Title

Analysis and Mitigation of Tropospheric Effects on Ka Band Satellite Signals and Estimation of Ergodic Capacity and Outage Probability for Terrestrial Links

Author

Enserink, Scott Warren

Publication Date

2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Analysis and Mitigation of Tropospheric Effects on Ka Band Satellite Signals and Estimation of Ergodic Capacity and Outage Probability for Terrestrial Links

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Electrical Engineering

by

Scott Warren Enserink

2012

ABSTRACT OF THE DISSERTATION

Analysis and Mitigation of Tropospheric Effects on Ka Band Satellite Signals and Estimation of Ergodic Capacity and Outage Probability for Terrestrial Links

by

Scott Warren Enserink

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2012

Professor Michael Paul Fitz, Chair

The first part of this work covers the effect of the troposphere on Ka band (20-30 GHz) satellite signals. The second part deals with the estimation of the capacity and outage probability for terrestrial links when constrained to quadrature amplitude modulations.

The desire for higher data rates and the need for available bandwidth has pushed satellite communications into the Ka band (20-30 GHz). At these higher carrier frequencies the effects of scintillation and rain attenuation are increased. In regards to the effects of scintillation, the first part of this work quantifies, through the use of a multiple phase screen simulation model, the benefits of using two receive antennas to mitigate tropospheric-induced scintillation on Ka band satellite downlinks. Two representative turbulence profiles are considered, and cumulative distribution curves for scintillation-induced attenuation are generated for selection and maximal ratio combining schemes and compared to those for a single antenna. The results indicate that there can be significant diversity gains achieved by combining two antennas separated by only a short distance. Also, a comparison of simulation results with the results predicted by the *basic Rytov* approximation shows that at elevation angles greater than 10 degrees, Rytov theory can accurately predict performance benefits of antenna combining, but at elevation angles less than 10 degrees it is better to use multiple phase screen simulations to make performance predictions. In addition, the effects of scintillation-induced phase perturbations on the output power of large

aperture antennas is examined. It is found that the output power degradation due to scintillation-induced phase perturbations is generally negligible and can be countered by the simple means of antenna tracking if necessary.

In regards to rain attenuation, this work developed simple methods for estimating the outage probability and outage capacity and ergodic capacity of satellite links due to rain fades. The rain-induced fades of a satellite link are often modeled with a log-log-normal distribution. Researchers have determined methods for calculating the outage probability for Shannon capacity for log-log-normal channels. However, in practical communications systems, the input signal is constrained to a discrete signalling set such as finite-size quadrature amplitude modulations. Under these conditions the outage probability with regards to the constrained capacity is a more accurate measure. A method is detailed in this work for tightly estimating the outage probability and outage capacity of satellite links with quadrature amplitude modulations. In addition this work derives a lower bound for the ergodic constrained capacity of log-log-normal channels. To date, no other method for calculating the outage probability, outage capacity, or a lower bound for the ergodic capacity for a log-log-normal channel with a finite-size quadrature amplitude modulation has been published. Also, this portion of the work quantifies the benefit of using receive diversity to mitigate rain fades, providing the gains in outage capacity due to the use of diversity for a tropical region and a fairly dry region under the constraint that practical constellations are transmitted. The above information and analysis methods provide useful tools for satellite system planners.

The second part of this work examines terrestrial communication links, which can suffer greatly from channel fading or shadowing. Two common statistical models for channels are the Rayleigh distribution and the log-normal distribution. The goal of this second part of the work was to develop a simple method for tightly estimating the ergodic capacity and outage probability of these two channel types when used with quadrature amplitude modulated signalling sets. Specifically an innovative method was developed for estimating the ergodic constrained capacity for Rayleigh and log-normal channels with and without antenna combining. The expressions facilitate straightforward computation of outage probability as well. Researchers have deter-

mined methods for calculating the ergodic Shannon capacity for log-normal and Rayleigh channels for single and multiple receive antenna systems. However, in practical communications systems, the input signal is constrained to a discrete signalling set such as finite-size quadrature amplitude modulation constellations. Under these conditions the ergodic constrained capacity is a more accurate measure. The method detailed in this work provides a uniform expression for computing the ergodic capacity, both Shannon and constrained, of Rayleigh and log-normal channels with and without antenna combining. The expressions facilitate straightforward computation of outage probability as well. Both the noise-limited and interference-limited cases are studied. To date, no other method for estimating the outage probabilities for the constrained capacity of Rayleigh or log-normal channels has been published for either the noise-limited case or interference-limited case. Also, no method for estimating the ergodic constrained capacity of a log-normal channel or of an interference-limited Rayleigh channel has appeared in the literature. The analysis methods and information for terrestrial links developed in the second part of this work provide useful tools for the designers of wireless communication systems in general and have particular application to cellular mobile and ultra-wideband systems.

The dissertation of Scott Warren Enserink is approved.

Kuo-Nan Liou

Richard D. Wesel

Yahya Rahmat-Samii

Michael Paul Fitz, Committee Chair

University of California, Los Angeles

2012

To my father, Egbert Enserink . . .

a man of science, perseverance, wisdom and compassion.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank Northrop Grumman for their support of my academic work with a fellowship, Tom Farrell for his consultation in implementing the multiple phase screen simulation model, Athanasios D. Panagopoulos for his review of aspects of my work on satellite links, and Jared Dulmage for his review of and insightful comments on all of my material. I also wish to thank all my colleagues in the UnWiReD Laboratory at UCLA for their encouragement and insightful discussions. These persons are: Daniel Liu, Michael Samuel, Jared Dulmage, David Browne, Cong Shen, Maged Barsoum, Weijun Zhu and Heechoon Lee.

VITA

1987	B.S. (Electrical Engineering), Calvin College, Grand Rapids, Michigan.
1987–1988	Graduate studies in electro-magnetics, UCLA, Los Angeles, California.
1988-1991	Radio engineer, Radio Lumiere, Port–au–Prince, Haiti
1994	M.S. (Electrical Engineering), Arizona State University (ASU), Tempe, Arizona.
1995	Graduate studies in wireless communications, ASU, Tempe, Arizona.
1995-1998	Communications systems engineer, Government Systems Technology Group, Motorola, Scottsdale, Arizona.
1998-1999	Communications systems engineer, Sicom Inc., Scottsdale, Arizona.
1999-2000	Communications systems engineer, Government Systems Technology Group, Motorola, Scottsdale, Arizona.
2000-2001	Communications systems engineer, Intersil Inc., Scottsdale, Arizona.
2001–present	Communications systems engineer, Northrop Grumman Aerospace Systems, Redondo Beach, California.

PUBLICATIONS AND PRESENTATIONS

S. Enserink and M.P Fitz, "Constrained Capacities of DVB-S2 Constellations in Log-Normal Channels at Ka Band." In *Advances in Satellite and Space Communications (SPACOMM), 2010*

Second International Conference on, pp.93-99, Jun., 2010.

S. Enserink and M.P Fitz, "Constrained Capacities in Log-Normal Channels with Site Diversity at Ka Band." In *Advances in Satellite and Space Communications (SPACOMM), 2010 Second International Conference on*, pp.100-105, Jun., 2010.

S. Enserink and M.P Fitz, "Quantification of the Performance Benefits of Receiver Diversity in Mitigating Scintillation on Ka Band Downlinks." In *Proceedings of the General Assembly of the International Union of Radio Scientists (URSI) General Assembly 2008*, session F04, Aug., 2008. (http://www.ursi.org/proceedings/procGA08/papers/ F04p4.pdf)

S. Enserink and M.P Fitz, "Use of Antenna Receive Diversity to Mitigate Scintillation on Ka band Satellite Links." In *Antennas and Propagation Society International Symposium, 2008. AP-S 2008*. Jan., 2008, pp.1-4.

S. Enserink and J. Liebetreu, "System Impairment Model: Power Amplifier Model," Presentation to IEEE 802.16 Broadband Wireless Access Working Group, Doc. Number 802.16.1pp-00/15, Aug. 2000.

S. Enserink and D. Cochran, "On Detection of Cyclostationary Signals," In *Acoustics, Speech, and Signal Processing, ICASSP-95, 1995 International Conference on* , vol. 3, pp.2004-2007.

S. Enserink and D. Cochran, "A Cyclostationary Feature Detector," In *Signals, Systems and Computers, 1994 Conference Record of the Twenty-Eighth Asilomar Conference on*, vol.2, pp.806-810.

Part I

Analysis and Mitigation of Tropospheric Effects on Ka Band Satellite Signals

CHAPTER 1

Analysis and Mitigation of Tropospheric Scintillation

Contributions:

- Quantified the benefits of using receiver diversity to mitigate tropospheric scintillation of Ka band satellite signals.
- Quantified the power output degradation of large antennas due to scintillation-induced phase perturbations of Ka band satellite signals.
- Determined guidelines for when analytic approximations can be used to accurately estimate the diversity gain obtained by two receive antennas, and when simulation modeling is necessary.
- Re-introduced the multiple phase screen method to characterize scintillation effects on satellite communications.

1.1 Introduction

The desire for higher data rates and the need for available bandwidth has pushed satellite communications into the Ka band (20-30 GHz). At these higher carrier frequencies the effects of scintillation due to turbulence-induced index of refraction irregularities in the troposphere is known to become more prevalent, particularly on slant paths. Mitigation of scintillation can be very important for certain links such as those operating at a low-margin along a low elevation path in a dry climate, where scintillation, rather than rain, is the key link factor. One of the goals of this work was to quantify the benefit of using two receive antennas to mitigate the

tropospheric-induced amplitude scintillation of a satellite signal in the Ka band. Another of the goals of this work was to characterize the effect of tropospheric-induced phase perturbations on the output power of large aperture antennas. A secondary goal was to establish some guidelines to determine in what situations *basic Rytov* theory, which is applicable in weak scattering conditions, can be used and in what situations simulations should be used in determining possible antenna combining benefits. In this work the multiple phase screen (MPS) model was used to simulate the effect of the troposphere on the Ka band satellite signal in order to study both amplitude and phase scintillation. In years past the MPS method has been used to model the effects of the ionosphere on the propagation of satellite signals at less than 10 GHz [Kne83]. In recent years it has been used extensively by the free-space optical communications community to model the effects of the troposphere on laser signals, and by one other researcher (other than the author of this work) to study tropospheric-induced amplitude scintillation of Ka band signals [VGW05]. Various spectrums for the index of refraction irregularities were used in conjunction with the MPS modeling to determine the effect of outer scale size on the tropospheric-induced amplitude and phase scintillation.

1.2 Background

This section presents some results derived from scintillation theory. The interested reader may wish to consult the classic text by Tatarski [Tat61] or the more recent text by Wheelon [Whe03] for more details.

1.2.1 The Origins Of Scintillation: Turbulence

Air has an index of refraction near 1, however, the index of refraction, at radio frequencies, is a function of temperature, pressure and water vapor density. Mixing of the air in the troposphere, caused by heating and cooling, creates wind eddies. Large eddies break apart into smaller and smaller eddies until finally the energy of the smallest eddies dissipates into the heat energy of random molecule movement. Different eddies can have slightly different values for temperature,

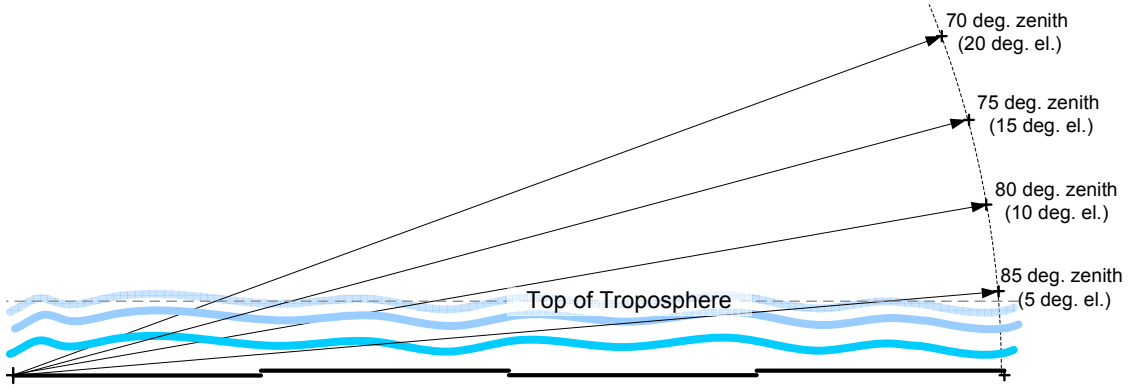


Figure 1.1: As the zenith angle approaches 90 degrees (elevation angle approaches 0 degrees) the distance traveled by the signal through turbulence greatly increases.

pressure and water vapor density. This creates inhomogeneities in these quantities, and thus in the index of refraction, or equivalently, in the dielectric constant ε [Tys00]. The dielectric constant will be written as $\varepsilon = 1 + \triangle\varepsilon$ to reflect the fact that it has a value that is nearly 1, but has minor random fluctuations about that value. The eddy-induced slight spatial variations in the index of refraction causes some parts of a propagating wave to travel more quickly or slowly then other parts. These phases changes lead to amplitude changes as constructive and destructive interference occurs as the wave progresses. This results in phase and amplitude variations at the receiver. This effect is referred to as scintillation. Since they pass through more of the turbulent troposphere, signals traveling along slant paths, defined as having a low elevation angle or equivalently a large zenith angle, have larger amounts of scintillation than do signals traveling along higher elevation path as shown in Fig. 1.1.

1.2.2 Spatial Correlation and Structure Function

Since the dielectric constant irregularities $\varepsilon(\mathbf{r},t)$ form a random process with a spatial correlation defined by

$$< \triangle\varepsilon(\mathbf{r},t), \triangle\varepsilon(\mathbf{r}',t) > \tag{1.1}$$

where, $< . >$ indicates an ensemble average.

A related quantity, one that cancels out terms common at both location $\mathbf{r}$ and at $\mathbf{r}'$ is the

structure function, defined as,

$$D_\varepsilon(\mathbf{r},\mathbf{r}') \doteq < [\triangle\varepsilon(\mathbf{r},t) - \triangle\varepsilon(\mathbf{r}',t)]^2 >$$

Assuming a homogeneous medium this simplifies to,

$$D_\varepsilon(\mathbf{r},\mathbf{r}') = 2[< \triangle\varepsilon >^2 - < \triangle\varepsilon(\mathbf{r},t), \triangle\varepsilon(\mathbf{r}',t) >] \tag{1.2}$$

Developments by Kolmogorov led to a theoretical expression for the above structure function. Working with wind velocity measurements he reasoned that the structure function for the velocity component parallel to the separation vector ρ must be a function of the rate of energy Υ going into the creation of eddies (and eventually being dissipated) and the distance between the two measurement points ρ.

$$D_v(\mathbf{r},\mathbf{r}+\rho) \doteq < [v(\mathbf{r},t) - v(\mathbf{r}+\rho,t)]^2 >= F(\Upsilon,\rho)$$

The energy dissipation rate Υ has units of $m^2\ s^{-3}$. The only combination of ρ and Υ that produces units of a squared velocity is,

$$D_v(\mathbf{r},\mathbf{r}+\rho) = constant(\Upsilon\rho)^{\frac{2}{3}} = C_v^2\rho^{\frac{2}{3}}$$

The 2/3 scaling law has been verified by numerous experiments. It works under a wide range of conditions, from small-scale turbulence near the earth's surface to enormous irregularities in the interstellar plasma. This formula holds only for separation distance larger than the smallest eddy size l_0 and smaller than the largest eddy size L_0. These quantities are referred to as the *inner scale* and *outer scale* respectively. The constant C_v^2 is called the *velocity structure constant*.

After Kolmogorov's development others realized that the same technique could be applied to describe the turbulent mixing of passive scalars, such as the index of refraction n. This led to the following relationship key to the study of the propagation of electromagnetic waves through

the turbulent troposphere.

$$D_n(\mathbf{r},\mathbf{r}+\rho) \doteq E[(n(\mathbf{r})-n(\mathbf{r}+\rho))^2] = C_n^2 \rho^{\frac{2}{3}}$$

The constant C_n^2, called the *refractive index structure constant*, is the key parameter in the analysis of the effects of random media on electromagnetic propagation. Its value is a measure of the strength of the turbulence. Its profile along the propagation path directly influences the variance of the received signal's amplitude and phase. This formula holds only for separation distance $l_0 < \rho < L_0$.

1.2.3 Models for the Spectrum of Index of Refraction Irregularities

With homogeneous medium, the spatial correlation of (1.1) can be expressed in terms of its Fourier wavenumber decomposition, $\Phi_\varepsilon(\kappa)$, analogous to how the temporal correlation of a random process can be expressed as the inverse Fourier transform of the power spectral density.

$$< \triangle\varepsilon(\mathbf{r},t)\triangle\varepsilon(\mathbf{r}',t) > \doteq \int \mathrm{d}^3\kappa \Phi_\varepsilon(\kappa) e^{\kappa\cdot(\mathbf{r}-\mathbf{r}')} \tag{1.3}$$

where the spectrum $\Phi_\varepsilon(\kappa)$ describes the relative energy contained in eddies of size l by their associated wavenumber $\kappa = \frac{2\pi}{l}$. This relationship, or the equivalent one for the index of refraction n, allows analysts to separate the wave-scattering features of the problem (i.e. the geometry of the problem and the type of wave) from the description of the turbulent medium.

Using the index of refraction versions of (1.2) and (1.3) it can be shown that the wavenumber spectrum of index of refraction irregularities will be [Whe01]

$$\Phi_n(\kappa) = 0.033 C_n^2 \kappa^{-\frac{11}{3}} \quad \text{for} \quad 0 < \kappa \tag{1.4}$$

This is the *unrestricted Kolmogorov* spectrum. Strictly speaking this formula only applies for $\kappa_0 \doteq \frac{2\pi}{L_O} < \kappa < \frac{2\pi}{l_O} \doteq \kappa_s$. This is referred to as the inertial range, over which larger eddies

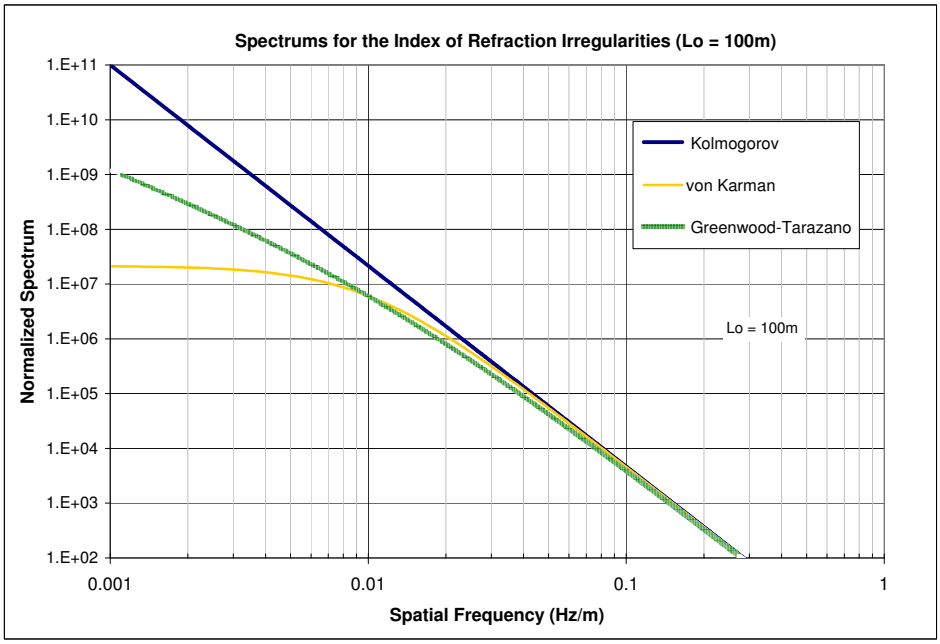


Figure 1.2: These three spectrums for the index of refraction irregularities were used in the simulations, each handles the energy-input region slightly differently.

transfer energy to smaller eddies. The range of $l > L_0$ is the energy-input region in which the translational motion of the wind is beginning to be turned into the rotational motion of large eddies. This is an ambiguous region, the behavior of which is not well known because it is difficult to distinguish between the effects of large eddies and basic changes in the weather. In the unrestricted Kolmogorov model the behavior of the energy-input region is assumed to be the same as that of the inertial range, i.e. it is assumed that the Kolmogorov spectrum can be extended into the energy-input region without modification. This simplifies analytical work, and is a sufficient model in some situations, but provides inaccurate results in other situation.

The following formulas have been used by researchers to attempt to model the energy-input portion of the spectrum. Fig. 1.2 provides graphs of these spectral models.

a.) *Kolmogorov model with Cutoff*: The Kolmogorov model with cutoff simply ignores the energy-input and energy-loss regions.

$$\Phi_n(\kappa) = \begin{cases} 0.033 C_n^2 \kappa^{-\frac{11}{3}} & : \quad \frac{2\pi}{L_O} < \kappa < \frac{2\pi}{l_O} \\ 0 & : \quad otherwise \end{cases} \tag{1.5}$$

In the simulation work, this model was used effectively in that the spatial span covered 1,000 m, thus creating an effective outer scale value of L_0 = 1,000 m.

b.) *von Karman* model: This spectral model is often used in analysis because it is tractable, and has been shown to fairly well agree with actual measurements. It also was used as one of the models in this effort's simulation work with an outer scale value of L_0 = 100 m which is a typical value for space-to-earth paths [Whe01].

$$\Phi_n(\kappa) = 0.033 C_n^2 (\kappa^2 + \kappa_0^2)^{-\frac{11}{6}} \quad \text{for} \quad 0 < \kappa < \kappa_s \tag{1.6}$$

c.) *Greenwood-Tarazano* model: Greenwood and Tarazano developed an expression for the spectrum that better matched the measured spectrums and structure functions then the above spectral expressions. However, the expression is not very tractable, and as result is not used in analytical developments. It also was used as one of the models in this effort's simulation work, again with L_0 = 100 m. This highlights another feature of MPS simulations, they can implement spectrums which are not analytically tractable.

$$\Phi_n(\kappa) = 0.033 C_n^2 (\kappa^2 + \kappa_0 \kappa)^{-\frac{11}{6}} \quad \text{for} \quad 0 < \kappa < \kappa_s \tag{1.7}$$

d.) *Exponential* model: The exponential model lends itself to analysis, but experimental results showed that it does not match measured data as well as the von Karman model. As a result it was not used as a model in this project's simulation work.

$$\Phi_n(\kappa) = 0.033 C_n^2 \kappa^{-\frac{11}{3}} \left[1 - e^{\left(-\frac{\kappa^2}{\kappa_0^2} \right)} \right] \quad \text{for} \quad 0 < \kappa \tag{1.8}$$

1.2.4 Amplitude Scintillation Theoretical Formula

Using the *basic Rytov approximation* (i.e. the approximation based on the first term in the series representation of the surrogate function as described in appendix B), it can be shown that 1) the received signal's amplitude has a log-normal distribution, and thus that the log of that

amplitude, χ, has a Gaussian distribution and that 2) the variance of χ, can be found for a plane wave from [Tat61, Whe03] for the Kolmogorov spectrum to be

$$\sigma_\chi^2 \;=\; 0.563k^{\frac{7}{6}}\csc^{\frac{11}{6}}(\theta)\int C_n^2(h)h^{\frac{5}{6}}\mathrm{d}h\,(\mathrm{Np})^2 \tag{1.9}$$

where, 1) $C_n^2(h)$ describes how C_n^2, the index of refraction structure function constant, varies with altitude h, 2) θ is the elevation angle of the slant path, 3) $k = \frac{2\pi}{\lambda}$ is the wave number and 4) λ the wave length of the signal. The value C_n^2 quantifies the turbulence strength, for this reason C_n^2 is also often simply termed the *turbulence strength*. Fig. 1.3 shows a typical turbulence profile when 1) clouds are present and 2) when clouds are not present. The cloud-present profile was determined from radio-sonde measurements taken in Belgium on a summer day (August 14, 1990) [VV98]. The no-cloud profile was developed in [YWL03]. It is based on the ITU-R C_n^2 profile model, with a modification to take into account water vapor. The version shown assumes a standard summer atmosphere for European latitudes. These two profiles represent what might be expected on a summer day in European latitudes.

The above formula from *basic Rytov* theory was used to calculate, for a given turbulence profile, the anticipated value of σ_χ^2. Which in turn was used along with a log-normal distribution to predict the benefits of using antenna combining. These values, based on the *basic Rytov* approximation, were then compared with the values generated from simulations.

1.2.5 Antenna Configurations

This study considered the following antenna configurations: 1) a single antenna, 2) maximal-ratio-combining (MRC), in which the received signals from two separate antennas are scaled by the conjugate of their channel gains and then added and 3) selection-combining, in which the signal from the antenna with the best signal-to-noise ratio (SNR) is used. It is well known that the MRC method of combining results in the best SNR, but that selection combining is less complicated to implement. Let $|a_i|$ be the magnitude of the signal at antenna i, then the relative output SNRs can be expressed as, 1) $\mathrm{SNR}_{\mathrm{single}} = |a_1|^2$, 2) $\mathrm{SNR}_{\mathrm{MRC}} = (|a_1|^2 + |a_2|^2)$ and 3)

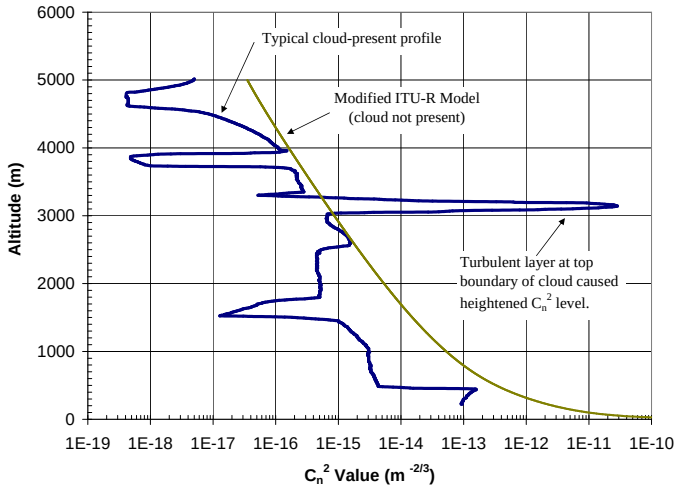


Figure 1.3: Two typical C_n^2 profiles during summer in European latitudes.

$\text{SNR}_{\text{SC}} = max(|a_1|^2, |a_2|^2)$.

1.3 Multiple Phase Screen Model

The MPS method was used to simulate the effects of the troposphere in creating scintillation. This section describes the MPS method. Additional information on the MPS method can be found in [Kne83], which deals with propagation through the ionosphere, but the principles are the same for the troposphere. In general, in modeling propagation through a random medium, MPS models have been found to give good agreement with theory and experimental measurements for both weak and strong scattering conditions (see [Kne83] and the references within).

The MPS technique models the propagation of an electro-magnetic (E-M) wave through the turbulent atmosphere by dividing the path into sections and then modeling each section as a 1) thin phase screen followed by a 2) propagator. Each phase screen models the relative delay of portions of the wavefront due to the variations in the index of refraction in the volume associated with that phase screen, as if there were no diffraction. The phase delays of each phase screen are selected with a probability distribution based on turbulence theory. Each propagator, in turn, models the E-M signal propagating through the section of the path as if there were no variations in the index of refraction. The MPS model and its relation to the actual physical model are

shown in Fig. 1.4.

In summary the MPS method consists of the following two parts:

1. Phase screen: At each stage each spatial sample of the incoming E-M signal is multiplied in the spatial domain by the value of the associated grid point of the randomly generated phase screen. This is done in order to cause a phase change between any two points on the wavefront equal to that which would be caused by the delta in the time delay between the two associated paths due to random differences in the index of refraction along the paths. The generation of the phase screens is described in appendix A.1.

2. Propagation: The resulting output is propagated to the next stage by convolving it with the propagation filter. The convolution is actually carried out by a multiplication in the spatial frequency domain, using equation (A.2), as described in appendix A.2, as the transfer function for the effective propagation filter.

The MPS model parameters such as the spacing of grid points, Δ_{grid}, the distance between the application of phase screens, Δz_{screen}, and the size of a $N \times N$ grid, must all be carefully selected to ensure that the turbulence effects are adequately modeled, while still allowing for a feasible implementation. In this work these parameters were determined in a three step process. First, estimates, $\Delta_{grid,e}$, $\Delta z_{screen,e}$, and N_e , of these simulation parameters were determined based on a calculation of the maximum angular change of any portion of the E-M wave produced by any phase screen. Second, simulations were performed with these MPS model parameters. Third, additional simulations were run with a higher model resolution (i.e. grid spacing less than $\Delta_{grid,e}$, screens closer than $\Delta z_{screen,e}$, more grid points than N_e, and total grid span larger than $N_e \Delta_{grid,e}$) in order to see if the results produced by the simulations would change. The resolution of the MPS model was increased until the results no longer changed. Generally the values estimated for the MPS model parameters in the first step were sufficient.

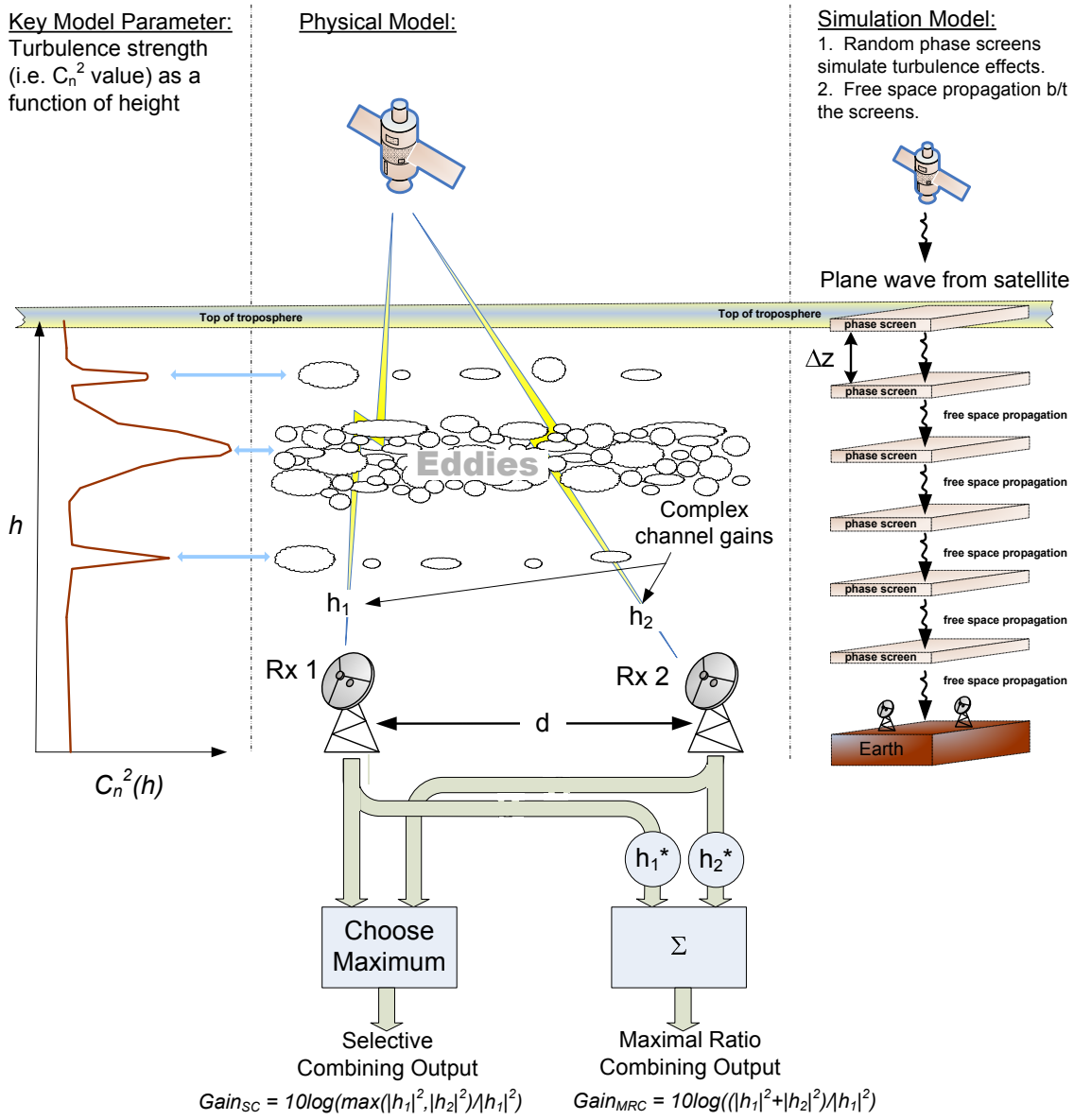


Figure 1.4: Illustration of the modeling approach showing: 1) the turbulence profile (left), 2) representation of physical channel (middle), and 3) the associated multiple phase screen model.

1.4 MPS Simulation Results

The MPS method was used to simulate the effects of the troposphere in creating amplitude and phase scintillation. In order to quantify the benefit of using antenna diversity to mitigate amplitude scintillation, simulations were run with frequencies of f_c= 20 GHz and f_c= 30 GHz assuming the presence of two receive antennas separated by various distances and using each

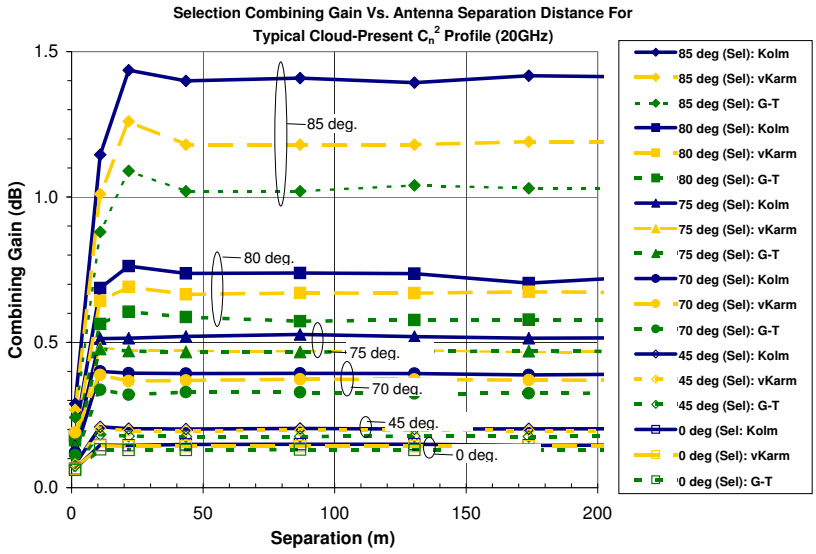
of the following spectrums for the index of refraction irregularities: 1) restricted Kolmogorov with L_0= 1000 m , 2) von Karman, with L_0= 100 m, and 3) Greenwood-Tarazano also with L_0= 100 m. The antennas were assumed to act like point receivers. (Simulations run assuming an antenna dish diameter of 3.4 m yielded almost identical results). Both 1) the typical cloud-present and 2) the cloud-not-present (i.e. the modified ITU-R model) turbulence profiles shown in Fig. 1.3 were used in the simulations. In order to determine the degradation of an antenna's output power caused by phase scintillation and determine the effect of the value of L_0 on that degradation, simulations were run with frequencies of f_c= 20 GHz and f_c= 30 GHz assuming various diameter apertures for the receive antenna using each of the aforementioned spectrums for the index of refraction irregularities. Simulation parameters of Δ_{grid}= 64λ for the grid spacing, Δz_{screen}= 100 meters for the distance between phase screens and $N \times N = 1024 \times 1024$ for the grid size were selected based on the method described in section 1.3.

1.4.1 Antenna Separation

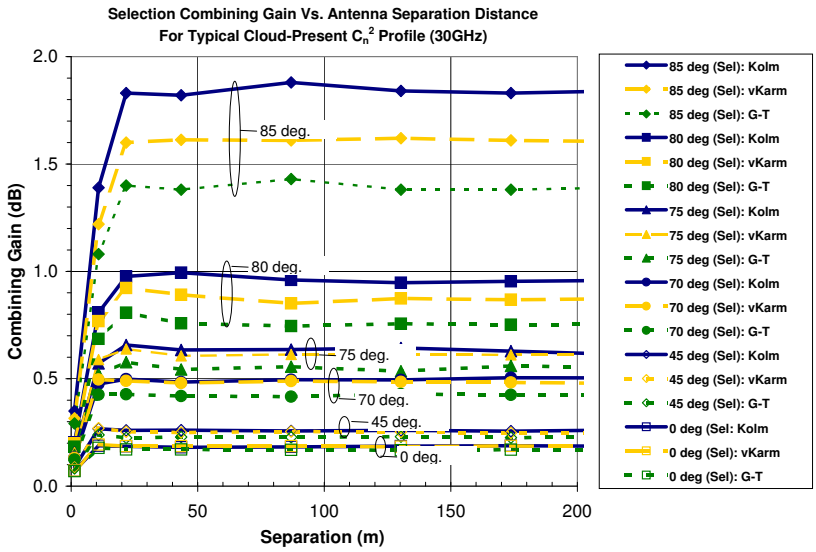
Data portrayed in Fig. 1.5 for 20 GHz and 30 GHz shows the average gain of selection combining, relative to a single antenna, in mitigating scintillation. From analyzing these diversity gain values for various antenna spacings it was determined that a separation of only 20 m is required to obtain a low correlation between the fades at the two antennas for any path elevation and Ka band frequency. Increasing the spacing beyond 20 m will lead to no additional gain. The determination of the required spacing is an important feature of MPS modeling in that it can be quite difficult to obtain by theoretical analysis for a general turbulence profile [Whe03, ch. 4.2].

1.4.2 Diversity Gain

The simulation results showed that the use of two independent antennas can provide substantial diversity benefits, particularly for slant paths of less than 20 degrees of elevation. For the typical cloud-present turbulence profile, Fig. 1.6 portrays the cumulative distribution function (CDF) curves, determined by simulation, for the scintillation-induced attenuation at 30 GHz for 1)



(a) 20 GHz



(b) 30 GHz

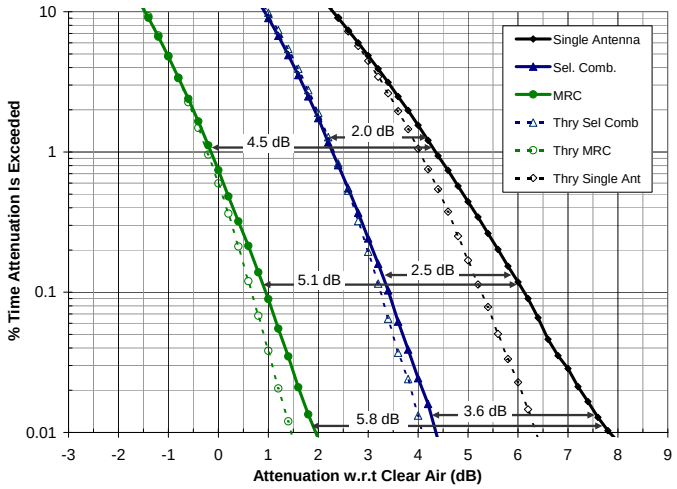
Figure 1.5: Selective combining gain versus antenna separation distances for 20 and 30 GHz for typical cloud-present C_n^2 profile with Kolmogorov, von Karman and Greenwood-Tarazano spectrums.

a single antenna, 2) selection combining, and 3) MRC combining for elevation angles of a) 10 and b) 5 degrees. The CDF curves generated using the *basic Rytov* theory results are also shown. The 0 dB level represents the attenuation level for a single antenna if no scintillation is present, in other words these attenuation values are with respect to clear air. The results shown are for the von Karman spectrum, with a realistic outer scale length of 100 m. The results

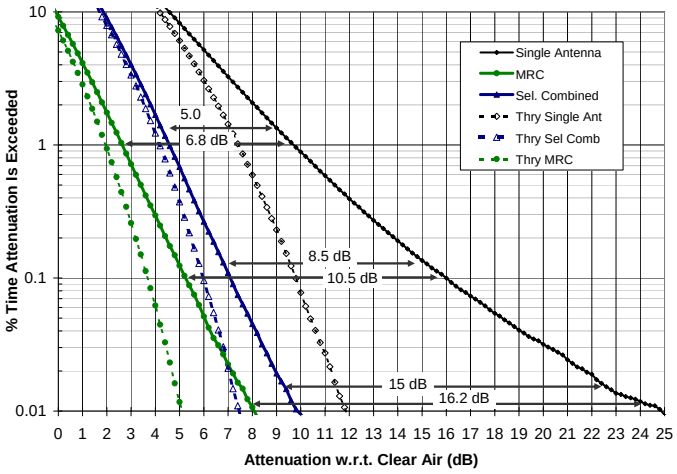
for the Kolmogorov spectrum were slightly more dramatic, but probably not as realistic since the Kolmogorov spectrum does not limit the outer scale length to a realistic value [Whe01]. These results show that there can be significant advantage to using receiver diversity to mitigate scintillation.

For an elevation angle of 10 degrees (Fig.1.6(a)), in this typical cloud-present example, the results show gains ranging from 2.0 dB (3.6 dB) for selection combining and 4.5 dB (5.8 dB) for MRC combining at a desired link availability of 99% (99.99%). Also, the *basic Rytov* theory accurately predicted the CDF curves for probability values greater than 1% (i.e. link availabilities less than 99%). At probabilities less than 1% the values obtained using the *basic Rytov* approximation produced values for attenuation that were too optimistic for the single antenna case. These single-antenna theoretical values would result in values that were 1 to 2 dB too low in estimating the mitigation benefit of antenna combining.

For an elevation angle of 5 degrees (Fig.1.6(b)), in this typical cloud-present example, the results show large gains ranging from 5.0 dB (15.0 dB) for selection combining and 6.8 dB (16.2 dB) for MRC combining at a desired link availability of 99% (99.99%). At this elevation, the scintillation is so strong that the *basic Rytov* theory fails to accurately predicted the CDF curves for any of the antenna configurations, showing that simulations need to be used to obtain accurate estimates. Although equation (1.9) from basic Rytov theory produced a value of σ_χ^2 in close agreement with that found by MPS simulation (0.13 compared to 0.12) for the single antenna case, the CDF curves were very different, indicating a substantial difference in amplitude distributions. In contrast for the no-cloud case the theoretical predictions were close to the results produced by MPS simulations (e.g. just 0.2 dB too low for the single antenna case at the 99.9 %availability level). As mentioned above, *basic Rytov* theory predicts that the received amplitude will have a log-normal distribution. Under weak scattering conditions this does hold true. However, in actuality, as the scintillation increases that distribution will become skewed towards lower amplitude values. The *second order Rytov* solution can predict this skew up to a point [Whe03, ch. 10], but the amount of skew is difficult to predict analytically and the analysis holds over a limited range, yet this skew can greatly influences the link margin values



(a) 10 deg.



(b) 5 deg.

Figure 1.6: CDFs from simulation and from Rytov theory for attenuation due to scintillation for cloud-present turbulence profile for a) 10 degree and b) 5 degree elevations.

(as seen in Fig.1.6(b)), which makes the MPS simulation method valuable when accurate values are needed.

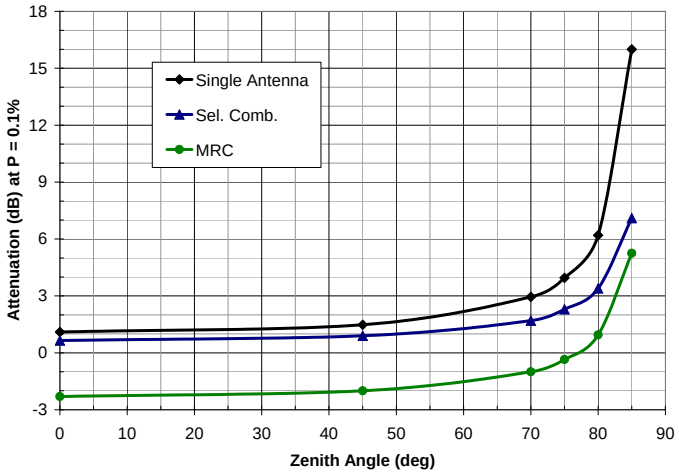
Fig. 1.7 portrays the attenuation values that were exceeded 0.1% of the time as a function of elevation angle for both the a) cloud-present and b) no-cloud-present profiles. In other words, these are the link margin values that would be required in order to obtain a link availability of 99.9% under the type of turbulent conditions pictured in Fig. 1.3. As the elevation angle decreases (zenith angle increases) the signal traverses more of the troposphere, and the scintillation-induced attenuation becomes more pronounced, drastically increasing, particularly

for the cloud-present profile, as the elevation angle decreases below 20 degrees (increases above 70 degrees zenith). At a given elevation angle, the vertical distance from the single antenna curve to the curve for either selection combining or MRC provides the value for combining gain that could be expected to be obtained under this type of turbulence profile when receiving a signal at that elevation angle. As expected, the no-cloud-present turbulence profile produces significantly less scintillation, which is evidenced by the much lower attenuation levels of Fig. 1.7(b) relative to those of Fig. 1.7(a). For the no-cloud-present case, it should be noted that the lower altitude portion of the profile contributes significantly to the scintillation and that this is a section of the model that is open to debate. For an elevation angle of 20 degrees (70 degree zenith) the results in Fig. 1.7 show a gain of 1.2 dB for selection combining and 4 dB for MRC for the cloud-present case and 0.5 dB for selection combining and 3.2 for MRC for the no-cloud-present case. For an elevation angle of 5 degrees (85 degree zenith) the results in Fig. 1.7b show a gain 2 dB for selection combining and 4.6 for MRC for the no-cloud-present case.

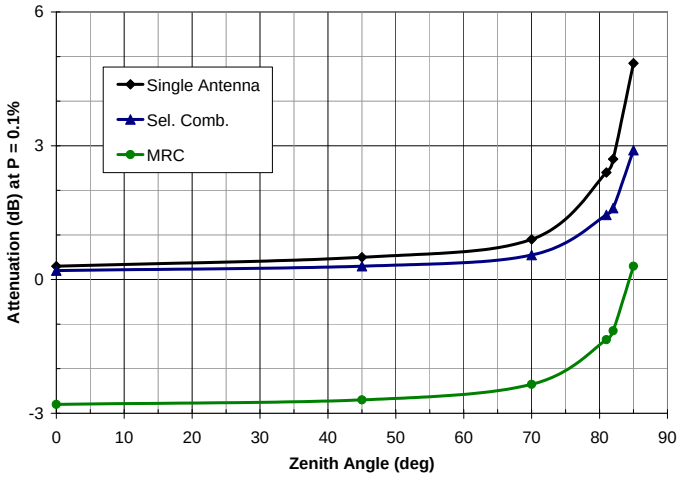
1.4.3 Scintillation-Induced Phase Fluctuations and Antenna Output Power

This part of the study quantified the possible benefits of applying some techniques from adaptive optics (AO) to the Ka band. Namely, the benefits that antenna tracking provides, which is analogous to tilting the mirror in an AO system, and the possible use of a changeable sub-reflector for large diameter antennas, which is analogous to the use of a deformable mirror in an AO system.

In the simulation the effect of antenna tracking, i.e. of pointing the antenna in the best possible direction, was implemented by 1) determining the least squares linear (i.e. planar) fit to the phase front over the aperture and then 2) removing this estimate from the phase front prior to combining. The least squares planar fit also determines the angle of arrival of the signal. In the reported results the angle of arrival (AOA) is defined as the angle by which the axis of the dish antenna would have to change from the nominal direction in order to best receive the incoming wavefront. This quantity will be called the *bore-sight* angle. Fig. 1.8 presents an example phase front generated by MPS simulations. Fig. 1.9 shows the least squares planar fit



(a) Results with cloud-present

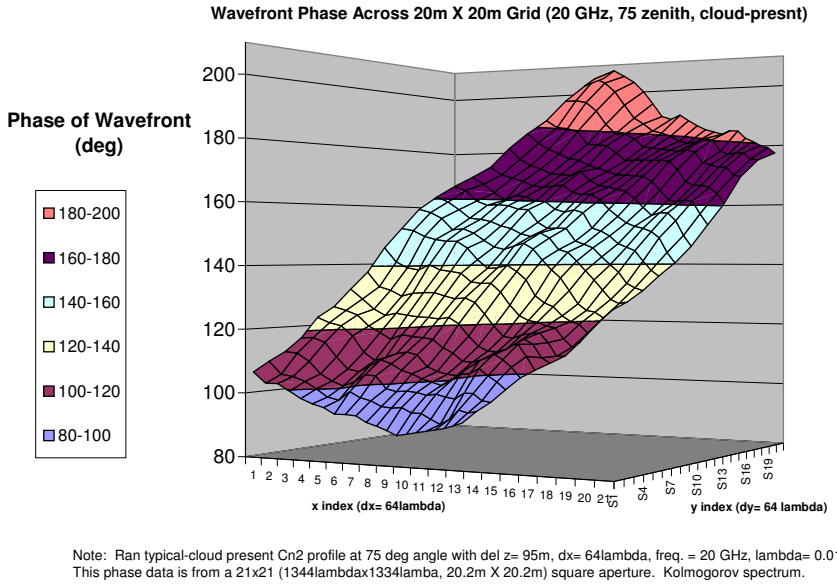


(b) Results with clouds not-present

Figure 1.7: Attenuation values exceeded 0.1% of the time with a 30 GHz signal for various zenith angles for the two turbulence profiles shown in Fig. 1.3: a) cloud-present and b) cloud-not-present (modified ITU-R model).

to the example phase front. The typically small residual error is shown in Fig. 1.10.

Antennas of the following diameters were considered: 1) 3.2 m, 2) 6 m, 3) 12 m, 4) 25 m, and 70 m. These antenna diameters were chosen because antennas of this size are used in actual satellite communication systems. Jet Propulsion Labs (JPL) is considering building a large antenna array using 6 m and 12 m antennas for its Deep Space Network (DSN). The proposed array will cover elevation angles down to 10 degrees. The Westerbork telescope in the Netherlands and the Expanded Very Large Array (EVLA) in New Mexico use arrays of 25 m



Note: Ran typical-cloud present Cn2 profile at 75 deg angle with del z= 95m, dx= 64lambda, freq. = 20 GHz, lambda= 0.015
This phase data is from a 21x21 (1344lambdax1334lamba, 20.2m X 20.2m) square aperture. Kolmogorov spectrum.

Figure 1.8: Example of the phase of the received signal over 20m by 20m grid.

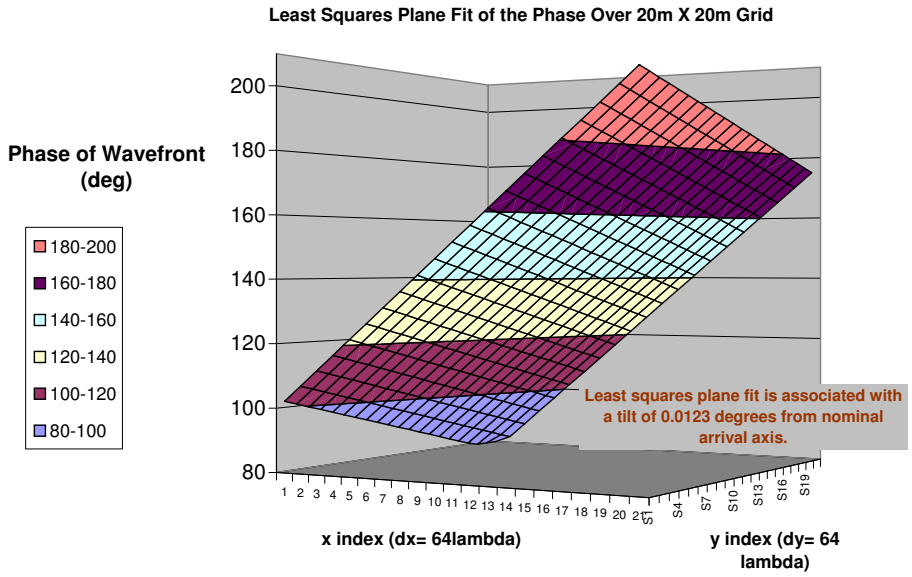


Figure 1.9: The least squares fit to the example phase front.

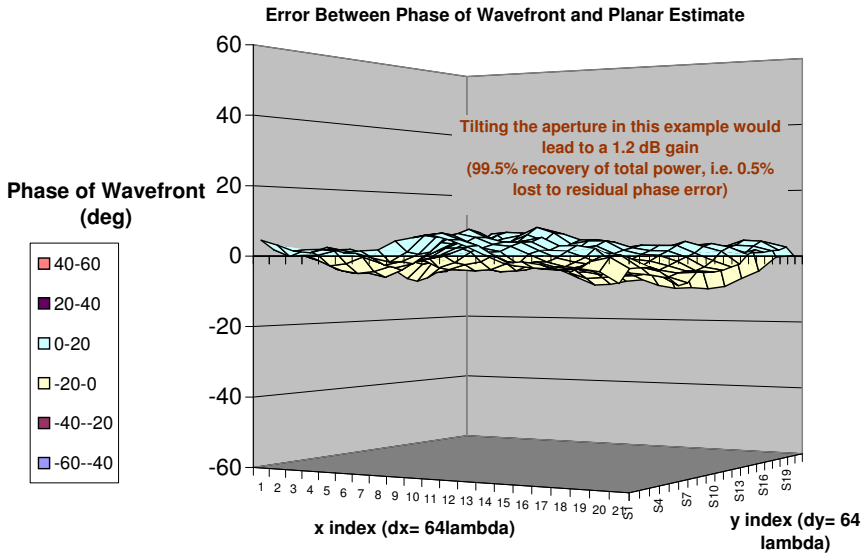


Figure 1.10: The residual error of the least squares fit to the example phase front.

antennas. The Goldstone Deep Space Communications Complex near Barstow, California uses antennas of 34 m and 70 m down to elevation angles of 20 degrees. For each antenna size, at each zenith angle, the following three methods were used in combining the signals within the dish's aperture:

- *Normal combining*: The signal values within the antenna's aperture were added together without any phase compensation.

- *De-tilted combining*: First the least squares planar fit to the phase was subtracted from the phase of each signal value within the aperture of the antenna. Then the resulting values were combined together. The generates the value that would occur if the antenna were pointed, or tilted, in the optimal direction. It is analogous to the use of a de-tilt mirror in adaptive optics.

- *Equi-phase combining*: The magnitude of the signal values within the antenna's aperture were added together. This generates the value that would occur if the antenna had a malleable subreflector that perfectly took out any phase distortion. This is analogous to the use of a deformable mirror in adaptive optics.

The average angle of arrival (AOA) as defined by the bore-sight angle $\phi_{bore} > 0$, is shown in

Fig. 1.11 for antenna diameters of 12m, 25m and 70m. Results are shown for frequencies of 20 GHz and 30 GHz as a function of zenith angle. The simulation used the typical cloud-present turbulence profile and the Kolmogorov spectrum with a L_0 cutoff of 1,000 m. The smaller antennas average the phase distortions over a smaller area and thus show larger average AOA values then those of larger antennas for the same phase fronts. It was surprising to see that the AOA values for a given antenna diameter and zenith angle were the same at 20 GHz as they were at 30 GHz. A larger value for the higher frequency had been anticipated. This behavior requires further examination for better understanding.

Larger antennas are more directional than smaller antennas and thus suffer a larger power loss for a given AOA error than do smaller antennas. Fig. 1.12 portrays this effect. It shows the associated improvement obtained in antenna output power, relative to normal combining, of equi-phase combining in which AOA error and all other phase distortions were removed. Most of the improvement, as will be seen later, is due to removing the AOA error through antenna tracking. The ratio of the power from equi-phase combining to that from normal combining is the inverse of the Strehl ratio commonly used in free-space optical communication systems to asses the signal power loss due to phase-distortion-induced de-focussing.

Next the effect of outer scale size L_0 was considered by running MPS simulations with the three aforementioned models for the spectrums of the index of refraction irregularities. The results show that the outer scale value has a large influence on AOA values. A 70m dish was considered since the effects of phase distortion are heightened for such large antennas. It is improbable that such a large dish would ever be used at low elevation angles but for the sake of studying a worst case situation, such angles were examined. Fig. 1.13 shows that large eddies contribute significantly to the average difference of the AOA relative to its nominal value. The restricted Kolmogorov spectrum, which in the MPS simulation had L_0 = 1,000 m, has much larger AOA values than either the Greenwood-Tarazano spectrum or the von Karman spectrum which both used L_0= 100 m.

Fig. 1.14 shows the improvements that would be obtained relative to normal combining by using 1) equi-phase combining (i.e. perfect phase compensation), and 2) de-tilt combining

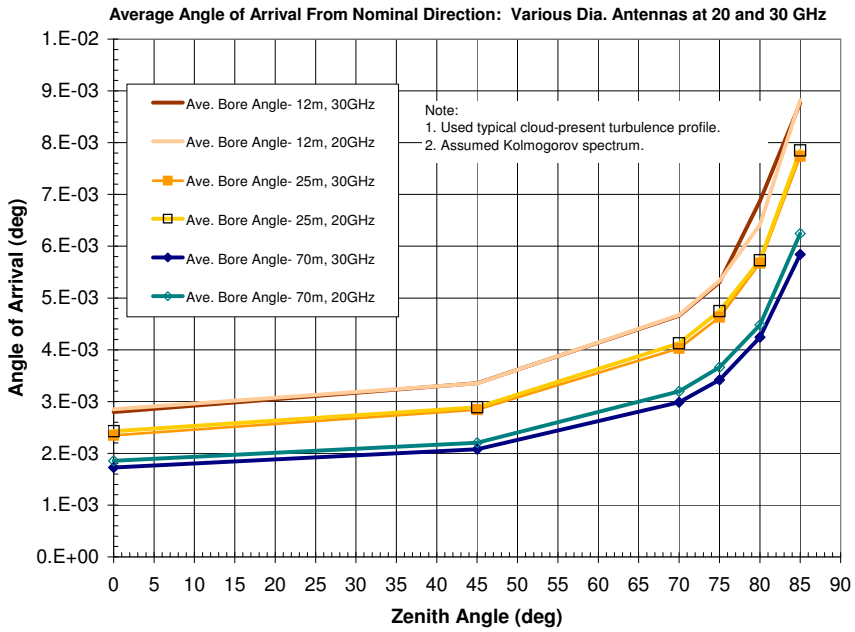


Figure 1.11: Average angle of arrival, relative to nominal bore-sight direction, for various antenna sizes for 20 GHz and 30 GHz (Kolmogorov Spectrum).

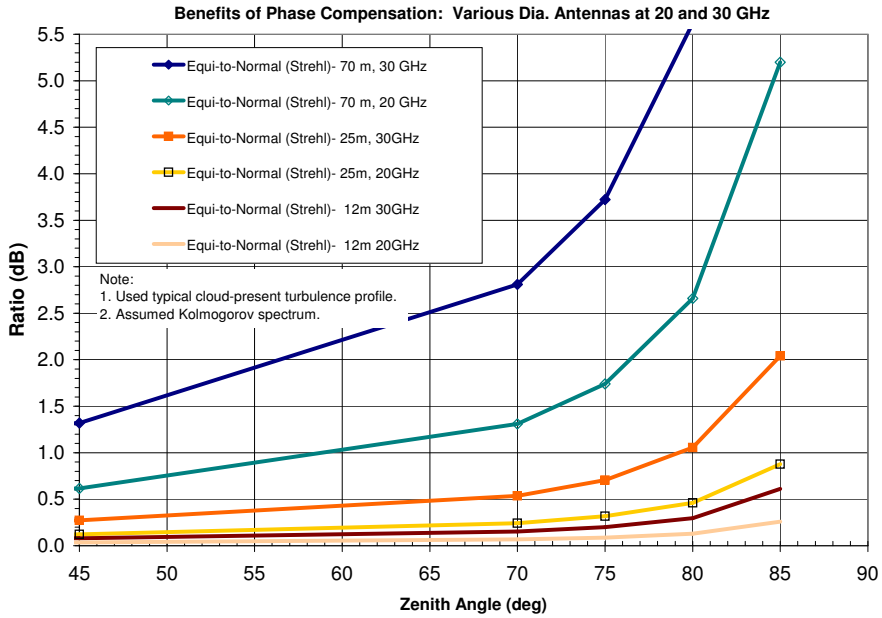


Figure 1.12: Benefits of phase compensation for various antenna sizes for 20 GHz and 30 GHz (Kolmogorov spectrum).

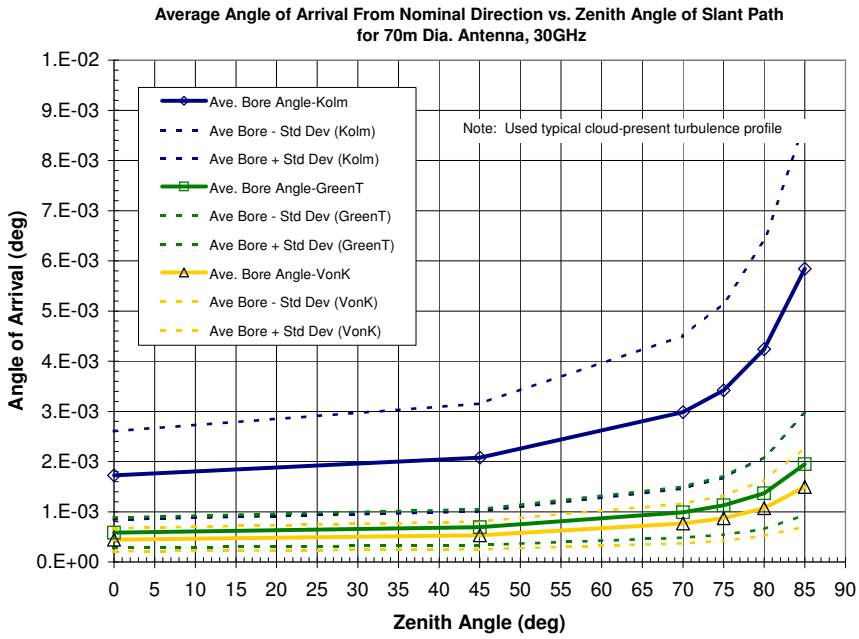


Figure 1.13: Angle of arrival, relative to the nominal bore-sight axis, for 70 m diameter antenna at 30 GHz, the Von Karman and Greenwood-Tarazano spectrums have fewer large eddies, which results in milder angle of arrival effects.

(i.e. perfect antenna tracking). Also shown is the percentage of the available energy that can be recovered through antenna tracking. The results show that antenna tracking can re-cover most of the available power. This was not surprising since in adaptive optics the de-tilt mirror typically provides 80% or more of the benefit in removing phase-distortion effects.

1.5 Conclusions

According to MPS simulations of tropospheric scintillation effects on Ka band satellite signals, used in conjunction with 1) a typical summer cloud-present turbulence profile and 2) typical summer no-cloud-present profile, there can be a substantial diversity benefit from using multiple antennas for links with elevation angles less than 20 degrees. At elevation angles greater than 10 degrees the potential benefit can be adequately estimated by *basic Rytov* theory for link availability percentages up to 99%. At elevations less than 10 degrees one should use simulations in order to adequately ascertain the potential benefits of using receive antenna diversity to mitigate scintillation at Ka band.

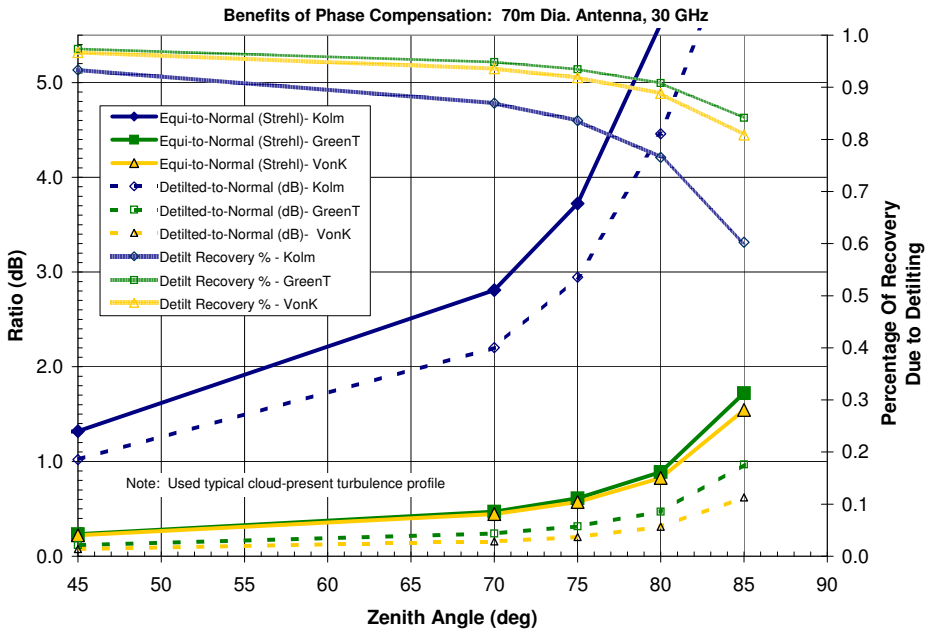


Figure 1.14: Benefits of phase compensation for 70 m diameter antenna at 30 GHz, the Von Karman and Greenwood-Tarazano spectrums have fewer large eddies, which results in less phase distortion and milder antenna output power degradation.

The results of the MPS simulations also show that the troposphere, under typical conditions, does not cause serious AOA fluctuations, or other forms of phase front distortion, for antennas with diameters of 25 m or less. For larger antennas, when the outer scale L_0 is larger than 100 m, a loss in antenna output power can be substantial. Tracking the direction of the incoming signal mitigates the loss, recovering almost all of the energy, with only a small fraction lost to higher-order phase distortions.

In terms of future applications of the method proposed in this chapter, MPS simulations could be used to confirm the estimates provided in [DA05] of the output power loss due to scintillation of JPL's proposed DSN array, which consists of 400 12-m diameter antennas spread out over a circular area with a diameter of 1.8 km.

CHAPTER 2

Analysis of Effects of Rain on the Capacity of Ka Band Satellite Signals

Contributions:

- Developed method for tightly estimating the outage probability of satellite links due to rain fades in regards to constrained capacity.
- Developed lower bound for the ergodic constrained capacity of log-log-normal channels.
- Developed constrained capacity analysis technique of using a piece-wise linear fit of the constrained capacity, selecting as the variable of fit either SNR in linear or logarithmic form to facilitate analysis.

2.1 Introduction

The rain fades, in decibels, for satellite channels are often modeled as having a log-normal distribution. The channel gain then has a log-log-normal (LLN) distribution. Researchers have determined methods for calculating the outage probabilities for LLN channels in regards to Shannon capacity [FV90, PLA09]. In these analyses the Shannon capacity assumes a continuously valued input signal with infinite support and Gaussian distribution. In practical satellite communications systems, such as those conforming to the European Telecommunication Standards Institute's (ETSI) second generation digital video broadcast-satellite (DVB-S2) standard, the input signal is constrained to a discrete signalling set such as finite-size quadrature amplitude modulation (QAM) constellations. At high SNR the Shannon capacity greatly overestimates the

capacity of these practical systems, particularly for low order constellations. For this reason a method is needed to evaluate the capacity and outage probability for LLN channels when the signal set is constrained to a finite alphabet. This chapter 1) introduces a new technique for estimating the outage probability, and outage capacity of LLN channels and 2) develops an efficient method for calculating a lower bound for the ergodic constrained capacity (ECC) of LLN channels. The outage probability method presented is applicable to both Shannon capacity and constrained capacity.

This chapter is organized as follows, section 2.2 gives the details of the LLN channel model and provides the constrained capacity formula that is to be evaluated for the LLN channel. Section 2.3 details the recent work done in determining estimates of the outage probabilities for the LLN channel. In addition, recent work done in determining lower bounds for the constrained capacity of fading channels is described. Section 2.4 describes the approach taken to develop a method applicable to the calculation of outage probabilities associated with the constrained capacity for the LLN channel. Section 2.5 details the derivation of a lower bound for the ECC of a LLN channel. Section 2.6 provides results and section 2.7 lists conclusions and possible future work.

2.2 System Model and Performance Measures

2.2.1 Channel Model

The PDF commonly used for the attenuation in decibels, A_{dB}, caused by rain fades is the log-normal distribution [GF04]

$$f_{A_{dB}}(A_{dB}) = \frac{1}{\sqrt{2\pi\sigma_c^2}A_{dB}} e^{-(\ln(A_{dB})-m_c)^2/(2\sigma_c^2)} \qquad A_{dB} \geq 0 \tag{2.1}$$

where m_c and σ_c are the log-normal parameters, i.e. the mean and standard deviation of the underlying normal process.

The associated CDF is

$$F_{A_{dB}}(A_{dB}) \;=\; 1-\mathrm{Q}\Big(\big(\ln(A_{dB})-m_c\big)/\sigma_c\big)\Big), \tag{2.2}$$

where

$$\mathrm{Q}(x) \;\doteq\; \frac{1}{\sqrt{2\pi}}\int_x^{\infty} e^{-t^2}\mathrm{d}t.$$

Substituting, $A_{dB}=\kappa\ln(g)$, where $\kappa=-20/\ln(10)$, the PDF of the channel gain g is obtained

$$f_G(g)= \frac{-1}{\sqrt{2\pi\sigma_c}\,g\ln(g)}\exp\left(\frac{-(\ln(\kappa\ln(g))-m_c)^2}{2\sigma_c^2}\right) \quad 0<g\leq 1. \tag{2.3}$$

This is a LLN distribution.

The received signal y through the channel is modeled as

$$y \;=\; \sqrt{E_s}\,g\,x+n \tag{2.4}$$

where E_s is the average transmitted signal energy, $x\in X$ is the transmitted signal with constellation X, g is the LLN distributed channel gain (2.3), and n is complex additive white Gaussian noise with zero mean and variance N_0.

The received SNR is $\gamma_{\mathrm{rx}}=g^2\gamma_{cs}$, where $\gamma_{cs}=E_s/N_0$ is the SNR received under clear sky conditions for which $g=1$. In decibels this can be expressed as

$$\gamma_{\mathrm{rx,dB}}=\log_{10}(g^2\gamma_{cs})=\gamma_{\mathrm{cs,dB}}-A_{dB}. \tag{2.5}$$

2.2.2 Capacity Formulas

The Shannon capacity as a function of the clear sky SNR and channel gain is

$$C_{Shan}(g,\gamma_{cs}) \quad = \quad \lambda \ln(1+g^2\gamma_{cs}) \tag{2.6}$$

where constant $\lambda = 1/\ln(2)$ ($\lambda = 1$) if the capacity is expressed in bits/s/Hz (nats/s/Hz).

The constrained capacity, with a QAM input that has a finite alphabet of M equiprobable values is [Bac99] (see also (C.6) in appendix C)

$$\begin{aligned} C_{con}(g,\gamma_{cs}) \doteq & \\ & \lambda \ln(M) - \frac{\lambda}{M}\sum_{j=1}^{M}\frac{1}{\pi N_0}\int_{y\in\mathcal{C}} e^{-\Delta^2(y,x_j)/N_0} \\ & \cdot \ln\left[\sum_{s=1}^{M} e^{(\Delta^2(y,x_j)-\Delta^2(y,x_s))/N_0}\right] \mathrm{d}y \end{aligned} \tag{2.7}$$

where x_i is the complex value of input i, $\Delta^2(y,x_i) \doteq |y - x_i\sqrt{g^2\gamma_{cs}N_0}|^2$ and $|\cdot|^2$ denotes squared Euclidean distance. The summation is over all the elements of the constellation X, and the integration is over all possible values of output Y, i.e. over the entire complex plane $\mathcal{C}$.

2.2.3 Outage Probability

Given a desired capacity value, c_o, there is an associated SNR, γ_o that must maintained in order for the channel to support c_o. Methods for determining the γ_o associated with c_o from (2.6) and (2.7) will be discussed in a following section. There is a non-zero probability at any point in time that the gain, g, of the LLN channel will be large enough to allow the received SNR to be larger than γ_o. The outage probability P_{out} is defined as the probability that the capacity, c, of the channel at a particular point in time will be less than c_o due to a channel fade. This is equal to the probability that γ_{rx} will be less than γ_o, which in turn is equal to probability that the

attenuation will exceed

$$A_0 = \gamma_{\mathrm{cs,dB}} - \gamma_{\mathrm{o,dB}} \tag{2.8}$$

The outage probability as a function of desired capacity c_o is

$$\begin{aligned} P_{out}(c_o) &\doteq P(c \le c_o) = P(\gamma_{\mathrm{rx}} \le \gamma_o) \\ &= P(A_{dB} > A_o) \doteq 1 - F_{A_{dB}}(A_0) \\ &= \mathrm{Q}\Big((\ln(A_o) - m_c)/\sigma_c\Big). \end{aligned} \tag{2.9}$$

A performance metric related to the outage probability is the outage capacity, $C_{out,P}(\gamma_{cs})$, which gives the capacity value available with an outage probability of P as a function of the clear sky SNR γ_{cs}.

2.2.4 Ergodic Constrained Capacity

The ergodic capacity at a specific γ_{cs} is, by definition,

$$E[C(\gamma)] \doteq \int_0^\infty x f_{C(\gamma_{\mathrm{cs}})}(x)\, \mathrm{d}x \tag{2.10}$$

where $f_{C(\gamma)}(\cdot)$ is the PDF for the capacity. Equivalently, (2.10) can be written in terms of the PDF of the channel gain,

$$E[C(\gamma)] = \int_0^\infty C(g, \gamma_{\mathrm{cs}}) f_g(g)\, \mathrm{d}g \tag{2.11}$$

where $f_g(g)$ is the PDF of the channel gain and $C = C_{con}$ ($C = C_{Shan}$) to determine the ergodic constrained (Shannon) capacity of the given channel for an SNR of γ.

For a LLN channel (2.11) becomes

$$E[C(\gamma)] = \frac{1}{\sigma_c^2}\int_0^1 C(g,\gamma_{cs})\frac{-1}{\sqrt{2\pi\sigma_c}\,g\ln(g)}e^{\frac{-(\ln(-\kappa\ln(g))-m_c)^2}{2\sigma_c^2}}\,dg. \tag{2.12}$$

2.3 State of the Art

2.3.1 Calculation of Outage Probability

Filip and Vilar calculated the outage probability for the Shannon capacity of a bent-pipe Ka band satellite links subject to a global fading process (i.e., rain attenuation and scintillation amplitude fluctuations) [FV90]. In their work they showed the advantage of looking at a link from a capacity point of view: features are highlighted that would not otherwise be considered, such as the fact that scintillation enhances the link approximately 50% of the time. Others have expanded their work by adding the noise temperature increase caused by rain attenuation to the global fading process [ABS97]. This work expands the calculation of outage probability for a given capacity to include the effects of using a finite alphabet input such as a QAM constellation rather than the Gaussian distributed input assumed by Shannon capacity calculations. Since rain attenuation is the dominant factor in determining outage probability and outage capacity, scintillation and noise temperature will be ignored and only rain fading will be considered.

As shown in (2.9) the outage probability of a given capacity value, c_o can be determined from the cumulative distribution function (CDF) of the channel attenuation by mapping c_o to the associated SNR γ_o, and in turn to the attenuation A_0 associated with γ_o. A means is required to map c_o to γ_o. In [FV90, PLA09] for the Shannon capacity, (2.6) provides such a mapping

$$\gamma_o = e^{c_o/\lambda} - 1 \tag{2.13}$$

A similar method is needed to map c_o to γ_o for constrained capacity. However, the SNR associated with a given constrained capacity is not readily obtained from (2.7) due to the 2-D

integration over the complex plane and to the summation over all the constellation points. An indirect method for obtaining the attenuation associated with a given constrained capacity value must be used. Such a method is introduced in section 2.4.

2.3.2 Lower Bound on Ergodic Constrained Capacity

For typical channel PDF's a closed form expression does not exist for (2.11) for either the Shannon or the constrained capacity and numerical means must be used. The evaluation of (2.11) for constrained capacity is further complicated by the 2-D integration over the complex plane required by (2.7). For these reasons a bound of (2.11) that is readily evaluated is required. The cutoff rate, R_0, provides a lower bound for the constrained capacity, $LB(g) \equiv R_0(g)$. From ([Wil96], 4.3.37) the cutoff rate is given by,

$$R_0 \equiv log_2(M) - log_2\left[1 + \frac{1}{M}\sum_{j=1}^{M}\sum_{s=1,s\neq j} e^{-\frac{g^2\|\alpha_j-\alpha_s\|^2}{4N_0}}\right]$$

By taking the expectation on both sides of the equation and applying Jensen's inequality, a lower bound, LB_1, for (2.12) for the constrained capacity case is obtained [Bac99],

$$\begin{aligned} E[LB] &\equiv E[R_0] \\ &= log_2(M) - \\ &\quad E\left[log_2\left[1 + \frac{1}{M}\sum_{j=1}^{M}\sum_{s=1,s\neq j} e^{-\frac{g^2\|\alpha_j-\alpha_s\|^2}{4N_0}}\right]\right] \qquad (2.14) \\ &\geq log_2(M) - \\ &\quad log_2\left[1 + \frac{1}{M}\sum_{j=1}^{M}\sum_{s=1,s\neq j} E_g\left[e^{-\frac{g^2\|\alpha_j-\alpha_s\|^2}{4N_0}}\right]\right] \\ &\equiv LB_1 \qquad (2.15) \end{aligned}$$

A lower bound, LB_1, for the channel-averaged (ergodic) constrained capacity can be obtained by the above formula for any channel. The crux is the evaluation of the expectation with respect to the channel gain, g, of the exponential term. In [Bac99] and [BF00] this expectation

was evaluated, and LB_1 expressions developed, for Rayleigh, Nakagami, and Rician channels. For these channels the LB_1 value was found to be within 1-5 dB of numerical evaluations of (2.11) under a variety of conditions. This paper will evaluate this expectation and develop an expression for LB_1 for the LLN channel, and thus a lower bound for (2.12), in Section V.

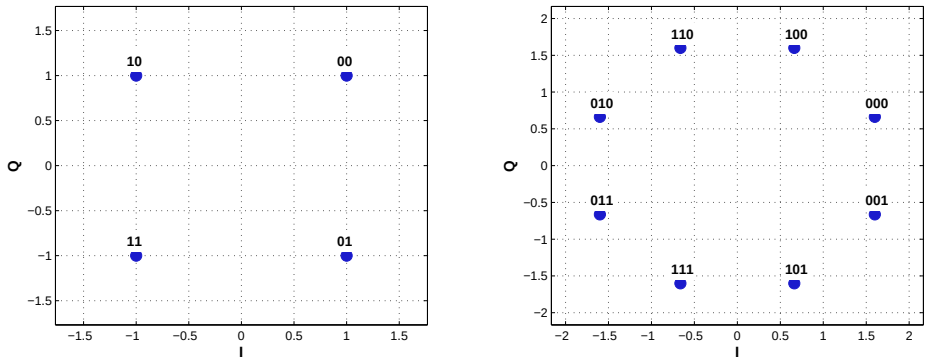
2.4 Method for Calculation Outage Probability for LLN Channel

This section describes an indirect method for obtaining the SNR associated with a given constrained capacity value. The long near linear sections and large radii of curvature of the constrained capacity function (2.7) make it an ideal candidate for piece-wise linear (PWL) curve fitting. A PWL fit can also be applied to the Shannon capacity (2.6) to obtain a uniform method for obtaining the outage probability associated with a given capacity value, whether it be associated with Shannon capacity or constrained capacity. Using the method of [Tom74], PWL fits were made to (2.6) and to (2.7) for DVB-S2 constellations, shown in Fig. 2.1, as a function of $\gamma_{\mathrm{rx,dB}}$

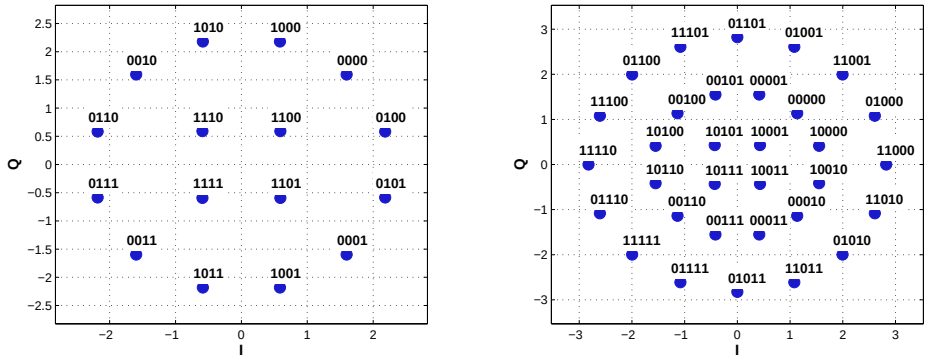
$$\begin{aligned} C(\alpha,\gamma) &\cong C_{pwl}(\alpha,\gamma) \\ &= m_k\,\gamma_{\mathrm{rx,dB}} + b_k \ \text{ for } c_{k-1} < C \le c_k, \end{aligned} \tag{2.16}$$

where m_k and b_k are the slope and intercept values for section k of the PWL fit, which apply to capacity values between c_{k-1} and c_k. Figure 2.2 (a) shows the Shannon capacity and its PWL fit and Figure 2.2 (b) shows the constrained capacity curves and their PWL approximations. In all cases the fits were made so as to have a maximum absolute error of 0.04 bits/s/Hz. Table 2.1 provides the corresponding slope and intercept values for each section of the PWL fits portrayed in these figures. Since the constrained capacity curves for the various ring ratios (RR) for 16APSK are within 0.1 dB of each other, only values for the RR= 2.7 constellation are provided. A similar situation holds for 32APSK.

The maximum absolute error can be reduced further if desired at the expense of the addition



(a) QPSK (b) 8PSK



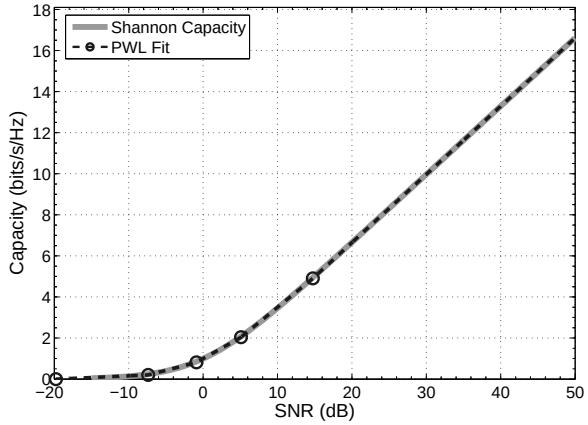
(c) 16APSK (d) 32APSK

Figure 2.1: Constellations of the DVB-S2 standard: (a) QPSK, (b) 8PSK, (c) 16APSK, and (d) 32APSK

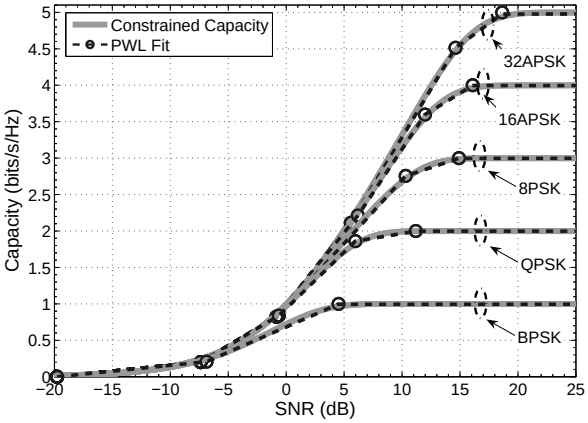
of more sections to the PWL fit. Conversely, the PWL fit can be made with even fewer sections with a relaxed error limit. Since as the SNR increases the Shannon capacity becomes linear as a function of the SNR in dB the PWL of the Shannon capacity has a finite number of sections. In addition, since the constrained capacity of an M-ary constellation goes to $c_{max} = \log_2(M)$ as the SNR becomes large the capacity curve plateaus and thus the number of sections needed for the PWL fit to a constrained capacity curve is also finite.

The outage probability associated with a given capacity value c_o can be evaluated by using the PWL fit of the capacity to estimate the SNR value $\gamma_{o,dB}$ needed to achieve $c_{k-1} < c_o \leq c_k$

$$\gamma_{o,dB} \cong \frac{c_o - b_k}{m_k}, \tag{2.17}$$



(a) Shannon Capacity



(b) Constrained Capacities

Figure 2.2: Piece-wise linear fits for (a) Shannon capacity and (b) constrained capacities of DVB-S2 constellations as a function of SNR in dB with a maximum error of 0.04 bits/s/Hz.

and then using (2.8) in conjunction with (2.9) to calculate the outage probability associated with c_o for an SNR of $\gamma_{cs,dB}$ resulting in

$$P_{out}(c_o) \cong \mathrm{Q}\left(\left(\ln\left(\gamma_{cs,dB} - \frac{c_o - b_k}{m_k}\right) - m_c\right)/\sigma_c\right) \tag{2.18}$$

(a) Shannon

i	c_i (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.20214	0.016307	0.32253
2	0.81754	0.095012	0.90353
3	2.04144	0.20398	1.00214
4	4.90771	0.29684	0.52905
5	>4.90771	0.33210	0.00884059

(b) BPSK

i	c_i (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.20729	0.016077	0.31777
2	1.0	0.069586	0.68548
3	>1	0	1

(c) QPSK

i	c_i (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.20221	0.016308	0.32253
2	0.83973	0.094504	0.89946
3	1.86128	0.15464	0.93746
4	2	0.026623	1.70223
5	>2	0	2

(d) 8PSK

i	c_i (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.20219	0.016308	0.32253
2	0.83160	0.094647	0.90059
3	2.75584	0.17427	0.95863
4	3	0.053184	2.20735
5	>3	0	3

(e) 16APSK (RR= 2.70)

i	c_i (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.20218	0.0163084	0.32254
2	0.82437	0.094817	0.90193
3	2.21297	0.19826	0.98655
4	3.59738	0.23927	0.73286
5	4	0.097632	2.42853
6	>4	0	4

(f) 32APSK (RR1= 2.64, RR2= 4.63)

i	c_i (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.20216	0.016308	0.32253
2	0.82156	0.094884	0.90250
3	2.11144	0.20082	0.99287
4	4.51479	0.26574	0.63127
5	5	0.11584	2.82188
6	>5	0	5

Table 2.1: Coefficients for the PWL fit for: (a) the Shannon capacity and (b)-(f) constrained capacities for DVB-S2 constellations.

2.5 Proposed Lower Bound on the Ergodic Constrained Capacity of LLN Channels

This section describes the evaluation of the LB_1 lower bound given in (2.15) for the ergodic constrained capacity of a LLN channel. In order to obtain LB_1 for a LLN channel the following expectation is needed,

$$E_g\left[e^{-\frac{g^2\|\alpha_j-\alpha_s\|^2}{4N_0}}\right] = \frac{1}{\sqrt{2\pi\sigma_c^2}}\int_{g=0}^{g=1}\frac{1}{g\ln(g)}e^{-\frac{(\ln(-\kappa\ln(g))-m_c)^2}{2\sigma_c}}e^{-\frac{g^2\|\alpha_j-\alpha_s\|^2}{4N_0}}\mathrm{d}g$$

By first substituting $x = -\kappa \ln(g)$, and then $u = \ln(x)$, the above integral becomes

$$E_g\left[e^{-\frac{g^2\|\alpha_j-\alpha_s\|^2}{4N_0}}\right] = \frac{1}{\sqrt{2\pi\sigma_c^2}}\int_{u=-\infty}^{u=\infty} e^{-\left(\frac{(u-m_c)^2}{2\sigma_c^2}+\frac{\|\alpha_j-\alpha_s\|^2}{4N_0}e^{(-2e^u/\kappa)}\right)}\mathrm{d}u \tag{2.19}$$

The term $f(u) = e^{-2e^u/\kappa}$ in (2.19) can be well approximated with the piecewise continuous function

$$g(u) = \begin{cases} 1 & -\infty < u \le b_1 \\ 1-\left(\frac{u-b_1}{b_2-b_1}\right)^2 & b_1 < u \le b_2 \\ 0 & b_2 < u < \infty \end{cases} \tag{2.20}$$

Figure 2.3 compares the original function, $f(u)$, with the above approximation (2.20). Use of (2.20) with values of b_1= -2.65 and b_2= 2.60 provide integration results that are within 2% of numerically integrated values for equation (2.19) for values of $0 < \| \alpha_j - \alpha_s \|^2 / 4N_0 < 2$, which is the range of interest, and values of the LLN parameters m_c and σ_c ranging from those for arid regions to those for tropical regions.

Using this piece-wise continuous approximation, a closed form expression can be developed (see appendix D) for (2.19) that can be evaluated for any LLN channel. This expression can be used in (2.15) to provide a lower bound expression for the ergodic constrained capacity of a

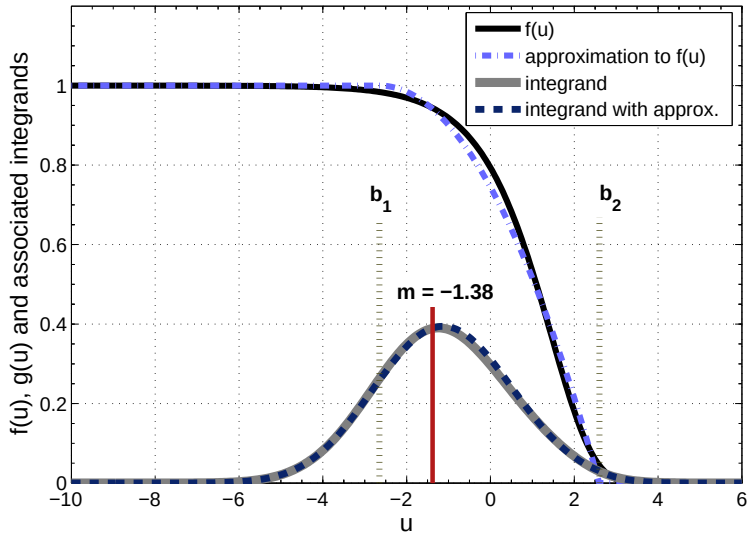


Figure 2.3: Comparison of the actual function and its associated integrand with the approximation and its associated integrand.

given QAM constellation in a LLN channel

$$
\begin{aligned}
LB_1 = & \, log_2(M) \\
& -log_2\Big(1+\frac{1}{M}\sum_{j=1}^{M}\sum_{s=1,s\neq j}^{M} E_g\Big[e^{-\frac{g^2\|\alpha_j-\alpha_s\|^2}{4N_0}}\Big]\Big) \\
= & \, log_2(M) - log_2\Big(1+ \\
& \frac{1}{M}\sum_{j=1}^{M}\sum_{s=1,s\neq j}^{M}\Big[e^{-\frac{\|\alpha_j-\alpha_s\|^2}{4N_0}} Q\left(\frac{m_c-b_1}{\sigma_c}\right) \\
& +e^{-\tau_{j,s}}\frac{1}{\sqrt{2\pi\sigma_c|a_{j,s}|}}(Q(v_{2_{j,s}})-Q(v_{1_{j,s}})) \\
& +Q\left(\frac{b_2-m_c}{\sigma_c}\right)\Big]\Big) \\
= & \, log_2(M) - log_2\Big(1+(M-1)Q\left(\frac{b_2-m_c}{\sigma_c}\right) \\
& +\frac{1}{M}\sum_{j=1}^{M}\sum_{s=1,s\neq j}^{M}\Big[e^{-\frac{\|\alpha_j-\alpha_s\|^2}{4N_0}} Q\left(\frac{m_c-b_1}{\sigma_c}\right) \\
& +e^{-\tau_{j,s}}\frac{1}{\sqrt{2\pi\sigma_c|a_{j,s}|}}(Q(v_{1_{j,s}})-Q(v_{2_{j,s}}))\Big]\Big)
\end{aligned}
\tag{2.21}
$$

where,

M: The number of elements in the modulation alphabet

α_i: The ith element of the modulation alphabet

m_c: The mean of the underlying normal process of the LLN channel

σ_c: The standard deviation of the underlying normal process of the LLN channel.

b_1: The boundary point between region 1 and region 2 of the piecewise approximation to $e^{-2e^u/\kappa}$, nominally -2.65

b_2: The boundary point between region 2 and region 3 of the piecewise approximation to $e^{-2e^u/\kappa}$, nominally 2.60

and,

$$a_{j,s} = \frac{1}{2\sigma_c^2} - \frac{\|\alpha_j-\alpha_s\|^2}{4N_0(b_2-b_1)^2}$$

$$\tau_{s,j} = \frac{m_c^2}{2\sigma_c^2} + \frac{\|\alpha_j-\alpha_s\|^2 b_2(b_2-2b_1)}{4N_0(b_2-b_1)^2} - \frac{(\|\alpha_j-\alpha_s\|^2-2N_0m(b_2-b_1)^2)^2}{4N_0(b_2-b_1)^2[2N_0(b_2-b_1)^2-\|\alpha_j-\alpha_s\|^2\sigma_c^2]}$$

$$v_{1_{s,j}} = \sqrt{2a'_{j,s}}(b_1-\gamma_{j,s})$$

$$v_{2_{s,j}} = v_{1_{j,s}} + \sqrt{2a'_{j,s}}(b_2-b_1)$$

$$\gamma_{j,s} = \frac{2N_0m(b_2-b_1)^2-\sigma_c^2b_1\|\alpha_j-\alpha_s\|^2}{2N_0(b_2-b_1)^2-\sigma_c^2\|\alpha_j-\alpha_s\|^2}$$

$$a'_{s,j} = \max(a_{s,j},0).$$

The evaluation of variables $a_{j,s}$, $\tau_{j,s}$, $v_{1_{j,s}}$, $v_{2_{j,s}}$ and $\gamma_{j,s}$ may be tedious but is straightforward. Equation (2.21) can be used to determine a lower bound for the ergodic constrained capacity of a given QAM constellation in a LLN channel with parameters (m_c,σ_c) without any 2-D integrations using only well known functions.

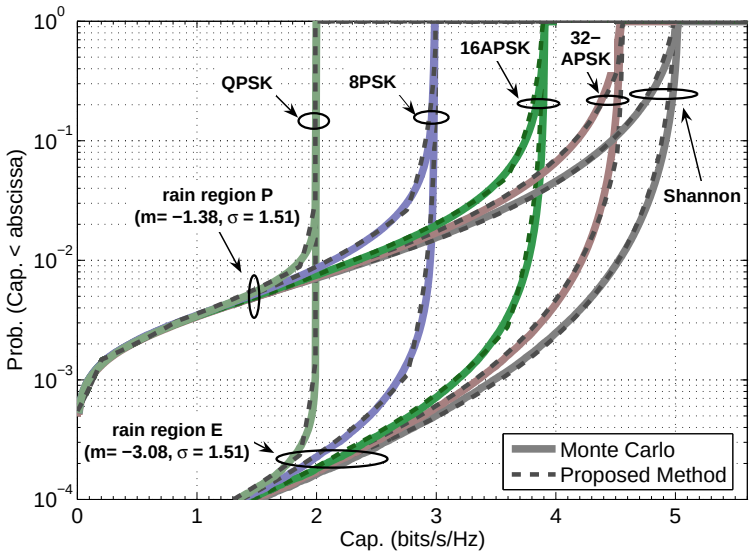


Figure 2.4: Comparison of Monte Carlo results and proposed expression for the outage probability for constrained capacity using DVB-S2 modulations in two LLN channels (rain regions E and P, freq= 20 GHz, el. angle = 20 degrees) for a SNR of 15 dB.

2.6 Results

2.6.1 Results Using the LLN Outage Probability Formula

Figure 2.4 compares Monte Carlo simulation results with values calculated using the proposed method, given in (2.18), for the outage probability of the constrained capacity for DVB-S2 modulations in two LLN channels associated with rain regions E (m_c= -3.08, σ_c= 1.51) and P (m_c= -1.38, σ_c= 1.51) for an elevation angle of 20 degrees and a frequency of 20 GHz for a clear sky SNR of 15 dB. At a given outage probability value, the estimate from the proposed expression agrees within the error of the PWL fit, 0.04 bits/s/Hz in this example, of the results from Monte Carlo simulations and numerical methods.

Figure 2.5 shows the outage constrained capacity, $C_{out,P}(\gamma)$, associated with outage percentages of P= 10%, 1%, 0.5% and 0.1% for DVB-S2 constellations for a LLN channel with parameters (m_c= -1.38, σ_c= 1.51). These values were determined by using (2.18) to generate outage probability curves for several clear sky SNR values and then determining from each of those curves the capacity value associated with the desired outage percentage. Also included in

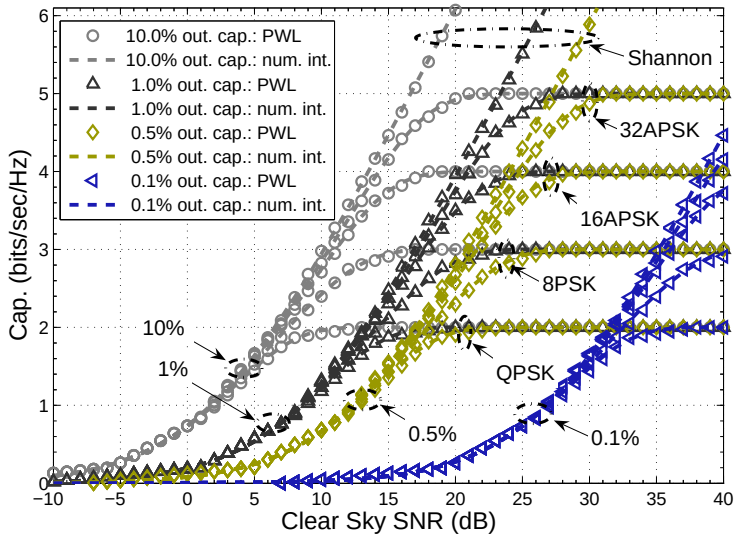


Figure 2.5: Outage capacity for DVB-S2 constellations for various outage rates for a LLN channel with parameters (m_c= -1.38, σ_c= 1.51).

the graph are the outage capacities determined using numeric integration.

2.6.2 Results Using the LLN Lower Bound Formula

Figure 2.6 shows the ergodic constrained capacity and the associated lower bound from Monte Carlo simulation for a LLN channel (m_c= -1.38, σ_c= 1.51) along with LB_1 values calculated by numeric integration of (2.15) and by the proposed expression (2.21) for DVB-S2 constellations in a LLN channel. The LB_1 curves vary from being 3.0 dB from the Monte Carlo constrained capacity curves at c= 0 to being within 2.1 dB at $c = 0.95\,c_{max}$. The match between the Monte Carlo LB curves and the LB_1 curves is very close, which shows that the application of Jensen's inequality in obtaining (2.15) from (2.14) had a negligible effect on the bound (0 to 0.2 dB).

2.7 Conclusion

This chapter presented a new method to quickly and accurately estimate the outage probabilities and outage capacities of a LLN channel, for links constrained to a finite constellation, without

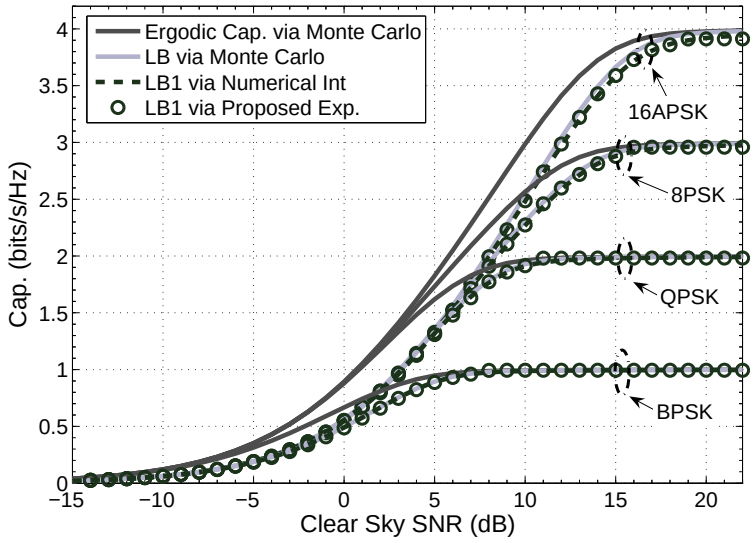


Figure 2.6: Ergodic constrained capacity and lower bound from Monte Carlo simulation for LLN channel along with LB_1 values calculated by numeric integration of (2.15) and by proposed expression (2.21). The LLN channel parameters were (m_c= -1.38, σ_c= 1.51).

resorting to lengthy Monte Carlo simulations or numerical integration. This new method was shown to match well with Monte Carlo simulation results. The method was shown to be applicable to Shannon capacity outage probability calculations as well. The proposed method is applicable to any QAM modulation once the parameters of the PWL fit of (2.7) are obtained via some method such as that of [Tom74]. For the DVB-S2 QPSK, 8PSK, 16APSK, and 32APSK constellations, Table 2.1 provides the required parameter values.

This chapter also presented a closed form expression for a lower bound for the ergodic constrained capacity of a LLN channel.

These methods can be extended to the antenna combining situation, and also to the Weibull distribution, which is also used to model the effects of rain fades on satellite channels.

CHAPTER 3

Mitigation of the Effects of Rain on the Capacity of Ka Band Satellite Signals Through Antenna Diversity

Contributions:

- Quantified benefit or receiver diversity in combating rain fades in order to improve ergodic capacity and outage probability in regards to constrained capacity.

3.1 Introduction

Given that modern coding methods have brought the performance of communication links close to theoretical limits there has been increased interest in determining theoretical capacity limits for a given link. With the large rain attenuation that Ka band satellite links are subject to there has been appreciable interest in using site diversity to mitigate the effects of rain fades. This chapter presents a method to determine the constrained capacity of a satellite link that uses site diversity to mitigate the effects of rain fades. The constrained capacity is defined as the maximum information rate that can be sent, error free, through a channel under the constraint that the input has a finite alphabet, which is the case when standard constellations are used. The constrained capacity stands in contrast to the Shannon capacity, which assumes a Gaussian distributed input for an additive white Gaussian noise (AWGN) channel. The method of this chapter determines the ergodic constrained capacity, the outage probability and the outage constrained capacity of a satellite link with antenna diversity that is subject to rain fades. The constrained capacity, when averaged over all fade realizations, is termed the ergodic constrained capacity. The constrained capacity at a given outage rate is known as the outage constrained

capacity. In the previous chapter, the ergodic constrained capacities, outage probabilities and outage constrained capacities of a standard satellite link consisting of a single satellite and a single ground station, i.e., a single-input-single-output (SISO) configuration, were calculated (see also [EF10]). In this chapter, the ergodic, outage probabilities and outage constrained capacities when site diversity is used will be calculated and compared against those of a SISO link.

Section 3.2 describes the state of the art of calculating the capacities of satellite links that use site diversity. Section 3.3 gives the channel models. Section 3.4 describes how the constrained capacities were computed. Section 3.5 provides the results of the constrained capacity calculations when the DVB-S2 constellations are used. Section 3.6 concludes the chapter and lists some related areas of future work.

3.2 State of the Art

Liolis, Panagopoulos, and Arapoglou considered the Shannon capacity (i.e., the capacity assuming a Gaussian distributed input signal) of satellite links that use site diversity to mitigate rain fading [PLA09]. They developed a closed form expression for the outage probability for Shannon capacity that is applicable to both single-input-multiple output (SIMO) and multiple-input-single-output (MISO) situations. The formula given in their paper is applicable to both the single-input-double-output and the double-input-to-single output cases, but could be directly expanded to cover other SIMO and MISO cases. They then numerically evaluated the expression for some cases of interest.

This chapter further expands on this site diversity capacity work by utilizing the constrained capacity, as opposed to the Shannon capacity. Practical constellations, such as those defined in ETSI's second generation Digital Video Broadcast-Satellite (DVB-S2) standard [ET05] will be used. Two types of rain regions, a tropical environment and a relatively dry European environment will be considered. In [PLA09], only outage probabilities were given; in this work ergodic capacities, outage capacities and outage probabilities will be provided.

A two-site receive-diversity scheme employing maximal ratio combining (MRC) was con-

sidered. Three different amounts of correlation between the rain fades on the two paths were examined. Although the constrained capacity results given in this chapter were calculated considering receive-diversity with two receive antennas, the results are applicable as well to a transmit-diversity situation with two transmit antennas for the reasons brought forth in [PLA09].

3.3 Channel Model

For the SISO case, the following channel model was used,

$$y_{SISO} = g_1 x + n \tag{3.1}$$

For the two Rx receive-diversity situation it was assumed that MRC combining of the signals from the two paths would be used. MRC combining was selected because it provides the maximum SNR of any linear combining method. In MRC combining, the signal from each path is weighted by the conjugate of its path gain, and then the two resulting signals are summed together. Thus, for the site diversity case, the following channel model was used,

$$y_{MRC} = \frac{1}{\sqrt{2}}(|g_1|^2 + |g_2|^2)x + n \tag{3.2}$$

where, in both of the above expressions, x is the transmitted signal, n is additive white Gaussian noise, y_{SISO} and y_{MRC} are the outputs of the SISO and MRC channels, respectively, and g_1 and g_2 are the channel gains of path one and path two, respectively. The variables g_1 and g_2 represents rain-induced fading and, when converted to dB values, have log-normal distributions. For the MRC channel, the division by $\sqrt{2}$ was done in order to create equal received power for the SISO and MRC cases in the clear sky situation, so that any increase in performance of the MRC case relative to that of the SISO case could be attributed to diversity.

In order to keep the problem description fairly generic, this chapter considers situations described by rain regions, as formerly used in the ITU-R rain models. This chapter considers both rain region P, a tropical environment, and also rain region E, a fairly dry region that covers a

large portion of Europe. It is commonly accepted that the rain attenuation values, when converted to dB values, have a log-normal distribution [GF04]. Therefore the probability density function of the attenuation of path i, $A_i = -10\log(g_i)$ (dB), is given by,

$$f(A_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}}\frac{1}{A_i}\exp\left[-\frac{(\ln A_i - m_i)^2}{2\sigma_i^2}\right] \qquad i = 1,2 \tag{3.3}$$

where m_i and σ_i are the log-normal parameters of path i and are the mean and standard deviation, respectively, of the underlying Gaussian process.

In this work, it was assumed that the attenuations of the two paths were jointly log-normal, having a joint probability density function given by,

$$\begin{aligned} f(A_1,A_2) = {} & \frac{1}{2\pi\sigma_1\sigma_2 A_1 A_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho)^2}\left(\frac{(\ln A_1 - m_1)^2}{\sigma_2^2}\right.\right. \\ & \left.\left. -\frac{2\rho(\ln A_1 - m_1)(\ln A_2 - m_2)}{\sigma_1\sigma_2} + \frac{(\ln A_2 - m_2)^2}{\sigma_2^2}\right)\right] \end{aligned} \tag{3.4}$$

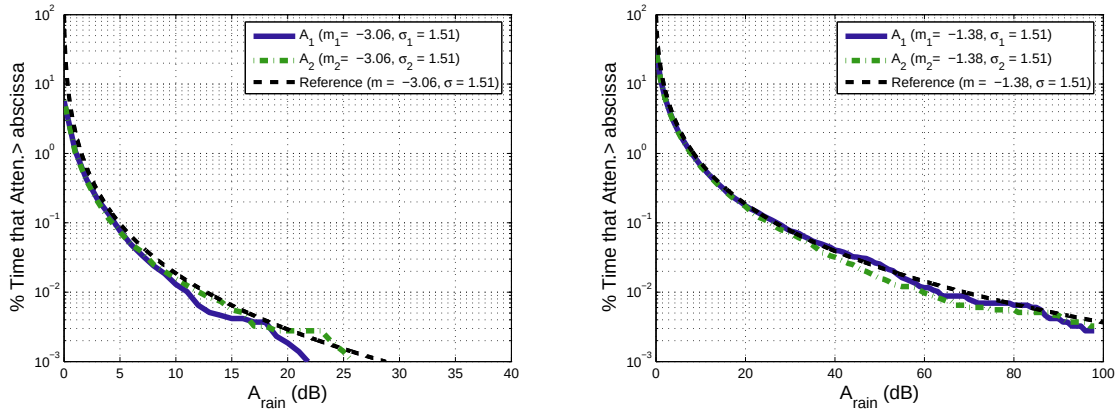
where m_i and σ_i are the log-normal parameters of path i and ρ is the correlation coefficient between A_1 and A_2. This chapter examined site diversity for three different values of ρ (0, 0.25 and 0.5) in order to quantify the effects of channel correlation on site diversity. It was assumed that the two paths had the same statistical parameters, i.e., $m_1 = m_2 = m$ and $\sigma_1 = \sigma_2 = \sigma$.

Fig. 3.1 plots the complementary cumulative distribution function (CCDF) for rain regions E and P with a 20 GHz signal and a 40 deg elevation angle, as specified by the ITU-R model. The log-normal parameters, (m,σ) for this scenario are (-3.06, 1.51) for region E and (-1.38,1.51) for region P.

3.4 Constrained Capacity Calculation

From information theory [CT91], the mutual information of a channel with input X and output Y is described by,

$$I(X;Y) = H(X) - H(X|Y) \tag{3.5}$$



(a) Region E (b) Region P

Figure 3.1: Complementary CDF of rain fade levels along with the empirical distribution of rain fade values used in determining the constrained capacities for that region for each of: (a) region E and (b) region P.

where $H(X)$ is the entropy of the input and the conditional entropy $H(X|Y)$ is given by,

$$H(X|Y) \quad = \quad \int_{\Omega_Y} f(y)H(X|Y=y)dy \tag{3.6}$$

where,

$$H(X|Y=y) = -\sum_{\Omega_X} p(x|Y=y)\log(p(x|Y=y)) \tag{3.7}$$

when X has a finite alphabet, and,

$$H(X|Y=y) = -\int_{\Omega_X} f(x|Y=y)\log(f(x|Y=y))dx \tag{3.8}$$

when X is continuously distributed.

In these equations, $f(y)$ is the probability density function of the channel output Y. In (3.7), $p(x|y)$ is the probability mass function of a finite-valued channel input, X, conditioned on a given value for the output, y, of the channel. In (3.8), $f(x|y)$ is the conditional probability density function of a continuously-valued channel input, X.

The mutual information is the upper limit on the rate of information that can be sent across a channel error-free. It is typically measured in terms of how many bits of information can be sent per channel usage, e.g., per transmission of a constellation symbol. The capacity of a channel is the value of mutual information which has been maximized over all possible distributions for the input, X. For an additive white Gaussian noise channel, subject to the constraint that the average power must by less than some value P_{sig}, the mutual information is maximized by a Gaussian distribution on the two-dimensional input X, resulting in the well known formula,

$$\begin{aligned} C &= max_{p(x):E[X^2]\leq P_{sig}} I(X;Y) \\ &= log_2(1+\frac{P_{sig}}{N}) \end{aligned} \tag{3.9}$$

where P_{sig}/N is the signal-to-noise ratio (SNR).

In practice, however, it is not feasible to use an input with a Gaussian distribution and so the input is usually drawn uniformly from a set of discrete values, i.e., from a constellation, such as the QPSK, 8PSK, 16APSK and 32APSK constellations of the DVB-S2 standard shown in Fig. 2.1. The constrained capacity is the value of the mutual information when the input is constrained to a given constellation and the distribution of the input values is set to maximize the value of the mutual information. As an alternative to altering the input distribution, constellation designers can hold the input distribution uniform, and alter the location of points within the constellation, such as by changing the ratios of the radii of the rings of the 16APSK or 32APSK constellations, to maximize the mutual information. This also can be called the constrained capacity.

A uniform distribution maximizes the entropy of a random variable with a finite alphabet [CT91]. An input, X, that consists of an M-ary constellation will have an entropy, $H(X)$, that is less than or equal to $\log_2(M)$. The capacity of a channel with such a constrained input can not exceed $\log_2(M)$. For example, the 8PSK constellation with a uniform input distribution, will have $H(X) = 3$, and thus a maximum mutual information of 3 bits per channel usage.

In this chapter, a Monte Carlo approach was used to determine the constrained capacity for

a given set of channel gains, g_1 and g_2. Each channel gain was calculated from a randomly generated attenuation value, A_i, i.e., $g_i = 10^{-A_i/10}$. The attenuation values were drawn according to the distribution of (3.4). The input, X, to the channel consisted of symbols drawn uniformly from a given DVB-S2 constellation. Noise values were drawn from a Gaussian distribution with a variance value determined by the desired clear sky SNR. The constrained capacity conditioned on a given set of channel gains, $C_{con}(g)$, was computed by applying information theoretic formulas (3.5), (3.6), and (3.7) to the resulting empirical distributions for the output of the channel, Y. At a given SNR, the resulting values of $C_{con}(g)$ from many channel realizations were 1) averaged together to determine the ergodic constrained capacity and 2) binned in order to create a histogram from which outage constrained capacities could be determined.

This was done for the situations in which the end-to-end channel was considered to send either 1) symbols, or 2) individual bits. The first case is referred to as the *joint capacity*. It is applicable for coding methods that work on a symbol-by-symbol basis, such as trellis-coded modulation methods. In the other chapters of this work only the joint capacity is considered and it is referred to simply as constrained capacity. The second case is applicable to coding methods that work on a bit-by-bit basis such as binary LDPC codes employed by the DVB-S2 standard. For the second case, each bit of a constellation symbol is considered as a separate binary channel. The capacities of these individual binary channels were calculated and then summed to yield what has been termed as the *parallel decoding* capacity, which will always be less than the symbol-based capacity. A discussion of joint capacity and parallel decoding capacity is given in [BJF07] and the formulas for each are given by (C.6) and (C.10) respectively in appendix C. It was found that the parallel decoding capacity was generally only slightly less (within ~ 0.1 bit/channel usage) than the joint capacity for the DVB-S2 constellations. Since this chapter is interested in DVB-S2 satellite links only the parallel decoding constrained capacity values will be shown.

3.5 Results

Using the method described in Section 3.4, both (1) the average of the constrained capacity and (2) the outage capacity for each of the constellations of the DVB-S2 standard were calculated for both rain region E and rain region P. A carrier frequency of 20 GHz and an elevation angle of 40 degrees were assumed.

The empirical CCDFs of all the realizations of rain attenuations, A_i $(i = 1,2)$, for each of the two paths that were used in the Monte Carlo calculations of the constrained capacity values are plotted in Fig. 3.1.

For the 16APSK constellation a value of 2.7 was used for the ratio of the outer ring radius to that of the inner ring. This ring ratio is the one defined in the DVB-S2 standard for a code rate of 5/6, which yields 3.33 information bits sent per constellation symbol. The constrained capacity results for 16APSK presented in this section could be improved possibly by a few tenths of bits per channel usage (bpcu) outside of the 3.3 bpcu zone by altering the ring ratio. A similar situation holds for the 32APSK results, for which the ring ratios associated with a code rate of 5/6 were used (4.17 information bits per symbol). A discussion on how the ring ratios of these constellations were selected for the DVB-S2 standard so as to maximize their constrained joint capacity is presented in [GFM06].

For the two-site MRC situation the transmitter power was cut in half. This was done so that any advantage observed could be completely attributed to diversity gains.

Fig. 3.2 shows the results for the ergodic parallel decoding capacity for the QPSK, 8PSK, 16APSK and 32APSK constellations of the DVB-S2 standard for SISO and MRC channels for rain region P. For the MRC channel the rain fades were assumed to be uncorrelated. The plot also includes the Shannon capacity (i.e. Gaussian distribution on the input) as well as the clear sky (CS) (i.e. no rain) constrained capacity values, for the sake of comparison. As expected, for each M-ary constellation, the constrained capacities reach a maximum of $log_2(M)$ bits per channel usage. The figure shows that the clear-sky constrained capacities are only slightly higher than the ergodic constrained capacities of the MRC and SISO links, which are practically

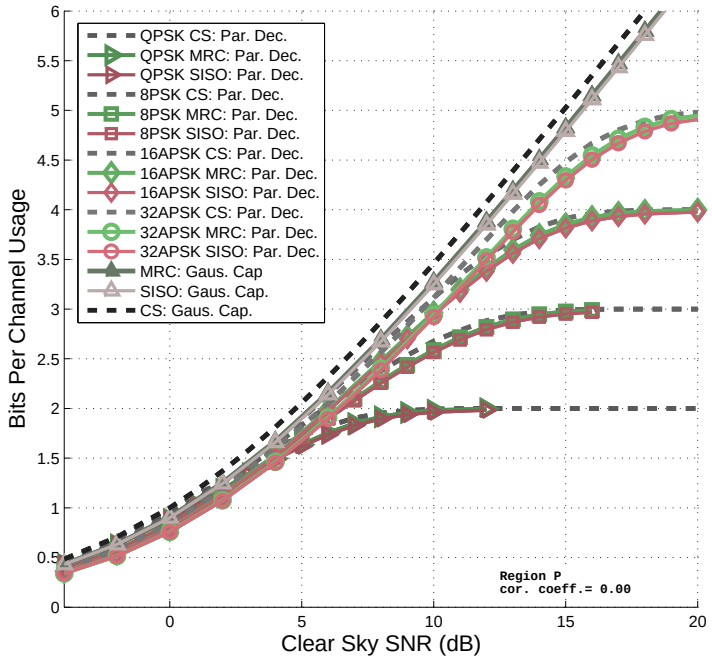


Figure 3.2: Average parallel decoding constrained capacities (*Par. Dec.*) in rain region P with QPSK, 8PSK, 16APSK and 32APSK. Clear sky (*CS*) counterparts and Shannon capacity curves (*Gauss. Cap.*) are also shown for reference.

the same. There is less than 1 dB difference between the CS values and the rain region P values. The ergodic capacity results for region E were found to be negligibly different to those of the clear sky situation.

The main advantage of using two-site receive diversity becomes evident when we consider the outage probability and outage capacity. Fig. 3.3 shows the outage probabilities for rain regions E and P when the channel input is constrained to QPSK or to 32APSK. For QPSK, the clear sky SNR was 6 dB and for 32APSK it was 16 dB. Both of these SNR values are at the upper range of the SNR values at which the DVB-S2 standard lists performance values for these constellations in a purely AWGN channel. The outage probabilities in regards to Shannon capacity are also shown in these figures for reference. The vertical axis provides the percentage of time that a given capacity value on the horizontal axis will be unobtainable due to fading. From these outage probability curves we see, as expected, the main advantage of using site diversity: it greatly increases the outage probability and capacity. For example, from Fig. 3.3

(c), the 1% outage capacity in region P with QPSK at a clear sky SNR of 6 dB is 0.6 bpcu for the SISO channel but increases significantly to 0.9 bpcu, 1.1 bpcu and 1.4 bpcu for the MRC channel with path fading correlation coefficients, ρ, of 0.25, 0.5 and 0 respectively. For the 0.1% outage capacity the situation is even more dramatic. For the situation just described, the SISO channel has an 0.1% outage capacity that is negligibly small, but the MRC channel has 0.1% outage capacities of 0.1 bpcu, 0.3 bpcu and 1.0 bpcu for correlation coefficients of 0.25, 0.5 and 0 respectively. These results show that site diversity greatly increases outage capacity, but that correlation of the rain fades on the two paths can significantly reduce the benefit, particulary for small outage percentages.

For the relatively dry rain region E, the outage capacity curves of Fig. 3.3 (a) show that there is no need for site diversity to increase the 1% outage capacity of QPSK, but that site diversity can help increase the 0.1% outage capacity.

The outage capacity curves for 32APSK of Fig. 3.3(b) and (d) show the same type of benefit obtained by using site diversity as that exemplified by the outage capacity curves of QPSK, i.e., significant increase in the outage capacity with site diversity, particularly at small outage percentages, and a fairly significant influence on those increases of the correlation of the fades of the two paths.

Outage probability curves were generated for the 8PSK and 16APSK constellations as well, but are omitted here due to space considerations. Outage capacity values derived from the outage probability curves for the four constellation types are provided in Table 3.1 for outage percentages of 0.1% and 1% at the upper range of the SNR values at which the DVB-S2 standard lists performance values for these constellations in a purely AWGN channel.

3.6 Conclusion and Future Work

The application of information theoretic analysis tools to the rain-faded Ka band satellite link has generated realistic capacity values that take into account the usage of practical constellations, as opposed to the Gaussian distributed inputs required to achieve the Shannon capacity.

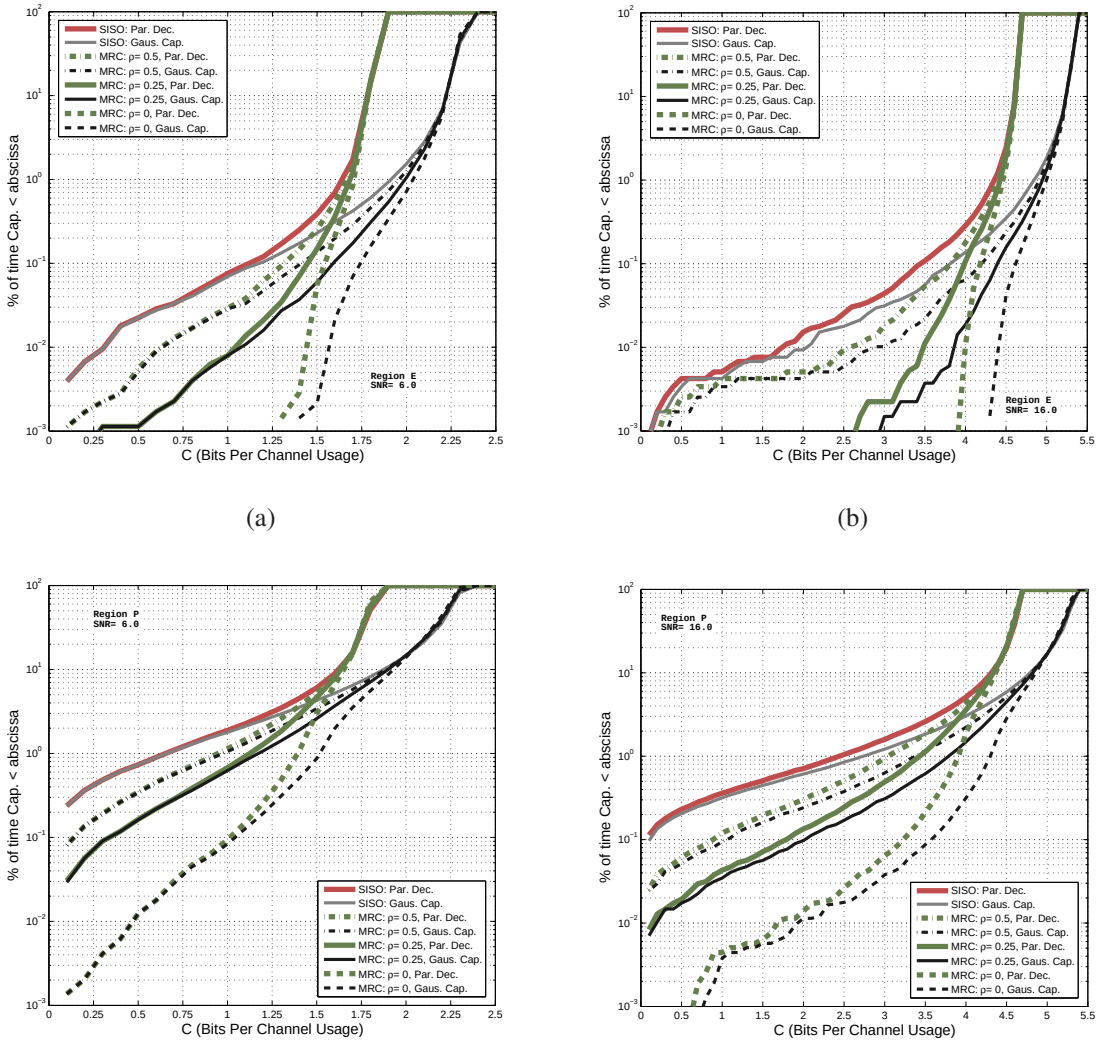


Figure 3.3: Outage percentages for parallel decoding capacity for 3 different correlation coefficients for (a) QPSK in region E, (c) QPSK in region P, (b) 32APSK in region E and (d) 32APSK in region P, with Shannon capacity (*Gaus. Cap.*) included for reference. The clear sky SNR was 6 dB for QPSK and 16 dB for 32APSK.

In order to quantify the benefits of diversity in mitigating rain fades, this was done for both the single site and the two-site diversity situations. This technique can be used by satellite link designers for a given set of link conditions (i.e. log-normal parameters m and σ) to examine the efficiency of a given satellite link in terms of how much of the constrained capacity the link is actually using, to determine the benefits of diversity methods, and to set realistic goals for

Table 3.1: Outage constrained capacities for various constellations.

Constell-ation	CS SNR (dB)	Channel Type	Cor. Coeff. ρ	0.1% Out. Cap. Reg. E (bpcu)	0.1% Out. Cap. Reg. P (bpcu)	1% Out. Cap. Reg. E (bpcu)	1% Out. Cap. Reg. P (bpcu)
QPSK	6	SISO	-	1.1	~ 0	1.6	0.6
QPSK	6	MRC	0.25	1.3	0.1	1.7	0.9
QPSK	6	MRC	0.5	1.4	0.3	1.7	1.1
QPSK	6	MRC	0	1.5	1.0	1.7	1.4
8PSK	11	SISO	-	2.1	~ 0	2.6	1.9
8PSK	11	MRC	0.25	2.3	0.4	2.7	2.1
8PSK	11	MRC	0.5	2.4	0.9	2.7	2.3
8PSK	11	MRC	0	2.5	1.9	2.7	2.4
16APSK	13	SISO	-	2.6	~ 0	3.4	1.7
16APSK	13	MRC	0.25	2.9	0.5	3.5	2.4
16APSK	13	MRC	0.5	3.1	1.2	3.5	2.6
16APSK	13	MRC	0	3.2	2.4	3.5	3.0
32PSK	16	SISO	-	3.4	~ 0	4.4	2.4
32APSK	16	MRC	0.25	3.8	0.9	4.5	3.0
32APSK	16	MRC	0.5	4.0	1.7	4.5	3.5
32APSK	16	MRC	0	4.2	3.2	4.5	3.9

adaptive modulation methods.

Future work in this area of studying the benefits of diversity in combatting rain fades may include the calculation of the constrained capacities with global fading processes, i.e., with the addition of scintillation effects and the increased noise temperatures due to rain fades.

Part II

Estimation of Ergodic Constrained Capacity and Outage Probabilities for Terrestrial Links

CHAPTER 4

Estimation of Constrained Capacity and Outage Probability of QAM Modulations in Rayleigh Channels

Contributions:

- Improved on past methods for approximating the ergodic constrained capacity (ECC) of noise limited Rayleigh channels by increasing the estimation accuracy while maintaining a low complexity.
- Extended the estimation of ECC of noise limited Rayleigh channels to include maximal ratio antenna combining (MRC).
- Introduced a new technique for estimating the outage probability of noise limited Rayleigh channels with and without MRC antenna combining.
- Extended the proposed method to the calculation of ECC and outage probability for interference limited Rayleigh channels that incorporate either MRC or optimum combining (OC).
- To date, no other method for calculating the ergodic constrained capacity of an interference limited Rayleigh channel or estimating the outage probabilities for the constrained capacity of a noise limited or interference limited Rayleigh channel has been published.

4.1 Introduction

The well-known Rayleigh fading channel model describes the gain of a non-line-of-sight multipath channel as a Rayleigh random variable. Researchers have determined methods for calculating the ergodic Shannon Capacity (ESC) for Rayleigh channels for single and multiple antenna systems [Tel99, AV08]. In these analyses the Shannon capacity assumes a continuously valued input signal with infinite support and Gaussian distribution. In practical communications systems, the input signal is constrained to a discrete signalling set such as a finite-size quadrature amplitude modulation (QAM) constellation. At high SNR the Shannon capacity greatly overestimates the capacity of these practical systems, particularly for low order constellations. For this reason a method is needed to evaluate the capacity and outage probability for Rayleigh channels when the signal set is constrained to a finite alphabet. This chapter 1) improves on past methods for approximating the ergodic constrained capacity (ECC) of noise limited Rayleigh channels by increasing the estimation accuracy while maintaining a low complexity, 2) extends the estimation of ECC of noise limited Rayleigh channels to include maximal ratio antenna combining (MRC), 3) introduces a new technique for estimating the outage probability of noise limited Rayleigh channels with and without MRC antenna combining, and 4) further extends the proposed method to the calculation of ECC and outage probability for interference limited Rayleigh channels that incorporate either MRC or optimum combining (OC). The method for the interference limited case can be used to estimate the throughput of cellular systems and determine channel reuse factors.

This chapter is organized as follows, section 4.2 gives the details of the noise limited Rayleigh channel model and provides the constrained capacity formula that is to be evaluated for the Rayleigh channel. Section 4.3 details the recent work done in determining estimates of, and bounds for, the ergodic constrained capacity for noise limited Rayleigh channels. Section 4.4 describes the approach taken to develop a method applicable to the calculation of 1) outage probabilities and 2) the ECC for noise limited Rayleigh channels with and without antenna combining and details the proposed method. Section 4.5 extends the proposed method to the interference limited Rayleigh channel in conjunction with OC and MRC combining. Section

4.6 provides results and section 4.7 lists conclusions and possible future work.

4.2 System Model and Performance Measures

The received signal y through the channel is modeled as

$$y = \sqrt{E_s} g\, x + n \tag{4.1}$$

where E_s is the average transmitted signal energy, $x \in X$ is the transmitted signal with constellation X, g is the Rayleigh distributed channel gain, and n is complex additive white Gaussian noise with zero mean and variance N_0.

The probability density function (PDF) of g is

$$f_G(g) = \frac{g}{\sigma_g^2} e^{-\frac{g^2}{2\sigma_g^2}} \tag{4.2}$$

where σ_g^2 is the variance of the Rayleigh random variable's underlying Gaussian process. Typically, the channel is modeled as having unit power gain. This requires that $\sigma_g^2 = 1/2$.

The Shannon capacity as a function of average SNR, $\gamma = E_s/N_0$ and g is

$$C_{Shan}(g,\gamma) = \lambda \ln(1 + g^2\gamma) \tag{4.3}$$

where constant $\lambda = 1/\ln(2)$ ($\lambda = 1$) if the capacity is expressed in bits/s/Hz (nats/s/Hz).

The constrained capacity, with a QAM input that has a finite alphabet of M equiprobable

values is [Bac99] (see also (C.6) in appendix C)

$$\begin{aligned} C_{con}(g,\gamma) &= H(X) - E_Y[H(X|y,g,\gamma)] \\ &= \lambda \ln(M) - \frac{\lambda}{M} \sum_{j=1}^{M} \frac{1}{\pi N_0} \int_{y \in \mathcal{C}} e^{-\Delta^2(y,x_j)/N_0} \\ &\quad \times \ln \left[\sum_{s=1}^{M} e^{(\Delta^2(y,x_j) - \Delta^2(y,x_s))/N_0)} \right] \mathrm{d}y \end{aligned} \tag{4.4}$$

where x_i is the complex value of input i, $\Delta^2(y,x_i) \doteq |y - x_i\sqrt{g^2\gamma N_0}|^2$ and $|\cdot|^2$ denotes squared Euclidean distance. The summation is over all the elements of discrete input alphabet X, and the integration is over all possible values of output Y, i.e. over the entire complex plane $\mathcal{C}$.

The ergodic capacity at a specific γ is, by definition,

$$E[C(\gamma)] \doteq \int_0^{\infty} x f_{C(\gamma)}(x)\, \mathrm{d}x \tag{4.5}$$

where $f_{C(\gamma)}(\cdot)$ is the PDF for the capacity.

For a Rayleigh channel (4.5) can be shown to be

$$E[C(\gamma)] = \frac{1}{\sigma_g^2} \int_0^{+\infty} C(g,\gamma) g e^{-\frac{g^2}{2\sigma_g^2}}\, \mathrm{d}g \tag{4.6}$$

where $C = C_{con}$ ($C = C_{Shan}$) to determine the ergodic constrained (Shannon) capacity.

Given a desired capacity value, c_o, there is a non-zero probability at any point in time that the gain, g, of the Rayleigh channel will be large enough to allow c_o to be achieved. The outage probability P_{out} is defined as the probability that the capacity, c, of the channel at a particular point in time will be less than c_o due to a channel fade. The outage probability as a function of desired capacity c_o is

$$P_{out}(c_o) \doteq P(c \le c_o) = P(\gamma_{\mathrm{rx}} \le \gamma_o) \doteq F_{\gamma_{rx}}(\gamma_o) \tag{4.7}$$

where γ_o is the received SNR required to achieve a capacity of c_o, $\gamma_{rx} = g^2\gamma$ is the received

SNR, and $F_{\gamma_{rx}}(\cdot)$ is its cumulative distribution function (CDF). A performance metric related to the outage probability is the outage capacity, $C_{out,P}(\gamma)$, which gives the capacity value available with an outage probability of P as a function of the average SNR γ.

4.3 State of the Art

Baccarelli developed upper and lower bounds for the ergodic constrained capacity of a Rayleigh channel (4.6) that remove the integration over the complex plane shown in (4.4). These bounds are readily evaluated for any QAM modulation type, but are relatively loose, i.e. 1 to 5 dB. The method presented in this chapter provides tight ($<$ 0.15 dB) estimations of the ergodic constrained capacity. He also looked at a problem related to estimating the outage probability of a Rayleigh channel for a given constrained capacity value [Bac99]. This chapter takes a direct approach to estimating the outage probability.

Wong developed a simple series representation for the ergodic constrained capacity of a Rayleigh fading channel (4.6) applicable only to BPSK and QPSK input constellations [Won01]. The method developed in this chapter is applicable to any QAM modulation type.

Slimane developed a simplified but accurate approximation of (4.4) for any QAM modulation that removes the integration over the complex plane [Sli06]. However, the subsequent evaluation of the ECC for the Rayleigh channel (4.6) still requires numeric integration. The method presented in this chapter replaces numerical integration with a finite sum.

None of the aforementioned techniques address antenna combining. The method presented in this chapter for calculating the ECC of Rayleigh channels works well for any QAM modulation, with or without antenna combining.

Regarding interference limited Rayleigh channels, no research has been published dealing with the estimation of ECC or of the outage probabilities associated with the constrained capacity for such channels. The method presented in this chapter provides a means for closely estimating both of these quantities for interference limited Rayleigh channels.

4.4 Proposed Method

The goal of this work was to develop a simple means to determine the outage probability of Rayleigh channels and to develop a readily evaluated expression that approximates (4.6) in conjunction with either (4.3) or (4.4). The existing literature substitutes (4.3) or (4.4) into (4.6) and, with the exception of a few special cases, resorts to numerical techniques for its evaluation. In this work an approximation to (4.4) is introduced which facilitates an accurate closed-form evaluation when inserted into (4.6).

4.4.1 Piece-wise Linear Fit of the Capacity

The long approximately linear sections and large radii of curvature of the constrained capacity function (4.4) make it an ideal candidate for piece-wise linear (PWL) curve fitting. Using $\gamma_{rx} \doteq g^2\gamma$, the received SNR, as the variable of the PWL fit facilitates the development of expressions for the ergodic capacity. Such a fit as a function of γ_{rx} can be expressed as

$$\begin{aligned} C(g,\gamma) &\cong C_{pwl}(g,\gamma) \\ &= m_k\gamma_{\text{rx}} + b_k \ \text{ for } c_{k-1} < C \leq c_k \end{aligned} \tag{4.8}$$

where m_k and b_k are the slope and intercept values for section k of the PWL fit, which apply to capacity values between c_{k-1} and c_k.

Using the method of [Tom74], PWL fits were made to the constrained capacity expression (4.4) for modulations commonly of interest. Figure 4.1 shows the constrained capacity and Shannon capacity curves and their PWL approximations. In all cases the fits were made so as to have a maximum absolute error of 0.01 bits/s/Hz. Since the constrained capacity of an M-ary constellation goes to $c_{max} = \log_2(M)$ as the SNR becomes large the capacity curve plateaus and thus the number of sections needed for the PWL fit to a constrained capacity curve is always finite.

Table 4.1 provides the corresponding slope and intercept values for each section of the PWL

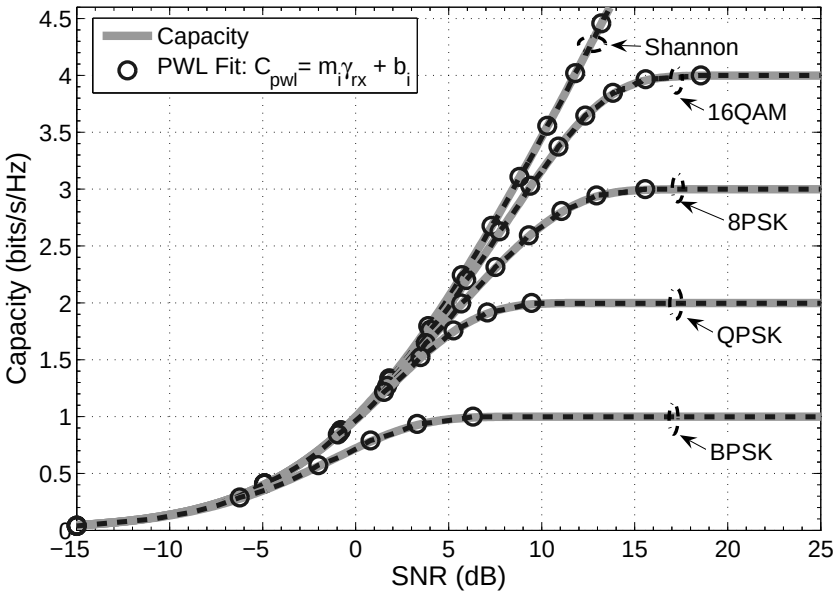


Figure 4.1: Piece-wise linear fits for BPSK, QPSK, 8PSK, and 16QAM constrained capacities and Shannon capacity.

fits of the constrained capacity curves portrayed in Figure 4.1. The maximum absolute error can be reduced further if desired at the expense of the addition of more sections to the PWL fit. Conversely, the PWL fit can be made with even fewer sections with a relaxed error limit.

4.4.2 PDF and CDF of the SNR

Consider the MRC combining of N receive antennas, of which the single antenna situation is a special case with N=1. The instantaneous SNR of the MRC signal, γ_{mrc}, is the sum of the SNRs, $\gamma_{rx,n}$, present at the individual antennas

$$\gamma_{mrc} = \sum_{n=1}^{N} \gamma_{rx,n} = \gamma \sum_{n=1}^{N} g_n^2 \tag{4.9}$$

where g_n is the Rayleigh-distributed channel gain associated with antenna n. Since each g_n is Rayleigh distributed the sum, $v \doteq \sum_n^N g_n^2$, is chi-squared distributed with $2N$ degrees of freedom

(a) BPSK

i	c_k (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.28982	1.2047	8.2634e-4
2	0.56367	0.72696	0.11543
3	0.77398	0.39564	0.31972
4	0.92035	0.17279	0.57558
5	1	0.042057	0.83644
6	–	0	1

(b) QPSK

i	c_k (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.41250	1.2729	5.9338e-4
2	0.83649	0.90071	0.12103
3	1.1961	0.61369	0.34902
4	1.4943	0.39361	0.65281
5	1.7255	0.22563	1.0120
6	1.8882	0.10740	1.3858
7	1.9808	0.033846	1.7299
8	2	0.0015203	1.9695
9	–	0	2

(c) 8PSK

i	c_k (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.41266	1.2734	5.9161e-4
2	0.84039	0.90865	0.11862
3	1.2353	0.63874	0.33302
4	1.5933	0.43337	0.62311
5	1.9236	0.28732	0.95006
6	2.2242	0.18514	1.2963
7	2.5025	0.11587	1.6435
8	2.7199	0.065848	2.0143
9	2.8773	0.03297	2.3666
10	2.9741	0.011397	2.7007
11	3	0.0008973	2.9526
12	–	0	3

(d) 16QAM

i	c_k (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.41283	1.2740	5.8967e-4
2	0.86003	0.91408	0.11704
3	1.2939	0.65116	0.33075
4	1.7161	0.45906	0.61489
5	2.1165	0.32499	0.93649
6	2.5208	0.23244	1.2726
7	2.9095	0.16247	1.6483
8	3.2375	0.11110	2.0471
9	3.5108	0.073085	2.4544
10	3.728	0.043057	2.8884
11	3.8863	0.021348	3.3117
12	3.9809	0.0068402	3.7022
13	4	0.00030194	3.9686
14	–	0	4

Table 4.1: Coefficients for the PWL fit (maximum error 0.01 bits/s/Hz) as a function of the SNR to constrained capacities (a) BPSK, (b) QPSK, (c) 8PSK, and (d) 16QAM.

and the PDF of γ_{mrc} is [Bre59],

$$f_{\gamma_{mrc,N}}(x) = \frac{x^{N-1}}{(2\gamma\sigma_g^2)^N \Gamma(N)} e^{-x/(2\gamma\sigma_g^2)} \tag{4.10}$$

where $\Gamma(\cdot)$ is the standard gamma function.

The associated CDF is

$$F_{\gamma_{mrc,N}}(x) = 1 - e^{-x/(2\gamma\sigma_g^2)} \sum_{n=0}^{N-1} \frac{x^n}{n!(2\gamma\sigma_g^2)^n}. \tag{4.11}$$

In order to examine diversity gains in isolation from aperture gains, the average transmitted

energy, E_s, and thus the average SNR, γ, can be divided by N. With this normalization (4.10) becomes

$$f_{\gamma_{mrc,N,div}}(x) = \frac{N^N x^{N-1}}{(2\gamma\sigma_g^2)^N \Gamma(N)} e^{-Nx/(2\gamma\sigma_g^2)} \tag{4.12}$$

and (4.12) becomes

$$F_{\gamma_{mrc,N,div}}(x) = 1 - e^{-x/(2\gamma\sigma_g^2)} \sum_{n=0}^{N-1} \frac{x^n}{n!(2\gamma\sigma_g^2)^n}. \tag{4.13}$$

The *div* subscript denotes that these distributions are applicable to examining diversity gains without the effect of aperture gains. As $N \to \infty$, (4.12) becomes a Dirac delta function centered at the average SNR γ (assuming $2\sigma_g^2 = 1$) and (4.13) becomes a step function at $x = \gamma$. Thus as $N \to \infty$ the Rayleigh channel with MRC and SNR γ/N approaches the unfaded AWGN case with SNR γ [Lee90]. This information will be helpful in examining the results in section VI.

4.4.3 Outage Probability Calculation Using PWL Fit

The outage probability associated with a given capacity value c_o can be evaluated by using the PWL fit of the capacity to estimate the SNR value γ_o needed to achieve c_o, and then using (4.11) to determine the probability that the actual SNR will be less than γ_o. The slope and intercept values can be found in Table 4.1 according to the modulation type. From these values γ_o can be estimated

$$\gamma_o \cong (c_o - b_k)/m_k \tag{4.14}$$

where the value of k is such that $c_{k-1} < c_o \leq c_k$. Using (4.7) and (4.14) the outage probability for an N antenna system with MRC combining can be estimated from

$$\begin{aligned} P_{out}(c_o) &= F_{\gamma_{mrc,N}}(\gamma_o) \\ &\cong F_{\gamma_{mrc,N}}((c_o - b_k)/m_k) \\ &= 1 - e^{-(c_o-b_k)/\beta_k} \sum_{n=0}^{N-1} \frac{(c_o-\beta_k)^n}{n!\beta_k^n} \end{aligned} \tag{4.15}$$

where $\beta_k = 2\sigma_g^2 m_k \gamma$.

4.4.4 Ergodic Capacity Calculation Using PWL Fit

Since C is a monotonic function of γ_{rx} the PDF of C over a Rayleigh channel with MRC antenna combining can be obtained from

$$f_C(c) = \left. \frac{f_{mrc,N}(\gamma_{rx})}{\left|\frac{dC}{d\gamma_{rx}}\right|} \right|_{\gamma_{rx}=C^{-1}(c)}. \tag{4.16}$$

For the constrained capacity case the derivative can be evaluated using (4.4), but this does not simplify the evaluation of (4.5) since the resulting expression for the derivative will still contain an integral over the complex plane and a double sum. Also, it is not apparent how the inverse $\gamma_{rx} = C^{-1}(c)$ can be obtained from (4.4). Using the piecewise linear fit (4.8), these two quantities can be approximated. For $c_{k-1} < c \leq c_k$

$$\left|\frac{dC}{d\gamma_{rx}}\right| \cong m_k \tag{4.17}$$

and

$$\gamma_{rx} = C^{-1}(c) \cong \frac{c - b_k}{m_k}. \tag{4.18}$$

Using (4.17) and (4.18) the PDF of (4.16) can be approximated for $c_{k-1} < c \leq c_k$ as

$$\begin{aligned} f_C(c) &\cong f_{C,N,k}(c) \\ &\doteq \frac{(c-b_k)^{N-1}}{(\beta_k)^N \Gamma(N)} e^{-\frac{c-b_k}{\beta_k}}. \end{aligned} \tag{4.19}$$

Using (4.19) in (4.5) results in a readily integrable expression. The PWL fit was made to γ_{rx} in order to take advantage of the well-known PDF of the SNR of the output of N-MRC combined antennas for a Rayleigh channel and to create an approximation for $f_C(\cdot)$ that facilitates the evaluation of (4.5). From (4.5) and (4.19) the ergodic capacity of a Rayleigh channel with N MRC combined antennas can be approximated as

$$\begin{aligned} E[C(\gamma)] &\cong \left(\sum_{k=1}^{K-1} \int_{c_{k-1}}^{c_k} c f_{C,N,k}(c)\, \mathrm{d}c \right) \\ &\quad + (1 - F_{\gamma_{mrc,N}}(\tilde{\gamma}_{max}))c_{max} \\ &= \left(\sum_{k=1}^{K-1} \int_{c_{k-1}}^{c_k} c \frac{(c-b_k)^{N-1}}{\beta_k^N \Gamma(N)} e^{-\frac{c-b_k}{\beta_k}}\, \mathrm{d}c \right) \\ &\quad + (1 - F_{\gamma_{mrc,N}}(\tilde{\gamma}_{max}))c_{max} \end{aligned} \tag{4.20}$$

where $c_0 = 0$, $c_{K-1} = c_{max}$, K is the number of segments in the PWL fit and $\tilde{\gamma}_{max}$ is an approximation to γ_{max} the SNR needed to achieve c_{max}, i.e. $\tilde{\gamma}_{max} = (c_{max} - b_{K-1})/m_{K-1}$. The $(1 - F_{\gamma_{rx}}(\tilde{\gamma}_{max}))c_{max}$ term accounts for a possible probability mass in the capacity PDF at $c = c_{max}$ due to the received SNR exceeding that required to achieve c_{max}. With an M-ary modulation $c_{max} = \log_2(M)$. The integral can be readily evaluated for various values of N through the use of the identity

$$\begin{aligned} \int_a^b x^n e^{-x}\, \mathrm{d}x &= -x^n e^x \Big|_{x=a}^{x=b} + n \int_a^b x^{n-1} e^{-x}\, \mathrm{dx} \\ &= -e^{-x} n! \sum_{i=0}^{n} \frac{x^i}{i!} \Bigg|_{x=a}^{x=b} \end{aligned} \tag{4.21}$$

giving

$$E[C(\gamma)] \cong \sum_{k=1}^{K-1} \phi_k(N) \; + \; c_{max} e^{-\tilde{\gamma}_{max}/(2\gamma\sigma_g^2)} \sum_{n=0}^{N-1} \frac{\tilde{\gamma}_{max}^n}{n!(2\gamma\sigma_g^2)^n} \tag{4.22}$$

where

$$\phi_k(N) = e^{-x}\left[\frac{\beta_k}{\Gamma(N)}x^N + (N\beta_k + b_k)\sum_{n=0}^{N-1}\frac{x^n}{n!}\right]\Bigg|_{x=\frac{c_{k-1}-b_k}{\beta_k}}^{x=\frac{c_k-b_k}{\beta_k}}.$$

Table 4.2 provides the evaluation of (4.22) for $1 \leq N \leq 4$.

Equations (4.15) and (4.22) can be evaluated for BPSK, QPSK, 8PSK and 16QAM using the coefficients of Table 4.1. These equations are also applicable to the evaluation of ergodic Shannon capacity with the use of a PWL fit of (4.3) over the range of SNR values of interest.

4.5 Interference Dominated Links

In wireless communication networks the presence of co-channel interference (CCI) will diminish link capacity. In this section the proposed method will be extended to determine the outage probability and constrained capacity of a CCI limited link in a Rayleigh channel.

An interference limited link over a Rayleigh channel with L co-channel interferers each of power P and $N \leq L$ optimally combined (OC) receiver antenna elements was examined in [SH98], and the MRC case was developed in [SH00]. As pointed out in [SH98], although the equal power interferers model is not as realistic as an exponential decay model its use lends itself to analysis and allows the development of closed form expressions. If $P = \max_{k \leq L}(P_k)$ ($P = \min_{k \leq L}(P_k)$), where P_k is the power of the kth interfering signal, then the results using this model will provide a lower (upper) bound for the performance of the link. The following

N	$E[C]$
1	$\sum_{k=1}^{K-1} e^{-x}[\beta_k(x+1)+b_k]\Big\|_{x=\frac{c_k-b_k}{\beta_k}}^{x=\frac{c_{k-1}-b_k}{\beta_k}}$ $+e^{-\frac{\tilde{\gamma}_{max}}{2\gamma\sigma_g^2}} c_{max}$
2	$\sum_{k=1}^{K-1} e^{-x}[\beta_k x^2+(2\beta_k+b_k)(x+1)]\Big\|_{x=\frac{c_k-b_k}{\beta_k}}^{x=\frac{c_{k-1}-b_k}{\beta_k}}$ $+e^{-\frac{\tilde{\gamma}_{max}}{2\gamma\sigma_g^2}}(\frac{\tilde{\gamma}_{max}}{2\gamma\sigma_g^2}+1)\, c_{max}$
3	$\sum_{k=1}^{K-1} e^{-x}[\frac{\beta_k}{2}x^3+(3\beta_k+b_k)(\frac{1}{2}x^2+x+1)]\Big\|_{x=\frac{c_k-b_k}{\beta_k}}^{x=\frac{c_{k-1}-b_k}{\beta_k}}$ $+e^{-\frac{\tilde{\gamma}_{max}}{2\gamma\sigma_g^2}}(\frac{1}{2}(\frac{\tilde{\gamma}_{max}}{2\gamma\sigma_g^2})^2+\frac{\tilde{\gamma}_{max}}{2\gamma\sigma_g^2}+1)\, c_{max}$
4	$\sum_{k=1}^{K-1} e^{-x}[\frac{\beta_k}{6}x^4+(4\beta_k+b_k)(\frac{1}{6}x^3+\frac{1}{2}x^2+x+1)]\Big\|_{x=\frac{c_k-b_k}{\beta_k}}^{x=\frac{c_{k-1}-b_k}{\beta_k}}$ $+e^{-\frac{\tilde{\gamma}_{max}}{2\gamma\sigma_g^2}}(\frac{1}{6}(\frac{\tilde{\gamma}_{max}}{2\gamma\sigma_g^2})^3+\frac{1}{2}(\frac{\tilde{\gamma}_{max}}{2\gamma\sigma_g^2})^2+\frac{\tilde{\gamma}_{max}}{2\gamma\sigma_g^2}+1)\, c_{max}$

Table 4.2: Evaluations of (4.20) for $N = 1$, 2, 3 and 4 based on a PWL fit to the capacity as a function of received SNR.

development deals with links that are CCI limited and so, to simplify analysis, the effect of thermal noise is assumed to be negligible and is ignored. The received signal y through the channel is modeled as

$$y = \sqrt{P_s}\, g_0 x_0 + \sqrt{P} \sum_{k=1}^{L} g_k x_k \tag{4.23}$$

where x_0 and x_k are the desired and the kth interfering signals, respectively. The power of the desired signal is P_s and the power of each of the cochannel interferers is P. The complex vectors g_0 and g_k, $k = 1, ..., L$, are independently identically distributed, with elements that have magnitudes which are Rayleigh distributed according to (4.2).

When $N > L$ there are sufficient degrees of freedom (DOF) to reject the interferers resulting, effectively, in a noise limited Rayleigh channel and the method of the previous section can be applied, taking into account the reduction in DOF after the interference cancelation to $N-L-1$. In this section, as noted above, the link is assumed to be CCI limited which implies that the

number of antenna elements N and the number of cochannel interferers L are related by $L \geq N$.

4.5.1 PDF and CDF of the SIR With OC and MRC

The PDF of the signal-to-interference ratio (SIR), μ, of such a CCI limited link with OC combining is [SH98]

$$\begin{aligned} f_{oc,\mu}(\mu) &= \frac{\Gamma(L+1)}{\Gamma(N)\Gamma(L-N+1)}\rho^{L-N+1} \\ &\quad \cdot\frac{\mu^{N-1}}{(\rho+\mu)^{L+1}},\ \forall\,\mu \geq 0,\ 1 \leq N \leq L, \end{aligned} \tag{4.24}$$

where N is the number of antennas combined, L is the number of equal power co-channel interferers and $\rho = P_s/P$.

The associated CDF is [SH98]

$$\begin{aligned} F_{oc,\mu}(\mu) &= \frac{\Gamma(L+1)}{\Gamma(N+1)\Gamma(L-N+1)}\left(\frac{\mu}{\rho}\right)^N \\ &\quad \cdot {}_2F_1(L+1,N;N+1;-\mu/\rho) \end{aligned} \tag{4.25}$$

where ${}_2F_1\left(a,b;c;x\right)$ is the hypergeometric function as defined in [GR65, eqn. 9.100]. If the hypergeometric function is available by a given software package, then (4.25) can be used. If not then the following, developed in this work, can be used in evaluating the outage probability

$$\begin{aligned} F_{oc,\mu}(\mu) &= 1 - \frac{\Gamma(L+1)\rho^{L-N+1}}{\mu(\rho+\mu)^L} \\ &\quad \cdot\sum_{k=0}^{N-1}\frac{\rho^k\mu^{N-k}}{\Gamma(N-k)\Gamma(L-N+k+2)}. \end{aligned} \tag{4.26}$$

The PDF of the SIR of such a CCI limited link with MRC is [SH00]

$$f_{mrc,\mu}(\mu) = \frac{\Gamma(L+N)}{\Gamma(L)\Gamma(N)}\rho^L\frac{\mu^{N-1}}{(\rho+\mu)^{L+N}}. \tag{4.27}$$

The associated CDF is [SH00]

$$F_{mrc,\mu}(\mu) = \frac{\Gamma(L+N)}{\Gamma(L)\Gamma(N+1)}\left(\frac{\mu}{\rho}\right)^N \cdot {}_2F_1(L+N,N;N+1;-\mu/\rho). \quad (4.28)$$

An alternate expression developed in this work is

$$F_{mrc,\mu}(\mu) = 1 - \frac{\Gamma(L+N)\rho^L}{\mu(\rho+\mu)^{L+N-1}} \cdot \sum_{k=0}^{N-1} \frac{\rho^k \mu^{N-k}}{\Gamma(N-k)\Gamma(L+k+1)}. \quad (4.29)$$

4.5.2 Outage Probability Calculation for Interference Limited Link Using PWL

In [LH02] it was demonstrated that the interference can be represented accurately as having a Gaussian distribution. Under this assumption the constrained capacity expression of (4.4) as well as the PWL fits of Table 4.1 are applicable to the interference limited case with SNR replaced by SIR.

The outage probability associated with a given capacity value c_o can be evaluated by using the PWL fit of the capacity to estimate the SIR value μ_o needed to achieve c_o, and then using (4.25) or (4.26) for the OC case and (4.28) or (4.29) for the MRC case to determine the probability that the actual SIR will be less than μ_o. The slope and intercept values can be found in Table 4.1 according to the modulation type. From these values μ_o can be estimated

$$\mu_o \cong (c_o - b_k)/m_k \quad (4.30)$$

where the value of k is such that $c_{k-1} < c_o \leq c_k$. Using (4.26) and (4.30) the outage probability for an interference limited link with L equal power interferers and N optimally combined

antennas can be estimated from

$$\begin{aligned} P_{oc,out}(c_o) &= F_{oc,\mu}(\mu_o) \\ &\cong F_{oc,\mu}((c_o-b_k)/m_k) \\ &= \frac{\Gamma(L+1)}{\Gamma(N+1)\Gamma(L-N+1)}\left(\frac{(c_o-b_k)}{m_k\rho}\right)^N \\ &\quad \cdot {}_2F_1\left(L+1,N;N+1;-\frac{(c_o-b_k)}{m_k\rho}\right) \\ &= 1-\Gamma(L+1)\rho^{L-N+1}\left(\rho+\frac{c_o-b_k}{m_k}\right)^{-L} \\ &\quad \cdot\sum_{k=0}^{N-1}\frac{\rho^k((c_o-b_k)/m_k)^{N-k-1}}{\Gamma(N-k)\Gamma(L-N+k+2)}. \end{aligned} \tag{4.31}$$

Following a similar development, using (4.29) and (4.30) the outage probability for the MRC case can be estimated from

$$\begin{aligned} P_{mrc,out}(c_o) &= F_{mrc,\mu}(\mu_o) \\ &\cong F_{mrc,\mu}((c_o-b_k)/m_k) \\ &= \frac{\Gamma(L+N)}{\Gamma(L)\Gamma(N+1)}\left(\frac{c_o-b_k}{m_k\rho}\right)^N \\ &\quad \cdot {}_2F_1\left(L+N,N;N+1;-\frac{c_o-b_k}{m_k\rho}\right) \\ &= 1-\frac{\Gamma(L+N)\rho^L}{(\rho+(c_o-b_k)/m_k)^{(L+N-1)}} \\ &\quad \cdot\sum_{k=0}^{N-1}\frac{\rho^k((c_o-b_k)/m_k)^{N-k-1}}{\Gamma(N-k)\Gamma(L+k)}. \end{aligned} \tag{4.32}$$

4.5.3 Ergodic Capacity Calculation for Interference Limited Link Using PWL Fit

The PDF of the constrained capacity C as a function of the SIR, μ, is

$$f_C(c) = \left.\frac{f_\mu(\mu)}{\left|\frac{dC}{d\mu}\right|}\right|_{\mu=C^{-1}(c)}. \tag{4.33}$$

Using (4.17) and (4.18) the PDF of (4.33) for the OC case can be approximated for $c_{k-1} < c \leq c_k$ as

$$\begin{aligned} f_C(c) &\cong f_{C,k}(c) \\ &\doteq \frac{\Gamma(L+1)}{\Gamma(N)\Gamma(L-N+1)}\,\alpha_k^{L-N+1} \\ &\quad \cdot\frac{(c-b_k)^{N-1}}{(\alpha_k+c-b_k)^{L+1}} \end{aligned} \tag{4.34}$$

where $\alpha_k = m_k P_s/P$.

From (4.5) and (4.34) the ergodic capacity of a Rayleigh channel with L equal powered interferers and N OC combined antenna elements is

$$\begin{aligned} E[C_{oc}(\rho)] &\cong \left(\sum_{k=1}^{K-1}\int_{c_{k-1}}^{c_k} c f_{C,k}(c)\,\mathrm{d}c\right) \\ &\quad +(1-F_\mu(\tilde{\mu}_{max}))c_{max} \\ &= \sum_{k=1}^{K-1}\int_{c_{k-1}}^{c_k} c\,\frac{\Gamma(L+1)}{\Gamma(N)\Gamma(L-N+1)}\,\alpha_k^{L-N+1} \\ &\quad \cdot\frac{(c-b_k)^{N-1}}{(\alpha_k+c-b_k)^{L+1}}\;\mathrm{d}c \\ &\quad +(1-F_\mu(\tilde{\mu}_{max}))c_{max} \end{aligned} \tag{4.35}$$

where $c_0 = 0$, $c_{K-1} = c_{max}$, K is the number of segments in the PWL fit and $\tilde{\mu}_{max}$ is an approximation to μ_{max} the SIR needed to achieve c_{max}, i.e. $\tilde{\mu}_{max} = (c_{max} - b_{K-1})/m_{K-1}$. The $(1-F_\mu(\tilde{\mu}_{max}))c_{max}$ term accounts for a possible probability mass in the capacity PDF at c_{max} due to the SIR exceeding that required to achieve c_{max}. The integral can be readily evaluated for various values of L and N through the use of the identity [GR65, eqn. 2.111.2]

$$\begin{aligned} \int \frac{x^n}{(a+bx)^m}\;\mathrm{d}x &= \frac{-x^n}{(a+bx)^{m-1}(m-n-1)b} \\ &\quad +\frac{na}{(m-n-1)b}\int \frac{x^{n-1}}{(a+bx)^m}\;\mathrm{d}x \\ &= \frac{-\Gamma(n+1)\,x^n}{(a+bx)^{m-1}}\sum_{i=0}^{n}\frac{\Gamma(m-n-1)x^{-i}}{\Gamma(n-i+1)\Gamma(m-n+i)}\,\frac{a^i}{b^{i+1}} \end{aligned} \tag{4.36}$$

giving, for $L > N$,

$$E[C_{oc}(\rho)] \cong \sum_{k=1}^{K-1} \varphi_k(N,L) + c_{max} \frac{\Gamma(L+1)\rho^{L-N+1}}{\tilde{\mu}_{max}(\rho+\tilde{\mu}_{max})^L} \cdot \sum_{k=0}^{N-1} \frac{\rho^k \tilde{\mu}_{max}^{N-k}}{\Gamma(N-k)\Gamma(L-N+k+2)} \tag{4.37}$$

where

$$\varphi_k(N,L) = \Gamma(L+1)\frac{v^N \alpha_k^{L-N+1}}{(\alpha_k+v)^L}\left[\frac{1}{(L-N)\Gamma(N)} \cdot \frac{1}{\Gamma(L-N+1)} + \left(\frac{N}{L-N}+\frac{b_k}{\alpha_k}\right) \cdot \left(\sum_{i=1}^{N} \frac{\alpha_k^i v^{-i}}{\Gamma(N-i+1)\Gamma(L-N+i+1)}\right)\right]\Bigg|_{v=c_k-b_k}^{v=c_{k-1}-b_k} \tag{4.38}$$

The case for $L = N$ has been derived in a similar fashion through the use of an alternate form of (4.36), given in [GR65, eqn. 2.111.3], but is omitted here.

Following a similar development, the ergodic capacity of a Rayleigh channel with L equal power interferers and N MRC combined antenna elements is, for $L \geq N$

$$E[C_{mrc}(\rho)] \cong \sum_{k=1}^{K-1} \psi_k(N,L) + c_{max} \frac{\Gamma(L+N)\rho^L}{\tilde{\mu}_{max}(\rho+\tilde{\mu}_{max})^{L+N-1}} \cdot \sum_{k=0}^{N-1} \frac{\rho^k \tilde{\mu}_{max}^{N-k}}{\Gamma(N-k)\Gamma(L+k+1)} \tag{4.39}$$

where

$$\psi_k(N,L) = \Gamma(L+N)\frac{v^N\alpha_k^L}{(\alpha_k+v)^{L+N-1}}\left[\frac{1}{(L-1)\Gamma(L)\Gamma(N)} + \left(\frac{N}{L-1}+\frac{b_k}{\alpha_k}\right) \cdot\left(\sum_{i=1}^{N}\frac{\alpha_k^i v^{-i}}{\Gamma(N-i+1)\Gamma(L+i)}\right)\right]\Bigg|_{v=c_k-b_k}^{v=c_{k-1}-b_k}. \quad (4.40)$$

Equations (4.31), (4.32), (4.37), and (4.39) can be evaluated for BPSK, QPSK, 8PSK and 16QAM using the coefficients of Table 4.1. These equations are also applicable to the evaluation of ergodic Shannon capacity with the use of a PWL fit of (4.3) over the range of SIR values of interest.

4.6 Results

Figure 4.2 shows the match between ergodic capacity values obtained from the proposed expression (4.22) as evaluated in Table 4.2 and from Monte Carlo simulations for BPSK, QPSK, 8PSK, 16QAM and Gaussian (i.e. the Shannon case) alphabets for $N = 1$, 2 and 4 antennas. The transmitted power was divided by N in order to study the diversity gains as a function of N in isolation from aperture gains. The match is within 0.01 bits/s/Hz at any SNR, and within 0.15 dB at any capacity value between 0.2 and $0.99c_{max}$ bits/s/Hz across all modulations types and values of N. As expected based on [Lee90] and the comments following (4.13), there is a decreasing marginal increase in ergodic capacity with the addition of antennas as the ECC approaches the AWGN constrained capacity.

Figure 4.3 shows the outage probability values, P_{out}, obtained from the proposed expression (4.15) and from Monte Carlo simulations for BPSK, QPSK, 8PSK, 16QAM and Gaussian input alphabets for $N = 1$, 2, 3, 4 and 100 antennas at an SNR of 15 dB. At a given outage probability value, the estimate from the proposed expression agrees within the error of the PWL fit, 0.01

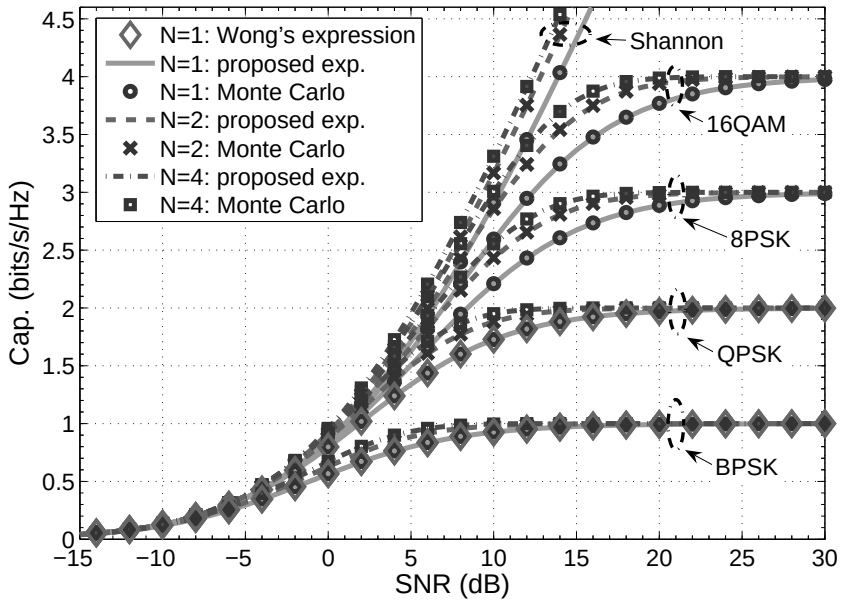


Figure 4.2: Ergodic constrained capacity for BPSK, QPSK, 8PSK and 16QAM in Rayleigh channel for single antenna and for MRC combining with two and four antennas comparing values from the proposed expression and values from Monte Carlo simulations.

bits/s/Hz in this example, of the results from Monte Carlo simulations and numerical methods. As $N \to \infty$ the outage probability approaches a step function at the constrained capacity associated with the average SNR γ as anticipated from the comments following (4.13).

Figure 4.4 shows the outage constrained capacity, $C_{out,P}(\gamma)$, for 16QAM associated with $P = 0.10\%$ and $P = 10\%$ outage rates with MRC antenna combining for $N = 1, 2, 3$ and 4. These values were determined by using (4.15) to generate outage probability curves for several average SNR values and then determining from each of those curves the capacity value associated with the desired outage percentage.

For the interference dominated case, ergodic capacity values were obtained using the proposed expressions (4.37) and (4.39) for OC and MRC combining, respectively, and by Monte Carlo simulations for QPSK, 8PSK, 16QAM and Gaussian (i.e. Shannon) modulations with values of $1 \leq N \leq 7$ antennas and $N+1 \leq L \leq 8$ equal power interferers. The results from the proposed expressions matched the Monte Carlo results within 0.01 bits/s/Hz at any P_s/P ratio, and within 0.4 dB of P_s/P at any capacity value between 0.2 and $0.99c_{max}$ bits/s/Hz. Figure

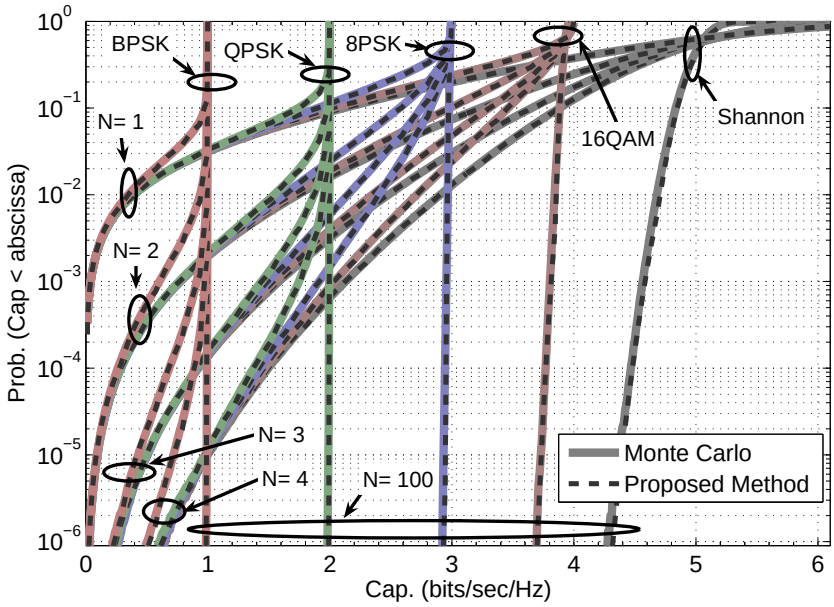


Figure 4.3: Outage probabilities for BPSK, QPSK, 8PSK and 16QAM constrained capacities and also Shannon capacity (Gaussian-distributed input) in Rayleigh channel with N= 1, 2, 3, 4, and 100 MRC-combined receive antennas obtained from the proposed expression and from Monte Carlo simulations (100 million trials). The average SNR, γ, out of the combiner was 15 dB in all cases.

4.5 shows the ergodic capacity values obtained from the proposed expressions and from Monte Carlo simulations for 16QAM with $L = 8$ interferers and $N = 1$, 3 and 6 antennas.

Figure 4.6 shows the outage probability values, P_{out}, obtained from the proposed expressions (4.31) and (4.32) for OC and MRC, respectively, and from Monte Carlo simulations for QPSK, 16QAM and Gaussian input alphabets for $N = 2$ and 4 antennas and $L = 6$ equal power interferers with $P_s/P = 20$ dB. At a given outage probability value, the estimate from the proposed expression agrees within the error of the PWL fit, 0.01 bits/s/Hz in this example, of the results from Monte Carlo simulations and numerical methods. This type of data can be useful to a system designer in determining channel reuse factors, the number of antennas required and whether the extra expense of implementing OC instead of MRC is merited.

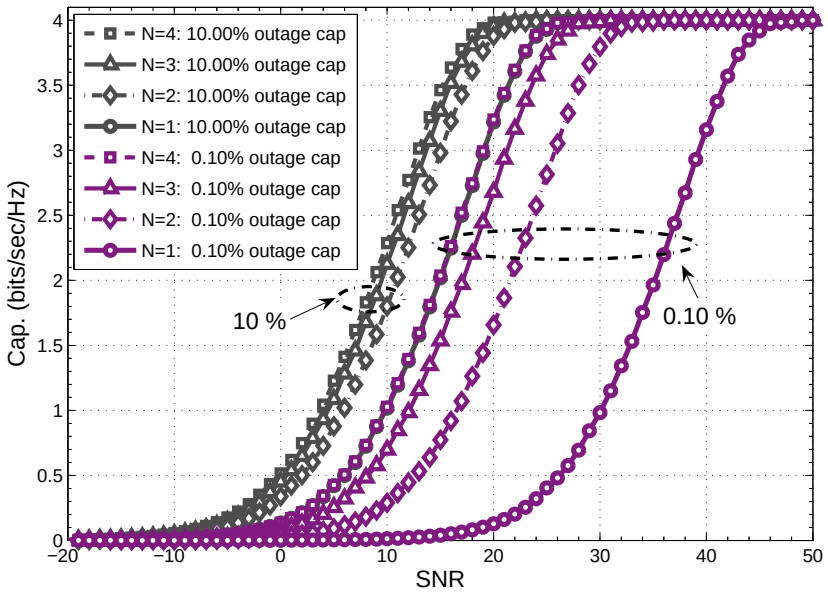


Figure 4.4: Outage capacity for 0.10% and 10% outages with 16QAM in Rayleigh channel with N= 1, 2, 3, and 4 MRC-combined receive antennas obtained from the proposed expression.

4.7 Conclusion

This chapter presented a new method to quickly and accurately estimate the ergodic capacity, outage probabilities, and outage capacity of a Rayleigh channel, with and without antenna combining, for links constrained to a finite alphabet, without resorting to lengthy Monte Carlo simulations or numerical integration. Both the noise limited case and the interference limited case were considered. This new method was shown to match well with Monte Carlo simulation results and with Wong's results for BPSK and QPSK. The method was shown to be applicable to ergodic Shannon capacity and outage probability calculations as well. The method for the interference limited case can serve as an analysis tool for a cellular system designer.

The proposed method is applicable to any QAM modulation once the parameters of the PWL fit of (4.4) are obtained via some method such as that of [Tom74]. For BPSK, QPSK, 8PSK and 16QAM, Table 4.1 provides the required parameter values.

The proposed method readily lends itself to the evaluation of the ergodic constrained capacity and outage probabilities for other channel types, e.g. MIMO systems with transmit-

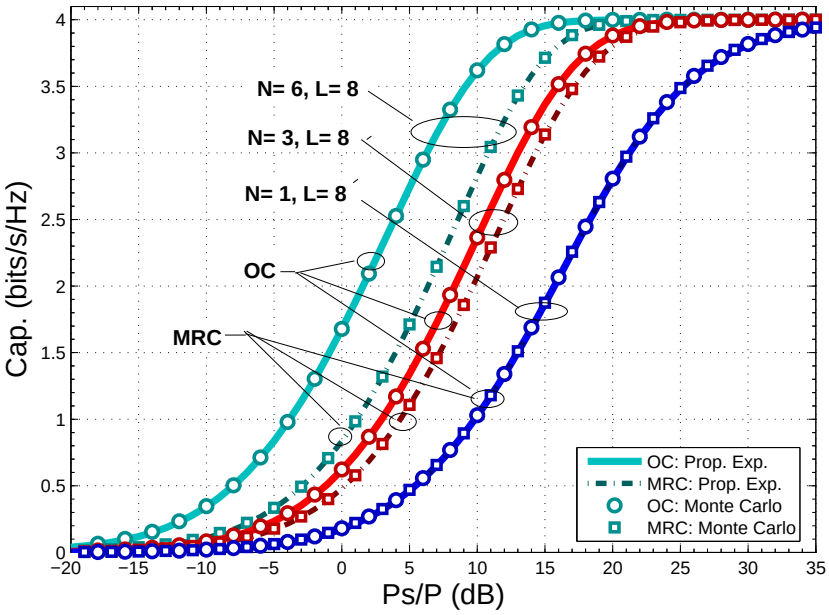


Figure 4.5: Ergodic constrained capacity for 16QAM in Rayleigh channel with OC and MRC combining of N antennas for interference limited case with $L = 8$ equal power interferers each of power P, the signal of interest has power P_s.

beamforming and MRC combining at the receiver over Rayleigh channels [MA05], by replacing the PDF and CDF of the SNR at the receiver in (4.10) and (4.11), respectively, with the appropriate expressions for the channel of interest. In addition, the PWL fit (4.8) of the constrained capacity and associated coefficients of Table 4.1 could be used in forming the objective function for a cellular system's power control algorithm similar to that in [QC99]. The application of the proposed method to other channel types and the use of (4.8) in a power control algorithm are two areas of possible future work.

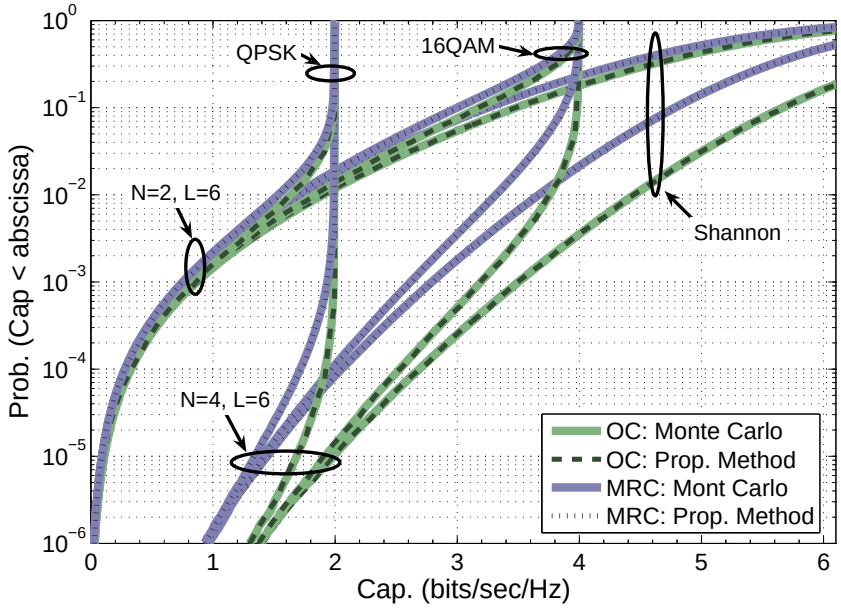


Figure 4.6: Outage probabilities for QPSK and 16QAM constrained capacities along with Shannon capacity in an interference limited Rayleigh channel with OC and MRC combining of $N = 2$ and 4 antennas with $L = 6$ equal power interferers obtained from the proposed expressions and from Monte Carlo simulations (100 million trials). The P_s/P ratio was 20 dB.

CHAPTER 5

Estimation of Constrained Capacity and Outage Probability of QAM Modulations in Log-Normal Channels

Contributions:

- Developed a simple but accurate method for calculating both 1) the ergodic Shannon capacity and 2) the ergodic constrained capacity of practical signals for log-normal channels.
- Introduced a new technique for estimating the outage probability of log-normal channels.
- To date, the ergodic constrained capacity and outage probability of practical signals for log-normal channels had not been dealt with in the literature.

5.1 Introduction

The log-normal distribution is often used in wireless communications to model mid-scale fading caused by shadowing in outdoor environments. It has also been applied more recently to describe small-scale fading of indoor ultra-wideband channels. Formulas for estimating the ergodic Shannon capacity of a log-normal (LN) channel have been recently presented in the literature [LSA07, HCH09]. In these analyses the Shannon capacity assumes a continuously valued input signal with infinite support and Gaussian distribution. In practical communications systems, the input signal is constrained to a discrete signalling set such as finite-size quadrature amplitude modulation (QAM) constellations. At high SNRs the Shannon capacity greatly over-estimates the capacity of these systems, particularly for low order constellations. For this

reason a method is needed to evaluate the ergodic capacity and outage probability for the LN channel when the signal set is constrained to a finite alphabet. This chapter proposes such a method.

Ergodic capacity and outage probability are important performance measures of communication systems in fading channels. For fast fading channels, for which the range of channel fades are experienced during the time period covered by a practical interleaver, the ergodic capacity determines the rate at which data can be sent error free. For slow fading channels, in which it is impractical to interleave over a time interval long enough to cover the range of channel fades, the ergodic capacity provides the average rate at which data can be transmitted error free using adaptive coded modulation (ACM) in which the coding rate and/or modulation order is adjusted in accordance with the channel conditions. The ergodic capacity also is used in determining the area spectral efficiency for mobile cellular systems [AG99]. The outage probability provides the percentage of the time the link will be unable to transmit error free at a given data rate, a particularly useful measure for slow fading channels.

This chapter is organized as follows, section 5.2 gives the details of the channel model and provides the Shannon and constrained capacity formulas that are to be evaluated for the LN channel. Section 5.3 details the recent work done in determining the ergodic Shannon capacity for LN channels. Section 5.4 describes the approach taken to develop a method applicable to the calculation of 1) outage probabilities and 2) ergodic values, for both Shannon and constrained capacity, of LN channels and details the proposed method. Section 5.5 provides results and section 5.6 lists conclusions and future work.

5.2 System Model and Performance Measures

The received signal y through the channel is modeled as,

$$y = \sqrt{E_s}\alpha x + n \tag{5.1}$$

where E_s is the average transmitted signal energy, x is the transmitted signal, α is the log-normally distributed channel power gain, and n is additive white Gaussian noise with a variance of N_0.

The probability density function (PDF) of the channel power gain is,

$$f(\alpha) = \frac{1}{\alpha\sqrt{2\pi\sigma_\alpha^2}} e^{-\frac{(\ln(\alpha)-\mu_\alpha)^2}{2\sigma_\alpha^2}} \tag{5.2}$$

where μ_α and σ_α^2 are the mean and variance of $\ln(\alpha)$. Typically the channel is normalized to have $E[\alpha]$= 1, which requires that $\mu_\alpha = -\sigma_\alpha^2/2$.

The Shannon capacity as a function of average SNR, $\gamma = E_s/N_0$, and α is,

$$C_{Shan}(\alpha,\gamma) = \lambda \ln(1+\alpha\gamma) \tag{5.3}$$

where constant λ= $1/\ln(2)$ (λ= 1) if the capacity is expressed in bits/s/Hz (nats/s/Hz).

The constrained capacity, with a QAM input that has a finite alphabet of M equiprobable values is [Bac99] (see also (C.6) in appendix C)

$$\begin{aligned} C_{con}(\alpha,\gamma) &= H(X) - H(X|Y,\alpha,\gamma) \\ &= \lambda\ln(M) - \frac{\lambda}{M}\sum_{j=1}^{M}\frac{1}{\pi N_0}\int_{y\in C} e^{-\Delta^2(y,x_j)/N_0} \\ &\quad \times \ln\left[\sum_{s=1}^{M} e^{(\Delta^2(y,x_j)-\Delta^2(y,x_s))/N_0)}\right] \mathrm{d}y \end{aligned} \tag{5.4}$$

where, x_i is the complex value of input element i, $\Delta^2(y,x_i) \doteq \| y - x_i\sqrt{\alpha\gamma N_0}\|^2$ and $\|\cdot\|^2$ denotes squared Euclidean distance. The summation is over all the elements of discrete input alphabet X, and the integration is over all possible values of output Y, i.e., over the entire complex plane C.

The ergodic capacity at a specific γ is, by definition,

$$E[C(\gamma)] \doteq \int_0^\infty x f_{C(\gamma)}(x)\mathrm{d}x \tag{5.5}$$

where $f_{C(\gamma)}(\cdot)$ is the PDF for the capacity at average SNR γ.

For a LN channel (5.5) can be shown to be,

$$E[C(\gamma)] = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \int_0^{+\infty} \frac{C(\alpha,\gamma)}{\alpha} e^{-\frac{(\ln(\alpha)-\mu_\alpha)^2}{2\sigma_\alpha^2}} \mathrm{d}\alpha \tag{5.6}$$

where $C = C_{Shan}$ ($C = C_{con}$) to determine the ergodic Shannon (constrained) capacity.

Given a desired capacity value, c_o, there is a non-zero probability at any point in time that the gain, α, of the LN channel will be large enough to allow c_o to be achieved. The outage probability P_{out} is defined as the probability that the capacity, c, of the channel at a particular point in time will be less than c_o due to a channel fade. The outage probability as a function of desired capacity c_o is,

$$P_{out}(c_o) \doteq P(c \le c_o) = P(\gamma_{\mathrm{rx}} \le \gamma_o) \doteq F_{\gamma_{rx}}(\gamma_o) \tag{5.7}$$

where γ_o is the received SNR required to achieve a capacity of c_o, $\gamma_{rx} = \alpha\gamma$ is the received SNR, and $F_{\gamma_{rx}}(\cdot)$ is its cumulative distribution function (CDF).

A performance metric related to the outage probability is the outage capacity, $C_{out,P}(\gamma)$, which gives the capacity value available with an outage probability of P as a function of the average SNR γ.

In wireless communications it is convenient to express the standard deviation of the channel gain in dB. Define parameters (μ_y, σ_y) to characterize the LN distribution, where $\sigma_y = \kappa\sigma_\alpha$ ($\mu_y = \kappa\mu_\alpha$) is the standard deviation (average value) in dB of the channel power gain and $\kappa = 10/\ln(10)$. Typical values for σ_y are, 6 to 12 dB for medium-scale fading, 3 to 5 dB for single-input single-output indoor ultra-wideband (UWB) channels and 1 to 4 dB for multiple-input multiple-output indoor UWB channels (see [HCH09] and references therein).

5.3 State of the Art

Upper and lower bounds for (5.6) were derived in [AG99] for Shannon capacity. As mentioned in [AG99] and quantified in [LSA07] these bounds are loose for low SNR values, e.g. at a SNR of 7 dB the upper bound is off by 3 dB and the lower bound is off by 1 dB. The bounds become looser as the SNR is decreased.

In [LSA07] Laourine developed an accurate approximation of (5.6) for the Shannon capacity,

$$\begin{aligned} E[C] \;\cong\; & \frac{1}{2} e^{-\frac{\hat{\gamma}_{dB}^2}{2\sigma_y^2}} \sum_{k=1}^{K} a_k \left[\operatorname{erfcx}\left(\frac{k\sigma_y}{\sqrt{2}\,\kappa} + \frac{\widehat{\gamma}_{dB}}{\sqrt{2}\,\sigma_y} \right) \right. \\ & \left. + \; \operatorname{erfcx}\left(\frac{k\sigma_y}{\sqrt{2}\,\kappa} - \frac{\widehat{\gamma} dB}{\sqrt{2}\,\sigma_y} \right) \right] \\ + \; & \frac{\sigma_y}{\sqrt{2\pi}\,\kappa} e^{-\frac{\hat{\gamma}_{dB}^2}{2\sigma_y^2}} + \frac{\widehat{\gamma}_{dB}}{2\kappa} \operatorname{erfc}\left(- \frac{\widehat{\gamma}_{dB}}{\sqrt{2}\,\sigma_y} \right) \end{aligned} \tag{5.8}$$

where $\widehat{\gamma}_{dB} = \gamma_{dB} + \mu_y$, $\gamma_{dB} = \kappa \ln(\gamma)$, $\operatorname{erfcx}(x) \doteq e^{x^2} \operatorname{erfc}(x)$, and the coefficients a_k are from Hasting's polynomial approximation of $\ln(1+x) = \sum_{k=1}^{i=K} a_k x^k$ for $0 \le x \le 1$. These coefficients are given in [AS72, (4.1.43-44)] for K=5 and K=8. In [LSA09] Laourine developed an additional formula that provides a generic solution for computing all the moments of the Shannon capacity for a LN channel. However, at a given complexity, this formula is not as precise as (5.8) in computing the first moment.

Héliot's work led to the following accurate closed form approximation of (5.6) for Shannon capacity [HCH09],

$$\begin{aligned} E[C] \;\cong\; & \lambda \left[\frac{\widehat{\gamma}_{dB}}{2\kappa} + \ln\left(2\cosh\left(\frac{\widehat{\gamma}_{dB}}{2\kappa} \right) \right) \right] \\ & + \lambda \phi\left(\sigma_y, \widehat{\gamma}_{dB} \right) \end{aligned} \tag{5.9}$$

The first term is the greatest lower bound (GLB) of the actual ergodic Shannon capacity and the second term, $\phi(\cdot)$, compensates for the difference between the actual ergodic capacity and the

GLB, and is given by,

$$\phi(\sigma_y, \widehat{\gamma}_{dB}) = \frac{\eta_0(\sigma_y)}{\cosh^{\eta_1(\sigma_y)}\left(\frac{\widehat{\gamma}_{dB}}{2\kappa\eta_2(\sigma_y)}\right)} \tag{5.10}$$

The parameters η_i were optimized via a piece-wise curve fitting method and each has the form,

$$\eta_i(\sigma_y) \quad = \quad \beta_i \sum_{k=1}^{N_i} b_{i,k}\sigma_y^k \tag{5.11}$$

The values for each β_i, N_i, and $b_{i,k}$ are given in [HCH09].

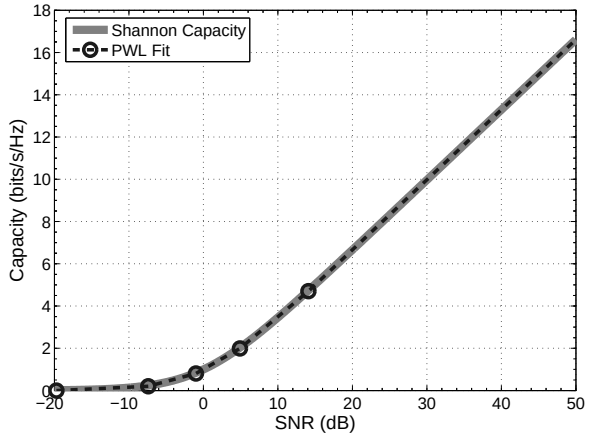
Each of these two methods for calculating the ergodic Shannon capacity employs approximations and the use of pre-calculated coefficients. In the following section a method of similar complexity is proposed that also uses an approximation and pre-calculated coefficients, but which is applicable not only to Shannon capacity but to constrained capacity as well.

5.4 Proposed Method

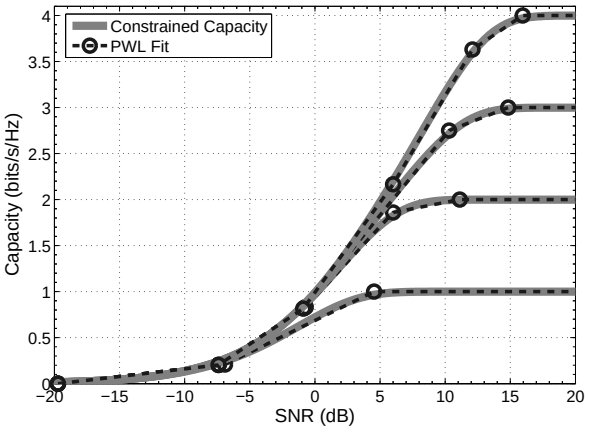
Reminiscent of Laourine's use of Hastings approximation and to Héliot's use of a piece-wise function based on curve fitting, this work will use an approximation to the capacity functions based on piece-wise linear (PWL) curve fitting. The Shannon capacity (5.3) and the constrained capacity (5.4) were each fit to a PWL curve as a function of $\gamma_{\text{rx,dB}}$, the received SNR in dB, using the easily implemented method of [Tom74] in which the maximum allowable error of the fit is specified as a parameter. The Shannon and constrained capacity functions can be approximated by the PWL fit

$$\begin{aligned} C(\alpha,\gamma) &\cong C_{pwl}(\alpha,\gamma) \\ &= m_k\gamma_{\text{rx,dB}} + b_k \text{ for } c_{k-1} < C \leq c_k \end{aligned} \tag{5.12}$$

where m_k and b_k are the slope and intercept values for section k of the PWL fit, which apply to capacity values between c_{k-1} and c_k. Figure 5.1 (a) shows the Shannon capacity and its PWL



(a) Shannon Capacity



(b) Constrained Capacities

Figure 5.1: Piece-wise linear fits for (a) Shannon capacity and (b) BPSK, QPSK, 8PSK, and 16QAM constrained capacities as a function of SNR in dB with a maximum error of 0.04 bits/s/Hz.

fit and Figure 5.1 (b) shows the constrained capacity curves and their PWL approximations. In all cases the fits were made so as to have a maximum absolute error of 0.04 bits/s/Hz. Table 5.1 provides the corresponding slope and intercept values for each section of the PWL fits portrayed in these figures. The maximum absolute error can be reduced further if desired at the expense of the addition of more sections to the PWL fit. Conversely, the PWL fit can be made with even fewer sections with a relaxed error limit.

The PWL fit was made using $\gamma_{rx,dB}$ as the independent variable in order to facilitate the

(a) Shannon

i	c_i (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.20214	0.016307	0.32253
2	0.81754	0.095012	0.90353
3	2.04144	0.20398	1.0021
4	4.7012	0.29684	0.52905
5	>4.90771	0.33210	0.00884059

(b) BPSK

i	c_i (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.20807	0.016205	0.31988
2	1.0	0.069184	0.68544
3	–	0	1

(c) QPSK

i	c_i (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.20348	0.016509	0.32565
2	0.83249	0.093881	0.89820
3	1.8600	0.15336	0.93984
4	2	0.027322	1.69610
5	–	0	2

(d) 8PSK

i	c_i (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.20350	0.016511	0.32568
2	0.82407	0.094025	0.89929
3	2.75000	0.17351	0.96288
4	3	0.055022	2.18330
5	–	0	3

(e) 16QAM

i	c_i (bits/s/Hz)	slope (m_i)	intercept (b_i)
1	0.20353	0.016513	0.32572
2	0.81540	0.094134	0.90012
3	2.16570	0.19570	0.99153
4	3.6309	0.24019	0.72459
5	4	0.09572	2.47270
6	–	0	4

Table 5.1: Coefficients for the PWL fit for: (a) the Shannon capacity and (b)-(e) constrained capacities for common modulations.

development of a convenient expression for the ergodic constrained capacity as will be seen shortly. Since α is log-normal, $\gamma_{\text{rx,dB}} = \kappa \ln(\alpha\gamma) = \kappa \ln(\alpha) + \gamma_{dB}$, is Gaussian distributed with PDF,

$$f_{\gamma_{rx,dB}}(x) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-(x-\widehat{\gamma}_{dB})^2/2\sigma_y^2} \tag{5.13}$$

The associated CDF is,

$$F_{\gamma_{rx,dB}}(x) = \frac{1}{2}\,\text{erfc}\Big(-(x-\widehat{\gamma}_{dB})\Big/\sqrt{2\sigma_y^2}\Big) \tag{5.14}$$

The outage probability associated with a given capacity value c_o can be evaluated by using

the PWL fit of the capacity to estimate the SNR value γ_0 needed to achieve c_o, and then using (5.14) to determine the probability that actual SNR will be less than γ_0. The slope and intercept values can be found in Table 5.1 according to the modulation type. From these values γ_o can be estimated,

$$\gamma_o \quad \cong \quad (c_o - b_k)/m_k \tag{5.15}$$

where the value of k is such that $c_{k-1} < c_o \leq c_k$. Using (5.7) and (5.15) the outage probability can be estimated from,

$$P_{out}(c_o) = F_{\gamma_{rx,dB}}(\gamma_o) \cong F_{\gamma_{rx,dB}}\big((c_o - b_k)/m_k\big) \tag{5.16}$$

Since C is a monotonic function of $\gamma_{rx,dB}$ the PDF of C over a LN channel can be obtained from,

$$f_C(c) \quad = \quad \left.\frac{f_{\gamma_{rx,dB}}(\gamma_{rx,dB})}{\left|\frac{dC}{d\gamma_{rx,dB}}\right|}\right|_{\gamma_{rx,dB}=C^{-1}(c)} \tag{5.17}$$

For the constrained capacity case the derivative could possibly be evaluated using an altered form of (5.4), but this does not simplify the evaluation of (5.5) since the resulting expression for the derivative will still contain an integral over the complex plane and a double sum. Also, it is not apparent how the inverse $\gamma_{rx} = C^{-1}(c)$ can be obtained from (5.4). Using the piecewise linear fit (5.12), these two quantities can be approximated. For $c_{k-1} < c \leq c_k$,

$$\left|\frac{dC}{d\gamma_{rx,dB}}\right| \quad \cong \quad m_k \tag{5.18}$$

$$\gamma_{rx,dB} = C^{-1}(c) \cong \frac{c - b_k}{m_k} \tag{5.19}$$

Using (5.18) and (5.19) the PDF of (5.17) can be approximated for $c_{k-1} < c \leq c_k$ as,

$$\begin{aligned} f_C(c) &\cong f_{C,k}(c) \\ &\doteq \frac{1}{\sqrt{2\pi\sigma_k^2}}\, e^{-\frac{(c-\mu_k)^2}{2\sigma_k^2}} \end{aligned} \tag{5.20}$$

where,

$$\begin{aligned} \mu_k &= b_k + m_k \widehat{\gamma}_{dB} \\ \sigma_k &= m_k \sigma_y \end{aligned}$$

Using (5.20) in (5.5) results in a readily evaluable expression. The PWL fit was made to $\gamma_{rx,dB}$ in order to take advantage of the Gaussian distribution of the received SNR as measured in dB and thus to create an approximation for $f_C(\cdot)$ that facilitates the evaluation of (5.5). From (5.5) and (5.20) the ergodic capacity of a LN channel can be approximated as,

$$\begin{aligned} E[C] &\cong \sum_{k=1}^{K-1} \left[\int_{c_{k-1}}^{c_k} c\, f_{C,k}(c) \mathrm{d}c \right] \\ &\quad + \big(1 - F_{\gamma_{rx,dB}}(\tilde{\gamma}_{max})\big) c_{max} \\ &= \sum_{k=1}^{K-1} \left[\frac{\sigma_k}{\sqrt{2\pi}} \left(e^{-\frac{(c_{k-1}-\mu_k)^2}{2\sigma_k^2}} - e^{-\frac{(c_k-\mu_k)^2}{2\sigma_k^2}} \right) \right. \\ &\quad \left. + \frac{\mu_k}{2} \left(\mathrm{erfc}\Big(\frac{c_{k-1}-\mu_k}{\sqrt{2}\,\sigma_k} \Big) - \mathrm{erfc}\Big(\frac{c_k-\mu_k}{\sqrt{2}\,\sigma_k} \Big) \right) \right] \\ &\quad + \frac{1}{2} \mathrm{erfc}\left(\frac{c_{max}-\mu_{K-1}}{\sqrt{2}\,\sigma_{K-1}} \right) c_{max} \end{aligned} \tag{5.21}$$

where $c_0 = 0$, c_{max} is the maximum achievable capacity, K is the number of segments in the PWL fit, and $\tilde{\gamma}_{max,dB}$ is an approximation to $\gamma_{max,dB}$ the SNR needed to achieve c_{max}, i.e. $\tilde{\gamma}_{max,dB} = (c_{max} - b_{K-1})/m_{K-1}$. The $(1 - F_{\gamma_{rx,dB}}(\tilde{\gamma}_{max}))c_{max}$ term accounts for a possible probability mass in the capacity PDF due to the received SNR exceeding that required to achieve

c_{max}. With an M-ary modulation $c_{max} = \log_2(M)$. For the Shannon case there is no probability mass at any value of c and so the probability mass term is ignored and the sum is taken over all K terms of the PWL fit instead.

Equations (5.16) and (5.21) provide a uniform computation of both Shannon and constrained capacity outage probabilities and ergodic capacities. These equations can be evaluated using the coefficients of Table 5.1 (a) for the Shannon case and the coefficients of Table 5.1 (b)-(e) for BPSK, QPSK, 8PSK and 16QAM constellations.

The outage probability and ergodic capacity estimation method described above can be expanded to the situation in which N antennas are combined via maximal ratio combing (MRC) by using the log-normal approximation, which states that the sum of LN random variables can be well approximated by another log-normal random variable (see [HCH09] and [LSA09] and the references therein). The SNR values of each branch of the MRC combiner have a log-normal distribution. The SNR of the output of the MRC combiner is the sum of the SNR's of the individual branches, and thus has a distribution that is well represented by a log-normal distribution.

The parameters for this log-normal distribution can be estimated by the Fenton-Wilkinson (F-W) moment-matching method for low to moderate values of σ_y (4 dB or less [AG99]) as described in [LSA09]. For large values of σ_y the method in [MWM07] can be used. For the combination of N branches, where $\ln(\alpha_i)$ of branch i has standard deviation σ_i, and a correlation coefficient with $\ln(\alpha_j)$ of branch j given by $\rho_{i,j}$, the F-W estimates of μ_{mrc} and σ_{mrc} of the effective log-normal channel are

$$\begin{aligned} \mu_{mrc} &= -\sigma_{MRC}^2/2 \\ \sigma_{MRC}^2 &= \ln\left(\frac{1}{N^2}\sum_{i,j=1}^{N}\exp\left(\rho_{i,j}\sigma_i\sigma_j\right)\right). \end{aligned} \tag{5.22}$$

These values for μ_{mrc} and σ_{MRC} can then be used in conjunction with (5.16) and (5.21) to find the outage probabilities and ergodic capacities of a LN channel with the MRC combining of N antennas.

The proposed outage probability and ergodic capacity estimation method also can be expanded to the interference dominated situation as well. As explained in [AG99] the signal-to-interference ratio (SIR) of an interference dominated LN channel has a LN distribution. The parameters for the LN distribution of the SIR can be found using the F-W method and (5.16) and (5.21) applied to find the outage probabilities and ergodic capacities of the interference dominated LN channel. The ergodic capacities of the interference dominated channel can in turn be used to calculate the area spectral efficiency of a cellular mobile radio system.

5.5 Results

As seen in Fig. 5.2 the proposed expression closely matches Laourine's method, Héliot's method and Monte Carlo results. The match between the proposed method's results and those of the previously published methods and of the Monte Carlo simulations is within 0.03 bits/s/Hz at any SNR, and within 0.4 dB at any capacity value > 0.2 bits/s/Hz across all values of σ_y.

With BPSK, QPSK, 8PSK and 16QAM modulations the match between results from the proposed method and Monte Carlo simulations was within 0.03 bits/s/Hz at any SNR, and within 0.4 dB at any capacity value between 0.2 and $0.99c_{max}$ bits/s/Hz with values of σ_y of 2, 6, 9 and 12 dB. Fig. 5.3 shows the results for QPSK and 16QAM.

Figure 5.4 shows the outage probability values, P_{out}, obtained from the proposed expression (5.16) and from Monte Carlo simulations for BPSK, QPSK, 8PSK, 16QAM and Gaussian (i.e. the Shannon case) input alphabets. At a given outage probability value, the estimate from the proposed expression agrees within the error of the PWL fit, 0.04 bits/s/Hz in this example, of the results from Monte Carlo simulations and numerical methods.

Using the approach described in the previous section for the antenna combining case, the proposed method was applied to evaluate the ergodic Shannon capacity and ergodic constrained capacity for a 16QAM transmission in a log-normal channel with the MRC combining of N= 1, 2 and 8 antennas. The correlation model used was exponential, i.e. $\rho_{ij} = \rho^{|i-j|}$ with $0 \leq \rho \leq 1$. Figure 5.5 shows that the proposed method agrees very well with Monte Carlo simulation

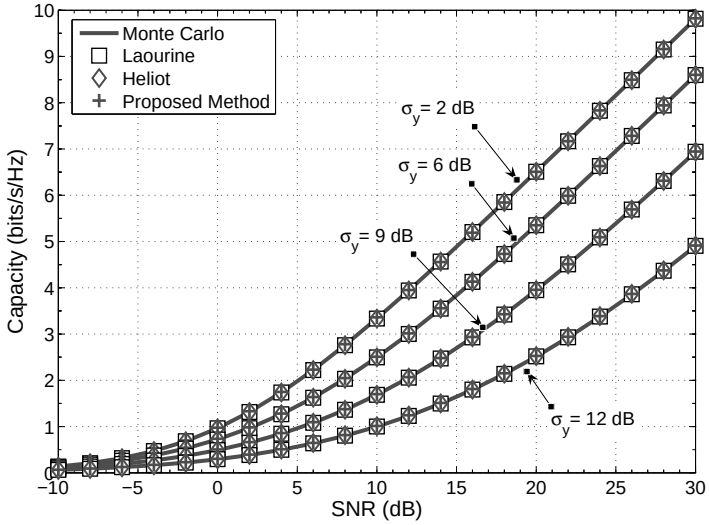


Figure 5.2: Ergodic Shannon capacity obtained by 1) Laourine's method (K=5), 2) Héliot's method 3) the proposed method (K=5) and 4) Monte Carlo simulations in log-normal channels of various σ_y values.

results. In generating the data for this figure a value of σ_y= 3 dB was used for the N=1 case, whereas for N= 2 and 8, half the channels used $\sigma_y = 3$ dB and the other half used $\sigma_y = 4$ dB. The proposed expression was within 0.4 dB of the Monte Carlo results at any capacity value between 0.2 and $0.99c_{max}$ bits/s/Hz for N= 1, 2, and 8 for $0 \leq \rho \leq 1$. When the values for σ_y were increased from 3 and 4 dB to 5 and 6 dB, the error margin for N= 2 and 8 increased to 0.8 dB, due to the inaccuracy of the Fenton-Wilkinson method for large values of σ_y. The diversity gain can be obtained by comparing the ρ= 0 and ρ= 1 cases. For the situation considered, it was found to be 0.5 dB (0.75 dB) for N= 2 antennas and 1.1 dB (1.5 dB) for N= 8 antennas for ergodic Shannon (16QAM constrained) capacity at 3 b/s/Hz.

Figure 5.6 shows the outage constrained capacity, $C_{out,P}(\gamma)$, for 16QAM, QPSK and Shannon associated with $P = 0.01\%$ and $P = 10\%$ outage rates with MRC antenna combining for $N = 1$ and 2 for a LN channel with σ_y= 5 dB. The SNR was normalized by the number of antennas in order to remove the aperture gain so that the effect of the diversity gain could be clearly seen. These values were determined by using (5.16) to generate outage probability curves for several average SNR values and then determining from each of those curves the capacity value associated with the desired outage percentage.

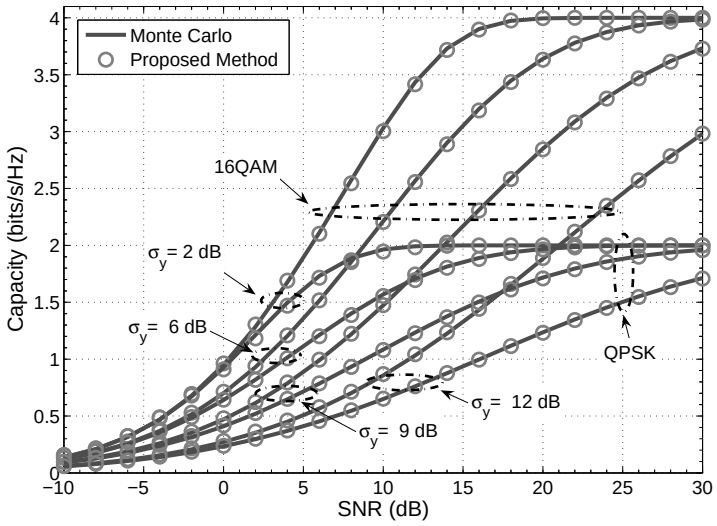


Figure 5.3: Ergodic constrained capacity for QPSK and 16QAM in log-normal channels of various σ_y values.

5.6 Conclusion

This chapter presented an innovative method to quickly and accurately estimate the ergodic capacity and outage probability of a log-normal channel for practical communication systems constrained to a finite alphabet, without resorting to lengthy Monte Carlo simulations or numerical integration. This new method is similar to two previously published methods that dealt solely with Shannon capacity in that it relies on pre-computed coefficients and the summation of a small number of terms. The new method has the advantage over the previous methods of also being applicable to the calculation of outage probabilities and ergodic constrained capacities for practical communication systems. In this chapter pre-computed coefficients have been provided for Shannon capacity and for the constrained capacities of BPSK, QPSK, 8PSK and 16QAM. This new method was shown to match well with Monte Carlo simulation results and with the two formerly published Shannon capacity methods.

The proposed method is applicable to any QAM modulation once the parameters of the PWL fit of (5.4) are obtained via some method such as that of [Tom74]. For BPSK, QPSK, 8PSK and 16QAM modulations the coefficients of Table 5.1 can be used.

This method also facilitates the evaluation of ergodic constrained capacities for other chan-

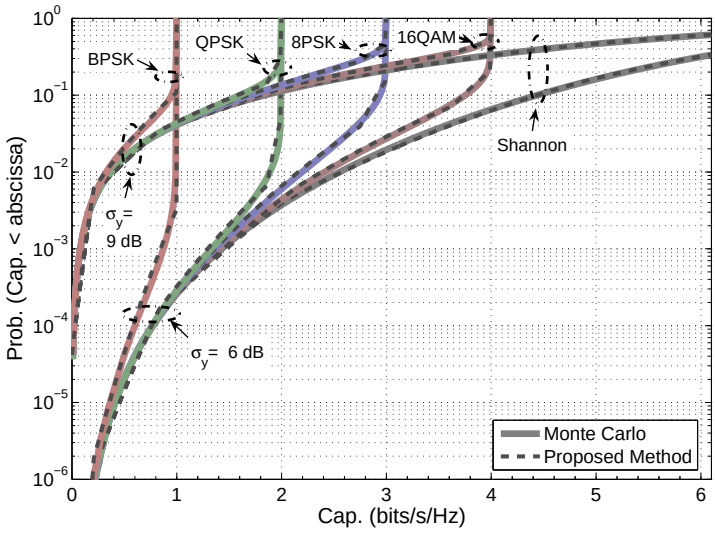


Figure 5.4: Outage probabilities for BPSK, QPSK, 8PSK and 16QAM constrained capacities and also Shannon capacity in log-normal channel obtained from the proposed expression and from Monte Carlo simulations (100 million trials). The average SNR, γ_{dB}, was 25 dB in all cases.

nel types, e.g. Nakagami channel. In addition, the PWL fit (5.12) of the constrained capacity and associated coefficients of Table 5.1 could be used in forming the objective function for a cellular system's power control algorithm. The usage of the highly accurate approximation to the constrained capacity (5.12) in place of a somewhat loose bound ($\pm$ 1 dB) on the bit error rate used in [QC99] may yield a more optimized power distribution and a higher system throughput. The application of the proposed method to other channel types and the use of (5.12) in a power control algorithm are two areas of possible future work.

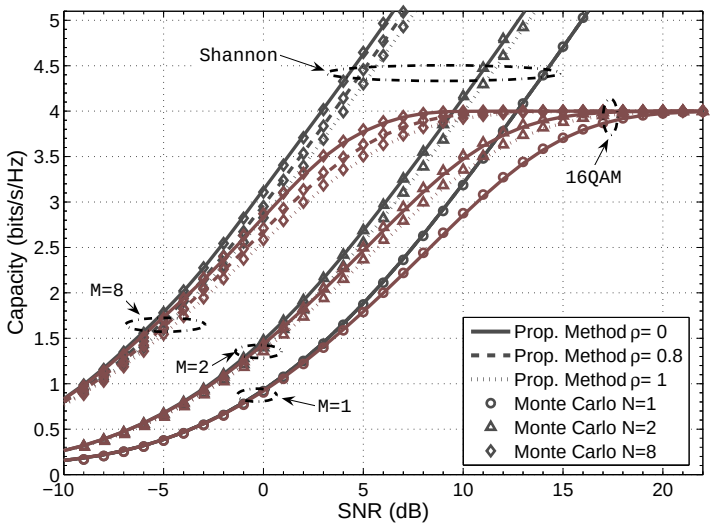


Figure 5.5: Ergodic Shannon capacity and ergodic constrained capacity for 16QAM in a log-normal channel with the MRC combining of N= 1, 2 and 8 antennas for correlation coefficients of ρ= 0, 0.8 and 1. For N=1, σ_y= 3 dB, whereas for N= 2 and 8, half the channels had $\sigma_y = 3$ dB and the other half had $\sigma_y = 4$ dB.

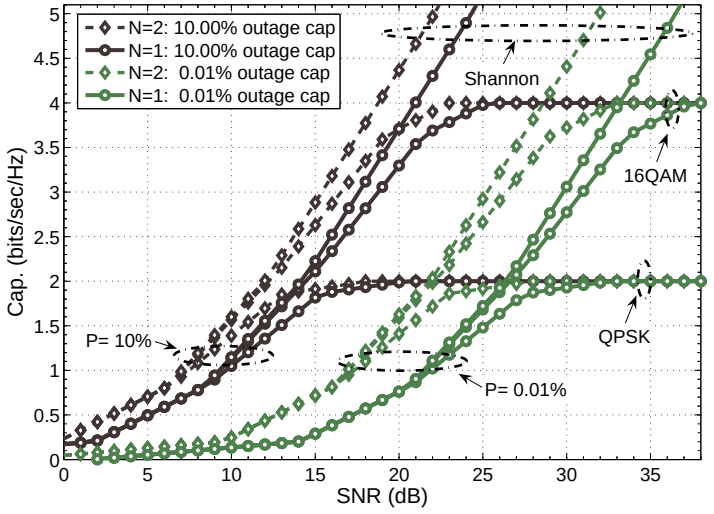


Figure 5.6: Outage capacities for QPSK and 16QAM constellations and Shannon (Gaussian-distributed input) in LN channel with $\sigma_y = 5$ dB for $N = 1$ and 2 antennas.

APPENDIX A

Multiple Phase Screen Model

A.1 Generation of Phase Screens

In brief, each phase screen is generated randomly according to the following process:

1. The phase at each (n_x, n_y) of the *NxN* grid is generated independently according to a Gaussian distribution of zero mean and unit variance. This generation is assumed to take place in the frequency domain, resulting in a *NxN* grid of values $\Phi_1(f_x, f_y)$.

2. The resulting values are then spectrally shaped according to the desired spectrum of index of refraction irregularities $\Psi(f_x, f_y)$ in order to introduce the same phase correlation that is observed in the troposphere. This is done in the frequency domain by multiplying each $\Phi_1(f_x, f_y)$ value by $\sqrt{\Psi(f_x, f_y)}$ producing $\Phi_2(f_x, f_y)$.

3. The spectrally shaped values $\Phi_2(f_x, f_y)$ are then scaled appropriately so that the phase difference between any two grid points $(n_{x,1}, n_{y,1})$ and $(n_{x,2}, n_{y,2})$ has the same statistical average as that of the phase difference between the two corresponding points of a plane E-M wave that has propagated a distance Δ_z in the z direction through a random medium with turbulence strength $C_n^2(z)$. This produces $\Phi_3(f_x, f_y)$.

4. The scaled and spectrally shaped grid of values $\Phi_3(f_x, f_y)$ is inverse Fourier transformed and the real (or imaginary) components taken in order to generate the spatial-domain values for the random phase screen $\phi_{ps}(n_x, n_y)$.

The resulting phase screen has the same spatial correlation and statistical properties as the phase differences between the corresponding points of a E-M plane wave that has propagated a dis-

tance Δ_z through an isotropic, homogeneous medium of turbulence strength $C_n^2(z)$. The Gaussian assumption is based on the central limit theorem. The delay, $t(|z_1 - z_0|, x_0, y_0)$,of an E-M wave along a path in the z-direction from (z_0, x_0, y_0) to (z_1, x_0, y_0), scaled by that through a vacuum, is given by

$$t(|z_1 - z_0|, x_0, y_0) = \int_{z_1}^{z_0} n(z, x_0, y_0) \mathrm{d}z$$

The variable of interest is the difference in phase across the wavefront, which is what the values of the phase screen represent. This is directly proportional to the difference in the delay, $\triangle t(|z_1 - z_0|, |x_1 - x_0|, |y_1 - y_0|)$, along two paths in the z-direction (z, x_0, y_0) and (z, x_1, y_1) which is given by

$$\begin{aligned} \triangle t(|z_1 - z_0|, |x_1 - x_0|, |y_1 - y_0| &= \int_{z_1}^{z_0} n(z, x_1, y_1) \mathrm{d}z - \int_{z_1}^{z_0} n(z, x_0, y_0) \mathrm{d}z \\ &= \int_{z_1}^{z_0} n(z, x_1, y_1) - n(z, x_0, y_0) \mathrm{d}z \end{aligned}$$

The difference in delay is an infinite sum (integral) of the random index of refraction, which is independent for sufficient separation distances. Consequently the central limit theorem applies, and the difference in delay, and thus phase, can be well represented by a Gaussian random variable. This assumption is supported by the material in [Whe03, ch. 10.1].

A.2 Propagation Between Screens

It can be shown [Goo05, sec. 3.10] that if $U(x, y, z_0)$ is the field at (x, y, z_0) and $A(f_x, f_y; z_0)$ is its Fourier transform with respect to (x, y) in the plane $z = z_0$, then $A(f_x, f_y; z_0)$ represents the spectrum of plane waves propagating at wave vector $\vec{k}$, where,

$$\vec{k} = \frac{2\pi}{\lambda}(\alpha \hat{a}_x + \beta \hat{a}_y + \gamma \hat{a}_z)$$

where,

$$\begin{aligned}\alpha &= \lambda f_x \\ \beta &= \lambda f_y \\ \gamma &= \sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2}\end{aligned}$$

It can also be shown that at some further point z, in the direction of propagation, the angular spectrum can be expressed in terms of the angular spectrum at z_0.

$$\begin{aligned}A(f_x,f_y;z_1) &= A\Big(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z_1\Big) \\ &= A\Big(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z_0\Big)\exp\big(j\mu\big|z_1-z_0\big|\big) \\ &= \begin{cases} A\Big(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z_0\Big)\exp\Big(j\frac{2\pi}{\lambda}\sqrt{1-\alpha^2-\beta^2}\,\Big|z_1-z_0\Big|\Big) & \text{for} \quad \alpha^2+\beta^2<1 \\ 0 & \text{for} \quad \alpha^2+\beta^2\geq 1 \end{cases} \\ &= A\Big(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z_0\Big)\exp\Big(j\frac{2\pi}{\lambda}\sqrt{1-\alpha^2-\beta^2}\,\big|z_1-z_0\big|\Big)\mathrm{circ}\Big(\sqrt{\alpha^2+\beta^2}\Big)\end{aligned}$$

Since,

$$U(x,y;z_1)=\mathcal{F}^{-1}\Big(A\Big(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z_1\Big)\Big) \tag{A.1}$$

We have,

$$\begin{aligned}U(x,y,;z_1) &= \mathcal{F}^{-1}\Big(A\Big(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z_1\Big)\Big) \\ &= \int\int_{-\infty}^{+\infty} A\Big(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z_1\Big)\exp\Big(j2\pi\Big(\frac{\alpha}{\lambda}x+\frac{\alpha}{\lambda}y\Big)\Big)\mathrm{d}\frac{\alpha}{\lambda}\mathrm{d}\frac{\beta}{\lambda} \\ &= \int\int_{-\infty}^{+\infty} A\Big(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z_0\Big)\exp\Big(j\frac{2\pi}{\lambda}\sqrt{1-\alpha^2-\beta^2}\Big|z_1-z_0\Big|\Big)\cdot \\ &\quad \mathrm{circ}\Big(\sqrt{\alpha^2+\beta^2}\Big)\exp(j2\pi(\frac{\alpha}{\lambda}x+\frac{\alpha}{\lambda}y))\mathrm{d}\frac{\alpha}{\lambda}\mathrm{d}\frac{\beta}{\lambda} \\ &= \mathcal{F}^{-1}\Big(A\Big(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z_0\Big)\exp\Big(j\frac{2\pi}{\lambda}\sqrt{1-\alpha^2-\beta^2}\Big|z_1-z_0\Big|\Big)\mathrm{circ}\Big(\sqrt{\alpha^2+\beta^2}\Big)\Big)\end{aligned}$$

Since multiplication in the spatial frequency domain is like convolution in the spatial domain we see that the wave propagation phenomena can be viewed as a linear filter where the spatial frequency domain expression of the filter is

$$H(f_X, f_Y; z) = \begin{cases} \exp\left[j2\pi \frac{z}{\lambda} \sqrt{1 - (\lambda f_X)^2 - (\lambda f_Y)^2} \right] & \text{for} \quad \sqrt{(\lambda f_X)^2 + (\lambda f_Y)^2} < 1 \\ 0 & \text{for} \quad \sqrt{(\lambda f_X)^2 + (\lambda f_Y)^2} \geq 1 \end{cases} \tag{A.2}$$

This is the formula that the MPS model used to perform the propagation of the wave a distance z between phase screens.

APPENDIX B

Derivation of the Random Wave Equation and Description of the *Rytov Approximation* Solution

B.1 Derivation of the Random Wave Equation

B.1.1 Maxwell's Equations

We will derive the random wave equation, closely following Wheelon's clearly-presented development in [Whe01] throughout, beginning with Maxwell's equations,

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \tag{B.1}$$

$$\nabla \times \mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t} + \frac{4\pi}{c}\mathbf{J} \tag{B.2}$$

$$\nabla \cdot \mathbf{D} = 4\pi\rho_e \tag{B.3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{B.4}$$

Where, **E** is the electric field, **D** is the displacement field, **H** is the magnetic field, **B** is the induction field, and ρ_e is the net charge density.

The displacement field is related to the electric field by,

$$\mathbf{D} = \varepsilon(\mathbf{r},t)\mathbf{E} \tag{B.5}$$

where $\varepsilon(\mathbf{r},t)$ is the permittivity of the medium at position $\mathbf{r}$ at time t.

The induction field is related to the magnetic field by,

$$\mathbf{B} = \mu(\mathbf{r},t)\mathbf{H} \tag{B.6}$$

where $\mu_m(\mathbf{r},t)$ is the permeability of the medium at position $\mathbf{r}$ at time t. In the units we will be using μ_m will have a value of 1.

B.1.2 Random Wave Equation

In this section we will provide a derivation for the scalar random wave equation.

First though, let us take a look at the balance of current and charge on the transmitting source. Taking the divergence of (B.2) gives

$$\nabla \cdot (\nabla \times \mathbf{H}) = \frac{1}{c}\frac{\partial \nabla \cdot \mathbf{D}}{\partial t} + \frac{4\pi}{c}\nabla \cdot \mathbf{J} \tag{B.7}$$

Using the identity, that for any vector $\mathbf{a}$,

$$\nabla \cdot (\nabla \times \mathbf{a}) = 0$$

and (B.3), the above becomes

$$0 = \frac{1}{c}\frac{\partial \rho_e}{\partial t} + \frac{4\pi}{c}\nabla \cdot \mathbf{J}$$

Thus we have,

$$\frac{1}{c}\frac{\partial \rho_e}{\partial t} + \frac{4\pi}{c}\nabla \cdot \mathbf{J} = 0 \tag{B.8}$$

This equation relates to the balance of current and charge on the transmitting antenna or laser source, elsewhere in the transmission region one can ignore these quantities, (i.e.$\mathbf{J} = 0$ and $\rho_e = 0$). In this situation (B.3) becomes

$$\nabla \cdot \mathbf{D} = 0 \tag{B.9}$$

We also know that

$$\mathbf{B} = \mu_m(\mathbf{r},t)\mathbf{H} = \mathbf{H} \tag{B.10}$$

since in the units we are using $\mu_m = 1$, in the earth's atmosphere.

We decompose $\varepsilon(\mathbf{r},t)$ into its average value and a small component that is a stochastic function of position and time,

$$\varepsilon(\mathbf{r},t) = \varepsilon_0(\mathbf{r}) + \triangle\varepsilon(\mathbf{r},t) \tag{B.11}$$

In the lower atmosphere (e.g. troposphere) $\varepsilon_0(\mathbf{r})$=1. The fluctuating component $\triangle\varepsilon(\mathbf{r},t)$ gives rise to electromagnetic scintillation. This is a key point.

Let us move on to the specifics of the random wave equation. To get the random wave equation we combine Maxwell's equation. Applying the curl operator to (B.1) and using (B.10) gives

$$\begin{aligned}
\nabla \times (\nabla \times \mathbf{E}) &= -\frac{1}{c}\nabla \times \frac{\partial \mathbf{H}}{\partial t} && \text{(B.12)}\\
&= -\frac{1}{c}\frac{\partial \nabla \times \mathbf{H}}{\partial t} \\
&= -\frac{1}{c}\frac{\partial}{\partial t}(\frac{1}{c}\frac{\partial \mathbf{D}}{\partial t} + 4\pi\mathbf{J}) \\
&= -\frac{1}{c^2}\frac{\partial^2(D)}{\partial t^2} - \frac{4\pi}{c}\frac{\partial \mathbf{J}}{\partial t} \\
&= -\frac{1}{c^2}\frac{\partial^2 \varepsilon\mathbf{E}}{\partial t^2} - \frac{4\pi}{c}\frac{\partial \mathbf{J}}{\partial t} && \text{(B.13)}
\end{aligned}$$

In the above we went from line 2 to line 3 by using (B.2).

The double curl operation can be simplified since the following holds for any vector,

$$\nabla \times \nabla \times \mathbf{E} = -\nabla^2 \mathbf{E} + \nabla(\nabla \cdot \mathbf{E}) \tag{B.14}$$

So (B.13) becomes,

$$-\nabla^2 \mathbf{E} + \nabla(\nabla \cdot \mathbf{E}) = -\frac{1}{c^2}\frac{\partial^2 \varepsilon \mathbf{E}}{\partial t^2} - \frac{4\pi}{c}\frac{\partial \mathbf{J}}{\partial t} \tag{B.15}$$

Now from (B.3) with ρ_e=0,

$$\begin{aligned} 0 &= \nabla \cdot \mathbf{D} \\ &= \nabla \cdot (\varepsilon \mathbf{E}) \\ &= \mathbf{E} \cdot \nabla \varepsilon + \varepsilon \nabla \cdot \mathbf{E} \end{aligned}$$

So,

$$\begin{aligned} \nabla \cdot \mathbf{E} &= -\frac{1}{\varepsilon}\mathbf{E} \cdot \nabla \varepsilon \\ &= -\mathbf{E} \cdot \frac{\nabla \varepsilon}{\varepsilon} \\ &= -\mathbf{E} \cdot \nabla(\ln \varepsilon) \end{aligned}$$

Where the last step comes because $\nabla \ln \varepsilon = \frac{1}{\varepsilon}\nabla \varepsilon$ by the chain rule.

So (B.15) becomes,

$$-\nabla^2 \mathbf{E} + \nabla \cdot (-\mathbf{E} \nabla(\ln \varepsilon)) = -\frac{1}{c^2}\frac{\partial^2 \varepsilon \mathbf{E}}{\partial t^2} - \frac{4\pi}{c}\frac{\partial \mathbf{J}}{\partial t} \tag{B.16}$$

Writing ε as $\varepsilon = 1 + \triangle\varepsilon$ we get

$$\nabla^2\mathbf{E} - \frac{1}{c^2}\frac{\partial^2(1+\triangle\varepsilon)\mathbf{E}}{\partial t^2} = \frac{4\pi}{c}\frac{\partial\mathbf{J}}{\partial t} - \nabla(\mathbf{E}\cdot\nabla(\ln(1+\triangle\varepsilon))) \tag{B.17}$$

The final term descries polarization changes induced by scattering in the random medium, this is supposedly negligible for atmospheric propagation (Assumption number 1, which is validated in [Whe03]], chapter 11), so that term is dropped leaving,

$$\nabla^2\mathbf{E} - \frac{1}{c^2}\frac{\partial^2(1+\triangle\varepsilon)\mathbf{E}}{\partial t^2} = \frac{4\pi}{c}\frac{\partial\mathbf{J}}{\partial t} \tag{B.18}$$

The dielectric fluctuations, $\triangle\varepsilon$ change slowly compared to the frequencies of interest for $\mathbf{E}$, i.e. $\triangle\varepsilon$ fluctuations are in the Hz range (1-10 Hz). As a result frequency mixing that occurs in the term $\triangle\varepsilon\mathbf{E}$ are not important (Assumption number 2).This means we can consider a single frequency at a time in (B.18), i.e. we can study $\mathbf{J}(\mathbf{r},t) = \mathbf{J}(\mathbf{r})e^{(-i\omega t)}$ and $\mathbf{E}(\mathbf{r},t) = \mathbf{E}(\mathbf{r})e^{(-i\omega t)}$ and get the solution for a general $\mathbf{J}(\mathbf{r},t)$ and $\mathbf{E}(\mathbf{r},t)$ from the superposition of the single frequency solutions.

With $\mathbf{J}(\mathbf{r},t) = \mathbf{J}(\mathbf{r})e^{(-i\omega t)}$ and $\mathbf{E}(\mathbf{r},t) = \mathbf{E}(\mathbf{r})e^{(-i\omega t)}$ (B.18) becomes,

$$\begin{aligned}
\nabla^2[\mathbf{E}(\mathbf{r})e^{(i\omega t)}] - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}[(1+\triangle\varepsilon)\mathbf{E}(\mathbf{r})e^{(-i\omega t)}] &= \frac{4\pi}{c}\frac{\partial\mathbf{J}\mathbf{r}e^{(-i\omega t)}}{\partial t} \\
e^{(i\omega t)}\nabla^2\mathbf{E}(\mathbf{r}) - \frac{1}{c^2}\mathbf{E}(\mathbf{r})\frac{\partial^2}{\partial t^2}[(1+\triangle\varepsilon)e^{(-i\omega t)}] &= \frac{-4\pi(i\omega)}{c}\mathbf{J}(\mathbf{r})e^{(-i\omega t)} \\
\nabla^2\mathbf{E}(\mathbf{r}) - \mathbf{E}(\mathbf{r})\{\frac{1}{c^2}e^{(+i\omega t)}\frac{\partial^2}{\partial t^2}[(1+\triangle\varepsilon)e^{(-i\omega t)}]\} &= \frac{-4\pi(i\omega)}{c}\mathbf{J}(\mathbf{r})
\end{aligned} \tag{B.19}$$

The effects of the random medium are concentrated in the second term of the right hand side. Let us look at the second term, and eventually make an assumption that will simplify it.

$$\frac{\partial}{\partial t}[(1+\triangle\varepsilon)e^{(-i\omega t)}] = \frac{\partial\triangle\varepsilon}{\partial t}e^{(-i\omega t)} + (1+\triangle\varepsilon)(-i\omega)e^{(-i\omega t)}$$

and thus,

$$\begin{aligned}\frac{\partial^2}{\partial t^2}[(1+\triangle\varepsilon)e^{(-i\omega t)}] &= \frac{\partial^2\triangle\varepsilon}{\partial t^2}e^{(-i\omega t)}+\frac{\partial\triangle\varepsilon}{\partial t}(-i\omega e^{(-i\omega t)})+ \\ &\quad (-i\omega)[\frac{\partial\triangle\varepsilon}{\partial t}e^{(-i\omega t)}+(1+\triangle\varepsilon)(-i\omega)e^{(-i\omega t)}] \\ &= -\omega^2(1+\triangle\varepsilon)e^{(-i\omega t)}-2i\omega\frac{\partial\triangle\varepsilon}{\partial t}e^{(-i\omega t)}+\frac{\partial^2\triangle\varepsilon}{\partial t^2}e^{(-i\omega t)}\end{aligned}$$

Using this equation the second term in (B.19) becomes,

$$\frac{1}{c^2}e^{(+i\omega t)}\frac{\partial^2}{\partial t^2}[(1+\triangle\varepsilon)e^{(-i\omega t)}] = -\frac{\omega^2}{c^2}(1+\triangle\varepsilon)-2i\frac{\omega}{c^2}\frac{\partial\triangle\varepsilon}{\partial t}+\frac{1}{c^2}\frac{\partial^2\triangle\varepsilon}{\partial t^2}$$

Since, $\frac{\omega}{c}=\frac{2\pi f}{c}=\frac{2\pi}{\lambda}\equiv k$, we have

$$\frac{1}{c^2}e^{(+i\omega t)}\frac{\partial^2}{\partial t^2}[(1+\triangle\varepsilon)e^{(-i\omega t)}] = -k^2(1+\triangle\varepsilon)-2i\frac{k}{c}\frac{\partial\triangle\varepsilon}{\partial t}+\frac{1}{c^2}\frac{\partial^2\triangle\varepsilon}{\partial t^2} \qquad \text{(B.20)}$$

Now we are at the point where the physics of the situation allow us to simplify this equation. The first derivative of with respect to time, $\frac{\partial\triangle\varepsilon}{\partial t}$, can be estimated from the wind speed v. The wind blows the eddies, with their differing values of $\triangle\varepsilon$ through space, thus causing a change in $\triangle\varepsilon$ with respect to time at a fixed spatial location. The first derivative of $\triangle\varepsilon$ w.r.t. time can be estimated as.

$$\frac{\partial\triangle\varepsilon(\mathbf{r},t)}{\partial t}\cong\frac{v}{l}\triangle\varepsilon$$

where l is the size of an eddy (1e-3 m to 100 m). The quantity v/l is just $1/\triangle t$ where $\triangle t$ is the time it takes for the eddy to pass the point, $\mathbf{r}$, under consideration.

So the second term on the r.h.s. of (B.20) compared to the portion of the first term involving $\triangle\varepsilon$ is,

$$
\begin{aligned}
-\frac{2ik}{c}\frac{\partial}{\partial t}\triangle\varepsilon(\mathbf{r},t) \quad & vs \quad -k^2\triangle\varepsilon(\mathbf{r},t) \\
\frac{2k}{c}\frac{v}{l}\triangle\varepsilon(\mathbf{r},t) \quad & vs \quad k^2\triangle\varepsilon(\mathbf{r},t) \\
\frac{2v}{lc} \quad & vs \quad k
\end{aligned}
$$

Since $\frac{v}{c}$ is very small and k is large we know that the second term of (B.20) is much, much less than the portion of the first term involving $\triangle\varepsilon$, and can be considered negligible in comparison. The third term, involving the second time derivative of $\triangle\varepsilon(\mathbf{r},t)$ will be even smaller, and it is multiplied by $1/c^2$, making it even more negligible.

Given the above qualitative argument that the second and third terms of (B.20) will be negligible relative to $k^2\triangle\varepsilon(\mathbf{r},t)$, we have

$$
\begin{aligned}
\frac{1}{c^2}e^{(+i\omega t)}\frac{\partial^2}{\partial t^2}[(1+\triangle\varepsilon)e^{(-i\omega t)}] &= -k^2(1+\triangle\varepsilon) - 2i\frac{k}{c}\frac{\partial\triangle\varepsilon}{\partial t} + \frac{1}{c^2}\frac{\partial^2\triangle\varepsilon}{\partial t^2} \\
\frac{1}{c^2}e^{(+i\omega t)}\frac{\partial^2}{\partial t^2}[(1+\triangle\varepsilon)e^{(-i\omega t)}] &\cong -k^2(1+\triangle\varepsilon) \qquad \text{(B.21)}
\end{aligned}
$$

where the approximation is quite good.

Plugging the above approximation into (B.19) for the second term of the r.h.s of (B.19), we get

$$
\begin{aligned}
\nabla^2\mathbf{E}(\mathbf{r}) - \{\frac{1}{c^2}e^{(+i\omega t)}\frac{\partial^2}{\partial t^2}[(1+\triangle\varepsilon(\mathbf{r},t))e^{(-i\omega t)}]\}\mathbf{E}(\mathbf{r}) &= \frac{-4\pi(i\omega)}{c}\mathbf{J}(\mathbf{r}) \\
\nabla^2\mathbf{E}(\mathbf{r}) + k^2[1+\triangle\varepsilon(\mathbf{r},t)]\mathbf{E}(\mathbf{r}) &= -4\pi ik\mathbf{J}(\mathbf{r}) \qquad \text{(B.22)}
\end{aligned}
$$

This is the vector wave equation for propagation through a random medium. The following assumptions were made in deriving this equation:

1. No polarization changes are induced by scattering in the medium.

2. ε changes slowly relative to $\mathbf{E}$, thus frequency mixing in the product $\varepsilon\mathbf{E}$ is negligible. This allows us to consider one frequency at a time.

3. $k\varepsilon(\mathbf{r},t) >> \frac{2}{c}\frac{\partial\triangle\varepsilon(\mathbf{r},t)}{\partial t} > \frac{1}{kc^2}\frac{\partial^2\triangle\varepsilon(\mathbf{r},t)}{\partial t^2}$.

Note that in (B.22) there is no interaction between the various components E_x, E_y and E_z of $\mathbf{E}$. Thus each component of the vector $\mathbf{E}$ must satisfy this equation, and we can now finally write the scalar wave equation [see also p.10 of [Whe01]].

$$\nabla^2 E(\mathbf{r}) + k^2[1+\triangle\varepsilon(\mathbf{r},t)]E(\mathbf{r}) \quad = \quad -4\pi i\omega k J(\mathbf{r}) \tag{B.23}$$

Where,

- $E(\mathbf{r})$ is either the x, y or z component of $\mathbf{E}$
- $\triangle\varepsilon(\mathbf{r},t)$ is the random fluctuation of the dielectric constant ε from it nominal value of 1 at position $\mathbf{r}$ at time t.
- $k = \frac{2\pi}{\lambda}$ is the wave number of the e-m wave under consideration.
- $J(\mathbf{r})$ is the x, y or z component of the current density vector $\mathbf{J}(\mathbf{r})$, depending on whether E_x, E_y or E_z is being considered.

If we are interested in solving this equation far from the transmitter than $J(\mathbf{r}) = 0$ and (B.23) becomes,

$$\nabla^2 E(\mathbf{r}) + k^2[1+\triangle\varepsilon(\mathbf{r},t)]E(\mathbf{r}) \quad = \quad 0 \tag{B.24}$$

This is the scalar wave equation that will be used with the Rytov approximation to obtain an expression for $E(\mathbf{r})$ that can be used to generate the statistics of the amplitude and the phase of $E(\mathbf{r})$ in the random media.

B.2 Rytov Solution to Random Wave Equation

To deal with weak scattering conditions, where ray optics-based analysis no longer holds, researchers use the Rytov approximation [Whe03] [Sas94] to solve the random wave equation. In this section we will again closely follow Wheelon's description provided in [Whe03, p.10-21].

There are two steps to the Rytov Approximation.

1. A transformation of the electric-field strength, E, that changes the random wave equation to one that can be solved in many interesting situations.

2. Expand the transformed field strength into a series of terms proportional to successively higher powers of the dielectric variations. The first term, which is proportional to $\triangle\varepsilon$, is the *basic Rytov* approximation, and provides a solution for E in weak scattering cases that well match experimental results.

Restating (B.23)

$$\nabla^2 E(\mathbf{r}) + k^2[1 + \triangle\varepsilon(\mathbf{r},t)]E(\mathbf{r}) \quad = \quad -4\pi i\omega k J(\mathbf{r})$$

This equation presents a difficulty in that the random quantity $\triangle\varepsilon(\mathbf{r},t)$ multiplies the function $E(\mathbf{r})$ we are trying to find. We would like to separate these two functions and have $\triangle\varepsilon(\mathbf{r},t)$appear an additional source term in B.23. Rytov's transformation achieves this goal as we shall below.

Let,

$$E(\mathbf{r}) = E_0(\mathbf{r})e^{\Psi(\mathbf{r})} \tag{B.25}$$

Where,

1. $E_0(\mathbf{r})$ is the coherent field strength that would be measured in the absence of irregularities (i.e. if $\triangle\varepsilon(\mathbf{r},t) = 0$ everywhere).

2. $\Psi(\mathbf{r})$is a *surrogate function* that must be discovered. It is assumed to be complex.

We will substitute $E(\mathbf{r}) = E_0(\mathbf{r})e^{\Psi(\mathbf{r})}$ into the random wave equation (B.23) and see what drops out, but first lets state the following identity which we will use in the substitution.

$$\nabla^2(FG) = F\nabla^2 G + 2\nabla F \cdot \nabla G + G\nabla^2 F \tag{B.26}$$

Which comes from,

$$\begin{aligned}
\nabla^2(FG) &\doteq \nabla \cdot \nabla(FG) \\
&= \nabla \cdot [F\nabla G + G\nabla F] \\
&= F\nabla^2 G + \nabla F \cdot \nabla G + G\nabla^2 F + \nabla G \cdot \nabla F \\
&= F\nabla^2 G + 2\nabla F \cdot \nabla G + G\nabla^2 F
\end{aligned}$$

Using this identity and the form $E(\mathbf{r}) = E_0(\mathbf{r})e^{(\Psi)}$ we have

$$\begin{aligned}
\nabla^2 E(\mathbf{r}) &= \nabla^2[E_0(\mathbf{r})e^{\Psi(\mathbf{r})}] \\
&= E_0(\mathbf{r})\nabla^2 e^{\Psi(\mathbf{r})} + 2\nabla E_0 \cdot \nabla e^{\Psi(\mathbf{r})} + e^{\Psi(\mathbf{r})}\nabla^2 E_0(\mathbf{r})
\end{aligned}$$

Putting this expression for $\nabla^2 E(\mathbf{r})$ and $E(\mathbf{r}) = E_0(\mathbf{r})e^{\Psi(\mathbf{r})}$ into (B.23) gives,

$$\begin{gathered}
\nabla^2 E(\mathbf{r}) + k^2[1 + \triangle\varepsilon(\mathbf{r},t)]E(\mathbf{r}) = -4\pi k j(\mathbf{r}) \\
\left[E_0(\mathbf{r})\nabla^2 e^{\Psi(\mathbf{r})} + 2\nabla E_0 \cdot \nabla e^{\Psi(\mathbf{r})} + e^{\Psi(\mathbf{r})}\nabla^2 E_0(\mathbf{r})\right] \\
+k^2\left[1 + \triangle\varepsilon(\mathbf{r},t)\right]E_0(\mathbf{r})e^{\Psi(\mathbf{r})} = -4\pi i k j(\mathbf{r})
\end{gathered}$$

Re-arranging gives,

$$E_0(\mathbf{r})\nabla^2 e^{\Psi(\mathbf{r})} + 2\nabla E_0(\mathbf{r})\cdot\nabla e^{\Psi(\mathbf{r})} + k^2\triangle\varepsilon(\mathbf{r},t)E_0(\mathbf{r})e^{\Psi(\mathbf{r})} = -4\pi ikj(\mathbf{r}) - e^{\Psi(\mathbf{r})}\nabla^2 E_0(\mathbf{r}) - k^2 E_0(\mathbf{r})e^{\Psi(\mathbf{r})} \quad \text{(B.27)}$$

Since by definition $E_0(\mathbf{r})$ is the solution of the wave equation without irregularities (i.e. with $\triangle\varepsilon = 0$), we know that,

$$\nabla^2 E_0(\mathbf{r}) + k^2 E_0(\mathbf{r}) = -4\pi ikj(\mathbf{r})$$

So the right hand side of (B.27) becomes

$$\begin{aligned} -4\pi ikj(\mathbf{r}) - e^{\Psi(\mathbf{r})}\nabla^2 E_0(\mathbf{r}) - k^2 E_0(\mathbf{r})e^{\Psi(\mathbf{r})} &= -4\pi ikj(\mathbf{r}) - e^{\Psi(\mathbf{r})}[\nabla^2 E_0(\mathbf{r}) + k^2 E_0(\mathbf{r})] \\ &= -4\pi ikj(\mathbf{r})[1 - e^{\Psi(\mathbf{r})}] \end{aligned}$$

And (B.27) itself becomes

$$E_0(\mathbf{r})\nabla^2 e^{\Psi(\mathbf{r})} + 2\nabla E_0\cdot\nabla e^{\Psi(\mathbf{r})} + k^2\triangle\varepsilon(\mathbf{r},t)E_0(\mathbf{r})e^{\Psi(\mathbf{r})} = 4\pi ikj(\mathbf{r})\left[e^{\Psi(\mathbf{r})} - 1\right]$$

Or,

$$\nabla^2 e^{\Psi(\mathbf{r})} + 2\frac{1}{E_0(\mathbf{r})}\nabla E_0(\mathbf{r})\cdot\nabla e^{\Psi(\mathbf{r})} + k^2\triangle\varepsilon(\mathbf{r},t)e^{\Psi(\mathbf{r})} = 4\pi ikj(\mathbf{r})\left[\frac{e^{\Psi(\mathbf{r})} - 1}{E_0(\mathbf{r})}\right] \quad \text{(B.28)}$$

Since $\nabla\ln(E_0(\mathbf{r})) = \frac{1}{E_0(\mathbf{r})}\nabla E_0(\mathbf{r})$ we can write (B.28) as

$$\nabla^2 e^{\Psi(\mathbf{r})} + 2\nabla\ln(E_0(\mathbf{r}))\cdot\nabla e^{\Psi(\mathbf{r})} + k^2\triangle\varepsilon(\mathbf{r},t)e^{\Psi(\mathbf{r})} = 4\pi ikj(\mathbf{r})\left[\frac{e^{\Psi(\mathbf{r})} - 1}{E_0(\mathbf{r})}\right] \quad \text{(B.29)}$$

Now,

$$\begin{aligned}
\nabla^2 e^{\Psi(\mathbf{r})} &\doteq \nabla \cdot \left[\nabla e^{\Psi(\mathbf{r})}\right] \\
&= \nabla \cdot \left[e^{\Psi(\mathbf{r})} \nabla \Psi(\mathbf{r})\right] \\
&= e^{\Psi(\mathbf{r})} \left[\nabla \Psi(\mathbf{r}) \cdot \nabla \Psi(\mathbf{r})\right] + e^{\Psi(\mathbf{r})} \nabla^2 \Psi(\mathbf{r}) \\
&= e^{\Psi(\mathbf{r})} \left[\nabla \Psi(\mathbf{r})\right]^2 + e^{\Psi(\mathbf{r})} \nabla^2 \Psi(\mathbf{r})
\end{aligned}$$

With this (B.29) becomes,

$$\begin{aligned}
&\left[e^{\Psi(\mathbf{r})} (\nabla \Psi(\mathbf{r}))^2 + e^{\Psi(\mathbf{r})} \nabla^2 \Psi(\mathbf{r})\right] \\
+2\nabla \ln\left(E_0(\mathbf{r})\right) \cdot e^{\Psi(\mathbf{r})} \nabla \Psi(\mathbf{r}) + k^2 \triangle \varepsilon(\mathbf{r},t) e^{\Psi(\mathbf{r})} &= 4\pi i k j(\mathbf{r}) \left[\frac{e^{\Psi(\mathbf{r})} - 1}{E_0(\mathbf{r})}\right]
\end{aligned} \tag{B.30}$$

When we are far from the transmitter $j(\mathbf{r}) = 0$, so the r.h.s. become zero. Canceling $e^{\Psi(\mathbf{r})}$ from every term we get,

$$\nabla^2 \Psi(\mathbf{r}) + (\nabla \Psi(\mathbf{r}))^2 + 2\nabla \ln\left(E_0(\mathbf{r})\right) \cdot \nabla \Psi(\mathbf{r}) = -k^2 \triangle \varepsilon(\mathbf{r},t) \tag{B.31}$$

We have achieved the goal! The random function $\triangle\varepsilon(\mathbf{r})$ no longer multiplies the function we are seeking, Ψ. Instead it appears as an artificial source term (i.e. it does not multiply the function for which we are solving).

The transmitter current density influences the solution of (B.31) through the unperturbed field strength E_0 which satisfies,

$$\nabla E_0(\mathbf{r}) + k^2 E_0(\mathbf{r}) = -4\pi i k j(\mathbf{r})$$

It will be notationally handy to represent E_0 as, $E_0 = e^{\psi_0(\mathbf{r})}$, where ψ_0 is a complex-valued function.

With this representation, the total field which travels through the medium can be written in

exponential form as,

$$E(\mathbf{r}) \doteq E_0(\mathbf{r})e^{\Psi(\mathbf{r})} = e^{\psi_0(\mathbf{r})+\Psi(\mathbf{r})}$$

Since,

$$\ln(E_0(\mathbf{r})) = \ln(e^{\psi_0(\mathbf{r})}) = \psi_0(\mathbf{r})$$

(B.31) becomes

$$\nabla^2\Psi(\mathbf{r}) + (\nabla\Psi(\mathbf{r}))^2 + 2\nabla\psi_0(\mathbf{r})\cdot\nabla\Psi(\mathbf{r}) = -k^2\triangle\varepsilon(\mathbf{r},t) \tag{B.32}$$

We see that the surrogate function Ψ must satisfy the the above *non-linear* partial differential equation. In order to make a solution possible to this non-linear equation we will expand $\Psi(\mathbf{r})$ into a series of terms, each proportional to a power of the random fluctuation in the permittivity, $\triangle\varepsilon$. This is step number 2 of the Rytov approximation. Let $\Psi(\mathbf{r})$ be expressed as,

$$\Psi(\mathbf{r}) = \psi_1(\mathbf{r}) + \psi_2(\mathbf{r}) + \psi_3(\mathbf{r}) + ... \tag{B.33}$$

where $\psi_n(\mathbf{r})$ is proportional to the n^{th} power of $\triangle\varepsilon$. Substituting $\Psi(\mathbf{r}) = \sum_{n=1}^{\infty}\psi(\mathbf{r})$ into (B.32) we get,

$$\begin{aligned}
\nabla^2\Psi(\mathbf{r}) + (\nabla\Psi(\mathbf{r}))^2 + 2\nabla\psi_0(\mathbf{r})\cdot\nabla\Psi(\mathbf{r}) + k^2\triangle\varepsilon(\mathbf{r},t) &= 0 \\
\left[\sum_{n=1}^{\infty}\nabla^2\psi_n(\mathbf{r})\right] + \left[\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\nabla\psi_n(\mathbf{r})\cdot\nabla\psi_k(\mathbf{r})\right] & \\
+\sum_{n=1}^{\infty}2\nabla\psi_0(\mathbf{r})\cdot\nabla\psi_n(\mathbf{r}) + k^2\triangle\varepsilon(\mathbf{r},t) &= 0
\end{aligned} \tag{B.34}$$

Grouping terms with like power of $\triangle\varepsilon$ in (B.34) we get a set of equations that can be used to obtain successive approximations (a.k.a terms ψ_n) to the surrogate function $\Psi(\mathbf{r})$. Note that $\nabla\psi_n\cdot\nabla\psi_k$ is proportional to $\triangle\varepsilon^n\triangle\varepsilon^k = \triangle\varepsilon^{n+k}$. Terms proportional to $\triangle\varepsilon$ provide the following

partial differential equation,

$$\begin{aligned} \nabla^2\psi_1 + 0 + 2\psi_0 \cdot \nabla\psi_1 + k^2 \triangle\varepsilon &= 0 \\ \nabla^2\psi_1 + 2\psi_0 \cdot \nabla\psi_1 + k^2 \triangle\varepsilon &= 0 \end{aligned} \tag{B.35}$$

The above equation is called the *basic Rytov equation* and it provides a means for solving for the *basic Rytov* solution ψ_1, which is the solution used in most estimates of signal variations due to turbulence effects, mainly in the weak scattering regime.

Terms proportional to $\triangle\varepsilon^2$ provide,

$$\begin{aligned} \nabla^2\psi_2 + \nabla\psi_1 \cdot \psi_1 + 2\psi_0 \cdot \nabla\psi_2 &= 0 \\ \nabla^2\psi_2 + 2\psi_0 \cdot \nabla\psi_2 + (\nabla\psi_1)^2 &= 0 \end{aligned}$$

The second order solution ψ_2 comes into play when a more precise estimate is needed for the surrogate function, Ψ, and thus for the field strength E, than that provided by ψ_1. This situation comes about as the turbulence increases in strength from the weak scattering regime.

Terms proportional to $\triangle\varepsilon^3$ give,

$$\nabla^2\psi_3 + 2\psi_0 \cdot \nabla\psi_3 + 2\nabla\psi_1 \cdot \nabla\psi_2 = 0 \tag{B.36}$$

Terms proportional to $\triangle\varepsilon^4$ provide,

$$\nabla^2\psi_4 + 2\psi_0 \cdot \nabla\psi_4 + 2\nabla\psi_1 \cdot \nabla\psi_3 + (\nabla\psi_2)^2 = 0 \tag{B.37}$$

In general $\triangle\varepsilon^n$ terms provide,

$$\nabla^2\psi_n + 2\psi_0 \cdot \nabla\psi_n + \sum_{p=1}^{n-1} \nabla\psi_p \cdot \nabla\psi_{n-p} = 0 \tag{B.38}$$

The vacuum field $E_0 = e^{\psi_0}$ influences each succeeding approximation through the coupling

term $2\nabla\psi_0 \cdot \nabla\psi_n$.

We know that $E_0 = e^{\Psi_0}$ satisfies

$$\begin{aligned} \nabla^2 E_0(\mathbf{r}) + k^2 E_0(\mathbf{r}) &= 0 \\ \nabla^2 e^{\Psi_0(\mathbf{r})} + k^2 e^{\Psi_0(\mathbf{r})} &= 0 \end{aligned} \tag{B.39}$$

From earlier, $\nabla^2 e^{\Psi_0} = e^{\Psi_0}(\nabla\psi_0)^2 + e^{\Psi_0}\nabla^2\psi_0$. So we have,

$$\begin{aligned} \left[e^{\Psi_0}(\nabla\psi_0)^2 + e^{\Psi_0}\nabla^2\psi_0\right] + k^2 e^{\Psi_0(\mathbf{r})} &= 0 \\ \nabla^2\psi_0 + (\nabla\psi_0)^2 + k^2 &= 0 \end{aligned} \tag{B.40}$$

This equation provides a means for finding the unperturbed field's surrogate, ψ_0, which can then be used in the above equation to get solutions for ψ_1, ψ_2, ...ψ_n, for some desired value of n. The complexity of the solutions increases as n increases. Fortunately, the $n = 1$ approximation provides an approximation that is quite good under weak scattering conditions. As noted previously, the $n = 1$ approximation provides the *basic Rytov solution*, which will be denoted E_1,

$$E_1(\mathbf{r}) \doteq E_0(\mathbf{r}) e^{\Psi_1(\mathbf{r})} = e^{\Psi_0(\mathbf{r}) + \Psi_1(\mathbf{r})} \tag{B.41}$$

The equation that allows us to solve for ψ_1 is (B.35),

$$\nabla^2\psi_1 + 2\psi_0 \cdot \nabla\psi_1 = -k^2 \triangle\varepsilon$$

In order to solve this equation we will make yet another substitution. Let,

$$\psi_1(\mathbf{r}) = Q(\mathbf{r}) e^{-\Psi_0(\mathbf{r})} \tag{B.42}$$

Putting this into (B.35) will result in a solvable expression for $Q(\mathbf{r})$. First note that using the

identity,

$$\nabla^2(FG) \;=\; F\nabla^2 G + 2\nabla F\cdot\nabla G + G\nabla^2 F$$

we have,

$$\begin{aligned}
\nabla^2\psi_1(\mathbf{r}) &= \nabla^2\big(Q(\mathbf{r})e^{-\psi_0(\mathbf{r})}\big)\\
&= Q(\mathbf{r})\nabla^2 e^{-\psi_0} + 2\nabla Q(\mathbf{r})\cdot\nabla e^{-\psi_0(\mathbf{r})} + e^{-\psi_0(\mathbf{r})}\nabla^2 Q(\mathbf{r})\\
&= e^{-\psi_0(\mathbf{r})}\nabla^2 Q(\mathbf{r}) - 2e^{-\psi_0(\mathbf{r})}\nabla Q(\mathbf{r})\cdot\nabla\psi_0(\mathbf{r})\\
&\quad + Q(\mathbf{r})e^{-\psi_0(\mathbf{r})}\Big[-\nabla^2\psi_0(\mathbf{r}) + \big(\nabla\psi_0(\mathbf{r})\big)^2\Big]
\end{aligned}$$

Putting our new expression for ψ_1 and the resulting expression for $\nabla^2\psi_1$ into (B.35) gives

$$\begin{aligned}
\nabla^2\psi_1 + 2\psi_0\cdot\nabla\psi_1 &= -k^2\triangle\varepsilon\\
e^{-\psi_0(\mathbf{r})}\Big[\nabla^2 Q(\mathbf{r}) - 2\nabla Q(\mathbf{r})\cdot\nabla\psi_0(\mathbf{r}) + Q(\mathbf{r})\big[-\nabla^2\psi_0(\mathbf{r}) + (\nabla\psi_0(\mathbf{r}))^2\big]&\\
+ 2\nabla\psi_0(\mathbf{r})\cdot(\nabla Q(\mathbf{r}) - Q(\mathbf{r})\nabla\psi_0(\mathbf{r}))\Big] &= -k^2\triangle\varepsilon\\
\nabla^2 Q(\mathbf{r}) - 2\nabla Q(\mathbf{r})\cdot\nabla\psi_0(\mathbf{r}) + Q(\mathbf{r})\left[-\nabla^2\psi_0(\mathbf{r}) + (\nabla\psi_0(\mathbf{r}))^2\right]&\\
+ 2\nabla\psi_0(\mathbf{r})\cdot(\nabla Q(\mathbf{r}) - Q\nabla\psi_0(\mathbf{r})) &= -k^2 e^{\psi_0(\mathbf{r})}\triangle\varepsilon\\
\nabla^2 Q(\mathbf{r}) + Q(\mathbf{r})\left[-\nabla^2\psi_0(\mathbf{r}) - (\nabla\psi_0(\mathbf{r}))^2\right] &= -k^2 e^{\psi_0(\mathbf{r})}\triangle\varepsilon
\end{aligned}$$

In the above we used the identity $(\nabla\psi_0)^2 \doteq \nabla\psi_0\cdot\nabla\psi_0$ to arrive at the final expression. This expression can be further reduced using, (B.40) to simplify the second term on the l.h.s. to obtain,

$$\nabla^2 Q(\mathbf{r}) + k^2 Q(\mathbf{r}) = -k^2 e^{\psi_0(\mathbf{r})}\triangle\varepsilon \tag{B.43}$$

This expression shows the benefit of using the $\psi_1 = Qe^{\psi_0}$ substitution. We now have an expression with only one term, Q, and it is located in a simplified wave equation for which the solution is known. Note that the quantity on the r.h.s is considered as an artificial source term.

Green's Function provides the solution to the above wave equation. Green's Function,

$G(\mathbf{R},\mathbf{r})$, is the solution to the partial differential equation,

$$\nabla_r G(\mathbf{R},\mathbf{r}) + k^2 G(\mathbf{R},\mathbf{r}) = \triangle(\mathbf{R}-\mathbf{r}) \tag{B.44}$$

The solution is,

$$G(\mathbf{R},\mathbf{r}) = \frac{e^{ik|\mathbf{R}-\mathbf{r}|}}{4\pi|\mathbf{R}-\mathbf{r}|} \tag{B.45}$$

which represents a spherical wave of unit amplitude at position $\mathbf{R}$ originating from a point source at $\mathbf{r}$. Note that (B.43) looks like (B.44) except that the source term (artificial) is not a point source but is distributed over some volume. The solution to (B.43) will be the superposition of Green's functions, where each Green's function is viewed as coming from a different part of the source, which is viewed as a composite of point sources.

$$Q(\mathbf{R}) = \int d^3\mathbf{r} \left[-k^2 e^{\psi_0(\mathbf{r})} \triangle\varepsilon(\mathbf{r},t)\right] G(\mathbf{R},\mathbf{r})$$

The term in brackets is the value of the artificial source at position $\mathbf{r}$, and $G(\mathbf{R},\mathbf{r})$ is the Greens function at $\mathbf{R}$ for a source located at $\mathbf{r}$.

Recall from (B.42) that we defined ψ_1 as,

$$\psi_1(\mathbf{r}) \doteq Q(\mathbf{r}) e^{-\psi_0(\mathbf{r})}$$

Also recall that,

$$E_0(\mathbf{r}) \doteq e^{-\psi_0(\mathbf{r})}$$

so,

$$\psi_1(\mathbf{R}) = \frac{Q(\mathbf{R})}{E_0(\mathbf{R})}$$

Thus,

$$\psi_1(\mathbf{R}) = -k^2 \int d^3\mathbf{r} G(\mathbf{R},\mathbf{r}) \triangle\varepsilon(\mathbf{r},t) \frac{E_0(\mathbf{r})}{E_0(\mathbf{R})} \tag{B.46}$$

This is the final form for the *basic Rytov* solution!

This type of equation, i.e. (B.35),

$$\nabla^2\psi_n + 2\psi_0 \cdot \nabla\psi_n \quad = \quad -(\textbf{some "source" term})$$

shows up repeatedly in solving for the higher order Rytov approximations. From the above we know that the solution will be,

$$\psi_n(\mathbf{R}) = -\int d^3\mathbf{r} G(\mathbf{R},\mathbf{r})(\textbf{some "source" term})\frac{E_0(\mathbf{r})}{E_0(\mathbf{R})} \tag{B.47}$$

Let's look at the amplitude and phase of the basic Rytov solution,

$$E_1(\mathbf{R}) = E_0(\mathbf{R})e^{\psi_1(\mathbf{R})} = e^{\psi_0(\mathbf{R})+\psi_1(\mathbf{R})} \tag{B.48}$$

Let's write this in amplitude and phase form,

$$E_1(\mathbf{R}) = A(\mathbf{R})e^{\phi_0(\mathbf{R})+\phi_1(\mathbf{R})} \tag{B.49}$$

The fluctuating part of the signal phase is, from (B.46)

$$\begin{aligned}\phi(\mathbf{R}) &= Im\left[-k^2\int d^3\mathbf{r} G(\mathbf{R},\mathbf{r})\triangle\varepsilon(\mathbf{r},t)\tfrac{E_0(\mathbf{r})}{E_0(\mathbf{R})}\right] \\ &= -k^2\int d^3\mathbf{r}\triangle\varepsilon(\mathbf{r},t) Im\left[G(\mathbf{R},\mathbf{r})\tfrac{E_0(\mathbf{r})}{E_0(\mathbf{R})}\right]\end{aligned} \tag{B.50}$$

The signal amplitude is,

$$\begin{aligned}A(\mathbf{R}) \qquad &= \mid E_1(\mathbf{R}) \mid = \mid E_0(\mathbf{R}) \mid\mid e^{\psi_1(\mathbf{R})} \mid \\ &= \mid E_0(\mathbf{R}) \mid e^{Re(\psi_1(\mathbf{R}))} \\ &= \mid E_0(\mathbf{R}) \mid e^{-k^2\int d^3\mathbf{r}\triangle\varepsilon(\mathbf{r},t)Re\left[G(\mathbf{R},\mathbf{r})\frac{E_0(\mathbf{r})}{E_0(\mathbf{R})}\right]}\end{aligned} \tag{B.51}$$

Normally, we deal with the normalized log-amplitude,

$$\chi(\mathbf{R}) \doteq \log \frac{A(\mathbf{R})}{|E_0(\mathbf{R})|} = -k^2 \int d^3\mathbf{r} \triangle\varepsilon(\mathbf{r},t) Re\left[G(\mathbf{R},\mathbf{r}) \frac{E_0(\mathbf{r})}{E_0(\mathbf{R})}\right] \tag{B.52}$$

Recall that $E_0(\mathbf{R})$ is what the field at $\mathbf{R}$ would be if there were no perturbations.

Using the notation,

$$G(\mathbf{R},\mathbf{r}) \frac{E_0(\mathbf{R})}{E_0(\mathbf{r})} = A(\mathbf{R},\mathbf{r}) + iB(\mathbf{R},\mathbf{r})$$

We can express the amplitude and phase fluctuations, from the basic Rytov solution, as

$$\begin{aligned} \chi(\mathbf{R}) \doteq a(\mathbf{R}) &= -k^2 \int d^3\mathbf{r} A(\mathbf{R},\mathbf{r}) \triangle\varepsilon(\mathbf{r},t) \\ \phi(\mathbf{R}) \doteq b(\mathbf{R}) &= -k^2 \int d^3\mathbf{r} B(\mathbf{R},\mathbf{r}) \triangle\varepsilon(\mathbf{r},t) \end{aligned}$$

Since $< \triangle\varepsilon(\mathbf{r},t) >= 0$, meaning that the average random fluctuation (positive and negative0 is zero, we have

$$\begin{aligned} < \chi(\mathbf{R}) >= 0 \\ < \phi(\mathbf{R}) >= 0 \end{aligned}$$

Keep in mind this is for the basic Rytov solution (i.e. ψ_1), once the weak scattering regime is left and other orders of the Rytov solution are needed to model the actual situation, we will get to the point where $< \chi > \neq 0$.

The variances of each, assuming $< \chi >= 0$ and $< \phi >= 0$, are given by,

$$\sigma_\chi^2 = < \chi^2(\mathbf{R}) >= k^4 \int d^3\mathbf{r} A(\mathbf{R},\mathbf{r}) \int d^3\mathbf{r}' A(\mathbf{R},\mathbf{r}') < \triangle\varepsilon(\mathbf{r},t) \triangle\varepsilon(\mathbf{r}',t) > \tag{B.53}$$

and,

$$\sigma_{\phi}^2 =< \phi^2(\mathbf{R}) >= k^4 \int d^3\mathbf{r} B(\mathbf{R},\mathbf{r}) \int d^3\mathbf{r}' B(\mathbf{R},\mathbf{r}') < \triangle\varepsilon(\mathbf{r},t)\triangle\varepsilon(\mathbf{r}',t) > \tag{B.54}$$

We see that the variances depend on the spatial correlation of the index of refraction irregularities

$$< \triangle\varepsilon(\mathbf{r},t)\triangle\varepsilon(\mathbf{r}',t) >$$

With homogeneous medium, we can express the spatial correlation in terms of its Fourier wavenumber decomposition analogous to how the temporal correlation of a random process can be expressed as the inverse Fourier transform (FT) of the power spectral density (PSD).

$$< \triangle\varepsilon(\mathbf{r},t)\triangle\varepsilon(\mathbf{r}',t) >\doteq \int d^3\kappa \Phi_{\varepsilon}(\kappa) e^{\kappa\cdot(\mathbf{r}-\mathbf{r}')} \tag{B.55}$$

This substitution separates the wave-scattering features of the problem (i.e. the geometry of the problem and the type of wave) from the description of the turbulent medium. When we put (B.55) into (B.53) we get,

$$\sigma_{\chi}^2 =< \chi^2(\mathbf{R}) >= k^4 \int d^3\kappa \Phi_{\varepsilon}(\kappa) \int d^3\mathbf{r} A(\mathbf{R},\mathbf{r}) e^{i\kappa\cdot\mathbf{r}} \int d^3\mathbf{r}' A(\mathbf{R},\mathbf{r}') e^{-i\kappa\cdot\mathbf{r}'} \tag{B.56}$$

The turbulence properties are located in $\Phi_{\varepsilon}(\kappa)$, while the geometric properties of the propagation (without turbulence-induced index of refraction irregularities) are located in $A(\mathbf{R},\mathbf{r})$. Expressing $< \triangle\varepsilon(\mathbf{r},t)\triangle\varepsilon(\mathbf{r}',t) >$ in terms of the FT of the power spectrum of the index of refraction irregularities allowed this separation. Note that this separation via the expression of the spatial correlation in terms of the F.T. of a spectrum is valid only for homogenous medium, strictly speaking, just like the autocorrelation of a random process in time can be expressed as the FT of the PSD only if the process is wide sense stationary.

Let us consider briefly how the inhomogenous situation might be handled. This will aid

in the analysis of a satellite signal's propagation down through the atmosphere, in which the turbulence activity changes as a function of height. When the medium is not homogenous, approximations can be made that reflect the inhomogeneity of the medium while still allowing the above separation. In one approximation the turbulence spectrum is made to depend on the average position of the two points, i.e.

$$\Phi_\varepsilon(\kappa, \frac{\mathbf{r}+\mathbf{r}'}{2})$$

is used as the spectrum expression. But, this does not allow the turbulence properties to be separated from the geometric properties (i.e. the κ integral is coupled to the $\mathbf{r}$ and $\mathbf{R}$ integrals).

In a second approximation for inhomogeneous media the spatial correlation is considered to be able to be written as

$$< \triangle\varepsilon(\mathbf{r})\triangle\varepsilon(\mathbf{r}') >= \triangle\varepsilon_{rms}(\mathbf{r})\triangle\varepsilon_{rms}(\mathbf{r}')C(\mathbf{r}-\mathbf{r}')$$

Where $C(\cdot)$ is a function of only the separation between the two point $\mathbf{r}$ and $\mathbf{r}'$. The function $\triangle\varepsilon_{rms}(\mathbf{a})$ is the rms value of $\triangle\varepsilon$ at position $\mathbf{a}$. The rms values at a position are proportional to the local value of the index of refraction structure constant, C_n^2, at that position. We now introduce a turbulence-profile function,

$$\wp(\mathbf{r}) = \sqrt{\frac{C_n^2(\mathbf{r})}{C_n^2(\mathbf{0})}} \tag{B.57}$$

Note that this expression can also be used for the homogenous expression, in which case $\wp(\mathbf{r}) = 1$. We can now represent the spatial correlation of $\triangle\varepsilon$, as,

$$< \triangle\varepsilon(\mathbf{r})\triangle\varepsilon(\mathbf{r}') >= \wp(\mathbf{r})\wp(\mathbf{r}') < \triangle\varepsilon^2 > C(\mathbf{r}-\mathbf{r}') \tag{B.58}$$

We can represent the $C(\mathbf{r}-\mathbf{r}')$ function by a wave number integral,

$$C(\mathbf{r}-\mathbf{r}') \doteq \frac{1}{<\triangle\varepsilon^2>}\int d^3\kappa\Phi_\varepsilon(\kappa)e^{\kappa\cdot(\mathbf{r}-\mathbf{r}')}$$

and thus,

$$<\triangle\varepsilon(\mathbf{r})\triangle\varepsilon(\mathbf{r}')>=\wp(\mathbf{r})\wp(\mathbf{r}')\int d^3\kappa\Phi_\varepsilon(\kappa)e^{\kappa\cdot(\mathbf{r}-\mathbf{r}')} \tag{B.59}$$

Using (B.59) for the spatial correlation in an inhomogeneous medium we can write the variance of $\chi(\mathbf{R})$ from (B.53) as

$$\begin{aligned}<\chi^2(\mathbf{R})> &= k^4\int d^3\mathbf{r}A(\mathbf{R},\mathbf{r})\int d^3\mathbf{r}'A(\mathbf{R},\mathbf{r}')<\triangle\varepsilon(\mathbf{r},t)\triangle\varepsilon(\mathbf{r}',t)> \\ &= k^4\int d^3\kappa\Phi_\varepsilon(\kappa)\int d^3\mathbf{r}A(\mathbf{R},\mathbf{r})\wp(\mathbf{r})e^{i\kappa\cdot\mathbf{r}}\int d^3\mathbf{r}'A(\mathbf{R},\mathbf{r}')\wp(\mathbf{r})e^{-i\kappa\cdot\mathbf{r}'}\end{aligned} \tag{B.60}$$

Once again, just as in the homogenous situation, given in (B.56), we have the desired separation. We can integrate over $\mathbf{r}$ and $\mathbf{r}'$, obtaining an expression in terms of κ, and then put in the expression for the spectrum of irregularities, $\Phi_\varepsilon(\kappa)$, and integrate over κ. This allows the analysis of various turbulence conditions to be analyzed, without redoing the entire problem, through the substitution of various expressions for $\Phi_\varepsilon(\kappa)$. As a result, this second expression for dealing with an inhomogeneous medium is often used. As mentioned previously, the profile function of C_n^2 is often a function of only elevation, i.e. $C_n^2(\mathbf{r}) = C_n^2(z)$.

APPENDIX C

Constrained Capacity Formula

C.1 Constrained Capacity

This appendix derives a formula for the constrained capacity, denoted C_{con}, used throughout this work. This quantity is also known as the *symmetric capacity* [Bac99] or as the *joint capacity* [BJF07]. It measures the channel capacity well for certain systems that are constrained to used QAM modulations with equiprobable input symbols. These systems include: 1) uncoded systems, 2) systems that use symbol-based error correction codes (e.g. trellis coded modulation), and 3) systems that use a binary error correction code (e.g. low density parity check (LDPC)) along with belief propagation iterations between the decoder and constellation demapper. If a binary error correction codes is used without belief propagation iterations between the decoder and demapper than the parallel decoding capacity, C_{PD}, described in the next section is a better measure of the system's capacity [BJF07].

Let X, be drawn uniformly from a finite alphabet of size M, i.e.,

$$X \in \{\alpha_1, \alpha_2, ..., \alpha_M\}$$

Let Y be the output of the AWGN channel. Let the constrained capacity of the channel be defined as the mutual information between X and Y when the input is constrained to be drawn from the finite alphabet in a uniform manner. The derivation of the constrained capacity of this channel, for a given channel gain g, begins with the basic definition,

$$C_{con}(g) \doteq H(X) - H(X|Y,g) \tag{C.1}$$

The * denotes that this is the capacity under the given constraints. Since the input distribution is uniform, $p(X) = \frac{1}{M}$, the first term on the right hand side of (C.1) is, $H(X) = \log_2(M)$.

The derivation of an expression for $H(X|Y,g)$ is more involved,

$$\begin{aligned}
H(X|Y,g) = & \\
& E_Y[H(X|Y=y,g)] \\
= & \int_{y\in C} f_Y(y) H(X|Y=y,g)\mathrm{d}y \\
= & \int_{y\in C} f_Y(y)\Big[\sum_{x\in X} p(X=x|Y=y,g) \\
& \cdot\log_2(\frac{1}{p(X=x|Y=y,g)})\Big]\,\mathrm{d}y \\
= & \int_{y\in C} f_Y(y)\Big[\sum_{j=1}^{j=M} p(X=\alpha_j|Y=y,g) \\
& \cdot\log_2(\frac{1}{p(X=\alpha_j|Y=y,g)})\Big]\,\mathrm{d}y \\
= & \sum_{j=1}^{j=M}\int_{y\in C} f_Y(y) p(X=\alpha_j|Y=y,g) \\
& \log_2\Big(\frac{1}{p(X=\alpha_j|Y=y,g)}\Big)\,\mathrm{d}y
\end{aligned} \tag{C.2}$$

The probability, $p(X=\alpha_j|Y=y,g)$, can be expressed, using Baye's rule as the starting point, as,

$$\begin{aligned}
p(X=\alpha_j|Y=y,g) = & \\
& f(Y|X=\alpha_j,g)\frac{p(X=\alpha_j)}{f(Y|g)} \\
= & \Big[\frac{1}{\pi N_0}\exp\Big(-\frac{\|y-g\alpha_j\|^2}{No}\Big)\Big]\Big[\frac{1}{M}\Big] \\
& \cdot\Big[\frac{1}{M}\sum_{i=1}^{i=M}\frac{1}{\pi N_0}\exp\Big(-\frac{\|y-g\alpha_i\|^2}{No}\Big)\Big]^{-1} \\
= & \Big[\sum_{i=1}^{i=M}\exp\Big(\frac{\|y-g\alpha_j\|^2-\|y-g\alpha_i\|^2}{No}\Big)\Big]^{-1}
\end{aligned} \tag{C.3}$$

Applying Baye's rule again, this time to $f_Y(y)p(X=\alpha_j|Y=y,g)$, gives,

$$\begin{aligned} & f_Y(y)p(X=\alpha_j|Y=y,g) \\ &= f(y|X=\alpha_j,g)p(X=\alpha_j) \\ &= \frac{1}{\pi N_0}\exp(-\frac{\|y-g\alpha_j\|^2}{No})\frac{1}{M} \end{aligned} \tag{C.4}$$

Inserting (C.3) and (C.4) into (C.2) gives,

$$\begin{aligned} & H(X|Y,g) = \\ & \quad \sum_{j=1}^{j=M}\int_{y\in C}\Big[\frac{1}{\pi N_0}\exp(-\frac{\|y-g\alpha_j\|^2}{No})\frac{1}{M}\Big] \\ & \quad \cdot\log_2\Big(\frac{1}{p(X=\alpha_j|Y=y,g)}\Big)\,\mathrm{d}y \\ &= \frac{1}{M}\frac{1}{\pi N_0}\sum_{j=1}^{j=M}\int_{y\in C}\exp\Big(-\frac{\|y-g\alpha_j\|^2}{No}\Big) \\ & \quad \cdot\log_2\left(\sum_{i=1}^{i=M}\exp\Big(\frac{(\|y-g\alpha_j\|^2-\|y-g\alpha_i\|^2)}{No}\Big)\right)\mathrm{d}y \end{aligned} \tag{C.5}$$

Placing (C.5) into (C.1) provides the desired formula,

$$\begin{aligned} & C_{con}(g) \doteq \\ & \quad H(X)-H(X|Y,g) \\ &= \log_2(M)-\frac{1}{M}\frac{1}{\pi N_0}\sum_{j=1}^{j=M}\int_{y\in C}\exp\Big(-\frac{\|y-g\alpha_j\|^2}{No}\Big) \\ & \quad \cdot\log_2\left(\sum_{i=1}^{i=M}\exp\Big(\frac{(\|y-g\alpha_j\|^2-\|y-g\alpha_i\|^2)}{No}\Big)\right)\mathrm{d}y \end{aligned} \tag{C.6}$$

C.2 Parallel Decoding Capacity

Systems that use an M-ary QAM modulation along with a binary code and have no belief propagation iterations between the decoder and the constellation demapper can be considered

to consist of $l = \log_2(M)$ parallel binary channels. The capacity of such a system is the parallel decoding capacity, given by

$$C_{PD} = \sum_{i=0}^{l-1} I(B_i;Y), \tag{C.7}$$

where B_i is the i^{th} bit of the l-bit transmitted symbol, Y is the received symbol, and $I(U;V)$ denotes the mutual information between random variables U and V.

From the definition of mutual information

$$I(B_i;Y) = H(B_i) - H(B_i|Y) \tag{C.8}$$

Given that B_i is a binary random variable with a uniform distribution (i.e. p(B_i=0) = p(B_i=1) = 1/2) inherited from the uniform distribution of the input symbols, the entropy of B_i is 1 bit: $H(B_i)$= 1.

From the definition for conditional entropy

$$\begin{aligned} H(B_i|Y) &= E_y\Big[H(B_i|Y=y)\Big] \\ &= E_y\Big[\sum_{b_i=\{0,1\}} p(b_i|y)\log_2(1/p(b_i|y))\Big] \\ &= \int_{y\in\mathcal{C}} \sum_{b_i=\{0,1\}} p(b_i|y)f(y)\log_2(1/p(b_i|y))\,\mathrm{d}y. \end{aligned}$$

Applying Baye's rule this becomes

$$\begin{aligned} H(B_i|Y) &= \int_{y\in\mathcal{C}} \sum_{b_i=\{0,1\}} f(y|b_i)p(b_i)\log_2\Big(\frac{f(y)}{f(y|b_i)p(b_i)}\Big)\mathrm{d}y \\ &= \frac{1}{2}\sum_{b_i=\{0,1\}}\int_{y\in\mathcal{C}} f(y|b_i)\log_2\left(\frac{\sum_{b_j=\{0,1\}} f(y|b_j)p(b_j)}{f(y|b_i)p(b_i)}\right)\mathrm{d}y \\ &= \frac{1}{2}\sum_{b_i=\{0,1\}}\int_{y\in\mathcal{C}} f(y|b_i)\log_2\left(\frac{\sum_{b_j=\{0,1\}} f(y|b_j)}{f(y|b_i)}\right)\mathrm{d}y. \end{aligned}$$

Let S_0 (S_1) be the set indices for symbols $\alpha \in X$ that have a value of zero (one) for the i^{th} bit. Then the above can be expressed as

$$H(B_i|Y) = \frac{1}{2} \sum_{b_i=\{0,1\}} \int_{y\in C} \Big(\sum_{j\in S_{b_i}} f(y|\alpha_j) \Big) \log_2 \left(\frac{\sum_{m=1}^{M} f(y|\alpha_m)}{\sum_{j\in S_{b_i}} f(y|\alpha_j)} \right) \mathrm{d}y.$$

Assuming an AWGN channel and allowing for a channel gain of g, this becomes

$$\begin{aligned} H(B_i|Y,g) = {} & \frac{1}{2\pi N_0} \sum_{b_i=\{0,1\}} \int_{y\in C} \left(\sum_{j\in S_{b_i}} \exp\Big(- \frac{\|y - g\alpha_j\|^2}{No} \Big) \right) \cdot \\ & \log_2 \left(\frac{\sum_{m=1}^{M} \frac{1}{\pi N_0} \exp\Big(- \frac{\|y-g\alpha_m\|^2}{No} \Big)}{\sum_{j\in S_{b_i}} \frac{1}{\pi N_0} \exp\Big(- \frac{\|y-g\alpha_j\|^2}{No} \Big)} \right) \mathrm{d}y. \end{aligned} \tag{C.9}$$

Using (C.9) and (C.8) in (C.7) results in an expression for the parallel decoding capacity for a system that uses an M-ary QAM constellation in a AWGN channel

$$\begin{aligned} C_{PD} = {} & \sum_{i=0}^{l-1} I(B_i;Y) \\ = {} & \sum_{i=0}^{l-1} \left[1 - \frac{1}{2\pi N_0} \sum_{b_i=\{0,1\}} \int_{y\in C} \left(\sum_{j\in S_{b_i}} \exp\Big(- \frac{\|y - g\alpha_j\|^2}{No} \Big) \right) \cdot \right. \\ & \left. \log_2 \left(\frac{\sum_{m=1}^{M} \frac{1}{\pi N_0} \exp\Big(- \frac{\|y-g\alpha_m\|^2}{No} \Big)}{\sum_{j\in S_{b_i}} \frac{1}{\pi N_0} \exp\Big(- \frac{\|y-g\alpha_j\|^2}{No} \Big)} \right) \mathrm{d}y \right]. \end{aligned} \tag{C.10}$$

APPENDIX D

Lower Bound On the Cutoff Rate in a LLN Channel

Expression for LB1 for LLN Channel

A closed form approximation for (2.19) can be made using the proposed piece-wise approximation, over each of the three regions of the approximation.

Evaluation of the integral on region 1 ($u \in [-\infty, b_1)$)

With the approximation that $e^{-2e^u/\kappa} = 1$ on the first region, the integral of (2.19) over this region becomes,

$$\begin{aligned}\Theta_1 &= \frac{1}{\sqrt{2\pi\sigma_c^2}} \int_{-\infty}^{b_1} e^{-(\frac{(u-m)^2}{2\sigma_c^2} + \frac{\|\alpha_j - \alpha_s\|^2}{4N_0})} \mathrm{d}u \\ &= e^{-\frac{\|\alpha_j - \alpha_s\|^2}{4N_0}} \frac{1}{\sqrt{2\pi\sigma_c^2}} \int_{-\infty}^{b_1} e^{-\frac{(u-m)^2}{2\sigma_c^2}} \mathrm{d}u\end{aligned}$$

Substituting $v = \frac{u-m}{\sigma_c}$ gives,

$$\begin{aligned}\Theta_1 &= e^{-\frac{\|\alpha_j - \alpha_s\|^2}{4N_0}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{m-b_1}{\sigma_c}} e^{-\frac{v^2}{2}} dv \\ &= e^{-\frac{\|\alpha_j - \alpha_s\|^2}{4N_0}} \frac{1}{\sqrt{2\pi}} \int_{\frac{m-b_1}{\sigma_c}}^{\infty} e^{-\frac{v^2}{2}} dv \\ &= e^{-\frac{\|\alpha_j - \alpha_s\|^2}{4N_0}} Q(\frac{m-b_1}{\sigma_c}) \qquad \text{(D.1)}\end{aligned}$$

Evaluation of the integral on region 2 ($u \in [b_1, b_2]$)

With the approximation that $e^{-2e^u/\kappa} = 1 - (\frac{u-b_1}{b_2-b_1})^2$ on the second region, the integral of (2.19) over this region becomes,

$$\begin{aligned}
\Theta_2 &= \frac{1}{\sqrt{2\pi\sigma_c^2}} \int_{b_1}^{b_2} e^{-(\frac{(u-m)^2}{2\sigma_c^2} + \frac{\|\alpha_j - \alpha_s\|^2}{4N_0}(1-(\frac{u-b_1}{b_2-b_1})^2))} \mathrm{d}u \\
&= \frac{1}{\sqrt{2\pi\sigma_c^2}} \int_{b_1}^{b_2} e^{-((\frac{1}{2\sigma_c^2} - \frac{k_{j,s}}{(b_2-b_1)^2})u^2 + (\frac{2k_{j,s}}{(b_2-b_1)^2} - \frac{m}{\sigma_c^2})u)} \\
&\quad \cdot e^{-(\frac{m^2}{2\sigma_c^2} + k_{j,s}(1 - \frac{b_1^2}{(b_2-b_1)^2}))} \mathrm{d}u
\end{aligned}$$

where $k_{j,s} = \frac{\|\alpha_j - \alpha_s\|^2}{4N_0}$. Completing the square in the argument of the exponential and making some variable substitutions for the sake of simplification, the argument becomes,

$$\begin{aligned}
&\Big(\frac{1}{2\sigma_c^2} - \frac{k_{j,s}}{(b_2-b_1)^2}\Big)u^2 + \Big(\frac{2k_{j,s}}{(b_2-b_1)^2} - \frac{m}{\sigma_c^2}\Big)u \\
&\qquad + \frac{m^2}{2\sigma_c^2} + k_{j,s}\Big(1 - \frac{b_1^2}{(b_2-b_1)^2}\Big) \\
&\qquad = a_{j,s}(u - \gamma_{j,s})^2 + \tau_{j,s}
\end{aligned}$$

where,

$$a_{j,s} \doteq \frac{1}{2\sigma_c^2} - \frac{k_{j,s}}{(b_2-b_1)^2}$$

$$\begin{aligned}
\gamma_{j,s} &\doteq \frac{m(b_2-b_1)^2 - 2k_{j,s}\sigma_c^2 b_1}{(b_2-b_1)^2 - 2k_{j,s}\sigma_c^2} \\
&= \frac{2N_0 m(b_2-b_1)^2 - \sigma_c^2 b_1 \parallel \alpha_j - \alpha_s \parallel^2}{2N_0(b_2-b_1)^2 - \sigma_c^2 \parallel \alpha_j - \alpha_s \parallel^2}
\end{aligned}$$

$$\begin{aligned}\tau_{j,s} &\doteq \frac{m^2}{2\sigma_c^2}+\frac{k_{j,s}b_2(b_2-2b_1)}{(b_2-b_1)^2} \\ &\quad -\frac{(2k_{j,s}\sigma_c^2 b_1 - m(b_2-b_1)^2)^2}{2\sigma_c^2(b_2-b_1)^2((b_2-b_1)^2-2k_{j,s}\sigma_c^2)} \\ &= \frac{m^2}{2\sigma_c^2}+\frac{\parallel \alpha_j-\alpha_s \parallel^2 b_2(b_2-2b_1)}{4N_0(b_2-b_1)^2} \\ &\quad -\frac{(\parallel \alpha_j-\alpha_s \parallel^2 -2N_0 m(b_2-b_1)^2)^2}{4N_0(b_2-b_1)^2[2N_0(b_2-b_1)^2- \parallel \alpha_j-\alpha_s \parallel^2 \sigma_c^2]}\end{aligned}$$

As a result the integral in region 2 becomes,

$$\Theta_2 = \frac{1}{\sqrt{2\pi\sigma_c^2}} e^{-\tau_{j,s}} \int_{b_1}^{b_2} e^{-a_{j,s}(u-\gamma_{j,s})^2} \mathrm{d}u \tag{D.2}$$

With the substitution $v_{j,s} = \sqrt{2a_{j,s}}(u-\gamma_{j,s})$ this becomes,

$$\begin{aligned}\Theta_2 &= \frac{1}{\sqrt{2a_{j,s}\sigma_c^2}} e^{-\tau_{j,s}} \frac{1}{\sqrt{2\pi}} \int_{v1_{j,s}}^{v2_{j,s}} e^{\frac{-v^2}{2}} dv \\ &= \frac{1}{\sqrt{2|a_{j,s}|\sigma_c^2}} e^{-\tau_{j,s}} (Q(v1_{j,s}) - Q(v2_{j,s}))\end{aligned} \tag{D.3}$$

where

$$v1_{j,s} = \sqrt{2a'_{j,s}}(b_1-\gamma_{j,s}),$$

$$\begin{aligned}v2_{j,s} &= \sqrt{2a'_{j,s}}(b_2-\gamma_{j,s}) \\ &= v1_{j,s} + \sqrt{2a'_{j,s}}(b_2-b_1),\end{aligned}$$

and,

$$a'_{s,j} = \max(a_{s,j}, 0). \tag{D.4}$$

The integral (D.2) is valid at all SNR values, however, the re-formulation of it in (D.3) is only valid for SNR values such that $a_{s,j}$¿0. Equation D.4 takes this in to account. Typically this

restriction has a negligible effect on the bound, lowering it slightly at high SNR's where the constrained capacity approaches $C_{max} = \log_2(M)$, where M is the order of the constellation.

Evaluation of the integral on region 3 ($u \in (b_2, \infty)$)

With the approximation that $e^{-2e^u/\kappa} = 0$ on the third region, the integral of (2.19) over this region becomes,

$$\Theta_3 = \frac{1}{\sqrt{2\pi\sigma_c^2}} \int_{b_2}^{\infty} e^{-\frac{(u-m)^2}{2\sigma_c^2}} \mathrm{d}u$$

Substituting $v = \frac{u-m}{\sigma_c}$ gives,

$$\begin{aligned} \Theta_3 &= \frac{1}{\sqrt{2\pi}} \int_{\frac{b_2-m}{\sigma_c}}^{\infty} e^{-\frac{v^2}{2}} dv \\ &= Q\left(\frac{b_2-m}{\sigma_c}\right) \end{aligned} \tag{D.5}$$

The overall integral

Using (D.1), (D.3) and (D.5) the above evaluations, we have,

$$\begin{aligned} & E_g[e^{-\frac{g^2\|\alpha_j-\alpha_s\|^2}{4N_0}}] \cong \Theta_1 + \Theta_2 + \Theta_3 \\ &= e^{-\frac{\|\alpha_j-\alpha_s\|^2}{4N_0}} Q\left(\frac{m-b_1}{\sigma_c}\right) \\ &\quad + e^{-\tau_{j,s}} \frac{1}{\sqrt{2\pi\sigma_c|a_{j,s}|}} \Big(Q(v_{1_{j,s}}) - Q(v_{2_{j,s}})\Big) \\ &\quad + Q\left(\frac{b_2-m}{\sigma_c}\right) \end{aligned} \tag{D.6}$$

Equation (D.6) was used in (2.15), resulting in the LB1 expression (2.21) for LLN channels.

REFERENCES

[ABS97] M. Alouini, S. Borsmiller, and P. Steffes. "Channel Characterization and Modeling for Ka-Band Very Small Aperture Terminals." *Proceedings of the IEEE*, **85**(6):981–997, June 1997.

[AG99] Mohamed-Slim Alouini and Andrea J. Goldsmith. "Area Spectral Efficiency of Cellular Mobile Radio Systems." *IEEE Trans. Veh. Technol.*, **48**(4):1047–1066, Jul. 1999.

[AS72] M. Abramowittz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover Press, 9th edition, 1972.

[AV08] Mehdi Adibi and Vahid Tabataba Vakili. "An Analytical Closed-form Expression for the Ergodic Capacity of Rayleigh Fading MIMO channels and Optimal Antenna Allocation." *2008 International Symposium on Telecommunications*, pp. 758–764, Aug. 2008.

[Bac99] Enzo Baccarelli. "Asymptotically Tight Bounds on the Capacity and Outage Probability for QAM Transmissions over Rayleigh-Faded Data Channels with CSI." *IEEE Trans. Commun.*, **47**:1273–1277, 1999.

[BF00] Enzo Baccarelli and Antonio Fasano. "Some Simple Bounds on the Symmetric Capacity and Outage Probability for QAM Wireless Channels Some Simple Bounds on the Symmetric Capacity and Outage Probability for QAM Wireless Channels With Rice and Nakagami Fading." *IEEE Journal on Selected Areas in Communications*, **18**:361–368, 2000.

[BJF07] M.F. Baroum, C. Jones, and M. Fitz. "Constellation Design via Capacity Maximization." *ISIT2007*, pp. 1821–1825, 2007.

[Bre59] D.G. Brennan. "Linear Diversity Combining Techniques." *Proc. IRE*, **47**(6):1075 –1102, Jun. 1959.

[CT91] T. Cover and J. Thomas. *Elements of Information Theory*. Wiley, 1991.

[DA05] L. R. D'Addario. "Estimates of Atmospheric-Induced Gain Loss for the Deep Space Network Array." *IPN Progress Report*, **42**(160):1–7, Feb. 2005.

[EF10] Scott Enserink and Michael P. Fitz. "Constrained Capacities of DVB-S2 Constellations in Log-Normal Channels at Ka Band." In *Advances in Satellite and Space Communications (SPACOMM), 2010, Second International Conference on*, pp. 93–99, June 2010.

[ET05] E.T.S.I. *Digital Vido Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other boradband satellite applications (ETSI EN 302 307)*. European Telecommunications Standards Institute, 2005.

[FV90] M. Filip and E. Vilar. "Optimum utilization of the channel capacity of a satellite link in the presence of amplitude scintillations and rain attenuation." *IEEE Trans. Commun.*, **38**:1958–1965, Nov. 1990.

[GF04] Boris Gremont and Miodrag Filip. "Spatio-Temporal Rain Attenuation Model For Application to Fade Mitigation Techniques." *IEEE Transactions on Antennas and Propaation*, **52**(5):1245–1256, May 2004.

[GFM06] Riccardo De Gaudenzi, Albert Guillen Fabregas, and Alfonso Martinez. "Performance Analysis of Turbo-Coded APSK Modulations over Nonlinear Satellite Channels." *IEEE Transactions on Wireless Communications*, **5**:2398–2407, September 2006.

[Goo05] Joseph W. Goodman. *Intro. to Fourier Optics*. Ben Roberts, 2005.

[GR65] I. S. Gradsheteyn and I. M. Ryzhik. *Table of Integrals and Products*. Academic Press, New York, 4th edition, 1965.

[HCH09] Fabien Héliot, Xiaoli Chu, Reza Hoshyar, and Ahim Tafazolli. "A Tight Closed-Form Approximation of the Log-Normal Fading Channel Capacity." In *IEEE Trans. Wireless Commun.* [LSA07], pp. 2842–2847.

[Kne83] D. Knepp. "Multiple Phase-Screen Calculation of the Temporal Behavior of Stochastic Waves." *Proceedings of the IEEE*, **71**(6):722–738, June 1983.

[Lee90] William C. Y. Lee. "Estimate of channel capacity in Rayleigh fading environment." *IEEE Trans. Veh. Technol.*, **39**(3):187–189, Aug. 1990.

[LH02] Debang Lao and A. M. Haimovich. "New error probability expressions for optimum combining with MPSK modulation." In *Vehicular Technology Conference, 2002, IEEE 56th*, pp. 2293 – 2297, Dec. 2002.

[LSA07] Amine Laourine, Alex Stephenne, and Sofiene Affes. "Estimating the Ergodic Capacity of Log-Normal Channels." *IEEE Commun. Lett.*, **11**(7):568–570, July 2007.

[LSA09] Amine Laourine, Alex Stephenne, and Sofiene Affes. "On the Capacity of Log-Normal Fading Channels." In *IEEE Trans. on Commun.* [LSA07], pp. 1603–1607.

[MA05] Amine Maaref and Sonia Aissa. "Closed-Form Expressions for the Outage and Ergodic Shannon Capacity of MIMO MRC Systems." *IEEE Trans. Commun.*, **53**(7):1092–1095, Jul. 2005.

[MWM07] Neelesh B. Mehta, Jingxian Wu, Andreas F. Molisch, and Jin Zhang. "Approximating a Sum of Random Variables with a Lognormal." *IEEE Trans. Wireless Commun.*, **6**(7):2690–2699, Jul. 2007.

[PLA09] A. D. Panagopoulos, K.P. Liolis, and P.-D.M Arapoglou. "An Analytical Unifying Approach for Outage Capacity Achieved in SIMO and MISO Broadband Satellite Channel Configurations." *EuCAP 2009, 3rd Conference on Antennas and Propogation (23-27 March 2009)*, pp. 2911–2915, March 2009.

[QC99] Xiaoxin Qiu and Kapil Chawla. "On the Performance of Adaptive Modulation in Cellular Systems." *IEEE Trans. Commun.*, **47**(6):884–895, Jun. 1999.

[Sas94] Richard J. Sasiela. *Electromagnetic Wave Propagation in Turbulence: Evaluations and Application of Mellin Transforms*. Springer Series on Wave Phenomena. Springer-Verlag, 1994.

[SH98] Amit Shah and Alexander M. Haimovich. "Performance Analysis of Optimum Combining in Wireless Communications With Rayleigh Fading and Cochannel Interference." *IEEE Trans. Commun.*, **46**(4):473–479, Apr. 1998.

[SH00] A. Shah and A. M. Haimovich. "Performance Analysis of Maximal Ratio Combining and Comparison with Optimum Combining for Mobile Radio Communications with Cochannel Interference." *IEEE Trans. Veh. Technol.*, **49**(4):1454–1463, Jul. 2000.

[Sli06] Ben Slimane. "Approximation to the Symmetric Capacity of Rayleigh Fading Channels with Multi-Level Signals." In *IEEE Commun. Lett.* [Won01], pp. 129–131.

[Tat61] V.I. Tatarski. *Wave Propagation In A Turbulent Medium*. Mc-Graw Hill, 1961.

[Tel99] I. E. Telatar. "Capacity of Multi-Antenna Gaussian Channels." *Eur. Trans. Telecomm.*, **10**(6):585–595, Nov 1999.

[Tom74] Ivan Tomek. "Two Algorithms for Piecewise-Linear Continuous Approximation of Functions of One Variable." *IEEE Trans. Comput.*, **C-23**(4):445–448, Apr. 1974.

[Tys00] R. K. Tyson. *Introduction to Adaptive Optics*. SPIE Press, 2000.

[VGW05] D. Vanhoenacker-Janvier, C. Gibbins, C. Walder, S. Ventouras, J. Spiegel, C. Oestgges, D. Mertens, and A. Martellucci. "Characterization and Modelling of Propagation Effects in the 20-50 GHz Band." *Proceedings of the XVIIIth General Assemably of International Union of Radio Science (URSI),*, 2005.

[VV98] H. Vasseur and D. Vanhoenacker. "Characterisation of Tropospheric Turbulent Layers From Radiosonde Data." *Electronics Letters*, **34**(4):318–319, Feb. 1998.

[Whe01] Albert D. Wheelon. *Electromagnetic Scintillation: I. Geometrical Optics*. Cambridge University Press, 2001.

[Whe03] Albert D. Wheelon. *Electromagnetic Scintillation: II. Weak Scattering*. Cambridge University Press, 2003.

[Wil96] S. G. Wilson. *Digital Modulation and Coding*. Prentice-Hall, 1996.

[Won01] Tan F. Wong. "Numerical Calculation of Symmetric Capacity of Rayleigh Fading Channel with BPSK/QPSK." *IEEE Commun. Lett.*, **5**(8):328–330, Aug. 2001.

[YWL03] R. Yang, Z. Wu, and Y. Li. "Analysis of tropospheric scintillation due to clear-air and meteorological elements on slant microwave links." *Proc. of SPIE*, **4899**:165–171, 2003.