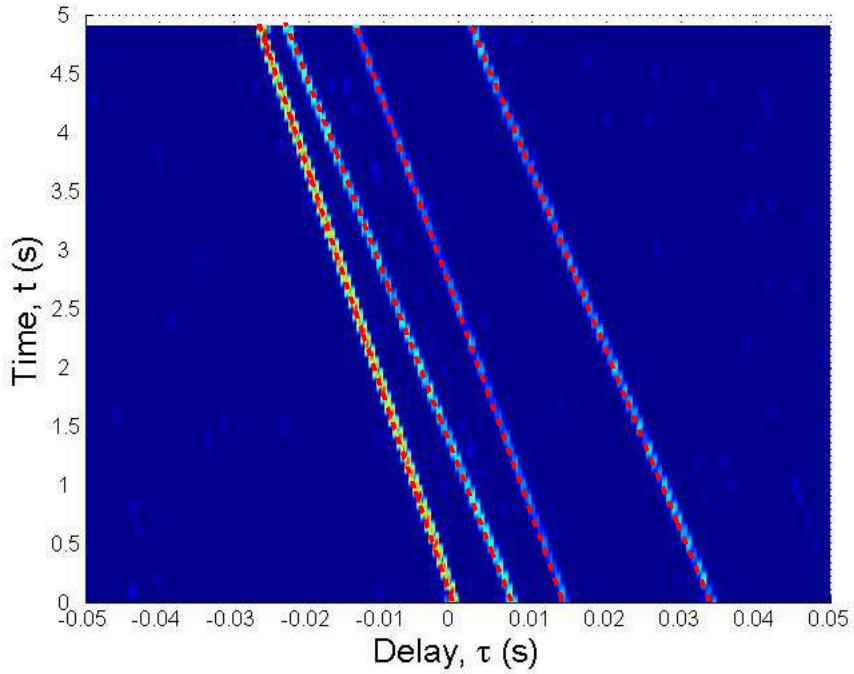


Figure 2.9: Output of the time-varying channel sounder for the deterministic channel implemented using the TVFDL with clusters defined in Table 2.1. The red dotted lines indicate the ideal path for each tap. The underlying channel impulse response is the normalized power of each tap. Dark blue indicates 0, dark red indicates 1.

fading in excess of the slight pseudo fading introduced by the channel sounder itself.

The last benefit of the TVFDL is that it is able to implement time-varying scale values. To show this, the last simulation that was run implemented a single cluster path that had a simple sinusoidal delay path, this corresponds to a time-varying time-scaling factor of

$$\alpha(t) = (1 + A_s \sin(2\pi f_s t)) \,. \tag{2.56}$$

For this simulation $A_s = 0.01$ and $f_s = 1/10$ and the cluster path was given a constant amplitude of 1. The same transmit channel sounding signal used above was again used here, however the length of the signal was extended to 10 seconds to account for a full period of the wave. The output CIR calculated using the time-varying channel sounder is shown in Figure 2.11.

So, while a time-varying fractional delay line has been proposed in the past for efficient

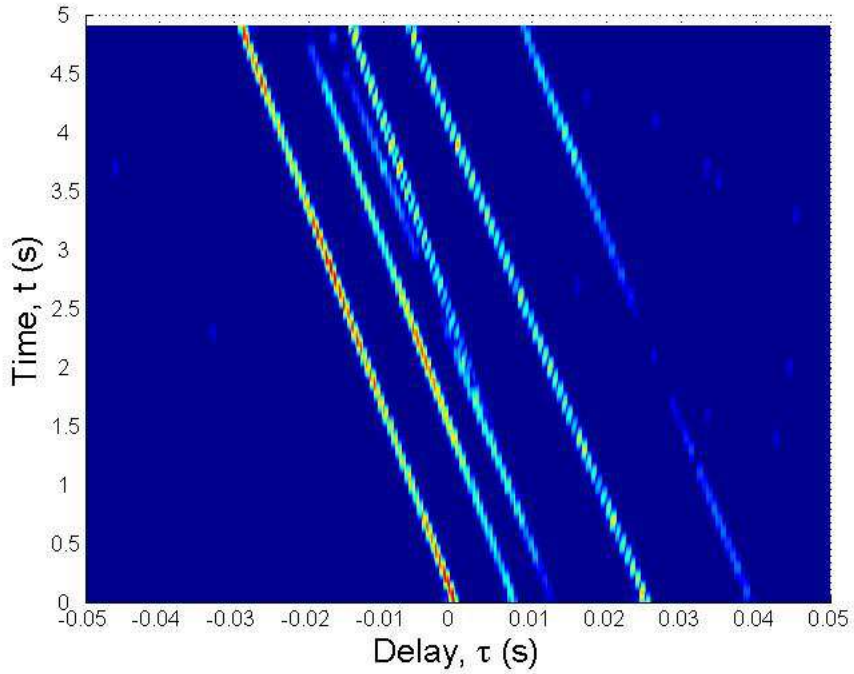


Figure 2.10: CIR of a time-varying channel produced using a TVFDL and channel parameters produced using the stochastic model presented in Section 2.2.2. This plot is again normalized to a power of 1. Dark blue indicates 0, dark red indicates 1.

resampling when the resampling rate is close to 1, [28], [29], it has not been used to implement a time-varying wideband channel model. The simulations above show that this structure is capable of implementing time-variance of the delay and amplitude associated with each cluster arrival, this makes the TVFDL a plausible technique for modeling a wideband channel. The ability to implement time-varying scaling factors also makes the TVFDL unique in that in can model changes in velocity, or acceleration, in a channel without having to change the model. This is a unique characteristic of the TVFDL that the author has not found in other channel models.

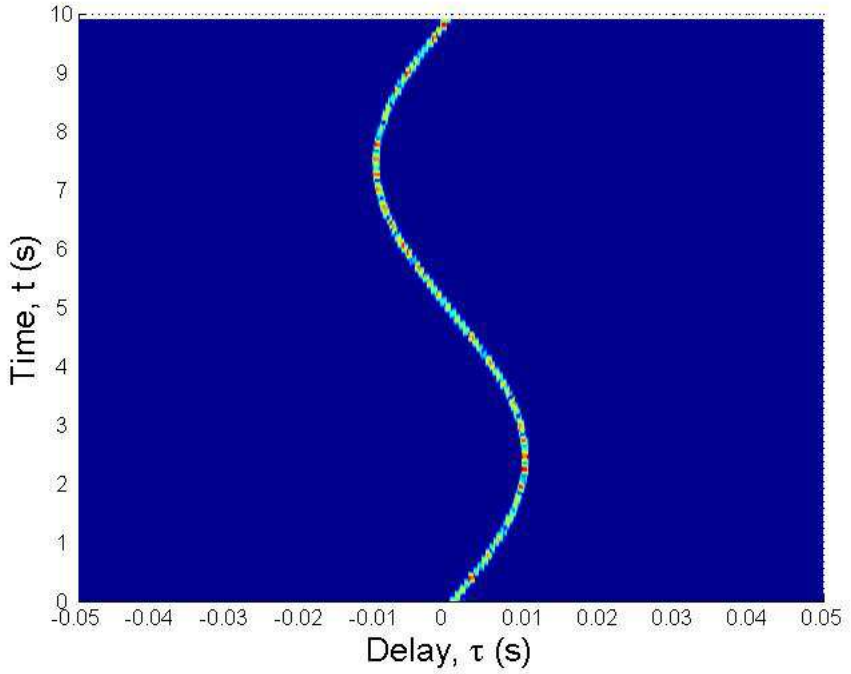


Figure 2.11: CIR of a time-varying channel produced using a sinusoidal time-scaling factor. This plot is also normalized to a power of 1. Dark blue indicates 0, dark red indicates 1.

Chapter 3

Wideband Beamforming Techniques

In Chapter 2 it was shown that the underwater acoustic channel introduces angular and temporal variations to a signal. Specifically, multipath components arrive from different angles and at different times with respect to the direct path arrival from the transmitter. If the angular characteristics of the impinging acoustic field can be captured then the angle-of-arrival can be recorded for each multipath component of the received signal.

This leads to a new source of diversity in addition to temporal diversity, angular diversity, which can be used to improve an underwater communication system as will be discussed in Chapter 4. One of the difficulties with trying to capture angular diversity in underwater communication systems is that most communications signals will be wideband in nature, as discussed in Chapter 2. This means a wideband beamformer is required. While classic techniques could be used to beamform the received data, there are some drawbacks to these techniques. Instead, wideband modal beamforming is presented as a unique technique that uses a uniform concentric circular array (UCCA) in order to record the angular diversity of the signal with 360° of unambiguous coverage.

This Chapter is broken down as follows. Section 3.1 introduces the geometry of the UCCA and looks at how both narrowband and wideband beamforming can be accomplished. Section 3.2 briefly reviews some well-known methods of wideband beamforming and then discusses why these techniques are not as suitable for underwater communication systems. Section 3.3 introduces modal beamforming and discusses

how a wideband modal beamformer can be implemented using both complex weights, for ease of mathematical analysis, and real weights for physically beamforming real received signals. This Section also discusses the physical limitations of a modal beamformer and discusses how modal weights can be designed for low order systems using Cardiod beamforming. Section 3.4 concludes the Chapter by presenting results from an underwater experiment carried out with a modal beamformer that demonstrates the desired wideband characteristics and Cardiod beam patterns.

3.1 Beamforming Theory

In this Section wideband beamforming theory is reviewed. Section 3.1.1 introduces the UCCA and discusses how the elemental excitation due to a plane wave can be modeled. Section 3.1.2 then presents the basics of narrowband beamforming and Section 3.1.3 builds upon this by introducing a method to beamform wideband signals. In Section 3.1.4, figures of merits for beamformers are briefly discussed in order to evaluate the performance of different beamforming techniques.

3.1.1 The Uniform Concentric Circular Array

Figure 3.1 shows a uniform concentric circular array (UCCA). Compared to a linear array, a circular array is attractive because it provides 360° of unambiguous coverage in the azimuthal plane as opposed to just 180°. A circular array also allows for non-conventional beamforming techniques that will be discussed in Section 3.3.

A UCCA is made up of $M+1$ elements, where M elements are uniformly spaced on the circumference of a circle of radius r, referred to as the circumferential elements, and there is a single element located in the center of the circle. Assuming that the array is located in the x-y plane, the angular spacing between each circumferential element is limited to an azimuthal angle and is described by

$$\phi_e = \frac{2\pi}{M}, \tag{3.1}$$

which corresponds to a circumferential element spacing of

$$d_e = r\phi_e = r\frac{2\pi}{M}. \tag{3.2}$$

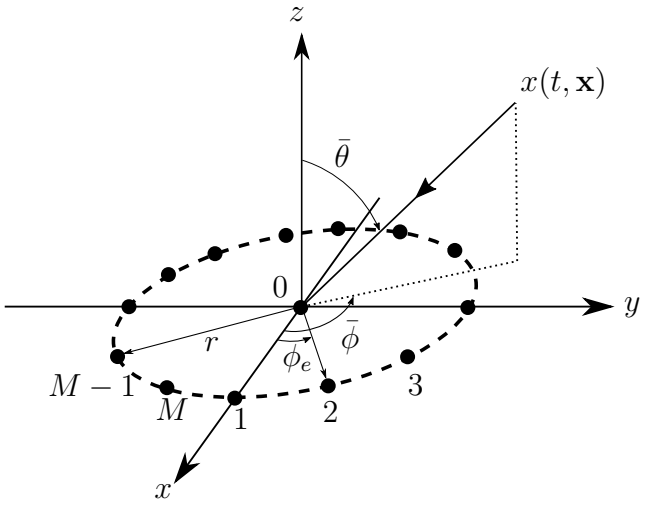


Figure 3.1: Layout of a UCCA located on the x-y plane. Spherical coordinates are used to describe angular positions with the polar angle measured off the z-axis and the azimuthal angle measured off the x-axis.

This distance is chosen to be less than $\frac{\lambda_{min}}{2}$ in order to conform to conventional narrowband beamformer design techniques [32]. Therefore, the radius is effectively set by the number of elements and the desired maximum frequency. The reason for this spacing in circular arrays will be discussed in more detail in Section 3.3.

The angular position of each element on the circumference of the circle then becomes

$$\phi_m = (m-1)\phi_e = (m-1)\frac{2\pi}{M}, \quad m = 1, \ldots, M. \tag{3.3}$$

The center element, $m = 0$, does not have an angular description and so is excluded.

The position of each of these elements can be described in cartesian coordinates using a vector from the origin of the x-y plane, $\mathbf{r}_m$, and the element numbering convention shown in Figure 3.1. Again, this does not apply to the center element that has $\mathbf{r}_0 = \mathbf{0}$.

$$\mathbf{r}_m = \begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix} = r \begin{bmatrix} \cos(\phi_m)\sin(\theta) \\ \sin(\phi_m)\sin(\theta) \\ \cos(\theta) \end{bmatrix} = r \begin{bmatrix} \cos(\phi_m) \\ \sin(\phi_m) \\ 0 \end{bmatrix}, \quad m = 1, \ldots, M. \tag{3.4}$$

If a plane wave arrives from a direction $(\bar{\phi}, \bar{\theta})$, as shown in Figure 3.1, with a frequency, ω, and spatial frequency vector, $\mathbf{k}$, then the signal that is recorded by an element in space depends on the position of the element and the time of the recording, $\mathbf{x}$ and t,

respectively.

$$x(t, \mathbf{x}) = e^{j(\omega t + \mathbf{k}^T \mathbf{x})}, \tag{3.5}$$

where

$$\mathbf{k} = k \begin{bmatrix} \cos(\bar{\phi}) \sin(\bar{\theta}) \\ \sin(\bar{\phi}) \sin(\bar{\theta}) \\ \cos(\bar{\theta}) \end{bmatrix} = \frac{\omega}{c} \begin{bmatrix} \cos(\bar{\phi}) \sin(\bar{\theta}) \\ \sin(\bar{\phi}) \sin(\bar{\theta}) \\ \cos(\bar{\theta}) \end{bmatrix}, \tag{3.6}$$

and k is the wave number of the plane wave.

When the plane wave impinges on the array at a time, t, each element in the UCCA receives the signal with a different phase delay. The signal received on array element m, $x_m(t)$, can be described using the element position vector, $\mathbf{r}_m$.

$$x_m(t) = e^{j(\omega t + \mathbf{k}^T \mathbf{r}_m)} = e^{j\omega t} e^{j\mathbf{k}^T \mathbf{r}_m} \tag{3.7}$$

The signal on each element contains the time varying term $e^{j\omega t}$. For notational simplicity this term will be excluded for the rest of this Chapter and only the phase portion of the term will be considered. For the circumferential elements, the phase term then simplifies to

$$x_m = e^{jkr \sin(\bar{\theta}) \cos(\bar{\phi} - \phi_m)}, \quad m = 1, \ldots, M, \tag{3.8}$$

and for the center element there is no phase shift, $x_0 = 1$.

Instead of defining the phase delay on the circumferential elements in terms of the wave number, k, and radius, r, as done in Eqn. (3.8), a circular beamformer can also be described in terms of the minimum desired wavelength, λ_{min}, and the number of circumferential elements, M [32]. This notation is helpful when the time domain is discretized. The radius can be rewritten in terms of λ_{min} and M as

$$\lambda_{min} = 2d_e = 2\frac{2\pi r}{M} \quad \rightarrow \quad r = \frac{\lambda_{min}}{2} \frac{M}{2\pi}. \tag{3.9}$$

Using Eqn. (3.9), the phase delay on each circumferential array element can be written

as

$$x_m = e^{j\frac{\omega}{c}\frac{\lambda_{min}}{2}\frac{M}{2\pi}\sin(\bar{\theta})\cos(\bar{\phi}-\phi_m)}, \quad m = 1, \ldots, M. \tag{3.10}$$

This equation can be simplified further if each array element is sampled in time using the Nyquist criterion with respect the maximum frequency of the beamformer, $\frac{c}{\lambda_{min}}$. If the sampling period is chosen to be $T_s = \frac{\lambda_{min}}{2c}$, then the phase delay on each circumferential element can be written as

$$x_m = e^{j\Omega\beta\sin(\bar{\theta})\cos(\bar{\phi}-\phi_m)}, \quad m = 1, \ldots, M, \tag{3.11}$$

where $\beta = \frac{M}{2\pi}$ and $\Omega = \omega T_s$, which is the normalized frequency with a range of $[-\pi, \pi]$. This choice of T_s means the maximum normalized frequency, π, is the maximum frequency for both the temporal and spatial recording.

3.1.2 Narrowband Beamforming

In order to determine the response of the array to a plane wave with a single frequency component, or the narrowband case, the weighted signals on each receive element need to be summed, as shown in Figure 3.2.

$$y = \sum_{m=0}^{M} x_m w_m^* = w_0^* + \sum_{m=1}^{M} e^{j\Omega\beta\cos(\bar{\phi}-\phi_m)} w_m^*. \tag{3.12}$$

It is often convenient to express the response of the beamformer to a signal arriving from a direction, $\bar{\phi}$, and with a frequency, ω, using vector multiplication. To do this, two length M vectors need to be created. The first is a steering vector, $\mathbf{d}(\omega, \bar{\phi})$, which models the phase delay of the signal recorded on each array element. The second is the weight vector, $\mathbf{w}$, which stores the weights applied to the signals on each array element [32]. The response of a beamformer when steered in a direction $\bar{\phi}$ can then be expressed as

$$P(\Omega, \phi) = \mathbf{w}^H \mathbf{d}(\Omega, \bar{\phi}), \tag{3.13}$$

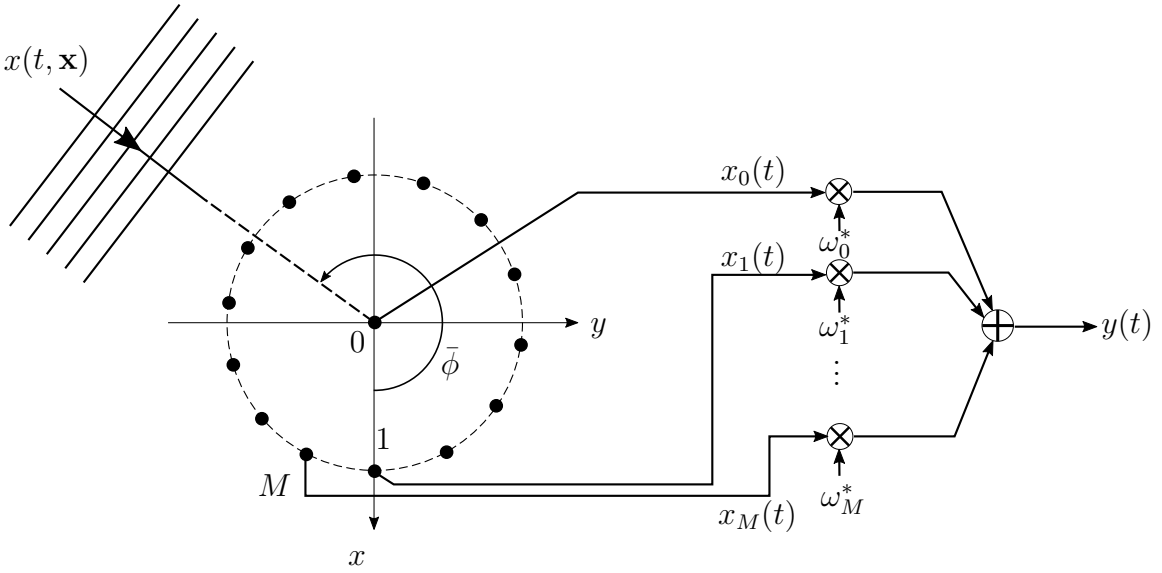


Figure 3.2: Narrowband beamforming structure for the UCCA with a plane wave arriving from $(\bar{\phi}, \frac{\pi}{2})$.

where

$$\mathbf{w} = [w_0 \quad w_1 \dots w_M]^T \tag{3.14}$$

and

$$\mathbf{d}(\omega, \bar{\phi}) = [1 \quad e^{j\Omega\beta\cos(\bar{\phi}-\phi_1)} \dots e^{j\Omega\beta\cos(\bar{\phi}-\phi_M)}]^T \tag{3.15}$$

The beamformer output due to a plane wave with a normalized frequency, Ω, and DOA, $\bar{\phi}$, can now be written in terms of $\mathbf{w}$ and $\mathbf{d}(\Omega, \bar{\phi})$.

$$y = xP(\Omega, \bar{\phi}) = x\mathbf{w}^H\mathbf{d}(\Omega, \bar{\phi}), \tag{3.16}$$

where x is the transmitted signal and y is the beamformed output at a single instant in time.

Therefore, in order to maximize gain in a direction of interest and minimize interference from others, a suitable weight vector needs to be created which implements a desired beam pattern.

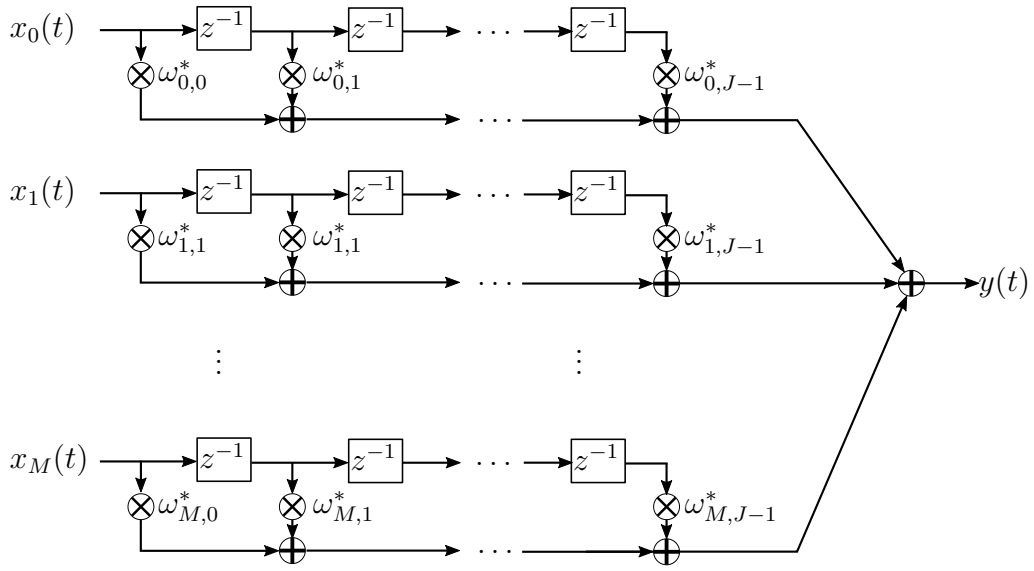


Figure 3.3: Wideband beamforming structure with FIR filters of length J attached to each array element output.

3.1.3 Wideband Beamforming

If the signal that is transmitted is narrowband, or approximately narrowband, then the difference between the signals that arrive on different array elements can be approximated by a phase shift [32]. This phase shift can then be compensated by a single complex weight applied to each element. However, if the bandwidth of the signal is too large, then this assumption no longer holds and different weights need to be applied for different frequencies in order to avoid distortion of the signal.

One method to create frequency dependent weights is to apply time-domain filters at the output of each array element. This filter equalizes the signal in the frequency domain and can be designed in order to maintain a desired beam pattern over the frequency range of interest. In this Chapter we focus on finite impulse response (FIR) filters due to their stability and linear phase response. The new beamforming structure is shown in Figure 3.3 and is often referred to as a filter-and-sum beamformer.

Similar to the narrowband case, the response of the beamformer to a signal arriving from a direction, $(\bar{\phi}, \bar{\theta})$, and at a frequency, ω, can be computed by summing each weighted element. However, with an FIR filter attached to the output of each element in the array, the response also has to be summed over the J delay elements associated

with each FIR filter.

$$y(t) = \sum_{m=0}^{M} \sum_{i=0}^{J-1} x_m(t - iT_s) w_{m,i}^*, \tag{3.17}$$

where the reference to time has been included in order to describe the time delay operation performed by the FIR filter.

For a plane wave impinging on a wideband circular array, the received signal on each circumferential element becomes

$$x_m(t - iT_S) = e^{j\omega(t-iT_s)} e^{j\mathbf{k}^T \mathbf{r}_m} = e^{j\omega t} e^{j(kr\sin(\bar{\theta})\cos(\bar{\phi}-\phi_m)-\omega i T_s)}. \tag{3.18}$$

Again, the time varying term is dropped and it is assumed that the signal only arrives from the x-y plane, $\bar{\theta} = \frac{\pi}{2}$. Eqn. (3.18) then simplifies to

$$x_{m,i} = e^{j(\Omega\beta\cos(\bar{\phi}-\phi_m)-i\Omega)}. \tag{3.19}$$

The response of the wideband beamformer to a plane wave can then again be expressed conveniently using the vector multiplication in Eqn. (3.13). However, the steering and weight vectors need to be modified to include the effects of the FIR filter and so the length of each increases to $(M+1)J$ elements for the UCCA. To do this, the weight vector is redefined as

$$\mathbf{w} = \begin{bmatrix} \mathbf{w}_0 & \mathbf{w}_1 \dots \mathbf{w}_{J-1} \end{bmatrix}^T, \tag{3.20}$$

where each internal weight vector, $\mathbf{w}_i$, contains the weights applied to the i^{th} FIR tap for elements 0 to M,

$$\mathbf{w}_i = [w_{0,i} \quad w_{1,i} \dots w_{M,i}]. \tag{3.21}$$

The steering vector also needs to be redefined to account for the phase delays due to the element spacing and due to the FIR time delays,

$$\mathbf{d}(\Omega, \bar{\phi}) = [1 \dots e^{j(\Omega\beta\cos(\bar{\phi}-\phi_M))} \quad e^{-j\Omega} \dots e^{j(\Omega\beta\cos(\bar{\phi}-\phi_M)-\Omega)} \tag{3.22}$$

$$e^{-j(J-1)\Omega} \dots e^{j(\Omega\beta\cos(\bar{\phi}-\phi_M)-(j-1)\Omega)}\Big]^T. \tag{3.23}$$

With these new vectors, the response of the wideband beamformer can now be com-

puted. The next Sections deal with how these weights can be designed in order to attain desired wideband, or frequency invariant, beam patterns.

3.1.4 Beamformer Comparison Criterion

Before methods to design wideband beamformers are discussed, it is important to highlight the criterion commonly used to describe their performance. A common beamformer criterion is the beam pattern, which is an analysis tool used to evaluate the beamformer response in the azimuthal plane as a function of the angle-of-arrival of the signal. For narrowband beamformers this is done over a very narrow band of frequencies, and so generally a single, central frequency can be used to estimate the response of the beamformer. For wideband beamformers, this response is calculated over a range of frequencies in order to determine the frequency response of the beamformer.

For a UCCA the beam pattern can be calculated by discretizing the angle of arrival, ϕ, between $[-\pi, \pi]$ using N_ϕ points, and by discretizing the normalized frequency of the signal, Ω, between $[\Omega_{min}, \Omega_{max}]$ using N_Ω points. The normalized frequencies Ω_{min} and Ω_{max} are limited to minimum and maximum values of 0 and π. The frequency dependent beam pattern of a weighted array can then be expressed as an $N_\Omega \times N_\phi$ matrix created using

$$\begin{aligned} \mathbf{P}_{N_\Omega \times N_\phi}(i,k) = P(\Omega_i, \phi_k) = \mathbf{w}^H \mathbf{d}(\Omega_i, \phi_k) \text{ where} \\ i = 0, \ldots, N_\Omega - 1 \text{ and } k = 0, \ldots, N_\phi - 1. \end{aligned} \tag{3.24}$$

For the wideband case the steering and weight vector are calculated as described in Section 3.1.3. Often, this beam pattern is displayed in terms of the power of the received signal, and is normalized so the gain is 0 dB at the main beam. An example of a narrowband beam pattern is shown in Figure 3.4.

Figure 3.4 shows two main criterion for a narrowband beamformer design, the 3 dB beamwidth and the sidelobe level. The 3 dB beamwidth is the width of the main beam in the steering direction, $\bar{\phi}$, of interest. The sidelobe level is a measure of the highest power received from a sidelobe. Often in beamformer design it is desirable to have a narrow 3 dB beamwidth and a low sidelobe level.

For a wideband beamformer it is also important to determine how these values change

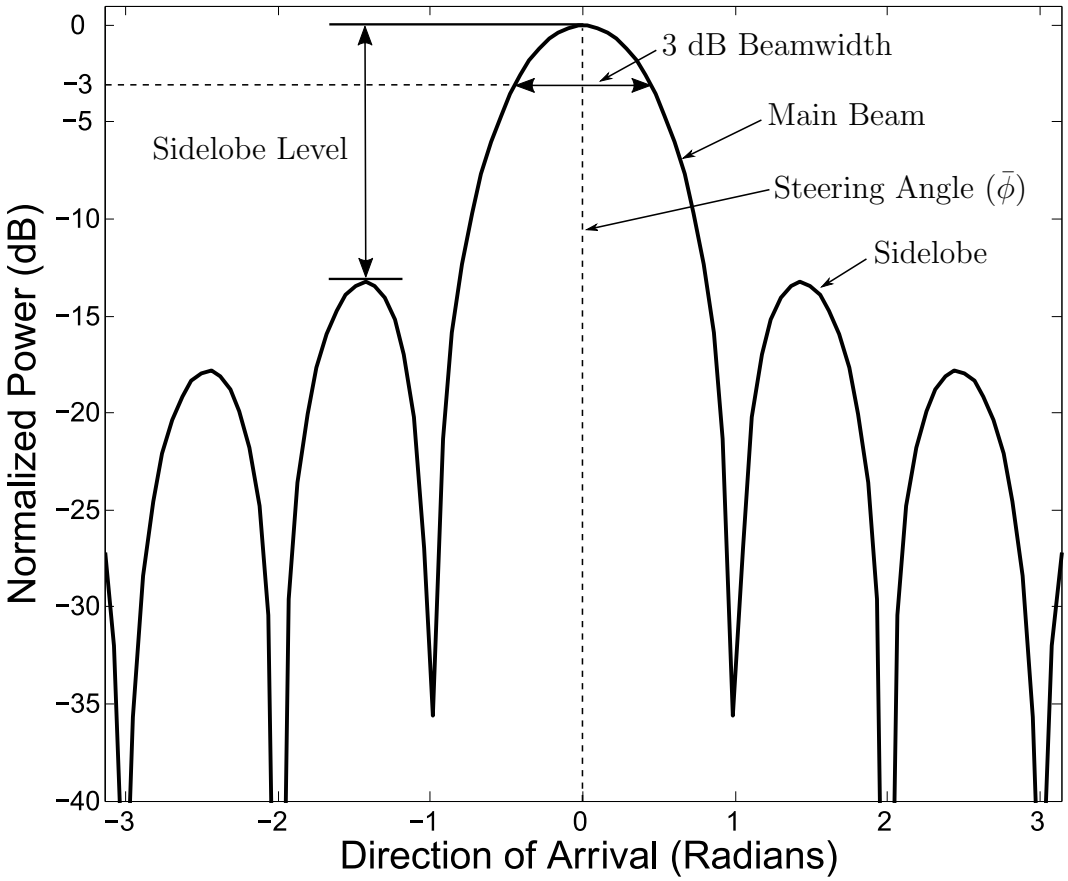


Figure 3.4: Narrowband beam pattern showing common beam pattern characteristic names and design criterion.

with frequency. Specifically, for a wideband beamformer, the beam pattern should remain the same, or be frequency invariant, over some bandwidth, B_{BF}. This implies that over this bandwidth the steering angle, the 3 dB bandwidth, and the sidelobe level and position should all remain constant, or very close to constant. While no quantitative measure is presented to determine B_{BF}, qualitative analysis of the wideband beam pattern provides good insight into this parameter, and in most cases theoretical upper and lower frequency bounds are presented depending on the method used to design the beamformer.

3.2 Review of Classic Wideband Beamformer Design Techniques

There are a number of non-adaptive beamforming techniques that can be used to design the FIR filters attached to each array element in order to produce a desired wideband beampattern. The two filter-and-sum techniques discussed briefly here are delay-and-sum and least-squares beamforming. In this Section the two techniques are only briefly reviewed, Appendix C and D contain more detailed descriptions of each.

Delay-and-sum beamforming is a wideband technique that makes use of the fact that the phase shift experienced by a signal between array elements is frequency dependent, but the time delay is constant as a function of frequency, as explained in [33]. Specifically, the phase-shift, Φ, experienced by a single frequency, ω, on the m^{th} sensor at a distance r from the origin can be described using

$$\Phi = kr\cos(\bar{\phi} - \phi_m) = \frac{\omega}{c} r\cos(\bar{\phi} - \phi_m), \quad m = 1, \ldots, M. \tag{3.25}$$

If this same delay is instead represented using a time delay,

$$x_m(t) = x(t + \tau_m(\bar{\phi})), \quad m = 1, \ldots, M, \tag{3.26}$$

then the time delay will be constant for all frequencies as

$$\tau_m(\bar{\phi}) = \frac{r}{c}\cos(\bar{\phi} - \phi_m). \tag{3.27}$$

which only depends on radius from the element to the origin and the speed of propagation and not on frequency. This does however assume a non-dispersive medium where the speed of propagation remains constant for all frequencies.

Delay-and-sum beamforming simply reverses the time-delay on a specific element caused by a signal coming from a desired direction, $\bar{\phi}$, and then sums the delay compensated element outputs together. This produces directionality as only signals coming from $\bar{\phi}$ will sum in perfect coherence, while signals coming from other directions will sum at-least partially incoherently [34]. This is discussed in more detail in Appendix C.

The main drawback of using delay-and-sum beamforming is that the pattern is not frequency invariant. As the frequency increase the main beam width decreases and the

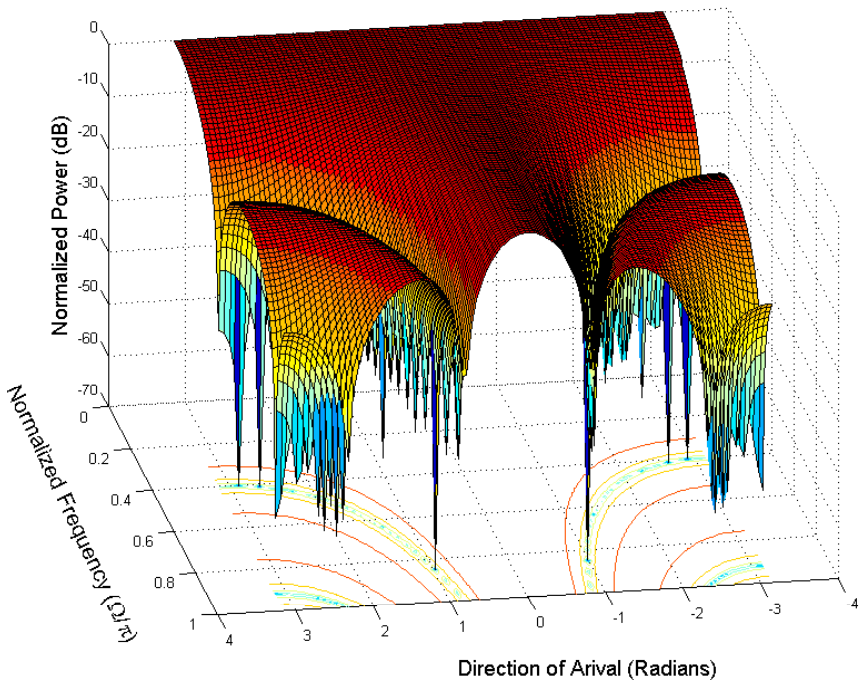


Figure 3.5: Ideal response of the Delay-and-Sum beamformer with $M = 7$ and a steering direction of $\bar{\phi} = 0$.

number of sidelobes increases. The beam pattern for a UCCA with 7 elements is shown in Figure 3.5 to demonstrate this. A wideband signal transmitted through a multipath channel could therefore be distorted through the delay-and-sum beamforming process.

The least-squares (LS) technique creates the weights by first specifying a desired beam pattern, and then selecting the array weights such that the actual beam pattern minimizes the squared error between the actual and desired beam pattern. More details about the LS technique can be found in Appendix D. An example beam pattern produced using the LS approach with the same 7 element UCCA used for the delay-and-sum beamformer is shown in Figure 3.6. The desired beam pattern for this plot was a simple delta function at $\phi = 0$ for all frequencies.

While the LS approach produces a frequency independent pattern, one of the main drawbacks of this technique when applied to the UCCA is that as the desired steering angle changes, the beam pattern changes as well. This is demonstrated in Figure D.3 in Appendix D by showing that the beam pattern produced by the same UCCA when steered to $\bar{\phi} = \frac{\pi}{6}$ has a much wider main beam with fewer, and more powerful sidelobes.

The non frequency-invariant beam pattern of the delay-and-sum beamformer and the

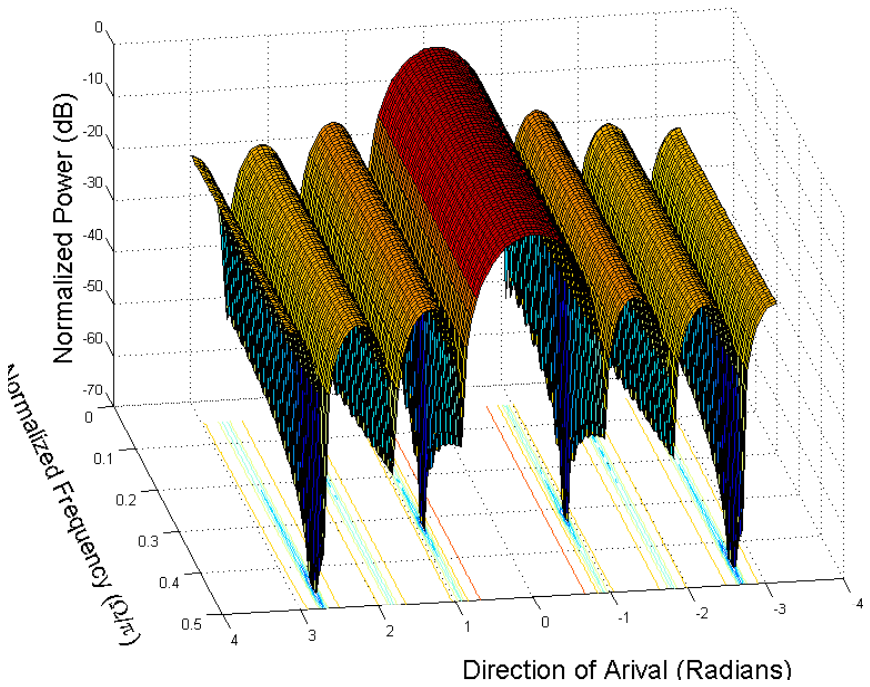


Figure 3.6: Beam pattern of the 7-element UCCA beamformer with FIR filters produced using a least-squares solution with $\bar{\phi} = 0$.

inability of the LS beamformer to maintain a constant beam pattern for all steering angles makes these two techniques less attractive wideband beamforming solutions for the UCCA. The other main drawback for the two techniques is that they require new FIR filters for each array element for each look direction of interest.

While for fixed beamforming with only a few steering angles of interest the need for multiple sets of wideband filters may not be a problem, for the directional estimation problem presented in Chapter 4, this could present a problem as each desired direction of arrival would require a wideband beamformer steering in that direction. In order to have a fine directional resolution this could mean implementing a number of beamformers. Modal beamforming provides a solution to this problem as wideband, frequency invariant beams steered in any direction can be produced easily. Modal beamforming is the focus of the next Section.

3.3 Modal Beamforming

Modal beamforming differs from the delay-and-sum and least-squares beamformers discussed in Section 3.2 in one important sense. Instead of using the filter-and-sum

techniques, a modal beamformer combines different weighted elements to first form modes. Each mode is an element in an orthogonal set, and once they are created a beamformed output can be produced by weighting and then summing the different mode outputs. This step of first producing modes then beamforming means that if each mode can be made frequency invariant, then a wideband beamformed output can be produced using only a single weight applied to each mode.

This Section first introduces phase modes in Section 3.3.1 as a possible mode that could be used with a circular apertures and discusses the limitations of using a sampled circular aperture for phase mode beamforming. Section 3.3.2 then introduces amplitude mode processing as the alternative to phase mode processing that can be used to beamform real signals and discusses how wideband amplitude modes can be implemented on a UCCA. Section 3.3.3 concludes by introducing cardiod beamforming as a method of designing the modal weights in order to produce a wideband beamformed output.

3.3.1 Phase Mode Processing

Phase mode processing is presented in this Section as an introduction into modal beamforming using complex modes. These modes are easier to work with and help to develop some of the basic theory of modal beamforming. First, the Fourier series is introduced as the orthogonal bases that can be used as the modes in a modal beamformer when a continuous circular aperture is used as the receiving element. The production of a mode using a continuous circular aperture and a sampled circular aperture, the UCCA, is then discussed and the disadvantages of mode production using a circular and sampled circular array are introduced.

3.3.1.1 The Complex Fourier Series Modes for Circular Apertures

To introduce phase modes, circular apertures are considered. A circular aperture is a circular array in which the interelement spacing of the circumferential elements approaches 0, in other words $d_e \to 0$ [35]. When a plane wave impinges on a circular aperture from a direction $\bar{\phi}$, the aperture is excited and each element of the aperture is excited differently depending on the spatial frequency of the wave. An important aspect of this excitation is that it is periodic around the aperture with a period of 2π.

This periodicity means that the excitation of the circular aperture can be expressed

using a Fourier series [36]. Specifically, a spatial Fourier series of the form

$$f(\phi) = \sum_{n=-\infty}^{\infty} f_n e^{jn\phi} \tag{3.28}$$

can be used, where $e^{jn\phi}$ are the orthogonal basis functions, or modes of the decomposition, and f_n are the complex mode weights. These modes are often referred to as circular harmonics for a 2-dimensional spatial decomposition [35].

The sum in Eqn. (3.28) is important as $f(\phi)$ can be thought of as a beamformed output if the modes of interest in the sum have been previously produced. The beam pattern of this output can then be varied by changing the mode weights, f_n. This method of beamforming by weighting different modes constitutes modal beamforming. However, in order for a wideband beam pattern to be produced, wideband modes must first be created from the received signal. Therefore, the method of producing these modes using both continuous and sampled arrays must be studied.

3.3.1.2 Complex Fourier Mode Creation Using a Circular Aperture

Above it was noted that if the modes of interest can be produced by the circular aperture, then a beam pattern, $f(\phi)$ can be produced by changing the mode weights, f_n. However, the mode creation itself makes use of beamforming and relies on the Fourier series applied to the circular aperture as well. So, to distinguish between the modal beamforming procedure used to produce the final beamformed output signal in Eqn. (3.28) and the beamforming to produce the required modes, a new Fourier series is defined as

$$g(\phi) = \sum_{b=-\infty}^{\infty} g_b e^{jb\phi}, \tag{3.29}$$

where now $g(\phi)$ represents an incoming signal or desired modal pattern, b represent the modes excited on the circular aperture and g_n represents the mode weights produced by the incoming signal or required to produce a desired modal pattern.

In order to beamform to create a desired mode, $g(\phi) = e^{jb\phi}$, it is easy to see from Eqn. (3.29) that this corresponds to setting $g_b = 1$ at the desired mode and $g_b = 0$ at all other modes. Each element of the aperture located at a different position,ϕ, will then be weighted by $e^{jb\phi}$ in order to produce the b^{th} mode. While in theory this should

produce the desired beamformed mode, it is important to look at the effect that a plane wave has on each mode as each cluster arrival in the UAC can be modeled as a plane wave.

To do this, it can be assumed that a plane wave arrives from a direction $\bar{\phi}$, as in Eq. (3.5). For this analysis $g(\phi)$ now represents a plane wave and the excitation caused on the b^{th} mode of the aperture can be found by calculating the mode weight produced by the plane wave as

$$g_b(kr, \bar{\phi}) = \int_{\phi=0}^{2\pi} e^{jkr\cos(\phi-\bar{\phi})} e^{-jb\phi} d\phi. \tag{3.30}$$

The mode weight is now a function of the spatial frequency k, radius, r, and angle-of-arrival, $\bar{\phi}$, as changing any of these parameters will effect the mode weight. Using a change of variables with $\gamma = \phi - \bar{\phi}$ [35], Eqn. (3.30) can be simplified to

$$g_b(kr, \bar{\phi}) = e^{-jb\bar{\phi}} \int_{\phi=0}^{2\pi} e^{jkr\cos\gamma} e^{-jb\gamma} d\phi. \tag{3.31}$$

This integral is the integral form of a Bessel function of the first kind [37], so this equation further simplifies to

$$g_b(kr, \bar{\phi}) = j^b J_b(kr) e^{-jb\bar{\phi}}. \tag{3.32}$$

This result is important as Eqn. (3.32) shows that when a plane wave arrives on a circular aperture with a frequency, k, the excitation of the b^{th} mode becomes frequency dependent due to the Bessel function, $J_b(kr)$. This has a two main consequences. First, the Bessel function imposes a high-pass like filter effect with an approximate spatial cutoff frequency of $k_b r \approx b$ [36], where b is the Bessel function order and k_b is the spatial cutoff frequency assuming a constant radius. Below this cutoff frequency, the b^{th} mode has a very small value due to the low value of the associated Bessel function.

This means that if the radius and desired maximum spatial frequency are set, then the number of modes that can be created that have a substantial value are set as well at some finite number. As an example, if the the radius is chosen to be $r = \lambda_{min}$ and the minimum wavelength is encountered, then the spatial frequency is $k_{max} = \frac{2\pi}{\lambda_{min}}$ and the maximum argument encountered by the Bessel function becomes $k_{max} r = 2\pi$. Figure 3.7 shows a plot of the Bessel functions with different orders. This Figure

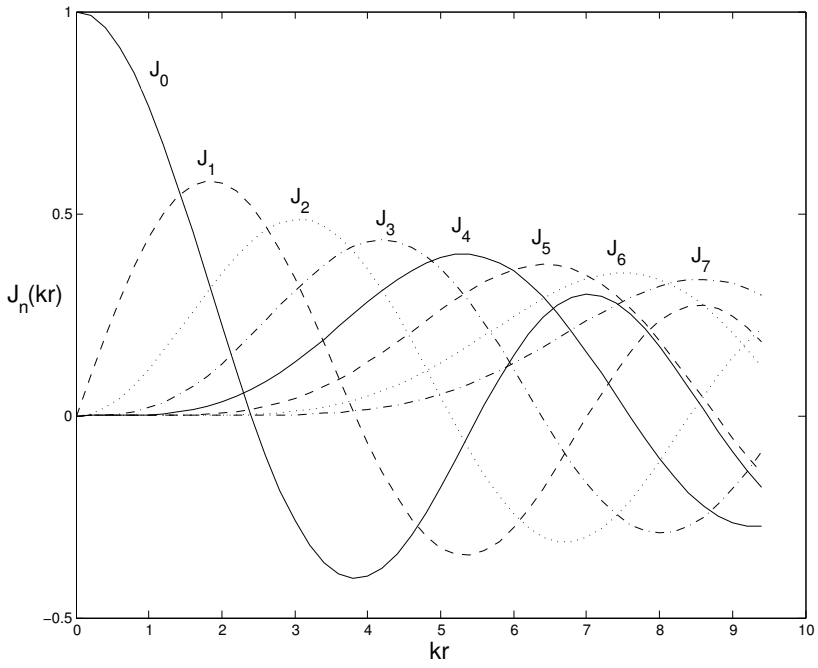


Figure 3.7: Plot of Bessel functions of the first kind with respect to the argument kr.

shows that when $kr = 2\pi$, the last mode that has a significant contribution is J_6. Therefore, a rule of thumb that is often used to calculate the highest mode that can be created using a circular aperture is [36], [38]

$$b_{max} \approx k_{max} r. \tag{3.33}$$

This implies that if the argument of the Bessel function is less than the order, then the output of the Bessel function will be negligible and does not need to be calculated.

The second problem introduced by the Bessel functions of different orders, n, is that even if the modes contain enough energy, some of these functions contain zero crossings as kr changes. This can again be seen in Figure 3.7. So, for an aperture with a constant radius, as the frequency increases and wavelength decreases, certain modes could encounter zeros causing the mode to have zero output at that value.

So when it is assumed that all arrivals will be plane waves, then the summation in Eqn. (3.29) can be rewritten using the mode weights in (3.32) as

$$g(\phi) = \sum_{b=-\infty}^{\infty} g_b(kr, \bar{\phi}) e^{jb\phi} = \sum_{b=-\infty}^{\infty} j^b J_b(kr) e^{jb(\phi - \bar{\phi})}, \tag{3.34}$$

where now $g(\phi)$ represents a desired beam pattern if it is assumed only plane waves are impinging on the array.

Now, to beamform an impinging plane wave to create a desired mode involves compensating for the frequency variant mode weight. Specifically, if the elemental weights are still chosen to be $e^{jb\phi}$ at the elemental position ϕ, the output of the array becomes

$$g(\phi) = j^b J_b(kr) e^{jb\phi} \tag{3.35}$$

when it is assumed the wave arrives from $\bar{\phi} = 0$. If the wave arrives from another direction $\bar{\phi} \neq 0$, then it will be weighted by the mode $e^{jb\phi}$ as desired. However, in order to remove the frequency dependency of the mode the output of the aperture with these element weights needs to be multiplied by

$$\frac{1}{j^b J_b(kr)} \tag{3.36}$$

The two problems discussed above will also effect the modes produced. The number of modes used will be limited by kr, and the frequency range over which the multiplication by $\frac{1}{j^b J_b(kr)}$ functions will be limited by the zeros-crossings of the modes. This is discussed in more detail in Section 3.3.2. Once the complex modes up to $\pm b_{max}$ are created and the frequency dependency is removed using the procedures mentioned above, then a wideband beamformed output can be created using

$$f(\phi) = \sum_{n=-b_{max}}^{b_{max}} f_n e^{jn\phi}. \tag{3.37}$$

In reality the number of elements in the circular aperture will be finite and so the aperture will need to be sampled. This leads to the use of the UCCA to produce the modes of interest which introduces its own challenges.

3.3.1.3 Complex Fourier Mode Creation Using a UCCA

In order to practically implement a phase mode beamformer the circular aperture needs to be sampled, and the desired modes need to be produced from this sampled circular array. However, in order to ensure that the beamformer functions for all modes up to $n = b_{max}$, the aperture needs to be sampled properly. In order to determine the number of uniformly distributed circumferential elements that need to

be used for a UCCA implementing modes up to b_{max}, an l^{th} mode beamformer can be implemented. To do this, the weight vector is chosen in the same way as for the circular aperture in the Section above with

$$\mathbf{w}_l = \frac{1}{M}[e^{jl\phi_1} \quad e^{jl\phi_2} \dots e^{jl\phi_M}]^T, \tag{3.38}$$

except now due to the sampling the weights are calculated at discrete angles of ϕ_m, rather than at continuous angles of ϕ. In (3.38) the center element is given a constant weight of zero and is excluded. The center element is only used to create the zeroth order mode, $n = 0$, with all other elements given a weight of zero. This will also be discussed in Section 3.3.2. The beam pattern of the l^{th} mode beamformer can then be described as

$$P_l(\omega, \phi) = \mathbf{w}_l^H \mathbf{d}(\omega, \phi) = \frac{1}{M} \sum_{m=1}^{M} e^{jl\phi_m} e^{jkr\cos(\phi - \phi_m)} \tag{3.39}$$

The plane wave in (3.39) can be expressed in terms of a spatial Fourier series by expanding the term describing the signal, $e^{jkr\cos(\phi-\phi_m)}$, on each array element using

$$e^{jkr\cos(\phi-\phi_m)} = \sum_{b=-\infty}^{\infty} j^b J_b(kr) e^{-jb\phi_m} e^{jb\phi} = \sum_{b=-\infty}^{\infty} j^b J_b(kr) e^{jb(\phi-\phi_m)}. \tag{3.40}$$

The beam pattern then becomes

$$P_l(\omega, \phi) = \frac{1}{M} \sum_{m=1}^{M} \sum_{b=-\infty}^{\infty} j^b J_b(kr) e^{jb(\phi-\phi_m)} e^{jl\phi_m}, \tag{3.41}$$

which can be rearranged into the form

$$P_l(\omega, \phi) = \sum_{b=-\infty}^{\infty} j^b J_b(kr) e^{jb\phi} \frac{1}{M} \sum_{m=1}^{M} e^{j(m-1)\frac{2\pi(l-b)}{M}}. \tag{3.42}$$

The second summation term with respect to m only has a non-zero value of M, when the $(l - b) = pM$, where p is an integer value. So the beam pattern of the l^{th} mode beamformer can be described as

$$P_l(\omega, \phi) = j^l J_l(kr) e^{jl\phi} + \sum_{q=1}^{\infty} \left(j^g J_g(kr) e^{jg\phi} + j^h J_h(kr) e^{-jh\phi} \right), \tag{3.43}$$

where $g = (qM - l)$ and $h = (qM + l)$ [36].

Discretizing the circular aperture using a UCCA and applying the weight vector $\mathbf{w}_l$ to its elements produces an approximation of an l^{th} mode. As with the l^{th} mode produced on the continuous circular aperture, the desired mode on the sampled circular array is multiplied by the frequency dependent term $j^l J_l(kr)$. However, the mode produced using the sampled circular array also has extra error terms represented by the summation term in (3.43) with respect to q. In order for these error terms to be negligible, the highest mode produced by the sampled circular array must be limited. To do determine this upper limit we consider the closest error term at $M - l$. For the the Bessel function $J_{M-l}(kr)$ to have a small value, the argument when $kr = k_{max}r$ should be less than the order $M - l$. This can be written as

$$k_{max}r < M - l. \tag{3.44}$$

It was also shown above that the maximum Bessel function order that can be used for an aperture depends on $k_{max}r$. Choosing $l = b_{max}$ means $l \approx k_{max}r$ from Eqn. (3.33), which leads to

$$k_{max}r < M - k_{max}r \quad \rightarrow \quad M > 2k_{max}r \approx 2b_{max}. \tag{3.45}$$

This indicates that the minimum number of elements required in order to produce an $l = b_{max}$ mode beamformer is $2b_{max}$. This can be expressed in terms of the circumferential element spacing as

$$M > 2\frac{2\pi r}{\lambda_{min}} \quad \rightarrow \quad \frac{2\pi r}{M} = d_e < \frac{\lambda_{min}}{2}. \tag{3.46}$$

If M is chosen so that the error term is small for all $|l| \leq b_{max}$, which corresponds to $M > 2b_{max}$, then the output of the circular array with the l^{th} mode applied will be

$$P_l(\omega, \phi) \approx j^l J_l(kr) e^{jl\phi}, \tag{3.47}$$

where this is still an approximation as the error term may still have a small effect.

By sampling the circular aperture properly, the modes produced by the UCCA will be very similar to those produced by the continuous aperture. This means the frequency effects can be removed in the same way and then the modes can be used for beamforming as shown in Eqn. (3.37) in the Section above. While phase modes have

appealing characteristics, unfortunately they require complex weights to be applied to the receive elements. So, when dealing with real received signals, this requires the extra step of first producing an imaginary version of the signal, which involves the use of a Hilbert transform. A Hilbert transform can be implemented practically using an FIR filter, however this adds computational overhead that can be avoided if real weights are instead used. In order to produce real weights, the focus shifts to amplitude mode processing.

3.3.2 Amplitude Mode Processing

With phase mode processing complex basis functions were used for the spatial decomposition as complex numbers are often easier to work with. However, if real weights are required for the beamformer, as is the case when working with real signals, then the basis functions used must also be real. One possible set of basis functions is the sine and cosine basis set, which when used leads to the Fourier series

$$f(\phi) = \frac{1}{2}d + \sum_{n=1}^{\infty}\Big(a_n \cos(n\phi) + b_n \sin(n\phi)\Big). \tag{3.48}$$

When the Fourier series is used, the weight vector applied to the array to create the n^{th} mode will always be real. However, in order replicate the complex mode both a sine and cosine mode of order n need to be implemented. To show that this series is equivalent to the exponential series in Eqn. (3.28), the sine and cosine terms can be written using Euler's formula as

$$\begin{aligned} f(\phi) &= \frac{1}{2}d + \sum_{n=1}^{\infty}\left(a_n \frac{(e^{jn\phi} + e^{-jn\phi})}{2} - jb_n \frac{(e^{jn\phi} - e^{-jn\phi})}{2}\right) \\ &= \frac{1}{2}d + \sum_{n=1}^{\infty}\left(e^{jn\phi}\frac{(a_n - jb_n)}{2} + e^{-jn\phi}\frac{(a_n + jb_n)}{2}\right). \end{aligned} \tag{3.49}$$

This sum can be made equal to the complex sum in Eqn. (3.28) by choosing the weights of the complex sum to be

$$w_n = \begin{cases} d, & n = 0 \\ (a_n - jb_n)/2, & n > 0 \\ (a_n + jb_n)/2, & n < 0 \end{cases} \tag{3.50}$$

This equality with the complex Fourier series means that the same results presented above for phase modes also hold for amplitude mode processing. Specifically, when amplitude modes of order n are applied to a circular aperture, the resulting beam patterns produced for both the cosine and sine modes can be represented as

$$P_{n,\cos}(\Omega,\phi) = j^n J_n(\Omega\beta)\cos(n\phi) \tag{3.51}$$

and

$$P_{n,\sin}(\Omega,\phi) = j^n J_n(\Omega\beta)\sin(n\phi), \tag{3.52}$$

respectively. Here, $\Omega\beta$ has been used to replace kr to indicate both temporal and spatial sampling has occurred.

With modal beamforming, first the modes of interest need to be created then the weighted combination of the modes can be used to produce a beamformed output. To produce the required cosine and sine modes of order n, the weight vector for the circumferential elements is chosen as

$$\mathbf{w}_{n,\cos} = [\cos(n\phi_1) \quad \cos(n\phi_2)\ldots\cos(n\phi_M)]^T \tag{3.53}$$

and

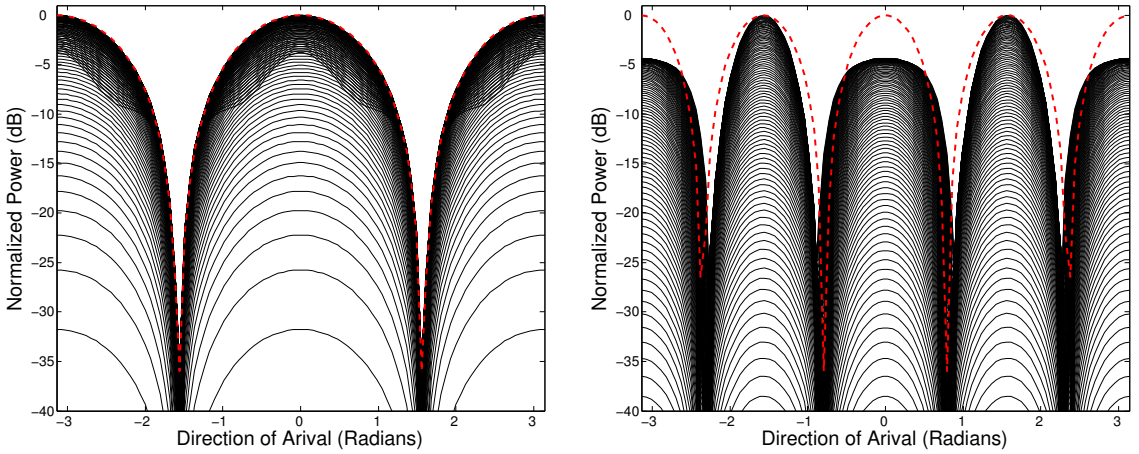
$$\mathbf{w}_{n,\sin} = [\sin(n\phi_1) \quad \sin(n\phi_2)\ldots\sin(n\phi_M)]^T, \tag{3.54}$$

respectively.

Where again, as long as $n < \frac{M}{2}$ the output of the circular array from these weights will closely approximate (3.51) and (3.52). These equations show that the outputs of the array for the n^{th} amplitude modes have frequency dependent gain terms, $j^n J_n(\Omega\beta)$. This dependence on frequency is shown in Figure 3.8 for cosine modes with $n = 1$ and $n = 2$ for a 7-element UCCA with 6 circumferential elements.

In order to produce the frequency-invariant modes required for wideband modal beamforming, the output of each mode can be filtered using a mode dependent output filter of the form [32]

$$H_n(\Omega) = \frac{1}{j^n J_n(\Omega\beta)}, \quad \text{for } n \geq 1. \tag{3.55}$$



(a) $n = 1$ (b) $n = 2$

Figure 3.8: Frequency response of cosine modes of order n produced using a 7-element UCCA. The black lines represent slices of the normalized frequency response, Ω, from 0 to 1. The dashed red line is the desired response.

Here n is chosen to be greater than 0 as the zeroth order mode, which is simply an omni-directional channel that records in all directions, can be produced using the center element of the UCCA, $m = 0$. If the center element is given a weight of 1 and all circumferential elements are given a weight of 0, then no filtering will be required as no frequency effects will be introduced through the summation of elements.

These filters do however pose some problems that lie in the frequency dependent Bessel functions. Specifically, the Bessel functions contain zeros, and so when they are used to produce an inverse filter, that filter will contain frequencies that require infinite gain. This effect can be limited by choosing the minimum and maximum normalized frequency, $[\Omega_{min}, \Omega_{max}]$, such that in this range no Bessel function has a zero crossing [32]. However, as more modes are included this range becomes harder to choose and may severely limit the frequency range over which an amplitude mode beamformer can be implemented.

When only a small number of modes are used, a modal beamformer produced using a UCCA can create a wideband beamformer that avoids the zero crossings of the Bessel functions. The general structure of an amplitude mode wideband beamformer is shown in Figure 3.9.

When FIR filters are designed to approximate the ideal mode dependent filters, the output of each mode becomes very close to frequency flat. This is shown in Figure 3.10 where FIR filters of length $J = 201$ were designed for each mode. Even with the

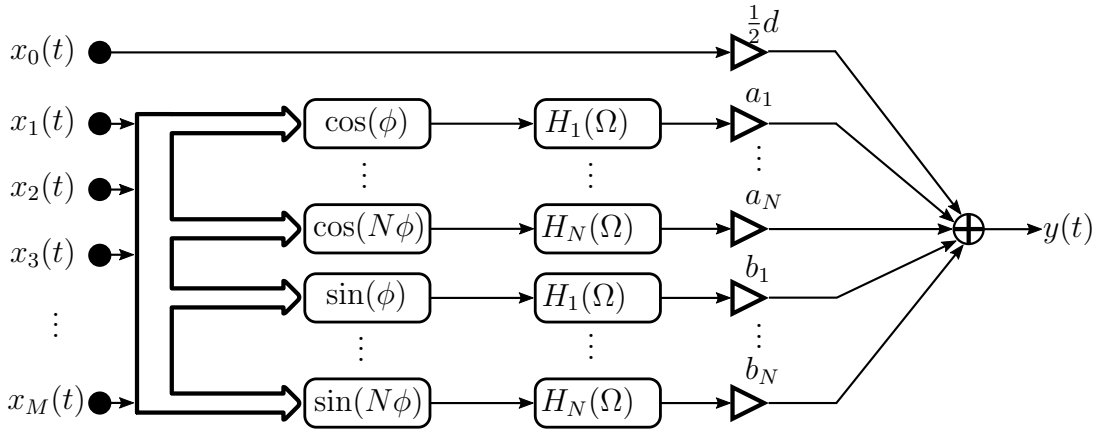
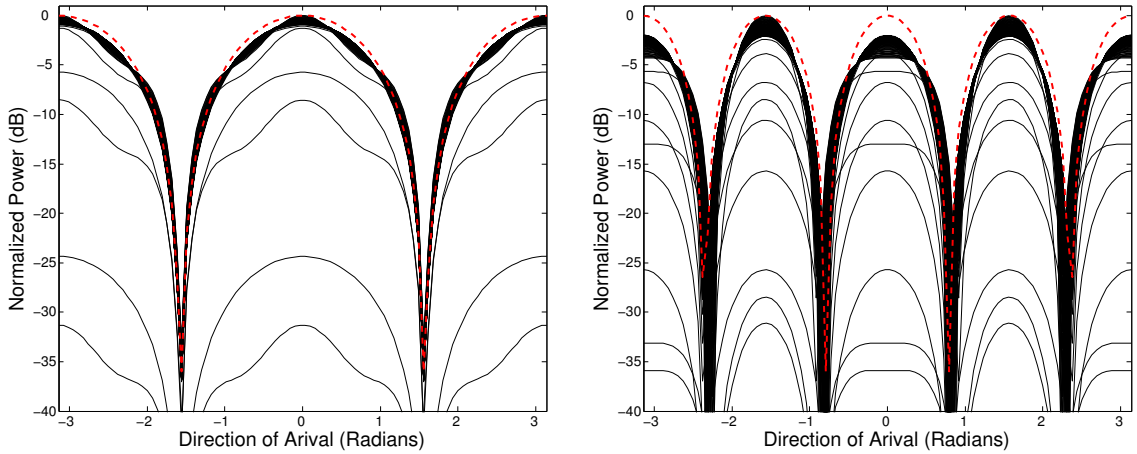


Figure 3.9: Wideband modal beamformer architecture using N modes. The center element, $m = 0$, is taken as the omnidirectional channel rather then creating it by summing all elements with a weights of 1. $H_n(\Omega)$ are the mode dependent filters used to produce frequency invariant modes.

filtering, Figure 3.10 shows that as the frequency increases there is distortion in the desired mode shape. This can be explained by modal aliasing [35].



(a) $n = 1$ (b) $n = 2$

Figure 3.10: Filtered frequency response of cosine modes produced using a 7-element UCCA. The black lines represent slices of the normalized frequency response, Ω, from 0 to 1. The dashed red line is the desired response.

It was shown that when the circular aperture is sampled by M elements and used to produce a desired mode, this mode can only be produced if $n < \frac{M}{2}$. Even when this condition is satisfied, there is still an error term that is present in the expression. The

output of a sampled circular aperture for a cosine mode of order n can be written as

$$P_n^s(\Omega,\phi) = \underbrace{j^n J_n(\Omega\beta)\cos(n\phi)}_{P_n(\Omega\beta,\phi)} + \underbrace{\sum_{q=1}^{\infty}\left(j^g J_g(\Omega\beta)\cos(g\phi) + j^h J_h(\Omega\beta)\cos(h\phi)\right)}_{P_n^e(\Omega\beta,\phi)}, \quad (3.56)$$

where $P_n^s(\Omega,\phi)$ is the sampled n^{th} order response, $P_n(\Omega,\phi)$ is the desired n^{th} mode response, and $P_n^e(\Omega,\phi)$ is the error caused by sampling. Here g and h are the same variables that were defined for Eq. (3.43).

$P_n^e(\Omega,\phi)$ is the term that causes modal aliasing or aliasing of higher order modes into the sampled mode of interest. In order to determine the effect of this aliasing, an error term, $\varepsilon_n(\Omega,\phi)$, can be derived that looks at the energy of the error term with respect to the energy of the sampled response for a specific order, n [35]. A plot of this error term for a UCCA with 7-elements for orders 1 and 2 when $\phi = 0$ is shown in Figure 3.11.

$$\varepsilon_n(\Omega,\phi) = \frac{|P_n^e(\Omega,\phi)|^2}{|P_n^s(\Omega,\phi)|^2} \quad (3.57)$$

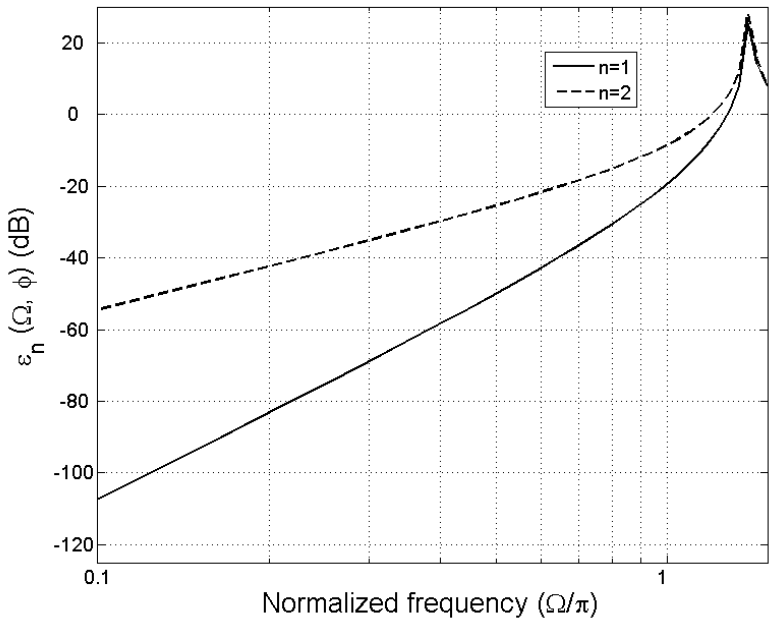
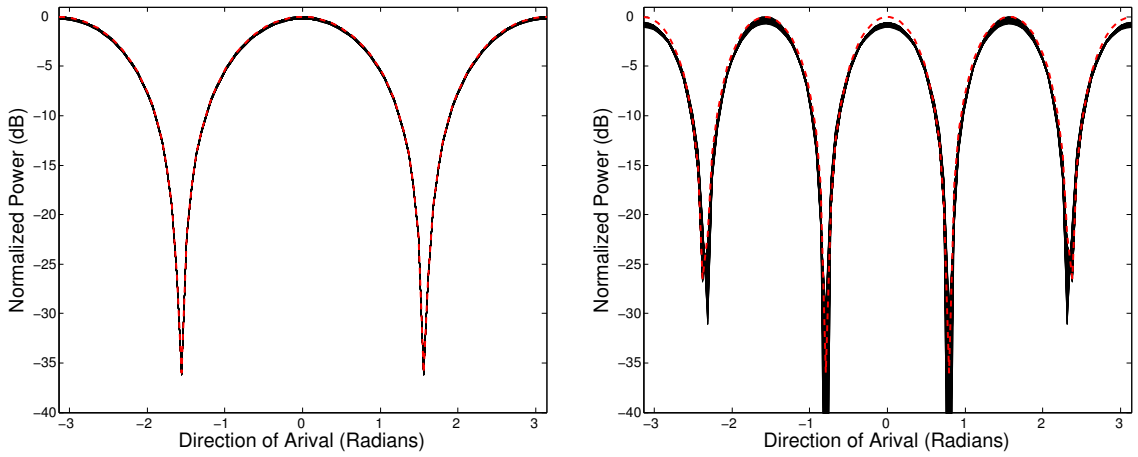


Figure 3.11: Plot of the energy of the aliasing term with respect to the total energy recorded, $\varepsilon_n(\Omega,\phi)$.

This error is always less than 1 when the normalized frequency is below 1. This means

that if each element is critically sampled in time with $T_s = \frac{\lambda_{min}}{2c}$, then the error term will never be dominant. However, the presence of the error term does cause distortion as can be seen in Figure 3.10. In order to reduce this distortion the bandwidth of the beamformer needs to be limited to ensure $\varepsilon(\Omega, \phi)$ is small for all modes. If the functional bandwidth is limited to $[0.1\Omega, 0.5\Omega]$, then the modal aliasing error is below -20 dB for both cosine modes and the frequency response remains much more constant as shown in Figure 3.12.

While the upper frequency limit is determined by the modal aliasing, the lower frequency is limited by the amount of gain that needs to be applied to invert the effect of the Bessel function for the highest mode included in the beamformer. The requirement to amplify the signal significantly to attain a unitary mode means that the noise will be greatly amplified, so the lowest frequency used in the modal beamformer will be set by the minimum SNR required.



(a) $n = 1$ (b) $n = 2$

Figure 3.12: Filtered frequency response of cosine modes produced using a 7-element UCCA limited to a functional bandwidth of $[0.1\Omega, 0.5\Omega]$. The dashed red line is the desired response.

Once all of the frequency flat amplitude modes up to order b_{max} are produced using the UCCA of order M, then the task of designing the beamformer becomes a task of choosing the modal weights, w_n, that implement a desired beam pattern, $f(\phi)$, as in Eqn. (3.37). To convert back to the cosine and sine coefficients required for amplitude mode beamforming, the real weights can be selected following Eqn. (3.50). Because the modes have already been made frequency invariant, the pattern will be implemented for all frequencies within the beandwidth of the modes. So for the example used above with the 7-element UCCA with $b_{max} = 2$, $f(\phi)$ would be implemented for

all frequencies within the normalized frequency range of $[0.1\Omega, 0.5\Omega]$.

3.3.3 Design of Modal Weights

Once the frequency invariant amplitude modes have been produced these modes can be combined with different weights in order to create different beam patterns. The relationship between Eqn. (3.37) and the inverse Discrete Fourier Transform mean that FIR filter design techniques could be used to design these weights. However, when the number of modes included in the summation is small a simpler solution is to use a method called Cardiod beamforming.

Cardiod beamforming creates a beamformed output through the weighted combination of a zeroth and first order term, or an omnidirectional and first-order pressure gradient element [39]. This weighted combination produces a cardiod beamformed output of the form

$$P_{\text{Cardiod}}(p) = p + (1-p)\cos(\phi) \quad 0 \leq p \leq 1 \tag{3.58}$$

where p controls the beam pattern produced from the combination.

When $p = 0.5$ this corresponds to a cardiod output with a null optimally placed π radians away from the steering direction. When $p > 0.5$ this corresponds to a wide, or subcardiod with a main beam wider than a cardiod, and when $p < 0.5$ this corresponds to a hypercardiod with a main beam narrower than a cardiod. Plots of these three types of beam patterns are shown in Figure 3.13.

The cardiod pattern can be steered by rotating the cosine term to some desired steering angle, $\bar{\phi}$. This rotation can be expressed as a weighted sum of a cosine and sine term of the same order [40],

$$\cos(\phi - \bar{\phi}) = \cos(\phi)\cos(\bar{\phi}) + \sin(\phi)\sin(\bar{\phi}). \tag{3.59}$$

The first-order beamformer steered to a desired direction can then be written as

$$P_{\text{Cardiod}}(p, \bar{\phi}) = p + (1-p)(\cos(\phi)\cos(\bar{\phi}) + \sin(\phi)\sin(\bar{\phi})) \quad 0 \leq p \leq 1. \tag{3.60}$$

For modal beamformers with $b_{max} > 1$ a cardiod beamformer can also be formed by increasing the order of the equation by raising the original beamformer to the desired

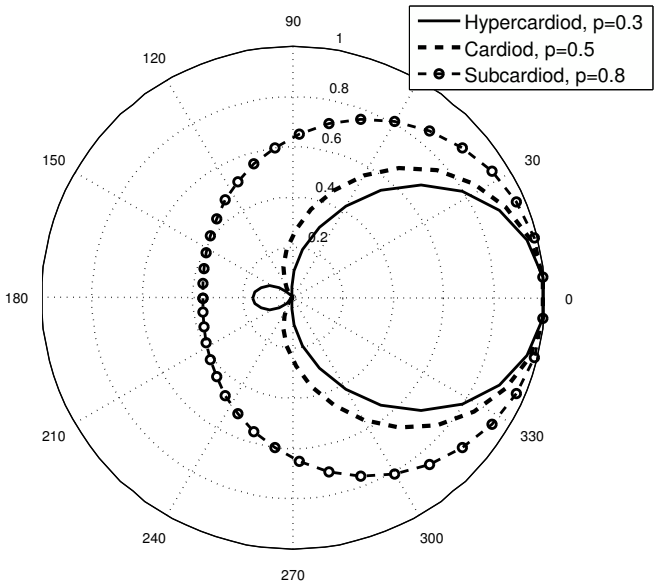


Figure 3.13: Polar plot of beam patterns formed using the first-order cardiod beamformer with different values of p.

order, b_{max} [39].

$$P_{\text{Cardiod},b_{max}}(p) = [p + (1-p)\cos(\phi)]^{b_{max}} \quad 0 \leq p \leq 1 \tag{3.61}$$

As an example, for the 7-element UCCA with $b_{max} = 2$, the beam pattern can be expressed as

$$\begin{aligned} P_{\text{Cardiod},2}(p) &= p^2 + 2p(1-p)\cos(\phi) + (1-p)^2\cos^2(\phi) \\ &= p^2 + 2p(1-p)\cos(\phi) + (1-p)^2\left(\frac{\cos(2\theta)+1}{2}\right). \end{aligned} \tag{3.62}$$

Beam patterns for this new beamformer using modes upto $b_{max} = 2$ are shown in Figure 3.14. Comparing these beam patterns to the ones formed using only modes upto $b_{max} = 1$ in Figure 3.13, it can be seen that adding the extra order helps to reduce the main beamwidth for all plots as well as reduce the power received from $\phi = \pi$, for the hypercardiod and subcardiod. This beamformer can be steered in the same way as the first order system by combining weighted cosine and sine terms of the same order.

While the Cardiod beamforming technique provides a solution for systems of any order, as the order increases the solution becomes more complicated. Therefore, this

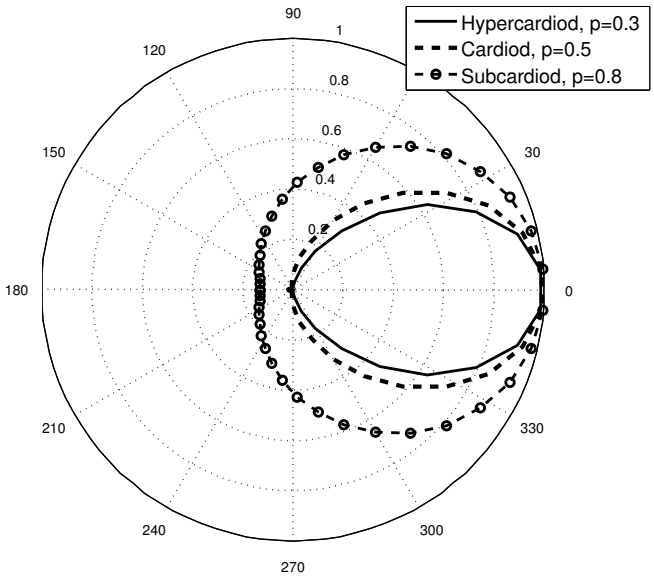


Figure 3.14: Polar plot of beam patterns formed using the first-order cardiod beamformer with different values of p.

technique remains a plausible beamforming solution for the low order systems used in the experiment presented in Section 3.4 and in Chapter 4. However, as the number of modes included increases, common FIR design techniques provide a better solution.

3.4 Modal Beamforming Experimental Results

In order to verify the theory presented in Section 3.3 on modal beamforming, experimental underwater trials were conducted. For these experiments a 5-element UCCA was used and was mounted onto the end cap of the TR-Orca, which is an autonomous underwater data acquisition system provided by Turbulent Research. The spacing between the elements in this array was 0.0381 m. The TR-Orca was used to record the 5-channels required to implement the modal beamformer and all processing was done offline in Matlab.

The experiment was carried out in the anechoic tank at Defence Research and Development Canada's Atlantic Research Centre (DRDC). This tank is cylindrical in shape with a 7.3 m diameter and with a depth of 4.5 m. DRDC Atlantic also provided a broadband piston-actuated acoustic source and the equipment required to drive it. A side view of the experimental deployment is shown in Figure 3.15 and a bottom view

of the experimental deployment is shown in Figure 3.16.

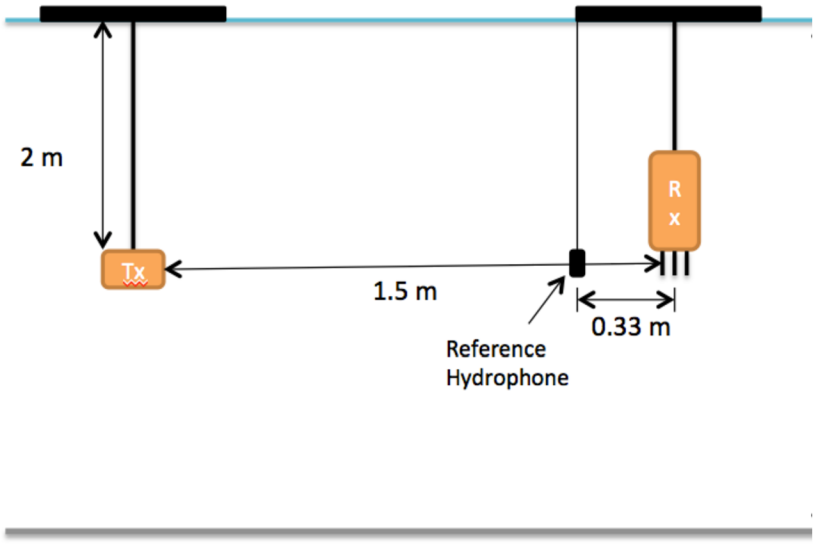


Figure 3.15: Side view of the experimental deployment.

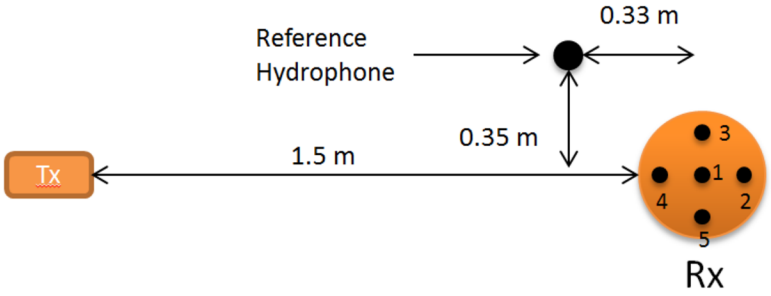


Figure 3.16: Bottom view of the experimental deployment.

The transmitter, Tx, (the piston actuated source) and the 5-element receiver, Rx, (the TR-Orca) were both placed at a depth of 2 m in order to maximize the time between the direct path arrival and the first multipath arrival. Both the receiver and transmitter were placed at the same depth to ensure that this was a test of the planar directionality and that no height information would effect the recorded data. The transmitter and receiver were placed 1.5 meters apart so that far field conditions were met.

The bottom view in Figure 3.16 shows the orientation of the hydrophones, which allows the proper hydrophones to be used to create the amplitude modes of interest.

Specifically, for this experiment hydrophone 4 was used as the $\phi = 0$ position and hydrophone 5 was used as the $\phi = \frac{\pi}{2}$ position. The omnidirectional channel, d, was produced using hydrophone 1, while the cosine and sine amplitude modes were produced using the 4 outer hydrophones, 2 to 5, as in Figure 3.9.

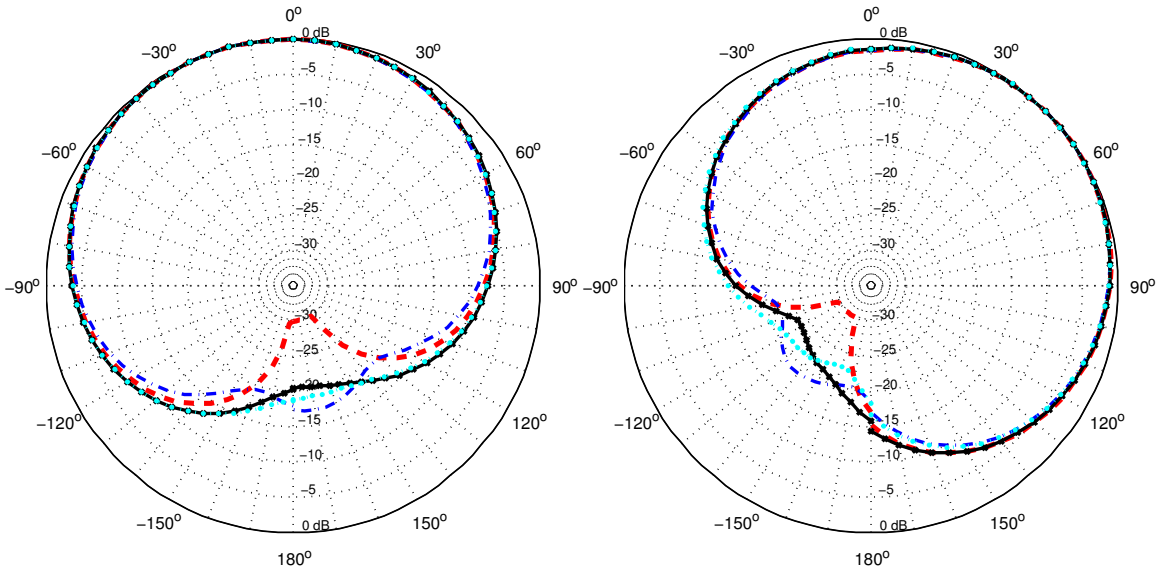
Specifically, because $M = 4$ for the 5-element UCCA, the maximum order that could be produced was $b_{max} = 1$, and so only 2 amplitude modes were produced, $\cos(\phi)$ and $\sin(\phi)$. The array element weights for these modes were produced using Eqns. (3.53) and (3.54), respectively. To implement wideband modes an FIR filter was designed to approximate the mode dependent output filter in Eqn. (3.55) with $n = 1$. Using a sampling frequency of 20 kHz, the filter designed had a bandwidth of 9 kHz from 0.4 to 9.4 kHz and so the beamformer was limited to these frequencies.

In the experiment the TR-Orca was connected to an actuator that could rotate it. The actuator controller was zeroed at the position shown in Figure 3.16. However, during post-processing it was found that the zeroed position was off by about -11° degrees, and so the uncorrected data was used.

In order to test the frequency response of the modal beamformer, a number of short, windowed constant frequency pulses were sent. Specifically, pulses were sent at 2 kHz, 5 kHz and 9 kHz, where the lower limit was set by the transmitter and the upper limit was determined by the beamformer. In order to determine the modal weights, first-order cardiod beamforming was applied.

Beam patterns showing the attenuation of the beamformed pulses with respect to the omnidirectional channel for the three frequencies were created for angles of arrival of 0° and 60° and for two different pattern parameters, $p = 0.5$ and $p = 0.3$. These plots are shown in Figure 3.18 and the subcaption indicates the original angle of signal arrival, ϕ, as well as the corrected angle of arrival, ϕ_c. The ideal response is plotted with the measured response as a reference to show the close match between the measured and theoretical polar patterns. These beam patterns were created by steering the beamformer to different angles by varying $\bar{\phi}$ between $[-180,\ 180]$ and calculating the power of the recorded pulse.

These plots demonstrate that the modal beamformer closely approximates the ideal Cardiod beamformer at all measured frequencies. This indicates that the theory presented in Section 3.3 is valid. Specifically this means that amplitude modes are being implemented properly and that the mode dependent output filter, $H_n(\phi)$ does in fact produce amplitude modes that are wideband and frequency invariant. This is



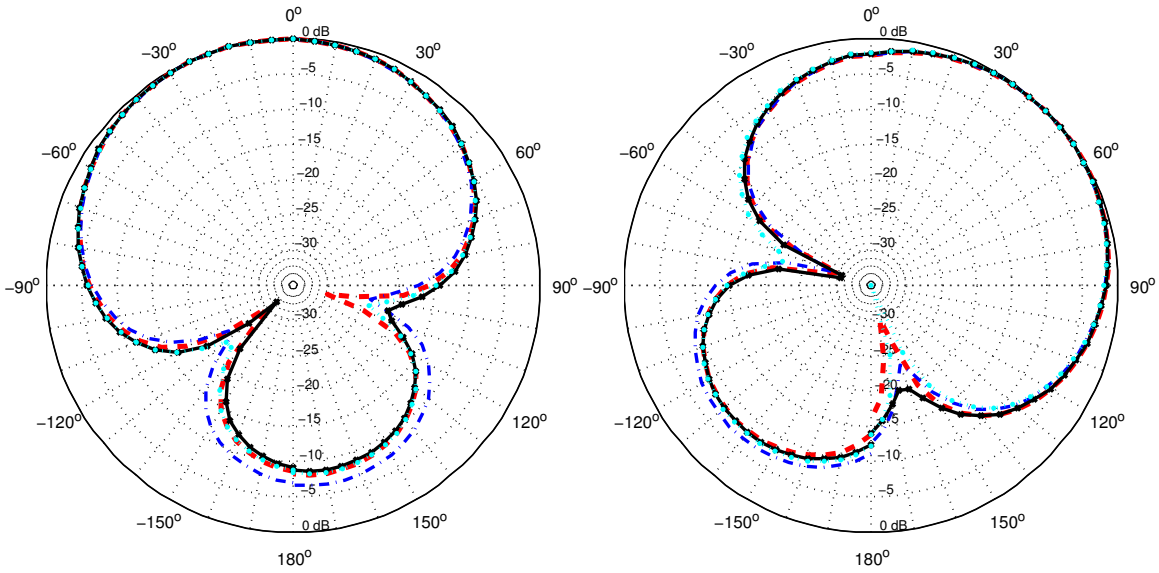
(a) p=0.5, $\phi = 0$, $\phi_c = -11$

(b) p=0.5, $\phi = 60$, $\phi_c = 49$

Figure 3.17: Polar patterns with p=0.5 measured using short, constant frequency pulses. Red is the ideal response, blue is the measured 2 kHz pattern, black is the measured 5 kHz pattern and turquoise is the measured 9 kHz pattern.

demonstrated by the constant beam pattern for all frequencies and the fact that they closely match the ideal beam pattern. These plots also demonstrate the ease with which a Cardiod pattern can be steered as all of the beam patterns were produced by electronically rotating the beamformer rather than physically rotating it. Using the first order Cardiod beamformer in Eqn. (3.60), this only required 4 multiplication and 2 addition operations.

While experiments were only carried out for a first order system due equipment limitations, if more circumferential elements were available then better beam patterns with narrower main beams and lower side lobe levels could be produced. However for the rest of this work, and specifically the simulations run in Chapter 4 that make use of the modal beamformer, b_{max} is limited to 1 as this is what has been experimentally demonstrated.



(a) p=0.3, $\phi = 0$, $\phi_c = -11$

(b) p=0.3, $\phi = 60$, $\phi_c = 49$

Figure 3.18: Polar patterns with p=0.3 measured using short, constant frequency pulses. Red is the ideal response, blue is the measured 2 kHz pattern, black is the measured 5 kHz pattern and turquoise is the measured 9 kHz pattern.

Chapter 4

Sparse Directional Channel Estimation using Orthogonal Frequency Division Multiplexing

This Chapter introduces a receiver design that makes use of a UCCA along with orthogonal frequency division mutliplexing (OFDM) in order to produce a sparse directional channel estimate of a time-varying channel that can be used to equalize the received signal. The effects of time-variance on the transmitted OFDM signal are studied and along with the modal beamformer are used to produce a representation of a received cluster arrival that has a delay, time-scaling factor and direction of arrival. This representation of the received signal is then used to create a sparse estimation problem and simulation results are presented to show the benefit of using this sparse approximation technique as well as the benefit of adding directionality to the channel estimate.

Section 4.1 presents the theory behind OFDM and discusses why it is a suitable choice for a delay-spread channel. The effect of time-variance on the received OFDM signal is also discussed. Section 4.2 introduces the compressed sensing problem and creates a compressed sensing dictionary based on the time-varying OFDM model for use in the sparse approximation environment. Section 4.3 concludes the Chapter by presenting simulation results to show the improvement of the proposed estimation technique over the classic least squares estimator.

4.1 The Effect of a Time-Varying Channel on OFDM Signals

The underwater acoustic channel so far has been dealt with only in terms of its physical effect on any transmitted signal. This Section focuses specifically on the effect of the UAC on an OFDM modulated signal and the distortions it produces on the received signal. Section 4.1.1 introduces OFDM and explains why OFDM is the preferred method of communication over delay-spread channels. Section 4.1.2 then looks at the specific effects of the time-varying UAC on an OFDM signal in order to determine how the received signal needs to be equalized.

4.1.1 Orthogonal Frequency Division Multiplexing

Single-carrier communication schemes use a number of different modulation techniques to transmit data over a channel centered around a single carrier frequency, f_c. When the channel consists of multiple path arrivals, or is spread in the delay domain, then the received signal first needs to be equalized to mitigate the effects of the channel before it can be demodulated to increase the chance of proper reception of the data. If we consider only a delay spread channel that is not time-varying, then the passband channel can be represented as a simplification of Eqn. (2.5),

$$\tilde{h}(\tau) = \sum_{c=1}^{N_c} a_c \delta(\tau - \tau_c). \tag{4.1}$$

If the delay spread, or maximum delay of the channel, τ_{max}, is smaller than the duration of a single transmitted symbol, T_{sym}, then the effect of the channel can be removed by a single complex multiplication. This assumes the received signal is sampled at the proper position. This means the channel is represented as a single complex tap, which is often referred to as frequency-flat fading because the Fourier transform of the single tap channel model leads to a flat response in the frequency domain. An example of a frequency flat channel is shown in Figure 4.1a.

However, as the data rate increases or the delay spread of the channel increases, then $T_{sym} < \tau_{max}$ and the channel can no longer be treated as a single tap and rather needs to be treated as a channel with multiple taps. This leads to frequency selective fading as the Fourier transform of shifted and scaled delta functions produces a varying

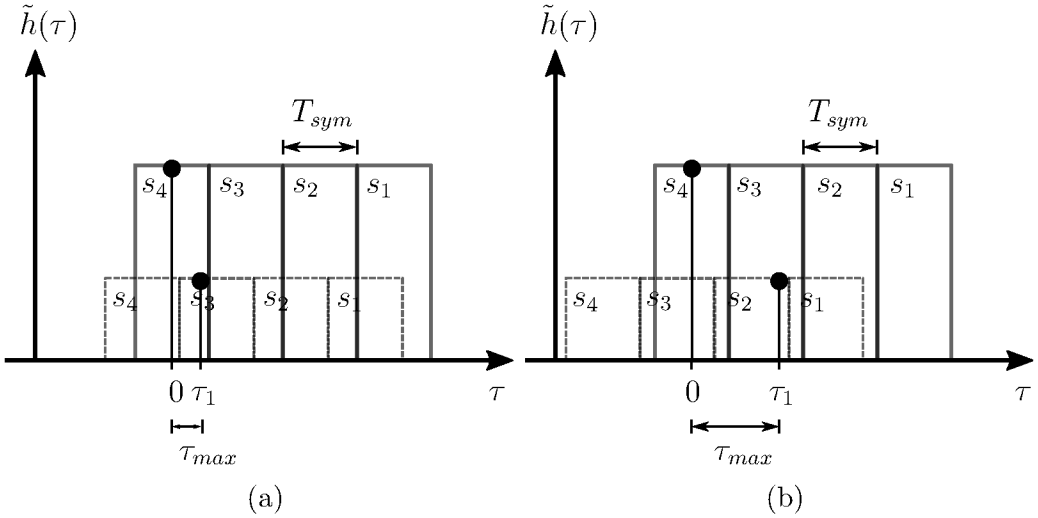


Figure 4.1: Plots showing the difference between a simple two-tap channel that causes frequency flat and frequency selective fading. 4.1a shows that when $\tau_{sym} > \tau_{max}$ and the symbol is sampled properly, the channel causes a delayed and scaled version of the same symbol to interfere with the current sample. This causes frequency flat fading and a single tap can be used to equalize. 4.1b shows that when $T_{sym} < \tau_{max}$ then other symbols interfere with the current symbol. This causes frequency selective fading and multiple taps are required to equalize.

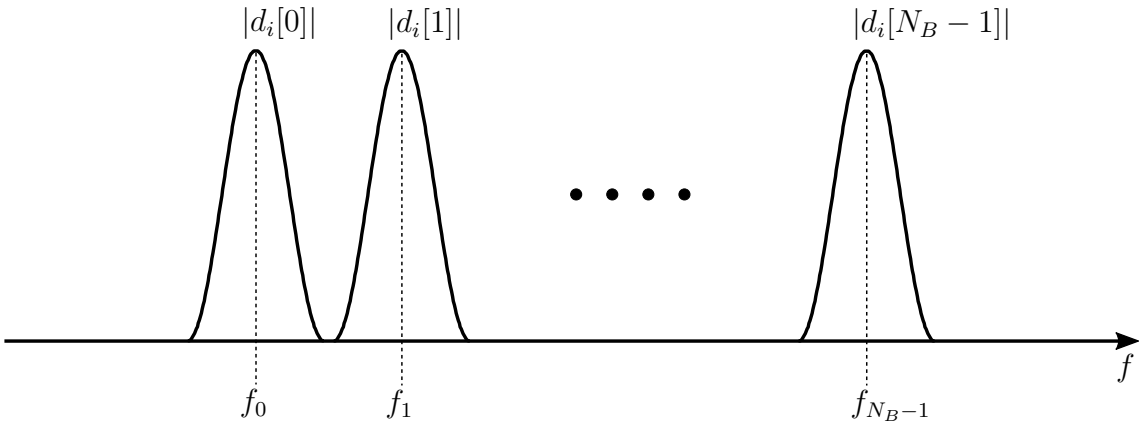
response in the frequency domain. Multiple taps in the channel cause a problem called inter-symbol interference (ISI), which causes previous symbols to interfere with the current symbol. This is shown in Figure 4.1b.

Another way of expressing this phenomenon is that as the symbol duration decreases and the transmission rate increases, the channel becomes longer with respect to the symbol rate. The equalizer therefore needs to compute more taps in order to have an accurate estimate of the channel, which can then be used to equalize the received data and remove the ISI. This can be quite computationally expensive in order to attain a good estimate if the transmission rate is high [5].

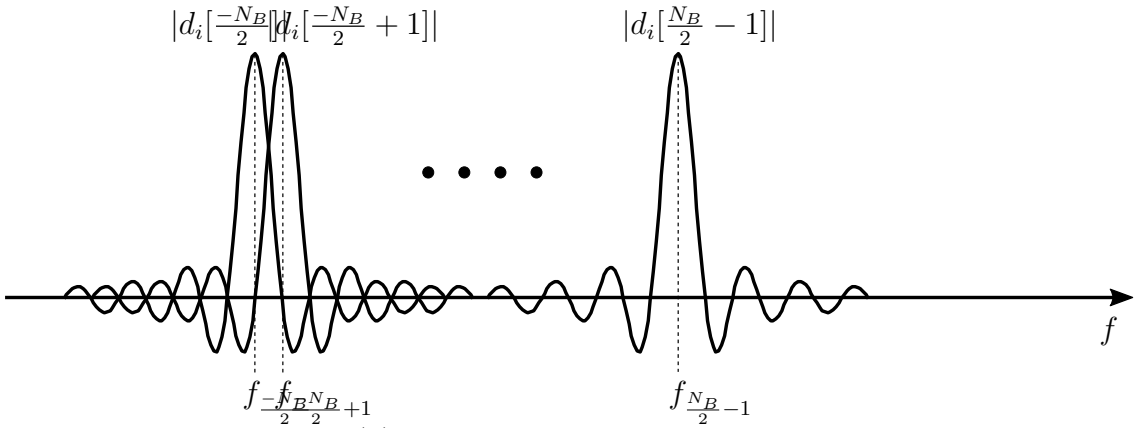
One method that can be used to reduce the number of taps required to equalize the received signal is to transmit the data over a number of carriers, often these carriers are referred to as sub-carriers. This reduces the data rate on each carrier and causes $T_{sym} > \tau_{max}$ for all sub carriers. This makes each carrier channel a narrowband channel that can be equalized using a single tap.

Two methods can be used to allocate spectrum for each channel, or in other words, choose the sub-carrier spacing. The first is frequency division multiplexing (FDM), which implements each channel such that the spectrum of no channel overlaps with

another, this is shown in Figure 4.2a. The drawback of this is that if a number of channels are to be implemented in parallel then the bandwidth required for this is quite high. A second method that can be used to choose the sub-carrier spacing is such that at a sub-carrier frequency of interest no other sub-carrier affects the value. Or, the effect of all other sub-carriers at the sub-carrier of interest is zero. Using this method, much narrower bandwidths can be obtained when compared to FDM [4].



(a) Spectrum of an FDM signal.



(b) Spectrum of an OFDM signal.

Figure 4.2: Comparison of the spectrum of an FDM and OFDM signal. This shows that by using orthogonal sub-carriers the bandwidth required for the transmitted signal can be reduced.

To show how OFDM is implemented, it can be assumed that the length of each block of data, $\mathbf{d}[i]$, to be transmitted is N_B, where i is the block index. So, N_B sub-carriers with frequencies f_k are required for each block of data to be transmitted and each sub-carrier is modulated using a single symbol $d_i[k]$. If the length of each block of data is chosen to be T_B, then the spectrum of each sub-carrier can be computed using the Fourier transform of the time-domain sub-carrier, $s_{i,k}(t)$, modulated using the

symbol $d_i[k]$. If $d_i[k]$ is represented using a scaled rect function and this rect function is modulated onto a sub-carrier with a frequency of f_k, then the resulting signal $s_{i,k}(t)$ can be written as

$$s_{i,k}(t) = d_i[k]\frac{1}{T_B}\text{rect}\Big(\frac{t}{T_B}\Big)e^{j2\pi f_k t}. \tag{4.2}$$

The frequency domain representation of $s_{i,k}(t)$ is then

$$S_{i,k}(f) = \mathscr{F}\left\{s_{i,k}(t)\right\} = d_i[k]\frac{\sin(\pi T_B(f-f_k))}{\pi T_B(f-f_k)} = d_i[k]\text{sinc}\big(T_B(f-f_k)\big). \tag{4.3}$$

This time domain to frequency domain conversion is shown in Figure 4.3.

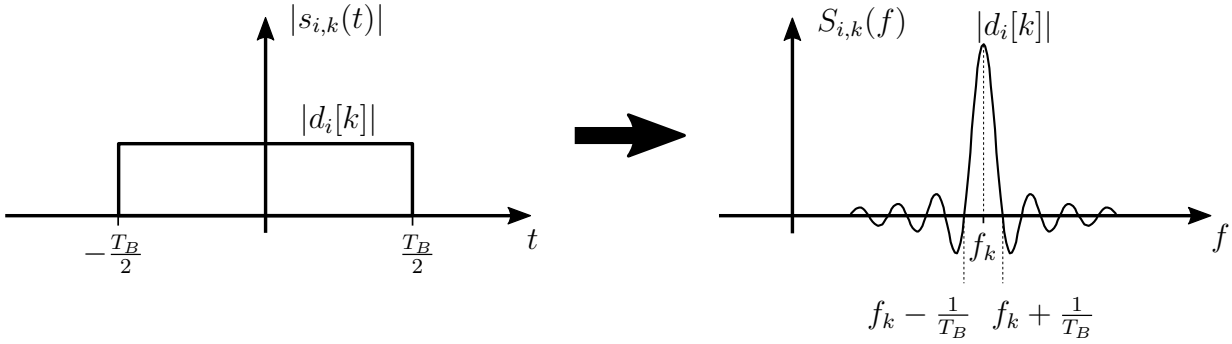


Figure 4.3: Conversion of the time domain sub-carrier, $s_{i,k}(t)$, to the frequency domain representation, $S_{i,k}(f)$.

So in order to ensure that no other sub-carrier interferes at the k^{th} sub-carrier, the sub-carrier frequencies can be chosen to be orthogonal over the period of interest, T_B. Using the frequency domain representation of the k^{th} sub-carrier in Figure 4.3, $S_{i,k}(f)$, it can be seen that this corresponds to a sub-carrier spacing of $\frac{1}{T_B}$. Therefore, the sub-carrier frequencies can be written as

$$f_k = f_0 + \frac{k}{T_B} \quad \text{where} \quad k = -\frac{N_B}{2}, \ldots, 0, \ldots, \frac{N_B}{2} - 1 \tag{4.4}$$

where f_0 can be chosen to place the center of the spectrum of the OFDM signal in different locations.

With this spacing, over a block duration of T_B all frequencies will be orthogonal to each other. The continuous spectrum of an OFDM transmitted signal can then be

written as

$$S_i(f) = \sum_{k=-N_B/2}^{N_B/2-1} d_i[k]\text{sinc}\big(T_B(f - f_k)\big). \tag{4.5}$$

and is shown in Figure 4.2b, which when compared to the FDM spectrum shows that an OFDM signal will have a smaller bandwidth.

The baseband transmitted time domain OFDM signal can be written as a sum of all the orthogonal sub-carriers

$$s_i(t) = \sum_{k=-N_B/2}^{N_B/2-1} d_i[k]e^{j2\pi f_k t} \quad \text{for} \quad 0 \leq t \leq T_B. \tag{4.6}$$

If $s_i(t)$ is sampled with a sampling frequency of $F_s = \frac{N_B}{T_B}$, then the time can be expressed as

$$t = \frac{n}{F_s} = \frac{nT_B}{N_B}, \quad \text{where} \quad n = 0, 1, 2, \ldots, N_B - 1. \tag{4.7}$$

The sampled OFDM signal can then be expressed as

$$s_i[n] = \sum_{k=0}^{N_B-1} d_i'[k]e^{j2\pi f_k' \frac{nT_B}{N_B}} = \sum_{k=0}^{N_B-1} d_i'[k]e^{\frac{j2\pi kn}{N_B}} \quad \text{for} \quad n = 0, 1, 2, \ldots, N_B - 1, \tag{4.8}$$

where f_k in (4.6) has been replaced with

$$f_k' = \frac{k}{T_B} \quad \text{where} \quad k = 0, 1, \ldots, N_B - 1, \tag{4.9}$$

and $d_i'[k]$ represents $d_i[k]$ circularly shifted to the left by $\frac{N_B}{2}$. f_k' still contains all of the same frequencies as f_k, assuming $f_0 = 0$, due to aliasing when the frequencies are sampled at the sampling frequency, F_s. The data symbols, $d_i[k]$, are circularly shifted to ensure the data matches the proper sub-carrier frequencies.

With the proper sampling of the OFDM baseband time domain signal the conversion from the frequency domain to time domain can be recognized as the inverse discrete Fourier transform (iDFT) of the data symbols in the i^{th} block. The transmitter can therefore be implemented efficiently using the inverse fast Fourier transform (iFFT). Similarly, the received time domain OFDM baseband signal can be received using the discrete Fourier transform (DFT), which can also be implemented efficiently using the

fast Fourier transform (FFT) [41]. A block diagram of a simple OFDM transmitter and receiver is shown in Figure 4.4.

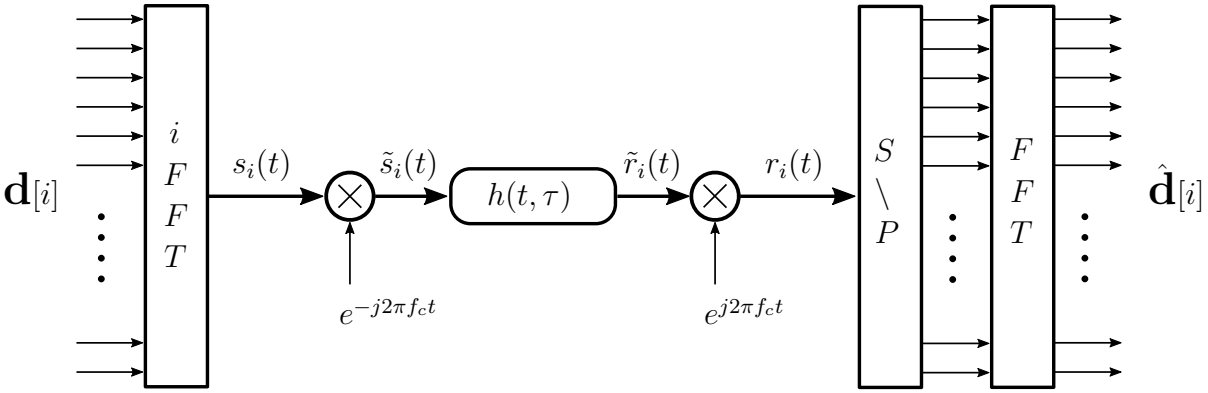


Figure 4.4: A complete OFDM system transmitting a single block of data, $\mathbf{d}[i]$, using a carrier frequency, f_c. S \P indicates a serial to parallel converter which takes the baseband time domain received signal, $r_i(t)$, and turns it into a parallel stream of data for the FFT operation.

For the Sections that follow it is also convenient to express the OFDM modulation and demodulation process in terms of matrix operations. The received baseband block of data in the frequency domain can be represented as

$$\hat{\mathbf{d}}[i] = \mathbf{F}_{N_B}\mathbf{H}^{(t)}\mathbf{F}^{\dagger}_{N_B}\mathbf{d}[i] + \mathbf{v}, \tag{4.10}$$

where $\hat{\mathbf{d}}[i]$ represents an estimate of the i^{th} data block. $\mathbf{F}_{N_B}$ is the $N_B \times N_B$ DFT matrix with the (i^{th}, j^{th}) entry

$$\frac{1}{\sqrt{N_B}} e^{-j\frac{2\pi ij}{N_B}} \quad \text{where} \quad 0 \leq i, j \leq N_B - 1. \tag{4.11}$$

$\mathbf{F}^{\dagger}_{N_B}$ is the Hermitian transpose of the DFT matrix which forms the iDFT. $\mathbf{v}$ is additive white noise, and $\mathbf{H}^{(t)}$ is the baseband time-domain channel matrix that represents the convolution of the transmitted baseband time domain signal with the channel.

One of the important features of OFDM is that as long as $\mathbf{H}^{(t)}$ is an $N_B \times N_B$ circulant matrix with (i^{th}, j^{th}) entries $h((i-j) \bmod N_B)$, then the channel matrix will be diagonalized in the frequency domain. Here, $h(n)$ is the baseband channel of length L_h sampled at the symbol rate and padded with $N_B - L_H$ zeros [42],

$$\mathbf{h} = [h(0), h(1), \ldots, h(L_h - 1), 0, \ldots, 0]^T. \tag{4.12}$$

If this is the case, then Eqn. (4.10) can be simplified using the fact that

$$\mathbf{F}_{N_B}\mathbf{H}^{(t)}\mathbf{F}^{\dagger}_{N_B} = \mathbf{D}_H, \tag{4.13}$$

where $\mathbf{D}_H$ is the diagonal matrix

$$\mathbf{D}_H = \begin{bmatrix} H(0) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & H(N_B-1) \end{bmatrix} \tag{4.14}$$

with entries

$$H(k) = \sum_{n=0}^{L_h-1} h(n)e^{-j\frac{2\pi nk}{N_B}}, \quad k = 0, 1, \ldots, N_B - 1. \tag{4.15}$$

Eqn. (4.10) then becomes

$$\hat{\mathbf{d}}[i] = \mathbf{D}_H\mathbf{d}[i] + \mathbf{v}. \tag{4.16}$$

This shows that when the channel matrix is circulant each sub-carrier can be treated as a narrowband channel with no other sub-carriers affecting its value. Therefore, each sub-carrier has a constant complex scaling term that can be equalized using a simple single tap equalizer. As an example, for the k^{th} sub-carrier the scaling factor would be $H(k)$, and so the equalization factor using a zero forcing equalizer would be $\frac{1}{H(k)}$ to obtain the original transmitted symbol [4]. Another way of stating this is that as long as the sub-carriers remain orthogonal when passed through the channel, each sub-carrier can be demodulated using a single tap equalizer.

If the channel matrix is not circulant in the time domain, then the received signal becomes

$$\hat{\mathbf{d}}[i] = \mathbf{H}\mathbf{d}[i] + \mathbf{v}, \tag{4.17}$$

where $\mathbf{H}$ is the channel matrix in the frequency domain that does not only contain diagonal entries and so accounts for the interference among sub-carriers.

While OFDM does provide a method to overcome delay-spread channels by extending the length of each symbol on a sub-carrier, the delay-spread as well as time-variance of the channel can still affect the performance of an OFDM system by destroying

the orthogonality of the sub-carriers. So in order to equalize effectively the effects of delay-spread and time-variance on an OFDM signal need to be understood.

4.1.2 The Effect of Time-Varying Channels on OFDM Signals

In the Section above it was shown that if the channel matrix can be considered to be circulant, then the task of equalizing a received OFDM signal is quite simple and involves a single tap equalizer for each sub-carrier. However, a time-varying and delay-spread channel ruins the orthogonality of the sub-carriers and leads to a non-circulant channel matrix. This channel can be represented in the passband as

$$\tilde{h}(t,\tau) = \sum_{c=1}^{N_c} a_c \delta(\tau - \tau_c(t)). \tag{4.18}$$

This is the same as the model in Chapter 2 (Eqn. (2.20)), with the time-scaling replaced by a time-varying delay so it can still be represented as a channel impulse response.

This channel model is often simplified using two assumptions. The first is that the length of a single OFDM block, T_B, is much smaller than the coherence time of the channel. So, for a single OFDM block $a_c(t)$ is approximately constant. The second assumption is that over a single block of data the change in delay is approximately linear, which is equivalent to saying the velocity along each path arrival remains constant. This is the same assumption made in Chapter 2 to produce the received signal in Eqn. (2.20). With these assumptions, the CIR becomes

$$\tilde{h}(t,\tau) = \sum_{c=1}^{N_c} a_c \delta(\tau - \tau_c(t)), \tag{4.19}$$

where

$$\tau_c(t) = \tau_c - \frac{v_c}{c_w}(t - \tau_c) = \tau_c + (1 - \alpha_c)(t - \tau_c). \tag{4.20}$$

The channel can also be expressed at baseband by accounting for the effect that the time-variance and delay has on the carrier frequency before down conversion [43]. The

baseband channel is then

$$h(t,\tau) = \sum_{c=1}^{N_c} a_c \delta(\tau - \tau_c(t)) e^{-j2\pi f_c \tau_c(t)}. \tag{4.21}$$

The time-varying channel with delay-spread causes two main problems for OFDM signals, inter-block interference (IBI) and inter-carrier interference (ICI). IBI is similar to ISI encountered with single carrier systems. If the symbols in Figure 4.1 are replaced with OFDM blocks, $\mathbf{d}[i]$, then any delay-spread will cause the OFDM blocks to interfere with each other. Again, the extent of this interference is dictated by the delay-spread of the channel, τ_{max}.

A simple way to mitigate IBI is to insert a guard interval between each transmitted OFDM block of length T_G seconds. As long as $T_G > \tau_{max}$, then no IBI will be encountered. While this is an effective way to mitigate IBI, ICI can still occur due to the delay spread and so the symbols inserted in this guard interval need to be chosen carefully to avoid this.

ICI occurs when the sub-carriers lose their orthogonality and the channel matrix loses it's circulant property in the time domain, and thus its diagonal property in the frequency domain. This happens both due to delay-spread as well as time-variance. The mechanism by which the orthogonality is lost is similar in both cases, however it is useful to separate the two as the solutions to overcome each are different.

Section 4.1.2.1 deals with how delay-spread causes ICI, and then presents two methods to overcome this ICI by choosing specific symbols to fill the guard interval and by decoding properly. In that discussion, it is assumed that the delays do not vary with time and $\tau_c(t) = \tau_c$ in Eqns. (4.19) and (4.21). Section 4.1.2.2 then presents how ICI is caused by the time-variance of the channel by allowing the delays to vary with time. While no solution is presented, a method of modeling this ICI is discussed. If the relevant parameters of the cluster arrivals can be estimated, then this model can be produced and used to equalize the ICI caused by time-variance.

4.1.2.1 ICI Due to Delay-Spread

Figure 4.5a shows the reception of an OFDM block through a delay spread channel at time $t = T_B$, where the reception began at time $t = 0$. In this simple two tap channel, the block has been fully received on the first tap, but has not yet been fully received

from the second tap.

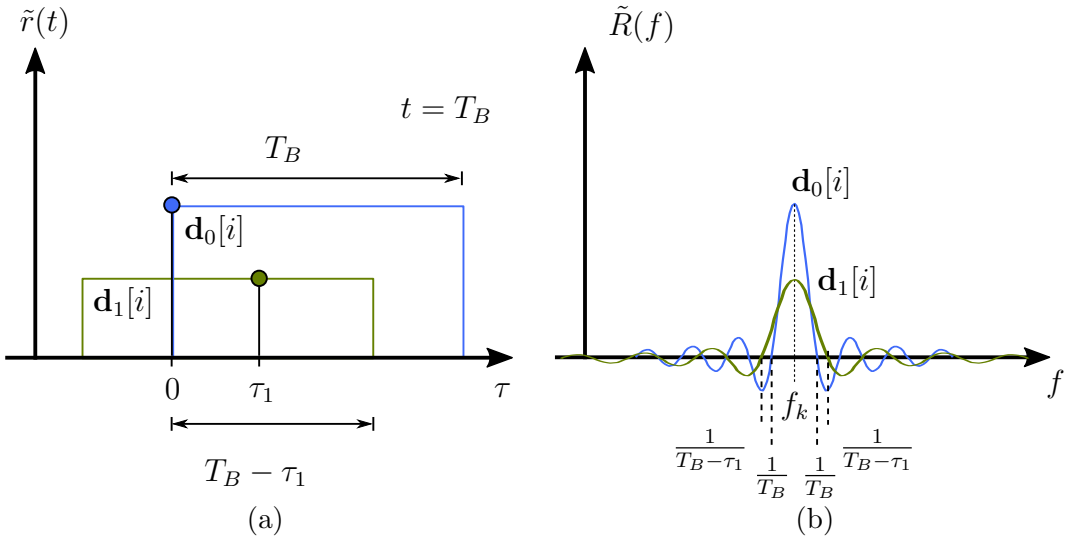


Figure 4.5: Reception of an OFDM block on a two-tap delay-spread channel. a) shows the delay between receptions and b) shows the effect of this delay in the frequency domain for a single sub-carrier, k, if the received time domain signal $\tilde{r}(t)$ only uses samples from $t = [0, T_B]$.

With an OFDM period of T_B, the first tap contains the full OFDM block period while the second tap only contains a period of $T_B - \tau_1$ over the same duration. Figure 4.5b shows the received signal in the frequency domain for a single sub-carrier, k. The frequency domain representation shows that while the first tap arrival maintains orthogonality with all other sub-carriers by maintaining the proper zero locations, the second tap arrival will not. This is due to the fact that the period of the received signal on the second tap is reduced and so the zeros are shifted to $\frac{1}{T_B-\tau_1}$. This means that the signal from the second tap will affect neighboring sub-carriers, causing ICI.

The ICI due to delay spread can be removed if the symbols that fill the guard interval are chosen properly. This means that the guard interval can be chosen to maintain the orthogonality of the received OFDM block. To do this, two methods can be used, cyclic prefixing or zero-padding. Both methods are discussed in detail in Appendix E and are summarized here.

Cyclic prefixing is the most common technique used and can be found in most textbooks that discuss OFDM [4], [41], [44]. In a cyclic prefixed signal the last N_{cp} baseband samples, corresponding to a length of T_G seconds, of the time domain signal are prepended to the beginning of the signal. This makes the transmitted signal

cyclical as long as $N_{cp} > M$, where

$$M = \frac{\tau_{max}}{T_s} \tag{4.22}$$

is the discrete delay-spread. If the received signal is sampled properly, then a full block reception will be received over each tap and the ICI due to delay spread will be eliminated. The main drawback with cyclic prefixing is that it involves filling the guard interval with symbols, which consumes power to transmit [42].

While this isn't a problem for channels with small delay spreads, for channels with large delay spread like the UAC this involves transmitting a substantial amount of excess power. A less common method to overcome this ICI is called zero-padding, which appends N_{zp} zeros to the end of the baseband time domain transmit signal.

If the the received signal is processed using the overlap-and-add (OLA) algorithm summarized in Appendix E [42], [45], then it can be shown that this will also remove the ICI introduced by the delay-spread channel without the need to transmit excess power. Again, N_{zp} must be larger than the discrete delay spread of the channel, M. The main drawback of using the OLA algorithm is that it increases the noise floor due to the extra number of time-domain samples that are required [45].

So by choosing the guard interval such that $T_G > \tau_{max}$ both IBI and ICI due to the delay-spread in the channel can be eliminated. In order to eliminate ICI, either cyclic prefixed OFDM or zero-padded OFDM can be used. This makes the channel matrix diagonal in the frequency domain leading to simple single-tap equalization for all sub-carriers. However, in channels such as the UAC where the delay-spread can be larger, zero-padded OFDM is often preferred as it requires less power for transmission. Throughout the rest of this chapter it is assumed zero-padded OFDM is used unless otherwise noted. This assumption means that ICI due to delay-spread should not be present, and only ICI due to time-variance needs to be accounted for and modeled in the channel estimation process.

4.1.2.2 ICI Due to Time-Variance

While ICI due to delay spread is removed using a guard interval, ICI due to time-varying channels cannot be removed as easily because the channel matrix can no longer be made circulant. The time-varying channel cannot be made circulant as the channel is no longer stationary, and so at each instant in time the delays have moved.

A shift in time will therefore be accompanied by a shift of the delays. To account for ICI from a time-varying channel, the effect of this time-variance on each sub-carrier needs to be taken into account.

The problem caused by a TV channel is best understood through the diagrams in figure 4.6. The wavefront of a propagating wave travels at the speed of sound in water, c_w. If both the receiver and transmitter are stationary and the transmitted signal has a period T_B, then the received signal will also have a period T_B with a time delay between transmission and reception. However, if there is movement between the receiver and transmitter represented by a constant velocity, v_c, then this period will change.

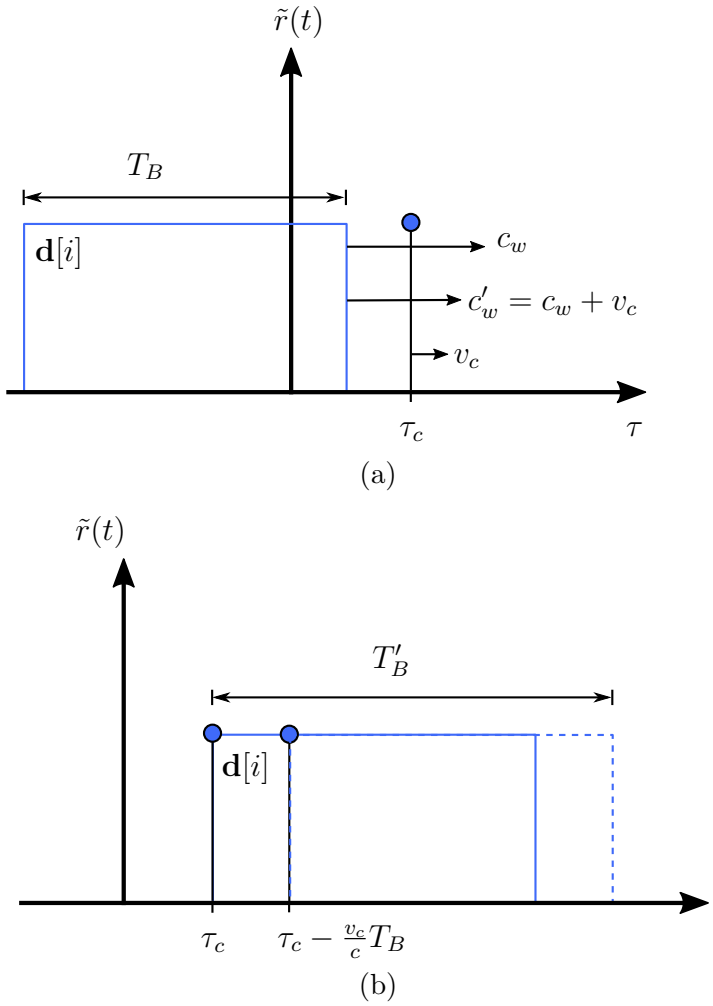


Figure 4.6: Images showing the effect of a time varying channel on a received OFDM signal. (a) shows the relative velocity of the path and the OFDM block. (b) shows how the received signal is dilated as the path moves away during reception.

Figure 4.6a shows that when the cluster path is moving, the effective speed of prop-

agation of the wave along the path is

$$c'_w = c_w + v_c. \tag{4.23}$$

In figure 4.6a the velocity is negative due to the increasing path length, following the definition in Chapter 1. When the wave front finally reaches the moving receiver then the reception period is increased, or dilated, due to the new velocity of the wavefront caused by the moving tap. This is shown in Figure 4.6b. The new period can be related to the cluster path time-scaling factor, α_c, by

$$T'_B = \frac{T_B}{\alpha_c}. \tag{4.24}$$

If the path length is decreasing and $\alpha_c > 1$ then the period of the received signal decreases. If the path length is increasing and $\alpha_c < 1$ then the period of the received signal increases. The effect of a channel with time-varying delays on an OFDM signal can be found by passing a baseband ZP-OFDM signal, $s_{zp,i}(t)$ through the baseband, time-varying channel in Eqn. (4.21) to form the received signal

$$r_{zp,i}(t) = s_{zp,i}(t) * h(t,\tau) = \sum_{c=1}^{N_c} A_c s_{zp,i}(\alpha_c(t-\tau_c))e^{-j2\pi f_c(1-\alpha_c)t}, \tag{4.25}$$

where

$$A_c = a_c e^{-j2\pi f_c \alpha_c \tau_c}. \tag{4.26}$$

To determine the effect that this channel has on each sub-carrier, it is informative to look at a single reception path in the frequency domain. The c^{th} path in the continuous frequency domain can be represented as

$$R_{zp,i,c}(f) = A_c \frac{1}{\alpha_c} S_i(f')e^{-j2\pi\tau_c f'} \quad \text{where} \quad f' = \frac{f + (1-\alpha_c)f_c}{\alpha_c} \tag{4.27}$$

and $S_i(f)$ is the spectrum of the transmitted OFDM signal for the i^{th} data block. This assumes no ICI has been introduced due to the delay-spread.

To determine the effect that the time-varying channel has on the received signal it is

useful to expand the argument of $S_i(f')$. Specifically,

$$S_i(f') = \sum_{k=-N_B/2}^{N_B/2-1} d_i[k]\text{sinc}\big(T_B(f' - f_k)\big). \tag{4.28}$$

and so the argument of the sinc function becomes

$$T_B(f' - f_k) = T_B\left(\frac{f + (1-\alpha_c)f_c}{\alpha_c} - f_k\right) = \frac{T_B}{\alpha_c}\left(f + (1-\alpha_c)f_c - \alpha_c f_k\right). \tag{4.29}$$

The time-varying channel therefore has two main effects on each sub-carrier which are shown in Figure 4.7. The first is that the sinc function has been frequency-scaled, or the zeros of each sub-carrier have been shifted from integer multiples of $\frac{1}{T_B}$, excluding $f_k = 0$, to integer multiples of $\frac{\alpha_c}{T_B}$. The second effect is that each sub-carrier has been shifted by $(1-\alpha_c)f_c - \alpha_c f_k$. This shift is due to the time-scaling of the carrier frequency after down conversion and the time-scaling of the sub-carrier itself.

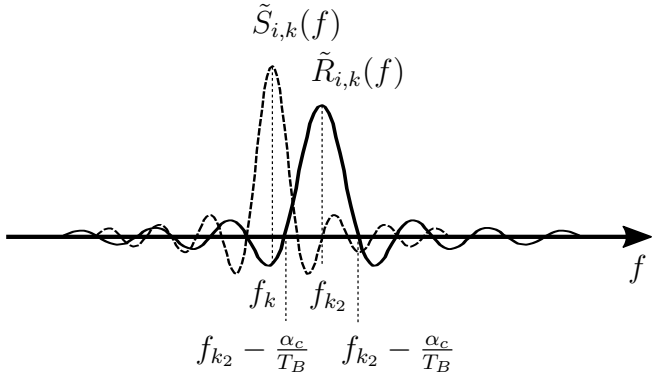


Figure 4.7: The effect of a time-varying channel on a single sub-carrier. The sub-carrier is shifted to $f_{k_2} = (1-\alpha_c)f_c - \alpha_c f_k$ and has been scaled with new zero crossings at $\frac{\alpha_c}{T_B}$.

The cumulative effect of a time-varying channel is that if the received signal in the frequency domain, $R_{zp,i}(f)$, is sampled at the original frequencies, f_k, due to the use of the FFT to receive the signal, the output will be a combination of neighboring data symbols weighted by their shifted and frequency-scaled sinc functions. This constitutes ICI. So, in order to produce the channel matrix in the frequency domain, $\mathbf{H}$, the effect of all sub-carriers on the sub-carrier of interest needs to be taken into account. The (m, k) entry of the frequency domain channel matrix for a single cluster

arrival can then be defined as

$$[\mathbf{H}_c]_{m,k} = A_c \frac{1}{\alpha_c} S_{i,k}(f'_m) e^{-j2\pi\tau_c f'_m} \quad \text{where} \quad f'_m = \frac{f_m + (1-\alpha_c) f_c}{\alpha_c}, \tag{4.30}$$

m is the sub-carrier of interest and k is the neighboring sub-carrier. The channel matrix produced by each path is then completely described by the complex amplitude, delay, and time-scaling factor associated with the path.

For a complete representation of the channel, the interference of all sub-carriers with the sub-carrier of interest can be calculated. This corresponds to $-\frac{N_B}{2} \leq k \leq \frac{N_B}{2} - 1$. However, if α_c is close to 1 then most of the entries in $\mathbf{H}_c$ will be very small with most of the energy concentrated on the main diagonal, and only a few off diagonal values of $\mathbf{H}_c$ will be of relevance. So, $\mathbf{H}_c$ can be approximated by considering only D off diagonals, which corresponds to $m - D \leq k \leq m + D$. When k is limited to contain only D off diagonals, the new matrix can then be expressed as $\mathbf{H}_{c,D}$.

With respect to the sub-carriers, this means that only the ICI from the D closest sub-carriers on either side of the sub-carrier of interest is considered when modeling the effect of the channel. The channel matrix can be extended to include all cluster arrivals in a multipath environment simply by summing over all cluster path arrivals

$$\mathbf{H}_D = \sum_{c=1}^{N_c} \mathbf{H}_{c,D}. \tag{4.31}$$

So, the UAC channel presents both IBI and ICI due to delay-spread and time-variance. The IBI and ICI due to delay-spread can be removed simply by inserting a guard interval between OFDM blocks and choosing the symbols that fill this guard interval properly. The ICI caused by the time-variance of the channel however can not be removed as easily. In order to accurately model the effect that a time-varying channel has on a received OFDM block the effect of the frequency scaled and shifted sub-carriers on the sub-carrier of interest need to be taken into account. This model can then be used to equalize the received symbol. The main difficulty now is how the parameters of the channel can be estimated so that this channel matrix can be produced and used to eliminate the ICI introduced by the time-varying channel.

4.2 Sparse Estimation and Its Application to OFDM in the Underwater Acoustic Channel

Sparse estimation, which is also commonly referred to as compressed sensing, is a method for solving underdetermined linear systems when it is assumed that the solution is sparse in some domain. The goal of this Section is to introduce sparse estimation and explore how it can be used to estimate the underwater acoustic channel. Section 4.2.1 explores the basic theory behind sparse estimation. Section 4.2.2 then presents a practical algorithm that can be used to recover sparse solutions. Lastly, Section 4.2.3 presents a formulation of a sparse estimation problem for the underwater acoustic channel.

4.2.1 Sparse Estimation Basics

Sparse estimation is one method of solving an underdetermined system where the number of measurements, m, is less than the number of parameters under consideration, n, having previous knowledge that the solution is sparse in some domain. The general formulation of a sparse estimation problem is

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}, \tag{4.32}$$

where $\mathbf{y} \in \mathbb{R}^m$ is the measurement or observation vector, $\mathbf{A} \in \mathbb{R}^{m \times n}$ is the sensing matrix, often called an overcomplete dictionary in estimation problems with each column referred to as an atom, $\mathbf{x} \in \mathbb{R}^n$ is the unknown sparse atom weight vector to be estimated and $\mathbf{v}$ is a noise vector of length m that is assumed to be independent and identically distributed (i.i.d) [46].

The problem with an underdetermined system is that generally there are infinitely many solutions, assuming $\text{rank}(\mathbf{A}) = m$. The null space of of the dictionary matrix $\mathbf{A}$ is therefore non-trivial and no solution is unique. There are methods to solve these problems, such as minimum-norm solutions that solve for the unknown vector $\mathbf{x}$ by minimizing the selected norm [46]. The minimum squared ℓ_2 norm is given by

$$\hat{\mathbf{x}}_{MN} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_2^2 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y}, \tag{4.33}$$

where $\|\mathbf{x}\|_2^2$ is the squared Euclidean norm and measures the squared distance from

the origin to the point defined by $\mathbf{x}$. The solution to this equation is the value of $\mathbf{x}$ that minimizes the squared distance and can be found using [46]

$$\hat{\mathbf{x}}_{MN} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{y}. \tag{4.34}$$

$\hat{\mathbf{x}}_{MN}$ is a best estimate with respect to the ℓ_2 norm, but does not take into account the known sparsity of the solution. Sparse approximation on the other hand produces a unique solution provided a noise free measurement using the prior information that the solution must be sparse. A signal is said to be sparse if most of the elements in the vector are zero, or if most of the energy of the signal is represented by only a few elements. So, for a vector $\mathbf{x}$ of length n, the number of non-zero elements in the vector must be much less than n. This can be expressed as

$$\|\mathbf{x}\|_0 = |\text{supp}(\mathbf{x})| \ll n, \tag{4.35}$$

where the ℓ_0 norm is simply the number of elements in, or cardinality of the set. A signal is said to be k-sparse if $\|\mathbf{x}\|_0 \leq k$. If the signal is known to be sparse, the best approach to finding the coefficients is through an ℓ_0 optimization

$$\hat{\mathbf{x}}_{\ell_0} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y}, \tag{4.36}$$

where the minimization is now with respect to the number of elements in $\mathbf{x}$ rather than the squared distance as in (4.33).

Provided a noise free measurement, sparse estimation should provide a sparse, unique solution, $\mathbf{x}$, which has k or fewer elements. In order to prove this generally involves studying the structure of the sensing, or dictionary matrix $\mathbf{A}$, which is covered in most introductory texts to sparse approximation or compressed sensing such as [46] and [47]. The focus instead now shifts to an algorithm that promotes a sparse solution.

4.2.2 Sparse Estimation Using Orthogonal Matching Pursuit

There are a number of algorithms used to solve sparse estimation problems. The one discussed in this Section and that is used to test the estimation technique presented in Section 4.2.3 is Orthogonal Matching Pursuit (OMP). OMP is a greedy algorithm

that attempts to solve

$$\hat{\mathbf{x}}_{\ell_0} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \leq \epsilon. \tag{4.37}$$

The term *greedy* indicates that at each step the algorithm makes an optimal choice with respect to some criterion without regard for future iterations [47].

It should be noted that for this estimator, the general problem that we are trying to solve is the approximation, $\hat{\mathbf{x}}$, of the sparse vector $\mathbf{x}$ given the noisy measurement vector $\mathbf{y}$, as shown in Eqn. (4.32). This explains the substitution of the condition $\mathbf{A}\mathbf{x} = \mathbf{y}$ with the ℓ_2 norm $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \leq \epsilon$ in (4.37) [47], where ϵ is some small number that allows the estimate to deviate slightly from the actual solution due to the noise in the measurement.

Orthogonal matching pursuit is an iterative algorithm that updates the estimation $\hat{\mathbf{x}}$ at each iteration by choosing the column of $\mathbf{A}$, $\mathbf{a}$, that maximally reduces the error between the approximation and the measurement. For a 1-sparse signal that contains only 1 element, the approximation of $\mathbf{y}$ corresponds to a single weighted column of the dictionary matrix $\mathbf{A}$. The error can then be written as

$$e(j) = \min_{z} \|z\mathbf{a}_j - \mathbf{y}\|_2^2 = \|\mathbf{y}\|_2^2 - \frac{(\mathbf{a}_j^T \mathbf{y})^2}{\|\mathbf{a}_j\|_2^2}, \quad \text{for} \quad j = 1, 2, ..., n, \tag{4.38}$$

where $\mathbf{a}_j$ is the j^{th} column in $\mathbf{A}$ and z is the columns weight that maximally reduces the error.

The index j is the column of $\mathbf{A}$ that reduces the error most significantly and is found using

$$j = \arg\min_{j} e(j) = \arg\max_{j} \frac{|\mathbf{a}_j^T \mathbf{y}|}{\|\mathbf{a}_j\|_2^2}. \tag{4.39}$$

The index j, or atom, in $\mathbf{A}$ that minimizes the residuals is then given a weight in the approximation vector $\hat{\mathbf{x}}$, while all the other elements of the estimation vector remain at zero [46]. For a 1-sparse solution this would constitute the end of the algorithm. However, for a k-sparse signal this procedure can be repeated k times. As this procedure is repeated the measurement vector can no longer be used to determine the error in Eqn. (4.39). Instead, a residual vector, $\mathbf{r}$, is created which shows the residual error between the estimation $\mathbf{A}\hat{\mathbf{x}}$ and $\mathbf{y}$. So the residual error is

$$\mathbf{r} = \mathbf{y} - \mathbf{A}\hat{\mathbf{x}}. \tag{4.40}$$

It should be noted that if the exact sparsity of the signal is unknown then the algorithm can continue to iterate until the power in the residual is below some predetermined value, $\|\mathbf{r}\|_2 \leq \epsilon$.

The OMP algorithm functions by first finding the column that minimizes the current residual vector as in (4.39). This column is then added to the support set of $\hat{\mathbf{x}}$, $\mathbf{\Lambda}$, which contains the indices that correspond to the non-zero elements in $\hat{\mathbf{x}}$. A matrix $\mathbf{A}_\Lambda$ is then constructed using the columns that correspond to the set $\mathbf{\Lambda}$, and the measurement vector is orthogonally projected onto this matrix. The weights from this projection, $\hat{\mathbf{x}}_\Lambda$ can then be mapped backed to their respective positions in $\hat{\mathbf{x}}$ and the residual vector can be updated to include the new element in $\hat{\mathbf{x}}$. The algorithm can then continue until $\|\mathbf{r}\|_2 \leq \epsilon_{OMP}$ or the number of elements required in $\hat{\mathbf{x}}$, k, is satisfied. The benefit of this technique is that the elements in $\hat{\mathbf{x}}$ are always orthogonal to the columns that have already been selected and so no column is ever selected twice [47].

The algorithm can be summarized as

Algorithm 1: Summary of OMP algorithm.

Data: Record $\mathbf{y}$

Initialize: $\mathbf{r} = \mathbf{y}$, $\hat{\mathbf{x}} = \mathbf{0}$, $\mathbf{\Lambda} = \emptyset$, $k = 0$

while *stopping criterion not met* **do**

$j = \arg\max_j \frac{|\mathbf{a}_j^T \mathbf{y}|}{\|\mathbf{a}_j\|_2^2}$

$\mathbf{\Lambda} = \mathbf{\Lambda} \cup j$

$\hat{\mathbf{x}}_{\mathbf{\Lambda}} = \arg\min_v \|\mathbf{A}_{\mathbf{\Lambda}}\mathbf{v} - \mathbf{y}\|_2^2$

$\mathbf{r} = \mathbf{y} - \mathbf{A}_{\mathbf{\Lambda}}\hat{\mathbf{x}}_{\mathbf{\Lambda}}$

$k = k + 1$

end

Return: $\hat{\mathbf{x}}$

Where in the algorithm $\hat{\mathbf{x}}_{\mathbf{\Lambda}}$ indicates the entries of $\hat{\mathbf{x}}$ at rows $\mathbf{\Lambda}$ and $\mathbf{A}_{\mathbf{\Lambda}}$ indicates the submatrix in $\mathbf{A}$ created using only the columns specified by the indices in $\mathbf{\Lambda}$.

The stopping criterion can be chosen to either be a set number of elements, k, in $\hat{\mathbf{x}}$, or it can also be chosen so that the residual error is minimized below a certain threshold, $\|\mathbf{r}\|_2^2 < \epsilon_{OMP}$. Using either stopping criterion, if the solution is in fact sparse than the OMP should produce a sparse estimate, $\hat{\mathbf{x}}$, provided the proper dictionary matrix is used.

4.2.3 Sparse Directional Underwater Acoustic Channel Estimation

With the theory of sparse estimation briefly introduced along with a basic algorithm that can be used to recover sparse estimates, in this Section it will be applied to a directional channel estimation problem using an OFDM system. Section 4.2.3.1 deals with how a non-directional sparse channel estimate can be produced using a known transmitted OFDM block. Section 4.2.3.2 then deals with how pilot, or known information can be inserted among the unknown data in order to estimate the channel on a block-by-block basis. Section 4.2.3.3 concludes by introducing how directional information can be attained from the channel estimate using the output channels from a modal beamformer.

4.2.3.1 Non-Directional Sparse Channel Estimate

A non-directional sparse channel estimate can be produced by first looking at the baseband representation of the received OFDM data block. This was described in Section 4.1.1 as

$$\hat{\mathbf{d}} = \mathbf{H}\mathbf{d} + \mathbf{v}, \tag{4.41}$$

where $\mathbf{d}$ and $\hat{\mathbf{d}}$ are the transmitted and received data blocks, respectively, $\mathbf{H}$ is the channel matrix in the frequency domain that implements the effects of a time-varying channel, and $\mathbf{v}$ is the additive noise. The block index, $[i]$, has been dropped from the data blocks for notational convenience as we only consider the estimation for a single block.

Ideally, we would like to create an estimate of $\mathbf{H}$, knowing the received data $\hat{\mathbf{d}}$ and some, or all of the data in $\mathbf{d}$. However, the current form of the problem doesn't show explicitly how the channel can be estimated, and so does not show the sparsity in the solution.

In Chapter 2 it was shown that due to the low speed of propagation of acoustic waves in water as well as the sparsity of reflectors in the environment, the UAC channel can be considered sparse with respect to delay since only a few, N_c, cluster path arrivals occur during a reception. In Section 4.1.2 it was also shown that in order to properly model a time-varying delay-spread channel three parameters of each cluster path

are required, the complex amplitude, A_c, the initial delay, τ_c, and the time-scaling factor, α_c. Therefore each cluster path can be represented as a triplet of information $\{A_c, \tau_c, \alpha_c\}$ [20], [21], [48].

If the channel is represented as a sum of all possible delay, time-scale pairs each with a complex amplitude, A_c, then only a few of these pairs should have non-zero amplitudes due to the sparsity of the channel. To make this problem tractable both the path delays and time-scaling terms need to be discretized [20], [21], [48], [19]. The time-delay can be discretized into N_τ values following

$$\tau \in \left\{0, \frac{T_B}{\lambda N_B}, \frac{2T_B}{\lambda N_B}, ..., T_G\right\}, \tag{4.42}$$

where $\frac{T_B}{N_B}$ is the baseband time resolution of the OFDM system, T_G is the length of the guard interval which corresponds to the maximum delay resolution, and λ is an integer oversampling factor, where $\lambda \geq 1$.

The time-scaling factor can also be discretized using N_α discrete values as

$$\alpha \in \{-\alpha_{max}, -\alpha_{max} + \Delta\alpha, ..., -\alpha_{max} - \Delta\alpha, \alpha_{max}\}, \tag{4.43}$$

where $\Delta\alpha = \frac{N_\alpha - 1}{2\alpha_{max}}$ and α_{max} is chosen based off of the maximum velocity expected to be encountered.

Discretizing the delay and time-scales as shown above leads to $N_\tau N_\alpha$ possible combinations, where $N_\tau N_\alpha \gg N_c$. Due to this sparsity, in the estimation process only a few values of A_c corresponding to specific delay, time-scale pairs should be identified. The received data symbol can then be expressed in terms of a sum of OFDM blocks being received along each possible cluster path as

$$\hat{\mathbf{d}} = \sum_{i=0}^{N_\tau - 1} \sum_{j=0}^{N_\alpha - 1} x[iN_\alpha + j]\mathbf{H}_{\{\tau_i, \alpha_j\}, D}\mathbf{d} + \mathbf{v}, \tag{4.44}$$

where $x[\cdot]$ represents an element in $\mathbf{x}$, which is the weight vector containing A_c for each cluster path and is assumed sparse. $\mathbf{H}_{\{\tau_i, \alpha_j\}, D}$ corresponds to the channel matrix estimate formed using Eqn. (4.30) with the values $\tau_c = \tau_i$, $\alpha_c = \alpha_j$, and using D off diagonals.

Using the representation of the received signal in Eqn. (4.44), Eqn. (4.41) can then

be restated as

$$\hat{\mathbf{d}} = \mathbf{A}\mathbf{x} + \mathbf{v}, \tag{4.45}$$

where $\mathbf{x}$ is the $N_\tau N_\alpha \times 1$ sparse weighting vector and $\mathbf{A}$ is the $N_B \times N_\tau N_\alpha$ dictionary matrix of the form [20], [21], [48],

$$\begin{aligned}\mathbf{A} = \Big[&\mathbf{H}_{\{\tau_0,\alpha_0\},D}\mathbf{d},\ \ldots\ ,\mathbf{H}_{\{\tau_0,\alpha_{N_\alpha-1}\},D}\mathbf{d},\ \ldots\ , \\ &\mathbf{H}_{\{\tau_{N_\tau-1},\alpha_0\},D}\mathbf{d},\ \ldots\ ,\mathbf{H}_{\{\tau_{N_\tau-1},\alpha_{N_\alpha-1}\},D}\mathbf{d}\Big].\end{aligned} \tag{4.46}$$

This is a sparse estimation problem as the goal becomes finding the few entries in $\mathbf{x}$ that are non-zero and estimating their value. Specifically, this sparse estimation problem corresponds to finding the few paths over which data has traveled that, when summed together, produce the best approximation of $\hat{\mathbf{d}}$. The entries in $\mathbf{x}$ provide the required triplet of information for each channel cluster path arrival as the actual value corresponds to A_c and the position of this value in the vector $\mathbf{x}$ corresponds to a specific delay, time-scale pair.

4.2.3.2 Insertion of Pilot Data Into an OFDM Block

An OFDM block is a grouping of N_B data symbols which are modulated and transmitted using OFDM. In OFDM systems there are two main forms of channel estimation, block estimation and comb estimation. Block estimation is used in slow fading channels where the channel remains almost the same for many blocks of transmitted data. Therefore, in block estimation a whole OFDM block is transmitted with known data, or pilot symbols, and then the channel estimation is done using this full block and the estimate is used to equalize the unknown data blocks that follow [49]. With block estimation, Eqn. (4.45) can be used directly.

Unfortunately, the UAC is fast-fading as the channel varies from block to block. So a new channel estimate is required for each block of data and this requires pilot symbols to be inserted with the unknown data in each OFDM block in order to estimate the channel. In most systems where the channel is considered time-invariant over the block duration, the optimal solution is simply to insert known data at equi-distantly spaced sub-carriers [50].

When this method is used in a time-varying channel that produces ICI, the main

problem that occurs is that the unknown data sub-carriers interfere with the pilot sub-carriers. The added interference on the pilot symbol reduces the estimation accuracy [11]. A second problem that occurs is that as the channel is time-varying, the ICI pattern needs to be estimated as well in order to estimate the time-scaling factor introduced by the channel. So the pilot design needs to be able to facilitate ICI estimation as well.

The pilot design method used here accounts for the interference from the data sub-carriers and the need to estimate the ICI pattern by inserting a guard band of zeros, or nulls, on sub-carriers on either side of the pilot sub-carriers [19]. The length of this guard interval is chosen based off of the degree of the ICI encountered in the channel, D, as presented in Section 4.1.2. Specifically, $2D$ zeros, or nulls, can be placed on either side of the pilot. In doing so, there are D zeros on either side of the pilot that should have no interference from the unknown data and so these nulls can be used to estimate the ICI pattern. This pilot method will be referred to as the zero-padded (ZP) pilot method and is shown in figure 4.8.

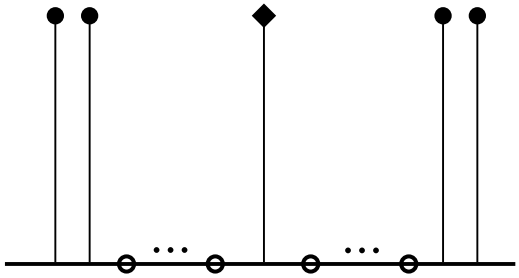


Figure 4.8: Block layout for the ZP pilot method. Diamonds represent pilot symbols, open circles represent nulls and closed circles represent the unknown data symbols.

The ZP method has a length of $4P_N + P_P$ symbols, where P_P is the number of pilot symbols, or tones, used and P_N is the number of null tones placed on either side of the pilot. P_N has been used to replace D as the number of nulls can be chosen to be more or less than the number of off diagonals in the channel matrix. To create the final OFDM block of data, the pilot symbols, nulls and unknown data symbols are combined to form a sub-frame.

This sub-frame is chosen to have a length N_{SF} and so the number of data symbols that can be included in each sub-frame is limited to $N_s = N_{SF} - (4P_N + P_P)$. To ensure that at least one symbol is included in each frame, N_{SF} should be chosen so that $N_{SF} > 4P_N + P_P$. Lastly, assuming N_B is chosen to be a power of 2, then N_{SF} should also be chosen to be a power of 2 so that an integer number of sub-frames are

used to produce the final transmit OFDM block. An example of a sub-frame with $N_{SF} = 16$ using the ZP pilot method with $P_P = 1$ and $P_N = 2$ is shown in figure 4.9

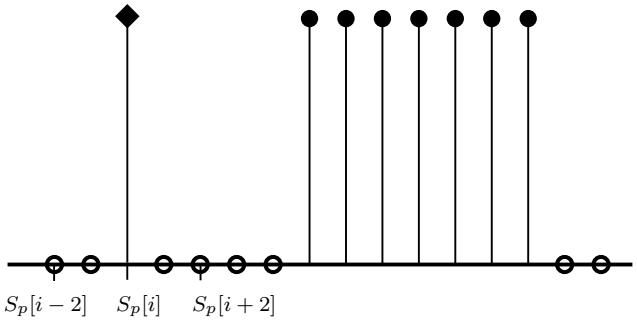


Figure 4.9: OFDM subframe with $N_{SF} = 16$ using the ZP pilot method with $D = 2$. The pilot indices, S_p, are included to show how the pilots are chosen for a ZP pilot symbol.

In order to use the final OFDM block for channel estimation, the pilots as well as the unknown data need to be incorporated into the channel estimation in Eqn. (4.45). To do this, the dictionary needs to be created so that it only relies on the pilot symbols. This is done by replacing the full block of transmitted data, $\mathbf{d}$, originally used in Eqn. (4.46) to create the dictionary matrix, with the OFDM block $\mathbf{d}_S$. $\mathbf{d}_S$ is the transmit OFDM block in which all the data symbols have been set to their average value of zero and only the pilot tones have been retained [50]. The nulled dictionary matrix is then produced as

$$\begin{aligned}\mathbf{A}_N = \Big[&\mathbf{H}_{\{\tau_0,\alpha_0\},D}\mathbf{d}_S,\ \dots\ ,\mathbf{H}_{\{\tau_0,\alpha_{N_{\alpha-1}}\},D}\mathbf{d}_S,\ \dots\ , \\ &\mathbf{H}_{\{\tau_{N_\tau-1},\alpha_0\},D}\mathbf{d}_S,\ \dots\ ,\mathbf{H}_{\{\tau_{N_\tau-1},\alpha_{N_{\alpha-1}}\},D}\mathbf{d}_S\Big],\end{aligned} \tag{4.47}$$

where the subscript N indicates the data has been nulled out.

The compressed sensing problem can then be expressed as

$$\hat{\mathbf{d}} = \mathbf{A}_N\mathbf{x} + \mathbf{v}, \tag{4.48}$$

where $\hat{\mathbf{d}}$ is still the measured block of data at the output of the FFT receiver used in Eqn. (4.45). Therefore, the only difference between (4.48) and (4.45) is the dictionary that is used.

4.2.3.3 Directional Sparse Channel Estimate

Until now, the directionality of the received signal has been ignored and the channel estimate has focused on estimating the complex amplitude, time delay and time-scaling factor associated with each cluster path. Adding directionality to the wideband channel estimate relies on using the outputs of the modal beamformer in order to add a fourth parameter to the channel estimation, the cluster path angle-of-arrival, ϕ_c.

Two observations are required in order to incorporate modal beamfroming easily into the channel estimate. The first is that as long as the amplitude mode outputs from the UCCA , d, $\cos(n\phi)$, and $\sin(n\phi)$, maintain their frequency invariance over the frequency range of interest then in the frequency domain the effect of each mode on an OFDM signal can be represented as a constant scaling term. The second observation is that with these frequency invariant modes, modal beamforming can be related to sensor based narrowband beamforming techniques that make use of linear arrays operating at the array design frequency [35].

This means that techniques derived for angle of arrival estimates with linear arrays can be translated to a UCCA through the use of modal beamforming. The use of modes rather than sensors also presents two benefits. One benefit of working with modes of the UCCA rather than discrete sensors is that while for a linear array the unambiguous range only covers π radians, for a circular array the unambiguous range covers the full 2π radians. A second benefit is that while for a linear array these modes apply only to small frequency ranges, with a well designed modal beamformer these direction of arrival estimates can be applied over large frequency ranges.

The directional channel estimate can be related to the non-directional channel estimate by adding a new search parameter to Eqn. (4.44), the angle-of-arrival, ϕ. In order to be able to estimate the angle-of-arrival the modal channels also need to be incorporated into this estimate. The addition of this search parameter can be written as

$$\hat{\mathbf{d}}_n = \sum_{i=0}^{N_\phi-1} \sum_{j=0}^{N_\tau-1} \sum_{k=0}^{N_\alpha-1} x[iN_\tau N_\alpha + jN_\alpha + k] e^{jn\phi_i} \mathbf{H}_{\{\tau_j,\alpha_k\},D}\mathbf{d} + \mathbf{v}_n, \tag{4.49}$$

where $e^{jn\phi_i}$ is the mode of interest, $\hat{\mathbf{d}}_n$ is the received OFDM block on the n^{th} mode in the frequency domain, $\mathbf{v}_n$ is the noise on the received signal from the n^{th} mode, N_ϕ is the number of angles included in the search, and ϕ_i is the discretized angle-of-arrival

with values of

$$\phi \in \{-\pi,\ -\pi + \Delta\phi\ \ \ldots\ , \pi - \Delta\phi\ ,\ \pi\}\,. \tag{4.50}$$

$\Delta\phi$ can be chosen so that the set contains N_ϕ values as $\Delta\phi = \frac{N_\phi - 1}{2\pi}$.

Eqn. (4.49) shows that by adding directionality to the estimate, a new dictionary needs to be created. This new dictionary can related back to the original dictionary in Eqn. (4.46) by

$$\mathbf{A}_n = \left[e^{jn\phi_0}\mathbf{A},\ \ldots\ e^{jn\phi_{N_\phi - 1}}\mathbf{A}\right], \tag{4.51}$$

where $-b_{max} \leq n \leq b_{max}$. The complex modes used here can be produced using the real amplitude modes as

$$e^{\pm jn\phi} = \begin{cases} \cos(n\phi) \pm j\sin(n\phi), & \text{when } |n| > 0, \\ d, & \text{when } n = 0. \end{cases} \tag{4.52}$$

Using the new modal dictionaries, $\mathbf{A}_n$, along with the received modal outputs, $\hat{\mathbf{d}}_n$, the compressed sensing problem becomes the same as that in Eqn. (4.45), with the angle-of-arrival as a new output parameter. However, this estimate only applies to the output of one mode.

To create the final directional channel estimate a new sparse estimation problem can be produced by stacking the received data and modal dictionaries for each mode, and treating the channel estimate as a single estimate using all of the modal channels. The new complete modal (CM) estimation problem then becomes

$$\hat{\mathbf{d}}_{CM} = \mathbf{A}_{CM}\mathbf{x}_{CM} + \mathbf{v}_{CM}, \tag{4.53}$$

where

$$\hat{\mathbf{d}}_{CM} = \left[\hat{\mathbf{d}}^T_{-n_{max}},\ \ldots\ , \hat{\mathbf{d}}^T_0,\ \ldots\ , \hat{\mathbf{d}}^T_{n_{max}}\right]^T \tag{4.54}$$

is the stacked received modal vector and

$$\mathbf{A}_{CM} = \left[\mathbf{A}^T_{-n_{max}},\ \ldots\ , \mathbf{A}^T_0,\ \ldots\ , \mathbf{A}^T_{n_{max}}\right]^T \tag{4.55}$$

is the stacked modal dictionary matrix. The entries of the sparse weight vector from

Eqn. (4.49), $\mathbf{x}_{CM}$, now correspond to a quadruplet of information for each cluster path arrival: the complex amplitude, the delay, the time-scaling factor, and the direction of arrival, $\{A_c, \tau_c, \alpha_c, \phi_c\}$, respectively.

The directional sparse channel estimator presented here is similar to the one presented in [12], with a few key differences. First, the time-variance in [12] is represented using a constant Doppler shift rather than a time-scaling operation which better represents a wideband signal with significant time-variance. Second, here modal beamforming is used while in [12] a linear array is used with each sensor simply treated as wideband. While modal beamforming can be made inherently wideband, maintaining a constant aperture with a linear array for a wideband signal can be difficult and is not dealt with in [12]. Third, while the work in [12] works in the time domain, this estimator works in the frequency domain using OFDM signals.

Similar to the non-directional estimate presented in Section 4.2.3.1, the directional estimate can make use of the pilot design in Section 4.2.3.2 as well. However, the dictionary matrix defined in (4.51) for each channel now needs to be produced using the nulled dictionary matrix for each channel, $\mathbf{A}_N$, rather than the full block dictionary matrix, $\mathbf{A}$. These nulled modal dictionaries can then be used to produce a nulled complete modal dictionary that can be used with the received signal.

4.3 Simulation Results

This Section presents results from simulations that were carried out in order to confirm the theory introduced in this Chapter. Specifically, this involves looking at different methods for estimating the channel and then equalizing the received data.

Section 4.3.1 introduces the simulation setup that was used for these simulations, any additional information pertaining to a specific simulation is contained within that simulations Section. Section 4.3.2 presents results that show the effect of accounting for time-variance when equalizing a signal that has passed through a time-varying channel. Section 4.3.3 then presents results that focus on the performance gains that are attained using the sparse non-directional channel estimator to estimate a time-varying channel for equalization when compared to more commonplace channel estimation techniques for OFDM. Section 4.3.4 concludes the Chapter by showing that the directional sparse channel estimator is capable of estimating directionality and by demonstrating the performance gains that can be attained when a directional

channel estimate is used over a non-directional channel estimate.

4.3.1 Simulation Parameters

In order to run simulations to test the performance of the channel estimators presented in this Chapter, an OFDM system as well as a time-varying channel needed to be implemented. For these simulations the OFDM system used the following parameters:

- OFDM block length, $N_B = 512$ symbols
- OFDM block duration, $T_B = 250$ ms
- OFDM bandwidth, $B_{OFDM} = 2048$
- Gaurd interval, $T_G = 62.5$ ms (filled using zero-padding)
- Carrier frequency, $f_c = 4096$ Hz
- QPSK modulation of the sub-carriers
- Overlap-and-add (OLA) receiver
- Pilot sub frame with
  - Length $N_{SF} = 8$
  - Number of pilot symbols, $P_P = 2$
  - Number of nulls, $P_N = 1$
  - 2 data symbols per frame

In these simulations either a predefined channel or a stochastic channel was used. In both cases the time-varying channel was implemented using the TVFDL presented in Chapter 2. In the case of the predefined channel, the parameters are presented in their respective Sections. If a stochastic channel was used then the stochastic model in Chapter 2 was used with the following parameters:

- Number of cluster arrivals, $N_c = 6$
- Number of paths per cluster arrival, $N_p = 10$
- RMS delay, $\tau_{RMS} = 0.005$ s
- Max delay, $\tau_{max} = 0.01$ s
- Coherence time, $T_c = 2$ s

- Minimum velocity, $v_{min} = 0$ m/s
- Maximum velocity, $v_{max} = 15$ m/s
- Minimum angle of arrival, $\theta_{min} = -\pi$
- Maximum angle of arrival, $\theta_{max} = \pi$

For the simulations that required a directional receiver the same UCCA that was used in the experiment presented in Section 3.4 was used. So the simulated UCCA had 5-elements with a radius of 0.0381 meters. The directional channel was also implemented using a TVFDL capable of implementing the delays required for the different receive elements due to the different angles-of-arrival of the cluster paths.

Lastly, the figure of merit that is used in these simulations to determine how well a receiver functions is the bit error rate (BER), which compares the equalized received data to the actual transmitted data. If a single bit of the received data is wrong, the BER is increased by 1. The final BER is then normalized by total number of unknown data bits transmitted to achieve a BER between 0 and 1.

4.3.2 Consideration of ICI in Equalization

The first simulation that was run looked at how taking ICI into consideration affected the performance of a receiver. To do this, a channel was implemented that had a single cluster arrival with an amplitude of 1 and a delay of 0. The time-scaling factor associated with this cluster arrival was then varied, creating different levels of ICI. In Section 4.1.2 it was noted that most of the energy of the channel, and thus the ICI, can be represented using only D off diagonals of the channel matrix in the frequency domain, $\mathbf{H}$.

So, in this simulation in order to determine the effect that ICI has on a receiver, OFDM blocks were transmitted through the single tap channel with different values of α. The received signals were then equalized using oracle equalizers that produced channel matrices with different numbers of off diagonals considered. An oracle equalizer is an equalizer in which the channel information is perfectly known. Therefore, the only variable that effected the equalization was the amount of ICI considered by the equalizer, D.

For this simulation a block of random data was sent rather than the pilot OFDM block outlined in Section 4.3.1 as channel estimation was not required. The simulation was

run for α values between 0.95 and 1.05, which corresponds to velocities between -75 and 75 m/s. The BER at the output of each oracle equalizer using a different value of D was then calculated and the results are shown in Figure 4.10.

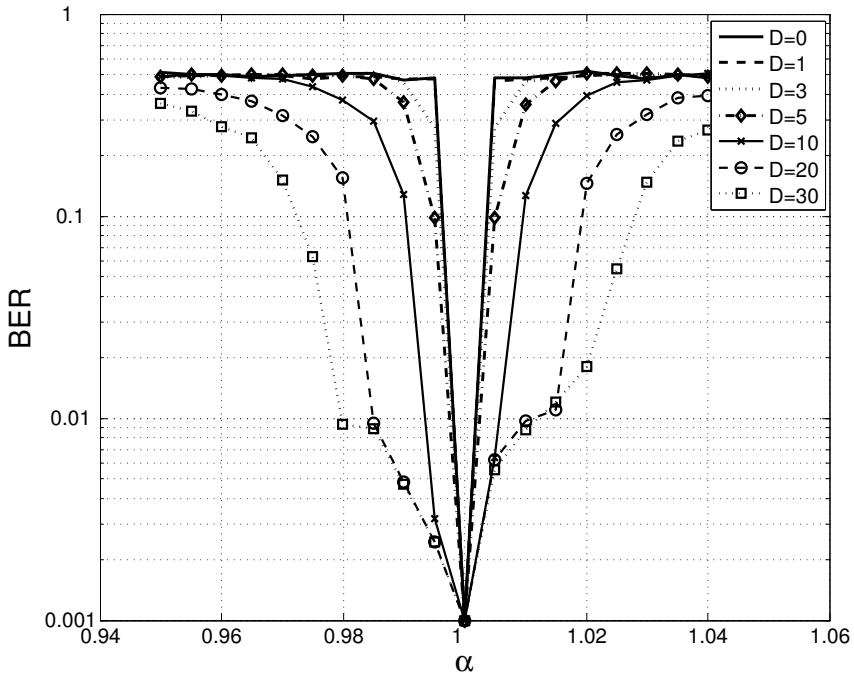


Figure 4.10: Plot of BER with respect to α to demonstrate the effect of considering ICI in the equalization process.

The plot in Figure 4.10 demonstrates that as the time-scaling increases or decreases away from 1, the number of off diagonals required to properly equalize the data also increases. This implies that as the time-scaling factor increases or decreases, the amount of ICI that needs to be accounted for in order to properly decode the received data increases as well.

While the speeds used in this simulation are high, Figure 4.10 shows that in order to maintain a BER below 0.1 even for alpha values between 0.99 and 1.01, at least 10 off diagonals should be used to equalize the channel. This plot also only shows the effect from one cluster path. More complex channel matrices formed by channels with more cluster arrivals may further degrade the performance of the receiver, and so more off diagonals may be useful. Lastly, no noise was added in this simulation which will further degrade performance depending on the amount of noise added.

The plot in Figure 4.10 also demonstrates that the model for ICI presented in Section 4.1.2 accurately represents the effect of time-variance of the channel on an OFDM signal as the channel mixing matrix was able to equalize the effects caused by the

single tap channel implemented using the TVFDL channel.

4.3.3 Sparse Non-Directional Channel Estimation

With the verification that the model for the effect of ICI due to time-variance on an OFDM signal was correct, the next simulation that was run looked at the performance of the non-directional sparse channel estimator, presented in Section 4.2.3.1, in a time-varying channel. To run this simulation the OFDM system and stochastic channel presented in Section 4.3.1 were used. In order for the the estimator to work however, a dictionary matrix needed to be produced.

The dictionary matrix that was used for this simulation had a delay oversampling factor of 4 and α values between 1 and 1.01 with a step size of 0.00025. This produced 82 τ values and 21 α values, leading to a dictionary matrix with 1722 columns.

The plot in Figure 4.14 shows the results from simulations that were run to determine the effect that noise has on a receiver using the non-directional sparse channel estimation for equalization. This method uses the OMP algorithm and the non-directional dictionary matrix presented above in order to estimate the channel. For each channel estimate a number of receivers were then implemented that produced the estimated channel matrix, $\mathbf{H}$, using a different number of off diagonals. The plot in Figure 4.14 therefore also shows the effect of using a different number of off diagonals in the equalization process.

For each signal-to-noise (SNR) value, 20 different channels were simulated and for each channel 20 noise iterations were run at the specified SNR. The BER for each receiver at each SNR value was then calculated as an average over all channels and noise realizations.

In Figure 4.11 a least-squares (LS) estimator as well as an oracle estimator are also included and used as receivers as a reference. The least-squares estimator, taken from [49], is most commonly used in frequency-selective channels and does not account for ICI due to time-variance. This estimate therefore acts an upper bound on the sparse-approximation estimators. The oracle estimator has perfect channel information and creates the channel matrix following (4.31). The oracle estimate therefore acts as a lower bound on the sparse approximation channel estimators.

Figure 4.11 shows three important things. The first is that the non-directional sparse channel estimation does improve the receiver performance dramatically by taking

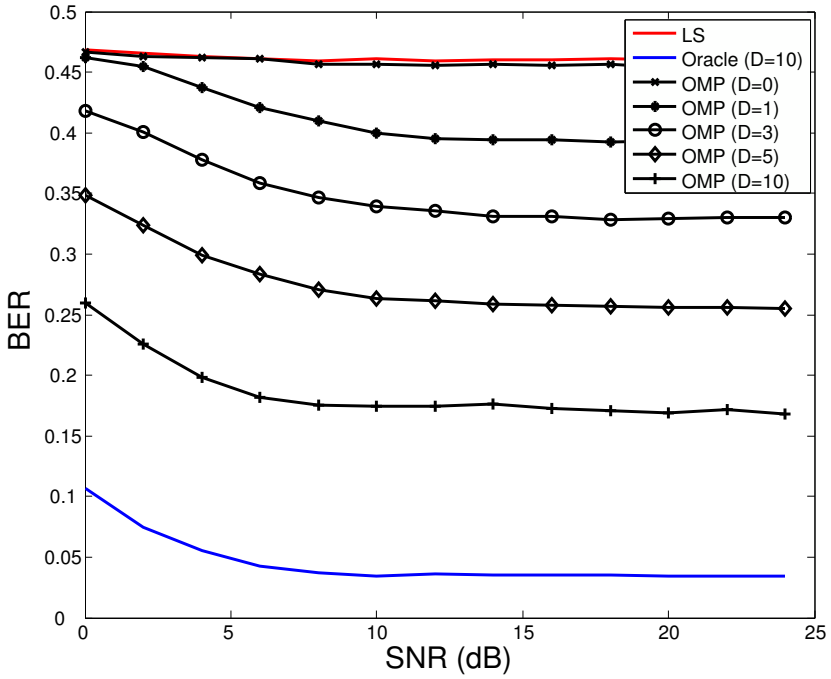


Figure 4.11: Plot of BER with respect to SNR to different number of off-diagonals considered. The least-squares estimator and oracle estimator are included as a reference.

into account ICI due to time-variance. The BER has been reduced from a value of 0.47 for the LS estimator, down to a value of close to 0.17 using the sparse estimate with $D = 10$. Secondly, similar to the results shown in Figure 4.10, this plot again shows that the performance of the system improves with the number of off diagonals considered. Although the BER is still above 0.1 with $D = 10$, this is uncoded data and with a proper forward error correction (FEC) code this BER could be reduced to 0 for each OFDM block.

Lastly, Figure 4.11 shows that a noise floor is encountered by both the oracle and sparse approximation receivers. In [50] it was noted that if the unknown data is close to the pilot symbols then the data symbols act as noise. For these simulations, the distance between the pilot and unknown data was only 2, which likely lead to the noise floor encountered in Figure 4.11.

So, the simulation results in Figure 4.11 show that taking into account ICI caused by the time-variance of the channel by using the non-directional sparse estimation technique improves the channel estimate. This leads to better performance of the receiver with respect to BER when compared to the classic LS technique.

4.3.4 Sparse Directional Channel Estimation

The last two simulations were run to show that the sparse directional channel estimator can estimate the angle-of-arrival of different cluster paths and also to show that the directional estimate can be used to further improve the performance of a receiver.

In order to produce directional information the UCCA outlined in Section 3.4.1 was used and the directional TVFDL was used to implement the channel. Using this UCCA, only modes up to order 1 could be produced. The OFDM system outlined in Section 4.3.1 was also used here in order to facilitate the directional channel estimate.

In order to use the directional sparse channel estimator a directional dictionary was created using Eqn. (4.55) and the non-directional dictionary specified in Section 4.3.2. The directional dictionary contained the -1, 0 and 1 complex modes and the direction of arrival, ϕ, was discretized between $-\pi$ and π in steps of $\frac{\pi}{36}$ radians.

For the first simulation a known channel was implemented using the directional TVFDL. Table 4.1 contains the parameters for each cluster arrival in the channel.

Table 4.1: Cluster parameters for the the directional channel implemented using a TVFDL.

Cluster Arrival	**Amplitude**	**Delay (ms)**	**Time-Scaling Factor (α)**	**Direction of Arrival (rads)**
1	1	0	1.0043	-0.2
2	0.5	4	1.00	0.2
3	0.4	6.5	1.0068	0
4	0.2	9	1.00	0.4

In order to determine the accuracy of the sparse directional channel estimator, a single OFDM frame was transmitted through the directional channel and a channel estimate was made. The SNR for this simulation was set to 20 dB, and using the channel estimate two plots were created. The first is a channel impulse response that shows the ideal response, using the values from Table 4.1, and the estimated response using 5 estimated paths. This plot therefore shows how accurate the estimate is with respect to amplitude and delay. This CIR is shown in Figure 4.12.

An interesting observation from Figure 4.12 is that while the delays associated with the estimated cluster paths are all similar to the actual delays, the amplitude of the first estimated cluster arrival is off and instead is broken into two arrivals. The reason for this is that while the delay of 0 for the first cluster falls into an exact discrete

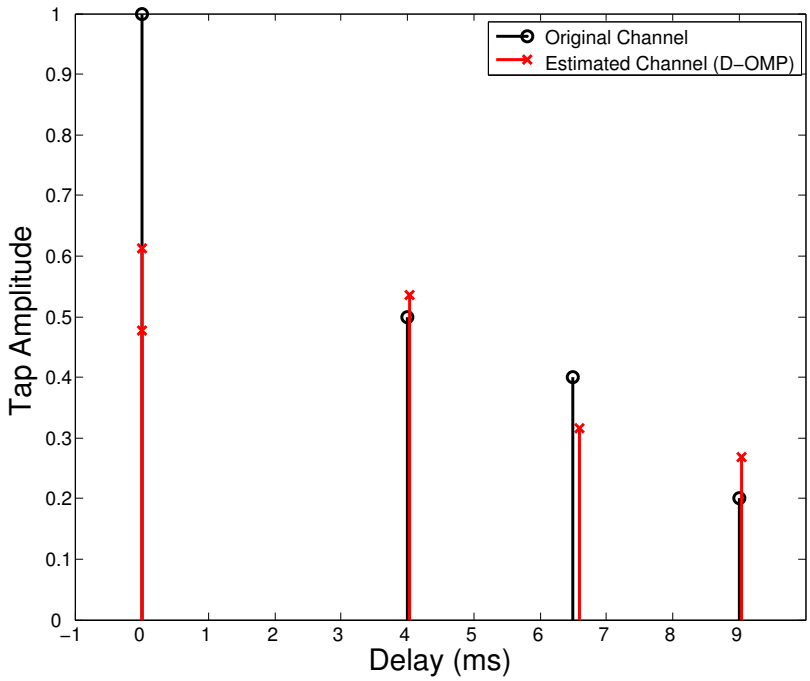


Figure 4.12: Channel impulse response of the actual cluster arrivals in the simulated directional channel and the estimated cluster arrival using the sparse directional channel estimation technique.

delay value, the angle-of-arrival and time-scaling factor do not. Therefore two taps are used with different angles-of-arrival and time-scaling factors in order to estimate the true angle-of-arrival and time-scaling values. To demonstrate this, a weighted average of the angle-of-arrival and time-scaling value was produced using the first two tap amplitudes, and the values came out to be -0.2127 and 1.0043, respectively. These values are very close, or exactly the same as the known channel values. It should also be noted that the estimated value of the time-scaling on the third channel tap, which is the only other tap to have time scaling, was 1.0070, which is close to the actual value of 1.0068.

While the CIR shows the accuracy of the estimate with respect to delay, the directional plot in Figure 4.13 shows the accuracy of the estimate with respect to angle-of-arrival. In this plot however, the error between the estimated angle-of-arrival and actual angle-of-arrival is larger than the error with respect to delay. This is due to the fact that the angle-of-arrival was discretized using steps of $\frac{\pi}{36}$, which only produces 73 discrete values to cover the full range of the UCCA from $-\pi$ to π. Better estimates would be produced if the step size was decreased at the expense of increasing the dictionary size.

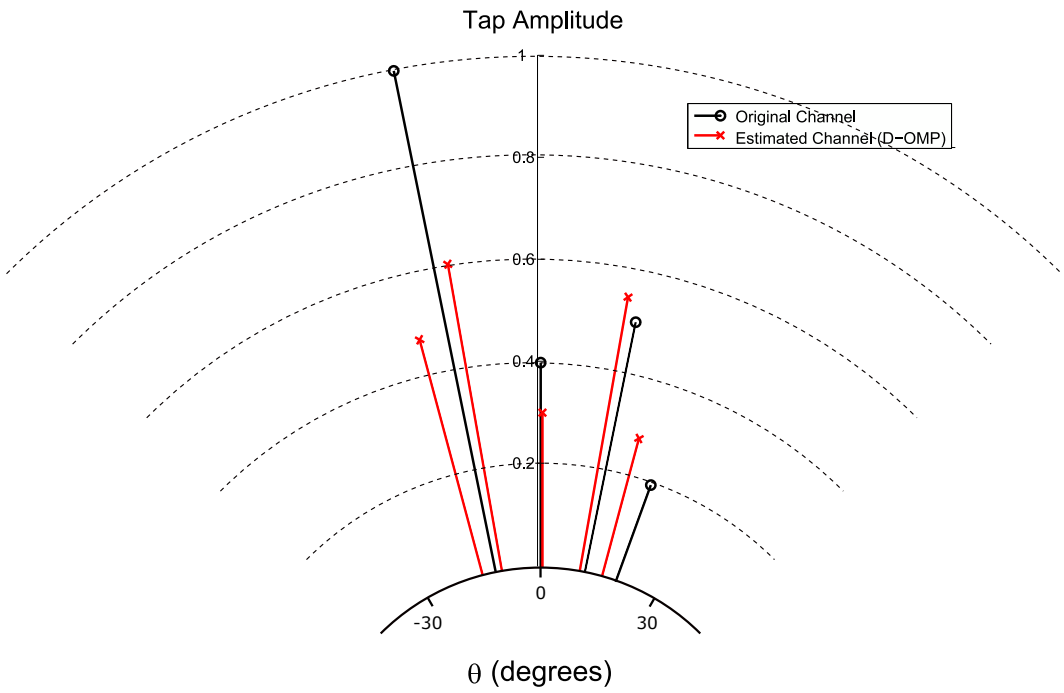


Figure 4.13: Plot of estimated cluster arrivals with respect to direction of arrival.

The simulations above show that the sparse directional estimation technique is capable of accurate estimates of the delay, time-scale factor and angle-of-arrival for different cluster arrivals. To show the benefit of adding the angle-of-arrival to the estimation using a modal beamformer, the same simulation that was run in Section 4.3.2 for the non-directional sparse channel estimation to determine the effect of noise on the BER was run again using both a directional and non-directional estimate.

Specifically, this last simulation looked at two receivers. The first estimated the channel using the non-directional sparse channel estimator and then equalized the data using this channel estimate with ten off diagonals. The second receiver estimated the directional channel and then beamformed in the direction of the strongest cluster arrival using a first-order Cardiod beamformer, presented in Section 3.3.3, with p=0.3. The beamformed output was then equalized using the directional channel estimate with the beam pattern overlayed in order to account for the effect of the beamformer. This channel estimate was then used to equalize the beamformed output using ten off diagonals as well. The results from this simulation are shown in Figure 4.14.

This simulation shows that by adding directionality to the estimate, at high SNR the directional receiver is able to perform better than the non-directional estimate by about 0.05. Adding directionality to the estimate has two benefits. First, by knowing

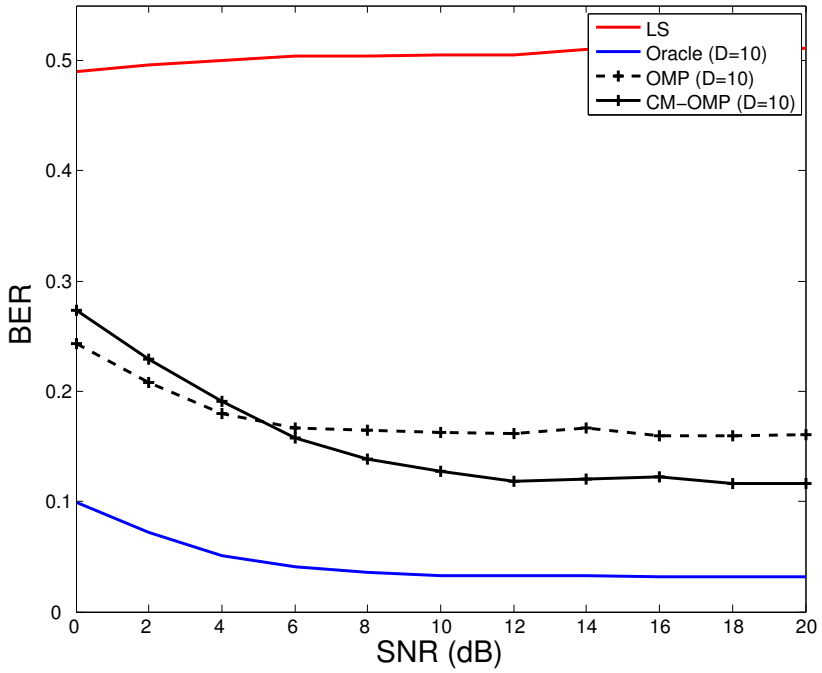


Figure 4.14: Plot of BER with respect to SNR for the non-directional and directional sparse channel estimate.

the direction of the strongest, and assumed direct, cluster arrival a beamformer can be steered in that direction to help reduce the effect of mulitpath interference. Secondly, adding the extra modes to produce the stacked received modal vector helps to refine the search through the dictionary as a higher correlation will be attained when a cluster arrival matches an atom in the directional dictionary.

The results of the directional channel estimate simulation could therefore be further improved by increasing the number of modes available by increasing the size of the UCCA. This would improve the beamforming ability of the UCCA, leading to a narrower main beam that further reduces the multipath interference. This would also help to improve the dictionary gain as the stacked received modal vector would increase in size as well.

These simulation therefore show that compared to classic techniques for OFDM equalization, accounting for the time-variance in the channel using sparse estimation techniques helps to improve the performance of the system. These simulations also showed that adding directionality to the estimator using a modal beamformer helps to further improve the channel estimate by reducing the multipath and increasing the diversity gain. The performance of a wideband OFDM receiver working in the UAC with mobility can therefore be greatly improved by accounting for time-variance and direc-

tionality at the receiver by using an appropriate sparse directional channel estimation technique.

Chapter 5

Conclusion

In this thesis an OFDM receiver design was proposed to enable reliable, high speed acoustic communication in time-varying underwater acoustic channels. Section 5.1 contains a summary of the results presented in this thesis that led to the final receiver design. Section 5.2 then discussed future work that could be done to build upon the work that has been presented. Section 5.3 concludes the thesis with final remarks.

5.1 Summary of Contributions

The three main contributions of this thesis were covered in Chapters 2, 3 and 4. The TVFDL was presented in Chapter 2 as a method of implementing a wideband UAC model that includes time-variance. While there are other options for implementing a stochastic channel model, Section 2.3 demonstrated analytically that the TVFDL can implement the time-scaling operation associated with the wideband UAC as well as a time-varying time-scaling operation. This was then confirmed through simulation in Section 2.4. While the channel model presented in Chapter 2 did not make use of a time-varying time-scaling factor, the ability to change this value during the simulation could make the channel model much more realistic for longer simulations. Specifically, it was noted that over longer time periods the channel will not be wide-sense stationary which could be modeled by changing the time-scaling factors and delays associated with each tap with respect to time. This is accomplished much more easily using the TVFDL than with other models.

Chapter 3 presented modal beamforming as a alternative method for beamform-

ing wideband underwater acoustic signals using a uniform concentric circular array (UCCA). Modal beamforming differs from the classic techniques outlined in Section 3.2 as first wideband orthogonal modes are produced, and then these modes are combined in order to produce the wideband beamformed output. The first benefit of this procedure is that once the wideband modes are produced any number of beams can be created with different beam patterns. The second benefit is that producing one beam only requires weighting each mode by a single value, and then summing these weighted modes.

In Section 3.3.2, amplitude mode processing was discussed as a method of weighting the elements of the UCCA using real values. This method produces the required modes for modal beamforming when real signals impinge on the array. Section 3.4 then presented experimental results that demonstrated that the required first-order modes could be produced underwater using a 5-element UCCA with amplitude mode processing. By deploying a receive array in a controlled tank, the theoretical patterns were validated using real measurements for different look directions, ϕ, and patterns, p. These experiments also demonstrated that the modes were wideband as the desired patterns were produced at 2, 5, and 9 kHz.

Chapter 4 combined the study of the wideband channel in Chapter 2 and the modal beamformer from Chapter 3 in order to produce a directional estimate of the time-varying channel that could be used to equalize a received OFDM signal. Section 4.3 then presented simulation results that demonstrated the benefit of accounting for time-variance and directionality in the channel when equalizing an OFDM signal.

In Section 4.2 an OFDM system was produced, and it was shown that by accounting for time-variance and using the angle-of-arrival of the signal, the performance of the system, measured using the bit error rate (BER) between the transmitted and received data, could be greatly increased when compared to classic OFDM equalization techniques such as least-squares (LS) equalization. It was also shown that accounting for different levels of ICI also improved the performance. A receiver using a sparse directional channel estimate in the frequency domain that considered the 10 closest sub-carriers on either side of the sub-carrier of interest to account for ICI ($D = 10$), produced a BER of approximately 0.12, while the LS receiver that used an LS estimate of the channel had a BER close to 0.47. This is an improvement in BER by a factor of 3 and with the proper forward error correction (FEC) method, this could be reduced further.

Using the results from Chapter 4 and the OFDM system parameters from Section 4.3.1, the rate of the system in a time-varying channel is only 410 bits/s. However, this system only uses 2 kHz of bandwidth. If the bandwidth of the signal is increased to 8 kHz by increasing the block size to $N_B = 2048$ then the rate increases by 4 to 1638 bits/s. If the carrier frequency is chosen to be around 5 kHz, then the maximum frequency of the signal is below 10 kHz and this can still be used with the UCCA outline in Section 3.4.

5.2 Future Work

There are a number of areas in which the work above can be expanded and improved upon and these are discussed briefly here. In Chapter 2 a stochastic channel model was presented for the time-varying channel. The drawback with this model is that while it makes use of measured results for some parameters, this measured data was attained using receivers and transmitters that were static. Therefore these parameters are extrapolated for use in the time-varying channel model. Currently, there has not been a comprehensive study of the underwater channel for wideband communication in a time-varying channel due to intentional velocity. This is one area where work could be done that would be relevant to a number of groups.

In Chapter 3 modal beamforming was discussed, but only in its 2-dimensional implementation using circular harmonics. Due to this, signals were limited to arrive in one plane. The work with circular harmonics can be extended to spherical harmonics [35], which would give the receiver an ability to estimate the angle-of-arrival of a signal in 3-dimensions. A spherical array would however be required in order to produce the desired spherical harmonics.

In Chapter 3 the frequency dependency of the modes was also discussed and it was shown that the modes encounter zeros at different frequencies. These zeros are impossible to equalize as it involves applying infinite gain. As a result, the bandwidth of the beamformer is limited. This could be solved two ways. First, by considering a baffled array, in which the elements are placed onto a rigid structure through which the impinging waves cannot propagate [35]. And second, by using directional elements in the array, such as vector sensors, rather than omni-directional elements [51].

In Chapter 4 there are three specific areas where improvements could be made. The first is in the sparse estimation algorithm used. In this work OMP was used, however

in [20] it was shown that basis pursuit algorithms may provide better results at the expense of algorithm complexity. The second area is in the dictionary matrix design and specifically how the sampling parameters are chosen for the delay, time-scaling factor and angle-of-arrival. For the simulation in Section 4.3 the sampling parameters were chosen in an ad-hoc fashion that showed good performance. However, for different OFDM parameters and channel conditions one would expect that an optimal choice in parameters could lead to a smaller dictionary size with optimal performance.

The last area where improvements could be made is in the design of the OFDM frame that is used for the block channel estimation. Specifically, this refers to how the pilots and nulls are distributed in the frame. In [50] a comparison is made between systematic and random pilot insertion into the OFDM block, however in that work no nulls are used. Therefore, there is room to explore the effect of placing the pilots and nulls in different fashions within the frame and with different densities. For different channels there should be a different number of pilot blocks required and in order to attain the best estimate there should also be an optimal way to place these pilots within the frame.

5.3 Concluding Remarks

In this thesis, a time-varying UAC channel was modeled and simulated and its effects on an OFDM system were studied. This study was carried out to continue to develop the knowledge base surrounding underwater communications, with a focus on high data rate communication when intentional transmitter and receiver velocity are involved in the channel. In this work it was shown that by accounting for the effects of time-variance on an OFDM signal and by including directionality in the channel estimate, the performance of the receiver can be greatly improved when compared to common OFDM equalization techniques. The data rate of the OFDM system simulated in this work could only be extrapolated to 1640 bits/s with a bandwidth of 8 kHz and a spectral efficiency of roughly 0.2 (bits/s/Hz). However, by addressing some of the future work outlined in Section 5.2, specifically increasing the usable bandwidth of the modal beamformer and optimizing the pilot insertion technique as well as the dictionary design, it can be expected that the data rate could be increased significantly.

Bibliography

[1] "Discovery of sound in the sea: The first practical uses of underwater acoustics," http://www.dosits.org/people/history/early1900/, Accessed: 2015-07-14.

[2] J. Heidemann, Wei Ye, J. Wills, A. Syed, and Yuan Li, "Research challenges and applications for underwater sensor networking," in *Wireless Communications and Networking Conference, 2006. WCNC 2006. IEEE*, April 2006, vol. 1, pp. 228–235.

[3] Ian F. Akyildiz, Dario Pompili, and Tommaso Melodia, "Underwater acoustic sensor networks: Research challenges," *AD HOC NETWORKS (ELSEVIER*, vol. 3, pp. 257–279, 2005.

[4] Tzi-Dar Chiueh and Pei-Yun Tsai, *OFDM Baseband Receiver Design for Wireless Communications*, Wiley Publishing, 2007.

[5] A. Molisch, *Wireless Communications*, Wiley, 2011.

[6] X Lurton, *An Introduction to Underwater Acoustics*, chapter 2, Springer Praxis Books, UK, 2 edition, 2010.

[7] R. J. Urich, *Principles of Underwater Sound*, vol. 1, chapter 5, McGraw Hill Book Company, 1 edition, 1975.

[8] J. Preisig, "Acoustic propagation considerations for underwater acoustic communications network development," in *Proc. ACM WUWNet.*

[9] Baosheng Li, Shengli Zhou, M. Stojanovic, L. Freitag, and P. Willett, "Multicarrier communication over underwater acoustic channels with nonuniform doppler shifts," *Oceanic Engineering, IEEE Journal of*, vol. 33, no. 2, pp. 198–209, April 2008.

[10] P.A. van Walree and R. Otnes, "Ultrawideband underwater acoustic communication channels," *Oceanic Engineering, IEEE Journal of*, vol. 38, no. 4, pp. 678–688, Oct 2013.

[11] C.R. Berger, Zhaohui Wang, Jianzhong Huang, and Shengli Zhou, "Application of compressive sensing to sparse channel estimation," *Communications Magazine, IEEE*, vol. 48, no. 11, pp. 164–174, November 2010.

[12] Sung-Hoon Byun, Woojae Seong, and Sea-Moon Kim, "Sparse underwater acoustic channel parameter estimation using a wideband receiver array," *Oceanic Engineering, IEEE Journal of*, vol. 38, no. 4, pp. 718–729, Oct 2013.

[13] A. Amar, Y. Buchris, and M. Stojanovic, "Angle-of-arrival-based detection of underwater acoustic ofdm signals," in *nternational Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2015.

[14] J. Presig and G. Deane, "Surface wave focusing and acoustic communications in the surf zone," *The Journal of the Acoustical Society of America*, 2004.

[15] *Wireless Communications Over Rapidly Time-Varying Channels*, Academic Press, 2011.

[16] E.J. Kelly and R.P. Wishner, "Matched-filter theory for high-velocity, accelerating targets," *Military Electronics, IEEE Transactions on*, vol. 9, no. 1, pp. 56–69, Jan 1965.

[17] L.G. Weiss, "Wavelets and wideband correlation processing," *Signal Processing Magazine, IEEE*, vol. 11, no. 1, pp. 13–32, Jan 1994.

[18] P.A. van Walree and R. Otnes, "Ultrawideband underwater acoustic communication channels," *Oceanic Engineering, IEEE Journal of*, vol. 38, no. 4, pp. 678–688, Oct 2013.

[19] S. Mason, C. Berger, Shengli Zhou, K. Ball, L. Freitag, and P. Willett, "An ofdm design for underwater acoustic channels with doppler spread," in *Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop, 2009. DSP/SPE 2009. IEEE 13th*, Jan 2009, pp. 138–143.

[20] Jianzhong Huang, C.R. Berger, Shengli Zhou, and Jie Huang, "Comparison of basis pursuit algorithms for sparse channel estimation in underwater acoustic ofdm," in *OCEANS 2010 IEEE - Sydney*, May 2010, pp. 1–6.

[21] C.R. Berger, Shengli Zhou, J.C. Preisig, and P. Willett, "Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing," *Signal Processing, IEEE Transactions on*, vol. 58, no. 3, pp. 1708–1721, March 2010.

[22] N.F. Josso, J.J. Zhang, A. Papandreou-Suppappola, C. Ioana, J.I. Mars, C. Gervaise, and Y. Stephan, "On the characterization of time-scale underwater acoustic signals using matching pursuit decomposition," in *OCEANS 2009, MTS/IEEE Biloxi - Marine Technology for Our Future: Global and Local Challenges*, Oct 2009, pp. 1–6.

[23] P. Hoeher, "A statistical discrete-time model for the wssus multipath channel," *Vehicular Technology, IEEE Transactions on*, vol. 41, no. 4, pp. 461–468, Nov 1992.

[24] P.A. van Walree, T. Jenserud, and M. Smedsrud, "A discrete-time channel simulator driven by measured scattering functions," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 9, pp. 1628–1637, December 2008.

[25] Cheng-Xiang Wang, M. Patzold, and Qi Yao, "Stochastic modeling and simulation of frequency-correlated wideband fading channels," *Vehicular Technology, IEEE Transactions on*, vol. 56, no. 3, pp. 1050–1063, May 2007.

[26] P. vna Walree, T. Jenserud, and R. Otnes, "Stretched-exponential doppler spectra in underwater acoustic communication channels," in *JASA Express Letters*.

[27] T. S. Rappaport, *Wireless Communications: Principles and Practice*, chapter 4, Prentice HAll PTR, Upper Saddle River, NJ, USA, 1996.

[28] S. Park, G. Hillman, and R. Robles, "A novel structure for real-time digital sample-rate converters with finite precision error analysis," in *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on*, Apr 1991, pp. 3613–3616 vol.5.

[29] S. Park, "Real-time pitch (frequency) shifting techniques," in *Consumer Electronics, 1991 IEEE International Conference on*, Jun 1991, pp. 296–297.

[30] T.I. Laakso, V. Valimaki, M. Karjalainen, and U.K. Laine, "Splitting the unit delay [fir/all pass filters design]," *Signal Processing Magazine, IEEE*, vol. 13, no. 1, pp. 30–60, Jan 1996.

[31] Julius O. Smith and P. Gossett, "A flexible sampling-rate conversion method," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84.*, Mar 1984, vol. 9, pp. 112–115.

[32] W. Liu and S. Weiss, *Wideband Beamforming: Concepts and Techniques*, Wiley, 2010.

[33] W. Liu, "Wideband beamforming for multipath signals based on frequency invariant transformation," *International Journal of Automation and Computing*, 2011.

[34] Q. Liu, *Wideband Digital Filter-and-Sum Beamforming with Simultaneous Correction of Dispersive Cable and Antenna Effects*, Ph.D. thesis, Virginia Polytechnic Institute, 2012.

[35] H Teutsch, *Wavefield Decomposition Using Microphone Arrays and Its Applications to Acoustic Scene Analysis*, Ph.D. thesis, Friedrich-Alexander-University, 2005.

[36] D.E.N Davies, *The Handbook of Antenna Design, Volume 2*, chapter 12, Peter Peregrinus Ltd., 1983.

[37] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic Press, 1999.

[38] Cherian P. Mathews and M.D. Zoltowski, "Eigenstructure techniques for 2-d angle estimation with uniform circular arrays," *Signal Processing, IEEE Transactions on*, vol. 42, no. 9, pp. 2395–2407, Sep 1994.

[39] J.A. McConnell and S.C. Jensen, "Forming first-and second-order cardioids with multimode hydrophones," in *OCEANS 2006*, Sept 2006, pp. 1–6.

[40] R.M.M. Derkx and K. Janse, "Theoretical analysis of a first-order azimuth-steerable superdirective microphone array," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 17, no. 1, pp. 150–162, Jan 2009.

[41] G. Li and G. Stuber, *Othogonal Frequency Division Multiplexing for Wireless Communication*, Springer, 2006.

[42] Baosheng Li, Shengli Zhou, M. Stojanovic, and L. Freitag, "Pilot-tone based zp-ofdm demodulation for an underwater acoustic channel," in *OCEANS 2006*, Sept 2006, pp. 1–5.

[43] David Tse and Pramod Viswanath, *Fundamentals of Wireless Communications*, Cambridge University Press, 2004.

[44] Andrea Goldsmith, *Wireless Communications*, Cambridge University Press, New York, NY, USA, 2005.

[45] B. Muquet, Zhengdao Wang, G.B. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions?," *Communications, IEEE Transactions on*, vol. 50, no. 12, pp. 2136–2148, Dec 2002.

[46] K. Hayashi, M. Nagahara, and T. Tanaka, "A user's guide to compressed sensing for communications systems," *IEEE Transactions on Communications*, vol. E96-B, no. 3, pp. 685–712, 2013.

[47] Yonina C. Eldar and Gitta Kutyniok, *Compressed Sensing: Theory and Applications*, Cambridge University Press, 2012.

[48] S.F. Mason, C.R. Berger, Shengli Zhou, K.R. Ball, L. Freitag, and P. Willett, "Receiver comparisons on an ofdm design for doppler spread channels," in *OCEANS 2009 - EUROPE*, May 2009, pp. 1–7.

[49] Yushi Shen and Ed Martinex, "Channel estimation in ofdm systems," Tech. Rep., Freescale Semiconductor, 2006.

[50] C.R. Berger, J. Gomes, and J.M.F. Moura, "Study of pilot designs for cyclic-prefix ofdm on time-varying and sparse underwater acoustic channels," in *OCEANS, 2011 IEEE - Spain*, June 2011, pp. 1–8.

[51] Nan Zou and Arye Nehorai, "Circular acoustic vector-sensor array for mode beamforming," *Signal Processing, IEEE Transactions on*, vol. 57, no. 8, pp. 3041–3052, Aug 2009.

[52] P. Walree, "Channel sounding for acoustic communications: techniques and shallow-water examples," Tech. Rep., Norwegian Defence Research Establishment (FFI), April 11, 2011.

[53] David Tse and Pramod Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, New York, NY, USA, 2005.

[54] J. Barry, E. Lee, and D. Messerschmitt, *Digital Communication*, chapter 2, Springer, 2005.

[55] Y. Zhao, W. Liu, and R. Langley, "A least squares approach to the design of frequency invariant beamformers," 2009, 17th European Signal Processing Conference.

Appendix A

Underwater Acoustic Doppler Profile

In mobile communications a common Doppler profile that is used to select f_p is Jake's profile, which assumes a uniformly distributed AoA about the receiver [15]. However, in underwater channels for a single cluster arrival this is not the case as the AoA for different paths within a cluster is in fact very similar. A Doppler profile suggested from underwater measurements carried out for a stationary receiver and transmitter in Norway is a double-sided stretched exponential function of the form [26]

$$S(f) = Ae^{-(\frac{|f|}{a})^B}, \tag{A.1}$$

where A is a scaling factor and a and B control how the exponential function spreads. When $B = 1$ a regular double sided exponential is obtained and with $B = 2$ a Gaussian shape is obtained.

One important thing to note is that these measurements look at surface reflections only. So ideally this Doppler spreading function should only be applied to paths that interact with the surface. However, due to the fact that mobility is assumed in this model, this is applied to the bottom reflection as movement of a cluster path along the bottom surface could produce similar results, as was justified in Section 2.1.

The main reason for this choice is that currently there has been no work done on quantifying how the bottom reflection varies when the receiver and transmitter are non-stationary. In fact no work to date has been performed on quantifying how the surface and bottom reflections change in channels that involve substantial transmitter-

receiver motion, and so largely unsupported assumptions need to be made.

Therefore, a double-sided stretched exponential is used as the Doppler spectrum for the time-variance of all clusters excluding the direct path, which has no interaction with a boundary. To make this function easier to work with, B can be set to 1 to produce a double-sided exponential function, and a can be chosen to produce a desired coherence time, T_c. The coherence time of a cluster arrival is a common measure of how long each fading tap in a channel remains correlated with itself with respect to time, and so is directly related to the Doppler spread of each cluster arrival which describes the frequencies involved in this fading. The longer the coherence time, the longer the tap amplitude remains relatively constant and the lower the frequency components involved in the fading process. When the coherence time is short, the amplitude of the cluster varies quickly and the frequencies involved in the fading process become larger. Therefore, the coherence time and width of the Doppler profile, or Doppler spread, are considered inversely proportional.

The coherence time is related to the correlation of the fading taps, and so is best described using the time correlation function for the amplitude of each cluster arrival, $R_c(t)$ [5]. The correlation function with respect to time is the Fourier pair of the Doppler spectra, or Doppler profile, and so can be expressed as

$$R(t) = \mathscr{F}^{-1}\{S(f)\}, \tag{A.2}$$

where the subscript indicating the functions dependence on the cluster arrival has been dropped as the same profile is assumed for all cluster arrivals. So, the correlation function for all clusters is then

$$R(t) = \frac{\frac{2A}{a}}{4\pi^2 t^2 + (\frac{1}{a})^2}. \tag{A.3}$$

In order to quantify the Doppler profile a coherence time, T_c, and a desired correlation, or coherence level at this time, $R(T_c)$, can be specified. Using these values the Doppler profile parameters a and A can be determined. To simplify this calculation, similar to the power delay profile the Doppler profile can be treated as a pdf as long as the integral of the function is 1. To satisfy this, A can be chosen as

$$A = \frac{1}{2a}, \tag{A.4}$$

which simplifies the correlation function so that it only depends on a,

$$R(t) = \frac{(\frac{1}{a})^2}{4\pi^2 t^2 + (\frac{1}{a})^2}. \tag{A.5}$$

Using the desired coherence time and coherence level, Eqn. (A.5) can be reformulated to solve for a,

$$a = \sqrt{\frac{1 + R(T_c)}{4R(T_c)\pi^2 T_c^2}}. \tag{A.6}$$

A common measure of coherence time is when the coherence power drops below 3 dB, or when $R(T_c) = 0.5$. The value of a given a 3 dB coherence time can then be expressed as

$$a = \sqrt{\frac{1.5}{2\pi^2 T_c^2}}. \tag{A.7}$$

Appendix B

Time-Varying Channel Sounding

Channel sounding is the process of completely determining the channel impulse response (CIR) by transmitting a known block of data and then properly processing the received signal. The transmit baseband signal, $s(t)$, is usually chosen to be a maximum length pseudo-random noise(PN) sequence so that correlative channel sounding can be used [5], [27].

A PN sequence is used rather than a linear frequency-modulated (LFM) chirp train due to the fact the ambiguity function of the PN sequence shows a more distinct peak. While more details about this can be found in [52], this means that a PN sequence shows a distinct correlation peak with respect to both delayed and Doppler shifted versions of itself. LFM's only have a distinct peak with respect to delay, but this peak broadens as the Doppler shift increases. Therefore, PN sequences are preferred over LFM sequences when Doppler shifts or time-scaling may be encountered.

The basic theory behind correlative channel sounding is that if a passband PN-sequence, $\tilde{p}_{Tx}(t)$, is transmitted through a passband channel, $\tilde{h}(t,\tau)$, and then the received signal, $\tilde{p}_{Rx}(t)$, is matched filtered by the original PN sequence, $\tilde{p}_{Tx}(t)$, then the output of this filter should be

$$\begin{aligned}\hat{\tilde{h}}(t,\tau) &= \tilde{p}_{Tx}(t) * \tilde{p}_{Rx}(t) = \tilde{p}_{Tx}(t) * (\tilde{p}_{Tx}(t) * \tilde{h}(t,\tau)) \\ &= (\tilde{p}_{Tx}(t) * \tilde{p}_{Tx}(t)) * \tilde{h}(t,\tau),\end{aligned} \tag{B.1}$$

where $\hat{\tilde{h}}(t,\tau)$ represents an estimate of the passband channel.

If $\tilde{p}_{Tx}(t)$ is a maximum length PN sequence then the auto correlation function, $(\tilde{p}_{Tx}(t)*$

$\tilde{p}_{Tx}(t)$), is closely approximated by a Delta function at $\tau = 0$. So, the output from the matched filter should be

$$\hat{\tilde{h}}(t,\tau) = \delta(0) * \tilde{h}(t,\tau) = \tilde{h}(t,\tau), \tag{B.2}$$

or the channel impulse response itself. A more detailed examination of this process can be found in [27].

In order to create a passband PN sequence, a complex baseband PN sequence can be created and then this sequence can be upconverted to the desired passband region. In order for this sequence to be able to properly sound a channel, there are a number of parameters that can be set.

The first is the baseband chip rate, R_{chip}, which dictates the delay resolution, $\Delta\tau$, of the channel sounder. The delay resolution is inversely proportional to the chip rate, and so increasing the chip rate will yield better delay resolution. This also increases the bandwidth of the sounding signal.

The length of the sounding sequence, T_{seq}, is another parameter that can be set. The longer the sequence, the longer the delay axis which yields better estimates of the length of the CIR. The trade off is that as T_{seq} increases and the length of the delay axis increases, the resolution with respect to time decreases as now a new channel estimate is only produced every T_{seq} seconds. Choosing the length of the signal to be just above the approximate delay-spread of the channel will produce the best results. The cluster arrivals can then be fully determined without aliasing and the time resolution of the channel will be the best it can be. Ideally, the sounding sequence will be shorter than the coherence time of the channel so that a smooth progression of the channel can be observed.

The last parameter that can be set is the carrier frequency of the signal, f_c. This frequency will determine where the sounding occurs in the spectrum and should correspond to the spectrum that will be occupied by a transmitted signal.

Once these parameters are chosen, a complex baseband PN sequence can be produced of length $R_{chip} * T_{seq}$ and then this signal can be modulated up to passband centered around f_c using quadrature phase shift keying (QPSK). The final signal, $\tilde{p}_{Tx}(t)$, can then be transmitted through the channel and received to be processed to produce the final CIR.

If the signal is passed through a non time-varying, or slowly time-varying channel,

then regular correlative channel sounding should produce the desired CIR. However, if there is substantial time-variance in the channel and the PN sequence is a wideband signal, which most will be, then the correlative channel sounding in Eqn. (B.1) cannot be applied directly. This is because correlative channel sounding with the original transmit signal only applies to stationary channels, or slowly varying channels that don't change drastically over T_{seq}.

If the channel does change during T_{seq} then the best way to analyze the channel is to account for these possible changes by creating a bank of correlative filters that are matched to the time-varying PN sequence rather than just the original PN sequence. Due to the reduced rate at baseband, often the baseband channel is easiest channel to work with. The channel model in Eqn. (2.6) can be used to model the baseband channel by accounting for the phase shift on the carrier frequency and the equivalent baseband channel becomes [53]

$$p_{Rx}(t) = \sum_{c=1}^{N_c} \sum_{p=1}^{N_p} a_p p_{Tx}(t - \tau_c(t) - \tau_p(t)) e^{-j2\pi f_c(\tau_c(t)+\tau_p(t))}. \tag{B.3}$$

Knowing $\tau_p(t)$ is small and leads to time variation in the amplitude means that the final baseband received signal can be written as

$$p_{Rx}(t) = \sum_{c=1}^{N_c} a_c p_{Tx}(t - \tau_c(t)) e^{-j2\pi f_c \tau_c(t)}, \tag{B.4}$$

where it is assumed that $T_{seq} << T_c$, and so the amplitude is roughly constant for a single PN sequence.

This baseband analysis shows that if the delay associated with a cluster varies quickly then the received baseband PN sequence will have a time-varying phase shift as well as time-scaling. This means that the ideal matched filter will not be the original transmitted PN sequence, but rather a time-scaled and phase shifted version of the original signal. So, unlike a stationary channel sounder where only one matched filter is required, for a wideband channel sounder a bank of matched filters is required with proper time-scaling and phase-shifting. Specifically, these matched filters can be written in terms of the transmitted PN sequence, $p_{Tx}(t)$, as

$$p_{Tx,\alpha}(t) = p_{Tx}(\alpha t) e^{-j2\pi f_c(1-\alpha)}. \tag{B.5}$$

In order to create the required bank, a minimum and maximum velocity of interest can be chosen leading to a minimum and maximum value *alpha*. The step size from α_{min} to α_{max} is then chosen so that

$$\alpha_{step} \leq \frac{1}{2T_{seq}R_{chip}}. \tag{B.6}$$

This states that there should be enough matched filters so that no time-scaled version of the transmitted PN sequence is ever off in time scaling by more than half of a chip duration. This ensures a minimum level of correlation is always attained for all time scaled signals.

With this bank of filters, the final CIR can be produced by filtering the received signal through all of the filters accounting for time-scaling and then combining their outputs to produce the final CIR. Unlike the simple single correlative channel sounder, this response can now contain time-variance, so the delays can drift with time and this channel sounder will be able to track the movement. It should be noted that the direct combination of the filtered outputs may not be the optimal method of combining the filtered outputs, however this is the method that is used here when channel sounding is completed.

While a direct combination of the filtered outputs may not be the ideal fashion of combination after filtering, no analyses was done to determine the benefits of other combination techniques and so direct combination will be used here.

Appendix C

Delay-and-Sum Beamforming

For broadband signals made up of infinitely many frequency components, a single phase shift cannot be used to model the difference in the signals between array elements as the phase shift is frequency dependent. This frequency dependence can be seen in the phase term for the received signal on the m^{th} circumferential array element for a signal arriving from a direction, $\bar{\phi}$.

$$\Phi = kr\cos(\bar{\phi} - \phi_m) = \frac{\omega}{c} r\cos(\bar{\phi} - \phi_m), \quad m = 1, \ldots, M \tag{C.1}$$

For a broadband signal, a more appropriate measure of how the signal changes between array elements is in terms of the frequency independent time delay with reference to the center of the UCCA. If a signal, $x(t)$, is transmitted, then the signal that arrives on the circumferential sensors can be described as

$$x_m(t) = x(t + \tau_m(\bar{\phi})), \quad m = 1, \ldots, M, \tag{C.2}$$

where

$$\tau_m(\bar{\phi}) = \frac{r}{c}\cos(\bar{\phi} - \phi_m). \tag{C.3}$$

For a single frequency this corresponds to a single phase delay, however for a wideband signal this time delay describes the phase shift experienced by all frequencies assuming a non-dispersive medium. If the received signal is sampled every T_s seconds, and T_s is chosen to satisfy the Nyquist criterion for the incoming signal, then the signal can

be described using the Whittaker-Shannon interpolation formula:

$$x(t) = \sum_{k=-\infty}^{\infty} x[k] \operatorname{sinc}\left(\frac{t - kT_s}{T_s}\right), \tag{C.4}$$

where the sinc function acts as an ideal low pass filter [54]. If the sampled signal is used in Equation [C.2], then the continuous signal on the m^{th} element becomes

$$x_m(t) = x(t + \tau_m) = \sum_{k=-\infty}^{\infty} x[k] \operatorname{sinc}\left(\frac{t + \tau_m(\bar{\phi}) - kT_s}{T_s}\right) \quad m = 1, \ldots, M. \tag{C.5}$$

If the continuous signal on the m^{th} element is now sampled as well at the same rate, T_s, then the result is [34]

$$x_m[n] = x_m(nT_s) = \sum_{k=-\infty}^{\infty} x[k] h_{m,\bar{\phi}}[n - k] \quad m = 1, \ldots, M, \tag{C.6}$$

where

$$h_{m,\bar{\phi}}[n] = \operatorname{sinc}\left(n + \frac{\tau_m(\bar{\phi})}{T_s}\right). \tag{C.7}$$

In order to produce a beamformed output at some desired angle, $\bar{\phi}$, the outputs of all of the array elements can be made to sum coherently at this angle. A simple way to do this is to reverse the time delay caused by the spatial sampling operation for a signal coming from the desired angle, $\bar{\phi}$. The ideal filter delay-and-sum for the m^{th} array element is then

$$f_{m,\bar{\phi}}[n] = \operatorname{sinc}\left(n - \frac{\tau_m(\bar{\phi})}{T_s}\right) \quad n = 0, \pm 1, \pm 2, \ldots, \pm\frac{(J-1)}{2} \tag{C.8}$$

This ideal time-delay filter can be implemented using an FIR filter if J is chosen to be a finite number. In Equation [C.8], it was assumed J was odd. The response of the delay-and-sum beamformer steered to an angle $\bar{\phi}$ then becomes

$$y[n] = \sum_{m=1}^{M} x_m[n] * f_{m,\bar{\phi}}[n] = \sum_{m=1}^{M} \sum_{k=-J}^{J} x_m[k] f_{m,\bar{\phi}}[n - k], \tag{C.9}$$

where the equation could be made causal by delaying the output, $y[n]$, by $\frac{(J-1)}{2}$

samples.

In order to determine the response of the beamformer to both frequency and angle-of-arrival, it can be assumed that there are an infinite number of elements along the circle and the ideal delay filter can be replaced by a frequency dependent phase shift operator.

$$D(\phi,\omega) = \frac{1}{2\pi}\int_0^{2\pi} F(\alpha)e^{j\omega\frac{r}{c}\cos(\alpha-\phi)}d\alpha \tag{C.10}$$

where $F(\phi)$ is a weighting function that implements the desired delay when steered in a direction $\bar{\phi}$,

$$F(\alpha) = e^{-j\omega\frac{r}{c}\cos(\alpha-\bar{\phi})}. \tag{C.11}$$

The resulting directional pattern can then be represented as

$$D(\phi,\omega) = J_0\left(2\omega\frac{r}{c}\sin\left(\frac{\phi-\bar{\phi}}{2}\right)\right), \tag{C.12}$$

where J_0 is a Bessel function of the first kind [36], and the center element has been given a constant weight of 0.

Plots of the theoretical and actual beam patterns for a 7-element UCCA array with $M = 6$, steered toward $\bar{\phi} = 0$ are shown in Figures C.1 and C.2, respectively. The theoretical beam pattern was created using Equation [C.12] and the frequency was limited to $\lambda_{min} = 2d_e$. The actual beam pattern was created using Equation [3.24] with filters of length $J = 121$ designed using [C.8].

The actual delay-and-sum beamformer does produce a beam pattern that is very similar to the theoretical one aside from a slight ripple on the main beam as $\frac{\Omega}{\pi} \to 1$. However, there are two main drawbacks. First, to change the steering angle of the beamformer involves implementing $M-1$ new FIR filters, one for each circumferential array element. So, to implement a second beam at a different steering angle, $\bar{\phi}_2$, would involve a second FIR filter applied to each element, and $M-1$ new FIR filters in total. The second problem is that the main beam has a beamwidth that is a function of frequency. Specifically, when the argument of the first order Bessel function, $J_0(z)$, is roughly 1.126 the 3 dB point of the main beam is encountered. So, the 3 dB

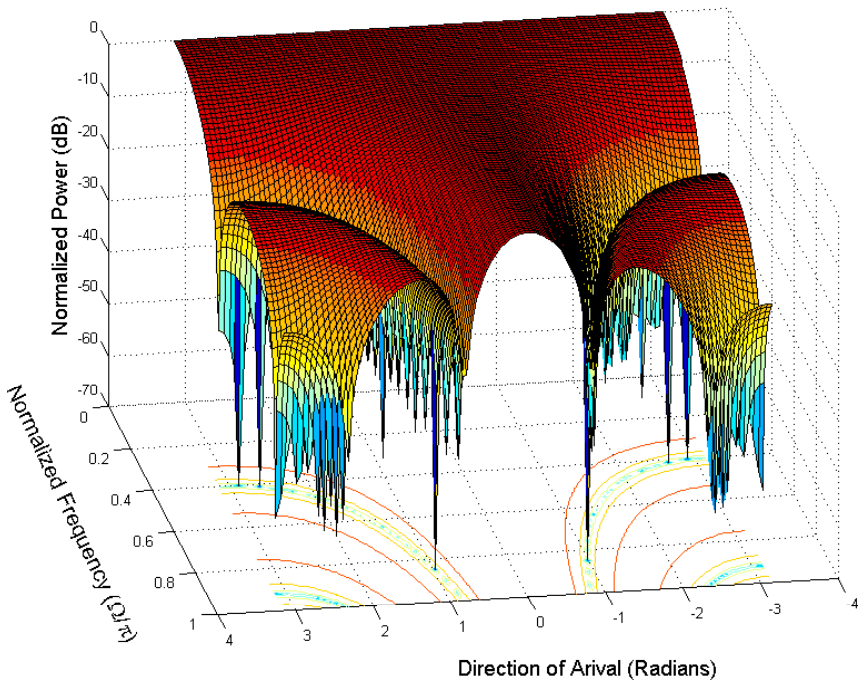


Figure C.1: Ideal response of the Delay-and-Sum beamformer with $M = 7$ and a steering direction of $\phi = 0$.

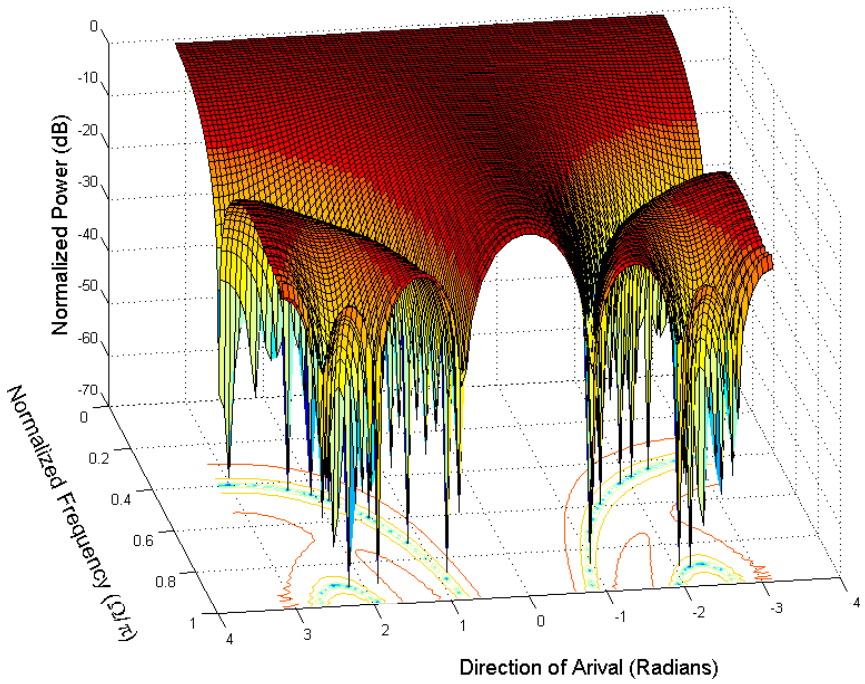


Figure C.2: Actual response of the Delay-and-Sum beamformer with $M = 7$, a steering direction of $\phi = 0$, and ideal interpolation filters of length $J = 121$ attached to each circumferential array element.

beamwidth of the delay-and-sum beamformer can be expressed as

$$B_{\text{Delay-and-Sum}} = 4\sin^{-1}\left(\frac{1.126c}{2\omega r}\right), \tag{C.13}$$

which is clearly a function that is inversely proportional to frequency. So as the frequency decreases, the 3 dB beamwidth increases until it encompasses the full angular range of the array, 2π. This can be seen in both the theoretical and actual response of the beamformer when the output begins to flatten across all angles of arrival. This dependence of the main beam on the frequency makes the delay-and-sum beamformer a poor choice for frequency invariant beamforming.

Appendix D

Least-Squares Beamforming

Least-squares beamforming is another method that can be used to design the FIR filters attached to each array element. In this method, the weight vector, $\mathbf{w}$, is designed so that it optimally fits a desired beam pattern, $D(\Omega, \phi)$. To do this a cost function needs to be designed and in this section the focus is on a weighted least-squares (LS) cost function of the form [55]

$$J_{LS} = \int_{\Omega_B} \int_{\phi} F(\Omega, \phi) |P(\Omega, \phi) - D(\Omega, \phi)|^2 d\phi d\Omega. \tag{D.1}$$

$P(\Omega, \phi)$ is the beamformer response, $|P(\Omega, \phi) - D(\Omega, \phi)|^2$ is the LS cost function, $F(\Omega, \phi)$ is a weighting function, Ω_B is the normalized frequency range of interest, $[\Omega_{min}\ \Omega_{max}]$, and ϕ is the angle range of interest, which for a circular array is the full azimuthal plane, $[-\pi\ \pi]$.

In order to practically implement this approach, both the angle and normalized frequency range need to be discretized. So the weighted LS cost function in Eqn. (D.1) can be written as the summation

$$J_{LS}(\mathbf{w}) = \sum_{\Omega_i} \sum_{\phi_k} F(\Omega_i, \phi_k) |\mathbf{w}^H \mathbf{d}(\Omega_i, \phi_k) - D(\Omega_i, \phi_k)|^2, \tag{D.2}$$

where $P(\Omega_i, \phi_k)$ has been replaced by the response of the beamformer, $\mathbf{w}^H \mathbf{d}(\Omega_i, \phi_k)$.

If desired response, $P(\Omega_i, \phi_k)$, is chosen to be 1 at a desired angle, or angles, of arrival, ϕ_{PB}, for all frequencies in Ω_B, and the response at all other angles of arrival, ϕ_{SB}, is

chosen to be zero then this equation can be simplified to

$$J_{LS}(\mathbf{w}) = \sum_{\Omega_i \in \Omega_B} \sum_{\phi_k \in \phi_{PB}} |\mathbf{w}^H \mathbf{d}(\Omega_i, \phi_k) - 1|^2 + \alpha \sum_{\Omega_i \in \Omega_B} \sum_{\phi_k \in \phi_{SB}} |\mathbf{w}^H \mathbf{d}(\Omega_i, \phi_k)|^2, \quad \text{(D.3)}$$

where the weighting function, $F(\Omega_i, \phi_k)$, is 1 in ϕ_{PB} and α in ϕ_{SB} for all frequencies. For real valued weights this can be written as a quadratic equation of the form [32]

$$J_{LS}(\mathbf{w}) = \mathbf{w}^T \mathbf{G}_{LS} \mathbf{w} - 2\mathbf{w}^T \mathbf{g}_{LS} + g_{LS}, \quad \text{(D.4)}$$

where

$$\mathbf{G}_{LS} = \sum_{\Omega_i \in \Omega_B} \sum_{\phi_k \in \phi_{PB}} \mathbf{D}_R(\Omega_i, \phi_k) + \alpha \sum_{\Omega_i \in \Omega_B} \sum_{\phi_k \in \phi_{SB}} \mathbf{D}_R(\Omega_i, \phi_k), \quad \text{(D.5)}$$

$$\mathbf{g}_{LS} = \sum_{\Omega_i \in \Omega_B} \sum_{\phi_k \in \phi_{PB}} \Big(\mathbf{d}_R(\Omega_i, \phi_k) \cos(D_G \Omega_i) + \mathbf{d}_I(\Omega_i, \phi_k) \sin(D_G \Omega_i) \Big), \quad \text{(D.6)}$$

and

$$g_{ls} = \sum_{\Omega_i \in \Omega_B} \sum_{\phi_k \in \phi_{PB}} 1. \quad \text{(D.7)}$$

$\mathbf{D}_R(\Omega_i, \phi_k)$ is the real part of $\mathbf{D}(\Omega_i, \phi_k) = \mathbf{d}(\Omega_i, \phi_k)\mathbf{d}(\Omega_i, \phi_k)^H$, and D_G is the group delay of the FIR filters used for each array element [55].

The quadratic in Eq. (D.4) can be minimized by taking the derivative with respect to $\mathbf{w}$ and setting this equal to zero, which yields the optimal least-squares solution

$$\mathbf{w}_{LS} = G_{LS}^{-1} \mathbf{g}_{LS}. \quad \text{(D.8)}$$

As an example, a design was implemented for a 7-element UCCA where the desired response was a delta function at 0 rads, $\phi_{PB} = 0$, and 0 at all other angles. The desired frequency range was from 0.05π to 0.5π, the weighting function variable was $\alpha = 0.6$ and the FIR filters had lengths of $J = 100$. The resulting beam pattern using the weights calculated in Equation [D.8] is shown in Figure D.1

The main beam of this beamformer remains constant with respect to gain and to 3 dB beamwidth, where a width of $\frac{\pi}{4}$ is maintained over the whole frequency range of

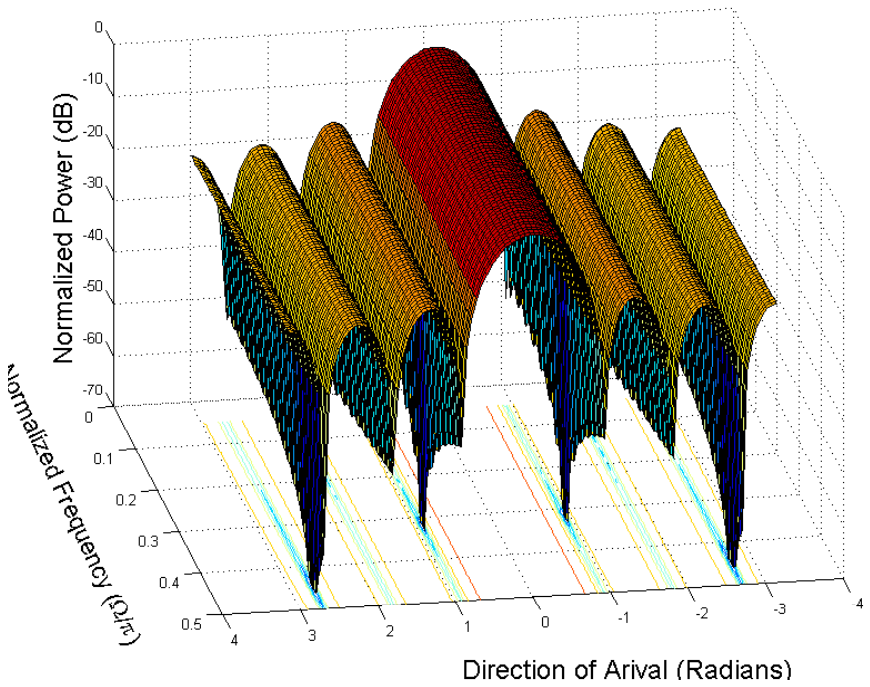


Figure D.1: Beam pattern of the 7-element UCCA beamformer with FIR filters produced using a least-squares solution with $\phi_{PB} = 0$.

interest, as shown in Figure D.2. The side beams also remain fairly constant and have a maximum sidelobe level of 13.5 dB.

The main drawback of this technique is that the beam pattern changes size as it is rotated. Specifically, for a 7-element UCCA when the main beam is located at integer multiples of $\phi_{PB} = \frac{\pi}{6}$, the beam pattern suffers its most severe deviations from the beam pattern in Figure D.1. As an example, the resulting beam pattern when the beamformer is steered to $\phi_{PB} = \frac{\pi}{6}$ is shown in Figure D.3, which shows that the 3 dB beamwidth of the main beam has increased to $\frac{\pi}{3}$.

So while the weighted LS approach provides some attractive wideband beam patterns, the fact that the beam patterns produced change with steering direction and that new FIR filters are required for each array element and for each steering direction still make this a relatively undesirable beamforming technique when multiple beams in different directions are required.

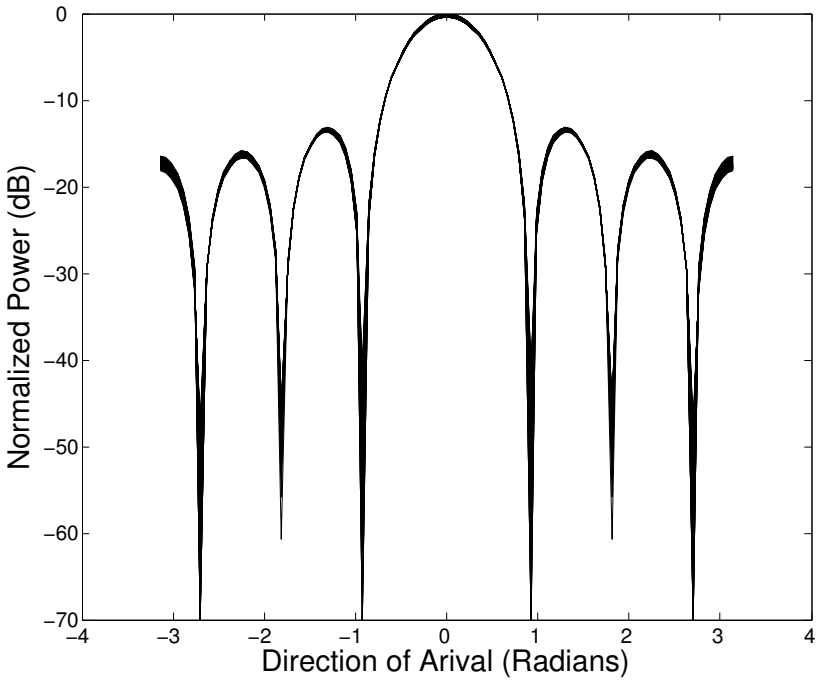


Figure D.2: Slice view of the beam pattern of the 7-element UCCA beamformer with FIR filters produced using a least-squares solution. Each slice represents a frequency sampling point.

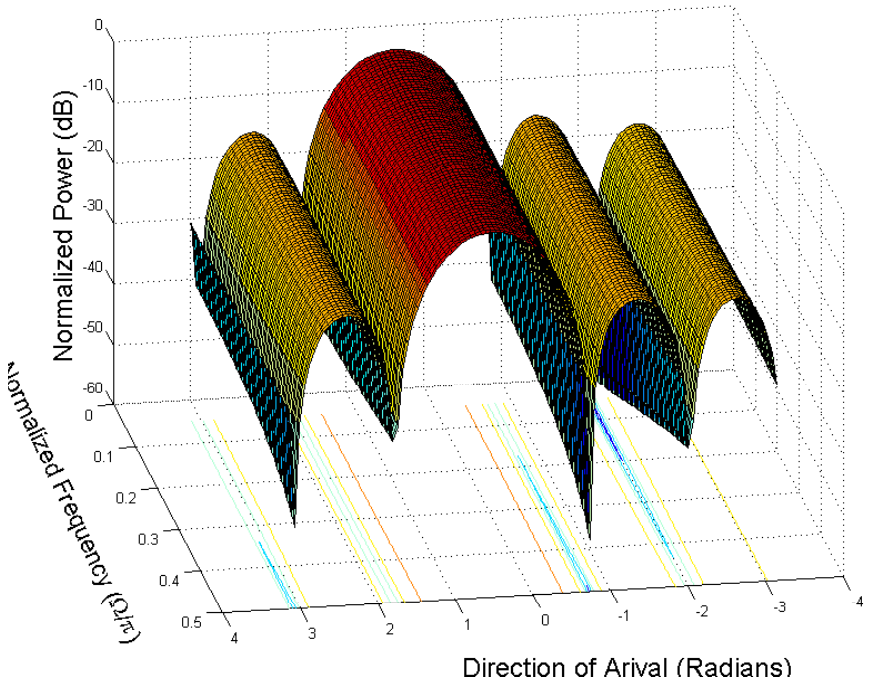


Figure D.3: Beam pattern of the 7-element UCCA beamformer with FIR filters produced using a least-squares solution with $\phi_{PB} = \frac{\pi}{6}$.

Appendix E

Cyclic Prefixing and Zero-Padding to Remove ICI cause by Delay-Spread Channels

There are two methods that are used to eliminate the ICI in OFDM systems due to delay-spread channels. The first is called a cyclic prefix (CP) and is the most commonly used method to reduce ICI due to delay spread. The idea behind the CP is that if the linear convolution with the channel can be made cyclical, then the channel matrix will be circulant and the sub-carriers will then be orthogonal.

To show this, the discrete baseband time-domain OFDM signal can be created using the OFDM block of interest, $\mathbf{d}[i]$, and Eqn. 4.8. The resulting signal is then

$$\mathbf{s}_i = \big[s_i[0], s_i[1], \ldots, s_i[N_B - 1]\big]. \tag{E.1}$$

Cyclic prefixing takes the last N_{cp} entries of $\mathbf{s}_i$, and prepends them to $\mathbf{s}_i$ to form the new time domain baseband transmit signal with the CP [44],

$$\mathbf{s}_{cp,i} = \big[s_i[N_B - N_{cp}], \ldots, s_i[N_B - 1], s_i[0], s_i[1], \ldots, s_i[N_B - 1]\big]. \tag{E.2}$$

The advantage of using a cyclic prefix is that the new transmit signal $s_{cp,i}[n]$ can be considered periodic with period N_B as long as $-N_{cp} \leq n \leq N_B - 1$. If this signal is convolved with a discrete baseband non-time-varying channel, $h[m]$, then the linear

convolution becomes

$$r_{cp,i}[n] = \sum_{m=0}^{M} h[m]s_{c,i}[n-m] \quad \text{for} \quad n = 0, 1, \ldots, N_B - 1. \tag{E.3}$$

This linear convolution can be represented using a circular convolution as long as $n - m \geq -N_{cp}$ so $s_{cp,i}[n]$ remains periodic. The lower limit occurs when $n = 0$ and $m = M$, where M is the discrete maximum delay, $M = \frac{\tau_{max}}{T_s}$. So as long as $N_{cp} \geq M$, then this linear convolution of the cyclic prefixed signal with the channel can be represented using a circular convolution [44]

$$r_{cp,i}[n] = \sum_{m=0}^{M} h[m]s_{cp,i}[n-m]_{N_B} = s_{cp,i}[n] \circledast h[n] \quad \text{for} \quad n = 0, 1, \ldots, N_B - 1, \tag{E.4}$$

where $s_{cp,i}[n]_{N_B}$ indicates $s_{cp,i}[n]$ repeats every N_B samples and so can be represented as $s_{cp,i}[n \bmod N_B]$.

This can also be restated using continuous variables. The linear convolution can be represented using a circular convolution as long as the length of the guard interval, T_g, filled with the the cyclic prefix is longer than the delay spread of the channel, τ_{max}. This circular convolution produces a circulant channel matrix and the desired diagonal channel matrix, $\mathbf{D}_H$, in the frequency domain.

The operation of cyclic prefixing is shown in Figure E.1. Due to periodic nature of the signal being transmitted as long as $T_g > \tau_{max}$ then a full OFDM block duration can be produced for all path arrivals, therefore maintaining the orthogonality of the carriers and avoiding the ICI problem presented in Figure 4.5.

Practically, demodulating the cyclic prefixed signal involves recording $N_B + N_C$ baseband samples past the synchronization point for the start of the received OFDM block. This recorded signal, $r_{cp,i}[n]$, is then truncated by removing the first N_c entries and only keeping the last N_B entries to form $r_i[n]$ [44]. The signal can then be passed to an FFT of length N_B for demodulation.

While cyclic prefixing is popular for channels with small delay spreads, as the delay spread increases cyclic prefixing becomes less attractive as it requires constantly sending symbols which consumes power [42]. A second alternative is zero-padding the signal. Zero padding is a simple operation where N_{zp} zeros are added to the end of

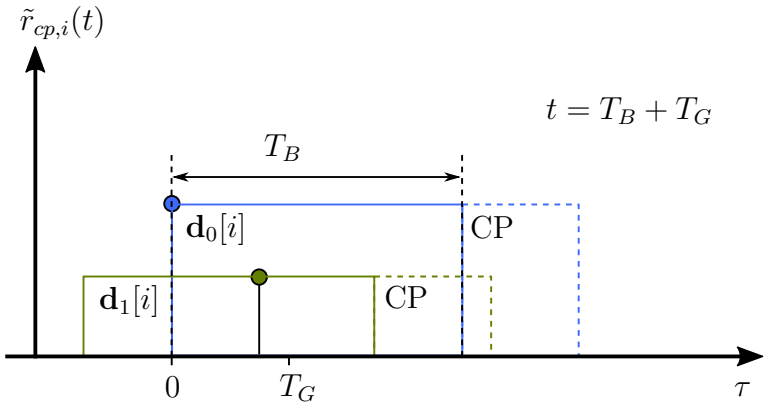


Figure E.1: Reception of a cyclic prefixed transmit signal, $s_{c,i}(t)$. The period T_B shown in the plot indicates the period over which the signal is recorded to produce the received signal. This period contains a full block length for all paths as $T_G > \tau_{max}$.

the time domain baseband signal to form the zero padded baseband signal

$$\mathbf{s}_{zp,i} = \big[s_i[0], s_i[1], \ldots, s_i[N_B - 1], 0, \ldots, 0\big]. \tag{E.5}$$

In order to demodulate a zero padded signal $N_B + N_{zp}$ baseband samples are recorded after the start of the first received OFDM block to create $r_{zp,i}[n]$. The last N_{zp} samples of $r_{zp,i}[n]$ are then added to the first N_{zp} samples to form the final received block $r_i[n]$ with only N_B samples. This operation is called an overlap-and-add (OLA) operation and can be represented as [45]

$$r_i[n] = \begin{cases} r_{zp,i}[n] + r_{zp,i}[n + N_B], & \text{if } 0 \leq n \leq N_{zp} - 1 \\ r_{zp,i}[n], & \text{if } N_{zp} \leq n \leq N_B - 1, \end{cases} \tag{E.6}$$

which can also be expressed in terms of the linear convolution of the original transmit signal and the channel as

$$r_i[n] = \begin{cases} \sum_{m=0}^{M} h[m]\big[s_{zp,i}[n-m] + s_{zp,i}[n + N_B - m]\big], & \text{if } 0 \leq n \leq N_{zp} - 1 \\ \sum_{m=0}^{M} h[m] s_{zp,i}[n-m], & \text{if } N_{zp} \leq n \leq N_B - 1. \end{cases} \tag{E.7}$$

This can be shown to be equivalent to the circular convolution in Eqn. (E.4). Specifically, the sum is broken up into two separate sums to account for the sections where the received signal is required to be periodic with period N_B and where it does not need to be. The first sum in Eqn. (E.7) requires the the signal to be peri-

odic, and $s_{zp,i}[n-m] + s_{zp,i}[n+N_B-m]$ can be represented as $s_{zp,i}[n-m]_{N_B} = s_{zp,i}[(n-m) \bmod N_B]$ due to the separation of the signals by N_B and the fact that $s_{zp,i}[n+N_B-m] = 0$ when $n-m > 0$.

The second sum however should require no periodicity and should always fall into a single period. So, $0 \leq n-m \leq N_B-1$ for all values of m and n. To ensure this, N_{zp} can be chosen such that $N_{zp} \geq M$. With this choice of N_{zp} the convolution can be considered circulant as in Eqn. (E.4). Again, this equates to choosing the guard interval such that $T_G > \tau_{max}$, except now this guard interval is at the end of the OFDM signal and is filled with zeros rather than a cyclic prefix. The zero padding demodulation operation is shown in Figure E.2.

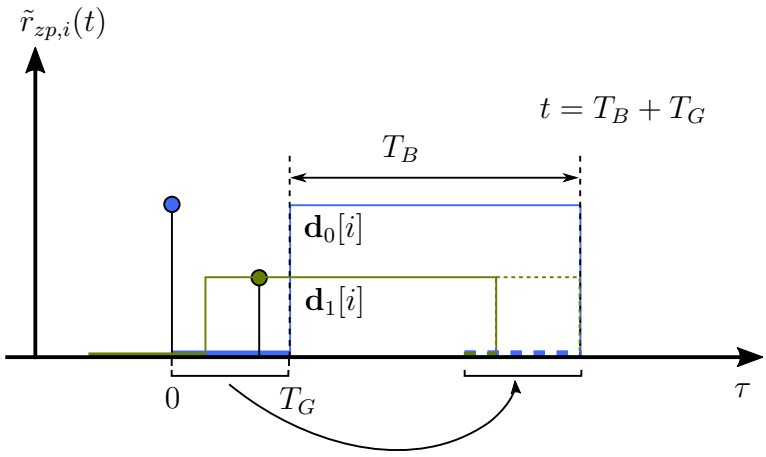


Figure E.2: Zero padded OFDM demodulation using the overlap and add method. This shows that by taking the last T_G seconds of the reception and adding them to the first T_G seconds ensures a full block reception for all paths if $T_G > \tau_{max}$.

Another way of looking at zero padded OFDM demodulation is that is integrates the received signal over a period of $t \in [0, T_B+T_G]$. So, the full OFDM block from all taps will be received leading to orthogonal sub-carriers for all taps. The main drawback of the zero padded OFDM is that due to the fact that receiver effectively integrates over a longer duration than a single OFDM block, the output of the OLA operation will have a higher noise floor [45].