On the Capacity of Bosonic Channels

by

Christopher Graham Blake

Submitted to the Department of Electrical Engineering and Computer
Science
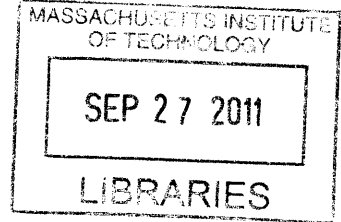in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2011

© Massachusetts Institute of Technology 2011. All rights reserved.

Author ....
Department of Electrical Engineering and Computer Science
August 31, 2011

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

Accepted by ....
Leslie A. Kolodziejski
Chairman, Department Committee on Graduate Theses

On the Capacity of Bosonic Channels

by

Christopher Graham Blake

Submitted to the Department of Electrical Engineering and Computer Science
on August 31, 2011, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering

Abstract

The capacity of the bosonic channel with additive Gaussian noise is unknown, but there is a known lower bound that is conjectured to be the capacity. We have quantified the gap that exists between this known achievable rate and rates achievable by the known methods of detection including direct, heterodyne, and homodyne detection. We have also quantified these capacities in the case of multiple independent spatial modes in terms of spectral and photon efficiency. Furthermore, we have considered the ergodic and outage capacities of fading channel models for far-field and near-field propagation through atmospheric turbulence. For the far field, good models for the transmissivity statistics are known. For the near field we establish bounds on these capacities, and we show that these bounds are reasonably tight. Finally, we extend the results for ergodic capacity to the case of multiple spatial modes where a turbulent atmosphere results in crosstalk between different spatial modes.

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

Acknowledgments

First, I would like to thank my advisor Professor Jeffrey Shapiro for his support during the writing of this thesis. Despite being the director of the Research Laboratory of Electronics and having countless other responsibilities during the past two years, he has always made time to meet with me to discuss my research and his advice and guidance are something without which I could not have completed this thesis.

I would also like to mention my labmates who have been helpful with many discussions during the writing of this thesis.

I would also like to thank Professor Alan Oppenheim who, although he did not actually directly help with the writing of this thesis, was a source of guidance for me during the past two years at MIT.

Finally, I think it is important to mention Professor Alan Edelman whose expertise on random matrix theory, and ability to answer all the questions I had in class about random matrix theory, was invaluable for me in completing the final chapter of this thesis.

Contents

Chapter 1

Introduction

Information theory is concerned with determining the ultimate limits on reliable transmission over noisy communication channels and finding practical means of approaching them. Communication at optical frequencies provides the highest data rates and, through the fiber-optic backbone, is the major thoroughfare for Internet traffic. The application of information theory to determine the ultimate limits on optical communication, however, must go beyond the standard classical-physics paradigm that is used, for example, in the study of wireless communication at microwave frequencies. This is because electromagnetic waves are intrinsically quantum mechanical—their energies are quantized into photons—and high sensitivity photodetection systems are limited by noise of a quantum mechanical origin, whereas microwave systems are generally limited by thermal noise that can be treated classically.

This thesis addresses a collection of interrelated problems associated with the classical information capacities of bosonic channels, i.e., optical communication channels treated quantum mechanically. Before proceeding to discussing the specific problems to be treated, it is useful to present a quick summary of the key ideas from quantum optics. In particular, we will will describe the basics of Dirac-notation quantum mechanics applied to the harmonic oscillator, which models a single mode of the electromagnetic field. This will be followed with quantum descriptions of the three fundamental photodetection techniques: direct, heterodyne, and homodyne detection. Finally, we will present short characterizations

of the optical channels that will be considered.

1.1   Quantum Optics Essentials

It is outside the scope of this thesis to review all the basics of quantum mechanics, but some important things will be reviewed. For a more detailed discussion of vector spaces, inner product spaces, and Hilbert spaces, see [13]. In Dirac notation quantum mechanics, the state of a system can be represented by a vector, called a ket vector, which will be denoted by $|\cdot\rangle$. See [13] for a complete description of the properties of a vector space. These ket vectors form an inner product space. If $|x\rangle$ is a ket vector, then its adjoint (the bra vector) is denoted by $\langle x|$ and the inner product between two vectors $|x\rangle$ and $|y\rangle$ is denoted by $\langle x \mid y \rangle$. Furthermore, a finite energy state has unit length, i.e.,

$$\langle \psi \mid \psi \rangle = 1. \tag{1.1}$$

It may be that a quantum state is a classical mixture of quantum states, i.e., there is some probability distribution for it to be in a particular set of pure quantum states. Suppose that a quantum system is in a classically random mixture of states. Denote those states by $|x_n\rangle$ and the probability that the system is in state $|x_n\rangle$ by p_n. Then we can define the density operator for this quantum system as:

$$\hat{\rho} = \sum_n p_n |x_n\rangle \langle x_n|. \tag{1.2}$$

1.1.1   Operators

Before we discuss the quantum harmonic oscillator, a quick review of quantum observables and operators is needed. See [13] for a more detailed discussion of operators and observables. For a detailed discussion of quantum optics, see [15].

If we let H_s be any Hilbert space, then an operator that maps H_s into H_s, denoted $\hat{T}$, has the property that for any $|x\rangle \in H_s$ there exists a $|y\rangle \in H_s$ such that $\hat{T}|x\rangle = |y\rangle$

10

An observable is an Hermitian operator that has a complete set of eigenkets. It is a measurable variable in a quantum system and can be represented by:

$$\hat{O} = \sum_n o_n \left| o_n \right\rangle \left\langle o_n \right|, \tag{1.3}$$

where $\{\left| o_n \right\rangle\}$ are the orthonormal eigenkets and a discrete eigenspectrum $\{o_n\}$ has been assumed. If the system is in state $\left| \psi \right\rangle$, a fundamental postulate of quantum mechanics says that a measurement of this observable gives outcome o_n with probability

$$\Pr\left(\text{outcome} = o_n\right) = \left| \left\langle o_n \mid \psi \right\rangle \right|^2, \tag{1.4}$$

when the o_n are distinct.

If a system is in a mixed state represented by a density operator $\hat{\rho} = \sum_n p_n \left| x_n \right\rangle \left\langle x_n \right|$ we can use classical probability theory combined with the probability law for a quantum measurement made on a pure state to show that

$$\Pr\left(\text{outcome} = o\right) = \sum_n p_n \Pr\left(\text{outcome} = o | \text{state} = \left| x_n \right\rangle\right) \tag{1.5}$$

$$= \sum_n p_n |\left\langle o \mid x_n \right\rangle|^2 = \sum_n p_n \left\langle o \mid x_n \right\rangle \left\langle x_n \mid o \right\rangle = \left\langle o \right| \left[\sum_n p_n \left| x_n \right\rangle \left\langle x_n \right| \right] \left| o \right\rangle = \left\langle o \right| \hat{\rho} \left| o \right\rangle. \tag{1.6}$$

1.1.2   Quantum Harmonic Oscillator

We can now discuss the quantum harmonic oscillator. The quantum harmonic oscillator represents a single mode of the electromagnetic field. In the semiclassical understanding of light, a single mode of light is represented by

$$ae^{-i\omega t}, \tag{1.7}$$

11

where a is some complex number. Then, in this system, the energy is given by

$$H = \hbar\omega |a|^2. \tag{1.8}$$

However, in the quantum case, a single mode of the harmonic oscillator is given by

$$\hat{a}e^{-i\omega t}, \tag{1.9}$$

where $\hat{a}$ is the annihilation operator. In this case, the energy is represented by the Hamiltonian operator and is given by

$$\hat{H} = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) \tag{1.10}$$

where the $\frac{1}{2}$ represents the zero-point energy and $\hat{a}^\dagger$ is the creation operator, i.e., the adjoint of $\hat{a}$. Annihilation and creation operators have the property that

$$\left[\hat{a}^\dagger, \hat{a} \right] = 1, \tag{1.11}$$

where $\left[\hat{a}, \hat{b} \right] = \hat{a}\hat{b} - \hat{b}\hat{a}$ is the commutator of $\hat{a}$ and $\hat{b}$.

It is now useful to define some states of interest. One such set of states are the number states, denoted $|n\rangle$, defined for all integers $n \geq 0$. They have the property that

$$\hat{N} |n\rangle = n |n\rangle, \tag{1.12}$$

where $\hat{N}$ is the number operator and is defined using the annihilation and creation operators by

$$\hat{N} = \hat{a}^\dagger \hat{a}. \tag{1.13}$$

The number states are a complete orthormal set of states. In other words,

$$\langle n \mid m \rangle = \delta_{nm} = \begin{cases} 1, \ n = m \\ 0, \ n \neq m \end{cases} \tag{1.14}$$

and the states resolve the identity

$$\hat{I} = \sum_n |n\rangle \langle n|. \tag{1.15}$$

Other states of interest are the coherent states, $|a\rangle$, which have the property that $\hat{a} |a\rangle = a |a\rangle$, for any complex number a. In other words, they are eigenkets of the annihilation operator. Unlike the number states, these states are not orthonormal, but they are overcomplete. Their inner product is

$$\langle a \mid a' \rangle = e^{-\frac{|a|^2}{2} - \frac{|a'|^2}{2} + a'a^*}, \tag{1.16}$$

and they resolve the identity as follows:

$$\hat{I} = \int \frac{d^2 a}{\pi} |a\rangle \langle a|. \tag{1.17}$$

Other operators we will define are the quadrature operators, viz., the real and imaginary parts of the annihilation operator,

$$\hat{a}_1 = \text{Re} (\hat{a}) = \frac{\hat{a} + \hat{a}^\dagger}{2}, \tag{1.18}$$

and

$$\hat{a}_2 = \text{Im} (\hat{a}) = \frac{\hat{a} - \hat{a}^\dagger}{2i}. \tag{1.19}$$

These operators lead to two more sets of states, the quadrature eigenstates. These states

have the following properties

$$\hat{a}_1 \left| a_1 \right\rangle_1 = a_1 \left| a_1 \right\rangle_1 \tag{1.20}$$

and

$$\hat{a}_2 \left| a_2 \right\rangle_2 = a_2 \left| a_2 \right\rangle_2 \tag{1.21}$$

where a_1 and a_2 are real numbers. The quadrature eigenkets have inner products given by

$$_1\!\left\langle a_1 \mid a_1' \right\rangle_1 = \delta\left(a_1 - a_1'\right) \tag{1.22}$$

and

$$_2\!\left\langle a_2 \mid a_2' \right\rangle_2 = \delta\left(a_2 - a_2'\right) \tag{1.23}$$

so they have infinite energy, but are orthogonal and complete.

1.1.3 Direct Detection Statistics

In a real photodetector, there are many sources of noise. In these systems, incoming light illuminates an optical filter, and the light emerging from this filter then strikes the photodetector, which converts the light into a light-induced current called photocurrent. Photodetectors have some current flow in the absence of light called dark current. High sensitivity photodetectors will amplify this current, and this multiplication has some random noise associated with it. There is also thermal noise that is injected into the detector's circuit, and there is also a limited bandwidth to the detector. However, for the purposes of this thesis, all of these noise elements will be stripped away, and only the fundamental quantum noise will remain. A diagram of the direct detection system we will use is given in Fig. 1-1. Ideal direct detection is equivalent to making a measurement of the $\hat{N}$ observable. If a coherent state illuminates the detector, then from the postulates of quantum mechanics, the probability of
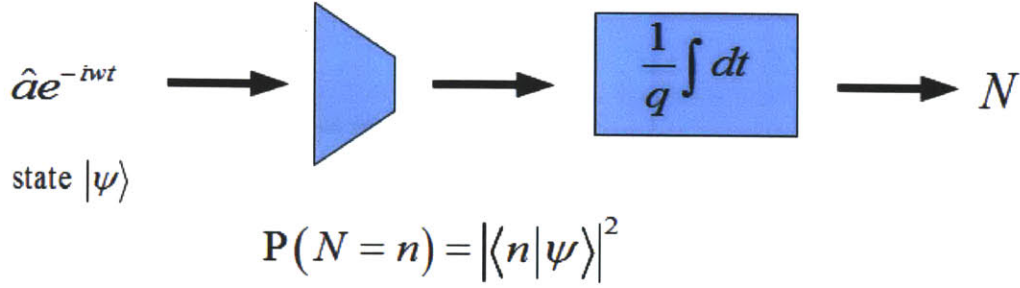
Figure 1-1: Diagram of a direct detection system.

outcome n being observed is

$$\mathrm{P}\left(N = n | \mathrm{state} = |a\rangle \right) = |\langle n \mid a\rangle|^2 = \frac{|a|^{2n} e^{-|a|^2}}{n!} \tag{1.24}$$

where we have used the number-state representation

$$|a\rangle = \sum_{n=1}^{\infty} \frac{a^n e^{-|a|^2/2}}{\sqrt{n!}} |n\rangle . \tag{1.25}$$

1.1.4 Heterodyne Detection Statistics

A description of balanced heterodyne detection is given in [13]. Figure 1-2 shows a diagram of a heterodyne detector. In this setup, a single-mode signal field of frequency ω is combined on a 50/50 beam splitter with a strong local-oscillator field frequency of $\omega - \omega_{IF}$, where ω_{IF} is an intermediate frequency.

In [13] it is shown that this setup realizes a measurement of the $\hat{a}$ positive operator-valued measurement. Supposing that the quantum system is in a particular coherent state $|\alpha_s\rangle$ then
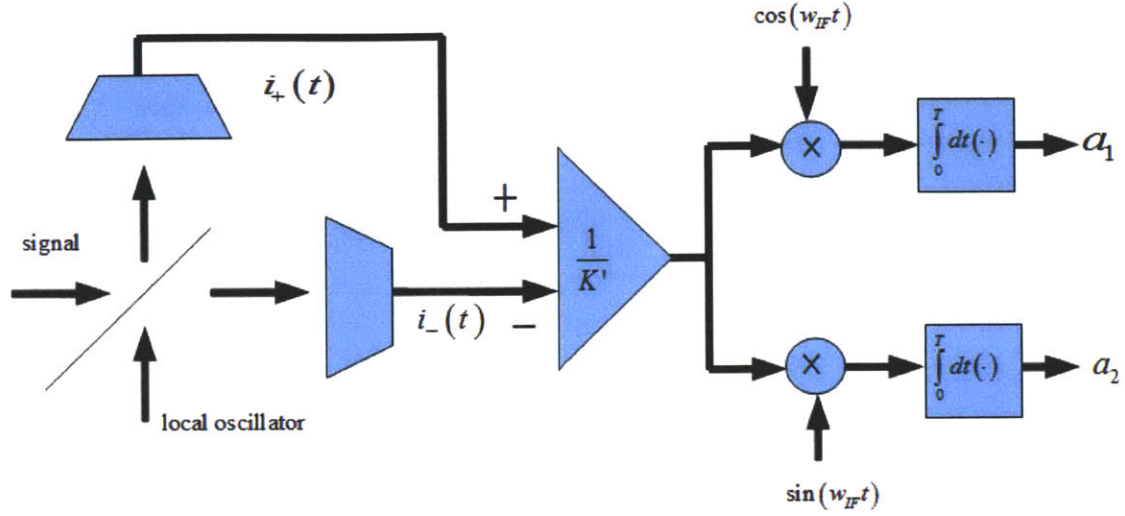
Figure 1-2: Diagram of a heterodyne detection system.

the probability density function for the heterodyne measurement to result in a value α is

$$p\left(\alpha|\text{state} = \alpha_s\right) = |\langle \alpha \mid \alpha_s \rangle|^2 = \frac{e^{-|\alpha-\alpha_s|^2}}{\pi}, \tag{1.26}$$

which is a complex Gaussian distribution with real and imaginary parts having statistically independent, variance $\frac{1}{2}$ Gaussian distributions. If the state we are in is a random mixture of coherent states with probability density $p\left(\alpha_s\right)$ for the state's eigenvalue, then the heterodyne detection measurement statistics will be given by:

$$p\left(\alpha\right) = \int d^2\alpha_s p\left(\alpha_s\right) \frac{e^{-|\alpha-\alpha_s|^2}}{\pi}. \tag{1.27}$$

1.1.5 Homodyne Detection Statistics

A detailed account of the quantum theory of photodetection comes from [13]. A diagram of homodyne detection is given in Fig. 1-3. It is similar to heterodyne detection, in that the

16

signal field is combined on a 50/50 beam splitter with a strong local-oscillator field, only now the signal and local oscillator are both at frequency ω. Homodyne detection corresponds to a measurement of

$$\hat{a}_\theta = \mathrm{Re}\left(\hat{a}_s e^{-j\theta}\right), \qquad (1.28)$$

where θ is the relative phase between the signal and the local oscillator. If $\theta = 0$, then this corresponds to measurement of the $\hat{a}_{s_1} = \mathrm{Re}\left(\hat{a}_s\right)$ observable. If $\theta = \frac{\pi}{2}$, then this corresponds to the $\hat{a}_{s_2} = \mathrm{Im}\left(\hat{a}_s\right)$ observable.
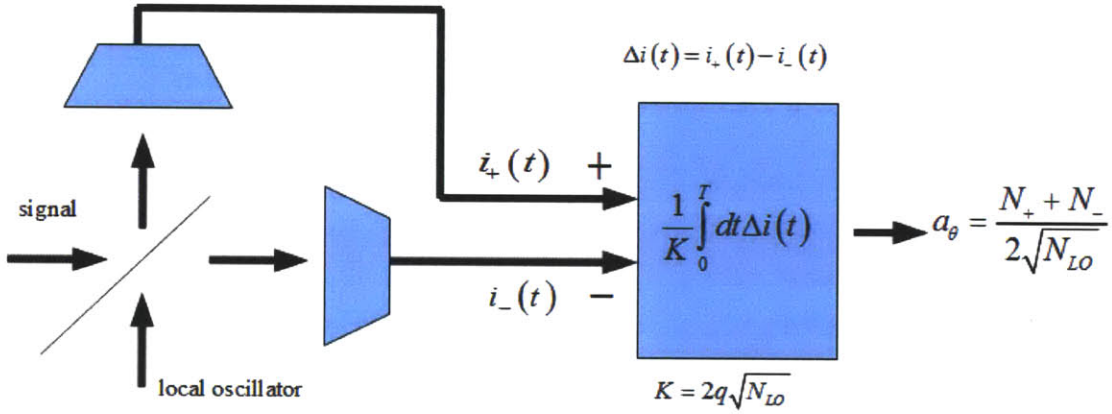
Figure 1-3: Diagram of a homodyne detection system.

If the signal is in the coherent state $|a_s\rangle$, then the homodyne detector's output, as calculated in [13], will be a Gaussian random variable with mean $\mathrm{Re}\left(a_s e^{-j\theta}\right)$ and variance $\frac{1}{4}$. For an arbitrary state $|\psi\rangle$, the homodyne detector's output will have probability density function

$$p\left(a_\theta | |\psi\rangle\right) = |_\theta \langle a_\theta \mid \psi\rangle|^2, \qquad (1.29)$$

where $\hat{a}_\theta |a_\theta\rangle_\theta = a_\theta |a_\theta\rangle_\theta$ specifies the eigenket-eigenvalue decomposition of $\hat{a}_\theta$.

1.2   Modal Theory of Diffraction in Vacuum
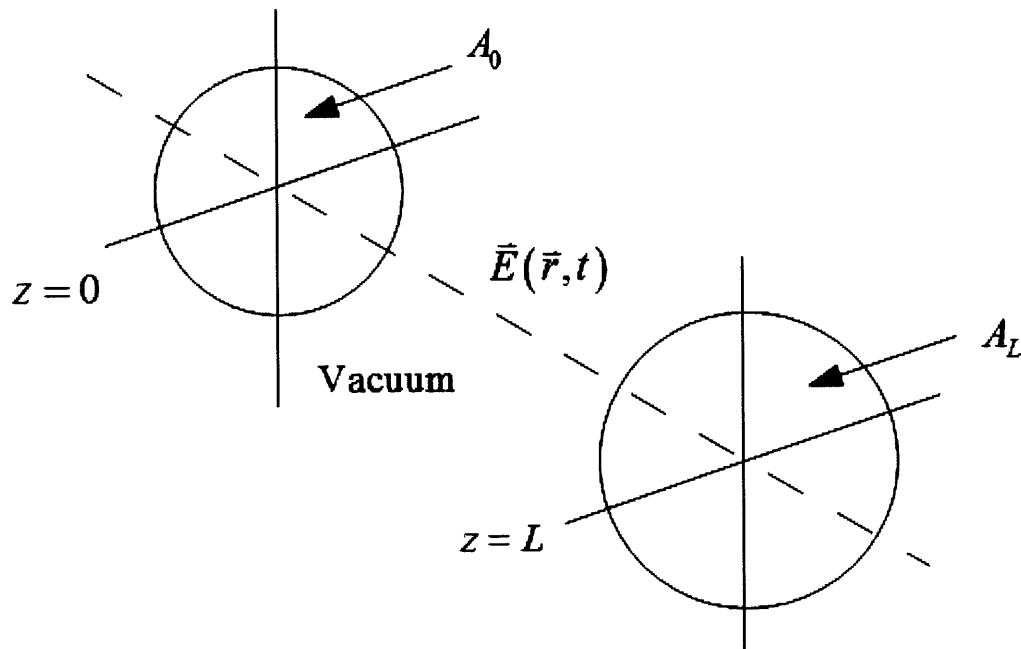
Figure 1-4: Diagram of prototypical line-of-site free-space propagation geometry.

Figure 1-4 gives a diagram of a prototypical line-of-sight free-space propagation geometry. For a linearly polarized transmitter, we can use scalar wave optics. For a quasimonochromatic, paraxial propagation transmitter we can employ the Huygens-Fresnel principle

$$E_L\left(\vec{\rho}',t\right) = \int_{A_0} d\vec{\rho}\, E_0\left(\vec{\rho},t - \frac{L}{c}\right)\frac{e^{ikL + i\frac{h|\vec{\rho}' - \vec{\rho}|^2}{2L}}}{i\lambda L} \tag{1.30}$$

for $\vec{\rho}' \in A_L$, ω_0 is the center frequency, $\lambda = 2\pi c/\omega_0$ is the wavelength, and $k = \omega_0/c$ is the

wave number. Also,

$$E_0\left(\vec{\rho}, t\right) = \left. E\left(\vec{r}, t\right)\right|_{\vec{r}=\vec{\rho}, z=0} \tag{1.31}$$

and

$$E_L\left(\vec{\rho'}, t\right) = \left. E\left(\vec{r}, t\right)\right|_{\vec{r}=\vec{\rho'}, z=L} \tag{1.32}$$

are the input and output complex-field envelopes. For convenience these are taken to have units $\sqrt{\text{photons/m}^2\text{s}}$, even though they are classical fields. For our quantum capacity analyses they will become the eigenfunctions of coherent states of the electromagnetic field.

We can perform a singular value decomposition for

$$h\left(\vec{\rho'} - \vec{\rho}\right) = \frac{e^{ikL+i\frac{k|\vec{\rho'}-\vec{\rho}|^2}{2L}}}{i\lambda L} \tag{1.33}$$

for $\vec{\rho} \in A_0, \vec{\rho'} \in A_L$, i.e.

$$h\left(\vec{\rho'} - \vec{\rho}\right) = \sum_{n=0}^{\infty} \sqrt{\eta_n}\phi_n\left(\vec{\rho'}\right)\Phi_n^*\left(\vec{\rho}\right) \tag{1.34}$$

for $\vec{\rho} \in A_0, \vec{\rho'} \in A_L$ with $\Phi_n\left(\vec{\rho}\right)$ a complete orthonormal set on A_0 and $\phi_n\left(\vec{\rho'}\right)$ a complete orthonormal set on A_L. We also have that

$$1 \geq \eta_1 \geq \eta_2 \geq ...\eta_n \geq 0. \tag{1.35}$$

Using this decomposition, we get

$$E_0\left(\vec{\rho}, t\right) = \sum_{n=0}^{\infty} E_{0_n}\left(t\right)\Phi_n\left(\vec{\rho}\right) \tag{1.36}$$

19

and

$$E_L\left(\vec{\rho}\,',t\right) = \sum_{n=0}^{\infty} \sqrt{\eta_n} E_{0_n}\left(t - L/c\right) \phi_n\left(\vec{\rho}\,'\right). \tag{1.37}$$

Thus η_n represents the fractional photon-flux transfer from A_0 to A_L when the $\Phi_n\left(\vec{\rho}\right)$ mode is transmitted from A_0. Expanding these $E_{0_n}\left(t\right)$ on a transmission interval $0 \le t \le T$ using a generalized Fourier series

$$E_{0_n}\left(t\right) = \sum_{m=1}^{\infty} a_{0_{nm}} \xi_m\left(t\right), \tag{1.38}$$

where $\{\xi_m\left(t\right)\}$ is a complete orthonormal set on $0 \le t \le T$, we get

$$E_{L_n}\left(t\right) = \sum_{m=1}^{\infty} a_{L_{nm}} \xi_m\left(t - L/c\right) = \sum_{m=1}^{\infty} \sqrt{\eta_n} a_{0_{nm}} \xi_m\left(t - L/c\right). \tag{1.39}$$

Classically the spatiotemporal mode transformation

$$a_{0_{nm}} \rightarrow \sqrt{\eta_n} a_{L_{nm}} \tag{1.40}$$

will be the foundation for our free-space propagation channel models. Because we are interested in quantum limits on capacity, we must convert this classical description to one involving modal photon annihilation operators. Here we have $a_{0_{nm}}$ being replaced by $\hat{a}_{0_{nm}}$ and $a_{L_{nm}}$ being replaced by $\hat{a}_{L_{nm}}$. However, in order to ensure commutator-bracket conservation the classical modal transformation from $a_{0_{nm}}$ to $a_{L_{nm}}$ is replaced by

$$\hat{a}_{L_{nm}} = \sqrt{\eta_n} \hat{a}_{0_{nm}} + \sqrt{1 - \eta_n} \hat{b}_{nm}, \tag{1.41}$$

where $\hat{b}_{nm}$ is the modal annihilation operator for an environmental (noise) mode.

1.3 Propagation through Atmospheric Turbulence

Optical communication through the earth's atmosphere is subject to many impairments that go well beyond the diffraction-induced propagation loss — represented by the modal transmissivities $\{\eta_n\}$—seen in the last section for vacuum propagation. There are a number of challenges in this type of propagation. Bad weather, for example fog, rain, or snow, and atmospheric molecular constituents cause absorption and scattering that degrade the performance of optical communication systems. We will focus on understanding clear weather propagation away from absorption lines. Temporal and spatial thermal inhomogeneities in the atmosphere under clear weather conditions cause random fluctuations in the refractive index at optical wavelengths. These refractive-index perturbations are usually referred to as atmospheric turbulence and lead to amplitude and phase fluctuations of light beams propagating through the atmosphere. Propagation through a turbulent atmosphere has been described in [16]. Furthermore, a discussion of the extended Huygens-Fresnel principle as it applies to atmospheric turbulence can be found in [7].

Consider a field propagating from A_0 in the $z = 0$ plane to A_L in the $z = L$ plane, where the duration of the pulse is much less than the atmospheric coherence time. In general, the coherence time for the atmosphere is on the order of milliseconds [7, 10]. In this case, the introduction of atmospheric turbulence in clear weather conditions can be described by the extended Huygens-Fresnel Principle

$$E_L\left(\vec{\rho}',t\right) = \int_{A_0} d\vec{\rho}\, E_0\left(\vec{\rho},t-\frac{L}{c}\right)\frac{e^{ikL+i\frac{k|\vec{\rho}'-\vec{\rho}|^2}{2L}}}{i\lambda L}e^{\chi(\vec{\rho}',\vec{\rho})+i\phi(\vec{\rho}',\vec{\rho})}, \tag{1.42}$$

where χ and ϕ are real-valued random turbulence-induced log-amplitude and phase perturbations respectively.

Within the weak-perturbation regime, χ and ϕ can be taken to be jointly Gaussian random processes. Energy conservation implies that the mean of χ will equal minus its variance; the mean of ϕ can be taken to be zero, and its variance much greater than one.

Expressions are available for the covariance functions

$$K_{\chi\chi}\left(\Delta\vec{\rho}',\Delta\vec{\rho}\right) = \left\langle \chi\left(\vec{\rho}' + \Delta\vec{\rho}', \vec{\rho} + \Delta\vec{\rho}\right)\chi\left(\vec{\rho}',\vec{\rho}\right)\right\rangle - \left\langle \chi\right\rangle^2, \tag{1.43}$$

$$K_{\phi\phi}\left(\Delta\vec{\rho}',\Delta\vec{\rho}\right) = \left\langle \phi\left(\vec{\rho}' + \Delta\vec{\rho}', \vec{\rho} + \Delta\vec{\rho}\right)\phi\left(\vec{\rho}',\vec{\rho}\right)\right\rangle - \left\langle \phi\right\rangle^2, \tag{1.44}$$

and

$$K_{\chi\phi}\left(\Delta\vec{\rho}',\Delta\vec{\rho}\right) = \left\langle \chi\left(\vec{\rho}' + \Delta\vec{\rho}', \vec{\rho} + \Delta\vec{\rho}\right)\phi\left(\vec{\rho}',\vec{\rho}\right)\right\rangle - \left\langle \chi\right\rangle\left\langle \phi\right\rangle, \tag{1.45}$$

but these will not be needed for our work.

1.4   Information Theory Background

Shannon defined the maximum rate for a given channel at which reliable communication is possible in his noisy channel coding theorem [3]. His results have been applied to determine the capacities of lightwave channels in which conventional photodetection schemes are employed, i.e., homodyne, heterodyne, and direct detection [1]. These studies primarily use semiclassical photodetection theory, in which light is treated classically and detector shot noise sets the fundamental performance limit. However, high sensitivity photodetection is fundamentally limited by quantum noise. Thus, light must be treated quantum mechanically and the receiver must be allowed to use arbitrary quantum measurements if we are to establish the ultimate capacities of lightwave channels. The quantum equivalent of Shannon's coding theorem is the Holevo-Schumacher-Westmoreland theorem [4]. The classical capacity of the single-mode lossy channel is established by random coding arguments akin to those employed in classical information theory. A set of symbols j is represented by a collection of input states that are chosen according to some prior distribution p_j. The Holevo information

associated with the priors $\{p_j\}$ and density operators $\{\hat{\sigma}_j\}$ is:

$$\chi\left(p_j, \hat{\sigma}_j\right) = S\left(\sum_j p_j \hat{\sigma}_j\right) - \sum_j p_j S\left(\hat{\sigma}_j\right), \tag{1.46}$$

where $S\left(\hat{\sigma}\right) = -\text{tr}\left(\hat{\sigma}\ln\left(\hat{\sigma}\right)\right)$ is the Von Neuman entropy of the density operator $\hat{\sigma}$. The Holevo-Schumacher-Westmoreland theorem then gives the capacity of the thermal-noise channel as:

$$C = \sup_n \left(C_n/n\right) = \sup_n \left(\max_{\{p_j, \hat{\rho}_j\}} \left\{\chi\left[p_j, \left(\text{E}_\eta^N\right)^{\otimes n}\left(\hat{\rho}_j\right)\right]/n\right\}\right). \tag{1.47}$$

Here, $\left(\text{E}_\eta^N\right)^{\otimes n}\left(\hat{\rho}_j\right)$ is the density operator at the output of a bosonic channel with transmittivity η and thermal noise with average photon number N when $\hat{\rho}$ is the density operator of its input. The $\otimes n$ superscript indicates a sequence of n independent uses of the channel; $\hat{\rho}_j$ is the joint density operator over those n channel uses. The capacity of this channel is taken as the supremum over n-channel-use symbols because it may be that the channel is superadditive.

What is unique about a quantum channel, as opposed to a classical channel, is that the statistics of this channel are determined both by the state that is sent and also the quantum measurement employed. The capacity of various types of these quantum channels is outlined in [1]. The capacity of a pure-loss bosonic channel is known. This is a channel in which signal photons may be lost in propagation and the channel injects the minimum (vacuum state) quantum noise needed to preserve the Heisenberg uncertainty principle. The thermal-noise channel is a lossy channel in which noise is injected from the external environment according to a Gaussian distribution. A lower bound on the capacity of this channel has been proven, and it is conjectured to be the ultimate quantum capacity of this channel. However, this capacity cannot be reached by the use of heterodyne, homodyne, or direct detection alone. Of interest is how big a gap exists between this conjectured quantum capacity and the known capacities for heterodyne and homodyne detection. In the case of direct detection, the capacity of a thermal-noise channel in which direct detection is used is not known, but

lower bounds can be computed. We have derived some asymptotic properties of the direct detection channel and quantified the gap that exists between the known detection techniques and the conjectured quantum capacity of the thermal-noise channel.

In this thesis we will outline some of the known capacities of various types of bosonic channels and also present the conjectured capacity for the thermal-noise channel. Next, we will present the statistics for a thermal-noise channel in the case of direct detection. We will then prove that in the limit of low average received photon number, the direct detection capacity is asymptotic to the conjectured thermal-noise channel capacity. Furthermore, we will discuss the capacity of a channel that experiences fading due to atmospheric turbulence. We will look at the capacity of fading channels operating in the far field, where exponential or lognormal fading can occur. Here we will focus on the outage capacity and the ergodic capacity of this channel. Then, we will discuss bounds on the capacities of the fading channels in the near field. Finally, we will show results on how to extend these calculations to cases in which multiple spatial modes are transmitted through a turbulent atmosphere, which can result in random interference.

We begin, in Chapter 2, by presenting the capacity of the pure-loss channel and a conjectured capacity for the thermal-noise channel. From this, we will derive some asymptotic properties of the thermal-noise direct detection channel. We will then calculate the capacities in various average received photon number regimes and quantify the gap that exists between the three known detection techniques and a theoretically optimal detection technique. Here we will find there is only a modest gap between capacities achieved with known technologies and the theoretical maximum. Nevertheless, if both spectral and photon efficiencies are to be optimized, high spectral and high photon efficiencies can be achieved using far fewer spatial modes if optimal detection is used as opposed to known techniques.

In Chapter 3, we will define the notion of outage and ergodic capacity and derive bounds on these capacities in the case of near and far field optical communication through atmospheric turbulence. We will do this assuming lognormal and exponential fading statistics, and also derive some bounds for ergodic capacity in the case of worst-case fading statistics.

Finally, we will investigate the capacity of a multiple input, multiple output fading channel in which there is intersymbol interference. We will show how to calculate bounds on the capacity of a particular sparse aperture system and show through simulation how tight these bounds are.

Chapter 2

Pure-Loss and Thermal-Noise Channels

The statistics of the thermal-noise channel model can be derived from the commutator preserving beam splitter relationship:

$$\hat{a}' = \sqrt{\eta}\hat{a} + \sqrt{1 - \eta}\hat{b}, \tag{2.1}$$

where $\hat{a}$, $\hat{b}$, and $\hat{a}'$ are modal photon annihilation operators and $0 \leqslant \eta \leqslant 1$ is the channel transmissivity. Here, $\hat{a}$ is the input mode whose information-bearing state is controlled by the transmitter, $\hat{b}$ is a mode injected by the channel, and $\hat{a}'$ is the output mode from which the receiver will attempt to retrieve the transmitted message.

For the pure-loss channel, the $\hat{b}$ mode is in its vacuum state. For the thermal-noise channel this mode is in a thermal state, i.e., a Gaussian mixture of coherent states with an average photon number $N_b > 0$ [1] whose density operator is therefore

$$\hat{\rho}_b = \int d^2\beta \frac{\exp\left(-|\beta|^2/N_b\right)}{\pi N_b} |\beta\rangle \langle\beta|. \tag{2.2}$$

For a channel with the beam splitter relation in (2.1), the average received photon number

27

that is due to the transmitter is related to the average transmitted photon number as follows:

$$\bar{N}_r = \eta \bar{N}_t. \tag{2.3}$$

Similarly, when the average photon number of the noise input into the beam splitter relation is N_b, the average photon number for the received noise, N_r is given by:

$$N_r = (1 - \eta) N_b. \tag{2.4}$$

For the sake of notational simplicity, in what follows $\bar{N}$ will be used to denote the average received signal photon number ($\bar{N}_r$), and N will denote the average received noise photon number (N_r) in what follows.

It has been shown that the pure-loss channel ($N = 0$) is not superadditive, and that its classical information capacity is achieved by coherent-state encoding. The capacity is given by [5]:

$$C = g(\bar{N}) \tag{2.5}$$

where $g(x)$ is the Shannon entropy of the Bose-Einstein distribution with mean x, i.e.,

$$g(x) = (x + 1) \log (x + 1) - x \log (x). \tag{2.6}$$

The base used for the logarithm in this expression sets the units for measuring information content, i.e., base e leads to nats and base 2 leads to bits.

It is conjectured that the capacity of the thermal-noise channel with average received signal photon number $\bar{N}$ and average received noise photon number N is [6]:

$$C_{conj}(\bar{N}, \ N) = g(\bar{N} + N) - g(N). \tag{2.7}$$

This rate is achievable with coherent-state encoding, and so is a lower bound of the true

capacity, but it has not been shown to be the capacity.

We also know the classical information capacities of two types of thermal-noise channels, when the quantum measurement at the receiver is constrained to be either heterodyne detection or homodyne detection and the transmitter uses coherent-state encoding. Their capacities are given below:

$$C_{het}\left(\bar{N}, N\right) = \log\left(1 + \frac{\bar{N}}{1 + N}\right) \tag{2.8}$$

$$C_{hom}\left(\bar{N}, N\right) = \frac{1}{2}\log\left(1 + \frac{4\bar{N}}{1 + 2N}\right). \tag{2.9}$$

In the remainder of this chapter we will address the following issues with respect to the capacity of the thermal-noise channel. First, in Section 2.1, we will study the asymptotic behaviors of the thermal-noise channel with coherent-state encoding and direct detection. Then, in Section 2.2 we will consider the photon and spectral efficiencies that can be achieved for the thermal-noise channel. Here, we will include the use of multiple spatial modes in a near-field propagation geometry.

2.1 Asymptotic Behaviors of the Direct Detection Channel Capacity

A closed-form expression for the capacity of the single-mode direct detection channel is not known. The measurement statistics for an ideal single-mode direct detection channel (in which the noise injected into the channel is a Gaussian mixture of coherent states), are given by the Laguerre distribution for the observed photon count n

$$\Pr(n) = e^{\frac{-|\alpha_s|^2}{(N+1)}} \frac{N^n}{(N+1)^{n+1}} L_n\left(\frac{-|\alpha_s|^2}{N(N+1)}\right), \tag{2.10}$$

where $|\alpha_s|^2$ is the average received photon number of a single coherent-state transmission $|\alpha_s\rangle$ and N is the average received noise photon number, and

$$L_n(x) = \sum_{m=0}^{n} (-1)^m \begin{pmatrix} n \\ n-m \end{pmatrix} \frac{x^m}{m!} \tag{2.11}$$

is the nth order Laguerre polynomial. More details of the derivation of these statistics can be found in Appendix B.

We can derive some bounds for the capacity of this direct-detection channel given a constraint on the average received signal photon number by considering an on-off keying transmitter. The direct-detection receiver can decide if a 0 was sent or a 1 was sent by simply saying whether or not a photon count of 0 was received. This is not a minimum probability of error decision rule, but we will show that in the low received photon number regime it is asymptotically optimal. In this channel, the probability (p) of sending a 1 must be such that

$$|\alpha_s|^2 p = \bar{N}. \tag{2.12}$$

With this arrangement we obtain a binary channel whose transition probabilities are shown in Fig. 2-1. We then note that for every particular value of p in this binary channel, the mutual information is given by:

$$I(X;Y) = H(Y) - H(Y|X) \tag{2.13}$$

$$= H_B \left(\frac{pe^{\frac{-\bar{N}}{p(N+1)}}}{N+1} + \frac{1-p}{N+1} \right) - pH_B \left(\frac{e^{\frac{-\bar{N}}{p(N+1)}}}{N+1} \right) - (1-p)H_B \left(\frac{1}{N+1} \right), \tag{2.14}$$

where $H(Y)$ is the Shannon entropy [3] and

$$H_B(x) = -x\log(x) - (1-x)\log(1-x) \tag{2.15}$$
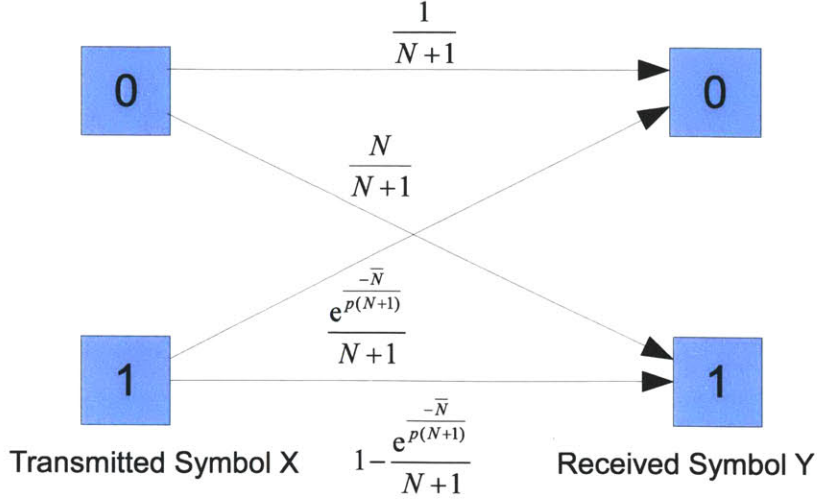
is the binary entropy function.

Figure 2-1: Binary channel that can be produced with on-off keying of coherent-state light and direct detection on the thermal-noise channel.

The capacity of a channel with Laguerre statistics is at least the maximum over p of the preceding expression, where p can range from 0 to 1. In particular, for every c and sufficiently small $\bar{N}$, we can use $p = c\bar{N}$ so that the mutual information is equal to:

$$I(X;Y) = H_B \left(\frac{c\bar{N}e^{\frac{-1}{c(N+1)}}}{N+1} + \frac{1-c\bar{N}}{N+1} \right) - c\bar{N}H_B \left(\frac{e^{\frac{-1}{c(N+1)}}}{N+1} \right) - (1-c\bar{N})H_B \left(\frac{1}{N+1} \right).$$

(2.16)

Let us explore the behavior of this expression for $\bar{N} \ll 1$.

Our focus on the $\bar{N} \ll 1$ regime stems from the fact that for fixed N heterodyne detection has a capacity that satisfies

$$\lim_{\bar{N} \to \infty} \frac{C_{het}\left(\bar{N}, N\right)}{C_{conj}\left(\bar{N}, N\right)} = 1,$$

(2.17)

whereas, as will be seen when we present capacity plots later in this section, both heterodyne and homodyne detection have capacities that fall well below $C_{conj}\left(\bar{N}, N\right)$ for $\bar{N} \ll 1$.

31

Consider first the case in which $\bar{N} = N$, i.e., when the average received signal photon number is equal to the average received noise photon number. We thus know that:

$$\lim_{\bar{N}\to 0} \frac{C_{dir}(\bar{N},\bar{N})}{-\bar{N}\log\bar{N}} \geqslant \lim_{\bar{N}\to 0} \frac{H_B\left(\frac{c\bar{N}e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} + \frac{1-c\bar{N}}{\bar{N}+1}\right) - c\bar{N}H_B\left(\frac{e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1}\right) - (1-c\bar{N})H_B\left(\frac{1}{\bar{N}+1}\right)}{-\bar{N}\log\bar{N}}.$$

(2.18)

In Appendix A we will show that the limit on the right equals $c - ce^{-\frac{1}{c}}$ and thus conclude:

$$\lim_{\bar{N}\to 0} \frac{C_{dir}(\bar{N},\bar{N})}{-\bar{N}\log\bar{N}} \geqslant c - ce^{-\frac{1}{c}}.$$

(2.19)

It is easy to show that

$$\lim_{c\to\infty} c - ce^{-\frac{1}{c}} = 1.$$

(2.20)

For every c, there is a sufficiently small $\bar{N}$ such that $p = c\bar{N}$ is a valid probability for all $\bar{N}$ less than this value. Thus, we conclude that

$$\lim_{\bar{N}\to 0} \frac{C_{dir}(\bar{N},\bar{N})}{-\bar{N}\log\bar{N}} \geqslant 1.$$

(2.21)

However, from [2] we know that

$$\lim_{\bar{N}\to 0} \frac{C_{dir}(\bar{N},0)}{-\bar{N}\log\bar{N}} = 1.$$

(2.22)

We also know that the addition of random noise can only decrease the capacity, thus

$$C_{dir}(\bar{N},\bar{N}) \leqslant C_{dir}(\bar{N},0).$$

(2.23)

So we get

$$\lim_{\bar{N} \to 0} \frac{C_{dir}(\bar{N}, \bar{N})}{-\bar{N} \log \bar{N}} \leqslant \lim_{\bar{N} \to 0} \frac{C_{dir}(\bar{N}, 0)}{-\bar{N} \log \bar{N}} = 1. \tag{2.24}$$

Combining our results, it must be that

$$\lim_{\bar{N} \to 0} \frac{C_{dir}(\bar{N}, \bar{N})}{-\bar{N} \log \bar{N}} = 1. \tag{2.25}$$

Now let us consider the case in which the noise is a constant multiple of the average received photon number. Here we find that

$$\lim_{\bar{N} \to 0} \frac{C_{dir}(\bar{N}, k\bar{N})}{-\bar{N} \log \bar{N}} \geqslant \lim_{N \to 0} \frac{H_B \left(\frac{c\bar{N} e^{\frac{-1}{c(k\bar{N}+1)}}}{k\bar{N}+1} + \frac{1-c\bar{N}}{k\bar{N}+1} \right) - c\bar{N} H_B \left(\frac{e^{\frac{-1}{c(k\bar{N}+1)}}}{k\bar{N}+1} \right) - (1 - c\bar{N}) H_B \left(\frac{1}{k\bar{N}+1} \right)}{-\bar{N} \log \bar{N}}. \tag{2.26}$$

We can then show that:

$$\lim_{N \to 0} \frac{C_{dir}(\bar{N}, k\bar{N})}{-\bar{N} \log \bar{N}} = \text{c} - \text{c}e^{-\frac{1}{c}}; \tag{2.27}$$

the details of this calculation are in Appendix A. From this result we can use the same argument employed previously to show that

$$\lim_{\bar{N} \to 0} \frac{C_{dir}(\bar{N}, k\bar{N})}{-\bar{N} \log \bar{N}} = 1. \tag{2.28}$$

In other words, the direct-detection capacity of the thermal-noise channel has a low average received signal photon number asymptote that is independent of the strength of the noise. Of course, convergence to this asymptote can be expected to depend on the noise strength.

We will now plot the various capacities for different values of the average noise photon number. Since we do not have a closed form expression for the on-off keying direct detection capacity, we instead numerically estimate the capacity by optimizing (2.14) over p. In Fig. 2-2 we show the capacities of the heterodyne, homodyne, and direct detection on-off keying

channels, as well as the conjectured quantum capacity when $N = \bar{N}$. From this figure we see that there is not a large difference between achievable capacities using the best of direct, heterodyne, and homodyne detection and the conjectured quantum capacity when the noise is equal to the average received signal photon number. However, we are more interested in the case in which the average received noise photon number is some realistic constant value. At near-infrared to visible wavelengths, the average noise photon number is typically much less than unity [8]. We set $N = 10^{-6}$ and plot the thermal-noise capacities for this case in Fig. 2-3.
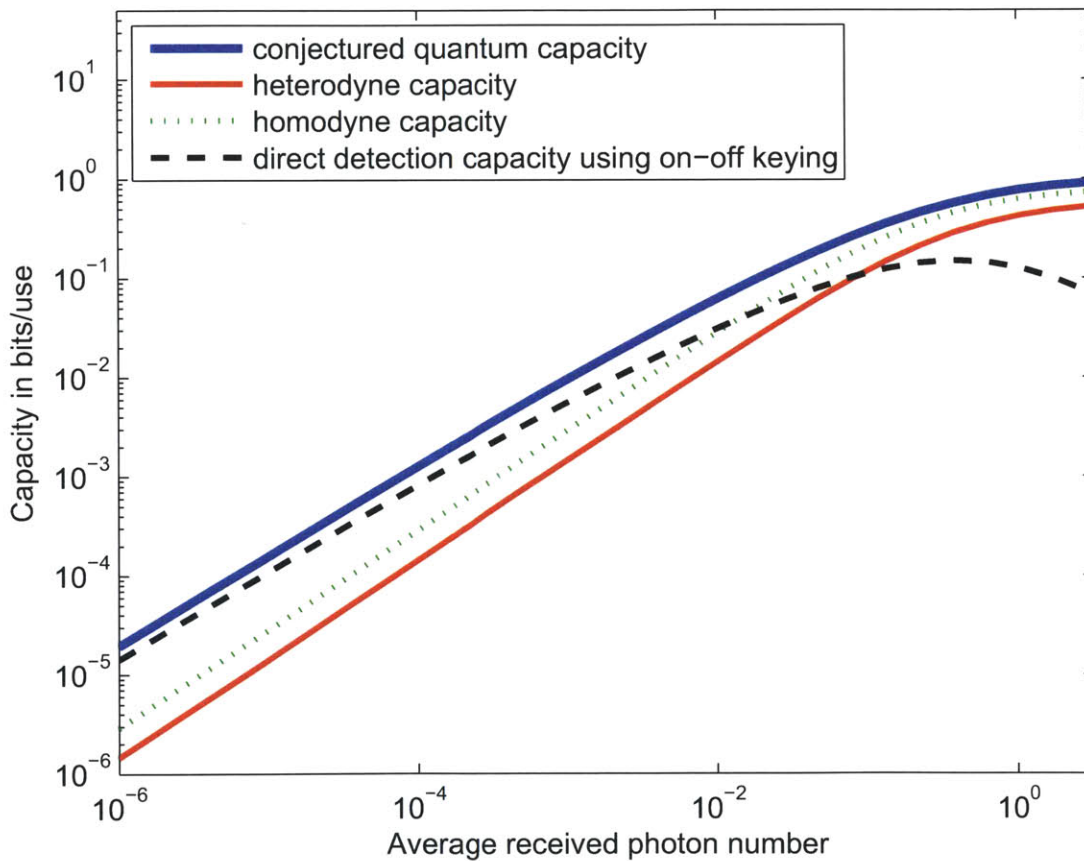
Figure 2-2: Capacity plots for various bosonic channels when the average received signal photon number equals the average received noise photon number ($\bar{N} = N$). Shown are plots of the conjectured quantum capacity from (2.7), and the heterodyne capacity from (2.8), and the homodyne capacity from (2.9). Also included is the direct detection capacity when on-off keying is used, which is calculated numerically by optimizing (2.14) over p.
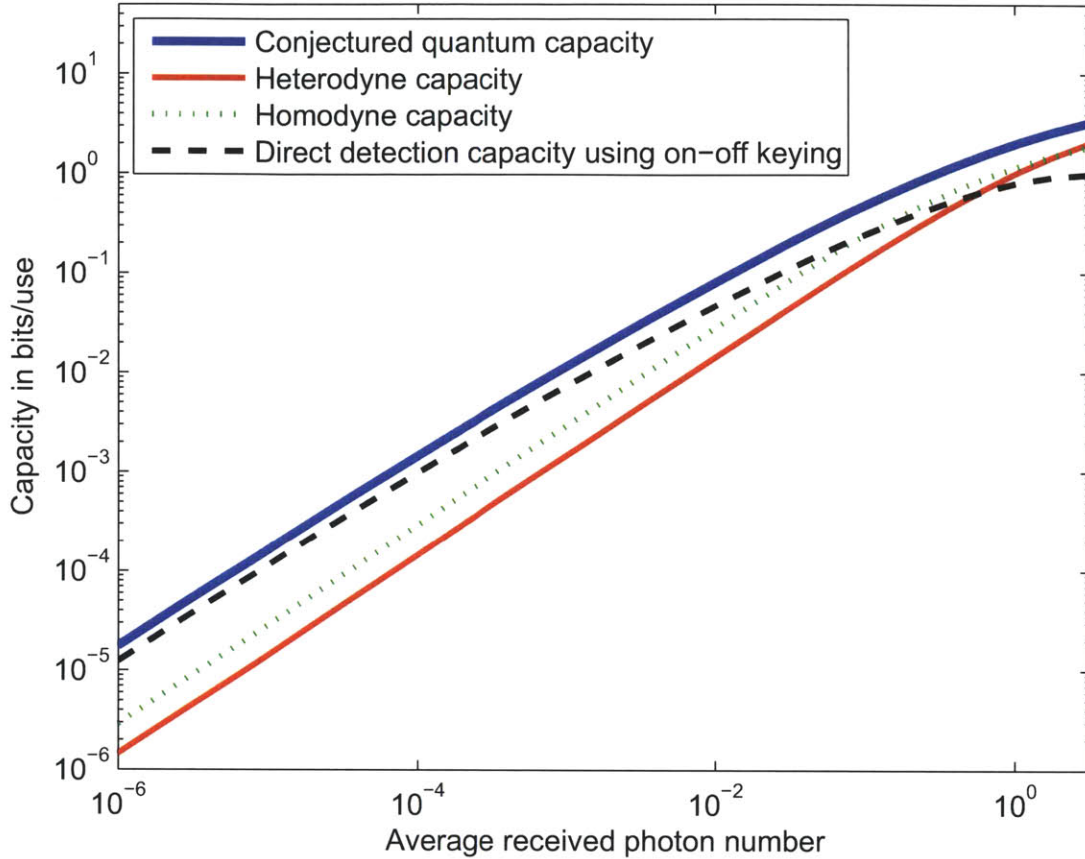
Figure 2-3: Plot of capacity of various bosonic channels when average received noise photon number is equal to 10^{-6}

2.2 Photon and Spectral Efficiencies

In this section we will discuss the achievable photon and spectral efficiencies for heterodyne, homodyne, and on-off keying direct detection, and compare them to the theoretically achievable photon and spectral efficiencies for optimal quantum detection in the case of the pure-loss channel. Furthermore, assuming the thermal-noise channel capacity conjecture is true, we will present curves for the photon and spectral efficiencies for an optimal detection thermal-noise channel, and compare that to a channel that uses direct detection on-off

keying. We will also demonstrate that heterodyne and homodyne detection cannot achieve photon efficiencies greater than 1 nat/photon and 2 nats/photon, respectively.

For every method of quantum detection and average received signal photon number, there is an associated capacity. In the single-mode case, each capacity has a photon efficiency:

$$\text{PE} = \frac{C\left(\bar{N}, N\right)}{\bar{N}}, \tag{2.29}$$

measured in bits/photon or nats/photon depending on the logarithm base chosen for evaluating $C\left(\bar{N}, N\right)$, and a spectral efficiency

$$\text{SE} = C\left(\bar{N}, N\right), \tag{2.30}$$

measured in bits/s/Hz, or nats/sec/Hz, depending on the logarithm base.

We will soon see that single-mode operation is incapable of achieving high photon efficiency, $\text{PE} \gg 1$ and high photon efficiency, $\text{SE} \gg 1$. If both are desired, there is a way to accomplish that goal when multiple spatial modes are available. Suppose that the transmitter can employ M spatial modes, whose annihilation operators are $\hat{a}_m$, for $1 \leq m \leq M$, and that each of these modes couples to the receiver by a beam splitter relation of the form (2.1), i.e.

$$\hat{a}'_m = \sqrt{\eta}\hat{a}_m + \sqrt{1 - \eta}\hat{b}_m, \tag{2.31}$$

where the noise modes $\hat{b}_m$ are in independent thermal states with average photon number N_b. As we did for the single-mode case, we shall use $\bar{N}$ and N for the average received signal and noise photon numbers, respectively, but now $\bar{N}$ represents the total over all M spatial modes, while N applies to each mode individually.

By symmetry, the maximum efficiencies are attained when the transmitted photon num-

36

ber is split evenly between all M modes of the quantum channel, resulting in

$$\text{PE} = M \frac{C\left(\frac{\bar{N}}{M}, N\right)}{\bar{N}} \tag{2.32}$$

$$\text{SE} = MC\left(\frac{\bar{N}}{M}, N\right), \tag{2.33}$$

where C denotes single-mode capacity.

We will now show that heterodyne and homodyne detection, in the case of pure-loss and thermal-noise channels, can never reach photon efficiencies above 1 nat/photon and 2 nats/photon, respectively. The homodyne and heterodyne capacities for the single-mode pure-loss channel are

$$C_{het}\left(\bar{N}\right) = \log\left(1 + \bar{N}\right) \tag{2.34}$$

$$C_{hom}\left(\bar{N}\right) = \frac{1}{2}\log\left(1 + 4\bar{N}\right). \tag{2.35}$$

Because the addition of random noise can only decrease capacity, we know that the photon efficiency achieved by heterodyne and homodyne detection on the thermal-noise channel will not exceed what these detection methods achieve for this metric on the pure-loss channel. For pure-loss we have

$$\text{P}E_{het} = \frac{MC_{het}\left(\frac{\bar{N}}{M}, 0\right)}{\bar{N}} \tag{2.36}$$

and

$$\text{PE}_{hom} = \frac{MC_{hom}\left(\frac{\bar{N}}{M}, 0\right)}{\bar{N}}. \tag{2.37}$$

The associated M spatial-mode spectral efficiencies are

$$\text{SE}_{het} = MC_{het}\left(\frac{\bar{N}}{M}, 0\right) \tag{2.38}$$

37

and

$$\text{SE}_{hom} = MC_{hom}\left(\frac{\bar{N}}{M}, 0\right). \tag{2.39}$$

If we choose a particular value for the spectral efficiency, for either heterodyne or homodyne detection, we can solve for $\bar{N}$ as a function of SE and M. We get

$$\bar{N} = M\left(e^{\frac{\text{SE}}{M}} - 1\right) \tag{2.40}$$

for heterodyne detection and

$$\bar{N} = \frac{M}{4}\left(e^{\frac{2\text{SE}}{M}} - 1\right) \tag{2.41}$$

for homodyne detection. Substituting these results into the PE expressions then yields:

$$\text{PE}_{het} = \frac{\text{SE}}{M\left(e^{\frac{\text{SE}}{M}} - 1\right)} \tag{2.42}$$

and for homodyne

$$\text{PE}_{hom} = \frac{\text{SE}}{\frac{M}{4}\left(e^{\frac{2\text{SE}}{M}} - 1\right)}. \tag{2.43}$$

With $x = \frac{\text{SE}}{M}$, we have

$$\text{PE}_{het} = \frac{x}{e^x - 1} \tag{2.44}$$

which is monotonically decreasing with increasing x, and approaches 1 as $x \to 0$. Likewise with $y = \frac{2\text{SE}}{M}$, we have

$$\text{PE}_{hom} = \frac{2y}{e^y - 1}, \tag{2.45}$$

which is monotically decreasing with increasing y, and approaches 2 as $y \to 0$. Thus, $\mathrm{PE}_{het} \leq$ 1 nat/photon and $\mathrm{PE}_{hom} \leq 2$ nats/photon, with equality in both cases being approached as $\bar{N} \to 0$, where we have made use of the fact that $\frac{\mathrm{SE}}{M} \to 0$ on the pure-loss channel is equivalent to $\bar{N} \to 0$. Because of these limits, we will not plot homodyne and heterodyne photon and spectral efficiency curves in what follows.

Figure 2-4 plots PE and SE for the pure-loss and thermal-noise ($N = 10^{-6}$) channels with $M = 1$, 10, 100, and 1000. In all cases we include the optimum-detection quantum capacity (conjectured capacity in the thermal-noise channel) and the on-off keying direct detection capacity lower bound. From this figure we see that for a given spectral efficiency, to achieve the same photon efficiency as is possible based on the thermal-noise channel lower bound, a factor of over 10 times more spatial modes may have to be used. This is a situation in which there is a substantial gap between direct detection and the conjectured quantum capacity, and where there might be some room for improvement. In particular, it may be possible to achieve a target spectral efficiency and photon efficiency using some optimal detection technique more easily than using direct detection.
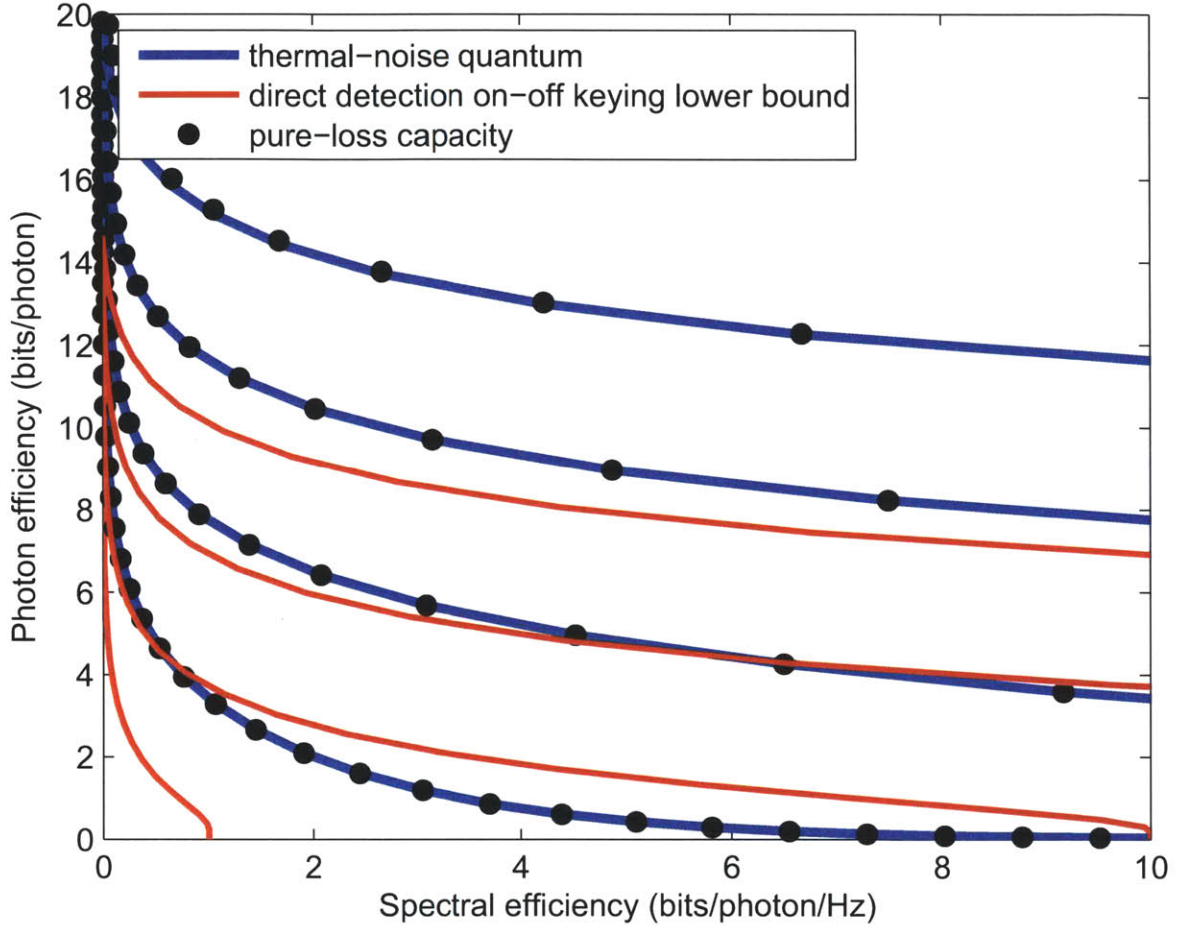
Figure 2-4: Photon efficiency versus spectral efficiency for a pure-loss ($N = 0$) and thermal-noise ($N = 10^{-6}$) channels with $M = 1$, 10, 100, and 1000 spatial modes. The lowest number of spatial modes ($M = 1$) corresponds to the lowest curves on the graph, and they increase sequentially until the $M = 1000$ highest curves on the graph. Quantum capacity (pure-loss) and the conjectured quantum capacity (thermal-noise), are compared with lower bounds for the on-off keying direct detection capacities. Also included is a plot of the photon efficiency versus spectral efficiency for the case of $N = 0$. We observe that the gap between the curves for the conjectured quantum capacity and this upper bound is not large in this region of spectral efficiency. We note that the number of spatial modes required to achieve a particular photon efficiency in the case of direct detection is much greater than in the conjectured quantum capacity case.

Chapter 3

Ergodic Capacity and Outage Capacity

The beam splitter channel models — pure-loss and thermal-noise — whose capacities we addressed in Chapter 2 represent idealizations that could be applied as first approximations to vacuum propagation and fiber-optic propagation. However, for line-of-sight propagation through the atmosphere in clear-weather conditions, they are insufficient because they fail to capture the severe time-dependent fading that is due to refractive index turbulence [7]. Fading environments have long been studied — in the classical domain — for wireless communication at microwave frequencies and with semiclassical photodetection models for optical communication through atmospheric turbulence [14]. Our main purpose, in this chapter, is to extend prior work on fading-channel capacities — specifically the ergodic and outage capacities — to quantum models for both far-field and near-field propagation through turbulence.

In Chapter 1 we reviewed the theory of optical propagation through turbulence. Because we are interested in high data-rate communication — say Gbps — and because turbulence multipath spread is on the order of psec and its coherence time is on the order of msec, it is appropriate to model a single channel use between a transmitter employing a fixed spatial

pattern and a receiver extracting a single spatial mode as a beam-splitter model

$$\hat{a}' = \sqrt{\eta}e^{i\theta}\hat{a} + \sqrt{1-\eta}\hat{b} \tag{3.1}$$

where, as in Chapter 2, $\hat{a}$, $\hat{b}$, and $\hat{a}'$ are modal photon annihilation operators for the input mode, the mode injected by the channel, and the output mode, respectively, $0 \leq \eta \leq 1$ is the channel's transmissivity, and θ is the channel's phase shift. Now, unlike Chapter 2, η is a random variable, and one that has very strong statistical dependence between different channel uses. In Section 3.1 we address the ergodic capacity for this fading beam-splitter channel, and in Section 3.4 we consider its outage capacity. In both cases we will only treat single-mode operation. The extension to multi-mode operation will be given in Chapter 4.

3.1 Ergodic Capacity

We assume that the channel state, i.e., the transmissivity η and the phase θ, are known for each channel use by both the transmitter and receiver. Because as many as 10^6 channel uses are achievable for coding within a single channel coherence time, viz., while η is fixed, the transmitter and receiver are able to achieve capacity

$$C\left(\eta\bar{N}_T, (1-\eta)N_B\right) \tag{3.2}$$

where $\bar{N}_T = \left\langle\hat{a}^\dagger\hat{a}\right\rangle$ is the average number of transmitted signal photons per channel use, $\bar{N}_B = \left\langle\hat{b}^\dagger\hat{b}\right\rangle$ is the average number of noise photons entering the channel, and $C\left(\eta\bar{N}_T, (1-\eta)N_B\right)$ is the channel capacity of the thermal-noise channel from Chapter 2. With $p(\eta)$ being the probability density for the fading channel transmissivity, we have that

$$C_{ergodic} = \int\limits_0^1 d\eta\, p(\eta) C\left(\eta\bar{N}_T, (1-\eta)N_B\right) \tag{3.3}$$

is the channel's ergodic capacity. Note that this formulation encompasses both conventional receivers — by using the heterodyne, homodyne, or direct detection results from Chapter 2 for $C\left(\eta\bar{N}_T, (1-\eta)N_B\right)$ — as well as the ultimate quantum form of the ergodic capacity, which follows from the quantum results in Chapter 2 for $C\left(\eta\bar{N}_T, (1-\eta)N_B\right)$.

We can see that the ergodic capacity depends on the probability distribution for the value of η. In the far field, there are a number of models for this distribution. For practical purposes, it is convenient to suppose that the channel stays stable on the order of msec, while a practical implementation of a high rate optical channel will transmit at rates on the order of GHz, so the ergodic capacity could be practically approached by dividing time into discrete msec intervals, measuring the channel, and then transmitting for that time at a rate close to the instantaneous capacity of that channel. We note that this requires channel knowledge, but if the channel is slowly varying in time, this may be practical to achieve. In this case, during each milliseconds-long time interval, the value of η is a random variable. We now need to know what the probability distribution of η is during each of these time intervals and then the ergodic capacity can be computed for various channels.

We will consider ergodic capacities in both the far field and the near field, i.e., when the average power transferred is a small fraction of what was sent in the case of far field, or when the power transferred is nearly all of what was sent in the case of the near field. In the next section we proceed to develop two far-field models — exponential and lognormal fading models — which correspond to earth-to-space communication and space-to-earth communication, respectively. These will be used in Section 3.3 to evaluate the ergodic capacities for those communication scenarios. Later, after we introduce the notion of outage capacity, we will develop tight bounds on the ergodic capacity of near-field operation.

3.2 Fading Models

Figure 3-1 shows a diagram of a bidirectional earth-space channel. For uplink communication, the transmitter is on the ground and its output is emitted from a diameter-D_0 exit pupil A_0. The uplink receiver is in a synchronous orbit and collects light through a
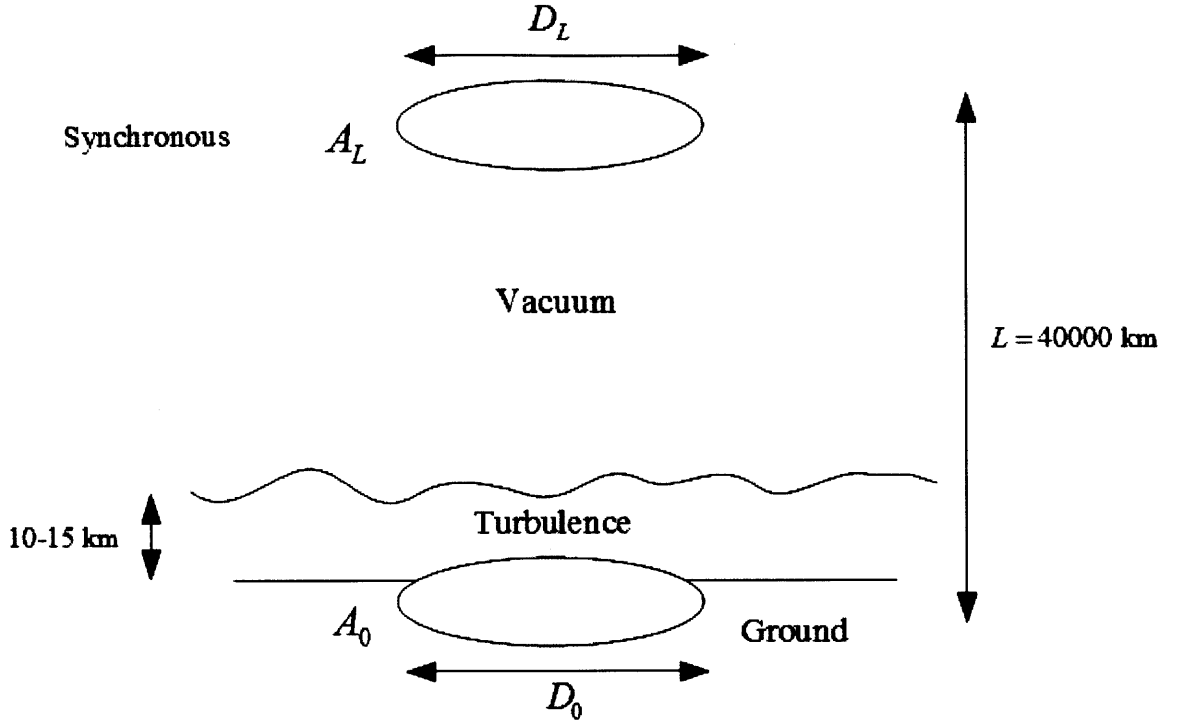
Figure 3-1: Diagram of earth to space communication geometry.

diameter-D_L entrance pupil A_L. The turbulence is all contained within a height of 10 to 15 km above ground near the transmitter, i.e., in the troposphere. In applying the extended Huygens-Fresnel principle to this setup, we have that $D_L \ll$ turbulence coherence length at the receiver and and that $D_0 \gg$ turbulence coherence length at the transmitter, whence

$$E_L\left(\vec{\rho}',t\right) = \int\limits_{A_0} d\vec{\rho} E_0\left(\vec{\rho},t-\frac{L}{c}\right) \frac{e^{ikL+i\frac{k|\vec{\rho}'-\vec{\rho}|^2}{2L}}}{i\lambda L} e^{\chi(\vec{0},\vec{\rho})+i\phi(\vec{0},\vec{\rho})}, \tag{3.4}$$

where $\vec{\rho}' = \vec{0}$ is the center of the A_L pupil. When $\frac{D_0^2}{\lambda L} \ll 1$ and $\frac{D_L^2}{\lambda L} \ll 1$, as will be the case for practical pupil diameters, then the extended Huygens-Fresnel principle can be approximated

as

$$E_L\left(\vec{\rho'},t\right) = \frac{e^{ikL}}{i\lambda L}\int\limits_{A_0} d\vec{\rho}\, E_0\left(\vec{\rho},t-\frac{L}{c}\right)e^{\chi(0,\vec{\rho})+i\phi(0,\vec{\rho})}. \tag{3.5}$$

We shall assume a collimated-beam transmitter

$$E_0\left(\vec{\rho},t\right) = \sqrt{\frac{4N_T}{\pi D_0{}^2}}s\left(t\right) \tag{3.6}$$

which achieves optimum ground-to-space power transfer in the absence of turbulence, where N_T is the average number of transmitted photon and $s\left(t\right)$ is a normalized modulation obeying $\int \left|s\left(t\right)\right|^2 dt = 1$. The extended Huygens-Fresnel Principle now yields

$$E_L\left(\vec{\rho'},t\right) = \frac{e^{ikL}}{i\lambda L}s\left(t-\frac{L}{c}\right)\sqrt{\frac{4N_T}{\pi D_0{}^2}}\int\limits_{A_0} d\vec{\rho}\, e^{\chi(0,\vec{\rho})+i\phi(0,\vec{\rho})}. \tag{3.7}$$

Decomposing the A_0 integral into statistically-independent coherence areas gives us

$$\int\limits_{A_0} d\vec{\rho}\, e^{\chi(0,\vec{\rho})+i\phi(0,\vec{\rho})} \approx \sum_q A_{coh}e^{\chi_q+i\phi_q} \tag{3.8}$$

where the $\{\chi_q\}$ and $\{\phi_q\}$ are the logamplitude and phase fluctuations for these coherence areas. The central limit theorem now implies that this summation has a zero-mean complex-Gaussian probability distribution because $\left\langle e^{\chi_n+i\phi_n}\right\rangle \approx 0$.

Equation (3.7) shows that the field received over A_L is a collimated beam, i.e., a single spatial mode. Thus if we extract the $s\left(t-\frac{L}{c}\right)$ temporal mode from this collimated beam spatial mode via

$$a' \equiv \int dt\, s^*\left(t-\frac{L}{c}\right)\int\limits_{A_L} d\vec{\rho'}\sqrt{\frac{4}{\pi D_L{}^2}}E_L\left(\vec{\rho'},t\right) \tag{3.9}$$

we get

$$a' = \sqrt{\eta}e^{i\theta}a \tag{3.10}$$

where

$$a \equiv \int dt s^*(t) \int_{A_0} d\vec{\rho} \sqrt{\frac{4}{\pi D_0{}^2}} E_L(\vec{\rho}, t) \tag{3.11}$$

is the corresponding spatiotemporal input mode and

$$\sqrt{\eta}e^{i\theta} \equiv \left(\frac{\pi D_0 D_L e^{ikL}}{i4\lambda L}\right) \frac{1}{Q} \sum_{q=1}^{Q} e^{\chi_q + i\phi_q} \tag{3.12}$$

where we have used

$$Q A_{coh} = \frac{\pi D_0{}^2}{4}. \tag{3.13}$$

Thus, we see that for $Q \gg 1$, we get $\sqrt{\eta}e^{i\theta}$ to be a complex-valued Gaussian random variable, which implies that η is exponentially distributed. We also see that

$$\langle \eta \rangle = \left(\frac{\pi D_0 D_L}{4\lambda L}\right)^2 \frac{1}{Q^2} \sum_{q=1}^{Q} \langle e^{2\chi_q} \rangle = \left(\frac{\pi D_T D_L}{4\lambda L}\right)^2 \ll 1 \tag{3.14}$$

where the first equality follows from our assumption that the fluctuations incurred on different coherence areas are statistically independent and the second equality follows from χ_q being Gaussian distributed with a mean value equal to minus its variance, and the definition $D_T = D_0/\sqrt{Q}$ of the turbulence-limited diameter for diffraction-limited propagation over the ground-to-space path. The final inequality is a consequence of $\frac{D_0{}^2}{\lambda L} \ll 1$, $\frac{D_L{}^2}{\lambda L} \ll 1$, and $Q \gg 1$. Physically, $\langle \eta \rangle \ll 1$ represents far-field propagation, i.e., only a very small fraction of the transmitted photons reach the receiver over the uplink to synchronous orbit. Note that $Q \gg 1$ implies that $\langle \eta \rangle \ll \left(\frac{\pi D_0 D_L}{4\lambda L}\right)^2$ which is the result that applies in the absence of

turbulence.

The preceding analysis of the uplink is entirely classical, although we have chosen to measure energy in units of photons at the operating wavelength. Because we will employ coherent-state encoding, we can take the beam-splitter model with exponential fading for the annihilation operator's input-output relation to be

$$\hat{a}' = \sqrt{\eta}e^{i\theta}\hat{a} + \sqrt{1 - \eta}\hat{b}. \tag{3.15}$$

Here, θ and η will be statistically independent with θ uniformly distributed on $[0, 2\pi]$ and η exponentially distributed with mean $\left(\frac{\pi D_T D_L}{4\lambda L}\right)^2 \ll 1$. Strictly speaking, we cannot use this model when $\eta > 1$, but the probability of $\eta > 1$ occurring is extremely small so, as we will see later, this restriction will not pose any problem.

Now let us consider a system in which the transmitter is in space and the receiver is on the ground. With the same propagation assumptions that were made for the uplink, we know that

$$E_L(\vec{\rho}', t) = \sqrt{\frac{4N_T}{\pi D_L{}^2}}s(t) \tag{3.16}$$

achieves optimum space-to-ground power transfer whether or not turbulence is present. The extended Huygens-Fresnel principle now gives us

$$E_0(\vec{\rho}, t) = \frac{e^{ikL}}{i\lambda L}\sqrt{\frac{\pi N_T D_L{}^2}{4}}s\left(t - \frac{L}{c}\right)e^{\chi(\vec{0}, \vec{\rho}) + i\phi(\vec{0}, \vec{\rho})}. \tag{3.17}$$

If we collect plane-wave spatial modes over each of the Q coherence areas in A_0, and extract the $s(t - L/c)$ temporal mode what results is

$$a_q{}' = \int dt s^*\left(t - \frac{L}{c}\right)\int_{A_q} d\vec{\rho}\sqrt{\frac{4}{\pi D_T{}^2}}E_0(\vec{\rho}, t) \tag{3.18}$$

$$= \left(\frac{\pi D_T D_L e^{ikL}}{i4\lambda L} \right) e^{\chi_q + i\phi_q} a \tag{3.19}$$

for $1 \leqslant q \leqslant Q$ where

$$a = \int dt s^*(t) \int_{A_L} d\vec{\rho} \sqrt{\frac{4}{\pi D_L^2}} E_0(\vec{\rho}, t). \tag{3.20}$$

Quantizing this classical relation yields the beam-splitter model

$$\hat{a}'_q = \sqrt{\eta_q} e^{i\theta_q} \hat{a} + \sqrt{1 - \eta_q} \hat{b}_q \tag{3.21}$$

for $1 \leqslant q \leqslant Q$, where $\{\eta_q, \theta_q\}$ is a set of independent identically distributed random variables, with η_q being lognormally distributed with mean equal to minus its variance and θ_q being uniformly distributed on $[0, 2\pi]$. Here we see that each coherence area in the ground receiver's entrance pupil has

$$\langle \eta_q \rangle = \left(\frac{\pi D_T D_L}{4\lambda L} \right)^2 \tag{3.22}$$

fractional energy transfer, so that the total average energy transfer is

$$Q \langle \eta_q \rangle = \left(\frac{\pi D_0 D_L}{4\lambda L} \right)^2 \tag{3.23}$$

which matches what is achieved in the absence of turbulence.

3.3   Computation of Ergodic Capacity

In Fig. 3-2 we plot the ergodic capacities of the pure-loss heterodyne, homodyne, direct detection on-off keying channels, as well as the optimal quantum detection channel for the beam-splitter model in which η is exponentially distributed with mean value $\langle \eta \rangle = 0.005$. In other words, plotted are the capacities given by:

$$C_{ergodic} = \int_0^1 d\eta\, p(\eta) C\left(\eta \bar{N}_T, (1-\eta) N_B\right) \tag{3.24}$$

where $p(\eta)$ is

$$p(\eta) = \frac{\frac{1}{\langle\eta\rangle} e^{\frac{-\eta}{\langle\eta\rangle}}}{\int_0^1 \frac{1}{\langle\eta\rangle} e^{\frac{-\eta}{\langle\eta\rangle}} d\eta} = \frac{\frac{1}{\langle\eta\rangle} e^{\frac{-\eta}{\langle\eta\rangle}}}{1 - e^{\frac{-1}{\langle\eta\rangle}}}, \quad \text{for } 0 \leqslant \eta \leqslant 1,\ 0 \text{ otherwise} \tag{3.25}$$

which is a standard exponential distribution that has been truncated at $\eta = 1$. In the end, this truncation does not appreciably change the value of the integral because $\langle\eta\rangle$ is so small. Exact expressions for this capacity cannot be obtained, but they have been numerically evaluated. We note that the ergodic capacity is very close to the upper bound on any ergodic capacity when evaluated at constant $\eta = \langle\eta\rangle$, indicating that although random fading hurts the channel capacity, it does not hurt very much.

In Fig. 3-3, we plot the ergodic capacity where $p(\eta)$ is taken to be a lognormal distribution, i.e., when we consider space-to-ground propagation with a single coherence area receiver on the ground. The lognormal distribution is given in terms of parameters μ and σ^2 as follows:

$$p(\eta) = \frac{\frac{1}{\eta\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln\eta - \mu)^2}{2\sigma^2}}}{\int_0^1 \frac{1}{\eta\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln\eta - \mu)^2}{2\sigma^2}} d\eta} = \frac{\frac{1}{\eta\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln\eta - \mu)^2}{2\sigma^2}}}{\frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{-\mu}{\sqrt{2\sigma^2}}\right)}, \quad \text{for } 0 \leqslant \eta \leqslant 1,\ 0 \text{ otherwise} \tag{3.26}$$

where again the distribution is truncated. Once again this truncation is insignificant for very small $\langle\eta\rangle$. In this case, the value of $\langle\eta\rangle$ can be given as:

$$\langle\eta\rangle = e^{\mu + \frac{\sigma^2}{2}}. \tag{3.27}$$

As we saw in the case of the exponential distribution, with the lognormal distribution the
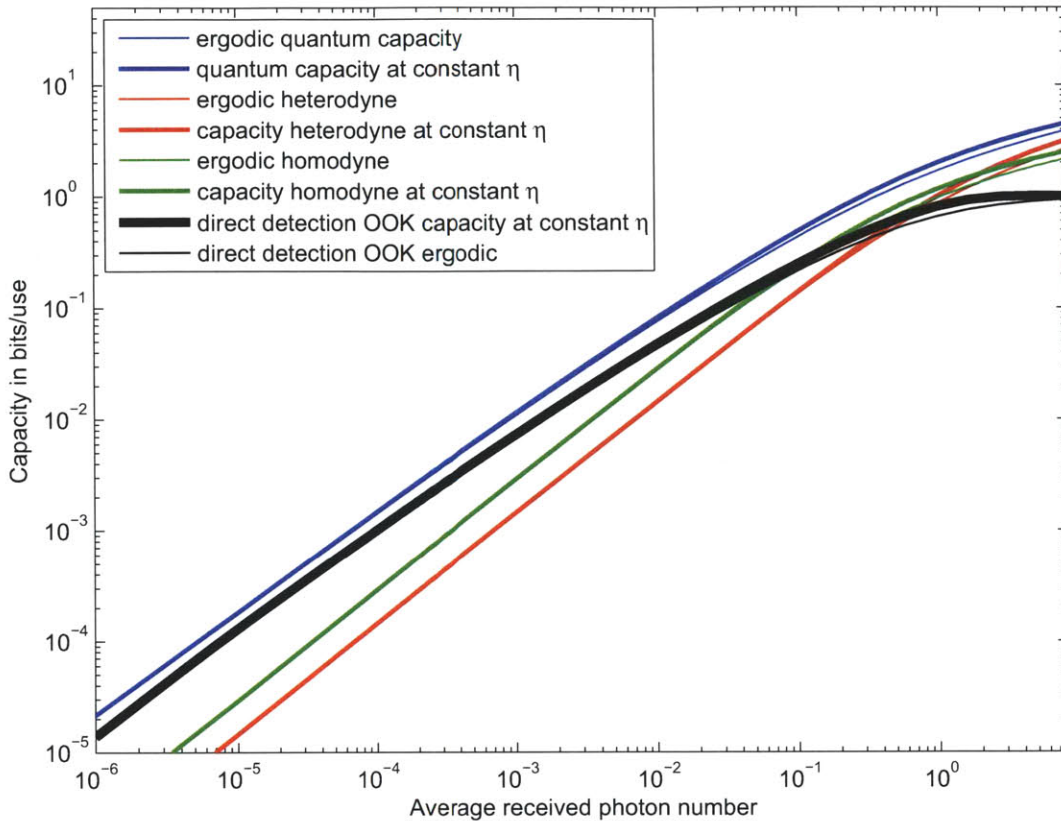
Figure 3-2: Ergodic capacity plots for various bosonic pure-loss channels as a function of average received photon number in the case of exponential fading with $\langle \eta \rangle = 0.005$, compared to the capacities when the channel transmissivity is always constant $\eta = 0.005$. We show in (3.42) that constant-η capacities are upper bounds on the corresponding ergodic capacities. Notably, the ergodic capacities are very close to their upper bounds for $\eta = 0.005$.

capacity is not significantly affected by random fading when $\langle \eta \rangle \ll 1$

3.4   Outage Capacity

Although the ergodic capacity of a channel is of interest to us, achieving this capacity may be difficult because it requires that channel knowledge be available to both the transmitter and receiver, and it implies that a continuum of different transmitting rates be used for a continuum of channel states. Far easier to implement is a transmitting structure that
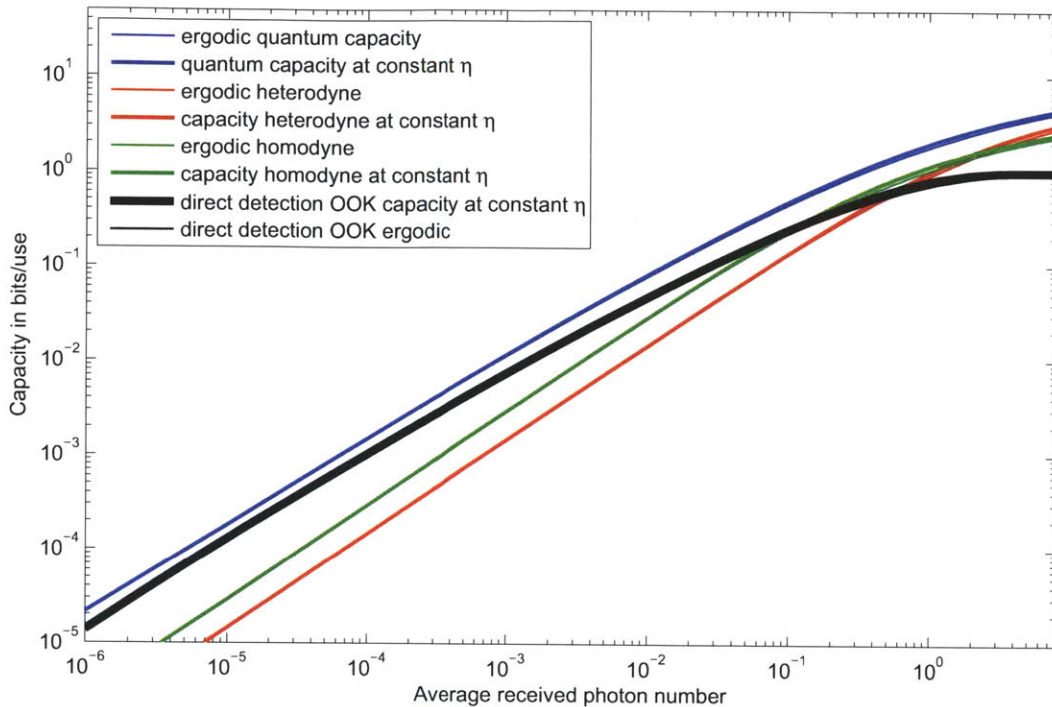
Figure 3-3: Ergodic capacities for various pure-loss bosonic channels as a function of average received photon number in the case of lognormal fading with parameter $\langle \eta \rangle = 0.01$ and $\sigma^2 = 0.5$, compared to the capacities when the channel transmissivity is a constant at $\eta = 0.01$. We show in (3.42) that constant-η capacties are upper bounds on the corresponding ergodic capacities. Notably, the ergodic capacities are very close to theor upper bounds for $\eta = 0.01$.

transmits at one rate if the channel is in a state that can support that rate, and when the channel is in a poor state, the transmitter does not transmit at all. The capacity of a fading channel that can be in a transmitting state a fraction p_t of the time is called the outage capacity. We note that over very long periods of time, the average rate of transmission for this type of channel is

$$R\left(p_t\right) = p_t C_t\left(p_t\right) \tag{3.28}$$

where $C_t\left(p_t\right)$, the outage capacity, is defined as the maximum rate that can be reliably trans-mitted for at least a fraction p_t of the time. For the pure-loss and thermal-noise channels, we

can further calculate the outage capacity as a function of the probability that η is above a particular value. We first define $\eta_{\max}$ as the maximum value that the channel transmissivity will equal or exceed with probability p_t or greater. In other words, $\eta_{\max}$ satisfies

$$\mathrm{Pr}\left(\eta \geqslant \eta_{\max}\right) \geqslant p_t. \tag{3.29}$$

Using this definition, and assuming that the receiver knows the channel phase θ, we can show that the outage capacity as a function of p_t is

$$C_t\left(p_t\right) = C\left(\eta_{\max}\bar{N}_T, \left(1 - \eta_{\max}\right)N_B\right) \tag{3.30}$$

for the thermal-noise channel, where, as in (3.2), $\bar{N}_T = \left\langle \hat{a}^\dagger \hat{a}\right\rangle$ is the average number of transmitted signal photons entering the channel, $N_B = \left\langle \hat{b}^\dagger \hat{b}\right\rangle$ is the average number of background photons entering the channel, and $C\left(\eta_{\max}\bar{N}_T, \left(1 - \eta_{\max}\right)N_B\right)$ is the thermal-noise capacity from Chapter 2.

Over long periods of time, the average rate at which information may be reliably communicated for a given p_t is therefore

$$R\left(p_t\right) = p_t C\left(\eta_{\max}\bar{N}_T, \left(1 - \eta_{\max}\right)N_B\right). \tag{3.31}$$

In the remainder of this section we shall use the dependence of η_{max} on p_t for the exponential and lognormal fading models of far-field propagation to maximize $R\left(p_t\right)$ as a function of p_t in the case of the pure-loss channel, for which

$$R\left(p_t\right) = p_t C\left(\eta_{\max}\bar{N}_T, 0\right) = p_t g\left(\eta_{\max}\bar{N}_T\right). \tag{3.32}$$

3.4.1 Outage Capacity for the Exponential-Fading Channel

Here we shall presume that η follows the truncated exponential distribution $p\left(\eta\right)$ given in Eq. (3.25). As explained in Section 3.2, this distribution models the fading statistics encountered

52

in ground-to-space communication. It is now easy to find $\eta_{\max}$ as a function of p_t. We have that

$$\int\limits_{\eta_{\max}}^{1} p\left(\eta\right) d\eta = p_t, \tag{3.33}$$

from which straightforward integration yields

$$\frac{e^{\frac{-\eta_{\max}}{\langle\eta\rangle}} - e^{\frac{-1}{\langle\eta\rangle}}}{\left(1 - e^{\frac{-1}{\langle\eta\rangle}}\right)} = p_t \tag{3.34}$$

and hence

$$\eta_{\max} = -\langle\eta\rangle \ln\left(e^{\frac{-1}{\langle\eta\rangle}} + \left(1 - e^{\frac{-1}{\langle\eta\rangle}}\right) p_t\right). \tag{3.35}$$

For $\langle\eta\rangle \ll 1$, which will be the case deep into the far field, we can safely neglect the truncation in (3.25), so

$$\eta_{\max} \approx -\langle\eta\rangle \ln\left(p_t\right). \tag{3.36}$$

Equipped with our expression for $\eta_{\max}$ as a function of p_t, we can now maximize

$$R\left(p_t\right) = p_t g\left(-\langle\eta\rangle \ln\left(e^{\frac{-1}{\langle\eta\rangle}} + \left(1 - e^{\frac{-1}{\langle\eta\rangle}}\right) p_t\right) \bar{N}_T\right) \approx p_t g\left(-\langle\eta\rangle \ln\left(p_t\right) \bar{N}_T\right) \tag{3.37}$$

for the pure-loss channel with exponential fading. In Fig. 3-4, we have plotted $\max\limits_{p_t} R\left(p_t\right)$ versus $\bar{N}_T$ for the pure-loss channel and several values of $\langle\eta\rangle$. For comparison, we also plot the corresponding values of $g\left(\langle\eta\rangle \bar{N}_T\right)$, the capacity when there is no fading, which is an upper bound on $R\left(p_t\right)$, as we now show. Let p_t^* be the value that maximizes $R\left(p_t\right)$, and let

η^* be the associated transmissivity value, i.e.,

$$p_t{}^* = \int\limits_{\eta^*}^{1} p\left(\eta\right)d\eta \tag{3.38}$$

We have that

$$\max_{p_t} R\left(p_t\right) = R\left(p_t{}^*\right) = p_t{}^* g\left(\eta^* \bar{N}_T\right) \tag{3.39}$$

$$\leqslant \int\limits_{\eta^*}^{1} g\left(\eta \bar{N}_T\right)p\left(\eta\right)d\eta \tag{3.40}$$

$$\leqslant \int\limits_{0}^{1} g\left(\eta \bar{N}_T\right)p\left(\eta\right)d\eta \tag{3.41}$$

$$\leqslant g\left(\langle\eta\rangle\,\bar{N}_T\right) = C\left(\langle\eta\rangle\,\bar{N}_T, 0\right) \tag{3.42}$$

where the first inequality follows from $g\left(x\right)$ being a monotonically increasing function of x, the second inequality follows from $g\left(x\right)$ being a non-negative function of $x \geqslant 0$, and the third inequality follows from $g\left(x\right)$ being a concave function of x for $x \geqslant 0$.

Interestingly, the ergodic capacity is also an upper bound on $R\left(p_t{}^*\right)$, as the following calculation shows:

$$C_{ergodic} = \int\limits_{0}^{1} p\left(\eta\right)C\left(\eta \bar{N}_T, \left(1 - \eta\right)N_B\right)d\eta \tag{3.43}$$

$$\geqslant \int\limits_{\eta^*}^{1} p\left(\eta\right)C\left(\eta \bar{N}_T, \left(1 - \eta\right)N_B\right)d\eta \tag{3.44}$$

$$\geqslant C\left(\eta^* \bar{N}_T, \left(1 - \eta^*\right)N_B\right)\int\limits_{\eta^*}^{1} p\left(\eta\right)d\eta \tag{3.45}$$

$$= p_t{}^* C\left(\eta^* \bar{N}_T, \left(1 - \eta^*\right)N_B\right) = R\left(p_t{}^*\right) \tag{3.46}$$

where the second inequality follows from $C\left(\eta \bar{N}_T, \left(1 - \eta\right)N_B\right)$ being a monotonically increas-

ing function of η.

Figure 3-4 shows that exponential fading causes appreciable performance degradation in terms of outage rates, i.e., $R(p_t^*)$ falls significantly below $g\left(\langle\eta\rangle \bar{N}_T\right)$ for the pure-loss channel.
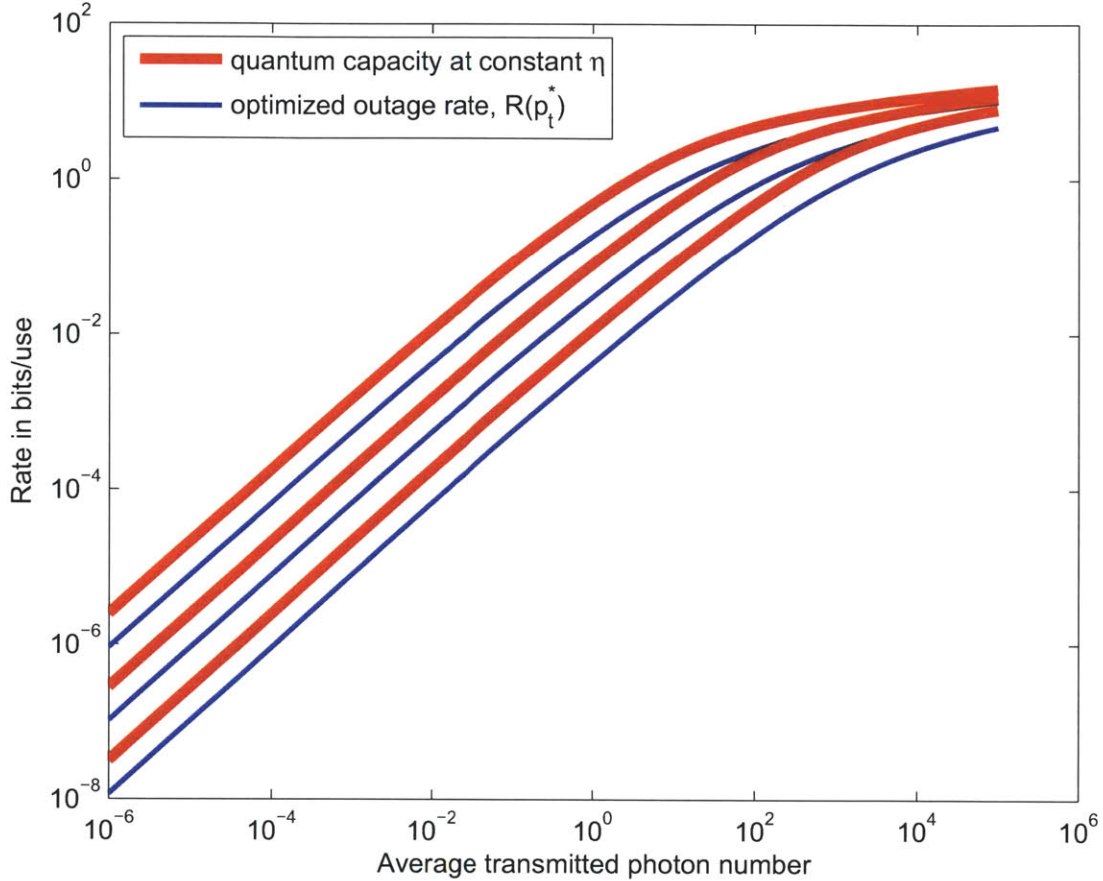

Figure 3-4: This figure shows the optimized outage rate, $R(p_t^*)$, of an exponential fading pure-loss channel as detailed in (3.37). In this case, the parameter $\langle\eta\rangle$ is varied to demonstrate how fading affects the outage capacity average rate. From top to bottom, $\langle\eta\rangle = 0.1$, 0.01 and 0.001. For comparison purposes, also shown are the corresponding capacities without fading, i.e., when $\eta = \langle\eta\rangle$.

We are also interested in the outage capacity. Let us assume that $\langle\eta\rangle \ll 1$ and employ

the untruncated exponential distribution. We then get

$$p_t = \Pr\left(\eta \geqslant \eta_{\max}\right) = e^{\frac{-\eta_{\max}}{\langle\eta\rangle}} \tag{3.47}$$

$$\tag{3.48}$$

so that

$$\eta_{\max} = -\langle\eta\rangle \ln\left(p_t\right). \tag{3.49}$$

Thus, the outage capacity of the pure-loss channel at $\langle\eta\rangle \ll 1$ for the exponential distribution is

$$C_t\left(p_t\right) = g\left(-\langle\eta\rangle \ln\left(p_t\right)\bar{N}_T\right). \tag{3.50}$$

If we want a very high availability, i.e., $p_0 = 1 - p_t \ll 1$ outage probability, we find that the outage capacity is much lower than the non-fading capacity. For example, for $p_0 = 0.05$, we have

$$C_t\left(0.95\right) = g\left(-\langle\eta\rangle \ln\left(0.95\right)\bar{N}_T\right) = g\left(0.0513\langle\eta\rangle \bar{N}_T\right). \tag{3.51}$$

3.4.2 Outage Capacity for the Lognormal-Fading Channel

In this section we shall assume that η follows the truncated lognormal distribution from (3.26), which we showed in Section 3.2 applies to a single coherence-area receiver on a space-to-ground far-field link. Moreover, by invoking $\langle\eta\rangle \ll 1$ we shall neglect the truncation in (3.26) without appreciable loss of accuracy. We thus can calculate $\eta_{\max}$ as a function of p_t as follows:

$$p_t = \Pr\left(\eta \geqslant \eta_{\max}\right) = 1 - \left(\frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\ln\left(\eta_t\right) - \mu}{\sqrt{2\sigma^2}}\right)\right) \tag{3.52}$$

which yields

$$\eta_{\max} = \exp\left[\sqrt{2\sigma^2}\left(\text{erf}^{-1}\left(1 - 2p_t\right) + \mu\right)\right]. \tag{3.53}$$

So, using this expression for $\eta_{\max}$ as a function of p_t, we can now maximize

$$R\left(p_t\right) = p_t g\left(\exp\left(\sqrt{2\sigma^2}\left(\text{erf}\left(1 - 2p_t\right)\right) + \mu\right)\bar{N}_T\right) \tag{3.54}$$

for the pure-loss channel with lognormal fading. In Fig. 3-5 we have plotted $\max_{p_t} R\left(p_t\right)$ versus $\bar{N}_T$ for the pure-loss channel and several values of σ^2 in which $\langle\eta\rangle$ has been kept constant. As we did for the exponontial fading channel, for comparison we also plot the corresponding values of $g\left(\langle\eta\rangle \bar{N}_T\right)$, the capacity when there is no fading. As in the case of exponential fading, we also note that the lognormal fading causes appreciable performance degradation in terms of outage rate. However, variation in the σ^2 parameter has little effect on the optimized outage rate.

3.5   Best Case and Worst Case Statistics Given $\langle\eta\rangle$

The exponential and lognormal models whose ergodic and outage capacities we addressed earlier in this chapter do not apply to near-field propagation. Also, they need not represent all possible fading situations that might be encountered on bosonic channels. Thus in this section we will seek bounds on these capacities that only require knowledge of $\langle\eta\rangle$, the average value of the channel's transmissivity. It turns out that these results will be of value for near-field operation, in which $1 - \langle\eta\rangle \ll 1$. We will begin our development with bounds that are relevant to outage capacity. Suppose we know the value of $\langle\eta\rangle$ but not the probability distribution of η. The best-case and worst case capacities for a given p_t are defined to be

$$C_{best}\left(\langle\eta\rangle, p_t\right) = \max_{p(\eta)} C_t\left(p_t\right) \tag{3.55}$$
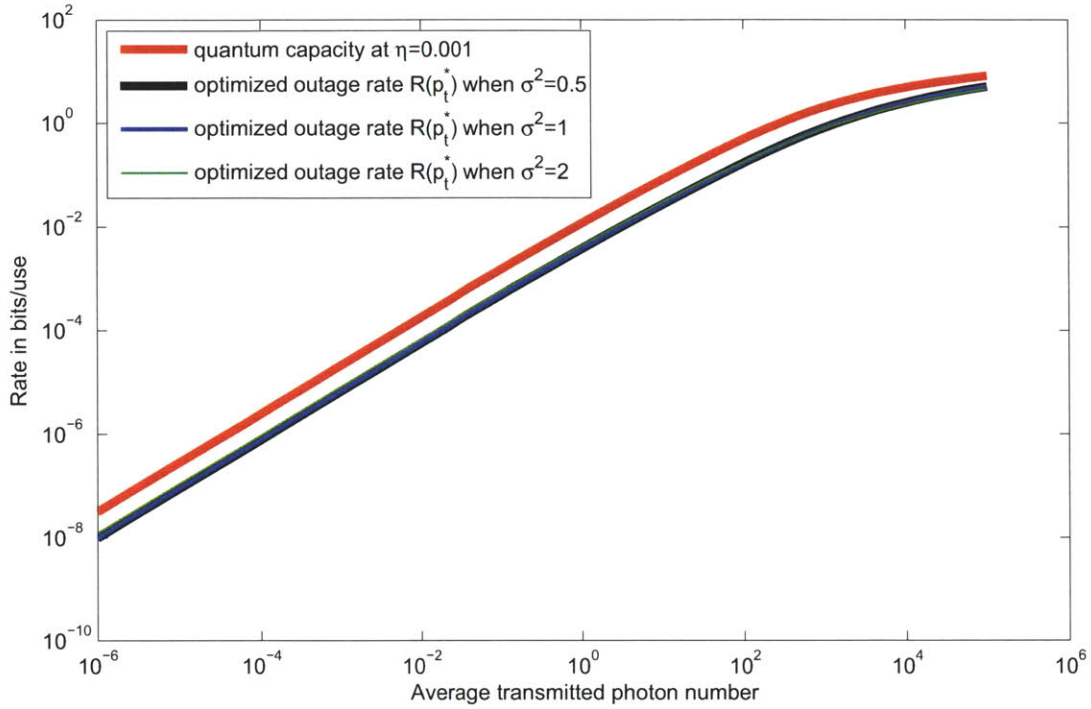
Figure 3-5: Plot of optimized outage rates, $R\left(p_t^*\right)$, for lognormal fading pure-loss channel, when $\langle \eta \rangle = 0.001$. Shown are plots of the optimized outage rate, optimized over p_t, as detailed in (3.54). As the parameter σ^2 is varied, we see little difference in the optimized outage rate.

and

$$C_{worst}\left(\langle \eta \rangle, p_t\right) = \min_{p(\eta)} C_t\left(p_t\right) \tag{3.56}$$

where $p\left(\eta\right)$ is a probability distribution on $0 \leqslant \eta \leqslant 1$ satisfying

$$\langle \eta \rangle = \int_0^1 \eta p\left(\eta\right)d\eta. \tag{3.57}$$

We know that

$$C_t\left(p_t\right) = C\left(\eta_{\max}\bar{N}_T, \left(1 - \eta_{\max}\right) N_B\right) \tag{3.58}$$

for the thermal-noise channel where

$$p_t = \int\limits_{\eta_{\max}}^{1} p(\eta)\, d\eta \qquad (3.59)$$

and $N_B = 0$ gives the pure-loss case. Because $C\left(\eta_{\max}\bar{N}_T, (1 - \eta_{\max}) N_B\right)$ is monotonically increasing with increasing $\eta_{\max}$, the distributions, $p_{best}(\eta)$ and $p_{worst}(\eta)$, that give the upper and lower bounds on $C_t(p_t)$ are found by choosing $p_{best}(\eta)$ to maximize the η^* value for which

$$\int\limits_{\eta^*}^{1} p_{best}(\eta)\, d\eta \geqslant p_t \qquad (3.60)$$

subject to

$$\int\limits_{0}^{1} \eta p_{best}(\eta)\, d\eta = \langle \eta \rangle. \qquad (3.61)$$

Likewise, $p_{worst}(\eta)$ minimizes the η^* value for which

$$\int\limits_{\eta^*}^{1} p_{worst}(\eta)\, d\eta \geqslant p_t \qquad (3.62)$$

subject to

$$\int\limits_{0}^{1} \eta p_{worst}(\eta)\, d\eta = \langle \eta \rangle. \qquad (3.63)$$

We will use η_{best} and η_{worst} to denote these maximum and minimum values of η^* as defined above. The following theorem gives our results for η_{best} and η_{worst}.

Theorem: For given values of $\langle \eta \rangle$ and p_t,

$$\eta_{best} = \begin{cases} \frac{\langle \eta \rangle}{p_t}, & \text{for } \langle \eta \rangle \leqslant p_t \leqslant 1 \\ 1, & \text{otherwise} \end{cases} \tag{3.64}$$

and

$$\eta_{worst} = \begin{cases} \frac{\langle \eta \rangle - p_t}{1 - p_t}, & \text{for } p_t \leqslant \langle \eta \rangle \\ 0, & \text{otherwise} \end{cases} \tag{3.65}$$

Proof: We will start with the η_{best} result. Consider the discrete probability distribution

$$\Pr(\eta = 0) = 1 - p_t$$
$$\Pr(\eta = x) = p_t \tag{3.66}$$

where $0 \leqslant x \leqslant 1$ is chosen to satisfy the $\langle \eta \rangle$ constraint,

$$x \Pr(\eta = x) = \langle \eta \rangle. \tag{3.67}$$

We conclude that

$$x = \frac{\langle \eta \rangle}{p_t}, \qquad \text{for } \langle \eta \rangle \leqslant p_t \leqslant 1 \tag{3.68}$$

and we assert that this distribution yields $\eta_{\max}$ for p_t in the range given above.

To prove that this distribution is indeed optimal, we first write

$$\langle \eta \rangle = \int_0^x \eta p(\eta) \, d\eta + \int_x^1 \eta p(\eta) \, d\eta \tag{3.69}$$

60

for $0 \leqslant x \leqslant 1$ and then use the lower limits in each integral to obtain

$$\langle \eta \rangle \geqslant \int\limits_0^x 0 p\left(\eta\right) d\eta + \int\limits_x^1 x p\left(\eta\right) d\eta = x p_t, \tag{3.70}$$

where we have assumed that x is such that

$$\int\limits_x^1 p\left(\eta\right) d\eta = p_t. \tag{3.71}$$

Therefore, any transmissivity x that is supportable a fraction p_t of the time must satisfy

$$x \leqslant \frac{\langle \eta \rangle}{p_t}. \tag{3.72}$$

The distribution from (3.66) saturates this bound, and so it must be optimal when $\langle \eta \rangle \leqslant p_t$. We can thus conclude that in this $\langle \eta \rangle$ range

$$C_t\left(p_t\right) \leqslant C\left(\frac{\langle \eta \rangle}{p_t} \bar{N}_T, \left(1 - \frac{\langle \eta \rangle}{p_t}\right) N_B\right) = C_{best}\left(\langle \eta \rangle, p_t\right). \tag{3.73}$$

and

$$R\left(p_t\right) \leqslant p_t C\left(\frac{\langle \eta \rangle}{p_t} \bar{N}_T, \left(1 - \frac{\langle \eta \rangle}{p_t}\right) N_b\right) \tag{3.74}$$

for the thermal-noise channel.

When $\langle \eta \rangle > p_t$ we cannot use the preceding $\eta_{\max}$ expression. In this regime, however, we have $\eta_{\max} = 1$, and

$$C_t\left(p_t\right) \leqslant C\left(\bar{N}_T, 0\right) = C_{best}\left(\langle \eta \rangle, p_t\right) \tag{3.75}$$

because $\eta_{\max}$ is a monotonically increasing function of $\langle \eta \rangle$, and for $\langle \eta \rangle \leqslant p_t$, $\lim\limits_{\langle \eta \rangle \to p_t} \frac{\langle \eta \rangle}{p_t} = 1$.

Now let us turn to the case of η_{worst}. Consider the distribution

$$\Pr\left(\eta = 1\right) = p_t$$
$$\Pr\left(\eta = x\right) = 1 - p_t \tag{3.76}$$

where $0 \leqslant x \leqslant 1$ is chosen to satisfy

$$\Pr\left(\eta = 1\right) + x \Pr\left(\eta = x\right) = \langle \eta \rangle \tag{3.77}$$

so that

$$x = \frac{\langle \eta \rangle - p_t}{1 - p_t}, \quad \text{for } 0 \leqslant p_t \leqslant \langle \eta \rangle. \tag{3.78}$$

To prove that (3.76) is the distribution that yields $\eta_{\min}$, we proceed as follows. As in our $\eta_{\max}$ proof, we start with

$$\langle \eta \rangle = \int_0^x \eta p\left(\eta\right) d\eta + \int_x^1 \eta p\left(\eta\right) d\eta. \tag{3.79}$$

This time, however, we use the upper limit on each integral to obtain

$$\langle \eta \rangle \leqslant \int_0^x x p\left(\eta\right) d\eta + \int_x^1 1 p\left(\eta\right) d\eta = x\left(1 - p_t\right) + p_t \tag{3.80}$$

so that

$$x \geqslant \frac{\langle \eta \rangle - p_t}{1 - p_t} \tag{3.81}$$

where we have assumed that x satisfies (3.71). Because the distribution in (3.76) satisfies

this bound with equality, it gives $\eta_{\min}$ when $\langle \eta \rangle \geqslant p_t$. In this p_t range we then have

$$C_t\left(p_t\right) \geqslant C\left(\frac{\langle \eta \rangle - p_t}{1 - p_t}\bar{N}_T, \frac{1 - \langle \eta \rangle}{1 - p_t}N_B\right) = C_{worst}\left(\langle \eta \rangle, p_t\right) \tag{3.82}$$

and

$$R\left(p_t\right) \geqslant p_t C\left(\frac{\langle \eta \rangle - p_t}{1 - p_t}\bar{N}_T, \left(\frac{1 - \langle \eta \rangle}{1 - p_t}\right)N_B\right) \tag{3.83}$$

for the thermal-noise channel. For $\langle \eta \rangle < p_t \leqslant 1$, monotonicity implies that $\eta_{\min} = 0$, and hence

$$C_t\left(p_t\right) \geqslant C\left(0, N_B\right) = 0 = C_{worst}\left(\langle \eta \rangle, p_t\right). \tag{3.84}$$

In Fig. 3-6, we plot outage rate bounds, (3.83) and (3.74), versus p_t for the pure-loss channel with $\bar{N}_T = 0.001$ and $\langle \eta \rangle = 0.99$, and in Fig. 3-7 we do the same for $\langle \eta \rangle = 0.9$. We see that our bounds are quite tight for p_t appreciably less than $\langle \eta \rangle$ in these near-field cases. We will not plot the outage-rate bounds for $\langle \eta \rangle \ll 1$, because the lower bound is quite bad in this far-field regime.

Note that (3.83) gives a lower bound on the average rate of transmission for a fading channel with average transmissivity $\langle \eta \rangle$ regardless of the details of the distribution. It is of interest to find what this rate is for various values of $\langle \eta \rangle$. In Fig. 3-8 we plot this bound for various values of $\langle \eta \rangle$, and compare them to the value of $C(\langle \eta \rangle, \bar{N}_T)$, the non-fading capacity.

On the other hand, we may want to set the probability of transmitting to a constant and then calculate what the best possible rate of transmission is for this fraction of the time. We can use (3.84) to calculate a lower bound on the capacity for a realistic p_t. Suppose we wish to transmit for fraction 0.95 of the time, and we are in the near field where $\langle \eta \rangle = 0.99$. Then we conclude from (3.84) that our outage capacity is guaranteed to be satisfy:

$$C_t\left(0.95\right) \geqslant C\left(\frac{\langle \eta \rangle - p_t}{1 - p_t}\bar{N}_T, 0\right) = g\left(0.8\bar{N}_T\right) \tag{3.85}$$
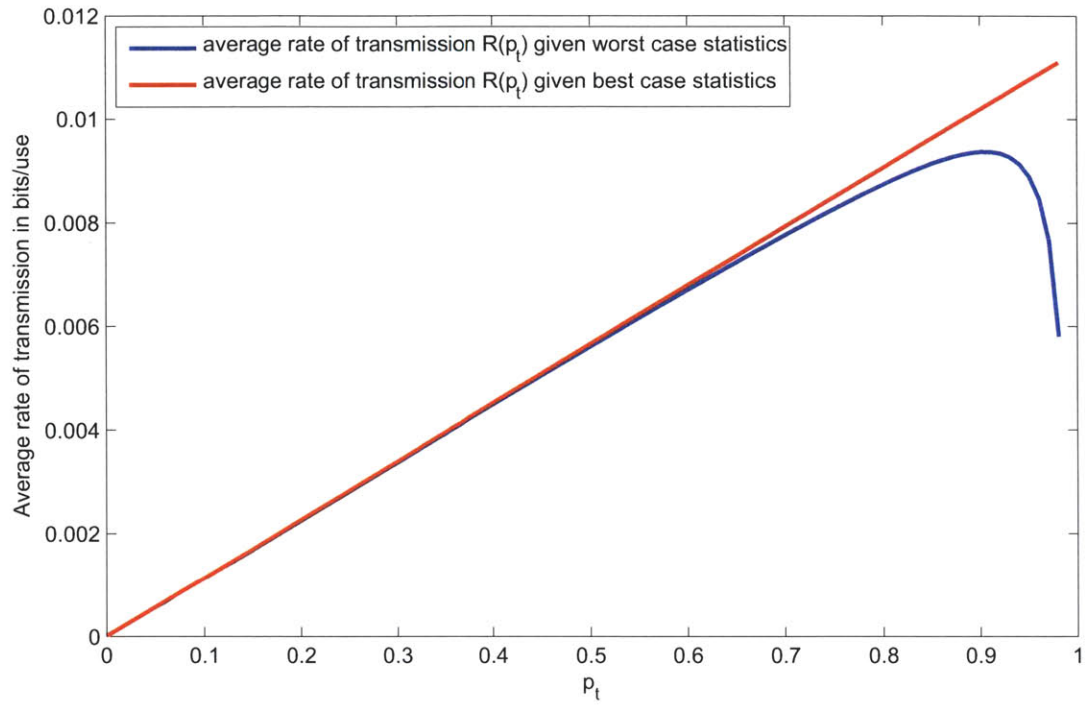
Figure 3-6: Upper and lower bounds on the outage rate, $R(p_t)$, of a fading pure-loss channel as a function of p_t when $\bar{N}_T = 0.001$ and $\langle \eta \rangle = 0.99$.

for the pure-loss channel. In other words, during each coherence interval we can transmit at a rate that is as if the transmissivity of the channel was 0.8, but we can do that for fraction 0.95 of the time. So, our long term average rate of transmission is given by

$$R = 0.95g\left(0.8\bar{N}_T\right). \tag{3.86}$$

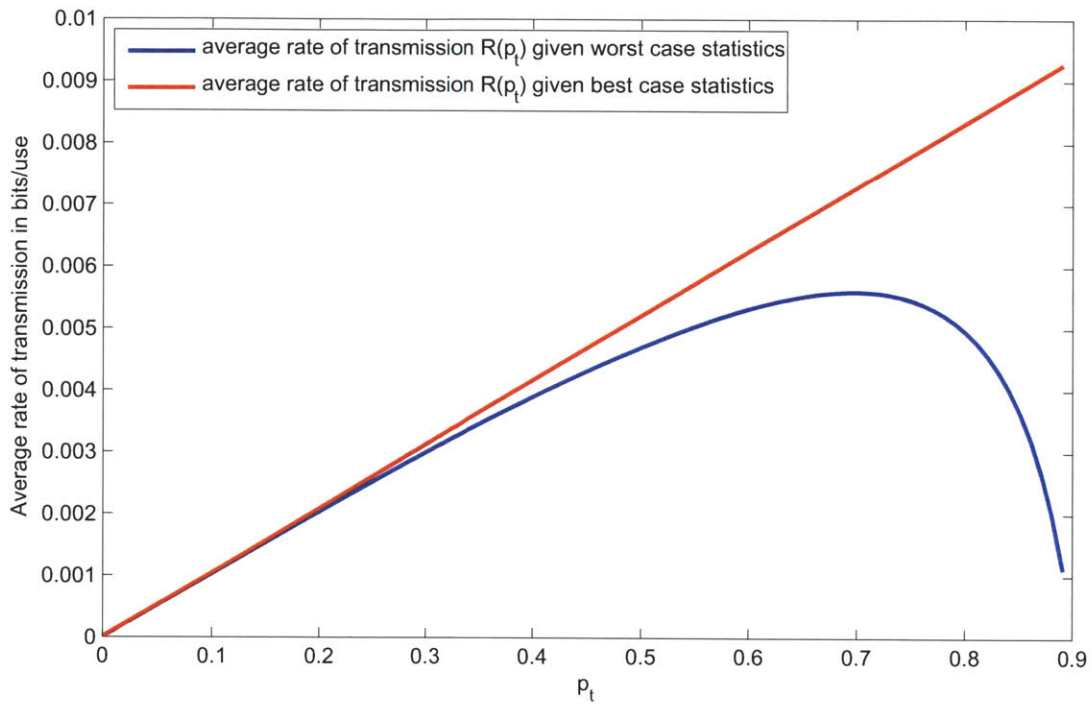for the $\langle \eta \rangle = 0.99$, $p_t = 0.95$, pure-loss case.

Figure 3-7: Upper and lower bounds on the outage rate, $R(p_t)$, of a fading pure-loss channel as a function of p_t when $\bar{N}_T = 0.001$ and $\langle \eta \rangle = 0.9$.
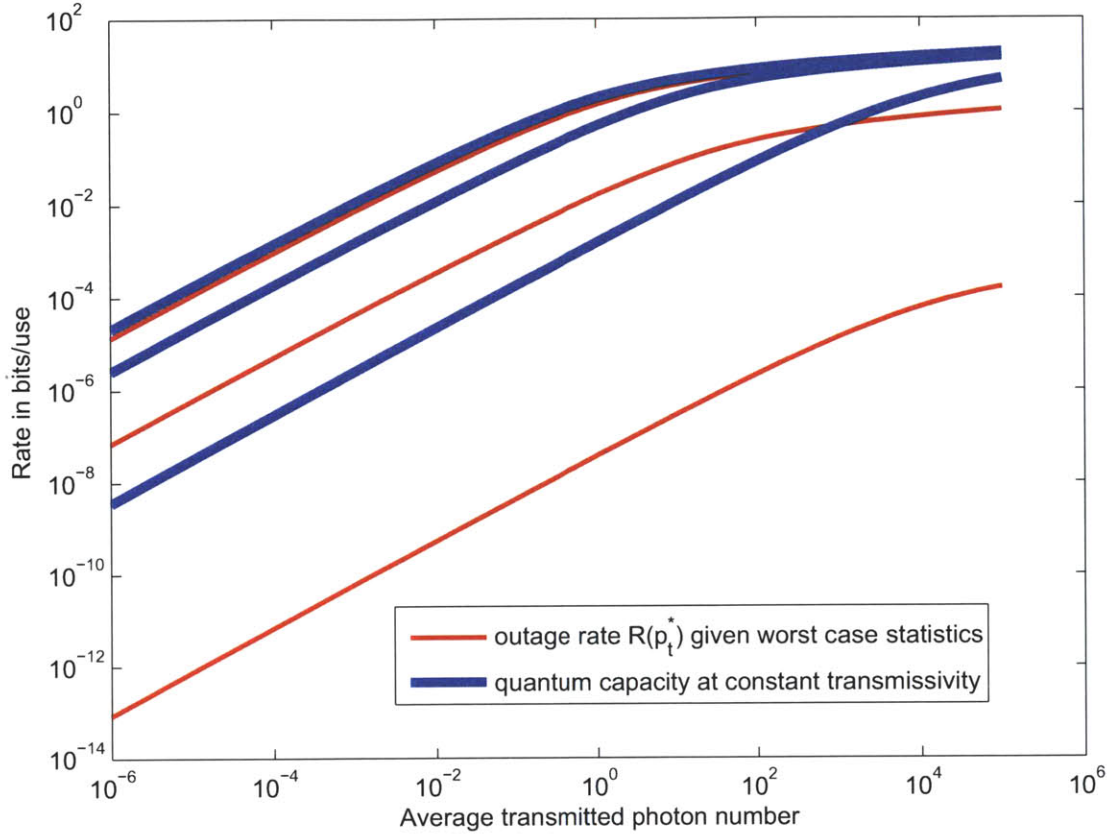
Figure 3-8: Lower bounds on outage rate, $R(p_t^*)$, for (from top to bottom) $\langle\eta\rangle = 0.95$, 0.5, and 0.001. We note that this bound is very tight for high $\langle\eta\rangle$, but it quickly becomes a very bad bound for low $\langle\eta\rangle$. Thus, this bound works very well in the near field, but poorly in the far field. Also plotted is the capacity when the transmissivity is constant. This acts as an upper bound on outage rate, and also on ergodic capacity. Interestingly, as we showed in (3.46), the $R(p_t^*)$ curve also is a lower bound on ergodic capacity. So, these curves also show upper and lower bounds on the ergodic capacity of a fading channel given $\langle\eta\rangle$, regardless of the specific distribution of $\langle\eta\rangle$.

Chapter 4

Multi-mode Fading Channel with Intermodal Interference

4.1 Definition of Channel

To model a multi-mode quantum fading channel with intermodal interference, we will denote each of the t transmitters by a vector of annihilation operators

$$\underline{\hat{a}} = \begin{pmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_t \end{pmatrix}.$$

(4.1)

The receiver's output corresponds to a vector of annihilation operators

$$\underline{\hat{b}} = \begin{pmatrix} \hat{b}_1 \\ \vdots \\ \hat{b}_r \end{pmatrix},$$

(4.2)

each one a random mixture of the input annihilation operators and a vector of noise modes

$$\hat{\underline{e}} = \begin{bmatrix} \hat{e}_1 \\ \vdots \\ \hat{e}_k \end{bmatrix} \tag{4.3}$$

given by:

$$\hat{\underline{b}} = H\hat{\underline{a}} + \mathrm{R}\,\hat{\underline{e}}, \tag{4.4}$$

where $r = t+k$. Here, H is an $r \times t$ matrix and R is an $r \times k$ matrix such that the annihilation operator commutator relationships are preserved:

$$\left[\hat{b}_i, \hat{b}_j\right] = 0 \tag{4.5}$$

$$\left[\hat{b}_i, \hat{b}_j^\dagger\right] = \delta_{ij}. \tag{4.6}$$

We will allow the noise modes to be in independent thermal states with average photon number $\left\langle \hat{e}_i^\dagger \hat{e}_i \right\rangle = N_{B_i}$.

Suppose that the transfer matrix H is given by

$$H = \begin{bmatrix} \sqrt{\eta_{11}} & \sqrt{\eta_{12}} & \cdots & \sqrt{\eta_{1t}} \\ \sqrt{\eta_{21}} & \ddots & \ddots & \sqrt{\eta_{2t}} \\ \vdots & \ddots & \ddots & \vdots \\ \sqrt{\eta_{r1}} & \sqrt{\eta_{r2}} & \cdots & \sqrt{\eta_{rt}} \end{bmatrix}, \tag{4.7}$$

where $\sum_{j=1}^{t} \eta_{ij} \leqslant 1$ for $1 \leqslant i \leqslant r$. It follows that

$$
R = \begin{pmatrix} \sqrt{1 - \sum_{j=1}^{t} \eta_{1j}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{1 - \sum_{j=1}^{t} \eta_{rj}} \end{pmatrix}
\tag{4.8}
$$

achieves the commutator preservation, making Fig. 4-1 the multi-mode generalization of the fading channel model from Chapter 3. In general, the matrix H will be random. In what follows, we will first calculate the capacity of the channel for deterministic H. After that we will calculate the ergodic capacity for random H. Finally, we will apply our results for random H to a sparse aperture system operating through atmospheric turbulence.
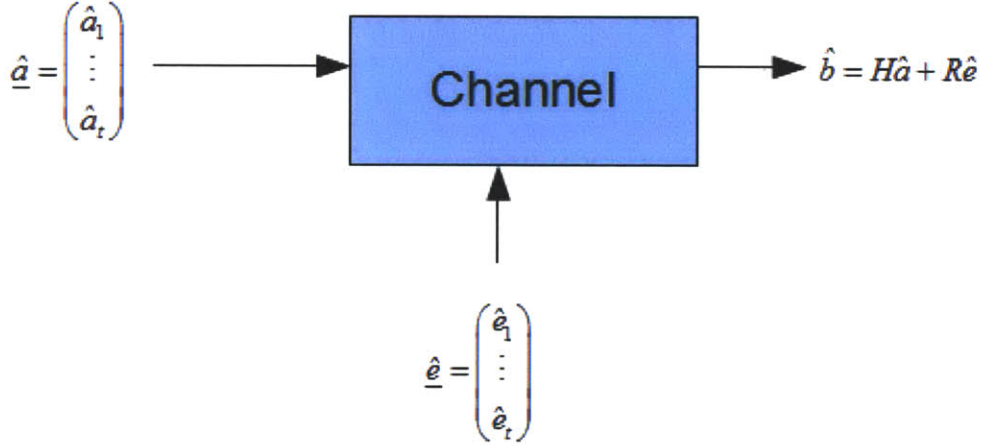

Figure 4-1: Model of multi-mode fading channel with intermodal interference. The entries of the matrices H and R are such that the output modes have the proper commutator relations, $\left[\hat{b}_i, \hat{b}_j\right] = 0$ and $\left[\hat{b}_i, \hat{b}_j^{\dagger}\right] = \delta_{ij}$.

4.2 Deterministic Transfer Matrix

To find the capacity of a multi-mode channel whose H matrix is deterministic, we follow the derivation from [9]. By the singular-value decomposition theorem, any matrix $H \in \mathbb{C}^{r \times t}$ can be written as

$$H = UDV^\dagger \tag{4.9}$$

where $U \in \mathbb{C}^{r \times r}$ and $V \in \mathbb{C}^{t \times t}$ are unitary. Furthermore, $D \in \mathbb{C}^{r \times t}$ is a diagonal matrix whose non-zero values are the non-negative square roots of the eigenvalues of $HH^\dagger$. We can thus express (4.4) as

$$\hat{\underline{b}} = UDV^\dagger \hat{\underline{a}} + R\hat{\underline{e}}. \tag{4.10}$$

We let $\hat{\underline{b}}^* = U^\dagger \hat{\underline{b}}$, $\hat{\underline{a}}^* = V^\dagger \hat{\underline{a}}$, and $R^* = U^\dagger R$. Because U and V are invertible, the original channel is equivalent to

$$\hat{\underline{b}}^* = D\hat{\underline{a}}^* + R^*\hat{\underline{e}}. \tag{4.11}$$

Because the rank of H is at most $\min\{r, t\}$, it will have at most $\min\{r, t\}$ non-zero singular values, which we will denote $\sqrt{\eta_i}$. Writing the input output relationship component-wise, we get

$$\hat{b}_i^* = \sqrt{\eta_i}\hat{a}_i^* + \sum_{j=1}^{k} r_{ij}^* \hat{e}_j, \quad \text{for } 1 \leqslant i \leqslant t^* \tag{4.12}$$

where $t^* = \text{rank}\,(H)$ and we have ignored any modes associated with zero eigenvalues. At this point we will specialize to a pure-loss channel, so that the $\{\hat{e}_j\}$ will all be in their vacuum states. We can then rewrite (4.12) in terms of a transformed set of vacuum state modes $\{\hat{e}_j^*\}$

such that

$$\hat{b}_i^* = \sqrt{\eta_i}\hat{a}_i^* + \sqrt{1 - \eta_i}\hat{e}_i^*, \quad \text{for } 1 \leqslant i \leqslant t^* \tag{4.13}$$

Having reduced our multi-mode channel to a collection of t^* independent channels with individual transmissivities η_i, it remains for us to determine its capacity, subject to the restriction that the total average transmitted photon number is less than or equal to $\bar{N}_T$. We will now provide a simple Lagrange multiplier derivation for this capacity. We first realize that it only makes sense that each individual, independent channel uses an input distribution that makes full use of its allocated average photon number. In doing so, the contribution to the capacity from the ith subchannel will be $g\left(\eta_i\bar{N}_i\right)$, where $\bar{N}_i$ is its average transmitted photon number. Thus, finding the overall capacity reduces to allocating the powers $\bar{N}_1, \ldots, \bar{N}_T$ to maximize

$$\sum_{i=1}^{t^*} g\left(\eta_i\bar{N}_i\right) \tag{4.14}$$

subject to the restriction that

$$\sum_{i=1}^{t^*} \bar{N}_i = \bar{N}_T. \tag{4.15}$$

The constrained maximization is equivalent to the unconstrained maximization of

$$\sum_{i=1}^{t^*} g\left(\eta_i\bar{N}_i\right) - \lambda\left(\sum_{i=1}^{t^*}\bar{N}_i - \bar{N}_T\right). \tag{4.16}$$

Using

$$g'\left(x\right) = \ln\left(x + 1\right) - \ln\left(x\right) \tag{4.17}$$

71

to take the partial derivative of (4.16) with respect to $\bar{N}_i$ gives

$$\frac{\partial}{\partial N_i} \left[\sum_{i=1}^{t^*} g\left(\eta_i \bar{N}_i\right) - \lambda \left(\sum_{i=1}^{t^*} \bar{N}_i - \bar{N}_T \right) \right] = \eta_i \left(\log\left(\eta_i \bar{N}_i + 1\right) - \log\left(\eta_i \bar{N}_i\right) \right) - \lambda. \quad (4.18)$$

A necessary condition for optimization will be when this equation is zero, implying an optimum value occurs when

$$\eta_i \left(\ln\left(\eta_i \bar{N}_i + 1\right) - \ln\left(\eta_i \bar{N}_i\right) \right) - \lambda = 0. \quad (4.19)$$

Solving for $\bar{N}_i$ we find that

$$\bar{N}_i = \frac{1}{\eta_i \left(e^{\frac{\lambda}{\eta_i}} - 1 \right)}. \quad (4.20)$$

We note that $\bar{N}_i$ is positive for all possible values of the parameter λ, indicating that all modes with non-zero path gain should be used if optimal transmission is to occur. This contrasts to the classical case of an additive white Gaussian noise channel with different path gains, in which the well known "water-filling" formula implies that some modes will not be used until the available power becomes sufficiently high to justify their use [17]. We now have the values of $\bar{N}_i$ parameterized by λ, where λ must be chosen so that

$$\sum_{i=1}^{t^*} \frac{1}{\eta_i \left(e^{\frac{\lambda}{\eta_i}} - 1 \right)} = \bar{N}_T. \quad (4.21)$$

With this value of λ our multi-mode channel capacity will be given by

$$C = \sum_{i=1}^{t^*} g\left(\eta_i \bar{N}_i\right) = \sum_{i=1}^{t^*} g\left(\frac{1}{e^{\frac{\lambda}{\eta_i}} - 1} \right). \quad (4.22)$$

4.3   Deterministic Transfer-Matrix Examples

The capacity of a pure-loss channel with an $r \times t$ transfer matrix H can be calculated by finding the eigenvalues of $HH^\dagger$ and using the results of the previous section. We will now perform such calculations for several transfer matrix examples.

Consider a transfer matrix in which all of the entries are idential and $r = t$, i.e., H is the $t \times t$ matrix

$$H = \begin{pmatrix} \sqrt{\langle \eta \rangle / t} & \cdots & \sqrt{\langle \eta \rangle / t} \\ \vdots & \ddots & \vdots \\ \sqrt{\langle \eta \rangle / t} & \cdots & \sqrt{\langle \eta \rangle / t} \end{pmatrix}. \tag{4.23}$$

For this H matrix we can take the R matrix to be the $t \times t$ matrix

$$R = \begin{pmatrix} \sqrt{1 - \langle \eta \rangle} & 0 & 0 & 0 \\ 0 & \sqrt{1 - \langle \eta \rangle} & 0 & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{1 - \langle \eta \rangle} \end{pmatrix}. \tag{4.24}$$

We then have that $HH^\dagger$ is the $t \times t$ matrix

$$HH^\dagger = \begin{pmatrix} \langle \eta \rangle & \cdots & \langle \eta \rangle \\ \vdots & \ddots & \vdots \\ \langle \eta \rangle & \cdots & \langle \eta \rangle \end{pmatrix}. \tag{4.25}$$

This matrix has one non-zero eigenvalue, namely $t \langle \eta \rangle$, hence, the optimal photon-number allocation is achieved when all the photons are used on used on the eigenmode $\hat{a} = \sum_{i=1}^{t} \frac{\hat{a}_i}{\sqrt{t}}$. The resulting pure-loss channel capacity for this H matrix is then

$$C = g\left(t \langle \eta \rangle \bar{N}_T \right). \tag{4.26}$$

As another example, suppose that H is the $t \times t$ diagonal matrix

$$H = \begin{pmatrix} \sqrt{\langle \eta \rangle} & 0 & \cdots & 0 \\ 0 & \sqrt{\langle \eta \rangle} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{\langle \eta \rangle} \end{pmatrix} \qquad (4.27)$$

and so:

$$HH^\dagger = \begin{pmatrix} \langle \eta \rangle & 0 & \cdots & 0 \\ 0 & \langle \eta \rangle & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \langle \eta \rangle \end{pmatrix} \qquad (4.28)$$

This matrix has t non-zero eigenvalues, each equal to $\langle \eta \rangle$, whose eigenmodes are the $\{\hat{a}_i\}$. Capacity is thus achieved when $\bar{N}_T$ is equally distributed across these eigenmodes, yielding,

$$C = \sum_{i=1}^{t} g\left(\langle \eta \rangle \frac{\bar{N}_T}{t} \right) = tg\left(\langle \eta \rangle \frac{\bar{N}_T}{t} \right). \qquad (4.29)$$

For the rest of this chapter the transfer matrix under consideration will not be as simple as these two deterministics examples. Instead, we shall treat H matrices that are random, and, in particular, focus on a model for a sparse-aperture system operating through atmospheric turbulence.

4.4 Transfer Matrix Statistics

Consider the sparse-aperture system for communication through atmospheric turbulence that is shown in Fig. 4-2 [11]. Here there are t small transmitter apertures in the $z = 0$ plane and r small receiver apertures in the $z = L$ plane. Each of the small transmitter apertures, A_{t_j}, for $1 \leqslant j \leqslant t$, is a diameter-$d_t$ circular pupil with center $\vec{\rho}_j$. Each of the

small receiver apertures, A_{r_j}, for $1 \leq j \leq r$ is a diameter-d_r circular pupil with center $\vec{\rho}_j'$. The center-to-center spacing between the adjacent transmitter apertures is δ_t, and that for the receiver apertures is δ_r. The center of the constellation of transmitter apertures is the origin $\vec{0}$ in the $z = 0$ plane, and similarly that for the receiver apertures is the origin in the $z = L$ plane.
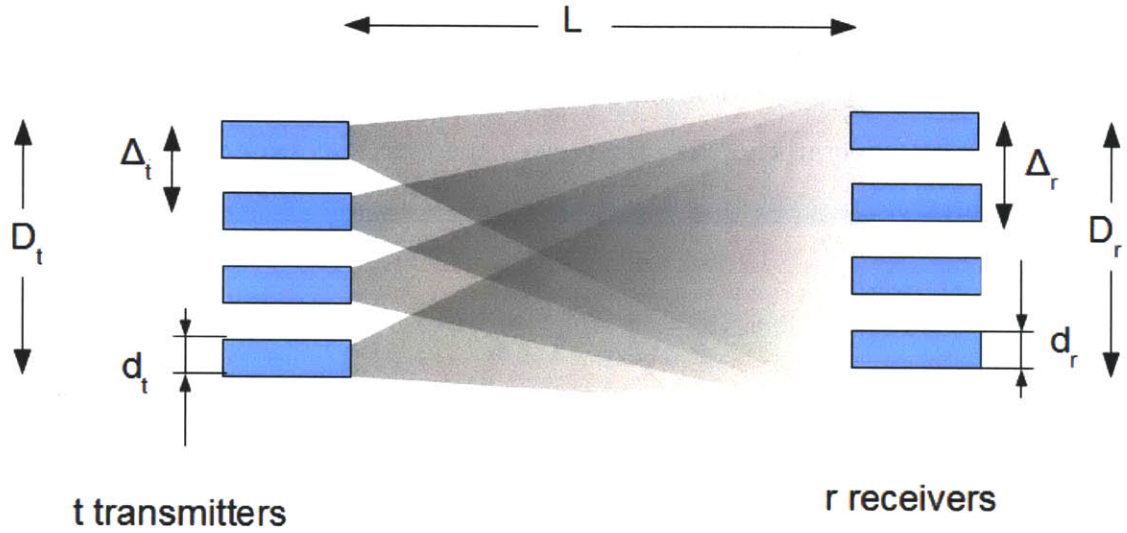

Figure 4-2: Diagram of sparse aperture system. Transmitters and receivers are arranged in a two-dimensional constellation.

We shall assume, similar to [11], that the photon-units classical complex field transmitted from the jth transmitter aperture is

$$E_{t_j}(\vec{\rho}, t) = \sqrt{\frac{4\bar{N}_j}{\pi d_t^2}} s_j(t) e^{\frac{-ik\vec{\rho}' \cdot \vec{\rho}_j}{L}}, \quad \text{for } |\vec{\rho} - \vec{\rho}_j| \leq \frac{d_t}{2} \tag{4.30}$$

where $\bar{N}_j$ is its average photon number, $d_t^2/\lambda L \ll 1$ is assumed, and the phase tilt has the effect of centering the resulting $z = L$ field pattern in the constellation of receiver apertures.

75

Here, $s_j(t)$ is the field's information-bearing modulation, normalized to satisfy

$$\int\limits_0^T |s_j(t)|^2 dt = 1 \tag{4.31}$$

for the $0 \leqslant t \leqslant T$ channel use. Note that because we will be concerned with the coherent state transmitters, it suffices, for now, to work with the classical fields that are the coherent-state eigenfunctions.

We will continue to parallel [11] by assuming that the kth receiver element only measures light arriving within the diffraction-limited field of view. Thus, if $\hat{E}_{r_k}(\vec{\rho}', t)$ is the photon-units, baseband field operator in the kth receiver aperture, the receiver associated with this aperture measures

$$\hat{y}_k(t) = \int\limits_{A_{r_k}} d\vec{\rho}' \sqrt{\frac{4}{\pi d_r{}^2}} \hat{E}_{r_k}(\vec{\rho}', t), \tag{4.32}$$

where we will presume $d_r{}^2 / \lambda L \ll 1$.

Applying the extended Huygens-Fresnel principle to this sparse-aperture configuration we have that

$$y_k(t) = \sum_{j=1}^t \int\limits_{A_{r_k}} d\vec{\rho}' \sqrt{\frac{4}{\pi d_r{}^2}} \int\limits_{A_{t_j}} d\vec{\rho} \sqrt{\frac{4N_j}{\pi d_t{}^2}} e^{\frac{-jk\vec{\rho}' \cdot \vec{\rho}_j}{L}} h(\vec{\rho}', \vec{\rho}) s_j(t) \tag{4.33}$$

where we have suppressed the L/c propagation delay, and

$$h(\vec{\rho}', \vec{\rho}) = \frac{e^{\left(ikl + \frac{ik|\vec{\rho}' - \vec{\rho}|^2}{2L}\right)}}{i\lambda L} e^{\chi(\vec{\rho}', \vec{\rho}) + i\phi(\vec{\rho}', \vec{\rho}')} \tag{4.34}$$

is the atmospheric impulse response, as in Chapter 3, gives the coherent state eigenfunction for $\hat{y}_k(t)$, conditional on knowledge of the logamplitude and phase fluctuations $\chi(\vec{\rho}', \vec{\rho})$ and $\phi(\vec{\rho}', \vec{\rho})$.

We next assume that d_t and d_r are smaller than the turbulence coherence lengths in the $z = 0$ and $z = L$ planes, respectively, thus allowing us to use

$$h\left(\vec{\rho}', \vec{\rho}\right) = \frac{e^{\left(ikl + \frac{ik|\vec{\rho}' - \vec{\rho}|^2}{2L}\right)}}{i\lambda L} e^{\chi\left(\vec{\rho}_k', \vec{\rho}_j\right) + i\phi\left(\vec{\rho}_k', \vec{\rho}_j\right)} \tag{4.35}$$

for $\vec{\rho} \in A_{t_j}$ and $\vec{\rho}' \in A_{r_k}$. Under this condition we have that

$$y_k(t) = \sum_{j=1}^{t} \left[\int_{A_{r_k}} d\vec{\rho}' \sqrt{\frac{4}{\pi d_r^2}} \int_{A_{t_j}} d\vec{\rho} \sqrt{\frac{4\bar{N}_j}{\pi d_t^2}} \frac{e^{\frac{-ik\vec{\rho}' \cdot \left(\vec{\rho} - \vec{\rho}_j\right)}{L}}}{i\lambda L} \right] e^{\frac{ik|\vec{\rho}'|^2}{2L}} e^{\chi\left(\vec{\rho}_k', \vec{\rho}_j\right) + i\phi\left(\vec{\rho}_k', \vec{\rho}_j\right)} s_j(t) \tag{4.36}$$

$$= \sum_{j=1}^{t} \left[\int_{A_{r_k}} d\vec{\rho}' \sqrt{\frac{d_t^2 \bar{N}_j}{d_r^2 (\lambda L)^2}} \frac{1}{\lambda L} \frac{J_1\left(\pi d_t |\vec{\rho}'|/\lambda L\right)}{\pi d_t |\vec{\rho}'|/2\lambda L} \right] e^{\chi\left(\vec{\rho}_k', \vec{\rho}_j\right) + i\phi\left(\vec{\rho}_k', \vec{\rho}_j\right)} s_j(t) \tag{4.37}$$

where we have used the fact that $\phi\left(\vec{\rho}_k', \vec{\rho}_j\right)$ is Gaussian with a variance that is much greater than one to neglect the receiver-aperture quadratic phase term and a $-\frac{\pi}{2}$ phase shift.

Our final assumptions are: (1) that $\frac{\lambda L}{d_t}$ exceeds the overall span of the receiver constellation ($\sim \sqrt{r}\Delta_r$); and (2) that Δ_t and Δ_r are larger than the turbulence coherence lengths in the $z = 0$ and $z = L$ planes, respectively. We then get

$$y_k(t) = \sum_{j=1}^{t} \frac{\pi d_t d_r}{4\lambda L} \sqrt{\bar{N}_j} s_j(t) e^{\psi_{kj}} \tag{4.38}$$

where the $\{\psi_{kj}\}$ are statistically independent, identically distributed complex-valued random variables with

$$\psi_{kj} = \chi\left(\vec{\rho}_k', \vec{\rho}_j\right) + i\phi\left(\vec{\rho}_k', \vec{\rho}_j\right). \tag{4.39}$$

We can now put our sparse-aperture channel model for pure-loss operation into quantum

77

form, with the following result:

$$\hat{\underline{b}} = \begin{bmatrix} \hat{b}_1 \\ \vdots \\ \hat{b}_r \end{bmatrix} = H\hat{\underline{a}} + R\hat{\underline{e}} \tag{4.40}$$

and

$$\hat{\underline{a}} = \begin{bmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_t \end{bmatrix} \tag{4.41}$$

are the vectors of output (receiver) and input (transmitter) annihilation operators for a single channel use, and the $\{\hat{e}_l\}$ are in their vacuum states. The transfer matrix H is $r \times t$ with kjth element

$$h_{kj} = \frac{\pi d_t d_r}{4\lambda L} e^{\psi_{kj}} \tag{4.42}$$

and, as explained earlier in this chapter, the R matrix ensures that $\hat{\underline{b}}$ has the proper annihilation-operator commutator brackets.

Now let us employ the singular-value decomposition for H — whose eigenvalues and right eigenvectors will be random quantities — as in the previous section to get an input-output relation of the form

$$\hat{b}_k^* = \sqrt{\eta_k}\hat{a}_k^* + \sqrt{1 - \eta_k}\hat{e}_k^* \tag{4.43}$$

for $1 \leqslant k \leqslant t^*$, where $\{\eta_k\}$ are the non-zero eigenvalues of $HH^\dagger$, and t^* is the rank of $HH^\dagger$. The $\{\hat{a}_k^*\}$, $\{\hat{b}_k^*\}$ and $\{\hat{e}_k^*\}$ are annihilation operators for input, output, and noise modes, with the noise modes being in their vacuum states.

As in Chapter 3, we will seek the ergodic capacity of the preceding sparse-aperture channel model under the assumption that the transmitter and receiver have perfect knowledge of the

eigenvalues and the eigenmodes. The value — from an analysis point of view — of the sparse-aperture setup we have specified is that random matrix theory will allow us to get results for the statistics of $\{\eta_k\}$, and it is the eigenvalue statistics that we will need to find the ergodic capacity.

Puryear and Chan [11] developed the sparse-aperture setup and studied its performance for receivers that used heterodyne detection. Our work will go beyond theirs by treating optimum quantum reception. To do so, we will build on the same theorem from random-matrix theory that they used. That key result is the following. Let A be a random $r \times t$ matrix whose entries are independent, identically distributed, complex-valued, zero mean, unity variance random variables. Then, the eigenvalues $\{\mu_k\}$ of the matrix $\frac{AA^\dagger}{t}$ have their empircal distribution given by the Marchenko-Pastur law [12]

$$f_\beta\left(\mu\right) = \left(1 - \beta^{-1}\right)^+ \delta\left(\mu\right) + \frac{\sqrt{\left(\mu - \left(1 - \sqrt{\beta}\right)^2\right)^+ \left(\left(1 + \sqrt{\beta}\right)^2 - \mu\right)^+}}{2\pi\beta\mu} \tag{4.44}$$

where $\beta = \frac{r}{t}$ and $(y)^+ = \max\left(y, 0\right)$, in the limit as $t \to \infty$ at constant β.

The entries in the transfer matrix for our sparse-aperture system are independent, identically distributed, zero-mean, complex-valued random variables, but their variance is

$$\left\langle |h_{ij}|^2 \right\rangle = \left(\frac{\pi d_t d_r}{4\lambda L}\right)^2 \left\langle e^{2\chi\left(\vec{\rho}_j, \vec{\rho}_i\right)} \right\rangle = \left(\frac{\pi d_t d_r}{4\lambda L}\right)^2 \tag{4.45}$$

because $2\chi\left(\vec{\rho}_j, \vec{\rho}_i\right)$ is a Gaussian random variable whose mean equals minus its variance. It follows that, for large t, the empirical distribution of the eigenvalues $\{\eta_k\}$ can be found by setting $\eta_k = \kappa t \mu_k$, where

$$\kappa = \left(\frac{\pi d_t d_r}{4\lambda L}\right)^2 \ll 1 \tag{4.46}$$

is the average single transmitter aperture to single receiver aperture transmissivity, and the $\{\mu_k\}$ follow the Marchenko-Pastur law. Note that $\kappa \ll 1$ follows because we assumed $\frac{\lambda L}{d_t}$ exceeds the overall span of the receiver constellation, which is itself much bigger than d_r.

Accounting for the preceding scaling, the Marchenko-Pastur law for the $\{\eta_k\}$ becomes

$$f_\beta(\eta) = \left(1 - \beta^{-1}\right)^+ \delta(\eta) + \frac{\sqrt{\left(\frac{\eta}{\kappa t} - (1 - \sqrt{\beta})^2\right)^+ \left((1 + \sqrt{\beta})^2 - \frac{\eta}{\kappa t}\right)^+}}{2\pi\beta\eta}. \qquad (4.47)$$

We have plotted this eigenvalue distribution in Figs. 4-3 and 4-4 and also a simulation for various values of t and r.
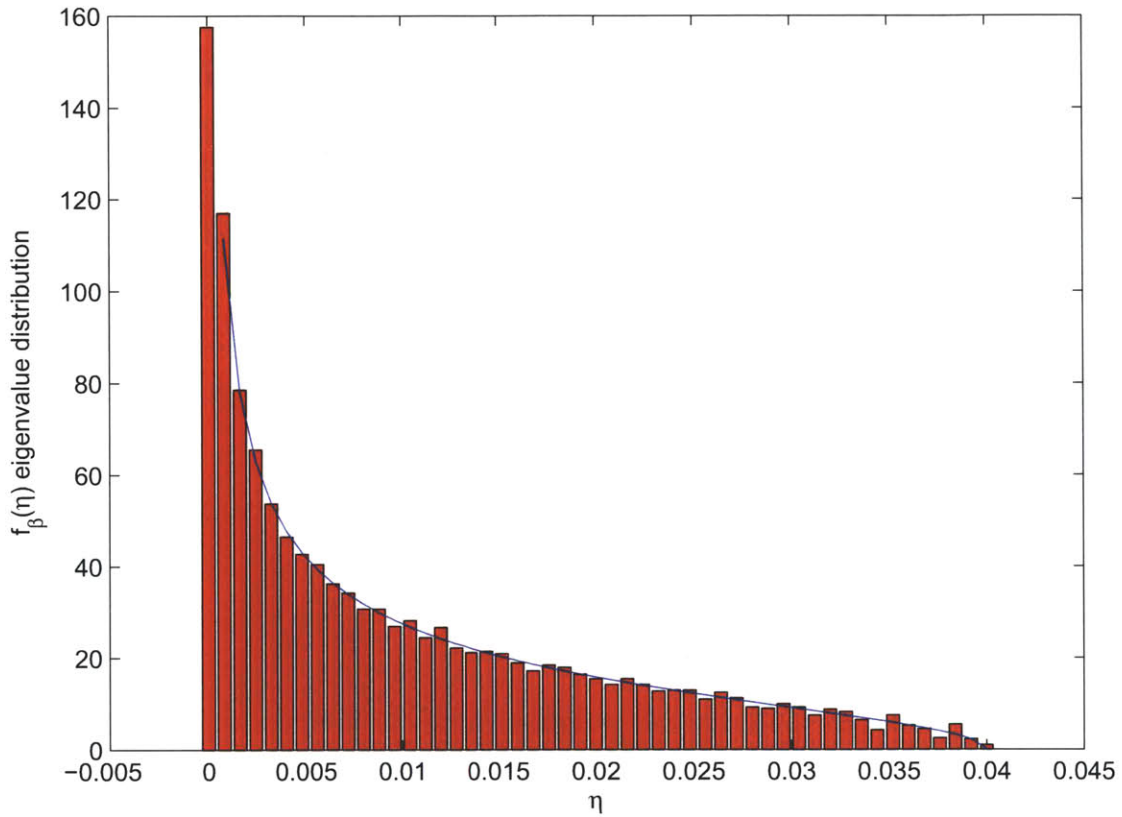

Figure 4-3: Blue curve is the Marchenko-Pastur eigenvalue distribution, from (4.47), of random H matrix when $t = 100$ and $r = 100$, $\kappa = \left(\frac{\pi d_t d_r}{4\lambda L}\right)^2 = 0.0001$. The histogram represents the distribution of eigenvalues obtained by computer simulation after 50 trials.

Another important result from random-matrix theory that we will use is that, in the
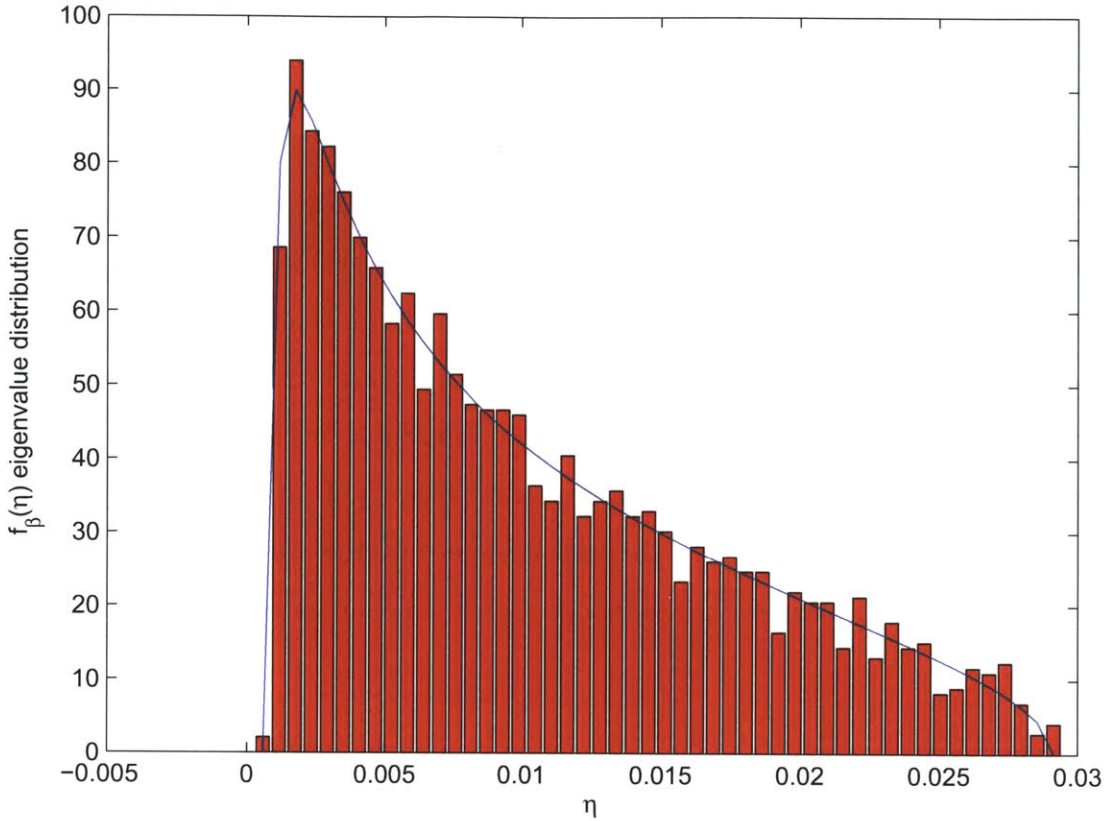
Figure 4-4: Blue curve is the Marchenko-Pastur eigenvalue distribution, from (4.47), of random H matrix when $t = 100$ and $r = 50$, $\kappa = \left(\frac{\pi d_t d_r}{4\lambda L}\right)^2 = 0.0001$. The histogram represents the distribution of eigenvalues obtained by computer simulation after 50 trials.

limit as $t \to \infty$ at constant β, the maximum eigenvalue of $HH^\dagger$ is:

$$\eta_{\max} = \kappa t \left(1 + \sqrt{\beta}\right)^2, \tag{4.48}$$

i.e., for our sparse-aperture system with r/t fixed and $t \gg 1$, the maximum transmissivity over all modes is this $\eta_{\max}$ value. It is important now to understand a limitation of this result. The turbulent atmosphere is passive, so it is not possible for any singular value of this system to have a value greater than 1. We need to recognize that the model used in this section is only an approximation of the atmosphere. In fact, the lognormal distribution model of the atmosphere is also only an approximation; the lognormal distribution has a non-

zero probability of having a value greater than 1. For our Marchenko-Pastur law eigenvalue distribution in our sparse aperture system, we can ensure $\eta_{\max}$ is much smaller than one by choosing $\beta \approx 1$ and $\kappa t \ll 1$, because we already have $\kappa \ll 1$.

We are now ready, with our random-matrix theory results, to calculate some ergodic capacity bounds for the sparse-aperture system. Note that to calculate the ergodic capacity of this system, we would need to know the joint distribution of the ordered eigenvalues. We do not have that distribution, but we can use the knowledge that we have to derive some bounds on the ergodic capacity. We suppose that the transmitter and receiver have full channel knowledge and can use this knowledge to optimally transmit and detect. One lower bound on the ergodic capacity follows from assuming that the transmitter devotes its entire photon budget to the eigenmode with the maximum eigenvalue. For t sufficiently large, the maximum eigenvalue is given by (4.48) with probability 1, and so the ergodic capacity is at least:

$$C_{ergodic} \geqslant g\left(\bar{N}_T \eta_{\max}\right) = g\left(\bar{N}_T \kappa t \left(1 + \sqrt{\beta}\right)^2\right). \tag{4.49}$$

Another lower bound on the ergodic capacity is obtained by assuming that the transmitter distributes its photon budget equally over all the eigenmodes with non-zero eigenvalues. This results in the lower bound

$$C_{ergodic} \geqslant E\left[\sum_{i=1}^{m} g\left(\frac{\bar{N}_T}{m}\eta_i\right)\right] = \sum_{i=1}^{m} E\left[g\left(\frac{\bar{N}_T}{m}\eta_i\right)\right] = mE\left[g\left(\frac{\bar{N}_T}{m}\eta_1\right)\right] \tag{4.50}$$

$$= \int mg\left(\frac{\bar{N}_T}{m}\eta\right) \frac{\sqrt{\left(\frac{\eta}{\kappa t} - \left(1 - \sqrt{\beta}\right)^2\right)^+ \left(\left(1 + \sqrt{\beta}\right)^2 - \frac{\eta}{\kappa t}\right)^+}}{2\pi\beta\eta} d\eta \tag{4.51}$$

where $m = \min(t, r)$. In Fig. 4-5 we plot various ergodic capacity results from the Marchenko-Pastur law along with simulation results for the ergodic capacity. We see that simulated results closely match the Marchenko-Pastur law. We also see that these bounds very closely approach a simulated ergodic capacity in which for each random transfer matrix the photon budget is optimally allocated over each of the eigenmodes according to (4.20).

Note that we have not included a closed-form expression for this capacity, because it requires a joint distribution for the ordered eigenvalues, which is unavailable.
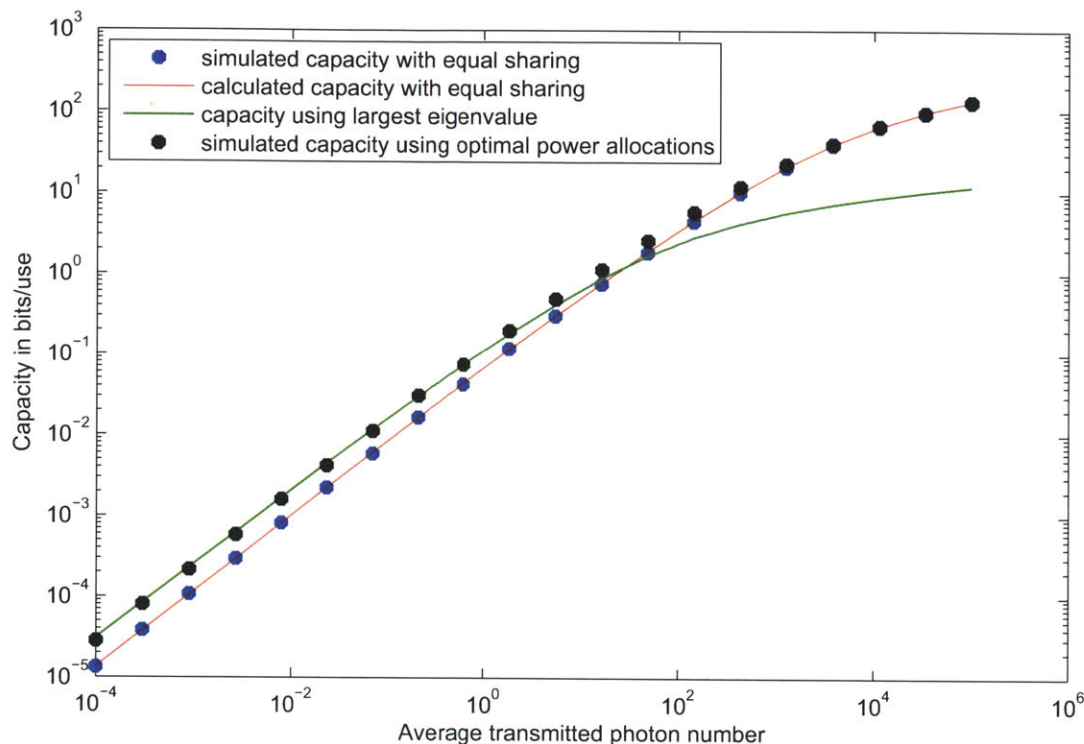

Figure 4-5: Ergodic capacity results for a sparse-aperture, pure-loss system with $\bar{N}_T$ varying, t=25 and r=50, and $\kappa = \left(\frac{\pi d_t d_r}{4\lambda L}\right)^2 = 0.0001$. The figure includes the ergodic capacity when the photons are distributed evenly among all eigenmodes of the system, and also a simulation of this equal-sharing ergodic capacity. Also shown is a simulated optimal power allocation ergodic capacity and a capacity obtained when the maximum eigenvalue is used. We observe that when the maximum of the two lower bounds is used, the result is close to the simulated optimal capacity. We also oberve that the Marchenko-Pastur distribution allows us to calculate very closely the capacity when equal power sharing among the eigenmodes is used because our simulations very closely match the theoretical result. We cannot be sure that this is true because Marchenko-Pastur is only a perfect distribution when the size of the matrix is infinite, but of course in any practical case it is not. This simulation demonstrates that this is not a problem for the chosen values of t and r.

4.5 Scaling Behavior of the Packed Sparse-Aperture System

Now, we would like to calculate how the addition of transmitters and receivers affects these capacity results when the maximum possible number of such transmitters and receivers are used while still preserving this sparse-aperture setup shown in Fig. 4-2. From this figure we see that the maximum number of transmitters that can be accommodated within a given D_t diameter region satisfies

$$t \leqslant k_t \frac{D_t{}^2}{\Delta_t{}^2} \tag{4.52}$$

where $k_t \sim 1$ is a fill factor that accounts for the areas of the transmitter apertures at the edge of the overall constellation. Similarly, the number of receivers is bounded by:

$$r \leqslant k_r \frac{D_r{}^2}{\Delta_r{}^2} \tag{4.53}$$

where $k_r \sim 1$ is defined analogously. We have already required that

$$\frac{\lambda L}{d_t} \gg D_r \tag{4.54}$$

and

$$\frac{\lambda L}{d_r} \gg D_t \tag{4.55}$$

so that each transmitter uniformly illuminates the receiver constellation and each receiver is uniformly sensitive to the entire transmitter constellation. Of course, the diameters of the transmitters and receivers cannot exceed the center-to-center spacings, i.e., it must be that

$$d_t \leqslant \Delta_t \tag{4.56}$$

and

$$d_r \leqslant \Delta_r.$$ (4.57)

We shall now assume that for every transmitter and receiver diameter, the maximum corresponding number of transmitters and receivers are used, and that we are in the region in which the bounds in (4.56) and (4.57) are satisfied.

Rearranging to solve for D_r in (4.53) (assuming this inequality is met with equality) and plugging into (4.54) we observe that:

$$d_t = \lambda L \sqrt{\frac{k_r}{r \Delta_r^2}}.$$ (4.58)

Similarly for d_r we conclude that:

$$d_r = \lambda L \sqrt{\frac{k_t}{t \Delta_t^2}}.$$ (4.59)

Substituting (4.58) and (4.59) into (4.46) gives us the following expression for κ in the case of a completely packed sparse-aperture system

$$\kappa = \left(\frac{\pi \lambda L}{4 \Delta_r \Delta_t} \right)^2 \frac{k_t k_r}{tr}.$$ (4.60)

We can now vary t and r, assuming a totally packed sparse-aperture setup, and observe how this changes the ergodic capacity of the channel, using (4.60) for κ and increasing D_t and D_r to accommodate increases in t and r, respectively. In Fig. 4-6 and Fig. 4-7 we plot the ergodic capacities for maximum eigenvalue and equal photon sharing assignments as we vary t and r. These plots demonstrate that a sparse-aperture setup, once filled, will lose ergodic capacity with the addition of more transmitters or receivers. This is because adding more receivers requires shrinking the transmitter diameters so that their outputs fully illuminate the larger receiver constellation. Similarly, the addition of transmitters requires

85

the shrinking of the receiver apertures so that each receiver is uniformly sensitive to the light from each transmitter. These changes come at the cost of a reduced average received photon number, which does not sufficiently compensate for the benefit of spatial diversity in the fading channels. However, if the transmitter and receiver diameters are constant, it does make sense to fill the D_t-diameter transmitter region and the D_r-diameter receiver region, as this will increase the capacity, cf. Fig. 4-5.
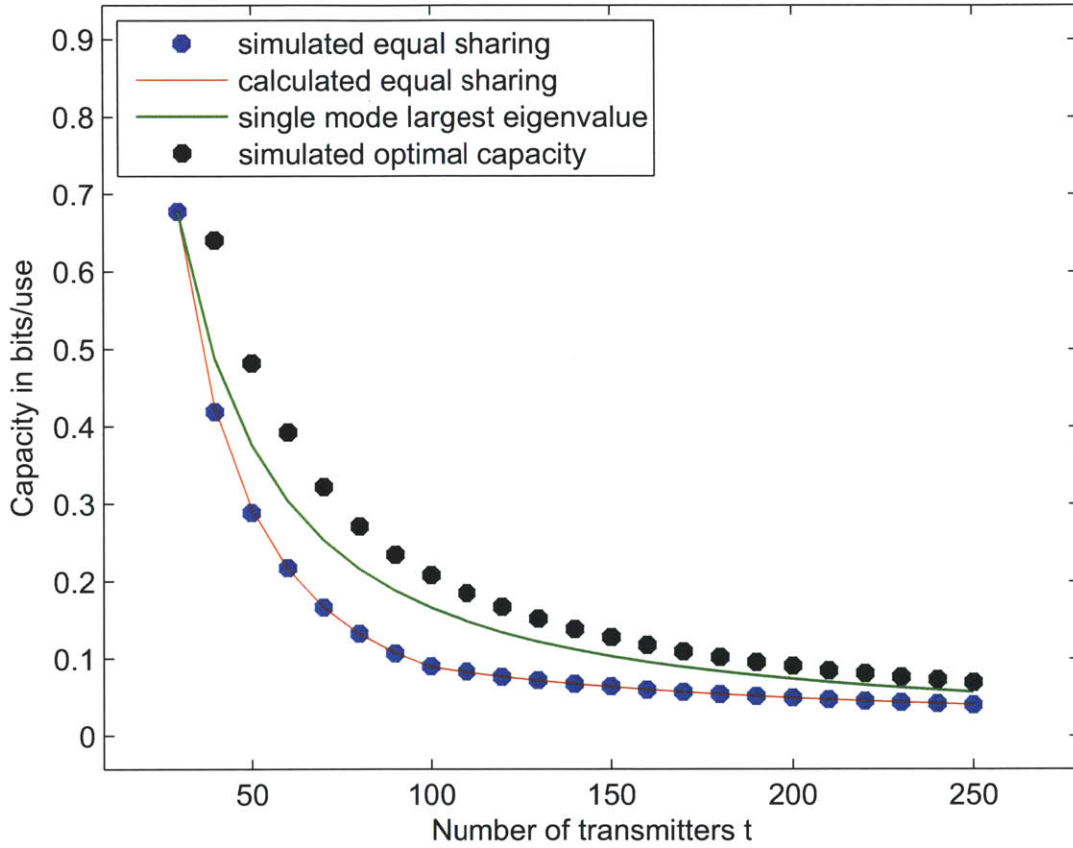

Figure 4-6: Pure-loss channel ergodic capacity results as a function of the number of transmitters t for a fully packed sparse-aperture system where $r = 100$, $\bar{N}_T = 10^{-2}$, $k_t = k_r = 0.9$, and $\kappa = \left(\frac{\pi\lambda L}{4\Delta_r\Delta_t}\right)^2 \frac{k_t k_r}{tr} = \frac{32}{tr}$.
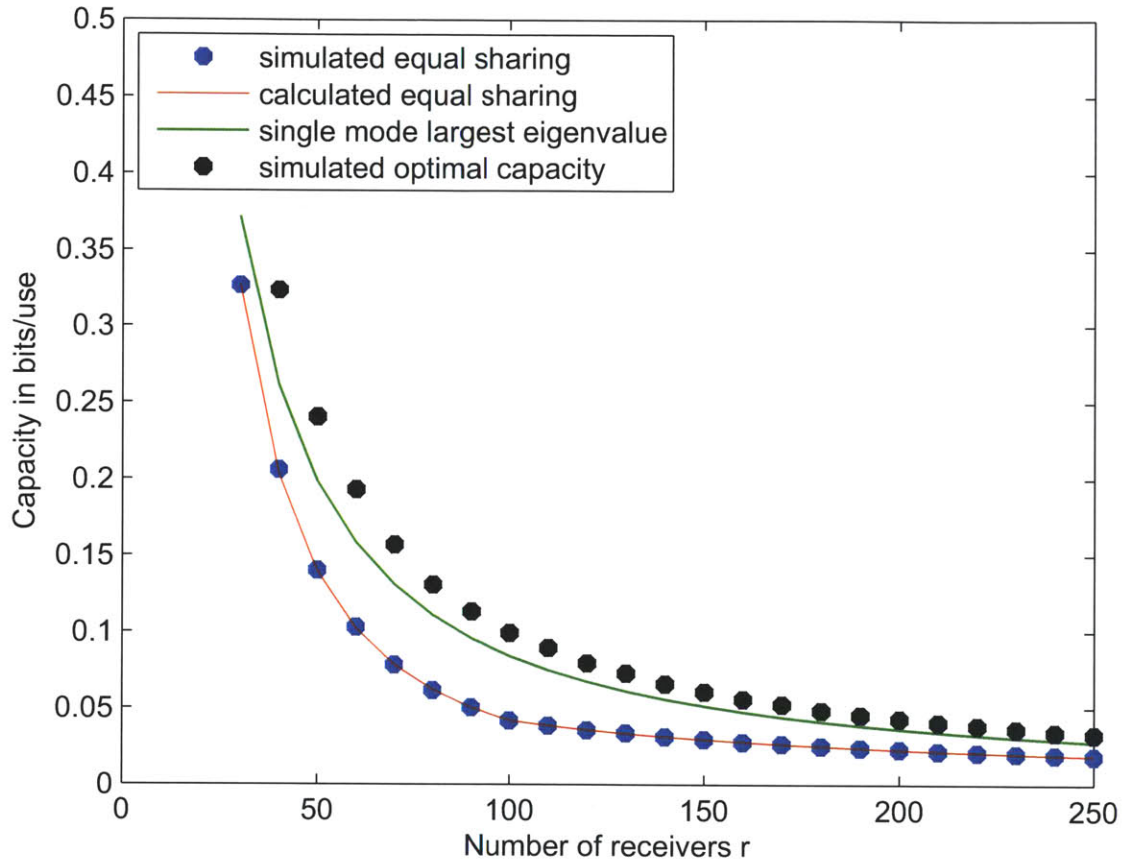
Figure 4-7: Pure-loss channel ergodic capacity results as a function of the number of receivers r for a fully packed sparse-aperture system where $t = 100$, $\bar{N}_T = 10^{-2}$, $k_t = k_r = 0.9$, and $\kappa = \left(\frac{\pi \lambda L}{4 \Delta_r \Delta_t} \right)^2 \frac{k_t k_r}{tr} = \frac{32}{tr}$.

Chapter 5

Conclusion

In this thesis we have investigated the capacities of several bosonic channels. We first introduced the fundamentals of quantum optics that are relevant to the study of classical communication over bosonic channels. Then we presented the Holevo-Schumacher-Westmoreland (HSW) theorem, which is the quantum equivalent of Shannon's noisy-channel coding theorem. Next we introduced two important bosonic channels: the pure-loss channel and the thermal-noise channel. The pure-loss channel has a known quantum capacity, and the thermal-noise channel has a lower bound to its capacity that is conjectured to be the capacity. We showed that when the average received noise photon number is proportional to average received signal photon number and much smaller than one, then direct-detection on-off keying is asymptotically tight to the conjectured quantum capacity of the thermal-noise channel. We extended the single-mode result by computing the spectral and photon efficiencies of the thermal-noise channel when multiple spatial modes are used. Here we found that heterodyne and homodyne detection cannot achieve high photon efficiencies, but that direct detection and optimal quantum detection can achieve high photon efficiency for a given spectral efficiency if enough spatial modes are employed. However, optimal quantum detection can reach target photon and spectral efficiencies with far fewer spatial modes than would be required with direct detection on-off keying.

The results summarized in the preceding paragraph were for non-fading channels. We

also treated fading bosonic channels, in which the transmissivity is random. In particular, we developed statistical models — exponential and lognormal fading — for ground-to-space and space-to-ground communication through atmospheric turbulence. Using these far-field models we computed their ergodic capacity and their outage capacity. For near-field communication in which the average transmissivity is quite high, the fading distribution is not known. Here we were able to find capacity bounds, based only on average transmissivity, that were very tight.

Our last effort addressed multiple-input, multiple-output communication through atmospheric turbulence. We reviewed the channel model for a sparse-aperture system that had previously been employed for heterodyne-detection analysis. We used its Marchenko-Pastur statistics to estimate its ergodic capacity when optimum quantum reception is employed and we showed that simulated results corroborated our analytical results. Finally, we considered the scaling behavior of the sparse-aperture system. For fixed diameters of the transmitter and receiver constellations, we showed that ergodic capacity increases as the number of receiving apertures is increased until the system is fully packed. However, once the system is fully packed, increasing the number of transmitters or receivers requires increasing the diameters of their respective constellations, and this results in a decreased ergodic capacity.

There remain many areas related to this thesis that can be further explored. First, a proof to the conjectured quantum capacity of a thermal-noise channel has still not been found. Furthermore, better bounds on the ergodic and outage capacities of near-field fading channels can be developed if a better understanding of their statistics is known. Finally, it may be of interest to see what other results from random matrix theory can be applied to finding capacity results for different multi-mode fading channels with intermodal interference. This thesis discussed the sparse-aperture setup, but it is not clear whether random matrix theory can be applied to other multi-mode setups.

Appendix A

Evaluation of Two Limits

To prove (2.19) we need to show that

$$\lim_{\bar{N}\to 0} \frac{H_B\left(\frac{c\bar{N}e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} + \frac{1-c\bar{N}}{\bar{N}+1}\right) - c\bar{N}H_B\left(\frac{e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1}\right) - (1-c\bar{N})H_B\left(\frac{1}{\bar{N}+1}\right)}{-\bar{N}\log\bar{N}} = c - ce^{-\frac{1}{c}}. \quad \text{(A.1)}$$

We expand the numerator using the definition of the binary entropy function:

$$H_B\left(\frac{c\bar{N}e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} + \frac{1-c\bar{N}}{\bar{N}+1}\right) - c\bar{N}H_B\left(\frac{e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1}\right) - (1-c\bar{N})H_B\left(\frac{1}{\bar{N}+1}\right)$$

$$= -\left(\frac{c\bar{N}e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} + \frac{1-c\bar{N}}{\bar{N}+1}\right)\log\left(\frac{c\bar{N}e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} + \frac{1-c\bar{N}}{\bar{N}+1}\right)$$

$$- \left(1 - \frac{c\bar{N}e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} - \frac{1-c\bar{N}}{\bar{N}+1}\right)\log\left(1 - \frac{c\bar{N}e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} - \frac{1-c\bar{N}}{\bar{N}+1}\right)$$

$$+ c\bar{N}\left(\frac{e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1}\right)\log\left(\frac{e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1}\right) + c\bar{N}\left(1 - \frac{e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1}\right)\log\left(1 - \frac{e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1}\right)$$

$$+ (1-c\bar{N})\left(\frac{1}{\bar{N}+1}\right)\log\left(\frac{1}{\bar{N}+1}\right) + (1-c\bar{N})\left(\frac{\bar{N}}{\bar{N}+1}\right)\log\left(\frac{\bar{N}}{\bar{N}+1}\right). \quad \text{(A.2)}$$

We can now employ a simple use of L'Hopital's rule to calculate the required limit on each one of these terms, the tedious details of which are omitted:

91

$$\lim_{\bar{N} \to 0} \frac{-\left(\frac{c\bar{N}e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} + \frac{1-c\bar{N}}{\bar{N}+1} \right) \log \left(\frac{c\bar{N}e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} + \frac{1-c\bar{N}}{\bar{N}+1} \right)}{-\bar{N} \log \left(\bar{N} \right)} = 0 \qquad \text{(A.3)}$$

$$\lim_{\bar{N} \to 0} \frac{-\left(1 - \frac{c\bar{N}e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} - \frac{1-c\bar{N}}{\bar{N}+1} \right) \log \left(1 - \frac{c\bar{N}e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} - \frac{1-c\bar{N}}{\bar{N}+1} \right)}{-\bar{N} \log \left(\bar{N} \right)} = c - ce^{-\frac{1}{c}} + 1 \qquad \text{(A.4)}$$

$$\lim_{\bar{N} \to 0} \frac{c\bar{N} \left(\frac{e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} \right) \log \left(\frac{e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} \right) + c\bar{N} \left(1 - \frac{e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} \right) \log \left(1 - \frac{e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} \right)}{-\bar{N} \log \left(\bar{N} \right)} = 0 \qquad \text{(A.5)}$$

$$\lim_{\bar{N} \to 0} \frac{\left(1 - c\bar{N} \right) \left(\frac{1}{\bar{N}+1} \right) \log \left(\frac{1}{\bar{N}+1} \right) + \left(1 - c\bar{N} \right) \left(\frac{\bar{N}}{\bar{N}+1} \right) \log \left(\frac{\bar{N}}{\bar{N}+1} \right)}{-\bar{N} \log \left(\bar{N} \right)} = -1. \qquad \text{(A.6)}$$

Adding the results above gives us

$$\lim_{\bar{N} \to 0} \frac{H_B \left(\frac{c\bar{N}e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} + \frac{1-c\bar{N}}{\bar{N}+1} \right) - c\bar{N} H_B \left(\frac{e^{\frac{-1}{c(\bar{N}+1)}}}{\bar{N}+1} \right) - (1 - c\bar{N}) H_B \left(\frac{1}{\bar{N}+1} \right)}{-\bar{N} \log \bar{N}} = c - ce^{-\frac{1}{c}}. \quad \text{(A.7)}$$

To prove (2.27), we need to show that:

$$\lim_{\bar{N} \to 0} \frac{H_B \left(\frac{c\bar{N}e^{\frac{-1}{c(k\bar{N}+1)}}}{k\bar{N}+1} + \frac{1-c\bar{N}}{k\bar{N}+1} \right) - c\bar{N} H_B \left(\frac{e^{\frac{-1}{c(k\bar{N}+1)}}}{k\bar{N}+1} \right) - (1 - c\bar{N}) H_B \left(\frac{1}{k\bar{N}+1} \right)}{-\bar{N} \log \bar{N}} = c - ce^{-\frac{1}{c}}.$$

$$\text{(A.8)}$$

This can be accomplished by expanding the numerator as we did in the previous paragraph

and then calculate each limit using L'Hopital's rule. The details of this calculation are omitted.

Appendix B

Laguerre Statistics Derivation

We would like to calculate the statistics of the photon-number measurement on a thermal-noise channel, i.e., a channel whose input-output relationship is given by

$$\hat{a}' = \sqrt{\eta}\hat{a} + \sqrt{1 - \eta}\hat{b}, \tag{B.1}$$

where $\hat{a}$, $\hat{b}$, and $\hat{a}'$ are modal photon annihilation operators and $0 \leqslant \eta \leqslant 1$ is the channel transmissivity. Here, $\hat{a}$ is the input mode whose information-bearing state is controlled by the transmitter, $\hat{b}$ is a mode injected by the channel, and $\hat{a}'$ is the output mode from which the receiver will attempt to retrieve the transmitted message. For the thermal-noise channel the $\hat{b}$ mode is in a thermal state, i.e., a Gaussian mixture of coherent states with an average photon number $N_b > 0$ whose density operator is therefore

$$\hat{\rho}_b = \int d^2\beta \frac{\exp\left(-|\beta|^2/N_b\right)}{\pi N_b} |\beta\rangle \langle\beta|. \tag{B.2}$$

We let a coherent state $|\alpha_s/\eta\rangle$ be input into this channel with eigenvalue α_s/η, so that it is equivalent to a coherent state with eigenvalue α_s being received. Let $a_s = \alpha_s + n_s$ where n_s is a complex-valued Gaussian random variable with statistically independent, zero-mean real and imaginary parts each having variance $N/2$. This corresponds to the thermal-noise channel described above in which n_s is the received noise injected into the channel by the

thermal mode. The statistics of a direct detection measurement given a particular value of a_s is

$$\Pr\left(N_s = n | a_s\right) = \frac{|a_s|^{2n}}{n!} e^{-|a_s|^2}. \tag{B.3}$$

because given a_s, we have that the $\hat{a}'$ mode is in the coherent state $|a_s\rangle$. We can use this (and the fact that the statistics of n_s are Gaussian distributed) to calculate the unconditional direct-detection measurement statistics as follows:

$$\Pr\left(N = n\right) = \int d^2 a_s \Pr\left(N_s = n | a_s\right) p\left(a_s\right) \tag{B.4}$$

$$= \int d^2 a_s \frac{|a_s|^{2n}}{n!} e^{-|a_s|^2} \frac{e^{-|a_s - \alpha_s|^2 / N}}{\pi N} \tag{B.5}$$

$$= \frac{e^{-\alpha_s^2 / N}}{n! N} \int_0^\infty dr 2 r^{2n+1} e^{-r^2(N+1)/N} \int_0^{2\pi} d\phi \frac{e^{2\alpha_s r \cos(\phi)/N}}{2\pi} \tag{B.6}$$

$$= \frac{e^{-\alpha_s^2 / N}}{n! N} \int_0^\infty dr 2 r^{2n+1} e^{-r^2(N+1)/N} I_0\left(2\alpha_s r / N\right) \tag{B.7}$$

$$= \frac{N^n}{(N+1)^{n+1}} e^{-a_s^2 / (N+1)} L_n\left(-\frac{a_s^2}{N(N+1)}\right). \tag{B.8}$$

In this computation, we used the fact that

$$\int_0^{2\pi} d\phi \frac{e^{\beta \cos(\phi)}}{2\pi} = I_0\left(2\beta\right) \tag{B.9}$$

where $I_0\left(2\beta\right)$ is a zeroth order modified Bessel function of the first kind, and

$$L_n\left(x\right) = \sum_{m=0}^{n} (-1)^m \begin{pmatrix} n \\ n-m \end{pmatrix} \frac{x^m}{m!} \tag{B.10}$$

is the nth order Laguerre polynomial.

Bibliography

[1] J. H. Shapiro, S. Guha, and B. I. Erkmen. Ultimate channel capacity of free-space optical communications. *J. Opt. Netw.*, 4(8):501-516, 2005.

[2] A. Lapidoth and S. M. Moser. On the capacity of discrete-time Poisson channel. *IEEE Trans. Inf. Theor.*, 55(1):303-322, 2009.

[3] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379-423, 623-656, 1948.

[4] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland and W. K. Wootters. Classical information capacity of a quantum channel. *Phys. Rev. A*, 54(3):1869-1876, 1996.

[5] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H Shapiro and H. P. Yuen. Classical capacity of the lossy bosonic channel: the exact solution. *Phys. Rev. Lett.*, 92(2):027902, 2004.

[6] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, and J. H. Shapiro. Minimum output entropy of bosonic channels: A conjecture. *Phys. Rev. A*, 70(3):032315, 2004.

[7] J.H. Shapiro, "Imaging and Optical Communication through Atmospheric Turbulence," Chapter 6 in J.W. Strohbehn (Ed.), Laser Beam Propagation in the Atmosphere (Springer Verlag, Berlin, 1978).

[8] N. S. Kopeika and J. Bordogna. Background noise in optical communication systems. *Proc. IEEE 58*, 1571-1577, 1970.

[9] I. E. Teletar. Capacity of multi-antenna gaussian channels. *European Trans. on Telecomm.*, 10(6):585-595, 1999.

[10] J. H. Shapiro and R. C. Harney. Simple algorithms for calculating optical communication performance through turbulence. *Proc. SPIE*, 295:41-54, 1981.

[11] A. L. Puryear and V. W. S Chan. Optical communication through the turbulent atmosphere with transmitter and receiver diversity, wavefront control, and coherent detection. *Proc SPIE*, 7464, 74640J, 2009.

[12] Z. Bai, B. Miao, and J. Yao. Convergence rates of spectral distributions of large sample covariance matrices. *SIAM J. Matrix Anal. Appl.* 25, 2003.

[13] J. H. Shapiro. Quantum Optical Communication. Fall 2008. (MIT OpenCourseWare: Massachusetts Institute of Technology). http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-453-quantum-optical-communication-fall-2008/lecture-notes/. (Accessed August 8, 2011). License: Creative commons BY-NC-SA.

[14] X. Zhu and J. M. Kahn, Free-space optical communication through atmospheric turbulence channels. *IEEE Trans. Commun.*, 50(8):1293-1300, August 2002.

[15] C. C. Gerry and P. L. Knight. Introductory Quantum Optics. Cambridge University Press, Cambridge, 2005.

[16] S. M. Haas. Capacity of and coding for multiple-aperture wireless optical communications. Ph.D. Thesis (Massachusetts Institute of Technology, 2003).

[17] T. M. Cover and J. A. Thomas. Elements of Information Theory. John Wiley and Sons, 1991.