MIMO COMMUNICATIONS OVER MULTI-MODE OPTICAL FIBERS

BY

PETER KAIROUZ

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Master's Committee:

Professor Andrew Singer
Professor Naresh Shanbhag

ABSTRACT

We consider multi-input multi-output (MIMO) communications over multi-mode fibers (MMFs). Current MMF standards, such as OM3 and OM4, use fibers with core radii of 50 μm, allowing hundreds of modes to propagate. Unfortunately, due to physical and computational complexity limitations, we cannot couple and detect hundreds of data streams. In order to circumvent this issue, two solutions were presented in the literature. The first is to design new fibers with smaller radii so that they can support a desired number of modes. The second is to design multi-core fibers with a reasonable number of cores. However, both approaches are expensive as they necessitate the replacement of installed fibers. In our work, we consider input-output coupling schemes that allow the user to couple and extract a reasonable number of signals from a fiber with many modes. This approach is particularly attractive as it is scalable; i.e., the fibers do not have to be replaced every time the number of transmitters or receivers is increased (which is likely to happen in the near future). In addition, fibers with large radii can support higher peak powers, relative to fibers with small radii, while still operating in the linear regime. However, the only concern is that fibers with more modes suffer from increased mode-dependent losses (MDLs). Our work addresses this last concern.

This thesis is divided into two parts. In the first part, we present a channel model that incorporates intermodal dispersion, chromatic dispersion, mode dependent losses, and mode coupling. We later extend this model to include the input and output couplers and provide an input-output coupling strategy that leads to an increase in the overall capacity. This strategy can be used whenever channel state information (CSI) is available at the transmitter and the designer has full control over the couplers. We show that the capacity of an $N_t \times N_t$ MIMO system over a fiber with $M \gg N_t$ modes can approach the capacity of an N-mode fiber with no loss. Moreover, we present a statistical

input-output coupling model in order to quantify the loss in capacity when CSI is not available at the transmitter or there is no control over the input-output coupler. It turns out that the loss, relative to N_t-mode fibers, is minimal (less than 0.5 dB) for a wide range of signal-to-noise ratios (SNRs) and a reasonable range of MDLs. This means that there is no real need to replace the already installed fibers and that our strategy is indeed a better approach to solving the above problem.

In the second part, we explore reduced complexity maximum likelihood sequence detection (MLSD) algorithms for single carrier MIMO systems. These algorithms can be used for optical as well as wireless communications. We show that a sphere decoding (SD)-like approach can be used to reduce the computational complexity of the vector Viterbi algorithm (VVA), an extension to the Viterbi algorithm for MIMO systems. Our combined SD-VVA approach is attractive because it provides substantial computational savings while solving an exact MIMO MLSD problem. Our results show a 50% reduction in complex multiplications and real additions, relative to the full VVA, for a 2×2 MIMO system using 16-QAM signal constellation and operating at an signal-to-noise ratio (SNR) of 10 dB. This figure is increased to 60% when the SNR is increased to 15 dB. We show that larger savings can be achieved for larger MIMO systems and higher order signal constellations. Finally, we show how our algorithm can be modified in order to further reduce the complexity of VVA while still achieving close to optimal performance.

"All that I am or ever hope to be, I owe to my angel mother."−Abraham Lincoln

To my mother, for her love and support.

ACKNOWLEDGMENTS

First, I would like to express my sincerest appreciation to my mother for her infinite love and support. Without her, I would not have been the successful person I am today.

I would also like to express my gratitude to my advisors, Professor Andrew Singer and Professor Naresh Shanbhag, for their patience and support. Their guidance helped me in the research that I conducted during my study at the University of Illinois at Urbana-Champaign (UIUC). Moreover, their openness to ideas and constant encouragement helped me shape my visions and interests. I recognize that this research would not have been possible without their financial assistance. Thank you.

A very special thank-you goes to my colleagues at UIUC. In specific, I would like to thank Aolin Xu, Eric Kim, Andrew Bean, and Thomas Riedl for the stimulating and exciting discussions that were balanced between research interests and personal pursuits. Together, we learned a lot about life, research, and how to tackle new problems and develop techniques to solve them.

Finally, I am grateful to everyone who helped with, and contributed to, the development of this work: my graduate and undergraduate professors, friends, and family members. Your continuous encouragement and support was an everlasting source of inspiration for this work.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 MIMO Communications over MMF

Since Shannon defined the notion of channel capacity as the fundamental limit on achievable transmission rates with vanishing probability of error, system designers have attempted to reach this limit by leveraging device technology advances and increasingly sophisticated algorithms and architectures. Moore's law, together with advances in signal processing, information theory, and coding theory have enabled us to essentially achieve this fundamental limit for a number of narrow-band wired and wireless communication links.

Because of their superior bandwidth-distance product, optical fibers have become extremely popular and have largely replaced traditional copper wire technologies. Optical communication links have serial data rates that are typically several orders of magnitude higher than their wired or wireless electrical counterparts, such as voice-band or cable modem technology or even high-speed chip-to-chip serial links.

Despite their superiority, optical links have limited capacity and the circuits, signal processing, and information theory communities need to completely re-think the design and analysis of communication systems in order to address the ever increasing demand for Internet bandwidth. Furthermore, Moore's law is reaching its limits, and device scaling can no longer readily provide the increases in electrical switching speed and transistor density to which we have become accustomed over the last several decades. Optical data rates are sufficiently high that we can no longer ignore constraints on receiver computational complexity and devices when considering achievable rates.

Multi-input multi-output (MIMO) communications over multi-mode fibers (MMFs) promises to boost the Internet bandwidth supply. However, the capacity of MIMO optical links has not been investigated properly due to the lack of accurate and mathematically tractable channel models. In the first part of this thesis, we present a detailed linear model for the MIMO multi-mode optical channel and analyze its capacity as a function of input-output coupling as well as other physical parameters. We also introduce an input-output coupling strategy and compare it to the uncontrolled coupling case in terms of achievable rate. In the second part, we develop reduced complexity and sub-optimal approximations to maximum likelihood sequence detection (MLSD) algorithms for frequency selective MIMO links. The algorithms we provide can be used for both optical and wireless communication systems. We show that computational complexity savings can be achieved if a sphere decoding-like algorithm is used to update the path metrics of all states at each stage in a vector Viterbi algorithm (VVA).

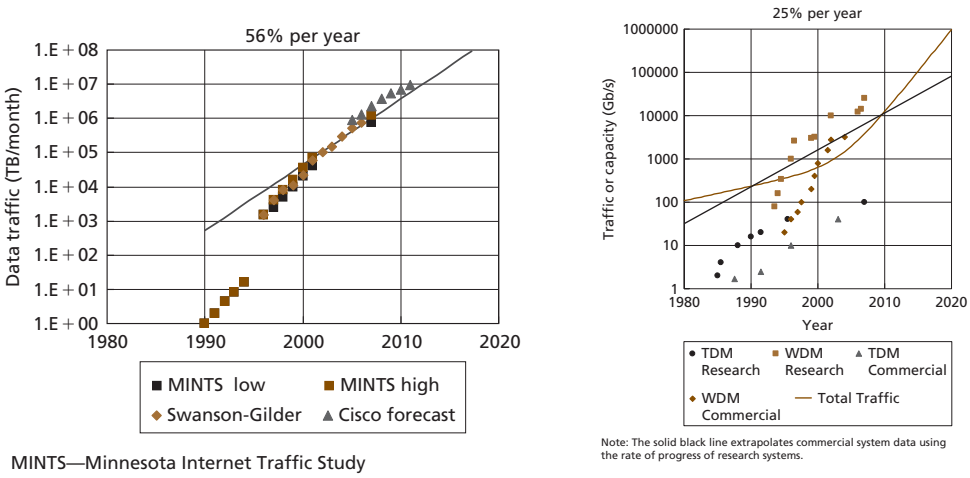
1.2 Historical Overview

Optical communications, broadly defined, is considered to be one of the oldest forms of long distance telecommunications. In fact, many centuries ago, people used beacons, fires, and smoke signals to communicate with one another. More specifically however, the development of optical fiber waveguide technologies is recent and dates back to the late 1960s. The early history of the first generation optical fiber communication (OFC) systems is summarized in Table 1.1. The invention of gallium arsenide (GaAs) semiconductor lasers in the early 1960s was the enabler of high speed OFC technologies. The second generation of optical fibers emerged in the early 1980s. It operated at 1.3 μm, as opposed to 0.8 μm in the first generation, and used indium gallium arsenide phosphide (InGaAsP) semiconductor lasers. In this generation, multi-mode dispersion (cf. Section 2.4) was the major limitation in OFC systems. In 1981, single-mode fibers (SMF, cf. Section 2.2) were developed to solve the multi-mode dispersion problem. This led to a boost in performance; however, practical connectors capable of working with SMF proved difficult to build. In 1987, optical systems were operating at rates close to 1.7 Gb/s with a repeater placed every 50 km. In 1988, TAT-8 was

Table 1.1: The early history of optical fiber communications

1792	Claude Chappe invented the first optical semaphore telegraph (optical telegraph)
1880	Alexander Graham Bell and Charles Sumner Tainter developed the Photophone, a device for transmitting sound on a beam of light
1841	Jean-Daniel Colladon demonstrated light guiding or total internal reflection for the first time
1920	John Logie Baird and Clarence W. Hansell patented the idea of using arrays of hollow pipes or transparent rods to transmit images for television
1966	Charles K. Kao and George Hockham proposed optical fibers at Standard Telecommunications Laboratories
1970	Corning Glass Works (now Corning Inc.) successfully developed optical fibers with an attenuation of 17-20 dB/km. A few years later they developed a fiber with only 4 dB/km.
1977	General Telephone and Electronics sent the first live telephone traffic through fiber optics at a speed of 6 Mbps

the first transatlantic telephone cable to use optical fiber based on Desurvire's optimized laser amplification technology. The third generation OFC systems operated at 1.55 μm and had losses of about 0.2 dB/km. Scientists overcame chromatic dispersion (cf. Section 2.4) by using dispersion-shifted fibers that are designed to have minimal spreading at 1.55 μm. The data rates of these systems was in the range of 2.5 Gb/s with repeaters placed in excess of 100 km. The fourth generation OFC systems used optical amplification and wavelength division multiplexing (WDM) to increase data capacity. In WDM systems, a number of lasers, with different wavelengths, are used to modulate independent data streams onto the same fiber. This technology led to doubling the capacity of optical systems every 6 months starting in 1992 until a bit rate of 10 Tb/s was reached in 2001 and then 14 Tb/s in 2006 (over a single 160 km line using optical amplifiers). However, as the technology matured the increase in capacity saturated in the late 2000s because almost all available *degrees of freedom* (cf. Section 2.3) had been already exploited by then. In this thesis, we show how we can make use of the single unexplored degree of freedom, the spatial modes of a fiber, to build MIMO systems that can help boost the data rates of optical links.



(a) Internet Demand (b) Internet Supply

Figure 1.1: Internet demand and supply (extracted from [1])

1.3 Motivation

In an information-intensive era, the demand for Internet bandwidth is increasing at a rate of 56% per year, while the increase in supply is falling behind at a rate of 25% per year [1]. The increase in demand is fueled by the boom in web-based data services such as cloud computing and real-time multimedia applications. Figure 1.1 predicts that the gap between the Internet demand and supply will only increase in the coming few years if no immediate solution is provided. As a result, optical fiber communication researchers are looking into new ways of boosting the transmission rate of optical links. Given that polarization division multiplexing (PDM) and wavelength division multiplexing (WDM) have already been exploited [3], the only remaining degree of freedom is space division multiplexing [4]. MIMO optical communication increases the transmission rates of MMF systems by multiplexing a number of independent data streams on different spatial modes. Note that, unlike WDM systems, all the laser sources in this case have the same wavelength. MMF is a dominant type of fiber used for high speed data communication in short-range links such as local area networks (LAN) and data centers [5]. It is usually favored over single-mode fibers because of its relaxed connector alignment tolerances and its reduced transceiver connector costs. Plastic optical fibers are great examples of MMFs with remarkably

low installation and operation costs [6]. However, they suffer from mode-dependent losses, mode coupling, intermodal dispersion, and chromatic dispersion (group velocity dispersion) [7]. All these phenomena will be explained in detail in Section 2.4. These limitations make the design and analysis of MIMO multi-mode systems challenging yet exciting.

1.4 Outline and Results

In Chapter 2, we present a MIMO channel propagation model that takes intermodal dispersion, chromatic dispersion, mode-dependent losses, and mode coupling into account. In Chapter 3, we compute the capacity of an M-mode fiber and demonstrate how mode-dependent losses and mode coupling affect it. In Chapter 4, we analyze the coupling of a reasonable number of laser sources to a fiber with hundreds of modes. We also propose an input-output coupling model and present a coupling strategy: using the input-output couplers to perform a special kind of beamforming. This strategy leads to an increase in the overall capacity as it allows the effective transmission of data along the least lossy subset of end-to-end eigenmodes. The resultant capacity is almost equal to that of a fiber with N_t modes and no modal losses, an ideal case which maximizes the capacity of an $N_t \times N_t$ MIMO system. This coupling strategy can only be used when channel state information (CSI) is available at the transmitter and there is full control over the input-output couplers. In the absence of these conditions, an appropriate random input-output coupling model is used in order to better model the behavior of the system and quantify the expected loss in the fiber's capacity. It turns out that the loss, relative to N_t-mode fibers, is minimal (less than 0.5 dB) for a wide range of SNRs and a reasonable range of MDLs. In Chapter 5, we present a reduced complexity maximum likelihood sequence detection (MLSD) algorithm for frequency selective MIMO channels. The algorithm we develop is a variant of the popular vector Viterbi algorithm (VVA). Complexity reductions are achieved by an approach similar to sphere decoding (SD) to update the path metrics and select the survived states at each stage in the VVA. We show that a 60% reduction in complex multiplications and real additions is achieved for 2×2 MIMO systems with 16-QAM signal constellations. More savings are achieved for higher order signal constellations and larger MIMO

systems.

In our thesis, we use results from random matrix theory, information theory, and MIMO Orthogonal Frequency Division Multiplexing (OFDM) communications. This is why we dedicate Appendices A, B, and C to discussing them in detail.

1.5 Notation

- $x(t)$ is a time domain signal
- $x(\omega)$ is the continuous time Fourier transform (CTFT) of $x(t)$
- $\mathbf{x}$ ($\mathbf{X}$) is a vector (matrix) of scalar entries
- $\mathbf{x}(t)$ ($\mathbf{X}(t)$) is a vector (matrix) of continuous time signals
- $\mathbf{x}(\omega)$ ($\mathbf{X}(\omega)$) is the entry-wise CTFT of $\mathbf{x}(t)$ ($\mathbf{X}(t)$)
- $\mathbf{I_N}$ denotes the $N \times N$ identity matrix
- $\mathbf{0_{L\times K}}$ represents the $L \times K$ zero matrix
- $x \star y(t)$ represents the convolution of $x(t)$ with $y(t)$
- $\mathbf{X}^*$ is the conjugate transpose of $\mathbf{X}$
- $\mathbf{X}^T$ is the transpose of $\mathbf{X}$
- $\det(\mathbf{X})$ and $\mathrm{tr}(\mathbf{X})$ denote the determinant and trace of $\mathbf{X}$ respectively
- $\mathrm{diag}(\mathbf{x})$ represents a diagonal matrix formed by the entries of $\mathbf{x}$
- $f'(x)$ is the first order derivative of $f(x)$
- $f''(x)$ is the second order derivative of $f(x)$

CHAPTER 2

FUNDAMENTALS AND MODELING

2.1 Fundamentals

Modern optical fiber systems transmit digital information (voice, video, or files) through an optical fiber using light waves. Optical fibers are cylindrical waveguides made from glass or plastic (for low cost systems). The transmission and reception processes in an optical communication system are depicted in Figure 2.1.

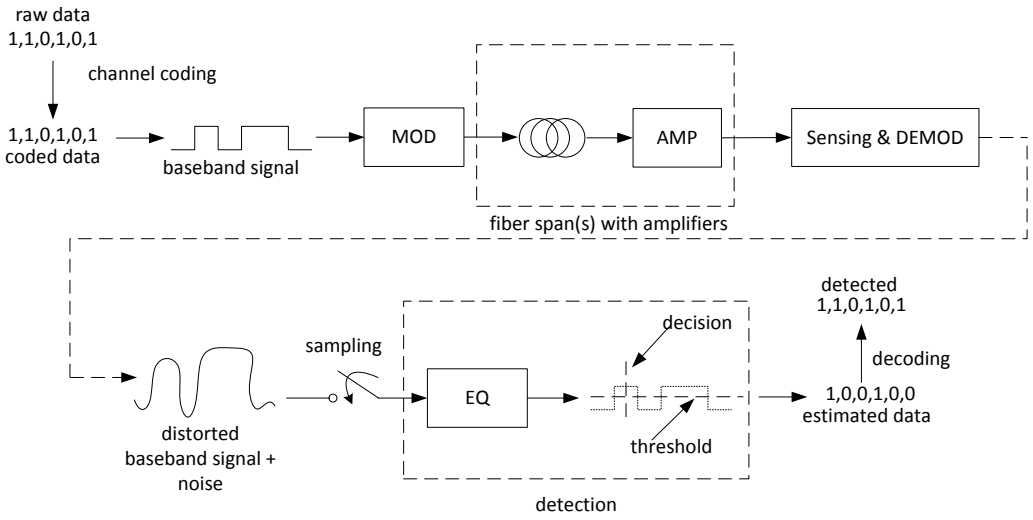
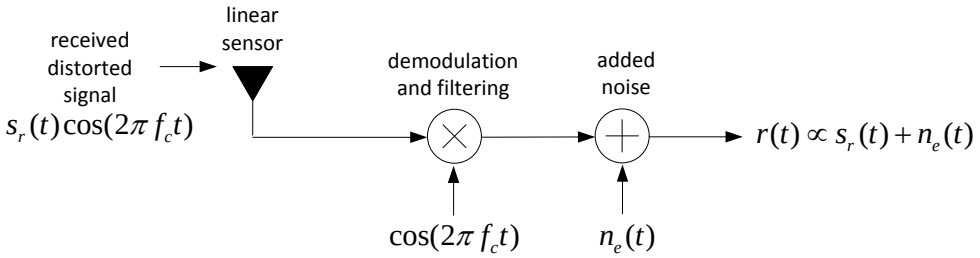


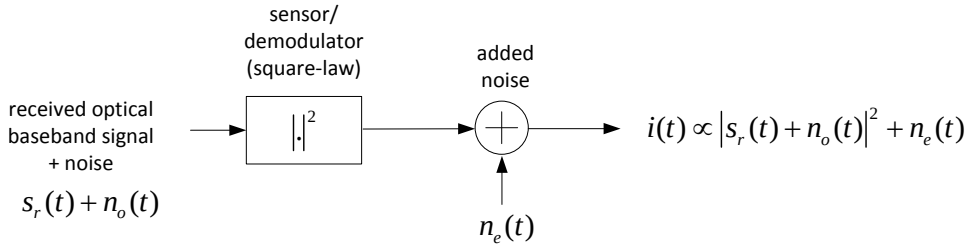
Figure 2.1: An optical communication system using single-mode fibers and on-off keying (reproduced from [2])

A block of raw data bits is first encoded via a channel code to yield a longer block of coded bits. The coded bits are mapped to a baseband signal that drives a laser source or a light emitting diode (LED). This essentially converts an electrical baseband signal to an optical passband signal. The optical

signal (electromagnetic field) is then transmitted through a cylindrical shaped fiber. If the fiber is sufficiently long, optical amplification is employed to boost the energy of the attenuated signal. At the other end of the fiber, the signal is detected using a detector such as a photo detector array (if an LED is used for transmission) or a MachZehnder interferometer (if a coherent laser is used for transmission). The combination of LED and photo detector array gives rise to a non-coherent optical communication system where the received signal is proportional to the energy of the transmitted field. Thus, any information transmitted through the phase is lost. On the other hand, the combination of coherent laser source and MachZehnder interferometer gives rise to a coherent optical communication system where the field's phase and amplitude are both recovered at the output of the fiber. In either case, the received signal is a distorted and attenuated version of the transmitted one with an additional, potentially signal dependent, noise component due to electronic processing and optical amplification. The task of the receiver is to apply equalization and detection techniques in order to recover the coded bits. Finally, a decoder is used to correct for detection errors and output the original block of information bits. Figure 2.2 summarizes the differences between coherent and non-coherent optical systems. In this figure, $s_r(t)$ represents the field at the output of the fiber. Coherent systems use well calibrated phase controlled laser sources and local oscillators operating well in the terahertz regime (hundreds of THz) to transmit and recover the phase and amplitude of an information-bearing signal. On the other hand, non-coherent systems use simple LEDs and photo detectors to transmit and detect the energy of an information signal. Thus, coherent systems are more complex, more expensive, and harder to build and maintain when compared to non-coherent systems. This is why the majority of currently deployed optical systems are non-coherent while a small percentage of the high end ones are coherent. However, coherent systems are becoming more popular as the optoelectronic devices are becoming more affordable. In fact, the state-of-the-art optical systems use both polarization and quadrature multiplexing to multiply the data rate by a factor of four. For example, OC-768 systems use the horizontal and vertical polarizations in addition to quadrature phase-shift keying (QPSK) to multiplex four independent data streams and transmit them all at the same time. An example of an OC-768 transmitter and receiver is shown in Figure 2.3.

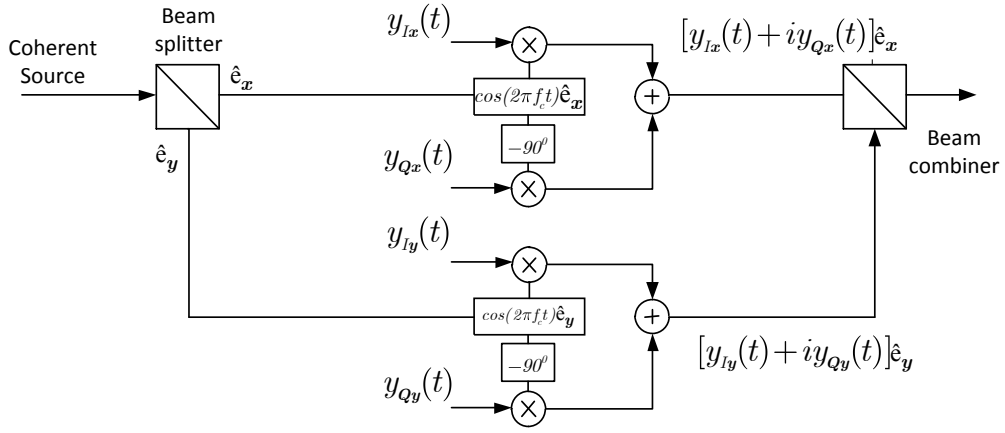


(a) Coherent optical communications

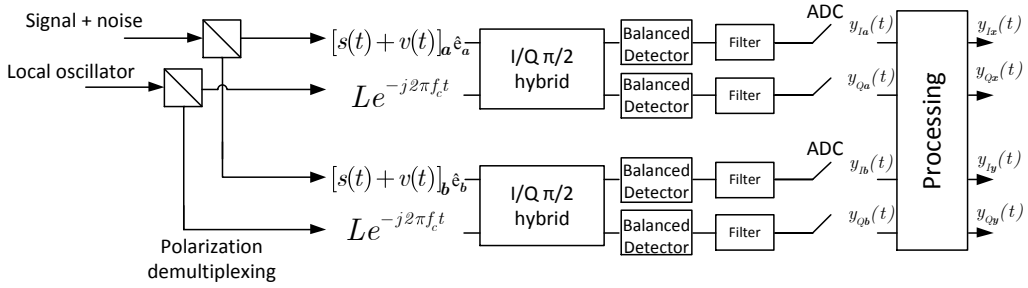


(b) Non-coherent optical communications

Figure 2.2: Optical fiber communications (reproduced from [2])



(a) Transmitter



(b) Receiver

Figure 2.3: Schematic of dual polarization QPSK optical transmitter and receiver (reproduced from [2])

The OC-768 network has transmission speeds of 40 Gbit/s. This means that in a system using QPSK and dual polarization, the transmitter operates at a frequency of about 10 GHz. Because coherent systems are becoming more popular and affordable, the capacity analysis we perform in Chapter 3 is exclusively applicable to coherent systems.

2.2 Single-Mode vs. Multi-Mode Fibers

Electromagnetic waves propagating inside the core of a fiber are characterized by Maxwell's equations. When the core radius is sufficiently small, only one solution to the wave equations is supported and the fiber is said to be a single-mode fiber. In multi-mode fiber systems, the core radius is relatively large and hence there is more than one solution (propagation mode) to the wave equation [7]. Ideally, the field inside the core would propagate in different orthogonal modes that do no interact with one another. However, due to manufacturing non-idealities and index of refraction inhomogeneities, the modes may couple. This phenomenon is called mode coupling and is modeled in Section 2.4.

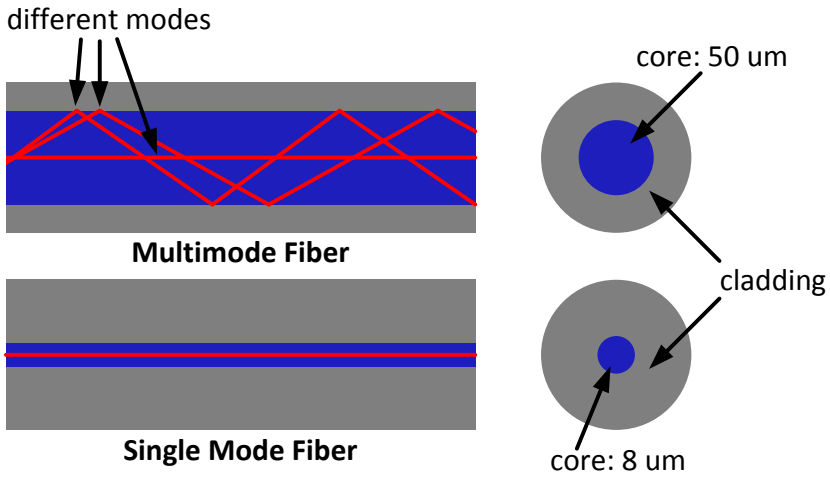


Figure 2.4: Single-mode vs. multi-mode fibers

A geometric optics visualization for modes is given in Figure 2.4. From a geometric optics point of view, the electromagnetic modes are represented by rays and their propagation through the fiber is described by a simple

ray tracing model. As can be seen in Figure 2.4, single-mode fibers are characterized by a single line of sight path from the input to the output of the fiber. However, in MMFs the light pulse takes several optical paths of different lengths. The fact that some paths are longer than others leads to pulse spreading at the output of the fiber. This phenomenon is called intermodal dispersion and will be formally presented in Section 2.4. In order to combat intermodal dispersion, the index of refraction of the fiber's core is graded to make its distribution non-uniform. A parabolic index profile, shown in Figure 2.5, is typically used in commercial multi-mode fibers.

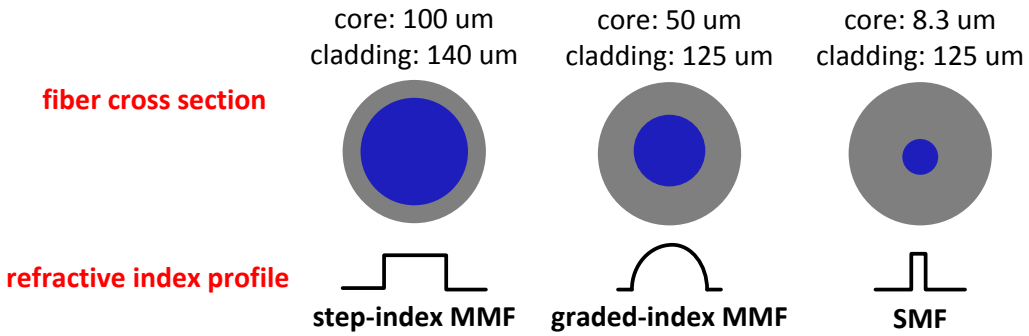


Figure 2.5: Step index vs. graded index single-mode and multi-mode fibers

Intuitively, as the optical signal moves into a region where the index of refraction is lower, its speed is increased and hence the total time needed for propagation is decreased. This reduces the delay spread, the maximum difference between the arrival times of different modes at the output of the fiber, and thus mitigates the effect of intermodal dispersion (cf. Section 2.4 for a formal definition of delay spread). Fibers that have uniform index of refraction are called step index fibers while fibers with graded index of refraction are called graded index fibers. The effect of index of refraction grading is depicted in Figure 2.6.

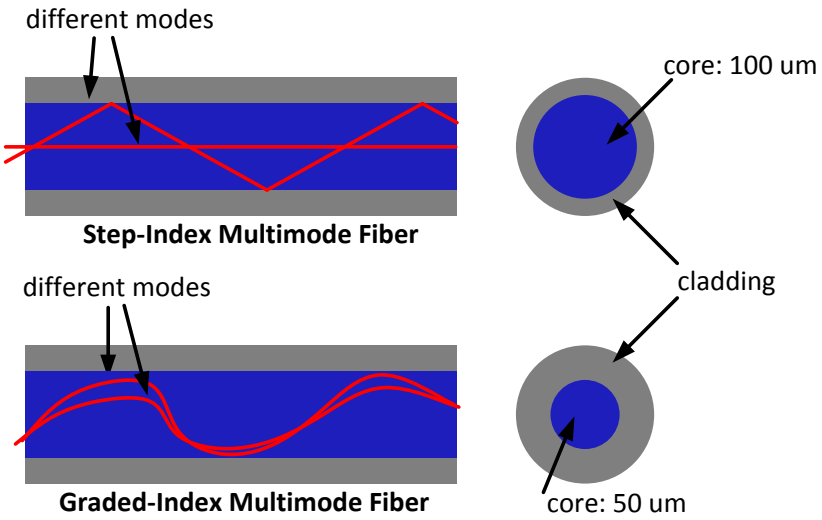


Figure 2.6: Step index vs. graded index fibers

2.3 Degrees of Freedom

Loosely speaking, the degrees of freedom (DoF) of a channel is an upper limit on the number of independent data streams that can be transmitted through the channel over a period of time. A more rigorous definition of DoF is given in [8] as the minimial dimension of the received signal space. The quadrature and in-phase components of a passband information signal are two familiar and commonly exploited degrees of freedom in wired and wireless communication systems. Figure 2.7 shows the degrees of freedom available in a multimode fiber communication system. Frequency, time, code, quadrature, and polarization states are all well explored and already utilized in commercial optical systems. However, the spatial degree of freedom, which is unique to MMFs, has not been exploited yet in commercial products and is still under research. In 2000, H. R. Stuart was the first to notice the similarity between the multipath wireless channel and the MMF optical channel and suggested using the spatial modes to multiplex several independent data streams onto the fiber [9]. Prior to this finding, single-mode fibers were always considered to be superior to MMFs because of their improved bandwidth-distance product (as SMFs do not suffer from intermodal dispersion). However, we will show in Chapter 3 that MMFs have advantages over single-mode fibers from

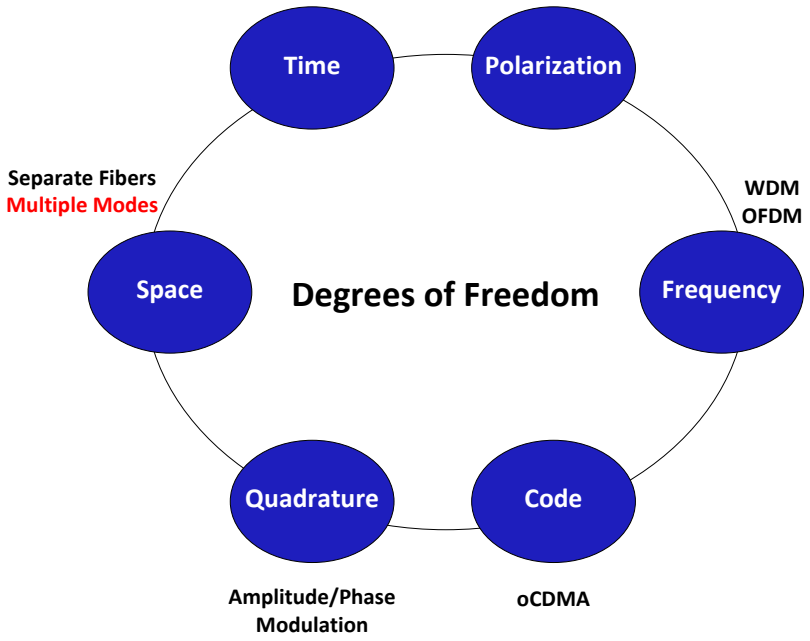


Figure 2.7: Degrees of freedom in multi-mode fibers (reproduced from [3])

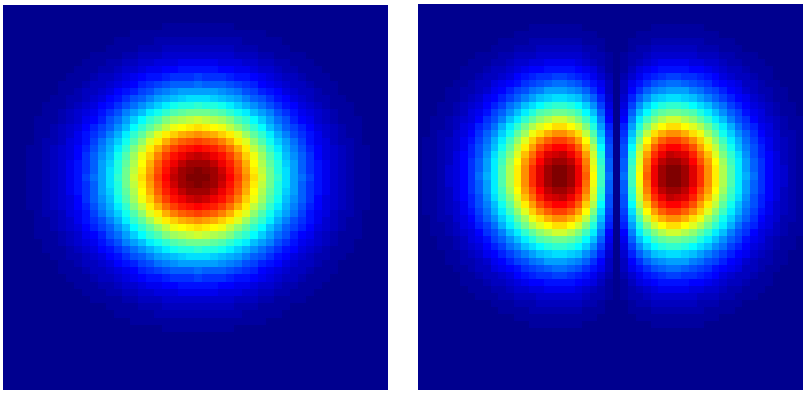
a capacity perspective. Therefore, MIMO over MMFs seems to be a better route to higher data rates.

2.4 Fiber Propagation Model

For coherent optical systems operating in the linear regime, the basic form of the baseband transfer function governing the input-output relationship of the i^{th} mode is given by

$$H_i(x, y, z, \omega) = \phi_i(x, y, \omega)\, e^{-\frac{\kappa_i z}{2}} e^{-j\beta_i(\omega+\omega_c)z} \tag{2.1}$$

where ω_c is the laser's center frequency, $\phi_i(x, y, \omega)$ is the transverse function (spatial pattern) of the i^{th} mode, κ_i is the mode-dependent attenuation factor, and $\beta_i(\omega + \omega_c)$ is the i^{th} mode's propagation constant [2]. Figure 2.8 shows the spatial patterns of two low order modes in a fiber with a perfect parabolic index of refraction profile. Expanding the function $\beta_i(\omega)$ around ω_c using its Taylor series expansion, and keeping the first and second order



(a) 0^{th} order mode (b) 1^{st} order mode

Figure 2.8: Lower order modes in an MMF

derivative terms, we get

$$H_i(x, y, \omega) \approx \phi_i(x, y)\, e^{\frac{g_i}{2}} e^{-j\theta_i} e^{-j\omega\tau_i} e^{-j\omega^2\alpha_i} \tag{2.2}$$

where $\phi_i(x, y) = (x, y, L, \omega_c)$, $g_i = -\kappa_i L$, $\theta_i = \beta_i(\omega_c) L$, $\tau_i = \beta_i^{'}(\omega_c) L$, and $\alpha_i = \beta_i^{''}(\omega_c) L$. Observe that z has been suppressed as it has been evaluated at L, the fiber's length. The function $\phi_i(x, y, \omega)$ generally depends on ω but since the signal spectrum (tens of GHz) is narrow around the laser's center frequency (hundreds of THz), we drop this dependency and evaluate it at ω_c. The model in (2.2) assumes that the propagation of the mode is completely characterized by a second order linear model where the only phenomena exhibited along the i^{th} mode are

- *mode-dependent loss* (MDL): $g_i = -\kappa_i L$
- *mode-dependent phase shift* (MDPS): $\theta_i = \beta_i(\omega_c) L$
- *group delay* (GD): $\tau_i = \beta_i^{'}(\omega_c) L$
- *group velocity dispersion* (GVD): $\alpha_i = \beta_i^{''}(\omega_c) L$

The mode-dependent losses (MDLs) are negative quantities describing the attenuation experienced by the modal fields. On the other hand, the mode-dependent phase shifts (MDPSs) represent phase shifts experienced by the modal fields. In general, modal fields propagate at different speeds and thus the group delays (GDs) characterize the arrival times of different modes. Therefore, if we transmit a narrow pulse through the fiber, it would appear

as a pulse having a width of $T_d = \max_{i,j}\{|\tau_i - \tau_j|\}$ at the output of the fiber. The quantity T_d is referred to as the channel's delay spread. Assume, without loss of generality, that the group delays are sorted in increasing order, τ_1 being the smallest and τ_M being the largest. In this case, T_d is given by

$$\begin{aligned} T_d &= \max_{i,j}\{|\tau_i - \tau_j|\} \\ &= \tau_M - \tau_1 \\ &= L\left(\beta_M^{'}(\omega_c) - \beta_1^{'}(\omega_c)\right) \end{aligned} \tag{2.3}$$

Thus, T_d is directly proportional to the length of the fiber. The pulse broadening phenomenon, due to nonzero T_d, is called intermodal dispersion and is a serious performance limitation in MMF systems. The group velocity dispersion (GVD), also called chromatic dispersion (CD), suggests that different frequencies coupled to the same mode propagate at different speeds and hence broadening occurs to the field propagating in a particular mode. This phenomenon is called intra-modal dispersion. In a first-order model, intermodal dispersion is assumed to dominate over intra-modal dispersion and the GVD term is typically neglected, especially for shorter lengths L. Furthermore, since we are not interested in analyzing the field at every point (x, y) in the fiber's core, we suppress this term to obtain the following expression:

$$H_i(\omega) \propto e^{\frac{g_i}{2}} e^{-j\theta_i} e^{-j\omega\tau_i} e^{-j\omega^2\alpha_i} \tag{2.4}$$

Ideally, the field at the output due to the i^{th} mode is given by $r_i(t) = s_i \star h_i(t)$, where $s_i(t)$ is the field at the input due to the same mode. Thus, the frequency domain vector representation of the modal fields at the output of the fiber is given by

$$\begin{bmatrix} r_1(\omega) \\ \vdots \\ r_M(\omega) \end{bmatrix} = \begin{bmatrix} H_1(\omega) & & \\ & \ddots & \\ & & H_M(\omega) \end{bmatrix} \begin{bmatrix} s_1(\omega) \\ \vdots \\ s_M(\omega) \end{bmatrix} \tag{2.5}$$

where the off-diagonal entries are zero because the modes are assumed to be orthogonal. This analysis neglects the existing fiber aberrations such as fiber bends, index of refraction inhomogeneities, and random vibrations, and is therefore incomplete. In fact, the modes interact with one another and

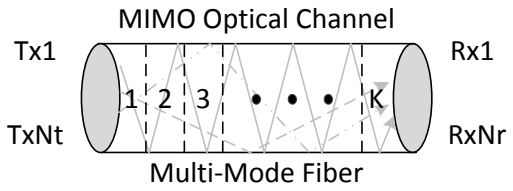


Figure 2.9: A multi-mode fiber with K propagation sections

exchange energy as they propagate along the fiber, complicating the analysis of the wave propagation. The treatment we present was first applied to polarization mode dispersion (PMD) in [10] and was then generalized to model mode coupling in [11]. In the regime of high mode coupling, for example when plastic optical fibers are used, an MMF with M modes[1] is split into $K \gg 1$ statistically independent longitudinal sections as depicted in Figure 2.9. The number of sections K is equal to L/l_c, where l_c represents the correlation length of the fiber. The frequency response of each section is given by

$$\mathbf{H}^k(\omega) = \mathbf{U}^k \mathbf{\Lambda}^k(\omega) \mathbf{V}^{k^*} \quad \texttt{for} \quad k = 1, ..., K \tag{2.6}$$

where $\mathbf{U}^k$ and $\mathbf{V}^k$ are $M \times M$ frequency-independent projection matrices (unitary matrices) describing the modal coupling via a phase and energy shuffling process at the input and output of each section and

$$\mathbf{\Lambda}^k(\omega) = \text{diag}\left(e^{\frac{1}{2}g_1^k - j\theta_1^k - j\omega\tau_1^k - j\omega^2\alpha_1^k}, ..., e^{\frac{1}{2}g_M^k - j\theta_M^k - j\omega\tau_M^k - j\omega^2\alpha_M^k}\right) \tag{2.7}$$

is the propagation matrix describing the ideal (uncoupled) field propagation in the k^{th} section. This model assumes that mode coupling occurs at the interface of different sections while the propagation in each section is ideal (and is described by $\mathbf{\Lambda}^k(\omega)$). In (2.7), the vectors $\mathbf{g}^k = \left(g_1^k, ..., g_M^k\right)$, $\boldsymbol{\theta}^k = \left(\theta_1^k, ..., \theta_M^k\right)$, $\boldsymbol{\tau}^k = \left(\tau_1^k, ..., \tau_M^k\right)$, and $\boldsymbol{\alpha}^k = \left(\alpha_1^k, ..., \alpha_M^k\right)$ represent the uncoupled MDL, MDPS, GD, and GVD coefficients in the k^{th} section. Here, $g_i^k = -\kappa_i^k l_c$, $\theta_i^k = \beta_i^k(\omega_c) l_c$, $\tau_i^k = \beta_i'^k(\omega_c) l_c$, and $\alpha_i^k = \beta_i''^k(\omega_c) l_c$ are not necessarily identical across the M modes and K sections and will be modeled as random variables in Section 2.5. The overall channel frequency response is equal to the product of the frequency responses of the K sections and is

[1]In this work, M refers to all the available spatial degrees of freedom including the x and y polarization states.

given by

$$\mathbf{H}(\omega) = \mathbf{H}^{(K)}(\omega) ... \mathbf{H}^{(1)}(\omega) \tag{2.8}$$

Alternatively, one could describe the input output relationship in time domain by

$$\mathbf{H}(t) = \mathbf{H}^{(K)} \star \mathbf{H}^{(K-1)} ... \mathbf{H}^{(2)} \star \mathbf{H}^{(1)}(t) \tag{2.9}$$

In (2.9), the operation $\mathbf{C}(t) = \mathbf{A} \star \mathbf{B}(t)$ represents a matrix convolution operation. Specifically, the $(i,j)^{th}$ entry of $\mathbf{C}(t)$ is given by

$$c_{ij}(t) = \sum_{l=1}^{M} a_{il} \star b_{lj}(t) \tag{2.10}$$

where M is the dimension of the square matrices $\mathbf{A}(t)$ and $\mathbf{B}(t)$.

2.5 Random Propagation Model

We now develop a random propagation model for the MIMO optical channel. The random model we introduce is an extended variant of what was presented in [11] and [12]. The per-section coupling matrices $\mathbf{U}^k$ and $\mathbf{V}^k$ are modeled as independent and identically distributed (i.i.d.) random unitary matrices with arbitrary distributions. Appendix A provides a brief introduction to random unitary matrices and some results that will prove useful when we compute the channel's capacity. We assume that the propagation characteristics g_i^k, θ_i^k, α_i^k, and τ_i^k are all independent random quantities. In addition, each of $\mathbf{g}^k$, $\boldsymbol{\theta}^k$, $\boldsymbol{\tau}^k$, and $\boldsymbol{\alpha}^k$ has zero mean identically distributed, but possibility correlated, entries. The zero mean assumption is not restrictive because the mean MDL, MDPS, GD, and GVD do not affect the capacity of the fiber. Even though the propagation characteristics are identically distributed within a particular section, they need not have the same distributions from one section to the other. We define σ_k to be the standard deviation of the uncoupled MDLs in the k^{th} section: $\sigma_k = \sqrt{\mathrm{Var}\left(g_i^k\right)} = l_c\sqrt{\mathrm{Var}\left(\kappa_i^k\right)}$. At any fixed frequency ω_0, the overall frequency response in (2.8) can be written as

$$\mathbf{H}(\omega_0) = \mathbf{U}_H \boldsymbol{\Lambda}_H \mathbf{V}_H^* \tag{2.11}$$

by the singular value decomposition (SVD) of $\mathbf{H}(\omega_0)$. In (2.11), all the matrices are random frequency dependent square matrices and

$$\mathbf{\Lambda}_H = \text{diag}\left(e^{\frac{1}{2}\rho_1}, ..., e^{\frac{1}{2}\rho_M}\right) \tag{2.12}$$

contains the *end-to-end eigenmodes*, singular values of $\mathbf{H}(\omega_0)$. We note that the end-to-end eigenmodes are not actual solutions to the wave equation, but rather they characterize the effective overall propagation through the fiber. The vector $\boldsymbol{\rho} = (\rho_1, \rho_2, ..., \rho_M)$ contains the *end-to-end mode-dependent losses*, the logarithms of the eigenvalues of $\mathbf{H}(\omega_\mathbf{0})\mathbf{H}^*(\omega_\mathbf{0})$. These quantities are obviously frequency dependent random variables as they are the logarithms of the eigenvalues of a frequency dependent random matrix. The *accumulated mode-dependent loss variance* is defined as

$$\xi^2 = \sigma_1^2 + \sigma_2^2 + ... + \sigma_K^2 \tag{2.13}$$

where ξ is measured in units of the logarithm of power gain and can be converted to decibels by multiplying its value by $10/\ln 10$ [11]. When all sections have identical distributions for the MDLs, Equation (2.13) reduces to $\xi^2 = K\sigma^2$ because $\sigma_k = \sigma$ for all k.

2.6 A First-Order Model

In order to get a feel for how mode coupling affects the field's propagation, we derive a first-order model for the end-to-end fiber impulse response. For the sake of simplicity, we ignore chromatic dispersion and assume that the propagation characteristics are deterministic and are identical for all sections. In other words, the random vector $\left(g_i^k, \theta_i^k, \alpha_i^k, \tau_i^k\right)$ is replaced by a deterministic vector $(g_i, \theta_i, 0, \tau_i)$ for all k. This is equivalent to replacing the random propagation matrices $\mathbf{\Lambda}^k(\omega)$ by a deterministic propagation matrix $\mathbf{\Lambda}(\omega)$. These assumptions are exclusive to this section and are relaxed in the chapters to follow. We now solve for $h_{l_0,l_{K+1}}(t)$, the $(l_0, l_{K+1})^{th}$ entry of $\mathbf{H}(t)$. Recall, from (2.9), that $\mathbf{H}(t)$ is the fiber's matrix impulse response and it contains M^2 impulse responses. Here, $h_{l_0,l_{K+1}}(t)$ is the impulse response of the effective channel between the input of the l_{K+1}^{th} mode and the output of the l_0^{th} mode. Therefore, each of l_0 and l_{K+1} can take values from 1 to M.

We use the indices (l_0, l_{K+1}) instead of (i, j) for notational convenience. Let $\mathbf{B}^k = \mathbf{V}^{k^*}\mathbf{U}^{(k-1)}$ for $2 \leq k \leq K$, $\mathbf{B}^{(K+1)} = \mathbf{U}^{(K)}$, and $\mathbf{B}^{(1)} = \mathbf{V}^{(1)^*}$, then Equation (2.8) can be written as

$$\begin{aligned} \mathbf{H}(\omega) &= \mathbf{H}^{(K)}(\omega)...\mathbf{H}^{(1)}(\omega) \\ &= \mathbf{U}^K\mathbf{\Lambda}^K(\omega)\mathbf{V}^{K^*}...\mathbf{U}^1\mathbf{\Lambda}^1(\omega)\mathbf{V}^{1^*} \\ &= \mathbf{B}^{(K+1)}\mathbf{\Lambda}^{(K)}(\omega)\mathbf{B}^{(K)}\mathbf{\Lambda}^{(K-1)}(\omega)\mathbf{B}^{(K-1)}...\mathbf{\Lambda}^{(1)}(\omega)\mathbf{B}^{(1)} \end{aligned} \tag{2.14}$$

We note that since $\mathbf{V}^{k^*}$ and $\mathbf{U}^{(k-1)}$ are independent random unitary matrices, their product $\mathbf{B}^k$ is a random unitary matrix. Taking the product of the $2K+1$ matrices in (2.14), one can show that the $(l_0, l_{K+1})^{th}$ entry of $\mathbf{H}(\omega)$ is given by

$$h_{l_0 l_{K+1}}(\omega) = \sum_{l_K=1}^{M} ... \sum_{l_1=1}^{M} \left[\prod_{m=0}^{K} b_{l_m, l_{m+1}}^{(K-m-1)} \right] e^{+\sum_{m=1}^{K}\left(\frac{1}{2}g_{l_m} - j\theta_{l_m} - j\omega\tau_{l_m}\right)} \tag{2.15}$$

where $b_{l_m, l_{m+1}}^{(K-m-1)}$ is the (l_m, l_{m+1}) entry of $\mathbf{B}^{(K-m-1)}$. At a first glance, it seems that all the K^M summands in (2.15) are distinct. However, the term $\frac{1}{2}g_{l_m} - j\theta_{l_m} - j\omega\tau_{l_m}$ can take one of at most M values because it has to belong to $\{\frac{1}{2}g_i - j\theta_i - j\omega\tau_i\}_{i=1}^{M}$. This means that the summation $\sum_{m=1}^{K}\left(\frac{1}{2}g_{l_m} - j\theta_{l_m} - j\omega\tau_{l_m}\right)$ can have up to M distinct summands. Thus, we can rewrite Equation (2.15) in a more compact and meaningful form as

$$h_{l_0 l_{K+1}}(\omega) = \sum_{k_1+...+k_M=K} B(k_1, ..., k_M)\, e^{+\frac{1}{2}\sum_{i=1}^{M} k_i g_i} e^{-j\sum_{i=1}^{M} k_i\theta_i} e^{-j\omega\sum_{i=1}^{M} k_i\tau_i} \tag{2.16}$$

where

$$B(k_1, ..., k_M) = \sum_{\{l_1,...,l_K | \#(l_i' s = j) = k_j, \forall j\}} \prod_{m=0}^{K} b_{l_m, l_{m+1}}^{(K-m-1)} \tag{2.17}$$

and the notation $\#(l_i' s = j) = k_j$ means that the number of l_i's equal to j is equal to k_j. This formulation is equivalent to choosing K elements from M distinct items with replacement, k_1 elements from item $\frac{1}{2}g_1 - j\theta_1 - j\omega\tau_1$, k_2 elements from item $\frac{1}{2}g_2 - j\theta_2 - j\omega\tau_2$, so on so forth. Therefore, the number of summands in (2.17) is given by the multinomial coefficient

$$\binom{K}{k_1, ..., k_M} = \frac{K!}{k_1!k_2!...k_M!} \tag{2.18}$$

Equation (2.16) can be further simplified and written as

$$h_{l_0 l_{K+1}}(\omega) = \sum_{k_1+...+k_M=K}^{M} A(k_1,...,k_M)\, e^{-j\omega\tau(k_1,...,k_M)} \tag{2.19}$$

where $\tau(k_1,...,k_M) = \sum_{i=1}^{M} k_i\tau_i$ is an effective group delay and

$$A(k_1,...,k_M) = B(k_1,...,k_M)\, e^{+\sum_{i=1}^{M} k_i\left(\frac{1}{2}g_i - j\theta_i\right)} \tag{2.20}$$

is an attenuation and phase shift factor. Alternatively, the time domain impulse response is given by

$$h_{l_0 l_{K+1}}(t) = \sum_{k_1+...+k_M=K}^{M} A(k_1,...,k_M)\,\delta(t-\tau(k_1,...,k_M)) \tag{2.21}$$

where the response has $\binom{M+K-1}{K}$ distinct impulse taps: $\delta(t-\tau(k_1,...,k_M))$. Furthermore, $\min\tau(k_1,...,k_M) = K\tau_1 = Kl_c\beta_1'(\omega_c)$ and $\max\tau(k_1,...,k_M) = K\tau_M = Kl_c\beta_M'(\omega_c)$ and hence mode coupling does not increase the delay spread $T_d = L\left(\beta_M'(\omega_c) - \beta_1'(\omega_c)\right)$. However, mode coupling spreads the energy of the signal in a non-unique way between $K\tau_1$ and $K\tau_M$. We define the energy of an impulse tap $\delta(t-\tau(k_1,...,k_M))$ as $\mathbb{E}[|A(k_1,...,k_M)|^2]$. When mode coupling is ignored, i.e. $K=1$, the impulse response has M impulse taps which represent the M modal group delays. In this case, the modes are perfectly orthogonal and do not interact with one another. As K increases, the number of taps increases and the energy distribution becomes concentrated around the center of the impulse response. This is because the average energy of an impulse tap $\delta(t-\tau(k_1,...,k_M))$ is directly proportional to the number of summands in Equation (2.17). When $k_1 = k_2 = ... = k_M = K/M$, assuming that K/M is an integer, the number of summands in (2.17) is maximized and hence the impulse tap corresponding to this assignment has the largest average energy. On the other hand, $\tau(k_1,...,k_M) = K\tau_M$ is achieved by setting $k_M = K$ and $k_1 = k_2 = ... = k_{M-1} = 0$ and hence there is only one term in the summation in (2.17). Similarly, the tap corresponding to $\tau(k_1,...,k_M) = K\tau_1$ has one term in the summation in (2.17) as well. This means that the energy of the impulse response taps for those two extreme points is very low when compared to the middle taps. This phenomenon is pictorially depicted in Figure 2.10. Even though the theoretical T_d is un-

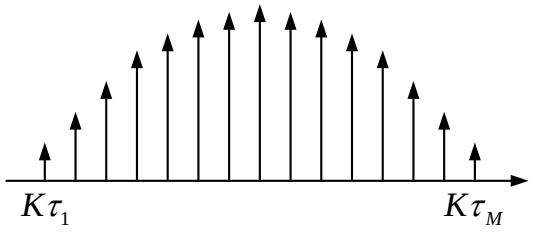


Figure 2.10: A sketch for $h_{l_0 l_{K+1}}(t)$

changed, the effective channel spread is reduced due to mode coupling. This is to be expected as mode coupling happens due to the exchange of energy between modes. Therefore, when K is large, a lot of energy mixing takes place (energy leaking from fast modes to slower modes and vice versa) and thus the middle tap will capture most of the signal energy.

CHAPTER 3

CAPACITY OF MULTI-MODE FIBERS

3.1 Literature Review

Following the work of Shannon [13], many information theorists investigated the capacity of different channels, including single-input single-output (SISO) channels with memory, channels with constrained input alphabet, and multiple-input multiple-output (MIMO) channels. In their seminal work, Telatar and Foschini *et al.*, independently showed that the capacity of a MIMO flat fading wireless channel, under the Raleigh fading model, scales linearly with respect to the minimum number of antennas at the transmitter and receiver [14, 15]. Since then, the wireless communications community has been focused on developing detection and coding schemes for MIMO systems in order to achieve the aforementioned capacity gains. Recent wireless technologies such as WLAN 802.11n and Long Term Evolution Advanced (LTE-A) are examples of MIMO systems deploying up to 8 transmitters and receivers. More importantly, as we discussed in Chapter 2, this MIMO technique is not limited to wireless systems. In fact, Stuart was the first to demonstrate the feasibility of a 2×2 MMF system and to show that there are indeed some capacity gains to be leveraged [9]. However, Stuart's analysis and experiments assumed a radio frequency sub-carrier ($\sim$ 1 GHz) instead of an optical carrier ($\sim$ 100 THz). This assumption was later relaxed in the work of Shah *et al.* but their treatment did not account for any intermodal dispersion, chromatic dispersion, or mode coupling [16]. Recently, the information theoretic capacity of coherent MMF systems has been studied in [4], where the authors ignored the frequency selectivity of the channel but incorporated the effects of mode coupling. In [11], Keang-Po *et al.* considered the capacity of a frequency selective MMF channel at a particular frequency.

They later studied the impact of frequency diversity on the channel's capacity for mutli-mode fibers with 10 modes [17]. However, their models did not incorporate the effect of mode-dependent phase shifts or chromatic dispersion.

In this chapter, we compute the capacity of coherent MMF systems under the presence of mode-dependent phase shifts (MDPSs), mode-dependent losses (MDLs), group delay (GD), chromatic dispersion (CD), and mode coupling. We review the information theoretic definition and operational meaning of channel capacity and derive the capacity of frequency flat and frequency selective MIMO channels in Appendix B. Table 3.1 summarizes the parameters governing the random propagation model presented in Section 2.5. Each of the vectors $\mathbf{g}^k$, $\boldsymbol{\theta}^k$, $\boldsymbol{\tau}^k$, and $\boldsymbol{\alpha}^k$ has zero mean identically

Table 3.1: Random propagation model

fiber's frequency response	$\mathbf{H}(\omega) = \mathbf{H}^K(\omega) ... \mathbf{H}^1(\omega)$
per-section response	$\mathbf{H}^k(\omega) = \mathbf{U}^k \mathbf{\Lambda}^k(\omega) \mathbf{V}^{k^*}$
per-section coupling matrices	$\mathbf{U}^k$ and $\mathbf{V}^k$
uncoupled MDL	$\mathbf{g}^k = \left(g_1^k, ..., g_M^k\right)$
uncoupled MDPS	$\boldsymbol{\theta}^k = \left(\theta_1^k, ..., \theta_M^k\right)$
uncoupled GD	$\boldsymbol{\tau}^k = \left(\tau_1^k, ..., \tau_M^k\right)$
uncoupled GVD	$\boldsymbol{\alpha}^k = \left(\alpha_1^k, ..., \alpha_M^k\right)$
uncoupled MDL variance	$\sigma_k^2 = \mathrm{Var}\left(g_i^k\right) = l_c^2 \mathrm{Var}\left(\kappa_i^k\right)$
accumulated MDL variance	$\xi^2 = \sigma_1^2 + \sigma_2^2 + ... + \sigma_K^2$

distributed, but possibility correlated, entries. Moreover, the vectors $\mathbf{g}^{k_1}$, $\boldsymbol{\theta}^{k_1}$, $\boldsymbol{\tau}^{k_1}$, and $\boldsymbol{\alpha}^{k_1}$ are independent of $\mathbf{g}^{k_2}$, $\boldsymbol{\theta}^{k_2}$, $\boldsymbol{\tau}^{k_2}$, and $\boldsymbol{\alpha}^{k_2}$ for $k_1 \neq k_2$. However, they can have the same statistical distributions. Recall, from Section 2.5, that the k^{th} section propagation matrix is given by

$$\begin{aligned} \mathbf{\Lambda}^k(\omega) &= \mathrm{diag}\left(e^{\frac{1}{2}g_1^k - j\theta_1^k - j\omega\tau_1^k - j\omega^2\alpha_1^k}, ..., e^{\frac{1}{2}g_M^k - j\theta_M^k - j\omega\tau_M^k - j\omega^2\alpha_M^k}\right) \\ &= \Theta^k \mathbf{T}^k \mathbf{A}^k \mathbf{G}^k \end{aligned} \tag{3.1}$$

where $\Theta^k = \mathrm{diag}\left(e^{-j\theta_1^k}, ..., e^{-j\theta_M^k}\right)$, $\mathbf{T}^k = \mathrm{diag}\left(e^{-j\omega\tau_1^k}, ..., e^{-j\omega\tau_M^k}\right)$, $\mathbf{A}^k = \mathrm{diag}\left(e^{-j\omega^2\alpha_1^k}, ..., e^{-j\omega^2\alpha_M^k}\right)$, and $\mathbf{G}^k = \mathrm{diag}\left(e^{\frac{1}{2}g_1^k}, ..., e^{\frac{1}{2}g_M^k}\right)$.

3.2 Frequency Flat Channel Capacity

We first study the capacity of the system when the channel's delay spread and CD are negligible. The more general frequency selective case is handled in Section 3.3. In this regime, $\max_{ij} |\tau_i^k - \tau_j^k| \approx 0$ and $\max_i |\alpha_i^k| \approx 0$ and hence $\tau_i^k = \tau^k$ and $\alpha_i^k = 0$ for all i and k. Therefore, the k^{th} section propagation matrix is given by

$$\begin{aligned} \mathbf{\Lambda}^k(\omega) &= \operatorname{diag}\left(e^{\frac{1}{2}g_1^k - j\theta_1^k - j\omega\tau_1^k - j\omega^2\alpha_1^k}, ..., e^{\frac{1}{2}g_M^k - j\theta_M^k - j\omega\tau_M^k - j\omega^2\alpha_M^k}\right) \\ &= e^{-j\omega\tau^k} \operatorname{diag}\left(e^{\frac{1}{2}g_1^k - j\theta_1^k}, ..., e^{\frac{1}{2}g_M^k - j\theta_M^k}\right) \\ &= e^{-j\omega\tau^k}\Theta^k\mathbf{G}^k \\ &= e^{-j\omega\tau^k}\mathbf{\Lambda}^k \end{aligned} \tag{3.2}$$

where $\mathbf{\Lambda}^k = \Theta^k\mathbf{G}^k$. Therefore, the overall response can be written as

$$\begin{aligned} \mathbf{H}(\omega) &= e^{-j\omega\sum_{k=1}^K \tau^k}\mathbf{U}^K\mathbf{\Lambda}^K\mathbf{V}^{K^*}...\mathbf{U}^1\mathbf{\Lambda}^1\mathbf{V}^{1^*} \\ &= e^{-j\omega\sum_{k=1}^K \tau^k}\mathbf{U}_H\mathbf{\Lambda}_H\mathbf{V}_H^* \end{aligned} \tag{3.3}$$

where $\mathbf{U}_H$, $\mathbf{\Lambda}_H$, and $\mathbf{V}_H^*$ are all frequency independent. The term $e^{-j\omega\sum_{k=1}^K \tau^k}$ is a delay term and can be neglected if we assume that the transmitter and receiver are synchronized. Thus, the channel is frequency flat and is given by

$$\mathbf{H} = \mathbf{U}_H\mathbf{\Lambda}_H\mathbf{V}_H^* \tag{3.4}$$

Consequently, the input-output relationship under this frequency flat channel model in (3.4) is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v} \tag{3.5}$$

where $\mathbf{x}$ and $\mathbf{y}$ represent the transmitted and received vectors, respectively, and $\mathbf{v}$ represents the modal noise which is modeled as additive white Gaussian noise (AWGN) with covariance matrix $N_0\mathbf{I_M}$, N_0 being the noise power density per Hz. This assumes that coherent optical communication is used and that electronic noise dominates over amplified spontaneous emission (ASE) noise. In addition, the fiber non-linearities are neglected under the assumption that the signal's peak to average power ratio (PAPR) and peak power are both low enough. This condition is not restrictive because MMFs have large radii and hence can support more power (relative to single mode fibers) while

still operating in the linear region. The input-output model in (3.5) may seem identical to the wireless MIMO flat fading one. However, the Rayleigh fading i.i.d. model does not hold in our case because $\mathbf{H}$ is a product of K terms, each containing a random diagonal matrix sandwiched between two random unitary matrices. Moreover, the entries of $\mathbf{H}$ are correlated. From Appendix B, the capacity of a single instantiation of the channel in (3.4), when channel state information (CSI) is not available at the transmitter, is given by

$$\begin{aligned} C(\mathbf{H}) &= \log\det\left(\mathbf{I_M} + \frac{\mathrm{SNR}}{M}\mathbf{H}\mathbf{H}^*\right) \\ &= \sum_{n=1}^{M}\log\left(1+\frac{\mathrm{SNR}}{M}\lambda_n^2\right) \quad \mathtt{b/s/Hz} \end{aligned} \tag{3.6}$$

where $\mathrm{SNR} = P/N_0W$, P representing the total power divided equally across all modes and W representing the available bandwidth in Hz. The λ_n^2's are the eigenvalues of $\mathbf{H}\mathbf{H}^*$. If CSI is available at the transmitter, the capacity could be further increased through waterfilling [18, 8]. In this case, the transmitter pre-processes the transmit vector $\mathbf{x}$ by allocating powers using a waterfilling procedure and then multiplies $\mathbf{x}$ by $\mathbf{V}_H$. On the other side, the receiver multiplies the received vector $\mathbf{y}$ by $\mathbf{U}_H^*$. This effectively turns the MIMO channel into a set of parallel AWGN channels. In optical communications, the beamforming process assumes that the designer can couple the fields of different sources onto the fiber exactly as determined by $\mathbf{V}_H$. This procedure, though beneficial, is complicated as it necessitates the design of sophisticated reconfigurable mode-selective spatial filters using coherent spatial light modulators [19, 20].

In the above analysis, we considered the capacity of (3.5) for a given instantiation of $\mathbf{H}$. However, since $\mathbf{H}$ is random, the channel capacity $C(\mathbf{H})$ is random. In the fast fading regime, the ergodic capacity, expected value of $C(\mathbf{H})$, is desired as it dictates the fastest rate of transmission [8]. On the other hand, in the slow fading regime, the cumulative distribution function (CDF) of $C(\mathbf{H})$ is desired as it determines the probability of an outage event for a particular rate of transmission [8]. In either case, the cumulative distribution and the expected value of $C(\mathbf{H})$ are both functions of the distribution of $\boldsymbol{\lambda} = (\lambda_1^2, ..., \lambda_M^2)$, the eigenvalues of $\mathbf{H}\mathbf{H}^*$. From (3.4), the matrix $\mathbf{H}\mathbf{H}^* = \mathbf{U}_H\boldsymbol{\Lambda}_H^2\mathbf{U}_H^*$ is Hermitian and its eigenvalues, the squares of the sin-

gular values of $\mathbf{H}$, are real non-negative quantities. Recall from Section 2.5 that the quantity $\lambda_n = e^{\frac{1}{2}\rho_n}$ refers to the n^{th} end-to-end eigenmode and the quantity ρ_n refers to the n^{th} end-to-end mode dependent loss (MDL). The distribution of the end-to-end MDL values was studied in [11] where it was conjectured that as M tends to infinity, the ρ_n's become independent and identically distributed on a semicircle. Their analysis and simulations, however, did not incorporate the effect of mode dependent phase shifts (MDPSs), θ_i^k's. The following proposition shows that the statistical distribution of the end-to-end MDL values is unchanged even when MDPSs are incorporated.

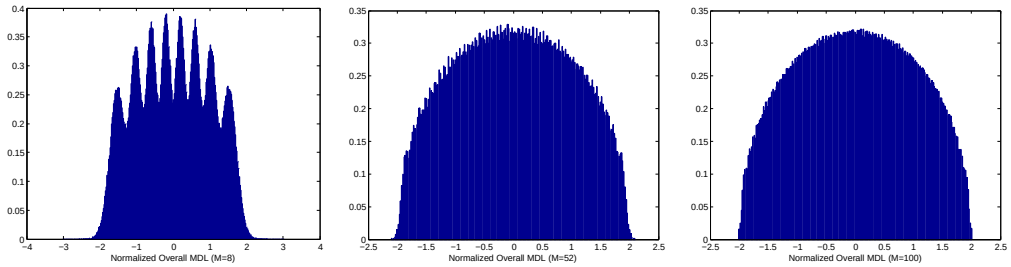
Proposition 3.2.1 *The statistics of* $\mathbf{H}$ *are unchanged when MDPSs are modeled.*

Proof 1 *In order to prove the above proposition, we show that the statistics of* $\mathbf{H}^k = \mathbf{U}^k\Theta^k\mathbf{G}^k\mathbf{V}^{k^*}$ *are the same as those of* $\mathbf{H}^k = \mathbf{U}^k\mathbf{G}^k\mathbf{V}^{k^*}$ *for all* $k = 1, ..., K$*. Observe that* Θ^k *is a unitary and random matrix because it has random orthonormal columns. However, the matrix does not necessarily belong to the class of random unitary matrices as it is not necessarily uniformly distributed over* $\mathbb{U}(M)$*, the space of* $M \times M$ *unitary matrices. Nonetheless, we note that the distribution of* $\mathbf{W} = \mathbf{U}^k\Theta^k$ *is the same as the distribution of* $\mathbf{U}^k$ *because*

$$\begin{aligned} f(\mathbf{W}) &= \int_{\Theta^k} f\left(\mathbf{W}|\Theta^k\right) f\left(\Theta^k\right) d\Theta^k \\ &= f\left(\mathbf{U}^k\right) \int_{\Theta^k} f\left(\Theta^k\right) d\Theta^k \\ &= f\left(\mathbf{U}^k\right) \end{aligned} \tag{3.7}$$

where the second equality holds because for a given instantiation of Θ^k*, the random matrix* $\mathbf{W}|\Theta^k$ *has the same distribution as* $\mathbf{U}^k$ *(cf. Lemma A.2.1 in Appendix A). Therefore, the statistics of* $\mathbf{H}$ *are unchanged when the MDPSs are incorporated, and thus the results in [11] carry over to this more general setting.*

Figure 3.1 shows that for $M = 100$ the distribution of ρ_n approaches a semicircle. The distributions in Figure 3.1 were obtained by generating a large sample of channel matrices $\mathbf{H}$ (for $M = 8$, $M = 52$, and 100) and estimating the distributions of the logarithm of their singular values. Appendix A



(a) M=8 (b) M=52 (c) M=100

Figure 3.1: Distribution of end-to-end MDL

explains how we can generate random unitary matrices, which are needed to create samples of $\mathbf{H}$, from matrices with i.i.d. complex Gaussian entries. The average capacity of $C(\mathbf{H})$ is given by

$$C_{\text{avg}} = \sum_{n=1}^{M} \mathbb{E}\left[\log\left(1 + \frac{\text{SNR}}{M} e^{\rho_n}\right)\right] \quad \texttt{b/s/Hz} \tag{3.8}$$

where the average is taken over the statistics of the end-to-end MDL values. Figure 3.2 shows the average capacity of MMFs for various values of M and $\xi = 4$ dB. The capacity of the system increases with an increasing number of modes. This is intuitive because as the number of modes increases, the fiber's spatial degrees of freedom are increased. Figure 3.3 shows the effect of accumulated MDLs on the average capacity. An increasing value of ξ results in a capacity equivalent to that of a fiber with fewer modes. This means that as ξ^2, the accumulated mode-dependent loss variance, increases the system loses its spatial degrees of freedom.

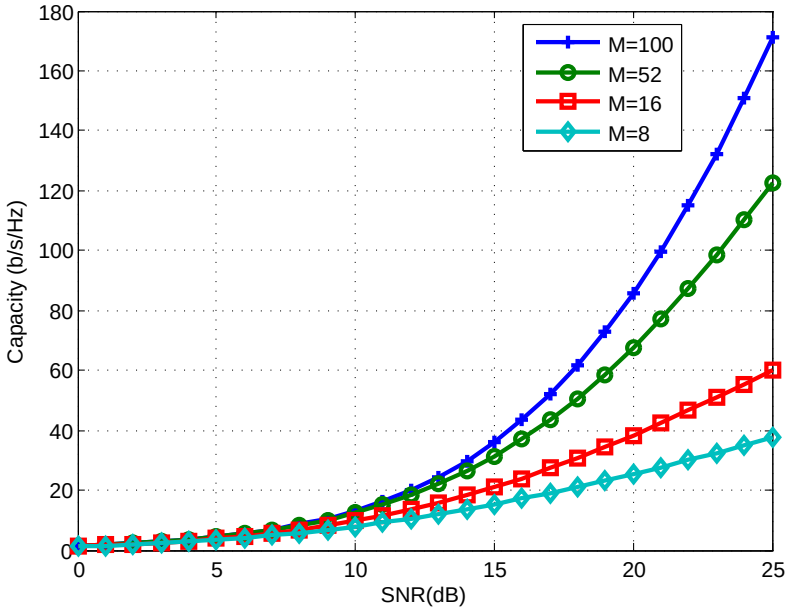


Figure 3.2: Capacity of MIMO MMF systems at $\xi = 4$ dB

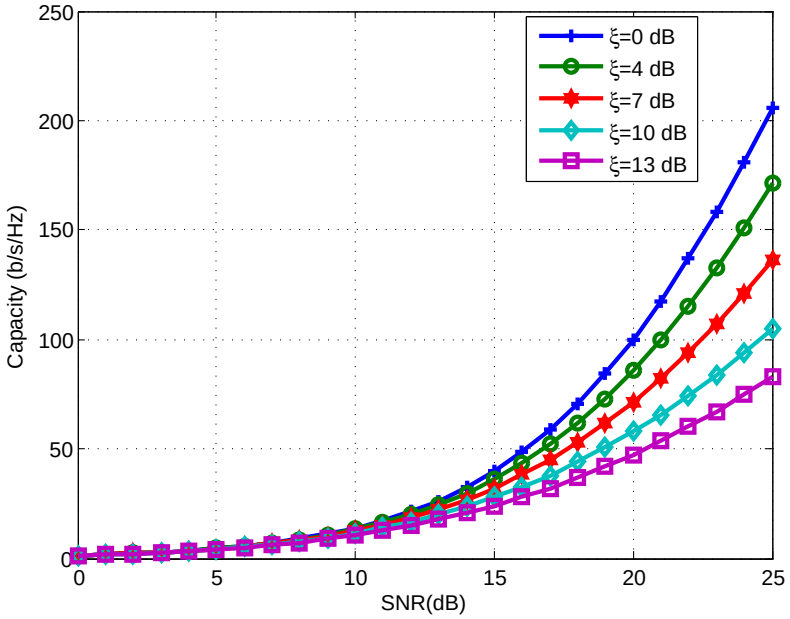


Figure 3.3: Effect of MDL for M=100

3.3 Frequency Selective Channel Capacity

When chromatic and intermodal dispersion are taken into account, the fiber's frequency response $\mathbf{H}(\omega)$ becomes frequency selective. Under the same linear

assumptions as in the previous section, the input-output relationship is given by

$$\mathbf{y}(t) = \mathbf{H}(t) \star \mathbf{x}(t) + \mathbf{v}(t) \tag{3.9}$$

where $\mathbf{v}(t)$ is a Gaussian vector process, $\mathbf{x}(t)$ is the input, and $\mathbf{y}(t)$ is the received signal. Recall, from Section 2.4, that $\mathbf{H}(\omega_0)$ can be written as

$$\mathbf{H}(\omega_0) = \mathbf{U}_H \mathbf{\Lambda}_H \mathbf{V}_H^* \tag{3.10}$$

where $\mathbf{U}_H$, $\mathbf{\Lambda}_H$, and $\mathbf{V}_H^*$ all depend on ω_0. From Appendix B, the capacity of a single instantiation of $\mathbf{H}(\omega)$, when CSI is not available at the transmitter, is equal to

$$C = \frac{1}{2\pi W} \int_0^{2\pi W} \log\det\left(\mathbf{I}_{\mathbf{N_r}} + \frac{\mathrm{SNR}}{M}\mathbf{H}(\omega)\,\mathbf{H}^*(\omega)\right) d\omega \quad \mathtt{b/s/Hz} \tag{3.11}$$

where W is the bandwidth of the system in Hz and $\mathrm{SNR} = P/N_0W$ [8]. This capacity can be achieved by Orthogonal Frequency Division Multiplexing (OFDM) with N sub-carriers (as N tends to infinity). MIMO OFDM modulation is a popular modulation scheme in wireless communications and is currently being developed for the next generation optical systems [21]. We describe this scheme in detail in Appendix C. The maximum achievable capacity of a MIMO-OFDM system with N sub-carriers is

$$\begin{aligned} C &= \frac{1}{N}\sum_{i=1}^{N} \log\det\left(\mathbf{I}_{\mathbf{N_r}} + \frac{\mathrm{SNR}}{M}\mathbf{H}_i\mathbf{H}^*{}_i\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\sum_{n=1}^{M} \log\left(1 + \frac{\mathrm{SNR}}{M}\lambda_{n,i}^2\right) \quad \mathtt{b/s/Hz} \end{aligned} \tag{3.12}$$

where $\mathbf{H}_i = \mathbf{H}(\omega_i)$ and $\lambda_{n,i}^2$ is the n^{th} eigenvalue of $\mathbf{H}_i\mathbf{H}^*{}_i$. When CSI is available at the transmitter, waterfilling can be performed to allocate optimal powers across sub-carriers and transmitters. This procedure is explained thoroughly in Appendix B.

In the above analysis, we considered the capacity of (3.9) for a given instantiation of $\mathbf{H}(\omega)$. We focus on analyzing the expected capacity of the

frequency selective system which is now given by

$$C_{\text{avg}} = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{n=1}^{M}\log\left(1+\frac{\text{SNR}}{M}\lambda_{n,i}^2\right)\right] \quad \texttt{b/s/Hz} \tag{3.13}$$

To begin with, if we assume that, in each section, all modes experience the same random loss (i.e., the entries of $\mathbf{g}^k$ are perfectly correlated), then $\mathbf{G}^k = e^{\frac{1}{2}g^k}\mathbf{I_M}$. Furthermore, assume that the K sections are statistically identical. Therefore, $\sigma_k = \sigma$ for all k and $\xi^2 = K\sigma^2$. In this case, the overall response is given by

$$\begin{aligned}\mathbf{H}(\omega) &= \mathbf{U}^K\mathbf{\Lambda}^K(\omega)\mathbf{V}^{K^*}...\mathbf{U}^1\mathbf{\Lambda}^1(\omega)\mathbf{V}^{1^*} \\ &= e^{\frac{1}{2}\sum_{k=1}^{K}g^k}\mathbf{U}^K\Theta^K\mathbf{T}^K\mathbf{A}^K\mathbf{V}^{K^*}...\mathbf{U}^1\Theta^1\mathbf{T}^1\mathbf{A}^1\mathbf{V}^{1^*}\end{aligned} \tag{3.14}$$

Observe that, even though $\mathbf{H}(\omega)$ is a function of ω, $\mathbf{H}(\omega)\mathbf{H}(\omega)^* = e^{\sum_{k=1}^{K}g^k}\mathbf{I_M}$ is independent of ω. This means that $\lambda_{n,i}^2 = e^{\sum_{k=1}^{K}g^k}$ is independent of the frequency index i and the mode number n. In the strong coupling regime (i.e., when K is large), the random variable $\sum_{k=1}^{K}g^k$ converges, in distribution, to a Gaussian random variable with variance equal to $\xi^2 = K\sigma^2$ by the central limit theorem. In other words, the distribution of the random variable $\lambda^2 = e^{\sum_{k=1}^{K}g^k}$ converges to a log-normal distribution. Thus, the average capacity of the fiber is given by

$$C_{\text{avg}} = M\mathbb{E}\left[\log\left(1+\frac{\text{SNR}}{M}\lambda^2\right)\right] \tag{3.15}$$

where the average is taken over the statistics of λ^2. Observe that average capacity scales linearly with M, the number of modes. In general, the distribution of $\lambda^2 = e^{\sum_{k=1}^{K}g^k}$ could converge to a log-normal distribution even if the K sections are not statistically identical. A sufficient condition for convergence is that the sequence $\{g^k\}_{k=1}^{K}$ satisfies Lindeberg's condition [22]. Thus, neither group delay nor chromatic dispersion affect the average capacity of the fiber.

We now derive the capacity for the general case (i.e., when the entries of $\mathbf{g}^k$ are potentially independent). The following proposition shows that the statistics of $\mathbf{H}(\omega)$ are independent of ω.

Proposition 3.3.1 *The statistics of* $\mathbf{H}(\omega)$ *are independent of* ω.

Proof 2 *Using the same proof as in proposition 3.2.1, we can show that the statistics of* $\mathbf{H}^k(\omega) = \mathbf{U}^k\Theta^k\mathbf{T}^k\mathbf{A}^k\mathbf{G}^k\mathbf{V}^{k^*}$ *are the same as the statistics of* $\mathbf{H}^k = \mathbf{U}^k\mathbf{G}^k\mathbf{V}^{k^*}$ *by showing that the distribution of* $\mathbf{W}^k = \mathbf{U}^k\Theta^k\mathbf{T}^k\mathbf{A}^k$ *is equal to the distribution of* $\mathbf{U}^k$*. Thus, the statistics of* $\mathbf{H}^k(\omega)$ *are independent of* ω*.*

This result shows that the statistics of the eigenvalues of $\mathbf{H}_i\mathbf{H}_i^*$ are identical for all i. Therefore, the average capacity expression can now be rewritten as

$$C_{\text{avg}} = \sum_{n=1}^{M} \mathbb{E}\left[\log\left(1 + \frac{\text{SNR}}{M}\lambda_n^2\right)\right] \quad \texttt{b/s/Hz} \tag{3.16}$$

which is identical to the average capacity of frequency flat optical MIMO systems. Therefore, the results of the previous section carry over to the frequency selective case.

CHAPTER 4

INPUT-OUTPUT COUPLING STRATEGIES

4.1 Input-Output Coupling

The capacity analysis presented in Chapter 3 is important, but it only serves as an upper limit on the achievable rate. This limit can only be achieved by making use of all available spatial modes. In theory, one can always design a fiber with a sufficiently small core radius such that a desired number of modes propagate through the fiber [7]. In reality, one has to rely on currently installed optical fibers and available technologies. The state-of-the art OM3 and OM4 MMF technologies have core radii of $50\,\mu$m with hundreds of propagation modes. Unfortunately, having a 100×100 MIMO system is neither physically nor computationally realizable at the moment. This means that a more careful look at the effective channel capacity has to be considered. This is why we now focus on the case when N_t transmit laser sources and N_r receivers are used. For most parts of this chapter, we assume that intermodal and chromatic dispersions are negligible. Even though this may seem like a restriction, this assumption actually serves to simplify the discussion and presentation of input-output coupling strategies. The results and procedures we present offer insight and can be extended to the more general frequency selective case.

4.2 Input-Output Coupling Model

The input coupling is described by $\mathbf{C_I}$, an $M \times N_t$ matrix, and the output coupling is described by $\mathbf{C_O}$, an $N_r \times M$ matrix. Here, M is much larger

than N_t and N_r and the overall response is given by

$$\begin{aligned} \mathbf{H}_t &= \mathbf{C_O}\mathbf{H}^{(K)}...\mathbf{H}^{(1)}\mathbf{C_I} \\ &= \mathbf{C_O}\mathbf{H}\mathbf{C_I} \end{aligned} \tag{4.1}$$

Therefore, for a single instantiation of $\mathbf{H}_t$, the capacity of the channel is given by

$$C\left(\mathbf{H_t}\right) = \log\det\left(\mathbf{I_{N_t}} + \frac{\mathrm{SNR}}{N_t}\mathbf{H_t}\mathbf{H_t^*}\right) \tag{4.2}$$

The input-output coupling coefficients (entries of $\mathbf{C_I}$ and $\mathbf{C_O}$) are complex quantities capturing the effect of both power and phase coupling into and out of the fiber. These coefficients are determined by the system geometry and launch conditions. For example, in order to study the input coupling profile of each light source one needs to specify its exact geometry and launching angle, and then solve the overlap integrals: two dimensional inner products between the laser's spatial patterns and those of each mode

$$c_{ij} = \int\int \phi_i\left(x,y\right)\phi_{s_j}\left(x,y\right)dxdy \tag{4.3}$$

where c_{ij} is the $(i,j)^{th}$ entry of $\mathbf{C_I}$, $\phi_i\left(x,y\right)$ is the i^{th} mode spatial pattern, and $\phi_{s_j}\left(x,y\right)$ is the j^{th} laser source spatial pattern. However, this procedure is cumbersome and offers little insight on the underlying channel physics. In what follows, we provide a simple condition on the input-output couplers. This condition will prove useful when we present an input-output scheme that maximizes the achievable rate of the overall system (Section 4.3) and impose a statistical model for $\mathbf{C_I}$ and $\mathbf{C_O}$ (Section 4.4).

Proposition 4.2.1 *If we neglect the power lost due to input coupling inefficiencies, then a necessary and sufficient condition for* $\mathbf{C_I}$ *to be an input coupling matrix is given by*

$$\left(\mathbf{c_i}, \mathbf{c_j}\right) = \delta_{ij} \tag{4.4}$$

where $\mathbf{c_i}$ *represents the* i^{th} *column of* $\mathbf{C_I}$ *and* $\left(\mathbf{a}, \mathbf{b}\right)$ *denotes the standard Euclidean inner product between the vectors* $\mathbf{a}$ *and* $\mathbf{b}$*. This means that the columns of* $\mathbf{C_I}$ *should form a complete orthonormal basis for* $\mathbb{C}^{N_t}$*. Similarly, if we neglect the power lost due to output coupling inefficiencies, then the rows of* $\mathbf{C_O}$ *should form a complete orthonormal basis for* $\mathbb{C}^{N_r}$*.*

Proof 3 *Satisfying the energy conservation principle requires that*

$$||\mathbf{C_I x}||^2 = ||\mathbf{x}||^2 \quad \forall \mathbf{x} \in \mathbb{C}^{N_t} \tag{4.5}$$

This means that the energy of the input vector should be equal to the energy of the mode vector at the input of the fiber. This condition holds whenever the mapping $\mathbf{C_I}$ *is a linear isometry mapping. In the special case where* $\mathbf{C_I}$ *is a square matrix, a classical result in linear algebra states that* $\mathbf{C_I}$ *has to be a unitary matrix [23]. However,* $\mathbf{C_I}$ *is a tall* $M \times N_t$ *(*$M \gg N_t$*) rectangular matrix. In this case, the condition in (4.5) can be rewritten as*

$$(\mathbf{C_I x}, \mathbf{C_I x}) \;\; = \;\; (\mathbf{x}, \mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{C}^{N_t} \tag{4.6}$$

or equivalently as

$$(\mathbf{x}, [\mathbf{C_I^* C_I} - \mathbf{I_{N_t}}]\, \mathbf{x}) = 0 \quad \forall \mathbf{x} \in \mathbb{C}^{N_t} \tag{4.7}$$

If $\mathbf{C_I^* C_I} = \mathbf{I_{N_t}}$*, the condition in (4.7) holds and* $\mathbf{C_I}$ *preserves the norm. This choice ensures that the columns of* $\mathbf{C_I}$ *form a complete orthonormal basis for* $\mathbb{C}^{N_t}$*. However, this only proves the sufficiency part of the proposition. To prove the necessity part, we consider* $\mathbf{B} = \mathbf{C_I^* C_I} - \mathbf{I_{N_t}}$ *and show that if (4.7) holds, then it is equal to zero. It can be easily verified that if* $\mathbf{B}$ *is a diagonal matrix, then* $(\mathbf{x}, \mathbf{Bx}) = 0 \quad \forall \mathbf{x} \in \mathbb{C}^{N_t}$ *implies that* $\mathbf{B} = \mathbf{0}$*. The same observation holds if* $\mathbf{B}$ *is diagonalizable. In this case, one can choose an orthonormal basis of eigenvectors and map* $(\mathbf{x}, \mathbf{Bx}) = 0 \quad \forall \mathbf{x} \in \mathbb{C}^{N_t}$ *to* $(\tilde{\mathbf{x}}, \mathbf{D}\tilde{\mathbf{x}}) = 0 \quad \forall \tilde{\mathbf{x}} \in \mathbb{C}^{N_t}$ *where* $\mathbf{D}$ *is a diagonal matrix. In our case,* $\mathbf{B}$ *is Hermitian and hence it is unitarily diagonalizable so that* $\mathbf{B} = \mathbf{0}$ *or alternatively,* $\mathbf{C_I^* C_I} = \mathbf{I_{N_t}}$ *as desired. A similar proof can be carried out to show that* $\mathbf{C_O C_O^*} = \mathbf{I_{N_r}}$*.*

We note that even though $\mathbf{C_I^* C_I} = \mathbf{I_{N_t}}$, $\mathbf{C_I C_I^*} \neq \mathbf{I_{N_t}}$ because $\mathbf{C_I}$ is of full column rank, but not of full row rank. We now use this property to show that $\mathbf{C_I}$ and $\mathbf{C_O}$ should have a special structure.

Proposition 4.2.2 *The input-output coupling matrices* $\mathbf{C_I}$ *and* $\mathbf{C_O}$ *can be expressed as*

$$\mathbf{C_I} = \mathbf{U_I} \begin{bmatrix} \mathbf{I_{N_t}} \\ \mathbf{0_{(M-N_t) \times N_t}} \end{bmatrix} \mathbf{V_I^*} \tag{4.8}$$

$$\mathbf{C_O} = \mathbf{U_O}\left[\mathbf{I_{N_r}} \mathbf{0_{N_r \times (M-N_r)}}\right] \mathbf{V_O^*} \tag{4.9}$$

where $\mathbf{U_I}$ *and* $\mathbf{V_O^*}$ *are* $M \times M$ *unitary matrices,* $\mathbf{V_I^*}$ *is an* $N_t \times N_t$ *unitary matrix, and* $\mathbf{U_O}$ *is an* $N_r \times N_r$ *unitary matrix.*

Proof 4 *By the singular value decomposition (SVD),* $\mathbf{C_I} = \mathbf{U_I \Lambda_I V_I^*}$ *and* $\mathbf{C_O} = \mathbf{U_O \Lambda_O V_O^*}$ *[23]. The non-zero singular values of* $\mathbf{C_I}$ *are the square roots of the eigenvalues of* $\mathbf{C_I^* C_I}$*, and* $\mathbf{C_I^* C_I} = \mathbf{I_{N_t}}$*. A similar argument holds for* $\mathbf{C_O}$*.*

4.3 Input-Output Coupling Strategies

In this section, we assume that $N_t = N_r$. When CSI is available at the transmitter and the design of $\mathbf{C_I}$ and $\mathbf{C_O}$ is affordable, an optimal choice for the input-output couplers is the one that maximizes the system's capacity

$$\begin{aligned}
\left(\mathbf{C_I^{opt}}, \mathbf{C_O^{opt}}\right) &= \arg \max_{(\mathbf{C_I}, \mathbf{C_O})} \log\det\left(\mathbf{I_{N_r}} + \frac{\mathrm{SNR}}{N_t}\mathbf{H_t H_t^*}\right) \\
&= \arg \max_{(\mathbf{C_I}, \mathbf{C_O})} \sum_{n=1}^{N} \log\left(1 + \frac{\mathrm{SNR}}{N_t}\lambda_n^2\right)
\end{aligned} \tag{4.10}$$

where the λ_n^2's are the eigenvalues of $\mathbf{H_t H_t^*}$. We note that $\mathbf{C_I^{opt}}$ and $\mathbf{C_O^{opt}}$ should have a structure compliant with (4.8) and (4.9), respectively. Instead of solving the above constrained optimization problem, we provide an intuitive choice for $(\mathbf{C_I}, \mathbf{C_O})$ and argue that it leads to a maximized overall capacity through simulations.

Proposition 4.3.1 *The capacity of the overall system in (4.2) is independent of the choice of* $\mathbf{V_I^*}$ *and* $\mathbf{U_O}$ *from (4.8) and (4.9).*

Proof 5 *The capacity of the overall system is given by*

$$\begin{aligned}
C\left(\mathbf{H_t}\right) &= \log\det\left(\mathbf{I_{N_t}} + \frac{\mathrm{SNR}}{N_t}\mathbf{H_t H_t^*}\right) \\
&= \log\det\left(\mathbf{I_{N_t}} + \frac{\mathrm{SNR}}{N_t}\mathbf{C_O H C_I C_I}^*\mathbf{H}^*\mathbf{C_O^*}\right) \\
&= \log\det\left(\mathbf{I_{N_t}} + \frac{\mathrm{SNR}}{N_t}\mathbf{U_O \Lambda_O V_O^* H U_I \Lambda_I \Lambda_I^* U_I^* H}^*\mathbf{V_O \Lambda_O^* U_O^*}\right) \\
&= \log\det\left(\mathbf{I_{N_t}} + \frac{\mathrm{SNR}}{N_t}\mathbf{\Lambda_O V_O^* H U_I \Lambda_I \Lambda_I^* U_I^* H}^*\mathbf{V_O \Lambda_O^*}\right)
\end{aligned} \tag{4.11}$$

Therefore, the capacity of the overall system is independent of $\mathbf{V_I^*}$ *and* $\mathbf{U_O}$ *and hence, without loss of generality, we will assume that they are both equal to the identity matrix.*

The following input-output coupling scheme is suggested

$$\mathbf{C_I} = \mathbf{V_H} \begin{bmatrix} \mathbf{I_{N_t}} \\ \mathbf{0_{(M-N_t)\times N_t}} \end{bmatrix} \tag{4.12}$$

$$\mathbf{C_O} = [\mathbf{I_{N_r}} \mathbf{0_{(M-N_t)\times N_t}}] \mathbf{U_H^*} \tag{4.13}$$

where $\mathbf{V_H}$ and $\mathbf{U_H}$ have been defined in (3.4). Choosing $\mathbf{V_O} = \mathbf{U_H}$, $\mathbf{U_O} = \mathbf{I_{N_t}}$, $\mathbf{U_I} = \mathbf{V_H}$, and $\mathbf{V_I} = \mathbf{I_{N_t}}$ leads to an overall response given by

$$\begin{aligned} \mathbf{H}_t &= \left[\mathbf{I_N} \mathbf{0_{(M-N)\times N}}\right] \mathbf{\Lambda_H} \begin{bmatrix} \mathbf{I_N} \\ \mathbf{0_{(M-N)\times N}} \end{bmatrix} \\ &= \operatorname{diag}(\lambda_1, ..., \lambda_N) \end{aligned} \tag{4.14}$$

Thus, the overall MIMO channel is transformed into a set of parallel AWGN channels. Moreover, since the SVD in (3.4) sorts the singular values in decreasing order, the signal energy has been restricted to the N_t (out of M) least lossy end-to-end eigenmodes. The capacity achieved by this choice of input-output coupling is

$$C(\mathbf{H}_t) = \sum_{n=1}^{N_t} \log\left(1 + \frac{\mathrm{SNR}}{N_t} e^{\rho_n}\right) \quad \mathtt{b/s/Hz} \tag{4.15}$$

As shown in Appendix B, this capacity could be further increased by preprocessing $\mathbf{x}$ via a diagonal power allocation matrix $\mathbf{K}$ using waterfilling. Our strategy is intuitive since we only have N_t degrees of freedom so it would be wise if we use the N_t least lossy end-to-end eigenmodes to transmit. We note that even though the effective end-to-end fiber response shows that we have used the N_t best end-to-end eigenmodes only, N_t signals were coupled to and collected from all the available physical modes at the input and output of the fiber. Nonetheless, achieving (4.15) requires, as discussed before, having CSI at transmitter and using adaptive spatial filters which is typically hard to implement. Even though we did not prove that the above strategy is capacity optimal, our simulations section will show that it appears to maximize the

capacity of the overall system.

4.4 Random Input-Output Coupling

The design of reconfigurable input-output couplers is expensive and assumes the availability of CSI at the transmitter (which is only feasible when the channel is varying slowly). More importantly, in many cases, the coupling coefficients are affected by continuous vibrations and system disturbances. Thus, full control over $\mathbf{C_I}$ and $\mathbf{C_O}$ is not always affordable. In this section, we analyze the capacity when the user does not have control over $\mathbf{C_I}$ and $\mathbf{C_O}$. This will give us better insight on the achievable capacity of MIMO MMF systems. We model the coupling coefficients as time varying random variables and impose a physically inspired distribution that respects both the fundamental energy preservation constraint and the maximum entropy principle. Even though we focus on describing the statistical model of $\mathbf{C_I}$, our discussion applies equally well for $\mathbf{C_O}$.

For an $M \times M$ square matrix $\mathbf{A}$, the energy conservation principle confirms that $\mathbf{A}$ should belong to the class of $M \times M$ unitary matrices, $\mathbb{U}(M) := \{\mathbf{U} \in \mathbb{C}^{M\times M} | \mathbf{U}^*\mathbf{U} = \mathbf{U}\mathbf{U}^* = \mathbf{I_M}\}$. It was proven in [24] that since a Haar measure exists over $\mathbb{U}(M)$, one could define a uniform distribution over $\mathbb{U}(M)$. Therefore, we choose a random setting where the input coupling matrix $\mathbf{C_I}$ has its N_t columns randomly selected from a square matrix $\mathbf{A}$ that is uniformly distributed over $\mathbb{U}(M)$. This distribution ensures that the columns of $\mathbf{C_I}$ form a complete orthonormal basis for $\mathbb{C}^{N_t}$ and gives equal probability measure for all such possible vectors. In other words, $\mathbf{C_I}$ is uniformly distributed over a Stiefel manifold $\mathbb{V}_{N_t}\left(\mathbb{C}^M\right)$. Appendix A shows how we can generate $\mathbf{C_I}$ and $\mathbf{C_O}$ from an $M \times M$ matrix with i.i.d. Gaussian entries. The ergodic capacity of the overall system can now be computed by averaging over the statistics of the input-output couplers and the statistics of the fiber response. Similarly, one could also compute the probability of an outage event by obtaining the cumulative distribution function (CDF) of the capacity, which now depends on the statistics of $\mathbf{C_I}$ and $\mathbf{C_O}$.

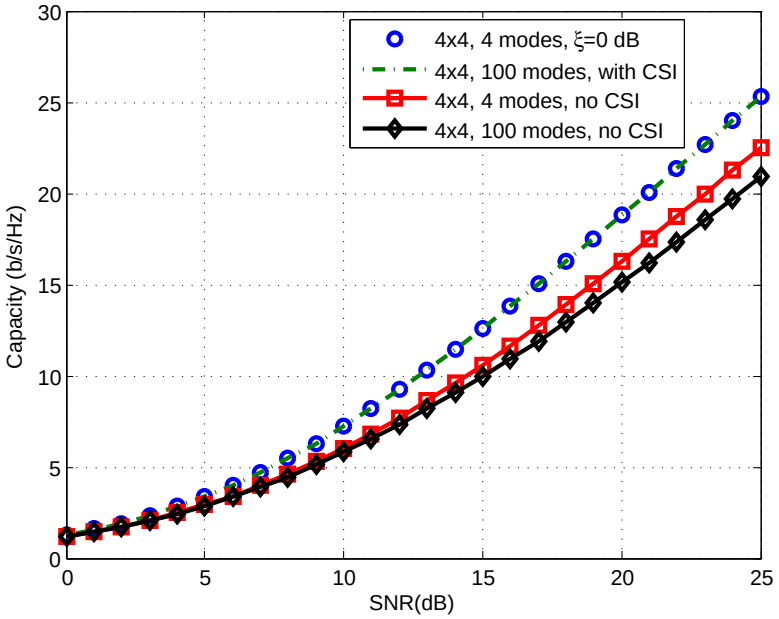


Figure 4.1: Achievable capacity of a 4×4 MIMO MMF system

4.5 Discussion

We have evaluated the capacity of both controlled and uncontrolled MIMO MMF systems. As discussed in Section 4.3, the controlled case refers to the case when CSI is available at the transmitter side and there is full control over the input-output couplers. The uncontrolled case refers to the random coupling model presented in Section 4.4. In our simulations, we numerically computed the ergodic capacity using (3.8), with $\mathbf{H}$ replaced by $\mathbf{H_t} = \mathbf{C_O H C_I}$. The average in (3.8) is taken over the statistics of the channel and the input-output couplers for the uncontrolled case. For reference, we included plots of the capacity when

- all mode dependent losses are equal to zero and there is no mode coupling ($K = 1$ and $\xi = 0$); hence the channel has unity eigenmodes. A fiber with such properties will be referred to as an *ideal fiber*.

- the fiber core radius is chosen so that only N_t modes can propagate. In this case the input and output coupling matrices are unitary matrices.

In this analysis, we consider $N_t = N_r = 4$, $K = 256$, $\xi = 4$ dB, and $M = 100$. Comparing Figures 4.1 and 3.3, we observe that the capacity of a 4×4 system over a 100-mode fiber is inferior to the intrinsic capacity of the fiber (i.e.,

when all the modes are used). This result is expected since we are using 4 out of 100 available degrees of freedom. At moderate SNR values the loss in capacity is about 6 dB. On the other hand, observe, from Figure 4.1, that the performance of an uncontrolled 4×4 system over a 100-mode fiber is close to that of a system with 4 modes. Thus, currently installed fibers could be used without significant loss in capacity. We also note that, by using the input-output coupling strategy presented in Section 4.3, performance equal to that of an ideal fiber can be achieved. This is explained by revisiting Figure 3.1 which shows the probability distribution of the end-to-end MDL values when $M = 100$. We observe that in this case we only use the best 4 eigenmodes to transmit the signal. As such, it is highly probable that these 4 (out of 100) modes will have close to zero end-to-end mode-dependent losses, and thus the performance is almost equal to that of an ideal fiber (even when ξ is large). The larger M is, the closer the capacity of a controlled fiber can get to that of an ideal fiber. Finally, one could argue that coupling a reasonable number of inputs to a fiber with hundreds of modes is advantageous since the fiber's peak power constraint is proportional to the number of modes (recall that more propagation modes means larger core radius). This means that compared to an N_t-mode fiber, a higher capacity could be achieved if we signal over an $M \gg N_t$-mode fiber since the total power budget can now be increased.

CHAPTER 5

REDUCED COMPLEXITY SEQUENCE DETECTION ALGORITHMS

5.1 Detection Algorithms

The vector Viterbi algorithm (VVA) extends the conventional Viterbi algorithm (VA) to make it operate over vector transmitted symbols [25, 26]. The emergence of multi-input multi-output (MIMO) Orthogonal Frequency Division Multiplexing (OFDM) reduced the complexity of the receiver while still achieving MIMO capacity gains [27]. This aided in the VVA being overlooked as a solution for MIMO systems. Let L, N, N_t, N_r, and $|\mathcal{A}|$ denote the channel memory, number of transmitted symbols per block, number of transmit elements, number of receive elements, and size of signal constellation, respectively. As shown in Appendix C, the computational complexity of a MIMO OFDM based detection scheme is $O\left((N_r + N_t)\, N \log(N) + N|\mathcal{A}|^{N_t}\right)$ in comparison to $O\left(N|\mathcal{A}|^{LN_t}\right)$ for the vector Viterbi algorithm. Thus, MIMO OFDM systems are computationally attractive when L is large. Nonetheless, this advantage comes at the following expenses:

1. OFDM requires the addition of a cyclic prefix which reduces the information rate of communication. If the channel is changing rapidly, the size of the block cannot be made long and hence the percentage overhead is even larger (up to 25% in some cases).

2. The transmitter complexity is increased because N_t N-point IFFTs have to be computed at the transmitter side.

3. The orthogonality between sub-carriers might be lost due to Doppler shifts and/or channel nonlinearities.

4. The peak to average power ratio (PAPR) of an OFDM system is significantly larger than that of a single carrier system.

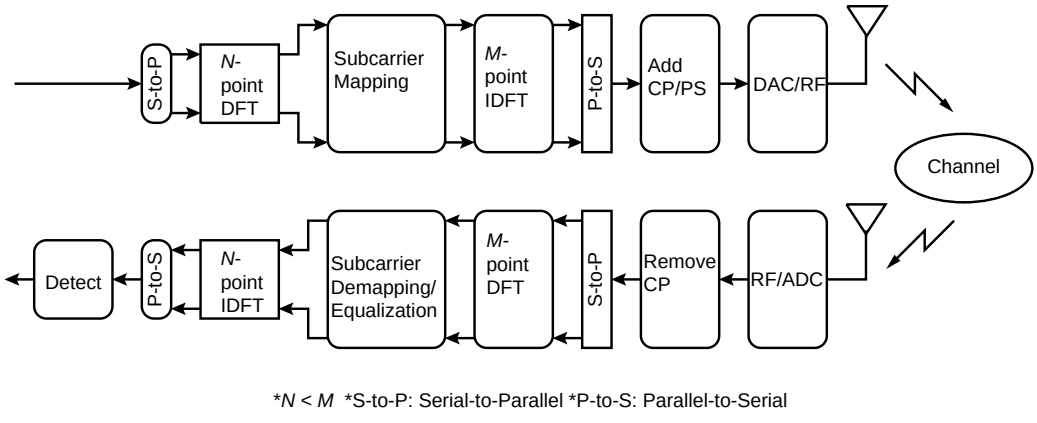


Figure 5.1: SC-FDMA modulation in LTE-A systems

In addition to the above disadvantages, some communication technologies, such as under-water acoustic systems and fiber optic systems, cannot easily make use of OFDM due to various transmitter and receiver limitations. For example, the laser sources in fiber optic systems cannot be easily modulated by arbitrarily shaped signals having high PAPR. Moreover, high PAPR signals excite fiber nonlinearities which destroy the orthogonality between sub-carriers. This is why state-of-the-art optical communication systems use simple modulation schemes such as quadrature phase shift keying. Despite the above challenges, the optical fiber communications community is interested in developing coherent optical OFDM for the next generation optical systems [21].

More importantly, single carrier systems are used in the latest wireless communication standards. For example, the Long Term Evolution Advanced (LTE-A) wireless standard does not use OFDM for the uplink channel due to its high PAPR [28]. Instead, LTE-A uplink systems use Single Carrier Frequency Division Multiple Access (SC-FDMA) modulation. Even though SC-FDMA divides the resources among users in the frequency domain, the transmission uses single carrier modulation. This process is shown in Figure 5.1. Therefore, it is natural to revisit single carrier systems and look for ways to reduce the computational complexity of sequence detection. Our sphere decoding (SD) approach for the VVA is a promising technique that reduces VVA's complexity significantly while preserving its optimality.

The remainder of this chapter is organized as follows. In Section 5.2, we review the sphere decoding algorithm. In Section 5.3, we show how SD can

be used to perform MIMO MLSD at reduced computational cost. In Section 5.4, we compare the complexity of the original naive implementation of the VVA to our scheme and show that for a 2×2 MIMO system with 16-QAM signal constellation, up to 60% of the complexity can be reduced using our approach.

5.2 Sphere Decoding

A frequency flat MIMO system is described by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v} \tag{5.1}$$

where $\mathbf{x}$ is an N_t-dimensional vector, $\mathbf{y}$ is an N_r-dimensional vector, $\mathbf{H}$ is an $N_r \times N_t$ matrix, and $\mathbf{v}$ is a zero mean complex Gaussian noise vector with a covariance matrix equal to $N_0\mathbf{I}_{\mathbf{N_r}}$. Given $\mathbf{x}$, we know that

$$f\left(\mathbf{y}|\mathbf{x}\right) \sim \mathcal{CN}\left(\mathbf{H}\mathbf{x}, N_0\mathbf{I}_{\mathbf{N_r}}\right) \tag{5.2}$$

where $f\left(\mathbf{y}|\mathbf{x}\right)$ denotes the conditional distribution of $\mathbf{y}$ given $\mathbf{x}$ and $\mathcal{CN}\left(\boldsymbol{\mu}, \mathbf{K}\right)$ represents a complex Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{K}$.

The goal of a MIMO detector is to separate out the original transmitted signal from the received signal. The detection process can be divided into two stages: a preprocessing stage, where the channel matrix $\mathbf{H}$ is learned, and a detection stage, where the transmitted symbols are recovered. Typically, the preprocessing stage is implemented at the channel variation rate, while the detection stage is implemented at the transmission rate. The optimal detector, in terms of minimizing the symbol error rate (SER), is the maximum likelihood (ML) detector and is given by

$$\begin{aligned} \hat{\mathbf{x}} &= \underset{\mathbf{x}\in\mathcal{A}^{N_t}}{\operatorname{argmax}} f\left(\mathbf{y}|\mathbf{x}\right) \\ &= \underset{\mathbf{x}\in\mathcal{A}^{N_t}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \end{aligned} \tag{5.3}$$

where $\mathcal{A}$ represents the signal constellation set. The ML detector finds the nearest neighbor to the received vector among all possible constellation points

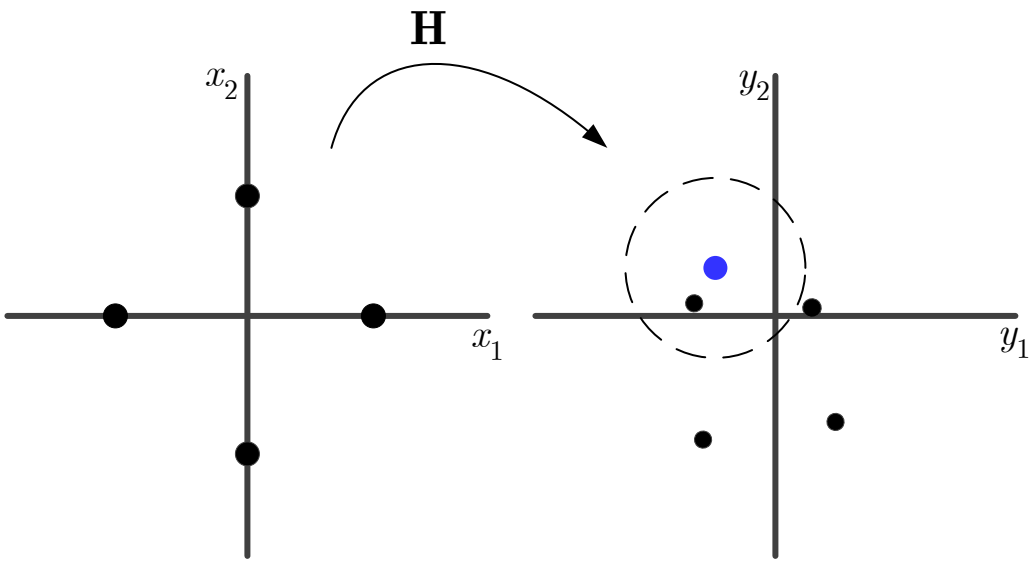


Figure 5.2: A 2×2 BPSK flat MIMO system

(lattice points) by performing an exhaustive search. Unfortunately, the computational complexity of this algorithm is exponential in N_t. Nonetheless, there exists a clever algorithm that performs ML detection at a substantially lower cost.

As shown in Figure 5.2, the basic idea of Sphere Decoding (SD) is to perform a search over the constellation points that lie within a sphere of radius r centered around the received vector [29]. This algorithm is noteworthy because its expected complexity is comparable to that of a DFE algorithm while its performance is superior to any suboptimal detection algorithm [30]. Observe that the closest lattice point to the received vector within the sphere is also the closest point among all the lattice points. This gives rise to two questions:

1. What should the sphere's radius r be? Determining the proper radius is important because, if the radius is too large, the number of points needed to search for will be large, while if the radius is too small, no lattice points may exist within the sphere.

2. Which points of the space lie within the sphere? This can be answered by testing the distance between each lattice point and the received vector. Without a simple method to determine which points lie in the sphere, SD will be no different from the original ML algorithm as every

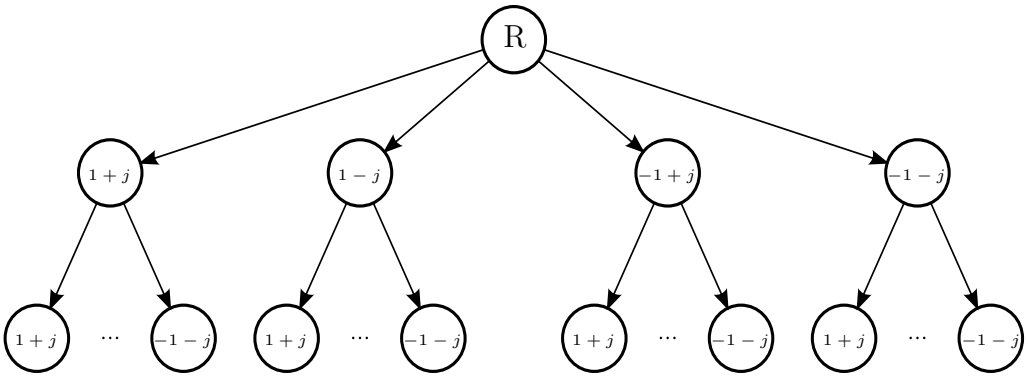


Figure 5.3: Tree representation of a 2-dimensional 4-QAM lattice

lattice point will need to be tested to determine if it lies within the sphere or not.

The SD algorithm we now present provides answers to both questions. The central idea is to represent the signal constellation via an $|\mathcal{A}|$-ary tree of depth N_t. Figure 5.3 shows the tree representation for a 4-QAM modulation, $\mathcal{A} = \{\pm 1 \pm j\}$, with 2 transmitters. The nodes at depth i correspond to instances of the $(N_t - i + 1)^{th}$ entry in $\mathbf{x}$. Thus, the tree has $|\mathcal{A}|^{N_t}$ leaves, each corresponding to an instance of $\mathbf{x}$. Assume that $N_r \geq N_t$, then by the QR decomposition $\mathbf{H} = \mathbf{Q}\left[\mathbf{R}^T \mathbf{0}_{\mathbf{N_t} \times (\mathbf{N_r} - \mathbf{N_t})}\right]^T$, where $\mathbf{Q}$ is an $N_r \times N_r$ unitary matrix and $\mathbf{R}$ is an $N_t \times N_t$ upper triangular matrix. As the norm is invariant to unitary transforms, the ML rule can be rewritten as

$$\begin{aligned} \hat{\mathbf{x}} &= \underset{\mathbf{x} \in \mathcal{A}^N}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \\ &= \underset{\mathbf{x} \in \mathcal{A}^N}{\operatorname{argmin}} \|\mathbf{Q}^*\mathbf{y} - \begin{bmatrix} \mathbf{R} \\ \mathbf{0}_{(\mathbf{N_r} - \mathbf{N_t}) \times \mathbf{N_t}} \end{bmatrix} \mathbf{x}\|^2 \\ &= \underset{\mathbf{x} \in \mathcal{A}^N}{\operatorname{argmin}} \|\tilde{\mathbf{y}} - \mathbf{R}\mathbf{x}\|^2 \end{aligned} \tag{5.4}$$

where $\tilde{\mathbf{y}}$ is a vector containing the first N_t entries of $\mathbf{Q}^*\mathbf{y}$. Due to the triangular structure of $\mathbf{R}$, the vector norm can now be rewritten as a sum of

scalar norms

$$\begin{aligned}\|\tilde{\mathbf{y}}-\mathbf{R}\mathbf{x}\|^2 &= \sum_{i=1}^{N_t}\left|\tilde{y}_i-\sum_{l=i}^{N_t}r_{i,l}x_l\right|^2 \\ &= \sum_{i=1}^{N_t}e_i\left(x_i,...,x_{N_t}\right) \\ &= e_1\left(x_1,...,x_{N_t}\right)+...+e_{N_t}\left(x_{N_t}\right) \qquad (5.5)\end{aligned}$$

where $e_i\left(x_i,...,x_{N_t}\right) = \left|\tilde{y}_i-\sum_{l=i}^{N_t}r_{i,l}x_l\right|^2$. Note that the last N_t-i+1 summands in (5.5) depend only on the last N_t-i+1 transmitted symbols and they are all non-negative. We define the partial Euclidean distance (PED) as $p_i = \sum_{j=N_t-i+1}^{N_t}e_j\left(x_j,...,x_{N_t}\right)$ for $i=1,...,N_t$. This sequence is computed recursively by traversing the tree from the root node down to a leaf node. For $i=1$, $p_1 = e_1\left(x_{N_t}\right)$ and for $i=2,...,N_t$, $p_i = p_{i-1}+e_{N_t-i+1}\left(x_{N_t-i+1},...,x_{N_t}\right)$. Notice that p_i is a non-negative and non-decreasing sequence and that $p_i \leq \|\tilde{\mathbf{y}}-\mathbf{R}\mathbf{x}\|^2$ for all i. Therefore, it is safe to drop all candidate vectors $\mathbf{x}^j$'s that end with the same i symbols if any p_i exceeds a specified radius r. This technique is referred to as tree pruning.

Tree pruning is a smart way of eliminating the lattice points that do not lie inside the sphere of radius r. However, we still do not know how to choose r. This can be done in a variety of ways. A simple scheme would set r to infinity and run a depth first search algorithm until the leftmost leaf node is reached. At this point, r is updated to become equal to the Euclidean distance of that particular instance of $\mathbf{x}$. The depth first search algorithm is then resumed and the aforementioned pruning process is applied whenever some p_i exceeds r. A leaf node is reached only if the distance between the received vector and that particular instance of $\mathbf{x}$ is less than r. In this case, the radius is updated to become equal to this new Euclidean distance and the process is continued until all leafs are either visited or pruned.

The complexity of SD is random as it depends on the quality of the channel realization which is a random variable. Moreover, the performance is a function of SNR. At high SNRs the savings are large because very few lattice points lie inside the sphere. However, marginal gains are achieved if the transformed lattice $\mathbf{Hx}$ happened to be such that all the points are close to each other. Therefore, SD still suffers from a worst case exponential

complexity. Nonetheless, this happens with low probability (especially when the SNR is high). It was shown in [31, 30] that the expected complexity of SD is polynomial in N_t for a wide range of SNRs and N_t. In fact for high SNRs, the expected complexity is cubic in N_t.

5.3 Frequency Selective Systems

A frequency selective MIMO system is described by

$$\mathbf{y}[n] = \sum_{k=0}^{L-1} \mathbf{H}[k]\mathbf{x}[n-k] + \mathbf{v}[n] \tag{5.6}$$

where $\mathbf{y}[n]$ and $\mathbf{x}[n]$ are the detected and transmitted symbol vectors, respectively. In (5.6), L represents the number of nonzero taps in $\mathbf{H}[n]$, the channel's matrix impulse response, and is given by T_d/T_s where T_d is the channel's delay spread and T_s is the sampling period. In our analysis, we assume that the channel's matrix impulse response $\mathbf{H}[n]$ is fixed for N consecutive transmissions. In a frequency selective MIMO system, each received signal is a noisy linear combination of current and previous symbols coming from all transmitted streams. The temporal mixing phenomenon is called inter-symbol interference (ISI) and the spatial mixing phenomenon is called inter-channel interference (ICI).

5.3.1 Prior Work

Sphere decoding has been recently introduced as a reduced complexity detection algorithm for single carrier MIMO frequency selective systems [32]. We define the following vectors:

$$\begin{aligned} \mathbf{y} &= [\mathbf{y}^T[1], \mathbf{y}^T[2], \ldots, \mathbf{y}^T[N+L-1]]^T \\ \mathbf{v} &= [\mathbf{v}^T[1], \mathbf{v}^T[2], \ldots, \mathbf{v}^T[N+L-1]]^T \\ \mathbf{x} &= [\mathbf{x}^T[1], \mathbf{x}^T[2], \ldots, \mathbf{x}^T[N]]^T \end{aligned} \tag{5.7}$$

Using (5.6) and (5.7), we can write the input-output relation for the dispersive channel in a matrix form as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v} \tag{5.8}$$

where $\mathbf{H}$ is an $N_r\,(N + L - 1) \times N_t N$ block Toeplitz matrix given by

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}[0] & & & & \\ \mathbf{H}[1] & \mathbf{H}[0] & & & \\ & \ddots & \ddots & & \\ & & & \mathbf{H}[L-1] & \mathbf{H}[L-2] \\ & & & & \mathbf{H}[L-1] \end{bmatrix}. \tag{5.9}$$

Given $\mathbf{x}$, we know that

$$f\,(\mathbf{y}|\mathbf{x}) \sim \mathcal{CN}\,(\mathbf{H}\mathbf{x}, N_0\mathbf{I}_{\mathbf{N+L-1}}) \tag{5.10}$$

where $\mathbf{H}\mathbf{x}$ is the mean vector and $N_0\mathbf{I}_{\mathbf{N+L-1}}$ is the covariance matrix of the complex Gaussian distribution. Therefore, the optimal detection rule is given by

$$\hat{\mathbf{x}} = \underset{\mathbf{x}\in\mathcal{A}^{N_t N}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \tag{5.11}$$

A straightforward implementation will perform an exhaustive search over all $|\mathcal{A}|^{N_t N}$ possible transmit vectors, which is stupendously expensive for large N_t or N. Observe that the problem in (5.11) is identical to the one in (5.4). Therefore, sphere decoding can be used to reduce the complexity of MLSD. However, the dimension of the equivalent frequency flat MIMO system is $N_r\,(N + L - 1) \times N_t N$ and thus, the expected complexity of this approach is at best polynomial in $N_t N$. This figure can be significantly larger than $O\left(N|\mathcal{A}|^{LN_t}\right)$ for large N or small L. Therefore, VVA seems to be more attractive for large N.

5.3.2 The Vector Viterbi Algorithm

The vector Viterbi algorithm (VVA) is the vector version of the popular Viterbi algorithm [25, 26]. The VVA uses the entire received sequence $\mathbf{y}$ to detect the transmitted sequence $\mathbf{x}$. Both $\mathbf{x}$ and $\mathbf{y}$ have been defined in (5.7).

It is convenient to define the mean vector $\boldsymbol{\mu}(\mathbf{x}) = \mathbf{H}\mathbf{x}$ and divide it into $N + L - 1$ sub-vectors where the k^{th} sub-vector is given by

$$\boldsymbol{\mu}_k(\mathbf{x}) = \sum_{l=0}^{L-1} \mathbf{H}[l]\mathbf{x}[k-l] \tag{5.12}$$

Similarly, we can divide the vector $\mathbf{y}$ into $N + L - 1$ sub-vectors where the k^{th} sub-vector is $\mathbf{y}_k = \mathbf{y}[k]$. The optimization problem in (5.11) can now be written as

$$\begin{aligned} \hat{\mathbf{x}} &= \underset{\mathbf{x}\in\mathcal{A}^{N_t N}}{\operatorname{argmin}} \|\mathbf{y} - \boldsymbol{\mu}(\mathbf{x})\|^2 \\ &= \underset{\mathbf{x}\in\mathcal{A}^{N_t N}}{\operatorname{argmin}} \sum_{k=1}^{N+L-1} \|\mathbf{y}_k - \boldsymbol{\mu}_k(\mathbf{x})\|^2 \\ &= \underset{\mathbf{x}\in\mathcal{A}^{N_t N}}{\operatorname{argmin}} \mathcal{P}_{N+L-1}(\mathbf{x}) \end{aligned} \tag{5.13}$$

where $\mathcal{P}_i(\mathbf{x}) = \sum_{k=1}^{i} \|\mathbf{y}_k - \boldsymbol{\mu}_k(\mathbf{x})\|^2$ is called the i^{th} path metric. The VVA performs the above minimization with complexity linear in N. Unlike the previously derived SD algorithm, VVA exploits the channel memory being limited to L. This is known as the Markovian property of the channel. We define the state $\mathcal{S}_k$ at time k to be

$$\mathcal{S}_k = (\mathbf{x}[k-1], \mathbf{x}[k-2], \ldots, \mathbf{x}[k-L+1]) \tag{5.14}$$

There are $|\mathcal{A}|^{N_t(L-1)}$ different states and $\mathcal{S}_k$ evolves as a Markov chain. Let $\mathcal{S}_k^j$ and $\mathbf{x}^i[k]$ represent instances of $\mathcal{S}_k$ and $\mathbf{x}[k]$, respectively. The observation $\mathbf{y}[k]$ is a function of $\mathcal{S}_k$ and $\mathbf{x}[k]$ corrupted by noise $\mathbf{v}[k]$, and thus $\mathbf{y}[k]$ forms a hidden Markov model (HMM). As shown in Figure 5.4, we can represent the state evolution in time using a trellis diagram. The paths will always converge to the zeroth state at time $N + L - 1$ since the transmission is stopped at time N. Finding the maximum-likelihood sequence estimate is equivalent to finding the shortest path through the trellis. Note that $\boldsymbol{\mu}_k(\mathbf{x})$ is only a function of $\mathbf{x}[k]$ and $\mathcal{S}_k$. We associate the following branch metric

$$\mathcal{B}\left(\mathbf{y}[k], \mathcal{S}_k^j, \mathbf{x}^i[k]\right) = \|\mathbf{y}_k - \boldsymbol{\mu}_k\left(\mathcal{S}_k^j, \mathbf{x}^i[k]\right)\|^2 \tag{5.15}$$

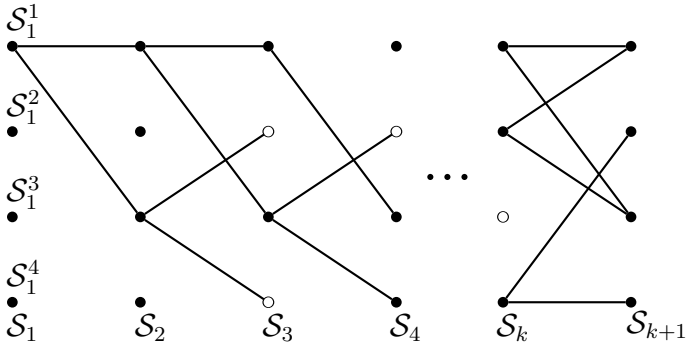


Figure 5.4: Trellis diagram for a 2×2 system with BPSK signalling and memory length of 2

with each branch emanating from $\mathcal{S}_k^j$ and terminating in $\mathcal{S}_{k+1}^i$. Note that the vectors $\mathbf{x}[k-1], \ldots, \mathbf{x}[k-L+2]$ are exactly the same for both states. Each state $\mathcal{S}_k^j$ can terminate in one of $|\mathcal{A}|^{N_t}$ states because the only new entry in $\mathcal{S}_{k+1}^i$ is $\mathbf{x}^i[k]$. The vector Viterbi algorithm uses dynamic programming to implement a breadth-first search on a trellis. The key observation is that the minimization could be solved recursively by noting that $\mathcal{P}_k = \mathcal{P}_{k-1} + \mathcal{B}\left(\mathbf{y}[k], \mathcal{S}_k^j, \mathbf{x}^i[k]\right)$. Therefore, to find the shortest path, it is sufficient to solve the following problem

$$\mathcal{P}_k^i = \min_{j \in \mathcal{F}} \mathcal{P}_{k-1}^j + \mathcal{B}\left(\mathbf{y}[k], \mathcal{S}_k^j, \mathbf{x}^i[k]\right) \tag{5.16}$$

for every $\mathcal{S}_k^i \in \mathcal{S}_k$ and $k = 1, ..., N+L-1$. In (5.16), $\mathcal{F}$ contains the indices of the states, at stage $k-1$, that are allowed to transition to $\mathcal{S}_k^i$. Observe that for $k = N+L-1$, the solution to $\min_i \mathcal{P}_{N+L-1}^i$ is the solution to the MIMO MLSD problem in (5.13).

5.3.3 Combined SD-VVA

The computational complexity of VVA is equal to the product of the number of computations required per state ($|\mathcal{A}|^{N_t}$), the number of states per stage ($|\mathcal{A}|^{N_t(L-1)}$), and the number of stages ($N+L-1$). As a result, the complexity grows linearly with the block length and exponentially with the number of transmitters and memory length. In what follows, we derive a new, lower complexity, optimal sequence detection algorithm. The aim is to break down the exponential number of computations required per state to something polynomial (often cubic) in N_t. This reduction in complexity is

Table 5.1: State and super state assignment at time k in a 2×2 system with BPSK signalling and memory length of 3

$\mathcal{S}_k$	$\{\mathbf{x}_{k-2}, \mathbf{x}_{k-1}\}$	$\underline{\mathcal{S}}_k$	$\mathcal{S}_k$	$\{\mathbf{x}_{k-2}, \mathbf{x}_{k-1}\}$	$\underline{\mathcal{S}}_k$
$\mathcal{S}^1$	$\{(0,0),(0,0)\}$	$\underline{\mathcal{S}}^1$	$\mathcal{S}^9$	$\{(1,0),(0,0)\}$	$\underline{\mathcal{S}}^1$
$\mathcal{S}^2$	$\{(0,0),(0,1)\}$	$\underline{\mathcal{S}}^2$	$\mathcal{S}^{10}$	$\{(1,0),(0,1)\}$	$\underline{\mathcal{S}}^2$
$\mathcal{S}^3$	$\{(0,0),(1,0)\}$	$\underline{\mathcal{S}}^3$	$\mathcal{S}^{11}$	$\{(1,0),(1,0)\}$	$\underline{\mathcal{S}}^3$
$\mathcal{S}^4$	$\{(0,0),(1,1)\}$	$\underline{\mathcal{S}}^4$	$\mathcal{S}^{12}$	$\{(1,0),(1,1)\}$	$\underline{\mathcal{S}}^4$
$\mathcal{S}^5$	$\{(0,1),(0,0)\}$	$\underline{\mathcal{S}}^1$	$\mathcal{S}^{13}$	$\{(1,1),(0,0)\}$	$\underline{\mathcal{S}}^1$
$\mathcal{S}^6$	$\{(0,1),(0,1)\}$	$\underline{\mathcal{S}}^2$	$\mathcal{S}^{14}$	$\{(1,1),(0,1)\}$	$\underline{\mathcal{S}}^2$
$\mathcal{S}^7$	$\{(0,1),(1,0)\}$	$\underline{\mathcal{S}}^3$	$\mathcal{S}^{15}$	$\{(1,1),(1,0)\}$	$\underline{\mathcal{S}}^3$
$\mathcal{S}^8$	$\{(0,1),(1,1)\}$	$\underline{\mathcal{S}}^4$	$\mathcal{S}^{16}$	$\{(1,1),(1,1)\}$	$\underline{\mathcal{S}}^4$

made possible by observing that the selection of the surviving path for each state can be computed via a tree based algorithm similar to the one used in sphere decoding. We define a super state $\underline{\mathcal{S}}_{k-1}$ to be the set of states $\mathcal{S}_{k-1}$ that differ only by $\mathbf{x}[k-L+1]$. Table 5.1 shows the state and super state assignments for a 2×2 system with $|\mathcal{A}| = 2$ and $L = 3$. Given this definition, we see that each super state contains $|\mathcal{A}|^{N_t}$ elements and that there are $|\mathcal{A}|^{N_t(L-2)}$ super states. Observe, from Figure 5.5, that there is a transition from each $\mathcal{S}_{k-1}^j \in \underline{\mathcal{S}}_{k-1}^l$ to one $\mathcal{S}_k^i \in \underline{\mathcal{S}}_k^m$. Furthermore, the first $L-2$ entries in $\mathcal{S}_{k-1}^j$ are identical to the last $L-2$ entries in $\mathcal{S}_k^i$. Thus, the following holds:

$$\begin{aligned} \mathcal{P}_k^i &= \min_{j \in \mathcal{F}} \mathcal{P}_{k-1}^j + \mathcal{B}\left(\mathbf{y}[k], \mathcal{S}_k^j, \mathbf{x}^i[k]\right) \\ &= \min_{j \in \mathcal{F}} \mathcal{P}_{k-1}^j + \|\mathbf{y}_k - \boldsymbol{\mu}_k\left(\mathcal{S}_k^j, \mathbf{x}^i[k]\right)\|^2 \\ &= \min_{j \in \mathcal{F}} \mathcal{P}_{k-1}^j + \|\mathbf{z}_k - \mathbf{G}\mathbf{x}^j\|^2 \end{aligned} \tag{5.17}$$

where $\mathbf{z}_k = \mathbf{y}_k - \sum_{l=0}^{L-2} \mathbf{H}[l]\mathbf{x}^i[k-l]$, $\mathbf{G} = \mathbf{H}[L-1]$, and $\mathbf{x}^j = \mathbf{x}^j[k-L+1]$. Had the term $\mathcal{P}_{k-1}^j$ not existed in (5.17), this minimization would have resembled the standard frequency flat MIMO ML detection problem in (5.4). In this case, the complexity can be reduced by solving for the surviving branch via a sphere decoding approach as detailed in Section 5.2. However, in our case every path is biased by a different quantity $\mathcal{P}_{k-1}^j$ that is only determined when we traverse the tree from the root node down to a leaf node. This problem is clearly more complicated than the frequency flat MIMO ML one and cannot be solved using the standard sphere decoding algorithm. Nonetheless, we

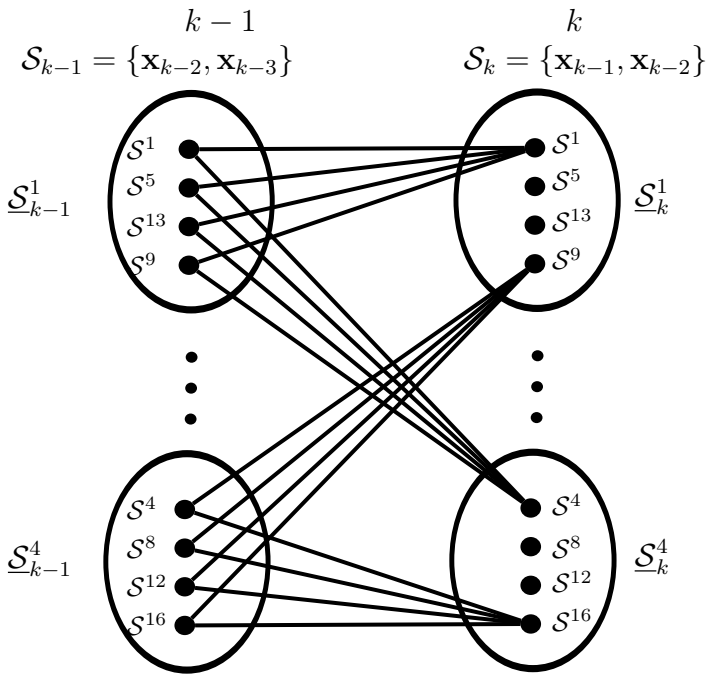


Figure 5.5: Trellis super state grouping in a 2×2 system with BPSK signalling and memory length of 3

suggest modifying the tree representation by appending the path metrics to the leaf nodes. This is depicted in Figure 5.6 where the tree has been extended to incorporate the effect of $\mathcal{P}_k^j$'s. Thus, we can now run a depth first search SD algorithm using the tree in Figure 5.6 to solve for the i^{th} state's path metric at k. The first time the radius is updated it will include both the branch and path metrics of the state that corresponds to the leftmost branch in the tree. An approach that would lead to larger computational savings would first rearrange the branches of the tree in Figure 5.6 so that the leftmost leaf node corresponds to the instance that has the smallest path metric and the rightmost leaf node corresponds to the instance that has the largest path metric. This leads to an improved performance as the radius is likely chosen to be small. However, this approach necessitates a sorted list of path metrics which complicates the implementation of the algorithm and increases the number of comparisons needed.

The performance of the combined SD-VVA approach depends on how large the path metrics are relative to the branch metrics. Little savings can be achieved if the $\mathcal{P}_k^j$'s are much larger than the weights (e_i^j's) shown in Figure 5.6. In this case, almost all leaf nodes would have to be visited. Therefore, large savings can be achieved if the trellis is shortened from N to $5L$. This

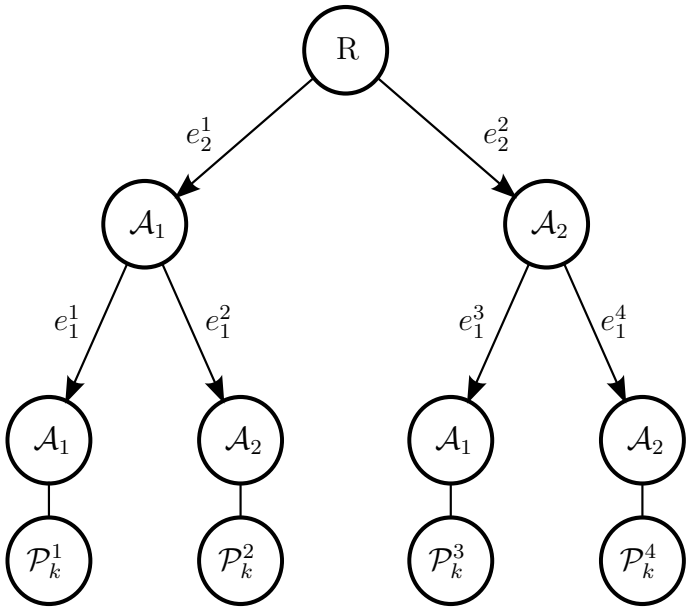


Figure 5.6: Tree representation for a 2×2 BPSK system at stage $k+1$

ensures that the path metrics do not accumulate and are still comparable to all other weights and thus, pruning will be more effective. However, this technique is sub-optimal. In addition, we will show in Section 5.4 that computational gains are large even for $N = 10^3$, and therefore this technique is not necessary.

5.4 Complexity Analysis & Results

The per state computational complexity of the VVA is given by

$$\begin{aligned} N_{add} &= 3N_t|\mathcal{A}|^{N_t} \\ N_{mult} &= N_t|\mathcal{A}|^{N_t} \\ N_{cmp} &= |\mathcal{A}|^{N_t} - 1 \end{aligned} \tag{5.18}$$

where N_{add} and N_{mult} represent the number of real additions and complex multiplications, respectively. For $N_t \geq 2$, the per state computational complexity of full tree search, without tree pruning, is given by

$$\begin{aligned} N_{add} &\approx 4|\mathcal{A}|^{N_t} \\ N_{mult} &\approx |\mathcal{A}|^{N_t} \end{aligned} \tag{5.19}$$

The exact expressions and derivation of (5.18) and (5.19) can be found in Appendix D. Observe that we can achieve computational gains even if we perform a naive tree search without pruning. For example, in a 2×2 MIMO system with 16-QAM signal constellation, a full tree search algorithm saves 33% of the real additions and 46% of the complex multiplications when compared to VVA. More importantly, the full tree search algorithm has a fixed (non-random) computational complexity. However, we can save more computions by using the combined SD-VVA algorithm described in the previous section. Unlike VVA or full tree search, the combined SD-VVA algorithm has a random complexity that depends on the SNR and channel statistics. In order to quantify the average computational savings, we computed, via simulations, the average complexity of the combined VVA-SD algorithm and compared it to VVA for various settings. In our experiments, we chose a 2×2 MIMO system with 16-QAM signal constellation, $L = 3$, and $N = 10^3$. The results are summarized in Table 5.2. As discussed in Section 5.2, the

Table 5.2: SD-VVA vs. VVA

constellation	N_{add}	N_{mult}
16-QAM (5 dB)	43%	49%
16-QAM (10 dB)	53%	54%
16-QAM (15 dB)	64%	62%

algorithm's performance improves with increasing SNR. For 16-QAM signal constellations, the computational complexity of VVA is reduced by 50% when the SNR is 10 dB and by 60% when the SNR is 15 dB.

5.5 Conclusion

Even though our approach provides substantial complexity gains, the number of states is still exponential in N_t and L. Therefore, for large N_t or L, performing exact MLSD might be expensive despite the reductions shown in the previous section. In this case, we can use a variety of techniques to further reduce the complexity. This, however, sacrifices optimality. For example, when L is large, a linear channel shortening filter can be used to reshape the channel's impulse response such that most of the signal's energy is concentrated in the first few L' taps, where $L' < L$. If N_t is large, we can

also save by keeping the best K states (states with the least path metrics) at each stage instead of keeping track of all $|\mathcal{A}|^{N_t(L-1)}$ states. The choice of K is determined by a reasonable performance-complexity tradeoff assessment.

The combined SD-VVA algorithm reduces the complexity of VVA while preserving its optimality. This algorithm is attractive whenever performance is not to be compromised. In LTE-A systems, the uplink can afford running expensive detection algorithms because the computations take place at the base station. Moreover, the combined SD-VVA algorithm can be easily modified to output likelihoods (soft decisions) that can be fed to the channel decoder. Future work could look at the architectural implementation and design of the combined SD-VVA algorithm.

CHAPTER 6

CONCLUSION

MIMO communications over optical fibers is an attractive solution to the ever increasing demand for Internet bandwidth. In the first part of this thesis, we presented a propagation model that takes input-output coupling into account for MIMO MMF systems. A coupling strategy was suggested and simulations showed that the capacity of an $N_t \times N_t$ MIMO system over a fiber with $M \gg N_t$ modes can approach the capacity of an ideal fiber with N_t modes. A random input-output coupling model was used to describe the behavior of the system when the design of the input-output couplers is not available. The results proved that, under random coupling, the capacity of an $N_t \times N_t$ MIMO system over a fiber with $M \gg N_t$ modes is almost equal to that of an N_t-mode fiber.

High speed multi-input multi-output (MIMO) communication systems suffer from inter-channel and inter-symbol interference (ICI and ISI). The vector Viterbi algorithm (VVA) is a maximum likelihood sequence detection (MLSD) algorithm for MIMO frequency selective channels. MLSD algorithms are attractive because they minimize the probability of sequence detection error. However, they suffer from high computational complexity. In the second part of this thesis, we showed how a sphere decoding-like algorithm can be used to reduce the complexity of VVA while preserving its optimality. For a 2×2 MIMO system with 16-QAM signal constellation, our algorithm appears to reduce the VVA complexity by 50% at an SNR of 10 dB and by 60% at an SNR of 15 dB.

APPENDIX A

RANDOM UNITARY MATRICES

This appendix provides a brief introduction to random matrix theory with particular emphasis on random unitary matrices. The results we show in the following sections are used in Chapters 3 and 4.

A.1 Random Matrices

An $M \times N$ random matrix $\mathbf{A}$ is described by the joint probability distribution function (pdf) of its entries

$$f(\mathbf{A}) = f(a_{i,j}; i = 1, ..., M; j = 1, ..., N) \tag{A.1}$$

A popular example of a family of random matrices is the Ginibre ensemble [24]. In this case, the entries of an $M \times M$ square matrix $\mathbf{A}$ are independent and identically distributed (i.i.d.) standard complex Gaussian random variables

$$f(\mathbf{A}) = \frac{1}{\pi^{M^2}} e^{-\sum_{j,k=1}^{M} |a_{jk}|^2} = \frac{1}{\pi^{M^2}} e^{-\operatorname{tr}(\mathbf{A}^*\mathbf{A})} \tag{A.2}$$

We define $\mathbb{U}(M) := \{\mathbf{U} \in \mathbb{C}^{M \times M} | \mathbf{U}^*\mathbf{U} = \mathbf{U}\mathbf{U}^* = \mathbf{I_M}\}$ to be the space of $M \times M$ unitary matrices. If the pdf of a random matrix is invariant to left (right) multiplication by any $M \times M$ ($N \times N$) deterministic unitary matrix, it is called left (right) rotationally invariant. In other words, if $\mathbf{A}$ is a random matrix, $\mathbf{A}$ is left rotationally invariant if

$$f_{\mathbf{UA}}(\mathbf{UA}) = f(\mathbf{A}) \tag{A.3}$$

where $\mathbf{U} \in \mathbb{U}(M)$, and $\mathbf{A}$ is right rotationally invariant if

$$f_{\mathbf{AV}}(\mathbf{AV}) = f(\mathbf{A}) \tag{A.4}$$

where $\mathbf{V} \in \mathbb{U}(N)$. A matrix that is left and right rotationally invariant is called isotropically invariant.

Lemma A.1.1 *The Ginibre ensemble is a family of isotropically invariant random matrices [33].*

Proof 6 *We will show that* $f(\mathbf{UA}) = f(\mathbf{A})$ *and the Jacobian of the mapping* $\mathbf{A} \to \mathbf{UA}$ *(a transformation in* $\mathbb{C}^{M^2}$*) is unity. From Equation (A.2), we have the following:*

$$\begin{aligned} f(\mathbf{UA}) &= \frac{1}{\pi^{M^2}} e^{-\operatorname{tr}((\mathbf{UA})^*(\mathbf{UA}))} \\ &= \frac{1}{\pi^{M^2}} e^{-\operatorname{tr}(\mathbf{AU}^*\mathbf{UA})} \\ &= f(\mathbf{A}) \end{aligned} \tag{A.5}$$

Furthermore, if we look at $\mathbf{A}$ *as a vector in* $\mathbb{C}^{M^2}$*, we can decompose the action of* $\mathbf{U}$ *as the direct sum* $\mathbf{U}' = \mathbf{U} \bigoplus ... \bigoplus \mathbf{U} \in \mathbb{U}(M^2)$*. Hence,* $|\det(\mathbf{U}')| = |\det(\mathbf{U}) ... \det(\mathbf{U})| = 1$*. Therefore,* $f_{\mathbf{UA}}(\mathbf{UA}) = f(\mathbf{A})$*. Right rotational invariance is shown in a similar way.*

A.2 Random Unitary Matrices

An $M \times M$ random unitary matrix $\mathbf{A}$ is one for which its pdf is given by

$$f(\mathbf{A}) = g(\mathbf{A})\,\delta(\mathbf{AA}^* - \mathbf{I_M}) \tag{A.6}$$

where $g(\mathbf{A})$ is some distribution function defined over $\mathbb{U}(M)$. Observer that $P\{\mathbf{A} \notin \mathbb{U}(M)\} = 0$. It turns out that $\mathbb{U}(M)$ forms a compact topological group, and thus a unique uniform measure (up to a scalar multiplication), called Haar measure, can be defined over $\mathbb{U}(M)$ [24, 33, 34]. In this case, $g(\mathbf{A})$ is independent of $\mathbf{A}$ and it is easy to show that $f(\mathbf{A})$ is given by

$$f(\mathbf{A}) = \frac{\prod_{i=1}^{M} \Gamma(i)}{\pi^{M(M+1)/2}} \delta(\mathbf{AA}^* - \mathbf{I_M}) \tag{A.7}$$

We provide two methods for generating uniformly distributed random unitary matrices. The first method is based on the Gram-Schmidt orthogonalization procedure. Suppose we have a sample $\mathbf{Z}$ form a Ginibre ensemble. We

can perform the Gram-Schmidt orthogonalization procedure on the column vectors $\mathbf{Z}_i$ of $\mathbf{Z}$ to obtain the column vectors $\mathbf{A}_i$ of $\mathbf{A}$, a random unitary matrix [34]. In other words, given $\mathbf{Z}$, $\mathbf{A}_i$ is given by

$$\mathbf{A}_i = \frac{\mathbf{Z}_i - \sum_{k=1}^{i-1} (\mathbf{Z}_i, \mathbf{A}_i)\, \mathbf{A}_k}{\|\mathbf{Z}_i - \sum_{k=1}^{i-1} (\mathbf{Z}_i, \mathbf{A}_i)\, \mathbf{A}_k\|} \tag{A.8}$$

where $(\mathbf{a}, \mathbf{b})$ denotes the standard inner product between the vectors $\mathbf{a}$ and $\mathbf{b}$. It is easy to verify that $\mathbf{A}$ is unitary for any sample $\mathbf{Z}$. Moreover, it can be shown that $\mathbf{A}$ has a uniform distribution over $\mathbb{U}(M)$ [34].

Lemma A.2.1 *The pdf of* $\mathbf{A}$ *is isotropically invariant [34].*

Proof 7 *We prove left rotational invariance. Right rotational invariance is shown in a similar way. The* i^{th} *column of* $\mathbf{UA}$ *is exactly* $\mathbf{UA}_i$ *and it is equal to*

$$\begin{aligned} \mathbf{UA}_i &= \frac{\mathbf{UZ}_i - \sum_{k=1}^{i-1} (\mathbf{Z}_i, \mathbf{Z}_i)\, \mathbf{UA}_k}{\|\mathbf{Z}_i - \sum_{k=1}^{i-1} (\mathbf{Z}_i, \mathbf{A}_i)\, \mathbf{A}_k\|} \\ &= \frac{\mathbf{UZ}_i - \sum_{k=1}^{i-1} (\mathbf{UZ}_i, \mathbf{UA}_i)\, \mathbf{UA}_k}{\|\mathbf{UZ}_i - \sum_{k=1}^{i-1} (\mathbf{UZ}_i, \mathbf{UA}_i)\, \mathbf{UA}_k\|} \end{aligned} \tag{A.9}$$

where the last equality holds because the standard inner product is invariant to unitary transformations. As shown in Lemma A.1.1, $\mathbf{UZ}$ *has the same distribution as* $\mathbf{Z}$*. Therefore, by induction,* $\mathbf{UA}_i$ *has the same distribution as* $\mathbf{A}_i$ *for all* i*. This proves the left rotational invariance property.*

The second method is based on the QR decomposition procedure [24]. In this case, $\mathbf{A}$ is constructed as follows:

1. Generate an $M \times M$ matrix $\mathbf{Z}$ with i.i.d. complex Gaussian entries.
2. Obtain the QR decomposition of $\mathbf{Z}$; $\mathbf{Z} = \mathbf{QR}$.
3. Form the following diagonal matrix:

$$\mathbf{\Lambda} = \begin{pmatrix} \frac{r_{11}}{|r_{11}|} & & & \\ & \frac{r_{22}}{|r_{22}|} & & \\ & & \ddots & \\ & & & \frac{r_{MM}}{|r_{MM}|} \end{pmatrix} \tag{A.10}$$

where $\{r_{ii}\}_{i=1}^{M}$ are the diagonal entries of $\mathbf{R}$.

4. Let $\mathbf{A} = \mathbf{\Lambda Q}$.

In the above construction, $\mathbf{A}$ is obviously unitary since $\mathbf{Q}$ is unitary. Furthermore, it can be shown that $\mathbf{A}$ has a uniform distribution over $\mathbb{U}(M)$.

A.3 Input-Output Coupling Matrices

To generate the input coupling matrix, the following method is used:

1. Generate an $M \times M$ unitary matrix $\mathbf{A}$ (as described above).

2. Choose N_t columns randomly from $\mathbf{A}$ to form $\mathbf{C_I}$.

A similar approach can be taken to generate $\mathbf{C_O}$. In this case, N_r columns are selected randomly from $\mathbf{A}$ to represent the rows of $\mathbf{C_O}$.

APPENDIX B

A BIT OF INFORMATION THEORY

This appendix discusses some fundamental results in information theory. We derive capacity expressions for the channel models used in Chapters 3 and 5. For an in-depth treatment of information theory we direct the reader to standard texts such as [35] and [18].

B.1 Entropy and Mutual Information

The entropy of a discrete random variable X that can take one of K possible values and has a probability mass function (pmf) $P(X)$ is defined as

$$\begin{aligned} H(X) &= \mathbb{E}\left[-\log P\left(X\right)\right] \\ &= -\sum_{i=1}^{K} P\left(x_i\right) \log P\left(x_i\right) \end{aligned} \tag{B.1}$$

where $-\log P\left(X\right)$ is defined as the information content of X. The base of the logarithm in (B.1) is usually 2 in which case the entropy is given in bits. For the discrete case, $H\left(X\right)$ is always non-negative and less than or equal to $\log K$. It is equal to zero if X is a deterministic random variable and equal to $\log K$ if X is uniformly distributed. The entropy can be interpreted as a measure of the amount of uncertainty in X.

Similarly, the joint entropy between two discrete random variables X and Y with a joint pmf $P\left(X,Y\right)$ is defined as

$$\begin{aligned} H(X,Y) &= \mathbb{E}\left[-\log P\left(X,Y\right)\right] \\ &= -\sum_{i,j=1}^{K_1,K_2} P\left(x_i,y_j\right) \log P\left(x_i,y_j\right) \end{aligned} \tag{B.2}$$

The conditional entropy of X given Y is defined as

$$H(X|Y) = -\sum_{i,j=1}^{K_1,K_2} P(x_i, y_j) \log P(x_i|y_j) \tag{B.3}$$

The conditional entropy can be interpreted as the average amount of uncertainty left in X after observing Y. If X and Y are independent $H(X|Y) = H(X)$. Using the above definitions, one can easily show the following:

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y) \tag{B.4}$$

Moreover, one can show that conditioning always reduces entropy: $H(X|Y) \leq H(X)$. The mutual information between two discrete random variables X and Y is defined as

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \tag{B.5}$$

The mutual information is symmetric with respect to X and Y and is always non-negative. It is equal to zero if and only if X and Y are independent.

For a continuous random variable x with a probability density function (pdf) $f_x(x)$, the differential entropy of x is defined as

$$h(x) = -\int_{-\infty}^{\infty} f_x(u) \log f_x(u)\, du \tag{B.6}$$

Similarly, the conditional differential entropy between x and y is defined as

$$h(x|y) = -\int_{-\infty}^{\infty} f_{x,y}(u,v) \log f_{x|y}(u|v)\, du dv \tag{B.7}$$

and the continuous mutual information is defined as

$$I(x;y) = h(x) - h(x|y) = h(y) - h(y|x) \tag{B.8}$$

whenever the quantities on the right-hand side are finite. Unlike the discrete case, the differential entropy can be negative. For example, if x is uniformly distributed between 0 and a, its differential entropy is equal to $\log a$, which is negative for $a < 1$. However, the continuous mutual information retains the same fundamental meaning as in the discrete case.

B.2 Channel Capacity

To convey information from one end to the other of a channel, a codebook $\mathcal{C} := \{\mathbf{c}_1, ..., \mathbf{c}_{|\mathcal{C}|}\}$ of block length N and size $|\mathcal{C}|$ is used. Here, $\mathbf{c}_k$ denotes the k^{th} codeword. Since there are $|\mathcal{C}|$ codewords, we are conveying $\log_2 |\mathcal{C}|$ bits per N channel uses. Therefore, the rate of communications is

$$R = \frac{\log_2 |\mathcal{C}|}{N} \tag{B.9}$$

The task of the decoder is to recover the index of the transmitted codeword given a block of N noisy symbols. Hence, an error occurs whenever the decoded codeword is not equal to the transmitted one. Clearly, the probability of error is a function of the codebook and the channel transition probability i.e., $P_e\left(\mathcal{C}, f_{y|x}\right)$. A rate R is said to be achievable if for every $\epsilon > 0$, $\exists N_\epsilon \in \mathbb{N}$ such that for all $N > N_\epsilon$, $P_e\left(\mathcal{C}, f_{y|x}\right) < \epsilon$. By definition, the capacity of a channel C, is the maximum achievable rate.

Theorem B.2.1 *For a discrete memoryless channel with input random variable X and output random variable Y. The capacity of the channel is*

$$C = \max_{P(X)} I\left(X;Y\right) \tag{B.10}$$

Proof 8 *See [13].*

Similarly, the capacity of continuous memoryless channels is given by $C = \max_{f_x(x)} I\left(x;y\right)$. However, this quantity could very well be infinite for many pdfs $f_x\left(x\right)$ with unbounded second moment. In digital communications, the average transmitted power is typically limited to some maximum value P. This is why the set of distributions that are valid for the mutual information maximization problem consists of those that have a second moment less than or equal to P. Therefore, the capacity of a continuous memoryless channel is given by

$$C = \max_{f_x(x):\ \ \mathbb{E}[x^2] \leq P} I\left(x;y\right) \tag{B.11}$$

B.2.1 AWGN Channels

We now apply the above results to the complex baseband additive white Gaussian noise (AWGN) channel. The input-output relationship in an AWGN channel is given by

$$y = x + v \tag{B.12}$$

where v is a zero mean complex Gaussian random variable with a variance equal to N_0. In this case, one can verify that $h(v) = \log(\pi e N_0)$. The capacity of the channel is thus given by

$$C = \max_{f_x(x):\ \mathbb{E}[x^2] \leq P} h(y) - \log(\pi e N_0) \tag{B.13}$$

Under the second moment constraint on x, the random variable y is constrained to have a total power of $P + N_0$. It was shown in [14] that a circular symmetric complex Gaussian random variable has the largest differential entropy among all complex random variable (under a second moment constraint). Therefore, the capacity of the complex baseband AWGN channel is

$$C = \log(\pi e(P + N_0)) - \log(\pi e N_0) = \log\left(1 + \frac{P}{N_0}\right) \tag{B.14}$$

and the distribution of x that achieves this capacity is a zero mean complex Gaussian with a variance equal to P.

B.2.2 Frequency Selective Channels

A linear time-invariant L-tap inter-symbol interference (ISI) channel is given by

$$y[n] = \sum_{l=0}^{L-1} h[l]x[n-l] + v[n] \tag{B.15}$$

with a second moment constraint on each $x[n]$. As explained in Appendix C, Orthogonal Frequency Division Multiplexing (OFDM) can be used to transform this frequency selective channel into the following set of parallel frequency flat channels:

$$\tilde{y}_n = \tilde{h}_n \tilde{x}_n + \tilde{v}_n \qquad\qquad n = 1, ..., N \tag{B.16}$$

where $\tilde{\mathbf{h}} = \left(\tilde{h}_1, ..., \tilde{h}_N\right)$ is the N-point discrete Fourier transform (DFT) of $h[n]$ scaled by $\sqrt{N}$, $\tilde{v}_n$ is a complex Gaussian symbol, and N is the number of sub-carriers per OFDM block . The power constraint on $x[n]$ translates to $\mathbb{E}\left[\|\tilde{\mathbf{x}}\|^2\right] \leq NP$, where $\tilde{\mathbf{x}} = (\tilde{x}_1, ..., \tilde{x}_N)$. The channel in (C.9) is known as a parallel AWGN channel. If the transmitter knows $h[n]$, it can assign a power budget P_n to the n^{th} sub-carrier so that the total capacity is maximized. Thus, the following problem has to be solved:

$$C = \frac{1}{N} \max_{P_1,...,P_N} \sum_{n=1}^{N} \log\left(1 + \frac{P_n |\tilde{h}_n|^2}{N_0}\right) \tag{B.17}$$

subject to $\sum_{n=1}^{N} P_n = NP$ and $P_n \geq 0$ for $n = 1, ..., N$. The solution to this problem is known as "waterfilling power allocation" [8]. Let $x^+ = \max(x, 0)$, the optimal power allocation is given by

$$P_n^* = \left(\frac{1}{\lambda} - \frac{N_0}{|\tilde{h}_n|^2}\right)^+ \tag{B.18}$$

where λ is chosen such that

$$\sum_{n=1}^{N} \left(\frac{1}{\lambda} - \frac{N_0}{|\tilde{h}_n|^2}\right)^+ = NP \tag{B.19}$$

A simple iterative numerical algorithm is used to solve (B.18) and obtain P_n^* for $n = 1, ..., N$. The solution has the name waterfilling because sub-channels with good conditions (large $|\tilde{h}_n|^2$) receive more power relative to sub-channels with bad conditions. As N goes to infinity, one can generalize the above discussion to arrive to the following conclusion:

$$C = \frac{1}{W} \int_0^W \log\left(1 + \frac{P^*(f)\,|\tilde{h}(f)\,|^2}{N_0}\right) df \tag{B.20}$$

where $\tilde{h}(f)$ is the discrete-time Fourier transform (DTFT) of $h[n]$ and W represents the bandwidth used in Hz. The optimal power allocation in this case is given by

$$P^*(f) = \left(\frac{1}{\lambda} - \frac{N_0}{|\tilde{h}(f)\,|^2}\right)^+ \tag{B.21}$$

where λ is chosen such that

$$\int_0^W \left(\frac{1}{\lambda} - \frac{N_0}{|\tilde{h}\,(f)\,|^2}\right)^+ df = P \tag{B.22}$$

When the transmitter does not know $h[n]$, it allocates power uniformly across all sub-carriers. Therefore, the capacity is given by

$$C = \frac{1}{W}\int_0^W \log\left(1 + |\tilde{h}\,(f)\,|^2\,\mathrm{SNR}\right) df \tag{B.23}$$

where $\mathrm{SNR} = P/N_0$.

B.2.3 MIMO Frequency Flat Channel

A time-invariant MIMO frequency flat channel is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v} \tag{B.24}$$

where $\mathbf{x} \in \mathbb{C}^{N_t}$ and $\mathbf{y} \in \mathbb{C}^{N_r}$ represent the transmitted and received vectors, respectively, and $\mathbf{v}$ is a complex Gaussian vector with a covariance matrix equal to $N_0\mathbf{I}_{\mathbf{N_r}}$. Without loss of generality, let $N_r \leq N_t$. To begin with, the channel matrix $\mathbf{H}$ is assumed to be known at the transmitter and receiver. We provide an intuitive proof for the channel capacity. A rigorous proof can be found in [14]. By the singular value decomposition (SVD), $\mathbf{H}$ can be decomposed as

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^* \tag{B.25}$$

where $\mathbf{U} \in \mathbb{C}^{N_r \times N_r}$ and $\mathbf{V}^* \in \mathbb{C}^{N_t \times N_t}$ are both unitary matrices and $\mathbf{\Lambda}$ is a rectangular matrix containing the singular values of $\mathbf{H}$. If the transmitter performs the following preprocessing $\tilde{\mathbf{x}} = \mathbf{V}\mathbf{x}$ and the receiver performs the following postprocessing $\tilde{\mathbf{y}} = \mathbf{U}^*\mathbf{y}$, then the input-output relationship becomes

$$\tilde{\mathbf{y}} = \mathbf{\Lambda}\tilde{\mathbf{x}} + \tilde{\mathbf{v}} \tag{B.26}$$

where $\tilde{\mathbf{v}} = \mathbf{U}^*\mathbf{v}$ is again a complex Gaussian vector with a covariance matrix equal to $N_0\mathbf{I}_{\mathbf{N_r}}$. Therefore, $\tilde{y}_i = \lambda_i\tilde{x}_i + \tilde{w}_i$ for $i = 1, .., N_t$. Thus, we have transformed the MIMO AWGN channel into a set of parallel AWGN channels and we can now readily use the results of the previous section to determine

the channel capacity. Note that one, or many, of the λ_i's might be equal to zero. As in the previous section, the capacity of the MIMO AWGN channel is given by

$$C = \sum_{k=1}^{N_t} \log\left(1 + \frac{P_k^* \lambda_k^2}{N_0}\right) \tag{B.27}$$

where the optimal power allocation is given by waterfilling

$$P_k^* = \left(\mu - \frac{N_0}{\lambda_k^2}\right)^+ \tag{B.28}$$

with μ chosen such that the total power constraint $\sum_k P_k^* = P$ is satisfied.

In general, one can show that the capacity is given by

$$C = \log\det\left(\mathbf{I}_{N_r} + \frac{1}{N_0}\mathbf{H}\mathbf{K}\mathbf{H}^*\right) \tag{B.29}$$

where $\mathbf{K} = \mathbf{Q}\,\mathrm{diag}\,(P_1, ..., P_{N_t})\,\mathbf{Q}^*$ is the covariance of the input vector $\mathbf{x}$. If the transmitter knows $\mathbf{H}$, it selects $\mathbf{Q} = \mathbf{U}^*$ and allocates power using the waterfilling procedure. However, if $\mathbf{H}$ is not known at the transmitter, then $\mathbf{Q}$ is chosen to be equal to the identity matrix and the powers are divided equally across all N_t transmit antennas. In this case, it is easy to verify that the capacity is given by

$$\begin{aligned} C &= \log\det\left(\mathbf{I}_{N_r} + \frac{1}{N_0}\mathbf{H}\,\mathrm{diag}\left(\frac{P}{N_t}, ..., \frac{P}{N_t}\right)\mathbf{H}^*\right) \\ &= \sum_{k=1}^{N_t} \log\left(1 + \frac{\mathrm{SNR}}{N_t}\lambda_k^2\right) \end{aligned} \tag{B.30}$$

B.2.4 MIMO Frequency Selective Channels

A frequency selective MIMO channel is described by

$$\mathbf{y}[n] = \sum_{k=0}^{L-1} \mathbf{H}[k]\mathbf{x}[n-k] + \mathbf{v}[n] \tag{B.31}$$

where $\mathbf{y}[n]$ and $\mathbf{x}[n]$ are the detected and transmitted symbol vectors, respectively. The channel's matrix impulse response $\mathbf{H}[n]$ is assumed to have L non-zero taps. As shown in Appendix C, if OFDM is used, the frequency

selective MIMO channel can be transformed into a set of frequency flat MIMO channels

$$\tilde{\mathbf{y}}_n = \tilde{\mathbf{H}}_n \tilde{\mathbf{x}}_n + \tilde{\mathbf{v}}_n \qquad n = 1, ..., N \tag{B.32}$$

where $\tilde{\mathbf{h}}_{i,j} = \left(\tilde{\mathbf{H}}_1^{i,j}, ..., \tilde{\mathbf{H}}_N^{i,j}\right)^T$ is the N-point discrete Fourier transform (DFT) of $h_{i,j}[n]$ scaled by $\sqrt{N}$. Here, $h_{i,j}[n]$ represents the impulse response from the j^{th} transmitter to the i^{th} receiver and $\tilde{\mathbf{H}}_l^{i,j}$ is the $(i,j)^{th}$ entry of $\tilde{\mathbf{H}}_l$. The power constraint on $\mathbf{x}[n]$ translates to $\mathbb{E}\left[\|\tilde{\mathbf{x}}\|^2\right] \leq NP$, where $\tilde{\mathbf{x}} = \left(\tilde{\mathbf{x}}_1^T, ..., \tilde{\mathbf{x}}_N^T\right)^T$. Using the result of the previous section, the capacity of a MIMO frequency selective channel is given by

$$C = \frac{1}{N} \sum_{n=1}^{N} \log\det\left(\mathbf{I}_{N_r} + \frac{1}{N_0} \tilde{\mathbf{H}}_n \mathbf{K}_n \tilde{\mathbf{H}}_n^*\right) \tag{B.33}$$

where $\mathbf{K}_n = \mathbf{Q}_n \operatorname{diag}\left(P_{1,n}, ..., P_{N_t,n}\right) \mathbf{Q}_n^*$ represents the covariance matrix of the vector $\tilde{\mathbf{x}}_n$ transmitted along the n^{th} sub-carrier. By the SVD, $\tilde{\mathbf{H}}_n = \mathbf{U}_n \Lambda_n \mathbf{V}_n^*$. If the transmitter knows $\mathbf{H}[n]$, it can set $\mathbf{Q}_n = \mathbf{V}_n$ and allocate power using waterfilling so that it maximizes the capacity. Therefore, the capacity of frequency selective MIMO channels is given by

$$\begin{aligned} C &= \frac{1}{N} \sum_{n=1}^{N} \log\det\left(\mathbf{I}_{N_r} + \frac{1}{N_0} \tilde{\mathbf{H}}_n \mathbf{V}_n \operatorname{diag}\left(P_{1,n}, ..., P_{N_t,n}\right) \mathbf{V}_n^* \tilde{\mathbf{H}}_n^*\right) \\ &= \frac{1}{N} \sum_{n=1}^{N} \sum_{i=1}^{N_t} \log\left(1 + \frac{P_{n,i}^* \lambda_{n,i}^2}{N_0}\right) \end{aligned} \tag{B.34}$$

where the optimal power allocation is given by waterfilling

$$P_{n,i}^* = \left(\mu - \frac{N_0}{\lambda_{n,i}^2}\right)^+ \tag{B.35}$$

with μ chosen such that the total power constraint $\sum_{n,i} P_{n,i}^* = NP$ is satisfied. As N goes to infinity, one can generalize the above discussion to arrive to the following conclusion:

$$\begin{aligned} C &= \frac{1}{W} \int_0^W \log\det\left(\mathbf{I}_{N_r} + \frac{1}{N_0} \tilde{\mathbf{H}}(f) \mathbf{K}(f) \tilde{\mathbf{H}}^*(f)\right) df \\ &= \frac{1}{W} \sum_{i=1}^{N_t} \int_0^W \log\left(1 + \frac{|\lambda_i(f)|^2 P_{i,i}^*(f)}{N_0}\right) df \end{aligned} \tag{B.36}$$

where $\mathbf{K}(f) = \mathbf{U}(f)\mathbf{P}(f)\mathbf{U}^*(f)$ and $\mathbf{P}(f)$ is a diagonal power allocation matrix. Here, $\tilde{\mathbf{H}}(f)$ is the point-wise discrete-time Fourier transform (DTFT) of $\mathbf{H}[n]$ and $\lambda_i(f)$ represents the i^{th} eigenvalue of $\tilde{\mathbf{H}}(f)$. The optimal power allocation in this case is given by

$$P_{i,i}^*(f) = \left(\frac{1}{\mu} - \frac{N_0}{|\lambda_i(f)|^2}\right)^+ \tag{B.37}$$

where μ is chosen such that

$$\sum_{i=1}^{N_t} \int_0^W \left(\frac{1}{\mu} - \frac{N_0}{|\lambda_i(f)|^2}\right)^+ df = P \tag{B.38}$$

When $\mathbf{H}[n]$ is not know at the transmitter, the power is divided equally across all transmit antennas and $\mathbf{Q}_n$ is chosen to be equal to an identity matrix for all n. In this case, the capacity becomes

$$C = \frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{N_t} \log\left(1 + \frac{\mathrm{SNR}\,\lambda_{n,i}^2}{N_t}\right) \tag{B.39}$$

As N goes to infinity, the capacity is given by

$$\begin{aligned} C &= \frac{1}{W}\int_0^W \log\det\left(\mathbf{I}_{N_r} + \frac{\mathrm{SNR}}{N_t}\tilde{\mathbf{H}}(f)\tilde{\mathbf{H}}^*(f)\right) df \\ &= \frac{1}{W}\sum_{i=1}^{N_t}\int_0^W \log\left(1 + \frac{\mathrm{SNR}\,|\lambda_i(f)|^2}{N_t}\right) df \end{aligned} \tag{B.40}$$

APPENDIX C

ORTHOGONAL FREQUENCY DIVISION MULTIPLEXING

In this appendix, we provide a brief overview on Orthogonal Frequency Division Multiplexing (OFDM). The results of this section are used in Chapters 3, 4, and 5.

OFDM is one of the most advanced modulation schemes and is widely used in today's wireless communication standards [27]. The trick is to use a modulation format that converts a frequency-selective channel into a set of parallel frequency flat channels to mitigate the effects of inter-symbol interference caused by the channel time spread. Blocks of inverse discrete Fourier transform (IDFT) symbols are typically preceded by a cyclic prefix (CP) and transmitted over a dispersive channel. Under this condition, a linear convolution of the transmitted sequence and the channel's impulse response is converted to a circular convolution. At the receiver, the discrete Fourier transform (DFT) is applied to the received block to transform the circular convolution to simple matrix product, thus eliminating the effects of ISI. Moreover, this approach enables the transmitter and receiver to use fast signal processing transforms such as the fast Fourier transform (FFT) and its inverse for the computation of the DFT and IDFT.

C.1 SISO OFDM

In a single-input single-output (SISO) communication system, the input-output relationship is given by

$$y[n] = \sum_{l=0}^{L-1} h[l]x[n-l] + v[n] \tag{C.1}$$

where $x[n]$ and $y[n]$ are the transmitted and received sequences, respectively. For every block of N symbols $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N)$, we first obtain the vector

$\mathbf{x} = (x[1], x[2], \ldots, x[N])$ by taking the IDFT of $\tilde{\mathbf{x}}$. Therefore, $\mathbf{x}$ is given by

$$\tilde{\mathbf{x}} = \mathbf{F}^* \tilde{\mathbf{x}} \tag{C.2}$$

where $\mathbf{F}$ is the DFT matrix and $\mathbf{F}^*$ is the IDFT matrix. We then add a cyclic prefix of $L-1$ symbols to form the following vector:

$$\mathbf{b} = (x[N-L+2], ..., x[N], x[1], x[2], \ldots, x[N]) \tag{C.3}$$

The symbols $b[n]$ are transmitted sequentially over the channel. With this input to the channel, the output is given by

$$d[n] = \sum_{l=0}^{L-1} h[l] b[n-l] + v[n] \tag{C.4}$$

The receiver discards the first $L-1$ symbols and forms the vector $\mathbf{y} = (d[L], ..., d[N+L-1])$. Under this condition the input-output vector relationship is given by $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$, where $\mathbf{v} = (v[L], ..., v[N+L-1])$ and

$$\mathbf{H} = \begin{bmatrix} h[0] & 0 & \ldots & 0 & h[L-1] & h[L-2] & \ldots & h[1] \\ h[1] & h[0] & 0 & \ldots & 0 & h[L-1] & \ldots & h[2] \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & h[L-1] & h[L-2] & \ldots & h[1] & h[0] \end{bmatrix} \tag{C.5}$$

is a circulant matrix. Let $\mathbf{h} = (h[0], ..., h[l-1], 0, ..., 0)$ be an $N \times 1$ vector. A classical result in linear algebra states that a circulant matrix can always be decomposed as $\mathbf{H} = \mathbf{F}^* \mathbf{\Lambda} \mathbf{F}$, where $\mathbf{\Lambda}$ is a diagonal matrix and $\mathbf{\Lambda}_{n,n} = \tilde{h}_n$, the n^{th} entry of the vector $\tilde{\mathbf{h}} = \sqrt{N}\mathbf{F}\mathbf{h}$. Thus, the following relation holds:

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{x} + \mathbf{v} \\ &= \mathbf{F}^* \mathbf{\Lambda} \mathbf{F} \mathbf{x} + \mathbf{v} \\ &= \mathbf{F}^* \mathbf{\Lambda} \tilde{\mathbf{x}} + \mathbf{v} \end{aligned} \tag{C.6}$$

If we let $\tilde{\mathbf{y}} = \mathbf{F}\mathbf{y}$, then the following holds:

$$\tilde{y}_n = \tilde{h}_n \tilde{x}_n + \tilde{v}_n \qquad n = 1, ..., N \tag{C.7}$$

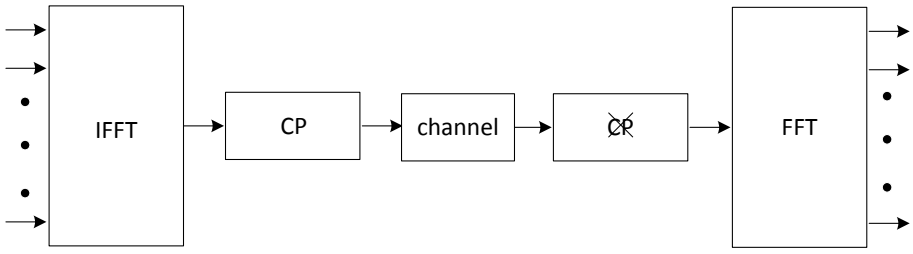


Figure C.1: OFDM modulation

and we have successfully transformed a frequency selective channel into a set of N frequency flat channels. The OFDM modulation scheme we just described is summarized in Figure C.1.

C.2 MIMO OFDM

A frequency selective MIMO channel is described by

$$\mathbf{y}[n] = \sum_{k=0}^{L-1} \mathbf{H}[k]\mathbf{x}[n-k] + \mathbf{v}[n] \tag{C.8}$$

where $\mathbf{y}[n]$ and $\mathbf{x}[n]$ are the detected and transmitted symbol vectors, respectively. The channel's matrix impulse response $\mathbf{H}[n]$ is assumed to have L non-zero taps. Multi-input multi-output (MIMO) OFDM is analogous to SISO OFDM. In fact, as shown in Figure C.2, a MIMO OFDM system applies OFDM modulation (see Figure C.1) for each transmit and receive antenna. Following the same derivation as in the previous section, one can show that the frequency selective MIMO channel is transformed into a set of frequency flat MIMO channels

$$\tilde{\mathbf{y}}_n = \tilde{\mathbf{H}}_n \tilde{\mathbf{x}}_n + \tilde{\mathbf{v}}_n \qquad\qquad n = 1, ..., N \tag{C.9}$$

where $\tilde{\mathbf{h}}_{i,j} = \left(\tilde{\mathbf{H}}_1^{i,j}, ..., \tilde{\mathbf{H}}_N^{i,j}\right)^T$ is the N-point discrete Fourier transform (DFT) of $h_{i,j}[n]$ scaled by $\sqrt{N}$. Here, $h_{i,j}[n]$ represents the impulse response from the j^{th} transmitter to the i^{th} receiver and $\tilde{\mathbf{H}}_l^{i,j}$ is the $(i,j)^{th}$ entry of $\tilde{\mathbf{H}}_l$.

The block length is crucial in OFDM systems as it dictates the size of FFTs and IFFTs and hence determines the overall computational complexity of the

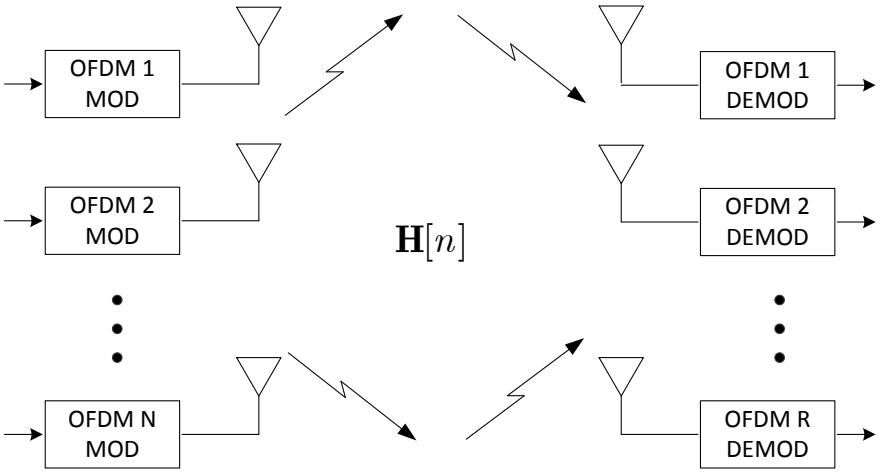


Figure C.2: MIMO OFDM modulation

system. For instance, if the block length is N, then for an $N_t \times N_r$ MIMO system, N_t N-point IFFTs must be computed at the transmitter and N_r N-points FFTs followed by N MIMO maximum likelihood (ML) estimates must be computed at the receiver. Therefore, the computational complexity of a MIMO-OFDM system is $O\left(\left(N_t + N_r\right) N \log\left(N\right) + N|\mathcal{A}|^{N_t}\right)$.

APPENDIX D

COMPUTATIONAL COMPLEXITY ANALYSIS

In this appendix, we derive the computational complexity expressions used in Chapter 5.

D.1 Computational Complexity of VVA

For VVA, the following operation needs to be performed for each state:

$$\mathcal{P}_k^i = \min_{j \in \mathcal{F}} \mathcal{P}_{k-1}^j + \|\mathbf{y}_k - \boldsymbol{\mu}_k \left(\mathcal{S}_k^j, \mathbf{x}^i[k]\right)\|^2,$$

where $\boldsymbol{\mu}_k \left(\mathcal{S}_k^j, \mathbf{x}^i[k]\right) = \sum_{l=0}^{L-2} \mathbf{H}[l]\mathbf{x}^i[k-l] + \mathbf{H}[L-1]\mathbf{x}^j[k-L+1]$ is precomputed for all i and j. There are $|\mathcal{A}|^{N_t}$ incoming branches for each state. To compute each branch metric, N_t complex additions, N_t complex multiplications, and $N_t - 1$ real additions are needed. Each computed branch metric has to be added to its corresponding path metric. This requires an additional real addition. Finally, to perform the min, $|\mathcal{A}|^{N_t} - 1$ comparisons are needed. Therefore, the per-state complexity of the VVA algorithm is given by

$$\begin{aligned} N_{add} &= 3N_t|\mathcal{A}|^{N_t} \\ N_{mult} &= N_t|\mathcal{A}|^{N_t} \\ N_{cmp} &= |\mathcal{A}|^{N_t} - 1. \end{aligned} \tag{D.1}$$

Here, N_{add} refers to the total number of real additions, N_{mult} refers to the total number of complex multiplications, and N_{cmp} refers to the number of comparisons. We assume that every complex addition is equivalent to two real additions.

D.2 Computational Complexity of Full Tree Search

For the full tree search algorithm, the following operation needs to be performed for each state:

$$\begin{aligned}\mathcal{P}_k^i &= \min_{j\in\mathcal{F}} \mathcal{P}_{k-1}^j + \|\mathbf{z}_k - \mathbf{G}\mathbf{x}^j\|^2 \\ &= \min_{j\in\mathcal{F}} \mathcal{P}_{k-1}^j + \|\tilde{\mathbf{z}}_k - \mathbf{R}\mathbf{x}^j\|^2, \end{aligned} \tag{D.2}$$

where $\mathbf{G} = \mathbf{Q}\left[\mathbf{R}^T \mathbf{0}_{\mathbf{N_t}\times(\mathbf{N_r}-\mathbf{N_t})}\right]^T$ by the QR decomposition and $\tilde{\mathbf{z}}_k$ corresponds to the first N_t entries of $\mathbf{Q}^*\mathbf{z}_k$. We assume that $\sum_{l=0}^{L-2}\mathbf{H}[l]\mathbf{x}^i[k-l]$ and $\mathbf{R}\mathbf{x}^j$ are precomputed for all i and j. First, to compute $\mathbf{z}_k = \mathbf{y}_k - \sum_{l=0}^{L-2}\mathbf{H}[l]\mathbf{x}^i[k-l]$, N_r complex additions are needed. The result has to be multiplied by $\mathbf{Q}^*$ to obtain $\tilde{\mathbf{z}}_k$. This requires $(N_r-1)\,N_t$ complex additions and N_rN_t complex multiplications. Next, we have to compute all the partial Euclidean distances. In a $|\mathcal{A}|$-ary tree of depth N_t, there are $\sum_{i=1}^{N_t}|\mathcal{A}|^i$ edges. Therefore, in order to compute the weights $e_l\left(x_l^j,...,x_{N_t}^j\right)$ of all edges, we need $\sum_{i=1}^{N_t}|\mathcal{A}|^i$ complex additions and $\sum_{i=1}^{N_t}|\mathcal{A}|^i$ complex multiplications. After having computed the weights of all edges in the tree, we need to traverse the tree from the root node to every leaf node to add the weights of all edges to each other and then add the result to the path metric. This requires $\sum_{i=2}^{N_t}|\mathcal{A}|^i + |\mathcal{A}|^{N_t}$ real additions. The number of comparisons that are needed is identical to VVA. Therefore, the per-state complexity of full tree search is given by

$$\begin{aligned} N_{add} &= \sum_{i=2}^{N_t}|\mathcal{A}|^i + |\mathcal{A}|^{N_t} + 2\sum_{i=1}^{N_t}|\mathcal{A}|^i + 2\,(N_r-1)\,N_t + 2N_r \\ N_{mult} &= \sum_{i=1}^{N_t}|\mathcal{A}|^i + N_rN_t \\ N_{cmp} &= |\mathcal{A}|^{N_t} - 1. \end{aligned} \tag{D.3}$$

REFERENCES

[1] R. W. Tkach, "Scaling optical communications for the next decade and beyond," *Bell Labs Technical Journal*, vol. 14, no. 4, pp. 3–9, 2010.

[2] G. C. Papen and R. E. Blahut, "Lightwave Communication Systems," unpublished draft, 2012.

[3] P. Winzer, "Modulation and multiplexing in optical communications," in *Lasers and Electro-Optics, 2009 and 2009 Conference on Quantum electronics and Laser Science Conference. CLEO/QELS 2009. Conference on*, June 2009, pp. 1 –2.

[4] P. J. Winzer and G. J. Foschini, "MIMO capacities and outage probabilities in spatially multiplexed optical transport systems," *Opt. Express*, vol. 19, pp. 16 680–16 696, 2011.

[5] A. F. Benner, M. Ignatowski, J. A. Kash, D. M. Kuchta, and M. B. Ritter, "Exploitation of optical interconnects in future server architectures," *IBM Journal of Research and Development*, vol. 49, no. 4.5, pp. 755 –775, July 2005.

[6] Y. Koike and S. Takahashi, "Plastic optical fibers: Technologies and communication links," in *Optical Fiber Telecommunications V A*, I. P. Kaminow, T. Li, and A. E. Willner, Eds. Burlington: Academic Press, 2008, pp. 593 – 603.

[7] G. P. Agarwal, *Fiber-Optic Communication Systems*, 3rd ed. Wiley, 2002.

[8] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[9] H. R. Stuart, "Dispersive multiplexing in multimode optical fiber," *Science*, vol. 289, no. 5477, pp. 281–283, 2000. [Online]. Available: http://www.sciencemag.org/content/289/5477/281.abstract

[10] R. Khosravani, I. T. Lima Jr., P. Ebrahimi, E. Ibragimov, A. Willner, and C. Menyuk, "Time and frequency domain characteristics of polarization-mode dispersion emulators," *Photonics Technology Letters, IEEE*, vol. 13, no. 2, pp. 127 –129, Feb. 2001.

[11] K.-P. Ho and J. M. Kahn, "Mode-dependent loss and gain: statistics and effect on mode-division multiplexing," *Opt. Express*, vol. 19, pp. 16 612–16 635, 2011.

[12] K.-P. Ho and J. Kahn, "Statistics of group delays in multimode fiber with strong mode coupling," *Lightwave Technology, Journal of*, vol. 29, no. 21, pp. 3119 –3128, Nov. 1, 2011.

[13] C. Shannon and W. Weaver, *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, 1949.

[14] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[15] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998, 10.1023/A:1008889222784. [Online]. Available: http://dx.doi.org/10.1023/A:1008889222784

[16] A. Shah, R. Hsu, A. Tarighat, A. Sayed, and B. Jalali, "Coherent optical MIMO (COMIMO)," *Lightwave Technology, Journal of*, vol. 23, no. 8, pp. 2410 – 2419, Aug. 2005.

[17] K.-P. Ho and J. Kahn, "Frequency diversity in mode-division multiplexing systems," *Lightwave Technology, Journal of*, vol. 29, no. 24, pp. 3719 –3726, Dec. 15, 2011.

[18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 2006.

[19] E. Alon, V. Stojanovic, J. Kahn, S. Boyd, and M. Horowitz, "Equalization of modal dispersion in multimode fiber using spatial light modulators," in *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, Nov.-3 Dec. 2004, pp. 1023 – 1029.

[20] H. Chen, H. van den Boom, and A. Koonen, "30gbit/s 3×3 optical mode group division multiplexing system with mode-selective spatial filtering," in *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference*, March 2011, pp. 1 –3.

[21] W. Shieh, H. Bao, and Y. Tang, "Coherent optical OFDM: theory and design," *Opt. Express*, vol. 16, no. 2, pp. 841–859, Jan 2008. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-16-2-841

[22] P. Billingsley, *Probability and Measure*. Wiley, 2012.

[23] G. Shilov, *Linear Algebra*. Prentice-Hall Inc, 1971.

[24] F. Mezzadri, "How to generate random matrices from the classical compact groups," Feb. 2007. [Online]. Available: http://arxiv.org/abs/math-ph/0609050

[25] W. van Etten, "Maximum likelihood receiver for multiple channel transmission systems," *Communications, IEEE Transactions on*, vol. 24, no. 2, pp. 276 – 283, Feb. 1976.

[26] G. D. Forney, "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268 – 278, March 1973.

[27] G. Stuber, J. Barry, S. McLaughlin, Y. Li, M. Ingram, and T. Pratt, "Broadband MIMO-OFDM wireless communications," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 271 – 294, Feb. 2004.

[28] S. Parkvall, E. Dahlman, A. Furuskar, Y. Jading, M. Olsson, S. Wanstedt, and K. Zangi, "LTE-Advanced - evolving LTE towards IMT-Advanced," in *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th*, Sept. 2008, pp. 1 –5.

[29] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Mathematics of Computation*, vol. 44, pp. 463–471, 1985.

[30] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm I. expected complexity," *Signal Processing, IEEE Transactions on*, vol. 53, no. 8, pp. 2806 – 2818, Aug. 2005.

[31] H. Vikalo and B. Hassibi, "On the sphere-decoding algorithm II. generalizations, second-order statistics, and applications to communications," *Signal Processing, IEEE Transactions on*, vol. 53, no. 8, pp. 2819 – 2834, Aug. 2005.

[32] H. Vikalo and B. Hassibi, "Maximum-likelihood sequence detection of multiple antenna systems over dispersive channels via sphere decoding," *EURASIP J. Appl. Signal Process.*, vol. 2002, pp. 525–531, January 2002.

[33] M. Ozols, "How to generate a random unitary matrix," 2009.

[34] D. Petz and J. Réffy, "On asymptotics of large haar distributed unitary matrices," *Periodica Mathematica Hungarica*, vol. 49, no. 1, pp. 103–117, 2004.

[35] R. Gallager, *Information Theory and Reliable Communication*. Wiley, 1968.