Performance Analysis for Near-Field Atmospheric Optical Communications

by

Baris Ibrahim Erkmen

Submitted to the Department of Electrical Engineering and Computer Science

in partial fulfillment of the requirements for the degree of
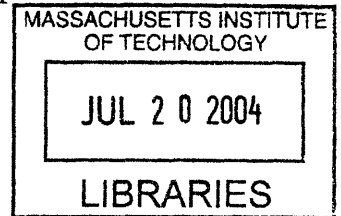
Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2003

©2003 Baris Ibrahim Erkmen. All rights reserved.

Author . . . .
Department of Electrical Engineering and Computer Science
August 06, 2003

Certified by. . . . .
Jeffrey H. Shapiro
Julius A. Stratton Professor of Electrical Engineering
Thesis Supervisor

Accepted by . . . .
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Performance Analysis for Near-Field Atmospheric Optical Communications

by

Baris Ibrahim Erkmen

Submitted to the Department of Electrical Engineering and Computer Science
on August 06, 2003, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

Establishing high data rate communication links over short distances via fiber-optic cables may prove costly and time consuming, particularly in urban areas. Through-the-air optical communication systems are a reliable and cost effective alternative for establishing high data rate links when the two terminals have a clear line-of-sight. Bad weather conditions, such as rain, snow or fog, degrade performance severely for such links. However, even in clear weather conditions, local refractive index fluctuations in the atmosphere known as atmospheric turbulence, may drastically impair performance. Most optical links are set up to operate in the far-field power transfer regime, in which diffraction spread is the dominant effect on the beam. Therefore, a very small portion of the beam is captured at the receiver, resulting in very weak power coupling between the transmitter and the receiver. However, it is also possible to establish geometries such that the link operates in the near-field regime, where, in absence of turbulence, it is possible to focus the beam onto the receiver with almost perfect power coupling. Work on performance of near-field atmospheric optical communication systems is scarce in existing literature, perhaps due to increased complexity in prescribed models. In this thesis, we analyze error probabilities of binary optical communication links operating in the near-field regime, utilizing on-off keying (OOK) or pulse position modulation (PPM) signaling techniques. We also obtain bounds on the capacity of a single-input, single-output (SISO) atmospheric optical communication link with a coherent-detection receiver operating in the near field.

Thesis Supervisor: Jeffrey H. Shapiro
Title: Julius A. Stratton Professor of Electrical Engineering

Acknowledgments

I wish to thank foremost, Prof. Jeffrey Shapiro. In the two years I worked with him as an Undergraduate Research Opportunities Program (UROP) student and the following two years as a graduate student, his guidance, support and insight have been invaluable to my research as well as my academic enrichment. I feel privileged to have met him early on in my academic career.

I also thank Dr. Franco Wong, from whom I've learned much about the intricacies of lab work. My office partner, Mohsen Razavi, has been incredibly helpful in answering all the random technical questions I have thrown at him throughout this work. Discussions with Shane Haas have proven very useful.

The greatest burden of my studying abroad lies on the shoulders of my parents, Aydan and Ismet Erkmen. In the six years I have spent here, their support, encouragement and most of all, their confidence in me have always been a driving force. The appreciation I can express in this short paragraph is only a small fraction of what I feel. I would like to extend special thanks to my sister Burcu Erkmen, whose presence in my life has always been heart-warming. The many cups of coffee she has delivered to my office only add to this warmth.

I also cannot thank Alexandra Chau enough for her support and encouragement. Finally, I would like to thank my friends Adnan Ilik, Andrew Lundgren, Ashish Koul, Borga Erbil, Emek Tuzun, Kerem Ozen and Oguz Ergin for proving that our friendship is greater than the oceans and higher than the mountains that may lie between us.

Contents

List of Figures

10

11

12

Chapter 1

Introduction

The increasing demand in the past decade for higher data-rate communication networks has been answered with the advances in fiber-optic communications. However, particularly in urban areas, replacing existing network infrastructure with fiber-optic communications remains as an expensive and time consuming challenge, and one that is disruptive to every-day life.

In the past decade, atmospheric optical communications has received much attention as an appealing alternative to fiber-optic and microwave communications. Improvements in semiconductor laser technology as well as optical amplification and detection technologies have significantly reduced the cost of hardware necessary to set up such links. Furthermore, avoiding the high construction costs and the need for difficult to obtain rights of way in urban areas associated with laying down fiber-optic cables provides strong incentive to consider atmospheric optical communications as a cost-effective alternative. Atmospheric optical links are also highly modular and can be installed fairly quickly. Hence they can serve as temporary solutions to maintain high data rate communications in the absence of a permanent infrastructure. Likewise, they may prove useful in emergency situations, such as in the immediate aftermath of natural disasters, when re-establishing global communications is vital, yet extremely challenging if the existing infrastructure is damaged. Through-the-air optical links also have unlicensed bandwidth, therefore can yield much higher data-rates than microwave communications. Finally, because laser beams can be much

15

narrower than microwave beams, atmospheric optical communications offer a more secure alternative to microwave communications.

Optical propagation through the atmosphere, however, is susceptible to meteorological events and atmospheric conditions. Bad weather conditions, such as snow, fog and rain, cause severe absorption and scattering, drastically reducing optical power transfer on a line-of-sight atmospheric path [1, 2, 3, 4]. Even in clear weather conditions, however, a phenomenon known as atmospheric turbulence significantly degrades the performance of atmospheric optical communication links. Atmospheric turbulence is what causes stars to twinkle at night and scenery to shimmer above an asphalt road on a hot day. Physically, atmospheric turbulence is due to parcels of air, referred to as "turbulent eddies", with slightly varying temperatures and therefore slightly varying refractive indices [5, 6, 7]. These refractive index fluctuations result in random distortions in the wavefront as well as constructive and destructive interference in the intensity of propagating optical waves, leading to undesirable effects such as beam spread, angular spread and fading.

A number of optical wave propagation theories and models have been developed to date which quantify these atmospheric propagation phenomena. One such model, which we shall utilize in this work, is the extended Huygens-Fresnel principle [8, 9]. This model treats the transmitter surface as a collection of infinitesimal point sources and uses the random spherical-wave Green's function of the turbulent atmosphere to describe the field at a point on the receiver surface via a superposition integral. This approach enables a linear system interpretation of atmospheric propagation, in which the wave at the transmitter aperture is considered the input and the wave entering the receiver aperture is considered the output of this system. The linear system interpretation provided by the extended Huygens-Fresnel principle is a starting point for a number of other theories which attempt to quantify wave propagation statistics and power transfer characteristics for the turbulent atmosphere. We will employ some of these theories in this thesis as well.

Free-space line-of-sight propagation between finite transmitter and receiver apertures may be divided into near-field and far-field power transfer regimes, which bracket

16

the power transfer characteristics of an atmospheric optical link. Qualitatively, in free-space propagation, the far field corresponds to the regime in which the minimum beam size at the receiver plane is much larger than the size of the aperture, due to diffraction spread [10, 11]. Thus, only a small fraction of the transmitted power is coupled into the receiver. On the other hand, the near field is the regime in which the minimum beam size is much less than the aperture area, yielding near-perfect power coupling when the beam is focused on the receiver. It turns out that these free-space power transfer characteristics can be summarized by a single parameter, the free-space Fresnel number. When this number is much smaller than unity, far-field power transfer prevails; when it is much larger than unity, power transfer is in the near field [11]. In the turbulent atmosphere, similar power-transfer regimes exist, with the added complication that the minimum beam size at the receiver depends on the random atmospheric state. Nevertheless, a parameterization similar to that for free-space is possible: for the turbulent atmosphere the operating regime is determined by comparing a random effective Fresnel number with 1.

Work in the field of atmospheric optical communications has mainly focused on the far-field power transfer regime [8, 5]. On the other hand, near-field analysis has been used mostly for imaging applications. However, it is possible to establish atmospheric communication links which operate in the near-field regime. There are a number of motivations for understanding the communication performance of optical links in the near field. As described earlier, the ability to focus a beam onto the receiver results in high power coupling, as opposed to a link operating in the far-field regime. Therefore, much less transmitter power will be required to achieve data rates and error probabilities comparable to those systems that operate in the far field. Furthermore, the reduced diffraction of near-field systems is attractive for high-security applications or situations in which privacy is at a premium.

In this thesis, we will analyze the performance of optical communication links operating in the near-field power transfer regime through atmospheric turbulence in clear weather conditions. Chapter 2 presents background material on atmospheric propagation and optical detection. Chapter 3 will be devoted to the maximum power

transfer statistics of the turbulent atmosphere. There we focus on bounding the mean and variance of the maximum power transfer. Through-the-air binary, atmospheric, optical communication systems with on-off keying (OOK) and pulse-position modulation (PPM) signaling techniques, in combination with direct detection or coherent detection receivers will be analyzed in Chapter 4. The results of Chapter 4 will be extended, in Chapter 5, to the case in which the maximum power transfer is assumed to be Beta distributed. Finally, Chapter 6 will analyze the ergodic and outage capacity of near-field atmospheric optical communication links employing coherent detection receivers.

Chapter 2

Background

2.1 Atmospheric Propagation

This section will discuss the linear system model for atmospheric optical wave propagation and its normal mode decomposition. The material will provide the necessary tools for analyzing the power transfer statistics of the atmosphere.

2.1.1 Wave Propagation Model

Optical waves propagating through the atmosphere are affected by atmospheric conditions and meteorological phenomena. In clear weather, turbulent mixing of air causes local temperature fluctuations in the atmosphere, generating spatially and temporally random air parcels with slightly varying refractive indices, as illustrated in Figure 2-1. Although these air parcels, referred to as "turbulent eddies," have refractive index variations on the order of 10^{-6}, their effect on wave propagation is profound. For optical communication systems, constructive and destructive interference results in fading, while distortion of the phase-front leads to beam spread and angular spread [5].

A number of theories have been developed to quantify wave propagation through the turbulent atmosphere, with considerable success in agreeing with experimental observations [6]. One such propagation model, widely utilized in atmospheric optical
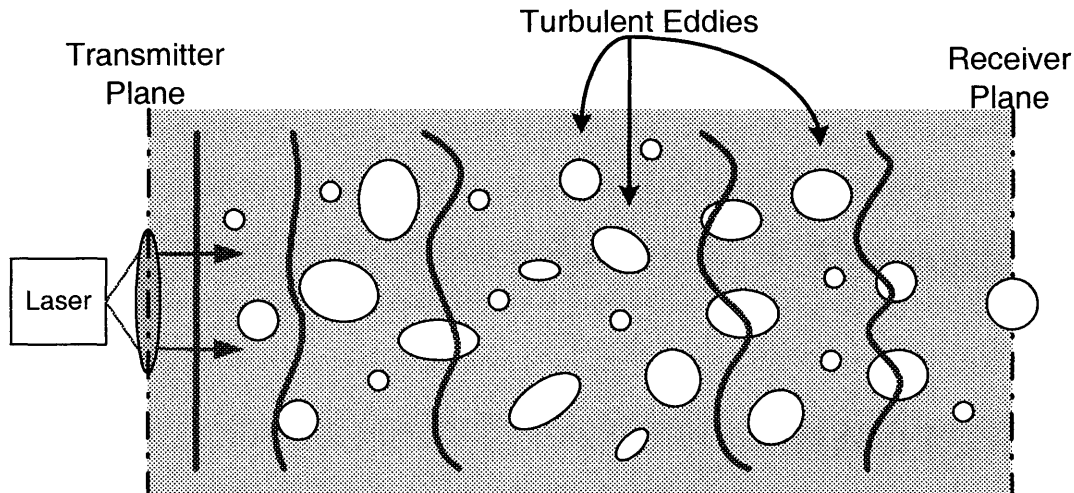
Figure 2-1: Propagation of a plane wave through the turbulent atmosphere.

communications studies, is the extended Huygens-Fresnel principle, which expresses the propagated field as a superposition integral of spherical waves originating from point sources that collectively constitute the source. This approach provides a system-level interpretation for atmospheric propagation, in which the field entering the medium can be considered the input to a stochastic, space and time-varying linear system with the atmosphere's spherical-wave Green's function as its impulse response. The output is considered to be the field at another plane along the propagation path.

Consider a line-of-sight propagation path where the transmitter and receiver apertures are placed in parallel, along some common principal axis, as shown in Figure 2-2. For definiteness, assume the propagation direction is chosen to be along the z-axis, with x and y axes as depicted in the figure. Furthermore, assume a quasimonochromatic and linearly polarized electric field with center frequency ω_o is transmitted from an aperture R_1 on the $z = 0$ plane.

For high data rate communications, it is feasible to assume that the duration of a transmitted symbol is much less than the atmospheric coherence time and the transmitted field has bandwidth much less than the atmospheric coherence bandwidth [5, 8, 10]. As a result, the atmosphere can be assumed *frozen* for the duration of transmission. In other words, temporal variations in the refractive index field can be suppressed, leaving only spatial variations. Furthermore, it has been shown that

20
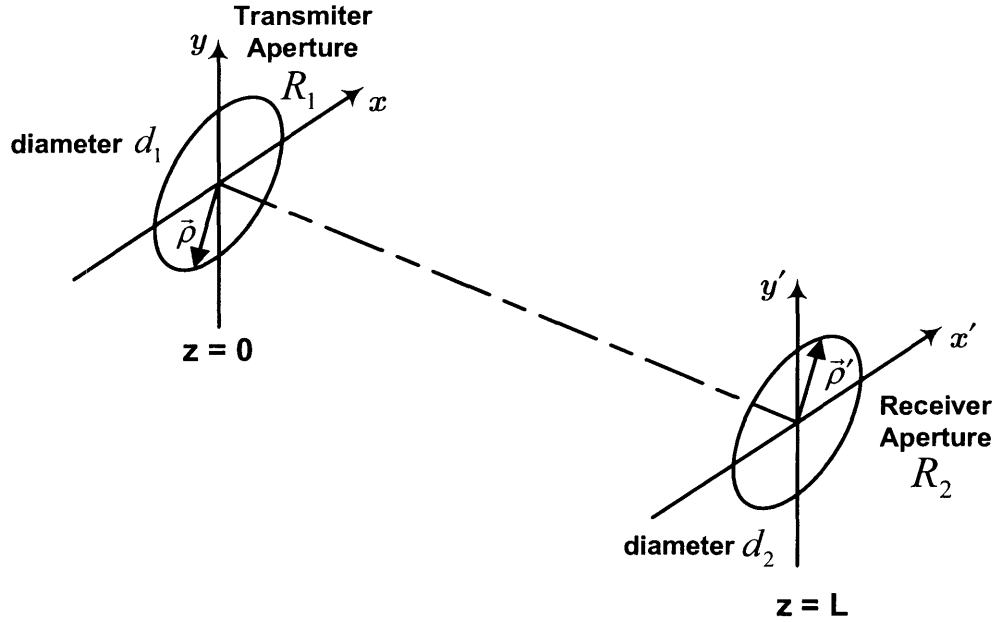
Figure 2-2: Line-of-sight propagation geometry.

turbulent atmosphere exhibits negligible depolarization [12], thus consideration of linearly polarized fields leads to scalar field equations.

Let us use $U_i(\boldsymbol{\rho}, t)$ to denote the scalar complex envelope of the transmitted field, where $\boldsymbol{\rho}$ is a point on the $z = 0$ plane; i.e., the point $(\boldsymbol{\rho}, z = 0) \in \mathbb{R}^3$. Under the assumptions specified in the previous paragraph, the extended Huygens-Fresnel principle states that the scalar complex envelope of the received field on the $z = L$ plane, $U_o(\boldsymbol{\rho}', t)$, is

$$U_o(\boldsymbol{\rho}', t) = \frac{1}{j\lambda L} \int_{R_1} U_i\left(\boldsymbol{\rho}, t - \tfrac{L}{c}\right) e^{jk\left(L + \frac{|\boldsymbol{\rho}' - \boldsymbol{\rho}|^2}{2L}\right) + \chi(\boldsymbol{\rho}', \boldsymbol{\rho}) + j\phi(\boldsymbol{\rho}', \boldsymbol{\rho})} \, d\boldsymbol{\rho} . \qquad (2.1)$$

Here, $U_i(\boldsymbol{\rho}, t)$ and $U_o(\boldsymbol{\rho}', t)$ are complex field envelopes[1] about optical (carrier) frequency $\omega_o = \frac{2\pi c}{\lambda}$, in terms of their center wavelength λ, $k = \frac{\omega_o}{c}$ is the wave number and $\chi(\boldsymbol{\rho}', \boldsymbol{\rho})$ and $\phi(\boldsymbol{\rho}', \boldsymbol{\rho})$ are turbulence-induced log-amplitude and phase fluctuations at $(\boldsymbol{\rho}', z = L)$, due to a monochromatic point source at $(\boldsymbol{\rho}, z = 0)$. Equation (2.1) can be

---

[1] The units of these fields are $\frac{(\text{Watts})^{1/2}}{m}$, so that $\int_{R_1} |U_i(\boldsymbol{\rho}, t)|^2 d\boldsymbol{\rho}$ is the short-time average power transmitted through the R_1 aperture at time t, etc.

interpreted as the input/output description of a linear system with impulse response,

$$h(\boldsymbol{\rho}', \boldsymbol{\rho}) = \frac{1}{j\lambda L} e^{jk\left(L + \frac{|\boldsymbol{\rho}' - \boldsymbol{\rho}|^2}{2L}\right) + \chi(\boldsymbol{\rho}', \boldsymbol{\rho}) + j\phi(\boldsymbol{\rho}', \boldsymbol{\rho})} .$$ (2.2)

$h(\boldsymbol{\rho}', \boldsymbol{\rho})$, referred to as the atmospheric propagation Green's function, is a multiplicatively perturbed version of the free-space propagation Green's function,

$$h_{\text{fs}}(\boldsymbol{\rho}', \boldsymbol{\rho}) = \frac{1}{j\lambda L} e^{jk\left(L + \frac{|\boldsymbol{\rho}' - \boldsymbol{\rho}|^2}{2L}\right)} .$$ (2.3)

In the weak perturbation regime for turbulence, $\chi(\boldsymbol{\rho}', \boldsymbol{\rho})$ and $\phi(\boldsymbol{\rho}', \boldsymbol{\rho})$ are real-valued, statistically homogeneous and isotropic, jointly Gaussian random fields [8, 5]. Therefore fully characterizing $\chi(\boldsymbol{\rho}', \boldsymbol{\rho})$ and $\phi(\boldsymbol{\rho}', \boldsymbol{\rho})$ requires specifying their mean functions, auto-covariance functions, and cross-covariance function, defined respectively as

$$m_\chi = E\big[\chi(\boldsymbol{\rho}', \boldsymbol{\rho})\big] ,$$ (2.4)

$$m_\phi = E\big[\phi(\boldsymbol{\rho}', \boldsymbol{\rho})\big] ,$$ (2.5)

$$K_{\chi,\chi}(\boldsymbol{\rho}', \boldsymbol{\rho}) = E\Big[\big\{\chi(\boldsymbol{\rho}_1' + \boldsymbol{\rho}', \boldsymbol{\rho}_1 + \boldsymbol{\rho}) - m_\chi\big\} \big\{\chi(\boldsymbol{\rho}_1', \boldsymbol{\rho}_1) - m_\chi\big\}\Big] ,$$ (2.6)

$$K_{\phi,\phi}(\boldsymbol{\rho}', \boldsymbol{\rho}) = E\Big[\big\{\phi(\boldsymbol{\rho}_1' + \boldsymbol{\rho}', \boldsymbol{\rho}_1 + \boldsymbol{\rho}) - m_\phi\big\} \big\{\phi(\boldsymbol{\rho}_1', \boldsymbol{\rho}_1) - m_\phi\big\}\Big] ,$$ (2.7)

$$K_{\chi,\phi}(\boldsymbol{\rho}', \boldsymbol{\rho}) = E\Big[\big\{\chi(\boldsymbol{\rho}_1' + \boldsymbol{\rho}', \boldsymbol{\rho}_1 + \boldsymbol{\rho}) - m_\chi\big\} \big\{\phi(\boldsymbol{\rho}_1', \boldsymbol{\rho}_1) - m_\phi\big\}\Big] .$$ (2.8)

Because propagation through clear turbulent air is both passive and lossless, average power must be conserved. This condition implies that the mean of the log-amplitude field must satisfy

$$m_\chi = -K_{\chi,\chi}(\mathbf{0}, \mathbf{0}) = -\sigma_\chi^2 .$$ (2.9)

Furthermore, the phase field is assumed to have zero-mean, in accordance with experimental observations. Assuming Kolmogorov spectrum for turbulence [13] and a constant turbulence strength profile $C_n^2(z) = C_n^2$, the auto-covariance and cross-

covariance expressions are given by [8]

$$K_{\chi,\chi}(\boldsymbol{\rho}', \boldsymbol{\rho}) = 0.132\,\pi^2 k^2 L\, C_n^2 \int_0^1 \int_0^\infty u^{-8/3}\, J_0(d\,u)\, \sin^2\!\left(\frac{u^2 z(1-z)}{2\,k/L}\right) du\, dz\ , \quad (2.10)$$

$$K_{\phi,\phi}(\boldsymbol{\rho}', \boldsymbol{\rho}) = 0.132\,\pi^2 k^2 L\, C_n^2 \int_0^1 \int_0^\infty u^{-8/3}\, J_0(d\,u)\, \cos^2\!\left(\frac{u^2 z(1-z)}{2\,k/L}\right) du\, dz\ , \quad (2.11)$$

$$K_{\chi,\phi}(\boldsymbol{\rho}', \boldsymbol{\rho}) = 0.066\,\pi^2 k^2 L\, C_n^2 \int_0^1 \int_0^\infty u^{-8/3}\, J_0(d\,u)\, \sin\!\left(\frac{u^2 z(1-z)}{k/L}\right) du\, dz\ , \quad (2.12)$$

where $d = |z\boldsymbol{\rho}' + (1-z)\boldsymbol{\rho}|$. Then it follows that the the log-amplitude variance is

$$\sigma_\chi^2 = 0.124\, k^{7/6}\, C_n^2\, L^{11/6} \qquad (2.13)$$

in the weak perturbation regime; i.e. when the right hand side is less than 0.5. The phase variance, on the other hand, satisfies $K_{\phi,\phi}(\mathbf{0}, \mathbf{0}) \gg 1$, so that, via (2.1),

$$E\big[U_o(\boldsymbol{\rho}')\big] \approx 0 \qquad (2.14)$$

and

$$E\big[U_o(\boldsymbol{\rho_1'})U_o(\boldsymbol{\rho_2'})\big] \approx 0\ , \qquad (2.15)$$

where we have suppressed the temporal modulation of the transmitter in order to focus our attention on the turbulence-induced spatial effects. Therefore, the lowest order non-zero moment of the output field in the weak perturbation regime is the mutual coherence function,

$$\begin{aligned} E\big[U_o(\boldsymbol{\rho_1'})U_o^*(\boldsymbol{\rho_2'})\big]& \\ = \int_{R_1}\int_{R_1} U_i\,(\boldsymbol{\rho_1})\, U_i^*\,(\boldsymbol{\rho_2})\, &E\,\big[h(\boldsymbol{\rho_1'}, \boldsymbol{\rho_1})h^*(\boldsymbol{\rho_2'}, \boldsymbol{\rho_2})\big]\, d\boldsymbol{\rho_1}\, d\boldsymbol{\rho_2}\ , \end{aligned} \qquad (2.16)$$

where the expectation is given by,

$$E\left[h(\boldsymbol{\rho}_1', \boldsymbol{\rho}_1)h^*(\boldsymbol{\rho}_2', \boldsymbol{\rho}_2)\right]$$

$$= \left(\frac{1}{\lambda L}\right)^2 E\left[e^{\chi(\boldsymbol{\rho}_1', \boldsymbol{\rho}_1)+\chi(\boldsymbol{\rho}_2', \boldsymbol{\rho}_2)+j\phi(\boldsymbol{\rho}_1', \boldsymbol{\rho}_1)-j\phi(\boldsymbol{\rho}_2', \boldsymbol{\rho}_2)}\right] e^{\frac{jk}{2L}\left(|\boldsymbol{\rho}_1'-\boldsymbol{\rho}_1|^2-|\boldsymbol{\rho}_2'-\boldsymbol{\rho}_2|^2\right)} \quad (2.17)$$

$$= \left(\frac{1}{\lambda L}\right)^2 e^{\frac{jk}{2L}\left(|\boldsymbol{\rho}_1'-\boldsymbol{\rho}_1|^2-|\boldsymbol{\rho}_2'-\boldsymbol{\rho}_2|^2\right)-\frac{1}{2}D\left(\boldsymbol{\rho}_1'-\boldsymbol{\rho}_2', \boldsymbol{\rho}_1-\boldsymbol{\rho}_2\right)} . \quad (2.18)$$

$D(\boldsymbol{\rho}', \boldsymbol{\rho})$ in (2.18) is the two-source spherical-wave wave structure function defined as

$$D(\boldsymbol{\rho}', \boldsymbol{\rho}) = D_{\chi,\chi}(\boldsymbol{\rho}', \boldsymbol{\rho}) + D_{\phi,\phi}(\boldsymbol{\rho}', \boldsymbol{\rho}) \quad (2.19)$$

$$= 2\left[K_{\chi,\chi}(\mathbf{0}, \mathbf{0}) - K_{\chi,\chi}(\boldsymbol{\rho}', \boldsymbol{\rho})\right] + 2\left[K_{\phi,\phi}(\mathbf{0}, \mathbf{0}) - K_{\phi,\phi}(\boldsymbol{\rho}', \boldsymbol{\rho})\right] , \quad (2.20)$$

which is equal to

$$= 2.91\, k^2\, C_n^2\, L \int_0^1 \left|s\boldsymbol{\rho}' + (1-s)\boldsymbol{\rho}\right|^{5/3} ds , \quad (2.21)$$

for Kolmogorov spectrum turbulence with a constant turbulence strength profile $C_n^2(z) = C_n^2$. Hence, the mutual coherence is now expressed in terms of known parameters and can be evaluated for different turbulence profiles. It is also relevant to note that although the mutual coherence function result is derived in the weak perturbation regime, it remains the correct mutual coherence function for the received field $U_o(\boldsymbol{\rho}')$ well into the strong perturbation regime [5].

2.1.2 Normal Mode Decomposition

If one regards the instantaneous state of the atmospheric propagation medium together with the input and output apertures as a passive resonator, then the normal mode decomposition refers to obtaining the input and output transverse modes, along with their associated diffraction losses for a single pass through the medium. An equivalent linear system interpretation consists of regarding propagation through turbulent atmosphere as a linear transformation from an input field on the transmitter

24

aperture to an output field on the receiver aperture. Then, the normal mode decomposition corresponds to obtaining the input and output eigenfunctions (diffraction modes) and associated singular values (single-pass diffraction losses).

For $h(\boldsymbol{\rho}', \boldsymbol{\rho})$, the atmospheric propagation channel impulse response, we define the propagation kernel

$$K(\boldsymbol{\rho_1}, \boldsymbol{\rho_2}) = \int_{R_2} h^*(\boldsymbol{\rho}', \boldsymbol{\rho_1}) \, h(\boldsymbol{\rho}', \boldsymbol{\rho_2}) \, d\boldsymbol{\rho}' \,, \qquad \boldsymbol{\rho_1}, \, \boldsymbol{\rho_2} \in R_1 \qquad (2.22)$$

which is Hermitian and positive semi-definite by construction. This kernel has a complete, orthonormal (CON) set of eigenfunctions on R_1, $\{\Phi_m(\boldsymbol{\rho}) \, : 1 \leq m < \infty\}$, and a set of associated eigenvalues, $\{\eta_m \, : 1 \leq m < \infty\}$, which satisfy [11, 8]

$$\int_{R_1} K(\boldsymbol{\rho_1}, \boldsymbol{\rho_2}) \, \Phi_m(\boldsymbol{\rho_2}) \, d\boldsymbol{\rho_2} = \eta_m \, \Phi_m(\boldsymbol{\rho_1}) \,, \qquad \boldsymbol{\rho_1} \in R_1 \,, \qquad (2.23)$$

where $\eta_m \geq 0$. We use this set of input eigenfunctions, $\{\Phi_m(\boldsymbol{\rho}) \, : 1 \leq m < \infty\}$, and the associated eigenvalues, $\{\eta_m \, : 1 \leq m < \infty\}$, to define a CON set of output eigenfunctions on R_2 via

$$\int_{R_1} h(\boldsymbol{\rho}', \boldsymbol{\rho}) \, \Phi_m(\boldsymbol{\rho}) \, d\boldsymbol{\rho} = \eta_m^{1/2} \, \phi_m(\boldsymbol{\rho}') \,, \qquad \boldsymbol{\rho}' \in R_2 \,. \qquad (2.24)$$

Without loss of generality, we shall assume that the $\{\eta_m\}$ are arranged in decreasing order. It is worthwhile noting that in (2.22)–(2.24), we have in effect obtained the singular value decomposition of the channel impulse response $h(\boldsymbol{\rho}', \boldsymbol{\rho})$.

Because $\{\Phi_m(\boldsymbol{\rho}) \, : 1 \leq m < \infty\}$ and $\{\phi_m(\boldsymbol{\rho}') \, : 1 \leq m < \infty\}$ are CON sets on their respective domains, every scalar input field $U_i(\boldsymbol{\rho})$, $\boldsymbol{\rho} \in R_1$ and output field $U_o(\boldsymbol{\rho}')$, $\boldsymbol{\rho}' \in R_2$ can be represented as a linear combination of the respective eigenfunctions, via

$$U_i(\boldsymbol{\rho}) = \sum_{m=1}^{\infty} u_{im} \Phi_m(\boldsymbol{\rho}) \,, \qquad \text{where} \quad u_{im} = \int_{R_1} U_i(\boldsymbol{\rho}) \Phi_m^*(\boldsymbol{\rho}) \, d\boldsymbol{\rho} \qquad (2.25)$$

25
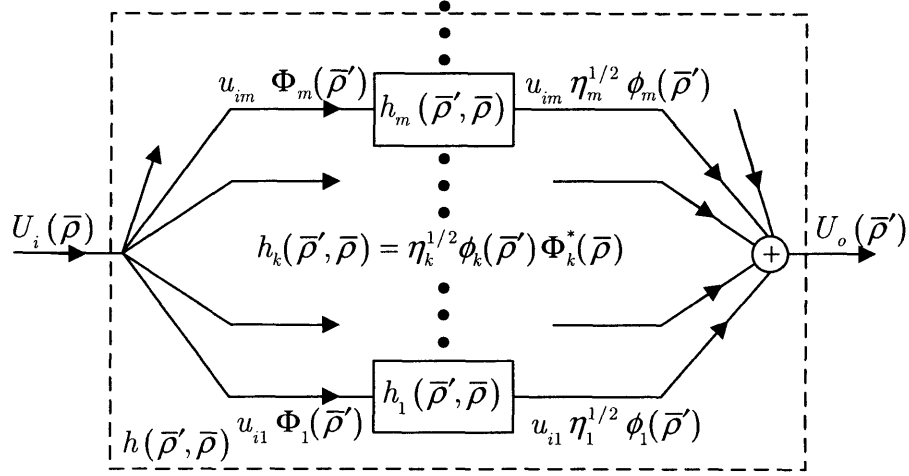
Figure 2-3: Parallel channel representation of the propagation impulse response.

and

$$U_o(\boldsymbol{\rho}') = \sum_{m=1}^{\infty} u_{om}\phi_m(\boldsymbol{\rho}'), \qquad \text{where} \quad u_{om} = \int_{R_2} U_o(\boldsymbol{\rho}')\phi_m^*(\boldsymbol{\rho}')\,d\boldsymbol{\rho}' . \qquad (2.26)$$

Considering (2.25) and (2.26) together with (2.24) leads to the conclusion

$$u_{om} = \eta_m^{1/2}\, u_{im} , \qquad\qquad 1 \leq m < \infty . \qquad (2.27)$$

Equation (2.27), in effect, shows that the channel $h(\boldsymbol{\rho}', \boldsymbol{\rho})$ can be equivalently represented as an infinite set of parallel channels, each of which transforms a particular input eigenfunction to the corresponding output eigenfunction and scales it by the associated singular value. This infinite set of parallel channels, depicted in Figure 2-3, is the normal mode decomposition for R_1–to–R_2 propagation. The atmospheric propagation impulse response is then,

$$h(\boldsymbol{\rho}', \boldsymbol{\rho}) = \sum_{m=1}^{\infty} \sqrt{\eta_m}\, \phi_m(\boldsymbol{\rho}')\, \Phi_m^*(\boldsymbol{\rho}) . \qquad (2.28)$$

Because the atmospheric propagation channel is random, i.e. the impulse response $h(\boldsymbol{\rho}', \boldsymbol{\rho})$ is a random field, the eigenvalues and the associated input and output eigenfunctions will, in general, be random as well. Nevertheless, because the atmosphere

26

is a passive medium, all its eigenvalues $\{\eta_m : 1 \leq m < \infty\}$ must lie in the interval $[0, 1]$. In other words, the received power can be less than the transmitted power due to losses during transmission, but cannot be greater than the transmitted power.

It can be deduced from equation (2.27) that the power transfer properties of the atmospheric channel are characterized by the eigenvalues of the normal mode decomposition. It has been shown [11, 8] that far-field and near-field power transfer regimes can be distinguished by comparing the effective instantaneous Fresnel number

$$
D_f = \int_{R_1} \int_{R_2} \left| h(\boldsymbol{\rho'}, \boldsymbol{\rho}) \right|^2 d\boldsymbol{\rho'} \, d\boldsymbol{\rho} \tag{2.29}
$$

with 1, where $D_f \ll 1$ is the far-field regime and $D_f \gg 1$ is the near-field regime. To gain insight regarding this parameterization and its consequences, let us consider the free-space situation with the circular apertures shown in Figure 2-2, for which the Fresnel number is a non-random value, $D_{fo} = \left(\frac{\pi d_1 d_2}{4\lambda L} \right)^2$. One can show via the Huygens-Fresnel principle that the minimum-area spot produced on the receiver plane $z = L$, by a beam transmitted from aperture R_1 is approximately

$$
\frac{(2\lambda L)^2}{\pi d_1^2} \, . \tag{2.30}
$$

When this minimum spot size, which is obtained by focusing the transmitted field onto the receiver, satisfies

$$
\frac{\left(\frac{\pi d_2^2}{4} \right)}{\left(\frac{(2\lambda L)^2}{\pi d_1^2} \right)} = \left(\frac{\pi d_1 d_2}{4\lambda L} \right)^2 \gg 1 \, , \tag{2.31}
$$

i.e., when the minimum spot size is much less than the receiver aperture diameter, we achieve near-perfect power coupling. Consequently, the condition $D_{fo} \gg 1$ corresponds to *near-field* power transfer. On the other hand, when the minimum spot size

obeys

$$\frac{\left(\frac{\pi d_2^2}{4}\right)}{\left(\frac{(2\lambda L)^2}{\pi d_1^2}\right)} = \left(\frac{\pi d_1 d_2}{4\lambda L}\right)^2 \ll 1 \, , \tag{2.32}$$

it is much larger than the receiver aperture, hence the received power will be only a small fraction of that transmitted. Correspondingly, the condition $D_{f_o} \ll 1$ indicates *far-field* power transfer.

The parameterization of the eigenvalues of turbulent atmosphere in terms of the instantaneous Fresnel number has analogous interpretations, with the added complication that the eigenvalues are random in the presence of turbulence. In particular, if we follow the minimum beam spot size analogy, the minimum spot size on the receiver plane is now dependent on the turbulent state of the atmosphere, thus, in general, the Fresnel number will be a different value in each turbulent state, i.e. it will be a random variable. Nevertheless, it is easily verified from (2.29) that the average value of the effective Fresnel number is D_{f_o}, the free-space Fresnel number. Therefore, the free-space behavior of an atmospheric optical link provides some insight into its performance in the presence of turbulence.

In the atmospheric propagation near-field regime it has been shown that, with high probability, there are D_f eigenvalues which are approximately equal to unity and all other eigenvalues are approximately zero [11]. In the atmospheric propagation far-field regime, there is one eigenvalue that is approximately equal to D_f with high probability, the other eigenvalues all being approximately zero [11]. Moreover, it has been shown that in the far-field regime, if the transmitter diameter is less than the phase coherence length of the medium, the input eigenfunction with non-zero eigenvalue is, with high probability, approximately the free-space eigenfunction. Likewise, in the near field, if the transmitter diameter is less than or equal to the receiver diameter ($d_1 \leq d_2$), the input eigenfunctions for the near-unity eigenvalues are, with high probability, approximately the free-space eigenfunctions.

2.2 Optical Detection

This section develops some basic tools that will be employed in future chapters to analyze the performance of atmospheric optical communication systems. We first develop a statistical description for a photodetector, then consider the output statistics of direct and coherent detection receivers. Throughout this discussion, we assume a quasimonochromatic, linearly polarized optical wave with center frequency ω_o is impinging upon the photosensitive surface of the photodetector and all observations start at time $t = 0$.

2.2.1 Photodetectors

In most general terms, a photodetector is a device that generates an electrical current proportional to the power of the optical wave that illuminates its photo-sensitive surface. There are a number of photodetection technologies available at present, each having their own benefits and shortcomings [14, 15, 16].

Despite the wide range of photodetector types available, it is possible to develop a single statistical model which captures the relevant features of photodetection for communication systems. The model we present here is often referred to as the semi-classical model of photodetection, as it rests upon assumptions that can be rigorously proven with quantum mechanical considerations, yet the quantities in question are classical, such as optical waves and electrical currents.

The conversion from optical power to electrical current in a photodetector is affected by two main factors. First, there are a number of noise sources that will result in fluctuations in the generated output current, and second, even in the absence of noise the conversion from optical power to electrical current does not take place with perfect efficiency.

The fraction of optical power that is converted to electrical power in the absence of any noise is specified by the parameter η_e and is referred to as the quantum efficiency of the detector. The quantum efficiency is an indication of the number of electrons
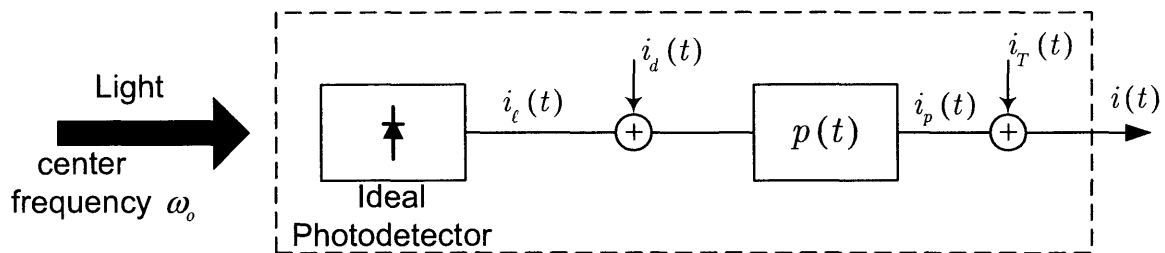
Figure 2-4: Block diagram representation of a photodetector.

emitted per absorbed photon[2] from light impinging on the photosensitive surface. Due to energy conservation, η_e can only take values between 0 and 1.

There are three main noise components affecting the output current of a general photodetector. First, there is shot-noise, which is the randomness in the number of incident photons for a light source with constant intensity. This noise is not due to any external fluctuation or lack of sensitivity of the detector, but a property inherent to light, imposed by quantum mechanics. However, even when all light impinging on the photosensitive area of a photodetector is blocked, there are random charge-carrier generations, referred to as dark current, which constitute another source of noise. Finally, the observable output current of a photodetector is subject to thermal noise from the load resistance through which it flows. The type of photodetector used, in general, will determine the dominant noise component in the output current.

Figure 2-4 is a block diagram representation of a photodetector, which captures the physical attributes relevant to communications analysis. The block referred to as an ideal photodetector in the figure, encompasses the quantum efficiency and shot-noise effects of photodetection [15]. In other words, it is a photodetector with quantum efficiency η_e, infinite optical and electrical bandwidth, and no thermal noise or dark current affecting its output. Therefore, the output current of the ideal detector, $i_\ell(t)$, is a sequence of randomly generated impulses corresponding to the generated electrons due to incoming light. Given $P(t)$, the total short-time-average power impinging upon the sensitive area of the detector, the generation times of these electrons constitute

---

[2]It is well known via quantum mechanical theory that energy from an optical wave at center frequency ω_o can be absorbed only in integer multiples of $\hbar\omega_o$. We will refer to this unit of energy as "a photon" in this discussion.

an inhomogenous conditional Poisson point process with rate function [15, 17]

$$\lambda_s(t) = \eta_e \frac{P(t)}{\hbar\,\omega_o}\,,$$

(2.33)

where $\hbar$ is Planck's constant divided by 2π.

Dark current, represented by $i_d(t)$, is also a sequence of randomly generated impulses, independent of the incoming light. The charge-carrier generation times constitute a homogeneous Poisson point process with rate

$$\lambda_d = \frac{I_d}{q}$$

(2.34)

expressed in terms of the mean dark current I_d and the electron charge q.

Due to electrical and optical bandwidth limitations as well as non-zero delays in current generation, the output current of a real detector will not be in the form of impulses, but rather in the form of current pulses with finite length. The physical attributes of a photodetector which have a 'smearing' effect on impulses due to individual electron releases are collectively modeled as a linear, time-invariant filter with impulse response $p(t)$. The output of this filter, $i_p(t)$, is the current that will be observed in the absence of significant thermal noise, which, in some very sensitive photodetectors is indeed the case.

Finally, thermal noise, $i_T(t)$, is statistically represented as a zero-mean, additive white Gaussian noise (AWGN) process independent of $i_s(t)$ and $i_d(t)$, and with covariance function $\frac{2KT_o}{R_L}\,\delta(t-u)$ [18]. T_o denotes the temperature of the load resistance, R_L is the load resistance seen at the output of the detector, and K is Boltzmann's constant.

It can be concluded from Figure 2-4 that the output current, $i(t)$, has infinite variance because thermal noise is not bandlimited. Therefore, it is necessary to include a low-pass filter after $i(t)$ to bandlimit this noise. However, if post-detection LTI processing is going to take place, it is convenient to assume that the statistics of $i(t)$ represents the output current of the detector, as long as post-detection filters are appropriately band-limiting.
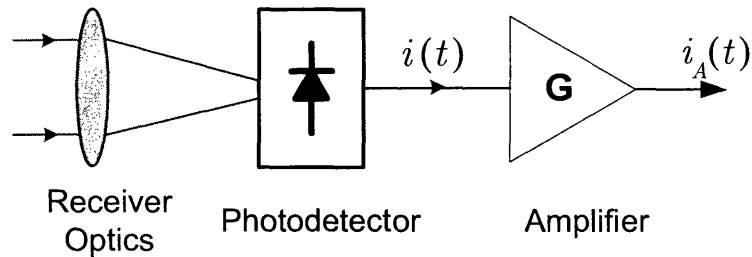
Figure 2-5: Block diagram representation of direct detection.

2.2.2 Direct Detection Receiver

A direct detection receiver, shown in Figure 2-5, is the most straightforward utilization of a photodetector in an optical detection circuit. It consists of optical elements to focus the incoming light onto the photosensitive area of a photodetector and an amplifier to increase the output current of the detector to compatible values with devices used in post-detection processing.

The pulse shape $p(t)$ of a typical photodetector is a short duration pulse, with therefore very large bandwidth. In many practical configurations, the amplifier follow-ing the detector will have much smaller bandwidth than that of $p(t)$, and furthermore, $p(t)$ will have a relatively flat frequency response over the pass-band of the amplifier [19]. Ignoring thermal noise for the moment, the cumulative effect of filtering $i_\ell(t)$ and $i_d(t)$ through $p(t)$ and the amplifier is equivalent to filtering these currents just through the amplifier. Therefore with no appreciable loss of generality, we will replace $p(t)$ with $\delta(t)$.

A further useful simplification in modeling the output current of the amplifier arises from results for filtered Poisson processes. If we assume a LTI filter has an impulse response of finite duration T, when a Poisson process with arrival rate much higher than $1/T$ is passed through the filter, the output process will be approximately a Gaussian random process [17]. Applying this result to the output of the amplifier in a direct detection system, the output $i_A(t)$ can be approximated as a Gaussian random process.

The output of the amplifier is clearly a bandlimited random process. However, in communication theory the eventual goal is to obtain a sufficient statistic for the

optimal decision rule needed to properly decode the message that was transmitted. Therefore, for analytical convenience we assume the output process from the amplifier is a *white* Gaussian process, bearing in mind that the overall effective bandwidth of the system cannot be greater than the bandwidth of the amplifier.

Merging all these approximations and using (2.33), (2.34) and the covariance function for thermal noise, we conclude that the output current from the amplifier is a Gaussian random process with conditional mean,

$$
\begin{aligned}
m_{i_A|\lambda_s}(t) &= E\big[i_A(t)|\lambda_s(t)\big] \\
&= \frac{q\,G\,\eta_e\,P(t)}{\hbar\,\omega_o} + G\,I_d
\end{aligned}
\tag{2.35}
$$

and conditional covariance function

$$
\begin{aligned}
K_{i_A i_A|\lambda_s}(t,u) &= E\left[\left(i_A(t) - m_{i_A|\lambda_s}(t)\right)\left(i_A(u) - m_{i_A|\lambda_s}(u)\right)\,\Big|\,\lambda(\cdot)\right] \\
&= G^2\left\{q^2\left(\lambda_s(t) + \lambda_d\right) + \frac{2KT_o}{R_L}\right\}\delta(t-u) \\
&= \left\{(q\,G)^2\,\eta_e\frac{P(t)}{\hbar\,\omega_o} + q\,G^2 I_d + G^2\frac{2KT_o}{R_L}\right\}\delta(t-u)\,.
\end{aligned}
$$

$$\tag{2.36}$$
$$\tag{2.37}$$

2.2.3 Coherent (Heterodyne) Detection Receivers

Direct-detection receivers simply collect power from the incoming light, convert it to an electronic signal and enhance that signal through electronic amplification. Coherent detection receivers, on the other hand, attempt to enhance the incoming optical wave by mixing it with a strong local oscillator field prior to detection, thereby obtaining an electronic signal proportional to the frequency translated optical field. If the center frequency of the local oscillator field is offset from that of the incoming signal by some intermediate (electronically detectable) frequency, then the detection scheme is called *heterodyne* detection, and if the center frequencies are the same, we refer to the process as *homodyne* detection.

In this section we present a statistical characterization of the output of a hetero-
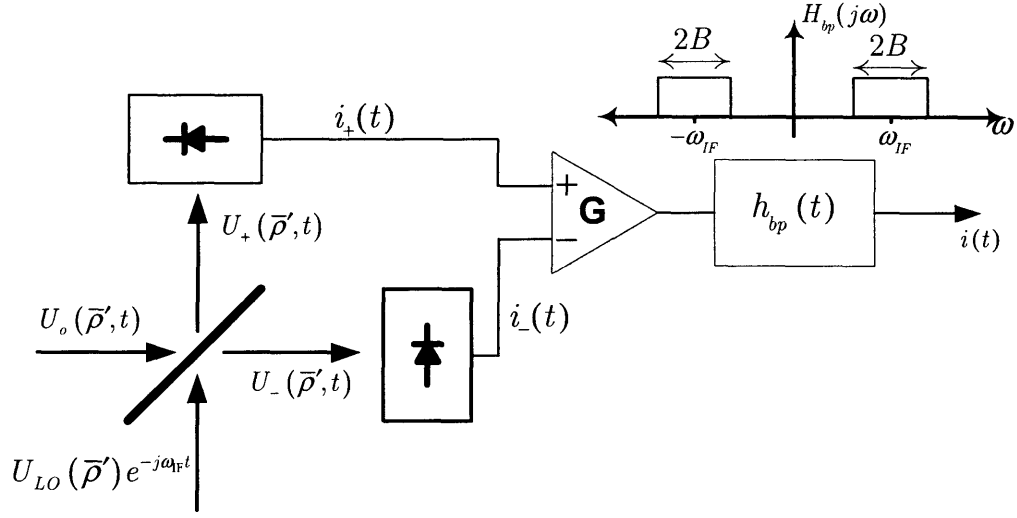
Figure 2-6: Block diagram of a coherent detection receiver.

dyne detection receiver, whose block diagram representation is shown in Figure 2-6. The differential amplifier in the figure is treated as an infinite bandwidth amplifier and a finite bandwidth is imposed explicitly via the LTI filter that performs bandpass filtering around the intermediate frequency, ω_{IF}.

Suppose the incoming optical wave has a normalized scalar complex envelope $U_i(\rho', t)$. We assume the local oscillator generates a monochromatic plane wave polarized linearly in the same direction as the incoming wave, and with center frequency $\omega_{IF} + \omega_o$. Then the normalized scalar complex field envelope of the local oscillator wave is expressed as

$$U_{LO}(\rho', t) = U_{LO}(\rho')e^{-j\omega_{IF} t} .$$ (2.38)

The 50/50 beam splitter equally mixes the two fields and can be modeled as a unitary transformation on the two input fields:

$$U_-(\rho', t) = \frac{1}{\sqrt{2}}\left(U_o(\rho', t) - U_{LO}(\rho')e^{-j\omega_{IF}t}\right) ,$$ (2.39)

$$U_+(\rho', t) = \frac{1}{\sqrt{2}}\left(U_o(\rho', t) + U_{LO}(\rho')e^{-j\omega_{IF}t}\right)$$ (2.40)

as the output field envelopes.

34

The output current of the first photodetector, $i_-(t)$, can be expressed as

$$i_-(t) = m_{i_-}(t) + \Delta i_-(t) , \qquad (2.41)$$

where

$$m_{i_-}(t) = E\left[i_-(t)\right] = \frac{q\,\eta_e}{\hbar\,\omega_o} P_-(t) + I_{d1} \qquad (2.42)$$

and $\Delta i_-(t)$ is a zero-mean random process with conditional covariance function

$$K_{\Delta i_- \Delta i_-}(t, u) = \left[\frac{q^2\,\eta_e}{\hbar\,\omega_o} P_-(t) + qI_{d1} + \frac{2kT_o}{R_L}\right]\delta(t - u) . \qquad (2.43)$$

In the equations above, I_{d1} is the average dark current from the first photodetector and $P_-(t)$ is the short-time average optical power illuminating the detector, i.e.,

$$
\begin{aligned}
P_-(t) &= \int_{R_2} |U_-(\boldsymbol{\rho}', t)|^2 \, d\boldsymbol{\rho}' && (2.44) \\
&= \int_{R_2} \frac{1}{2} \left|U_o(\boldsymbol{\rho}', t) - U_{\mathrm{LO}}(\boldsymbol{\rho}')e^{-j\omega_{\mathrm{IF}}t}\right|^2 \, d\boldsymbol{\rho}' && (2.45) \\
&= \left[\frac{1}{2}P_o(t) + \frac{1}{2}P_{\mathrm{LO}} - \Re\left\{e^{j\omega_{\mathrm{IF}}t}\int_{R_2} U_o(\boldsymbol{\rho}', t)U_{\mathrm{LO}}^*(\boldsymbol{\rho}') \, d\boldsymbol{\rho}'\right\}\right] && (2.46)
\end{aligned}
$$

where $P_o(t) = \int_{R_2} |U_o(\boldsymbol{\rho}', t)|^2 \, d\boldsymbol{\rho}'$ and $P_{\mathrm{LO}} = \int_{R_2} |U_{\mathrm{LO}}(\boldsymbol{\rho}')|^2 \, d\boldsymbol{\rho}'$ are signal and the local oscillator short-time average powers, respectively.

The characterization of the first and second moments of $i_+(t)$ parallels the development for $i_-(t)$. We write the current in terms of its deterministic part and noise as

$$i_+(t) = m_{i_+}(t) + \Delta i_+(t) , \qquad (2.47)$$

where

$$m_{i_+}(t) = E\left[i_+(t)\right] = \frac{q\,\eta_e}{\hbar\,\omega_o} P_+(t) + I_{d2} \qquad (2.48)$$

and $\Delta i_+(t)$ is a zero-mean random process with conditional covariance function,

$$K_{\Delta i_+ \Delta i_+}(t, u) = \left[\frac{q^2\,\eta_e}{\hbar\,\omega_o} P_+(t) + q\,I_{d2} + \frac{2kT_o}{R_L}\right]\delta(t - u) . \qquad (2.49)$$

35

I_{d2} in equations (2.48) and (2.49) is the average dark current from the second photodiode and $P_+(t)$ is the short-time average optical power illuminating this detector, namely

$$P_+(t) = \int_{R_2} |U_+(\boldsymbol{\rho}',t)|^2 \, d\boldsymbol{\rho}' \qquad (2.50)$$

$$= \int_{R_2} \frac{1}{2} |U_o(\boldsymbol{\rho}',t) + U_{\mathrm{LO}}(\boldsymbol{\rho}')e^{-j\omega_{\mathrm{IF}}t}|^2 \, d\boldsymbol{\rho}' \qquad (2.51)$$

$$= \left[\frac{1}{2}P_o(t) + \frac{1}{2}P_{\mathrm{LO}} + \Re\left\{e^{j\omega_{\mathrm{IF}}t}\int_{R_2} U_o(\boldsymbol{\rho}',t)U_{\mathrm{LO}}^*(\boldsymbol{\rho}') \, d\boldsymbol{\rho}'\right\}\right] . \qquad (2.52)$$

The noise terms $\Delta i_-(t)$ and $\Delta i_+(t)$ are due to two physically separate photodetectors. Therefore, given the signal field envelope, $U_o(\boldsymbol{\rho}',t)$, they are independent random processes. The output current $i(t)$ expressed in terms of the currents $i_-(t)$, $i_+(t)$ and the band-pass filter impulse response $h_{bp}(t)$ is

$$i(t) = G\left(i_+(t) - i_-(t)\right) \star h_{bp}(t) = m_i(t) + \Delta i(t) , \qquad (2.53)$$

where $m_i(t)$, the deterministic part of the output current, is

$$m_i(t) = E\left[i(t)\right]$$

$$= G\int \left(m_{i_+}(\tau) - m_{i_-}(\tau)\right) h_{bp}(t - \tau) \, d\tau$$

$$= \frac{2qG\eta_e}{\hbar\omega_o} \Re\left\{\int e^{j\omega_{\mathrm{IF}}\tau} \left[\int_{R_2} U_o(\boldsymbol{\rho}',\tau)U_{\mathrm{LO}}^*(\boldsymbol{\rho}')d\boldsymbol{\rho}'\right] h_{bp}(t - \tau) \, d\tau\right\} \qquad (2.54)$$

and $\Delta i(t)$ is the noise portion of the output current. Heterodyne receivers are operated in the strong local oscillator regime wherein the local oscillator shot-noise contribution is significantly larger than any other noise component, i.e.,

$$\frac{q^2\,\eta_e\,P_{\mathrm{LO}}(t)}{\hbar\,\omega_o} \gg \frac{q^2\eta_e P_o(t)}{\hbar\,\omega_o} , \quad q(I_{d1} + I_{d2}) \text{ and } \frac{4KT_o}{R_L}. \qquad (2.55)$$

Then via filtered Poisson statistics [17], $\Delta i(t)$ is approximately a Gaussian random

$$w(t)$$

$$\sqrt{P_o}\,\mathrm{Re}\Big\{\xi\,s(t)\,e^{jw_{\mathrm{IF}}t}\Big\} \longrightarrow \bigoplus \longrightarrow \boxed{h_{bp}(t)} \longrightarrow r(t)$$

$$H_{bp}(j\omega)$$
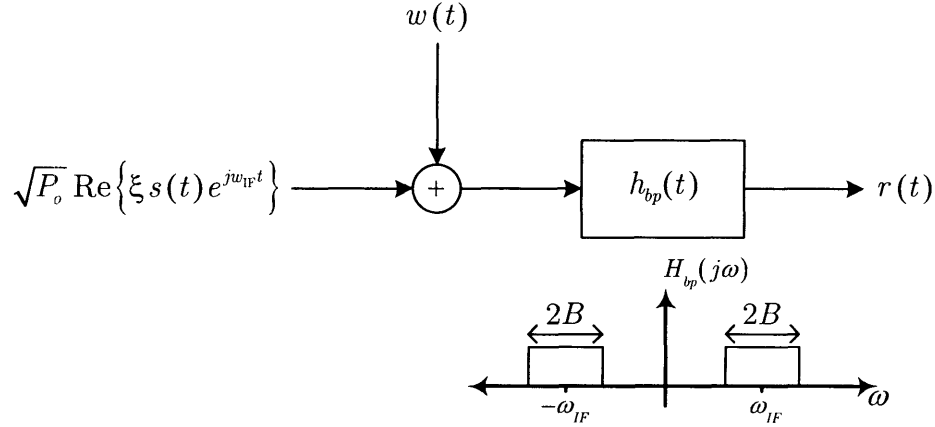
Figure 2-7: An equivalent representation of a coherent detection receiver.

process with zero-mean and covariance function

$$K_{\Delta i \Delta i}(t, u) = \left(\frac{q^2\,G^2\,\eta_e}{\hbar\,\omega_o}\right) P_{\mathrm{LO}} \int h_{bp}(t - \tau)\,h_{bp}(u - \tau)d\tau \ . \qquad (2.56)$$

Let us assume the signal field envelope is separable into space and time functions over the receiver aperture, i.e.

$$U_o(\boldsymbol{\rho'}, t) = \sqrt{P_o}\,s(t)\,\widetilde{U}_o(\boldsymbol{\rho'}) \ , \qquad \boldsymbol{\rho'} \in R_2 \ , \qquad (2.57)$$

where $\int_{R_2} |\widetilde{U}_o(\boldsymbol{\rho'})|^2\,d\boldsymbol{\rho'} = 1$. Furthermore, assume that the local oscillator field envelope has normalized spatial pattern $\widetilde{U}_{\mathrm{LO}}(\boldsymbol{\rho'})$; i.e.,

$$U_{\mathrm{LO}}(\boldsymbol{\rho'}) = \sqrt{P_{\mathrm{LO}}}\,\widetilde{U}_{\mathrm{LO}}(\boldsymbol{\rho'}) \ , \qquad \boldsymbol{\rho'} \in R_2 \ . \qquad (2.58)$$

Then, if we define a normalized output current $r(t)$ as

$$r(t) = \frac{i(t)}{\left(\frac{2q\,G\eta_e\sqrt{P_{\mathrm{LO}}}}{\hbar\omega_o}\right)} \ , \qquad (2.59)$$

it follows directly from (2.54) and (2.56), together with (2.57) and (2.58) that $r(t)$ can be written as

$$r(t) = m_r(t) + \Delta r(t) \ , \qquad (2.60)$$

37

where

$$m_r(t) = \sqrt{P_o}\, \Re\left\{ s(t)e^{j\omega_{\mathrm{IF}}t} \int_{R_2} \widetilde{U}_o\left(\boldsymbol{\rho}'\right) \widetilde{U}_{\mathrm{LO}}^*\left(\boldsymbol{\rho}'\right) d\boldsymbol{\rho}' \right\} \star h_{bp}(t). \qquad (2.61)$$

and $\Delta r(t)$ is a zero-mean Gaussian random process with power spectral density

$$S_{\Delta r \Delta r}(\omega) = \mathcal{F}\Big\{ E\left[\Delta r(t)\,\Delta r(u)\right]\Big\} = \begin{cases} \frac{\hbar\omega_o}{4\eta_e} & , \quad ||\omega| - \omega_{\mathrm{IF}}| < B \\ 0 & , \qquad \text{otherwise .} \end{cases} \qquad (2.62)$$

We note that the integral in (2.61) is independent of time and has magnitude between 0 and 1. We will refer to this number as the spatial correlation coefficient between the input field and the local oscillator field, and denote it with ξ.

Equations (2.61) and (2.62) imply that we can consider a coherent detector as the black-box abstraction shown in Figure 2-7, where $w(t)$ is zero-mean white Gaussian noise with spectral density $\frac{\hbar\omega_o}{4\eta_e}$. However, because our input signal is the complex scalar envelope of an optical wave, it is often even more convenient to work with a complex envelope representation for the normalized output current $r(t)$.

So, by defining a complex envelope $\mathbf{r}(t)$ via

$$r(t) = \Re\left\{ \mathbf{r}(t)e^{j\omega_{\mathrm{IF}}t} \right\} , \qquad (2.63)$$

we can show that $\mathbf{r}(t) = m_{\mathbf{r}}(t) + \Delta\mathbf{r}(t)$, where

$$m_{\mathbf{r}}(t) = \sqrt{P_o}\,\xi\, s(t) \star h(t) \qquad (2.64)$$

is the mean value of $\mathbf{r}(t)$, and $\Delta\mathbf{r}(t)$ is a stationary, zero-mean, complex Gaussian random process whose real and imaginary parts are independent and identically distributed with spectral density

$$S_{\Delta\mathbf{r}\Delta\mathbf{r}^*}(\omega) = \begin{cases} \frac{\hbar\omega_o}{\eta_e} & \text{if } |\omega| < B , \\ 0 & \text{otherwise .} \end{cases} \qquad (2.65)$$

The complex envelope equivalent of a heterodyne receiver is shown in Figure 2-8
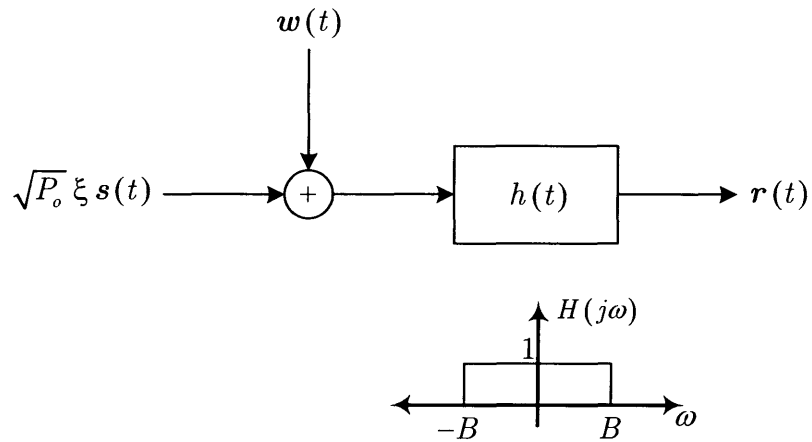
Figure 2-8: Complex representation of a coherent detection receiver.

where $w(t)$ is a zero-mean complex Gaussian random process with spectral density $\frac{\hbar\omega_o}{\eta_e}$. In Chapter 4 we shall assume that the post-detection filter imposes a more stringent bandwidth constraint than $H(j\omega)$ from Figure 2-8. Thus the model we shall use for heterodyne detection will omit the filter $h(t)$ in Figure 2-8.

Chapter 3

Statistical Considerations for Maximum Power Transfer

3.1 Introduction

Using the normal mode decomposition of the extended Huygens-Fresnel principle, the input field transmitted from the R_1 pupil in the $z = 0$ plane can be represented as a weighted sum of the input eigenfunctions:

$$U_i(\boldsymbol{\rho}) = \sum_{m=1}^{\infty} u_{im} \Phi_m(\boldsymbol{\rho}) , \qquad\qquad u_{im} = \int_{R_1} U_i(\boldsymbol{\rho}) \Phi_m^*(\boldsymbol{\rho}) \, d\boldsymbol{\rho} , \qquad (3.1)$$

where $\sum_{m=1}^{\infty} |u_{im}|^2 = P_i$ is the transmitted power. The resulting output field over the R_2 pupil in the $z = L$ plane is then,

$$U_o(\boldsymbol{\rho}') = \sum_{m=1}^{\infty} u_{im} \sqrt{\eta_m} \, \phi_m(\boldsymbol{\rho}') , \qquad (3.2)$$

by virtue of the normal-mode behavior. The output power collected over R_2 is therefore $\sum_{m=1}^{\infty} |u_{im}|^2 \eta_m$. Thus the fractional power transfer from R_1 to R_2 that is achieved by the transmitter field pattern $U_i(\boldsymbol{\rho})$ is

$$\frac{\sum_{m=1}^{\infty} |u_{im}|^2 \eta_m}{\sum_{m=1}^{\infty} |u_{im}|^2} \leq \eta_1 \leq 1 , \qquad (3.3)$$

41

where the first inequality follows from our having arranged the eigenvalues in decreasing order and the second inequality is due to the passive nature of the propagation medium.

It is clear from equations (3.1), (3.2) and (3.3) that the transmitter field $U_i(\boldsymbol{\rho}) = \sqrt{P_i}\,\Phi_1(\boldsymbol{\rho})$ maximizes the R_1–to–R_2 power transfer. For free-space propagation this approach has proven fruitful [20]. In the turbulent atmosphere however, this strategy is undermined by the fact that the channel is random, thus yielding random eigenfunctions and eigenvalues. The extended Huygens-Fresnel principle and its associated statistics have been used to study certain asymptotic characteristics of the maximum eigenvalue and its input and output eigenfunctions [11], but complete results for normal-mode statistics have yet to be obtained.

The goal of this chapter is to develop tight bounds on the first and second moments of the maximum eigenvalue for R_1–to–R_2 propagation through atmospheric turbulence. To simplify some of the results to the point that their numerical evaluation becomes feasible, we replace the circular (2-D) pupils with slit apertures (1-D pupils). As will be seen, we have succeeded in bounding the mean of the maximum eigenvalue, but our bound on the mean-square of this eigenvalue has proven to be too formidable to evaluate.

3.2 One-dimensional Formulation

In Chapter 2 we reviewed free-space and atmospheric propagation for a 3-D medium, i.e., 2-D apertures, plus one propagation dimension. These results reduce almost identically to a two-dimensional propagation medium, i.e., 1-D apertures, plus one propagation dimension. We will only present the fundamentals of the one-dimensional formulation in this section. A more complete compilation is given in Appendix A.

Consider the propagation geometry in Figure 3-1. The extended Huygens-Fresnel principle for turbulent atmospheric propagation in this geometry is

$$U_o(\rho',t) = \int_{R_1} U_i\left(\rho, t - \tfrac{L}{c}\right) \frac{1}{\sqrt{j\lambda L}}\, e^{jk\left(L + \frac{(\rho'-\rho)^2}{2L}\right) + \chi(\rho',\rho) + j\phi(\rho',\rho)}\, d\rho\,, \qquad (3.4)$$

42
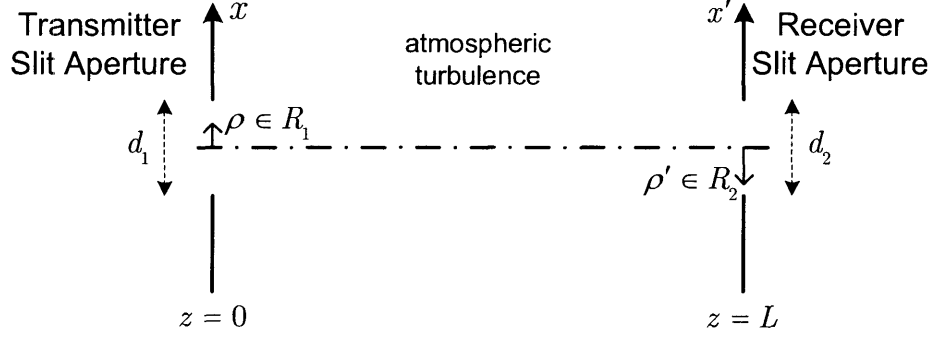
Figure 3-1: Propagation geometry of a two-dimensional propagation medium with one-dimensional apertures.

where $U_i(\rho, t)$ is the input field, $U_o(\rho', t)$ is the output field and the impulse response of the channel is defined as

$$h(\rho', \rho) = \frac{1}{\sqrt{j\lambda L}}\, e^{jk\left(L + \frac{(\rho'-\rho)^2}{2L}\right) + \chi(\rho',\rho) + j\phi(\rho',\rho)} .$$ (3.5)

All terms in (3.4) and (3.5) represent the same quantities as in the two-dimensional formulation. Note however, now ρ and ρ' are scalar quantities thus the integral in (3.4) is one-dimensional, rather than two dimensional. The free-space propagation impulse response is

$$h_{\text{fs}}(\rho', \rho) = \frac{1}{\sqrt{j\lambda L}}\, e^{jk\left(L + \frac{(\rho'-\rho)^2}{2L}\right)} .$$ (3.6)

We begin our analysis by obtaining an expression for the propagation kernel $K(\rho_1, \rho_2)$, introduced in (2.22):

$$K(\rho_1, \rho_2) = \int_{R_2} h^*(\rho', \rho_1)\, h(\rho', \rho_2)\, d\rho'$$

$$= \frac{1}{\lambda L}\, e^{j\frac{k}{2L}\left(\rho_2^2 - \rho_1^2\right)} \int_{-\frac{d_2}{2}}^{\frac{d_2}{2}} e^{-j\frac{k}{L}(\rho_2 - \rho_1)\rho' + \chi(\rho',\rho_2) + \chi(\rho',\rho_1) + j\phi(\rho',\rho_2) - j\phi(\rho',\rho_1)}\, d\rho' .$$ (3.7)

The normal mode decomposition for this kernel is

$$K(\rho_1, \rho_2) = \sum_{m=1}^{\infty} \eta_m\, \Phi_m(\rho_1)\, \Phi_m^*(\rho_2) ,$$ (3.8)

43

where $\{\Phi_m(\rho)\}$ are a CON set on the R_1 slit so that

$$\eta_m = \iint\limits_{R_1\,R_1} \Phi_m^*(\rho_1)\, K\left(\rho_1, \rho_2\right) \Phi_m(\rho_2)\, d\rho_2\, d\rho_1 \; . \tag{3.9}$$

It should be clear from (3.9) that the quadratic phase factors preceding the integral in (3.7) can be included in the eigenfunction definitions as $\Phi_m(\rho) = \widetilde{\Phi}_m(\rho)\, e^{-j\frac{k}{2L}|\rho|^2}$ so that we need only seek the eigenvalue/eigenfunction structure $\{\eta_m, \widetilde{\Phi}_m(\rho)\}$ of the reduced kernel

$$\widetilde{K}\left(\rho_1, \rho_2\right) = \frac{1}{\lambda L} \int_{-\frac{d_2}{2}}^{\frac{d_2}{2}} e^{-j\frac{k}{L}(\rho_2 - \rho_1)\rho'} e^{\chi(\rho',\rho_2)+\chi(\rho',\rho_1)+j\phi(\rho',\rho_2)-j\phi(\rho',\rho_1)}\, d\rho' \; . \tag{3.10}$$

We adopt this modified kernel expression henceforth, but we will drop the tilde superscripts to avoid notation clutter.

Ideally, one would like to obtain the probability distribution function for the maximum eigenvalue, i.e. $p(\eta_1)$. However, this is not feasible in the near-field regime. Instead, we will attempt to characterize a tight lower bound on the average of the maximum eigenvalue. Let $U_i(\rho)$ denote the complex scalar envelope of some input field with unity power, and let $U_o(\rho')$ be the corresponding output field obtained via (3.4)[1]. The output power collected at the receiver aperture is

$$\gamma = \int_{R_2} \left|U_o(\rho')\right|^2 d\rho' \tag{3.11}$$

$$= \iint\limits_{R_1\,R_1} U_i^*(\rho_1)\, K\left(\rho_1, \rho_2\right) U_i(\rho_2)\, d\rho_2\, d\rho_1 \; . \tag{3.12}$$

Note that γ is a random variable. By means of equation (3.3), (3.12) satisfies

$$\gamma \leq \eta_1 \; , \tag{3.13}$$

with equality if $U_i(\rho) = \Phi_1(\rho)$, i.e., the transmitter uses the maximum-eigenvalue

---

[1]The time dependence of the input and output fields are suppressed to focus on the turbulence-induced spatial properties.

input eigenfunction as its field pattern. It follows that the statistical average of γ is also less than the average of η_1:

$$\overline{\gamma} \le \overline{\eta} , \tag{3.14}$$

where we have also dropped the subscript for the maximum eigenvalue, because in all that follows our focus will be restricted to the maximum eigenvalue. Thus, by choosing different input fields of unity power, we can obtain various lower bounds on the mean value of the maximum eigenvalue.

3.3 Deterministic Input Fields

If the input field is restricted to deterministic functions, then via (3.12), $\overline{\gamma}$ can be expressed as

$$\overline{\gamma} = \iint\limits_{R_1 \, R_1} U_i^*(\rho_1) \, \overline{K(\rho_1, \rho_2)} \, U_i(\rho_2) \, d\rho_2 \, d\rho_1 , \tag{3.15}$$

where, with the help of the atmospheric mutual coherence function,

$$\overline{K(\rho_1, \rho_2)} = K_{\text{fs}}(\rho_1, \rho_2) \, e^{-\frac{1}{2} D(0, \rho_2 - \rho_1)} \tag{3.16}$$

is the average kernel, viz., the ensemble average of the instantaneous kernel over all possible turbulence states. In equation (3.16), $K_{\text{fs}}(\rho_1, \rho_2)$ is the free-space propagation kernel,

$$K_{\text{fs}}(\rho_1, \rho_2) = \frac{d_2}{\lambda L} \frac{\sin\left(\frac{\pi d_2}{\lambda L} |\rho_2 - \rho_1|\right)}{\frac{\pi d_2}{\lambda L} |\rho_2 - \rho_1|} , \qquad \rho_1, \rho_2 \in R_1 , \tag{3.17}$$

and $D(0, \rho_2 - \rho_1)$ is the spherical-wave, wave structure function introduced in (2.21)[2].

We see from (3.15) and the normal mode decomposition of $\overline{K(\rho_1, \rho_2)}$ that the highest achievable power transfer by a deterministic input field is the maximum eigenvalue of the average kernel. Furthermore, through (3.14), the maximum eigenvalue of

---

[2]See (A.7) for the 1-D equivalent.
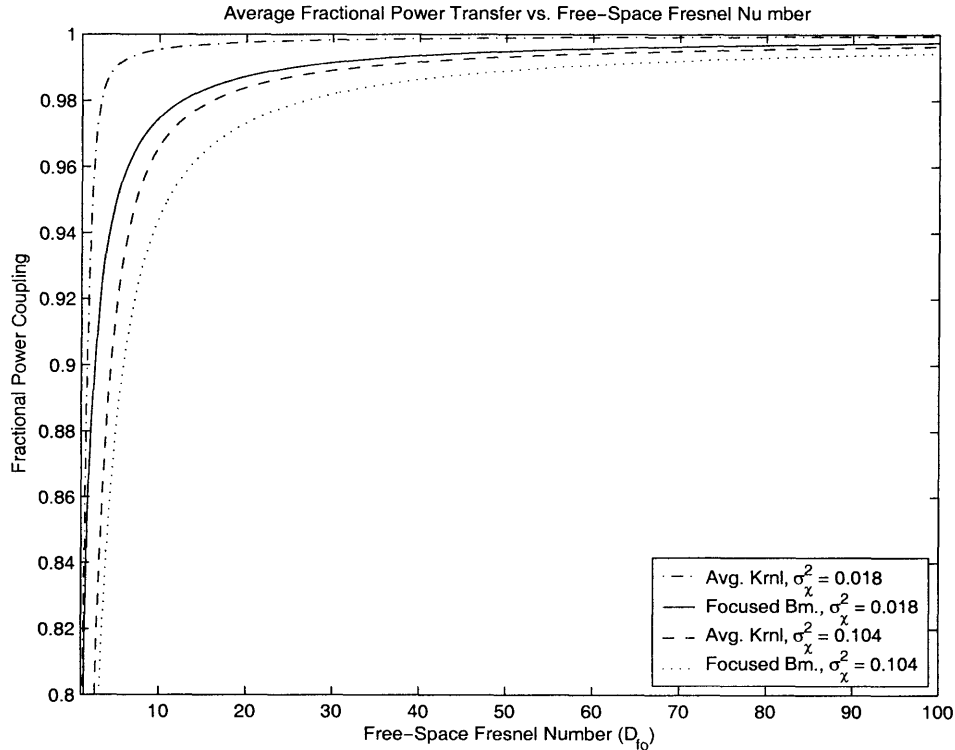
Figure 3-2: Maximum eigenvalue of the average kernel and mean power coupling of a focused beam in mild and strong turbulence conditions.

the average kernel is also a lower bound on the average maximum eigenvalue of the turbulent-atmosphere kernel [11, 8].

The maximum eigenvalue of $\overline{K\left(\rho_1, \rho_2\right)}$ for various turbulence strengths is shown in Figure 3-2. It has been shown in [11] that the worst-case power coupling, for a given Fresnel number, occurs when the apertures are equal size, an assumption adopted for the plots as well. As one expects, the maximum eigenvalue of the average kernel rapidly approaches 1 as the free-space Fresnel number is increased. On the other hand, as turbulence strength increases there is a reduction in the power coupling due to increased levels of turbulence-induced scattering.

Focused Beam

In free-space, the eigenfunctions of the normal mode decomposition are found to be prolate spheroidal wave functions [20]. However, in the limit of $D_{fo} \gg 1$ and $D_{fo} \ll 1$ the maximum-eigenvalue input eigenfunction becomes approximately a uniform-

46

intensity focused beam. It has also been shown that the maximum-eigenvalue input eigenfunction of the atmospheric propagation kernel converges asymptotically to a focused beam [11]. Therefore, one would expect that a focused beam input field would yield a tight bound on the average of the maximum eigenvalue. Notice, however, because this field is deterministic, it cannot yield a tighter bound than the maximum eigenvalue of the average kernel.

A focused beam input field with unity power is expressed as,

$$U_i(\rho) = \sqrt{\tfrac{1}{d_1}}\, e^{-j\frac{k}{2L}\rho^2}\,, \qquad \rho \in \left[-\frac{d_1}{2}, \frac{d_1}{2}\right]. \tag{3.18}$$

By substituting this expression in (3.15), we obtain,

$$\overline{\gamma} = D_{fo} \int_{-1}^{1} e^{-\frac{1}{2}D(x)} \left(1 - |x|\right) \frac{\sin\left(\pi\, D_{fo}\, x\right)}{\pi\, D_{fo}\, x}\, dx \tag{3.19}$$

$$= 2D_{fo} \int_{0}^{1} e^{-\frac{1}{2}D(x)} \left(1 - x\right) \frac{\sin\left(\pi\, D_{fo}\, x\right)}{\pi\, D_{fo}\, x}\, dx \tag{3.20}$$

where $D(x) = 75.952\,\sigma_\chi^2\, D_{fo}^{5/6}\, |x|^{5/3}$ is the spherical-wave wave structure function for the one-dimensional case with equal aperture sizes ($d_1 = d_2 = d_o$) and a uniform C_n^2 profile, written in terms of the normalized (dimensionless) coordinate[3] $x = \frac{\rho'}{d_o}$. From (3.20) it can be deduced that the average power coupling for the focused beam depends on two quantities: the variance of the spherical-wave log-amplitude fluctuations σ_χ^2 and the free-space Fresnel number D_{fo}.

Figure 3-2 compares the average focused beam power transfer through turbulence to the maximum eigenvalue of the average kernel. The plots are versus the free-space Fresnel numbers corresponding to the near-field regime. It is seen that the power coupling is indeed asymptotically close to the average kernel maximum eigenvalue, at large values of D_{fo}. Furthermore, increased strength of atmospheric turbulence, as expected, reduces average power transfer of both the focused beam and the maximum-eigenvalue eigenfunction of the average kernel.

---

[3]See (A.10).

3.4 Tighter Bounds via Adaptive Optics

In an effort to improve the overall performance of optical atmospheric communication systems, adaptive correction methods can be implemented to compensate for the phase and amplitude fluctuations in an optical wave propagating through the turbulent atmosphere.

Early work on adaptive optics for communication systems mostly concentrated on eliminating turbulence-induced phase distortions [5, 7]. The main principle underlying most phase-compensated systems is phase conjugation: overcoming the deleterious effects of phase fluctuations by applying the conjugate phase to the optical field. Such compensation systems can be implemented at the transmitter end to overcome beam spread, as well as at the receiver end to overcome angular spread. Recent work has also considered intensity correction for imaging and optical atmospheric communication applications [7, 21].

In the previous section, we limited the input to be a deterministic field and deduced that the maximum average power coupling between the transmitter and receiver apertures is the maximum eigenvalue of the average kernel, $\overline{K(\rho_1, \rho_2)}$, which is less than the average of the maximum eigenvalue of $K(\rho_1, \rho_2)$. This suggests that we could use adaptive techniques to construct an input field which is pre-distorted such that at least some of the deleterious effects of the turbulent atmosphere are compensated for in a manner that yields higher power transfer on average. It should be emphasized that via (3.12), it is still the case that an input field constructed using some adaptive optics technique will, in general, yield a lower bound on the maximum eigenvalue of the atmospheric propagation kernel, $K(\rho_1, \rho_2)$. We will begin by considering a particular adaptive scheme and derive its average power transfer. Afterwards, we will briefly consider a hypothetical limiting case, when we assume all phase fluctuations can be removed, leaving only log-amplitude fluctuations. We do not consider the feasibility of realizing our adaptive schemes with existing technologies. However, should there be a substantial performance gap between the power coupling predicted by these theoretical studies and those achievable with existing equipment, then investing more

effort in developing technologies to bridge such a gap might well be warranted.

3.4.1   Maximizing On-Axis Power Density

Consider a unity-power focused beam with additional phase compensation that depends on the turbulent state of the atmosphere; i.e.,

$$U_i(\boldsymbol{\rho}) = \sqrt{\frac{4}{\pi \, d_1^2}} \, e^{-j\frac{k}{2L}|\boldsymbol{\rho}|^2 - j\hat{\phi}(\boldsymbol{\rho})} \, , \qquad \boldsymbol{\rho} \in R_1 \, , \qquad (3.21)$$

where $\hat{\phi}(\boldsymbol{\rho})$ is the phase field imposed at the transmitter. This additional phase term pre-distorts the transmitted beam such that the deleterious effects of the atmospheric Green's function on power transfer are minimized. Note that implementing this scheme requires the ability to estimate those aspects of the atmospheric state that are relevant to determining the appropriate phase compensation dynamically; i.e. the estimation process must be able to track the changes in the turbulent state of the atmosphere, which occur on msec time scales [5]. In addition, the transmitter must be equipped with the technology necessary to impose the determined phase compensation in real-time.

To get the tightest lower bound on $\bar{\eta}$, we should choose $\hat{\phi}(\boldsymbol{\rho})$ to maximize the instantaneous R_1-to-R_2 power transfer and then average that instantaneous power transfer over the atmospheric statistics. Unfortunately, that approach is impossible to follow. So, instead, we will choose $\hat{\phi}(\boldsymbol{\rho})$ to maximize the instantaneous on-axis power density in the $z = L$ plane. In particular, for the full 3-D propagation geometry, the

49

output field intensity at $\boldsymbol{\rho'} = \mathbf{0}$ is

$$\left| U_o(\boldsymbol{\rho'} = \mathbf{0}) \right|^2 = \left| \int_{R_1} U_i(\boldsymbol{\rho})\, h(\mathbf{0}, \boldsymbol{\rho})\, d\boldsymbol{\rho} \right|^2 \tag{3.22}$$

$$= \frac{4}{\pi\, d_1^2\, (\lambda L)^2} \left| \int_{R_1} e^{-j\frac{k}{2L}|\boldsymbol{\rho}|^2 - j\hat{\phi}(\boldsymbol{\rho})}\, e^{jk\left(L + \frac{|\boldsymbol{\rho}|^2}{2L}\right)}\, e^{\chi(\mathbf{0},\boldsymbol{\rho}) + j\phi(\mathbf{0},\boldsymbol{\rho})}\, d\boldsymbol{\rho} \right|^2 \tag{3.23}$$

$$\leq \frac{4}{\pi\, d_1^2\, (\lambda L)^2} \left[\int_{R_1} \left| e^{\chi(\mathbf{0},\boldsymbol{\rho}) + j\phi(\mathbf{0},\boldsymbol{\rho}) - j\hat{\phi}(\boldsymbol{\rho})} \right| d\boldsymbol{\rho} \right]^2 \tag{3.24}$$

$$= \frac{4}{\pi\, d_1^2\, (\lambda L)^2} \left[\int_{R_1} e^{\chi(\mathbf{0},\boldsymbol{\rho})}\, d\boldsymbol{\rho} \right]^2 \tag{3.25}$$

with equality if, and only if, $\hat{\phi}(\boldsymbol{\rho}) = \phi(\mathbf{0},\boldsymbol{\rho})$, $\forall \boldsymbol{\rho} \in R_1$. Therefore the input field of the class (3.21) that maximizes the power density at the center of the receiver aperture is

$$U_i(\boldsymbol{\rho}) = \sqrt{\frac{4}{\pi d_1^2}}\, e^{-j\frac{k}{2L}|\boldsymbol{\rho}|^2 - j\phi(\mathbf{0},\boldsymbol{\rho})}\,, \qquad \boldsymbol{\rho} \in R_1\,. \tag{3.26}$$

For the one-dimensional aperture (two-dimensional propagation medium) case the equivalent of (3.26) is

$$U_i(\boldsymbol{\rho}) = \frac{1}{\sqrt{d_1}}\, e^{-j\frac{k}{2L}|\rho|^2 - j\phi(\mathbf{0},\boldsymbol{\rho})}\,, \qquad \rho \in R_1\,. \tag{3.27}$$

The resulting instantaneous power transfer γ for this phase-compensated focused beam is thus

$$\gamma = \int_{R_2} \left| U_o(\rho) \right|^2 d\rho' \tag{3.28}$$

$$= \frac{1}{d_1 \lambda L} \iiint_{R_1\, R_1\, R_2} e^{-j\frac{k}{L}(\rho_2 - \rho_1)\rho'}\, e^{X(\rho',\rho_1,\rho_2)}\, d\rho'\, d\rho_2\, d\rho_1\,, \tag{3.29}$$

50

where

$$X\left(\rho', \rho_1, \rho_2\right) = \chi(\rho', \rho_2) + \chi(\rho', \rho_1) + j\phi(\rho', \rho_2)$$
$$- j\phi(\rho', \rho_1) - j\phi(0, \rho_2) + j\phi(0, \rho_1) \ . \quad (3.30)$$

The fields in (3.30) are jointly Gaussian, from which it can be shown, using $e^{\overline{X}} = e^{E[X] + \frac{\mathrm{var}(X)}{2}}$,

$$\overline{e^{X(\rho', \rho_1, \rho_2)}} = \exp\Big\{ -\tfrac{1}{2}D_{\chi,\chi}(0, \rho_1 - \rho_2) - D_{\phi,\phi}(0, \rho_1 - \rho_2) - D_{\phi,\phi}(\rho', 0)$$
$$+ \tfrac{1}{2}D_{\phi,\phi}(\rho', \rho_2 - \rho_1) + \tfrac{1}{2}D_{\phi,\phi}(\rho', \rho_1 - \rho_2)$$
$$+ jK_{\chi,\phi}(\rho', \rho_2 - \rho_1) - jK_{\chi,\phi}(\rho', \rho_1 - \rho_2)\Big\} \ , \quad (3.31)$$

where $D_{\chi,\chi}(\rho', \rho)$ and $D_{\phi,\phi}(\rho', \rho)$ are the log-amplitude and phase structure functions, and $K_{\chi,\phi}(\rho', \rho)$ is their cross-covariance function.

For brevity, in the next few steps of our analysis, we will use $f(\rho', \rho_1 - \rho_2)$ to denote the exponent on the the right-hand side of (3.31). Assume the transmitter and receiver apertures are equal length, d_o. The mean value of the power transfer is then,

$$\overline{\gamma} = \int\limits_{-\frac{d_o}{2}}^{\frac{d_o}{2}} \int\limits_{-\frac{d_o}{2}}^{\frac{d_o}{2}} \int\limits_{-\frac{d_o}{2}}^{\frac{d_o}{2}} e^{-\frac{jk}{L}(\rho_2 - \rho_1)\rho'} e^{f(\rho', \rho_1, \rho_2)} \, d\rho' \, d\rho_2 \, d\rho_1 \ . \quad (3.32)$$

By change of coordinates from ρ_1, ρ_2 to $\rho_d = \rho_2 - \rho_1$, $\rho_c = \frac{\rho_1 + \rho_2}{2}$, the integral over ρ_c can be done analytically. Converting to dimensionless coordinates $x' = \frac{\rho'}{d_o}$, $x = \frac{\rho_c}{d_o}$, we then obtain

$$\overline{\gamma} = 4D_{fo} \int_0^1 \int_0^{\frac{1}{2}} \cos\Big(2\pi D_{fo}\, x'\, x - f_2(x', x)\Big)(1 - x)e^{-f_1(x', x) + f_3(x', x)} \, dx' \, dx \ , \quad (3.33)$$

51

where,

$$f_1(x', x) = \begin{cases} 0 & x = x' = 0 \\ 75.952\, \sigma_\chi^2\, D_{fo}^{5/6} \left(\frac{(x'x)^{5/3}}{(x'+x)(x'^{2/3}+(x'x)^{1/3}+x^{2/3})} \right) & \text{otherwise}, \end{cases}$$

(3.34)

$$f_2(x', x) = K_{\chi,\phi}(d_o\, x', d_o\, x) - K_{\chi,\phi}(d_o\, x', -d_o\, x)$$

(3.35)

and

$$f_3(x', x) = \tfrac{1}{2} D_{\chi,\chi}(0, d_o\, x) + D_{\chi,\chi}(d_o\, x', 0) - \tfrac{1}{2} D_{\chi,\chi}(d_o\, x', d_o\, x) - \tfrac{1}{2} D_{\chi,\chi}(d_o\, x', -d_o\, x).$$

(3.36)

We have performed the double integral in (3.33) numerically. The result for mild turbulence strength is plotted in Figure 3-3 along with the focused beam result. For ease of comparison, the plots show $1 - \overline{\gamma}$ on a semi-logarithmic scale. Although the average fractional power transfer for the given turbulence strength does not differ significantly for small free-space Fresnel numbers, as the aperture sizes are increased, the difference between the average power coupling of the two schemes grow larger. Consequently, for moderate values of D_{fo}, the phase-compensated focused beam field predicts a tighter lower bound on the mean of the maxiumum eigenvalue. Note however, that the average kernel maximum eigenvalue achieves higher average power coupling than the phase-compensated focused beam for the free-space Fresnel numbers shown in the plot.

The benefits of adaptive optics become more apparent in strong turbulence, for which the average power coupling results are plot in Figure 3-4. Although, the increased turbulence strength has reduced the overall average power coupling for all beams, the phase-compensated beam has been reduced the least, now yielding higher average power coupling than the maximum eigenvalue of the average kernel. It is also clear from the plot that in strong turbulence, the average power coupling of the the phase-compensated focused beam is distinctly higher than the average coupling given by a focused beam with no adaptive optics, throughout the near field regime.
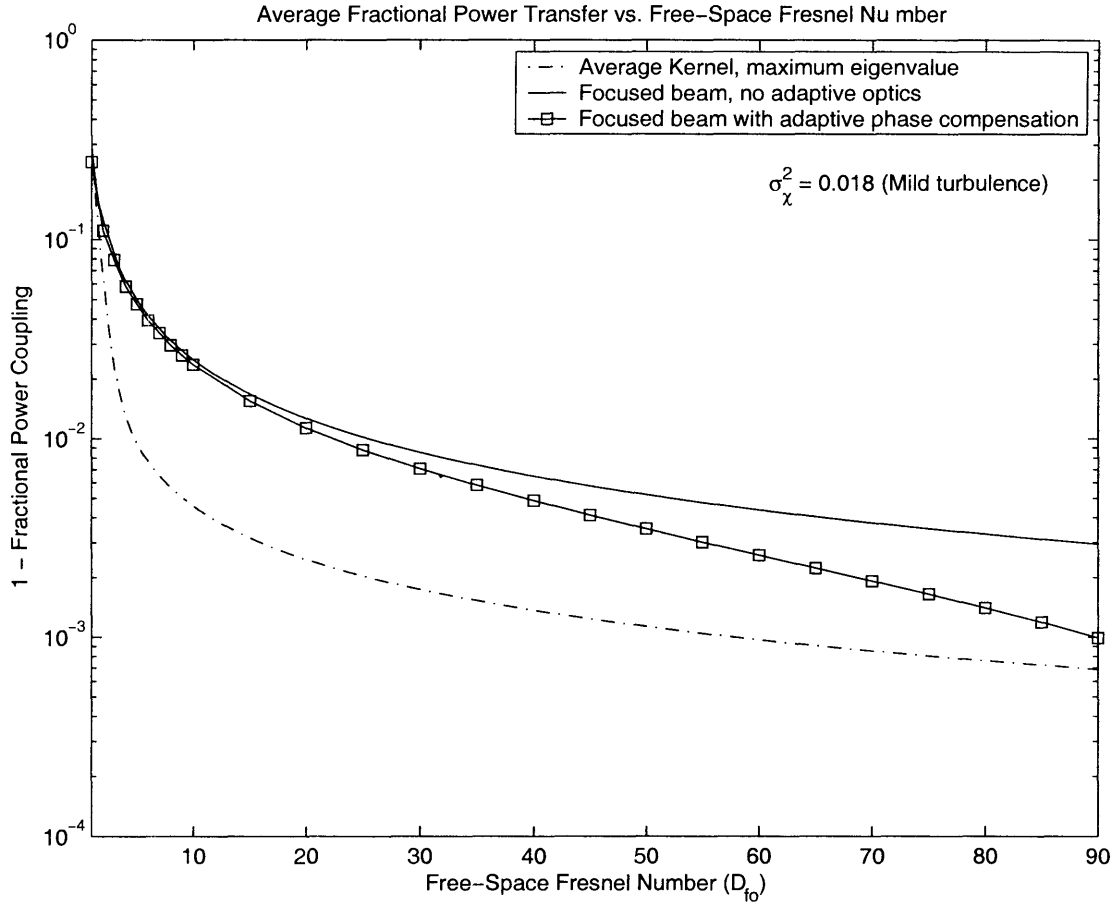
Figure 3-3: Average power transfer of a focused beam in mild turbulence conditions, with and without phase compensation. Dashed line shows the maximum eigenvalue of the average kernel.

3.4.2 Perfect Phase Cancellation

Assume a hypothetical scenario in which the turbulence-induced phase fluctuation in the atmospheric Green's function satisfies,

$$\phi(\rho', \rho) = \phi_1(\rho) + \phi_2(\rho') . \tag{3.37}$$

The atmospheric propagation kernel (with the quadratic phase factor suppressed) then becomes,

$$K\left(\rho_1, \rho_2\right) = \frac{1}{\lambda L} e^{j\phi_1(\rho_2) - j\phi_1(\rho_1)} \int_{R_2} e^{-j\frac{k}{L}(\rho_2 - \rho_1)\rho'} e^{\chi(\rho', \rho_1) + \chi(\rho', \rho_2)} \, d\rho' . \tag{3.38}$$

53
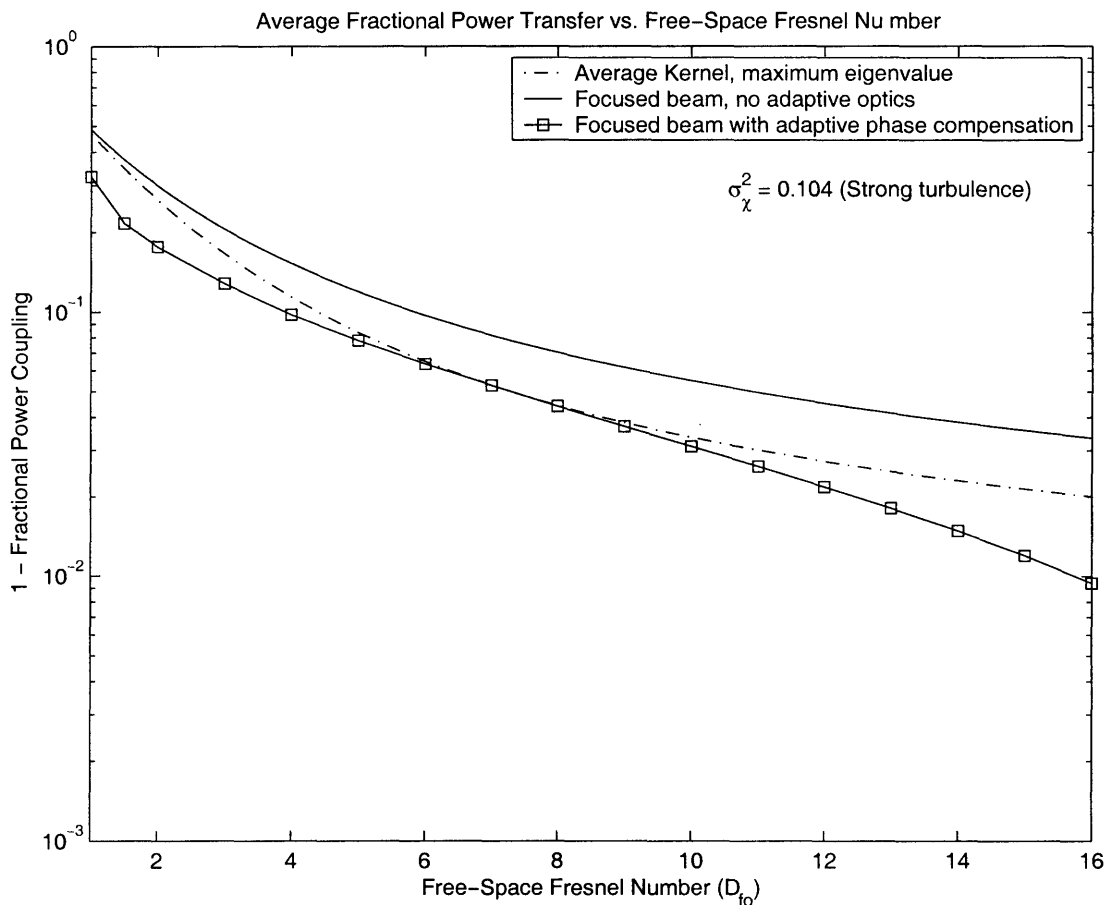
Figure 3-4: Average power transfer of a focused beam in strong turbulence conditions, with and without phase compensation. The maximum eigenvalue of the average kernel is also shown as a dashed line.

The eigenfunctions of this kernel, $\{\Phi_m(\rho) : 1 \leq m \leq \infty\}$, satisfy

$$\Phi_m(\rho) = \widetilde{\Phi}_m(\rho)e^{-j\phi_1(\rho)} \,, \tag{3.39}$$

where $\{\widetilde{\Phi}_m(\rho)\}$ are the eigenfunctions of the reduced kernel

$$\widetilde{K}(\rho_1, \rho_2) = \frac{1}{\lambda L}\int_{R_2} e^{-j\frac{k}{L}(\rho_2-\rho_1)\rho'}e^{\chi(\rho',\rho_1)+\chi(\rho',\rho_2)}\,d\rho' \,. \tag{3.40}$$

Therefore the maximum instantaneous power coupling is the maximum eigenvalue of the kernel with only log-amplitude fluctuations, i.e. $\widetilde{K}(\rho_1, \rho_2)$. On the other hand, if adaptive optics are employed to construct an input field of the form $U_D(\rho)e^{-j\phi_1(\rho)}$,

where $U_D(\rho)$ is deterministic, the highest average power coupling will be the maximum eigenvalue of the log-amplitude only average kernel

$$E\left[\widetilde{K}(\rho_1, \rho_2)\right] = e^{-\frac{1}{2}D_{\chi,\chi}(0,\rho_2-\rho_1)}K_{\text{fs}}(\rho_1, \rho_2) \ , \tag{3.41}$$

and will be achieved when $U_D(\rho)$ is chosen to be the maximum-eigenvalue eigenfunction of this average kernel.

The adaptive scheme presented in section 3.4.1, would estimate a phase compensation field

$$\hat{\phi}(\boldsymbol{\rho}) = \phi_1(\rho) + \phi_2(0); \tag{3.42}$$

and therefore would correspond to an input field where $U_D(\rho)$ is chosen to be a focused beam (with some absolute phase, $-\phi_2(0)$). Note that, because a focused beam is not an eigenfunction of (3.41), the adaptive scheme will not, in general, yield the maximum achievable average power coupling with input fields in the class (3.21). For an equal aperture geometry with slit width d_o, the average power coupling achieved by this focused beam is

$$\overline{\gamma} = \iint\limits_{R_1\,R_1} \overline{U_i^*(\rho_1)\,K(\rho_1, \rho_2)\,U_i(\rho_2)}\,d\rho_2\,d\rho_1 \tag{3.43}$$

$$= \frac{1}{\lambda L} \iiint\limits_{R_1\,R_1\,R_2} e^{-j\frac{k}{L}(\rho_2-\rho_1)\rho'}\,\overline{e^{\chi(\rho',\rho_1)+\chi(\rho',\rho_2)}}\,d\rho'\,d\rho_2\,d\rho_1 \tag{3.44}$$

$$= 2D_{fo} \int_0^1 e^{-\frac{1}{2}D_{\chi,\chi}(0,d_o\,x)}\,(1 - x)\,\frac{\sin\left(\pi\,D_{fo}\,x\right)}{\pi\,D_{fo}\,x}\,dx \ , \tag{3.45}$$

in terms of the normalized difference coordinate $x = \frac{\rho_2-\rho_1}{d_o}$.

Figure 3-7 plots all the average fractional power transfer results discussed until now, along with some free-space results, for mild turbulence strengths. It is clearly seen that the complete elimination of phase, combined with the maximum-eigenvalue eigenfunction yields the highest average power coupling. On the other hand, in the absence of complete phase elimination, the maximum-eigenvalue eigenfunction of the average kernel yields the highest average power coupling. The focused beam propa-

gating in free-space achieves fractional power transfer,

$$\overline{\gamma}_{\text{fs}} = 2D_{fo} \int_0^1 \left(1 - x\right) \frac{\sin\left(\pi D_{fo}x\right)}{\pi D_{fo}x} \, dx \ , \tag{3.46}$$

which is seen from the figure to be slightly higher than its average power coupling in the presence of only log-amplitude fluctuations. This in turn, is slightly greater than the average power coupling of the focused beam propagating in turbulent atmosphere. Nevertheless, the results appear very close to each other. The convergence of the focused beam results for limiting cases of D_{fo} can be verified by analyzing the behavior of the integrands in equations (3.20), (3.45) and (3.46). All three integrands consist of a weighting factor $1 - x$, a unity area sinc function expression $D_{fo}\frac{\sin(\pi D_{fo}x)}{\pi D_{fo}x}$ and, in the presence of turbulence, a decaying exponential term. The sinc function converges to an impulse at $x = 0$ as D_{fo} diverges to infinity, asymptotically yielding perfect power coupling on average for all three cases. However, in the presence of turbulence, the convergence of the sinc function competes against the decaying exponential term, which has a decay rate proportional to $\sqrt{D_{fo}}$ for full turbulence statistics, and a much smaller decay rate when only log-amplitude fluctuations are present[4]. This exponential term reduces the overall area underneath the integrand (which is mostly located underneath the main lobe of the sinc function), causing the slight reduction in average power transfer for intermediate values of D_{fo}. When D_{fo} is very small (yet greater than 1), the dominant effect in the integrand is the $1 - x$ weighting factor, thus all three results yield similar average power coupling in this regime. Note that the sinc function and decaying exponential terms are also present in the average kernel expressions of the turbulent atmosphere and the log-amplitude-only case. Thus as $D_{fo} \to \infty$, the maximum eigenvalue of these kernels must also tend to 1. However, comparing the eigenvalue behavior in these three cases from the plots, it is seen that the exponential term reduces the maximum eigenvalue of the average kernel to a much greater extent, than the average power transfer of the focused beam.

---

[4]The width of the main lobe of the sinc function expression is $\frac{2}{D_{fo}}$, thus the sinc function converges to an impulse faster than the exponential term decays to zero almost everywhere. The limiting behavior as $D_{fo} \to \infty$ is therefore determined by the impulsive behavior of the sinc function and not by the exponential decaying to zero.

Although the phase-compensated focused beam yields average fractional power transfer results similar to that of other focused beams, the plot in the figure shows that it follows a different slope in approaching 1 as D_{fo} is increased. To establish better insight as to why this occurs, we rewrite the average power transfer integral expression in (3.33) as

$$\bar{\gamma} = 2D_{fo} \int_0^1 (1-x)\, e^{-\frac{1}{2}D_{\chi,\chi}(0,\, d_o x)}$$
$$\left[\int_0^{\frac{1}{2}} 2\cos\Big(2\pi D_{fo}\, x'\, x - f_2(x',x)\Big) e^{\tilde{f}_1(x',x)}\, dx' \right] dx\ , \quad (3.47)$$

where

$$\tilde{f}_1(x',x) = -D_{\phi,\phi}(0, d_o\, x) - D_{\phi,\phi}(d_o\, x', 0) + \tfrac{1}{2}D_{\phi,\phi}(d_o\, x', d_o\, x) + \tfrac{1}{2}D_{\phi,\phi}(d_o\, x', -d_o\, x)\ ,$$
$$(3.48)$$

and $f_2(x',x)$ is given in (3.35). Comparing this expression to the average power transfer expressions for the focused beam, we see that the integral in brackets has replaced the sinc function expression in (3.45). Note that, in the absence of turbulence, f_2 and $\tilde{f}_1$ are zero, and the integral reduces to a sinc function. Figures 3-5 and 3-6 compare this function to the corresponding sinc function at a variety of D_{fo} values and turbulence strengths. It can be seen from these plots that the adaptive optics term has a slightly different main lobe and appreciably smaller side lobes. Therefore, although the average power coupling for large values of D_{fo} is still determined by the convergence of this term towards an impulse, the convergence rate or pattern need not be similar to that of a sinc function.

Strong turbulence results for the input fields discussed in this chapter are illustrated in Figure 3-8. Note that all average power coupling results have suffered, when compared to the mild turbulence results. However, the relative changes due to increased turbulence strength are quite different. The focused beam and maximum-eigenvalue eigenfunction of the average kernel suffer the most attenuation. The complete elimination of phase slightly improves power coupling of a focused beam, however both perform worse than a focused beam propagating in free-space. For the

57
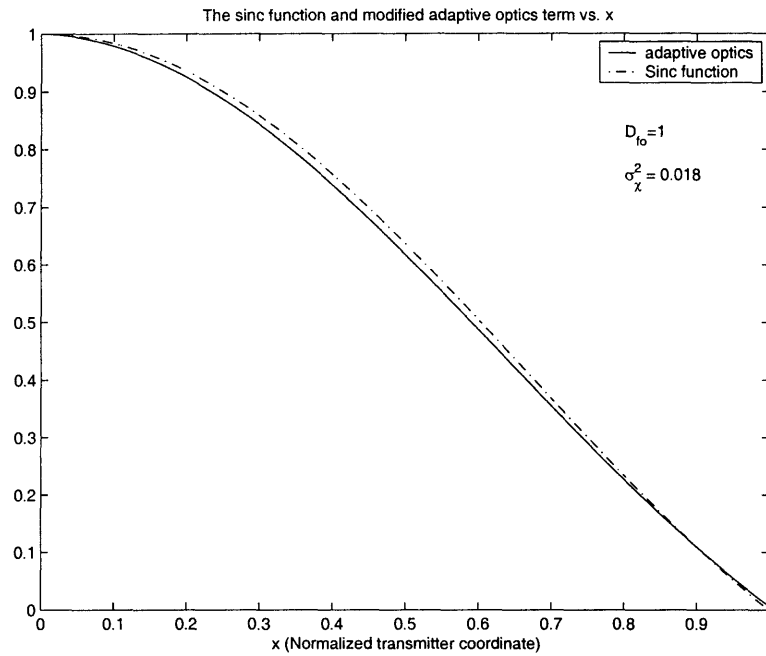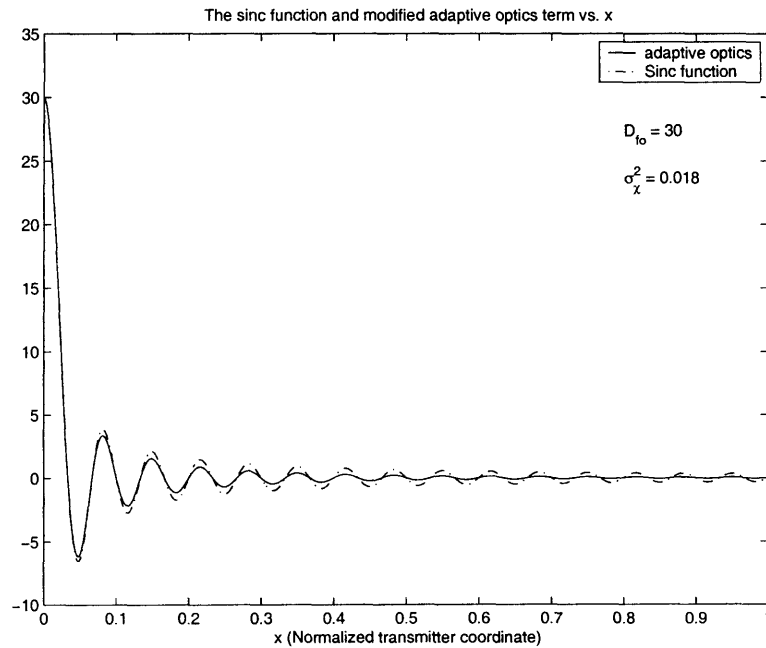
(a) $D_{fo} = 1$

(b) $D_{fo} = 30$

Figure 3-5: Plot of the sinc function vs. the integral expression obtained when adaptive phase compensation is used in mild turbulence conditions. $D_{fo} = 1$ compares the main lobes, while $D_{fo} = 30$ enables comparison of side-lobe behavior.
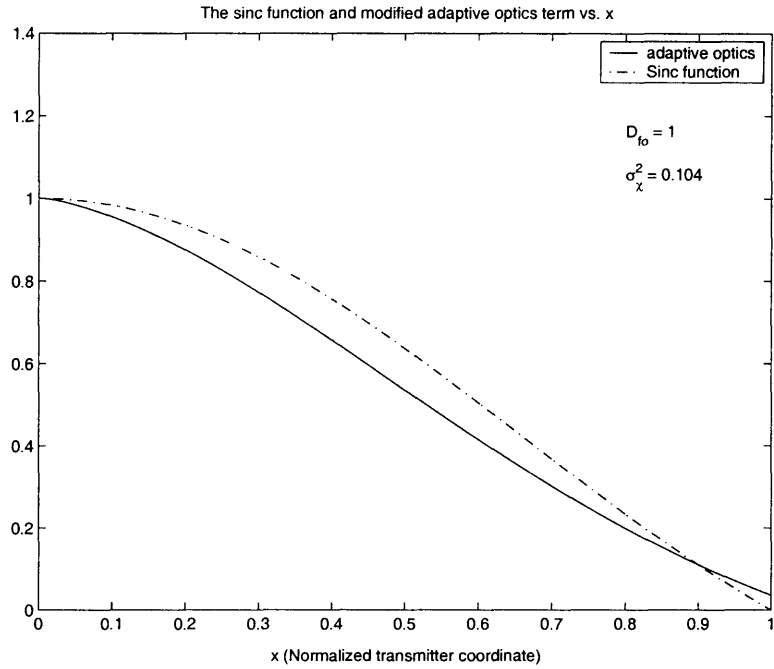
(a) $D_{fo} = 1$
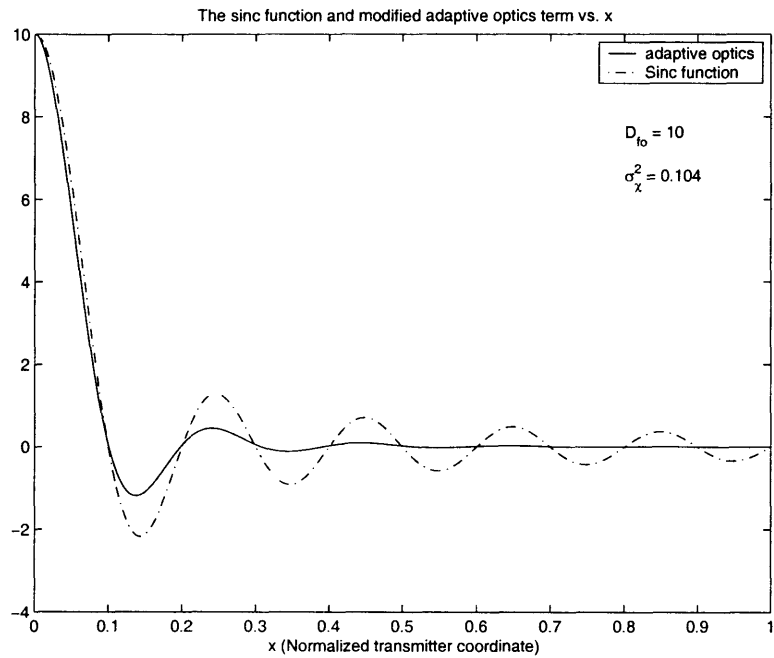
(b) $D_{fo} = 10$

Figure 3-6: Plot of the sinc function vs. the integral expression obtained when adaptive phase compensation is used in strong turbulence conditions. $D_{fo} = 1$ compares the main lobes, while $D_{fo} = 10$ enables comparison of side-lobe behavior.

particular turbulence strength, the phase compensated focused beam yields higher power transfer than the maximum-eigenvalue eigenfunction of the average kernel and eventually yields higher coupling than all other focused beam results. The maximum-eigenvalue eigenfunction of the log-amplitude only average kernel still yields the highest average power transfer.
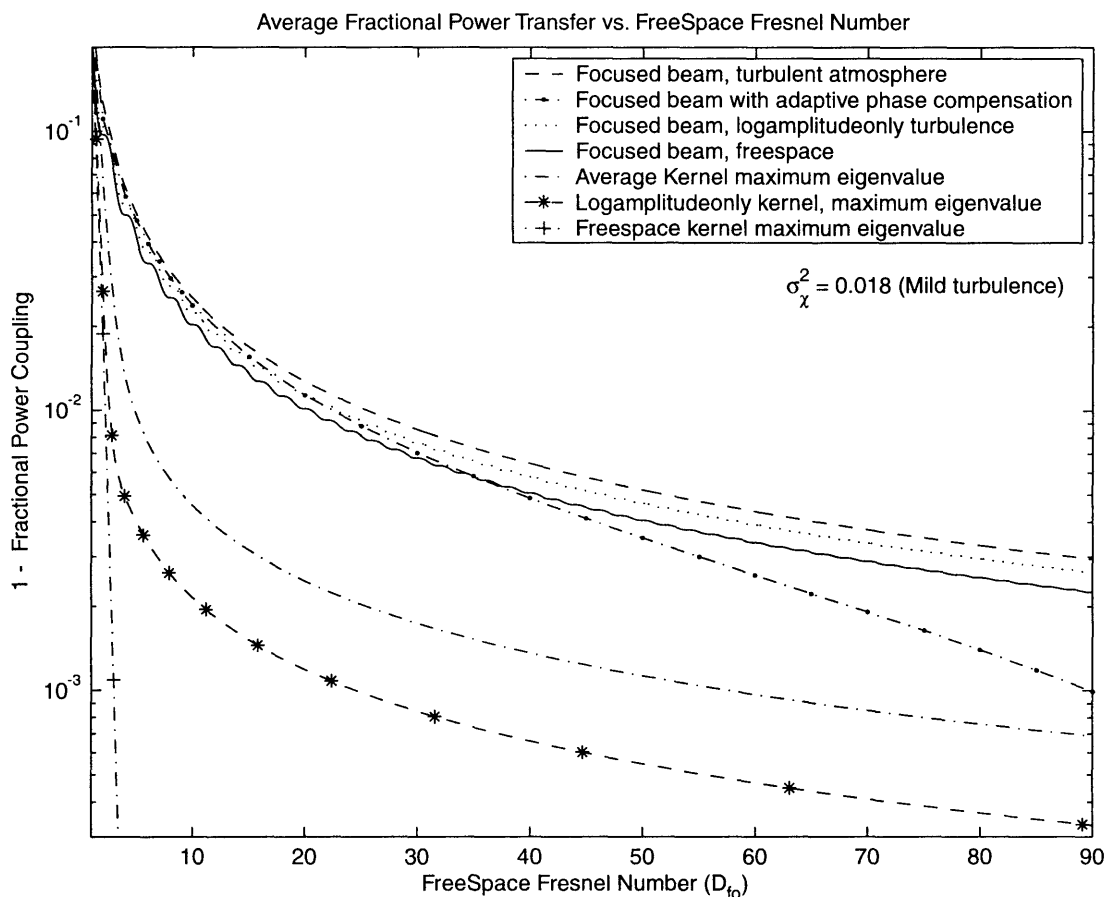
Figure 3-7: Average power transfer of the focused beam and the phase-compensated focused beam in mild turbulence conditions. The average power coupling of a hypothetical atmospheric condition with no phase fluctuations is also plotted, as are the maximum eigenvalues of the corresponding propagation kernels.

The figures show that in the equal aperture case, which is also the worst case, the free-space maximum eigenvalue approaches unity exponentially, while the focused beam results and the maximum eigenvalues of the average propagation kernels with or without phase fluctuations converge towards unity much slower. Figure 3-9 is a plot of the maximum eigenvalue of the average kernel and the log-amplitude-only

60
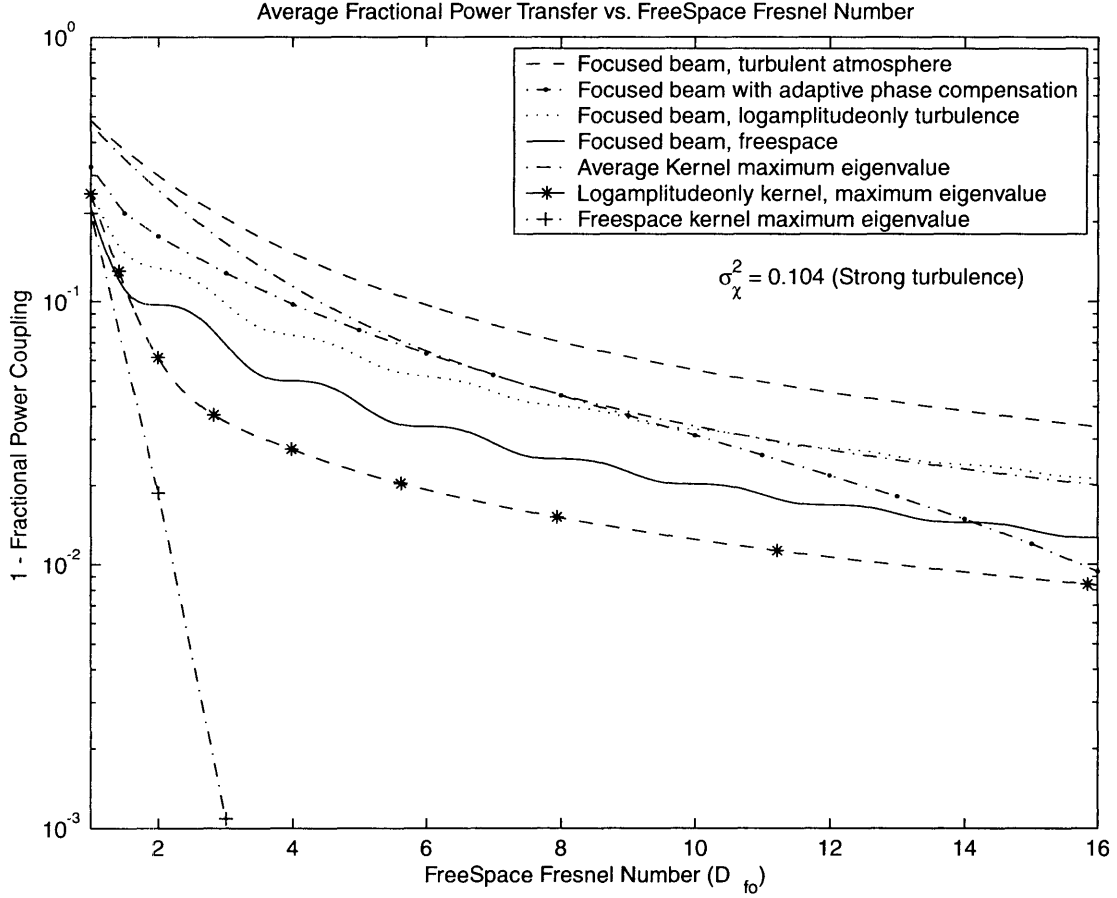
Figure 3-8: Average power transfer of the focused beam and the phase-compensated focused beam in strong turbulence conditions. Log-amplitude only focused beam results and maximum eigenvalues of the corresponding propagation kernels are also plotted.

kernel when unequal apertures are considered in mild turbulence and Figure 3-10 is the same plot for strong turbulence conditions. In both figures, $d_1 = \sqrt{\lambda L}$, so that the equal aperture case corresponds to $D_{fo} = 1$. Then, d_2 is varied to achieve free-space Fresnel numbers greater than 1. Because the transmitter aperture width is held constant, the exponential terms in the kernel expressions do no vary with D_{fo}, hence the maximum eigenvalues approach 1 faster than in the equal aperture case. Note however, that the convergence is still not exponential.

61
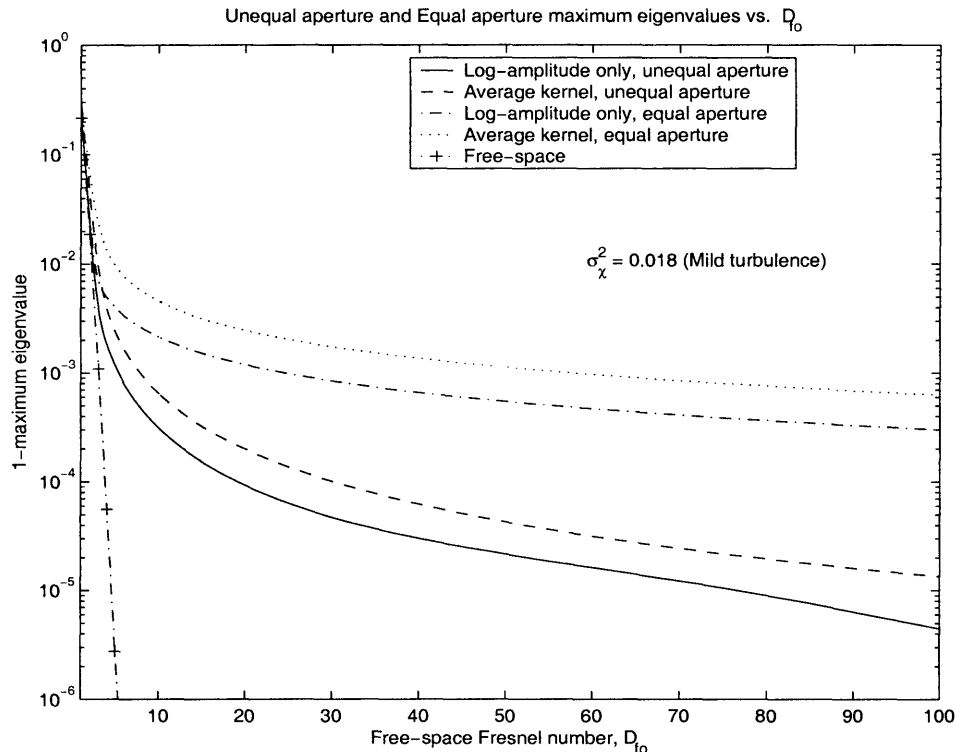
Figure 3-9: Comparison of the maximum eigenvalue of the average kernel as well as the log-amplitude-only case, when the transmitter and receiver aperture sizes are chosen to be equal versus unequal. Mild turbulence results shown.

3.5 Variance of Maximum Eigenvalue

We have shown that the mean value of the maximum eigenvalue of the atmospheric propagation kernel is very close to one when transmitter/receiver geometries yield free-space Fresnel numbers on the order of tens. Thus, we suspect that the variance of the maximum eigenvalue would be very small in this regime of operation.

In this section we obtain an expression for the second moment of the power transfer of a focused beam, $\overline{\gamma^2}$. Note that,

$$\gamma^2 = \left(\sum_{m=1}^{\infty} |u_{im}|^2 \eta_m \right)^2 \tag{3.49}$$

$$\leq \sum_{m=1}^{\infty} |u_{im}|^2 \sum_{m=1}^{\infty} |u_{im}|^2 \, \eta_m^2 \tag{3.50}$$

$$\leq \eta_1^2 \, , \tag{3.51}$$

62
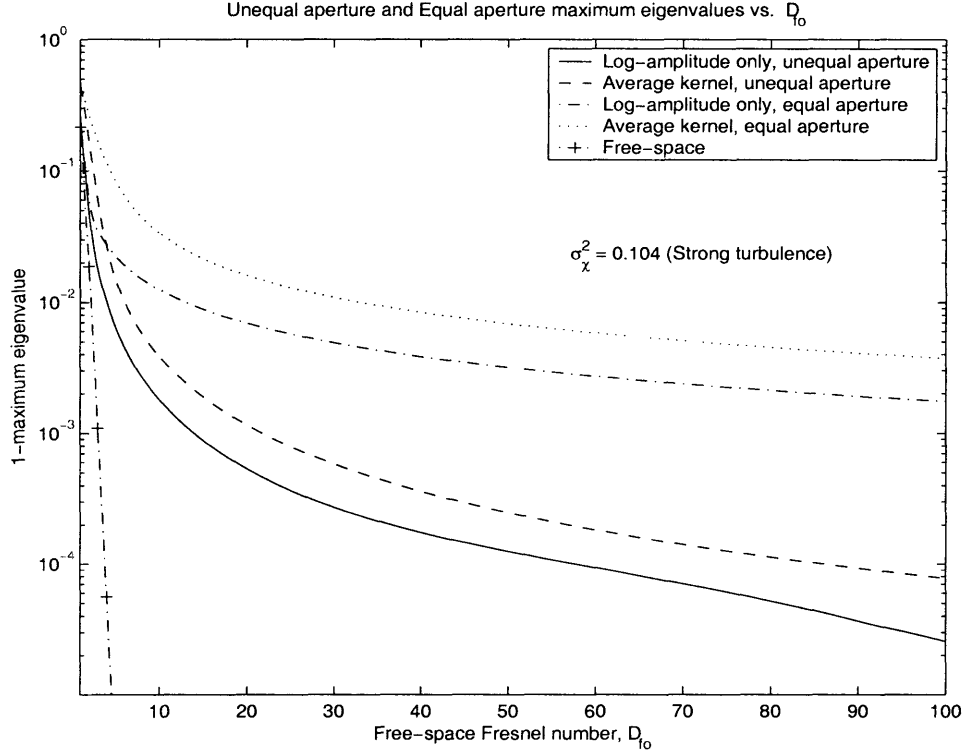
Figure 3-10: Comparison of the maximum eigenvalue of the average kernel as well as the log-amplitude-only case, when the transmitter and receiver aperture sizes are chosen to be equal versus unequal. Strong turbulence results shown.

where (3.50) follows from the Cauchy-Schwarz inequality and (3.51) assumes an input field with unity power, i.e., $\sum_{m=1}^{\infty} |u_{im}|^2 = 1$. Averaging (3.51) then yields $\overline{\gamma^2} \leq \overline{\eta_1^2}$.

Using the definition of γ in (3.11) and the focused beam input field from (3.18),

$$\gamma^2 = \frac{1}{(d_1 \lambda L)^2} \iiint_{R_2 R_2 R_1 R_1} \iiint_{R_1 R_1} \exp\left\{ -j\frac{k}{L}(\rho_{22} - \rho_{12})\rho_2' + j\frac{k}{L}(\rho_{21} - \rho_{11})\rho_1' \right. $$
$$\left. + g(\rho_{11}, \rho_{21}, \rho_{12}, \rho_{22}, \rho_1', \rho_2') \right\} d\rho_{22}\, d\rho_{12}\, d\rho_{21}\, d\rho_{11}\, d\rho_1'\, d\rho_2' \ . \quad (3.52)$$

where,

$$g(\rho_{11}, \rho_{21}, \rho_{12}, \rho_{22}, \rho_1', \rho_2') = \chi(\rho_2', \rho_{22}) + \chi(\rho_2', \rho_{12}) + \chi(\rho_1', \rho_{21}) + \chi(\rho_1', \rho_{11})$$
$$+ j\phi(\rho_2', \rho_{22}) - j\phi(\rho_2', \rho_{12}) - j\phi(\rho_1', \rho_{21}) + j\phi(\rho_1', \rho_{11}) \ . \quad (3.53)$$

Averaging over all χ and ϕ fields in (3.52), we obtain $\overline{\gamma^2}$. The expected value of

$\exp\{g(\rho_{11}, \rho_{21}, \rho_{12}, \rho_{22}, \rho_1', \rho_2')\}$, where g is a complex Gaussian random field, is given by $\exp\{g_2(\rho_{11}, \rho_{21}, \rho_{12}, \rho_{22}, \rho_1', \rho_2')\}$, with

$$
\begin{aligned}
g_2(\rho_{11}, \rho_{21}, \rho_{12}, \rho_{22}, \rho_1', \rho_2') = &-\tfrac{1}{2} D(0, \rho_{22} - \rho_{12}) - \tfrac{1}{2} D(\rho_2' - \rho_1', \rho_{22} - \rho_{21}) \\
&- \tfrac{1}{2} D(\rho_2' - \rho_1', \rho_{12} - \rho_{11}) - \tfrac{1}{2} D(0, \rho_{21} - \rho_{11}) - \tfrac{1}{2} D_{\chi,\chi}(\rho_2' - \rho_1', \rho_{22} - \rho_{11}) \\
&- \tfrac{1}{2} D_{\chi,\chi}(\rho_2' - \rho_1', \rho_{12} - \rho_{21}) + \tfrac{1}{2} D_{\phi,\phi}(\rho_2' - \rho_1', \rho_{22} - \rho_{11}) + \tfrac{1}{2} D_{\phi,\phi}(\rho_2' - \rho_1', \rho_{12} - \rho_{21}) \\
&+ 4\sigma_\chi^2 + 2j\, K_{\chi,\phi}(\rho_2' - \rho_1', \rho_{22} - \rho_{11}) - 2j\, K_{\chi,\phi}(\rho_2' - \rho_1', \rho_{12} - \rho_{21}) . \quad (3.54)
\end{aligned}
$$

With the assumption that both aperture sizes are equal to d_o, performing the transformations

$$
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \frac{1}{d_o}
\begin{bmatrix} 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{bmatrix}
\begin{bmatrix} \rho_{11} \\ \rho_{21} \\ \rho_{12} \\ \rho_{22} \end{bmatrix}
\qquad (3.55)
$$

and

$$
\begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{1}{d_o}
\begin{bmatrix} -1 & 1 \\ 0.5 & 0.5 \end{bmatrix}
\begin{bmatrix} \rho_1' \\ \rho_2' \end{bmatrix} ,
\qquad (3.56)
$$

allows x_4 and y' to be integrated out, hence reducing the original expression to the following four-dimensional integral:

$$
\overline{\gamma^2} = D_{fo}^2 \int_{-1}^{1} \int_{-1}^{1} \int_{-1}^{1} \int_{-1}^{1} g_3(x_1, x_2, x_3) \, \exp\{-j\pi D_{fo}(x_1 + x_2)x' + g_4(x_1, x_2, x_3, x')\}
$$
$$
\cdot \frac{\sin\left(\pi D_{fo}\left[1 - |x'|\right]\left[x_1 - x_2\right]\right)}{\pi D_{fo}(x_1 - x_2)} \, dx_1\, dx_2\, dx_3\, dx' , \quad (3.57)
$$

where

$$
g_3(x_1, x_2, x_3) = \max\Big\{0, 1 - \max\{0, x_2, x_2 + x_3, x_2 + x_3 - x_1\}
$$
$$
+ \min\{0, x_2, x_2 + x_3, x_2 + x_3 - x_1\}\Big\} \quad (3.58)
$$

64

and g_4 is just g_2 (defined in (3.54)) expressed in terms of the new variables. In particular,

$$
\begin{aligned}
g_4(x_1, x_2, x_3, x') = {}& -\tfrac{1}{2} D(0, x_1) - \tfrac{1}{2} D(x', x_3) \\
& - \tfrac{1}{2} D(x', x_3 + x_2 - x_1) - \tfrac{1}{2} D(0, x_2) + \tfrac{1}{2} D(x', x_2 + x_3) \\
& + \tfrac{1}{2} D(x', x_3 - x_1) - D_{\chi,\chi}\big(d_o\, x', d_o\, (x_2 + x_3)\big) - D_{\chi,\chi}\big(d_o\, x', d_o\, (x_3 - x_1)\big) \\
& + 4\sigma_\chi^2 + 2j \left(K_{\chi,\phi}\big(d_o\, x', d_o(x_2 + x_3)\big) - K_{\chi,\phi}\big(d_o\, x', d_o\, (x_3 - x_1)\big) \right), \quad (3.59)
\end{aligned}
$$

where $D(x', x)$ is the normalized spherical-wave wave structure function given in (A.10).

Attempts to evaluate this integral have not been successful with available resources for computation. However, the integrand in (3.57) is continuous in $\mathbb{R}^4$, and therefore is also integrable. Nevertheless, the integrand is highly oscillatory and therefore requires dense sampling — consequently vast resources — to converge to a valid result.

Chapter 4

Performance of Optical Binary Communication Systems

4.1   Introduction

In this chapter we will obtain bounds on probability of error for binary data transmission via near-field atmospheric optical communications. On-off keying (OOK) and pulse position modulation (PPM) will be considered. In each case we will treat both direct detection receivers and heterodyne detection receivers.

Figure 4-1 is a generic block diagram for the binary, atmospheric optical communication channel. A continuous-wave laser at some fixed and known wavelength (or center frequency) is OOK or PPM modulated by a binary sequence of data. In OOK, the laser is turned off during the bit interval if a 'zero' bit is being sent and turned on if a 'one' bit is being sent. An alternative to turning the laser on or off to communicate a bit is to choose two equal-energy, orthogonal waveforms and send one when $m = 0$ and the other when $m = 1$. The simplest two such waveforms are a pair of pulses with no overlap in time. PPM signal pulses fall in this category: The laser is turned on for the first half of the bit interval and turned off for the second half to transmit a 'zero' bit. If a 'one' bit is being sent, then the laser is turned off for the first half of the bit interval and turned on for the second half.

In all systems that are analyzed, we assume we have knowledge of the normal
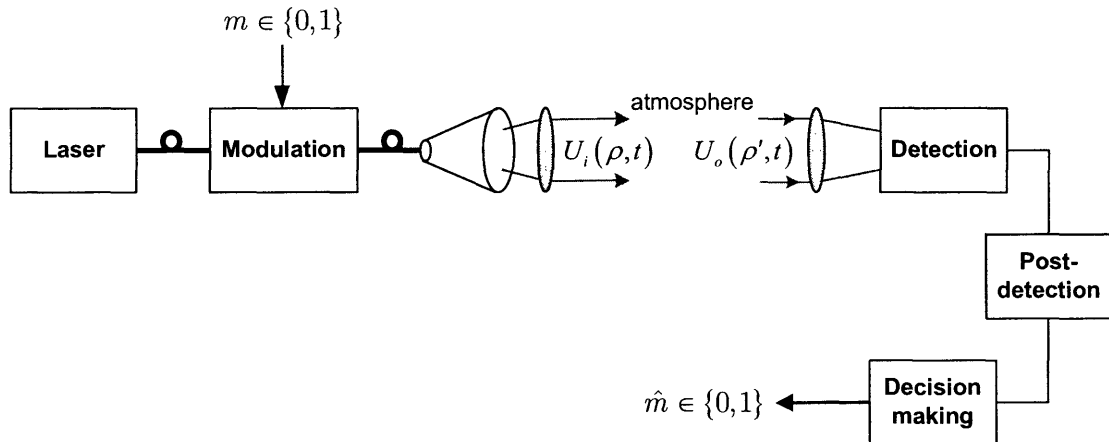
Figure 4-1: Block-diagram representation of binary, atmospheric optical communication systems.

mode decomposition of the atmosphere, or at least knowledge of the eigenfunctions associated with the maximum eigenvalue of the atmospheric propagation kernel. The scalar complex envelope of the transmitted field, $U_i(\rho, t)$, is assumed to factor into the product of spatial and temporal components and the space-dependent part is chosen to match the maximum-eigenvalue input eigenfunction of the atmospheric channel, $\Phi(\rho)$, to maximize the fractional power coupling between the transmitter and receiver. The transmitted field with power P_s is therefore expressed as

$$U_i(\rho, t) = \sqrt{P_s}\, s(t)\, \Phi(\rho) \,, \tag{4.1}$$

where the temporal modulation, $s(t)$, is chosen as shown in Figure 4-2 for OOK and as in Figure 4-3 for PPM during the bit transmission interval $0 \le t < T$.

4.2 Conditional Error Probabilities

4.2.1 On-Off Keying, Coherent Detection Receiver

Figure 4-2 shows a binary communication system, employing OOK modulation and a heterodyne receiver. For optimum heterodyne detection, when the transmitter's

68
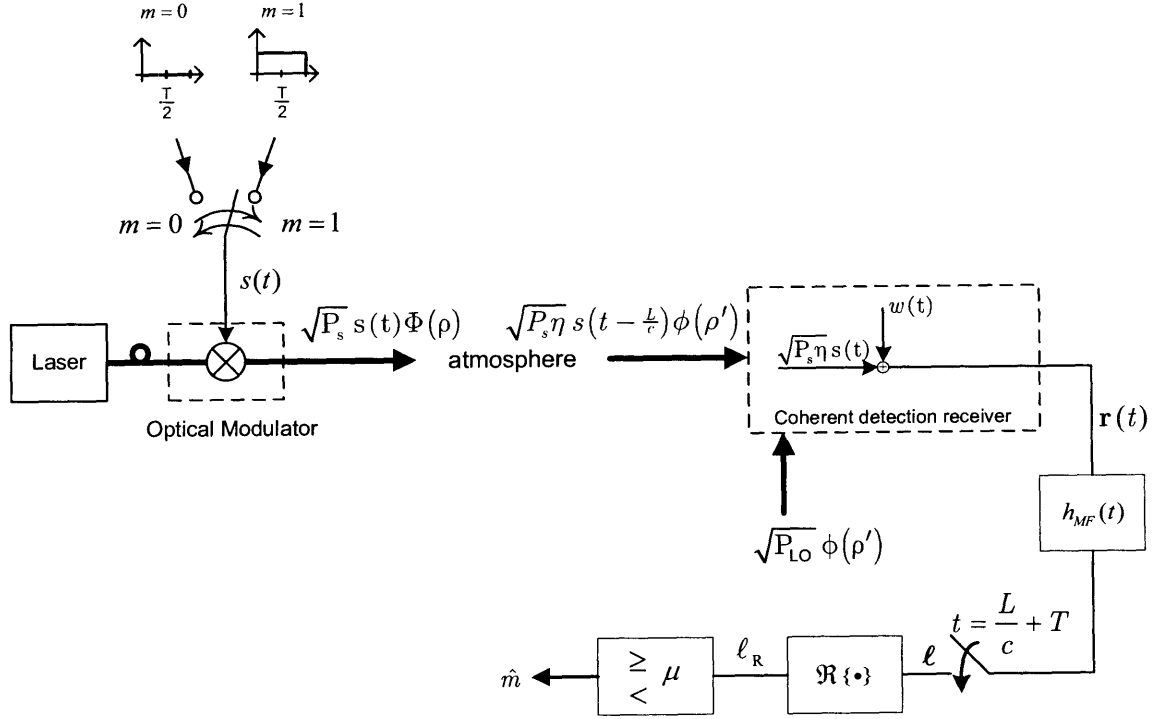
Figure 4-2: Block diagram of a binary (OOK) communication system with coherent detection.

spatial field pattern is $\Phi(\rho)$, the local oscillator field should be chosen to have the corresponding output eigenfunction $\phi(\rho')$; i.e. $U_{\mathrm{LO}}(\rho') = \sqrt{P_{\mathrm{LO}}}\,\phi(\rho')$. This leads to $\xi = 1$ for $m = 1$ and $\xi = 0$ for $m = 0$.

The normalized complex envelope of the output current from the coherent detector, $\mathbf{r}(t)$, will be a complex Gaussian random process[1]

$$\mathbf{r}(t) = m_{\mathbf{r}}(t) + \Delta\mathbf{r}(t) , \tag{4.2}$$

where $m_{\mathbf{r}}(t)$, the conditional mean, is given by

$$\left.\begin{array}{lll}\underline{m = 0}: & m_{\mathbf{r}}(t) & = & 0 \\[2mm] \underline{m = 1}: & m_{\mathbf{r}}(t) & = & \sqrt{P_s\,\eta}\end{array}\right\} \text{ for } \tfrac{L}{c} \le t < \tfrac{L}{c} + T , \tag{4.3}$$

and $\Delta\mathbf{r}(t)$ is a zero-mean, stationary, complex Gaussian random process with independent and identically distributed real and imaginary parts and spectral density $\frac{\hbar\omega_o}{\eta_e}$

---

[1]See Section 2.2.3.

69

in both cases.

It is well known from introductory systems analysis that the optimal processing for signals corrupted by stationary, additive white Gaussian noise is matched filtering followed by sampling and a threshold comparison test for decision making. Thus, we process the received signal $\mathbf{r}(t)$ with a LTI filter whose impulse response is

$$
h_{mf}(t) = \begin{cases} \sqrt{\frac{1}{T}} & , \quad \text{for } 0 < t \leq T \\ 0 & , \quad \text{otherwise} . \end{cases} \tag{4.4}
$$

The particular choice of the constant value in $h_{mf}(t)$ is to maintain a unit energy pulse, but the constant may be chosen to be any non-zero value, in general.

Sampling at the output of the matched filter yields the complex random variable

$$
\ell = \int_{-\infty}^{\infty} \mathbf{r}(\tau) \, h_{mf}\left(\tfrac{L}{c} + T - \tau\right) d\tau \tag{4.5}
$$

$$
= \frac{1}{\sqrt{T}} \int_{\frac{L}{c}}^{\frac{L}{c}+T} \mathbf{r}(\tau) \, d\tau . \tag{4.6}
$$

Note that, ℓ is simply the integral of the received signal $\mathbf{r}(t)$ over a single-symbol transmission interval. It is easily seen that ℓ is a complex Gaussian random variable, given m, with independent real and imaginary parts, each with variance $\sigma^2 = \frac{\hbar \omega_o}{2\eta_e}$. The mean of ℓ is 0 if $m = 0$ and $\sqrt{P_s \eta T}$ if $m = 1$. Because the mean is real both when $m = 0$ and $m = 1$, the real part of ℓ, ℓ_{R}, is a sufficient statistic for minimum probability of error detection.

The optimal decision threshold between two Gaussian distributions with equal variance, but different mean values is the mid-point between the two means; i.e. $\mu = \frac{\sqrt{P_s \eta T}}{2}$. Adopting this threshold, we can derive the probability of error, conditioned on

70

the knowledge of the maximum eigenvalue, η, of the atmospheric propagation kernel:

$$\Pr(\text{error}|\eta) \;=\; \frac{1}{2}\left\{ \Pr\left(\ell_{\text{R}} < \tfrac{\sqrt{P_s}\eta T}{2} \Big| m = 1 \right) + \Pr\left(\ell_{\text{R}} > \tfrac{\sqrt{P_s}\eta T}{2} \Big| m = 0 \right) \right\} \quad (4.7)$$

$$= \; Q\left(\sqrt{\frac{\eta_e P_s T}{2\hbar\omega_o}}\,\eta \right) \qquad\qquad\qquad (4.8)$$

$$= \; Q\left(\sqrt{\frac{n_s}{2}}\,\eta \right) \qquad\qquad\qquad\qquad (4.9)$$

where, $n_s\eta$ is the average number of detected signal photons over an $m = 1$ bit interval, and the Q-function is defined as,

$$Q\left(x \right) = \int_x^\infty \frac{1}{\sqrt{2\pi}}\, e^{-\frac{t^2}{2}}\, dt \;. \qquad\qquad (4.10)$$

We note that this particular detection and decoding scheme depends on the value of η to set the decision threshold optimally. Thus optimal decoding over an extended period of time is possible only if the threshold can be adjusted adaptively as the atmospheric kernel changes. Because fading fluctuations in the atmosphere occur on msec time scales and high rate communication systems have nsec bit intervals, it is feasible to assume feedback systems can be used to estimate and track the changes in power transfer over time.

Finally, this derivation assumes the bandwidth of the detector B is greater than the bandwidth of the post-detection processing filters. This assumption is equivalent to $B \gg \frac{2\pi}{T}$, which is the case in most applications[2].

4.2.2   Pulse-Position Modulation, Coherent Detection Receiver

Figure 4-3 illustrates a binary communication setup with PPM modulation and heterodyne detection. Noting that the local oscillator field is chosen as in the OOK case, the complex envelope of the output current from the coherent detector, $\mathbf{r}(t) =$

---

[2]B is in radians-per-second.
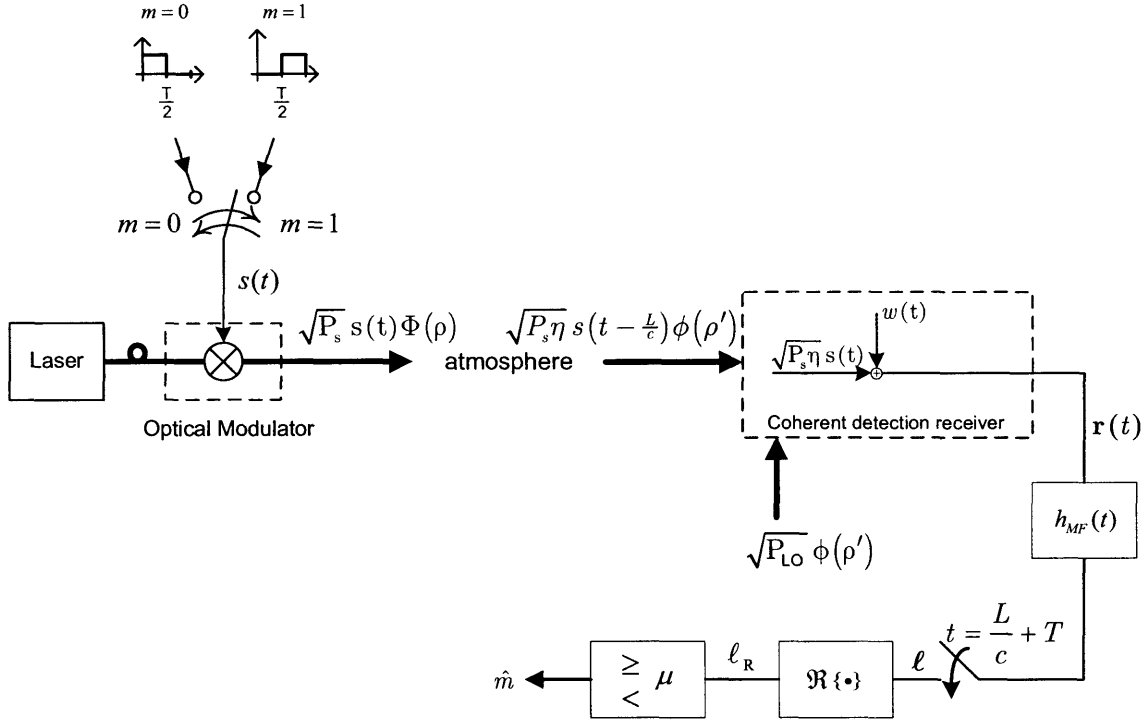
Figure 4-3: Block diagram of a binary PPM communication system with coherent detection.

$m_{\mathbf{r}}(t) + \Delta\mathbf{r}(t)$, is a complex Gaussian random process, given m, with mean function

$$
m_{\mathbf{r}}(t) = \begin{cases} \sqrt{P_s\,\eta} & \text{for } \frac{L}{c} \leq t < \frac{T}{2} + \frac{L}{c} \\ 0 & \text{for } \frac{T}{2} + \frac{L}{c} \leq t < T + \frac{L}{c} \end{cases} \tag{4.11}
$$

for $m = 0$, and

$$
m_{\mathbf{r}}(t) = \begin{cases} 0 & \text{for } \frac{L}{c} \leq t < \frac{T}{2} + \frac{L}{c} \\ \sqrt{P_s\,\eta} & \text{for } \frac{T}{2} + \frac{L}{c} \leq t < T + \frac{L}{c} \end{cases} \tag{4.12}
$$

for $m = 1$. The noise component $\Delta\mathbf{r}(t)$ is a zero-mean, complex white Gaussian random process with independent, identically distributed real and imaginary parts and spectral density $\frac{\hbar\omega_o}{\eta_e}$. Matched filtering $\mathbf{r}(t)$, followed by sampling and threshold comparison of the real part of the sample will yield a minimum probability of error decision rule [22]. The matched filter impulse response is the difference of the mean

72

processes for $m = 1$ and $m = 0$, normalized to unit energy for convenience; i.e.

$$h_{mf}(t) = \begin{cases} \frac{1}{\sqrt{T}} & , \quad 0 \leq t < \frac{T}{2} \\ -\frac{1}{\sqrt{T}} & , \quad \frac{T}{2} \leq t < T \, , \end{cases} \tag{4.13}$$

Sampling at time $t = \frac{L}{c} + T$ yields the complex Gaussian random variable

$$\ell = \frac{1}{\sqrt{T}} \left(\int_{\frac{L}{c}+\frac{T}{2}}^{\frac{L}{c}+T} \mathbf{r}(\tau) \, d\tau - \int_{\frac{L}{c}}^{\frac{L}{c}+\frac{T}{2}} \mathbf{r}(\tau) \, d\tau \right) . \tag{4.14}$$

The real part of $\boldsymbol{\ell}$ is a sufficient statistic, which can be written in regards to the two transmitted bit values as

$$\begin{aligned} \underline{m = 0}: \quad \ell_{\mathrm{R}} &= -\frac{\sqrt{P_s \eta T}}{2} + \Delta\ell \\ \underline{m = 1}: \quad \ell_{\mathrm{R}} &= \frac{\sqrt{P_s \eta T}}{2} + \Delta\ell \, , \end{aligned} \tag{4.15}$$

where $\Delta\ell$ is a zero-mean Gaussian random variable with variance $\sigma^2 = \frac{\hbar\omega_o}{2\eta_e}$.

Thus, the optimal decision threshold is $\mu = 0$. The probability of error, conditioned on the knowledge of the maximum eigenvalue η is then,

$$\begin{aligned} \Pr(\text{error}|\eta) &= \frac{1}{2}\Big\{ \Pr\big(\ell < 0|m = 1\big) + \Pr\big(\ell > 0|m = 0\big) \Big\} \tag{4.16} \\ &= Q\left(\sqrt{\frac{\eta_e P_s T}{2\hbar\omega_o}}\, \eta \right) \tag{4.17} \\ &= Q\big(\sqrt{n_s \eta}\big) \tag{4.18} \end{aligned}$$

where, $n_s\eta$ represents the average number of detected signal photons during a bit interval.

We see from equations (4.18) and (4.9) that at equal n_s values, the conditional error probability of OOK is 3 dB worse than that for PPM with a coherent detection receiver. This is because in OOK photons are transmitted only if $m = 1$. On the other hand, in PPM, photons are transmitted both when $m = 1$ and $m = 0$. Hence, on average, twice as many photons are being transmitted in PPM than what is the case for OOK. PPM also offers a significant advantage in implementation of post-

73
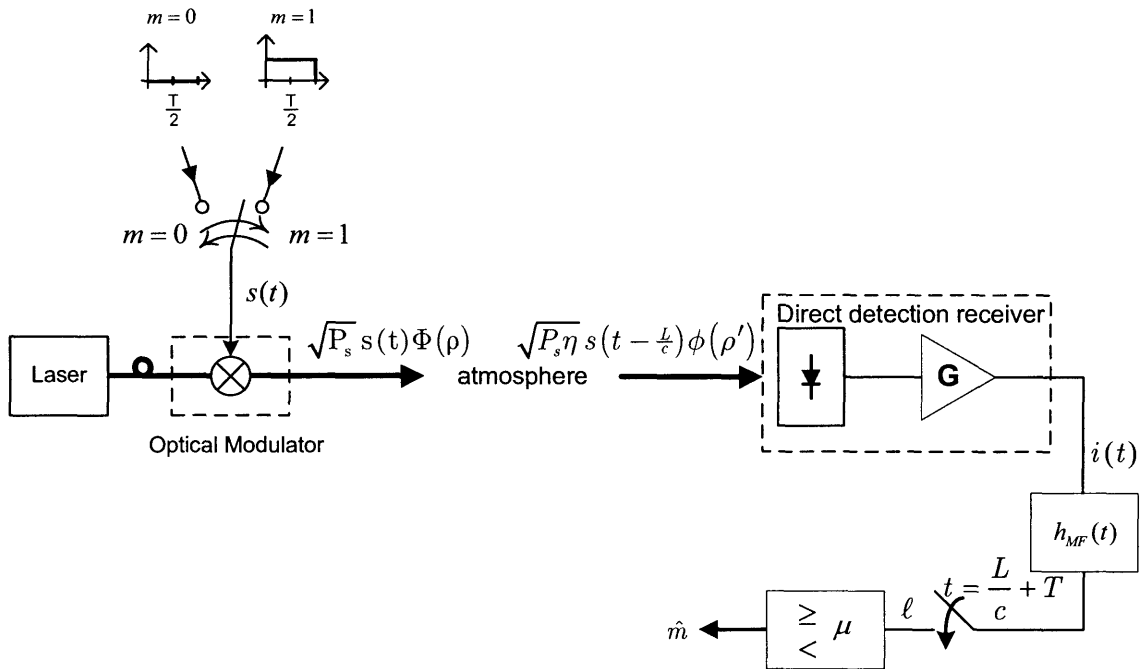
Figure 4-4: Block diagram of a binary OOK communication system with direct detection.

detection processing because the decision threshold is 0, regardless of the state of atmospheric turbulence. Thus this decision scheme is insensitive to time dependent changes in the atmosphere.

4.2.3 On-Off Keying, Direct Detection Receiver

Figure 4-4 shows a block diagram of a binary communication system using OOK modulation at the transmitter and a direct detection receiver. The complex envelope of the received field on the $z = L$ plane is

$$U_o(\boldsymbol{\rho}', t) = \sqrt{P_s \eta} \; s\!\left(t - \tfrac{L}{c}\right) \phi(\boldsymbol{\rho}') \,, \qquad (4.19)$$

where $s(t)$ represents the temporal OOK modulation.

The output current of the direct detection receiver, denoted by $i(t)$, is a non-stationary, white, Gaussian random process[3]. The mean and covariance functions of

---

[3]See section 2.2.2.

74

$i(t)$ are given by (2.35) and (2.37) respectively, where $P(t) = P_s \eta$ for $m = 1$ and $P(t) = 0$ for $m = 0$, over the bit interval $\frac{L}{c} \leq t, u < \frac{L}{c} + T$.

Because the covariance function of $i(t)$ depends on the received signal power, it is non-stationary. In addition, the noise statistics depend on whether a 'zero' or 'one' bit has been transmitted. This results in a minimum probability of error decision rule which is nonlinear and hard to implement. It is common, therefore, to implement a decision rule that would be optimal if the noise statistics were stationary and independent of the bit value. This condition is approached when the signal shot-noise is a minor contributor to the output current noise. Thus, this decision rule will be optimal for direct detection receivers operating in the thermal-noise dominant regime.

The post-detection processing involves filtering the output current through a LTI filter with impulse response

$$
h_{mf}(t) = \begin{cases} \frac{1}{\sqrt{T}} & , \quad \text{for } 0 < t \leq T \\ 0 & , \quad \text{otherwise} \end{cases} , \tag{4.20}
$$

and sampling the output of the matched filter at time $\frac{L}{c} + T$. This gives a Gaussian random variable

$$
\ell = \frac{1}{\sqrt{T}} \int_{\frac{L}{c}}^{\frac{L}{c}+T} i(\tau) \, d\tau , \tag{4.21}
$$

which is fully characterized by its mean and variance. When $m = 0$, ℓ has mean

$$
m_0 = \sqrt{T} \, G \, I_d , \tag{4.22}
$$

and variance

$$
\sigma_0^2 = qG^2 I_d + G^2 \frac{2KT_o}{R_L} . \tag{4.23}
$$

On the other hand, when $m = 1$, ℓ has mean

$$
m_1 = \sqrt{T} \left[\frac{\eta_e \, q \, G}{\hbar \, \omega_o} P_s \, \eta + G \, I_d \right] , \tag{4.24}
$$

75

and variance,

$$\sigma_1^2 = \frac{\eta_e \, q^2 \, G^2}{\hbar \, \omega_o} P_s \, \eta + q \, G^2 \, I_d + G^2 \frac{2KT_o}{R_L}. \tag{4.25}$$

The optimal decision threshold is not trivial to obtain for two Gaussian random variables with different means and variances. Hence the threshold is instead chosen such that the *false-alarm* probability, $\Pr(\hat{m} = 1 | m = 0)$, is equal to the *miss* probability, $\Pr(\hat{m} = 0 | m = 1)$. This condition yields a threshold

$$\mu = \frac{\sigma_0 m_1 + \sigma_1 m_0}{\sigma_0 + \sigma_1} . \tag{4.26}$$

Then the probability of error conditioned on the knowledge of η is

$$\Pr(\text{error}|\eta) = Q\left(\frac{m_1 - m_0}{\sigma_1 + \sigma_0}\right) \tag{4.27}$$

$$= Q\left(\frac{\frac{\eta_e P_s T}{\hbar \omega_o}\eta}{\sqrt{\frac{\eta_e P_s T}{\hbar \omega_o}\eta + \lambda_d T + \frac{2KT_o T}{R_L q^2}} + \sqrt{\lambda_d T + \frac{2KT_o T}{R_L q^2}}}\right), \tag{4.28}$$

which, after some simplification, can be rewritten as

$$\Pr(\text{error}|\eta) = Q\left(\sqrt{n_s \, \eta + n_N} - \sqrt{n_N}\right) \tag{4.29}$$

where, $n_s \eta = \frac{\eta_e P_s T}{\hbar \omega_o}\eta$ is the average number of detected signal photons over a $m = 1$ bit interval and $n_N = \lambda_d T + \frac{2KT_o T}{R_L q^2}$ is an effective average number of noise photons over a bit interval, i.e., the shot-noise variance of n_N matches the dark current plus thermal noise variance.

When thermal noise is dominant, i.e. $n_N \gg n_s$, the conditional error probability expression simplifies to

$$\Pr(\text{error}|\eta) = Q\left(\frac{n_s \, \eta}{2\sqrt{n_N}}\right) . \tag{4.30}$$

4.2.4 Pulse-Position Modulation, Direct Detection Receiver

The final system that we consider is a communication system utilizing PPM and a direct detection receiver, shown in Figure 4-5. The received field on the $z = L$ plane

76
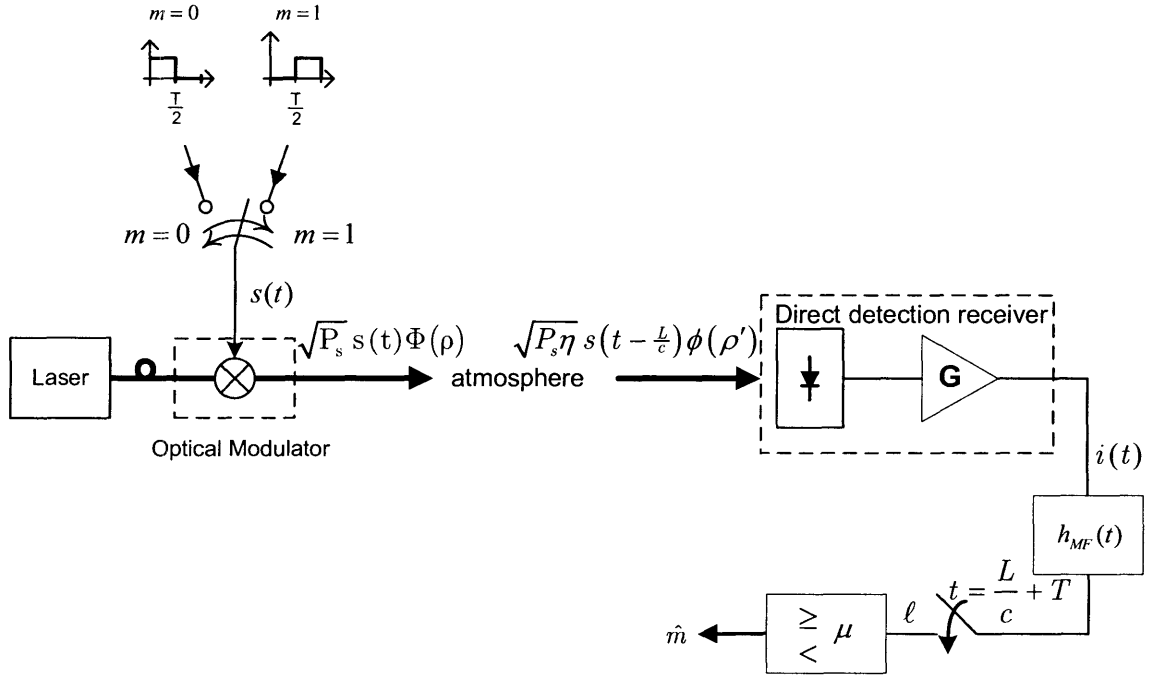
Figure 4-5: Block diagram of a binary PPM communication system with direct detection.

is $U_o(\rho', t) = \sqrt{P_s\,\eta}\; s(t - \frac{L}{c})\, \phi(\rho')$, with $s(t)$ shown in Figure 4-5.

The output current of the direct detection receiver, $i(t)$, is a Gaussian random process with mean and covariance functions given in (2.35) and (2.37) respectively. For $m = 0$, the received power function in these equations is

$$P(t) = \begin{cases} P_s\,\eta & \text{if } \frac{L}{c} \leq t < \frac{L}{c} + \frac{T}{2}\,, \\ 0 & \text{if } \frac{L}{c} + \frac{T}{2} \leq t < \frac{L}{c} + T\,, \end{cases} \tag{4.31}$$

and for $m = 1$,

$$P(t) = \begin{cases} 0 & \text{if } \frac{L}{c} \leq t < \frac{L}{c} + \frac{T}{2}\,, \\ P_s\,\eta & \text{if } \frac{L}{c} + \frac{T}{2} \leq t < \frac{L}{c} + T\,. \end{cases} \tag{4.32}$$

The post-processing scheme is identical to that of an OOK modulated system with a direct-detection receiver. The impulse response of the matched filter for PPM is

$$h_{mf}(t) = \begin{cases} \frac{1}{\sqrt{T}} & ,\; 0 \leq t < \frac{T}{2} \\ -\frac{1}{\sqrt{T}} & ,\; \frac{T}{2} \leq t < T\,. \end{cases} \tag{4.33}$$

Then, the sample at $t = \frac{L}{c} + T$ is a conditionally Gaussian random variable,

$$\ell = \frac{1}{\sqrt{T}} \left(\int_{\frac{L}{c}+\frac{T}{2}}^{\frac{L}{c}+T} i(\tau)\, d\tau - \int_{\frac{L}{c}}^{\frac{L}{c}+\frac{T}{2}} i(\tau)\, d\tau \right) , \qquad (4.34)$$

with mean

$$m_0 = -\frac{\sqrt{T}}{2} \frac{\eta_e\, q\, G}{\hbar\,\omega_o} P_s \eta , \qquad (4.35)$$

if $m = 0$, and

$$m_1 = \frac{\sqrt{T}}{2} \frac{\eta_e\, q\, G}{\hbar\,\omega_o} P_s \eta , \qquad (4.36)$$

if $m = 1$. The variances in both cases are equal and given by

$$\sigma_0^2 = \sigma_1^2 = \frac{\eta_e\, q^2\, G^2}{2\hbar\,\omega_o} P_s\, \eta + q\, G^2 I_d + G^2 \frac{2KT_o}{R_L} . \qquad (4.37)$$

Since ℓ has antipodal means and equal variances under the two hypotheses, the optimal threshold is $\mu = 0$. Then the error probability conditioned on η is

$$\Pr(\text{error}|\eta) = Q\left(\frac{m_1}{\sigma_1}\right) \qquad (4.38)$$

$$= Q\left(\frac{\frac{\eta_e P_s T}{2\hbar\omega_o}\eta}{\sqrt{\frac{\eta_e P_s T}{2\hbar\omega_o}\eta + \lambda_d T + \frac{2KT_o T}{R_L q^2}}} \right) , \qquad (4.39)$$

$$\qquad (4.40)$$

which can be rewritten as

$$\Pr(\text{error}|\eta) = Q\left(\frac{n_s \eta}{\sqrt{n_s\, \eta + n_N}} \right) , \qquad (4.41)$$

where $n_s \eta = \frac{\eta_e P_s T}{2\hbar\omega_o}\eta$ is the average number of detected signal photons over a bit interval.

In the thermal noise dominant regime, i.e. when $n_N \gg n_s$, the conditional error

78

probability expression simplifies to

$$\text{Pr}(\text{error}|\eta) = Q\left(\frac{n_s\,\eta}{\sqrt{n_N}}\right),$$

(4.42)

which is 3 dB better than the conditional error probability result for a OOK modulated communication system with direct detection receivers operating in the thermal noise dominated regime. This is because, at equal n_s values, the expected number of photons transmitted in OOK is half of the expected number of transmitted photons in PPM. As is the case for coherent detection receivers, the implementation of the PPM direct detection decision rule is easier than its OOK counterpart because the threshold is independent of the turbulent state of the atmosphere.

4.3 Unconditional Error Probabilities

4.3.1 The Markov Moment-Matching Problem

In the previous section we obtained the error probabilities *conditioned* on knowledge of the atmospheric propagation kernel's maximum eigenvalue. The unconditional error probabilities follow from the law of total probability:

$$\text{Pr}(\text{error}) = \int_0^1 \text{Pr}(\text{error}|\eta)\,p(\eta)d\eta$$

(4.43)

where $p(\eta)$ is the probability density function[4] of η. Thus, finding the unconditional error probability requires knowledge of the distribution of η. However, using the log-amplitude and phase statistics introduced in Section 2.1 to derive the distribution of η appears impossible. Part of the challenge is due to not knowing the input eigenfunction corresponding to the maximum eigenvalue. Thus, alternatively, one may be interested in obtaining the probability of error when a particular input field, such as any of the fields considered in Chapter 3, is used. The fractional power transfer in this case, γ, is a random variable which is less than the maximum eigenvalue, η,

---

[4]Through out this thesis, we assume η has a probability density function, possibly including generalized functions, such as delta functions.

i.e., $0 \leq \gamma \leq \eta \leq 1$. It follows immediately that $\Pr(\gamma \leq x) \geq \Pr(\eta \leq x)$, for any x between 0 and 1. Consider the difference between the unconditional error probability when γ is the power coupling between the transmitter and receiver, and the error probability when η is so, i.e., consider

$$\Pr_\gamma(\text{error}) - \Pr_\eta(\text{error}) = \int_0^1 \Pr_e(x)\, p_\gamma(x)\, dx - \int_0^1 \Pr_e(x)\, p_\eta(x)\, dx \,, \qquad (4.44)$$

where $\Pr_e(x) = \Pr(\text{error} \mid \text{power transfer} = x)$.

Using integration by parts on both integrals separately, we obtain

$$\Pr_\gamma(\text{error}) - \Pr_\eta(\text{error}) = \Pr_e(x)\, \Pr(\gamma \leq x) \Big|_0^1 - \int_0^1 \frac{\partial \Pr_e(x)}{\partial x}\, \Pr(\gamma \leq x)\, dx$$

$$- \Pr_e(x)\, \Pr(\eta \leq x) \Big|_0^1 + \int_0^1 \frac{\partial \Pr_e(x)}{\partial x}\, \Pr(\eta \leq x)\, dx \,, \quad (4.45)$$

which can be further simplified to

$$\Pr_\gamma(\text{error}) - \Pr_\eta(\text{error}) = \frac{1}{2}\left[\Pr(\gamma = 0) - \Pr(\eta = 0)\right]$$

$$+ \int_0^1 \frac{\partial \Pr_e(x)}{\partial x}\Big(\Pr(\eta \leq x) - \Pr(\gamma \leq x)\Big)\, dx \geq 0 \,, \quad (4.46)$$

where the inequality is true because the conditional error probability is a decreasing function of increasing x. Thus, we have shown via (4.46) that the unconditional error probability obtained using a sub-optimal transmitter field must be an upper bound to the unconditional error probability achieved when the maximum-eigenvalue eigenfunction is the transmitted spatial field pattern. However, this approach also requires the knowledge of the distribution of γ, which appears impossible to derive from first principles. Nevertheless, in Chapter 3 we have obtained the mean of γ for a variety of input fields, and with more computational power, we might obtain its second moment when the transmitted field is a focused beam.

Thus, as an alternative to deriving the probability density function of η, we consider the set of all probability distributions on $[0, 1]$ which satisfy a mean $\bar{\eta}$ and

variance σ_η^2 constraint. Within this set of possible distributions for η, we attempt to find those which result in maximum and minimum error probabilities for each of the four communication systems analyzed in the previous section. Note that, if we choose $\bar{\eta}$ and σ_η^2 equal to the first and second moments of the power coupling of a sub-optimal field, the maximum error probability derived from those constraints will also be an upper bound on the error probability achieved with the maximum-eigenvalue eigenfunction.

Much of the work in this chapter rests upon Markov's studies of the extremal values of integrals which are of the form given in (4.43) [23, 24]. The results relevant to our problem are compiled in the following lemma:

Lemma 4.3.1. *Let $\Omega(t)$ be a three-times differentiable function on the closed interval $[a, b]$. Also, let $\mathcal{V}(p(t); s_1, s_2)$ denote the set of all probability densities $p(t)$ on $[a, b]$, which satisfy*

$$\int_a^b t\, p(t)\, dt = s_1 \tag{4.47a}$$

and

$$\int_a^b t^2\, p(t)\, dt = s_2 \,. \tag{4.47b}$$

Furthermore, suppose the set of moments $\{s_1, s_2\}$ satisfy,

$$s_2 - s_1^2 > 0 \tag{4.48}$$

and

$$(a + b)s_1 - a\,b - s_2 > 0 \,.[5] \tag{4.49}$$

81

Finally, assume

$$\frac{\partial^3 \Omega(t)}{\partial t^3} > 0 \,, \quad t \in [a, b] \,. \tag{4.50}$$

Then, the minimum value of the integral $\int_a^b \Omega(t) p(t)\, dt$ is achieved by a unique distribution, $\sigma_{min}(t)$, which has mass concentrated at point $t = a$ and an interior point t_0 $\bigl(i.e.\ t_0 \in (a, b)\bigr)$[6].

Likewise, the maximum value of the integral $\int_a^b \Omega(t)\, p(t)\, dt$ is attained with a unique distribution, $\sigma_{max}(t)$, which has mass concentrated at point $t = b$ and an interior point $t_1 \in (a, b)$.

We shall draw upon the results of this lemma several times in the work that follows. Because a variety of closed intervals will appear in our analysis, it is convenient to derive the maximizing and minimizing distributions over an arbitrary interval $[a, b]$ before proceeding with the unconditional error probability results.

Assuming the conditions stated in the lemma are satisfied, the distribution with support $[a, b]$ which minimizes the integral expression has two mass concentration points: one at $t = a$ and another interior point we denote with t_0. Then the probability density function will be two impulses,

$$p_{min}(t) = p_0 \, \delta(t - a) + (1 - p_0) \, \delta(t - t_0) \,, \tag{4.51}$$

where p_0 and t_0 are obtained from the first and second moment constraints given by

$$s_1 = \int_a^b t \, p_{min}(t)\, dt = a\, p_0 + t_0 \,(1 - p_0) \tag{4.52}$$

$$s_2 = \int_a^b t^2 \, p_{min}(t)\, dt = a^2 \, p_0 + t_0^2 \,(1 - p_0) \,. \tag{4.53}$$

---

[5]Note that there exists at least one probability distribution with mean s_1 and mean-square s_2 if, and only if, the expressions in (4.48) and (4.49) are non-negative. This is because (4.48) and (4.49) are equivalent to $\mathrm{var}(X) > 0$ and $E[(X - a)(X - b)] < 0$, for X a random variable on $[a, b]$ with probability density $p(x)$. Therefore the two strict inequalities presented in the text guarantee $\mathcal{V}(\sigma(t); s_1, s_2)$ is non-empty.

[6]If either of (4.48) or (4.49) are not strict, there may exist other distributions with mass concentrated at two points which achieve the extremum values. If (4.50) is not strict, there may exist multiple distributions in $\mathcal{V}(\sigma(t); s_1, s_2)$ that attain the extremum values of the integral given by $\sigma_{min}(t)$ and $\sigma_{max}(t)$.

Solving these two equations yields

$$p_0 = \frac{s_2 - s_1^2}{(s_2 - s_1^2) + (s_1 - a)^2} = \frac{\sigma^2}{\sigma^2 + (s_1 - a)^2} \qquad (4.54)$$

and

$$t_0 = s_1 + \frac{\sigma^2}{s_1 - a} \, , \qquad (4.55)$$

where $\sigma^2 = s_2 - s_1^2$ is the prescribed variance of the distribution.

The distribution with support $[a, b]$ which maximizes the integral expression also has two mass concentration points. However, now one is at $t = b$ and the other is at an interior point denoted by t_1. The probability density is then

$$p_{max}(t) = p_1 \, \delta(t - b) + (1 - p_1) \, \delta(t - t_1) \, . \qquad (4.56)$$

Once again, p_1 and t_1 are obtained by solving the mean and mean-square constraints

$$s_1 = \int_a^b t \, p_{max}(t) \, dt = b \, p_1 + t_1 \, (1 - p_1) \qquad (4.57)$$

$$s_2 = \int_a^b t^2 \, p_{max}(t) \, dt = b^2 \, p_1 + t_1^2 \, (1 - p_1) \, , \qquad (4.58)$$

which results in

$$p_1 = \frac{\sigma^2}{\sigma^2 + (b - s_1)^2} \qquad (4.59)$$

and

$$t_1 = s_1 - \frac{\sigma^2}{(b - s_1)} \, . \qquad (4.60)$$

It is worthwhile to note two facts about the results derived above. First, the maximizing and minimizing distributions, $p_{min}(t)$ and $p_{max}(t)$, somewhat surprisingly do not depend on the function $\Omega(t)$, so long as $\Omega(t)$ is a three times differentiable

function with a positive third derivative. Second, if $\Omega(t)$ has a third derivative that is negative, then one can apply the above lemma and our subsequent results to $-\Omega(t)$, thereby concluding that (4.56) yields the minimizing distribution and (4.51) yields the maximizing distribution.

Through straightforward calculus, one can verify that the coherent detection conditional probability of error expressions, (4.29) and (4.41) are three times continuously differentiable on $(0, 1]$ and have strictly negative third derivatives. On the other hand, the direct detection results (4.9) and (4.18), although three times continuously differentiable on $[0, 1]$, do not have sign-preserving third derivatives. Therefore we start our analysis with communication systems employing coherent detection receivers. Later we show that the direct detection conditional error probability results can be closely approximated by functions that have non-positive third derivatives and thus the same distributions that yield extremum unconditional error probabilities in coherent detection systems, yield approximately the extremum unconditional error probabilities for direct-detection systems as well.

4.3.2 Coherent Detection Unconditional Probability of Error Bounds

Because OOK and PPM conditional error probability results differ only by a factor of 2 in the coefficient that precedes η, we will only derive the results for OOK in this thesis. The results for PPM with coherent detection can be obtained by simply substituting $2n_s$ in place of n_s in all of the results that follow in this section.

To emphasize that $\Pr(\text{error}|\eta)$ is a function of the random variable η, we use $f_c(\eta)$ to denote the conditional error probability for coherent detection, i.e.,

$$f_c(\eta) = Q\left(\sqrt{\frac{n_s}{2}}\,\eta\right).$$

(4.61)

The third derivative of $f_c(\eta)$ is

$$\frac{\partial^3 f_c(\eta)}{\partial \eta^3} = -\frac{1}{16}\sqrt{\frac{n_s}{\pi}}\left[\frac{n_s}{\eta^{3/2}} + \frac{3}{\eta^{5/2}} + \frac{n_s^2}{4\,\eta^{1/2}}\right]e^{-\frac{n_s\eta}{4}} ,\tag{4.62}$$

$\forall \eta \in (0,1]$, which is clearly strictly negative for all $\eta \in (0,1]$. Therefore, we can apply the Markov moment-matching lemma to $-f_c(\eta)$ and obtain the probability distributions of η that yield the maximum and minimum unconditional error probabilities [7]. Then, from (4.51), for a given mean $\overline{\eta}$ and mean-square $\overline{\eta^2}$, the distribution that maximizes the unconditional error probability, $\int_0^1 f_c(\eta)\,p(\eta)\,d\eta$, has mass concentrated at two points; i.e.

$$p_{max}(\eta) = p_0\,\delta(\eta) + (1 - p_0)\,\delta(\eta - \eta_0) ,\tag{4.63}$$

where

$$p_0 = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \overline{\eta}^2} = 1 - \frac{\overline{\eta}^2}{\overline{\eta^2}}\tag{4.64}$$

and

$$\eta_0 = \overline{\eta} + \frac{\sigma_\eta^2}{\overline{\eta}} = \frac{\overline{\eta^2}}{\overline{\eta}} ,\tag{4.65}$$

with σ_η^2 denoting the variance of η. The maximum error probability is then

$$\Pr(\text{error}) = \int_0^1 Q\left(\sqrt{\frac{n_s}{2}}\,\eta\right)p_{max}(\eta)\,d\eta\tag{4.66}$$

$$= \frac{1}{2}\left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \overline{\eta}^2}\right) + Q\left(\sqrt{\frac{n_s\,\overline{\eta^2}}{2\,\overline{\eta}}}\right)\left(\frac{\overline{\eta}^2}{\sigma_\eta^2 + \overline{\eta}^2}\right) .\tag{4.67}$$

The probability density function that minimizes the unconditional error probability also has mass concentrated at two points. Setting $b = 1$ in (4.56), this probability

---

[7]Notice, the Markov moment-matching theorem requires that the function $f_c(\eta)$ be three times differentiable in the closed interval $[0,1]$. However $f_c(\eta)$, although continuous and defined at $\eta = 0$, strictly speaking, is not differentiable there. Nevertheless, we can still apply the Markov moment-matching theorem in this case. To see this, consider applying Markov moment-matching theorem on the closed interval $[\epsilon, 1]$, where $f_c(\eta)$ is differentiable, then take the limit as $\epsilon \to 0$.

density function can be expressed as

$$p_{min}(\eta) = p_1\,\delta(\eta - 1) + (1 - p_1)\,\delta(\eta - \eta_1)\,, \tag{4.68}$$

where

$$p_1 = \frac{\sigma_\eta^2}{\sigma_\eta^2 + (1 - \overline{\eta})^2}\,, \tag{4.69}$$

and

$$\eta_1 = \overline{\eta} - \frac{\sigma_\eta^2}{(1 - \overline{\eta})} = \frac{\overline{\eta} - \overline{\eta^2}}{1 - \overline{\eta}}\,. \tag{4.70}$$

The minimum error probability is then

$$\Pr(\text{error}) = \int_0^1 Q\left(\sqrt{\frac{n_s}{2}}\eta\right) p_{min}(\eta)\,d\eta \tag{4.71}$$

$$= Q\left(\sqrt{\frac{n_s}{2}}\right)\frac{\sigma_\eta^2}{\sigma_\eta^2 + (1 - \overline{\eta})^2} + Q\left(\sqrt{\frac{n_s(\overline{\eta} - \overline{\eta^2})}{2(1 - \overline{\eta})}}\right)\frac{(1 - \overline{\eta})^2}{\sigma_\eta^2 + (1 - \overline{\eta})^2}\,. \tag{4.72}$$

4.3.3 Direct Detection Unconditional Probability of Error Bounds

For analytic tractability, let us assume we are in the regime in which thermal noise is the dominant noise component. Again, OOK and PPM conditional error probability expressions in this regime differ only by a factor of 2 in the coefficient preceding η, hence we only derive the results for OOK, and state that the PPM results can be obtained by substituting in $2n_s$ for n_s in all results that follow in this section.

Analogous to the coherent detection case, we use $f_d(\eta)$ to denote the conditional error probability expression given in (4.30), to emphasize that it is a function of η.

The third derivative of $f_d(\eta)$ is

$$\frac{\partial^3 f_d(\eta)}{\partial \eta^3} = \frac{n_s^3}{8\,n_N^{3/2}\,\sqrt{2\pi}}\left[1 - \frac{n_s^2}{4n_N}\eta^2\right]e^{-\frac{n_s^2}{8\,n_N}\eta^2}\,, \tag{4.73}$$

which satisfies,

$$\frac{\partial^3 f_d(\eta)}{\partial \eta^3} > 0 \quad , \quad \text{if} \quad \eta < \frac{2\sqrt{n_N}}{n_s} \qquad \text{and} \tag{4.74}$$

$$\frac{\partial^3 f_d(\eta)}{\partial \eta^3} \leq 0 \quad , \quad \text{if} \quad \eta \geq \frac{2\sqrt{n_N}}{n_s} . \tag{4.75}$$

Note that the low conditional error probability regime corresponds to $\frac{n_s \eta}{2\sqrt{n_N}} \gg 1$, and in this regime the transition point for the third derivative, $\eta_o = \frac{2\sqrt{n_N}}{n_s}$, is guaranteed to fall between $[0, 1]$[8].

Because, in general, the third derivative of $f_d(\eta)$ does not maintain a single algebraic sign for all $\eta \in [0, 1]$, it is not possible to directly apply the Markov moment-matching lemma. However, consider the following piecewise defined function, $\tilde{f}_d(\eta)$, as an approximation to $f_d(\eta)$ over $[0, 1]$:

$$\tilde{f}_d(\eta) = \begin{cases} f_d(\eta_o) + f_d'(\eta_o)(\eta - \eta_o) + \frac{f_d''(\eta_o)}{2}(\eta - \eta_o)^2 & , \quad \text{if } \eta < \eta_o \\ f_d(\eta) & , \quad \text{if } \eta \geq \eta_o . \end{cases} \tag{4.76}$$

Notice $\tilde{f}_d(\eta)$ is the second order Taylor-series approximation to $f_d(\eta)$ for $\eta \leq \eta_o$, so that it maintains a unipolar third derivative for all $\eta \in [0, 1]$. Figure 4-6 illustrates this approximation and the resulting derivatives for $\eta_o = 0.2$.

Substituting the explicit expression for η_o in (4.76), we obtain

$$\tilde{f}_d(\eta) = \begin{cases} Q(1) - \frac{1}{\eta_o\sqrt{2\pi}} e^{-\frac{1}{2}}(\eta - \eta_o) + \frac{1}{2\eta_0^2\sqrt{2\pi}} e^{-\frac{1}{2}}(\eta - \eta_o)^2 & , \quad \text{if } \eta < \eta_o \\ f_d(\eta) & , \quad \text{if } \eta \geq \eta_o . \end{cases} \tag{4.77}$$

Considering the sign of the error term in the Taylor-series expansion, it can be verified that $\tilde{f}_d(\eta) \geq f_d(\eta), \forall \eta \in [0, 1]$. Furthermore, for any η_o value, the maximum error between $\tilde{f}_d(\eta)$ and $f(\eta)$ is 0.0216 and it occurs at $\eta = 0$.

Because, the third derivative of $\tilde{f}_d(\eta)$ is unipolar, Markov moment-matching theory can be applied to the integral $\int_0^1 \tilde{f}_d(\eta) p(\eta) d\eta$ to find its extremal values. The

---

[8]Note that the transition point for PPM is half that of OOK, i.e. $\frac{\sqrt{n_N}}{n_s}$. Therefore, if the OOK transition point falls within the interval $[0, 1]$, the PPM transition point will also be in this interval. Hence, all analysis that follows applies to the PPM case as well.
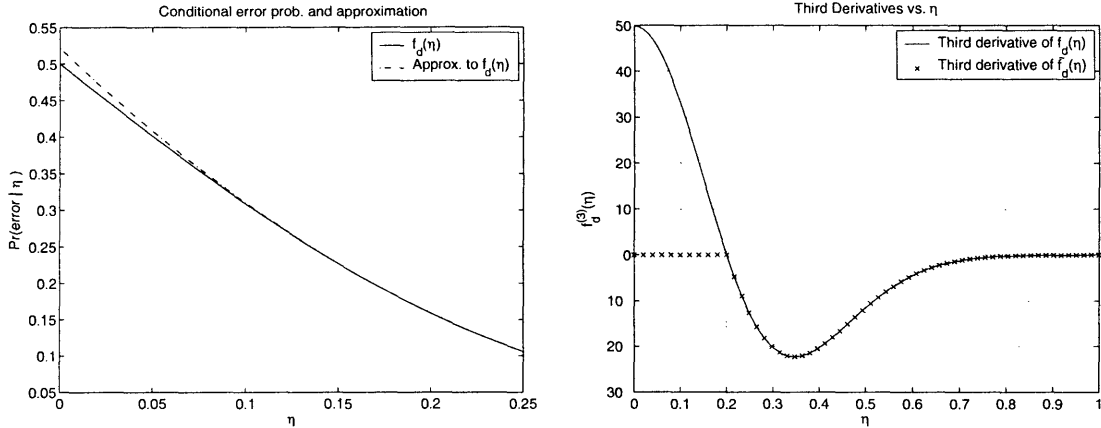
(a) $f_d(\eta)$ and its approximation for $\eta_0 = 0.2$. (b) The third-derivatives of $f_d(\eta)$ and $\tilde{f}_d(\eta)$. Functions are identical for $\eta \geq \eta_0$, thus not The derivative of the approximate function is a shown in plot. truncation of $\frac{\partial^3 f_d(\eta)}{\partial \eta^3}$.

Figure 4-6: Approximation to the direct detection conditional error probability function to obtain a unipolar function on which Markov's moment matching theory may be applied.

inequality

$$\int_0^1 f_d(\eta)\, p(\eta)\, d\eta \leq \int_0^1 \tilde{f}_d(\eta)\, p(\eta)\, d\eta \,, \tag{4.78}$$

implies that the maximum value of the right hand side is an upper bound on the direct detection unconditional error probability. The minimum of the right hand side, however, is not a lower bound on the error probability. Nevertheless, because the error between $\tilde{f}_d(\eta)$ and $f_d(\eta)$ is bounded and eventually decreasing towards 0 for large n_s values, we claim that the minimum of the left hand side of (4.78) is well approximated by the minimum of the right hand side. Therefore, we will obtain the distributions of η that yield the maximum and minimum values of the right hand side of (4.78) and use these as bounds for the unconditional error probability.

It follows from the Markov moment matching lemma results that the extremal values of $\int_0^1 \tilde{f}_d(\eta)\, p(\eta)\, d\eta$ are given by the same distributions that yield the extremal values of the unconditional error probability for coherent detection. Thus, (4.63) and (4.68) are maximizing and minimizing distributions for η in direct-detection systems, respectively. It is worthwhile to note that the third derivative of $\tilde{f}_d(\eta)$ is not strictly negative, thus minimum and maximum achieving distributions may not be unique.

88

Error probability bounds achieved by these distributions for coherent and direct detection cases are plotted in Figures 4-11 and 4-13 when σ_η^2 is much smaller than $\bar{\eta} - \bar{\eta}^2$, and in Figures 4-10 and 4-12, when σ_η^2 is close to $\bar{\eta} - \bar{\eta}^2$.[9] The worst-case distribution has an impulse at $\eta = 0$ regardless of the values of $\bar{\eta}$ and σ_η^2,[10] which implies that in the worst atmospheric state there is a positive probability that a field with power P_s will be transmitted and none of it will be collected at the receiver aperture. Therefore, for $n_s \gg 1$, the error probability is dominated by this zero-transmission event. On the other hand, the minimum error probability distribution always has an impulse at $\eta = 1$. This corresponds to a positive probability of receiving all the power that has been transmitted. For $n_s \gg 1$, the error probability with perfect power transfer is negligibly small, thus the error probability in this regime is dominated by the impulse at η_1. Hence, as can be verified from the Chernoff bound, the minimum error probability decays exponentially in the high transmit power regime.

4.3.4 Unimodal Distribution Error Probability Bounds

Based on physical observations and general understanding of atmospheric turbulence, it is highly unlikely that the maximum power coupling between the transmitter and receiver would behave as a two-level switch. Therefore, although the distributions found in the previous section yield the extremal values for the unconditional error probability, they are not realistic distributions for the maximum eigenvalue of the atmospheric propagation kernel. Hence, it is desirable to impose further constraints on the probability distribution function of η, so that distributions consisting of a finite collection of impulses are eliminated. One such constraint is to assume that η has a unimodal probability distribution over $[0, 1]$.

Although it is not possible to rigorously justify that η must have a unimodal

---

[9] For a distribution on $[0, 1]$ with mean $\bar{\eta}$, the maximum possible variance is $\bar{\eta}(1 - \bar{\eta})$.

[10] Because η is the maximum eigenvalue of the atmospheric propagation kernel, $\eta = 0$ implies that the kernel is *equal* to zero, which in turn, requires that $\chi(\bar{\rho'}, \bar{\rho}) = -\infty$ for all $\bar{\rho} \in R_1$ and $\bar{\rho'} \in R_2$, which is a zero probability event. However, it is still possible to have an impulse in $p(\eta)$ at a location ϵ as $\epsilon \to 0$. Hence this upper bound is strictly not achievable, yet possible to approach in the limit.

(a) Lognormal probability density function **(b)** I-K probability density function
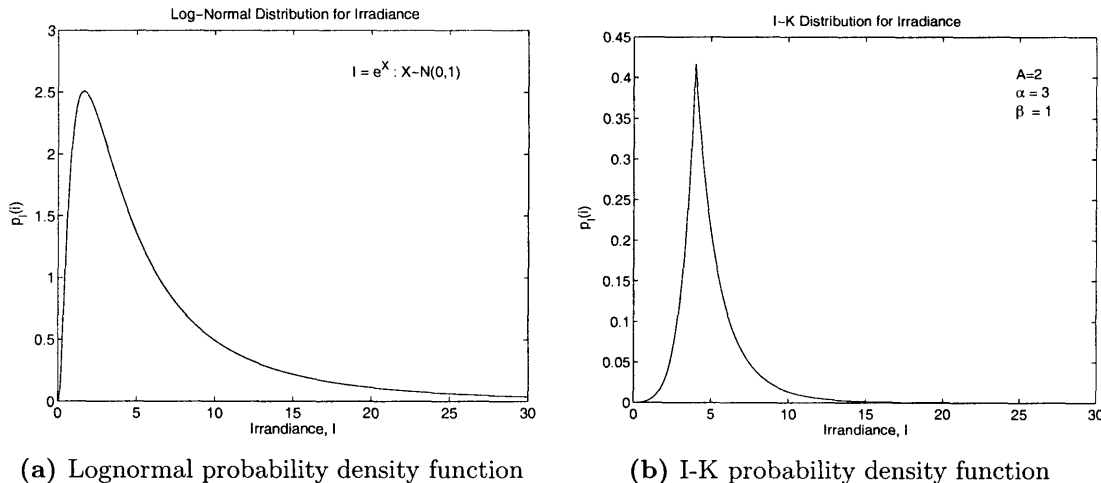
Figure 4-7: Some unimodal probability distributions, used as models for far-field irradiance.

distribution, consulting models characterizing far field power transfer statistics, it is possible to argue its plausibility. Many theoretical and heuristic models have been developed to characterize turbulent effects in far-field atmospheric propagation. The most well-known analytic theory developed for the weak turbulence regime is the Rytov theory of atmospheric turbulence [6]. This model predicts a log-normal distribution for the irradiance of the wave at the receiver plane, which is a unimodal probability distribution. Other models have also been developed for the strong fluctuation and weak fluctuation regimes [6]. The I-K distribution, which obtains its name from the presence of modified Bessel functions of type I and K in the probability density function, is valid in both regimes and yields a unimodal distribution. Some other theories predict a lognormally modulated exponential distribution in the strong fluctuation regime, which is also unimodal. Some of these probability distributions are shown in Figure 4-7. The fact that far-field maximum eigenvalue probability distribution models, heuristic or analytic, predict unimodal distributions, is additional motivation to examine the behavior of the error probabilities in the near field regime when the maximum eigenvalue is assumed to have a unimodal distribution.

A commonly accepted definition of unimodality, provided by Khintchine, is as follows [25]:

Definition 4.3.1 (Unimodal distribution). A real random variable X or its dis-

tribution function $F_X(x) = \Pr(X \le x)$, is *unimodal* about a *mode* ν if F is convex on $(-\infty, \nu)$ and concave on (ν, ∞).

Further work in the field, has led to very useful results in characterizing unimodal distributions [25, 26]. We present one result relevant to our analysis, due to Shepp:

Theorem 4.3.2 (Representation of a unimodal random variable). *A distribution function $F(x)$ on $\mathbb{R}$ is unimodal about 0, if, and only if, there exist independent random variables U and Z such that U is uniform on $[0, 1]$ and the product UZ has distribution function $F(x)$.*

A detailed treatment of the development leading to this theorem and its proof is presented in [25, Chapter 1]. It follows trivially that if a random variable X is unimodal about a mode ν, then $X - \nu$ is unimodal about 0. Hence, from Theorem 4.3.2, $X = \nu + UZ$, where U and Z are as explained above.

Assume that η has a unimodal distribution with a given mode $\nu \in [0, 1]$, as well as a given mean $\bar\eta$ and variance σ_η^2. Then,

$$\eta = \nu + UZ , \qquad (4.79)$$

where U is a uniform random variable over $[0, 1]$ and independent of Z. Z can be any random variable such that $\nu + UZ$ yields a distribution for η satisfying the support, mean and variance constraints on η. These three constraints on η imply three corresponding constraints on the distribution of Z. First, because $\eta \in [0, 1]$, Z must also have finite support. In particular,

$$Z \in [-\nu, 1 - \nu] . \qquad (4.80a)$$

Second, Z must have mean

$$\overline{Z} \triangleq E[Z] = 2(\bar\eta - \nu) . \qquad (4.80b)$$

91

Finally, Z must have variance

$$\sigma_Z^2 \triangleq E\left[(Z - \overline{Z})^2\right] = 3\sigma_\eta^2 - (\overline{\eta} - \nu)^2 \ , \tag{4.80c}$$

or, equivalently second moment

$$\overline{Z^2} \triangleq E\left[Z^2\right] = 3\left(\overline{\eta^2} - 2\nu\,\overline{\eta} + \nu^2\right) \ . \tag{4.80d}$$

It is evident from (4.80a – 4.80d) that the range of possible modes for η is not independent of the prescribed support, mean and variance constraints. In the discussion that follows, we will assume that the mode has been chosen such that equations in (4.80) constitute a consistent set of constraints. The interval a mode must lie in for a given set of constraints will be determined later.

Coherent detection systems

For an OOK, coherent detection system, the unconditional error probability is[11]

$$\Pr(\text{error}) = E_\eta\left[Q\left(\sqrt{\frac{n_s}{2}}\eta\right)\right] \ , \tag{4.81}$$

where, for clarity, the subscript in the expectation notation indicates the random variable in consideration. Combining this result and (4.79), we have that

$$\Pr(\text{error}) = E_{U,Z}\left[Q\left(\sqrt{\frac{n_s}{2}(\nu + UZ)}\right)\right] \tag{4.82}$$

$$= \int_{-\nu}^{1-\nu} \int_0^1 Q\left(\sqrt{\frac{n_s}{2}(\nu + uz)}\right) p(z)\,du\,dz \tag{4.83}$$

$$= \int_{-\nu}^{1-\nu} \left\{\int_0^1 Q\left(\sqrt{\frac{n_s}{2}(\nu + uz)}\right) du\right\} p(z)\,dz \tag{4.84}$$

$$= \int_{-\nu}^{1-\nu} g_c(z;\nu)\,p(z)\,dz \ . \tag{4.85}$$

---

[11]To obtain results for PPM, coherent detection systems, substitute $2n_s$ for n_s throughout the results in this subsection.

The parametric dependence of $g_c(z; \nu)$ on ν is made explicit by including ν in the list of arguments. With (4.85) and (4.80a) – (4.80d), we have re-defined our maximization and minimization problem in terms of the random variable Z. The objective can now be stated as finding the probability distributions for Z on $[-\nu, 1 - \nu]$ which yield the extremal values of (4.85), given $\overline{Z}$ and σ_Z^2.

This formulation warrants, once again, the use of the Markov moment matching problem if it is true that $g_c(z; \nu)$ has a non-positive or non-negative third derivative. Referring back to (4.61), we can write g_c as

$$ g_c(z; \nu) = \int_0^1 f_c(\nu + uz) \, du \, , \tag{4.86} $$

and note that

$$ \frac{\partial f_c(\nu + uz)}{\partial z} = u \left. \frac{\partial f_c(x)}{\partial x} \right|_{x = \nu + uz} . \tag{4.87} $$

Therefore, the third derivative of g_c is

$$ \frac{\partial^3 g_c(z; \nu)}{\partial z^3} = \int_0^1 u^3 \left. \frac{\partial^3 f_c(x)}{\partial x^3} \right|_{x = \nu + uz} du \quad < 0 \, . \tag{4.88} $$

The strict negativity follows from the positivity theorem for integrals in standard analysis [27, Chapter 19].

Having satisfied all conditions of Lemma 4.3.1, we conclude that the unique distribution of Z which yields the maximum unconditional error probability for given mean and variance values, specified in (4.80b) and (4.80c) respectively, is

$$ p_{max}(z) = p_0 \, \delta(z + \nu) + (1 - p_0) \, \delta(z - z_0) \, , \tag{4.89} $$

where

$$ p_0 = \frac{\sigma_Z^2}{\sigma_Z^2 + (\overline{Z} + \nu)^2} \tag{4.90} $$

and

$$z_0 = \overline{Z} + \frac{\sigma_Z^2}{\overline{Z} + \nu} \, , \tag{4.91}$$

for a given valid mode ν.

Using (4.79) and (4.89), we can derive the corresponding distribution for the unimodal density of η given a mode ν, satisfying mean $\overline{\eta}$ and variance σ_η^2. The probability density function for η depends on whether z_0 is positive, negative or equal to zero. The three possibilities for the density of η are shown in Figure 4-8 and the analytic expressions are given by

$$p_{max}(\eta) = \begin{cases} \frac{p_0}{\nu} & 0 \leq \eta < \nu \, , \\ \frac{1-p_0}{z_0} & \nu \leq \eta < z_0 + \nu \, , \\ 0 & \text{otherwise} \, , \end{cases} \tag{4.92a}$$

for $z_0 > 0$, and

$$p_{max}(\eta) = \begin{cases} \frac{p_0}{\nu} & 0 \leq \eta \leq \nu + z_0 \, , \\ \frac{p_0}{\nu} + \frac{1-p_0}{|z_0|} & \nu + z_0 \leq \eta < \nu \, , \\ 0 & \text{otherwise} \, , \end{cases} \tag{4.92b}$$

for $z_0 < 0$. When $z_0 = 0$, the distribution contains an impulse at $\eta = \nu$ of area $1 - p_0$, shown in Figure 4-8(c).

The maximum unconditional error probability achieved by this distribution is computed to be

$$\Pr(\text{error}) = \frac{p_0}{\nu} \int_0^\nu Q\left(\sqrt{\frac{n_s t}{2}}\right) dt + \frac{1 - p_0}{z_0} \int_\nu^{\nu+z_0} Q\left(\sqrt{\frac{n_s t}{2}}\right) dt \, , \tag{4.93}$$

where the second term is replaced by $(1 - p_0) Q\left(\sqrt{\frac{n_s \nu}{2}}\right)$ if $z_0 = 0$.

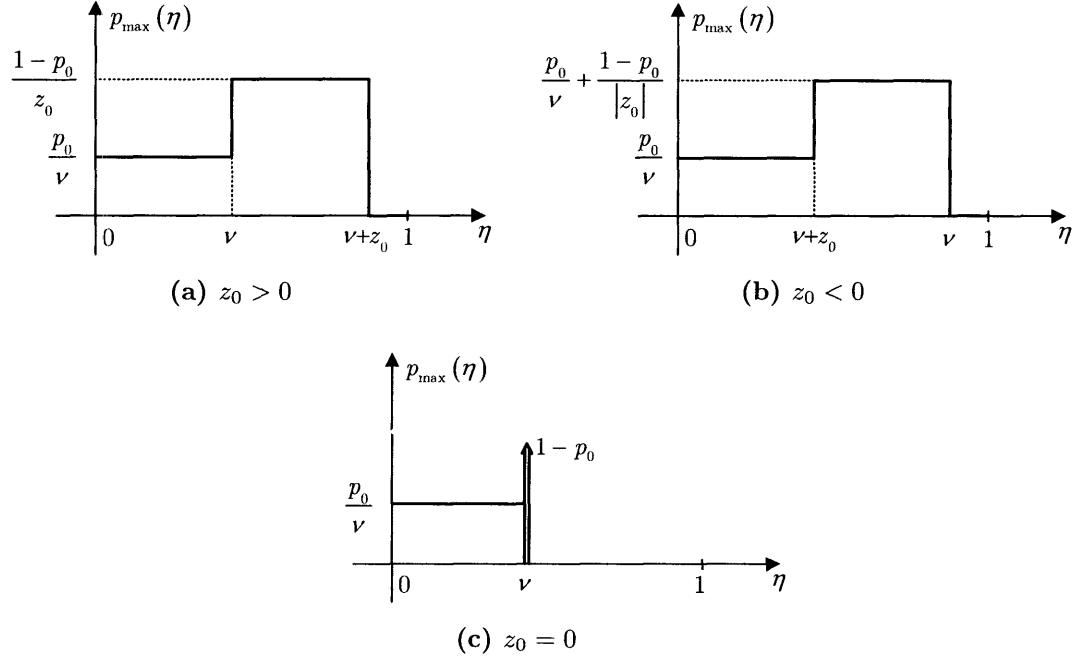Referring to (4.56), the unique distribution for Z which minimizes the uncondi-

(a) $z_0 > 0$ **(b)** $z_0 < 0$

(c) $z_0 = 0$

Figure 4-8: The unimodal probability density function of η satisfying a mean constraint $\bar{\eta}$ and a variance constraint σ_η^2, which maximizes the probability of error. ν is assumed to be some known valid mode.

tional error probability expression is also a two-impulse distribution given by

$$p_{min}(z) = p_1\,\delta(z - 1 + \nu) + \left(1 - p_1\right)\delta(z - z_1)\,, \tag{4.94}$$

where

$$p_1 = \frac{\sigma_Z^2}{\sigma_Z^2 + (1 - \nu - \overline{Z})^2} \tag{4.95}$$

and

$$z_1 = \overline{Z} - \frac{\sigma_Z^2}{1 - \nu - \overline{Z}}\,, \tag{4.96}$$

for a given mode ν.

The corresponding distribution for η, again, depends on an algebraic sign, this time that of z_1. The probability density function of η for the three possible cases are
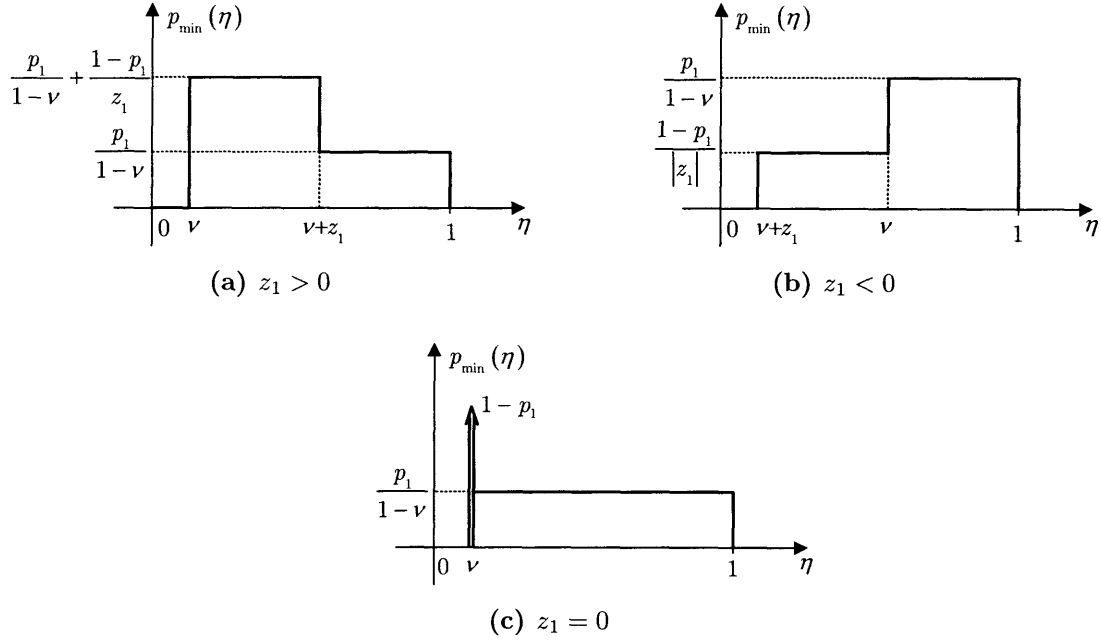
95

(a) $z_1 > 0$

(b) $z_1 < 0$

(c) $z_1 = 0$

Figure 4-9: The unimodal probability density function of η satisfying a mean constraint $\bar{\eta}$ and a variance constraint σ_η^2, which minimizes the probability of error. ν is assumed to be some known valid mode.

plotted in Figure 4-9, and expressed as

$$p_{min}(\eta) = \begin{cases} \frac{p_1}{1-\nu} + \frac{1-p_1}{z_1} & \nu \leq \eta < \nu + z_1 \, , \\ \frac{p_1}{1-\nu} & \nu + z_1 \leq \eta < 1 \, , \\ 0 & \text{otherwise.} \end{cases} \qquad (4.97a)$$

for $z_1 > 0$, and

$$p_{min}(\eta) = \begin{cases} \frac{1-p_1}{|z_1|} & \nu + z_1 \leq \eta < \nu \, , \\ \frac{p_1}{1-\nu} & \nu \leq \eta < 1 \, , \\ 0 & \text{otherwise.} \end{cases} \qquad (4.97b)$$

for $z_1 < 0$. As with the maximizing distribution, when $z_1 = 0$, the distribution contains an impulse at $\eta = \nu$ of area $1 - p_1$, shown in Figure 4-9(c).

96

The resulting minimum unconditional probability of error is then

$$\text{Pr(error)} = \frac{p_1}{1-\nu} \int_\nu^1 Q\left(\sqrt{\frac{n_s t}{2}}\right) dt + \frac{1-p_1}{z_1} \int_\nu^{\nu+z_1} Q\left(\sqrt{\frac{n_s t}{2}}\right) dt , \qquad (4.98)$$

where the second term is replaced by $(1-p_1)\, Q\left(\sqrt{\frac{n_s \nu}{2}}\right)$ if $z_1 = 0$.

Note that these distributions are also the maximum and minimum achieving distributions for the PPM, coherent detection case. However, the maximum and minimum error probability expressions are not the same. In particular, PPM extremal error probabilities are obtained by replacing n_s with $2n_s$ in (4.93) and (4.98).

The extremal unconditional error probability results above are for a particular mode, ν. To find probability error bounds valid for all modes, we simply maximize and minimize the results above over all possible modes. For a given mean $\bar\eta > 0.5$ (which, in general, holds true in the near-field regime) and variance $\sigma_\eta^2 > 0$, the mode ν must satisfy[12]

$$\max\left\{0\,,\bar\eta - \sqrt{3}\sigma_\eta\,, \frac{3\overline{\eta^2} - 2\bar\eta}{2\bar\eta - 1}\right\} \le \nu \le \min\left\{1\,,\bar\eta + \sqrt{3}\sigma_\eta\right\} . \qquad (4.99)$$

This maximization and minimization is carried out numerically and results are shown in the figures at the end of this chapter.

Direct detection systems

The derivation of the maximizing and minimizing unimodal probability densities of η for direct detection receivers parallels that of the coherent detection case. We continue to assume that the dominant noise source in the receiver is thermal noise. The derivation in this section is carried out for OOK systems, but PPM results can be obtained by substituting $2n_s$ for n_s.

Using (4.79), we write the unconditional probability of error in terms of the statis-

---

[12]See Appendix B for the derivation of the range of possible modes for a prescribed support, mean and variance.

tics of Z:

$$\Pr(\text{error}) = \int_{-\nu}^{1-\nu} g_d(z;\nu)\,p(z)\,dz\,, \tag{4.100}$$

where

$$g_d(z;\nu) = \int_0^1 f_d(\nu + uz)\,du \tag{4.101}$$

and $f_d(\eta)$ is the conditional probability of error, given in (4.30). This gives

$$\frac{\partial^3 g_d(z;\nu)}{\partial z^3} = \int_0^1 \frac{\partial^3 f_d(\nu + uz)}{\partial z^3}\,du \tag{4.102}$$

$$= \int_0^1 u^3 \left.\frac{\partial^3 f_d(x)}{\partial x^3}\right|_{x=\nu+uz}\,du \tag{4.103}$$

which is not unipolar for all $z \in [-\nu, 1 - \nu]$. However, the approximation to $g_d(z;\nu)$,

$$\tilde{g}_d(z;\nu) = \int_0^1 \tilde{f}_d(\nu + uz)\,du\,, \tag{4.104}$$

has a non-positive third derivative[13], the inequality being strict if $\nu > \eta_o$. In near-field applications $\nu \approx 1$, and for any acceptable level of input signal power $\eta_o \ll 0.5$, therefore the third derivative of (4.104) is, in general, strictly negative.

Applying the Markov moment-matching lemma to $\int_{-\nu}^{1-\nu} \tilde{g}_d(z;\nu)\,p(z)\,dz$, we can obtain the distributions of Z which will maximize and minimize this expression. Because $\tilde{g}_d(z;\nu) \geq g_d(z;\nu)$, the maximum of $\int_{-\nu}^{1-\nu} \tilde{g}_d(z;\nu)\,p(z)\,dz$ will also be an upper bound on (4.100). On the other hand, because the error between g_d and its approximation is small, the minimum obtained via Markov's moment matching lemma is a good approximation to the minimum unconditional error probability. The distributions which achieve this upper bound and approximate minimum are *identical* to those that achieve the extremal values of the coherent detection unconditional error probability [14]. Thus the maximizing density for Z is given by (4.89) and the minimizing density is given by (4.94). Consequently, Figure 4-8 and (4.92) give the maximum eigenvalue unimodal distribution which maximizes the error probability,

---

[13] $\tilde{f}_d$ is defined in (4.77)
[14] See concluding remarks in Lemma 4.3.1.

while Figure 4-9 and (4.97) are the unimodal distributions of η that minimize the unconditional error probability.

The resulting maximum unconditional probability of error for the OOK case is given by

$$\Pr(\text{error}) = \frac{p_\nu}{\nu} \int_0^\nu Q\left(\frac{n_s t}{2\sqrt{n_N}}\right) dt + \frac{1 - p_\nu}{z_0} \int_\nu^{\nu + z_0} Q\left(\frac{n_s t}{2\sqrt{n_N}}\right) dt , \qquad (4.105)$$

the second term being replaced by $(1 - p_0) Q\left(\frac{n_s \nu}{2\sqrt{n_N}}\right)$ if $z_0 = 0$. The corresponding minimum unconditional OOK error probability is

$$\Pr(\text{error}) = \frac{p_1}{1 - \nu} \int_\nu^1 Q\left(\frac{n_s t}{2\sqrt{n_N}}\right) dt + \frac{1 - p_1}{z_1} \int_\nu^{\nu + z_1} Q\left(\frac{n_s t}{2\sqrt{n_N}}\right) dt , \qquad (4.106)$$

in which the second term is replaced by $(1 - p_1) Q\left(\frac{n_s \nu}{2\sqrt{n_N}}\right)$ if $z_1 = 0$.

The maximum and minimum error probability expressions for a PPM, direct detection communication system are obtained from (4.105) and (4.106) respectively, by substituting in $2n_s$ for n_s.

The maximum and minimum probability of error curves for coherent and direct detection receiver systems are shown in Figures 4-10 – 4-13. The probability distributions for η that generate these error probability bounds are shown in Figures 4-14 and 4-15. Notice, direct detection systems require approximately three orders of magnitude more signal photons than coherent detection systems to achieve comparable sensitivity. This difference is due to the mixing with the strong local oscillator in coherent detection, which amplifies the incoming signal levels above thermal noise levels. The trade off, however, is the increased complexity and implementation cost of the coherent detection receiver. Determining the optimum local oscillator field as well as efficient mixing of the signal and local oscillator fields are some of the challenges to overcome.

It is seen from the figures that the two-impulse worst case distribution reaches a floor for $n_s \gg 1$. This is because, in this worst-case scenario, there is a positive probability that none of the transmitted power will couple into the receiver aperture.

99

Therefore, no matter how many signal photons are transmitted, the probability of this event will dominate the overall error probability. The worst-case unimodal distribution on the other hand, is a two-piece, piecewise constant distribution with a non-zero probability density out to $\eta = 0$. Thus, it can be shown that for $n_s \gg 1$, when most of the area underneath the conditional error probability expressions are concentrated around $\eta = 0$, the unconditional error probability results will be proportional to $\frac{1}{n_s}$.

The best-case error probability results, on the other hand, are dramatically different than the worst-case results. The error probability results when no fading is present decrease exponentially with n_s. The best-case two-impulse distribution has impulses at 1 and some other intermediate point, and therefore the error probability consists of two exponentially decaying terms. As n_s gets large the contribution to the probability of error due to the impulse at $\eta = 1$ will become negligible and the error probability will be dominated by the intermediate point. Thus, for $n_s \gg 1$, the error probability will decay exponentially, at a fraction of the rate of the non-fading case. The minimizing unimodal distribution is a two-piece piecewise constant distribution (with the possible exception of one of the pieces degenerating to an impulse), with a density that is zero for η values below some positive threshold. Thus, when n_s is large enough, the area underneath the tail of the conditional probability expressions will be very small, and the error probability will decay exponentially, at again a fraction of the non-fading rate.

As a consequence of the discussion in the previous paragraph, we observe from the figures that for a given mean of η, the error probabilities of the best-case distributions converge to the non-fading error probability as the variance becomes smaller. This is because a small variance combined with a very high mean will result in almost identical best-case unimodal and two-impulse distributions, with all probability concentrated around $\bar{\eta} \approx 1$.
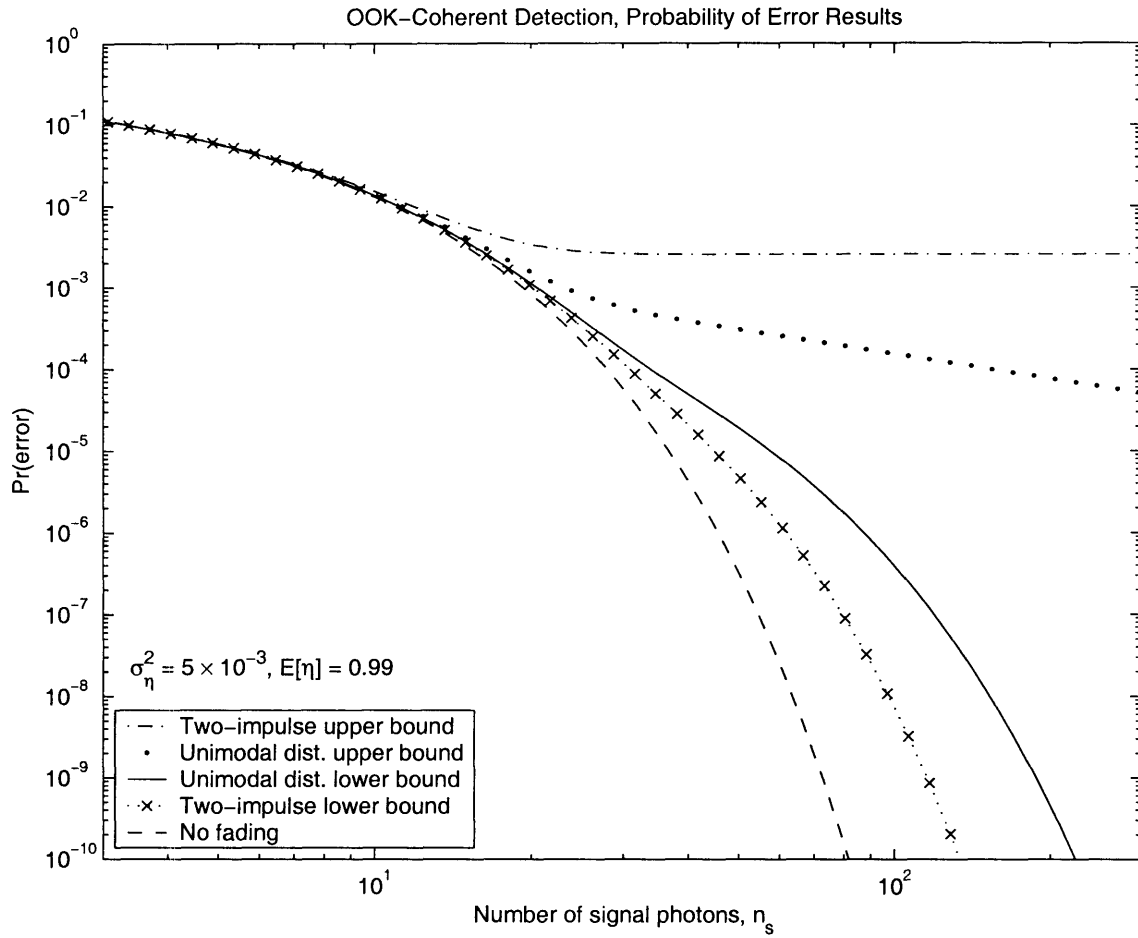
Figure 4-10: Coherent detection maximum and minimum probability of error results when the variance of η is close to its maximum, $\overline{\eta} - \overline{\eta}^2$. Plot shows error probability curves for OOK. PPM results are obtained by shifting all curves 3 dB to the left.
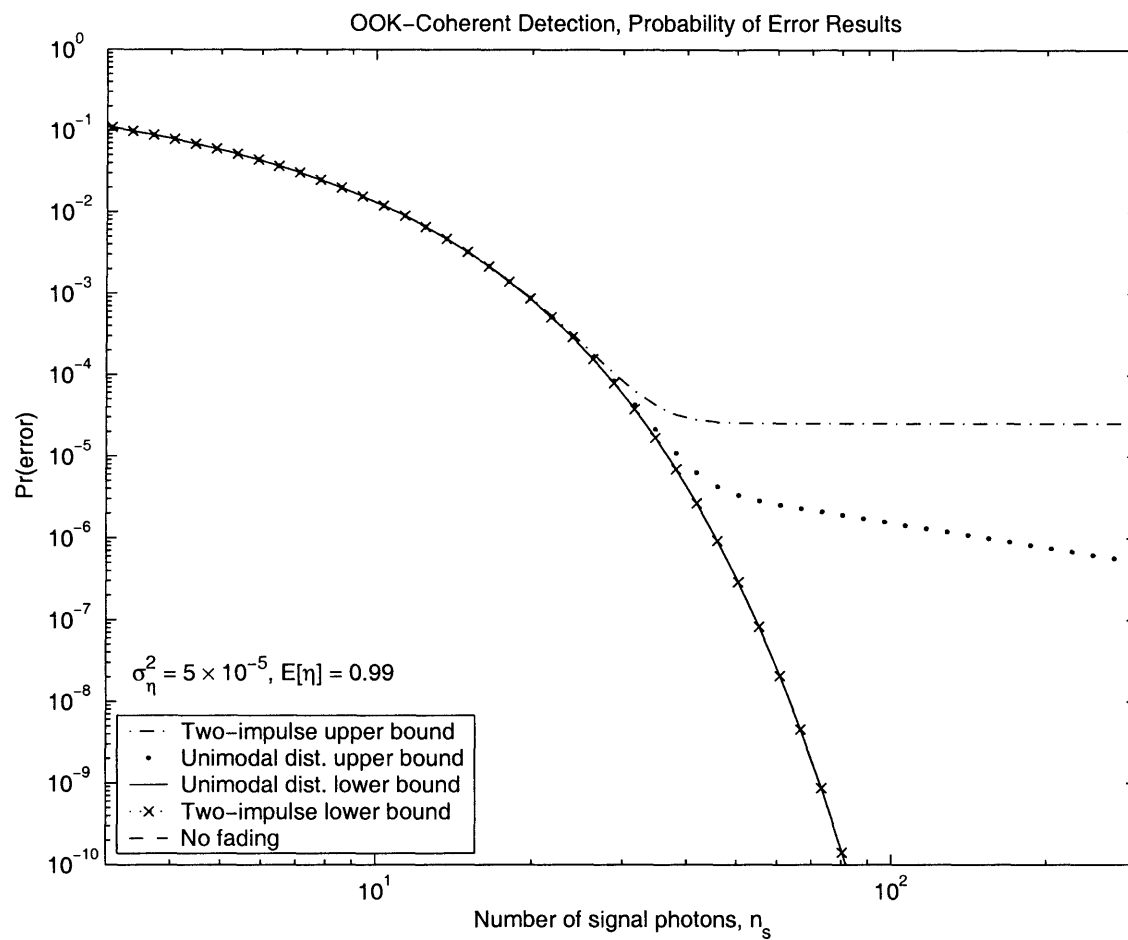
Figure 4-11: Coherent detection maximum and minimum probability of error results when the variance of η is small relative to its maximum, $\bar{\eta} - \bar{\eta}^2$. Plot shows error probability curves for OOK. PPM results are obtained by shifting all curves 3 dB to the left.
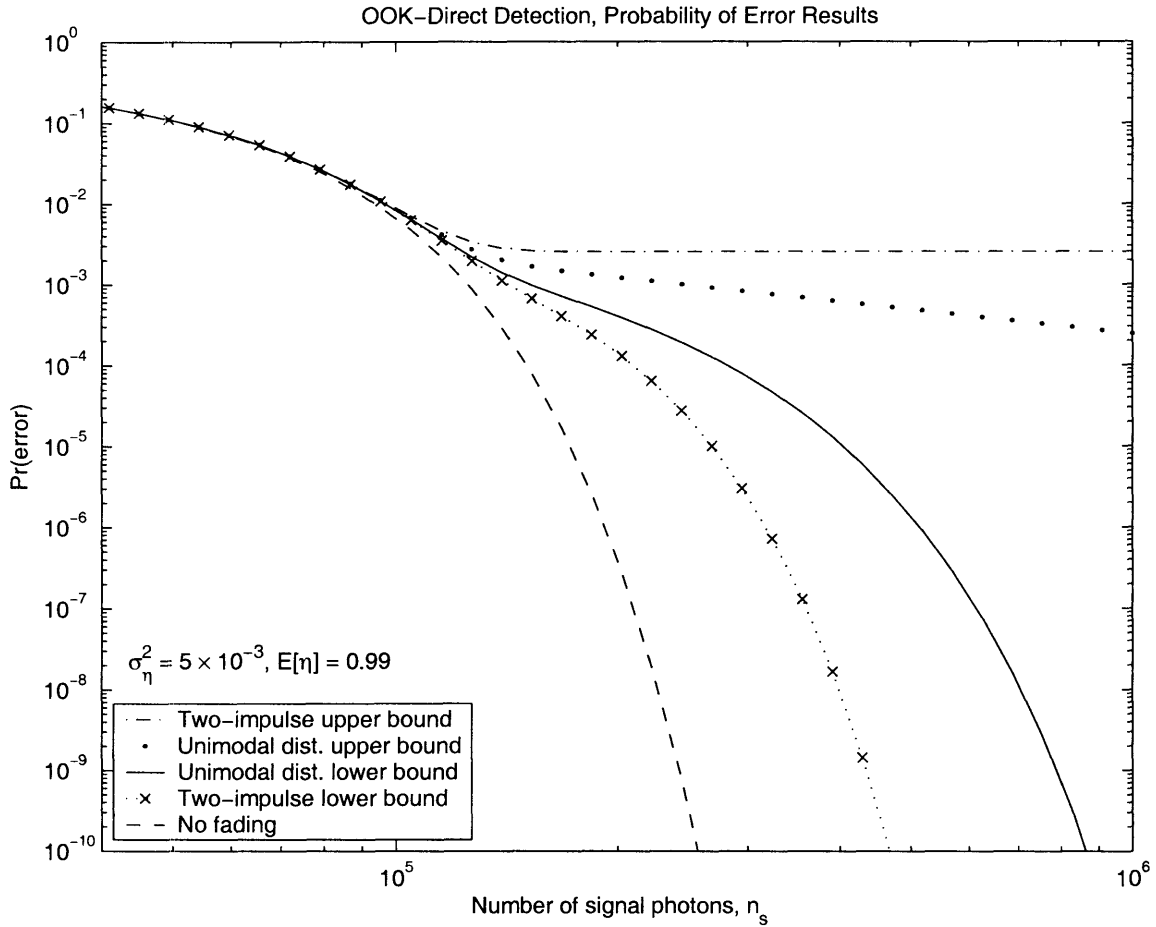
Figure 4-12: Direct detection maximum and minimum probability of error results when the variance of η is close to its maximum, $\overline{\eta} - \overline{\eta}^2$. $n_N = 4 \times 10^8$, an approximate value for gigabit communication rates when the detector output sees a 50Ω load resistance at room temperature. Plot shows error probability curves for OOK. PPM results are obtained by shifting all curves 3 dB to the left.

Figure 4-13: Direct detection maximum and minimum probability of error results when the variance of η is small relative to its maximum, $\overline{\eta} - \overline{\eta}^2$. $n_N = 4 \times 10^8$, an approximate value for gigabit communication rates when detector output sees 50Ω load resistance at room temperature. Plot shows error probability curves for OOK. PPM results are obtained by shifting all curves 3 dB to the left.
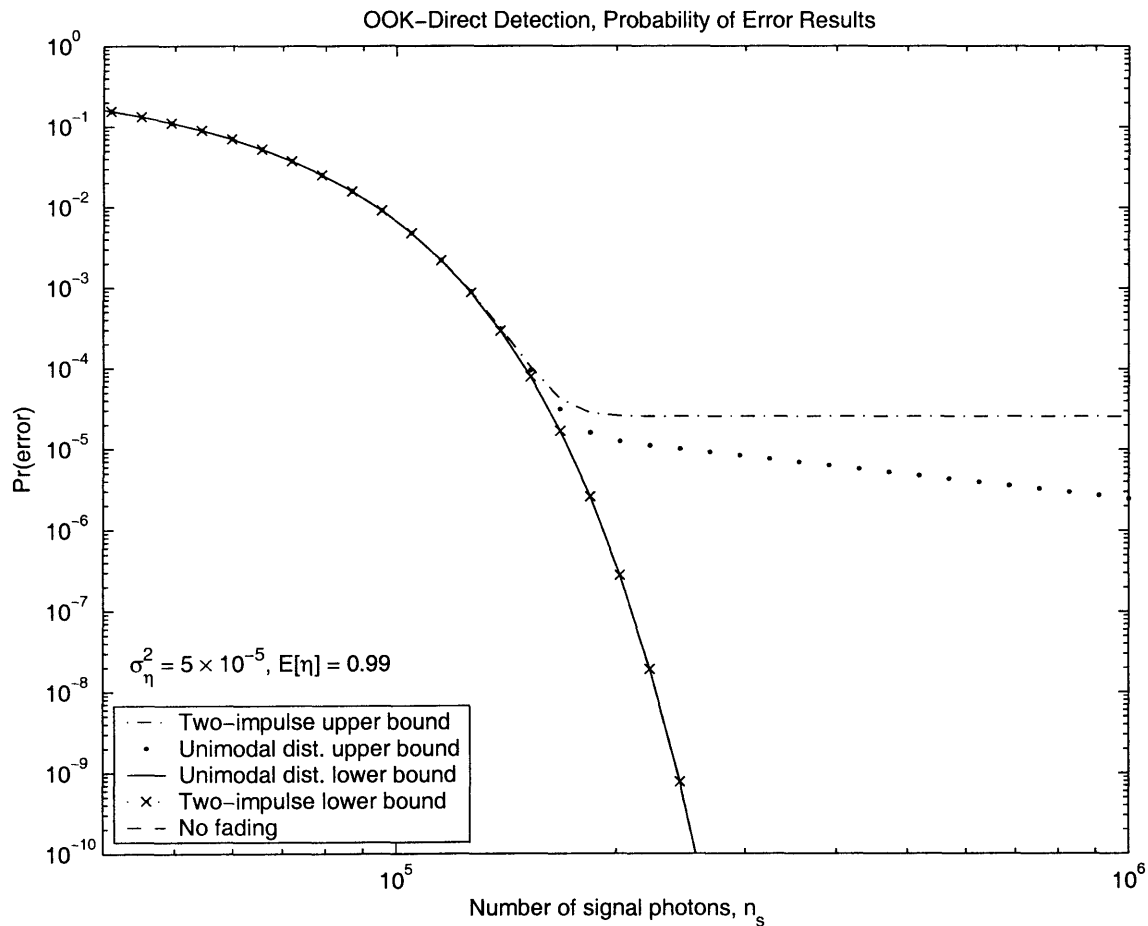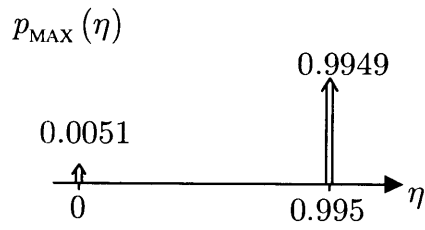
(a) Maximizing distribution.
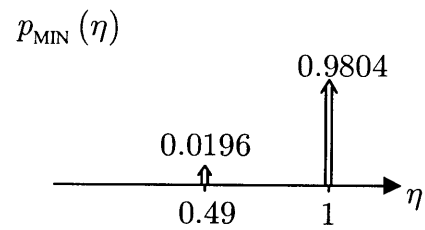
(b) Minimizing distribution.

(c) Maximizing unimodal distribution.

(d) Minimizing unimodal distribution.

Figure 4-14: Extremum achieving distributions of η in previous figures, when $\bar{\eta} = 0.99$ and $\sigma_\eta^2 = 5 \times 10^{-3}$.
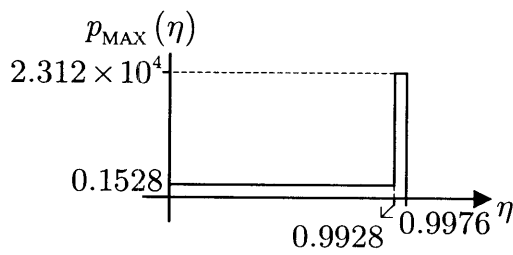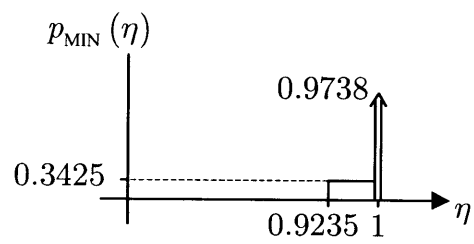
(a) Maximizing distribution.
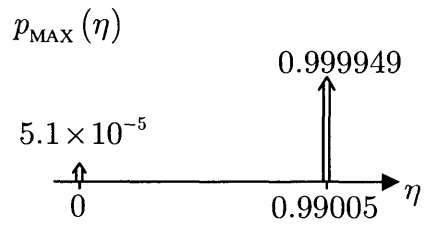

(b) Minimizing distribution.
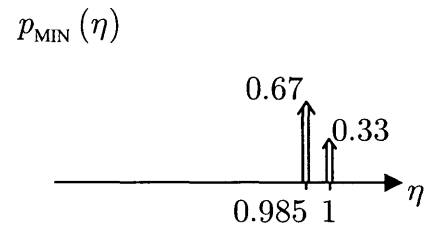

(c) Maximizing unimodal distribution.
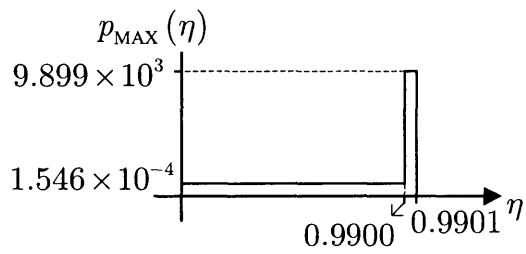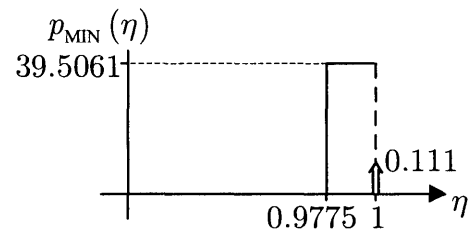

(d) Minimizing unimodal distribution.

Figure 4-15: Extremum achieving distributions of η in previous figures, when $\bar{\eta} = 0.99$ and $\sigma_\eta^2 = 5 \times 10^{-5}$.

Chapter 5

A Heuristic Model for Maximum Power Transfer

5.1 Introduction

In Chapter 3 various lower bounds on the mean of η were obtained and the possibility of obtaining a bound on the second moment was discussed. Chapter 4 built upon these results by deriving the distributions for η which achieved maximum and minimum unconditional error probabilities given the mean and variance. In this chapter, we will extend this work by considering a specific distribution, namely the Beta distribution.

The Beta distribution is of particular interest as a distribution of the maximum eigenvalue of the propagation kernel for a number of reasons. First, it is a continuous distribution on $[0, 1]$ which is completely characterized by two non-negative parameters. Therefore, mean and variance constraints on η are sufficient to fully specify the probability density function. Second, the Beta distribution is unimodal if either of its two parameters are greater than 1 and otherwise it becomes a continuous approximation to a pair of impulses at $\eta = 0$ and $\eta = 1$. Hence, there are grounds for comparison to the bounds computed in the previous chapter.

5.2   The Beta distribution

The probability density function of a Beta random variable with parameters a and b is given by

$$p(\eta) = \frac{\eta^{a-1} (1 - \eta)^{b-1}}{\beta(a, b)} \qquad \text{for } a > 0,\, b > 0 , \tag{5.1}$$

where $\beta(a, b)$ is the Beta function defined as [28],

$$\beta(a, b) = \int_0^1 \eta^{a-1} (1 - \eta)^{b-1} \, d\eta , \qquad a > 0,\, b > 0 . \tag{5.2}$$

It is evident from (5.1) that if $a < 1$ $p(\eta)$ has a singularity at $\eta = 0$, if $a = 1$ $p(0) = \frac{1}{\beta(a,b)}$, and if $a > 1$, $p(0) = 0$. Similar behavior with respect to b occurs at $\eta = 1$. Figure 5-1 shows the general shape of a Beta probability density function for various values of the parameters a and b.

The moments of a Beta distribution are given in terms of Beta functions by

$$E\left[\eta^k\right] = \frac{\beta(a + k, b)}{\beta(a, b)} . \tag{5.3}$$

After simplifications, the mean becomes

$$\bar{\eta} = \frac{a}{a + b} = \frac{1}{1 + \frac{b}{a}} , \tag{5.4}$$

and the variance,

$$\sigma_\eta^2 = \frac{a\,b}{(a + b)^2 \,(a + b + 1)} = \frac{\frac{b}{a}}{\left(1 + \frac{b}{a}\right)^2 (a + b + 1)} . \tag{5.5}$$

The second equalities in (5.4) and (5.5) are convenient because, given the mean and variance of η, they are easily solved for $\frac{b}{a}$ and $(a + b)$. We find that

$$a = \bar{\eta} \left(\frac{\bar{\eta} - \bar{\eta}^2}{\sigma_\eta^2} - 1 \right) \tag{5.6}$$
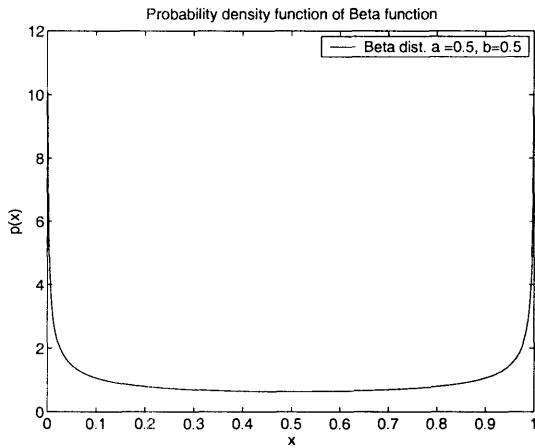
(a) $a < 1$ and $b < 1$.
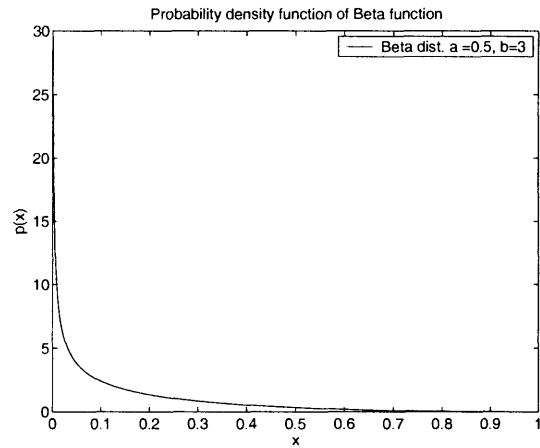
(b) $a < 1$ and $b > 1$.
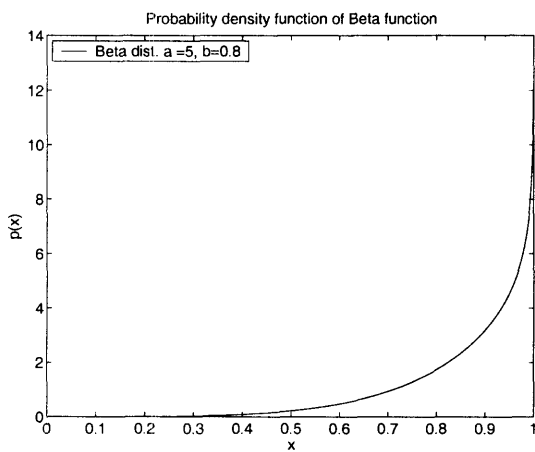
(c) $a > 1$ and $b < 1$.
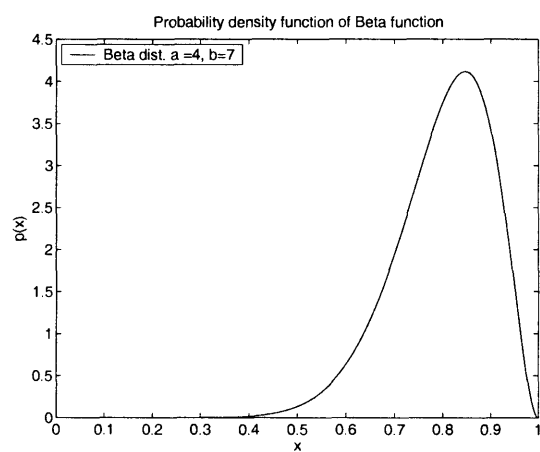
(d) $a > 1$ and $b > 1$.

Figure 5-1: Some Beta probability density function plots.

and

$$b = (1 - \bar{\eta}) \left(\frac{\bar{\eta} - \bar{\eta}^2}{\sigma_\eta^2} - 1 \right) . \tag{5.7}$$

Note that the necessary conditions $a > 0$ and $b > 0$ will be satisfied for $0 < \sigma_\eta^2 < (\bar{\eta} - \bar{\eta}^2)$. In other words, there exists a Beta distribution which satisfies any prescribed mean and variance except the limiting cases, $\sigma_\eta^2 = \bar{\eta} - \bar{\eta}^2$ (which is achieved by a two-impulse distribution with impulses at $\eta = 0$ and $\eta = 1$) and $\sigma_\eta^2 = 0$ (which is achieved by a unity impulse at $\eta = \bar{\eta}$).

As a final remark, the mode of the Beta distribution for $a > 1$ and $b > 1$ is [29]

$$\nu = \frac{a - 1}{a + b - 2} . \tag{5.8}$$

5.3   Unconditional Error Probability Results

Chapter 4 concluded that the unconditional error probability for large n_s values is strongly dependent on the probability of observing deep fades; i.e. the relevant part of the probability density function of η is the lower end of the $[0, 1]$ interval. Consequently, when η is Beta distributed, we expect that the unconditional probability of error results for large n_s will mostly depend on whether the Beta distribution has a singularity at $\eta = 0$ or not.

Because we have been unsuccessful in numerically evaluating the variance of the maximum eigenvalue of the atmospheric propagation kernel, we will resort to considering the performance of a set of Beta distributions with equal mean, yet different variance values within the open interval $\left(0, \bar{\eta} - \bar{\eta}^2\right)$. Using the expressions for the two parameters of the Beta distribution in (5.6) and (5.7), it can be concluded that, if

$$\sigma_\eta^2 > \frac{\bar{\eta}^2 (1 - \bar{\eta})}{1 + \bar{\eta}} = \psi_a , \tag{5.9}$$

the Beta distribution will have a pole at $\eta = 0$ and if

$$\sigma_\eta^2 > \frac{\overline{\eta}\,(1 - \overline{\eta})^2}{2 - \overline{\eta}} = \psi_b \,, \tag{5.10}$$

the beta distribution will have a pole at $\eta = 1$. It can be verified from these equations that $\psi_a, \psi_b \in (0, \overline{\eta} - \overline{\eta}^2)$ and when $\overline{\eta} > 0.5$, $\psi_a > \psi_b$. The results in Chapter 3 indicate that, without appreciable loss of generality, $\overline{\eta} > 0.5$ is true in the near-field regime. Then, as the variance is decreased from its maximum towards zero, the Beta distribution transitions through three phases: For $(\overline{\eta} - \overline{\eta}^2) > \sigma_\eta^2 > \psi_a$, the probability density function has poles both at $\eta = 0$ and $\eta = 1$ (Figure 5-1(a)). For $\psi_a \geq \sigma_\eta^2 > \psi_b$, there is only a pole at $\eta = 1$ (Figure 5-1(c)), and for $\sigma_\eta^2 \leq \psi_b$, the Beta distribution has no poles (Figure 5-1(d)). The Beta distribution is unimodal in the latter two regions, thus the unconditional error probability will lie within the unimodal bounds derived in Chapter 4. When $\sigma_\eta^2 > \psi_a$, because the distribution is not unimodal, the unconditional error probability may not fall within the unimodal bounds. However, the results must still lie between the two-impulse bounds.

Figure 5-2 illustrates a set of Beta cumulative distribution functions with equal mean and a variety of σ_η^2 values. It is seen that there is appreciably larger probability close to $\eta = 0$ (greater than two orders of magnitude up to $\eta = 0.1$ for the particular choice of parameters in the figure), when the density has a singularity at $\eta = 0$. Thus we expect that the performance will be closer to the worst-case distribution performance in this regime. On the other hand, suppose that the mean and variance are such that the probability density of η does not have any singularity or has a singularity only at $\eta = 1$. Then, because the deep fade probability is small, we expect that the unconditional probability of error results will be closer to the unimodal lower bound.

This transition in Beta-distribution performance, from worst-case towards best-case performance as the variance of η is decreased while the mean is kept constant, is demonstrated in the unconditional error probability plots in Figures 5-3 – 5-6. The mean of η is 0.99 in all of the plots, which yields $\psi_a = 4.925 \times 10^{-3}$ and $\psi_b = $
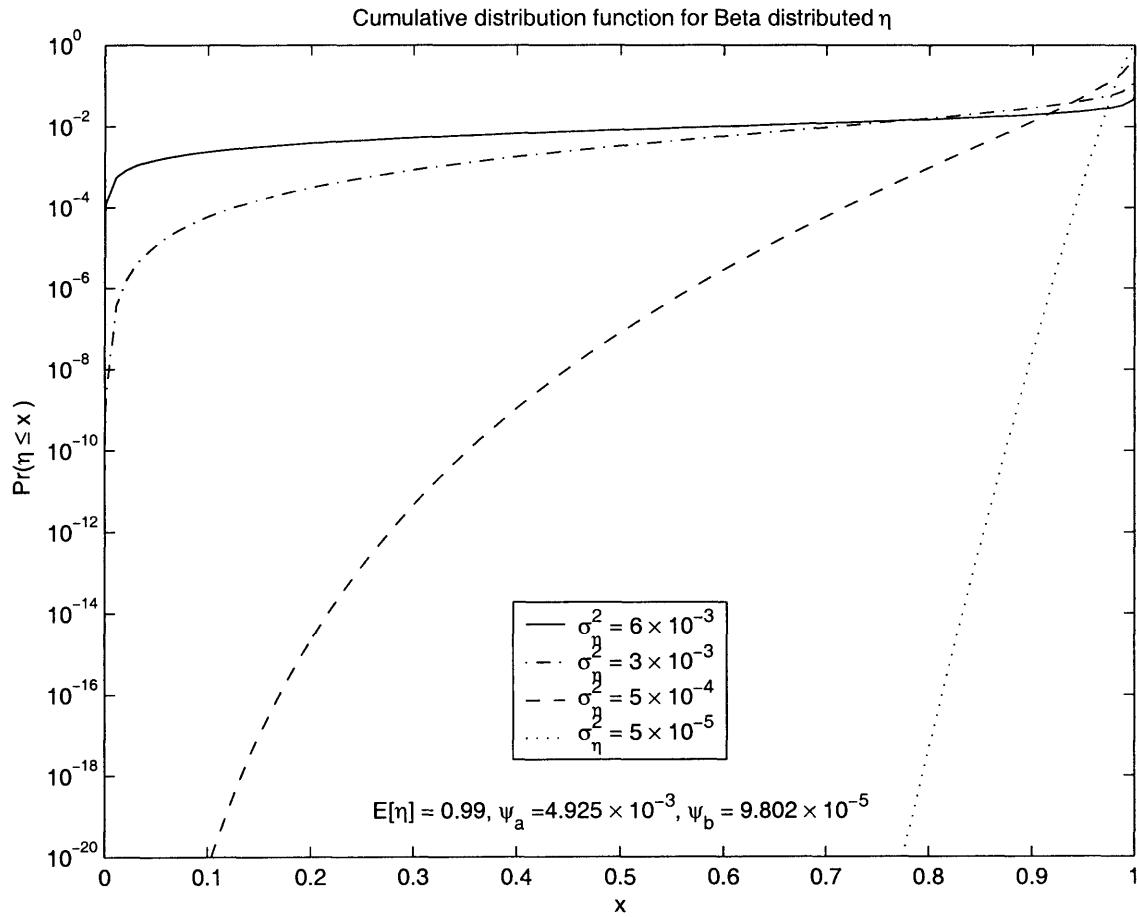
Figure 5-2: The cumulative distribution function of Beta-distributed η for several σ_η^2 values. For $\sigma_\eta^2 > \psi_a$, the CDF is much higher on the lower end of the $[0, 1]$ interval than it is for $\sigma_\eta^2 < \psi_a$.

9.802×10^{-5}. In Figure 5-3 the variance for the Beta distribution is 6×10^{-3}, which yields a density with poles at $\eta = 0$ and $\eta = 1$. The resulting unconditional error probability is closer to the upper bound. The error probability associated with this Beta distribution is higher than that of any unimodal probability distribution with the same mean and variance.

Figure 5-4 and 5-5 plot the unconditional error probability for a Beta distribution with $\sigma_\eta^2 = 3 \times 10^{-3}$ and $\sigma_\eta^2 = 5 \times 10^{-4}$ respectively, both of which have a singularity at $\eta = 1$ but none at $\eta = 0$. In Figure 5-4, the Beta distribution error probability is approximately half-way between the best-case and worst-case unimodal bounds. In addition, there is a noticeable difference between the minimum achievable error probability and the probability of error with no fading, for large signal photon numbers. On the other hand, in Figure 5-5 the Beta distribution is very close to the minimum achievable probability of error, which is approximately equal to the no-fading probability of error.

Figure 5-6 plots the results for $\sigma_\eta^2 = 5 \times 10^{-5}$, for which the Beta distribution has no poles. Comparing this figure to the previous figure, it is observed that further reducing the variance changes little in terms of the Beta distribution error probability and the lower bounds. However, because the worst case probability of error result is proportional to σ_η^2 for high n_s, the worst-case probability of error bounds have decreased an order of magnitude from the upper bounds given in Figure 5-5.
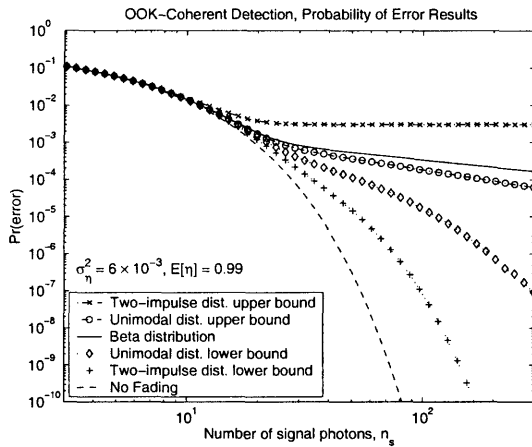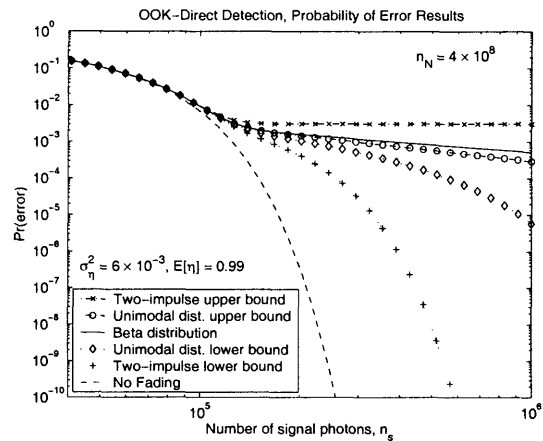
(a) Coherent detection   (b) Direct detection

Figure 5-3: Probability of error bounds from Chapter 4 plotted together with the error probability curve when η is Beta distributed with $\overline{\eta} = 0.99$ and $\sigma_\eta^2 = 6 \times 10^{-3}$. The Beta distribution is singular at both end-points. Plots are for OOK-modulation. PPM results are obtained by shifting all curves to the left by 3 dB.

(a) Coherent detection   (b) Direct detection

Figure 5-4: Probability of error bounds from Chapter 4 plotted together with the error probability curve when η is Beta distributed with $\overline{\eta} = 0.99$ and $\sigma_\eta^2 = 3 \times 10^{-3}$. The Beta distribution is singular at the right end-point if its support. Plots are for OOK-modulation. PPM results are obtained by shifting all curves to the left by 3 dB.

(a) Coherent detection   **(b)** Direct detection

Figure 5-5: Probability of error bounds from Chapter 4 plotted together with the error probability curve when η is Beta distributed with $\overline{\eta} = 0.99$ and $\sigma_\eta^2 = 5 \times 10^{-4}$. The Beta distribution is singular at the right end-point if its support. Plots are for OOK-modulation. PPM results are obtained by shifting all curves to the left by 3 dB.

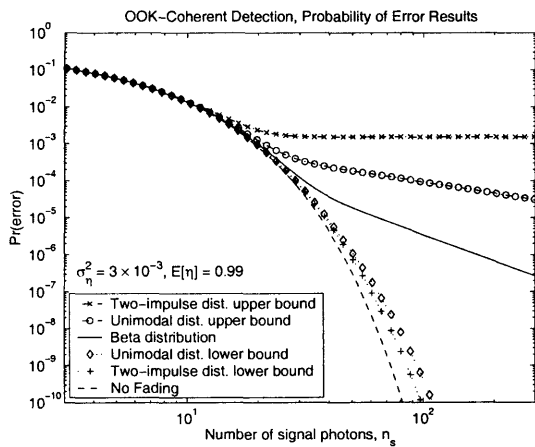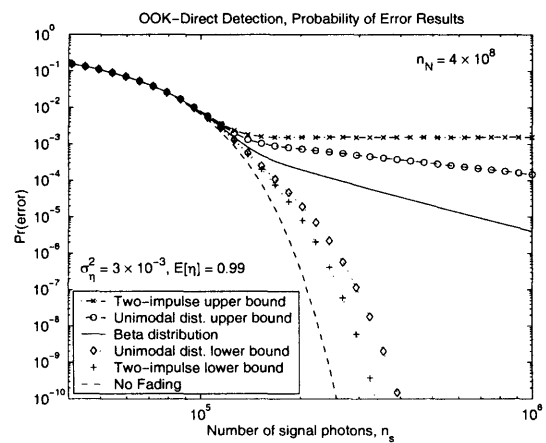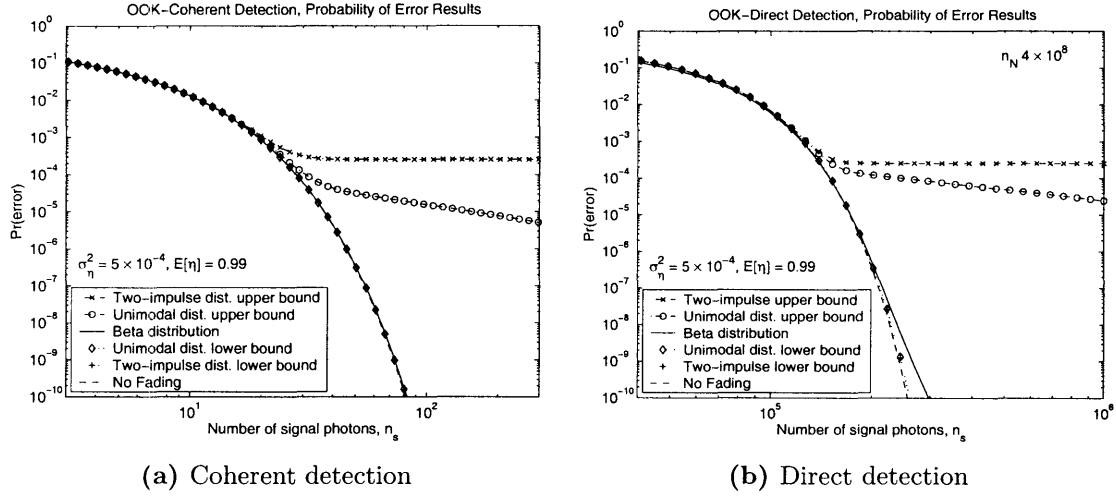**(a)** Coherent detection   **(b)** Direct detection
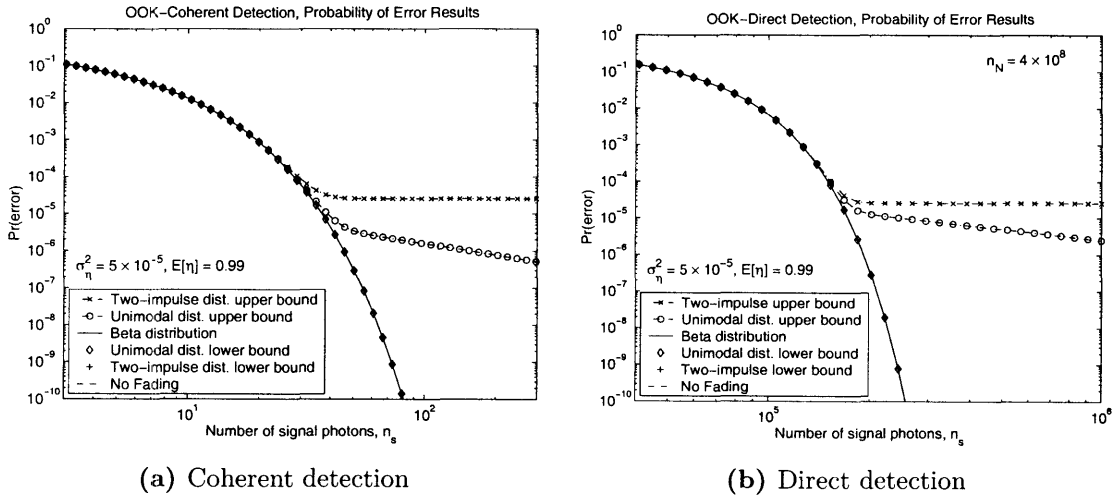
Figure 5-6: Probability of error bounds from Chapter 4 plotted together with the error probability curve when η is Beta distributed with $\overline{\eta} = 0.99$ and $\sigma_\eta^2 = 5 \times 10^{-5}$. The Beta distribution has no singularities. Plots are for OOK-modulation. PPM results are obtained by shifting all curves to the left by 3 dB.

115

Chapter 6

Capacity of SISO Optical Links with Coherent Detection Receivers

So far, the performance criterion in our analysis has been the probability of error in communicating a binary symbol (i.e., a bit) using particular encoding and decoding schemes at some chosen data rate. A more informative measure of the ultimate performance of near-field atmospheric optical communications is the Shannon capacity of the link. In most general terms, the capacity of a communication channel is the maximum rate (in bits-per-transmission or bits-per-second) at which information can be transmitted and decoded with arbitrarily small probability of error.

In this chapter, we will analyze the capacity of a single-input, single-output (SISO) atmospheric optical communication link operating in the near-field regime and utilizing a coherent detection receiver. We will show that a mean constraint on the maximum eigenvalue is sufficient to obtain a tight lower bound on the ergodic capacity of the channel. In addition, we consider performance when η is assumed to have a Beta distribution.

Consider the block diagram in Figure 6-1, showing the general structure of a communication channel. The set of all possible messages[1]$\mathcal{M}$ are mapped to a set of length-n codewords, denoted by $\mathbf{X}^n$ in the figure. The channel has a known statistical

---

[1]With no loss of generality assume the messages are the set of index numbers; $\mathcal{M} = \{1, 2, 3, \ldots, M\}$
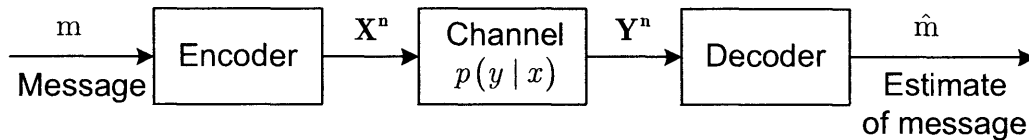
Figure 6-1: Block diagram of a general communications system.

transformation relating input symbols to output symbols, shown here as $\mathbf{X}^n \to \mathbf{Y}^n$, which we assume to be memoryless; i.e., each symbol transmission is independent of all other transmissions. The received codeword $\mathbf{Y}^n$ is mapped, via some deterministic decision rule, to an estimate of the transmitted message, $\hat{m}$. Therefore the probability of error for message m is defined as $\Pr(\hat{m} \neq m), m \in \mathcal{M}$. The maximum error probability within the set of all messages is then $\max_{m \in \mathcal{M}} \Pr(\hat{m} \neq m)$.

We refer to a code that maps M messages to codewords of length n, as a (M, n) code. The rate of transmission of a (M, n) code is defined as

$$R = \frac{\log_2 M}{n} \, , \tag{6.1}$$

and is referred to as being *achievable* if the maximum error probability goes to 0 as $n \to \infty$. The capacity of the channel is defined as the supremum of all achievable rates.

For most elementary communication channels, such as discrete memoryless channels or Gaussian channels, it has been shown that the channel capacity is equal to the maximum mutual information between the input and output symbols, where the maximization is over all possible distributions on the input [30]:

$$C = \max_{p(x)} I(X;Y) \, , \tag{6.2}$$

with the mutual information given by $I(X;Y) = \iint_{\mathcal{X},\mathcal{Y}} p(x,y) \log_2 \left(\frac{p(x,y)}{p(x)p(y)} \right) dx dy$.
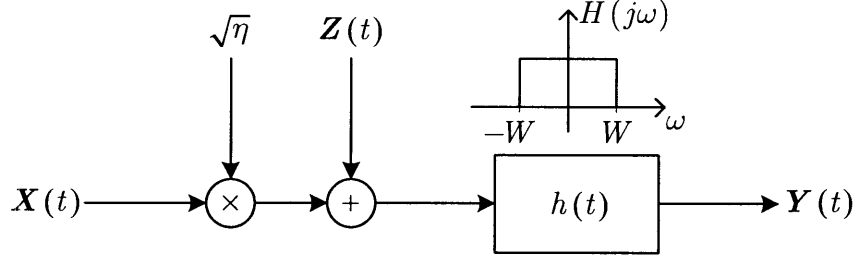
Figure 6-2: A block diagram representation of an atmospheric optical communication link using a coherent detection receiver.

6.1 Capacity of an Atmospheric Optical Link with Coherent Detection

Consider the simplified block diagram of a near-field atmospheric optical communication link employing a coherent detection receiver, shown in Figure 6-2. $\mathbf{X}(t)$, denotes the complex time modulation on the transmitted optical wave with an average power constraint of P; i.e.,

$$E\left[|\mathbf{X}(t)|^2\right] = P , \qquad t \in [0, T] . \tag{6.3}$$

Correspondingly, $\mathbf{Y}(t)$ represents the received waveform after optical detection. The spatial pattern of the transmitted waveform is assumed to be $\Phi(\boldsymbol{\rho})$, and $\phi(\boldsymbol{\rho}')$ is the local oscillator spatial field. η is the maximum eigenvalue of the atmospheric propagation kernel representing the power loss during atmospheric propagation and $\mathbf{Z}(t)$ is the local oscillator shot noise in coherent detection, which is a stationary, zero-mean, complex white Gaussian noise process with independent, identically distributed real and imaginary parts, each with double-sided power spectral density $\frac{N_0}{2} = \frac{\hbar\omega_o}{2\eta_e}$. The filter $h(t)$ is a bandlimiting filter with bandwidth W.

Assume η is known both at the transmitter and receiver ends. Then the capacity of the channel shown in Figure 6-2 is the maximum of the sample mutual information,

$$i(X, Y|\eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y|\eta) \log_2 \left(\frac{p(x, y|\eta)}{p(x|\eta)p(y|\eta)} \right) dy\, dx , \tag{6.4}$$

which is the well known capacity of a band-limited complex Gaussian channel with

119

average input power constraint $P\eta$ [30]:

$$C_I = 2W \log_2 \left(1 + \frac{P\eta}{2N_0W}\right) \qquad \text{bits / second.} \tag{6.5}$$

X in (6.4) is the transmitted symbol over $[0, T]$ and Y is the corresponding received symbol. The capacity in (6.5) has subscript "I" to emphasize that this capacity is "instantaneous"; i.e. for a particular realization of η.

This capacity result suggests that if η were constant for all time and known to both ends, the transmitter would simply use a codebook suitable for a transmission rate $R < C_I$, which would yield arbitrarily small probability of error as $t \to \infty$. However, due to the evolution of turbulent eddies and the drift effect caused by winds, the turbulent state of the atmosphere will alter over long time intervals, changing the value of η as well. Therefore the instantaneous capacity cannot be achieved over long periods of time. A more insightful measure of capacity is to consider the long term average of the instantaneous capacities; i.e.

$$C_E = \lim_{T' \to \infty} \left(\frac{\int_0^{T'} C_I\big(t\,; \eta(t)\big)\, dt}{T'}\right), \tag{6.6}$$

where $\eta(t)$ appears as a parameter of $C_I(t)$ to emphasize that the time dependent changes of the instantaneous capacity are due to the changes in the power coupling between the transmitter and the receiver. Now, given the knowledge of $\eta(t)$ for all time, the transmitter generates a codebook for each distinct value of $\eta(t)$ and encodes a symbol over all time intervals in which $\eta(t)$ takes the same value, hence achieving the instantaneous capacity at every time instant t. Then, as T' goes to infinity, the capacity of the channel becomes C_E. If $\eta(t)$ is assumed to be an ergodic process[2],

---

[2] In most general terms, if the limiting time-average properties of the sample outcomes of a random process are equal to the limiting ensemble-average properties of the random process, with probability 1, then the process is called ergodic. See [31].

then C_E will be equal to the ensemble average of the instantaneous capacity, i.e.,

$$C_E = \lim_{t \to \infty} E_\eta\big[C_I(t)\big] \tag{6.7}$$

$$= 2W \int_0^1 \log_2\left(1 + \frac{P\eta}{2N_0W}\right) p(\eta)\, d\eta \quad \text{bits/seconds.} \tag{6.8}$$

It is worthwhile to mention that the ergodic capacity C_E is equivalent to the maximizing the conditional mutual information $I\big(X, Y|\eta\big)$ subject to an average power constraint on the input.

When the channel bandwidth W is sufficient to ensure that $\frac{P}{2N_0W} \ll 1$, C_E simplifies to the power-limited result

$$C_E \approx \frac{P}{N_0}\, \bar{\eta} \log_2 e \ . \tag{6.9}$$

Therefore, in this regime of operation, the capacity is linearly proportional to the input power and average of the maximum eigenvalue. The lower bounds on $\bar{\eta}$ derived in Chapter 3 suggest that although establishing a link geometry with larger free-space Fresnel numbers will increase the ergodic capacity, there are diminishing returns to such an approach. Therefore, for links operating deep in the near-field regime, increasing average input power rather than the free-space Fresnel number may prove more rewarding in terms of capacity.

For a link that is not sufficiently broadband to enjoy power-limited capacity, obtaining its ergodic capacity requires the knowledge of the probability distribution of η. In the absence of this knowledge we derive bounds to the capacity, using the integral expression in (6.8). It is possible to obtain tight bounds by just imposing a mean constraint on η, hence we resort to the Markov moment matching theorem with a single constraint:

Lemma 6.1.1. *Let $\Omega(\eta)$ be a two-times differentiable and strictly convex function[3] on the closed interval $\big[a, b\big]$. Also, let $\mathcal{V}\big(p(\eta); \bar{\eta}\big)$ denote the set of all probability distributions on $\big[a, b\big]$ which satisfy*

---

[3]i.e. $\frac{\partial^2 \Omega(\eta)}{\partial \eta^2} > 0$

$$\int_a^b \eta \, p(\eta) \, d\eta = \overline{\eta} \, , \tag{6.10}$$

where $a < \overline{\eta} < b$. Then, the minimum value of the integral $\int_a^b \Omega(\eta) \, p(\eta) \, d\eta$ is achieved by a unique distribution[4] $p_{min}(\eta)$, which has all its mass concentrated at $\eta = \overline{\eta}$.

Likewise, the maximum value of this integral is attained with a unique distribution $p(\eta)_{max}$, which has mass concentrated at the two end-points, $\eta = a$ and $\eta = b$.

$C_I(\eta) = 2W \log_2 \left(1 + \frac{P\eta}{2N_0 W} \right)$ in (6.8) is twice differentiable, and strictly concave; i.e. $\frac{\partial^2 \Omega(t)}{\partial t^2} < 0 \, , \forall \eta \in [0, 1]$. Thus applying Lemma 6.1.1 to $-C_I(\eta)$, we conclude that the distribution of η which yields the lowest capacity is a two-impulse distribution with impulses located at $\eta = 0$ and $\eta = 1$. The probability density function for η is then

$$p_{min}(\eta) = (1 - \overline{\eta}) \, \delta(\eta) + \overline{\eta} \, \delta(\eta - 1) \, . \tag{6.11}$$

The resulting lower bound on the ergodic capacity is

$$C_E^{(min)} = \overline{\eta} \, 2W \log_2 \left(1 + \frac{P}{2N_0 W} \right) . \tag{6.12}$$

Not surprisingly, the maximum capacity is achieved when there is no fading[5]; i.e.

$$p_{max}(\eta) = \delta(\eta - \overline{\eta}) \, . \tag{6.13}$$

The corresponding upper bound on ergodic capacity is

$$C_E^{(max)} = 2W \log_2 \left(1 + \frac{P \overline{\eta}}{2N_0 W} \right) . \tag{6.14}$$

The normalized difference between these ergodic capacity bounds and zero transmission loss capacity, i.e. $\frac{C_I(\eta=1) - C_E}{C_I(\eta=1)}$, are plotted versus the free-space Fresnel number for moderate turbulence in Figure 6-3 and for strong turbulence in Figure 6-4 when a phase-compensated focused beam and the maximum-eigenvalue eigenfunction of the

---

[4]If $\Omega(\eta)$ is convex, but not strictly convex, then there may be multiple distributions in $\mathcal{V}(p(\eta); \overline{\eta})$ that achieve the maximum and minimum values given with $p_{max}(\eta)$ and $p_{min}(\eta)$ respectively.

[5]The upper bound on capacity also follows from Jensen's inequality.

average kernel are used as the transmitter spatial field patterns.

With this normalization, the lower-bound plots correspond to one minus the average power coupling for the field in consideration, while the upper bound plots are $1 - \dfrac{\log_2\left(1+\frac{P\,\bar{\eta}}{2N_0 W}\right)}{\log_2\left(1+\frac{P}{2N_0 W}\right)}$. Because the mean of the maximum eigenvalue approaches 1 very rapidly in the near field, regardless of the distribution of η, links operating in the near-field with moderately large free-space Fresnel numbers achieve near zero-fading ergodic capacity. Figure 6-3 suggests that the use of a phase-compensated focused beam in mild turbulence will result in 99% of zero-loss ergodic capacity with $D_{fo} \geq 25$, while the maximum-eigenvalue eigenfunction will fall in this range when $D_{fo} \geq 5$. The results for strong turbulence indicate that the ergodic capacity will lie within the 99-percentile of the zero-loss capacity with adaptive optics if $D_{fo} \geq 16$, while the maximum-eigenvalue eigenfunction requires $D_{fo} \geq 33$ (not shown in plot). Note that, because the average power coupling given by any sub-optimal field is a lower bound on $\bar{\eta}$, the lower bounds obtained for these sub-optimal fields are also lower bounds on the ergodic capacity when the maximum-eigenvalue eigenfunction is used. Thus, with the maximum-eigenvalue eigenfunction used at the transmitter, the ergodic capacity is guaranteed to achieve the 99-percentile for geometries with free-space Fresnel numbers no larger than 5 in mild turbulence, and 16 in strong turbulence conditions.

Achieving ergodic capacity requires full aprori knowledge of $\eta(t)$ and the ability to encode and decode over many atmospheric turbulence states, which are not feasible conditions in some cases. An alternative measure of capacity is therefore the highest data rate at which reliable communication is possible with probability $1 - p_{\mathrm{OUT}}$, i.e.,

$$\Pr\big(C_{OUT} > C_I(\eta)\big) = p_{\mathrm{OUT}} , \qquad (6.15)$$

which is equivalently,

$$\Pr\left(\eta < \frac{\left(2^{\frac{C_{OUT}}{2W}} - 1\right) 2N_0 W}{P}\right) = p_{\mathrm{OUT}} , \qquad (6.16)$$

123

Figure 6-3: Bounds on the percentile distance to zero-loss channel capacity when the maximum-eigenvalue eigenfunction is used as the transmitter spatial field pattern in mild turbulence. Bounds are obtained via the phase-compensated focused beam and maximum-eigenvalue eigenfunction of the average propagation kernel. The parameters values for the plot are $P = 0.1\text{mW}$, $W = 1.8$ GHz and $N_o = 2.5 \times 10^{-19}$ W/Hz.

124

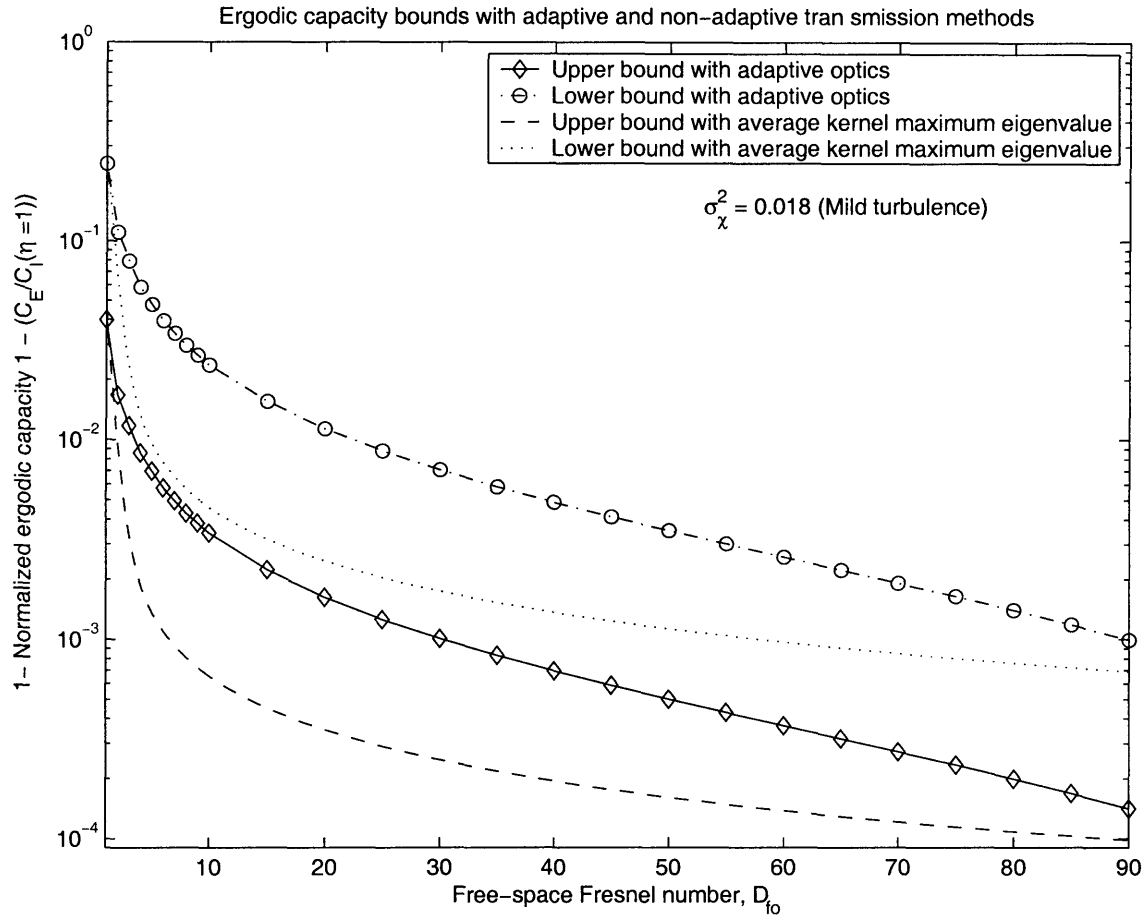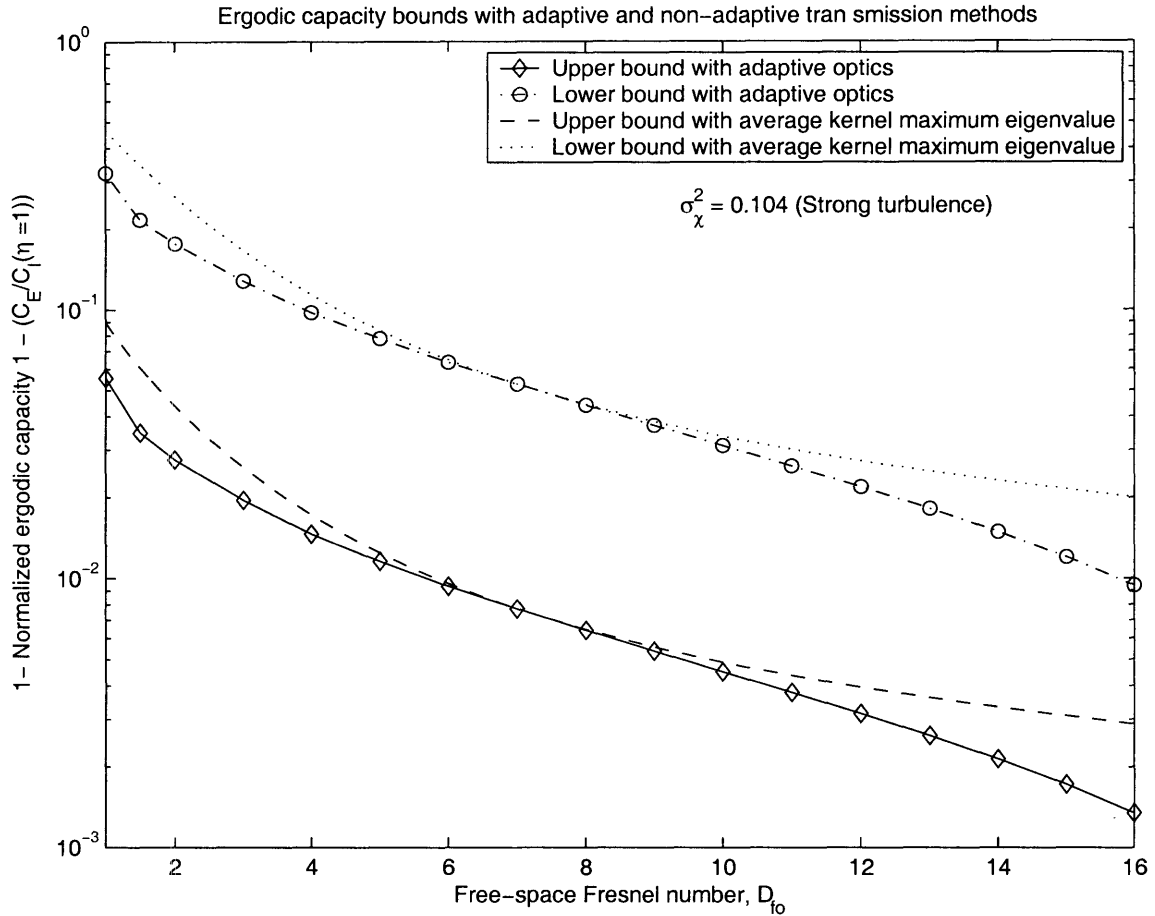Figure 6-4: Bounds on the percentile distance to zero-loss channel capacity when the maximum-eigenvalue eigenfunction is used as the transmitter spatial field pattern in strong turbulence. Bounds are obtained via the phase-compensated focused beam and maximum-eigenvalue eigenfunction of the average propagation kernel. The parameters values for the plot are $P = 0.1$mW , $W = 1.8$ GHz and $N_o = 2.5 \times 10^{-19}$ W/Hz.

where C_{OUT} is the outage capacity associated with the outage probability p_{OUT}. In other words, with probability p_{OUT}, the fade will be so deep that reliable communication at the chosen rate C_{OUT} will not be possible. For all data rates higher than C_{OUT}, the outage probability will exceed p_{OUT}, and for lower data rates, the outage probability will be less than or equal to p_{OUT}.

This definition is easily applied to the two-impulse distribution which yields minimum ergodic capacity for a given $\bar{\eta}$. For this distribution of η, the instantaneous capacity is 0 with probability $1 - \bar{\eta}$. Thus for any data rate R satisfying

$$2W \log_2 \left(1 + \frac{P}{2N_0 W} \right) \geq R > 0 \,, \tag{6.17}$$

there will be an *outage* with probability $1 - \bar{\eta}$. Then the outage capacity, when η is assumed to have a two-impulse distribution, is

$$C_{\text{OUT}} = \begin{cases} 2W \log_2 \left(1 + \frac{P}{2N_0 W} \right) & \text{if } p_{\text{OUT}} \geq (1 - \bar{\eta}) \\ 0 & \text{if } p_{\text{OUT}} < (1 - \bar{\eta}) \,. \end{cases} \tag{6.18}$$

Thus, for a given free-space Fresnel number, the minimum outage probability that yields non-zero capacity — and also the maximum capacity — is $1 - \bar{\eta}$. Figure 6-5 plots the minimum allowable outage probabilities for adaptive and non-adaptive communication schemes in moderate and strong fading conditions. It can be seen that there is approximately an order of magnitude difference between minimum outage probabilities for moderate and strong fading conditions when the maximum-eigenvalue eigenfunction of the average kernel is used. The change in the minimum outage probability is less for the phase-compensated focused beam. In strong turbulence conditions, the minimum outage probability with the adaptive scheme is lower than that with the non-adaptive scheme, however, in mild turbulence conditions, the non-adaptive scheme yields appreciably lower minimum outage probabilities than what is obtained with the phase-compensated focused beam. Nevertheless, because the mean of the maximum eigenvalue approaches 1 asymptotically, for any outage probability $p_{\text{OUT}} = \epsilon > 0$, there exists a free-space Fresnel number threshold D_ϵ such that, any
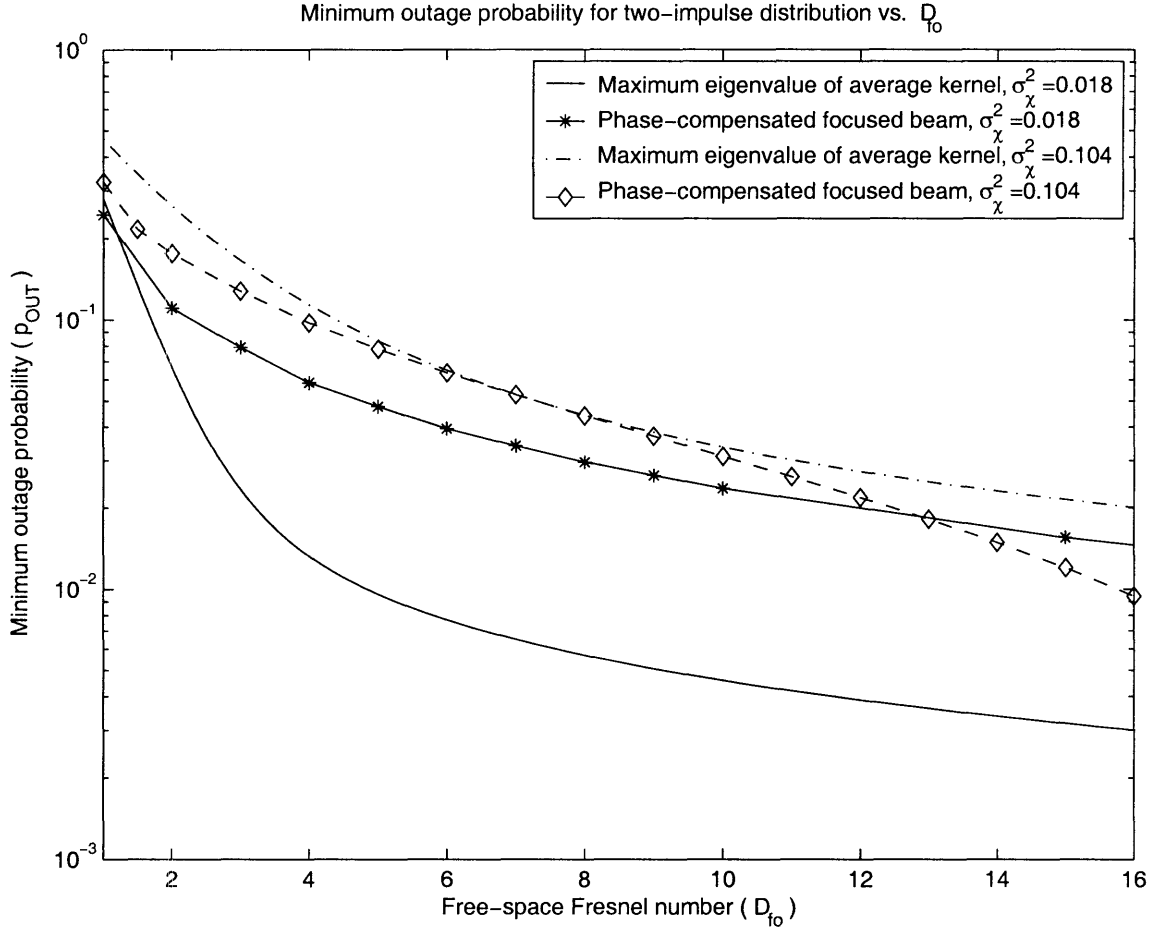
Figure 6-5: Minimum outage probability that will yield non-zero capacity when η has the worst-case two-impulse distribution. The plots correspond to the phase-compensated focused beam and the maximum-eigenvalue eigenfunction of the average kernel. Both moderate and strong fading conditions plotted.

link geometry with a free-space Fresnel numbers in excess of D_ϵ will have an outage capacity equal to that of a perfect transmission link.

If η is assumed to be Beta distributed, (6.16) yields an incomplete Beta function; i.e,

$$p_{\text{OUT}} = \frac{1}{\beta(a,b)} \int_0^{\eta_o} x^{a-1} (1-x)^{b-1} \, dx \qquad (6.19)$$

where,

$$\eta_o = \frac{\left(2^{\frac{C_{\text{OUT}}}{2W}} - 1\right) 2 N_0 W}{P}, \qquad (6.20)$$

and the two parameters for the Beta distribution are determined by the mean and

variance constraints on η. The shape of the Beta distribution varies dramatically depending on whether it has poles at $\eta = 0$ and/or $\eta = 1$. This leads to some interesting behavior in terms of the outage capacity of the link. From (6.20), the outage capacity increases as η_o increases. If the outage probability p_{OUT} is chosen to be very small, then the Beta distributions with singularities at $\eta = 0$ will yield the smallest outage capacity. Furthermore, as the variance is decreased, the outage capacity will increase. However, as the outage probability is chosen larger, this trend changes. In particular, the Beta distributions with singularities at $\eta = 1$ eventually yield higher η_o values than those with no singularities because more of the probability mass is concentrated around the pole at $\eta = 1$ (See Figure 5-2). Therefore, when the outage probability is larger than a threshold value, the Beta distribution with the largest variance (i.e. with poles at $\eta = 0$ and $\eta = 1$) will yield the highest outage capacity. This is not surprising given the following interpretation. Probability distributions with small variances have most of the probability density concentrated around their means, whereas densities with large variances are more spread out. Thus, as the outage probability p_{OUT} is increased, the threshold will vary only slightly around the mean if the probability density has small variance, while the threshold will sweep a much broader range if the distribution has large variance. That is to say, if we can tolerate large outage probabilities, then a channel with larger deviations from its mean will be able to support higher data rates on those rare occasions where transmission is anomalously good. On the other hand, if we require transmission most of the time (i.e. a small outage probability), then the channel with large fluctuations will have to use a much lower rate to maintain a steady link.

Figure 6-6 is a plot of outage capacities, normalized to the perfect-transmission capacity, for Beta distributed η with $\bar{\eta} = 0.99$ and variance values in the range $5 \times 10^{-5} - 6 \times 10^{-3}$. For $p_{\text{OUT}} < 0.014$ the Beta distribution with no poles, thus smallest variance, yields highest outage capacity. However, for $p_{\text{OUT}} > 0.028$, the Beta distribution with the largest variance will achieve the highest outage capacity and at $p_{\text{OUT}} = 0.05$, η_o is close enough to 1 that the outage capacity of this Beta distribution is almost equal to that of a zero-loss link. Thus, the plots confirm that

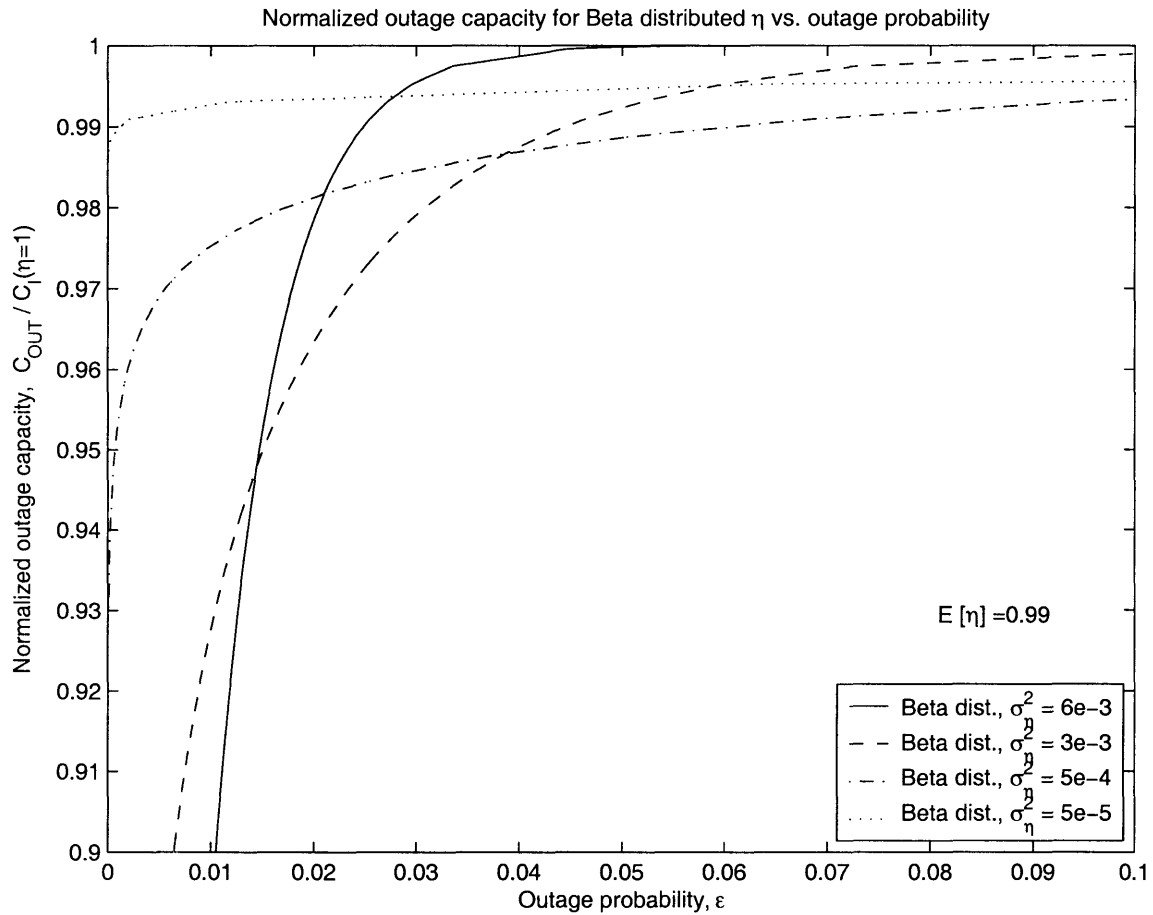Figure 6-6: Outage capacity results for Beta distributed η with different variance values. The outage capacity is normalized by zero-loss capacity; i.e. $\frac{C_{OUT}}{C_I(\eta=1)}$ is plotted.

larger variance may help improve the outage capacity when η is assumed to be Beta distributed. Note however, that larger values of p_{OUT}, imply that there is a larger probability of not being able to communicate reliably.

Chapter 7

Discussion and Future Work

In this thesis, we have attempted to establish a better understanding of the performance of atmospheric optical communication links operating in the near field power transfer regime. Our efforts were concentrated in two areas: increasing our understanding of the first and second moment behavior of maximal power coupling due to atmospheric propagation and using these results to obtain probability of error and capacity bounds for near-field, atmospheric optical links.

Chapter 3 addressed the atmospheric propagation aspects of near-field communication. In particular, we considered the normal mode decomposition of the turbulent atmosphere and utilized deterministic and adaptive input fields to obtain lower bounds on the average of its maximum eigenvalue. One striking observation is the rather invariant nature of the average power transfer of a focused beam in free-space, the turbulent atmosphere, and when all phase distortions are hypothetically assumed to be eliminated. Applying phase compensation to a focused beam which maximizes on-axis power density, yields slight improvement over other focused beam results. In mild turbulence conditions, the maximum eigenvalue of the average kernel yields highest average power coupling after the maximum eigenvalue of the log-amplitude-only average kernel, yet all bounds converge towards 1 much slower than the free-space maximum eigenvalue. A similar gap is present in strong turbulence conditions. However, deterministic fields appear to have a clear disadvantage over phase-compensated input fields in strong turbulence. The convergence of the maximum eigenvalue of the

average kernel and that of the log-amplitude-only average kernel towards 1 is appreciably faster when unequal apertures are considered. However, the convergence is still not comparable to that of the free-space maximum eigenvalue. Future work in this area might concentrate on obtaining the eigenfunction structure of the phase-compensated average kernel, thereby establishing a more rigorous understanding of the ultimate average power transfer benefits from phase compensation that maximizes the on-axis power density.

Another area in need of further attention is the second moment results for average power transfer. Although an expression in the form of a four dimensional integral is obtained in this thesis, computational resources have proven insufficient to calculate a valid numerical result from this expression. Much of the work that follows, however, assumes knowledge of the variance for η. Therefore future work should focus on simplifying the variance expression to facilitate a numerical solution. Furthermore, a number of numerical challenges have been encountered in the process of computing the phase compensated power transfer results. Additional work in numerics might resolve some of these challenges.

The second part of this thesis treated the communication aspects of atmospheric optical links. Chapter 4 developed unconditional error probability bounds for communication links with OOK or PPM modulation and coherent or direct detection receivers. Employing the lower bounds on the average of the maximum eigenvalue and assuming a variance is obtainable, we have shown that the distributions achieving the best and worst-case error probabilities are two-impulse distributions, which yield drastically different performance. Limiting the set of possible distributions to unimodal distributions with prescribed mean and variance values, we slightly improved the bounds on error probability. However, the most relevant conclusion from this analysis is that the unconditional probability of error of a near-field communication link does not depend as much on the particular shape of the distribution, as it does on the overall deep-fade probability. Therefore, some statistical characterization of deep-fade events in near field propagation could lead to much tighter probability of error bounds, hence a more precise characterization of binary communication per-

132

formance. Such a statistical characterization of deep-fade events could be pursued experimentally, by relying on the ergodicity of the time dependent changes in fractional power transfer. However, the need for precise aiming and focusing onto the receiver aperture in near field communications is a major deterrent to performing such experiments.

One distribution that is of particular interest, due to its convenient analytical properties is the Beta distribution. Chapter 5 assumed the maximum eigenvalue is Beta distributed and compared the resulting unconditional error probabilities to the bounds obtained in the preceding chapter. In this chapter we showed that, if η is Beta distributed with some known mean, then the unconditional probability of error results transition from the worst-case bounds towards the best-case bounds as the variance is decreased from its maximum towards zero.

In Chapter 6 we analyzed the capacity of an atmospheric optical link with a coherent detection receiver. In particular, we considered the ergodic capacity and the outage capacity of such a link. We concluded that a mean constraint is sufficient to obtain an accurate estimate of the ergodic capacity of the channel by deriving tight lower and upper bounds which rapidly converge to each other with increasing free-space Fresnel numbers. The minimum ergodic capacity distribution was shown to be a two-impulse distribution with one impulse at $\eta = 0$ and the other at $\eta = 1$ and the outage capacity was derived for this distribution. We also considered the outage capacity when η is Beta distributed and concluded that Beta distributions with large variance values can yield higher outage capacity than Beta distributions with smaller variance and can approach zero-loss outage capacity at high outage probabilities.

A natural extension to the work in Chapter 6 is to consider the capacity of direct detection links. If we assume that the detectors are operating in the shot-noise limited regime, then *given* η, the capacity of the channel is the capacity of a Poisson channel, for which an analytical expression has been derived in [32, 33]. However, the expression for the capacity of a Poisson channel cannot be directly structured to fit into the context of a moment matching problem. Therefore, the tools developed in this thesis do not apply to shot-noise limited direct detection. Nevertheless, the capacity of a

direct detection link and its comparison to coherent detection links remain a topic of interest for future research.

Appendix A

One-Dimensional equivalents of atmospheric propagation results

This appendix gives a brief summary of the two-dimensional propagation medium (one-dimensional aperture) results. The free-space fresnel number in this geometry is

$$D_{fo} = \frac{d_1 \, d_2}{\lambda L} \, . \tag{A.1}$$

Kolmogorov's development of a model for atmospheric turbulence in a 3-D propagation medium [13] is paralleled for a two-dimensional propagation medium, from which the Kolmogorov turbulence spectrum is found to be

$$\Phi_2(u) = 0.056 \, C_n^2 \, u^{-8/3} \, , \qquad u \geq 0 \, , \tag{A.2}$$

for constant turbulence profile strength $C_n^2(z) = C_n^2$.

The mutual coherence function can be derived following Ishimaru's derivation of the mutual coherence function [34], which gives the following covariance function

expressions for the log-amplitude and phase fluctuations:

$$K_{\chi,\chi}(\rho',\rho) = 0.224\pi k^2 L\, C_n^2 \int_0^1 \int_0^\infty u^{-8/3}\, \cos(d'u)\, \sin^2\!\left(\frac{u^2 z(1-z)}{2\,k/L}\right) du\, dz \ , \quad (A.3)$$

$$K_{\phi,\phi}(\rho',\rho') = 0.224\pi k^2 L\, C_n^2 \int_0^1 \int_0^\infty u^{-8/3}\, \cos(d'u)\, \cos^2\!\left(\frac{u^2 z(1-z)}{2\,k/L}\right) du\, dz \ , \quad (A.4)$$

$$K_{\chi,\phi}(\rho',\rho') = 0.112\pi k^2 L\, C_n^2 \int_0^1 \int_0^\infty u^{-8/3}\, \cos(d'u)\, \sin\!\left(\frac{u^2 z(1-z)}{k/L}\right) du\, dz \ , \quad (A.5)$$

where $d' = |z\,\rho' + (1-z)\,\rho|$. Consequently, the log-amplitude variance is

$$\sigma_\chi^2 = K_{\chi,\chi}(0,0) = 0.067\, k^{7/6}\, L^{11/6}\, C_n^2. \qquad (A.6)$$

The wave structure function for Kolmogorov-spectrum turbulence is given as

$$D(\rho',\rho) = 2.94\, k^2\, C_n^2\, L \int_0^1 \left|s\rho' + (1-s)\rho\right|^{5/3} ds \ , \qquad (A.7)$$

which can be analytically evaluated to

$$D(\rho',\rho) = \begin{cases} 1.1025 k^2\, C_n^2\, L\dfrac{\left(\mathrm{sgn}(\rho')\,|\rho'|^{8/3} - \mathrm{sgn}(\rho)\,|\rho|^{8/3}\right)}{(\rho'-\rho)} & \text{if } \rho \neq \rho' \\[4mm] 2.94\, k^2\, C_n^2\, L\, |\rho|^{5/3} & \text{if } \rho = \rho' \ , \end{cases} \qquad (A.8)$$

where $\mathrm{sgn}(x)$ is the signum function defined as

$$\mathrm{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\[2mm] 0 & \text{if } x = 0 \\[2mm] -1 & \text{if } x < 0 \ . \end{cases} \qquad (A.9)$$

Note that, in (A.8), the limit as $\rho \to \rho'$, gives value of the function at $\rho' = \rho\,, \forall \rho \in R_1$ and $\rho' \in R_2$. Thus, $D(\rho',\rho)$ is continuous.

With the equal aperture assumption, the terms in the wave structure function can

136

be regrouped such that

$$D(x', x) = 202.539 \, \sigma_\chi^2 \, D_{fo}^{5/6} \int_0^1 |sx' + (1-s)x|^{5/3} \, ds \qquad \text{(A.10)}$$

where d_o is the aperture size and $x d_o = \rho$, $x' d_o = \rho'$ are the normalized coordinates; i.e. $x, x' \in [-\frac{1}{2}, \frac{1}{2}]$.

Appendix B

Modes of a unimodal distribution with given mean and variance

Let X be a unimodal distribution on $[0, 1]$ with mode ν, mean m_X and variance σ_X^2. Then, via Theorem 4.3.2, there exist independent random variables U and Z, with U uniform on $[0, 1]$, such that UZ has the same distribution as $X - \nu$. It immediately follows from the support condition on X that,

$$0 \leq \nu \leq 1 \,. \tag{B.1}$$

We now look at the problem somewhat backwards. The set of all unimodal distributions for X with mode ν, given the support, mean and variance constraints for X, is equivalently described with the set of all probability distributions for Z satisfying the support, mean and variance constraints specified in equations (4.80a) – (4.80d) in Chapter 4. Therefore, ν can be any real number in the closed interval $[0, 1]$, so long as there exists a distribution for Z on $[-\nu, 1 - \nu]$, with mean given by (4.80b) and variance by (4.80c).

Our main interest here is to determine the interval in which ν must lie, given mean and variance constraints on a unimodal X. To this end, we refer to a theorem presented in [23], which specifies the existence conditions for a distribution with prescribed moments. A simplified version of this theorem, combined with some

corollaries, is presented below:

Lemma B.0.2 (Markov). *s_1 is the mean and s_2 is the mean-square of some probability distribution on $[\nu, 1 - \nu]$ if, and only if,*

$$s_1^2 \leq s_2 \tag{B.2}$$

and

$$(1 - 2\nu)s_1 + \nu(1 - \nu) - s_2 \geq 0 . \tag{B.3}$$

From the lemma, it immediately follows that Z is a valid random variable if, and only if,

$$\sigma_Z^2 \geq 0 \tag{B.4}$$

and

$$(1 - 2\nu)\, m_Z + \nu(1 - \nu) - E\left[Z^2\right] \geq 0 . \tag{B.5}$$

The first condition applied to (4.80c) gives

$$m_X - \sqrt{3}\,\sigma_X \leq \nu \leq m_X + \sqrt{3}\,\sigma_X , \tag{B.6}$$

and the second condition leads to

$$\nu(2\,m_X - 1) + 2\,m_X - 3E\left[X^2\right] \geq 0 . \tag{B.7}$$

This expression can be simplified to

$$\nu \geq \frac{3E\left[X^2\right] - 2m_X}{2m_X - 1} \tag{B.8}$$

for $m_X > 0.5$, and

$$\nu \leq \frac{3E[X^2] - 2m_X}{2m_X - 1} \qquad (B.9)$$

for $m_X < 0.5$. In near-field applications we are interested in $m_X > 0.5$, thus we assume (B.8) is true. Then the three conditions on ν, presented in (B.1), (B.6) and (B.8), are the necessary and sufficient conditions on ν to ensure X is an appropriate unimodal distribution on $[0, 1]$ with the prescribed first and second moments. If all three intervals do not have an intersection, then there is no unimodal distribution for X on $[0, 1]$ which will have the prescribed mean and variance. However, if the three conditions do indeed have an intersection, the interval in which ν must exist is

$$\max\left\{0, \bar{\eta} - \sqrt{3}\sigma_X, \frac{3E[X^2] - 2m_X}{2m_X - 1}\right\} \leq \nu \leq \min\left\{1, m_X + \sqrt{3}\sigma_X\right\}. \qquad (B.10)$$

This expression is presented in (4.99) in Chapter 4 with X replaced by η, m_X by $\bar{\eta}$, and $E[X^2]$ by $\overline{\eta^2}$.

Bibliography

[1] E. J. McCartney. *Optics of the Atmosphere : Scattering by Molecules and Particles.* Wiley and Sons, New York, 1976.

[2] C. F. Bohren and D. R. Huffman. *Absorption and Scattering of Light by Small Particles.* Wiley and Sons, New York, 1983.

[3] R. S. Kennedy. Communication through optical scattering channel: An introduction. *Proceedings of the IEEE, Special Issue on Optical Communication,* October 1970.

[4] Lerner and Holland. The optical scatter channel. *Proceedings of the IEEE, Special Issue on Optical Communication,* October 1970.

[5] J. H. Shapiro. Optics of the turbulent atmosphere. Unpublished notes, March 1980.

[6] L. C. Andrews and R. L. Phillips. *Laser Beam Propagation through Random Media.* SPIE Optical Engineering Press, Bellingham, Washington, 1998.

[7] R. K. Tyson. *Principles of Adaptive Optics.* Academic Press, Boston, 1991.

[8] J. H. Shapiro. Imaging and optical communication through atmospheric turbulence. In J.W. Strobehn, editor, *Laser Beam Propagation in the Atmosphere,* pages 171–222. Springer-Verlag, New York, 1978.

[9] R. F. Lutomirski and H. T. Yura. Propagation of a finite optical beam in an inhomogenous medium. *Applied Optics,* 10(7):1652–1658, July 1971.

[10] M. Young. *Optics and Lasers*. Springer, New York, 5th edition, 2000.

[11] J. H. Shapiro. Normal mode approach to wave propagation in the turbulent atmosphere. *Applied Optics*, 13(11):2614–2619, November 1974.

[12] A. A. M. Saleh. An investigation of laser wave depolarization due to atmospheric transmission. *IEEE Transactions on Quantum Electronics*, QE-3(11):540–543, November 1967.

[13] V. I. Tatarski. *Wave Propagation in a Turbulent Medium*. Translations of Mathematical Monographs. McGraw-Hill, New York, 1961. Translated by Silverman, R.A.

[14] D. Silvano. *Photodetectors : devices, circuits, and applications*. Prentice Hall, New Jersey, 2000.

[15] J. H. Shapiro. Lecture notes for optical detection and communication. Class notes, 1982.

[16] R. H. Kingston. *Detection of Optical and Infrared Radiation*. Springer-Verlag, New York, 1978.

[17] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. Electrical and Electronic Engineering Series. McGraw-Hill, New York, 3rd edition, 1991.

[18] G. R. Osche. *Optical Detection Theory for Laser Applications*. Wiley Series in Pure and Applied Optics. John Wiley and Sons, New Jersey, 2002.

[19] R. M. Gagliardi and S. Karp. *Optical Communications*. Wiley and Sons, New York, 1995.

[20] D. Slepian. Analytic solution of two apodization problems. *Journal of the Optical Society of America*, 55(9):1110–1115, September 1965.

[21] V. P. Lukin. *Adaptive Beaming and Imaging in the Turbulent Atmosphere*. SPIE Press, Washington, 2002. Translated by A. B. Malikova.

[22] A. S. Willsky, G. W. Wornell, and J. H. Shapiro. Stochastic processes, detection and estimation. 6.432 Class Notes, 1999.

[23] M. G. Krein. *The Ideas of P. L. Chebyshev and A. A. Markov in the Theory of Limiting Values of Integrals and Their Further Development*, volume 12 of *American Mathematical Society Translations, 2*. American Mathematical Society, Ann Arbor, Michigan, 1959.

[24] M. G. Krein and A. A. Nudelman. *The Markov Moment Matching Problem and Extremal Problems*, volume 50 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, Rhode Island, 1977.

[25] S. Dharmadhikari and J. Kumar. *Unimodality, Convexity and Applications*. Probability and Mathematical Statistics. Academic Press, Boston, 1988.

[26] E. M. J. Bertin, I. Cuculescu, and R. Theodorescu. *Unimodality of Probability Measures*. Kluwer Academic Publishers, Boston, 1997.

[27] A. Mattuck. *Introduction to Analysis*. Prentice Hall, Upper Saddle River, New Jersey, 1997.

[28] L. C. Andrews. *Special functions of Mathematics for Engineers*. SPIE Optical Engineering Press, Bellingham, Washington, 2nd edition, 1998.

[29] M. Evans. *Statistical Distributions*. Wiley series in Probability and Statistics. Wiley and Sons, New York, 2000.

[30] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley Series in Telecommunications. John Wiley and Sons, New York, 1991.

[31] R. G. Gallager. *Discrete Stochastic Processes*. Kluwer Academic Publishers, Boston, 1996.

[32] A. D. Wyner. Capacity and error exponent for the direct detection photon channel – part i. *IEEE Transactions on Information Theory*, 34(6):1449–1457, November 1988.

[33] M. H. A. Davis. Capacity and cutoff rate for poisson-type channels. *IEEE Transactions on Information Theory*, IT-26(6):710–715, November 1980.

[34] A. Ishimaru. *Wave Propagation and Scattering in Random Media*, volume 2. Academic Press, New York, 1978.