Arun K. Majumdar
Jennifer C. Ricklin

OPTICAL AND FIBER COMMUNICATIONS REPORTS 2

Free-Space Laser Communications

Principles and Advances

Springer

OPTICAL AND FIBER COMMUNICATIONS REPORTS

OPTICAL AND FIBER COMMUNICATIONS REPORTS

The Optical and Fiber Communications Reports (OFCR) book series provides a survey of selected topics at the forefront of research. Each book is a topical collection of contributions from leading research scientists that gives an up-to-date and broad-spectrum overview of various subjects.
The main topics in this expanding field will cover for example:

- specialty fibers (periodic fibers, holey fibers, erbium-doped fibers)
- broadband lasers
- optical switching (MEMS or others)
- polarization and chromatic mode dispersion and compensation
- long-haul transmission
- optical networks (LAN, MAN, WAN)
- protection and restoration
- further topics of contemporary interest.

Including both general information and a highly technical presentation of the results, this series satisfies the needs of experts as well as graduates and researchers starting in the field. Books in this series establish themselves as comprehensive guides and reference texts following the impressive evolution of this area of science and technology.
The editors encourage prospective authors to correspond with them in advance of submitting a manuscript. Submission of manuscripts should be made to one of the editors. See also http://springeronline.com/series/4810.

Arun K. Majumdar Jennifer C. Ricklin

Free-Space Laser Communications

Principles and Advances

Springer

Arun K. Majumdar
LCResearch, Inc.
30402 Rainbow View Drive
Agoura Hills, CA 91301
a.majumdar@IEEE.org

Jennifer C. Ricklin
AMSRL-CI-CN
2800 Powder Mill Rd.
Adelphi, MD 20783
jricklin@darpa.mil

ISBN-13: 978-0-387-28652-5 e-ISBN-13: 978-0-387-28677-8

Library of Congress Control Number: 2006939943

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

Preface

Free-space laser communications, also referred to as optical communications, is a popular subject in today's technological marketplace. A number of conferences on this subject have been organized by professional societies such as SPIE (the International Society of Photo Optical and Instrumentation Engineering), OSA (Optical Society of America), and IEEE (Institute of Electrical and Electronics Engineers). The evolving technology of free-space laser communications is emerging as an appealing alternative to RF communications for links between satellites, as well as a promising addition to terrestrial applications such as video or computer linkups between buildings. There is a pressing need for more information on laser communications that is comprehensive enough to provide in-depth knowledge of free-space communications, and that can satisfy the current demands of the research and commercial needs.

This book has been designed to provide a comprehensive, unified tutorial to further understanding of the fundamental techniques for laser communications through the earth's atmosphere. The driving force behind free-space laser communications is the continuous demand for higher bandwidth to deliver high-capacity voice, data, and images to the customer. Free-space propagation distances include ranges that encompass a few millimeters (for example between optical interconnects in a computer using photonics to replace metal interconnects), a few meters (such as indoor communications), a few kilometers (between buildings, campuses, and hospitals), and even up to thousands of kilometers (such as from an aircraft or satellite to the ground). Although the concept of free-space laser communications systems has been around as long as the laser itself, it has not been as widely implemented in fielded systems as one might expect. There have been several reasons for this, the principal one being the lack of a demonstratable need for extremely high-bandwidth wireless systems, and the cost, complexity, and reliability of the technology involved. There have also been complications with several of the attempts to field large-scale operational systems. But with the unprecedented growth of the internet, multimedia, and related data services and image transfer, there has been a tremendously growing appetite for immediately available information at high bandwidth. This has given free-space laser communications a unique niche in the communications market.

The need for bandwidth for both commercial and military applications is growing faster than the ability of the local infrastructure to provide them, and also faster than the rate at which the infrastructure can be augmented. Technology development in the areas of higher modulation capability of laser sources, fast and sensitive detectors, and optical components such as DWDM (Dense Wavelength Division Multiplexing) and optical amplifiers, all have helped to achieve higher bandwidth. However, at optical wavelengths, the interaction of electromagnetic waves with the atmosphere is much stronger than that at microwave frequencies, and hence weather conditions such as turbulence and fog limit the performance of optical links.

This book provides a comprehensive overview for readers who require information about the basics of free-space laser communications, as well as up-to-date advanced knowledge of the state-of-the-art technologies available today. We have tried to make this book as useful as possible as a resource for engineers, scientists, researchers, and students interested in understanding and designing laser communication systems for atmospheric channels. The reader will further benefit from the knowledge of how adaptive optics and related technologies such as MEMS (Micro-Electro-Mechanical Systems) can be integrated into free-space laser communications. There is continuity between chapters so that the information and learning from any one chapter is connected to that of the other chapters.

Chapter 1 is an introduction to atmospheric laser communication and discusses the overall laser communications architecture. This chapter serves as a general background for the concepts presented in subsequent chapters.

Chapter 2 describes the impact that atmospheric effects have on laser communication systems. This chapter describes the physical processes underlying optical extinction and optical turbulence-induced signal losses. Recent advances in the modeling of optical scintillation under weak-to-strong fluctuations associated with both terrestrial links and satellite/ground links have been addressed. This and other information given in this chapter enable calculation of the anticipated system performance over a wide range of channel characteristics for ground-to-ground, ground-to-air/air-to-ground, and satellite uplink/downlink optical communication scenarios.

Chapter 3 provides a theoretical treatment and analysis of the performance of a free-space laser communication system under a given communication format scheme. This chapter develops a communication link budget model relating the atmospheric channel effects and the main parameters of the transmitter and receiver stations. Numerous examples are given for specific system parameters in order to design a specific system. Various scenarios of practical importance are described: horizontal (terrestrial) links, and slant paths including downlink (satellite-to-ground) and uplink (ground-to-satellite). This system performance analysis will be useful for designers in optimizing the optical station main parameters in order to establish a data link with adequate performance under various atmospheric conditions.

Chapter 4 provides an overview of state-of-the-art, as well as future optical transmitter and receiver, designs that are particularly well suited for free-space laser communications. Emphasis is placed on high-performance photon-efficient systems and technologies suitable for operation in both deep space and terrestrial optical links. This chapter focuses on practical design options for the transmitter and receiver blocks. Practical tradeoffs and implementation issues that arise from using various technologies are presented, and the interplay between modulation format, transmitter waveform, and receiver design are discussed.

Wave-front distortions of optical waves can be mitigated, in principle, with adaptive optics. *Chapter 5* discusses an adaptive optics approach that has significant advantages for free-space laser communication terminals. This technique does not require wave-front measurements, which are difficult to obtain under the strong scintillation conditions typical for many communication scenarios. It is based on the direct optimization of a performance quality metric, e.g., the communication signal strength. An experimental adaptive optics system is also described.

Chapter 6 provides a broadband optical wireless solution for closing the "last mile" connectivity gap throughout metropolitan networks. This chapter discusses various network architectures and the usage of "optical repeaters." Point-to-point and point-to-multipoint solutions are explained for constructing different network architectures. Terrestrial applications for short and long ranges, as well as space applications, are described. Important design criteria for connecting the user to the "backbone" by free space optics techniques are explained.

Error-control codes can help to mitigate turbulence-induced signal fading in laser communications through atmospheric turbulence. *Chapter 7* provides an in-depth treatment of this critical subject through the use of techniques based on the statistical properties of turbulence-induced signal intensity fading as a function of both temporal and spatial coordinates. An optimal maximum-likelihood detection scheme is described for correlated spatial-diversity reception. This chapter also describes an experimental demonstration of the effectiveness of two temporal domain techniques in a terrestrial link using on–off keying.

The three remaining chapters cover the recent development of longer-wavelength free-space laser communications. A longer wavelength source can provide significantly improved transmission through the atmosphere, as compared to shorter wavelengths. *Chapter 8* discusses optical parametric oscillation devices used for mid-IR (MWIR) data communication through low visibility conditions. This chapter also describes the design of a MWIR transceiver that provides high peak power and short pulses for the Multiple Integrated Laser Engagement System (MILES) program.

Chapter 9 discusses the high-speed modulation properties of a quantum cascade laser (QCL) suitable for free-space laser communications. A comparison of experiments between a QCL and a near-infrared laser-based link

in controlled low-visibility experiments is discussed. A recent application of a QCL-based link used for transmission of satellite TV data streams is also presented in this chapter.

Chapter 10 discusses technology for all-weather long-wavelength infrared (LWIR) free-space optical communications. Superior transmission through atmospheric phenomena such as fog, clouds, and smoke, coupled with improvements in LWIR laser and modulator design, makes possible reliable optical replacements for radio and microwave communication links for many applications. This chapter summarizes the components and techniques developed for high-speed, full-duplex all-weather infrared communications systems.

We express our sincere appreciation to Dr. Hans J. Koelsch, the Springer Editorial Director for Physics, for his encouragement and interest from the very start of this project. We also thank all of the contributors, as well as Ms. Virginia Lipscy, Associate Editor for Physics, and Ms. Vasudha Gandhi, Editorial Assistant, for their continuous help throughout the preparation of this book. A.K. Majumdar wishes to thank his daughter Sharmistha for her help in checking the texts of his sections, his wife Gargi for her encouragement and support, and also his feline friends, Rocky and Sasha, who provided unconditional companionship for the entire duration of his work in completing this project. He also wishes to pay tribute to the memory of his brother, Dr. Nirmal Majumdar of London, who was a constant source of inspiration throughout his writing of this book. J.C. Ricklin thanks her husband Mikhail for his support and encouragement, as well as the gray family cat who kept the monitor's mouse icon under continuous strong observation and control.

Arun K. Majumdar
Agoura Hills, California
Jennifer C. Ricklin
Laurel, Maryland

Contributors

Dr. Ulla Birnbacher
Department of Communications and Wave Propagation
Graz Technical University
A-8010 Graz, Inffeldgasse 12, Austria
birnbacher@inw.tu-graz.ac.at
43-316-873-7442

Dr. David Caplan
MIT Lincoln Laboratory
Optical Communications Technology
244 Wood St., Mail Stop C-245, Lexington, MA 02420
doc@ll.mit.edu
(781) 981-5025; Cell: (978) 852-3888

Dr. Frank D. Eaton
AFRL/DESA, Air Force Research Laboratory
3550 Aberdeen Ave. SE, Kirtland AFB, NM 87117
frank.eaton@kirtland.af.mil
(505) 853-1091

Dr. Michael Gebhart
Department of Communications and Wave Propagation
Graz Technical University
A-8010 Graz, Inffeldgasse 12, Austria
gebhart@inw.tu-graz.ac.at
43-316-873-7442

Dr. Stephen Hammel
Atmospheric Propagation Branch
2558 Space and Naval Warfare Systems Center
San Diego, CA 92152
hammel@spawar.navy.mil
(619) 553-4578

Dr. Donald P. Hutchinson
Engineering Science and Technology Division
Oak Ridge National Laboratory
Bethel Valley Road, Oak Ridge, TN 37831
hutchinsondp@ornl.gov
(865) 574-4730; Cell: (865) 382-8155

Prof. Joseph M. Kahn
Department of Electrical Engineering
Stanford University
372 Packard Building, 350 Serra Mall, Stanford, CA 94305-9515
jmk@ee.stanford.edu
(650) 724-9584

Dr. Svetlana L. Lachinova
Institute for Systems Research
University of Maryland
College Park, MD 20742
sllachin@mail.umd.edu

Dr. Erich Leitgeb
Department of Communications and Wave Propagation
Graz Technical University
A-8010 Graz, Inffeldgasse 12, Austria
leitgeb@inw.tu-graz.ac.at
43-316-873-7442

Dr. Arun K. Majumdar
LCResearch, Inc.
30402 Rainbow View Drive, Agoura Hills, CA 91301
a.majumdar@IEEE.org
(818) 706-8788; Cell: (818) 620-7918

Prof. Rainer Martini
Department of Physics and Engineering Physics
Stevens Institute of Technology
Hoboken, NJ 07030
rmartini@stevens-tech.edu
(201) 216-5634

Dr. Narasimha S. Prasad
NASA Langley Research Center, 468/LEOB
5 North Dryden Street, B1202 Hampten, VA 23681-2199
n.s.prasad@larc.nasa.gov
(757) 864-9403

Dr. Roger K. Richards
Engineering Science and Technology Division
Oak Ridge National Laboratory
Bethel Valley Road, Oak Ridge, TN 37831
richardsrk@ornl.gov

Dr. Jennifer C. Ricklin
Defense Advanced Research Projects Agency (DARPA), Advanced Technology Office
3701 N. Fairfax Drive, Arlington, VA 22203
Jennifer.Ricklin@darpa.mil
(703) 526-4751

Dr. Mikhail A. Vorontsov
Army Research Laboratory
AMSRL-CI-S
2800 Powder Mill Road, Adelphi, MD 20783-1197
mvorontsov@arl.army.mil
(301) 394-0214

Dr. Thomas Weyrauch
Intelligent Optics Laboratory
Institute for Systems Research
University of Maryland
A.V. Williams Bldg., Mail Stop 1103, College Park, MD 20742
weyrauch@umd.edu

Prof. Edward A. Whittaker
Department of Physics and Engineering Physics
Stevens Institute of Technology
Hoboken, NJ 07030
ewhittak@stevens-tech.edu

Dr. Xiaoming Zhu
QUALCOMM, Inc.
675 Campbell Technology Parkway
Suite 200 Campbell, CA 95008
xzhu@qualcomm.com
(408) 423-0346

Contents

Preface v
Arun K. Majumdar and Jennifer C. Ricklin

Contributors ix

Introduction 1
Arun K. Majumdar

Atmospheric channel effects on free-space laser communication 9
Jennifer C. Ricklin, Stephen M. Hammel, Frank D. Eaton and Svetlana L. Lachinova

Free-space laser communication performance in the atmospheric channel 57
Arun K. Majumdar

Laser communication transmitter and receiver design 109
David O. Caplan

Free-space laser communications with adaptive optics:
Atmospheric compensation experiments 247
Thomas Weyrauch and Mikhail A. Vorontsov

Optical networks, last mile access and applications 273
E. Leitgeb, M. Gebhart, and U. Birnbacher

Communication techniques and coding for atmospheric turbulence channels 303
Xiaoming Zhu and Joseph M. Kahn

Optical communications in the mid-wave IR spectral band 347
Narasimha S. Prasad

Quantum cascade laser-based free space optical communications 393
R. Martini and E.A. Whittaker

All-weather long-wavelength infrared free space optical communications 407
D.P. Hutchinson and R.K. Richards

Introduction

Arun K. Majumdar

LCResearch, Inc.
30402 Rainbow View Drive
Agoura Hills, California 91301
Email: a.majumdar@IEEE.org

Abstract. This book covers various topics relevant to free-space laser communications (FSO lasercom). The purpose of this initial chapter is to provide the technical background needed to understand and appreciate the following chapters, as well as to establish the connections and continuity among them. This chapter will present the most fundamental physical equations and parameters needed to characterize and analyze free-space laser communications.

1. Introduction

Recent rapid progress in information and communication technologies has exceeded our expectations for meeting the requirements of multimedia society in the 21st century. Free-space optical communication is considered to be one of the key technologies for realizing very-high-speed multi-gigabit-per-second (multi-Gb/s) large-capacity aerospace communications. Using lasers as signal carriers, FSO lasercom can provide a line-of-sight, wireless, high-bandwidth, communication link between remote sites. Rapidly growing use of the Internet and multimedia services has created congestion in the telecommunications networks and placed many new requirements on carriers. Laser transmitters offer an intermediate, low-risk means to introduce desired network functionalities with extremely high bandwidth. The wireless aspect of FSO lasercom can be a crucial advantage, particularly in local area networks (LANs) and metropolitan area networks (MANs) where in cities the laying of fiber optic cables is expensive. When the World Trade Centers collapsed on September 11th, 2001, so too did the networks of all of the corporations located within the buildings. Hundreds of thousands of dollars were lost due to the down-time that these companies experienced. With the help of a few wireless broadband providers using FSO technology, these

companies were able to get their data networks up and running in a fraction of the time and considerably less money that it would have cost them to reinstall their wire lines, be it a fiber or cable network.

FSO lasercom offers substantial advantages over conventional RF wireless communications technology, including higher data rates, low probability of intercept, low power requirements, and much smaller packaging. FSO lasercom systems have proven to be a viable alternative to optical fiber based systems in several applications, as the technology comes closer and closer to providing the 5-nines (99.999%) service that many corporations require of their data networks.

1.1. Brief History

Communications using light is not a recent science. Old Roman records indicate that polished metal plates were occasionally used as mirrors to reflect sunlight for long-range signaling. The U.S. military used similar sunlight-powered devices to send telegraph information from mountain top to mountain top in the early 1800's. Blinking lights have been used to send messages from one ship to another for many centuries. Alexander Graham Bell performed an experiment with his "Photophone" where he used sunlight reflected off of a vibrating mirror and a selenium photo cell to send telephone-like signals over a range of 600 feet. A complete history of optical and laser communications can be found in the literature [1–2].

In 1895, Sir J.C. Bose of Calcutta, India gave his first public demonstration of radio communication, using electromagnetic waves to ring a bell remotely and to explode some gunpowder [3–5]. As reported by the Daily Chronicle of England: "The inventor (J.C. Bose) has transmitted signals to a distance of nearly a mile and herein lies the first and obvious and exceedingly valuable application of this new theoretical marvel." Invited by Lord Rayleigh, Bose reported on his microwave/millimeter-wave experiments to the Royal Institution and other societies in England [6]. The first successful wireless signaling experiment by Marconi was not until 1897. Thus, the inventor of Radio Communication was not Marconi but Sir J.C. Bose. The author appreciates this opportunity to acknowledge a fellow scientist from Calcutta for his pioneering work leading to the development of modern wireless laser communications. Interested readers can read about Bose's contributions further in [7].

During the years 1930–1950, increased carrier frequencies were considered to achieve greater information capacity/bandwidth: Microwaves of 3–30 GHz, and Millimeter Waves of 30–300 GHz.

Following the first demonstration of a working laser in 1960 at the Hughes Research Laboratories in Malibu, California, the first serious development of FSO lasercom was called the "Lightphone". NASA and the Air Force were especially interested and forced some new technological developments. During the boom period of optical fiber installation, civil FSO lasercom technology lay dormant; but in military and space laboratories the development never really stopped.

During the last thirty years, great advances have been made in electro-optics and opto-electronics, and incorporated into today's laser communications systems, mostly for defense applications. The aerospace/defense activity established a strong foundation upon which today's commercial optical wireless systems are based. Systems have been developed for ground-to-ground, ground-to-aircraft, ground-to-satellite, satellite-to-satellite, and even satellite-to-submarine applications.

1.2. Applications

Some typical scenarios where FSO lasercom can be used are

- LAN-LAN connections on campuses at Fast-Ethernet or Gigabit-Ethernet Speeds
- LAN-to-LAN connection in a city
- Speedy service delivery of high-bandwidth access to fiber networks
- Temporary network installation (for special events or other purposes)
- Re-establishing high-speed connection quickly (Disaster Recovery/Emergency Response)
- Communications between ground and spacecraft, or between spacecrafts, including elements of a satellite constellation
- Interstellar communication
- Ship-to-ship communications with high data rates providing complete security

Figure 1 illustrates free-space laser communications architectures for ground-to-ground, ground-to-space, space-to-ground and space-to-space links. These links may include, for example, a metropolitan network using air-to-air laser communication between office buildings, a long-distance optical link between an optical ground station (with adaptive optics) and a building or satellite platform, or an inter-satellite link. Another area of interest is the broadband wireless solution for closing the so-called last mile connectivity gap throughout metropolitan networks. Different network architectures and the usage of "Optical Repeaters", Point-to-Point, and Point-to-Multipoint solutions are explained in the sixth chapter ["Optical networks, last mile access and applications", by E. Leitgeb, M. Gebhart and U. Birnbacher (DOI 10.1007/s10297-004-0025-x)].

1.3. Advantages and Challenges

- High data rates
- High transmission security
- No Federal Communications Commission (FCC) license or frequency allocation required
- Light weight, small volume, and lower power consumption, providing a potential edge over RF communication
- Portability and quick deployment
- Increased security due to the laser's narrow beam—ideal for the wireless transfer of financial, legal, military, and other sensitive information

1.4. Limitations

For many cases of practical interest, the constant presence of optical turbulence in the atmospheric channel is the limiting factor in reliable free-space laser communication performance [8]. Other atmospheric effects that can degrade performance include

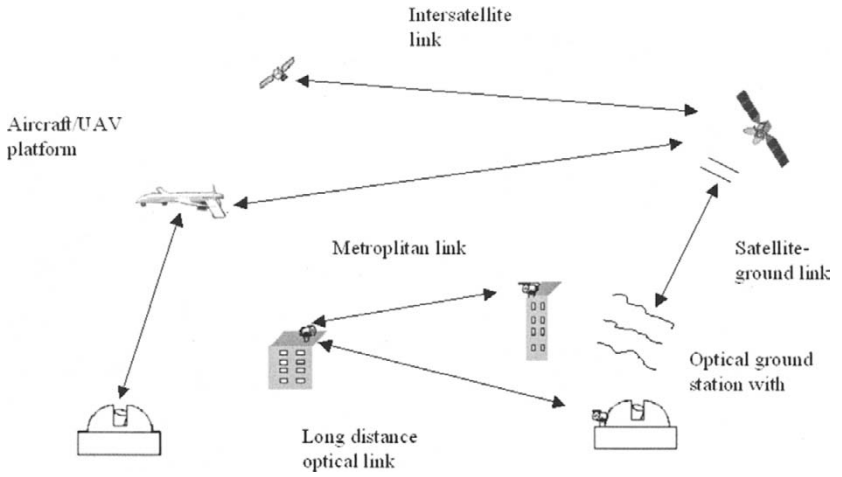


Fig. 1. Free-Space Laser Communications Architectures for ground-to-ground, ground-to-space, space-to-ground, and space-to-space terminal links.

absorption and scattering of the laser beam. Another limitation of light beam communications is the requirement of a clear line-of-sight path between the laser transmitter and the receiver, since light cannot penetrate trees, hills, buildings, etc. A further limitation, sometimes overlooked, applies to the position of the sun relative to the laser transmitter and receiver, to avoid having the receiver look into the sun.

1.5. Basics of Operation

The object of the communication system is to transfer data or information from a transmitter at one point to a receiver at another through the intervening atmospheric channel with an acceptable error rate while providing high reliability. The transmitter consists basically of a laser, a modulator (with a selected modulation scheme), a laser driver (to provide necessary power supply), and a telescope. Depending on the modulation format, the modulator converts bits of information into an electrical signal and modulates the laser to generate an optical signal. The telescope expands the optical beam to reduce diffractive effects that spread the beam out.

The optical signal propagates through the atmospheric channel and is collected by the receiver to recover the data. The receiver consists of a telescope, a detector, and a decoder to get the signal out. The telescope collects the optical signal and contracts the beam size small enough to fit onto the detector. The detector converts the optical signal back to an electrical signal. Generally for a direct detection system, the receiver consists of an avalanche photodiode (APD) photodetector, a matched filter implemented in the form of a moving integrator, and a clocked comparator. The clocked comparator performs a threshold test at bit interval boundaries.

2. Understanding Free-Space Laser Communications Systems Performance

The chapters of this book describe and explain different aspects of lasercom; but all address one common question: How can we use optical wavelengths to accomplish high data rate performance under different scenarios and atmospheric conditions?

2.1. Bit Error Rate

One standard measure for laser communication system performance is the available post-detection signal-to-noise ratio (SNR). While this is most relevant to describe analog systems, another measure is more meaningful for digital communications. When we send information by digital transmission, the desired message is converted to binary symbols (bits). The transmission of digital bits over the optical link is then done on a bit-by-bit basis (binary encoding). Note that optical bit modulation can be accomplished using either a pulsed laser beam or a continuous-wave (CW) laser. In either case, the real performance measure in digital communications is provided not directly by the SNR, but rather by the probability of error, also referred to as bit error rate (BER). We can define the BER as the probability that an error may occur in a bit in the pulse train, i.e., a "1" bit turns into a "0" bit or vice versa. For commercial telecommunications applications BER = 10^{-9} is a typical required value.

In general, the BER is given by

$$\mathrm{BER} = P_0/2 + P_1/2, \tag{1}$$

where P_0 is the probability of mistaken "0" for "1", and P_1 is the probability of mistaken "1" for "0". The 1/2 multiplicative factor comes from the fact that, in a general digital communication system, "0" and "1" are equally likely.

2.2. Fundamental Limit of Light Detection

One of the principal limitations in optical communications is the power sensitivity of the receivers. As detection devices based on semiconductor materials are technologically improved, the minimum amount of light power required in the detection process is reduced. A natural question is, what is the minimum amount of photons we can have so that the detector detects a "1" bit with acceptable precision? Even if we could grow absolutely perfect semiconductor materials and could suppress any possible noise in the detection circuit, there is still a fundamental limitation which is entirely independent of the detection device or system: the quantum fluctuation of light. This manifests itself as random or pseudorandom changes in the number of photons present in the optical signal, yielding intensity fluctuations which can produce a bit error. The quantum fluctuations, as do any other noise source present in the system, set a lower limit on the BER of the overall communication link.

We can use Eq. (1) to estimate the BER due to quantum fluctuation. The photon occupancy number $P(n)$ gives the probability of detecting n photons. For low optical signal levels, we can assume that the photon occupancy number distributes with Poisson statistics expressed as

$$P(n) = N_p^n \exp(-N_p)/n!, \tag{2}$$

where N_p is the average number of photons sent in the signal for a single bit of data. In the case of "0", $N_p = 0$, and in the case of "1", $N_p = 1$ or more. The Poisson distribution assumption is reasonable because the light being detected is being produced by a stimulated emission which distributes as a coherent field state. The probability of detecting zero photons provided that an average of N_p photons have been sent is therefore:

$$P(n = 0) = \exp(-N_p). \tag{3}$$

Note that $P(n > 1) = 0$ for the case of $N_p = 0$, since a total lack of photons can never give anything other than a "0". We are considering the best possible detection systems where thermal fluctuations have been neglected. Solving for N_p from Eqs. (3) and (1), and using a typical required BER, yields

$$N_p = -\ln(\text{BER}) = -\ln(10^{-9}) \approx 20 \text{ photons}. \tag{4}$$

This is the minimum number of photons that have to reach the detector to decide "1" with an acceptable BER of 10^{-9}. This is a fundamental limitation that comes directly from the statistical nature of the radiation and sets a limit to any improvement in semiconductor detectors.

2.3. *BER for a Random Stochastic Communication Channel*

When the communication channel is a turbulent atmosphere with random variations (both spatial and temporal) of refractive index, system performance is degraded by turbulence-induced beam spreading, angle-of-arrival fluctuations, and scintillation. In the presence of these temporal fluctuations, the BER is considered a conditional probability that must be averaged over the probability density function (PDF) of the random signals to determine the unconditional BER. For example, in the case of on-off keying (OOK) for pulsed modulation in binary direct detection receiver, the BER is given by

$$\text{BER} = \frac{1}{2}\int_0^\infty p_I(s)\text{erfc}\left(\frac{\langle \text{SNR}\rangle s}{2\sqrt{2}\langle i_s\rangle}\right) ds, \tag{5}$$

where $\langle \text{SNR}\rangle$ is the mean SNR in the presence of atmospheric turbulence, $\langle i_s\rangle$ is the mean output signal current , and $p_I(s)$ is the PDF of the received fluctuating signal at the receiver that depends on the level of turbulence strength and the propagation path. In the above equation, $\text{erfc}(x)$ is the complementary error function. Often-realistic theoretical models for the PDF proposed by various researchers include Log-Normal and Gamma-Gamma distributions. Recently, a method of reconstructing an unknown PDF of intensity fluctuations directly from the measured higher-order moments has been reported [9], eliminating the need to assume applicability of theoretical PDF models. The mean $\langle \text{BER}\rangle$ for a turbulent atmospheric channel is of course always larger than that of a no-turbulence channel:

$$\langle \text{BER}\rangle_{\text{Turbulence}} > \langle \text{BER}\rangle_{\text{NoTurbulence}}. \tag{6}$$

We describe below some of the techniques (described further in the book) which have been used to mitigate the atmospheric turbulence effects and hence to improve BER, that is, to decrease BER.

2.4. *Wavelength Selection Criteria*

The choice of the transmitting laser wavelength will depend upon the atmospheric propagation characteristics, optical background noise, and the technologies developed for lasers, detectors, and spectral filters. For a long atmospheric channel, the wavelength will generally need to be restricted to spectral regions of very low atmospheric absorption. Another consideration is intensity fluctuation (scintillation) due to turbulence. The strength of intensity fluctuations decreases as $\lambda^{-7/6}$; thus, scintillation and hence BER can be decreased by using a longer wavelength. Free-space laser communications using longer IR wavelengths are described in the eighth ["Optical communications in the mid-wave IR spectral band", by Narasimha S. Prasad (DOI 10.1007/s10297-005-0057-x)], nineth ["Quantum cascade laser-based free space optical communications", by R. Martini and E.A. Whittaker (DOI 10.1007/s10297-005-0052-2)], and tenth ["All-weather long-wavelength infrared free space optical communications", by D.P. Hutchinson and R.K. Richards (DOI 10.1007/s10297-006-0078-0)] chapters as follows: 3–5 μm range using a solid-state, compact, optical parametric oscillator device; 8 μm using a Quantum Cascade Laser (QCL); and around 10.6 μm using a compact, RF-driven waveguide CO_2 laser.

2.5. *Adaptive Optics for Free-Space Laser Communications*

Refractive-index inhomogeneities of the turbulent atmosphere cause wave-front distortions of optical wave and are responsible for severe scintillation which can lead to power losses at the receiver and eventually to fading of the received signal below a prescribed threshold. In principle, wavefront distortions can be mitigated with adaptive optics, thereby improving BER. The fifth chapter ["Free-space laser communications with adaptive optics: Atmospheric compensation experiments", by Thomas Weyrauch and Mikhail A. Vorontsov (DOI 10.1007/s10297-005-0033-5)] discusses a non-conventional adaptive optics approach that has certain advantages with respect to its incorporation into the free-space optical communication channel. This technique is based on direct optimization of a performance quality metric, e.g., the communication signal strength, with a stochastic parallel gradient descent (SPGD) algorithm. Thus, BER can be improved for a laser communication system with adaptive optics.

2.6. *Coding for Atmospheric Channel*

Error control coding can be used over free-space laser communication links to mitigate turbulence-induced fading. The techniques are generally based on the statistical properties of turbulence-induced signal intensity fading, as functions of both temporal and spatial coordinates. The seventh chapter ["Communication techniques and coding for atmospheric turbulence channels", by Xiaoming Zhu and Joseph M. Kahn (DOI 10.1007/s10297-006-0077-1)] discusses in detail the error-probability bound for coded OOK free-space optical communication through atmospheric turbulence channels. This is also a very effective way to improve BER through a random channel.

3. Summary

In this chapter, a general introductory discussion is outlined to help the readers understand the overall concepts and the objectives for free-space laser communications. This background material will be useful in making smooth connections among all of the chapters in this book.

Acknowledgement

The author would like to thank Dr. Bradley Bobbs for reviewing this chapter and his useful comments.

References

1. http://www.eee.strath.ac.uk/ug-info/19984/comoh1.pdf
2. http://www.ballaerospace.com/lasercomm_history_1950.html
3. "The Work of Jagadis Chandra Bose: 100 years of mm-wave Research," by D.T. Emerson of the National Radio Astronomy Observatory, Tucson, Arizona. It is based on material presented at the IEEE-MTT-S International Microwave Symposium in Denver, Colorado, June 8–13, 1997.
4. IEEE-MTT-S International Microwave Symposium Digest, Volume 2, ISSN 0149-645X, pp. 553–556, 1997.
5. D.T. Emerson, IEEE Trans. Microwave Theory Tech. **45** (12), 2267–2273 (December 1997).
6. *J.C. Bose, Collected Physical Papers* (New York: Longmans, Green and Co., 1927).
7. Probir K. Bondyopadhya, "Sir J.C. Bose's Diode Detector Received Marconi's First Transatlantic Wireless Signal of December 1901," (The "Italian Navy Cohere" Scandal Revisited), Proc. IEEE. **86**, 1 (January 1998).
8. Arun K. Majumdar and Jennifer C. Ricklin, "Effects of the atmospheric channel on Free-space laser Communications," Proc. SPIE, Free-Space Laser Communication V, Vol. 5892, 2005.
9. Arun K. Majumdar, Carlos E. Luna, and Paul S. Idell, "Reconstruction of Probability Density Function of Intensity Fluctuations Relevant to Free-Space Laser Communication through Atmospheric Turbulence", Proc. SPIE, Free- Space Laser Communications VI, Vol. 6304, 2006.

DOI: 10.1007/s10297-005-0056-y
Originally published in J. Opt. Fiber. Commun. Rep. 3, 111–158 (2006)

Atmospheric channel effects on free-space laser communication

Jennifer C. Ricklin[1], Stephen M. Hammel[2], Frank D. Eaton[3], and Svetlana L. Lachinova[4]

[1] Defense Advanced Research Projects Agency, Advanced Technology Office
3701 N. Fairfax Drive, Arlington, VA 22203
Email: Jennifer.Ricklin@darpa.mil
[2] Atmospheric Propagation Branch, 2858
Space and Naval Warfare Systems Center, San Diego, CA 92152
Email: stephen.doss-hammel@navy.mil
[3] AFRL/DESA, Air Force Research Laboratory
3550 Aberdeen Ave. SE, Kirtland AFB, NM 87117
Email: frank.eaton@kirtland.af.mil
[4] Intelligent Optics Laboratory, University of Maryland
2107 Technology Ventures Building, College Park, MD 20742
Email: sllachin@mail.umd.edu

Abstract. Free-space laser communication offers an attractive alternative for transferring high-bandwidth data when fiber optic cable is neither practical nor feasible. However, there are a variety of deleterious features of the atmospheric channel that may lead to serious signal fading, and even the complete loss of signal altogether. Physical obstructions—such as birds, insects, tree limbs, or other factors—can temporarily or permanently block the laser line-of-sight. Platform/building motion due to wind, differential heating and cooling, or ground motion over time can result in serious misalignment of fixed-position laser communication systems. But most importantly of all, absorption and scattering due to particulate matter in the atmosphere may significantly decrease the transmitted optical signal, while random atmospheric distortions due to optical turbulence can severely degrade the wave-front quality of a signal-carrying laser beam, causing intensity fading and random signal losses at the receiver.

1. Introduction

Fog, rain, dust, snow, smoke, and other aerosol particulate matter primarily attenuate the signal-carrying laser beam, which to a certain extent can be compensated by

increasing the signal gain. Molecular absorption may be minimized by appropriate selection of the optical wavelength. In contrast, the random fluctuations in the atmosphere's refractive index that cause optical turbulence always result in increased system bit error rates, especially for near-horizontal propagation paths through long stretches of the Earth's atmosphere. Optical turbulence-induced signal losses increase as the transmitter to receiver distance is increased, there is no wavelength "window" where these effects are avoided (although longer wavelengths fare better), and turning up the signal gain does not necessarily improve laser beam quality. For many cases of practical interest, the constant presence of optical turbulence in the atmospheric channel is the limiting factor in reliable wireless optical communication link performance.

In this paper we attempt to provide some understanding of the physical processes underlying optical extinction and optical turbulence-induced signal losses, with an eye towards providing the reader with the necessary tools to evaluate and predict anticipated laser communication system performance. Optical extinction is caused by absorption and scattering by molecules and aerosol particles, and we begin with a description of the basic physical principles behind these processes. There are different models in use for molecular extinction and for aerosol extinction, and we survey those that are the most useful in the prediction of signal loss. Next, atmospheric properties that lead to the formation of optical turbulence are discussed. A comprehensive set of optical turbulence models is provided to enable estimation of optical turbulence strength in all levels of the atmosphere. Behavior of the propagating signal beam in optical turbulence is discussed, which leads to expressions for estimating optical communication system bit error rates as a function of both communication system and optical turbulence parameters. Sufficient information is provided to allow both system designers and users to calculate anticipated system performance over a wide range of channel characteristics for ground-to-ground, ground-to-air/air-to-ground, and satellite uplink/downlink optical communication scenarios.

2. Beam Extinction Due to Atmospheric Aerosols and Molecules

Extinction is defined as a reduction in the intensity of a propagating laser beam. There are two processes that cause extinction: absorption, and scattering. Both of these processes remove energy from the forward propagating beam direction. Absorption and scattering processes can be further subdivided into two size domains: a molecular regime, and a regime of larger particles (aerosols). Here we examine the physics that drives these two processes, and review existing models for describing the effects of absorption and scattering on laser beam propagation in the atmospheric channel.

2.1. Extinction

Extinction describes the attenuation of a laser beam as it passes through a medium containing atoms, molecules, and particles. Goody and Yung [1] define the fundamental law of extinction as that of Lambert, which states that the extinction process is linear in both the intensity of radiation and in the amount of matter, provided that the physical state (i.e., temperature, pressure, composition) is held constant.

For the interaction of a monochromatic propagating beam with a non-vacuous medium over an incremental path length ds, the change in intensity is given by

$$dI_\nu = -I_\nu \beta_{\text{ext},\nu} ds, \tag{1}$$

where I_ν is intensity and $\beta_{\text{ext},\nu}$ an extinction coefficient, both at the frequency ν. Since the extinction process is linear, this implies that

$$\beta_{\text{ext},\nu} = \beta_{\text{sca},\nu} + \beta_{\text{abs},\nu}, \tag{2}$$

i.e., the total extinction β_ν consists of the linear sum of extinction coefficients $\beta_{\text{ext},\nu}$ due to absorption and scattering. Furthermore, the various terms resulting from different constituents of a medium can be summed. As an example, the absorption due to specific absorption lines is expressed as

$$\beta_{\text{abs},\nu} = \sum_{i=1}^{n} \beta_{\text{abs},\nu}(i). \tag{3}$$

The optical thickness, or optical depth, of a medium between two points h_1 and h_2 is defined as

$$\tau_\nu(h_1, h_2) = \int_{h_1}^{h_2} \beta_{\text{ext},\nu}(s) ds. \tag{4}$$

With knowledge of the optical thickness τ_ν and the initial intensityI_0, the resultant reduction in intensity after propagation between the two points h_1 and h_2 is given by

$$I_\nu = I_0 e^{-\tau_\nu}. \tag{5}$$

It is important to determine the explicit absorption $\beta_{\text{abs},\nu}$ for a given frequency ν. Molecules absorb energy in discrete quanta, and this process alters the translational, vibrational, and/or rotational state of the molecule. The resulting spectral line for the absorption coefficient can be described by its central frequency, with a shape that is dependent on several line-broadening features.

2.2. *Molecular Extinction*

Molecular absorption is a primary factor in beam attenuation. At wavelengths above 1 μm, atoms do not couple strongly with the electromagnetic field and their contributions to the extinction $\beta_{\text{ext},\nu}$ can be neglected. Furthermore, molecular extinction at these wavelengths is predominantly due to absorption of incident radiation with only minor scattering contributions.

Molecules absorb energy in discrete quanta, and this process alters the electronic, vibrational, and/or rotational state of the molecule. The absorption spectrum of molecules therefore consists of a series of discrete absorption lines. The bands for rotational spectra are from 10 cm wavelengths to 100 μm; vibrational spectra are typically from 100 to 1 μm. Electronic transitions are in the visible and ultraviolet bands. [2] For wavelengths of interest to free-space laser communication the vibrational spectra are of the greatest relevance.

The nominal frequency and strength of a discrete vibrational energy transition in a molecule can be calculated with quantum mechanical models. However, this discrete transition is rarely observed since the vibrational transition is generally coupled with

rotational transitions, which results in considerable broadening of the line. The broadened line can be described as the sum of individual transitions between a rotational state of the initial vibration state, and a rotational state of the final vibrational state. The approach to describe such a broadband line is then to determine the individual line-profile functions as a function of frequency. These line-profile functions have been represented as a variety of analytic expressions in different theories. The 'natural' line shapes apply to the vacuum. In the atmosphere, the line-profile functions are modified due to mechanisms such as Doppler broadening, or pressure broadening. Additional broadening occurs if molecules become aggregated or dissolve in water droplets.

As a consequence, determination of the shape and strength of the absorption spectrum, and hence the extinction of molecules, from first principles is a complicated task. For many molecules, these absorption spectra (both the broadband vibration-rotational transitions and the individual underlying rotational fine structure) have been measured experimentally under laboratory conditions and have been made available in tabular form. With the use of these tables, the extinction $\beta_{\text{ext},\nu}$ for a specific mixture of molecules (such as encountered in the atmosphere) may be evaluated.

2.3. *Molecular Transmittance Codes*

The Air Force Geophysical Laboratory (AFGL) has developed three models that describe transmittance losses due to molecular absorption: LOWTRAN, MODTRAN, and FASCODE. [3,4] LOWTRAN and MODTRAN are both band models, and FASCODE is a line-by-line model.

The line-by-line model FASCODE determines the optical depth as a function of either wave number or wavelength. An absorption line has three primary characteristics: the central position of the line ν_0; the strength or intensity of the line S; and the shape function for the absorption line. The natural line shape (or Lorentz profile) is given by [1]

$$f(\nu - \nu_0) = \frac{\delta}{\pi[(\nu - \nu_0)^2 + \delta^2]}, \tag{6}$$

where δ is the line width. This profile is most important in the lower atmosphere under conditions of relatively high pressure.

A line-by-line model computes the line for each absorbing gas, which necessarily includes the wings of lines at various distances from the monochromatic line of interest. The result of such a calculation is the layer optical depth, or an optical thickness. The transmittance T of a region is the ratio of incident power to emergent power, and the relation between transmittance and optical depth is implicitly defined by Eq. (5):

$$T = I_\nu / I_0 = e^{-\tau_\nu}.$$

Total transmittance for a path of interest is then determined by finding the transmittances for each homogeneous layer or region that composes the path, and then multiplying them together to get the resultant total transmittance.

A band model such as MODTRAN takes a fundamentally different approach to the determination of transmittance. A band model calculates the spectrally averaged atmospheric transmittance. Such a model does not obey the Lambert law, except for the very narrow spectral windows over which an absorption coefficient can be considered to be constant, which is much less than a line width. This characteristic of spectrally

averaged models is reflected in the fundamental output of the MODTRAN model: the full-path transmittance, rather than the optical depth or an extinction coefficient.

The MODTRAN code has evolved over the last 30 years from an initial low-resolution transmittance code called LOWTRAN. It has been upgraded and augmented by numerous auxiliary programs, including aerosol modules to predict the extinction induced by aerosol particles. MODTRAN is available as a FORTRAN code, but using it without a user interface requires extensive knowledge of the input field formats and output conventions. Many user interfaces have been developed to make the MODTRAN code more accessible. [5]

2.4. *Aerosol Extinction*

The previous section considered the problem of modeling molecular extinction. Here we consider the effects of particles. An aerosol particle is larger than a molecule, but still small enough to remain suspended in the atmosphere for an extended period. The diameter range d of atmospheric aerosols is roughly $0.002\ \mu\text{m} < d < 100\ \mu\text{m}$.

There is a large variability in the source of aerosols in the atmosphere as can be seen from Table 1.

Table 1. Estimates of particles smaller than 20 micron radius emitted into or formed in the atmosphere (10^6 metric tons/year) [SMIC, 1971] (reproduced from [3]).

Natural	
Soil and rock debris*	100–500
Forest fires and slash-burning debris*	3–150
Sea salt	300
Volcanic debris	25–150
Particles formed from gaseous emissions:	
Sulfate from H_2S	130–200
Ammonium salts from NH_3	80–270
Nitrate from NO_8	60–430
Hydracarbons from plant exudations	75–200
Subtotal	773–2200
Man-made	
Particles (direct emissions)	10–90
Particles formed from gaseous emissions:	
Sulfate from SO_2	130–200
Nitrate from NO_8	30–35
Hydracarbons	15–90
Subtotal	185–415
Total	958–2615

*Includes unknown amounts of indirect man-made contributions.

Atmospheric aerosols are generally size-classified in two primary modes, each mode having a different genesis. The fine mode consists of particles with diameters less than $d < 2.5\ \mu\text{m}$. This mode is again divided into an Aitken mode, with particles $d < 0.1\ \mu\text{m}$ created by condensation or nucleation of molecules that grow larger by

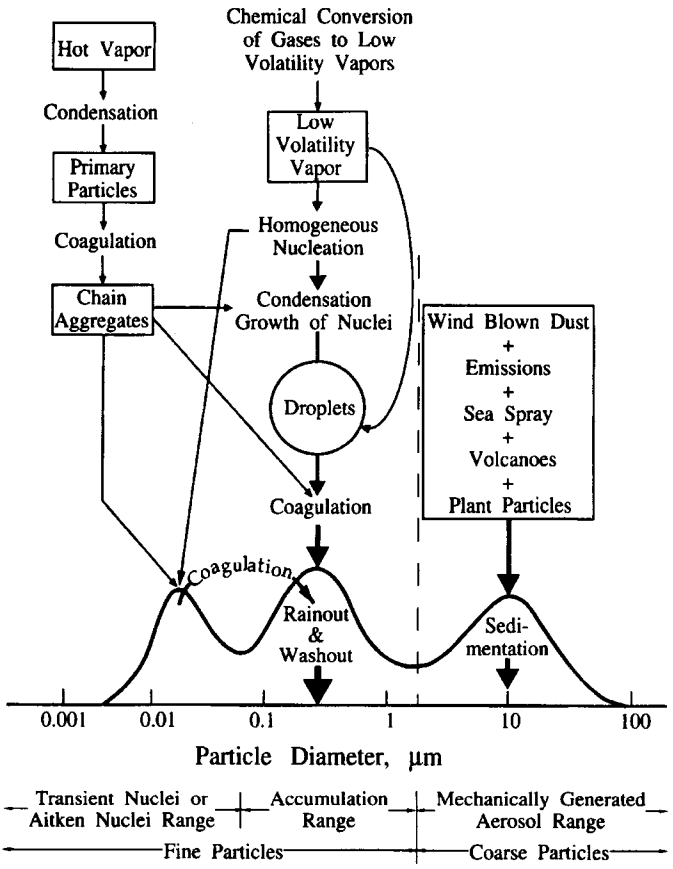


Fig. 1. Definition and genesis of fine mode and coarse mode for aerosol particle size [6].

coagulation; and an accumulation mode, with particle diameters 0.1 μm $< d <$ 2.5 μm that do not evolve into larger particles. The coarse mode with diameter $d >$ 2.5 μm consists primarily of particles generated by mechanical processes, such as wind blowing over soil, sand, or water (sea spray).

Particle lifetimes are dependent upon particle size. If the effects of turbulence are neglected and the only forces on a particle are the gravitational force and the drag force, then a steady-state or terminal settling velocity can be determined. Let T_{100} represent the time for a particle of diameter d to fall 100 m. Then $T_{100} \approx 3.6$ years for $d = 0.1$ μm, $T_{100} \approx 33$ days for $d = 1.0$ μm, $T_{100} \approx 8.5$ hours for $d = 10.0$ μm, and $T_{100} \approx$ 7 minutes for $d = 100.0$ μm. Thus within the first 10 m above the sea or land surface one can find strong gradients in large particles since the current particle population is heavily determined by the production mechanisms of the very recent past (see Fig. 1).

Aerosol composition for open-ocean conditions can be classified into two types: sea-salt particles, and a smaller (residual) concentration of continental aerosol particles. Sea-salt particles are produced from sea-spray droplets that have evaporated, with

the particles then serving as condensation nuclei within the high humidity oceanic environment. These particles can grow in size in regions of high humidity, and shrink by evaporation as humidity decreases.

As will be seen in the next section, beam extinction caused by a single aerosol particle depends on the particle's size and composition. Thus in order to evaluate aerosol extinction for a laser beam propagating through the atmosphere, we need to know the number of aerosols present, as well as their sizes and composition. The aerosol particle size distribution (PSD) describes the number of particles as a function of size. The PSD can be measured by a number of instruments utilizing a variety of mechanical, aerodynamic, and electromagnetic principles. It is more difficult to infer the composition of the particles. One technique analyzes the particles once they have passed through the instrument measuring the PSD, but can only yield general information about the composition. For chemical characterization, it is more customary to collect material on filters and analyze the contents in the laboratory. This yields a detailed composition, but usually with only a very coarse size resolution.

There are considerable experimental difficulties that may be encountered in carrying out measurements of PSD and composition, which makes these experimental techniques less suitable in an operational environment. Hence, models have been developed describing aerosol concentration in the atmosphere as a function of meteorological or climatological parameters. A few of these models are detailed in a subsequent section, but first we turn our attention to Mie theory, which relates aerosol size and composition to extinction.

2.5. *Mie Theory*

The models and measurement techniques described in the previous section result in an aerosol particle size distribution. To determine the effect of atmospheric aerosols in the channel on laser beam propagation, it is necessary to use particle size distribution to determine extinction for a given wavelength. Beam extinction is the sum of the absorption and scattering of laser energy, and both can be calculated using Mie theory.

Extinction characteristics of aerosol particles are computed almost exclusively for either spherical, or (oblate or prolate) spheroid particles. Although other simple geometries can be addressed (e.g., ellipsoids, cylinders, and disks), the fundamental Mie calculation determines the absorption and scattering of an incident plane wave by a homogeneous sphere [7]. In the lower marine atmosphere and at wavelengths of interest, major contributions to aerosol extinction result from sea-spray droplets. These droplets are spherical, implying that Mie theory can be used reliably. For particles that are roughly the same size or larger than the wavelength of interest, scattering is usually the dominant term in the extinction equation. For particles substantially smaller than the wavelength, absorption is dominant.

In Mie theory, the critical input parameters are wavelength and particle size. A key parameter is the dimensionless size parameter x:

$$x = kr = \frac{2\pi r}{\lambda} \tag{7}$$

for a particle of radius r, wave number k, and radiation wavelength λ. The dimensionless size parameter x illustrates the nature of the Mie calculation: it is only the ratio between the particle size and radiation wavelength that matters. For incident light

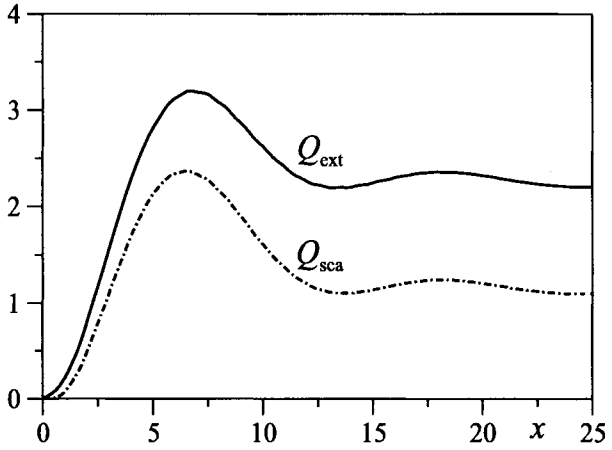


Fig. 2. Extinction efficiency computed for the refractive index $n = 1.29 + i0.05$. This value represents a wet sea salt particle near the 10.6 micron region.

of intensity I_0, the total scattered energy due to a single particle with radius r and refractive index n can be determined. This total energy can be equated to the incident energy on a defined area C_{sca}, the scattering cross-section. Likewise, the total incident energy absorbed by the particle can be equated to an area C_{abs}, called the absorption cross-section. It follows that the beam extinction C_{ext}, or the total energy removed from the beam, is given by

$$C_{\text{ext}} = C_{\text{sca}} + C_{\text{abs}}. \tag{8}$$

The particle size, beam wavelength, and extinction characteristics are related by a quantity called the extinction efficiency. A spherical particle of radius r has a geometric cross-section of πr^2 so that

$$Q_{\text{ext}} = \frac{C_{\text{ext}}}{\pi r^2} \tag{9}$$

defines the effective extinction cross-section. [7]

Next we examine the relation between the size parameter x and the effective cross-section Q_{ext}. In Fig. 2 we show a Mie calculation for the refractive index $n = 1.29 + i0.05$. Computations were done with a program for Mie scattering coefficients called BHMIE written by Bohren and Huffman [8] and adapted for the IDL language by P. J. Flatau. This refractive index is typical for a wet sea salt particle near a wavelength of $\lambda = 10.6$ μm. The relative value of Q_{abs} can be deduced in this figure from $Q_{\text{abs}} = Q_{\text{ext}} - Q_{\text{sca}}$; absorption is a substantial fraction of the overall extinction for this combination of real and imaginary parts of the refractive index.

Plots of the extinction cross-section vs. the size parameter are useful to display the relative change in cross-section as size parameter increases. But it is important to understand a point emphasized by Bohren and Huffman [8]: refractive index and size parameter are not truly independent variables in the sense that if we vary the wavelength λ to effectively change x, the refractive index changes as well; the refractive index of a given substance is a function of wavelength. On the other hand, the particle size r can be changed to vary x without difficulty.

2.6. *Aerosol Models*

It is important to be able to assess the aerosol particle concentration with some accuracy before invoking a Mie calculation. There are a number of aerosol models that have been developed to study extinction phenomena. Technical challenges render a detailed experimental assessment in operational scenarios difficult, so aerosol models are used instead to predict particle concentration on the basis of climatology or actual meteorological conditions. There is no aerosol model that can accurately predict extinction under all conditions. This is particularly the case when local meteorological conditions dominate the aerosol source function. We present modeling possibilities that permit the use of local geographical and meteorological information to more closely approximate the actual aerosol conditions. These models are presented in order of increasingly detailed requirements for local environmental conditions as input data.

The aerosol particle number concentration $c(r)$ is defined as the number of particles per unit volume per radius increment: $c(r) = dN/dr$. Deirmendjian [9] proposed a modified gamma distribution that models the number of particles per radius increment and per unit volume as

$$c(r) = \frac{dN}{dr} = ar^{\alpha} \exp(-br^{\gamma}) \tag{10}$$

with

$$b = \frac{\alpha}{\gamma r_{\text{mod}}^{\gamma}}. \tag{11}$$

Here r_{mod} is the mode radius (in μm) that corresponds to the radius at which $c(r)$ is a maximum, and N is the total number density in particles per cm^3. Note that a, b, α, and γ are four adjustable parameters that allow the function to be fit to a large variety of models. $N(r)$ is a continuous function, but if data from particle size counters are available (generally consisting of counts per discrete bin size), a histogram of N as a function of particle size r can be constructed. The concentration N is found by integrating Eq. (10), and the values of α and γ can be estimated by a curve fit, leaving the value of b to be found from Eq. (11).

The Junge function [10] is another well-known distribution function:

$$c(r) = \frac{dN}{d\log r} = pr^{-v}, \tag{12}$$

where $c(r)$ is now defined as the number of particles per unit volume and per logarithm of the radius. Determination of both p and v is dependent upon the particular characteristics of the aerosol. The exponent v can vary over the range $2 < v < 4$, with the lower value $v \approx 2$ typically found in fog conditions. If field experiments provide particle concentrations as a function of size, then p can be determined from Eq. (12). It is important to realize that Eqs. (10) and (12) have been suggested for aerosol concentrations over land, with only a minimal contribution from sea-salt particles. For the Junge function, these sea-salt particles cause a deviation from the predicted profile for the size range 1 – 10 μm, which includes many of the wavelengths of interest for laser communication applications.

Four models (rural, urban, maritime, and tropospheric) developed by Shettle and Fenn [3] are based upon either a single log-normal distribution, or the sum of two log-normal distributions:

$$c(r) = \frac{dN(r)}{dr} = \sum_{i=1}^{2} N_i f_i(r), \tag{13}$$

where $f_i(r)$ is the log-normal distribution defined by

$$f_i(r) = \frac{1}{r\sigma_i\sqrt{2\pi}} \exp\left\{-\frac{[\ln(r/\rho_i)]^2}{2\sigma_i^2}\right\}, \tag{14}$$

where ρ_i is the mode radius (radius for which the distribution is maximized), and σ_i is a distribution width parameter. These four models are available in the MODTRAN 4.0 code.

The tropospheric model contains only a single log-normal distribution, while the other three models combine two log-normal distributions. The normalized number density N_i is constrained so that $N_1 + N_2 = 1$, and actual distributions can be scaled to provide extinction coefficients or visibilities in accord with observations.

The rural model [3] is designed to represent aerosols that have no urban or industrial sources. The aerosol is a mixture of water soluble components and "dust-like" components. The complex refractive index is determined by the relative weighting of the components in the mixture. The urban model [3] modifies the aerosol mix by including combustion products (e.g., soot) and other industrial products. Note that the refractive index for soot has a relatively large imaginary part (especially in comparison with other aerosol types for the visible wavelengths).

Aerosols created over the open ocean are different from continental aerosols. Marine aerosols consist to a large extent of sea-salt particles that are produced from sea-spray and bursting bubbles, which then aggregate and condense in the high-humidity environment of the near-sea surface. The maritime aerosol model has two components: (1) a sea-salt component, and (2) a continental component. The continental component is identical with the mixture defined in the rural aerosol model except that the largest particles are eliminated due to assumed fallout by the time the air mass has transported out to the open ocean.

These four models represent the large range in variability of aerosol attenuation. It is often desirable to tie a model more closely to the local environmental conditions. The Navy Aerosol Model (NAM) [11] represents an effort to develop a maritime aerosol model that accounts for the production of near-surface aerosol by white-capping and sea spray, and which will permit particle size changes due to changes in relative humidity. NAM is widely used to model the effects of aerosol extinction in marine environments, and because NAM is available in the MODTRAN environment it has become one of the default marine aerosol models.

Similar to the four models developed by Shettle and Fenn based upon the log-normal distribution, in NAM [11,12] the aerosol distribution is assumed to be representable as a sum over four lognormally distributed modes:

$$\frac{dN}{dr} = \sum_{i=0}^{3} N_i f_i(r), \tag{15}$$

where N_i is amplitude, and the shape function $f_i(r)$ is a lognormal distribution normalized to unity:

$$f_i(r) = \frac{1}{\sqrt{2\pi}\sigma_i \exp(\sigma_i^2/2)\rho_i} \exp\left(-\frac{[\ln(r/\rho_i)]^2}{2\sigma_i^2}\right). \tag{16}$$

For historical reasons the amplitudes for each mode are defined by $\{A_i\}_{i=0,\ldots 3}$, where $A_i = \log_{10} N_i$. Each shape function $f_i(r)$ is characterized by the width σ_i and the center radius ρ_i. For $i = 0, \ldots, 3$ $\sigma_i = \sqrt{0.5}$ and the center radii ρ_i are relative humidity-based corrections to the nominal center radii:

$$\rho_i = \frac{R_i(G_i^A - \mathrm{RH}_{10}/100)}{G_i^B(1 - \mathrm{RH}_{10}/100)}, \tag{17}$$

where RH_{10} is relative humidity measured at 10 m above the surface, $R_0 = 0.03$, $R_1 = 0.03$, $R_2 = 0.24$, and $R_3 = 2.0$. The coefficients G_i^A and G_i^B are the Gerber coefficients. [13]

NAM has some well-known limitations. NAM was designed to determine extinction at a single height above the ocean surface, namely 10 m. In addition, it was designed to represent aerosol extinction conditions that occur over open-ocean conditions, and not necessarily the actual conditions (i.e., coastal environments). Because of the need for an accurate marine aerosol model that depends upon local meteorological conditions, the Advanced Navy Aerosol Model (ANAM) [12] was developed to correct some of the limitations and deficiencies found in NAM [11]. ANAM introduces a height dependence for aerosol particle concentrations from the surface up to several tens of meters.

The ANAM provides an additional height-dependent mode (index $i = 4$) for larger particles ($r > 5$ μm):

$$\frac{dN}{dr} = \sum_{i=0}^{4} N_i f_i(r), \tag{18}$$

where f_i is defined in Eq. (14), $\sigma_4 = 0.5$, $R_4 = 8.0$, and $N_4 = 10^{A_4}$ with

$$A_4 = -2.5 + 0.07 v_{10} - 0.04h, \tag{19}$$

where h is height above the water surface, and v_{10} is wind speed measured at a 10-m height.

Additional resources are available to permit the definition of a custom aerosol scenario and thence calculate aerosol extinction. The software package Optical Properties of Aerosols and Clouds (OPAC) provides optical properties of various aerosol species within the spectral range of 0.25 to 40.0 μm. OPAC [14] contains a database of microphysical properties, as well as the resulting optical properties of cloud and aerosol components for different wavelengths and humidity conditions. The optical properties that are provided as output include absorption, scattering, and extinction coefficients. OPAC also contains a FORTRAN program that allows the user to extract data from this database, to calculate additional optical properties, and to calculate optical properties of mixtures of the stored clouds and aerosol components. This model includes aerosol components drawn from a variety of sources [14–16]. OPAC is available on the World Wide Web at http://www.meteo.physik.uni-muenchen.de/strahlung/aerosol/aerosol.html.

2.7. *Atmospheric Attenuation of Laser Power*

The transmittance $T(s)$ is given by

$$T_\nu(s) = \frac{I_\nu(s)}{I_{\nu,0}} = e^{-\beta_\nu s}, \tag{20}$$

where the total extinction coefficient β_ν comprises the aerosol scattering, aerosol absorption, molecular scattering, and molecular absorption terms:

$$\beta_\nu = \beta^{\text{aer}}_{\text{sca},\nu} + \beta^{\text{aer}}_{\text{abs},\nu} + \beta^{\text{mol}}_{\text{sca},\nu} + \beta^{\text{mol}}_{\text{abs},\nu}. \tag{21}$$

The total extinction coefficient is defined in units of km^{-1}.

In a link budget calculation it is necessary to determine the loss L_a due to atmospheric extinction:

$$L_a = 10 \log_{10}[T(s)]. \tag{22}$$

For typical clear air conditions, $\beta_\nu \approx 0.1$; for haze conditions (with a visibility of roughly 4 km), $\beta_\nu \approx 1.0$; in fog conditions, $\beta_\nu \geq 10$, and for a dense fog with a visibility of 10 m, $\beta_\nu \approx 391$. Assume the path length of the desired optical length is 2 km, and assume the atmospheric conditions are hazy with $\beta_\nu = 1.0$ at the wavelength of interest. Transmittance is then given by $T = e^{-\beta_\nu s} = e^{-1\times 2} = e^{-2}$, so that the loss is $L_a = 10 \log_{10}(e^{-2}) \approx 8.7$ dB.

3. Channel Effects Due to Optical Turbulence

Random variations in the refractive index of the Earth's atmosphere—commonly referred to as optical turbulence—are responsible for random fluctuations in the signal-carrying laser beam intensity (irradiance) called 'scintillations'. Related turbulence-induced effects include beam spreading beyond the spreading predicted by diffraction (which decreases spatial power density incident on the receiver), and a continuous random motion of the beam centroid about the receiver ('beam wander'). These effects work in tandem to create random signal losses at the receiver, resulting in increased system bit error rates due to signal fading. They can also result in the complete loss of signal altogether. We review the atmospheric phenomenology giving rise to optical turbulence, and discuss the various models available to describe and quantify optical turbulence in the Earth's atmosphere.

3.1. Refractive Index Structure Parameter C_n^2

Optical turbulence occurs when turbulent winds in the Earth's atmosphere mix the always-present vertical moisture and temperature gradients caused by the Sun's heating of the Earth's surface. This creates irregularities in the refractive index of the atmosphere in the form of eddies, or cells, called 'optical turbules'. In 1941 Kolmogorov and Obukhov [17–20] simultaneously applied dimensional analysis to an isotropic velocity field, and derived inertial-subrange predictions for the forms of the velocity spectra and velocity structure functions that now bear Kolmogorov's name [21]. They found that a subclass of optical turbules within the inertial subrange, defined as the range of optical turbule sizes bounded above by the turbulent outer scale L_o and below by the turbulent inner scale l_o, share a degree of statistical consistency that permits a meaningful theoretical treatment. Here we briefly summarize these results.

The stochastic field describing fluctuations of the refractive index $n(\boldsymbol{\rho})$ at a vector location $\boldsymbol{\rho}$ in the turbulent atmosphere is characterized by a structure function. Let

$n(\boldsymbol{\rho}_1)$ and $n(\boldsymbol{\rho}_2)$ be values of the refractive index at the vector locations $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$, respectively. We can describe fluctuations in the refractive index using the refractive index structure function D_n, defined as

$$D_n(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = \left\langle |n(\boldsymbol{\rho}_1) - n(\boldsymbol{\rho}_2)|^2 \right\rangle, \tag{23}$$

where $\langle \cdots \rangle$ represents statistical averaging. For locally homogeneous and isotropic turbulence the refractive index structure function depends only on the modulus of the vector separation $\boldsymbol{\rho} = \boldsymbol{\rho}_1 - \boldsymbol{\rho}_2$ so that $D_n(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = D_n(\rho)$.

Within the inertial subrange the refractive index structure function is described by the Kolmogorov-Obukhov two-thirds power law [22]:

$$D_n(\rho) = C_n^2(h)\rho^{2/3}, \quad l_o < \rho < L_o, \tag{24}$$

where the proportionality constant $C_n^2(h)$, called the refractive index structure parameter, is a function of wavelength, atmospheric pressure, and atmospheric temperature, and varies strongly with the height above ground h. The "strength" of optical turbulence is characterized by the value of the refractive index structure parameter, which can vary from 10^{-17} $\mathrm{m}^{-2/3}$ or less when optical turbulence is "weak", up to 10^{-12} $\mathrm{m}^{-2/3}$ or more when it is "strong".

The refractive index $n(\rho)$ can be characterized by the spatial power spectral density $\Phi_n(\kappa)$, where κ is a three-dimensional wave vector. The turbulence spectrum links random fluctuations in the Earth's refractive index with optical turbulence effects on laser propagation. From the two-thirds power law, it can be deduced that the associated power spectral density for refractive index fluctuations is given by the well-known Kolmogorov spectrum:

$$\Phi_n(\kappa) = 0.033\, C_n^2\, \kappa^{-11/3}, \quad \kappa_o < \kappa < \kappa_\ell, \tag{25}$$

where $\kappa_\ell = 2\pi/l_o$ and $\kappa_o = 2\pi/L_o$ are the boundaries of the spatial spectrum corresponding to the inertial subrange. The spectrum wavenumber κ (not to be confused with the wavenumber $k = 2\pi/\lambda$ associated with electromagnetic propagation) is defined as $\kappa = 2\pi/\ell$, where ℓ represents turbulent eddy size. Small turbulent eddies correspond to large spectrum wavenumbers, and large turbulent eddies correspond to small spectrum wavenumbers.

To perform calculations using the Kolmogorov model requires assuming that the turbulent outer scale L_o is infinite while the turbulent inner scale l_o is negligibly small (typically l_o is on the order of mm, while L_o is on the order of meters). Note also that the Kolmogorov spectrum has a singularity at $\kappa = 0$. Tatarskii [23] suggested using a Gaussian function to truncate the Kolmogorov spectrum in the dissipation range $\kappa > \kappa_o$:

$$\Phi_n(\kappa) = 0.033\, C_n^2\, \kappa^{-11/3} \exp\left(-\kappa^2/\kappa_m^2\right), \quad \kappa_o < \kappa < \infty, \tag{26}$$

where $\kappa_m = 5.92/l_o$ is the cutoff at high wavenumbers. However, the Tatarskii spectrum also has a singularity at $\kappa = 0$.

In the modified von Karman spectrum, the Tatarskii spectrum is altered so that it is finite and isotropic for $\kappa < \kappa_o$:

$$\Phi_n(\kappa) = 0.033\, C_n^2\, \frac{\exp\left(-\kappa^2/\kappa_m^2\right)}{\left(\kappa^2 + \kappa_o^2\right)^{11/6}}, \quad 0 \leq \kappa < \infty. \tag{27}$$

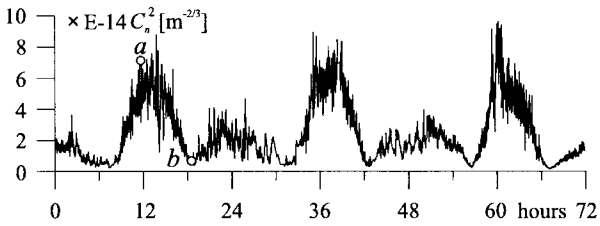


Fig. 3. Diurnal C_n^2 profile over a three-day period in August 2002. Typical daytime value is represented by 'a', and neutral event (turbulence minimum) by 'b' [courtesy Mikhail Vorontsov].

Other atmospheric spectral models have been proposed that account for a "bump" seen at high wavenumbers near $1/l_o$ that appears in measurements of temperature data [24,25]. Since the refractive index obeys the same spectral law as temperature, it follows that the bump should also appear in the spectrum of refractive-index fluctuations. However, for many engineering applications the Kolmogorov model, which does not include consideration of turbulent inner and outer scale effects, is sufficient.

For purposes of understanding the behavior of optical turbulence it is convenient to consider several distinct atmospheric layers with unique features: the surface boundary layer, the convective scaling layer, the tropopause region, and the lower stratosphere. The surface boundary layer is influenced by the aerodynamic roughness of the Earth's surface features, and temperature gradients due to solar heating and radiative cooling. Energy driving the optical turbulence mechanism in this layer comes from two sources: mechanical energy due to wind shear, and buoyant energy from temperature gradients. The vertical extent of the surface boundary layer is determined by the interplay between mechanical and buoyancy contributions to the turbulent energy budget, and varies diurnally from tens to hundreds of meters at night, to perhaps as high as a kilometer or more on warm, sunny days. In the surface boundary layer mechanical contributions dominate and the greatest extremes in optical turbulence behavior are found. In the tropopause there is often a peak in turbulence as a result of wind shear and the strong inversion found in this region. The coupling of gravity waves and turbulence from 2 to 20 km above mean sea level, and a climatology of C_n^2 profiles over the same altitude region, were described by Nastrom and Eaton [26,27].

In the convective scaling layer, mechanical energy no longer dominates the turbulent energy budget and instead buoyancy becomes the driving mechanism. Optical turbulence is therefore less dependent on the aerodynamic roughness of the Earth's surface in this layer. The surface boundary layer and the convective scaling layer together define the planetary boundary layer, which extends upward to the height of the lowest temperature inversion. Above the planetary boundary layer ($h \sim 10$ km) the Earth's surface features no longer influence optical turbulence. In the tropopause there is often a peak in turbulence as a result of wind shear. Above this level turbulence strength rapidly decreases; turbulence effects disappear at altitudes around 25 km.

Figure 3 illustrates typical diurnal behavior for C_n^2 over a three-day period in August 2002. This data was obtained at the Army Research Laboratory's A_LOT testbed in Adelphi, Maryland over a path representative of many free-space laser communication scenarios: a laser transmitter located on top of a two-story building, with the receiver

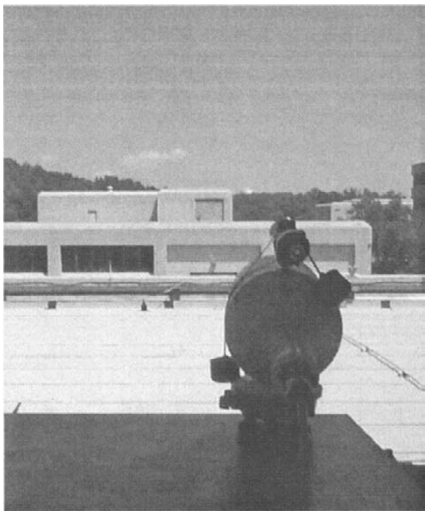


Fig. 4. Propagation path for C_n^2 data shown in Fig. 3. The receiver system is on a water tower, seen here as the small white dot above the tree line located above the laser transmitter.

located on a water tower 2.5 km away (see Fig. 4). Atmospheric turbulence measurements taken at this site over a one-year period indicate that atmospheric turbulence strengths for such a path typically vary between 10^{-13} and 10^{-14} $m^{-2/3}$.

3.2. Optical Turbulence Models

For most terrestrial sites within the first 100 m of the Earth's atmosphere, optical turbulence as described by $C_n^2(h)$ is typically within the range 10^{-13} – 10^{-15} $m^{-2/3}$. However, it is often desirable to have a more accurate estimate for local turbulence strength. The refractive index structure parameter $C_n^2(h)$ not only varies as a function of altitude, but also according to local conditions such as terrain type, geographic location, cloud cover, and time of day. Several dozen turbulence profile models have been developed from experimental measurements made at a variety of locations (e.g., see [28,29]). Most of these predict similar results, although with the exception of the PAMELA model they generally represent average values with little or no modification for local conditions. Most are designed based on the structure of the optical turbulence profile, with the units of measurement for h and $C_n^2(h)$ being m and $m^{-2/3}$, respectively. Some of the more commonly used models are presented below.

3.2.1. PAMELA Model

The PAMELA model provides estimates of C_n^2 within the surface boundary layer. The required inputs are latitude, longitude, date, time of day, percent cloud cover, and terrain type, as well as a single measurement (or estimate) of atmospheric temperature, pressure and wind speed at the desired height h of the C_n^2 estimate. For slant paths

with a variable height, measurements corresponding to the height of the laser receiver should be used for the most accurate C_n^2 estimate [30]. PAMELA was adapted from earlier, more complicated similarity-based optical turbulence models [31–33]. C_n^2 estimates predicted by these earlier models were found to compare reasonably well with experimental data taken at different sites under widely varying environmental conditions [34]. The theory and algorithm upon which the PAMELA model is based are provided here, and a sample version of PAMELA in Mathcad is given in Appendix A.

Mechanical and buoyancy contributions to the turbulent energy budget are described by the process [35]

$$\frac{dE}{dt} = M + B - \varepsilon, \tag{28}$$

where E is mean kinetic energy of turbulence per unit mass, M is rate of energy production due to wind shear, B is rate of energy production due to buoyancy, and ε is the energy dissipation rate. From the Kolmogorov hypothesis of universal equilibrium, for steady-state conditions [35]

$$\varepsilon = M + B. \tag{29}$$

In Eq. (29) the rate of energy production due to wind shear is given by

$$M = K_m \left| \frac{\partial \bar{v}}{\partial h} \right|^2 , \tag{30}$$

where K_m is the turbulent exchange coefficient for momentum, and $\bar{v}$ is average wind speed. The rate of energy production due to buoyancy can be obtained from

$$B = -K_h \left(\frac{g}{\theta} \right) \frac{\partial \theta}{\partial h} . \tag{31}$$

Here g is acceleration due to the Earth's gravity (9.8 m/s^2), K_h is the turbulent exchange coefficient for heat, and θ is mean potential temperature in degrees Kelvin. The potential temperature θ and temperature T are related by the expression $\theta = T + 0.0098\,h$.

The flux profile relationships for the vertical gradients of average wind velocity and mean potential temperature have similar forms, and are given by [36–38]

$$\frac{\partial \bar{v}}{\partial h} = \left(\frac{u_*}{k_v\, h} \right) \Phi_m \left(\frac{h}{L} \right) \qquad \frac{\partial \theta}{\partial h} = \left(\frac{T_*}{k_v\, h} \right) \Phi_h \left(\frac{h}{L} \right) , \tag{32}$$

where k_v is von Karman's constant taken to be 0.4, u_* is friction velocity, T_* is characteristic (or scaling) temperature, $\Phi_m\,(h/L)$ is dimensionless wind shear, $\Phi_h\,(h/L)$ is dimensionless temperature gradient (lapse rate), and L is the Monin-Obukhov length. The Monin-Obukhov length is fundamental in describing the onset of turbulence and the vertical structure of turbulence-induced flux profiles in the Earth's surface boundary layer.

The form of the flux profile relationships varies depending on atmospheric stability as characterized by the Pasquill stability category P. Stable atmospheric conditions occur when the Earth's surface begins to cool. Potential temperature increases with height, so that a parcel of air shifted vertically will return to its original position. Neutral

atmospheric conditions are associated with the absence of a temperature gradient. If a parcel of air is displaced it will remain in its new position. Unstable atmospheric conditions exist when the Earth's surface is being heated so that potential temperature decreases with height. An air parcel displaced upwards will continue to rise.

The Pasquill stability category P can be estimated using [31,39,40]

$$P = -\frac{1}{2}\,(\,4 - c_w + c_r\,)\ , \tag{33}$$

where c_r is radiation class, and the wind speed class c_w is obtained from

$$\begin{aligned} c_w &= \tfrac{\bar{v}_o}{2}, \qquad \bar{v}_o \leq 8\ \mathrm{m/s}, \\ c_w &= \ 4, \qquad \bar{v}_o > 8\ \mathrm{m/s}\ . \end{aligned} \tag{34}$$

In Eq. (34) $\bar{v}_o$ is the average wind speed measured (or estimated) at the reference height h. During the day the radiation class c_r is defined by

$$c_r = \frac{R}{300}, \tag{35}$$

where R is the solar irradiance incident on a unit horizontal area of the Earth's surface in W/m^2. A technique for estimating the solar irradiance R from latitude, longitude, date, time of day, and percent cloud cover is provided in Appendix B. At night, if the cloud cover is greater than 50% $c_r = -1$; otherwise, $c_r = -2$.

Once the Pasquill stability parameter P is determined it is possible to estimate the Monin-Obukhov length L [31,39,40]:

$$L = \left[\left(0.00435P + 0.0037P^3\right)\ h_r^{-\left(0.5 - 0.23\ |P| + 0.0325P^2\right)}\ \right]^{-1}\ . \tag{36}$$

The surface roughness length h_r appearing in Eq. (36) is related to the effective aerodynamic roughness of the Earth's surface features. Shear forces arising from wind flow over the Earth's surface introduce a vertical gradient into the wind velocity. Above a height of approximately ten times the surface roughness length the vertical wind velocity profile is logarithmic. The surface roughness length is related to the average height in cm of the Earth's surface features h_f by [41]

$$h_r \cong \exp\ [-2.85 + 1.2\ \ln(\,h_f)]\ , \quad h_f < 700\,\mathrm{cm}\ . \tag{37}$$

Table 2 lists measured surface roughness length values for typical urban and rural conditions. [42–45]

In the unstable atmosphere (negative P) flux profiles for the dimensionless wind shear $\Phi_m\ (h/L)$ and the dimensionless temperature gradient $\Phi_h\ (h/L)$ are given by [38]

$$\Phi_m\left(\frac{h}{L}\right) = \left[1 - 16\left(\frac{h}{L}\right)\right]^{-1/4} \qquad \Phi_h\left(\frac{h}{L}\right) = 0.74\left[1 - 9\left(\frac{h}{L}\right)\right]^{-1/2}, \tag{38}$$

and when the atmosphere is stable (positive P) they take the form

Table 2. Surface roughness length values for typically encountered surface types.

Surface Type	Roughness Length (cm)
Village	40
Town	55
Light density residential	108
Park	127
Office	175
Central business district	321
Heavy density residential	370
Grass (5-6 cm)	0.75
Alfalfa	2.7
Long grass	3
Grass (60-70 cm)	11.4
Open brush or scrub	16
Wheat	22
Dense brush or scrub	25
Forest clearings or cutovers	32-48
Corn	74
Coniferous forest	110
Citrus orchard	198
Fir forest	283

$$\Phi_m\left(\frac{h}{L}\right) = 1 + 5\left(\frac{h}{L}\right) \qquad \Phi_h\left(\frac{h}{L}\right) = 0.74 + 4.7\left(\frac{h}{L}\right). \tag{39}$$

The friction velocity u_* is given by [38]

$$u_* = \frac{k_v\,\bar{v}_o}{\ln\left(\frac{h}{h_r}\right) - \psi_m}, \tag{40}$$

with the diabatic influence function for momentum ψ_m obtained from

$$\psi_m = \begin{cases} -5(h/L), & P > 0, \\ \ln\left[\left(\frac{1+y^2}{2}\right)\left(\frac{1+y}{2}\right)^2\right] - 2\arctan(y) + \frac{\pi}{2}, y = [1 - 16(h/L)]^{1/4}, & P \leq 0. \end{cases} \tag{41}$$

The characteristic temperature T_* is defined as

$$T_* = \frac{-H}{c_\rho\,\rho\,u_*}. \tag{42}$$

In Eq. (42) c_ρ is the specific heat at constant pressure given by 1004 J/(kg°K), and ρ is atmospheric density in kg/m^3. The atmospheric density can be obtained from ρ = Pa/(2.9T) where Pa is atmospheric pressure in millibars and T is atmospheric temperature in degrees Kelvin, both measured (or estimated) at the height h. An algorithm for estimating the sensible heat flux H is provided in Appendix B. Note that the characteristic temperature becomes undefined at very low wind speeds. Equating wind speeds of less than 3 to 3 m/s avoids this singularity in the model while still retaining reasonable accuracy [30].

With knowledge of the friction velocity u_* it is possible to estimate the turbulent exchange coefficient for heat K_h [38]:

$$K_h = \frac{k_v u_* h}{\Phi_h\left(\frac{h}{L}\right)}. \tag{43}$$

Now consider the atmospheric refractive index, which at optical frequencies depends on four quantities: atmospheric temperature, atmospheric pressure, optical wavelength, and specific humidity. For optical frequencies, humidity over dry land generally contributes less than 1% to the refractive index fluctuations and is typically ignored. The refractive index of air is often written in the form $n = \bar{n} + \Delta n$, where the average value of the refractive index $\bar{n}$ is slightly greater than one and Δn represents random deviations from this average value. Optical turbulence-induced deviations of the refractive index of air from its average value at optical frequencies can be approximated by [46]

$$\Delta n = \frac{77.6 \times 10^{-6}\text{Pa}}{T}\left(1 + \frac{7.52 \times 10^{-3}}{\lambda^2}\right), \tag{44}$$

where λ is wavelength in microns. For standard atmospheric temperature and pressure at sea level ($T = 288\,^{\circ}$K, Pa = 1013 millibars) and a wavelength of one micron, variation in the refractive index dn is on the order of 3×10^{-4}.

By assuming that C_n^2 has a representation analogous to that of the potential temperature structure function we can write [35]

$$C_n^2 = \frac{bK_h}{\varepsilon^{1/3}}\left(\frac{\partial n}{\partial h}\right)^2, \tag{45}$$

where b is a constant commonly approximated by 2.8. [47]

By ignoring the small contribution to the total differential from fluctuations in atmospheric pressure, we can differentiate Eq. (44) with respect to the potential temperature θ and obtain

$$\frac{\partial n}{\partial \theta} = \frac{\left(-77.6 \times 10^{-6}\text{Pa}\right)\left(1 + 7.52 \times 10^{-3}\lambda^{-2}\right)}{\left(\theta\text{-}0.0098\, h\right)^2}. \tag{46}$$

From the chain rule

$$\frac{\partial n}{\partial h} \approx \frac{\partial n}{\partial \theta}\frac{\partial \theta}{\partial h}, \tag{47}$$

and using Eq. (32) while ignoring the small contribution due to wavelength it follows that

$$\frac{\partial n}{\partial h} \cong \frac{\left(-77.6 \times 10^{-6}\text{Pa}\right) T_* \Phi_h\left(\frac{h}{L}\right)}{k_v h T^2}. \tag{48}$$

This completes the set of equations required to estimate the refractive index structure parameter C_n^2 given by Eq. (45). An example of the PAMELA model programmed in Mathcad is given in Appendix A.

In Fig. 5(a–f) diurnal C_n^2 values estimated using the PAMELA model are compared with scintillometer values measured over the path depicted in Fig. 4 at ARL's A_LOT facility [30]. For each diurnal plot an averaged value of atmospheric pressure was used.

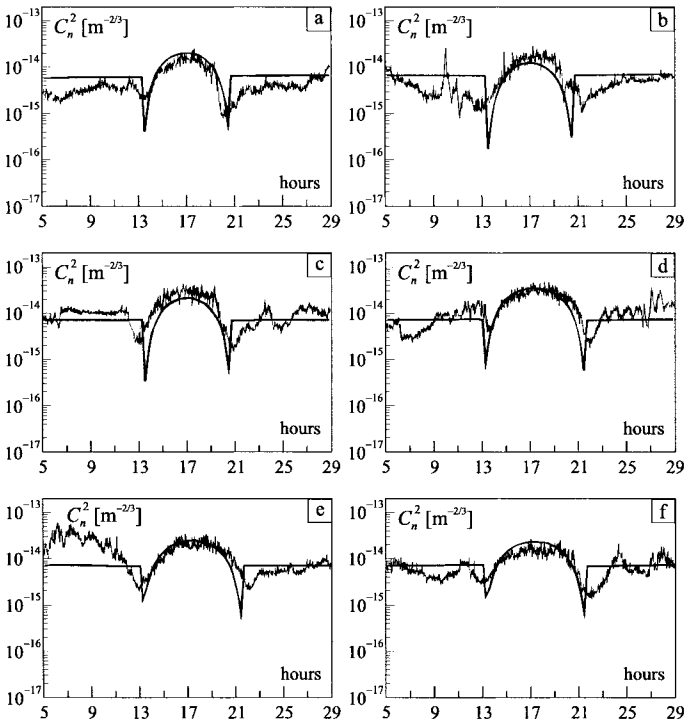


Fig. 5. Comparison of measured and predicted C_n^2 values for (a) 15 December 2002, (b) 16 December 2002, (c) 17 December 2002, (d) 8 February 2003, (e) 12 February 2003, and (f) 13 February 2003.

Three averaged values of wind speed were used, one for after midnight, one for daytime conditions, and one for late evening. Together with a smoothed running average of the atmospheric temperature measurements, these formed the necessary meteorological inputs to the PAMELA model. In each case the predicted average C_n^2 values obtained using the PAMELA model compare favorably with the measured C_n^2 values.

3.2.2. Gurvich Model

As shown in many experiments, the dependence of C_n^2 on the altitude h in the surface boundary layer follows the power law $C_n^2 \propto h^{-q}$, where q is 4/3, 2/3, or 0 for unstable, neutral, or stable atmospheric conditions, respectively [48]. Based on this law, Gurvich et. al. [49] suggested the following phenomenological model to obtain profiles of C_n^2 in the atmosphere. In this model profiles are classified into four groups dependent on the measured (or estimated) value of C_n^2 at the height $h = 2.5$ m. Analytical expressions that approximate the experimental data are given as follows [48,50]:

1. For $C_n^2|_{2.5\text{ m}} > 10^{-13}$ m$^{-2/3}$ (strong turbulence):

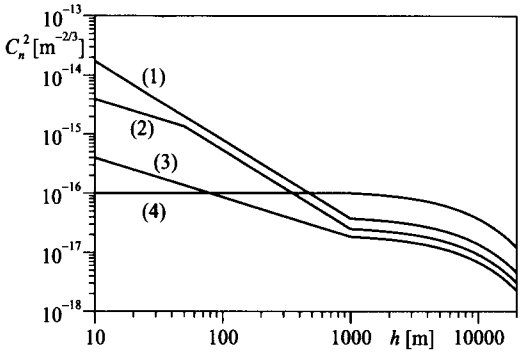


Fig. 6. Atmospheric turbulence profiles for the Gurvich model for four different values of $C_n^2|_{2.5\text{ m}}$. Top to bottom: (1) $C_n^2|_{2.5\text{ m}} = 1.1 \times 10^{-13}$ m$^{-2/3}$; (2) $C_n^2|_{2.5\text{ m}} = 10^{-14}$ m$^{-2/3}$; (3) $C_n^2|_{2.5\text{ m}} = 10^{-15}$ m$^{-2/3}$; (4) $C_n^2|_{2.5\text{ m}} = 10^{-16}$ m$^{-2/3}$.

$$C_n^2(h) = \begin{cases} C_n^2|_{2.5\text{ m}}\,(h/2.5)^{-4/3}, & 2.5\text{ m} \le h \le 1000\text{ m} \\ C_n^2|_{1000\text{ m}} \exp\left[-(h-1000)/9000\right], & 1000\text{ m} < h \end{cases} \tag{49}$$

2. For 10^{-13} m$^{-2/3} \ge C_n^2|_{2.5\text{ m}} > 6.5 \times 10^{-15}$ m$^{-2/3}$,

$$C_n^2(h) = \begin{cases} C_n^2|_{2.5\text{ m}}\,(h/2.5)^{-2/3}, & 2.5\text{ m} \le h \le 50\text{ m} \\ C_n^2|_{50\text{ m}}\,(h/50)^{-4/3}, & 50\text{ m} \le h \le 1000\text{ m} \\ C_n^2|_{1000\text{ m}} \exp\left[-(h-1000)/9000\right], & 1000\text{ m} < h \end{cases} \tag{50}$$

3. For 6.5×10^{-15} m$^{-2/3} \ge C_n^2|_{2.5\text{ m}} > 4.3 \times 10^{-16}$ m$^{-2/3}$,

$$C_n^2(h) = \begin{cases} C_n^2|_{2.5\text{ m}}\,(h/2.5)^{-2/3}, & 2.5\text{ m} \le h \le 1000\text{ m} \\ C_n^2|_{1000\text{ m}} \exp\left[-(h-1000)/9000\right], & 1000\text{ m} < h \end{cases} \tag{51}$$

4. For 4.3×10^{-16} m$^{-2/3} \ge C_n^2|_{2.5\text{ m}}$ (weak turbulence):

$$C_n^2(h) = \begin{cases} C_n^2|_{2.5\text{ m}}, & 2.5\text{ m} \le h \le 1000\text{ m} \\ C_n^2|_{1000\text{ m}} \exp\left[-(h-1000)/9000\right], & 1000\text{ m} < h \end{cases} \tag{52}$$

Sample calculations for C_n^2 profiles corresponding to the Gurvich model are shown in Fig. 6.

In subsequent studies, Gracheva and Gurvich [51] developed the so-called 'simple' models of turbulence, obtained from a large number of experimental measurements:

A. Model for the 'best' atmospheric conditions (weak turbulence):

$$\begin{aligned} &\log_{10}\left[C_{n\,\min}^2(h) - 5.19 \times 10^{-16} \times 10^{-0.00086\,h}\right] \\ &= -18.34 + 2.9 \times 10^{-4} h - 2.84 \times 10^{-8} h^2 + 7.43 \times 10^{-13} h^3 \end{aligned} \tag{53}$$

B. Model for the 'worst' atmospheric conditions (strong turbulence):

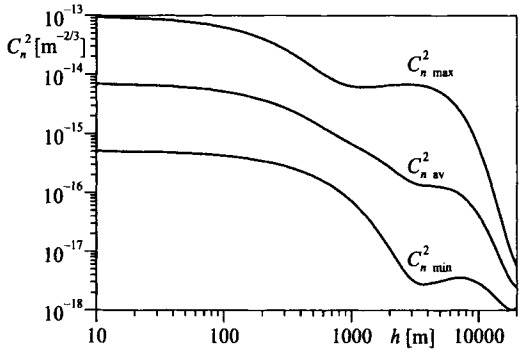


Fig. 7. Atmospheric turbulence profiles for the Gurvich 'simple' models.

$$\log_{10}\left[C_{n\,\max}^2(h) - 9.5\times10^{-14}\times10^{-0.00209h}\right] = -14.39 + 1.7\times10^{-4}h - 3.48\times10^{-8}h^2 + 9.59\times10^{-13}h^3 \tag{54}$$

C. Model for 'average' atmospheric conditions (geometrical mean value of the two previous profiles):

$$\log_{10}\left[C_{n\,\mathrm{av}}^2(h)\right] = \frac{1}{2}\left\{\log_{10}\left[C_{n\,\max}^2(h)\right] + \log_{10}\left[C_{n\,\min}^2(h)\right]\right\}. \tag{55}$$

Sample calculations for $C_{n\,\min}^2$, $C_{n\,\max}^2$, and $C_{n\,\mathrm{av}}^2$ are presented in Fig. 7.

3.2.3. SLC–Day Model

The Submarine Laser Communication–Day (SLC–Day) model given by [52,53]

$$C_n^2(h) = \begin{cases} 0 & 0\text{ m} < h < 19\text{ m} \\ 4.008\times10^{-13}h^{-1.054} & 19\text{ m} < h < 230\text{ m} \\ 1.300\times10^{-15} & 230\text{ m} < h < 850\text{ m} \\ 6.352\times10^{-7}h^{-2.966} & 850\text{ m} < h < 7000\text{ m} \\ 6.209\times10^{-16}h^{-0.6229} & 7000\text{ m} < h < 20000\text{ m} \end{cases} \tag{56}$$

is frequently used to describe daytime turbulence conditions at inland sites.

3.2.4. Hufnagel–Valley Model

The Hufnagel–Valley (HV) model [54,55] is one of the most popular models for inland sites and daytime viewing conditions since it permits variations in high-altitude wind speed and near-ground turbulence levels. In this model $C_n^2(h)$ is assigned a sum of three exponential decay terms corresponding to a surface boundary layer, a strong layer caused by the high-altitude jet stream, and a background tropopause layer, i.e.,

$$C_n^2(h) = A\exp(-h/100) + 5.94\times10^{-53}(v/27)^2h^{10}\exp(-h/1000) + 2.7\times10^{-16}\exp(-h/1500), \tag{57}$$

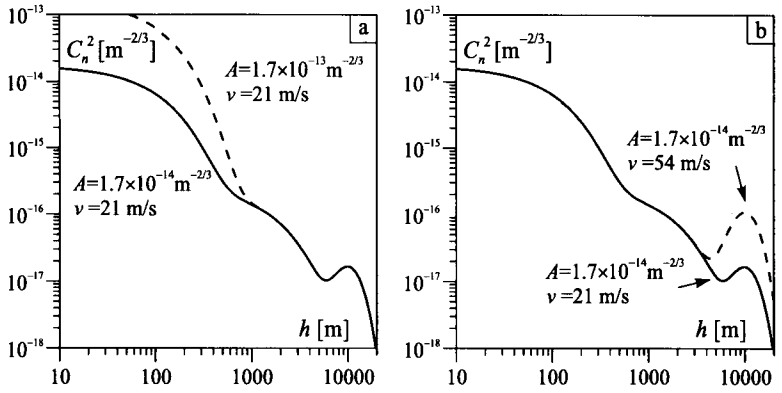


Fig. 8. Altitude turbulence profiles associated with the Hufnagel–Valley model for different values of (a) structure parameter A, and (b) high-altitude rms wind speed v.

where A is the nominal value of C_n^2 at the ground and v is the estimated high-altitude rms wind speed in m/s in the 5–20-km altitude range [29]. Commonly used values are $A = 1.7\times10^{-14}$ m$^{-2/3}$ and $v = 21$ m/s ("HV–21 model"). To illustrate the effect of changing these two parameters, two values of nominal ground-turbulence levels and two values of moderate-to-strong high-altitude wind speeds were used to calculate the C_n^2 profiles shown in Fig. 8. Note that the ground-turbulence level has no visible effect above 1 km.

3.2.5. HV–Night Model

Night-time turbulence conditions at good seeing locations have been modeled using a version of the Hufnagel–Valley profile modified to fit observations made during experiments at the Air Force Maui Optical Station [56]. The modified Hufnagel–Valley (or HV–Night) profile is given by [53]

$$C_n^2(h) = 1.9\times10^{-15}\exp(-h/100) + 8.16\times10^{-54}h^{10}\exp(-h/1000) + 3.02\times10^{-17}\exp(-h/1500)\,. \quad (58)$$

3.2.6. Greenwood Model

Another nighttime turbulence model developed for astronomical imaging from mountaintop sites is given by [28,57]

$$C_n^2(h) = \left[2.2\times10^{-13}(h+10)^{-1.3} + 4.3\times10^{-17}\right]\exp(-h/4000)\,. \quad (59)$$

The C_n^2 profile models discussed above are compared in Fig. 9. As shown, turbulence in the surface layer is typically stronger for the day models than for the night models.

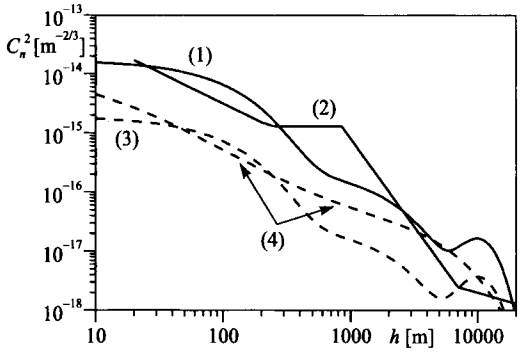


Fig. 9. Atmospheric turbulence profiles associated with two day (solid lines) and two night (dashed lines) models: (1) HV–21 model; (2) SLC–Day model; (3) HV–Night model; (4) Greenwood model.

3.2.7. Other Turbulence Models

Van Zandt et al. [58,59] developed a sophisticated theoretical Aeronomy Laboratory model based on relationships proposed by Tatarski [23] to calculate profiles of C_n^2 from archival rawinsonde data:

$$C_n^2 = b\,\alpha'\,L_o^{4/3}\,M^2\,, \tag{60}$$

where b is a constant commonly approximated by 2.8, α' is a ratio of eddy diffusivities usually taken to be unity, and M is the generalized index of refraction involving atmospheric pressure, absolute temperature, potential temperature, and specific humidity-all of which can be obtained from rawinsonde data. The motivation was to create a model that could be used anywhere rawinsonde data exists. One difficulty with this approach is to relate or parameterize small-scale fluctuations in terms of large-scale variables that can be measured. In [58] L_o was treated unrealistically as a constant within turbulent layers. Within each slab defined by the vertical resolution of rawinsonde data, the value of C_n^2 was calculated using $L_o = 10$ m. Another important part of the model is to specify the probability of turbulence occurrence for the fraction F of the slab that is expected to be turbulent, where F is determined from a statistical calculation that accounts for stability and mean shear of each slab. Since 1978 the model has evolved considerably, and now contains a more realistic probabilistic treatment of L_o and F [60]. Model results have shown good agreement with radar measurements.

The AFRL Radiosonde Model [61] is similar in many respects to the Aeronomy Laboratory Model in that it uses the relationship between C_n^2 and L_o, and the Richardson's number. However, it is simpler in construction. It assumes that microshears cause turbulence, but the shear used in the Richardson number calculation is not data-dependent as in the Aeronomy model calculation. Instead, in the AFRL model a "model atmosphere" is used for calculation. Two separate models are used by the AFRL model, one for the troposphere and one for the stratosphere. These models are based on high-resolution wind profiles obtained using measurements of rocket trails.

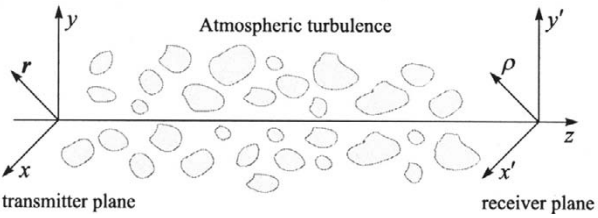


Fig. 10. Propagation geometry.

3.3. Free-Space Laser Communication in Optical Turbulence

With the use of any of the previously described optical turbulence models it is possible to estimate the anticipated range of optical turbulence levels for a particular geographic location and time. But to fully characterize the effects of optical turbulence on free-space laser communication, it is also necessary to consider the actual propagation of the signal-carrying laser beam through this turbulent atmosphere. We begin here by examining the propagation of a laser beam in free space, and then use this framework to develop the necessary expressions for predicting laser communication system performance in a turbulent channel for various communication scenarios. The approach presented here enables system users and developers to consider the various design and site location trade-offs involved in maximizing system performance.

3.3.1. Free-Space Laser Beam Propagation

Consider a simple free-space Gaussian beam wave model with parabolic wavefront phase envelope that allows complete characterization of the focusing or diverging characteristics of a laser beam [62,63]. Since it was first proposed this beam wave model has been used in a number of studies concerning laser beam propagation through random media. [64–67]

At $z = 0$ the free-space electric field of a unit amplitude, lowest-order paraxial Gaussian beam propagating predominantly along the z-axis can be represented in the form

$$U(r,\ 0) = \exp\left[-\left(\frac{1}{w_o^2} + \frac{jk}{2R_o}\right) r^2\right], \tag{61}$$

where w_o is the transmitter beam radius (beam size), R_o is the radius of curvature of the phase front, $k = 2\pi/\lambda$ is the optical wave number, and $r = (x^2 + y^2)^{1/2}$ is transversal distance from the beam center in the transmitter plane. The propagation geometry is shown in Fig. 10.

After propagating a distance z from the transmitter the optical field becomes

$$U(\rho,\ z) = \frac{\exp(jkz)}{\hat{r} + j\,\hat{z}} \exp\left[-\frac{1}{\hat{r} + i\,\hat{z}}\left(\frac{1}{w_o^2} + \frac{jk}{2R_o}\right)\rho^2\right], \tag{62}$$

where $\rho = (x'^2 + y'^2)^{1/2}$ is transversal distance from the beam optical axis in the receiver plane. In Eq. (62) we have used the transmitter beam parameters [66,67][5]

$$\hat{r}(z) = \frac{R_o - z}{R_o}, \quad \hat{z} = \frac{z}{\hat{z}_d}, \tag{63}$$

where the normalized focusing parameter $\hat{r}$ characterizes focusing properties of the beam in terms of deviation of the wavefront curvature from the condition of optimal focusing $R_o = z$, and $\hat{z}_d = kw_o^2/2$ is diffractive distance. In this notation convergent (focused) beams are indicated by positive R_o and divergent beams by negative R_o. At the beam waist R_o is infinite and the beam size takes its smallest possible value. In terms of the focusing parameter convergent beams are defined by $\hat{r} < 1$, collimated beams by $\hat{r} = 1$, and divergent beams by $\hat{r} > 1$. In the following text, the functional dependence of $\hat{r}$ and $\hat{z}$ on propagation distance z is to be understood.

Free-space laser beam characteristics at the receiver are defined by

$$\hat{r}_{\text{rec}}(z) = \frac{\hat{r}}{\hat{r}^2 + \hat{z}^2} = \frac{R(z) + z}{R(z)}, \quad \hat{z}_{\text{rec}}(z) = \frac{\hat{z}}{\hat{r}^2 + \hat{z}^2} = \frac{z}{0.5\, k\, w^2(z)}. \tag{64}$$

As with the transmitter beam parameters, the functional dependence of $\hat{r}_{\text{rec}}$ and $\hat{z}_{\text{rec}}$ on z is to be understood.

The free-space beam size $w(z)$ and phase front radius of curvature $R(z)$ at the receiver plane can be expressed in terms of the transmitter beam parameters as

$$w(z) = w_o(\hat{r}^2 + \hat{z}^2)^{\frac{1}{2}}, \quad R(z) = \frac{z\,(\hat{r}^2 + \hat{z}^2)}{\hat{r}(1 - \hat{r}) - \hat{z}^2}. \tag{65}$$

For a collimated beam with the beam waist at the transmitter aperture $z = 0$ and $\hat{r} = 1$, so that we obtain the well-known expressions describing beam size and phase front radius of curvature as a function of z:

$$w(z) = w_o\left[1 + (\lambda z/\pi w_o^2)\right]^{1/2}, \quad R(z) = z\left[1 + (\pi w_o^2/\lambda z)\right]^{1/2}. \tag{66}$$

In terms of the beam size $w(z)$ the average intensity at the receiver is expressed as

$$I(\rho, z) = \frac{w_o^2}{w^2(z)} \exp\left(\frac{-2\rho^2}{w^2(z)}\right). \tag{67}$$

Other relations involving these parameters are discussed in [63].

3.3.2. Laser Beam Propagation in Optical Turbulence

Using the notation developed above, the beam size (radius) after propagating a distance z in optical turbulence is given by [66,67]

$$w(z) = w_o(\hat{r}^2 + \zeta\,\hat{z}^2)^{1/2}, \quad \zeta = \zeta_S + \frac{2w_o^2}{\rho_o^2}, \tag{68}$$

[5] Other notations used for these quantities include $\hat{z} = \Omega$ or $\hat{z} = \Lambda_o$, and $\hat{r} = \Omega_o$ or $\hat{r} = \Theta_o$.

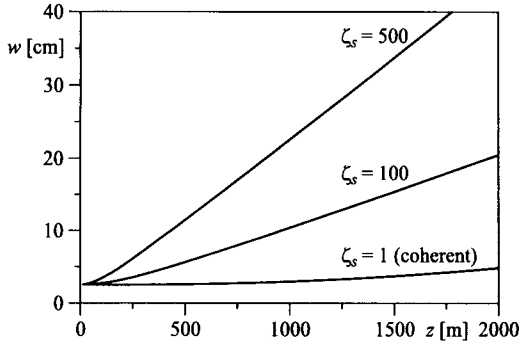


Fig. 11. Collimated beam size as a function of range for several choices of phase diffuser ($\lambda = 0.785$ μm, $w_o = 2.5$ cm, and $C_n^2 = 10^{-14}$ m$^{-2/3}$).

where $\rho_o(z) = (0.55\, C_n^2 k^2 z)^{-3/5}$ is the coherence length of a spherical wave propagating in optical turbulence, ζ is the global coherence parameter, and ζ_s describes spatial coherence properties of the signal-carrying laser beam as it exits the transmitter ($\zeta_S = 1$ for a coherent beam, $\zeta_S > 1$ for a partially coherent beam).

Equation (68) provides a simple way to calculate the desired transmitter diffuser characteristics in order to obtain the optimal beam footprint at the receiver for a set of anticipated turbulence conditions. For example, as shown in Fig. 11 after propagating 2 km through moderate optical turbulence a collimated beam ($\lambda = 0.785$ μm) with a transmitter beam size of 2.5 cm will have a beam footprint (diameter) of only 9.6 cm. With such a small beam footprint the probability of pointing errors is quite high. By placing a phase diffuser directly in front of the laser transmitter aperture it is possible to increase the beam size at the receiver to accommodate potential misalignment or pointing errors. If we choose $\zeta_S = 100$ the beam footprint after propagating 2 km through moderate optical turbulence will be 41 cm, and for $\zeta_S = 500$ the beam footprint will be 90 cm. However, since increasing the beam footprint also reduces the power incident on a fixed-size receiver, it is important to not significantly overestimate the necessary receiver beam size.

The average intensity for a unit amplitude laser beam in optical turbulence is provided by

$$\langle I(\rho) \rangle = \frac{w_o^2}{w^2(z)} \exp\left(\frac{-2\rho^2}{w^2(z)}\right), \tag{69}$$

where $w(z)$ is given by Eq. (68).

The phase front radius of curvature $R(z)$ for a laser beam propagating in optical turbulence is

$$R(z) = \frac{z(\hat{r}^2 + \zeta\,\hat{z}^2)}{\phi\hat{z} - \zeta\hat{z}^2 - \hat{r}^2}\,, \quad \phi \equiv \frac{\hat{r}}{\hat{z}} - \hat{z}\,\frac{w_o^2}{\rho_o^2} \tag{70}$$

and the receiver beam parameters are given by

$$\hat{r}_{\rm rec}(z) = \frac{R(z) + z}{R(z)}\,, \quad \hat{z}_{\rm rec}(z) = \frac{z}{0.5\,k\,w^2(z)}. \tag{71}$$

Finally, the complex degree of coherence for a laser beam in turbulence is available from

$$\mu(\rho, z) = \exp\left\{-\frac{\rho^2}{\rho_o^2}\left(1 + \frac{\rho_o^2}{2w_o^2\,\hat{z}^2} - \frac{\phi^2\rho_o^2}{2\,w^2(z)}\right)\right\} = \exp\left\{-\frac{\rho^2}{\rho_C^2}\right\}, \quad (72)$$

where ρ_C is coherence length of the optical field at the receiver:

$$\rho_C = \rho_o\left(1 + \frac{\rho_o^2}{2w_o^2\,\hat{z}^2} - \frac{\phi^2\rho_o^2}{2\,w^2(z)}\right)^{-1/2}. \quad (73)$$

3.3.3. Scintillation Index and Aperture Averaging

The scintillation index $\sigma_{\ln S}^2$ describes fluctuations (scintillations) in optical power as measured by a point receiver. It is related to the log-amplitude variance σ_χ^2 through [68]

$$\sigma_{\ln S}^2 = \frac{\langle I^2(\rho, z)\rangle}{\langle I(\rho, z)\rangle^2} - 1 = \exp\left(4\sigma_\chi^2\right) - 1. \quad (74)$$

In Eq. (74) $I(\rho, z)$ is the intensity received at a point on the receiver aperture after propagating a distance z through the turbulent channel, and $\sigma_\chi^2(\rho, z)$ is the log-amplitude variance. In the weak fluctuation regime where $4\sigma_\chi^2(\rho,z) < 1$ we can approximate

$$\sigma_{\ln S}^2 = \exp\left(4\sigma_\chi^2\right) - 1 \approx 4\sigma_\chi^2. \quad (75)$$

The use of Eq. (75) leads to the following expression for the scintillation index of a Gaussian beam wave [62,64]:

$$\begin{aligned}\sigma_{\ln S}^2 \cong\ & 4.42\,\sigma_1^2\,\hat{z}_{\rm rec}^{5/6}\frac{\rho^2}{w^2(z)} + 3.86\,\sigma_1^2\Big\{0.4\left[(1 + 2\,\hat{r}_{\rm rec})^2 + 4\hat{z}_{\rm rec}^2\right]^{5/12} \\ & \times\cos\left[\tfrac{5}{6}\arctan\left(\frac{1 + 2\hat{r}_{\rm rec}}{2\hat{z}_{\rm rec}}\right)\right] - \tfrac{11}{16}\,\hat{z}_{\rm rec}^{5/6}\Big\},\end{aligned} \quad (76)$$

where $\sigma_1^2(z) = 1.23\,C_n^2 k^{7/6} z^{11/6}$ is the Rytov variance for a plane wave. The receiver beam parameters are given above in Eq. (71). Since the effects of optical turbulence are accounted for in Eq. (76) by the Rytov variance σ_1^2, only the source coherence parameter ζ_S (not the global coherence parameter ζ) should be used when calculating $R(z)$ and $w(z)$ for use in Eq. (76). When $4\sigma_\chi^2(\rho,z) >> 1$ (strong fluctuation regime), results obtained using Eq. (76) tend to overestimate the level of signal fading and the resulting bit error rates.

Figure 12 illustrates the scintillation index $\sigma_{\ln S}^2\big|_{\rho=0}$ as a function of channel distance z for two optical turbulence strengths C_n^2. Note that as turbulence strength increases by half an order of magnitude from 10^{-14} to 5×10^{-14} $\mathrm{m}^{-2/3}$, the scintillation index increases dramatically.

After propagating through optical turbulence in the channel, the quality of the beam footprint at the receiver deteriorates and begins to break up into random regions

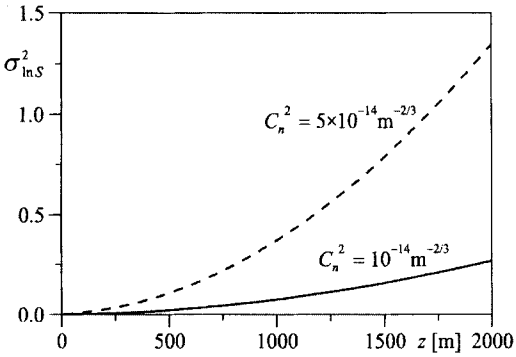


Fig. 12. Scintillation index for a collimated ($\hat{r} = 1$), partially coherent ($\zeta_S = 100$) laser beam as a function of channel length for different optical turbulence strengths ($\lambda = 1.55$ μm, $w_o = 2.5$ cm).

Fig. 13. Typical beam footprint of a laser beam after propagation through optical turbulence.

of high and low intensity (see Fig. 13). For sufficiently long path lengths and/or large optical turbulence strengths, the loss in beam quality at the receiver causes random, fluctuating regions of low (or no) signal intensity, leading to noticeable signal fades. Increasing the size of the receiver aperture relative to the size of these regions of low intensity "averages" the signal fluctuations, and thus decreases signal fading. This phenomenon is known as "aperture averaging".

To describe the effects of aperture averaging on the received signal it is necessary to multiply the scintillation index by the aperture-averaging factor A [23,67]:

$$\sigma'^2_{\ln S} = A\sigma^2_{\ln S}, \tag{77}$$

where A is given by [67]

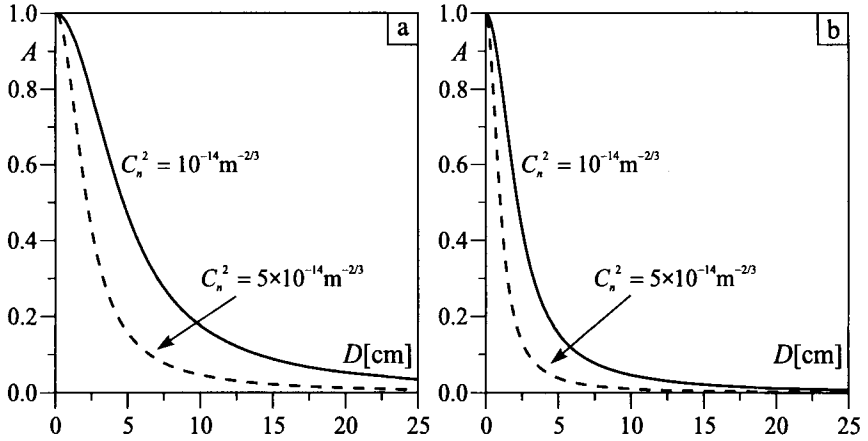


Fig. 14. Aperture-averaging factor as a function of the receiver lens diameter for (a) $\lambda = 1.55$ μm and (b) $\lambda = 0.785$ μm ($\hat{r} = 1$, $\zeta_S = 100$, $w_o = 2.5$ cm, $z = 2000$ m).

$$A = \frac{16}{\pi} \int_0^1 x\,dx \exp\left\{ -\frac{D^2 x^2}{\rho_o^2} \left[2 + \frac{\rho_o^2}{w_o^2\,\hat{z}^2} - \frac{\rho_o^2\,\phi^2}{w^2(z)} \right] \right\} \times \left[\arccos(x) - x\sqrt{1-x^2} \right]. \tag{78}$$

In Eq. (78) $x = \rho/D$, and D is the receiver aperture diameter. An example of how to calculate the aperture-averaged scintillation index using MathCad is given in Appendix C.

The aperture-averaging factor A is shown as a function of the receiver lens diameter D for different beam wavelengths and optical turbulence strengths in Figs. 14 (a, b). Note that the degree of beam divergence does not significantly affect the aperture-averaging factor A. Also note that more aperture averaging occurs when optical turbulence is stronger (compare solid lines for $C_n^2 = 10^{-14}$ m$^{-2/3}$ with the dashed lines for $C_n^2 = 5\times10^{-14}$ m$^{-2/3}$). This is because for stronger optical turbulence the coherence length ρ_o decreases relative to the size of the receiver aperture, resulting in additional aperture averaging. Similarly, the smaller wavelength in Fig. 14 results in a smaller coherence length ρ_o relative to the size of the receiver aperture, and thus additional aperture averaging occurs.

The effects of aperture averaging on the scintillation index can be seen from Figs. 15 (a, b). The scintillation index for a point corresponds to a lens diameter of zero. As receiver aperture size increases, the scintillation index decreases significantly. However, for the case considered here where $\lambda = 1.55$ μm, aperture sizes larger than about 6 cm do not lead to any significant addition in aperture averaging. Decreasing the wavelength by a factor of two to $\lambda = 0.785$ μm roughly halves the aperture size required to "average" the majority of scintillations. It is clear that aperture averaging can provide a powerful advantage in reducing atmospheric turbulence-induced signal fades. However, depending on the channel distance z, optical turbulence strength, and

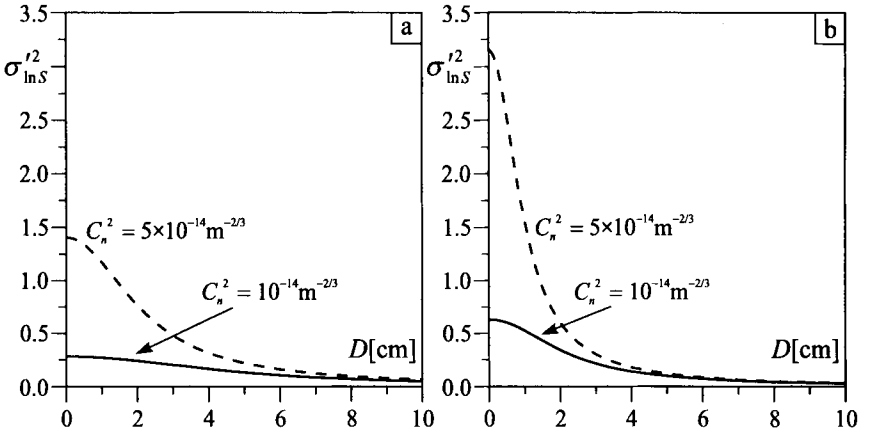


Fig. 15. Aperture-averaged scintillation index as a function of receiver lens diameter [$\hat{r} = 1$, $\zeta_S = 100$, $w_o = 2.5$ cm, $z = 2000$m; (a) $\lambda = 1.55$ μm, and (b) $\lambda = 0.785$ μm].

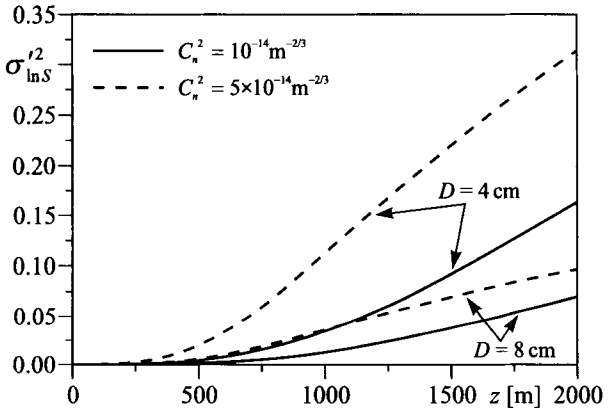


Fig. 16. Aperture-averaged scintillation index as a function of channel distance for different optical turbulence strengths and receiver lens diameters [$\hat{r} = 1$, $\zeta_S = 100$, $w_o = 2.5$ cm, $\lambda = 1.55$ μm].

operating wavelength, there is a maximum size beyond which increasing the receiver aperture does not lead to significant increases in scintillation reduction.

The aperture-averaged scintillation index $\sigma^{\prime 2}_{\ln S}\big|_{\rho=w(z)/2}$ is shown in Fig. 16 as a function of channel distance for 4 cm and 8 cm receiver diameters. In both cases scintillations increase with path length, and for the 4 cm diameter receiver these can be quite severe. Doubling the receiver aperture size decreases scintillations by about a factor of two.

3.3.4. Beam Wander

When a laser beam propagates through optical turbules that are smaller than the diameter of the beam, the quality of the beam wavefront is degraded resulting in laser beam scintillation (see Fig. 13). However, optical turbules larger than the beam diameter instead tend to cause the entire beam to be deflected, resulting in a random wander of the beam centroid about the aimpoint called "beam wander". Beam wander and scintillation are both important factors for laser communication satellite uplinks, since the laser beam diameter is smaller than many of the optical turbules encountered. Satellite downlinks are less affected by beam wander because the laser beam diameter will have spread somewhat before entering the Earth's atmosphere [69]. For near-horizontal paths beam wander can be a significant factor in the weak fluctuation regime, which is defined by the condition

$$\sigma_1^2 < 1 \quad \text{and} \quad \sigma_1^2 \, \hat{z}_{\text{rec}}^{5/6} < 1 \, . \tag{79}$$

Beyond the weak fluctuation regime, the magnitude of beam wander quickly begins to diminish as the beam wavefront loses coherence due to the accumulation of wavefront distortions. The effects of beam wander can be compensated by the use of an adaptive tip-tilt mirror system as discussed in "Free-space laser communications with adaptive optics: Atmospheric compensation experiments" by Weyrauch and Vorontsov (DOI: 10.1007/s10297-005-0033-5) in this publication.

If the beam footprint is recorded over a sufficiently long period of time, the resulting beam width is the summation of the short-term beam width in turbulence $w(z)$ given by Eq. (66), plus any additional widening due to random beam motion about the centroid position [70]:

$$w(z)_{\text{long}} = w(z) + \langle \beta_w^2 \rangle \, , \qquad \langle \beta_w^2 \rangle = 1.52 \, C_n^2 \, z^3 \, w_o^{-1/3} . \tag{80}$$

In Eq. (80) the term $\langle \beta_w^2 \rangle$ is the mean square lateral displacement of the beam from its centroid position.

3.3.5. Bit Error Rate Determination for a Direct-Detection Binary Optical Communication Link

The most important question for a laser communication system operated in the presence of turbulence is, how does optical turbulence affect the bit error rate (BER) of the system? Here we review a model for a direct-detection binary optical communication link consisting of a laser transmitter, atmospheric channel containing optical turbulence, and a maximum likelihood receiver [67]. The effects of the atmospheric channel are described by the aperture-averaged scintillation index $\sigma_{\ln S}'^2$. The maximum likelihood receiver consists of an avalanche photodiode (APD) photodetector, a matched filter implemented in the form of a moving integrator, and a clocked comparator. The clocked comparator performs a threshold test at the bit interval boundaries.

Assume the photodetector is an APD that uses current gain to reduce the effects of thermal (Boltzmann) noise, which correspondingly increases noise of the output photocurrent above the Poisson shot noise limit. For an average current gain G this excess noise is characterized by the factor F [71]:

$$F = k_{\text{eff}}G + (1 - k_{\text{eff}})(2 - 1/G), \tag{81}$$

where k_{eff} is the ratio of the ionization coefficient for holes to electrons. If the APD were instead a p-i-n photodiode, the value of both the current gain G and excess noise factor F would be unity.

When more than a few hundred photons are absorbed during a single bit interval, the APD output photocurrent can be modeled as a Gaussian stochastic process. The probability density function (PDF) for the output of the moving integrator at $t = T_b$ is described by the Gaussian PDF:

$$p(y_{T_b}) = \frac{1}{\sqrt{2\pi\sigma^2_{y_{Tb}}}} \exp\left(\frac{-(y_{T_b} - < y_{T_b} >)^2}{2\sigma^2_{y_{Tb}}}\right). \tag{82}$$

In the absence of atmospheric fading the mean $\langle y_{T_b}\rangle$ and variance $\sigma^2_{y_{T_b}}$ in Eq. (82) are given by [71]

$$\begin{aligned} \langle y_{T_b}\rangle &= \tfrac{e\eta}{hf} G P_{\text{opt}} = eGn_\phi, \\ \sigma^2_{y_{T_b}} &= 2B\left(\tfrac{e^2\eta}{hf} F G^2 P_{\text{opt}} + \tfrac{2K_B T}{R_L}\right), \end{aligned} \tag{83}$$

where $n_\phi = \eta P_{\text{opt}}/hf$ is photon flux absorbed by the detector in photons per second, $2B = 1/T_b$ is effective bandwidth (data rate) of the moving integrator, η is quantum efficiency (probability that an incident photon is absorbed), e is electric charge in coulombs, hf is the photon energy, and P_{opt} is the instantaneous optical power incident on the receiver in watts. The quantity $2K_BT/R_L$ represents thermal noise of the load resistor as seen by the APD, where $K_B = 1.38 \times 10^{-23}$ J/K is Boltzmann's constant, T is effective noise temperature in K, and R_L is load resistor input to the amplifier in ohms. For simplicity, background light and photodetector dark current are neglected.

Both experimental and theoretical evidence have shown that, in the weak fluctuation regime, turbulence-induced optical power fluctuations can be considered log-normal with a PDF given by [68,72]:

$$p(S)dS = \frac{1}{\sqrt{2\pi\sigma'^2_{\ln S}}} \exp\left(\frac{-(\ln S + \frac{1}{2}\sigma'^2_{\ln S})^2}{2\sigma'^2_{\ln S}}\right)\frac{1}{S}dS, \tag{84}$$

where S is a dimensionless random variable describing optical turbulence-induced fading in the received signal, and the average of $\ln S$ is $\langle \ln S\rangle = -1/2\,\sigma'^2_{\ln S}$.

For on-off signaling, the optical power levels at the receiver can be represented under the two signaling hypotheses as either a digital 1 (ON state) given by

$$P^1_{\text{opt}} = S\langle P_{\text{opt}}\rangle, \tag{85}$$

or a digital 0 (OFF state) given by

$$P^0_{\text{opt}} = \varepsilon S\langle P_{\text{opt}}\rangle, \tag{86}$$

where ε represents residual laser light in the OFF state due to the source modulator having an imperfect extinction ratio. Background light and photodetector dark current noise have been ignored. The mean and variance of the moving integrator output representing the Gaussian APD photocurrent under the H1 (ON state) are

$$\begin{aligned}\langle y_{T_b}^1 \rangle &= eGS\tfrac{\eta}{hf}\langle P_{\text{opt}}\rangle = eGSn_\phi \\ \sigma_1^2 \equiv \text{Var}(y_{T_b}^1) &= 2B\left(\tfrac{e^2\eta}{hf}FG^2S\langle P_{\text{opt}}\rangle + \tfrac{2K_BT}{R_L}\right),\end{aligned} \tag{87}$$

and under H0 (OFF state):

$$\begin{aligned}\langle y_{T_b}^0 \rangle &= eG\varepsilon S\tfrac{\eta}{hf}\langle P_{\text{opt}}\rangle = eG\varepsilon Sn_\phi \\ \sigma_0^2 \equiv \text{Var}(y_{T_b}^0) &= 2B\left(\tfrac{e^2\eta}{hf}FG^2\varepsilon S\langle P_{\text{opt}}\rangle + \tfrac{2K_BT}{R_L}\right).\end{aligned} \tag{88}$$

In Eqs. (87) and (88) the variance considers the effects of both excess shot noise and thermal noise from the APD load resistor.

The maximum likelihood receiver [73] consists of a circuit that implements the decision rule

$$p(y_{T_b}|H1) \underset{\text{Decide } H0}{\overset{\text{Decide } H1}{\gtrless}} p(y_{T_b}|H0), \tag{89}$$

which yields the following quadratic equation for the decision threshold value y_{Th}:

$$\begin{aligned}&\left(\frac{\sigma_1^2}{\sigma_0^2}-1\right)y_{Th}^2 + 2\left(\langle y_{T_b}^1\rangle - \frac{\sigma_1^2}{\sigma_0^2}\langle y_{T_b}^0\rangle\right)y_{Th} \\ &- \sigma_1^2\ln\left(\frac{\sigma_1^2}{\sigma_0^2}\right) + \frac{\sigma_1^2}{\sigma_0^2}\langle y_{T_b}^0\rangle^2 - \langle y_{T_b}^1\rangle^2 = 0.\end{aligned} \tag{90}$$

In the absence of fading, the average receiver probability of error (bite error rate, or BER), is given by the well-known expression:

$$\begin{aligned}P_r\,(\text{error}|\text{no fading}) &= \frac{1}{2}\left(P_{\text{FalseAlarm}} + P_{\text{Miss}}\right) \\ &= Q\left(\frac{y_{Th} - \langle y_{T_b}^0\rangle}{2\sigma_0}\right) + Q\left(\frac{y_{Th} - \langle y_{T_b}^1\rangle}{2\sigma_1}\right)\end{aligned} \tag{91}$$

where $Q(x)$ is the Gaussian tail integral defined as

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\left(\frac{-t^2}{2}\right)dt = \frac{1}{2}\text{erfc}\left(\frac{x}{\sqrt{2}}\right). \tag{92}$$

The effects of optical turbulence-induced log-normal fading on receiver BER are included by averaging Eq. (91) over the log-normal PDF (85):

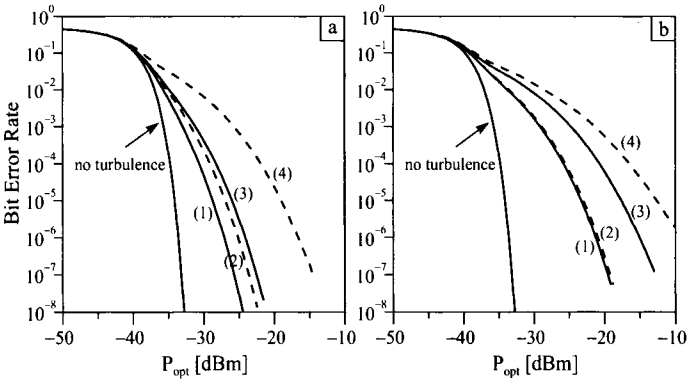


Fig. 17. BER as a function of received power for; (a) $\lambda = 1.55$ μm, $w_o = 2.5$ cm, $z = 1000$ m; and (b) $\lambda = 1.55$ μm, $w_o = 2.5$ cm, $z = 2000$ m. Left to right: (1) $D = 8$ cm, $C_n^2 = 10^{-14}$ m$^{-2/3}$; (2) $D = 8$ cm, $C_n^2 = 5\times10^{-14}$ m$^{-2/3}$; (3) $D = 4$ cm, $C_n^2 = 10^{-14}$ m$^{-2/3}$; (4) $D = 4$ cm, $C_n^2 = 5\times10^{-14}$ m$^{-2/3}$.

$$P_{\mathrm{FA}} = \int_0^\infty Q\left(\frac{y_{Th} - eG\varepsilon n_\phi}{\sqrt{2B\,(n_\phi \varepsilon e^2 F G^2 + 2K_B T/R_L)}}\right)\frac{1}{S\sqrt{2\pi\sigma_{\ln S}'^2}} \times \exp\left(\frac{-(\ln S + \frac{1}{2}\sigma_{\ln S}'^2)^2}{2\sigma_{\ln S}'^2}\right) dS$$

$$P_{\mathrm{MISS}} = \int_0^\infty Q\left(\frac{y_{Th} - eG n_\phi}{\sqrt{2B\,(n_\phi e^2 F G^2 + 2K_B T/R_L)}}\right)\frac{1}{S\sqrt{2\pi\sigma_{\ln S}'^2}} \times \exp\left(\frac{-(\ln S + \frac{1}{2}\sigma_{\ln S}'^2)^2}{2\sigma_{\ln S}'^2}\right) dS\,. \tag{93}$$

Figures 17(a,b) illustrate typical effects of optical turbulence on the average BER for a direct- detection binary optical communication link. These calculations assumed a current gain of $G = 20$, $k_{\mathrm{eff}} = 0.03$, $2B = 1/T_b = 1$ Gbps, load resistance $R_L = 50$ ohms, $T = 300$ K, quantum efficiency $\eta = 0.8$, and a 1000:1 modulator extinction ratio ($\varepsilon = 0.001$). The aperture-averaged scintillation index $\sigma_{\ln S}'^2$ characterizing optical turbulence effects for this receiver system was obtained using eqns. (76)–(78), where we assumed an average value for $\sigma_{\ln S}^2|_{\rho=w(z)/2}$. The average optical power at the receiver P_{opt} is expressed in dB milliwatts, with a 1 milliwatt optical power corresponding to 0 dBm. The optical power is obtained from the photon flux n_ϕ using the relationship $P_{\mathrm{opt}} = hf\,n_\phi\,/\eta$. We assume there are no pointing errors or beam motion/wander, and that the transmitter/receiver optical system is perfectly aligned. Once the average optical power is determined, the optical power in dBm can be obtained from

$$P_{\mathrm{opt}}\ [\mathrm{dBm}] = 10\,\log_{10}\left(P_{\mathrm{opt}}\ [\mathrm{mW}]\right). \tag{94}$$

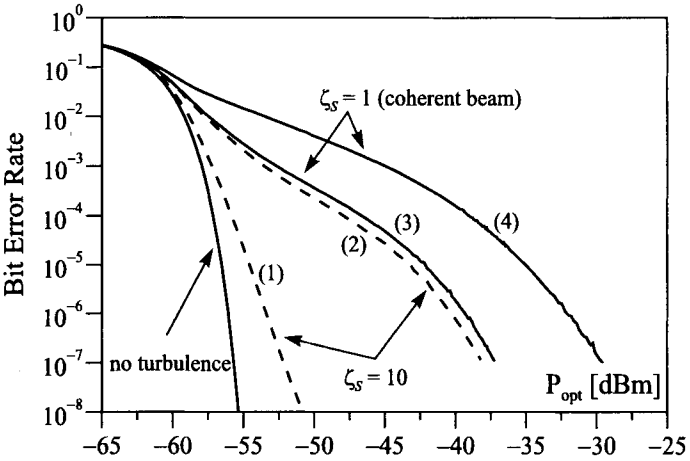


Fig. 18. Effect of optical turbulence on BER as a function of received optical power for collimated ($\hat{r} = 1$) coherent ($\zeta_S = 1$) and partially coherent ($\zeta_S = 10$) beams ($\lambda = 0.785$ μm, $w_o = 2.5$ cm, $z = 2000$ m, $D = 10$ cm). Left to right: (1) $\zeta_S = 10$, $C_n^2 = 10^{-15}$ m$^{-2/3}$ ($A\sigma_{\ln S}^2 = 0.013$); (2) $\zeta_S = 10$, $C_n^2 = 1.2\times10^{-14}$ m$^{-2/3}$ ($A\sigma_{\ln S}^2 = 0.05$); (3) $\zeta_S = 1$, $C_n^2 = 10^{-15}$ m$^{-2/3}$ ($A\sigma_{\ln S}^2 = 0.045$); (4) $\zeta_S = 1$, $C_n^2 = 1.2\times10^{-14}$ m$^{-2/3}$ ($A\sigma_{\ln S}^2 = 0.118$).

3.3.6. Reducing Optical Turbulence Effects

There are a variety of techniques available to mitigate the effects of optical turbulence. For example, increasing the size of the receiver aperture offers an effective and simple way to reduce turbulence-induced signal fades. A related approach is to use a phase diffuser as an inexpensive beam expander, thus reducing scintillations and signal fades through an "artificial" aperture averaging effect [67]. Both of these techniques reduce the correlation radius of the laser beam with respect to the receiver aperture size. As seen in Fig. 18, scintillation reduction due to "artificial" aperture averaging can directly translate to a significant reduction in BER. Another related approach is the use of multiple transceiver apertures to reduce intensity fluctuations [74]. While these three techniques can be helpful in mitigating BER losses in weak-to-moderate optical turbulence regimes, they begin to lose their effectiveness as turbulence strength increases. It then becomes necessary to incorporate adaptive optical control into the laser communication system itself [75–78]. This approach is discussed in greater detail in "Free-space laser communications with adaptive optics: Atmospheric compensation experiments" by Weyrauch and Vorontsov (DOI: 10.1007/s10297-005-0033-5) in this publication.

Acknowledgments

The authors thank Arnold Tunick for his careful review of the PAMELA model, as well as his many helpful suggestions. Jennifer C. Ricklin was an employee of the Army Research Laboratory when this work was performed.

Appendix A: Mathcad Version of PAMELA Model

PAMELA inputs

lat – latitude in degrees (positive north of the equator)
long – longitude in degrees (positive west of Greenwich)
Nd – number of day in the year [for 1 Jan, Nd = 1, etc.]
gmt – Greenwich mean time (24 hour clock – 15.20 is twenty minutes after 3 pm gmt)
tstart – define start of 24 hour period for diurnal calculations, in terms of GMT (24 hour clock)
cc – percent cloud cover in eighths (2 indicates 2/8, etc.)
Pa – atmospheric pressure in millibars
Tf – atmospheric temperature in degrees Fahrenheit
vo,v – average wind speed in m/s
hr – surface roughness length in meters, describes terrain features
h – height above ground in meters (height of wind, temperature, pressure measurements). This height will also be the height of the Cn2 estimate

Calculate Solar Insolation I, Irradiance R, Sensible Heat Flux H

d – angular fraction of a year for a particular date (degrees)
M – true solar noon (time in hours of the meridian passage, 24 hour clock)
ht – solar hour angle (degrees)
δ – solar declination
α – solar elevation angle
slz – solar zenith angle (degrees)
Tc – transmission coefficient of the atmosphere
I – Solar Insolation in W/m^2
R – Solar Irradiance in W/m^2
H – Sensible Heat flux in W/m^2

Calculate Pasquill Stability Class

cr – radiation class
cw – wind speed class
P – Pasquill stability category
L – Monin-Obukhov length in meters

Calculate Flux Profile Relationships

Φm – dimensionless wind shear
Φh – dimensionless temperature gradient
Ψm – diabatic influence function for momentum
Ustar – friction velocity
ρ – atmospheric density
Tstar – characteristic temperature
Kh – turbulent exchange coefficient for heat

Calculate Cn2

dndz – gradient for refractive index fluctuations
ε – eddy dissipation rate
M – rate of energy production due to wind shear
B – rate of energy production due to buoyancy
Cn2 – refractive index structure parameter, in m^-2/3

Model Inputs

Date, time, and location inputs

$$\text{Nd} := 151 \qquad \text{lat} := 32 \qquad \text{long} := 106 \qquad \text{tstart} := 6$$

A Greenwich mean time converter can be found at http://greenwichmeantime.com/. http://www.infoplease.com/ipa/A0001769.html gives latitudes and longitudes of major world cities.

Meteorological and Terrain Inputs

$$\text{cc} := 3 \qquad \text{v} := 2 \qquad \text{hr} := 0.1 \qquad \text{h} := 4 \qquad \text{Pa} := 1000 \qquad \text{Tf} := 85$$

$$\text{vo} := \begin{cases} 3 \text{ if } \text{v} < 3 \\ \text{v otherwise} \end{cases}$$

Change degrees Fahrenheit to Kelvin

$$\text{T} := \frac{5}{9} \cdot (\text{Tf} - 32) + 273.15$$

$$\text{T} = 302.594$$

Loop over one diurnal cycle

$$\text{gmt} := \text{tstart}, \text{tstart} + 0.1 .. \text{tstart} + 24.0$$

$$\text{tstart} = 6 \qquad \text{tstart} + 0.1 = 6.1 \qquad \text{tstart} + 24 = 30$$

Calculate Solar Insolation I (W/m^2)

$$\text{d} := (\text{Nd} - 1) \cdot \frac{360}{365.24}$$

$$\gamma := 279.93 + \text{d}$$

$$\text{M} := 12 + 0.12357\sin(\text{d}\cdot\text{deg}) - 0.004289\cos(\text{d}\cdot\text{deg}) + .153809\sin(2\cdot\text{d}\cdot\text{deg}) + 0.060783\cos(2\cdot\text{d}\cdot\text{deg})$$

$$\text{ht}(\text{gmt}) := 15(\text{gmt} - \text{M}) - \text{long}$$

$$\beta := \gamma + 0.4087\sin(\gamma\cdot\text{deg}) + 1.8724\cos(\gamma\cdot\text{deg}) - 0.0182\sin(2\cdot\gamma\cdot\text{deg}) + 0.0083\cos(2\cdot\gamma\cdot\text{deg})$$

$$\delta := \text{asin}(0.39785\sin(\beta\cdot\text{deg}))$$

$$\alpha(\text{gmt}) := \text{asin}(\sin(\text{lat}\cdot\text{deg})\cdot\sin(\delta\cdot\text{rad}) + \cos(\text{lat}\cdot\text{deg})\cdot(\delta\cdot\text{rad})\cdot\cos(\text{ht}(\text{gmt})\cdot\text{deg}))$$

$$\text{slz}(\text{gmt}) := 90.57 - \alpha(\text{gmt})\cdot 57.29$$

$$\text{Tc}(\text{gmt}) := 0.57 + 0.0045\,\text{slz}(\text{gmt})$$

$$\text{Itemp}(\text{gmt}) := 1353\cos(\text{slz}(\text{gmt})\cdot\text{deg})\cdot\text{Tc}(\text{gmt})^{\sec(\text{slz}(\text{gmt})\cdot\text{deg})}$$

$$\text{I}(\text{gmt}) := \begin{cases} 0 \text{ if } \text{Itemp}(\text{gmt}) < 0 \\ \text{Itemp}(\text{gmt}) \text{ otherwise} \end{cases}$$

$$\text{I}(18) = 819.659$$

Calculate Solar Irradiance R and Sensible Heat Flux H (W/m^2)

$B_{0,0} := 1.07$ $B_{5,0} := 0.67$
$B_{1,0} := 0.89$ $B_{6,0} := 0.59$
$B_{2,0} := 0.81$ $B_{7,0} := 0.45$
$B_{3,0} := 0.76$ $B_{8,0} := 0.23$
$B_{4,0} := 0.72$

b:= for $i \in 0..cc$
$x \leftarrow B_{i,0}$ if $i = cc$
x
b=0.76

R(gmt):= b·I(gmt)
H(gmt):= 0.4(R(gmt) – 100)
H(gmt):= 0.4(R(gmt) – 100) if I(gmt) > 0
–40 otherwise
H(18) = 209.176

Calculate Pasquill Stability Class

Calculate radiation class Cr

crtemp(gmt):= $x \leftarrow \frac{R(gmt)}{300}$ if H(gmt) > 0
$x \leftarrow -1$ if $H(gmt) \leq 0 \wedge cc \geq 4$
$x \leftarrow -2$ if $H(gmt) \leq 0 \wedge cc < 4$
x
cr(gmt):= 3 if crtemp(gmt) ≥ 0
crtemp(gmt) otherwise
cr(18) = 2.076

Calculate wind speed class cw

cw:= $x \leftarrow \frac{vo}{2}$ if vo ≤ 8
$x \leftarrow 4$ if vo > 8
x
Pt(gmt):= $x \leftarrow 0.5 \cdot (4 - cw + |cr(gmt)|)$ if H(gmt) ≤ 0
$x \leftarrow -0.5 \cdot (4 - cw + |cr(gmt)|)$ if H(gmt) > 0
$x \leftarrow 0$ if cc = 8
x
P(18) = –2.288

Calculate Flux Profile Relationships

Calculate Monin-Obukhov Length L

L(gmt):= $(0.004349P(gmt) + 0.003724P(gmt)^3) \cdot hr^{-(0.5034 - 0.231|P(gmt)| + 0.0325 \cdot P(gmt)^2)^{-1}}$
$\eta(gmt) := \frac{h}{L(gmt)}$
L(18) = – 13.124

Estimate dimensionless wind shear Φm

$$\Phi m(gmt):= \begin{cases} x \leftarrow (1 - 16\cdot\zeta(gmt))^{-.25} & \text{if } P(gmt) \le 0 \\ x \leftarrow 1 + 5\cdot\zeta(gmt) & \text{if } P(gmt) > 0 \\ x \end{cases}$$

$\Phi m(18)=0.642$

$$\Psi m(gmt):= \begin{cases} y \leftarrow (1 - 16\cdot\zeta(gmt))^{.25} & \text{if } P(gmt) \le 0 \\ x \leftarrow \ln\frac{1+y^2}{2} \cdot \frac{1+y^2}{2} - 2\cdot \operatorname{atan}(y) + \frac{\pi}{2} & \text{if } P(gmt) \le 0 \\ x \leftarrow -5\cdot\zeta(gmt) & \text{if } P(gmt) > 0 \\ x \end{cases}$$

$\Psi m(18)=0.6$

Estimate dimensionless temperature gradient Φh

$$\Phi h(gmt):= \begin{cases} x \leftarrow .74(1 - 9\cdot\zeta(gmt))^{-.5} & \text{if } P(gmt) \le 0 \\ x \leftarrow .74 + 4.7\cdot\zeta(gmt) & \text{if } P(gmt) > 0 \\ x \end{cases}$$

$\Phi h(18)=0.382$

Estimate friction velocity Ustar

$$Ustar(gmt):= \frac{0.4vo}{\ln\frac{h}{hr} - \Psi m(gmt)}$$

$Ustar(18) = 0.389$

Estimate characteristic temperature Tstar

$$\rho:= \frac{Pa}{2.87\cdot T}$$

$$Tstar(gmt):= \frac{-H(gmt)}{1004\rho\cdot Ustar(gmt)}$$

$Tstar(18) = 0.466$

Estimate turbulent exchange coefficients for heat Kh

$$Kh(gmt):= \frac{0.4\cdot Ustar(gmt)\cdot h}{\Phi h(gmt)}$$

$Kh(18) = 1.625$

Calculate Refractive Index Structure Parameter Cn2

Estimate gradient for refractive index fluctuations

$$dndz(gmt):= \frac{-(77.6\cdot 10^{-6})\cdot Pa\cdot Tstar(gmt)\cdot\Phi h(gmt)}{0.4h\cdot T^2}$$

$dndz(18) = 9.435 \times 10^{-8}$

Estimate eddy dissipation rate ϵ

$$\epsilon(\text{gmt}) := \frac{\text{Ustar(gmt)}^3}{0.4\text{h}} \cdot (\Phi\text{m}(\text{gmt}) - \zeta(\text{gmt}))$$

$$\epsilon(18) = 0.035$$

$$\text{Cn2(gmt)} := \frac{2.8 \cdot \text{Kh(gmt)} \cdot \text{dndz(gmt)}^2}{\epsilon(\text{gmt})^{\frac{1}{3}}}$$

$$\text{Cn2}(18) = 1.242 \times 10^{-13}$$

Appendix B: Calculating Solar Iirradiance and Sensible Heat Flux

The incident solar irradiance R and sensible heat flux H can be estimated from the latitude, longitude, date, time of day, and cloud cover as follows. The solar elevation angle (sun's zenith distance) α can be obtained from [79]

$$\sin\alpha = \sin\phi \ \sin\delta + \cos\phi \cos\delta \cos h_s, \tag{B1}$$

where ϕ is the latitude, δ is the solar declination, and h_s is the solar hour angle. The following method for calculating the solar declination angle accounts for the slight asymmetry due to the ellipticity of the Earth's orbit:

$$\sin\delta = \sin(23.4^\circ)\sin\beta, \tag{B2}$$

where β is given in degrees by

$$\beta = \gamma + 0.4087\sin\gamma + 1.8724\cos\gamma - 0.0182\sin(2\gamma) - 0.0083\cos(2\gamma). \tag{B3}$$

In (B3)

$$\gamma = 279.93 + d\,. \tag{B4}$$

The angle d is the angular fraction of a year for a particular date, and is given by

$$d = (\text{number of days in a year} - 1) \times \left(\frac{360}{365.24}\right)\,. \tag{B5}$$

For example, on January 1st $d = 0$, and on March 21st $d = 79 \times (0.985) = 77.815$. The solar hour angle h_s is a measure of the longitudinal distance from the sun to the point of interest on the ground, and can be obtained from

$$h_s = 15(T - M) - \eta. \tag{B6}$$

where T is Greenwich mean time (GMT), η is longitude (positive west of Greenwich), and M is the time in hours of the meridian passage (true solar noon) given by

$$M = 12 + 0.12357\sin d - 0.004289\cos d + 0.1538\sin(2d) + .06078\cos(2d). \tag{B7}$$

For time periods sufficiently short that all quantities can be considered constant, the solar irradiance per unit time incident on a unit horizontal area can be found from [79]

$$I = 1353 \text{ W/m}^2 \cos(90.57 - \alpha) a^{\sec(90.57 - \alpha)}, \tag{B8}$$

where a is the transmission coefficient of the atmosphere:

$$a = 0.57 + 0.0045(90.57 - \alpha). \tag{B9}$$

The value of the solar irradiance R is obtained from the incident solar irradiance I weighted to account for cloud cover [80]:

$$R = bI, \tag{B10}$$

where

$$\begin{array}{ll} \% \text{ cloud cover}: & 0 \quad 1/8 \quad 2/8 \quad 3/8 \quad 4/8 \quad 5/8 \quad 6/8 \quad 7/8 \quad 1 \\ \text{b}: & 1.07 \quad 0.89 \quad 0.81 \quad 0.76 \quad 0.72 \quad 0.67 \quad 0.59 \quad 0.45 \quad 0.23 \end{array}. \tag{B11}$$

The sensible heat flux in W/m^2 can then be estimated using [81]

$$H = 0.4\,(bI - 100). \tag{B12}$$

Appendix C: Calculation of Aperture-Averaged Scintillation Index Using Mathcad

This program calculates the aperture-averaging factor A as a function of the receiver aperture diameter D, the scintillation index as a function of the radial distance from the beam center ρ, and the aperture-averaged scintillation index for fixed D and ρ.

This approach includes effects due to source coherence. [see J. C. Ricklin and F. M. Davidson, "Atmospheric optical communication with a Gaussian Schell beam," J. Opt. Soc. Am. A, 20(5), 856–866(2003)].

L:= 1000 path length (m)
λ:= 785· 10^{-6} wavelength (m)
wo:= 0.025 transmitter beam size (radius) (m)
k:= $\frac{2\cdot\pi}{\lambda}$ wavenumber
ζs:= 1 source coherence: 1 for coherent beam, >1 for partially coherent beam
Cn2:= 1.0·10^{-14}

Rytov variance:
σ1:= 1.23·Cn2·$k^{\frac{7}{6}}$·$L^{\frac{11}{6}}$
σ1 = 0.44

focusing parameter:
r_hat:= 1 r_hat < 1 focused beam
r_hat = 1 collimated beam, plane wave
r_hat > 1 divergent beam (spherical wave is limiting case at infinity)

diffractive parameter:

$z\_hat := \frac{2 \cdot L}{k \cdot wo^2}$

$D := 0.001, 0.0015..0.5$ diameter (m) of receiving aperture

$\rho o := (0.55 \cdot Cn2 \cdot k^2 \cdot L)^{\frac{-3}{5}}$ coherence length of a propagating spherical wave

$\phi := \frac{r\_hat}{z\_hat} - \frac{z\_hat \cdot wo^2}{\rho o^2}$

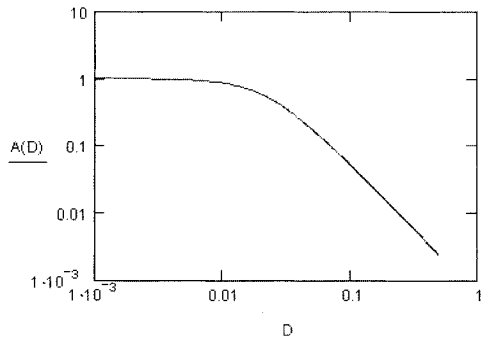
$\zeta := \zeta s + \frac{2 \cdot wo^2}{\rho o^2}$ global coherence parameter

beam size (radius) at the receiver:

$w := wo \cdot (r\_hat^2 + \zeta \cdot z\_hat^2)^{0.5}$

Aperture-averaging Factor A:

$$A(D) := \frac{16}{\pi} \int_0^1 \exp \frac{-D^2 \cdot x^2}{\rho o^2} \cdot 2 + \frac{\rho o^2}{wo^2 \cdot z\_hat^2} - \frac{\phi^2 \cdot \rho o^2}{w^2} \cdot \operatorname{acos}(x) - x \cdot (1 - x^2)^{0.5} \cdot x dx$$



Scintillation Index as a Function of ρ

We will need to calculate the beam parameters without considering optical turbulence effects because these are included by the Rytov variance. Only partial coherence effects are included. We need to use the receiver parameters rather than the transmitter parameters.

$\phi s := \frac{r\_hat}{r\_hat}$ no turbulence effects

$RS := \frac{L \cdot (r\_hat^2 + \zeta s \cdot z\_hat^2)}{z\_hat \cdot \phi s - r\_hat^2 - \zeta s \cdot z\_hat^2}$ beam focusing parameter at L (receiver)

$ws := wo \cdot (z\_hat^2 + \zeta s \cdot z\_hat^2)^{0.5}$ beam size at the receiver

$ws = 0.027$

$\Lambda := \frac{2 \cdot L}{k \cdot w^2}$ beam diffractive parameter (receiver)

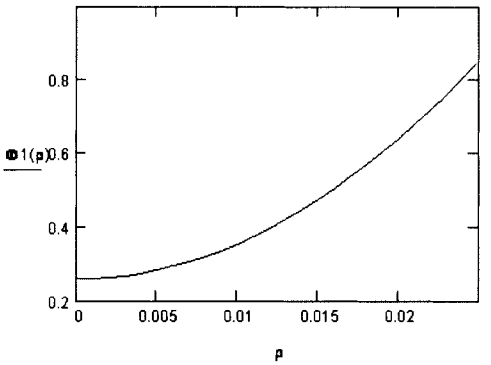
$\Theta := 1 + \frac{L}{Rs}$ beam focusing parameter (receiver)

Calculate Log-Intensity Variance as a function of radial distance from the beam center (ρ=0)

[from W. B. Miller, J. C. Ricklin, and L. C. Andrews, "Log-amplitude variance and wave structure function: a new perspective for Gaussian beams," JOSA A 10(4), 661-672 (1993); modified as described above in Ricklin et al].

$$\rho := 0, 0.001..0.025$$

$$\Phi 1(\rho) := 4.42 \cdot \sigma 1 \cdot \Lambda^{5/6} \cdot \frac{\rho^2}{ws^2} + 3.86 \cdot \sigma 1 \cdot 0.4 \cdot (1 + 2 \cdot \Theta)^2 + 4 \cdot \Lambda^{25/12} \cdot \cos\frac{5}{6} \cdot \operatorname{atan}\frac{1+2\cdot\Theta}{2\cdot\Lambda} - \frac{11\cdot\Lambda^{5/6}}{16}$$



Calculation of Aperture-Averaged Log-Intensity Variance for Fixed Receiver Aperture size

D := 0.01 receiver aperture diameter in m

$$A := \frac{16}{\pi}\int_0^1 \exp\frac{-D^2\cdot x^2}{\rho o^2}\cdot 2 + \frac{\rho o^2}{wo^2\cdot z\_hat^2} - \frac{\phi^2\cdot\rho o^2}{w^2}\cdot \operatorname{acos}(x) - x\cdot(1-x^2)^{0.5}\cdot x\,dx$$

$\rho := 0.25 \cdot ws$ radial location on beam (how far from the beam's center)

$$\Phi 1 := 4.42 \cdot \sigma 1 \cdot \Lambda^{5/6} \cdot \frac{\rho^2}{ws^2} + 3.86 \cdot \sigma 1 \cdot 0.4 \cdot (1 + 2 \cdot \Theta)^2 + 4 \cdot \Lambda^{25/12} \cdot \cos\frac{5}{6} \cdot \operatorname{atan}\frac{1+2\cdot\Theta}{2\cdot\Lambda} - \frac{11\cdot\Lambda^{5/6}}{16}$$

A = 0.855

$\Phi 1 = 0.302$

Aperture-averaged scintillation index:

$A \cdot \Phi 1 = 0.258$

References

1. R.M. Goody and Y.L. Yung, *Atmospheric Radiation: Theoretical Basis*, 2nd ed. (Oxford University Press, 1989).
2. M.E. Thomas and D.D. Duncan, Atmospheric Transmission, *The Infrared and Electro-Optical Systems Handbook*, Vol. 2 (SPIE Press, 1993).
3. A.S. Jursa, Ed., *Handbook of Geophysics and the Space Environment* (Air Force Geophysics Laboratory, 1985).
4. G.P. Anderson, F.X. Kneizys, J.H. Chetwind, J. Wand, M.L. Hoke, L.S. Rothman, L.M. Kimball, R.A. McClatchey, E.P. Shettle, S.A. Clough, W.O. Gallery, L.W. Abreu, J.E.A. Selby, "FASCODE/MODTRAN/LOWTRAN: Past/Present/Future", 18th Annual Review Conference on Atmospheric Transmission Models, 1995.
5. ONTAR Corporation (http://www.ontar.com/).
6. K.T. Whitby and B. Cantrell, "Atmospheric aerosols—characteristics and measurements," Proc. Intern. Conf. on Environ. Sensing and Assessment, Las Vegas, Nevada, 14–19 September (IEEE, New York, 1976).

7. H.C. van de Hulst, *Light Scattering by Small Particles* (John Wiley & Sons, 1957).

8. C.F. Bohren and D.R. Huffman, *Absorption and Scattering of Light by Small Particles* (John Wiley & Sons, 1983).

9. D. Deirmendjian, *Electromagnetic Scattering on Spherical Polydispersions* (Elsevier, 1969).

10. E.J. McCartney, *Optics of the Atmosphere; Scattering by Molecules and Particles,* (John Wiley & Sons, 1976).

11. S.G. Gathman, "Optical properties of the marine aerosol as predicted by the Navy aerosol model," Opt. Eng. **22**, 57–62 (1983).

12. A.M.J. van Eijk, "The ANAM-3.0 development," (TNO-The Netherlands Organisation for Applied Scientific Research Report FEL-01-C187, 2001).

13. A. Deepak and H.E. Gerber, "Report of the experts meeting on aerosols and their climatic effects", WCP-55, 1983.

14. M. Hess, P. Koepke, I. Schult, "Optical Properties of Aerosols and Clouds: The Software Package OPAC", Bull. Amer. Met. Soc. **79**, 831–844 (1998).

15. E.P. Shettle and R.W. Fenn, "Models for the aerosols of the lower atmosphere and the effects of humidity variations on their optical properties", AFGL-TR-79-0214, 1979.

16. G.A. d'Almeida, P. Koepke, and E.P. Shettle, *Atmospheric Aerosols: Global Climatology and Radiative Characteristics* (A. Deepak Publishing, 1991)

17. A.N. Kolmogorov, "The local structure of turbulence in an incompressible viscous fluid for very large Reynolds numbers," C. R. Acad. Sci. U.R.S.S. **30**, 301–305 (1941).

18. A.N. Kolmogorov, "Dissipation of energy in the locally isotropic turbulence," C. R. Acad. Sci. U.R.S.S. **32**, 16–18 (1941).

19. A.M. Obukhov, "On the energy distribution in the spectrum of a turbulent flow, C. R. Acad. Sci. U.R.S.S. **32**, 19–21 (1941).

20. A.M. Obukhov, "On the energy distribution in the spectrum of a turbulent flow," Izv., Akad. Nauk SSSR, Ser. Geogra. Geofiz. **5**, 453–466 (1941).

21. A.M. Yaglom, "Laws of small-scale turbulence in atmosphere and ocean (in commemoration of the 40th anniversary of the theory of locally isotropic turbulence)," Izv., Atmos. Oceanic Phy. **17**, 919–935 (1981).

22. A.N. Kolmogorov, "A refinement of previous hypotheses concerning local structure of turbulence in a viscous incompressible fluid at high Reynolds number," J. Fluid Mech. **13**, 82–85 (1962).

23. V.I. Tatarskii, *The Effects of the Turbulent Atmosphere on Wave Propagation* (National Technical Information Service, Springfield, Va., 1971).

24. R.J. Hill and S.F. Clifford, "Modified spectrum of atmospheric temperature fluctuations and its application to optical propagation," J. Opt. Soc. Am. **68**, 892–899 (1978).

25. L.C. Andrews, "An analytical model for the refractive index power spectrum and its application to optical scintillations in the atmosphere," J. Mod. Opt. **39**, 1849–1853 (1992).

26. G.D. Nastrom and F.D. Eaton, "The coupling of gravity waves and turbulence and turbulence at White Sands, New Mexico from radar observations," J. Appl. Meteorol., **32** (1), 81–87 (1993).

27. G.D. Nastrom and F.D. Eaton, "Variations of winds and turbulence seen by the 50-MHz radar at White Sands Missile Range, New Mexico," J. Appl. Meteorol., **34** (10), 2135–2148 (1995).

28. M.C. Roggemann and B.M. Welsh, *Imaging Through Turbulence* (CRC Press, Boca Raton, FL, 1996), p. 320.

29. J.W. Hardy, *Adaptive Optics for Astronomical Telescopes* (Oxford University Press, New York, NY, 1998), p. 448.

30. Y. Han Oh, J.C. Ricklin, E. Oh, S. Doss-Hammel and F.D. Eaton, "Estimating optical turbulence effects on free-space laser communication: modeling and measurements at ARL's A_LOT facility," SPIE, **5550**, 247–255 (2004).

31. W.B. Miller, J.C. Ricklin and W.J. Stewart, *An Optical Turbulence Code for the Surface Boundary Layer* (Army Research Laboratory, ASL-TR-0220, 1987).

32. W.B. Miller and J.C. Ricklin, *A Module for Imaging Through Optical Turbulence* (Army Research Laboratory, ASL-TR-0221-27, 1990).

33. H. Rachele and A. Tunick, "Energy balance model for imagery and electromagnetic propagation," J Appl Meteorol, **33**, 964–976 (1994).

34. R.W. Smith, J.C. Ricklin, K.E. Cranston and J.P. Cruncleton, "Comparison of a model describing propagation through optical turbulence (PROTURB) with field data," SPIE, **2222**, 780–789 (1994).

35. A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978), Vol. II, Appendix C.

36. J.A. Businger, "Turbulent Transfer in the Atmospheric Surface Layer," in *Workshop on Micrometeorology*, D. Haugen, Ed. (American Meteorological Society, Boston, MA, 1973).

37. N. Busch, "On the Mechanics of Atmospheric Turbulence," in *Workshop on Micrometeorology*, D. Haugen, Ed. (American Meteorological Society, Boston, MA, 1973).

38. H.A. Panofsky and J.A. Dutton, *Atmospheric Turbulence: Models and Methods for Engineering Applications* (John Wiley & Sons, New York, 1984).

39. M. Liu, D. Durran, P. Mundkur, M. Yocke and J. Ames, *The Chemistry, Dispersion and Transport of Air Pollutants Emitted From Fossil Fuel Plants in California: Data Analysis and Emission Impact Model* (National Technical Information Service, PB-264-822, 1976).

40. F.V. Hansen, *A Fractional Stability Category Scheme* (Atmospheric Sciences Laboratory Technical Report ASL-TR-0275, 1990).

41. E.C. Kung, "Climatology of aerodynamic roughness parameter and energy dissipation in the planetary boundary layer of the northern hemisphere," in *Studies of the Effects of Variations in Boundary Conditions on the Atmospheric Boundary Layer* (Annual Report, Contract DA-36-039-AMC-00878, Dept. of Meteorology, Univ. of Wisconsin, Madison, Wisconsin, 1963).

42. *Characterization of Wind Speed in the Lowest Layers of the Atmosphere Near the Ground: Strong Winds* (Engineering Sciences Data Unit Ltd, London, UK, EDSU 72026, 1972).

43. L.O. Myrup and J. Ranziere, *A Consistent Scheme for Estimating Diffusivities to be Used in Air Quality Models* (NTIC, Springfield, VA, PB-272-284, 1976).

44. A.S. Smedman-Hogstrom and U. Hogstrom, "Practical method of determining wind frequency distributions for the lowest m from routine meteorological data," J. Appl. Meteorol., **17** (7), 942–954 (1977).

45. C.D. MacArthur and P.A. Hains, *The Roughness Length Associated With Regions of Heterogeneous Vegetation and Elevation* (Army Research Laboratory, ASL-CR-82-0206-1, 1982).

46. S.F. Clifford, "The classical theory of wave propagation in a turbulent medium," in *Laser Beam Propagation in the Atmosphere*, J. Strohbehn, ed. (Springer, New York, 1978).

47. A.S. Monin and A.M. Yaglom, *Statistical Fluid Mechanics* (MIT Press, Cambridge, MA, 1971). V.L. Mironov, *Laser beam propagation in turbulence atmosphere* (Nauka, Moscow, 1981), p. 246.

48. V.L. Mironov, *Laser beam propagation in turbulence atmosphere* (Nauka, Moscow, 1981), p. 246.

49. A.S. Gurvich, A.I. Kon, V.L. Mironov, and S.S. Khmelevtsov, *Laser radiation in turbulent atmosphere* (Nauka, Moscow, 1976), p. 227.

50. V.A. Banakh, ed., *High-power laser beams in a randomly inhomogeneous atmosphere* (SO RAN Publ., Novosibirsk, 1998), p. 341.

51. M.E. Gracheva and A.S. Gurvich, "Simple models of turbulence," Izv. Akad. Nauk SSSR, Fiz. Atm. Okeana **16**(10), 1107–1111 (1980).

52. M.G. Miller and P.L. Zieske, *Turbulence environmental characterization* (Rome Air Development Center, Griffiss Air Force Base, NY, RADC-TR-79-131, 1979).

53. R.R. Parenti and R.J. Sasiela, "Laser-guide-star systems for astronomical applications," J. Opt. Soc. Am. A, **11**(1), 288–309 (1994).

54. R.E. Hufnagel, "Variations of atmospheric turbulence," in *Digest of topical meeting on optical propagation through turbulence* (Optical Society of America, Washington, D.C., 1974), p. WA1.

55. G.C. Valley, "Isoplanatic degradation of tilt correction and short-term imaging systems," Appl. Opt. **19**(4), 574–577 (1980).

56. B.G. Zollars, "Atmospheric-turbulence compensation experiments using synthetic beacons," in *Selected Papers on Adaptive Optics for Atmospheric Compensation* (SPIE Press, 1994), Vol. MS92, pp. 564–588.

57. D.P. Greenwood, "Bandwidth specification for adaptive optics systems," J. Opt. Soc. Am., **67**(3), 390–393 (1977).

58. T.E. Van Zandt, , J.L. Green, K.S. Gage, and W.L. Clark, Vertical profiles of refractivity turbulence structure constant: Comparison of observations by the Sunset radar with a new theoretical model, Radio Science, **13**, 819–829 (1978).

59. T.E. Van Zandt, K.S. Gage, and J.M. Warnock, An improve model for the calculation of profiles of wind, temperature, and humidity, Preprints Twentieth Conference on Radar Meteorology, American Meteorological Society, Boston MA, 129–135 (1981).

60. J.M. Warnock, T.E. Van Zandt, and J.L. Green, A statistical model to estimate mean values of parameters of turbulence in the free atmosphere, Preprints 7th symposium on turbulence and diffusion, Nov 12–15, 1985, Boulder, CO, AMS, Boston, MA, 156–159.

61. E.M. Dewan, R.E. Good, B. Beland, and J. Brown, *A model for Cn2 (optical turbulence) profiles using radiosonde data*, (PL-TR-93-2043, Environmental Research Papers, No. 1121, Phillips Laboratory, Hanscom AFB, MA, 1993).

62. W.B. Miller, J.C. Ricklin and L.C. Andrews, "Log-amplitude variance and wave structure function: a new perspective for Gaussian beams," J. Opt. Soc. Am. A, **10** (4), 661–672 (1993).

63. L.C. Andrews, W.B. Miller and J.C. Ricklin, "Geometrical representation of Gaussian beams propagating through complex optical systems," Appl. Opt., **32** (30), 5918–5929 (1993).

64. L.C. Andrews and R.L. Phillips, *Laser Beam Propagation Through Random Media* (SPIE Press, Bellingham, 1998).

65. L.C. Andrews, R.L. Phillips, and C.Y. Hopen, *Laser Beam Scintillation With Applications* (SPIE Press, Bellingham, 2001).

66. J.C. Ricklin and F.M. Davidson, "Atmospheric turbulence effects on a partially coherent Gaussian beam: implications for free-space laser communication," J. Opt. Soc. Am. A, **19** (9), 1794–1803 (2002).

67. J.C. Ricklin and F.M. Davidson, "Atmospheric optical communication with a Gaussian Schell beam," J. Opt. Soc. Am. A, **20** (5), 856–866 (2003).

68. V.I. Tatarskii, *Wave Propagation in a Turbulent Medium* (McGraw-Hill Book Company, Inc., New York, 1961).

69. F. Dios, J.A. Rubio, A. Rodriguez and A. Comeron, "Scintillation and beam-wander analysis in an optical ground station-satellite uplink," Appl. Opt., **43** (12), 3866–3873 (2004).

70. R.L. Fante, "Electromagnetic beam propagation in turbulent media: an update," Proc. IEEE, **68** (11), 1424–1443 (1980).

71. R.J. McIntyre, "The distribution of gains in uniformly multiplying avalanche photodiodes," IEEE Trans. Electron. Dev., **19** (6), 703–713 (1972).

72. J.H. Churnside and C.M. McIntyre, "Averaged threshold receiver for direct detection of optical communications through the lognormal atmospheric channel," Appl. Opt., **16** (10), 2669–2676 (1977).

73. R.M. Gagliardi and S. Karp, *Optical Communications* (Wiley-Interscience, New York, 1995).

74. A.R. Weeks, J. Xu, R.R. Phillips, L.C. Andrews, C.M. Stickley, G. Sellar, J.S. Stryjewsky and J.E. Harvey, "Experimental verification for an eight-element multiple-aperture equal-gain coherent laser receiver for laser communications," Appl. Opt., **37** (21), 4782–4788 (1998).

75. R.K. Tyson, "Adaptive optics and ground-to-space laser communications," Appl.d Opt., **35** (19), 3640–3646 (1996).

76. B.M. Levine, E.A. Martinsen, A. Wirth, A. Jankevics, M. Toledo-Quinones, F. Landers and T.L. Bruno, "Horizontal line-of-sight turbulence over near-ground paths and implications for adaptive optics corrections in laser communications," Appl. Opt., **37** (21), 4553–4560 (1998).

77. R.K. Tyson, "Bit-error rate for free-space adaptive optics laser communications," J. Opt. Soc. Am. A, **19** (4), 753–758 (2002).

78. R.K. Tyson, "Indirect measurement of a laser communications bit-error-rate reduction with low-order adaptive optics," Appl. Opt., **42** (21), 4239–4243 (2003).

79. R.J. List, *Smithsonian Meteorological Tables* (Smithsonian Institute, Washington, DC, 1951).

80. F.B. Smith, "A Scheme for estimating the vertical dispersion of a plume from a source near the ground," *Proc. of the Third Meeting of the Expert Panel on Air Pollution Modeling* (NATO Committee on the Challenges of Modern Society, Paris, France, 2–3 October, 1972).

81. F.B. Smith, "The relation between Pasquill stability P and Kazanski–Monin stability (in neutral and unstable conditions," Atmos. Environ., **13** (6), 879–881 (1979).

DOI: 10.1007/s10297-005-0054-0
Originally published in J. Opt. Fiber. Commun. Rep. 2, 345–396 (2005)

Free-space laser communication performance in the atmospheric channel

Arun K. Majumdar

LCResearch, Inc.
30402 Rainbow View Drive
Agoura Hills, California 91301
Email: a.majumdar@IEEE.org

Abstract. In spite of the tremendous technical advancement of available components, the major limitation of free-space laser communication (lasercom) performance is due to the atmosphere, because a portion of the atmospheric path always includes turbulence and multiple scattering effects. Starting from a fundamental understanding of the laser communications system under diverse weather conditions, this chapter provides a comprehensive treatment of the evaluation of parameters needed for analyzing system performance. The significance of higher-order statistics of probability density functions of irradiance fluctuations due to turbulence to performance analysis is explained. Starting from link analysis, the necessary expressions relating link margin, bit-error-rate, signal-to-noise-ratio, and probability of fade statistics are presented. Results for laboratory-simulated atmospheric turbulence and multiple scattering are presented. Example numerical results for simulations of lasercom systems operating under various atmospheric conditions are presented for various scenarios such as uplink-downlink (e.g., between ground and satellite, aircraft or UAV) and horizontal (terrestrial) link. Both turbulence and multiple scattering effects have been included in the analysis with both on-off keying and pulse-position modulation schemes. Statistical estimation and computation of communication parameters presented in this chapter will be useful in designing and optimizing lasercom systems that are reliable under all weather conditions.

1. Introduction

There has been tremendous technical advancement of available components such as laser/LED transmitter, high sensitivity optical receiver offering extremely high bandwidth, efficient modulation techniques, improvement in low power consumption,

weight, and size. Inspite of many such technological development, the major limitation of free-space laser communication (lasercom) performance is the atmosphere. Atmospheric condition ultimately determines the laser communications systems performance not only of terrestrial applications but also for uplink-downlink (e.g., between ground and satellite, aircraft or UAV terminals), because a portion of the atmospheric path always includes turbulence and multiple scattering effects.

This paper provides a comprehensive theoretical treatment and analysis of the performance of a laser communication (hereafter in short, "lasercom") system under a given communication format scheme. Numerous examples are given for specific system parameters, and the necessary tools are developed to design a system fulfilling specified requirements. The performance of a lasercom system is generally quantified by the "link margin", i.e., the ratio, expressed in dB, of the signal power received to the signal power required to achieve a specified data rate with a specified acceptable probability of error. The link margin calculation is therefore essential to design an acceptable system. Atmospheric conditions affect system performance and so need to be accounted for in the calculation. A link budget model has been developed that includes dependence on the atmospheric channel that the signal passes through as well as on the transmitter and receiver stations. This model aids designers in optimizing the optical station main parameters to be able to establish a data link with adequate performance.

The main advantage in communicating, i.e., transferring information from one point to another, with optical frequencies is the potential increase in information transfer rate. This rate is directly related to the bandwidth of the modulated carrier, which is generally limited to a fixed fraction of the carrier frequency itself. Therefore, increasing carrier frequency from that of, say, RF or microwaves to that of optical waves increases the information capacity of a communication system by many orders of magnitude. A usable bandwidth at an optical frequency of $\sim$200 THz will be about 10^5 times that of a carrier in the RF range. A second advantage of using optical frequencies is the station compactness it allows. A 6-inch-diameter lens "antenna" for use with an optical frequency of 6×10^{14} Hz has an effective gain of 122 dB. At an RF frequency of 1 GHz to generate about 60 dB gain, it would require an antenna with a diameter of over 200 feet. This is due to the fact that the beam divergence is roughly proportional to λ/D, where λ is the carrier wavelength and D is the aperture diameter, and varies inversely with antenna gain. Thus, a longer wavelength necessitates a proportionately larger antenna to achieve the same gain.

Because of these advantages, optical carrier frequencies can accommodate very high data rates of 100 GHz or more. With the advent of the new technology of wavelength division multiplexing (WDM), we can further increase the data rate per optical beam. As optical technology matures, the advantages of practical free-space laser communications systems can be realized with much architecture.

We here divide lasercom systems into several distinct application classes: (a) paths between terrestrial locations, e.g. between buildings in metropolitan areas, mountain-to-mountain, horizontal link between two ground stations, (b) paths between ground and near space, e.g., ground-to-satellite (Uplink) or satellite-to-ground (Downlink), (c) paths between airborne aircraft, (d) Space-shuttle-to-ground, (e) UAV-to-ground, (f) indoor communications with paths $<\sim$100 m, usually using infrared frequencies, and (g) free-space communications in optical interconnects on optical chips, with path lengths ranging from hundreds of microns up to $\sim$ 1 cm. Besides the differences in

path lengths, these application classes also differ in data-rate regimes (multi-Gbit/s to a few Mbit/s or less), line-of-sight requirements for the links, and the impact, if any, of weather and other atmospheric effects.

The design of a complex engineering system such as a lasercom system requires a thorough understanding of the technology building blocks, taking into account of the state-of-the-art development of the subsystems and devices, as well as an understanding of the complex phenomena of the physics of optical wave propagation through the atmosphere. The ability to predict system performance is essential for the design of a practical, optimal, cost-effective system. The lasercom system must be able to establish a communication link between transmitting and receiving stations with a specified data rate and a probability of error lower than a specified bit-error rate (BER).

Section 2 of this paper describes the procedure for computing link budget and explains the equations for data rate, link margin, and BER for an atmospheric channel with absorption and scattering. Numerical results are presented to show how the relevant communication parameters depend on range and required power. Section 3 describes the theoretical background of scintillation effects on signal-to-noise-ratio (SNR), and the procedures for computing BER for on-off keying (OOK) and pulse-position modulation (PPM) schemes. This section also discusses the probability distribution of fade statistics for lasercom system. In section 4, example numerical results for different communication scenarios such as uplink, downlink, and horizontal (terrestrial) link in the presence of atmospheric turbulence and multiple scattering medium are provided. Section 5 discusses other types of examples of free-space optical and laser communications which include indoor optical communications and free-space optical interconnects. Multiple transmitters/receivers approach for lasercom is the subject of section 6, and section 7 provides some concluding remarks.

2. Basics of Laser Communication Link Analysis

2.1. Communication Channel Characterization

The primary factors characterizing an atmospheric communication channel include atmospheric attenuation (from both scattering and absorption) and scintillation. Attenuation of the optical wave (i.e., reduction of its power) by atmosphere can be caused in several ways, including absorption of light by gas molecules and Rayleigh or Mie scattering by gas molecules or aerosol particles suspended in the air. Optical transmission spectrum of the U.S. Standard Atmosphere over the various communication windows is available in the literature. [1] The individual absorption lines are caused by the vibrational-rotational absorption lines of water vapor, CO_2, CH_4, and other gases of minor constituents in the atmosphere. Scattering from aerosols and particulates such as in fog, clouds, smoke, and dust also contribute to the total attenuation of the optical beam. Some typical values of near-ground (terrestrial path) attenuations with corresponding visibilities are [2,3]:

Very clear: 0.19 dB/km (50 km visibility), fog: 6.9 dB/km (1.9 km),
moderate fog: 28.9 dB/km (500 m visibility).

Scintillation is the fluctuations in the detector signal as a result of random variations in the refractive index of the turbulent atmosphere along the channel. The amount of scintillation can change by more than an order of magnitude during the course of a day, with the most severe effects around midday. The noise produced by scintillation in the received power results in increased BER and reduced performance. Scintillation-induced fades can ultimately cause loss of the communication signal altogether. Research for the past 40 years has allowed us to estimate the effects of turbulence on optical beam irradiance fluctuations by relating them to the refractive-turbulence structure parameter, C_n^2. The turbulence strength of the atmosphere is usually expressed in terms of a scintillation index, explained in a later section. We will cover communication performance for three different types of scenarios as examples later on in this paper: horizontal (terrestrial) link, downlink (satellite-to-ground), and uplink (ground-to-satellite).

2.2. *Transmitter and Receiver System*

There are several atmospheric transmission windows within the 700–10,000-nm wavelength range that have an attenuation of <0.2 dB/km and so are nearly transparent. The majority of free-space lasercom systems are designed to operate in the windows of 780–850 and 1520–1600 nm. Around 850 nm, reliable, inexpensive, high-performance transmitter and detector components are readily available. These are often used in network and transmission equipment. A highly sensitive silicon (Si) avalanche photodiode (APD) and an advanced vertical-cavity surface-emitting laser (VCSEL) are available for operation near 850 nm. For the wavelength range of 1520–1600 nm, high quality transmitter and detector components are also readily available. InGaAs is the most commonly used detector material for this longer wavelength; it exhibits high bandwidth capability combined with a high spectral response. This material is extremely sensitive because of an internal amplification (avalanche) process. These wavelengths are also suitable with erbium-doped fiber amplifier (EDFA) technology for high-power generation (>500 mW) and high-data rate (>2.5 Gbit/s). This wavelength region is also important from a laser eye safety point of view. The American National Standard for Eye Safe Use of Lasers Institute (ANSI) standards for the maximum permissible exposure level (MPE) for a 10-sec exposure is about 1 mW/cm^2 for an 800-nm wavelength and about 100 mW/cm^2 for a 1550-nm wavelength. The longer wavelength system is thus readily designed to be eye safe. Other benefits of the 1550-nm wavelength region include a reduced solar background and reduced scattering attenuation, because of the Rayleigh dependence of inverse wavelength to the fourth power.

2.3. *Link Analysis*

The overall system performance of a lasercom is quantified using a link budget derived from the range equation, which combines attenuation and geometrical aspects to calculate the received power. The process of finding the link margin through the system link calculation is the topic of this section. The system link calculation is carried out as shown in the Fig. 1. We will show some examples of link calculations for different communication scenarios of practical importance for various communication parameters

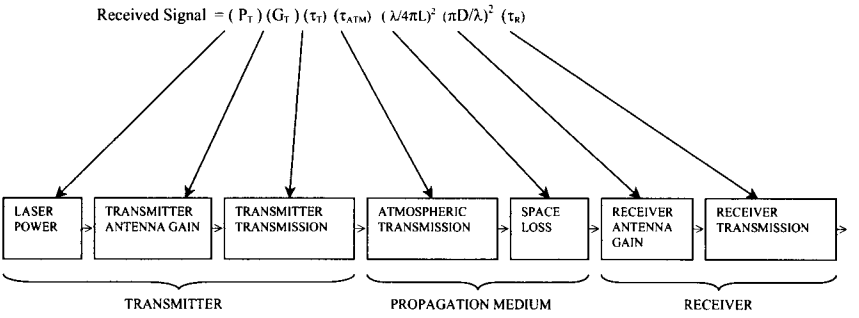


Fig. 1. Optical signal link calculations.

such as transmitter power, propagation losses, receiver sensitivity. The receiver's sensitivity determines the amount of received optical power needed to achieve the required signal-to-noise ratio (SNR) for a given expected communication performance.

The purpose of this section is to develop the parameters necessary to calculate the performance of an optical communication link. We shall consider the situation of optical propagation between points in free-space. Consider a laser transmitter antenna with gain G_T transmitting a total power P_T at the wavelength. The signal power received at the communications detector can be expressed (from the range equation) as

$$\text{Received Signal}, P_{\text{REC}} = P_T G_T \tau_T \tau_{\text{ATM}} S G_R \tau_R, \tag{1}$$

where τ_T is the transmitter optical efficiency, τ_{ATM} is the value of the atmospheric transmission at the laser transmitter wavelength, S is the free-space loss, G_R is the receiver antenna gain, and τ_R is the receiver optical efficiency. The transmitter gain, free-space loss, and receiver antenna gain are given by $G_T = 16/\theta_T^2$ (where θ_T is the full transmitting divergence angle), $S = (\lambda/4\pi L)^2$ (where L is the range), and $G_R = (\pi D/\lambda)^2$ (where D is the receiver diameter). τ_{ATM} may be written in terms of the atmospheric attenuation factor α given by $-10\log(\tau_{\text{ATM}})/L$.We can then express the received signal as

$$P_{\text{REC}}(\text{Received Signal}) = P_T G_T \tau_T \tau_{\text{ATM}} (\lambda/4\pi L)^2 (\pi D/\lambda)^2 \tau_R. \tag{2}$$

Equation (2) can be re-written as

$$P_{\text{REC}} = P_T (D^2/\theta_T^2 L^2) \tau_T \tau_{\text{ATM}} \tau_R. \tag{3}$$

If we write τ_{ATM} in terms of atmospheric attenuation factor, α (dB/km) at the wavelength, λ, the received power is then given by

$$P_{\text{REC}} = P_T (D^2/\theta_T^2 L^2) \tau_T 10^{(-\alpha L/10)} \tau_R. \tag{4}$$

Normally an optical link typically consists of two transceivers, each made up of one (or more) transmitting laser(s) and receiving photo detector(s). Transmitting

optics (telescope, lenses, mirrors) shape the transmitted laser beam which is collected by the receiver optics so that the received signal is focused onto the photo detector. The parameters of the communications are chosen so that sufficient signal from the lasers on one transceiver reaches the photo detector on the other transceiver through the atmosphere to differentiate ones (signal) and zeros (no signal) with negligible error (i.e., very small specified BER).

2.3.1. Data Rate

Given a laser transmitter power P_t, with transmitter divergence of θ_t, receiver telescope area A, transmit and receive optical efficiency τ_{opt}, the achievable data rate R can be obtained from

$$R = \frac{P_t \tau_{\text{opt}} \tau_{\text{ATM}} A}{\pi(\theta_t/2)^2 L^2 E_p N_b}, \tag{5}$$

where $E_p = hc/\lambda$ is the photon energy and N_b is the receiver sensitivity in # photons/bit.

2.3.2. Link Margin

Another important parameter in optical communications link analysis is "Link Margin", which is the ratio of available received power to the receiver power required to achieve a specified BER at a given data rate. Note that the "required" power at the receiver P_{REQ} (watts) to achieve a given data rate, R (bits/sec), and receiver sensitivity, N_b (# photons/bit), is related by $P_{\text{REQ}} = N_b R h\nu = N_b R hc/\lambda$, where ν is the frequency of the laser light of wavelength (h = Planck's constant, c = velocity of light). The parameter N_b is usually defined as receiver sensitivity and can be expressed as # photons/bit, or in dBm (taking the logarithm to the base 10 and multiplied by 10 with base as 1 mW). Finally, we can define the link margin M as

$$M = \text{Received Power/Required Power}.$$

Combining the above equations yields the expression

$$M = [P_T/(N_b R h\nu)](D^2/\theta_T^2 L^2)\tau_T 10^{(-\alpha L/10)}\tau_R. \tag{6}$$

2.3.3. Bit Error Rate in Presence of Atmospheric Absorption and Scattering

A laser beam propagating through the atmosphere is attenuated by absorption and scattering due to the presence of aerosols, dust, smoke, fog, clouds, rain, snow, and atmospheric molecules. In this section, we address only the optical propagation causing absorption and scattering. The detailed analysis of lasercom performance due to turbulence-induced atmospheric scintillation will be discussed in later section. We first calculate the received signal power for a lasercom system for a given range and extinction (combined absorption and scattering). The detector noise that comes from the background power and the inherent detector noise are then evaluated. For the transmission of both a binary 1 and a binary 0, we can then calculate the distribution of signal plus additive Gaussian noise. Finally, the error probability associated with binary detection (bit error rate) can be determined by the procedure described below.

The photocurrent induced by the received optical signal is given by

$$i = \frac{GP_{\mathrm{REC}}\lambda q}{hc}, \tag{7}$$

where P_{REC} is the received laser power given by Eq.(4), $q = 1.601 \times 10^{-19}$ C is the electronic charge, and G is the avalanche photo detector (APD) gain. For a PIN Photodiode, we take $G = 1$. The background power and the detector noise power can be written as [4]

$$P_{\mathrm{bkg}} = H_{\mathrm{bkg}} BWif A_{\mathrm{rec}} \phi_{\mathrm{det}} \eta_{\mathrm{recv}}, \tag{8}$$

where H_{bkg} is the background radiance, $BWif$ is the receiver filter's optical bandwidth, A_{rec} is the receiver area, ϕ_{det} is the detector's solid angle field-of-view (FOV), and η_{recv} is receiver optical efficiency. The detector noise power can be expressed as

$$P\,\mathrm{det} = \sqrt{B_W if} NEP\,\mathrm{det}, \tag{9}$$

where the values for the bandwidth and the noise-equivalent power (NEP) of the detector are typically supplied by the manufacturer. The mean-square fluctuations in the signal current i_{sig} can be expressed as [5,6]

$$\langle i_{\mathrm{sig\ noise}}^{\ 2} \rangle = 2qMFi_{\mathrm{sig}} BW_{\mathrm{det}}, \tag{10}$$

where F is the excess-noise factor for the APD ($F = 1$ for a PIN photodiode). We can similarly write the expressions for the mean-square fluctuations in the currents due to background and the detector noise. Assuming the mean-square fluctuations in the photo detector current due to signal, background, and detector are all independent and uncorrelated, we can calculate the total rms noise for a binary 1 or 0 in the following way:

$$\text{For transmission of binary 1 :}\sigma_1 = \sqrt{\langle i_{\mathrm{sig\ noise}}^{\ 2} \rangle + \langle i_{\mathrm{bkg\ noise}}^{\ 2} \rangle + \langle i_{\mathrm{det\ noise}}^{\ 2} \rangle}, \tag{11}$$

$$\text{For transmission of binary 0 :}\sigma_0 = \sqrt{\langle i_{\mathrm{sig\ noise}}^{2}(\eta) \rangle + \langle i_{\mathrm{bkg\ noise}}^{\ 2} \rangle + \langle i_{\mathrm{det\ noise}}^{\ 2} \rangle}, \tag{12}$$

where $i_{\mathrm{sig\ noise}}(\eta)$ reflects the extinction ratio of the laser signal and thus takes into account of the incomplete modulation of the laser from an ON state to an OFF state with a typical value [4] of modulation depth of $\sim 90\%$.

The probability of detecting either a 1 or a 0 in error can be expressed as

$$\mathrm{error1} = \int_{-\infty}^{0.5I} \frac{\exp[-(i-I)^2/2\sigma_1^2]}{\sqrt{2\pi\sigma_1^2}} di, \tag{13}$$

$$\mathrm{error0} = \int_{0.5I}^{\infty} \frac{\exp[-(i)^2/2\sigma_0^2]}{\sqrt{2\pi\sigma_0^2}} di, \tag{14}$$

The overall BER is the mean of these two probabilities:

$$\mathrm{BER} = (\mathrm{error1} + \mathrm{error0})/2. \tag{15}$$

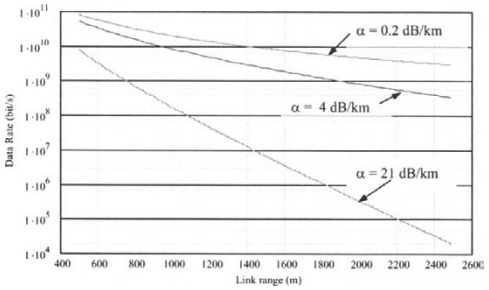


Fig. 2. Communication data rate as a function of range for three atmospheric conditions: clear ($\alpha = 0.2$ dB/km), haze ($\alpha = 4$ dB/km) and fog ($\alpha = 21$ dB/km).

2.3.4. Example Numerical Results

Data Rate Versus Range The range equation can be used to generate the communications data rate versus range for varying atmospheric conditions. Suppose a 2.5-km link is desired at 1.25 Gbit/s. Furthermore, the system has to be eye safe, so that we use a 1.55-μm laser with 70 mW power. We assume a 13-cm receiver aperture and transmitting beam divergence of 8.5 mrad FWHM. Receiver sensitivity for 1.25 Gbit/s was taken to be -36 dBm, which is equivalent to 1568 photons/bit at this wavelength and data rate. Based on these assumptions, link margins were calculated for signal attenuation due to weather events [3] such as clear (visibility = 23 km, $\alpha = 0.2$ dB/km), haze (visibility = 2 km, $\alpha = 4$ dB/km), and fog (visibility = 0.5 km, $\alpha = 21$ dB/km). The transmitter and receiver optics efficiencies are both assumed to be equal to 0.5. Data rates versus range is shown in Fig. 2. As can be seen, a data rate of 3 Gbit/s can be achieved for a range of 2.4 km under clear condition, for 1.4 km under haze condition, and for 0.6 km under fog condition. The system is desired to have communication capability for the above ranges and atmospheric conditions. Under night-time conditions, the ranges would be increased further. Increased data rate is possible if a higher power laser, or a larger aperture, or both, is used.

Available Optical Link Margin (dB) Versus Link Range Figure 3 illustrates available optical link margin (dB) versus link range for the same system parameters such as laser power, transmitter beam divergence, aperture diameter, and operating wavelength of 1.55 μm and the above visibility conditions. As seen in this figure, 10 dB of link margin is available for clear condition for the system operating over a 1.2-km span, for haze over a 0.885 km span, and for fog just under a 0.5-km span for the above conditions of 1.25 Gbit/s at 1.55 μm. The curve shows how much margin a given system has at a given range to compensate for scattering, absorption, and scintillation losses. Under system operation, parameters defining channel capacity can be varied to accommodate link availability under varying atmospheric conditions. The link margin curves fully encompass almost all basic aspects of laser communication system performance.

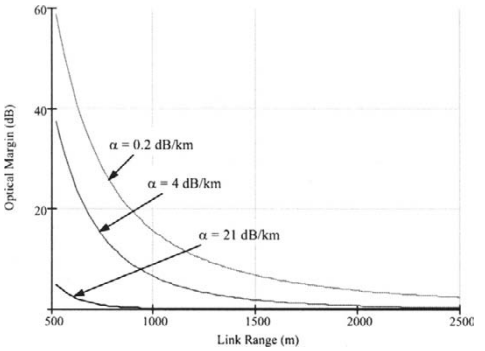


Fig. 3. Link margin versus link range for three atmospheric conditions: clear, haze and fog.

Link Budget Computation To give an idea of the various components of link budget, Table 1 shows an overall calculation of the link performance of a downlink (satellite-to-ground) laser communication system The link analysis is performed for a satellite-to-ground range of 4.83×10^5 m, transmitter laser operating at wavelength $\lambda = 0.635$ m at the data rate of 3 Gbit/s using a detector (receiver) sensitivity $N_b = 100$ photons/bit and a receiver of diameter of 1.4 m. The table shows the laser power required for a specified overall link margin of 6.0 dB for this link.

Table 1. Link analysis example of a satellite-to-ground laser communication system.

Parameter	Value/Factor	dB
Wavelength (λ)	0.635 μm	
Range (L)	4.83×10^5 m	
Data Rate	3 Gbps	
Receiver Diameter (D)	1.4 m	
Transmitter Divergence Angle (θ_T)	2.07×10^{-4} rad	
Transmitter Antenna Gain ($G_T = 16/(\theta_T)^2$)	3.73×10^8	+85.72
Transmitter Optical Loss	0.1	−10.0
Space Loss ($S = (\lambda/4\pi L)^2$)	1.09×10^{-26}	−259.61
Receiver Antenna Gain ($G_R = (\pi D/\lambda)^2$)	47.974×10^{12}	+136.81
Receiver Optical Loss	0.1	−10.0
SYSTEM LOSS		−57.08
Atmospheric Turbulence Margin		−11.30
Clear Air Transmission Loss		−2.08
TOTAL LINK LOSS		−70.46
LINK MARGIN		−6.00
DESIGN LOSS		−76.46
Required Received Signal at 3 Gbps	9.36×10^{-8} W	−70.29 (= $10 \log_{10} 9.36 \times 10^{-8}$)
Required Laser Power at 3 Gbps = Required received signal - Design Loss	4.14 W (= $10^{6.17/10}$)	−70.29 + 76.46 = 6.17

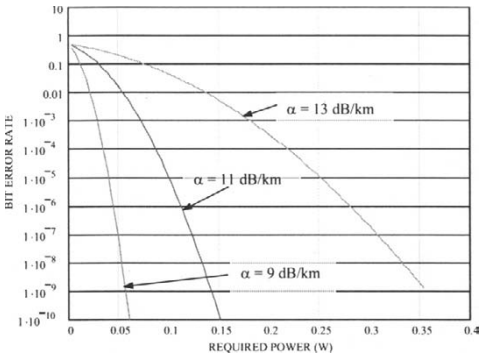


Fig. 4. Bit error rate (BER) as a function of required laser power for three atmospheric conditions : thin fog (= 9 dB/km), light fog (= 11 dB/km) and moderate fog (= 13 dB/km).

Bit Error Rate Versus Required Power for Absorption And Scattering Numerical results for laser communication system operating under various atmospheric conditions are depicted in Fig. 4. The BER of a high-bandwidth (570 Mbits/s) lasercom system as a function of required laser power (to achieve that BER) for a one-way path of 2 km is shown in Fig. 4. Some of the parameters were similar to the ones used for a retro-reflected experiment reported by other researchers [4]: laser wavelength $\lambda =$ 810 nm, transmitter efficiency $\tau_T = 0.5$, receiver optical efficiency $\tau_R = 0.5$, receiver diameter $D = 13$ cm, beam divergence (full) $\theta_T = 1$ mrad, detector field-of-view (FOV) ϕ det $= 200$ mrad, background radiation $H_{\text{bkg}} = 0.2\,\text{W m}^{-2}\,\text{nm}^{-1}\,\text{sr}^{-1}$, filter optical bandwidth $BWif = 20$ nm, detector bandwidth $BW_{\text{det}} = 422$ MHz, detector NEP: $\text{NEP}_{\text{det}} = 0.19\,\text{pW} / \text{Hz}^{1/2}$, APD gain $G = 135$, effective ionization coefficient $\kappa_{\text{eff}} = 0.005$. In this work presented in this paper, we varied the required power from about 4 to 350 mW and calculated the corresponding BER for three different foggy atmospheric conditions: the atmospheric attenuation factor, $\alpha = 9$ dB/km (thin fog, about 1.5 km visibility), $\alpha = 11$ dB/km (light fog, about 1.3 km visibility) , and $\alpha = 13$ dB/km (lightly moderate fog, 1 km visibility). It is seen from the Fig. 4 that to achieve a BER of 10^{-9} we need the laser powers of about 60, 140, and 350 mW, respectively, for operating under the above various foggy conditions specified by 9, 11, and 13 dB/km attenuation, respectively. There are also other parameters one can vary such as transmitter beam divergence angle, receiver diameter, etc. to change the required power for achieving a given BER. The designer can then trade-off these parameters in designing the lasercom system to operate optimally under various atmospheric scattering conditions.

Low-Visibility Atmospheric Optical Communication Link Using Pulse Position Modulation (PPM) In a low visibility atmosphere, multiple scattering due to particulates such as haze, rain, snow, clouds, fog, dust, and smoke can seriously degrade the performance of the atmospheric communication system by limiting the percentage of time the communication link would be available for data transmission. By adapting an optical receiver to increase the field-of-view (FOV), the performance degrada-

tion can be somewhat compensated because multiple scattering increases the beam divergence. In the example below we have analyzed the communication link using multiple-forward-scattered (MFS) radiation. The analysis is based on simulating various physical factors that are encountered in the actual atmospheric channel. Parameters that affect signal-to-noise fluctuation at the receiver end, such as multiple scattering of the laser beam off particles in clouds and the atmosphere, sky background, and the effects of direct illumination by the solar flux, are evaluated. Typically, a turbid atmosphere exhibits strong forward-scattering peaks with significant wide-angle scatter and backscattered components. Therefore, we can exploit scattered radiation in addition to the un-scattered component to improve communication through a low-visibility atmosphere (by additionally collecting some portion of the scattered signal with increased field-of-view). The atmosphere is characterized by the optical depth $\tau = \beta_L > 1$, where β is the atmospheric attenuation coefficient and L is the propagation path length. The available signal power received through the entire atmospheric link is given by utilizing the MFS propagation model [7]:

$$\langle P_A \rangle = \frac{P_t(\rho_R/\theta_T L)^2 \exp(-\beta_a' L)}{[1 + (\theta_0/\theta_T)^2][1 + (\theta_0/\theta_R)^2]}, \tag{16}$$

where P_t = cw laser power at wavelength λ, ρ_R = receiver aperture radius, θ_T = transmitter initial beam divergence (half-cone angle), β_a' = effective attenuation coefficient = $\beta - \beta_s'$, with $\beta = \beta_a + \beta_s$ (β_a being the absorption coefficient and β_s being the scattering coefficient of the medium), $\beta_s' = \beta_s \phi$, ϕ being the forward-scattering efficiency, $\theta_0 = 2^{1/2}/k\rho_0$ (k is the wave number = $2\pi/\lambda$ and the atmospheric channel coherence length $\rho_0 = (3/\beta_s' L k^2 \theta_F^2)$), θ_F = rms forward scatter angle, θ_R = field-of-view half cone angle, i.e., FOV = $2\theta_R$.

The unscattered component of the received signal is given by

$$\langle P_A \rangle_{\text{UNS}} = P_t(\rho_R/\theta_T L)^2 \exp(-\beta L). \tag{17}$$

Note that for the multiple-scattering regime, $[(P_A)_{\text{SC}}/(P_A)_{\text{UNS}} > 1]$,where $(P_A)_{\text{SC}}$ is the scattered component of radiation, given by $\langle P_A \rangle$ from the Eq. (16). Photo detector noise plus fluctuations in the channel and background noise cause errors to some of the demodulated information. The background noise power can be written as

$$P_B = \pi^2 \tau_a \tau_r \Delta\lambda \theta_R^2 \rho_R^2 N(\lambda), \tag{18}$$

where τ_a and τ_r are atmospheric and receiver transmissivities, $\Delta\lambda$ is the input optical filter bandwidth (which is the same as $BWif$ used in earlier equation), and $N(\lambda)$ is the spectral radiance of the background at wavelength λ. A typical value of $N(\lambda)$in the daytime is 10 W/m^2 sr μm. The ratio of scattered-to-unscattered power with the same transmitter and the same propagation distance can be expressed as

$$\frac{(P_A)_{\text{scattered}}}{(P_A)_{\text{unscattered}}} = \frac{\exp(\phi\beta_S L)}{\left[1 + \frac{2}{3}\frac{\beta_S \phi\, L\theta_F^2}{\theta_T^2}\right]\left[1 + \frac{2}{3}\frac{\beta_S \phi\, L\theta_F^2}{\theta_R^2}\right]}. \tag{19}$$

Figure 5 shows the result of SNR in dB as a function of FOV for a monodisperse and polydisperse (containing a number of different monodisperse particles) medium from an experiment described in detail elsewhere. [8,10] A scattering cell containing

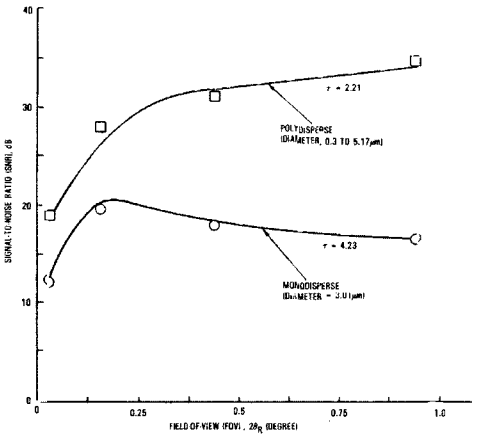


Fig. 5. Signal-to-noise-ratio as a function of receiver FOV for a monodisperse and polydisperse medium (τ = optical depth) (reproduced with permission from Optical Society of America).

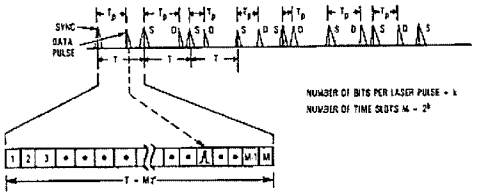


Fig. 6. Pulse position modulation (PPM) format (M is the total number of time slots and τ is the slot width) (reproduced with permission from Optical Society of America).

both polydisperse and monodisperse particles of polystyrene latex spheres (of diameters ranging from 0.3 to 5.17 μm) in water suspension simulated the low-visibility (cloudy) atmosphere a few kilometers thick in the laboratory using a 12.7-cm-long scattering cell. A narrowband interference filter with $\Delta\lambda \approx 15$ Å was used. A diode laser transmitter ($\lambda = 0.8486$ μm) in the experiment had a cw power output of 7 mW and a full-angle divergence of 15°. The optical thickness τ characterizing the medium was determined by extrapolating to the ratio of I_0 (intensity without medium) and I (intensity including medium) a zero FOV using the relation

$$\tau = [\ln(I_0/I)]_{FOV \to 0^-}.$$

In Fig. 5 notice that by increasing the FOV beyond 0.25 degrees, we will not gain much in SNR of an atmospheric optical communication system for a polydisperse medium, and actually lose SNR for a monodisperse medium.

Signal requirements for MFS-modeled communication channel When an airborne laser transmitter terminal (previously called "station") passes between the sun and the receiver terminal (which may be another airborne system), the signal becomes masked to a degree by the spectral component of the sun's background near the laser wavelength. By varying the receiver FOV, it is possible to control both the amount of information signal impinging on the detector and background noise. For a given data rate and background condition, there exists an optimal FOV to maximize the system margin. Pulse position modulation (PPM) is a highly efficient modulation technique for communication through clouds, haze, dust, smoke, rain, snow, etc. in which multiple scattering is involved with or without direct solar background. The PPM format employs a synch pulse to determine a major time division often called a window, which is then divided into time slots defined by the particular data value depending on which slot the pulse falls in. Figure 6 illustrates the PPM format. The SNR (assuming no pulse stretch effect in time domain) is derived from [9]

$$P_{\mathrm{BE}} = P_{\mathrm{FA}} + (1 - P_D), \tag{20}$$

$$P_{\mathrm{FA}} = \frac{1}{\sqrt{2\pi}\xi} \exp(-\xi^2/2), \tag{21}$$

$$P_D = \frac{1}{\sqrt{2\pi}\zeta} \exp(-\zeta^2/2), \tag{22}$$

where

$$\xi = \frac{q_T}{\sqrt{\langle q_N^2 \rangle}}, \quad \zeta = \frac{q_S - q_T}{\sqrt{\langle q_N^2 \rangle}}.$$

In the above equation, P_{BE} = probability of bit error, P_{FA}= probability of false alarm, P_D= probability of pulse detection, q_S= signal charge, q_T= charge threshold for pulse detection, $\langle q_N^2 \rangle$= rms noise charge. The voltage signal-to-noise ratio is then given by

$$\left(\frac{S}{N}\right)_V = \xi + \zeta = \frac{q_S}{\sqrt{\langle q_N^2 \rangle}}. \tag{23}$$

Power SNR is obtained from SNR $= \left[\left(\frac{S}{N}\right)_V\right]^2$.

Let us denote k = number of bits per laser pulse, and the number of time slots is $M = 2^k$. In this analysis, $M = 32$, and $k = 5$ were assumed. Assuming that the probability of error P_{BE} is 10^{-6}, and since thirty-one slots can have a false alarm rate, then

$$31P_{\mathrm{FA}} + (1 - P_D) = 10^{-6}.$$

If we assume equal probability so that

$$31P_{\mathrm{FA}} = \frac{1}{2}10^{-6} = 5 \times 10^{-7},$$

$$(1 - P_D) = \frac{1}{2}10^{-6} = 5 \times 10^{-7}.$$

Solving for P_D and P_{FA} and from Eqs. (??) and (22), the parameters ξ and ζ can now be determined as, in this example, ξ=5.5341 and ζ= 4.8991, so that $\xi + \zeta$=10.4332 yielding power SNR= 108.852 (= 20.37 dB). Following the method as outlined in an

earlier paper, [10] the optical power P_r required at the receiver for a specified bandwidth and SNR is given by

$$P_r = \left(\frac{h\nu Mm}{2\eta k}\right)\left[(\mathrm{SNR}) + \sqrt{(\mathrm{SNR})^2 + 4(N_B + N_D)\left(\frac{k}{M}\right)(\mathrm{SNR}),}\right] \quad (24)$$

where m is the bit rate.

The system margin is defined as the ratio of the available power received power to the signal power required to achieve a given bit rate m. Note that for the PPM coding, the bandwidth B is given by $B = 1/(2\times$ slot width). The slot width τ is then given by $\tau = k/(Mm)$, so that the bandwidth can be written as

$$B = \frac{1}{2\tau} = \frac{Mm}{2k}. \quad (25)$$

The system margin S_M can be now expressed in dB as a function of total FOV (= $2\theta_R$):

$$(S_M)\mathrm{dB} = 10\ \log\frac{P_A(\theta_R)}{P_r(\theta_R)}. \quad (26)$$

The system margin was determined for a concentration of the latex spheres suspended in water for optical depth $\tau = 4.23$, and two values of the spectral radiance generated by a quartz-halogen lamp were used: $N(\lambda) = 700$ W/m^2 sr μm and $N(\lambda) = 7000$ W/m^2 sr μm at the wavelength $\lambda = 0.8486$ μm. Figure 7 is a plot of the system margin vs. FOV for $\tau = 4.23$ and background radiances of $N(\lambda) = 700$ W/m^2 sr μm and $N(\lambda) = 7000$ W/m^2 sr μm. For a given data rate, the system margin increases initially with FOV, but levels off, so that at some point further increasing FOV does not increase the system margin further. Notice that the system margin is zero under some conditions, resulting in excess error rate (for example, in the ranges of FOV of 0.043° to 0.18° for the data rate of 327.86 Mbps and for the range of 0.043° to 0.27° for 655.36 Mbps). This implies that for the system margin above zero, communication can be maintained only when the system margin is positive. The data also suggests that the limiting FOV is really a minimum FOV, but from practical point of view, the system margin for large-angle FOV would be of greatest interest.

The multiple-forward scattering (MFS) model propagation model is applicable to provide parameter values needed for communication system design under most atmospheric conditions such as haze, dust, smoke, clouds, rain, snow etc. with or without direct solar background. Although the results from a laboratory-simulation scattering medium are presented here, the results and conclusions are useful in analyzing communication performance for a real scattering atmosphere. The multiple-forward-scattered radiation can be utilized for the potential use of optical communications such as between aircraft or ground-to-UAV in presence of scattering medium such as dust, fog, or smoke.

2.4. *Optical Link Reliability*

As we have seen that connection in free space is accomplished by means of a narrow optical beam with low divergence. For a successful and reliable installation of optical link, it is therefore necessary to know the steady parameters for standard atmosphere

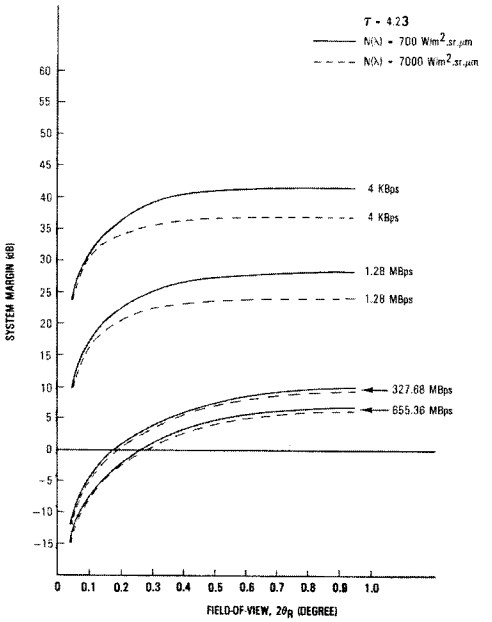


Fig. 7. System margin vs. FOV for optical depth = 4.23 for various data rates and two different background noise radiances (reproduced with permission from Optical Society of America).

and statistical character of the weather in a given locality. The link budget includes all average losses of optical power P_t [dBm] transmitted by the laser source, the received power P_r [dBm], receiver sensitivity N_b [dBm] and propagation loss L_p [dBm]. If we express link margin LM in dBm, we can write

$$\mathrm{LM} = P_t - L_p - N_b. \tag{27}$$

LM is a basic link parameter that serves to express the reliability of optical link. We can define the dynamic range as the interval of acceptable power in which the link function is guaranteed with a definite error rate. The receiver is saturated when $P_r > P_S$ (a specified saturation value), and the required signal-to-noise ratio is not provided when $P_r < N_b$. Link reliability is quantified by the availability, which is the percentage of time $T_{\mathrm{av}}(\%)$, when the data transmission bit rate is more than its required value. The link availability can equivalently be defined as the probability that additional power losses L_A caused by atmospheric effects (including absorption and scattering) are less (in dB) than the link margin LM. The link availability can now be expressed by means of a probability density $p(\alpha_A)$ of an attenuation coefficient α_A (dB/km) as

$$T_{\mathrm{av}} = 100\% \int_0^{\alpha_A \mathrm{lim}} p(\alpha_A) d\alpha_A, \tag{28}$$

where $\alpha_A \mathrm{lim}$ is the limiting attenuation coefficient value, given by

$$\alpha_A \text{lim} = \text{LM}(L)1000/L \tag{29}$$

for a range L. The probability distribution of $p(\alpha_A)$ can be determined from long-time monitoring of the received signal level from a real measuring link, or using the data collected over a local area in the past for a long time. In airports or meteorological stations, the visibility V(km) is monitored and is then converted to attenuation coefficient as [3]

$$\alpha_A(\text{dB/km}) = 4.434\frac{3.91}{V}\left(\frac{\lambda}{550}\right)^{-q} \tag{30}$$

[the factor 4.34 comes from the relationship α (dB/km) = (10 $\log_{10}e$) α (km^{-1}) = 4.34 α (km^{-1})] where the parameter $q = 1.6$ @ $V > 50$, $q = 1.3$ @$6 < V < 50$, $q = 0.16$ V+0.34 @ $1 < V < 6$, $q = V$- 0.5 @ $0.5 < V < 1$, q=0 @$V < 0.5$. By evaluating the above equation for T_{av} for given limiting attenuation coefficients, the link availability can be obtained, if we know the probability distribution of α_A for a local area.

3. Laser Communication Performance Prediction and Analysis Under Scintillation Conditions

Laser radiation propagating through turbulence develops both temporal and spatial fluctuations of irradiance, which is defined as scintillation. A lasercom system generally integrates the signal spatially at the plane of a receiver, but will still see temporal fluctuations in intensity. Scintillation is a serious issue for laser communications for both terrestrial as well as ground-to-space or space-to-ground data links, as it can produce large transient dips in the signal. The fading of the received signal below a prescribed threshold temporarily degrades, or even annihilates, the link performance. The probability density function (pdf) of the randomly fading irradiance signal due to atmospheric turbulence determines the reliability of a communication link. Over the years, many pdf models have been predicted for both weak and strong turbulence cases. In this section we will summarize some of the proposed PDFs that are relevant to optical communications. We will evaluate the performance of lasercom systems that operate under various atmospheric conditions. Each turbulent regime will be assumed homogeneous and isotropic (can be spatially varying) with refractive-index structure parameter C_n^2 and inner scale size l_0.Results from the literature for the pdfs for both plane wave and spherical wave will be considered.

We now discuss some of the pdf models for irradiance fluctuations which will be used for evaluating the performance of various lasercom systems such as horizontal (terrestrial) link and slant paths involving downlink (from space-to-ground) and uplink (ground-to-space) through atmospheric turbulence. One other parameter of interest is log-intensity variance (scintillation index). Expressions for point receiver and aperture-averaged (for finite size detector) irradiance variance will be outlined.

For laser communications system applications it is desirable to have available quantitative estimates of various statistical quantities that are associated with atmospheric turbulence-induced irradiance fluctuations (scintillation)

Most common measure of the magnitude of scintillation is the scintillation index defined by

$$\sigma_I^2 = \frac{\langle I^2\rangle - \langle I\rangle^2}{\langle I\rangle^2} = \frac{\langle I^2\rangle}{\langle I\rangle^2} - 1, \tag{31}$$

where I is the irradiance of the optical wave and $\langle\ \rangle$ denotes the ensemble average which is also equal to long-time average assuming the process to be ergodic. In weak fluctuation theory, the scintillation index is proportional to the Rytov variance given by

$$\sigma_1^2 = 1.23 C_n^2 k^{7/6} L^{11/6}. \tag{32}$$

C_n^2 is the strength of turbulence, $k = 2\pi/\lambda$, is the optical wave number, and L is the path length between the communication transmitter and the receiver. For a horizontal path of up to a few kilometer, the value of C_n^2 is taken to be a constant, whereas for downlink (space-to-ground) or uplink (ground-to-space), the C_n^2 profile model as a function of altitude h has to be used. For vertical or slant path, the so-called Hufnagel-Valley (HV5/7) model is generally regarded as representative of continental conditions. The value of 5 and 7 in the HV5/7 profile refer to the values of the atmospheric coherence length (r_0) in cm, and the isoplanatic angle (θ_0) in microradians, respectively at the wavelength of 0.55 μm. At an wavelength of 1.315 μm, the corresponding values are 14 cm and 20 μrad, respectively. [11]

3.0.1. Hufnagel-Valley Model

$$\begin{aligned} C_n^2(h) = {} & 0.00594(v/27)^2 (10^{-5}h)^{10} \exp(-h/1000) \\ & + 2.7 \times 10^{-16} \exp(-h/1500) + A \exp(-h/100), \end{aligned} \tag{33}$$

where $A = C_n^2\ (0)$ is the ground-level value of C_n^2, and v is rms wind speed modeled by

$$v = \left[\frac{1}{15 \times 10^3} \int_{5\times 10^3}^{20\times 10^3} \left(v_s h + v_g + 30 \exp\left[-\left(\frac{h - 9400}{4800} \right)^2 \right] \right)^2 dh \right]^{1/2}. \tag{34}$$

In Eq. (??), v_g is the ground wind speed and v_s is the beam slew rate. For our numerical calculations for various examples in this paper, we have assumed $v = 21$ m/s and $A = 1.7 \times 10^{-14}$ for the $C_n^2(h)$ model.

In some special cases such as tactical directed energy applications, it is necessary to have a maritime $C_n^2(h)$ model supported by long-term data. In contrast to HV5/7 model, there is no "standard" *maritime* $C_n^2(h)$ model. A turbulence profile that closely approximates the low-altitude component of a maritime model can be written as the following form:

$$C_n^2(h) = c_1 + c_2 \exp(-h/c_3) + c_4 \exp(-h/c_5)(o \le h \le 6000\text{m}), \tag{35}$$

where the values of the constants are given [11] in Table 2 for three maritime cases (best, median and worst):

In many practical cases either the HV5/7 or median maritime profile is used for computational purposes.

3.1. Scintillation Index: Point Receiver and Aperture Averaged Variance

After the optical beam propagates a sufficiently long path and/or due to large atmospheric turbulence strengths, beam quality at the receiver tends to break up into regions

Table 2. Values of the constants for maritime C_n^2 profiles.

Case	$C_n^2(h=0)$ $(m^{-2/3})$	c_1 $(m^{-2/3})$	c_2 $(m^{-2/3})$	c_3 (m)	c_4 $(m^{-2/3})$	c_5 (m)
Best − 10%	$1.0E-16$	$9.8286E-18$	$7.1609E-17$	100	$1.9521E-17$	1500
Median − 50%	$8.0E-16$	$9.8583E-18$	$4.9877E-16$	300	$2.9228E-16$	1200
Worst − 10%	$1.0E-14$	$9.2002E-18$	$9.4387E-15$	800	$6.7328E-16$	1000

of high and low intensity which leads to significant signal fades. This is the case typically happens for a point receiver. If we now increase the size of the receiver aperture relative to the size of theses "patches" of high and low intensity, this will basically "average" these signal fluctuations over the aperture size. This is called "aperture averaging". It is therefore necessary to multiply the scintillation index for a point aperture (i.e., diameter, $D = 0$) by the aperture averaging factor A to obtain the scintillation index for a specific finite aperture (of size D):

$$A = \sigma_I^2(D)/\sigma_I^2(0). \tag{36}$$

The factor A is given by

$$A = \frac{16}{\pi D^2} \int_0^D b_I(\rho) \left[\cos^{-1}\left(\frac{\rho}{D}\right) - \frac{\rho}{D}\left(1 - \frac{\rho^2}{D^2}\right)^{1/2} \right] \rho d\rho, \tag{37}$$

where ρ is the separation distance between two points and $b_I(\rho)$ is the normalized covariance function.

In this paper, we will consider basically three types of optical waves for generating numerical examples of evaluating communication performance: Plane Wave, Spherical Wave and Beam Wave. For the space-to-ground communications downlink, the beam is distorted as it enters down through the atmosphere. The ground-level scintillation near the center of the received wave can be accurately modeled by a plane wave. [12] For an uplink, where the atmospheric turbulence begins just outside the transmitting aperture, we can assume a spherical wave. For a horizontal path communication link, we can assume that the C_n^2 is constant (for shorter paths, 10 km or less), and we assume Gaussian beam wave.

We follow the treatment given in [13] for the expressions of scintillation index for plane and spherical waves including the aperture averaging effect for a finite size aperture. For Gaussian beam wave, we will use the expressions given by [14].

Let us assume that the received irradiance I can be modeled as a product of two terms, x (arises from large-scale turbulent eddies, and y (arises from small-scale eddies): $I = xy$. We assume that the two processes x and y to be statistically independent. The received irradiance of the optical wave can thus be modeled as a modulation process in which small-scale (diffractive) fluctuations are multiplicatively modulated by large-scale (refractive) fluctuations. For normalized irradiance, $\langle I \rangle = 1$, and we can write the scintillation index as

$$\sigma_I^2 = \sigma_x^2 + \sigma_y^2 + \sigma_x^2\sigma_y^2. \tag{38}$$

Next we take into account of the finite size aperture effect, which is the factor A as defined earlier. We can then finally write the expressions for aperture-averaged scintillation index for negligible inner scale as given by [13].

3.1.1. Plane Wave

$$\sigma_I^2(D) = \exp\left[\frac{0.49\sigma_1^2}{\left(1+0.65d^2+1.11\sigma_1^{12/5}\right)^{7/6}} + \frac{0.51\sigma_1^2\left(1+0.69\sigma_1^{12/5}\right)^{-5/6}}{1+0.90d^2+0.62d^2\sigma_1^{12/5}}\right] - 1, \tag{39}$$

where σ_1^2 is the Rytov variance for plane wave. Note that for constant C_n^2, the Rytov variance is given by Eq. (??), but for slant path (downlink or uplink), altitude-dependence of $C_n^2(h)$ has to be taken into account. The parameter d is defined by

$$d = \sqrt{\frac{kD^2}{4L}}, \tag{40}$$

where L is the path length.

3.1.2. Spherical Wave

Similar expressions for the aperture-averaged scintillation index for spherical wave are given by [13]

$$\sigma_I^2(D) = \exp\left[\frac{0.49\sigma_2^2}{\left(1+0.18d^2+0.56\sigma_2^{12/5}\right)^{7/6}} + \frac{0.51\sigma_2^2\left(1+0.69\sigma_2^{12/5}\right)^{-5/6}}{1+0.90d^2+0.62d^2\sigma_2^{12/5}}\right] - 1 \tag{41}$$

where σ_2^2 is the Rytov variance for spherical wave (note: $\sigma_2^2 = 0.4\,\sigma_1^2$). Again note that for slant path (downlink or uplink), altitude-dependence of $C_n^2(h)$ has to be taken into account in calculating σ_2^2. We will use these scintillation index values for analyzing the communication performance examples later in this paper.

3.1.3. Gaussian beam wave [14]

For the Gaussian beam wave, the beam size (radius) at the receiver is given by

$$w(L) = w_0(\hat{r} + \zeta\,\hat{z}^2)^{1/2}, \tag{42}$$

where w_0 is the transmitter beam size (radius), $\hat{r}(L) = (R_0 - L)\,/\,R_0$ is the normalized transmitter focusing parameter, R_0 is the phase front radius of curvature at the transmitter, $\hat{z} = L\,/\,\hat{z}_d$ is the normalized distance, $\hat{z}_d = kw_0^2/2$ is diffractive distance, L is the communication channel propagation distance. The global coherence parameter ζ is defined by

$$\zeta = \zeta_S + \frac{2w_0^2}{\rho_0^2}. \tag{43}$$

In the above equation ζ_S is the coherence of the source beam ($\zeta_S = 1$ for coherent beam) and $\rho_0(L) = (0.55C_n^2k^2L)^{-3/5}$ is the coherence length of a spherical wave propagating in the atmospheric turbulence.

The scintillation index for the Gaussian beam wave for a point receiver is given by [14]

$$\sigma_I^2 \cong 4.42\sigma_1^2\hat{z}_{\rm rec}^{5/6}\frac{\rho^2}{w^2(L)} + 3.86\sigma_1^2\Big\{0.4\left[(1+2\hat{r}_{\rm rec})^2 + 4\hat{z}_{\rm rec}\right]^{5/12} \times \cos\left[\frac{5}{6}\tan^{-1}\left[\frac{1+2\hat{r}_{\rm rec}}{2\hat{z}_{\rm rec}}\right] - \frac{11}{16}\hat{z}_{\rm rec}^{5/6}\right]\Big\}, \tag{44}$$

where $\sigma_1^2 = 1.23C_n^2k^{7/6}L^{11/6}$ is the Rytov variance for a plane wave. The receiver beam parameters are given by [15]

$$\hat{r}_{\rm rec}(L) = \frac{R(L)+L}{R(L)} \quad \text{and} \quad \hat{z}_{\rm rec}(L) = \frac{L}{0.5kw^2(L)}. \tag{45}$$

The phase front radius of curvature at the receiver is defined by [14]

$$R(L) = \frac{L(\hat{r}^2 + \zeta\hat{z}^2)}{\phi\hat{z} - \zeta\hat{z}^2 - \hat{r}^2} \quad \text{and} \quad \phi \equiv \frac{\hat{r}}{\hat{z}} - \hat{z}\frac{w_0^2}{\rho_0^2}. \tag{46}$$

As before, the aperture-averaged scintillation index for Gaussian beam wave case can be obtained by multiplying the point-receiver scintillation index by the aperture average factor, A_G for Gaussian beam wave case given by [15]

$$A_G = \frac{16}{\pi}\int_0^1 x\,dx\,\exp\left\{\frac{-D^2x^2}{\rho_0^2}\left(2 + \frac{\rho_0^2}{w_0^2\hat{z}^2} - \frac{\rho_0^2\phi^2}{w^2(L)}\right)\right\} \times \left[\cos^{-1}(x) - x(1-x^2)^{1/2}\right]. \tag{47}$$

In the above equation $x = \rho/D$ with D being the aperture diameter.

3.2. *Probability Density Function (pdf) Models*

Free-space laser communications have the potential for offering extremely high data rate links and are very suitable where fiber optic links are not practical, such as between buildings across cities and space-to-ground or ground-to-space data links. To establish high reliable free-space laser communications links quantitative estimates of various statistical quantities that are associated with atmospheric turbulence-induced irradiance fluctuations (scintillation) are absolutely necessary. The signal fading severely

degrades the link performance and the bit error probability of the received information signal is increased above an unacceptable limit. Scintillation can thus cause losses in received power which can lead to ultimate break in communications. The form of the probability density function (pdf) of the randomly fading irradiance signal determines the extent of reliability of an optical communication system. An accurate pdf is therefore necessary in designing a reliable lasercom system. Understanding optical propagation through turbulence is extremely complex because of the complicated nature of the generation of atmospheric turbulence occurring in different situations. There is no one single pdf which can describe correctly the random fluctuations through every type of turbulent condition. Most of the treatment available for optical propagation are based on statistical theory, and is valid for isotropic or homogeneous turbulence following Kolmogorov's law. Assumptions of stationarity and ergoticity help to simplify the calculation (because the ensemble average in that case is the same as long-time average). It was pointed out that the higher-order moments can give more accurate information concerning the contributions from the tails of the probability density function.[16] Detection and fade probabilities are primarily based on the tails of the pdf ,and therefore underestimating this region significantly affects the accuracy of communication performance analysis.

The moments, central moments and other higher-order coefficients up to eighth order (non-dimensional) in statistical analysis of irradiance fluctuations with a given pdf were obtained using the equations [17,18]:

$$k\text{th moments} : m_k = \int_0^\infty I^k p(I) dI, \tag{48}$$

$$k\text{th central moments} : \mu_k = \langle (I - \langle I \rangle)^k \rangle, \tag{49}$$

where I stands for irradiance, $p(I)$ is the pdf of I ,and $\langle\rangle$ indicates an ensemble average.

$$\text{Skewness} : \Gamma_3 = \mu_3 / \mu_2^{3/2}, \tag{50}$$

$$\text{Excess} : \Gamma_4 = \mu_4 / \mu_2^2 - 3, \tag{51}$$

$$\text{Superskewness} : \Gamma_5 = \mu_5 / \mu_3 \mu_2 - 10, \tag{52}$$

$$\text{Superexcess} : \Gamma_6 = \mu_6 / \mu_2^3 - 15, \tag{53}$$

$$\text{Hyperskewness} : \Gamma_7 = \mu_7 / \mu_3 \mu_2^2 - 105, \tag{54}$$

$$\text{Hyperexcess} : \Gamma_8 = \mu_8 / \mu_2^4 - 105. \tag{55}$$

Note that, for a Gaussian pdf, the odd-order coefficients Γ_5 and Γ_7 do not exist and become indeterminate, whereas the even-order coefficients $\Gamma_4 = \Gamma_6 = \Gamma_8 = 0$. Therefore any nonzero value of the even-order parameters indicates that the random process deviates from a Gaussian distribution.

The higher-order coefficients (up to eighth order) defined above can give much more detailed information concerning the contribution from the tails of non-Gaussian distributions and were utilized in [19] to compare model distributions (such as K, log-normal, Rice-Nakagami and a universal statistical model distribution) with the measured data of 0.6328 μm laser-irradiance fluctuations propagated through a laboratory-simulated turbulence. [16,17] The laboratory atmospheric turbulence generator is described in details elsewhere. [20,21]

From practical point of view, it is desirable to have available a tractable pdf model for the irradiance fluctuations so that we can predict the communication performance with an acceptable accuracy. Log-normal distribution is the most widely used model for the probability density function of the irradiance due to its simplicity : this pdf however is only applicable to weak turbulence regime. As the strength of turbulence increases, multiple scattering effects become important and significant deviations from log-normal statistics are exhibited in experimental data. Log-normal pdf underestimates the behavior in the tails as compared with measurement results. For communication systems, accurate detection and fade probabilities primarily depend on the tails of the pdf, hence underestimating this region significantly affects the accuracy of the communication performance. As mentioned earlier, many statistical models have been proposed over the years to describe scintillation characteristics under a wide range of turbulence conditions, e.g., K distribution, $I-K$ distribution, and log-normal Rician channel. In a recent series of paper a gamma-gamma pdf has been proposed [13,22] based on work by Nakagami. [23] This model is a two-parameter distribution which is based on a doubly stochastic theory of scintillation, and assumes that small-scale irradiance fluctuations are modulated by large-scale irradiance fluctuations of the propagating wave, both governed by independent gamma distributions. The gamma-gamma distribution can be directly related to the atmospheric parameters defining turbulence conditions. In this chapter, we will consider mainly the two pdfs, namely, log-normal and gamma-gamma: log-normal distribution is applicable to weak turbulence case and the gamma-gamma distribution is assumed to valid for both weak and strong turbulence regimes. These two pdfs which are used in this paper for computational purpose of evaluating communication performance are given by

Log-normal distribution:

$$p(I) = \frac{1}{I\sigma_I(D)\sqrt{2\pi}} \exp\left\{ -\frac{\left[\ln\left(\frac{I}{\langle I\rangle}\right) + \frac{1}{2}\sigma_I^2(D)\right]^2}{2\sigma_I^2(D)} \right\}, \quad I > 0, \tag{56}$$

where $\sigma_I(D)$ is the aperture-averaged scintillation index as defined earlier, and $\langle I\rangle$ is the mean intensity.

Gamma-gamma probability density function:

$$p(I) = \frac{2(\alpha\beta)^{(\alpha+\beta)/2}}{\Gamma(\alpha)\Gamma(\beta)} I^{(\alpha+\beta)/2-1} K_{\alpha-\beta}\left(2\sqrt{\alpha\beta I}\right), \tag{57}$$

where $K_a(\,)$ is the modified Bessel function of the second kind of order a. Here $\alpha$and β are the effective number of small-scale and large scale eddies of the turbulent environment. These parameters are directly related to the atmospheric turbulence conditions as [13,22]

Plane Wave with aperture-averaged scintillation index:

$$\alpha = \left\{ \exp\left[\frac{0.49\sigma_1^2}{\left(1 + 0.65d^2 + 1.11\sigma_1^{12/5}\right)^{7/6}} \right] - 1 \right\}^{-1}, \tag{58}$$

$$\beta = \left\{ \exp\left[\frac{0.51\sigma_1^2 \left(1 + 0.69\sigma_1^{12/5}\right)^{-5/6}}{1 + 0.90d^2 + 0.62d^2\sigma_1^{12/5}} \right] - 1 \right\}^{-1}. \tag{59}$$

Spherical Wave with aperture-averaged scintillation index:

$$\alpha = \left\{ \exp\left[\frac{0.49\sigma_2^2}{\left(1 + 0.18d^2 + 0.56\sigma_2^{12/5}\right)^{7/6}} \right] - 1 \right\}^{-1}, \tag{60}$$

$$\beta = \left\{ \exp\left[\frac{0.51\sigma_2^2 \left(1 + 0.69\sigma_2^{12/5}\right)^{-5/6}}{1 + 0.90d^2 + 0.62d^2\sigma_2^{12/5}} \right] - 1 \right\}^{-1}, \tag{61}$$

where the parameters σ_1, $\sigma_2$and d were defined earlier.

3.3. Received Signal-to-Noise-Ratio (SNR) and Bit-Error-Rate (BER)

In almost every area of measurements, the ultimate limit to the detectability of a weak signal is set by noise—unwanted signals that obscure the desired signal. The same is true for free-space laser communications systems where the shot noise, background noise and thermal noise contribute to the total noise at the receiver. The goal of a lasercom system for digital communication is to transmit the maximum number of bits per second over the maximum possible range with the fewest errors. Electrical data signals are converted to optical signals via a modulator. A "1" is transmitted as a pulse of light while a "0" has no light output. The number of "1's" and "0's" transmitted per second determines the speed of the link (bit rate). At the receiving end of the link, the optical signal is detected by an optical-to-electrical converter (e.g., a photo detector). A decision circuit then identifies the "1's" and "0's" in the signal, and thus recovers the information sent.

For digital communication system, information is sent over an optical link as digital symbols. This is accompanied by encoding the source information into binary symbols (bits) and transmitting the bits as some type of coded optical field, for example, by encoding on a bit-by-bit basis (binary encoding). Each bit is then sent individually by transmitting one of two optical fields to represent each bit. In this paper we will consider only direct detection (DD) system in which the standard binary procedure is to pulse an optical source (e.g., a laser or an LED) on or off depending on data bit. This encoding is referred to as on-off keying (OOK). At the receiver OOK decoding is based on a decision as to whether the pulse slot time has high enough field energy or not. The selected threshold determines the best performance in decoding the correct signal with the lowest probability of making a bit decision error and thus the bit error rate (BER) can be obtained. The other modulation format of pulse position modulation (PPM) will also be considered.

3.3.1. Relationship between SNR and BER

The performance of the lasercom system can be predicted by computing the bit-error-rate (BER) of the system which depends on the modulation format, and the signal-to-noise ratio (SNR) where the noise contributions come from all possible sources which

include signal shot noise, dark current noise, thermal/Johnson noise in the electronics following the photo detector, and the background noise. Assuming a Gaussian distribution of noise, the SNR at the output of the photodetctor in the absence of turbulence is given by [13]

$$\mathrm{SNR}_0 = \frac{P_S}{\sqrt{\left(\frac{2h\nu B}{\eta}\right)(P_S + P_B) + \left(\frac{h\nu}{\eta e}\right)^2 \left(\frac{4kT_N B}{R}\right)}}, \tag{62}$$

where P_S is the signal power of the optical transmitter and P_B is the background noise, both in watts, η is the detector quantum efficiency, e is the electronic charge in coulombs, h is the Planck's constant, ν is the optical frequency in hertz, k is the Boltzmann constant, B is the bandwidth of the (detector) filter, T_N is the effective noise temperature and R is the effective input resistance to the amplifier of the detector. For shot noise limited operation, the background noise and the thermal noise can be ignored.

In presence of turbulence, the SNR is a fluctuating term (i.e., an instantaneous value) and the average (mean) value is to be taken The mean SNR can be expressed as [13]

$$\langle \mathrm{SNR} \rangle = \frac{\mathrm{SNR}_0}{\sqrt{\frac{P_{S0}}{\langle P_S \rangle} + \sigma_I^2(D)\mathrm{SNR}_0^2}}, \tag{63}$$

where SNR_0 is the signal-to-noise-ratio in absence of turbulence defined earlier, P_{S0} is the signal power in the absence of atmospheric effects, the mean input signal power $\langle P_S \rangle$ (i.e., the mean of the instantaneous input signal power P_S), and $\sigma_I^2(D)$is the aperture-averaged scintillation index previously discussed.

Let us now understand the communication performance criteria defined by the parameter BER. The signal entering the decision circuit fluctuates due to the various noise mechanisms. The bit-error-rate (BER) is the probability of incorrect bit identification by the decision circuit. The design of optical communication systems relies on the estimation of performance figure of merit such as bit-error-rate (BER). One of the most commonly used parameters is the receiver sensitivity defined as the minimum averaged received optical power required to achieve a given BER. In order to simulate the characteristics of optical links or sub-systems, it is therefore essential to be able to compute the BER in an effective and reliable way. It is important to investigate the degree to which the optical communication may be expected to be degraded by atmospheric turbulence, by calculating the expected increase in bit error probability due to turbulence, including the effect of a finite aperture receiver.

3.3.2. Why Atmospheric Turbulence Increases the Bit Error Rate?

For any type of detector, we can assume that the bit error probability (i.e., the relative frequency of errors) is functionally related to the signal strength (denoted by s) which can be defined as the number of signal photons per bit incident on the detector. The bit error probability can be written as $E(s)$ for a signal pulse of strength s. If the signal strength fluctuates randomly due to atmospheric turbulence, s becomes a random variable with a probability distribution $P(s)$. Therefore the bit error probability will also become a random variable. Most of the statistical information about E is contained

in the first two moments of the distribution (the mean and the variance), which can be expressed as

$$\langle E\rangle = \int_0^\infty E(s)P(s)ds, \tag{64}$$

$$\mathrm{var}(E) = \int_0^\infty (E - \langle E\rangle)^2 ds, \tag{65}$$

where $\langle E\rangle$ denotes ensemble average and "var" indicates the variance. If we express $E(s)$ as a finite Taylor formula about $\langle s\rangle$, with a second degree error term, we can write

$$\langle E\rangle = E(\langle s\rangle) + 1/2 \int_0^\infty E^{(2)}(\langle s\rangle + \theta(s)(s-\langle s\rangle)(s - \langle s\rangle))^2 P(s)ds, \tag{66}$$

where $E^{(2)}$ is the second derivative of $E(s)$, and $\theta(s)$ is between 0 and 1 for each value of s, but is otherwise completely unknown. It is evident from the above equation that $\langle E\rangle$ is always greater than $E(\langle s\rangle)$ if the second derivative of $E(s)$ is positive for all s. *Therefore a sufficient condition for atmospheric turbulence to cause an increase in the bit error rate is that the second derivative of the bit error probability function, $E(s)$ be everywhere positive.* This can be understood intuitively by observing the typical "BER" curves vs. signal strength which is decreasing with s, and highly nonlinear, with a positive second derivative. As the signal s fluctuates about its mean $\langle s\rangle$, value of E given by $E(s)$ oscillates along the curve, to either side of the point $(\langle s\rangle, E(\langle s\rangle))$. But since the curve is decreasing with s, but more slowly as s increases, the increase in E as the point moves to the left is greater in magnitude than the decrease in E as the point moves to the right. Thus the average value of E is biased upward from its value at $\langle s\rangle$.

In this paper we will assume lognormal and gamma-gamma distributions (depending on the strength of turbulence and the communication link scenario) arising from atmospheric turbulence.

3.3.3. Bit-Error Rate Computation for OOK Modulation

One of the primary considerations in the design of a communication system is the performance of the system when it is corrupted by noise. We have discussed the various sources of noise in a lasercom system including fluctuating random noise introduced by the turbulence environment. In a binary on-off keying (OOK) system, the transmitter sends a pulse of light into a channel to represent a "1" and does not send any light for a "0". When signal plus noise is present at the receiver input, there are two ways in which errors can occur. The receiver decides a "0" was sent when in fact a "1" was transmitted. Also, the receiver may decide a "1" was transmitted when a "0" was actually sent. The probability of error or the bit-error rate (BER) can be expressed as

$$P_e = P(1|0)P(0) + P(0|1)P(1), \tag{67}$$

where $P(0)$ and $P(1)$ are the probability of a binary "0" and "1" respectively, and $P(1|0)$ and $P(0|1)$ are the conditional probabilities. Assuming the transmitter is

sending "1"s and "0"s with equal probability, each has a probability equaled to $1/2$, so the probability of bit error is then given by

$$P_e = \frac{1}{2}P(1|0) + \frac{1}{2}P(0|1). \tag{68}$$

We also assume the Gaussian distribution for both noise as well as signal plus noise, the probability of error in absence of turbulence is given by

$$\mathrm{BER}_0 = \frac{1}{2}\left(\frac{\mathrm{SNR}_0}{2\sqrt{2}}\right). \tag{69}$$

In presence of turbulence, the probability of error is given by [13,24]

$$\mathrm{BER} = \frac{1}{2}\int_0^\infty p_I(s)\mathrm{erfc}\left(\frac{\langle \mathrm{SNR}\rangle s}{2\sqrt{2}\langle i_S\rangle}\right) ds, \tag{70}$$

where $p_I(s)$ is the probability distribution of irradiance, i_S is the instantaneous signal current whose mean value is given by $\langle i_S\rangle = \eta e\langle P_S\rangle/h\nu$, $\langle P_S\rangle$ being the mean signal power, and $\langle \mathrm{SNR}\rangle$ is the mean SNR in presence of turbulence defined earlier. Depending on the level of turbulence and the communication link scenario, various forms of the probability distributions can be modeled. In this chapter we will compute BER using lognormal and gamma-gamma probability distributions. Note that the probability distribution contains the parameter s which is related to the aperture-averaged scintillation index. Some of the numerical results showing BER computations will be presented later in the paper.

3.3.4. Bit-Error Rate Computation for Pulse Position Modulation(PPM): Some Basics

We presented the concept of pulse position modulation (PPM) in the earlier section to evaluate performance of laser communication system which has to operate under scattering medium. In that case we assumed a non fluctuating medium, i.e. no turbulence medium was assumed. In many situations of practical importance, we need to take into account of the turbulence environment. It has been shown that M-ary pulse-position modulation is the optimum modulation format for background-limited optical communication. In this type of modulation, the number of bits sent per PPM word or pulse is $\log_2 M = k$ and the PPM assumes a mechanism for achieving time synchronization having an average data rate equal to

$$R = \frac{\log_2 M}{(M+N)\Delta T}. \tag{71}$$

In a basic time period called the interpulse interval, one has to determine in which one of the M slots, each of width ΔT seconds, the received pulse resides. The difference between the interpulse interval and $M\Delta T$, the word time, is called the dead time and is denoted $N\Delta T$, where N in most situations is not an integer. M is chosen in order to maximize the average data rate, so that the maximum data rate is given by [25]

$$R_{\max} = R_{P_{\max}}\left[\frac{\log_2 M_{\mathrm{opt}}}{1+\alpha M_{\mathrm{opt}}}\right], \tag{72}$$

where $R_{P_{\max}}$ is the maximum pulse repetition rate of the system which can be expressed in terms of pulse repetition rate R_p as follows

$$R_P = \frac{1}{T_D(1+\alpha M)} \quad \text{for} \quad \alpha = \frac{\Delta T}{T_D}$$
$$R_{P_{\max}} = \frac{1}{T_D} \quad \text{for} \quad \alpha = 0, \tag{73}$$

where ΔT is the slot width, T_D is the dead time, α for maximum data rate is given by

$$\alpha = \frac{\log_2 e}{M \log_2(M/e)}. \tag{74}$$

In order to transmit at a lower data rate, one can use a lower $R_{P_{\max}}$ or a smaller value of M. If one has to operate at a higher data rate, one would require either a higher value for $R_{P_{\max}}$ or the use of multiple frequencies to increase M.

The probability of making an error in demodulating the PPM symbol, P_E, is expressed as [25]

$$P_E = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-(x - \mathrm{SNR})^2/2\right\} (1 - \mathrm{erfc}(x)) M^{-1} dx, \tag{75}$$

where $\mathrm{erfc}(x)$ is a complimentary error function.

The above equation can be approximated as

$$P_E \cong \frac{(M-1)}{(\pi \mathrm{SNR})^{1/2}} \exp(-\mathrm{SNR}/4), \tag{76}$$

where the signal-to-noise ratio, SNR is $\mathrm{SNR} = \frac{K_S^2}{K_n}$ with K_S is the signal count and K_n is the noise count. Finally, the bit error rate, BER can be related to P_E in the following way [25]:

$$\mathrm{BER} = \frac{2^{k-1}}{2^k - 1} P_E, \tag{77}$$

where k is the number of bits sent per PPM word or pulse, as defined earlier.

Note that in the above derivations, we have assumed that we have ignored pulses which overlap more than one slot, i.e., pulse broadening by the optical scatter channel is assumed less than ΔT, with perfect synchronization. However, we have also studied the effect of pulse stretching and will be discussed in details in numerical example later on in this paper.

3.4. Probability of Fade

The object of designing any communication system is to obtain uninterrupted communications between the transmitter and the receiver during the exchange of information of sending and receiving data. We have discussed earlier that because of the presence of randomly varying atmospheric channel, the received signal fluctuates so that it is possible for the received signal to fall below an acceptable detection level and error rate. The fading probability for an optical communication link can be determined from

the knowledge of statistics of intensity fluctuations (pdf) and the scintillation index. The reliability of a communication system thus depends on the fading probability. Let us denote $I(t)$ as the instantaneous value of irradiance normalized to its mean. We need to obtain the fraction of time $I(t) \geq I_T$, where I_T is the threshold level of intensity specified by the communication system. Because of the assumption of an ergodic process for the statistics of optical propagation in our case, the time averages are equal to ensemble averages, so that we can also write the fraction of time of a fade where $I \leq I_T$ as

$$\text{Fraction}(I \leq I_T) = CP_I(I \leq I_T) = \int_0^{I_T} p_I(I)dI, \tag{78}$$

where CP_I is the cumulative probability for irradiance and is assumed to be either a lognormal (for weak turbulence) or a Gamma-Gamma distribution (for weak to strong turbulence) model, for the numerical examples described in this papter. The fade threshold parameter F_T of a signal (in decibels) can be defined by [13,26]

$$F_T = 10\log_{10}(\langle I(0,L)\rangle / I_T), \tag{79}$$

where $\langle I(0,L)\rangle$ is the on-axis mean intensity at a communication distance L. We can also write

$$\ln\left[\frac{I_T}{\langle I(0,L)\rangle}\right] = -0.23F_T. \tag{80}$$

In the numerical examples described in the following section, we show the probability of fade as a function of the threshold parameter F_T (given in dB below the mean on-axis intensity at the receiver) for various communication scenarios, such as downlink, uplink and horizontal links and for two model pdfs (lognormal and Gamma-Gamma).

4. Example Numerical Results

In this section we present numerical results generated by computer programs to predict the communication performance for various specific scenarios of laser communication links.

4.1. Example 1: Uplink (Spherical Wave/OOK)

For an uplink lasercom link, i.e., from a ground-based terminal transmitting to a satellite, the atmospheric turbulence begins just outside the transmitter aperture, and we assume a spherical wave. The aperture-averaged scintillation index $\sigma_I^2(D)$ is calculated from the parameters of the propagation path (propagation model, satellite altitude, and zenith angle) and the wavelength. For computational purpose, we assume a Hufnagel-Valley (H-V) turbulence profile discussed earlier with $A = 1.7\times10^{-14}$ m$^{-2/3}$, and $v = 21$ m/s to represent the H-V5/7 model conditions. We predict lasercom system performance by first computing the BER. The parameters used for the OOK modulation format for uplink are the following: wavelength $\lambda = 1.55$ μm, satellite height $H = 3.8\times10^5$ m, transmitter divergent angle = 17.45 μrad, and a data rate of 2 Gbit/s.

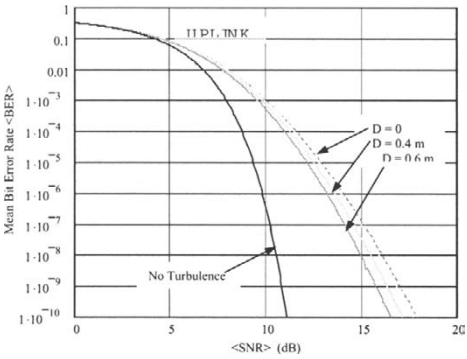


Fig. 8. Uplink Mean bit-error-rate for on-off-keying direct detection as a function of mean signal-to-noise-ratio for various receiver aperture diameters [$D = 0$ (point receiver), and $D = 0.4$ m, 0.6 m]. The zenith angle is 45 degrees. The no-turbulence case is also shown.

The BER calculation is based on a Gamma-Gamma probability density function of the intensity fluctuations and the procedure was outlined earlier.

Figure 8 shows the mean BER as a function of average signal-to-noise ratio ($\langle \text{SNR} \rangle$) for a point receiver ($D = 0$) and a receiver with diameter $D = 40$ cm and $D = 60$ cm, with zenith angle $= 45°$, and gamma-gamma PDF. The effect of atmospheric turbulence on BER is clearly apparent: at $\langle \text{SNR} \rangle$ of about 11 dB, the BER increases from a value of 10^{-9} for no turbulence case by several orders of magnitude to about 1.9×10^{-4} for a point receiver uplink communication with turbulence. Also seen from the figure is that increasing the aperture diameter can improve BER as predicted (for example, for a BER of 10^{-9} one needs average $\langle \text{SNR} \rangle$ of approximately 16.97 dB for point receiver, 16.32 dB for a diameter of $D = 40$ cm, and 15.74 dB for $D = 60$ cm). For a communication system with available $\langle \text{SNR} \rangle$ of 16.35 dB, the BER is decreased from about 3.6×10^{-9} for a point receiver to 10^{-9} for a receiver diameter of 40 cm. Increasing the aperture diameter can thus improve BER as predicted.

Figure 9 shows the mean BER as a function of mean SNR for three zenith angles: 0, 45 and 60 degrees. A gamma-gamma distribution with spherical wave model was assumed and the effects of inner scale were ignored. The effect of turbulence on mean bit-error-rate is clear: for example for a mean signal-to-noise-ratio of 10 dB the mean bit-error-rate increases from a value of 2.5×10^{-7} (no turbulence) to about 5×10^{-4} (several orders of magnitude) for a zenith angle of 60 degrees. The figure shows that an increase in zenith angle requires a greater SNR for a given BER For example, to achieve a BER of 10^{-9}, the required SNR is increased from 15.42 dB (0 degree) to 16.06 dB (60 degree).

Next we calculate the probability of fade for the same uplink spherical wave model to the satellite under similar conditions for two PDFs, namely, gamma-gamma and log-normal distributions, and for both a point receiver ($D = 0$) and for a number of receivers with diameters of $D = 20$, 40, and 60 cm. Figure 10 shows the probability-of-fade curves as a function of threshold level F_T below the mean irradiance. A lower probability of fade is predicted for the log-normal PDF than for gamma-gamma. The fade

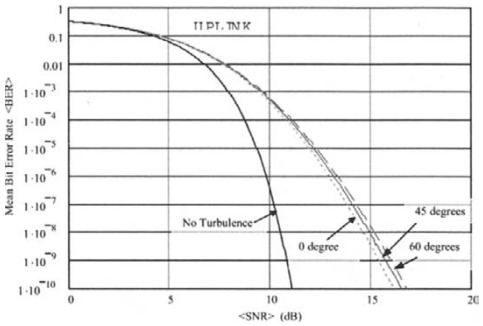


Fig. 9. Uplink Mean bit error rate for on-off keying direct detection as a function of mean signal-to-noise-ratio for three zenith angles: 0, 45 and 60 degrees with aperture diameter $D = 60$ cm, and for the no-turbulence case.

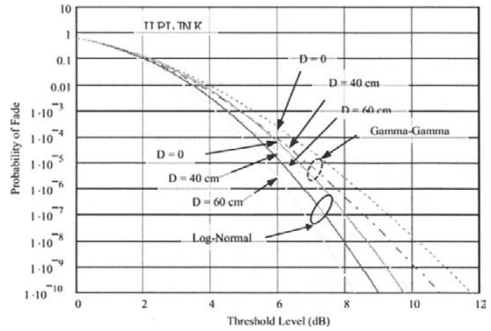


Fig. 10. Uplink Probability of fade vs. threshold level for two PDFs: Gamma-Gamma and Log-Normal distributions and for three receiver diameters D and zenith angle of 45 degrees. The case $D = 0$ represents a point receiver.

probability is also reduced by increasing the aperture. For gamma-gamma distribution with a 6-dB threshold level, for example, the fade probability is decreased from 2.2×10^{-4} for a point receiver to about 2×10^{-5} for an aperture diameter of 60 cm. The similar conclusion is drawn for the log-normal distribution where the fade probability is decreased from 4.5×10^{-5} for a point receiver to about 1.8×10^{-6} for a diameter of 60 cm.

Decreasing the zenith angle also improves the predicted fade probability in the presence of atmospheric turbulence. Figure 11 depicts the probability of fade as a function of threshold level for three zenith angles: 0, 45 and 60 degrees, for both gamma-gamma and log-normal distributions. To achieve a probability of fade of 10^{-9} with a gamma-gamma distribution, for example, the threshold levels required increase from 8.96 dB for 0 degree to 9.7 dB for 60 degrees. The similar conclusion is valid also

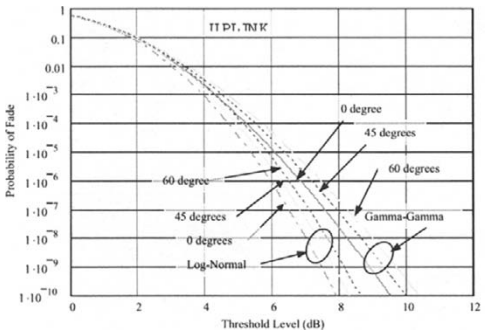


Fig. 11. Uplink Probability of fade vs. threshold level for three zenith angles: 0, 45 and 60 degrees, and for two PDFs: Gamma-Gamma and Log-Normal distributions.

for log-normal distribution. The log-normal distribution needs lower threshold levels than those for gamma-gamma distribution for a fixed probability of fade value.

4.2. Example 2: Downlink (Plane Wave/OOK)

For a downlink path from a satellite to a ground terminal, the ground-level scintillation near the center of the received wave can be accurately modeled by a plane wave [12]. The Rytov variance in this case depends mostly on high-altitude turbulence, and is consistent with weak-fluctuation theory except for the case of very large zenith angle. We evaluated the system performance by again calculating the average bit error rate $\langle \text{BER} \rangle$ for a plane wave using OOK modulation scheme. The same values of the parameters A and v for the H-V atmospheric model as used for the uplink case were taken. The fixed system parameters were also taken to be the same as for the uplink case, except for transmitter divergence = 5.1 μrad, data rate = 2 Gbit/s, receiver aperture diameters $D = 0$ (point receiver), $D = 10$ and 40 cm. Figure 12 shows the numerical results of the mean bit error rate as a function of average signal-to-noise-ratio $\langle \text{SNR} \rangle$. As in the uplink case, it is again apparent that the atmospheric turbulence degrades the system performance: for $\langle \text{SNR} \rangle$ of 10.8 dB, the BER increases from 10^{-9} (no turbulence case) to about 10^{-4} with $D = 40$ cm, and to about 2×10^{-4} with a point receiver ($D = 0$). Increasing the aperture helps in improving the system BER, for example, one needs approximately 16.21 dB available $\langle \text{SNR} \rangle$ while 16.77 dB of $\langle \text{SNR} \rangle$ is required for a point receiver to achieve the same mean bit error rate $\langle \text{BER} \rangle$ of 10^{-9}. Figure 13 shows the values of $\langle \text{BER} \rangle$ as a function of $\langle \text{SNR} \rangle$ for different zenith angles: $\xi = 0°$, $45°$ and $60°$. A fixed aperture diameter of $D = 40$ cm was assumed. The BER is increased from $10^{-9}$to approximately 2.8×10^{-9} if we increase the zenith angle from $0°$ to $60°$ for available $\langle \text{SNR} \rangle$ of 15.75 dB. Also in order to achieve the same $\langle \text{BER} \rangle$, say, 10^{-9}, the $\langle \text{SNR} \rangle$ has to be increased from 15.75 dB at $0°$ zenith angle to 16.38 dB at $60°$ zenith angle.

The probability of fade variation is similar to the uplink case. Figure 14 shows the probability of fade as a function of threshold level in dB for two aperture diameters,

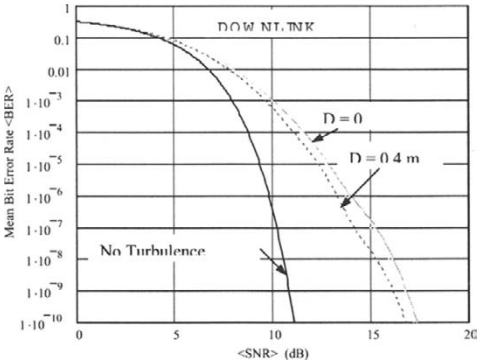


Fig. 12. Downlink Mean bit-error-rate for on-off keying direct detection as a function of mean signal-to-noise-ratio for two receiver aperture diameters and the zenith angle of 45 degrees. The no-turbulence case is also shown.

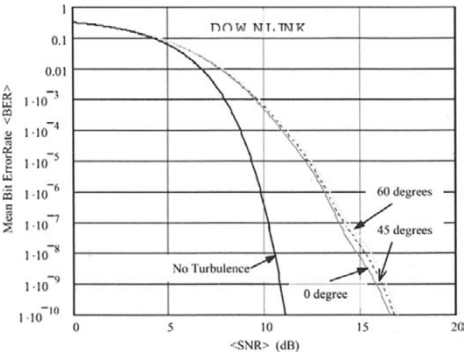


Fig. 13. Downlink Mean bit-error-rate for on-off keying direct detection vs. mean signal-to-noise-ratio for three different zenith angles: 0, 45 and 60 degrees, and aperture diameter $D = 40$ cm. The no-turbulence case is also shown.

$D = 0$ cm (point receiver), $D = 40$ cm. A fixed value of the zenith angle $\xi = 45°$ was assumed. The two probability density functions for irradiance fluctuations were assumed: Gamma-Gamma and log-normal. The bit error rate values seem to be lower for log-normal distribution than the Gamma-Gamma distribution. For example, for a threshold level of 8 dB and for a 40-cm-diameter aperture, the mean bit error rate is increased from 2×10^{-9} for log-normal distribution to the value 1.6×10^{-7} if one assumes Gamma-Gamma distribution. Increasing the aperture diameter also decreases the probability of fade, as expected, for example using a Gamma-Gamma distribution model, and for a fixed threshold level of 9.6 dB, the probability of fade is decreased from 3×10^{-8} for a point receiver ($D = 0$) to a value of 10^{-9} if we use a

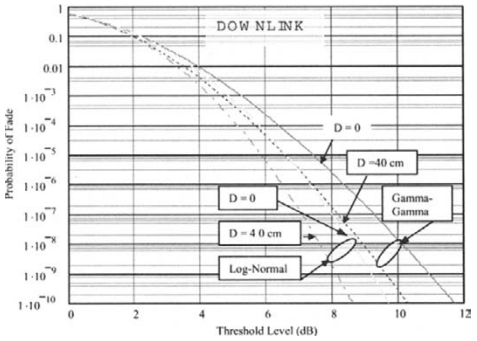


Fig. 14. Downlink Probability of fade vs. threshold level for two PDFs: Gamma-Gamma and Log-Normal distributions, two receiver diameters and zenith angle 45 degrees.

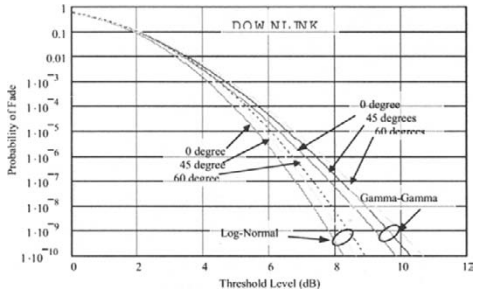


Fig. 15. Downlink Probability of fade vs. threshold level for three zenith angles: 0, 45 and 60 degrees, and for two PDFs: Gamma-Gamma and Log-Normal distributions. The receiver diamter $D = 40$ cm.

diameter of 40 cm. The similar trend follows for log-normal distribution. Finally, Fig. 15 shows the probability of fade as a function of threshold level in dB for three zenith angles $\xi = 0°, 45°$ and $60°$. The effect of using both Gamma-Gamma and log-normal distributions are also depicted in the same figure. The similar trends are also observed: for a fixed threshold level, increasing the zenith angle causes the probability of fade to increase (approximately even by an order of magnitude) for both the distributions, thus degrading the system performance at higher zenith angles.

4.3. Example 3: Horizontal: Terrestrial Link (Gaussian Beam Wave/OOK)

We provide here results for a numerical example of the average bit error rate (BER) of a horizontal link such as between buildings or mountaintops. This direct-detection binary lasercom link consists of a laser transmitter, atmospheric channel containing clear-air turbulence, and a maximum-likelihood receiver as described by other researchers.

[14] The maximum-likelihood receiver consists of an avalanche photodiode (APD) detector, a matched filter implemented in the form of a moving integrator, and a clock comparator which performs a threshold test at bit interval boundaries. The excess noise for the APD is characterized by a factor, F given by

$$F = k_{\text{eff}} \cdot G + (1 - k_{\text{eff}})(2 - 1/G), \tag{81}$$

where G is the average current gain, k_{eff} is the ratio of the ionization coefficient for holes to electrons. The expression for BER for no turbulence is given by [13]

$$\begin{aligned} \text{NoTurbulence} : \text{BER} &= \frac{1}{2}\left(P_{\text{FAlSEAlarm}} + P_{\text{MISS}}\right) \\ &= Q\left(\frac{y_{\text{Th}} - < y^0_{\text{Tb}} >}{2\sigma_0}\right) + Q\left(\frac{y_{\text{Th}} - < y^1_{\text{Tb}} >}{2\sigma_1}\right), \end{aligned} \tag{82}$$

where $Q(x)$ is the Gaussian tail integral defined by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(\frac{-t^2}{2}\right) dt = \frac{1}{2}\text{erfc}\left(\frac{x}{\sqrt{2}}\right). \tag{83}$$

The decision threshold value y_{Th} above is solved from a quadratic equation:

$$\begin{aligned} \left(\frac{\sigma_1^2}{\sigma_0^2} - 1\right) y_{\text{Th}}^2 + 2\left(\langle y^1_{\text{Tb}}\rangle - \frac{\sigma_1^2}{\sigma_0^2}\langle y^0_{\text{Tb}}\rangle\right) y_{\text{Th}} - \sigma_1^2 \ln\left(\frac{\sigma_1^2}{\sigma_0^2}\right) \\ + \frac{\sigma_1^2}{\sigma_0^2}\langle y^0_{\text{Tb}}\rangle^2 - \langle y^1_{\text{Tb}}\rangle^2 = 0, \end{aligned} \tag{84}$$

where $\langle y^1_{\text{Tb}}\rangle$ and σ_1^2 are the mean and variance respectively of the moving integrator output of the Gaussian APD photocurrent under ON state, and $\langle y^0_{\text{Tb}}\rangle$ and σ_0^2 are the same parameters for OFF state. The BER for turbulence is given by [14]

$$\text{Turbulence} : \text{BER} = \frac{1}{2}\left(P_{\text{FAlSEAlarm}} + P_{\text{MISS}}\right),$$

where

$$\begin{aligned} P_{\text{FAlSEAlarm}} &= \int_0^\infty Q\left(\frac{y_{\text{Th}} - eG\varepsilon n_\phi}{\sqrt{2B(n_\phi \varepsilon e^2 F G^2 + 2K_B T/R_L)}}\right) \frac{1}{S\sqrt{2\pi A\sigma^2_{\ln S}}} \\ &\quad \times \exp\left(\frac{-(\ln S + \frac{1}{2}A\sigma^2_{\ln S})^2}{2A\sigma^2_{\ln S}}\right) dS \\ P_{\text{MISS}} &= \int_0^\infty Q\left(\frac{y_{\text{Th}} - eG\varepsilon n_\phi}{\sqrt{2B(n_\phi e^2 F G^2 + 2K_B T/R_L)}}\right) \frac{1}{S\sqrt{2\pi A\sigma^2_{\ln S}}} \\ &\quad \times \exp\left(\frac{-(\ln S + \frac{1}{2}A\sigma^2_{\ln S})^2}{2A\sigma^2_{\ln S}}\right) dS. \end{aligned} \tag{85}$$

Figures 16 and 17 depict the effects of turbulence on the BER for a direct-detection binary optical detection link. The following parameters were chosen for generating

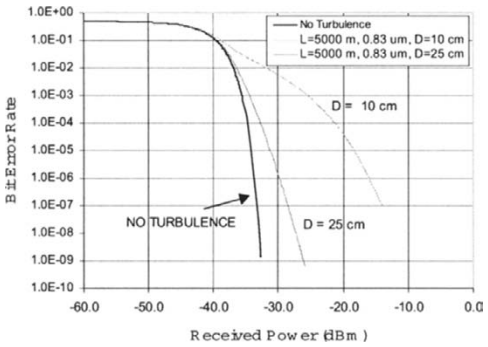


Fig. 16. BER vs. received power for $L = 5$ km, $= 0.83$ m, and for two aperture diameters $D =$ 10 cm, $D = 25$ cm.

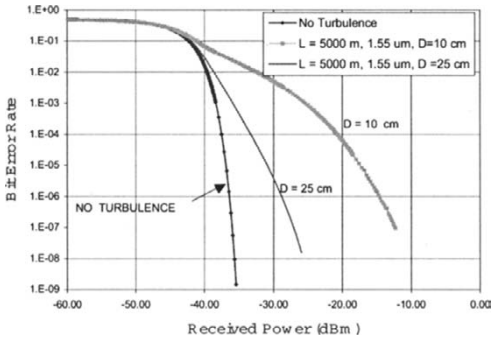


Fig. 17. BER vs. received power for $L = 5$ km, = 1.55 m, and for two aperture diameters $D =$ 10 cm, $D = 25$ cm.

the numerical results shown in the Figures 16 and 17: two laser wavelengths $\lambda = 0.830\mu$m and $\lambda = 1.55\mu$m, two aperture diameters $D = 10$ cm and $D = 25$ cm, communication range = 5 km, transmitter beam size ω_0 = 2.5 cm, data rate =1.25 Gbit/sec, efficiency η= 0.8, the strength-of-turbulence parameter $C_n^2 = 1.2 \times 10^{-14}$ $m^{-2/3}$ (assumed constant over the horizontal propagation path), current gain $G = 20$, $\kappa_{\text{eff}} = 0.03$, load resistance $R_L = 50$ ohm, temperature $T = 300$ K, and 1000:1 modulator extinction ratio ($\varepsilon = 0.001$). The aperture-averaged scintillation index $\sigma^2_{\ln S}$ was calculated from the equations discussed in earlier section. The figures show the BER versus the average optical power P_{opt} at the receiver. As seen previously for uplink and downlink examples, the BER is again clearly increased by atmospheric turbulence and by decreasing the aperture diameter. For example, at the wavelength of 0.830 μm, the BER increases from approximately 10^{-9} (no turbulence value) to 10^{-4}

for a 25-cm-diameter aperture, and even to about 10^{-2} for a 10-cm-aperture diameter. These values are taken for the received power of -30 dBm. For the wavelength of 1.55 μm, compared to no turbulence BER of 10^{-9}, the BER increases to approximately 5 $\times 10^{-4}$ for a 25-cm-diameter aperture, and 1.5 $\times 10^{-2}$ for a 10-cm-diameter aperture, all the values at the received power of -35 dBm. Atmospheric turbulence is therefore the dominant cause for the system degradation affecting BER to increase several order of magnitudes. The curves will be helpful for designing the communication system to operate under atmospheric turbulent conditions.

4.4. Example 4: Downlink PPM Laser Communication in Presence of Atmospheric Turbulence and Multiple Scattering Media

4.4.1. Theory and Formulation of BER for PPM with Stretched Pulses for Pulse Position Modulation

In this example, numerical examples of degradation in the bit-error rate (BER) of a multi-Gigabit, pulse position modulation (PPM) laser downlink due to turbulence and multiple scattering media is presented. Multiple scattering of optical wave occurs for a variety of atmospheric conditions, such as cloud, haze, fog and rain. The greatest degradations are scintillation which causes temporal fluctuations of the signal count at the receiver, pulse stretching of the initially transmitted short pulse and random propagation time delays. Severe turbulence and multiple scattering in the communication channel may lead to unacceptable symbol error probabilities. The impulse responses of both the turbulent and multiple scattering channels have been computed and their effects on the data rate are developed. An example of a solid state laser transmitter and a high-speed detector shows the engineering feasibility and discusses practical limitations and problems associated with a multi-Gigabit data rate for downlink atmospheric laser communication channel.

Pulse position modulation (PPM) in atmospheric optical communication has been used in the past. [25,27–29] The PPM's primary advantage is the ability to communicate at a high data rate over noisy channels. For multi-Gigabit downlink, very short, highly peaked laser pulses are transmitted which encounter the following effects after propagating through the atmosphere: pulse spreading (temporal broadening), scintillation and the random arrival times. These will result in increasing BER for data encoded in the PPM format in a space-to-ground downlink. The receiver on the ground has to decode the pulses by detecting the correct slot position in a given word. Figure 18 shows a M'ary PPM format where there are M different waveforms, each is represented by a pulse in each of M time slots corresponding to distinct symbols. The stretched (and distorted) pulse is shown in the slot q. In vacuum, the pulses should fall exactly within each time slot without any overlap into other slots or frames. Because of the spread due to atmospheric effects, an overlap into other time slots may thus produce inter-symbol interference, and an overlap into other frames may produce inter-frame interference. The data rate, R is given by

$$R = \frac{\log_2 M}{MT_S + T_d}, \tag{86}$$

where T_S is the duration of any of the M time slots, $T_d$is the frame dead time after the last slot. In order to estimate the correct slot location of the transmitted pulse and corresponding symbol, the receiver detects a "greatest of" photoelectron counts in each

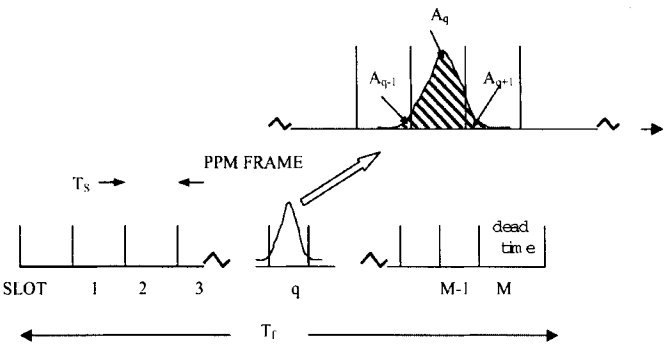


Fig. 18. Expected distorted receive pulse in the qth time slot of a PPM frame, showing the expected slot pulse areas.

of the M time slots over this time T_S. When the pulse is stretched or distorted, the counts in the qth slot will decrease and the counts in the adjacent slots will be increased. This will cause the receiver to incorrectly decode transmitted symbols resulting in an increase in BER.

We now evaluate the BER performance for such PPM lasercom system. The exact procedure and the necessary equations are described in details elsewhere. [30,31] With a matched filter at the receiver to be a pure slot integrator, the random sampled output count for the jth slot is

$$r_j = \int_{(j-1)T_S}^{jT_S} P(t)dt + n_j, \tag{87}$$

where $P(t)$ is the instantaneous temporal pulse shape of the received signal at the detector output, and n_j is the combined background and receiver noise count in this slot for a particular sampling. Assuming the pulse is always at the qth time slot, a long series of frames is sent. Neglecting inter-frame and edge effects at the first and Mth slots, the result will be equivalent to randomly placing the pulses in individual frames. Each slot will have some expected count that will depend on the expected pulse shape, pulse amplitude and arrival time. If we denote these expected slot counts be

$$\overline{m_j}, \quad j = 1, 2, \ldots, q, \ldots, M. \tag{88}$$

When the turbulence and scattering effects are significant, these expected counts are time varying, giving rise to conditional counting processes in each time slot. Thus the expected counts m_j will be random variables, in general. In particular frames, the counts in individual slots will usually deviate from their expected values. A symbol error will occur if

$$r_j > r_q, \quad j = 1, 2, \ldots, M, \quad j \neq q. \tag{89}$$

The probability of symbol error (PSE) in a frame can then be expressed as

$$\text{PSE} = \sum_{r_j = 1, j \neq q}^{M} \text{Prob}(r_j > r_q). \tag{90}$$

Assuming the mean count is a random variation in all the slots and is continuous over the range $[0, \infty]$, the total probability of the count in a slot j exceeding the count in slot q is

$$\text{Prob}(r_j > r_q) = \int_0^\infty P(m_q) \int_0^\infty P(m_j) \sum_{r_j=1}^{\infty} P(r_j \,|m_j) \sum_{r_q=0}^{r_j-1} P(r_q \,|m_q)\, dm_j dm_q. \tag{91}$$

In the above equation, $P(m_q)$ and $P(m_j)$ are the probability densities of the expected slot counts over the continuous range. Similar conditional Poisson count was used by Churnside and Hill. [32] Substituting Eq. (??) into Eq. (??), we obtain

$$\text{PSE} = \sum_{j=1, j\neq q}^{M} \int_0^\infty P(m_q) \int_0^\infty P(m_j) \sum_{r_j=1}^{\infty} P(r_j \,|m_j) \sum_{r_q=0}^{r_j-1} P(r_q \,|m_q)\, dm_j dm_q. \tag{92}$$

The BER can now be obtained from

$$\text{BER} = \frac{M \log_2 M}{2(M-1)} \text{PSE}. \tag{93}$$

If the discrete counting process is a conditional Poisson, given by

$$P(r_k \,|m_k) = \frac{m_k^{r_k}}{r_k!} e^{-m_k} \quad k = 1, 2, \ldots, M. \tag{94}$$

We can now write the expression for BER as

$$\text{BER} = \frac{M \log_2 M}{2(M-1)} \sum_{j=1, j\neq q}^{M} \times \int_0^\infty P(m_q) \int_0^\infty P(m_j) \sum_{r_j=1}^{\infty} \frac{m_j^{r_j}}{r_j!} e^{-m_j} \sum_{r_q=0}^{r_j-1} \frac{m_q^{r_q}}{r_q!} e^{-m_q} dm_j dm_q. \tag{95}$$

Next we determine the functional forms of the mean count probability densities resulting from turbulence and multiple scattering effects and then evaluate communication performance for optical downlink.

In order to evaluate the temporal spreading of the initially very short transmitted pulse, we need to know the transfer function characteristics of the atmospheric channel. This transfer function can actually be obtained from the convolution of $G_1(t - z/c)$ with $G_2(t - z/c)$ where $G_1(t - z/c)$ corresponds to the *high-frequency* component of the transfer function, and $G_2(t - z/c)$ corresponds to the *low-frequency* one. The reason for the separation of the G function into these two components comes from the fact that the output pulse can be expressed in terns of the two-frequency mutual coherence function (MCF), Γ, given by [33]

$$\Gamma = \Gamma(\omega_1, \omega_2). \tag{96}$$

The impulse response function of the random medium and the MCF is related by

$$G(t - t') = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Gamma(\omega_d) e^{-j\varpi_d(t-t^-)} d\omega_d, \tag{97}$$

where ω_d represents the difference between the angular frequencies, ω_1 and ω_2.

The expected instantaneous output power resulting from a given input pulse, $P_i(t)$ is

$$P_0(t) = \int_{-\infty}^{\infty} P_i(t')G(t-t')dt'. \tag{98}$$

The expected impulse response can be expressed as (for optical propagation along z direction)

$$G(t-z/c) = G_1(t-z/c) \otimes G_2(t-z/c) = \int_{-\infty}^{\infty} G_1(t-z/c-t')G_2(t')dt', \tag{99}$$

where c is the velocity of light, and $\otimes$ represents a convolution.

The first function is [31,33]

$$G_1(t-z/c) = \frac{\pi}{4T_1}\sum_0^{\infty}(-1)^n(2n+1)\exp\left\{-(2n+1)^2\frac{\pi^2}{16}\frac{(t-z/c)}{T_1}\right\}, \tag{100}$$

where $T_1 = \left(\frac{1}{1.28c}\right)C_n^{12/5}k_0^{2/5}L^{11/5}$, with L being the propagation length with turbulence, $k_0 = 2\pi/\lambda$ is the wave number of the optical signal.

The second function to be convolved to get the impulse response is

$$G_2(t-z/c) = \frac{1}{\sqrt{\pi}T_2}\exp\left\{-\frac{(t-z/c)^2}{T_2^2}\right\} \tag{101}$$

where

$$T_2 = \frac{1}{c}1.2056C_nL_0^{5/6}L^{1/2} \tag{102}$$

with L_0 = outer scale size of turbulence. The parameters T_1 and T_2 were calculated for a given turbulence strength C_n^2. Thus G_1 and G_2 functions were evaluated. We will discuss the impulse response function for multiple scattering media later in this section.

4.4.2. BET-Error-Rate (BER) and Slot Count Statistics

The received pulse power over the effective aperture area of the telescope is a random function which depends upon the impulse response $G(t)$and the background noise of the receiver. Let the received pulse power per unit area be

$$P(t) = IS(t), \tag{103}$$

where I is a random intensity function and $S(t)$ is defined as

$$S(t) = \frac{\gamma}{\langle I\rangle}\int_{-\infty}^{\infty} P_i(t')G(t-t')dt', \tag{104}$$

where $P_i(t')$ is the input pulse power, γ is a parameter that depends on the full width transmitter beam angle, θ_T, and the range, L, and is defined as

$$\gamma = \frac{1}{\theta_T^2 L^2}. \tag{105}$$

$\langle I \rangle$ is the average or expected mean value of the intensity random process, I. The expected instantaneous count for the slot q, which contains the information signal will then be

$$m_q = \alpha A_{\text{rec}} \int_{(q-1)T_S}^{qT_S} I\, S(t)\, dt + n_q, \tag{106}$$

where $\alpha = \frac{\eta}{h\nu}$ is a constant, η is the detector quantum efficiency, h is the Planck's constant, ν is the optical frequency, A_{rec} is the receiver aperture area, and n_q is the combined background and receiver noise count at the qth slot. Following the procedure as shown in [30], we can obtain the following expression for the pdf of m_j (if we know the pdf for P_i with the noise mean pdf given by $P_{nq}(n_q)$ as shown below):

$$P(m_q) = P_I\left(\frac{m_q}{\alpha A_{\text{rec}} A_q}\right) \otimes P_{nq}(n_q). \tag{107}$$

The term A_q can be expressed as [31]

$$A_q = \frac{\gamma}{\langle I \rangle} \int_{(q-1)T_S}^{qT_S} \int_{-\infty}^{\infty} P_i(t') G(t - t') dt' dt. \tag{108}$$

After some mathematical manipulation the area A_q can be simplified to

$$A_q = \frac{\gamma P_t\, t_p}{2\langle I \rangle}, \tag{109}$$

where t_p is the duration of the pulse with a constant power P_t within this pulse duration. Assuming the expected noise count, n_q, in each slot is constant, then the pdf, $P_{nq}(n_q)$ must be a delta function of the form $\delta(m_q - n_q)$, so that we can finally write

$$P(m_q) = \int_0^{\infty} P_I\left(\frac{m_q - m_q'}{\alpha A_{\text{rec}} A_q}\right) \delta(m_q' - n_q) dm_q', \tag{110}$$

which implies that

$$P(m_q) = P_I\left(\frac{m_q - n_q}{\alpha A_{\text{rec}} A_q}\right). \tag{111}$$

4.4.3. The Impulse Response Function in the Case of Multiple Scattering

While the turbulent medium is considered to be a continuous random medium, multiple scattering occurs for a discrete random medium whereas the medium consists of scatterers of random sizes, shapes, orientations, velocities and densities. For multiple scattering processes, the mutual coherence function can be written as

$$\Gamma(w_d) = \Gamma_1(w_d) \exp\left\{-\left[\frac{1 - w_0}{w}\right]\tau\right\}, \tag{112}$$

where w_0 is the albedo of a scatterer, $\tau = \rho_n \sigma_s z$ is the optical distance for a path z, and σ_s is the scattering cross section. Note that albedo is defined as $w_0 = \frac{\sigma_s}{\sigma_t}$ where σ_t is the total cross section of the random medium which is equal to $(\sigma_s + \sigma_a)$ where σ_a is the absorption cross section of the randomly distributed particles. Since $G_1(t - z/c)$ depends on $\Gamma_1(w_d)$ only, it remains same for multiple scattering. $G_2(t - z/c)$, however, depends on the exponential term of the above equation, so that we can write

$$G_2(t - z/c) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left\{ -\left[\frac{1 - w_0}{w_0}\right] \tau \right\} e^{-iw_d(t - z/c)} dw_d, \tag{113}$$

$$G_2(t - z/c) = \delta(t - z/c) \exp\left\{ -\left[\frac{1 - w_0}{w_0}\right] \tau \right\} \tag{114}$$

for discrete random media involving multiple scattering. [33,34]

The impulse response function, $G(t - z/c)$ applicable to multiple scattering media can be finally written as

$$G(t - z/c) = \frac{\pi}{4T_1} \exp\left\{ -\left[\frac{1 - w_0}{w_0}\right] \tau \right\} \sum_0^{\infty} (-1)^n (2n + 1) \times \exp\left\{ -(2n + 1)^2 \frac{\pi^2}{16} \frac{(t - 2z/c)}{T_1} \right\}. \tag{115}$$

The expected pulse area will be different for this multiple scattering case and is given by

$$A_q = \frac{\gamma P_t \, t_p}{\langle I \rangle} \exp\left\{ -\left[\frac{1 - w_0}{w_0}\right] \tau \right\}. \tag{116}$$

The BER is obtained as outlined in [31] and is given by

$$\text{BER} = \frac{M \log_2 M}{2(M - 1)} \sum_{j=1, j \neq q} \exp\left[-(\alpha A_{\text{rec}} I_{\text{av}} (A_j + A_q)) \right] \times \sum_{r_j = 1, j \neq q}^{\infty} \frac{(n_j + I_{\text{av}} \alpha A_{\text{rec}} A_j)^{r_j}}{r_j!} \sum_{r_q = 0}^{r_j - 1} \frac{(n_q + I_{\text{av}} \alpha A_{\text{rec}} A_j)^{r_q}}{r_q!}, \tag{117}$$

where $I_{\text{av}} = I_q = I_j$ is the average constant power per unit area amplitude of the intensity random process. The equation for BER was evaluated numerically and the results are shown below.

4.4.4. Numerical Results

The G_1, G_2, and G impulse response functions for both turbulence and multiple scattering media and BER computation. Based on the theoretical background described earlier, the parameters T_1 and T_2 were calculated for both weak ($C_n^2 = 10^{-18}$ $m^{-2/3}$) and strong ($C_n^2 = 10^{-14}$ $m^{-2/3}$) turbulence cases. Using these values, the G_1 and G_2 functions were evaluated for both turbulence and multiple scattering media. Figure 19 shows the functions G_1, G_2 and G for strong turbulence case and Fig. 20 shows the function G_1 functions for various atmospheric conditions of dense rain, light fog, fog and cloud. The corresponding T_1 values are also shown in the figures. To calculate G_2 for multiple scattering media, we need to know the albedo w_0 and the optical depth $\tau$of the scattering media whose values were assumed as follows: fog (w_0= 0.6,τ= 4) and cloud (w_0= 0.8, τ= 20).

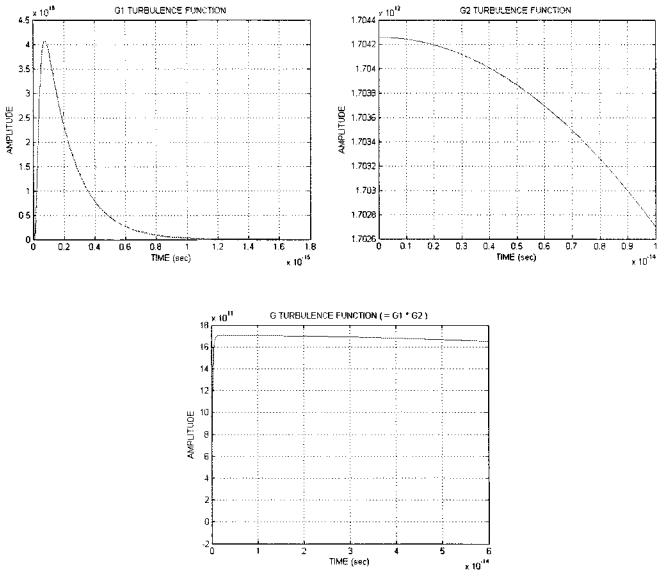


Fig. 19. The functions G_1, G_2 and G for a strong turbulence case ($C_n = 10^{-7}$ m$^{-2/3}$).

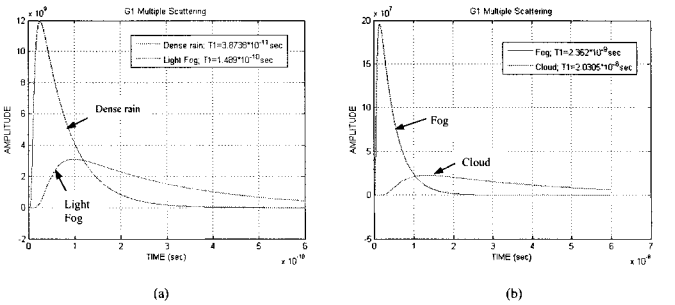
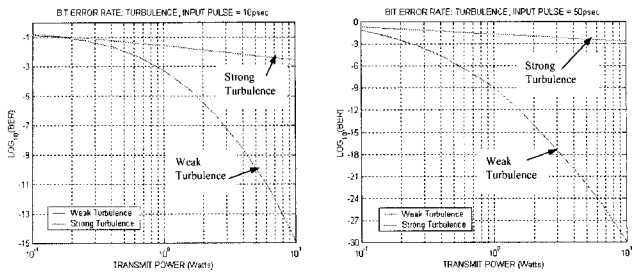


Fig. 20. The G_1 functions for multiple scattering media. The figures show for various atmospheric conditions of (a) dense rain and light fog, (b) fog and cloud.

4.4.5. BER Computation for Both Turbulence and Multiple Scattering Media

The following parameters were used for calculating BER for both turbulence and multiple-scattering media: laser wavelength $\lambda = 1.06$ μm, detector efficiency $\eta = 0.2$, effective receiver aperture area $A_{\text{rec}} = 0.78554$ m^2, $T_s = 167$ ps for 2 Gbit/s data rate, transmitter power $P_t = 1$ W, optical path transmissivity $\tau_a = 0.8$, transmitter optical efficiency $\tau_T = 0.9$, receiver optical efficiency $\tau_R = 0.9$, pointing loss $\tau_L = 1$, transmit beam angle $\theta_t = 15$ μrad, range $L = 3.6 \times 10^7$ m. The input Gaussian



(a) (b)

Fig. 21. BER vs. required transmitter power for weak and strong turbulence cases and for two input laser pulses: (a) 10 psec, (b) 50 psec. For both turbulence cases, the parameters used for the calculations are: $\alpha = 1.0668 \times 10^{18}$, $A_{\text{rec}} = 0.7854$, $\theta = 15 \times 10^{-6}$ rad, range $L = 3.6 \times 10^{7}$ m.

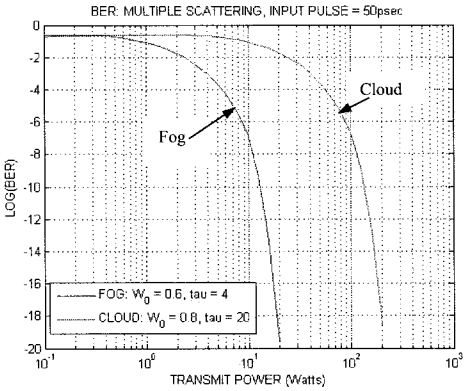


Fig. 22. BER vs. required transmitter power for multiple scattering media (fog and cloud) for an input laser pulse of 50 psec. The parameters used for the calculations are $\alpha = 1.0668 \times 10^{18}$, $A_{\text{rec}} = 0.7854$, $\theta = 15 \times 10^{-6}$ rad, $L = 3.6 \times 10^{7}$ m.

pulse with FWHM = 10 ps and 50 ps for two turbulence cases and 50 ps for multiple scattering media were assumed. Figure 21 shows the BER as a function of required transmitter power for both weak and strong turbulence, and Fig. 22 shows the same plots for fog and cloud using an input laser pulse of 50 ps. For the weak turbulence case using a 10-ps laser pulse, the required laser transmitter power to achieve a BER of 10^{-7} is approximately 3 W of laser power while for strong turbulence case the same power will achieve much increased BER of 2×10^{-3}. When a laser input pulse of 50 ps is used, it requires about 700 mW of laser power for weak turbulence case, whereas with this power, a BER of 3×10^{-4} can be achieved for strong turbulence case. For multiple scattering media with an input laser pulse of 50 ps, it requires a laser power

of approximately 12 W for fog and 120 W for cloud to achieve the same BER of 10^{-8}. A high-data-rate lasercom system downlink with transmitting short laser pulse of 50 ps and reasonable laser power is thus feasible with the PPM scheme described in this scheme.

5. Other Examples/Scenarios of Free-Space Optical and Laser Communications

5.1. Ground to Space Shuttle Link

Researchers have presented results from experimental demonstration using a very lightweight optical wavelength communication without a laser in space (LOWCAL) [35] between ground based telescope and a Space Shuttle. This link established an up-link/downlink carrier at 852-nm wavelength and 852-nm signal beam. Both the transmitting laser and the return link receiver were located on the ground terminal, while the optical elements located on the spacecraft were the liquid crystal retro-modulator and a simple forward link receiver. Some of the specifications for the experiment were: range = 640 km, data rate = 10 kbps, telescope diameter = 0.6 m, modular weight = 2–4 kg, and the retro-reflector area of 70–180 cm^2. The diode laser transmitter power was 0.2 W. The transmitted linearly polarized laser beam was directed through a quarter-wave plate converting the beam into a circularly polarized beam. The beam is then directed to the retro-reflector type modulator which directs the beam incident on the spacecraft directly back to the transmitting telescope. For downlink a differential circular polarization keying (DCPK) format was utilized where the binary data was encoded in the polarization of the beam. The uplink communication link used sub-carrier FSK modulation. The downlink receiver achieved a SNR of 138, and a BER $< 1.67 \times 10^{-7}$, whereas the uplink receiver had SNR of $\sim$ 1000, and BER $< 1.67 \times 10^{-7}$. The system thus had unique features such as no laser in space required, all pointing was on the ground and full-duplex operation with one laser beam.

5.2. UAV-to-Ground Lasercom Link

A group of researchers reported the design and development of acquisition, tracking and pointing subsystem for Unmanned-Aerial-Vehicle (UAV) to ground free-space optical communications link. [36]. The communication link was developed from a UAV to stationary ground stations located at Wrightwood, California and Maui, Hawaii. The range of the UAV was 50 km. The downlink laser transmitter wavelength was 1550 nm and had a power of 200mW for developing a 2.5 Gigabit per second data rate communication for a BER of $1E-9$. Analysis presented to demonstrate that the system met the requirements of a jitter error of 18.6 μrad and a bias error of 2.2 μrad .

5.2.1. Indoor Optical Communication

All of the design and performance characteristics described so far for long distance apply equally to short distance communication systems with one obvious exception—the effect of atmospheric loss is almost none at all for indoor systems. The link budget

for power is therefore determined almost entirely by the transmitter power, free space loss, and receiver sensitivity. The basic operating principles of indoor point-to-point systems are not different from the outdoor communication links, the designs are very different to accommodate various requirements. For example, the optical source must be eye safe, which generally means that it is an LED, which in turn limits the data transfer capacity to typically a few Mbit/s. On the other hand, indoor systems do not have to face atmospheric effects as the outdoor systems require, so that they can be produced very cost effectively. Finally, the indoor applications are basically confined to short distance spans and hence are appropriate to optical wireless systems that utilize wide diverging beams rather than the narrow beams of point-to-point systems. Each "cell" created by an optical "telepoint" base station can be shared by many users. In some cases, diffuse paths are utilized, where the beams, radiating over a wide angle, are also allowed to reflect off walls, ceiling, floor, furniture, etc. The penalty of diffuse transmission is a much reduced data rate capacity compared to line-of-sight systems.

The short-range indoor wireless communication is receiving increasing attention because of its role in the emerging technologies involving portable computing and multimedia terminals at work and everyday environment. Some of the typical portable devices may include laptop computers, personal digital assistants, and portable telephones, while the base stations are usually connected to a computer with other networked connections. Many indoor communication systems employing infrared LED wireless links have been reported. [37–39] Using pure diffuse links, a high-speed, power-efficient indoor wireless infrared communication using code combining is reported [40] where a multiple transmitter link design was used with a narrow field-of-view direction diversity receiver. The design goal was to eliminate the effect of inter-symbol interference (ISI) so that power-efficient signaling schemes such as PPM can be employed at very high data rates. Recently, a portable transceiver for indoor wireless link is reported. [41] The system employs a transmitter of eye-safe infrared LEDs and a receiver of photodiode arrays with multi-channel trans-impedance-summer architecture. A wide field of view (up to 55 deg half-angle) and high speed (up to 35 MHz) was attained to support different intensity modulation schemes. The received signal achieved a BER of 10^{-4} at a plane of 2 m away from the transmitter, even at a point $\pm$ 50 deg off the transmitter's vertical axis. The bit rate of the transceiver was up to 40 Mbit/sec in an indoor non-directed infrared wireless link to be extended to 100 Mbit/s using LEDs with higher cut frequency. A receiver angle diversity design for high-speed diffuse indoor wireless communication is reported [42] where the effects of multipath distortion was compensated but at the expense of an increased path loss. A (6m $\times$ 3m $\times$ 3.4 m) upper illuminated room was used with a receiver placed 1 m above the floor was used. Their receiver structure consisted of 25 non-overlapping 18-degree FOV receiver elements, each element having an effective collection area of 40 cm^2, using 1 cm^2 detectors. With a 60 mW diffuse optical source, and utilizing non-return-to-zero (NRZ) on-off keying modulation scheme at 50 Mbit/s, the performance achieved was BER of $> 10^{-4}$ for normal illumination and BER of $< 10^{-7}$ for upper illumination across most of the coverage area.

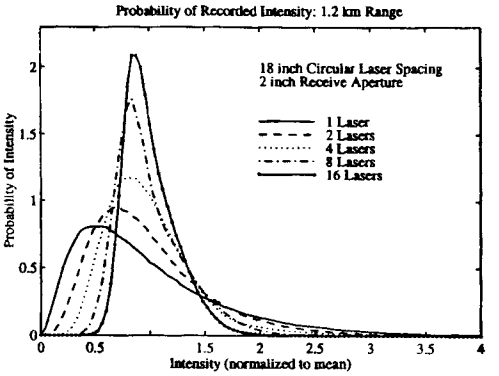
5.2.2. Free-space Optical Interconnect

This paper will be incomplete without the mention of one other type of extremely short range optical communication which uses free-space, such as data links between

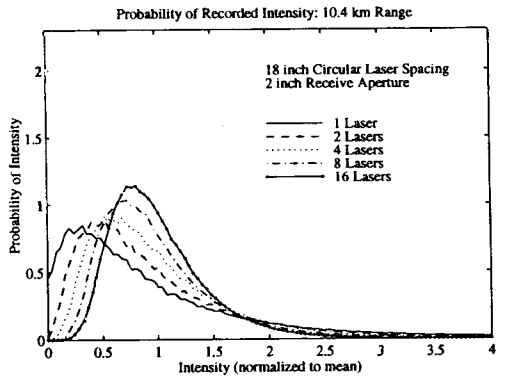
chips, multi-modules and boards. Optics is becoming a mature technology for interconnects and free-space optical interconnects are seen as well-suited for the chip-level and board-level interconnects. Several interconnect architectures based on arrays of VCSELs have been proposed. Some of the problems to overcome include minimization of optical channel crosstalk arising from the laser beam diffraction, and the misalignment between VCSELs, microlenses and photo detectors. Chip-to-chip optical communications through the air could boost data throughput between chips by a factor of a thousand. [43] Optical communications will the computer, connecting one circuit board to another, and chip-to-chip communications will the market. Free space optical interconnects links operating at 2.5 Gbit/s between VCSEL arrays and suitable detector arrays have been demonstrated. [44] The researchers developed and demonstrated the operation of a fully packaged optical interconnect system for multi-chip interconnections capable of sustaining channel data rates as high as 800 Mbit/s.

6. Multiple Transmitters/Receivers Approach for Lasercomm

The importance of using multiple transmitters and receivers was pointed out [45] many years ago, with an example of laser communications between aircraft showing that the system performance can be improved by using wide divergent beams. Wide divergent beams can be used to eliminate the need for complex gimbaled pointing and tracking mountings. For some very long range communication systems, the light from one single source (LED or a laser diode) may not be enough to have the necessary link margin to achieve the required BER. A large lens used in conjunction with a single light source may result in a light beam that is too narrow to be practical in some cases. The divergence angle may be so small that keeping the transmitted light aimed at the distant receiver may not be practical. To launch more light at the distant receiver, sometimes it is desirable to use multiple light sources. The multichannel approach was chosen because it can increase the range of communications in presence of background radiation more effectively than using a single-channel system with the same total FOV. [45] The multichannel approach can also help in the alignment procedure in a situation such as laser communications between aircraft where multiple LEDs were used, each with its own lens. With multiple light sources and lenses, it is also easier to aim at the distant receiver. For laser communication between a satellite and a ground-based terminal, uplink scintillation is a serious issue. Fluctuations on the uplinked beacon and the communication lasers can be minimized by transmitting multiple independent lasers from separate apertures which then sum incoherently at the satellite. In order to mimic the atmospheric effects for an uplink slant path to a satellite, experimental results from a horizontal link (for distances 1.2 and 10.4 km) were reported. [46] 1, 2, 4, 8, and 16 laser transmitters were located on an 18-inch-diameter circle at the two ranges of 1.2 and 10.4 km. For all laser separation and ranges, a reduction in received intensity fluctuations due to turbulence were observed as the number of laser transmitters was increased from 1 to 16. Figure 23 shows the histograms of the probability of intensity curves vs. intensity for the two paths (integrated probability was normalized to the value 1 and the mean intensity was taken as 1). As the number of laser transmitters was increased, the reduction in fluctuations was observed as can be seen from the decreases in variances (FWHM) of the individual histograms.



(a)



(b)

Fig. 23. Probability density function of intensity for varying number of laser transmitters (1, 2, 4, 8 and 16 lasers), and for two ranges (a) 1.2 and (b) 10.4 km. ([46], reprinted with permission from SPIE).

The reduction of scintillation by the use of multiple uplink beams, each incoherent with the rest, was demonstrated by a group of researchers. [47] The objective of the experiment was to establish a 1-Mbps optical communication link between a satellite and a ground terminal. The link was established by first transmitting a beacon from the ground station to the satellite and the scintillation effects were investigated by launching multiple beams during uplink. Figure 24 shows the expected pdf when the total laser power was equally distributed in 1, 2, 4, 8, or 16 beams. With increasing number of beams, the mean varied little, but the variance drops significantly with additional beams. The method of dividing the laser beam into a number of beams will help in the presence of strong scintillation to avoid deep fades and to improve BER statistics.

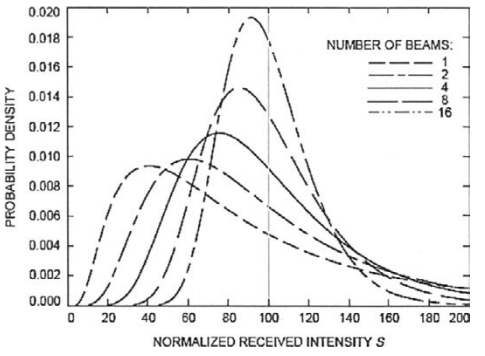


Fig. 24. The probability density functions of the received signal intensity for varying number of beams showing the improvement in variance with multiple uplink beams. The total power is equally distributed among 1, 2, 4, 8, or 16 beams and hence the PDFs are normalized by the total power.([47], reprinted with permission from JPL).

The advantage of using multiple receivers was already pointed out earlier to demonstrate the optical communications concept between aircraft. [45] The advantage of using a number of detectors each having small FOVs, and each with very narrowband optical filters effectively reduce the overall background noise as compared to using a single large FOV with a single detector. Thus multiple detectors can increase the received SNR. A theoretical analysis of an optical array receiver also showed [48] that an array receiver concept was a viable concept alternative to using a single large-aperture to collect the optical signal energy for high data rate applications. In presence of atmospheric turbulence and background interference, the multi-element telescope arrays observing pulse position modulated optical signals will be very desirable.

7. Conclusion

In this paper, we have provided the basics of free-space laser/optical communication system performance with a comprehensive treatment of the theoretical analysis needed for system design. We have investigated various communication scenarios of practical importance: horizontal (terrestrial) link and slant paths including downlink (from satellite-to-ground), and uplink (ground-to-satellite). This chapter discusses with detailed theoretical analysis the basic parameters for designing a laser communication system that has to operate under various atmospheric conditions, such as fog and turbulence. Numerical examples for various communication links under various conditions are given to illustrate the trends in how design parameters affect performance and to demonstrate how system design optimization is accomplished.

This paper has not considered adaptive optics for atmospheric turbulence mitigation or coding schemes, which can also significantly enhance system performance in many cases, but are beyond the scope of this chapter. We have also not covered the areas of heterodyne detection schemes and pointing and tracking.

It is obvious that the laser/optical communication is here to stay. With our increased knowledge of atmospheric effects and the physics of turbulence generation and interactions with optical waves, and the continuous push for developing more sensitive receivers, laser communication system performance will continue to improve, with increasingly higher data rates, accuracy, and availability under all-weather atmospheric conditions.

Acknowledgments

The author wishes to thank Jennifer C. Ricklin for providing data for bit-error-rate calculations for Gaussian beam wave. The author would like to acknowledge Larry C. Andrews, Ronald L. Phillips and Robert K. Tyson for very helpful technical discussions to clarify uplink-downlink theoretical equations. The author is deeply indebted to Fotios P. Kourouniotis for assistance in the computation and simulation of data. Helpful discussions with Laurence Chen and a critical reading of the manuscript by Bradley Bobbs are also acknowledged.

References

1. D.K. Killinger, J.H. Churnside, and L.S. Rothman, Atmospheric Optics, *OSA Handbook of Optics*, edited by M. Bass (1995), Chapter 44.
2. E.J. McCartney, *Optics of the Atmosphere* (Wiley, New York, 1969).
3. I.I. Kim, B. McArthur, and E. Korevaar, Comparison of laser beam propagation at 785 nm and 1550 nm in fog and haze for optical wireless communications, Proc. SPIE, Vol. 4214, p. 26–37, Optical Wireless Communications III, edited by Eric J. Korevaar (2000).
4. Brian R. Strickland, Michael J. Lavan, Eric Woodbridge, and Victor Chan, Effects of fog on the bit-error rate of a free-space laser communication system, Appl. Opt., **38**, 424–431 (1999).
5. S.G. Lambert and W.L. Casey, *Laser Communications in Space* (Artech, Boston, 1995).
6. R.H. Kingston, *Optical Sources, detectors, and Systems: Fundamentals and Applications* (Academic, San Diego, 1995).
7. W.S. Ross, W.P. Jaeger, J. Nakai, T.T. Nguyen and J.H. Shapiro, Atmospheric Optical Propagation—an Integrated Approach, Opt. Eng., **21**, 775 (1982).
8. A.K. Majumdar, Laboratory-Simulation Experiment for Optical Communication through Low-Visibility Atmosphere using a diode laser, IEEE J. Quantum Electron., **QE-20**, 919 (1984).
9. R.W. Svorec, Parametric Performance Analysis of Spaceborne Laser Communication Systems, Proc. SPIE, Vol. 295, 66 (1981).
10. A.K. Majumdar, Optical Communication between aircraft in low-visibility atmosphere using a diode laser, Appl. Opt., **24** (21), 3659–3665 (1985).
11. Alan J. MacGovern, David A. Nahrstedt, and Michael M. Johnson, Atmospheric propagation for tactical directed energy application, Proc. SPIE, Vol. 4034, pp. 128–139, Laser Weapons Technology, Todd D. Steiner, Paul H. Merritt, Eds. (2000).

12. V.I. Tatarskii, The Effects of the Turbulent Atmosphere on Wave Propagation (translated from Russian by IPST) (Available from the U.S. Dept. of Commerce, NTIS, Springfield, VA 22151), 1971.

13. Larry C. Andrews, Ronald L. Phillips, and Cynthia Y. Hopen, Laser Beam Scintillation with Applications, SPIE Press, Bellingham, Washington (2001).

14. Jennifer C. Ricklin, Stephane Bucaille, and Frederic M. Davidson, Performance loss factors for optical communication through clear air turbulence, Proc. SPIE, **5160**, 1–12 (2003).

15. J.C. Ricklin and F. M. Dadidson, Atmospheric turbulence effects on a partially coherent Gaussian beam: implications for free-space laser communications, J. Opt. Soc. Am. A, **19**(9), 1794–1803 (2002).

16. Arun K. Majumdar, Higher-order skewness and excess coefficients of some probability distributions applicable to optical propagation phenomena, J. Opt. Soc. Am., **69** (1), 199–202 (1979).

17. Arun K. Majumdar, Uniqueness of statistics derived from moments of irradiance fluctuations in atmospheric optical propagation, Opt. Commun., **50** (1),1–7 (1984).

18. Arun K. Majumdar, Higher-order statistics of laser-irradiance fluctuations due to turbulence, J. Opt. Soc. Am. A, **1**, 1067–1074 (1984).

19. Arun K. Majumdar and Hideya Gamo, Statistical measurements of irradiance fluctuations of a multipass laser beam propagated through laboratory-simulated atmospheric turbulence, Appl. Opt., **21** (12), 2229–2235 (1982).

20. Hideya Gamo and Arun K. Majumdar, Atmospheric turbulence chamber for optical transmission experiment: characterization by thermal method, Appl. Opt., **17** (23), 3755–3761 (1978).

21. Arun K. Majumdar, John A. DiUbaldo, and Alenka Brown-VanHoozer, Laboratory Simulation of Atmospheric Turbulence for Laser propagation: design and characterization, Proc. SPIE. Vol. 3432 (1998).

22. M.A. Al-Habash, L.C. Andrews and R. L. Phillips, Mathematical model for the irradiance probability density function of a laser beam propagating through turbulent media, Optical Eng., **40** (8), 1554–1562 (2001).

23. M. Nakagami, The m distribution—a general formula of intensity distribution of rapid fading, in *Statistical Methods in Radio Wave Propagation*, edited by W.C. Hoffman, 3-36 (Pergamon, New York, 1960).

24. Robert K. Tyson, Bit-error-rate for free-space adaptive optics laser communication, J. Opt. Soc. Am. A, **19** (40), 753–758 (2002).

25. Sherman Karp, Robert M. Gagliardi, Steven E. Moran and Larry B. Stotts, *Optical Channels* (Plenum Press, New York, 1988).

26. H.T. Yura and W.G. McKinley, Optical Scintillation Statistics for IR ground-to-space laser communication systems, Appl. Opt., **22** (21), 3353–3358 (1983).

27. H.G. Salfren, Effect of atmospheric turbulence on the bit error probability of a space to ground neaf infrared laser communications link using binary pulse position modulation and an avalanche photodiode detector, NASA Technical Memorandum 100699, Scientific and Technical Information Branch, 1987.

28. R.M. Gagliardi and S. Karp, M-ary Poisson detection and optical communication, IEEE Trans. Commun. Tech., **COM-17**, 208–216 (1969).

29. Robert M. Gagliardi and Sherman Karp, *Optical Communications* (Robert E. Krieger Publishing Company, Malabar, Florida, 1988).

30. Arun K. Majumdar and William C. Brown, Atmospheric Turbulence effecys on the Performance of Multi-Gigabit Downlink PPM Laser Communications, Proc. SPIE, Vol. 1218, 568-584, Free-Space Laser Communication Technologies II (1990).

31. Fotios Panagiotis Kourouniotis, Atmospheric Turbulence and Multiple Scattering effects on the Bit Error Rate of the Optical Receiver for PPM MultiGigabit Laser Communications, Master of Science Thesis, University of Colorado (1991).

32. J.H. Churnside and R. J. Hill, Probability density of irradiance scintillations for strong path-integrated refractive turbulence, J. Opt. Soc. Am. A., **4** (4), 727–733 (1987).

33. Akira Ishimaru, *Wave Propagation and Scattering in Random Media, Vol.2* (Academic Press, 1978).

34. Shin Tsy Hong, I. Sreenivasiah, and Akira. Ishimaru, Plane Wave Pulse Propagatioin Through Random Media, IEEE Trans. Antennas Propagation, **AP-25**, 6 (1997).

35. T.M. Shay, D.A. Hazzard, J. MacCannell, G. Lee, C.D. Garrett, J. A. Payne, N. Dahlstrom and S. Horan, First Experimental Demonstration of Full-Duplex optical Communications on a Single Laser Beam, AIAA 15th Annual AIAA/Utah State University Small Satellite Conference, August 13–16, 2001.

36. TGerardo G. Ortiz, Shinhak Lee, Steve Monacos, Malcom Wright and Abhijit Biswas, Design and development of a robust ATP subsystem for the Altair UAV-to-Ground Lasercomm 2.5 Gbps Demonstration, In Free-Space Laser Communication Technologies XV, G. Stephen Mecherle, Editor, Proc. SPIE, Vol. 4975 (2003).

37. J.M. Kahn and J. R. Barry, Wireless infrared communications, Proc. IEEE. **85** (3), 265–298 (1997).

38. G.W. Marsh and J. M. Kahn, Performance evaluation of experimental 50-Mbpss diffuse infrared wireless link using on-off keying with decision-feedback equalization, IEEE Trans. Commun., **44** (11), 1496–1504 (1996).

39. A.P. Tang and J. M. Kahn, Wireless infared communication links using multi-beam transmitter and imaging receivers, Proc. IEEE Intl. Conf. Commun., pp. 180–186 (1996).

40. Koorosh Akhavan, Mohsen Kavehrad and Svetla Jivkova, High-speed power-efficient indoor wireless infrared communication using code combining-Part I, IEEE Trans. Commun., **50** (7), 1098–1109 (2002).

41. Lijun Jiang, George Chung Kit Chen, Ye Yang, New Wei Lee, Allen Yeo and Hongga Li, Light-emitting-diode-based eide-field-of-view transceiver for indoor optical infrared wireless communication, Opt. Eng., **43** (4), 918–923 (2004).

42. Soo Hee Khoo, Wenwei Zhang, Grahame E. Faulkner, Dominic C. O'Brien and David J. Edwards, Receiver angle diversity design for high-speed diffuse indoor wireless communications, Proc. SPIE, Vol. 4530, 116–124, Optical Wireless Communication IV, edited by Eric J. Korevaar (2001).

43. Neil Savage, Linking with light, IEEE Spectrum, **39** (8), 32–35 (2002).

44. Sadik Esener and Philippe Marchand, Present and future needs of free-space optical interconnects, internet website: http://ipdps.eece.unm.edu/2000/wocs/18001109.pdf

45. Arun K. Majumdar and George H. Fortescue, Wide-beam atmospheric optical communication for aircraft application using semiconductor diodes, Appl. Opt., **22** (16), 2495–2504 (1983).

46. Isaac I. Kim, Harel Hakakha, Prasanna Adhikari, Eric Korevaar and Arun K. Majumdar, Scintillation reduction using multiple transmitters, in Free-Space Laser Communications Technologies IX, Proc. SPIE, Vol. 2990, 1997.

47. M. Jeganathan, K.E. Wilson, and J.R. Lesh, Preliminary analysis of fluctuations in the received Uplink-Beacon-Power data obtained from the GOLD experiments, JPL's TDA Progress Report 42-124, 20–32 (1996).

48. V. Vilnrotter, C.W. La, M. Srinivasan, R. Mukai, and K. Andrews, An optical array receiver for deep-space communication through atmospheric turbulence, JPL Publication: IPN Progress Report, Report 42-154, 1–21 (2003).

DOI: 10.1007/s10297-006-0079-z
Originally published in J. Opt. Fiber. Commun. Rep. 4, 225–362 (2007)

Laser communication transmitter and receiver design

David O. Caplan

MIT Lincoln Laboratory, Lexington, Massachusetts

Abstract. Free-space laser communication systems have the potential to provide flexible, high-speed connectivity suitable for long-haul intersatellite and deep-space links. For these applications, power-efficient transmitter and receiver designs are essential for cost-effective implementation. State-of-the-art designs can leverage many of the recent advances in optical communication technologies that have led to global wide-band fiber-optic networks with multiple Tbit/s capacities. While spectral efficiency has long been a key design parameter in the telecommunications industry, the many THz of excess channel bandwidth in the optical regime can be used to improve receiver sensitivities where photon efficiency is a design driver. Furthermore, the combination of excess bandwidth and average-power-limited optical transmitters has led to a new paradigm in transmitter and receiver design that can extend optimized performance of a single receiver to accommodate multiple data rates.

This paper discusses state-of-the-art optical transmitter and receiver designs that are particularly well suited for average-power-limited photon-starved links where channel bandwidth is readily available. For comparison, relatively simple direct-detection systems used in short terrestrial or fiber optic links are discussed, but emphasis is placed on mature high-performance photon-efficient systems and commercially available technologies suitable for operation in space. The fundamental characteristics of optical sources, modulators, amplifiers, detectors, and associated noise sources are reviewed along with some of the unique properties that distinguish laser communication systems and components from their RF counterparts. Also addressed is the interplay between modulation format, transmitter waveform, and receiver design, as well as practical tradeoffs and implementation considerations that arise from using various technologies.

1. Introduction

Optical communications has provided unprecedented capacity in modern networks. The rapid growth of the Internet has led to investment in wide-band fiber-optic networks that now span the planet. The optical communications revolution has also led

to the development of new and enabling technologies that can be applied outside of conventional telecom applications. Free-space laser communications, also known as lasercom, is one area that can benefit from these recent advances. Free-space optical networks have the potential to surround the planet with flexible and agile wide-band lasercom that could extend to the moon, Mars, [1,2] and beyond—capable of bridging billion kilometer links at Megabit/sec (Mbps) data rates. Such capabilities require photon-efficient transmitters and receivers, the main thrust of this paper, which can extend unrepeatered links over ultra-long-haul distances exceeding the net fiber deployed on Earth today.

Free-space laser communication performance directly depends on the efficiency and sensitivity of optical transmitters and receivers. However, until recently, these factors were not the driving factors in the buildup of fiber-optic networks. But as the apparent demand for bandwidth has approached the limitations of deployed fiber-optic links, more sensitive receivers have become a means of improving network performance in terms of power and bandwidth efficiency. Consequently fiber- and free-space-based technology requirements have started to converge. High-sensitivity photon-efficient transmitter and receiver designs can reduce mid-span amplifier requirements, diminish nonlinear impairments, and extend link distances in fiber networks, but they are especially beneficial for free space optical (FSO) communication, since improvements in receiver sensitivity directly reduce transmitted power requirements which in turn, can lead to significant size, weight, power (SWAP) and cost reductions.

The availability of modular high-reliability high-performance commercial-off-the-shelf (COTS) technologies already developed for and widely used by the telecommunication industry is of great value to the development of future free-space laser communication systems [3]. Not only can the tremendous telecom investment in technology development be leveraged in many cases, but the field-proven heritage of such technologies can provide increased confidence in reliability and life-time estimates, as well as the component manufacturing process. This knowledge can be used to accelerate the process of integrating cutting-edge technologies into reliable designs, and reduce the expense of starting costly qualification programs from scratch. This is especially important in space-based systems, where repair is impractical, and both reliability and performance are critical design drivers. The value added from heritage telecom technologies which require Telcordia (formerly known as Bellcore) qualification, is of additional value since these standards (e.g., [4–9]) often test to mechanical and thermal levels that are similar to many of the environmental requirements for space-based platforms.

Free-space optical links have some notable advantages over both radio-frequency (RF) and fiber-optic links that allow greater flexibility in transmitter and receiver design and optimization. These include the absence of channel dispersion and nonlinearities, and virtually unlimited channel bandwidth, especially in space-based applications where atmospheric absorption is nonexistent. Such characteristics allow FSO transmitter (TX) and receiver (RX) designs to leverage the average-power-limited properties of optical transmitters to generate optimized signaling waveforms [10–15], which, when used in combination with photon-efficient modulation formats, can contribute substantially to overall link efficiency, η_{link}. A practical measure of η_{link} is energy required per bit-received, [J/bit].

1.1. *Background*

The theories that govern both RF and optical communications over a noisy channel are well established, dating back to 1949 to the work of Claude Shannon [16,17] in which he established the channel capacity theorem which states that error-free communications are possible up to rate C (bits/s) over a channel of bandwidth B (Hz) with a signal of average power, P (W) perturbed by additive white Gaussian noise (AWGN) of power spectral density of $N_0/2$ (W/Hz). The capacity of such a link is given in Eq. (1), which highlights the interplay between the three key system parameters: bandwidth, average signal power, and noise power spectral density [17–19]:

$$C = B\log_2(1+\mathrm{SNR}) = B\log_2\left(1+\frac{P}{N_0B}\right) = B\log_2\left(1+\frac{rE_b}{N_0}\right) \quad [\mathrm{bits/s}]. \tag{1}$$

The SNR is simply the ratio of the average signal power (P), to noise power, (N_0B). For a data rate R_b (bits/s), the average signal power is R_bE_b, where E_b is the energy-per-bit (J/bit). Substituting spectral efficiency $r = R_b/B$ [(bit/s)/Hz], we obtain the right-hand side of Eq. (1), where E_b/N_0 is the energy-per-bit to noise power spectral density ratio. For shot-noise-limited optical signals the E_b/N_0 term can be replaced by the average number of signal photons-per-bit (PPB) in most cases (see section 2.3) and corresponds to the SNR over the observation window, with the exception of coherent-homodyne RXs, where the quantum-limited SNR = PPB/2 [20]. The Shannon Limit shown in Fig. 1, achieved when $R_b = C$, is the optimum tradeoff between spectral-efficiency, r, and photon-efficiency. In the limiting case of optimum bandwidth expansion, with modulation and coding, Shannon-limited sensitivity of 0.7 PPB (or ~0.35 PPB for homodyne) corresponding to a –1.6-dB SNR can be achieved for the AWGN channel [17,18,21–23]. A detailed discussion of Shannon-limited capacities for optical communication systems is given in [23].

Also highlighted in Fig. 1 is uncoded performance at the 10^{-9} bit-error-rate (BER) and Shannon-limited capacities for on-off-keying (OOK), differential-phase-shift-keying (DPSK), and M-ary orthogonal-keying modulation formats such as M-ary pulse-position-modulation (M-PPM) and frequency-shift-keying (M-FSK) with hard-decision coding. DPSK has received significant interest from the FSO community and the telecom industry since it is both energy-efficient and spectrally-efficient. For a quantum-limited optically preamplified DPSK receiver with optimal coding, Shannon-limited performance approaches 3 PPB with ~0.5 bit/s/Hz efficiency. High-rate optically-preamplified receivers using 24.6% and 7% low-overhead forward error correction (FEC) with 0.8 and 0.935 bits/s/Hz efficiencies have demonstrated 7 and 9 PPB receiver sensitivities at 10 and 40 Gbit/s data rates, respectively [24,25].

While spectral-efficiency has long been a key design parameter in the telecom industry, in many optical communication links excess channel bandwidth is available, and can be used to improve performance where photon efficiency is the design driver. In such links, the available bandwidth can be used to improve receiver sensitivities by ~5 dB by using, for example, coded M-ary orthogonal modulation formats such as M-PPM. As shown in Fig. 1, the quantum-limited sensitivity (pre-amplified, direct detection) for uncoded 1024-PPM in which each symbol carries 10 bits of information, is ~6 dB-PPB (~4 PPB), with $100\times(M/\log_2 M)$ bandwidth expansion. With the addition of ~50% overhead (OH) hard-decision FEC, ~1.5 PPB sensitivity can be

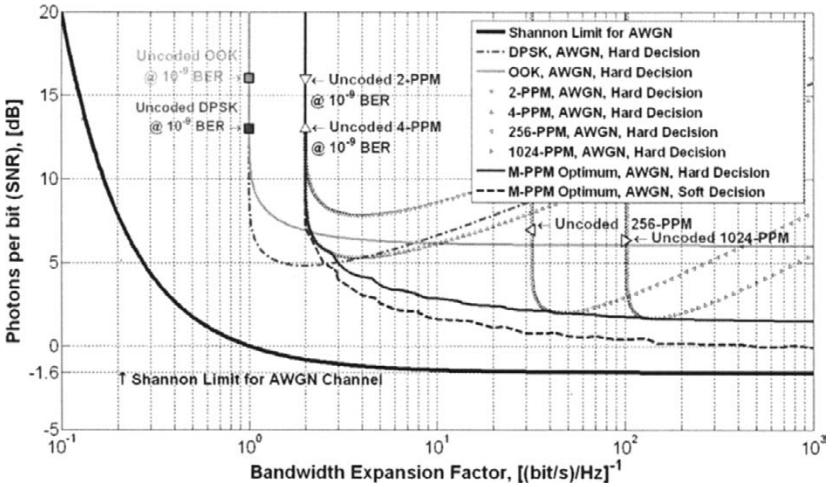


Fig. 1. The tradeoff between sensitivity (photon-efficiency) as measured in photons per bit, and bandwidth expansion (spectral-efficiency) for OOK, DPSK, and M-ary orthogonal, as well as the ultimate Shannon limit for AWGN systems [22]. Reprinted with permission of IEEE. (©2006 IEEE.)

achieved, and this can be extended to nearly 1 PPB by implementing optimal soft-decision decoding [22].

For the Poisson channel dominated by quantum noise, the Shannon-limited sensitivity per-photon can be made arbitrarily low, with a lower bound given by the ratio of thermal-noise-photon energy to signal-photon energy [26–28]:

$$\frac{B}{C_{\mathrm{ph\ Quantum-limit}}} = \frac{k_B T \ln(2)}{h\nu} \quad [\mathrm{photons/bit}], \tag{2}$$

where B is the receiver bandwidth in Hz, C_{ph} is the capacity in bits/sec per photon, k_B is Boltzmann's constant (1.38×10^{-23} J/K), h is Planck's constant (6.63×10^{-34} J s), and T is temperature in Kelvin (K). For the case of 1.5 μm signal photons at 300K, this yields a limit of 0.022 PPB (–16.6 dB SNR) or 46 bits/photon [27,28], which in principle, can be further improved by reducing the receiver temperature to 3K. Although this is not practically realizable, the promise of such capabilities has made the research and development of efficient wide-band photon-counting detector technologies an area of active research [1,2,29–32], especially since Gigabit-per-second (Gbps) class high-sensitivity communications employing super-conducting photon-counting detectors have recently been demonstrated [33–35].

A block diagram of the communication link is shown in Fig. 2, and the focus of this paper will be to optimize overall link performance (η_{link}) through the design of elements within the transmitter and receiver subsystems.

Functionally the free-space optical transmitter typically includes a laser source, a modulator that can impart a variety of modulation formats, a desired bit rate, coding, and a high power amplifier such as an Erbium-doped fiber amplifier (EDFA) with

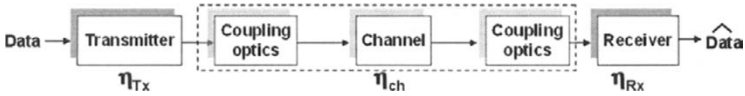


Fig. 2. Block diagram of the optical communications link, with the input being the data and the output being an estimate of the input data. Overall link performance $\eta_{\text{link}} = \eta_{\text{TX}}\eta_{\text{ch}}\eta_{\text{RX}}$.

an electrical to optical conversion efficiency which is often the dominant component of the net transmitter efficiency (η_{TX}). For transmission through the channel (η_{ch}), we can include transmitter optical (telescope) transmission efficiency (η_{TXOpt}), free space transmission efficiency (η_{space}), which is reduced by diffraction losses, atmospheric absorption, scattering, turbulence and scintillation, and interference from background radiance, e.g., loss and additive noise collectively ($1-\eta_{\text{atm}}$). Also included are pointing, acquisition and tracking efficiency (η_{PAT}), and lastly the receiver coupling efficiency (η_{RXOpt}). The net receiver efficiency (η_{RX}) includes the quantum-limited number of photons-per-bit to achieve the desired bit-error-rate (PPB_{QL}), associated implementation penalty (η_{QL}), and potential coding gain (G_{code}) from forward error correction (FEC). These terms are dependent on many factors such as the detection scheme, which may be direct, optically preamplified, homodyne or heterodyne detection, and the demodulation approach, which can be either incoherent or coherent and can occur in either the optical or electrical domain.

High-rate optical communication systems generally utilize PIN photodiodes which are commercially available with electrical bandwidths exceeding 50 GHz. Avalanche photodiodes are detectors with internal gain that can be used to improve RX sensitivity, although this generally comes at the expense of bandwidth and limited dynamic range of operation. A decoder circuit may be employed to process received FEC encoded data, and the final figure of merit is the BER, which is determined by comparing the input data (D) to the output estimate ($\hat{D}$). For systems which use FEC, the BER is often readily available and can be used as a feedback parameter to improve overall performance. Typically receiver sensitivities are measured in terms of the power received or the number of photons-per-bit required to achieve a particular BER, typically 10^{-9} (1 error per billion bits received), but the target can change depending on the application and system requirement.

Optical receiver sensitivities are ultimately limited by choice of modulation format, coding, and fundamental quantum fluctuations. This quantum noise is commonly referred to as "shot noise", which comes from the randomness in photon arrivals, and imposes theoretical limits on receiver performance. For high-rate high-sensitivity optical communication systems, optically preamplified receivers are the most practical, being widely used in the telecom industry and demonstrating the best performance at Gbit/s rates and beyond (see sections 4 and 5).

The theoretical (quantum limited) sensitivity is dependent on the modulation and receiver type and is often measured in terms of the number of PPB required to achieve a 10^{-9} BER. For optically-preamplified binary intensity modulation (IM) formats such as OOK and binary-PPM, the quantum limited sensitivity is ~40 photons-per-bit [20,36–39]. More complex orthogonal signaling formats such as M-PPM or M-FSK

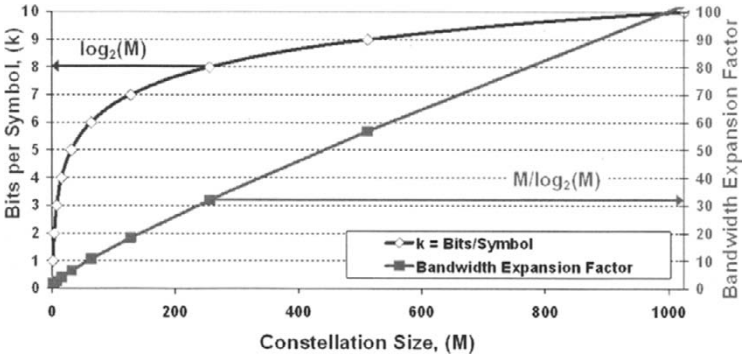


Fig. 3. Bits per symbol (left axis) and bandwidth expansion factor (right axis) for M-ary orthogonal signaling.

can be used to further improve the theoretical sensitivity at the expense of bandwidth, which degrades spectral efficiency by a factor of $(M/\log_2 M)$ as shown in Fig. 3.

For example 256-PPM improves receiver sensitivity by about a factor of 6, to ~6 PPB, but requires a factor of 32 expansion in bandwidth (see also Fig. 1). For high-rate bandwidth-limited applications, optically-preamplified Differential-Phase-Shift-Keying (DPSK) has demonstrated the best sensitivity, providing a ~3dB uncoded benefit over OOK and 2-PPM, corresponding to a theoretical sensitivity of ~20 PPB, although this comes at the expense of a more complicated receiver design.

Forward error correction coding is another powerful tool available to the system designer that is often a cost-effective means of improving receiver sensitivity [17,18,40]. FEC effectively operates by converting the BER of an input signal which contains redundant code bits to an improved output BER. As illustrated in Fig. 1, coding is needed in order to extend uncoded-quantum-limited performance to the ultimate Shannon-limited sensitivities. The use of FEC reduces spectral efficiency and increases TX and RX electronic complexity, requiring the additional encoding and decoding hardware, but can provide significant coding gain with relatively little overhead. For example, with commonly used enhanced Reed-Solomon 255/239 FEC coding, a ~6 dB sensitivity improvement is achieved by converting a $\sim 3 \times 10^{-3}$ input BER to a 10^{-9} output BER with only 7% overhead, i.e., additional bandwidth expansion relative to the rate. Given an optically preamplified DPSK receiver with near-quantum-limited performance this could improve sensitivities to ~6 PPB, within ~3 dB of the AWGN Shannon limit for DPSK. The benefits of such low-complexity FEC codes have been demonstrated at 10 [41] and 40 Gbps [25] data rates. They have also become commonly used in the telecom industry as a means of diminishing nonlinear impairments and increasing overall system throughput.

More powerful turbo codes with 100% overhead (rate $\frac{1}{2}$) have been developed that can closely approach Shannon-limited performance [42], however the complexity of implementing such serially concatenated codes has presently limited real-time application to low data rates <~50 Mbps [2,43–45]. But with less overhead (25%, rate 0.8) and complexity, block turbo codes have demonstrated ~10 dB coding gain at rates up

to 10 Gbps [24]. For the AWGN channel, with high-constellation orthogonal modulation formats such as M-PPM with optimal soft-decision decoding, ~1 photon/bit sensitivity can be achieved in principle [22].

1.2. Scope

This paper will expand upon the topics introduced above, providing an overview of state-of-the-art and future optical transmitter and receiver designs that are particularly well suited for FSO communication. For comparison, relatively simple direct detection systems used in short terrestrial or fiber optic links will be discussed, but emphasis will be placed on mature high-performance photon-efficient systems and technologies suitable for operation in deep space optical links.

While overall system level performance is dependent on many factors such as TX and RX aperture size, wavefront quality, and pointing acquisition and tracking (PAT) [3,46–48] this section will focus on practical design options for the TX and RX blocks of the communication link shown in Fig. 2. The remaining link elements—the coupling optics and channel considerations are discussed in detail in "Atmospheric channel effects on free-space laser communication", by Ricklin et al. (DOI 10.1007/s10297-005-0056-y) and "Free-space laser communication performance in the atmospheric channel", by Arun K. Majumdar (DOI 10.1007/s10297-005-0054-0) in this publication.

The fundamental characteristics of optical sources, modulators, amplifiers, detectors, and associated noise sources will be discussed along with some of the unique properties that distinguish laser communication systems and components from their RF counterparts. Practical tradeoffs and implementation issues that arise from using various technologies, and the interplay between modulation format, transmitter waveform, and receiver design and optimization will be presented.

This paper is intended to complement the rich subject of transmitter and receiver design, by highlighting practical design considerations and recent developments in the state-of-the-art high-performance FSO systems. While introductory and background material is included for clarity, the reader is directed towards the numerous papers, patents, and texts referenced at the end of this paper for a broader view of this subject.

1.3. Historical Perspective

There are many examples of free-space optical communications ranging from the use of mirrors and sunlight by the ancient Greeks, use of fire beacons by the Chinese (~800 BC) and later by Romans, and smoke signals by American Indians. Other notable examples include lighthouses, flags on sailing vessels, and Paul Revere's famous use of lanterns, "One if by land two if by sea" on April 18, 1775. Such early demonstrations of free-space optical communications were enabled by the use of widely available terahertz electromagnetic receivers, namely the human eye with its ability to detect the electromagnetic visible spectrum at wavelengths from about 400 to 750 nm. Alexander Graham Bell and Sumner Tainter demonstrated the first FSO telephone message using the patented "photophone" [49–53], illustrated Fig. 4. The photophone used sunlight as a source, modulated by reflecting off a vibrating mirror, and photoconductive selenium as the receiver on June 3, 1880.

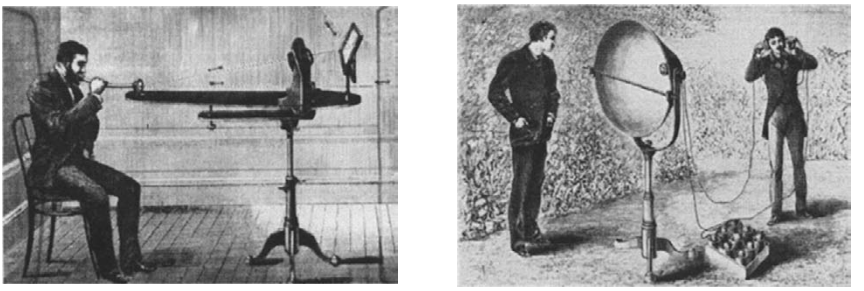


Fig. 4. (Left), The Bell photophone transmitter and (right) photophone receiver. *Rerinted with permission of Lucent Technologies, Inc./Bell Labs.*

However, despite these early advances, optical communications was soon overtaken by rapid development in wireline and wireless RF communications pioneered by the likes of Hertz [54], Bell[55], Edison [56], Marconi [57], Loomis [58], and Fessenden [59] in the late 1800s and early 1900s. During the following 100 years RF communications became a refined commodity with applications ranging from radio, television, cellular, and satellite communications. As the established means of wireless communications, the maturity and sophistication of existing RF capabilities make them worthy of comparison to up-and-coming FSO systems.

Present day FSO systems have significant potential to improve upon RF capabilities in areas where bandwidth or regulatory limitations exist, applications where security is important, and in ultra-long-haul space-based links where size, weight and power (SWAP) are at a premium and diffraction losses dominate. Since the invention of the laser in 1960, there have been significant developments in the critical technologies needed to tap the potential of photonic communications summarized in the list below.

Significant lasercom milestones

1960 Invention of the laser, [60–63]
1962 Invention of semiconductor diode laser, (GaAs) [64,65]
1964 Invention of fiber optical amplifier (Nd:glass 1μm) [66,67]
1970 First continuous-wave room-temperature semiconductor lasers [68]
1970 Development of low loss (<20 dB/km) glass fibers [69–74]
1980 First commercial optical fiber system at 45 Mb/s (AT&T)
1987 Development of Erbium-doped fiber laser amplifier EDFA (1.5μm) [75–77]
1988 First trans-Atlantic fiber cable (280 Mb/s, 1.3 μm) [78]
1989 First undersea tests of optically amplified (EDFA) fiber system
1992 First trans-Atlantic fiber cable using 1.5μm technology (560 Mb/s) [79]
1993 MCI purchases 500 EDFAs and begins installation in terrestrial network
1995 Bidirectional Ground-to-Orbit Lasercom Demonstration (GOLD), 1Mbps up- and down-link transmissions @ 0.514 and 0.830 μm on Engineering Test Satellite-VI (ETS-VI) in elliptical GEO transfer orbit [80–83]

1996 First trans-Atlantic fiber cable using EDFAs (5 Gb/s) [84]
1996 1 Tbit/s demonstrated over fiber [85–88]
2000 360 million km of fiber world wide [89]
2001 Capacity of commercial WDM systems exceeds 1.6 Tb/s [89]
2001 10 Tb/s demonstrated over fiber [90–94]
2001 Geosynchronous Lightweight Technology Experiment (GeoLITE) [95-97], successful demonstration of bidirectional laser communications between a satellite in geosynchronous orbit (GEO), ground, and aircraft
2001 Semiconductor Intersatellite Link Experiment (SILEX) [98–101], demonstrating bidirectional GEO-LEO and GEO-ground laser communications between the Advanced Relay Technology Mission Satellite (ARTEMIS) in GEO, with a 10mW (average) 2 Mbps (2-PPM) directly-driven semiconductor laser TX @ 0.8 μm and a 50 Mbps Si-APD based NRZ-OOK RX; the SPOT-4 imaging satellite in low earth orbit (LEO), and the optical ground station (OGS) [102,103].
2003 Mars Laser Communication Demonstration (MLCD) program initiated (targeting the ~300,000,000 km Mars to Earth link, ~1–30 Mbit/s) using 5W MOPA TX, coded M-PPM ($M = 16 - 64$) and photon-counting detectors @ 1.064 μm [1,2,104]. Critical technologies demonstrated but program cancelled in 2005.
2005 Successful bidirectional inter orbit lasercom link between Optical Inter-orbit Communications Engineering Test Satellite (OICETS) in LEO and ARTEMIS in GEO at 2 and 50 Mbps [105]

It is likely that within the next decade, we will see the first deployments of operational free-space laser communications systems with global and interplanetary reach.

2. General Wavelength Considerations

2.1. Carrier Characteristics

Electromagnetic (EM) waves can be described as an oscillating field with frequency, νs that travel at the speed of light, c. The inverse relationship between the EM carrier frequency (ν) and wavelength (λ) is

$$c = \lambda\nu \tag{3}$$

where the constant $c = 3.0 \times 10^8$ m/s is the speed of light in vacuum. Quantized energy is proportional to frequency, and is given by Planck's relation, which defines the energy-per-photon to be

$$E = h\nu = hc/\lambda \quad [\text{J}], \tag{4}$$

where $h = 6.63 \times 10^{-34}$ J s is Planck's constant. The illustration of the electromagnetic spectrum in Fig. 5 spans 15 orders of magnitude from kilometer long RF waves to picometer long gamma radiation and gives perspective into the broad range of wavelengths, frequencies, and photon-energies that can be used for communications.

As discussed in further detail in section 2.3, these basic EM properties profoundly impact the transmitter and receiver designs. For example, as we transition from using

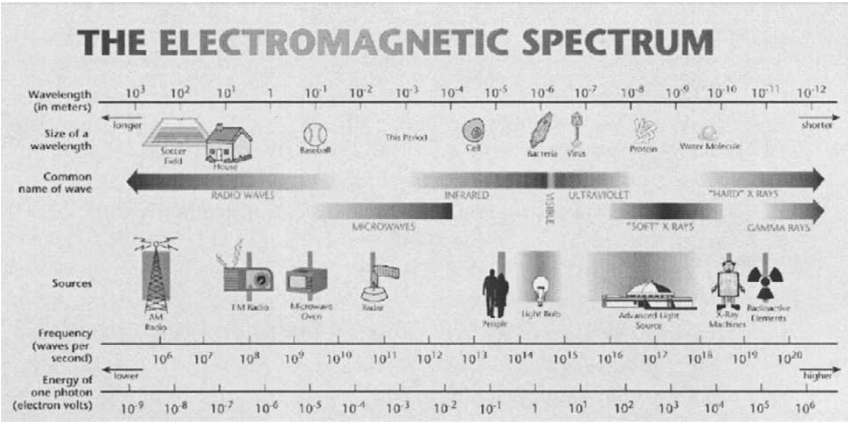


Fig. 5. The electromagnetic spectrum is shown in terms of wavelength along with corresponding frequency, and energy-per-photon. Conventional names of the spectral bands and example sources are also depicted. *Courtesy of the Advanced Light Source, Lawrence Berkeley National Laboratory.*

RF carriers towards higher frequency optical waves for communication, the carrier wavelength shortens. This directly impacts both transmit and receive antenna (aperture) size and diffraction, which is proportional to wavelength. The diffraction limited transmitter beam angle is approximately [47,106]

$$\theta = \lambda / D_{\mathrm{TX}} \quad [\mathrm{rad}], \tag{5}$$

where D_{TX} is the transmitter aperture diameter. Thus using shorter wavelengths allows for smaller divergence angles, or equivalently, better directivity for a given aperture size. Note that this also increases the precision required to hit the target receiver and therefore comes at the expense of increased difficulty in pointing, acquisition, and tracking. From Eq. (5), the far-field on-axis intensity ($\mathrm{W/m^2}$) can be estimated, and for a receiver with collection area (A_{RX}) the power delivered to the RX can be calculated. Diffraction losses reduce the free space transmission efficiency, η_{space}, defined as the ratio of on-axis received power and transmitted power, which in the far-field ($\lambda L \gg D_{\mathrm{TX}}^2$) is approximately given by [106],

$$\eta_{\mathrm{space}} \approx \frac{A_{\mathrm{RX}} A_{\mathrm{TX}}}{\lambda^2 L^2} = \left(\frac{\pi D_{\mathrm{RX}} D_{\mathrm{TX}}}{4 \lambda L} \right)^2 = 1 - \text{Diffraction Loss}, \tag{6}$$

and decreases with the square of the carrier wavelength (λ) and link distance (L).

Carrier frequency directly impacts the fractional bandwidth available for modulation, and therefore limits the maximum rate at which that information can be transmitted. For instance, if the modulation bandwidth is limited to 10% of the carrier frequency, this corresponds to 30 THz available bandwidth for a 1μm optical carrier versus 3 GHz bandwidth for a 1cm RF carrier. This highlights a key difference between high-rate RF and FSO communications, since RF systems need to be spectrally

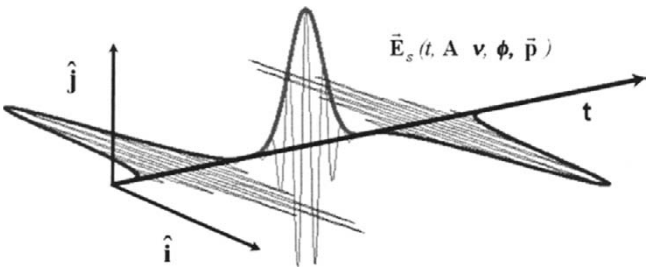


Fig. 6. A time varying electro-magnetic signal, $\vec{\mathbf{E}}_s(t, \boldsymbol{A}, \boldsymbol{\nu}, \boldsymbol{\phi}, \vec{\mathbf{p}})$, depicting the potential for amplitude, frequency, phase, and/or polarization modulation.

efficient in high-rate applications at the expense of energy or photon efficiency (see Fig. 1), whereas FSO systems can often use bandwidth as a flexible design parameter.

Channel properties also vary substantially with wavelength. Clouds, fog, rain, turbulence, scattering (e.g., Raleigh scattering $\propto 1/\lambda^4$), atmospheric absorption, etc. all have a strong wavelength dependence. Owing to the longer wavelength, RF communications are generally more robust in atmospheric channels where these factors are a limitation [107–109].

Lastly, photon energy is proportional to the quantum or shot noise in the transmitted signal, and therefore can impact the receiver sensitivity. As the photon energy (in Joules), increases, the number of photons per second decreases for a given power (in W, or J/s). This in turn increases the "graininess" or shot noise of the signal, which drives high-sensitivity optical receiver design as described further in section 2.3.

2.2. *Electromagnetic Signaling Options*

A general expression for a time varying electro-magnetic (EM) signal $\vec{\mathbf{E}}_s(t, \boldsymbol{A}, \boldsymbol{\nu}, \boldsymbol{\phi}, \vec{\mathbf{p}})$ is given in Eq. (7) and illustrated in Fig. 6 where variables $\boldsymbol{A}, \boldsymbol{\nu}, \boldsymbol{\phi}$, and $\vec{\mathbf{p}} = \mathbf{a}\hat{i} + \mathbf{b}\hat{j}$ represent the carrier amplitude, frequency, phase, and polarization field properties respectively, and can be time-dependent, e.g., $\boldsymbol{\phi} = \boldsymbol{\phi}(t)$:

$$\vec{\mathbf{E}}_s(t, \boldsymbol{A}, \boldsymbol{\nu}, \boldsymbol{\phi}, \vec{\mathbf{p}}) = A(t)\cos(2\pi\boldsymbol{\nu}t + \boldsymbol{\phi})\vec{\mathbf{p}} \quad [\mathrm{V/m}], \tag{7}$$

where $P(t) = |A(t)|^2/(2z_0)$ is the power in the field envelope in W, and z_o is the impedance of the medium, which is 300 Ω in free-space. Each of these basic field properties can be modulated either alone or in combination depending on the system needs. The additional dimensions can carry information, for use in multi-dimensional symbol constellations, or simply improve the modulation format.

For example, in narrowband intensity modulated optical systems where stimulated Brillouin scattering (SBS) nonlinearities are a limitation, carrier phase, polarization, and/or frequency can be modulated to expand the optical spectrum beyond the SBS bandwidth of ~50 MHz to mitigate this effect [110,111]. Even if unused, these parameters may need to be controlled, measured, or known at the receiver to achieve optimum performance. For instance, FSO receivers are often single polarization with

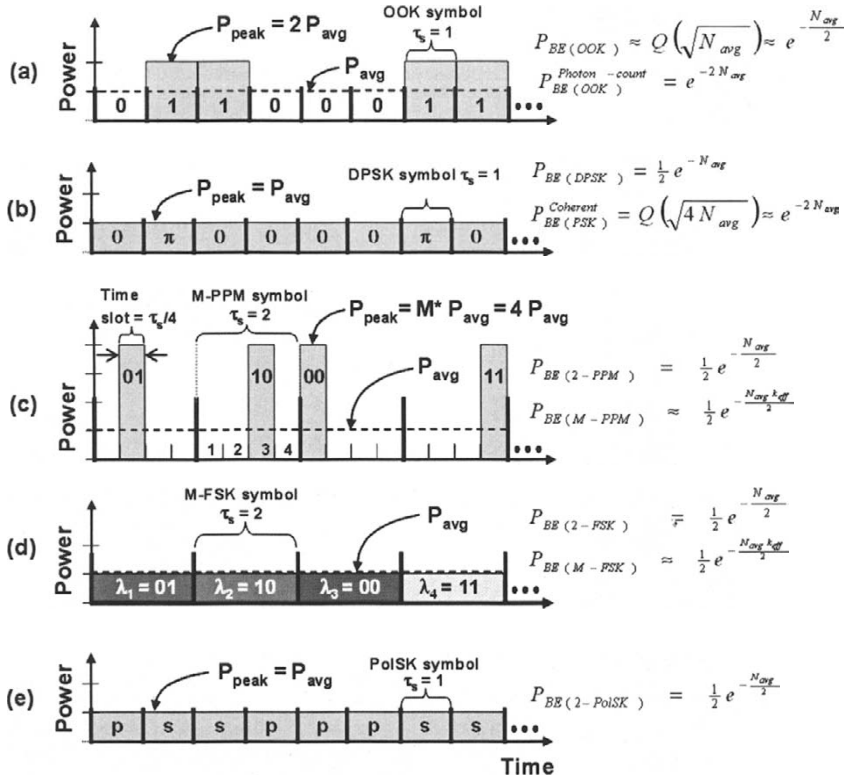


Fig. 7. Commonly used modulation formats (a) OOK, (b) DPSK, (c) 4-PPM, (d) 4-FSK, (e) PolSK with same data, data rate and average power. Also shown on the right-hand side is the approximate bit-error probably for the modulation formats (a)–(e) with ideal optically preamplified receivers [except for photon-counting OOK in (a) and coherent PSK in (b)].

tight spectral filtering to allow for efficient diplexing of transmit and receive signals [46], and/or to reduce detected noise. This requires proper alignment of transmit and receive wavelength and polarization, either by stable open loop control of both the TX and RX, tracking and compensation by the RX, feedback between the TX and RX, or a combination thereof.

2.2.1. Overview of FSO Modulation Formats and Sensitivities

While there are many modulation possibilities, the most common modulation formats considered for FSO links include on-off-keying (OOK), differential-phase-shift-keying (DPSK), phase-shift-keying (PSK), and orthogonal modulation formats such as M-ary pulse-position modulation (M-PPM) and frequency-shift-keying (M-FSK) illustrated in Fig. 7.

2.2.1.1. *On-Off-Keying (OOK)*

The predominant form of signaling used in optical communication systems today is on-off key (OOK) because the transmitter and receiver hardware are relatively simple and fiber optic networks generally operate at high signal-to-noise ratios with small dynamic range requirements and well controlled signal levels at the receivers. Also known as binary amplitude-shift-keying, OOK is a form of intensity modulation (IM) in which binary information is represented by the presence or absence optical signal energy within the symbol. At the receiver, the "1" or "0" logical decision is determined by the received symbol energy being above or below a predetermined threshold. The optimum threshold balances the probability of 0 and 1 errors, and is dependent on the received signal power and noise statistics, requiring adaptive thresholds for best performance over a fluctuating channel [20,39,112,113]. Regardless of the accuracy of optimum OOK threshold predictions based on theoretical calculations, the threshold and other operational parameters can be optimized with near-real-time BER feedback that is readily available with commercial FEC hardware [114,115]. Alternatively, modulation formats that inherently establish a threshold, such as antipodal or orthogonal signaling (e.g., PSK, DPSK, FSK, PPM) can be used to eliminate the need for adaptive thresholding.

For an ideal shot-noise-limited optically-preamplified receiver, the OOK error probability is approximately given by [18,20,23,37–39]

$$P_{\mathrm{BE(OOK)}} \approx Q\left(\sqrt{N_{\mathrm{avg}}}\right) \approx \frac{1}{2} e^{-N_{\mathrm{avg}}/2}, \tag{8}$$

where N_{avg} is the SNR, equal to the average number of photons-per bit (PPB) received, and the Q-function is

$$Q(z) = \frac{1}{\sqrt{2\pi}} \int_{z}^{\infty} e^{-x^2} dx = \frac{1}{2} \mathrm{Erfc}\left(\frac{z}{\sqrt{2}}\right) = \frac{1}{2}\left(1 - \mathrm{Erf}\left(\frac{z}{\sqrt{2}}\right)\right). \tag{9}$$

The Erf( ) and Erfc() functions are the error function and complementary error function, respectively [18, 20]. The 10^{-9} bit-error rate corresponds to 36 PPB, slightly lower than the right-hand approximation in Eq. (8) which results in ~40 PPB. When implemented with non-return-to-zero (NRZ) waveforms with an equal probability of 1's and 0's, the peak power is equal to twice the average. Pulsed return-to-zero or RZ waveforms can be used with all the modulation formats mentioned above, with the peak-to-average power ratio varying inversely with the duty-cycle.

Note that the analytical BER expressions in Eqs. (8)–(11) and (13)–(14) assume Gaussian noise statistics [18,39,116–118], which do not accurately reflect the Poisson noise statistics of the received input signal (see section 2.3.5), nor the statistics of the preamplified signal detected by the square-law photodetector. The statistics for a sinusoidal electric field with narrow-band additive white Gaussian noise after square-law detection for 0 and 1 signals are no longer Gaussian, they are Rayleigh and Rician, respectively [116,119,120]. However, detailed numerical evaluation of preamplified OOK performance assuming these distributions yields ~38 PPB at 10^{-9} BER [36, 38], compared to ~36 PPB with the Gaussian approximation. An additional correction includes the exact statistics based on the quantum-mechanical description of the

physical photon-detection process which are the degenerate Bose–Einstein probability distribution for 0's and the noncentral negative binomial distribution (NNB) also called the Laguerre distribution for 1's [47,121–126]. Nevertheless, the Gaussian approximation yields straightforward analytic BER expressions that provide a reasonable estimate of receiver sensitivity, with SNR estimates for a particular BER accurate to within ~1 dB relative to calculations based on with exact statistics [126,127].

2.2.1.2. Differential-Phase-Shift-Keying (DPSK)

Differentially encoded PSK [DPSK, see Fig. 7(b)] has received considerable attention by the FSO community and the telecom industry due to a ~3 dB sensitivity improvement over commonly used on-off-keying, and reduced peak power which mitigates nonlinear effects [24,25,41,128–133]. NRZ-DPSK can be implemented with a constant envelope, so that the peak power is equal to the average. Binary information is conveyed with two orthogonal symbols represented by the relative phase between two differentially encoded bits: a "0" represented by no-phase change and a "1" by a π phase difference, (or visa-vera). This phase modulation also tends to suppress the optical carrier, which can be advantageous in systems limited by slow optical nonlinearities such as SBS (see section 3.5.5). Typically, adjacent bits are differentially encoded with a time-separation τ_d of one symbol duration or a bit period, τ_{bit}, but this can generally be extended to an integer number n of symbol periods, i.e., $\tau_d = n\tau_{\text{bit}}$. This can provide some flexibility in implementing simplified multi-rate and multi-channel receivers (see, e.g., [134–136]), a subject discussed further in section 5.2.

DPSK's utility has been established with many long-haul fiber-optic experiments demonstrating multiple Tbit/s over ~10,000 km fiber spans with hundreds of WDM-DPSK channels (λs), e.g., [137–141]. Despite significant benefits, DPSK has not yet been widely deployed in operational systems, in part because the benefits of DPSK come at the cost of increased complexity over OOK, requiring a phase modulator and differential precoding in the transmitter, and an optical delay-line interferometer (DI) demodulator and balanced detection in the receiver in order to derive maximum benefit. The DPSK RX design is discussed further in section 5.2. Performance can also be degraded significantly by differential chirp [132,142]. However, DPSK is much easier to implement than coherent PSK, which requires a stable, narrow-linewidth local oscillator phase-locked to the received signal necessary to implement the homodyne receiver. In contrast, DPSK can be demodulated through a self-homodyne process of comparing the relative phase of the differentially encoded symbols. This eliminates the need for the LO and associated phase-locking challenges, and relaxes the need for narrow linewidth laser sources, since the self-homodyne process only requires phase-coherence for the time duration between the differentially encoded symbol. The theoretical error-rate for optically preamplified DPSK is

$$P_{\text{BE(DPSK)}} = \frac{1}{2} e^{-N_{\text{avg}}}, \tag{10}$$

corresponding to 20 PPB @ 10^{-9} BER [38]. The nearly 3 dB sensitivity benefit of DPSK over OOK can be viewed as a result of the differential encoding, which utilizes energy from the two symbols to determine the relative phase for one bit of information.

Optical differential quadrature phase shift keying (DQPK) [132,143–147] and duobinary modulation [148–152] are spectrally efficient modulation formats related

to DPSK that have recently been considered for use over long-haul fiber links, where dispersion is often a dominant limitation. However, dispersion and channel bandwidth are generally not limiting factors in free-space links, so the combination of reduced photon-efficiency and increased complexity make these modulation formats less attractive than DPSK for use in free-space applications.

2.2.1.3. Phase-Shift-Keying (PSK)

As noted above, coherent PSK has challenging LO and signal alignment requirements [20,153–155] needed to determine the absolute phase of the optical field. In addition, there are practical tradeoffs between PSK RX bandwidth and sensitivity, which are especially noticeable since, unlike preamplifier receivers, coherent receivers are not easily scaled to higher data rates via wavelength division multiplexing (WDM). For the additional complexity, coherent (binary) PSK provides among the best theoretical RX sensitivity, with

$$P_{\mathrm{BE(PSK)}} = Q\left(\sqrt{4N_{\mathrm{avg}}}\right) \approx e^{-2N_{\mathrm{avg}}}, \tag{11}$$

corresponding to 9 PPB RX sensitivity @ 10^{-9} BER [20,156]. But in practice the high-sensitivity potential of optical PSK has not been realized. As seen in Fig. 61, the best PSK demonstration reported uncoded sensitivity about 2.5 dB from theory at 4 Mbps [157]. In the Gbit/s regime the best PSK performance degraded to ~6 dB from theory at ~6 Gbps and ~9 dB from theory at ~8 Gbps [158], providing little benefit over optically preamplified DPSK, which is a WDM-scalable approach with numerous demonstrations of comparable or better performance in terms of sensitivity and data rate [24,25,41,128–130,133,159]. The nearly 4× sensitivity benefit PSK has over OOK is in part due to binary PSK being an antipodal format, in which a "0" is represented by signal field of −1 (π phase) and a "1" is represented by a 1 (phase of 0), so that 2-PSK has twice the signal-distance of OOK. The other 2× advantage can be viewed as a result the coherent detection process, which detects the peak power of the received optical carrier, that is twice the average power measured in intensity modulation formats.

2.2.1.4. M-ary Orthogonal Modulation

To improve receiver sensitivity, high-order M-ary orthogonal formats [18] such as M-PPM, M-FSK [160–162], and hybrid combinations can be used (see also section 5.3). While not spectrally efficient, these formats can significantly improve photon-efficiency. Two symbols m and n are considered orthogonal if

$$\int_0^{\tau_s} s_m(t)s_n(t)dt = \delta_{mn}, \tag{12}$$

where τ_s is the symbol period and δ_{mn} is the Kronecker delta function. The parameter M refers to the symbol alphabet or constellation size, with the information per symbol increasing as $k = \log_2(M)$ and bandwidth increasing with M. This results in a spectral efficiency [Bit/s/Hz] that varies as $\log_2(M)/M$ shown in Fig. 3, which clearly illustrates the trade between bandwidth (spectral efficiency) and SNR (photon

efficiency) noted by Shannon (see Fig. 1). For large M, ideal performance approaches Shannon-limited capacity [17,18,23]. Orthogonal modulation formats also establish their own threshold, with the optimum decision based on the largest sample within the symbol set. This enables the system to operate optimally over wide and rapid changes in signal level, without the need for adaptive thresholds that are necessary for OOK.

The bit-error rate for optically-preamplified binary-orthogonal modulation, e.g., 2-PPM is given by

$$P_{\mathrm{BE(2-ORTH)}} = \frac{1}{2} e^{-\frac{N_{\mathrm{avg}}}{2}}, \tag{13}$$

with sensitivity comparable to OOK and exactly twice that of DPSK, corresponding to 40 PPB at 10^{-9} BER. For M-ary orthogonal modulation, no simple closed form expression exists. An accurate but computationally challenging expression for the ideal optically preamplified symbol-error-rate is given by[163]

$$P_{\mathrm{SE(M-ORTH)}} = \int_0^\infty f_0(x_0) \left\{ 1 - \left[1 - \int_{x_0}^\infty f_1(x_1) dx_1 \right]^{M-1} \right\} dx_0, \tag{14}$$

where the Rician distribution [119]

$$f_j(x_j) = \frac{1}{N_{\mathrm{sp}}} e^{-\frac{x_j + N_{\mathrm{avg}j}}{N_{\mathrm{sp}}}} I_0 \left(2 \frac{\sqrt{N_{\mathrm{avg}j} x_j}}{N_{\mathrm{sp}}} \right), \tag{15}$$

is the probability density function for sample x_j, where x_1 represents a sample with both signal and noise, and x_0 represents the samples with noise only, and N_{sp} is the spontaneous emission factor related to the noise figure of the amplifier. For an ideal amplifier, $N_{\mathrm{sp}} = 1$. The function I_0 is the modified Bessel function of the first kind of zero order. Robinson, in [163] provides an excellent derivation of the optically-preamplified symbol-error rate in Eq. (14) along with simplified techniques for evaluating the expression.

For each symbol error, there is still a chance that some of the bits within the deocoded symbol are correct. Assuming that each symbol is transmitted with equal probability, the bit-error probability is

$$P_{\mathrm{BE(M-ORTH)}} = \left(\frac{M/2}{M-1} \right) P_{\mathrm{SE}} \quad \underset{M \to \mathrm{large}}{\Longrightarrow} \quad \frac{P_{\mathrm{SE(M-ORTH)}}}{2}. \tag{16}$$

Qualitatively, the M-ary orthogonal bit-error probability can be related to the binary BER expression in (13) by

$$P_{\mathrm{BE(M-ORTH)}} \approx \frac{1}{2} e^{-\frac{N_{\mathrm{avg}} k_{\mathrm{eff}}}{2}}, \tag{17}$$

where k_{eff} is related to the bits-per-symbol given in Eqs. (18)–(20):

$$k_{\mathrm{eff}} = k \left(1 - \frac{k-1}{k N_c} \right), \tag{18}$$

$$N_c = N_{\mathrm{avg}} \left(1.75 - \frac{1}{1 - e^{-T_k (N_{\mathrm{avg}} - 1)}} \right), \tag{19}$$

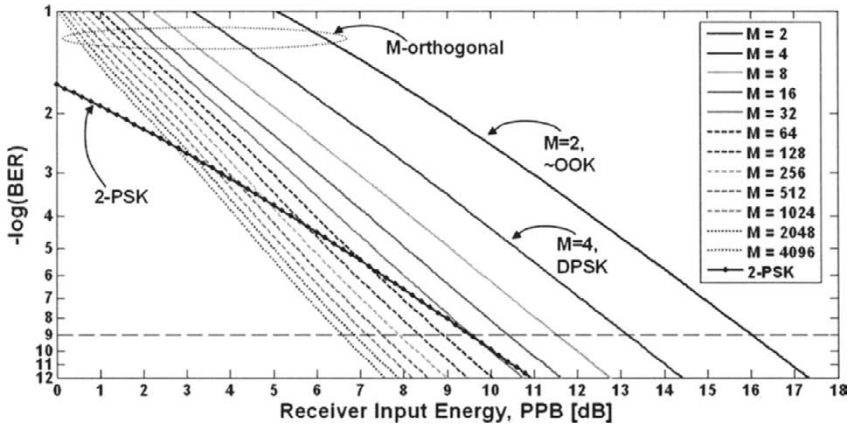


Fig. 8. BER curves for 2-PSK and optically preamplified M-ary orthogonal modulation.

$$T_k = 0.4\left(1 - \frac{4}{1 - e^{-(k-5)}}\right). \tag{20}$$

For small $M$and large received signal (N_s), $k_{\text{eff}} \to k$, yielding a receiver sensitivity of about 40 PPB/$\log_2 M$ or 40/k at 10^{-9} BER. This simple approximation overestimates the RX sensitivity by less than ~1.25 PPB, which for $M < 64$ at 10^{-9} BER is accurate to $\sim \frac{1}{2}$ dB with the accuracy improving at lower error rates. For increased accuracy without much computational complexity, Eq. (17) can be used to calculate M-ary othogonal BER performance with ~0.1 dB accuracy.

Bit-error rate curves as a function of photons/bit (in dB) are shown in Fig. 8 for ideal 2-PSK and optically premplified M-ary orthogonal modulation. Note that theoretical preamplified OOK and DPSK BER performance is similar to binary and 4-ary othogonal modulation, respectively. Without additional coding, 32-orthogonal modulation has comparable performance to ideal 2-PSK at 10^{-9} BER. At higher 10^{-3} BER, near the FEC threshold for enhanced 255/329 Reed-Solomon codes, 2-PSK RX sensitivity is as good as to 1024-orthogonal. Above 10^{-2} BER, ideal 2-PSK has a clear sensitivity advantage over preamplified M-ary orthogonal modulation. However in practice, at high rates above a Gbit/s (see Fig. 61), reported coherent-PSK [158,164–166] suffers ~5 dB greater implementation penalties than preamplified RXs [13,15,128–130,133,135,159], eliminating much of the theoretical benefit. While improved PSK designs may achieve the anticipated sensitivities, based on existing demonstrations, preamplified RX designs have exhibited superior sensitivity and data-rate, as well as WDM scalability [135,138,139,149,167].

2.2.1.5. M-ary Pulse-Position Modulation (M-PPM)

Pulse-Position Modulation (PPM) is a form of orthogonal signaling that uses the same transmitter and receiver hardware used for OOK. In M-PPM [2,15,26,43,168–171], k-bits of information are encoded by the position of an optical pulse within an M-slot symbol, (Fig. 7c). As a result, M-PPM waveforms have a low duty cycle, equal to $1/M$,

making it well suited for use with EDFA-based average-power-limited transmitters, but a poor choice for use with semiconductor-based peak-power-limited transmitters [172]. For M-PPM, the bandwidth expansion directly impacts the modulation bandwidth and associated transmit/receive electronic bandwidth

$$B_{\mathrm{E(M-PPM)}} = \frac{MR_{\mathrm{Data}}}{\log_2 M} = \frac{MR_{\mathrm{Data}}}{k} = S, \tag{21}$$

with associated symbol rate

$$R_{\mathrm{sym(M-PPM)}} = \frac{R_{\mathrm{Data}}}{\log_2 M} = \frac{S}{M}. \tag{22}$$

While electrical bandwidth limitations may limit the maximum slot-rate S for a single M-PPM channel, the high-speed 10 and 40 Gbps electronics developed for the telecom industry make it relatively easy to implement at reasonably high data-rates. For example, by transmitting 16-PPM at a slot-rate S = 10 Gslot/s, 2.5 Gbps can be delivered with symbol rate of 625 Msym/sec.

The low duty cycle of M-PPM waveforms can also lead to optical nonlinearities, which can limit the peak transmit power, a subject discussed further in section 3.5.5. On the receive side, M-PPM requires two clocks to be acquired, a symbol clock (R_{sym}) and slot clock at rate $S = MR_{\mathrm{sym}}$ [173]. Clock acquisition can be challenging for large M since received power at the clock frequencies varies as $\sim(1/M^2)$, which may require embedded synchronization bits within the data as M gets large [44].

M-PPM has several practical advantages over OOK and DPSK. As a result of the low duty cycle, PPM has less spectral content at low frequencies and consequently has a smaller fractional bandwidth than OOK and DPSK. Fractional bandwidth (B_f) is a factor used to classify signals as narrowband, wideband, or ultra-wideband and is often defined by the ratio of bandwidth at –10 dB points to center frequency [174] given by

$$B_f = \frac{\mathrm{BW_{10dB}}}{(f_h + f_l)/2} = \frac{2(f_h - f_l)}{(f_h + f_l)}, \tag{23}$$

where f_h and f_1 are the highest and lowest cutoff frequencies (at the –10 dB point) of the spectrum. For RF signals, $B_f > 20\%$ is considered ultra-wide band. However for optical communications signals, the spectrum of the RF drive spectrum typically has $B_f > \sim 200\%$ as shown in Fig. 9. For such broadband signals, a more relevant parameter is the high-low bandwidth ratio,

$$B_r = f_h/f_l. \tag{24}$$

For a pseudo-random bit sequence (PRBS) of 10 Gbps NRZ-OOK waveforms, B_r extends from a practical lower bound of ~10 kHz to ~10 GHz, or six decades. In contrast, the spectra for M-PPM waveforms operating at the same data rate span less than two decades, reducing B_r by over 4 orders of magnitude, despite having more high-frequency content (Fig. 9).

This relaxes the performance requirements on wide-band electronic amplifiers and drivers. In addition, since the longest string of consecutive 1's is two (from two adjacent PPM symbols), pattern-dependencies in transmit and receive hardware are reduced, making it easier to generate and receive high-quality waveforms. These benefits can also extend to the ultra-high-speed domain, where the reduced pattern dependence

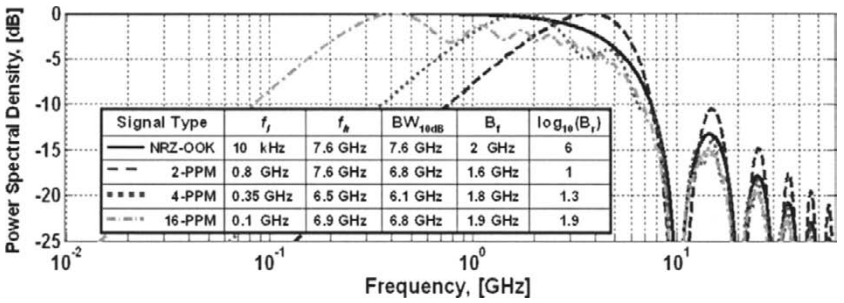


Fig. 9. Calculated spectra for square waveforms used in OOK and 2-PPM, 4-PPM, and 16-PPM modulation for with a fixed 100 psec pulse width [175]. The M-PPM waveforms have a smaller fractional-bandwidth (B_f) and a significantly smaller bandwidth ratio, ($B_r = f_h/f_l$) than OOK. For the OOK spectrum, f_l is set to 10 kHz, which is a common low-frequency specification for applicable broad-band electronics. Note that for a constant data rate, the M-PPM spectra are broadened by a factor of $M/\log_2(M)$, which increases $\mathrm{BW}_{10\mathrm{dB}}$ but does not impact B_r.

has improved the performance of all-optical demultiplexers [163,176]. M-PPM also benefits from the sequential nature of the symbol set, which enables a single-chain of drive electronics and associated filters to generate and receive the complete symbol set. This simplifies and improves decision process, since it is easier to make a fair comparison of the M-samples within a symbol to determine which is the largest.

2.2.1.6. *M-ary Frequency-Shift Keying (M-FSK)*

The FSK symbol set consists of M frequencies or wavelengths and like PPM, conveys $k = \log_2(M)$ bits of information per symbol (Fig. 7d). Since FSK transmits a different frequency for each symbol, modulation bandwidth and associated transmit/receive electronic bandwidth requirements are actually less than or equal to the data rate:

$$B_{\mathrm{E(M-FSK)}} = \frac{R_{\mathrm{Data}}}{\log_2 M} = R_{\mathrm{sym(M-FSK)}}. \tag{25}$$

The bandwidth expansion in M-FSK instead comes from the spread of M-frequencies used, which generally requires a parallel M-channel receiver design. This adds some complexity over the serially generated and received PPM waveforms mentioned above, since any imbalances in power or waveform fidelity in the M TX or RX parallel channels can degrade performance. As with DPSK, FSK waveforms can have a 100% duty cycle, making it well suited for use with average or peak power-limited transmitters. It can be implemented using a single frequency-modulated (FM) source with binary- [162,177–179] or M-ary modulation [160,180], or using external modulation and subcarrier multiplexing (SCM) [181,182]. These approaches generally require relatively tight wavelength spacing due to transmitter or receiver limitations [161]. However, with the use of independently-modulated wavelength-multiplexed-sources, or fast-tunable wideband DBR sources [180], arbitrarily wide wavelength spacing can be used. For parallel TX and RX designs, the duty-cycle per channel is

roughly $1/M$, assuming all symbols are equally likely, which leads to a similar reduction in the bandwidth-ratio and associated benefits that were described for M-PPM above.

2.2.1.7. *Polarization-Shift-Keying (PolSK)*

Binary-polarization-shift keying (2-PolSK) [183] can be used to extend the symbol set in the M-ary orthogonal formats described above by a bit (one bit per symbol), i.e., $M \rightarrow 2M$ and $k \rightarrow k + 1$, without increasing peak power or electrical bandwidth requirements per channel. The PolSK symbols encode binary information in two-orthogonal polarization bases, such as left and right-hand-circular polarizations or "s" and "p" linear polarizations, which can be implemented using combinations of phase modulator and parallel intensity modulators [184]. The performance of generalized M-PolSK has recently been evaluated for hybrid combinations with M-FSK [185] and multidifferential phase (DPSK) and amplitude modulation (ASK) [186,187]. However, for optimal sensitivity, an orthogonal M-ary signal set as defined in Eq. (12) is required. Like FSK and DPSK, PolSK is well suited for use with average or peak power limited transmitters, since it can operate with a 100% duty cycle. Further discussion of hybrid-modulation formats, performance, and implementation benefits is given in section 5.3.

2.3. *Comparison of RF and Optical Properties*

Until this point, the description of the EM field has been qualitative. However, there are many distinguishing characteristics between the RF carrier and the optical carrier that become apparent as the specific wavelengths are compared. A summary of values for a 1-cm microwave carrier and a 1-μm optical carrier at the same power level is shown in Table 1.

Table 1. Comparison of 1 cm RF and 1μm optical carrier characteristics.

	Microwave-RF	Optical	Units
λ	1.0×10^{-2}	1.0×10^{-6}	m
ν	3.0×10^{10}	3.0×10^{14}	Hz
$h\nu$	2.0×10^{-23}	2.0×10^{-19}	J
Power	1.0×10^{-9}	1.0×10^{-9}	W
r_{ph}	5.0×10^{13}	5.0×10^{9}	Photons/s
$r_{\mathrm{ph}} \times 10^{-9}$	5.0×10^{4}	5.0	Photons/ns
$r_{\mathrm{ph}} \times 10^{-10}$	5.0×10^{3}	0.5	Photons/100 ps

2.3.1. Diffraction

The factor of 10,000 between RF and optical wavelength leads to a substantial difference (10^8 or 80 dB) in free-space (vacuum) diffraction loss estimated in Eq. (6). Since diffraction loss increases with the square of the link distance-wavelength product, for equal TX and RX areas, the optical carrier can propagate a factor of 10,000 further before incurring the same loss, a primary reason for considering optical communications

for ultra-long-haul free-space communications. The optical signal is much is more directional and concentrated in the far field, delivering more signal to the target RX. But it does this at the expense of more challenging PAT, in contrast with RF systems which are more omni-directional.

The propagation losses in free-space grow as the square of the distance [see Eq. (6)] in contrast with fiber, in which the losses are exponential. The link loss for 12%, 100%, and 1000% free-space distance multipliers are 1, 6, and 20 dB, respectively. The channel loss for some potential planetary links, relative to an Earth to geosynchronous orbit link), are shown in Table 2 below. To highlight the difference between square-law diffraction and exponential losses, a link from Earth to Mars connected using low loss fiber with attenuation coefficient 0.2 dB/km would result in a loss of 80 million dB [104].

Table 2. Approximate maximum distance from Earth and relative diffraction loss for free-space propagation between the Earth and geosynchronous orbit, the Moon, Mars, Jupiter, and Saturn.

Orbit	Distance [km]	Relative Channel Loss [dB]
Geo	40,000	0
Moon	400,000	20
Mars	400,000,000	80
Jupiter	1,000,000,000	88
Saturn	1,600,000,000	92

Note that while pointing, acquisition, and tracking (PAT) is outside the scope of this paper, the TX and RX design can greatly facilitate PAT functionality. For example, prior to acquisition, the received SNR is generally poor due to low signal strength and since the benefits of filtering may need to be established. Such filtering may occur in several independent domains, e.g., spatial, temporal, spectral, and polarization; and acquisition or alignment in each may be necessary in order to acquire the link. Since the spatial acquisition generally poses the most fundamental challenge and time to acquire in FSO systems, streamlining acquisition or alignment in the other domains by providing fast clock recovery [188] recombined with stable TX polarization and wavelength that can be quickly aligned with the RX, can greatly expedite the overall acquisition process. PAT can also be simplified by increasing the number of RX spatial modes [1,2,47,189], by reducing tracking requirements and/or sensitivity to platform jitter, although this generally comes at the cost of increased RX noise.

2.3.2. Optical Detection

In contrast with RF electronics and detectors, which can directly detect the EM field and generally have bandwidths exceeding that of the carrier frequency, optical detectors are generally "square law detectors" with functional bandwidths that are a small fraction of the carrier frequency [190]. This means output is proportional to the square of the field, (proportional to the power intensity), with little direct dependence on optical phase, frequency, or polarization. Optical detectors can be extremely high speed, with bandwidths from DC to > 50 GHz. Because of optical detector bandwidth limitations, practical detectors have a response time which is generally much slower than the

~femtosecond (10^{-15}s) period of the optical field, so these detectors respond to the time average square of the field. The detectable low-frequency optical power is given by

$$P_s(t) = \frac{\left\langle \left| \vec{\mathbf{E}}_s(t) \right|^2 \right\rangle}{Z_0} = \left\langle \left| \frac{A(t)\cos(2\pi\nu t + \phi)\vec{\mathbf{p}}}{Z_0} \right|^2 \right\rangle$$
$$= \frac{|A(t)|^2 \left\langle \cos^2(2\pi\nu t + \phi)\vec{\mathbf{p}} \right\rangle}{Z_0^2} = \frac{|A(t)|^2}{2Z_0^2}, \quad [\mathrm{W/m^2}] \qquad (26)$$

where the time average of $\langle \cos^2() \rangle = \frac{1}{2}$. In high-speed designs, Z_0 is typically 50Ω in order to match RF waveguide impedance to minimize reflections.

In view of the intensity or power dependent response of optical detectors, the most commonly used optical modulation formats use intensity-modulated direct-detection (IM-DD) modulations. An example on-off-keying (OOK), in which the information is encoded by the presence (a logical "1") or absence (a logical "0") of the optical signal. In order to provide access to the phase, frequency or polarity of the electric field, more complex receivers which employ optical preprocessors are required. Preprocessors include wavelength dependent filters, interferometers, polarization beam splitters, and coherent measurements enable measurement of the electric field [153,157,160,161,179,191–195]. Further discussion of detector technologies, performance, and impact of noise sources is given in sections 2.3.5–2.3.7 and 4.

2.3.3. Technology Limitations

For some of the same reasons that fiber-optic communications presently dominates long-haul high-rate guided-wave applications, namely low channel losses and wide available bandwidth (with little or no regulatory limitations), FSO communications has a bright future. As summarized in Table 3 the technical advantages are considerable. For instance, there are roughly 4 orders of magnitude more bandwidth accessible via readily available WDM and ultra-wide-band, high-gain optical amplifiers. These optical technologies enable the trade of spectral-efficiency for improved photon-efficiency (as suggested in section 1.1), while maintaining the ability for scalable high-rate communications.

FSO terminals can also leverage robust wideband fiber-optic connectivity to permit modular construction [46] that is insensitive to electromagnetic interference (EMI).

On the other hand, RF technologies are generally more mature and have extensive ability to integrate. In terms of average TX power efficiency, RF transmitters have a ~3 to 6 dB advantage over amplifier-based optical transmitters, although this deficit is mitigated by the average-power-limited properties of optical amplifiers that enable the flexible use of variable-duty-cycle waveforms without reducing TX efficiency. Free-space RF communication is better suited for omnidirectional applications and provide improved reliability for links where clouds may be a limitation, However, for long-haul-high-rate free-space applications, especially space-based links, such benefits are overshadowed by the many advantages of optical technologies summarized above. Example link budgets in section 2.4, Table 4 based on fundamental carrier characteristics and in section 5.5, Table 15 based on currently available technologies illustrate these considerations.

Table 3. Comparison of RF and optical wavelength-dependent characteristics and technical limitations. *Note other optical bands are available, e.g., 0.85, 1 and 1.3 μm. The table includes references for efficient optical amplifiers at 1 μm [198, 204], which are generally ~1.5 to 2 times more power efficient than their 1.5 μm counterparts.

Parameter	RF	Optical
Wavelength (frequency)	~ 1 cm, (30 GHz)	1.5μm (200 THz)*,
RX noise limitation [dBm/Hz]	Thermal (KT = –174)	Shot ($h\nu$ = –159)
Bandwidth limitations	Government Regulated	Unregulated
	$\ll$ 30 GHz	$\ll$ 200 THz
State-of-the-art amplifiers		
Gain	10 dB	50 dB
Bandwidth	~ 50 GHz	~ 50 THz
NF	5 dB	3 dB
Transmitter Characteristics	Peak power limited	Average power limited
Integrability	Massive	Limited
Modulation Formats	Sophisticated	Basic
Efficiency (peak)	~ 50–80% [196]	~ 5–20% [172,197–200]
Efficiency (average)	~ 5%–46% [196,201–203]	~ 5–13% , 21% [172,197–200]
Waveguide Characteristics		
Type	Coax	Fiber
Bandwidth	~ 20 GHz	~ 50 THz
Loss	100 dB/km	0.2 dB/km
Free-space channel characteristics		
Diffraction Angle	10^{-2}/D	10^{-6}/D
Cloud penetration	Good	Poor

2.3.4. Average and Peak Power Limited Transmitters

RF transmitters are generally peak power limited, which means that there is an upper limit on the peak output power they can deliver (e.g., a maximum voltage). As a result, in order for RF systems to deliver maximum average power (the relevant metric for assessing communication performance), they need to operate with high ~100% duty cycles, which places limitations on the choice of modulation formats and waveforms. Furthermore, in order to maintain maximum transmitted power and operate at multiple rates, peak-power-limited systems need to alter the transmitted pulse width and shape so that the high duty cycle is preserved. In order to maintain optimum communication performance with this constraint, different receiver filters are necessary for each bit-rate, increasing RX complexity.

In contrast, optical transmitters are typically average power limited, which means that the average output power is independent of the transmitted waveform shape or duty-cycle. This provides significant advantages over peak-power limited RF systems, enabling optical communications systems to use aggressive pulse shaping techniques, and variable-duty-cycle modulation formats that can approach optimal performance

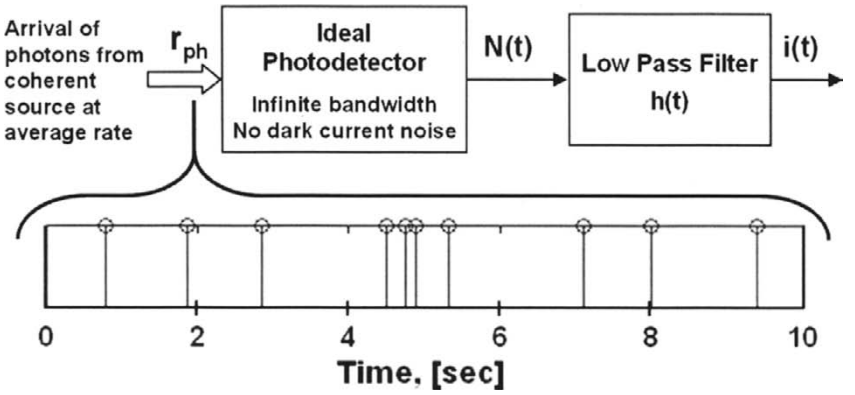


Fig. 10. Photodetection modeled as a fast noiseless ideal photon counting detector that generates a stream of detection events which form a discrete count $N(t)$ that is a function of the photon rate $r_{\rm ph}$, detection efficiency, and observation time-window described below. The detected impulse stream is then followed by an electrical filter with impulse response $h(t)$, that converts the detected events to a photocurrent.

over a wide range of rates while using simple high-sensitivity multi-rate receiver designs [10–15]. This subject is discussed further in sections 3 and 5.

2.3.5. Quantum Noise Limitations

When considering RF and optical quantum characteristics, the granularity of the photon flux is much more noticeable at optical frequencies, with rate corresponding to

$$R_{\rm ph} = P/(h\nu) \quad [\text{photons/s}], \tag{27}$$

where P is the optical power in W, and $h\nu$ is the photon energy in Joules. For a large number of photons per observation interval, the shot noise is relatively small relative to the average, and the photon flux appears continuous. However, for low photon flux, the shot noise or graininess in the interarrival times becomes apparent, as illustrated in Fig. 10. For the case of 1 nW of optical power and a 100-ps observation window in Table 1, on average, only $\frac{1}{2}$ a photon is detected, which means that statistically, no photons are detected in half the observations. Such quantum fluctuations in the detected photon counts, are fundamentally part of the optical signal, and therefore impose the ultimate limitation on optical receiver performance. This is in sharp contrast with RF receivers, which are typically dominated by thermal noise in the receiver. The impact of quantum and thermal noise on communication performance is discussed in further detail in references [20,121,127,190,205–207].

The statistics of photon arrivals for a coherent state, e.g., a laser source operating far above threshold, are Poisson [121,127] —with characteristics described in Eqs. (28)–(32) and illustrated in Fig. 11. The Poisson distribution in (28) gives the probability of detecting n photons in a time interval T given the average number of photon arrivals per interval is $< n >= N_{\rm avg}$:

$$P[n\,|N_{\text{avg}}] = \frac{(N_{\text{avg}})^n\, e^{-N_{\text{avg}}}}{n!} \underset{N\to\text{large}}{=} \frac{e^{-\frac{(n-N_{\text{avg}})^2}{2N_{\text{avg}}}}}{\sqrt{2\pi N_{\text{avg}}}}. \tag{28}$$

The right-hand expression in Eq. (28) is a discrete Gaussian approximation of the Poisson distribution which is accurate for large N_{avg}. As shown in Eq. (29), N_{avg} can be interpreted as the product of the rate of detected photons r_{ph} over observation interval T,

$$N_{\text{avg}} = r_{\text{ph}}T = \frac{\eta P}{h\nu}T = \eta\text{PPB}, \tag{29}$$

where η is detection efficiency, P is incident optical power, $h\nu$ is energy-per-photon, and PPB is the incident number of photons/bit when $T$is the bit period. For a Poisson distribution, the mean number of photon arrivals during the observation window, m, is equal to the variance, σ^2, as shown in Eqs. (30) and (31), which can be derived by twice differentiating the Taylor expansion of e^N [208].

$$m = \langle n\rangle = \sum_0^\infty nP[n\,|N_{\text{avg}}] = \sum_0^\infty \frac{n\,(N_{\text{avg}})^n\, e^{-N_{\text{avg}}}}{(n)!} = N_{\text{avg}}, \tag{30}$$

$$\sigma_{\text{sh}}^2 = \left\langle (n-\langle n\rangle)^2\right\rangle = \sum_0^\infty n^2 \frac{(N_{\text{avg}})^n\, e^{-N_{\text{avg}}}}{n!} - N_{\text{avg}}^2 = N_{\text{avg}}. \tag{31}$$

Accordingly, the SNR (in terms of received electrical power, proportional to detected photocurrent squared), given by[1]

$$\text{SNR} \equiv \frac{\langle n\rangle^2}{\langle \Delta n^2\rangle} = \frac{m^2}{\sigma_{\text{sh}}^2} = N_{\text{avg}} = \eta\text{PPB} \tag{32}$$

is also equal to the mean number of photons received. For the case where the observation interval is one bit duration, N_{avg} is equal to the number of photons/bit. As the mean number increases, the Poisson distribution becomes more symmetric and converges towards a Gaussian, also with equal mean and variance. While not noise free, the statistics of laser light are relatively quiet when compared to single-mode thermal light sources that follow Bose-Einstein statistics [121] shown also in Fig. 11.

From Eq. (32), the SNR for a shot-noise-limited signal can be interpreted as having a signal with N_{avg} photons accompanied by 1 noise photon with energy $h\nu$ Joules, or equivalently with a noise power spectral density:

$$N_{\text{shot}}(\nu) = h\nu \quad [\text{W/Hz}], \tag{33}$$

the shot noise power in 1 Hz of observed bandwidth. For an observation bandwidth of B Hz, the total shot noise power is

$$P_{\text{shot}}(\nu) = h\nu B \quad [\text{W}]. \tag{34}$$

For an ideal noiseless photodetector, the continuous-time photocurrent $i(t)$ generated by detecting a discrete photon stream is described by a random process with mean (signal) and variance (noise):

[1] Note that the SNR can alternatively be defined in terms of received electrical current or voltage rather than power. This alternative definition, commonly used in photon counting applications, yields $\text{SNR}_v \equiv \text{Mean}/\sqrt{\text{Variance}} = \sqrt{N_{\text{avg}}}$.

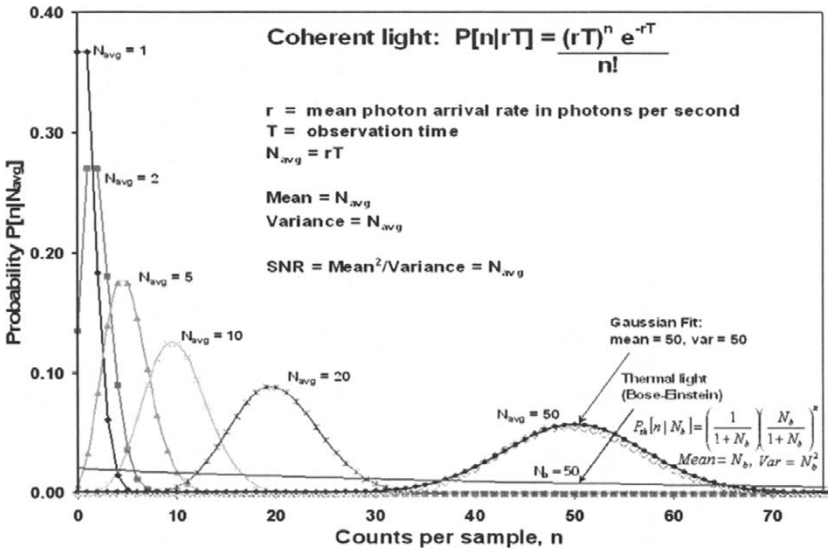


Fig. 11. The discrete Poisson probability distribution function (PDF) gives the probability of detecting n photons given a photon rate $R_{\rm ph}$ and observation window T, or equivalently the average number of photons in the interval $N = R_{\rm ph}T$, shown for $N_{\rm avg} = 1, 2, 5, 10, 20$, and 50. Also shown are a Bose-Einstein PDF (representing the statistics of thermal light in a single mode) with a mean of 50, and Gaussian PDF with mean and variance of 50 ($m = \sigma^2 = 50$).

$$i_{\rm avg}(t) = \langle i(t) \rangle = \frac{\eta q}{h\nu} \int_{-\infty}^{\infty} P(t)h(t-\tau)d\tau \quad [\text{amp}], \tag{35}$$

$$\sigma_{\rm sh}^2(t) = \left\langle (i(t) - i_{\rm avg})^2 \right\rangle = \frac{\eta q^2}{h\nu} \int_{-\infty}^{\infty} P(t)h^2(t-\tau)d\tau \quad [\text{amp}^2], \tag{36}$$

where $h(t)$ is the impulse response of the filtered photocurrent with integrated area normalized to a unit charge [20,127,208]. For example, given a photodetector with a simple exponential electrical impulse response,

$$h(t) = \frac{qe^{-t/\tau}}{\tau} u(t) \longleftrightarrow H(f) = \frac{q}{1 + j2\pi f\tau}, \tag{37}$$

where τ is the filter time constant and $u(t)$ is the unit step function, then with constant incident optical power P,

$$i_{\rm avg} = \frac{\eta q}{h\nu} P = r_{\rm ph} q \quad \text{and} \quad \sigma_{\rm sh}^2 = \frac{\eta q}{h\nu} P \left(\frac{q}{2\tau} \right) = 2q i_{\rm avg} B_e. \tag{38}$$

The far right-hand term in Eq. (38) is the well known general expression for the shot noise power (in units of amps^2) associated with the average photocurrent, where B_e is the noise equivalent power bandwidth of the electrical response over which the noise is observed, i.e.,

$$B_e = \int_0^\infty \left(|H(f)|^2 / H_{\max}\right) df. \tag{39}$$

For the case of the exponential filter in Eq. (37), $B = 1/(4\tau)$. We can solve for the quantum-limited SNR for direct detection by substituting Eq. (38) into Eq. (32), yielding

$$\mathrm{SNR_{DD}} = \frac{m^2}{\sigma_{\mathrm{sh}}^2} = \frac{\eta P}{2h\nu B_e} = \frac{r_{\mathrm{ph}}}{2B_e} = r_{\mathrm{ph}}\tau_{\mathrm{eff}} = N_{\mathrm{eff}}, \tag{40}$$

where τ_{eff} is the effective integration time of the electronic response, equal to 2τ in the example above. Note the similarities with the expressions for photon count in Eqs. (32) and (29) above. As with photon-counting, the direct-detected photocurrent SNR is proportional to the average number of photons N_{eff} received per effective observation interval. For communications, this is often the number of received photons-per-bit, which is reduced by any losses preceding detection. Examples of the impact of discrete photon arrivals and the resulting shot noise generated in the received photocurrent as a function of RX bandwidth are given in Fig. 12.

In practice, it is challenging to achieve quantum-limited performance due to the presence of other noise sources such as thermal noise (section 2.3.7) and dark current, effects that were neglected in the analysis above. Techniques of overcoming these limitations are discussed in sections 4 and 5.

2.3.6. Quantum-limited Direct Detection (DD)

One of the simplest optical receivers is an On-off-keyed (OOK), direct detection receiver. With knowledge of the photon statistics, the impact of shot noise on optical communication performance can be readily observed and quantified for the ideal intensity-modulation TX and direct-detection (IM/DD) photon-counting RX shown in Fig. 13. On-off-intensity-modulated-photons from the transmitter representing a logical "1" and no-photons (e.g., perfect modulation extinction) representing a logical "0" are received by a noiseless photon-counting receiver, (i.e., a Geiger Mode photon counting detector with no dark counts, after pulsing [30–32,209,210] background noise, or other parasitics such as timing jitter and inter-symbol-interference (ISI)).

Without background noise, the optimum decision rule is: if no photons are detected, assume that data $D =$ "0" was sent; and if photons were detected ($\hat{D}$= "1"), assume that $D =$ "1" was sent. For these rules, bit errors only occur if a photon is not detected, e.g., $\hat{D} =$ "0" when $D =$ "1" was sent, which is an "erasure" error. The photon-counting OOK probability of bit error is given by

$$\begin{aligned} P_{\mathrm{BE(OOK)}} &= P[D=0]P[\hat{D}=1|D=0] + P[D=1]P[\hat{D}=0|D=1] \\ &= \frac{1}{2}P[n=0|N] = \frac{1}{2}e^{-r_{\mathrm{ph}}T} = \frac{1}{2}e^{-2N_{\mathrm{avg}}}, \end{aligned} \tag{41}$$

where the number of photons during the observation interval (N) when a "1" is received is twice the average number received ($N = 2N_{\mathrm{avg}}$), which is true for equal probabilities of transmitting "0" or "1" data. From Eq. (41), the 10^{-9} BER is achieved for $N_{\mathrm{avg}} = 10$ photons/bit and corresponds to the "quantum limit" for the ideal IM-DD photon counting OOK, since the only limiting noise source is signal shot noise.

A similar analysis can be performed for ideal photon counting binary PPM, in which each symbol is determined by the location of the optical pulse, with "0" corresponding to photons detected in the first half of the symbol and "1" from the second

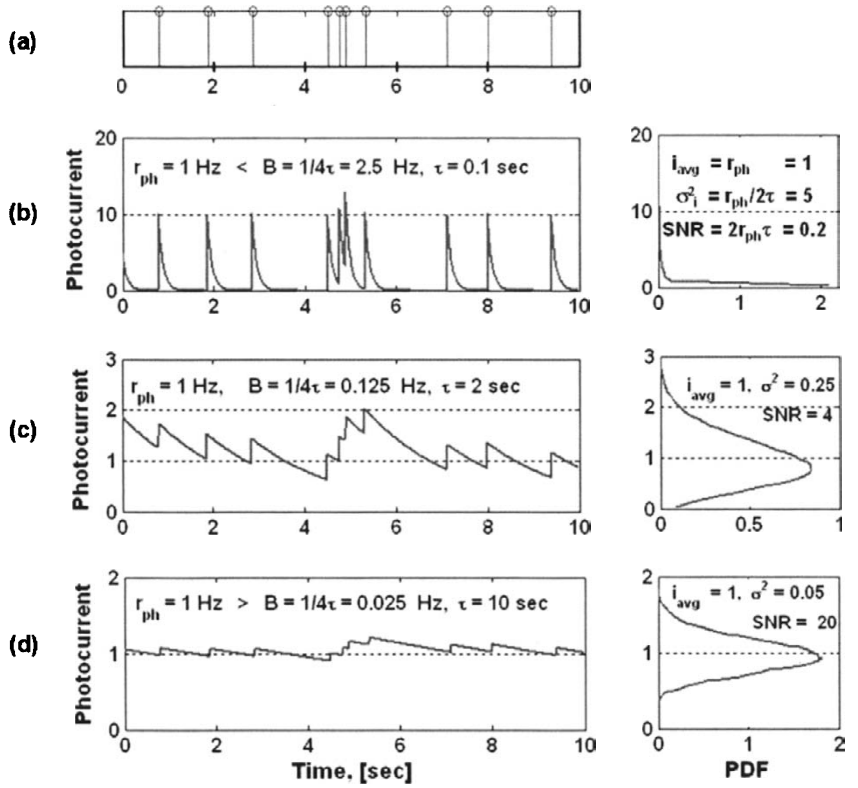


Fig. 12. Photocurrent and shot noise. The granularity of a stream of photons and the resulting variance in the signal is shown as a function of the effective receiver filter bandwidth B_e, where $B_e = 1/4\tau$ for the low-pass filter described in Eq. (37). (a) A discreate Poisson impulse stream of photons with arrival rate $r_{\rm ph} = 1$ photon/s. (b)–(d) left show the continuous-time photocurrent as filter bandwidth B_e is reduced two orders of magnitude from 25 to 0.025 Hz, and (b)–(d) right show the corresponding probability distribution functions (PDFs) of the received photocurrent for each filter bandwidth. The SNR improves as the bandwidth decreases or equivalently as the filter integration time increases. But even with significant filtering in which multiple photons are detected within the filter time constant, the (shot) noise added by the random arrival of discrete photons is apparent. Note that for simplicity, the units of photocurrent have been normalized per unit charge.

half. Bit errors in this ideal photon-counting system only occur when no photons are received. In this case, a "coin flip" is used to decide (arbitrarily) which slot is the winner (since both slots are identically zero in the ideal noiseless case) [163,211]. The coin flip gives the correct answer $\frac{1}{2}$ the time, yielding

$$P_{\rm BE(2-PPM)} = \frac{1}{2}e^{-r_{\rm ph}T/2} = \frac{1}{2}e^{-N_{\rm avg}}. \tag{42}$$

This is similar to the OOK result but note the factor of $\frac{1}{2}$ in the exponent that comes from the fact that the number of photons during any observation interval (either

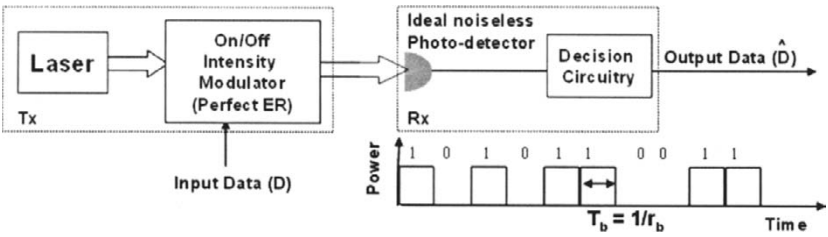


Fig. 13. Ideal photon-counting OOK TX and RX.

"0" or "1") is the same and equal to the average number received ($N = N_{\text{avg}}$). Thus, the quantum-limited photon counting sensitivity for 2-PPM is N_{avg}= 20 photons/bit at 10^{-9} BER.

Extending the binary case to M-ary PPM, in which each symbol is divided into M positions (slots) and carries $\log_2(M)$ bits per symbol, the probability of error within a symbol is

$$P_{\text{SE(M-PPM)}} = \left(\tfrac{M-1}{M}\right) e^{-N_{\text{avg}}}, \tag{43}$$

where the coefficient term is the (uniform) probability of symbol error when no photons are received [211], and the exponential term is the probability of detecting no photons.

Substituting Eq. (16) into Eq. (43) and accounting for the number of bits per symbol, yields the general expression for ideal photon-counting M-PPM:

$$P_{\text{BE(M-PPM)}} = \frac{1}{2} e^{-N_{\text{avg}} \log_2 M} = \frac{1}{2} e^{-N_{\text{avg}} k}. \tag{44}$$

As M increases, the PPB required to achieve a particular BER is reduced by $\log_2 M$, the number of bits per symbol, which can extend to arbitrarily low values in theory without additional coding. However, as noted by Yamamoto and Haus [27,28], a practical theoretical limit of ~0.02 PPB would ultimately be imposed by the ratio of photon to thermal noise energy. In practice, photon-counting receivers are not ideal, having parasitic effects such as dark counts, after pulsing, timing jitter, and limited bandwidth [1,2,30–32,209,210], which impose additional constraints on performance. Furthermore, background noise and transmitter imperfections can further degrade receiver performance [21,212].

To illustrate this point, given a TX with imperfect modulation extinction (ER = power off/power on), the probability of bit-error P_{BE} for the OOK photon-counting expression in Eq. (41) becomes

$$\begin{aligned} P_{\text{BE(OOK,ER)}} &= P[\hat{D} = 0 | D = 1] + P[\hat{D} = 1 | D = 0] \\ &= \frac{1}{2} P[n = 0 | N] + \frac{1}{2} P[n > 0 | N \cdot \text{ER}] \\ &= \frac{1}{2} P[n = 0 | N] + \frac{1}{2} \left(1\text{-}P[n = 0 | N \cdot \text{ER}]\right) \\ &= \frac{1}{2} \left(e^{-(2N_{\text{avg}} - N_{\text{avg}}\text{ER})} + 1 - e^{-N_{\text{avg}}\text{ER}} \right). \end{aligned} \tag{45}$$

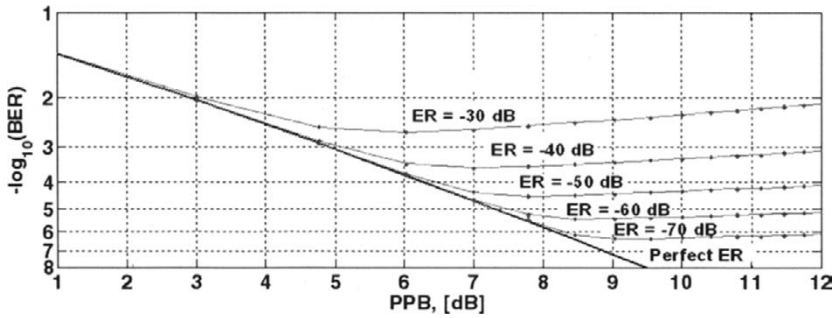


Fig. 14. Photon-counting OOK BER curves for varying TX extinction ratio (ER).

The $(1-e^{-N_{\text{avg}}\text{ER}})$ term is the modulator seepage penalty due to the N_{avg}ER (signal-dependent) background noise photons that arrive during the "0s". As shown in Fig. 14, the background noise resulting from the imperfect ER imposes an error floor that is a few times higher than the ER, (i.e., a 10^{-3}ER $\Rightarrow\sim 2\times 10^{-3}$ error floor). The error rate continues to worsen as the signal is increased since the optimum threshold is a function of the signal dependent "0" noise, not zero as assumed in Eq. (45). This illustrates a potential limitation of photon-counting using this approach in the presence of background noise, and the need for powerful FEC[42,213] that can generate error-free performance with input error-rates of $\sim 10^{-2}$[33,214] to overcome such limitations:

$$P_{\text{BE(2-PPM)}} = \frac{1}{2}e^{-\frac{N_{\text{avg}}^2}{4N_b}}, \quad (N_{\text{avg}} \ll N_b). \tag{46}$$

Binary PPM is similarly impacted by the presence of background-noise photons, although an adaptable decision can be easily implemented by comparing the first-half and second-half of the symbol. For the case where the background N_b is much larger than the signal, the BER expression becomes [212].

For average power limited transmitters, extending the modulation format to M-PPM is a technique for overcoming background noise (improving SNR) without increasing the average power. As shown in Fig. 15, both OOK and 8-PPM waveforms have the same average power, the peak of the OOK waveform is $\frac{1}{2}$ the background, whereas the peak of the 8-PPM waveform is twice the background, making it easier to distinguish (improving the SNR over the observation window), while conveying more information (3-bits per symbol), at 1.5 times the bit-rate. A more detailed discussion of M-PPM and other orthogonal formats is given in section 5.3.

2.3.7. Thermal Noise

There are two fundamental noise limitations in RF communications systems. One is thermal noise in the receiver, which is due to the temperature-dependent random-thermal-motion of charge carriers in resistive elements. This is often referred to as

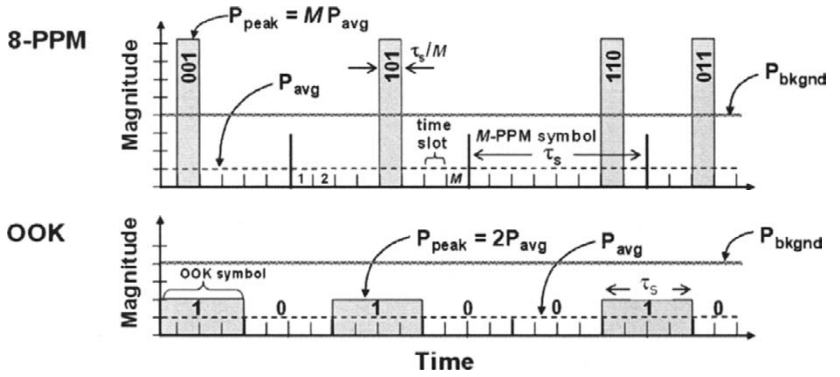


Fig. 15. On-off-keyed (OOK) and 8-ary Pulse Position Modulation (8-PPM) waveforms are shown for the same average power (P_{avg}) and background noise level. For a given average power, the peak power of the M-PPM waveform increases with m, improving the ability to distinguish the signal from the noise.

Johnson-Nyquist noise after Johnson [215], who first measured thermal noise in conductors, and Nyquist [216], who derived a theoretical expression to fit Johnson's measurements in 1928. The available thermal noise power spectral density delivered to a matched load can be described by [216,217]

$$N_{\text{th}}(\nu, T) = \frac{h\nu}{e^{h\nu/k_B T} - 1} \quad \underset{\nu<100\text{GHz}}{\longrightarrow} \quad \approx k_B T \quad [\text{W/Hz}], \tag{47}$$

where k_B is Boltzmann's constant (1.38×10^{-23} J/K), T is temperature of the resister in Kelvin, and the approximation $k_B T = -174$ dBm/Hz at room temperature. Note that in contrast with shot noise, thermal noise is dependent on temperature, and for practical bandwidths ($<\sim$100 GHz), is independent of signal power or wavelength. (See [20,156,218] for additional detail.)

The second noise source is received thermal blackbody background radiation, which is also described by Eq. (47) for a single mode and polarization, where T in this case is the average temperature within the field-of-view of the receiver. This shared expression for the two thermal noise sources arises from the fact that the resistive thermal noise can be viewed as the blackbody radiation emitted into a single mode. The contributions from these two thermal noise sources can be combined in a single effective temperature

$$T_{\text{eff}} = T_{\text{RX}} + T_{bb}. \tag{48}$$

The general expression for noise density that incorporates both thermal and shot noise is given by [217]

$$N(\nu, T_{\text{eff}}) = N_{\text{th}}(\nu, T_{\text{eff}}) + N_{shot}(\nu) = \frac{h\nu}{e^{h\nu/k_B T_{\text{eff}}} - 1} + h\nu \quad [\text{W/Hz}], \tag{49}$$

which is plotted in Fig. 16 for a receiver operating at four effective temperatures, 6000K (sun) [47,219], 290K (room temperature), 77K (liquid Nitrogen), and 2K (liquid Helium).

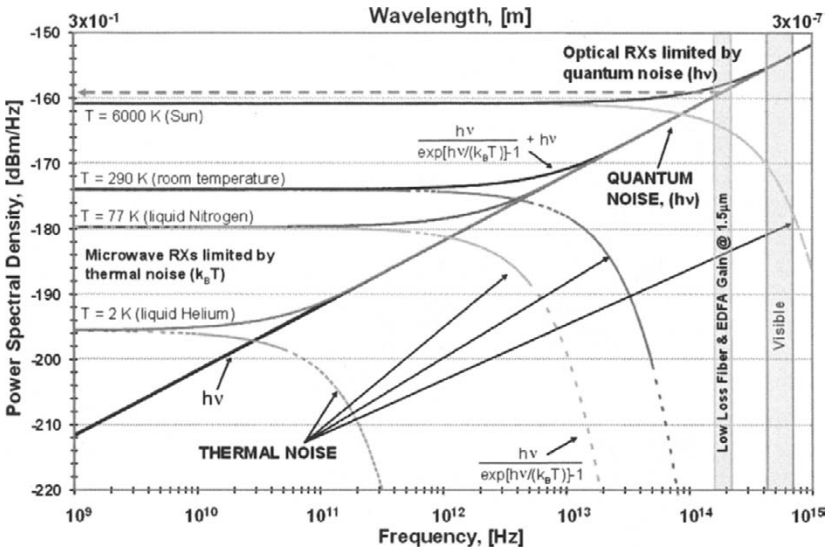


Fig. 16. The fundamental limits of receiver noise density at various temperatures. At high frequencies (e.g., optical), quantum noise (shot noise) dominates. At low frequencies, (e.g., RF or lower), temperature dependent thermal noise dominates. *Adapted from S.B. Alexander* [20].

Note that at 1.5 μm wavelengths, the shot noise dominated noise density, N_{shot}, is –159 dBm/Hz, 15 dB larger than the thermal noise component at room temperature.

Expanding on an example by Alexander [20], consider the fundamental noise limitations of a 30-GHz (λ=1 cm) microwave and a 300- THz (λ=1 μm) optical receiver, both with 1 GHz bandwidth operating at operating in room temperature (290K), looking into deep space with a conservative estimate of background temperature of 70K [220]. Thermal noise dominates the RF receiver with a noise power of

$$P_{\text{th}} = k_B T_{\text{eff}} B = (1.38 \times 10^{-23})(360)(1 \times 10^9) = 5 \times 10^{-12}\ \text{W} = -113\ \text{dBW}, \tag{50}$$

whereas quantum noise dominates the optical receiver with a noise power of

$$P_{\text{shot}} = h\nu B = (6.63 \times 10^{-34})(3 \times 10^{14})(1 \times 10^9) = 2 \times 10^{-10}\ \text{W} = -97\ \text{dBW}. \tag{51}$$

Thus the RF receiver limited by thermal noise can be be 40 times (16 dB) more sensitive than the shot-noise limited optical receiver, and this advantage can potentially be increased to 200× and ~2000× by lowering the effective temperature of the RF receiver to 77K (using LN_2) or ~8K, respectively. Note that in order to achieve the full limits of RX cooling, it must be looking directly into deep space (without channel loss), where the blackbody contribution to the effective background temperature is ~4K [220,221]. It is important to take into account such factors when making comparisons between RF and optical link designs, especially when nearly ideal superconducting photon-counting detectors [33,34,210,222] are considered.

This illustration highlights some additional benefits of free-space optical receivers over RF receivers—they can be much less sensitive to background noise, potentially allowing for a wider range of use. By using, for example, a super-cooled optical receiver including a narrow-band optical filter, thermal noise can be virtually eliminated from the receiver, since the incoming blackbody radiation at optical frequencies is significantly lower than that of shot noise for practical operational temperatures (i.e., $T_{bb} < 300\text{K}$). Even when pointing directly at the sun, where $T_{bb} = \sim 6000\text{K}$ (assuming the telescope doesn't melt), the blackbody noise is still $\sim$10 dB lower than the shot noise (at 1 μm)[2] for a single-mode receiver (single spatial-temporal and polarization mode), degrading a shot-noise limited signal by only $\sim$0.5 dB (see [47,223] for detailed analysis of background radiation and its impact on free-space optical links). On the other hand, for an RF receiver, this would degrade performance in excess of 13 dB. This illustrates potential benefits of using true single-mode optical receivers. While coupling into a single spatial mode can be more challenging, the reduced background due to strong spatial-filtering and available low-loss matched spectral filters can enable operation that is insensitive to background noise levels.

2.4. *Example Sensitivities and Link Budget*

Based on the fundamental noise limitations described above, Table 4 shows a rudimentary link budget comparing Optical and RF free-space performance over a 40,000 km link (e.g., distance to geosynchronous orbit) with the same transmit power-aperture product.

Table 4. Comparison of simplified FSO and RF link budgets based on fundamental noise-limited receiver sensitivities for a common 1-GHz RX bandwidth (at room temperature), 0.3-m TX antenna diameter and a 1.5-m RX antenna diameter, and 40,000 km link distance (Earth to GEO). *Adapted from S. Alexander*[20].

Item	$\lambda = 1 \times 10^{-6}$ m	$\lambda = 3 \times 10^{-2}$ m	Units
Transmitter			
1) Avg. TX Power	0	0	dBW
2) TX Losses	–2	0	dB
Channel			
3) Space Diffraction Loss	–41	–131	dB
Receiver			
4) RX Loss	–2	0	dB
Power at Comm. Detector	–57	–131	dBW
RX Sensitivity at 1 Gbps	–97	–114	dBW
Link Margin	50	–17	dB

Given these constraints, and some practical implementation considerations, the overall optical link has a 67-dB advantage, largely attributed to the optical carrier

[2] Note, at 1 μm, the kT approximation no longer holds, and blackbody radiation needs to be calculated with the exact expression on the left-hand side of Eq. (47).

size, and corresponding spectral, temporal, and spatial advantages. Furthermore, the 1000× bandwidth advantage of FSO systems, which can be used to scale data rates or improve RX sensitivity [22], is not accounted for in the basic accounting in Table 4. A more detailed link analysis which incorporates realizable modulation and coding performance suggested in Fig. 1 is given in section 5.5.

It's important to note that some of the RF deficit can be improved by increasing the RF aperture, which is practically easier due to reduced wavefront quality requirements. Also, RF systems usually have better transmitter efficiencies, providing another ~3–7 dB benefit, but ~60 dB is still an enormous shortfall to overcome. For instance, if both TX and RX aperture diameters were increased by a factor of ten, this would improve the RF link by 40 dB, still leaving a factor of 100 (20 dB) deficit. Thus, the net power efficiency advantage combined with the fundamental ability to efficiently scale to higher data rates, make free-space laser communications a promising technology for future long-haul communication links.

3. Transmitter Technologies

Factors that impact optical communication transmitter performance include the following: modulation bandwidth and extinction ratio, waveform generation and fidelity (control of shape and parasitic chirp), average and peak output power, output polarization state, wavelength, spatial profile and perhaps most important, wall-plug (electrical-to-optical conversion) efficiency (Joules required per photon transmitted). In contrast with transmitters for fiber-optic networks, in which the average and peak power that can be delivered over the fiber channel are generally limited by channel nonlinearities, the FSO link has no such constraint. While there are power limitations within the FSO transmitter, they are several orders of magnitude higher than those used in fiber networks.

Present lasercom transmitter designs incorporating high-power optical amplifiers, can reliably output in excess of 10W average and ~1 kW peak power in a single spatial mode with little waveform distortion or power loss due to nonlinear spectral broadening [21,198] with potential to increase to 100 kW levels [224]. The lack of dispersion over the FSO channel allows for more flexibility in wavelength selection and greatly reduced need for spectral efficiency. There are ~10 THz of optical bandwidth in 1.5 μm Erbium band [92,93,126], and other bands are available for the FSO link such as 0.5 and 0.8 μm [81–83,98–101,225], 0.98 nm [226], 1.06 μm [1,2,223,227–230], 1.3 and 10.6 μm [223]. While still subjected to practical availability and reliability constraints, the FSO transmitter and receiver designer has considerable leeway in optimizing the overall link efficiency,

$$\eta_{\text{link}} = \eta_{\text{TX}}\eta_{\text{ch}}\eta_{\text{RX}} \quad [\text{J/bit}], \tag{52}$$

the Joules-transmitted per bit-received. Here, η_{TX} is the transmitter (electrical to optical) efficiency [J/photon], η_{ch} is the net channel loss, and η_{RX} is the net receiver efficiency [photons/bit]. While the TX contribution η_{TX} is accounted for explicitly in the efficiency with which it can generate the photons, the TX indirectly influences channel loss, which is wavelength dependent, and receiver sensitivity, which is dependent upon modulation format and TX waveform fidelity.

3.1. *Direct Modulation and Semiconductor Laser Sources*

Present-day semiconductor lasers are robust, compact, efficient sources of coherent light. For Telcordia-qualified laser diodes [4,5], the mean-time-to-failure (MTTF) can exceed 100 years ($\sim 1 \times 10^6$ h) [231]. A practical feature of semiconductor lasers is that the output power can be directly modulated, at speeds approaching the relaxation oscillation frequency [232,233], f_{RO}, which is given by

$$f_{\mathrm{RO}} = \frac{1}{2\pi}\sqrt{\frac{\frac{c}{n}\frac{dg}{dN}S}{\tau_{\mathrm{ph}}}}, \tag{53}$$

where c/n is the velocity of light in the semiconductor material, dg/dN is the differential gain, S is the photon density in the cavity, and τ_{ph} is the photon lifetime. At output power levels of 10–25 mW, f_{RO} can exceed $\sim$10 GHz, which is particularly useful for subcarrier-multiplexing many narrow-band MHz class analog signals onto a single laser source, a capability that is commonly used by the cable television (CATV) industry [234,235]. However, for wide-band digital applications, direct modulation rates are practically limited to lower values, e.g., a few GHz, by parasitic capacitance in drive electronics or chirp-induced penalties [236].

As discussed in section 1, one of the simplest forms of digital modulation is on-off-keying (OOK), in which the logical information is imparted on the optical carrier by switching the light on and off. While direct modulation is an easy means of applying the intensity modulation suitable for OOK, there is a residual wavelength (or frequency) modulation or chirp ($d\phi/dt$) due to fast current density fluctuations and a slower temperature dependence of the refractive index in the active layer which causes a shift in the laser wavelength [233,237]. In fact, the fast change in laser frequency can be used for directly generating FM waveforms [160,162,177,178,180,238]. For the case of large frequency deviations, the residual IM due to current changes may need to be compensated. This can be achieved via dynamic compensation in the time domain [239] or relatively static gain/wavelength flattening in the frequency domain.

Chirp is often defined in terms of the α-parameter [240], which is the ratio of amplitude modulation to phase modulation, where

$$\alpha = \frac{d\phi}{dt}\left(\frac{1}{\frac{1}{E}\frac{dE}{dt}}\right) = \frac{d\phi}{dt}\left(\frac{1}{\frac{1}{2I}\frac{dI}{dt}}\right), \tag{54}$$

and I, E and ϕ, are instantaneous intensity, field amplitude and phase.

The chirped-output waveforms from a direct-drive laser have a broadened spectrum, that can be much larger than the modulating signal spectrum, i.e., the waveforms are not transform limited[3] (see, e.g., [236]). An example of a chirped waveform is shown in the spectrogram in Fig. 17.

The spectrogram is measured using a time-resolved technique [241–244] with time and frequency resolution of 100psec and 5 GHz, respectively. A direct-drive $\sim$75 mA current pulse causes a $\sim$20 GHz frequency shift in the laser center frequency, yielding a fast wavelength tuning response $d\lambda/di_{\mathrm{ac}} = -280$ MHz/mA. In addition, transient

[3] The transform limit (or Fourier limit), can be described as the minimum pulse duration which is possible for a given optical spectrum of a pulse. A transform-limited pulse has no chirp, a minimum time-bandwidth product, and a spectrum that is the Fouier-Transform of the pulse envelope.

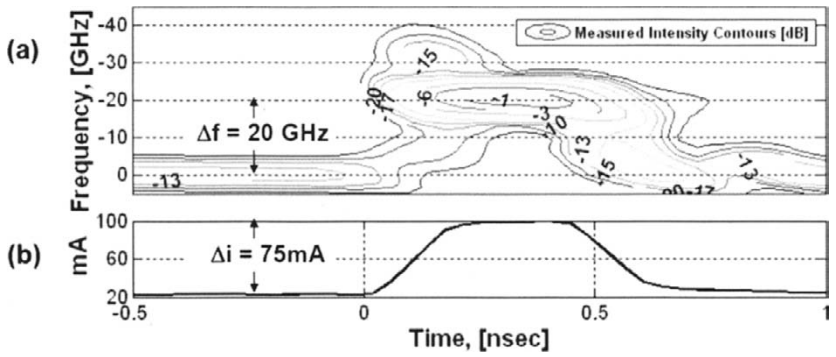


Fig. 17. (a) Measured time-frequency spectrogram (top view) of a directly-driven DFB laser biased at 25 mA, about 5 mA above the ~20 mA threshold current. Contours represent lines of constant power in dB, with the peak normalized to 0 dB. Due to the low bias current, the laser is operating near threshold, yielding relatively good ~13 dB ER (b) Measured input drive-current, with ~400 ps pulse width and ~75 mA peak current change. During fast current transitions, spectral components are generated that far exceed the transform-limited waveform spectrum, especially at the leading edge—which produces the characteristic '*h*' pattern above. This effect grows along with timing jitter on the leading edge as the laser bias approaches the laser threshold.

spectral generation, especially during the leading edge in semiconductor lasers biased near threshold, can be very wide band, exceeding 100 GHz. This leads to increased dispersion penalties in fiber-optic links and can limit receiver sensitivity in nearly-quantum-limited free-space links [245,246]. Furthermore, since chirp is a function of modulation depth, communication links which employ direct modulation need to balance conflicting extinction ratio and dispersion penalties.

In fiber-optic links, such considerations practically limit the extinction ratio (ER) to ~10 dB, which in turn can degrade RX sensitivity with ~3 dB penalties in preamplified RXs [see Eq. (64) in section 3.5.2], and limit modulation rates to less than 2.5 Gbps [233,247]. Since transmitter cost is generally a bigger concern than performance in telecom applications, directly-modulated lasers have widely been used for lower-rate (<~2.5 Gbps) links. But in FSO links, receiver sensitivity is often a design driver, and due to the limitations noted, conventional direct modulation is generally not considered the preferred approach. Channel capacity and bit-error-rate (BER) performance are degraded significantly by poor ER and chirp-induced spectral broadening, which increases the mismatch penalty between the TX signal and RX filter, a subject discussed in more detail in section 5.1.

3.1.1. Spectral Shaping

Direct-drive laser chirp and extinction ratio characteristics can be modified via subsequent optical filtering, a process that converts the frequency-shifted laser output to intensity modulated waveforms. When followed by an average power limited amplifier, this combination can be used as a power efficient source of high-fidelity transmit waveforms [10] despite the loss of the filter (see section 3.5.5). Reduced chirp and

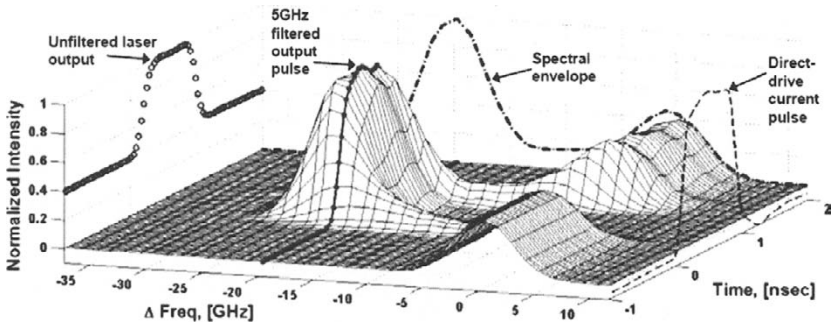


Fig. 18. Measured time-frequency spectrogram (oblique view) of a directly-driven DFB laser biased at 75 mA, with a ~20 mA threshold current. Time and frequency resolution are approximately 100 ps and 5 GHz, respectively. The response to a 1-ns, ~75-mA current pulse starting at time=0, increases output power by 2.4 times (~4 dB) and shifts the laser center frequency by 20 GHz.

improved ~13 to 14 dB ER have been demonstrated using this approach [244,248]. More recently, direct-drive chirp-managed lasers (CMLs) [249], which incorporate periodic filters into the laser package, have demonstrated impressive 10 Gbps long-haul fiber-communication performance without dispersion compensation [250,251]. Due to the complex interaction of pulse shape, dispersion, nonlinearities, and power launched over the fiber channel, these demonstrations showed better overall link performance with poor ~8 dB ER. However, these results are contrary to the impact of ER over the FSO channel, where good ER, generally > 30 dB is necessary to avoid performance penalties, especially in low-duty-cycle applications.

With narrower optical filtering, however, these approaches show promise for achieving low-chirp high-ER performance with compact power-efficient direct drive lasers, especially when used in conjunction with a subsequent TX amplifier (a subject discussed further in section 3.5.5). As shown in the spectrogram in Fig. 18, the (unfiltered) ~1 ns laser output has a ~4 dB ER and a ~20 GHz of spectrum. After narrow-band optical filtering, the ER is improved to ~30 dB and the spectrum is reduced to ~5GHz. As the external optical filter spectrum approaches the match to the desired signaling waveform, nearly transform-limited waveforms can be generated.

Another limitation with direct-drive laser modulation arises due to the non-uniform FM response that results from competing thermal and carrier modulation effects [238,241,252–256] that can lead to waveform distortion and pattern dependent errors. This can be compensated by pre-equalization of the modulation signal with a passive network in some cases [160,177,238]. However, this solution reduces the FM response and consequently leads to increased drive requirements. Another means of mitigating these effects is to limit the low-frequency drive content through pattern coding/scrambling [160,162,248], which can be achieved by driving the laser with low-duty-cycle waveforms such M-ary PPM or FSK modulation formats (see Fig 9). In addition, for low-duty-cycle waveforms, directly-modulated DFB lasers can be used in cascade with an external modulator and/or narrow-band optical filters to achieve

time/frequency windowed direct modulation. This can be a power-efficient technique to obtain high modulation extinction, while maintaining stable wavelength and near-chirp-free waveforms during transmission, a subject discussed further in section 3.5.2.

3.2. Semiconductor Laser Structures

There are several semiconductor laser design structures that are commonly available today. Fabry-Perot lasers have the simplest structure, but usually output multiple longitudinal (frequency) lasing modes, and therefore, are not well suited as signal sources for high-performance FSO links. But they are well suited as pump laser sources for commonly used Erbium-doped fiber amplifiers (EDFAs), Ytterbium-doped fiber amplifiers (YDFAs) and codoped Er:Yb EYDFAs. Here the increased spectral bandwidth is not a liability owing to broad absorption bands in the gain media [126] and since the effective reduction in spectral density lowers the risk of SBS penalties [110,111]. In order to provide some spectral selectivity and stability, Fabry-Perot-based pump lasers are often used in conjunction with narrow-band reflectors that are part of the laser cavity, or external stabilization filters such fiber-Bragg-grating (FBG) reflectors [257–259] to provide feedback that efficiently lock the center of the lasing spectrum (composed of 5 to 10 frequency modes) to that of the FBG.

Semiconductor lasers with integrated distributed Bragg reflectors (DBR) have been used to provide agile, tunable, single lasing modes in simultaneous time-division-multiplexed and frequency-division-multiplexed (TDM/FDM) digital communications [180,260]. In such lasers, great care needs to be taken in the control of the drive current and chip temperature in order to provide accurate and stable wavelength control to prevent mode-hopping. The use of such tunable sources is gaining traction in the telecom industry for flexible reconfigurable WDM networks, and especially since a single widely tunable laser could be used as a spare for any other channel.

For stable single-frequency operation, distributed-feedback (DFB) lasers are most commonly used throughout the telecom industry. A representative COTS DFB laser is shown in Fig. 19. The DFB laser has an index grating written into the gain medium which allows for only one lasing frequency mode within the gain-bandwidth of the laser. This typically provides a side-mode-suppression ratio (SMSR) exceeding ~45 dB, which precludes mode-hopping to other Fabry-Perot resonances. The DFB laser linewidth ($\Delta\nu$) is a function of output power and can be influenced by external reflections [233,261–263]. Typically $\Delta\nu <\sim 1$ MHz for well-isolated 20 mW DFB lasers, making them well suited for DPSK communication rates exceeding ~1 Gbps (see section 5.2), since coherence length penalties are incurred as the laser linewidth exceeds ~1% of the bit rate[4] [264].

[4] Linewidth penalties for DPSK are negligible when the interference between the differential bits is complete. This occurs when the coherence time of the signal is sufficiently long, i.e., $\tau_c \gg \tau_d$, where the laser coherence time $\tau_c = 1/\Delta\nu$, and τ_d, is the time delay between the differentially encoded bits. This ensures that coherent interference is achieved over the entire bit duration. Typically, DPSK data are encoded using adjacent bits, so that $\tau_d = \tau_b$, the bit period, with bit rate $R = 1/\tau_b$. However, DPSK data can be differentially encoded with a multiple-bit delay, e.g., $\tau_d = m\tau_b$, where m is a positive integer. For this general case linewidth penalties are avoided when $R \gg m\Delta\nu$, and thus, multi-bit differential-encoding imposes stricter linewidth requirements.

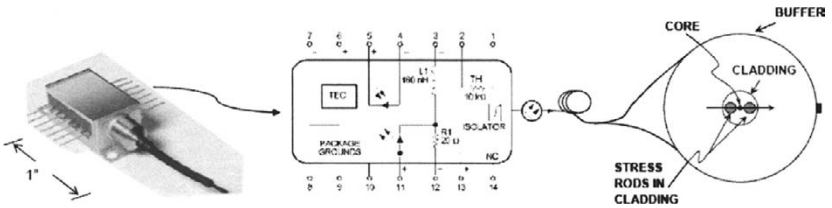


Fig. 19. Typical semiconductor DFB lasers come packaged in a 14-pin butterfly package with built-in thermo-electric cooler (TEC) and thermister for temperature control, and an internal optical isolator for increased stability. Since most external modulators require a polarized input, DFBs often come with polarization maintaining (PM) output fiber. In applications where tight wavelength stability is required, an internal etalon and monitor diode can be included within the package to provide an independent measure of wavelength that can provide sub-GHz wavelength stability when combined with external feedback [265].

3.3. Laser Wavelength Control

For optimum communication performance, narrow optical receiver filters that are well matched to the transmit signal spectrum are used (see section 5), requiring precise wavelength alignment of TX signal and RX filters. This can achieved via RX tracking of the incoming signal, but this approach is not always desirable from the system perspective. For instance, providing open-loop TX and RX control speeds acquisition and enables an FSO system to coast through fades, without having to re-acquire wavelength. In WDM applications, tight control of TX wavelength is valuable for both FSO and fiber-optic links, allowing for closer channel spacing, which improves spectral efficiency and overall link capacity. Precise open-loop control of the transmitter signal wavelength can provide additional benefits in WDM-DPSK links by facilitating acquisition and enabling cost-effective multi-channel DPSK receiver designs [266] as discussed further in section 5.2.

The steady-state laser wavelength is dependent on the optical path length within the laser cavity, which is a function of both laser temperature and average drive current. At 1.55 μm, these relations are given by

$$\frac{d\lambda}{dT} \approx 0.1\frac{\text{nm}}{^\circ\text{C}} = 12.5\frac{\text{GHz}}{^\circ\text{C}}, \tag{55}$$

$$\frac{d\lambda}{di_{\text{dc}}} \approx 0.004\frac{\text{nm}}{\text{mA}} = 0.5\frac{\text{GHz}}{\text{mA}}. \tag{56}$$

For a $\pm 10^\circ$C range with 10-bit control (corresponding to $\pm 0.01^\circ$C precision), the typical DFB laser can be temperature tuned over a range of $\pm$1nm ($\pm$125 GHz at 1.55μm) with $\pm$125 MHz resolution. Similar resolution can be obtained with current control better than $\pm$0.25 mA as shown in Fig. 20. Note that the wavelength dependence on drive current is complex (see, e.g., [233,238,253,256]), having fast components dependent on current density and slower components that are a function of carrier and thermal diffusion resulting from current transients.

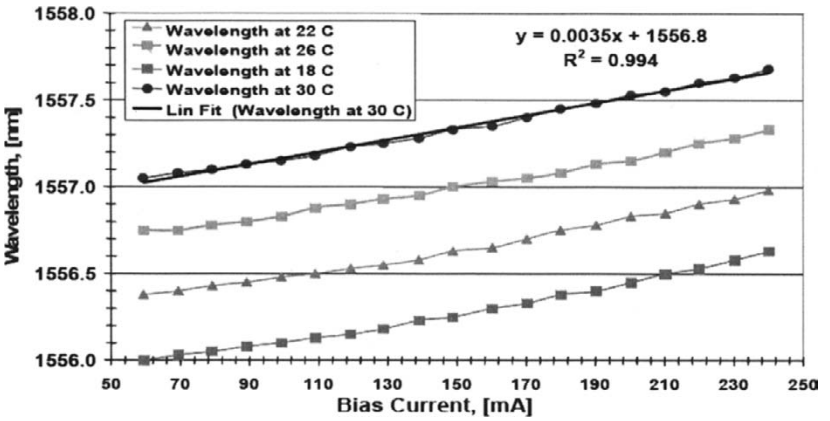


Fig. 20. Steady-state DFB laser wavelength measurements as a function of drive current for various temperature settings. The linear fits to the data yield a current dependence of $d\lambda/di_{\mathrm{dc}}$= 0.0035 nm/mA or ~440 MHz/mA and a temperature dependence of 0.14 m/°C. Note that this value is of larger magnitude and opposite sign of the fast wavelength dependence on current $d\lambda/di_{\mathrm{ac}} = -280$ MHz/mA observed in Figs. 17 and 18.

Over the long term, Telcordia-qualified COTS continuous-wave (CW) DFB lasers have demonstrated sub-GHz wavelength stability and repeatability over years of operation in the field. Precise temperature and current control combined with occasional feedback to correct for potential long term drifts in laser wavelength tuning characteristics $s\lambda(i,T)$, could enable ~100 MHz class wavelength tuning over the ~25+ year life of the laser. Such characteristics are useful since they allow for the option of open-loop TX wavelength alignment to RX filters, a capability which is particularly desirable for FSO links, especially those employing RX interferometers (e.g., DPSK—see section 5.2), where the need for spectral acquisition and tracking can significantly increase overall acquisition times.

The cost of DFB laser wavelength tunability includes two digital-to-analog converters (DACs) for adjustable current and temperature control, a feedback mechanism (e.g., a wave meter) to correct for long-term drifts, and the electrical power needed to drive the TECs to maintain a stable temperature. However, most applications require some degree of laser wavelength control, especially those employing DWDM, so some form of fixed current and temperature control is implemented anyway. The only additional cost for tunability becomes the additional DACs and the feedback mechanism. In addition to using calibrated current and temperature to control DFB laser wavelength, integrated etalon-based wavelength monitors are commercially available with Telcordia qualification [5], providing compact, reliable, locally-resident sub-GHz resolution feedback that can be used to align the laser wavelength over a 20+ year lifetime [231,265]. In some cases, the feedback information on laser wavelength (and power) can be remotely located, for instance at the receive terminal, and periodically be communicated back to the laser control via low-bandwidth back channel [266–268] (see section 5.2.2).

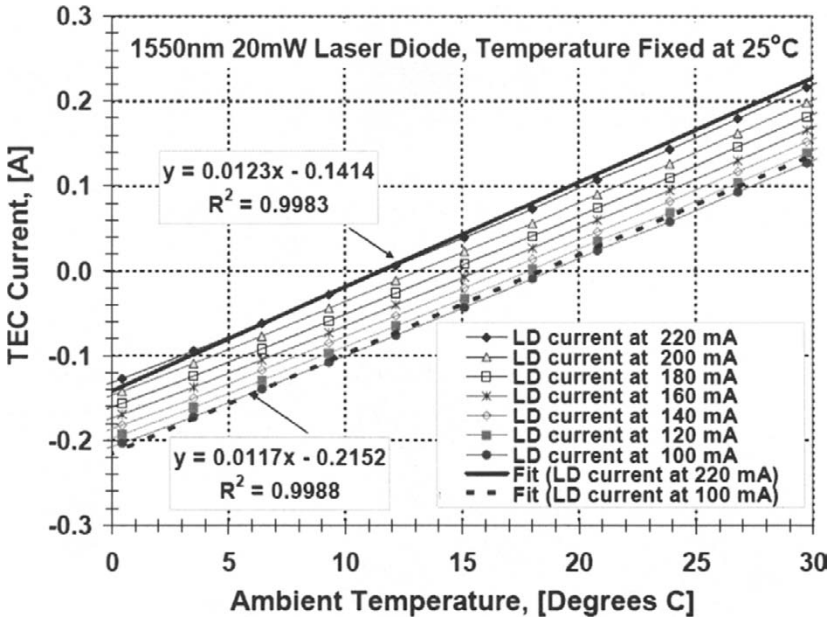


Fig. 21. TEC current versus ambient temperature for several drive currents while the laser is fixed at 25°C. The measured linear fits over the ±15°C range correspond to a TEC current slope $\frac{di_{\mathrm{TEC}}}{dT} \approx 12$ mA/°C and $\frac{di_{\mathrm{TEC}}}{di_{\mathrm{LD}}} \approx 0.4$ mA/mA. The 'zero-TEC-current' temperature is about 15°C.

As shown Fig. 21, The TEC power can be minimized if the laser is selected to output the desired wavelength near the 'zero-TEC-current' point, which is ~15°C in this case. The TEC current required to change the laser temperature, or equivalently, compensate for fluctuations in the ambient, is ~12 mA/°C. Over a range of drive currents, at 15°C ± 5°C ambient temperature, the TEC current is bounded by about 100 mA, which for a 3-V power supply limits the power requirement to < 300 mW. If power is not a driver in the design, TECs can be used to stabilize to temperature differentials up to ~40°C, although this may consume several watts of power. Note that the zero-TEC-current ambient temperature is about 10°C below the set point for the laser in Fig. 21, due to self heating from the laser drive current. This differential is typically in the range of 5°C to 10°C, dependent on the optical-electrical conversion efficiency of the laser.

3.4. Cavity-Dumped and Q-Switched Lasers

Cavity-dumped and Q-switched laser sources are capable of efficiently generating extremely high peak-power levels (e.g., tens of MW), and are especially useful for generating periodic pulse streams, used in Laser Radar (LADAR) and Laser Imaging Detection and Ranging (LIDAR) applications [269–273]. However, they are not partic-

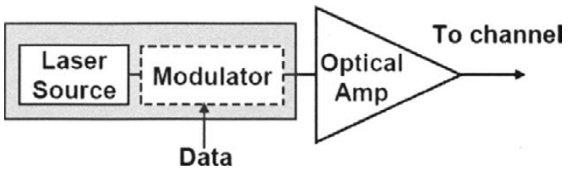


Fig. 22. Master oscillator power amplifier (MOPA) transmitter, consisting of a modulated laser source followed by a power amplifier. For high-rate applications, the MOPA is often uses a CW laser source and external modulation.

ularly well suited for high-sensitivity laser communications due to limited bandwidth and modulation options , difficulty in pulse shaping for optimized receiver sensitivity, and timing jitter [227], all of which tend to degrade communications performance. For these reasons, the MOPA transmitter discussed in the next sSection is most commonly used for high-rate-high-sensitivity communications applications.

3.5. Master Oscillator Power Amplifier (MOPA)

The Master Oscillator Power Amplifier (MOPA) transmitter shown in Fig. 22 is a flexible and scalable approach, compatible with both TDM and WDM communications at Mbit/s rates to tens of Gbit/s and beyond.

The MOPA design can be modular, allowing for a more flexible choice of waveforms, and independent design and optimization of the laser, modulator, and power amplifier. The design is commonly used for high rate optical communications in the telecom industry and is well suited for free-space lasercom platforms, where high-fidelity waveforms are required.

3.5.1. Modulation

As discussed in detail in section 3.1, direct-modulation of laser current can be used as a relatively simple method of generating a modulated optical signal at low rates. In order to transmit at high rates (e.g., 10 or 40 Gbps) or use more sophisticated modulation formats with carefully tailored pulse shapes (e.g., M-PPM, RZ-DPSK, etc.), external modulators that can generate nearly transform-limited chirp-free waveforms are generally used. Two kinds of modulators are commonly used for high-rate intensity modulation: Mach-Zehnder modulators (MZMs), which are interferometric based devices; and semiconductor-based electro-absorption modulators (EAMs) (see, e.g., [274–277]). EAMs are compact, offer ease of integration with other elements such as lasers and semiconductor optical amplifiers (SOAs), and can be cost-effective for wide-band intensity modulation transmitters. While EAMs are usually very wide band, they can introduce parasitic chirp, and often have high-loss (5–10 dB) and poor extinction at practical drive voltages. Therefore EAMs are generally not the first choice for use in high-sensitivity FSO applications.

MZMs on the other hand, are extremely versatile wide-band devices, able to modulate both intensity and phase at rates exceeding 40 Gbps. They also can be used to

generate a controlled amount of chirp depending upon the MZM design [240], which can be used to pre-compensate for the effects of dispersion [278–280]. In addition, MZMs can be used to make high-rate high-fidelity pulse-carved return-to-zero (RZ) waveforms that facilitate optimized (nearly matched) receiver design (see section 5) commonly used in series with data-modulating MZMs. For these reasons, MZMs are commonly used in high-performance fiber-optic and FSO links.

3.5.2. Mach-Zehnder Modulator (MZM)

The Mach-Zehnder interferometer structure (Fig. 23) named after its two inventors[5] [281,282] is commonly used in many optical devices including modulators, DPSK receivers, and wavelength interleavers. Active MZMs are most often fabricated in lithium niobate ($LiNO_3$) waveguides (see, e.g., [283–290]). To improve fabrication yields, they are typically constructed from y-junction couplers so that there is usually only single input (E_{in1}) and output ports (E_{out1})—with the secondary input and output ports being unguided waste ports. MZMs are also fabricated in semiconductor materials such as GaAs [143], InP [291,292], Si [293,294], and optical polymers [295]. For ultra-fast applications, nonlinear interferometric switches constructed from fibers, semiconductor optical amplifiers, and other nonlinear materials [296–302], have demonstrated 100 GHz class all-optical logic, wavelength conversion, and clock-recovery. Non-interferometric switching has also been demonstrated using self-phase modulation (SPM) and four-wave-mixing (FWM) [303–305] nonlinear processes.

The basic MZM geometry is shown in Fig. 23, composed of two directional couplers (ideally 50/50), the first splitting an incident electric field (e.g., E_{in1}) into two equal top and bottom propagating waves.

The two fields are then recombined in the second 50/50 splitter after experiencing controlled phase delays ϕ_{top} and ϕ_{bottom}. The two outputs (Eout_1 and Eout_2) are determined by the coherent addition (interference) of top and bottom fields incident on BS_2. The interference is a function of their amplitudes and relative phase ($\Delta\phi = \phi_{\text{top}} - \phi_{\text{bottom}}$). Chirp-free intensity modulation is achieved for balanced field amplitudes when the MZM operates in push-pull mode, with ϕ_{top} and ϕ_{bottom} having equal magnitude and opposite sign [240], which removes all residual phase modulation.

The MZM time-domain field impulse responses are given in Eqs. (57) and (58) for the two MZM outputs,

$$\begin{aligned} E\text{out}_1(t) &= \sqrt{R_1R_2}\delta(t+\tau/2) + \sqrt{T_1T_2}\delta(t-\tau/2) \\ &= \frac{1}{2}\left[\delta(t+\tau/2) + \delta(t-\tau/2)\right]_{\text{ideal}}, \end{aligned} \tag{57}$$

$$\begin{aligned} E\text{out}_2(t) &= \sqrt{R_1T_2}\delta(t+\tau/2) - \sqrt{T_1R_2}\delta(t-\tau/2) \\ &= \frac{1}{2}\left[\delta(t+\tau/2) - \delta(t-\tau/2)\right]_{\text{ideal}}, \end{aligned} \tag{58}$$

where factors of constant phase have been dropped, and the ideal impulse responses occur when BS_1 and BS_2 are 50/50 splitters, i.e., $R_1R_2 = T_1T_2 = 1/4$ and $R_1T_2 = T_1R_2 = 1/4$. These correspond to the sum and difference of two delta functions

[5] L. Mach and L. Zehnder independently invented in 1891/1892 what has become known as the Mach–Zehnder interferometer which could monitor changes in refractive index, and hence density, in compressible gas flows

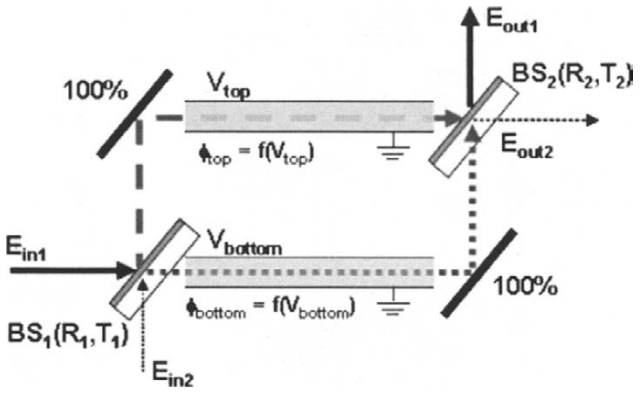


Fig. 23. The basic structure for a Mach-Zehnder modulator consists of two beam splitters $BS_{1,2}$ with intensity reflection and transmission coefficients $R_{1,2}$ and $T_{1,2}$, respectively. Ideally, the $BS_{1,2}$ are 50/50 splitters, i.e., $R_{1,2} = T_{1,2}$= 50%. The phase difference between the two arms: $\Delta\phi = \phi_{\text{top}} - \phi_{\text{bottom}}$, determines the output splitting ratio based on the constructive or destructive interference of the top and bottom arms at BS_2. For electro-optic MZMs, the phase change is proportional to the applied drive voltages V_{top} and V_{bottom}.

separated by time difference τ. In the frequency domain, the Mach-Zehnder outputs are periodic sinusoids given in Eqs. (59) and (60):

$$E\text{out}_1(f) = \left(\sqrt{R_1R_2}e^{j\pi f\tau} + \sqrt{T_1T_2}e^{-j\pi f\tau}\right) = \cos(\pi f\tau)_{\text{ideal}}, \quad (59)$$

$$E\text{out}_2(f) = \left(\sqrt{R_1T_2}e^{j\pi f\tau} - \sqrt{T_1R_2}e^{-j\pi f\tau}\right) = \sin(\pi f\tau)_{\text{ideal}}. \quad (60)$$

The sinusoidal MZM field and intensity transfer functions for $E\text{out}_1$(f) and $I\text{out}_1(f) = |E\text{out}_1(f)|^2$ are shown in Fig. 24, and highlight many attractive features. The voltage required to generate the π phase shift is referred to as the switching voltage, V_π, and corresponds to the voltage required to go from an intensity transmission peak to a null. By biasing the modulator at the 50% transmission point (referred to as the quadrature point), the MZM can generate intensity modulation (IM) waveforms when driven with a differential voltage $\Delta V = \pm V_\pi/2$, as illustrated for the superimposed OOK and PPM drive waveforms.

While it is not a requirement to operate the modulator from transmission null all the way to the peak as shown, doing so maximizes the optical transmission and reduces the impact of waveform distortion such as transient ringing on the input electrical signal. This is due to the nonlinear $\cos^2()$ transmission characteristics which tend to compress fluctuations at the transmission peak and null, and steepen transitions in between.

MZMs can also be used for digital phase modulation. By driving the MZM around the transmission null, the sign of the electric field output transitions from positive to negative, which corresponds to an abrupt π phase change suitable for generating DPSK and DQPSK waveforms. These digital characteristics offer advantages over

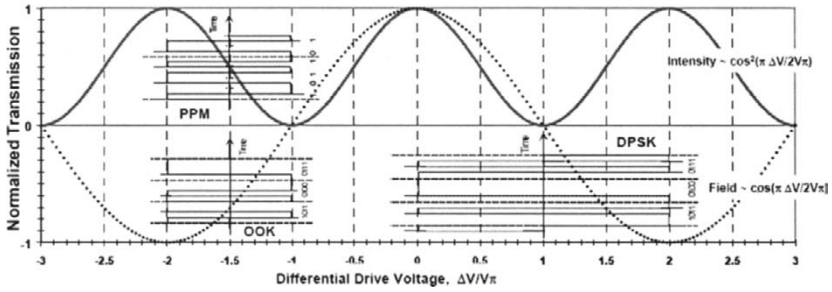


Fig. 24. MZM transfer function for the cos() arm versus differential drive voltage normalized to $V\pi$, the switching voltage. Superimposed are digital modulation drive waveforms (drive voltage versus time) for OOK, PPM, and DPSK formats, which are achieved by simply varying the drive bias position and amplitude.

linear phase modulation in which energy is transmitted during the phase transition (as the modulator slews between 0 and πp, which degrades the SNR. In order to extract maximum phase-modulated output power, the MZM is biased at a transmission null and is driven between two successive transmission peaks, requiring a drive voltage of $\pm V_\pi$, twice that of IM, corresponding to four times the drive power. As in the IM case, digital phase modulation can be achieved using a small drive voltage around the null to conserve power, at the expense of increased insertion loss and waveform fidelity that is sensitive to the input drive waveform. Furthermore, for non-ideal MZMs, this can also lead to increased chirp penalties for differential phase-modulation formats.

3.5.2.1. MZM Phase Elements

The time delay τ in Eqs. (57)–(60), has components of distinct origin

$$\tau = \tau_{\text{static}} + \tau_{\text{random}} + \tau_{\text{drive}}, \tag{61}$$

which map to optical phase, $\phi(\lambda)$ given by

$$\phi(\lambda) = \phi_{\text{static}} + \phi_{\text{random}} + \phi_{\text{drive}} = \tau\left(\frac{2\pi c}{\lambda n}\right) \quad \text{[radians]}. \tag{62}$$

The static delay, ϕ_{static}, corresponds to a fixed difference between the two arms due to manufacturing, and is typically many optical periods and much larger than the other components. Thus, τ_{static} determines the interferometer periodicity commonly referred to as the free-spectral range (FSR), where

$$\text{FSR} = 1/\tau_{\text{static}}. \tag{63}$$

Modulation bandwidth is limited to a small fraction of the FSR, and therefore the design target for τ_{static} is typically near zero, which enables a wide range of operation. For example, if $\tau_{\text{static}} \ll 8$ ps, the modulation bandwidth is $>\sim$125 GHz, or 1nm at

1550nm wavelengths. Note that this time is long relative to the optical period, corresponding to many hundreds of cycles of the ~5 fs carrier. For a modulator to operate simultaneously over the entire EDFA C-band (i.e., 1535–1565nm), $\tau_{\text{static}} \ll 0.2$ ps, which can be achieved in practice with precision manufacturing. This is important for wide-band WDM applications in which a single modulator may be used to modulate and/or shape (pulse carve) many channels (see, e.g., [93,138]).

3.5.2.2. MZM Bias Control

The MZM random phase term, ϕ_{random}, represents a slow drift ($>\sim$1 s to $>$ 1 day time-constant depending on the mechanism) in the relative phase between the two modulator arms. This is a well studied phenomenon that can be attributed to mechanisms such as thermal gradients [306] and charge buildup [307]. Consequently, closed-loop drive bias control is generally required to maintain the proper bias position while compensating for modulator drift [180,308]. Advances in MZM designs have significantly reduced the amount and rate of drift [309–313], but even small drifts can degrade extinction ratio which is important for low-duty-cycle applications [2,11,15,21,314] and lead to deleterious chirp in DPSK applications [132,142]. As long as the drift rate remains slower that the closed-loop control bandwidth, and is contained (within a few π), the random phase can be compensated indefinitely with closed-loop control (see, e.g., [315–317]) which applies a voltage typically up to a few V_π to counter the drift over a range of several interferometer fringes. If the random-phase drift continues to move in one direction beyond the range of the control loop, a reset is required, in which the control loop jumps back to the center of its operating range and re-optimizes the bias position. Such resets can be completed within ~1 ms, depending on the speed of the control system, during which the modulator is uncontrolled. With present modulator designs, such resets are becoming less frequent, and as long as the average reset rate is $>$ 1 hour this corresponds to better than 99.9999% availability. Present modulator designs have estimated reset-free operation times exceeding 20 years [311,313,318]. Further, since such resets can be anticipated as the bias slowly approaches a rail, steps can be taken to mitigate the effects of infrequent reset outages.

3.5.2.3. Extinction Ratio (ER)

Modulator extinction ratio (ER), defined as the power off-to-on ratio, impacts communication performance in many ways. It determines the initial distance between logical '0' and '1' power levels in OOK modulation and leads to multiple interferometric cross talk [319], which adversely impacts both TDM [320] and WDM systems. Incomplete ER can significantly degrade the sensitivity in optically preamplified links [15,20,314,321,322]. For ideal optically preamplified OOK, a slight modification of the upper-bound estimate for the SNR penalty given in [20], yields

$$\Delta\gamma = \frac{1 - 2\text{ER}}{1 + 2\text{ER} + 2\sqrt{\text{ER}}}\left(1 + \frac{2}{3}\sqrt{\text{ER}}\right), \tag{64}$$

which is accurate to about $\pm$ 0.2 dB for all SNR levels. This corresponds to a ~0.7 dB penalty for a 20-dB ER (i.e., ER = 0.01) and ~2.8 dB for a 10-dB ER. In addition, poor ER limits the dynamic range of data-rates for variable duty-cycle multi-rate optical communications and transmitter power efficiency low-duty-cycle modulation formats

such as M-PPM [1,10,11,15,21,22,198,214,314], due to "power robbing" from the signal at the transmitter output, (discussed further below and in section 3.5.3). Such low-duty-cycle modulation formats are particularly attractive for use in high-sensitivity FSO applications.

In order to have complete extinction, the fields interfering at BS_2 must have equal amplitudes, same polarization and π phase difference. For this reason, MZMs with high extinction typically have polarizing elements which force the interfering waves to be co-polarized. For non-ideal interferometer splitting ratios, the ER is given by

$$\mathrm{ER} = \left(\frac{\rho - 1}{2(\rho + 1)}\right)^2, \tag{65}$$

where $\rho = R_1R_2/(T_1T_2)$ for $E\mathrm{out}_1$ and $\rho = R_1T_2/(T_1R_2)$ for $E\mathrm{out}_2$ in Eqs. (57)–(60). For perfect extinction ρ = 1. A 20-dB ER results from ρ = 2/3, which can be achieved with BS_1 = 50/50 and BS_2 = 60/40 coupling ratios. For BS_2 = 47/53, a 30-dB ER can be achieved.

Since common telecom specifications for ER requirements are only ~13 dB, there are few vendors that market good-ER (ER < –30 dB) MZMs, despite the fact that such devices should be relatively easy to produce given the relations above. Nevertheless, commercially-available MZMs typically have ER better than –20 dB (at frequency), and MZMs with internal high-extinction polarizers have demonstrated modulation ER < –45 dB individually and ER < –60 dB in cascaded operation [15]. By actively balancing the input powers to the MZM output coupler, a commercial MZM was configured with ER = ~–70 dB using a single modulator [323]. Note that for ultra-wideband applications (e.g., >~30 nm), modulation bandwidth and ER may also be limited by coupler/waveguide wavelength dependence.

For low-duty-cycle waveforms, windowed direct modulation may be used as a power efficient technique for achieving high modulation extinction, while maintaining stable wavelength and chirp-free waveforms during transmission as illustrated in Fig. 25. In this manner, the power required for driving the laser and external modulator can potentially be reduced by a factor approaching the duty cycle. Such windowing could be performed in both the frequency-domain via narrow-band optical filtering and the time-domain, which could, for example, leverage a single pulse-carving modulator to efficiently time-window many WDM channels simultaneously.

3.5.2.4. Extinction Ratio Characterization and Optimization

While measurement of modulator ER is relatively straightforward for CW signals, measurement of ER for high-frequency waveforms can be challenging. Owing to the frequency dependence of transmitter components, the ER can vary as a function of the bit-rate and data pattern, so CW measurements of ER are often insufficient to accurately characterize the performance of wide-band modulators. Consequently, high-speed detection equipment is generally required to directly measure the ER for wide-band waveforms using time-domain techniqus [324,325]. But these methods can have limited accuracy and range of measurement, typically <~20 dB, or impose symmetry requirements on the waveforms being measured. Therefore they are not well suited for measuring RZ, low-duty-cycle or high extinction ratio waveforms.

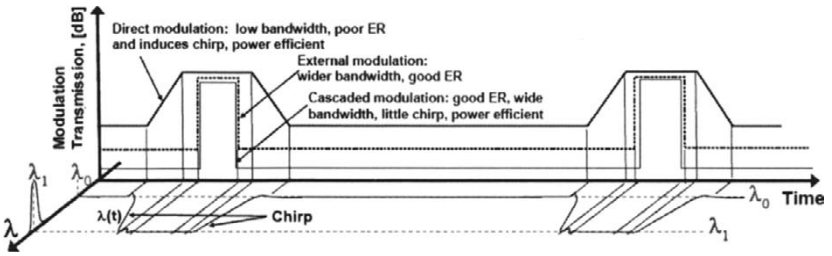


Fig. 25. Illustration of cascaded direct and external modulation waveforms showing both intensity and wavelength changes in time. The external modulator can improve extinction, waveform and spectral fidelity via a windowing function, but this need not be square as depicted. Other functions such as a variety of Gaussian-like RZ waveforms (discussed at the end of section 3.5.2) can be efficiently implemented without transmitter power penalty as long as it is followed by an APL amplifier. In a similar manner, spectral filtering can be used to improve ER, spectral fidelity, and reduce power requirements, by, for example, using a passive narrow band filter to pass the desired signal λ_1 and reject the unwanted λ_2.

However, with knowledge of the duty cycle[6] (DC), e.g., the ratio of *ones* to *zeros* transmitted, and the ability to accurately measure the average power ($\bar{P}$) and peak power ($P_{\rm pk}$), modulator ER performance can be characterized with the use of Eq. (66) [11,314]:

$$\mathrm{ER} = \frac{\bar{P}/P_{\rm pk} - \mathrm{DC}}{1 - \mathrm{DC}} \underset{\lim \mathrm{DC}\to 0}{\Rightarrow} \frac{\bar{P}}{P_{\rm pk}}. \tag{66}$$

For average-power-limited measurements, such as of a saturated MOPA TX output, the peak power is proportional to the average power and grows as the duty-cycle is reduced, as long as ER $\ll$ DC. This follows by solving Eq. (66) for $P_{\rm pk}$ which yields

$$P_{\rm pk} = \frac{\bar{P}}{\mathrm{DC} + \mathrm{ER}(1 - \mathrm{DC})} \underset{\lim \mathrm{DC}\to 0}{\Rightarrow} \frac{\bar{P}}{\mathrm{ER}}. \tag{67}$$

A plot of $P_{\rm pk}$ versus duty cycle for various ER is shown in Fig. 26. As the duty cycle approaches the ER, the peak power is reduced by half as the denominator in Eq. (67) approaches 2. For very low duty cycles, (i.e., DC $\ll$ ER) the normalized peak output power assymtotically approaches the ER. By fitting normalized peak power measurements to the calculated curves, the ER can be determined.

For peak-power-limited measurements, such as for the modulator output, the peak power is fixed and equal to the average power for a 100% duty cycle. In this case,

$$\mathrm{ER} = \frac{\bar{P}_{\rm norm} - \mathrm{DC}}{1 - \mathrm{DC}} \underset{\lim \mathrm{DC}\to 0}{\Rightarrow} \bar{P}_{\rm norm}, \tag{68}$$

where $\bar{P}_{\rm norm} = \bar{P}(\mathrm{DC})/\bar{P}(\mathrm{DC}) = 100\%$), and the measurement of ER can be achieved entirely with CW measurements as DC gets small relative to ER, without

[6] The duty cycle of a periodic or re-occurring intensity waveform, $s_I(t)$ is the percentage of '*on*' time during the period T, i.e., DC= $\frac{1}{T}\int_0^T (s_I(t)/\max[s_I(t)])dt$.

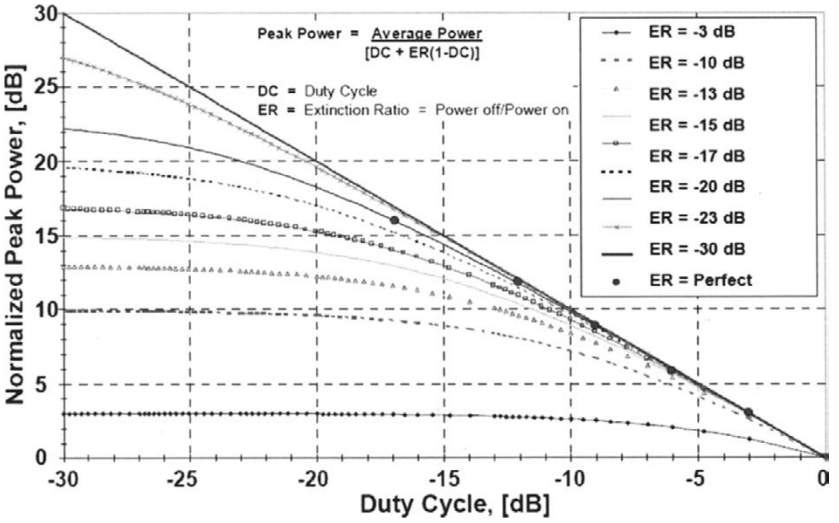


Fig. 26. Peak output power for an average power limited transmitter versus duty cycle for several values of modulator extinction [314]. Reprinted with permission. (©2000 IEEE.)

the need for high-speed measurement capabilities. Both the peak and average-power-limited measurement approaches described above have been used to characterize modulator ER with good agreement, accurately characterizing modulators with extinction > 60 dB [15,214].

Modulator control and ER optimization is often accomplished by minimizing the power-off transmission directly after the modulator [308,326]. However, for low duty-cycle or variable-rate average-power limited (APL) systems [11], another potentially more sensitive method of minimizing the ER is to maximize P_{pk} given in Eq. (67) at the output of the MOPA TX, since this value becomes increasingly sensitive to ER as the duty-cycle is reduced [314]. The peak power is maximized when the ER is minimized, and this effect is exaggerated significantly at low duty-cycles, by an amount that is approximately inversely proportional to the duty-cycle. Thus, this approach serves to optimize ER performance by maximizing the peak signal power delivered from the TX.

3.5.2.5. MZM Drive Power and Chirp Considerations

As shown in Table 5, typical values of V_π range from 3 to 5 V. Since operating power is often a primary consideration in many FSO applications, lower drive voltage ($V_\pi < 2$ V) is desirable and provides compatibility with COTS high-frequency low-voltage drive electronics. For a given electro-optic coefficient, V_π can be lowered by maximizing device interaction length (d), or increasing electric field (E) across the waveguide. Since $E = V/d$, decreasing d allows the voltage to be lowered while maintaining the same electric field. Both of these approaches tend to increase insertion loss and capacitance, which in turn reduces device bandwidth. Traveling wave designs can be

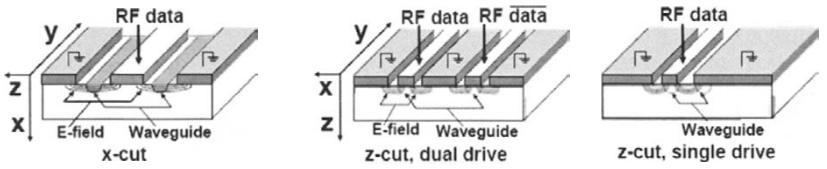


Fig. 27. Common $LiNb0_3$ modulator electrode configuration and crystal orientation with electric-field lines: (left) x-cut, (center) dual-drive z-cut, (right) single-drive z-cut.

used to mitigate the bandwidth issue [327–329], but RF-optical velocity matching constraints make it challenging to obtain a uniform response over the wide fractional bandwidth (e.g., 40 KHz to 40 GHz—6 orders of magnitude) desired in broad-band digital applications. This results in a frequency dependent $V_\pi(f)$, which can lead to pattern dependence in the transmitted waveforms.

There are two crystal orientations commonly used to make MZMs in $LiNbO_3$, x-cut and z-cut [283,330], shown in Fig. 27. X-cut designs have a single electrical modulation drive input that is split within the modulator to provide the push-pull action required for switching. If designed properly, the drive voltages arrive on each arm of the interferometer simultaneously and impart the same phase shift (for all frequencies of interest), resulting in chirp-free modulation. However, since there is only a single electrical input, there is no means of compensating with the drive waveform if there are fabrication errors.

Z-cut modulators come in dual- and single-drive configurations. Dual-drive z-cut MZMs can be used to adjust chirp by controlling the instantaneous imbalance of the input drive voltages (V_{top} and V_{bottom}). The resulting chirp parameter defined in Eq. (3) is given by [278]

$$\alpha = \frac{V_{\text{top}} + V_{\text{bottom}}}{V_{\text{top}} - V_{\text{bottom}}}. \tag{69}$$

The ability to adjust the chirp can be used, for example, to compensate for modulator fabrication errors or optimize transmission through a dispersive and nonlinear fiber-optic link [278–280,331,332], dynamically if necessary.

For the same V_π, dual-drive z-cut modulators require half the drive power of x-cut modulators, since each of the two drivers needs only $V_\pi/2$, which corresponds to a net power of $V_\pi^2/(2R_\Omega)$ for the z-cut, versus $V_\pi^2/2$ for the x-cut, where R_Ω is the modulator impedance. The dual-drive design also distributes the power load over a larger area, reducing the heat density which can be advantageous in the carefully controlled thermal environment of space-based platforms.

For single-drive z-cut MZMs, (Fig. 27-right), the electric field is stronger in the optical waveguide centered under the RF data electrode, with only $\sim$15% of the field driving the secondary waveguide. This fixed asymmetry leads to a phase-imbalance and residual chirp, with $\alpha = -0.7$. A summary of the above characteristics and representative (minimum) drive powers is given in Table 5.

Table 5. Summary of $LiNbO_3$ modulator drive and chirp characteristics for common crystal orientations and drive configurations.

	x-cut	z-cut single-drive	z-cut dual-drive
Typical V_π	5 V	4 V	4 V
Drive Power R_Ω = 50 ohms	V_π^2/R_Ω = 0.5 W	V_π^2/R_Ω =0.32W	$V_\pi^2/(2R_\Omega)$ = 0.16W
Chirp (α)	~0 (fixed by design)	–0.7 (fixed by design)	$-\infty < \alpha < \infty$ (in principle) $-2 < \alpha < 2$ (used in practice)
Layout size	smaller	smaller	larger
Heat density	higher	higher	lower

3.5.2.6. Pulsed Waveform Generation

Pulse carving using a MZM is a subject that has been discussed extensively in the literature [10,13,14,25,41,132,138,140,141,159,333–336] as a means of generating high-fidelity waveforms that are well suited for long-haul high-rate fiber and free-space optical links. Sinusoidal-drive techniques have also been used for ultra-high-speed serial-parallel conversion in optical time-division multiplexed (OTDM) systems [337] and photonic analog-to-digital converters [338]. Pulse carving reduces the influence of imperfections in the drive electronics, such as transient ringing, pattern dependence and inadequate bandwidth, on waveform quality. Broad-band high-fidelity electronics are not essential, since a variety of shaped-RZ waveforms with desirable characteristics can be generated by simple narrow-band sinusoidal waveforms with appropriate bias and drive amplitude and frequency [132,336]. Of particular interest are Gaussian-like waveforms that can be generated in this manner that can facilitate robust nearly-matched optical communication performance [13,14] a subject discussed further in section 5. A general expression for the MZM-pulse-carved power is given by

$$P(t) \propto |E(t)|^2 = \cos^2\left(\frac{V_a}{V_\pi}\frac{\pi}{2}\sin(2\pi f t) + \frac{\pi}{2}\frac{V_b}{V_\pi}\right), \tag{70}$$

where V_π is the switching voltage, V_a and f are the sinusoidal drive amplitude and frequency, respectively, and V_b is the bias voltage, which centers the drive input at locations summarized in Table 6. Note that the static and random phase offsets of Eq. (11) are not shown explicitly in Eq. (19) since it is assumed that these offsets are compensated for by bias-control.

Table 6. Summary of MZM bias points and corresponding bias voltages.

Bias Point	V_b/V_π	
Peak	m,	$[m \in \text{even}]$
Quadrature	$\frac{1}{2}m$,	$[m \in \text{odd}]$
Offset	$m \pm 1/4$,	$[m \in \text{odd}]$
Null	m,	$[m \in \text{odd}]$

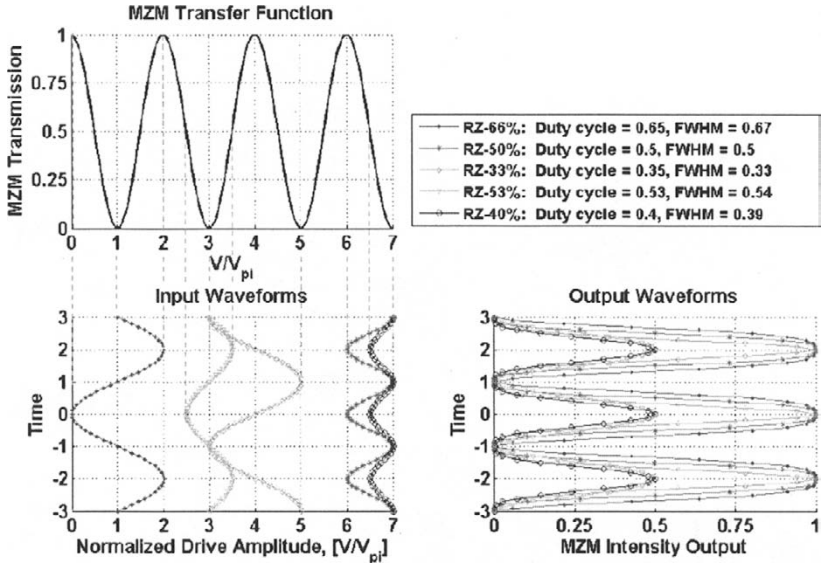


Fig. 28. Pulse carving waveform generation: (top) MZM Transfer Function, (bottom left) sinusoidal input waveforms, and (bottom right) output waveforms for RZ-66% (CSRZ) in red, RZ-50% in pink stars, RZ-33% (~Gaussian) in green diamonds, RZ-53% (Raised Cosine) in light blue triangles, and RZ-40% (~Gaussian) in blue circles, respectively. Note the differences in drive amplitude, frequency, and bias position.

Depending on the drive amplitude and bias position, pulsed-waveforms of duty-cycles ranging from 33% to 66% can be generated as shown in Figs. 28 and 29. A summary of waveform spectral and temporal characteristics and drive sensitivities are given in the pulse carving trade space matrix of Table 7.

The RZ-33% (Gaussian), RZ-66% (carrier-suppressed), and RZ-53% (raised cosine) waveforms are biased at either a transmission peak or null, and output a pulse stream at twice the drive frequency. However, when used with DPSK or DQPSK modulation, these waveforms require accurate drive amplitude or phase balance across the two arms of the MZM in order to avoid significant chirp-induced penalties [132,142].

In contrast, the 50%-RZ is attractive for use with DPSK for several reasons. The drive voltage is reduced by a factor of two, which widens the availability of driver amplifiers and reduces the power requirements. Also, the sensitivity to drive amplitude (V_a) and bias position (V_b) is reduced, making it easier to achieve stable bias control and long-term waveform fidelity, as well as diminished chirp penalty [132,142], since the MZM carver does not traverse the transmission null, the point at which the output phase is inverted. These characteristics in turn, make it easier to generate desired transmitter waveforms and optimize receiver sensitivity [133].

A representative pulse-carving MOPA TX is shown in Fig. 30. The data source can include an FEC encoder, which outputs the encoded data and corresponding channel clock at a rate equal to the data clock plus the coding overhead. The data formatter is

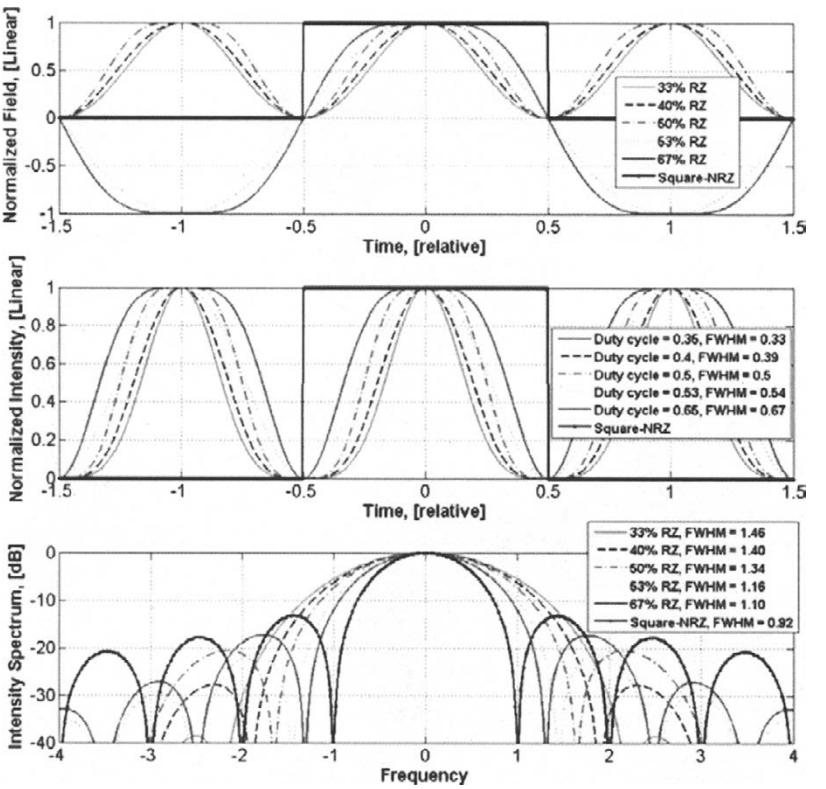


Fig. 29. Pulsed RZ waveform generation using periodically driven Mach-Zehnder modulator. (Top) Electric field; (middle) power, (bottom) power spectrum.

not needed for OOK but is needed to precode other modulation formats, such as DPSK or PPM.

3.5.3. High Power Optical Amplifier

Efficient high-power optical transmitters are necessary in optical communications scenarios such as broadband distribution networks where electrical power is increasingly becoming a valuable commodity. Efficiency also benefits large-scale WDM transmission where TX power is distributed over many channels (λs), and FSO links, where transmitter power is often a limiting design driver (see e.g., [2,21,339]). High power optical amplifier needs have historically been met by erbium-doped fiber amplifiers (EDFAs) operating near 1550nm [126,340–342]. When compared with RF amplifiers, EDFAs have many desirable characteristics including high gain (> 40 dB), large average output power (> 10 W), and ultra-wide bandwidth > 10 THz. However, as noted in Table 3 and discussed later in this section, for high-duty-cycle waveforms

Table 7. Pulse carving characteristics and trade space matrix highlighting drive frequency and drive amplitude (V_a) requirements, bias position, waveform duty cycle (DC_{pulse}), time-bandwidth-product, and sensitivities. *Adapted from Stevens in* [142].

#	Waveform Type	Time-Bandwidth Product			Waveform Characteristics			Drive Characteristics						
		3 dB $\Delta t_i * \Delta f_i$	10 dB $\Delta t_i * \Delta f_i$	20 dB $\Delta t_i * \Delta f_i$	FWHM	Duty Cycle	Carver Loss, [dB]	Drive Type	Drive Freq., [fRate]	Drive Amp. (V/Vπ)	Rel. Drive Power	Bias Point	Sensitivities Bias	Drive
1	NRZ Square	0.89	1.49	3.6	100%	100%	0	Sq.	na	0.5	8	Quad	Moderate	Moderate
2	RZ-Square (50%)	0.89	1.49	3.6	50%	50%	-3	Sq.	na	0.5	4	Quad	Moderate	Moderate
3	RZ-33% (Gaussian)	0.49	1.54	2.7	33%	35%	-4.8	Sin	0.5	1	16	Peak	High	High
4	RZ-66% (CSRZ)	0.73	1.62	2.9	67%	65%	-1.7	Sin	0.5	1	16	Null	High	High
5	RZ-53% (Raised Cosine)	0.63	1.62	2.4	54%	53%	-5.7	Sin	0.5	0.5	4	Null	High	Low
6	RZ-50%	0.66	1.62	4.0	50%	50%	-3	Sin	1	0.5	4	Quad	Moderate	Moderate
7	RZ-40% (Gaussian)	0.55	1.57	2.6	39%	40%	-7.1	Sin	1	0.25	1	Offset	Very High	Low

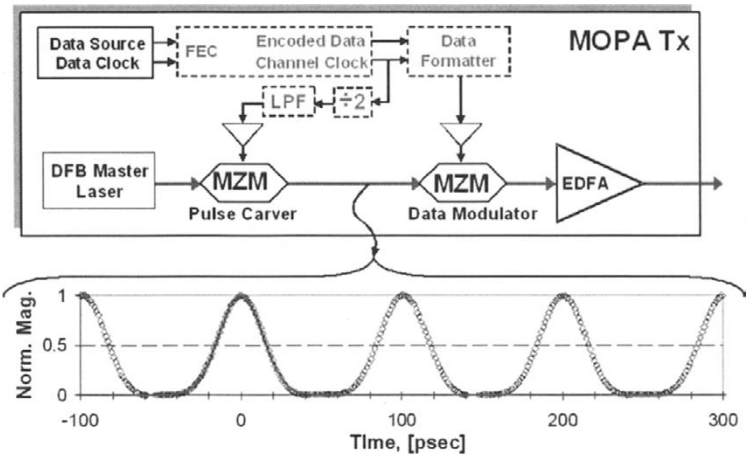


Fig. 30. (Top) A pulse-carving MOPA TX with optional FEC and Data Formatting elements. 33%-RZ pulse carving requires sinusoidal drive at $\frac{1}{2}$ the channel clock rate. (Bottom) Measured 10 Gbps 33%-RZ waveforms nearly indistinguishable from a superimposed 35-ps full-width-half-max (FWHM) Gaussian fit centered at $t = 0$.

optical amplifiers are typically less power efficient than their RF counterparts, but for low-duty-cycle waveforms the efficiencies are more comparable.

The ability to efficiently transmit a variety of signaling waveforms at high average and peak output power levels in a single spatial and polarization mode is of particular importance to free space applications. Single spatial mode operation is nec-

essary for efficient power delivery to a far-field FSO target. In fiber-optic networks, a single-polarization TX simply reduces the amount of unnecessary (orthogonally polarized) amplified spontaneous emission (ASE) noise entering the network. In FSO applications, single-polarization transmitters can improve TX-RX isolation and allow for simple polarization diplexing of TX/RX signals within an FSO terminal [46], without the need for active polarization tracking. Furthermore, single-polarization receivers can offer improved receiver sensitivity [20,41,77] since orthogonally polarized amplified spontaneous emission (ASE) noise can be eliminated at both RX and TX.

The ability to maintain transmission efficiency over a wide range of duty cycles, which is facilitated by high-gain and average-power limited properties of optical amplifiers [11,343–346], allows for aggressive pulse shaping at the transmitter. This can improve receiver sensitivity and provide flexible multi-rate capabilities with simplified receiver design options [11,12,15], a subject discussed below and further in section 5.

3.5.3.1. Average Power Limited (APL) Properties

EDFAs are average-power limited (APL) amplifiers, meaning the amplifier gain and average output power are fixed for a given average input power, independent of the shape or duty cycle of the incoming signal. This holds whenever the power-off time of the incoming signal is much less than upper state life time of the gain-medium [172]. For EDFAs, which have a slow $\sim$1 ms time constant, this is true for practical optical communications rates above $\sim$1 Mbps. Detailed experimental analysis and simulations of amplifier transients and average power limited amplifier dynamics are given in [126,340,347].

For APL amplifiers, the peak to average power ratio varies inversely with the duty cycle - as it decreases, the peak power grows while average power is maintained (see Figs. 35 and 36). This is in sharp contrast with RF amplifiers, which are generally peak-power limited (PPL), meaning there is a maximum (peak) voltage or output power they can deliver. In order to maximize the average power transmitted, which is important for power-starved communication links, a PPL TX must operate with nearly a 100% duty-cycle—a constraint that does not apply to APL optical transmitters.

3.5.3.2. Amplifier Gain, Saturation, and Noise

The functional dependence of amplified signal output power (P_{out}) and gain G as a function of average signal input power (P_{in}) is given by

$$P_{out} = P_{in}G(P_{in}) = P_{in}\left(1 + \frac{g_o}{1 + g_o\,(P_{in}/P_{sat})}\right) \underset{P_{in}g_o \gg P_{sat}}{\Rightarrow} P_{in} + P_{sat}, \tag{71}$$

where g_o is the small signal gain at the signal wavelength, and P_{sat} is the average output power when the amplifier is driven deep into saturation[7]. Note that the amplifier gain can have a strong wavelength dependence [126, 340], but for this analysis, the gain is assumed to be relatively constant over the narrow-band surrounding the signal. The dependence of amplifier gain, average output power, and noise figure on average input power are illustrated in Fig. 31.

[7] Note, that we refer to P_{sat} as the maximum useful amplified output power, which is different from the term P_{sat}^{out}, which is conventionally defined as the output power for which the EDFA gain has dropped by 3 dB from the small signal gain G_o [126].

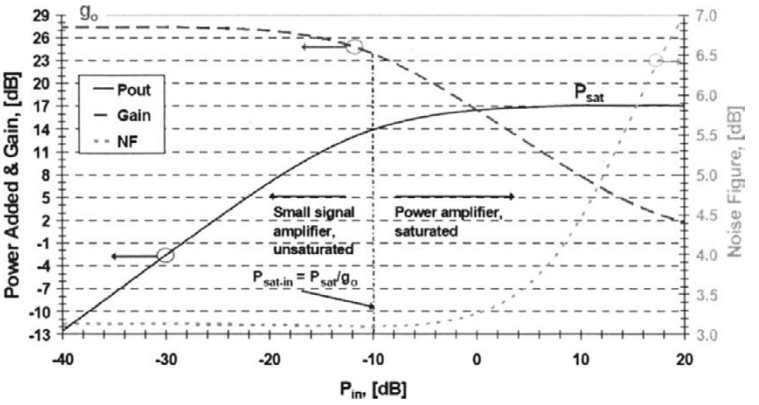


Fig. 31. Typical EDFA average power added *(blue, solid line)* and gain *(red, dashed line)* both on the left-hand axis, and noise figure *(green, dotted line)* on the right-hand axis versus average input power ($P_{\rm in}$). For this example, g_o = 27 dB and $P_{\rm sat}$ = 17 dBm, yielding $P_{\rm sat-in}$ = –10 dBm.

The 3-dB gain compression point occurs when $P_{\rm in} = P_{\rm sat}/g_o$, which is referred to as the saturated input power ($P_{\rm sat-in}$). To efficiently extract maximum output power, power amplifiers are typically run with $P_{\rm in} \gg P_{\rm sat-int}$. In this regime, the gain $G(P_{\rm in})$ is significantly compressed and the amplifier operates in a 'power added' mode, contributing most of Psat to the total output power. For $P_{\rm in} > P_{\rm sat-in}$ + 10dB, over 90% of the available power is extracted corresponding to an efficiency penalty of ~0.5 dB. This penalty drops to < 0.1 dB for $P_{\rm in} > P_{\rm sat-in}$ + 15dB.

Note that as the input power exceeds $P_{\rm sat-in}$, and the amplifier gain is reduced, the noise figure increases as a result of incomplete population inversion in the amplifying medium [126]. However this generally is of little consequence since the SNR at the amplifier input is typically large enough such that even a several-dB SNR degradation is not sufficient to impact communications performance. This holds as long as

$$GN_{\rm sp}\eta_{\rm ch} \ll 1, \tag{72}$$

and

$$\mathrm{SNR_{TX}} \cong \frac{N_{\rm sig-in}}{2N_{\rm sp}} > N_{Rx-\rm min} \gg 1, \tag{73}$$

where $\eta_{\rm ch}$ is the net channel loss, G is gain, $\mathrm{SNR_{TX}}$ is the transmitter output SNR[8], $N_{\rm sig-in}$ is the input signal PPB, $N_{Rx-\rm min}$ is the minimum required RX SNR[9], and $N_{\rm sp}$ is the spontaneous emission factor, equivalent to the number of noise-photons-per-mode at the input of the amplifier. For high-gain and a large signal input,

[8] This definition of $\mathrm{SNR_{TX}}$ assumes that $N_{\rm sig-in}$ is large enough that out-of-band ASE × ASE noise terms can be neglected. While this will be accurate for narrow-band or matched RX filtering, it will not be accurate for unfiltered-broad-band RXs.

[9] Note that for shot-noise-limited direct-detected signals, SNR = PPB so SNR and PPB can be used interchangeably. This is not the case for optically amplified signals.

$$N_{sp} = \frac{NF}{2}, \tag{74}$$

where NF is the linear noise figure. For a fully inverted amplifier *NF* approaches the quantum limit of 2 (3 dB) [126], corresponding to $N_{sp} = 1$. When the condition of Eq. (72) is satisfied, the received TX ASE noise is small relative to the shot noise of the signal and the minimum ASE noise added by an optically preamplified RX. If Eq. (72) is not satisfied, communication performance then becomes dominated by the received TX ASE [348], which limits SNR_{TX}. However, in this case, the conditions of Eq. (73) are generally easy to obtain, and as long as the $SNR_{TX} > N_{RX-min}$, the impact on communication performance is negligible.

Consider, for example, a TX amplifier operating at 10 Gbps with $P_{in} = -10$ dBm, 27 dB gain, 10 dB noise figure, and $P_{out} = 17$ dBm. The shot-noise-limited input SNR during the 100 psec bit period is ~49 dB and the corresponding $SNR_{TX} = N_{sig}/(2N_{sp})$ = 39 dB. Recall that the quantum limit receiver sensitivity (at 10^{-9} BER) for an uncoded optically preamplified OOK RX is ~40 PPB (16 dB SNR). Thus, in the regime where the SNR_{TX} dominates, this link would have 23 dB margin. The large SNR_{TX} indicates that the noise added by the amplifier is still small relative to the signal, corresponding to ~2500 ASE PPB out of 3.9×10^7 PPB transmitted in the single-spatial-temporal-polarization mode of the signal. Subsequently passing both the signal and ASE photons through a 60-dB attenuating channel results in a –43 dBm received signal that is nearly shot-noise limited, comprised of 38.9 signal PPB and 0.003 ASE PPB, with an SNR of 15.9 dB. For an optically preamplified RX, as long as the received ASE PPB $\ll N_{sp}$ of the RX, the TX ASE has a negligible impact on communication performance. A more detailed analysis including the impact of broad-band ASE on TX performance is given later in this section.

While we have shown above that the noise figure for a power amplifier is not a driving design concern, it is worth noting that with careful design, power amplifiers can deliver high gain, efficiency, and output power. By incorporating mid-stage isolation and out-of-band ASE filtering, multi-stage amplifiers have been designed with > 50 dB gain and near quantum-limited noise figures of ~3.1 dB, (see, e.g., [349,350]). This class of performance can be obtained from commercially available EDFAs, with small-signal-gain > 45 dB and NF < 3.5 dB readily available. When cascading additional stages, the net noise figure is given by [126]

$$NF_{net} = NF_1 + \frac{NF_2 - 1}{G_1} \underset{G_1 \gg NF_2 - 1}{\Longrightarrow} NF_1, \tag{75}$$

where NF_i and G_i represent the gain and noise figure of the ith amplifierstage. Clearly, with high-gain in the first amplifier, the net cascaded noise figure can be dominated by the (low) noise figure of the first. Since high-gain and low-NF can be achieved for relatively low (<~100mW) pump powers [350], reduced power efficiency in the first stage amplifier should have little impact on the overall efficiency for > 1 W high-power EDFAs.

3.5.3.3. Amplifier Efficiency

The net power amplifier electrical–to–optical conversion efficiency ($\eta_{E/O}$) is an important design consideration, especially for space-based transmitters since it generally consumes a significant portion of the overall power budget. For example, the design

target for the 5W Mars Lasercom YDFA transmitter power amplifier [1,2] transmitter YDFA power amplifier $\eta_{E/O}$ was ~15% [198], requiring > 33W of electrical power, more than 20% of the total power budget.

The main factors that contribute to $\eta_{E/O}$, include pump laser electrical-to-optical efficiency (η_{pump}), pump coupling efficiency ($\eta_{coupling}$), and pump-to-amplified-signal optical-to-optical conversion efficiency ($\eta_{O/O}$), with

$$\eta_{E/O} = \frac{P_{optical-out}}{P_{electrical}} = \eta_{pump}\eta_{coupling}\eta_{O/O}\eta_{overhead}. \quad (76)$$

Here, $\eta_{overhead}$ is included to account for any cooling or control that is not included in the other terms.

Pump laser efficiency includes the electrical power required to drive the pump as well as temperature and current control needed to reliably operate the pump at the proper wavelength. Improvements in pump laser efficiencies in the 0.9 μm absorption bands have occurred in several areas, including significant advancements in the reduction of heating/cooling power required to keep the pump wavelength within the absorption band of the gain medium. In single-mode pump lasers, FBG stabilized designs have been used to efficiently lock the laser wavelength near the absorption peak over an extended temperature range [257–259,351,352], with η_{pump} >~40% and > 200 mW coupled into single-mode fiber for individual single-mode sources. Single-mode coupling techniques often incorporate polarization and/or wavelength multiplexing to combine pump and signal. To increase single-mode pump power levels, both polarization and dense WDM of stabilized pump-lasers can be used to efficiently combine multiple sources into a single-mode [353]. These techniques can scale multi-stage single-mode EDFAs to Watt-class power levels [354]. In high-power cladding-pumped amplifier designs [75], ER:Yb codoped gain fiber is often used to broaden the absorption band, which can extend from 850 to 1100 nm [355–357]. This provides greater flexibility in selection of pump wavelengths, and allows for efficient absorption of uncooled pump wavelengths as they drift over an extended temperature range. Codoping also permits high erbium doping concentrations in fused silica fibers with without quenching induced degradation in gain or efficiency [358].

Cladding-pumped amplifier designs permit efficient coupling of many multi-Watt multimode pump lasers into a double-clad gain fiber, which consists of a large multimode outer-diameter waveguide with high numerical aperture (NA), and a smaller doped-core single-mode waveguide (required for FSO applications) with lower-NA [359]. Pump power coupled into the low-loss outer waveguide is eventually absorbed by the gain-media within the inner single-mode core. The ability to accept output from multimode pump lasers in the large-NA outer waveguide, significantly increases pump coupling efficiency and available pump power levels, enabling the use of efficient high-power (multi-Watt) sources with η_{pump} =~50% [360–362] and a variety of low-loss pump-coupling and combining strategies [75,349,363–372].

Optical-to-optical conversion or power conversion efficiency [126], defined as the ratio

$$\eta_{O/O} = \frac{P_{out} - P_{in}}{P_{in}^{Pump}} \leq \frac{\lambda_{pump}}{\lambda_{sig}} \quad (77)$$

is limited by the quantum-optical conversion efficiency in which a single pump-photon at λ_{pump} is converted to a single signal-photon at λ_{signal}. For 1.55 μm EDFAs with

0.98 μm pumping, $\eta_{O/O} \leq 63\%$, whereas for 1.06 μm YDFAs with 0.98 μm pumping, $\eta_{O/O} \leq 90\%$. Optical-to-optical conversion efficiency is further reduced by other losses in the amplifier such as incomplete pump absorption, insertion loss of filters, isolators and splices, excited state absorption, scattering, and other parasitic loss mechanisms [373].

The state-of-the-art high-efficiency high-power amplifier demonstrations are based cladding-pumped designs and include a high-reliability single-polarization EYDFA at 1550 nm with output power > 10W and $\eta_{E/O} = \sim 13\%$ [199]. With the combined use of PM components and internal polarizers, a 30-dB polarization extinction ratio (PER) was achieved. At 1064 nm, a high-efficiency YDFA design based on high-reliability PM components but without polarizing elements demonstrated output power > 8W, $\eta_{E/O} = \sim 21\%$, and ~13 dB PER [204]. Efficient laboratory-grade YDFAs with $\eta_{O/O}$ approaching 80% [374] and $\eta_{E/O}$ up to 40% [200] have been reported. We note that Nd-doped amplifiers have also been considered for use in space-based applications, however reported efficiencies are lower and they are less tolerant to the effects of radiation [228].

In section 1.1, we introduced net transmitter efficiency (η_{TX}), which represents the energy required to generate TX photons. In absolute terms, this is limited by $\eta_{E/O}$ and TX wavelength as summarized in Table 8 for the amplifiers mentioned above. Note that while 1μm amplifiers generally have better $\eta_{E/O}$ in terms of power efficiency due to a smaller quantum defect, this advantage is reduced by a comparable amount when TX efficiency η_{TX} is normalized in terms of energy per photon as shown below. This is a direct consequence of Eq. (4) with photon-energy being inversely proportional to wavelength.

Table 8. Minimum TX photon-generation efficiency in units of Joules/photon for demonstrated high-efficiency power amplifiers. Note, 1 attoJoule (aJ) = 10^{-18}J.

λ, [μm]	$\eta_{E/O}$	$\eta_{TX} \geq \frac{h\nu}{\eta_{E/O}} = \frac{hc}{\lambda\eta_{E/O}}$, [aJ/photon]
1.55	13%	1.0
1.06	21%	0.9
1.06	40%	0.5

Another important efficiency metric is the amplifier wall-plug efficiency, which includes conversion and conditioning ($\eta_{E/E}$) required to adapt the available power source from the "wall", (e.g., 120 V_{ac} for terrestrial applications or 28 V_{dc} for space-based applications) to voltage levels suitable for driving the amplifier hardware, primarily the pumps, which typically require a ~3 V_{dc} supply to drive lasers with bias voltage < 2.5 V:

$$\eta_{wall} = \eta_{E/E}\eta_{E/O}. \tag{78}$$

Typical values for $\eta_{E/E}$ are between ~70% and 90% depending on the voltage conversion levels, a factor that can influence the amplifier electrical design. In [204], for example, the pump lasers for each stage are wired in series to maximize electrical power conversion efficiency, yielding a YDFA with $\eta_{E/E} = \sim 85\%$, and $\eta_{wall} = \sim 19\%$.

3.5.3.4. Polarization-Maintaining (PM) Fiber Amplifier Designs

As noted at the beginning of section 3.5.3, optical amplifiers that can maintain the polarization of the optical signal are desirable for free-space applications since they provide additional capabilities including polarization diplexing of TX and RX signals and improved communication performance since orthogonally polarized amplifier and background noise can be eliminated in the receiver. They are also necessary for transmitting polarization modulated signals and expand the options for implementing multi-access FSO terminals that incorporate both wavelength and polarization multiplexing [48].

The conventional method of maintaining a single-polarization through a fiber amplifier uses passive PM components and gain media. However, it is challenging to achieve good polarization extinction ratio (PER) with this approach due to polarization cross-coupling of multiple cascaded components with limited PER. Amplifiers using this design typically can achieve ~12–17 dB PER as in [198,204]. To maintain good PER, polarizing elements, such as single-polarization isolators, can be introduced to the PM optical path to remove the orthogonal polarization seepage before it becomes excessive. As shown in [199], this approach is consistent with achieving good polarization extinction (PER = ~30 dB) without degrading electrical-to-optical conversion efficiency. However, this approach precludes the use of polarization modulation.

Active polarization control [375–377] can be used to compensate for polarization fluctuations in a non-PM amplifier at the expense of additional complexity and increased size, weight, and power. Alternatively, passive compensation can be achieved through a double-pass amplifier design incorporating a mid-stage Faraday rotator mirror [353,354,378]. This approach can compensate for both inherent and time varying birefringence to maintain either single or dual-polarization states with good PER unlike the passive designs based entirely on cascaded PM components noted above.

The basic amplifier configuration is shown in Fig. 32. A single-mode PM optical signal enters the amplifier through the input port of the PM I/O coupling element, which could be a polarization beam splitter (PBS) in a single-polarization design or a PM circulator for dual-polarization applications. Unlike conventional PM designs, the input and output ports of the I/O coupler are the only ones that must be polarization maintaining. This greatly simplifies the amplifier design since all other elements can be non-PM. The signal makes a first pass through the gain media and filter, and is then reflected by the Faraday rotator mirror (FRM) which reflects the amplified-filtered signal and rotates the polarization orthogonal to that of the input. After the second pass, the twice-amplified signal polarization is orthogonal to that of the PM input, and is directed towards the amplifier PM output. The PER performance for the entire amplifier is primarily a function of the PM I/O coupler PER and the 90° rotation of the FRM, which can exceed 30 dB with commercially available components. Pump coupling to the gain media can be achieved through a variety of single or multi-mode, WDM, side- or end-coupled elements mentioned in the Amplifier Efficiency Section above and related references.

A key feature of the double-pass design is that it re-uses the same gain medium, giving it twice the comparable small-signal gain of the equivalent single-pass design, but with fewer components in the critical path. However, with the double-pass design, the forward gain is equal to the backward gain and pump power that is not converted during the forward pass can be extracted during the backward pass. When operating

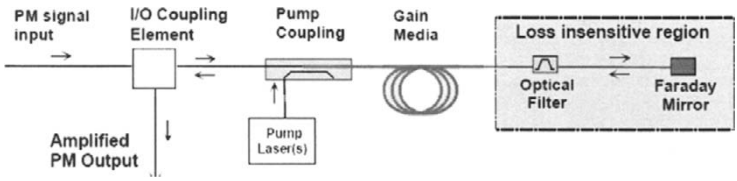


Fig. 32. Schematic for double-pass polarization-maintaining optical amplifier.

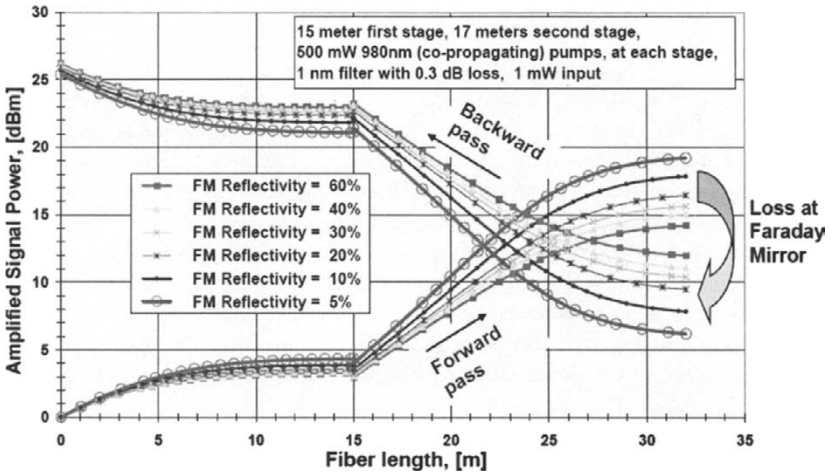


Fig. 33. Simulation of amplified signal power as a function of length and Faraday mirror reflectivity in a 2-stage double-pass EDFA with forward-propagating pumps. For FRM losses between 2 and 13 dB, simulated results show $<\sim$0.5 dB loss in output power, consistent with experimental measurements [354].

in saturation, the average gain adjusts over the length of the double pass EDFA until the amplified signal extracts all the available converted pump power. As a result, the double pass design is less sensitive to the distribution of pump power, in contrast to single pass designs, where the pump power of the final stage is critical. Furthermore, while excess loss near the amplifier output directly reduces the output power by 1 dB per dB loss, additional attenuation in the loss-insensitive region of the double–pass has little effect on net output power, as long as the amplifier continues to operate in saturation [354] as shown in Fig. 33.

Consequently, advantageous but lossy elements such as optical filters, additional pump-coupling elements, WDMs, etc., can be inserted into the design in the loss insensitive region with little power penalty at the output [379]. The optical filtering, for example, can perform double-pass pulse-shaping and remove out-of-band ASE that competes with the signal for gain, enabling the amplifier to saturate at lower

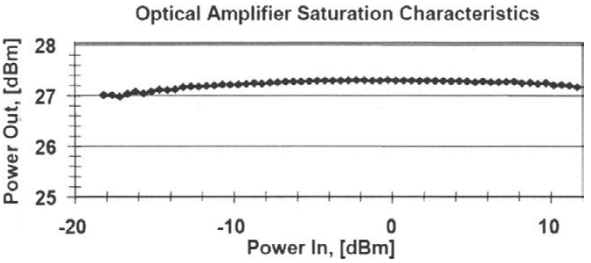


Fig. 34. Measured output saturation characteristics for a high-gain two-stage double-pass ~0.5W EDFA with mid-stage narrow-band optical filter [354].

input power levels. The filter also makes the amplifier more stable by increasing the threshold of lasing at the ASE peak, thus reducing the risk of potentially damaging Q-switched pulses. This is illustrated in Fig. 34, for a high-gain double-pass EDFA with mid-stage narrow-band optical filter. Without any internal isolation, the output power is stable to better than 0.5 dB over a 30-dB dynamic range of input power and 45-dB gain. In addition to stable PM performance, this amplifier design has demonstrated good power efficiency, with optical-to-optical conversion efficiency of ~45% at 1.55 μm [354]. Such characteristics combined with radiation tolerance of ~0.1 dB/krad [380], are attractive for use in space-based lasercom systems [21], especially those with variable-duty-cycle variable-rate capabilities discussed further in section 3.5.5

3.5.4. High-Efficiency Semiconductor Optical Amplifiers

While power-amplifier needs historically have been met by rare-earth-doped fiber amplifiers, alternative semiconductor-based optical amplifier technologies show promise at Watt-class power levels [197,381]. Efficiency, single-polarization, and extened wavelength range are some performance advantages that semiconductor optical amplifiers (SOAs) have over fiber amplifiers. More importantly, SOAs have the potential for improved size, weight, and power which, along with to manufacturability and integration into more comprehensive subsystems, can significantly reduce cost. Since SOAs are electrically pumped, the quantum-defect and other losses associated with optical-optical conversion are absent, so in principle, $\eta_{E/O}$ could be greater than $>$ 25%. Slab-coupled optical waveguide amplifiers (SCOWAs) are of particular interest since they have demonstrated Watt-class (chip) output powers with large mode size compatible with efficient coupling to single-mode fiber [197].

In comparison with EDFAs, the dynamics in SOAs, like semiconductor lasers discussed in section 3.1, are more complex and faster by over six orders of magnitude. The slow millisecond upper-state lifetime in erbium yields attractive kHz-class high-pass characteristics that are well suited for high-speed optical communications, enabling EDFAs to deliver a constant average power limited (APL) output when driven by high-frequency time-varying input signals above ~10 kHz [126]. For signals with frequency content below the cut-off frequency, the average amplifier gain and output power track the input signal. Output power and efficiency are reduced for low-frequency, low-duty

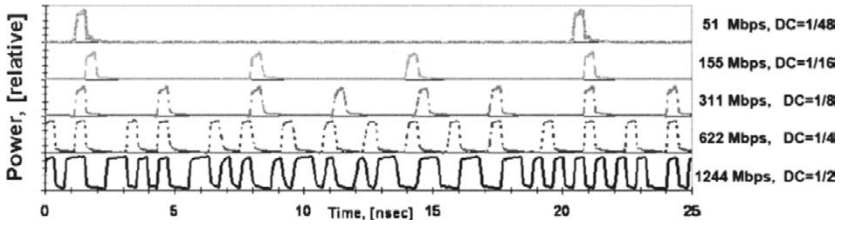


Fig. 35. Measured peak-power-limited (PPL) variable-duty-cycle 2-PPM waveforms at the input to a high-gain average-power-limited EDFA [11].

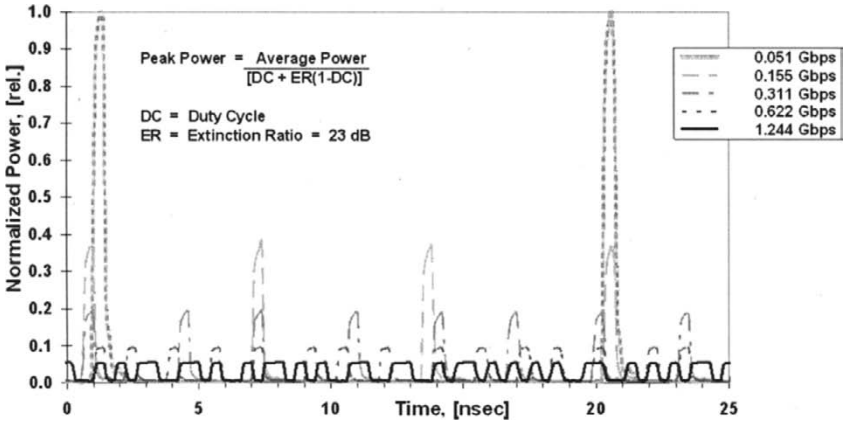


Fig. 36. Measured variable-duty-cycle 2-PPM waveforms at the output of an average-power-limited high-gain EDFA [11]. Reprinted with permission of IEEE. (©1999 IEEE.)

cycle waveforms, i.e., the amplifier acts peak-power-limited in this low frequency regime. Since SOAs have a nanosecond-class upper-state lifetime, the high-pass corner frequency is shifted up towards ~1 GHz, comparable to the Gbit/sec data rates of interest. Owing to these high-pass characteristics, maximum TX power-efficiency is attained whenever the transmitted waveforms have little (power off) spectral content below ~1 GHz, or equivalently, the power-off time is short relative to the upper-state lifetime [172]. For this class of amplifier, power efficiency is optimized for modulation formats that have high ~100% duty-cycles, such as DPSK and FSK, or formats with constrained power-off time such as binary-PPM at rates >~5 Gbps or RZ-DPSK/FSK at rates >~10 Gbps. Outside of these efficiency considerations, communication experiments using an SOA-based MOPA TX have achieved near-quantum-limited receiver sensitivities for binary-PPM, OOK, and DPSK modulation formats, showing that SOA-based transients have little impact on RX sensitivity [172].

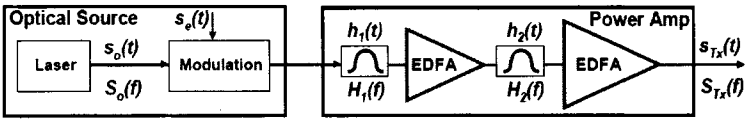


Fig. 37. An average-power-limited MOPA TX. Net optical filtering can comprise multiple pulse-shaping and mid-stage ASE rejection filters, i.e., $h_{\text{net}}(t) = h_1(t) * h_2(t)$, or equivalently, $H_{\text{net}}(f) = H_1(f) \cdot H_2(f)$.

3.5.5. Arbitrary Waveforms and Variable-Duty-Cycle Signaling

The characteristics of an average-power-limited high-gain saturated-EDFA described above provide a flexible and powerful platform for waveform optimization. Once a power amplifier is driven far into saturation, the output power added by the amplifier becomes insensitive to the average input power, making it a stable source of average power. Higher amplifier gain is beneficial since it widens the range of input power levels that can extract the maximum saturated output power. This makes the TX less sensitive to insertion loss changes in the preceding elements that can arise due to modulation or degradation [354], and allows the use of arbitrary waveforms over a wide range of duty-cycles without sacrificing transmitter power efficiency [10].

For the constant average output power delivered by a saturated EDFA, the peak output power is described by

$$P_{\text{peak}} = \frac{\bar{P}}{\text{DC}_{\text{eff}}} \quad \underset{go P_{\text{in}} \gg P_{\text{sat}}}{\Longrightarrow} \quad \frac{P_{\text{sat}}}{\text{DC}_{\text{eff}}} \tag{79}$$

and grows as the duty cycle is reduced. Here we have introduced the term, $\text{DC}_{\text{eff}} = \text{DC}_{\text{mod}}\,\text{DC}_{\text{pulse}}$ to identify the combined influence of both modulation and pulse shape on the effective duty cycle. The impact of variable duty cycle input waveforms (Fig. 35) on the output of a saturated high-gain EDFA is shown in Fig. 36.

Given these basic properties, an average-power-limited MOPA TX can be used to efficiently deliver high-power high-fidelity waveforms of arbitrary type [10], within limitations described below. For instance, a narrow-band electrical waveform $s_e(t)$ can simply be imparted on the CW laser source $s_o(t)$ in Fig. 37 and subsequently amplified without distortion as long as the bandwidth of the net filter $H_{\text{net}}(f)$ is sufficiently wide.

To generate high-fidelity optical waveforms that exceed the available electrical bandwidth, pulse carving techniques described in section 3.5.2 can be used. In this case, the electrical drive waveform could perform additional shaping or serve as a lower-frequency windowing function to select which pulses are transmitted. Additional pulse-shaping can also be performed in the optical domain through subsequent optical filtering, a process that can efficiently generate shaped signaling-waveforms with THz of spectral content. For example $s_o(t)$, could be a windowed short-pulse laser source. The TX output waveform $s_{\text{TX}}(t)$ becomes

$$s_{\text{TX}}(t) = [s_e(t)s_o(t)] * h_{\text{net}}(t) \quad \overset{\text{FT}}{\longleftrightarrow} \quad S_{\text{TX}}(f) = [S_e(f) * S_o(f)]\, H_{\text{net}}(f). \tag{80}$$

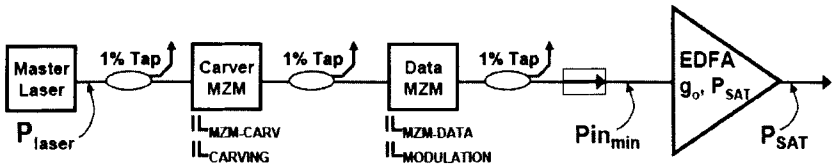


Fig. 38. Variable-duty-cycle pulse-carved MOPA-TX.

For sufficiently short optical pulses, s_o(t) looks like a delta function $\delta(t)$, and $s_{\mathrm{TX}}(t) \rightarrow h_{\mathrm{net}}(t)$. These and other wide-band pulse-shaping techniques (see, e.g., [382]), could, in principle, be incorporated into an APL TX with little impact on overall TX power efficiency.

When using a MOPA configuration with a wide range of variable-duty-cycle waveforms such as those shown in Figs. 35 and 36, the final power amplifier stage needs to run deeply saturated in order to efficiently deliver maximum average power. Efficiency is ultimately limited by the power of the master laser, mid-stage optical filtering and losses, and amplifier gain and noise figure. The impact on transmitter design is illustrated in the representative TX power budget given in Table 9 for the MOPA-TX shown in Fig. 38.

Table 9. Variable-duty-cycle, pulse-carved MOPA-TX power budget.

	Low Duty Cycle	High Duty Cycle
Master Laser Power (P_{laser})	15.0 dBm	15.0 dBm
99:1 PM tap	–0.2 dB	–0.2 dB
Static carver loss ($IL_{MZM-CARV}$)	–3.5 dB	–3.5 dB
Pulse-carving loss	–3.0 dB	–3.0 dB
99:1 PM tap	–0.2 dB	–0.2 dB
Static data mod. loss($IL_{MZM-DATA}$)	–3.5 dB	–3.5 dB
(VDC) Modulation loss	–13.0 dB	–3.0 dB
Isolator	–0.6 dB	–0.6 dB
99:1 PM tap	–0.2 dB	–0.2 dB
Other loss	–0.8 dB	-0.8 dB
Σ = Minimum power delivered to Power Amp PA (Pin_{min})	–10.0 dBm	0.0 dBm
Amplifier small-signal gain (g_o)	50.0 dB	50.0 dB
Efficient power extraction for Psat < Pinmin + go –10 dB	30.0 dBm	40.0 dBm

In this example, an amplifier with a small signal gain of 50 dBm can provide a 10-dB range in duty cycle, and deliver up to 1W (30 dBm) of average power with less than $\frac{1}{2}$ dB loss in efficiency. In order to maintain maximum output power and extend the range of duty cycles, additional gain stages along with inter-stage filtering may be required as discussed in the following Sections. With proper design, a dynamic range of operation in excess of 20 dB can be achieved [354].

3.5.5.1. *Low Duty Cycle Limitations*

As shown in the previous section, careful TX design is necessary to maintain TX efficiency over a wide range of duty-cycles. Several other factors can limit the useful range of duty cycles of a MOPA TX. These include a) limited modulation extinction, b) buildup of ASE, and c) nonlinear impairments.

3.5.5.2. *A) Limited TX Modulation Extinction*

Imperfect extinction can reduce the effective transmitted signal power for low-duty-cycle waveforms (see also section 3.5.2). As seen in Eq. (67), the peak power is a function of the average power, effective duty cycle and extinction ratio. For the case of an EDFA operating deep in saturation, Eq. (67) becomes [314]

$$P_{\text{peak}} = \frac{P_{\text{sat}}}{\text{DC}_{\text{eff}} + \text{ER}(1 - \text{DC}_{\text{eff}})} \quad \underset{\text{ER} \ll \text{DC}}{\Longrightarrow} \quad \frac{P_{\text{sat}}}{\text{DC}_{\text{eff}}}. \tag{81}$$

As the duty cycle becomes low, the average power seeping through the modulator due to imperfect extinction can become significant. For example, for a 1% duty cycle signal and a modulator with a –20 dB ER, the net signal power out of the modulator during the 1% transmission will be about the same as the seepage during the remaining 99%. The subsequent average power limited amplifier will output equal power in both signal and seepage, effectively reducing the transmitted signal power by ~3 dB. The extinction *power robbing* penalty is given by

$$\eta_{\text{TxER}} = \frac{\text{DC}_{\text{eff}}}{\text{DC}_{\text{eff}} + ER(1 - \text{DC}_{\text{eff}})} \quad \underset{\text{ER} \ll \text{DC}_{\text{eff}}}{\Longrightarrow} \quad 1. \tag{82}$$

In order to extract maximum amplified signal power from a saturated EDFA the modulation extinction must be much less than the duty cycle. For less than 0.1 dB "power robbing" penalty due to accumulated signal seepage during the (1-DC_{eff}) *off* portion of the waveform, the ER target should be 15 dB less than the duty cycle, i.e., ER $\leq \text{DC}_{\text{eff}}$ –15 dB. Thus, for 1024-PPM with an 0.1% duty cycle (DC_{eff} = –30dB), an ER better than –45 dB is required. A plot of ER measurements and calculated fit (from Eqs. and associated TX power penalty for a modulator with about –52 dB extinction is shown in Fig. 39. Using this modulator with 1024-PPM waveforms would result in negligible <~0.05 dB TX power penalty.

3.5.5.3. *B) Transmitter ASE*

Transmitter ASE buildup can degrade communication performance in two ways: by limiting the transmitted SNR (SNR_{TX}), and by reducing the transmitter efficiency and transmitted-signal output power. The first relates to the ratio of signal photons to ASE photons within the signal-mode and is independent of TX filtering, whereas the second relates to the ratio of output signal to total ASE power ($\text{TXPowerRatio}_{\text{sig/ASE}}$), and has a strong dependence on optical filtering.

As discussed earlier in section 3.5.3, the conditions of Eq. (73) can be met by simply providing sufficient input signal to the power amplifier. This results in sufficient SNR_{TX}, so that the TX ASE has little impact on communication performance. This

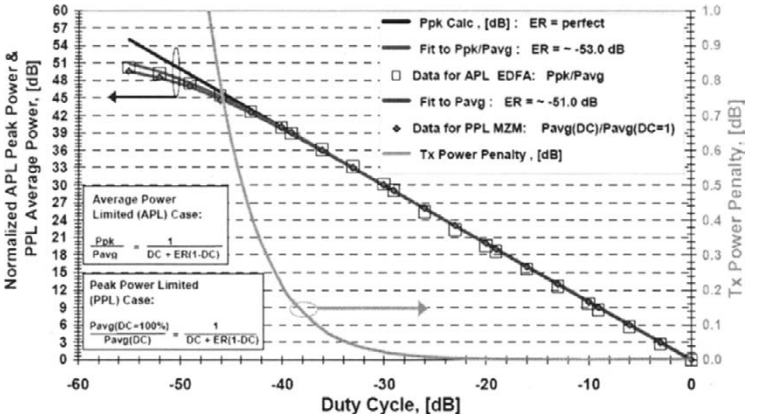


Fig. 39. (Left axis) measurements of modulator extinction using variable duty-cycle techniques described in section 3.5.2. Right axis) TX power penalty [15].

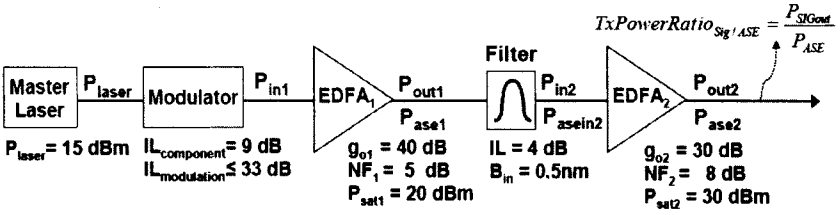


Fig. 40. MOPA transmitter configuration used for modeling the ASE buildup described by Eqs. (83)–(88).

also holds for a cascade of amplifiers, since the net noise figure (and N_{sp}) for the cascade is dominated by the first stage.

However, high-gain multi-stage power amplifiers, can generate significant amounts of broad-band ASE which can reduce the TX power efficiency. With proper design of both the TX power budget and mid-stage optical filtering the $\mathrm{TXPowerRatio}_{\mathrm{sig/ASE}}$ can be optimized. To avoid ASE power robbing, a $\mathrm{TXPowerRatio}_{\mathrm{sig/ASE}} > 20$ dB is desired, meaning more than 99% of the output power is the signal and less than 1% of the TX power is ASE. An example MOPA TX shown in Fig. 40 with a wide operational range of input powers is modeled using Eqs. (83)–(88) [126,142].

$$\mathrm{TxPowerRatio}_{\mathrm{sig/ASE}} = \frac{P_{\mathrm{SIGout}}}{P_{\mathrm{ASE}}}, \tag{83}$$

$$P_{\mathrm{SIGout}} = G\frac{P_{\mathrm{laser}}}{\mathrm{IL}_{\mathrm{component}}\mathrm{IL}_{\mathrm{modulation}}} = GP_{\mathrm{SIGin}}, \tag{84}$$

$$P_{\mathrm{in}} = P_{\mathrm{SIGin}} + P_{\mathrm{ASEin}}, \tag{85}$$

$$P_{\mathrm{ASE}} = (G-1)(N_{\mathrm{sp}}mBh\nu) + GP_{\mathrm{ASEin}} \cdot \min(B/B_{\mathrm{in}}, 1), \tag{86}$$

$$N_{\mathrm{sp}} = \frac{\mathrm{NF}}{2}\left(\frac{G}{G-1}\right) \underset{G\gg 1}{\approx} \frac{\mathrm{NF}}{2}, \tag{87}$$

$$G(P_{\mathrm{in}}, P_{\mathrm{sat}}) = 1 + \left(\frac{g_o}{1 + P_{\mathrm{in}}g_o/P_{\mathrm{sat}}}\right). \tag{88}$$

Here, B is the optical filter bandwidth and m is the number of polarization modes. These equations can be iteratively applied to solve for the cascaded amplifier output signal to ASE power ratio, which depends directly on the signal input power, and varies with the modulation duty cycle. For the first amplifier stage, $P_{\mathrm{ASEin}} = 0$ in Eq. (85). The amplified signal P_{SIGout} and ASE power P_{ASE} from the first amplifier are then used as inputs to the second amplifier stage. Variable-duty-cycle M-ary PPM waveforms are simulated for the binary case ($M = 2$) through 1024-PPM with 50%-RZ pulse carving. The corresponding modulation insertion loss, $\mathrm{IL}_{\mathrm{modulation}} = 1/\mathrm{DC} = 2M$ varies from 6 to 33 dB (assuming perfect ER). Representative elements contributing to the component insertion loss ($\mathrm{IL}_{\mathrm{component}}$) are given in Table 9.

Simulation results for the two-amplifier cascade with a narrow $\frac{1}{2}$ nm (62 GHz) mid-stage filter are shown in Fig. 41. Over the 27-dB range of input powers, the average output power is stable to ~0.5 dB. For M = 1024, with 10 bits of information carried per PPM symbol, the duty-cycle and input power are the lowest, but under these conditions, the APL amplifier generates the highest peak power of 62.5 dBm (~1.8 kW). In this case, the minimum $\mathrm{SNR}_{\mathrm{TX}}$ is ~27 dB and has more than 99% of TX output power in the signal with $\mathrm{TXPowerRatio}_{\mathrm{sig/ASE}}$ >22 dB.

3.5.5.4. *C) Nonlinear Impairments*

Optical nonlinearities are another factor that can impact low duty cycle transmission through a saturated average-power-limited transmitter. Fast nonlinear effects such as self-phase modulation (SPM), four-wave mixing (FWM), and Stimulated Raman Scattering (SRS) tend to shift usable signal power out of band, an effect that reduces transmitter efficiency and makes an average power limited amplifier appear peak power limited. Nonlinear effects occur when the peak power exceeds a threshold

$$P_{\mathrm{th}} \propto \frac{A_{\mathrm{eff}}}{g_{\mathrm{NL}}L_{\mathrm{eff}}}, \tag{89}$$

in which A_{eff} is the effective area, G_{NL} is the nonlinear gain coefficient (a material property), and L_{eff} is the effective propagation length [110,111]. Standard steps to increase the nonlinear threshold include using larger core fiber (which reduces the power density), minimizing amplifier lengths [383], and distributing amplified signal energy in time to lower the effective peak power through the amplifier. A summary of high power amplifier design options and tradeoffs is given in Table 10.

Practically, P_{th} can be defined as the peak output power when greater than ~1% of the average amplifier output is contained in the nonlinear induced spectrum. Typically,

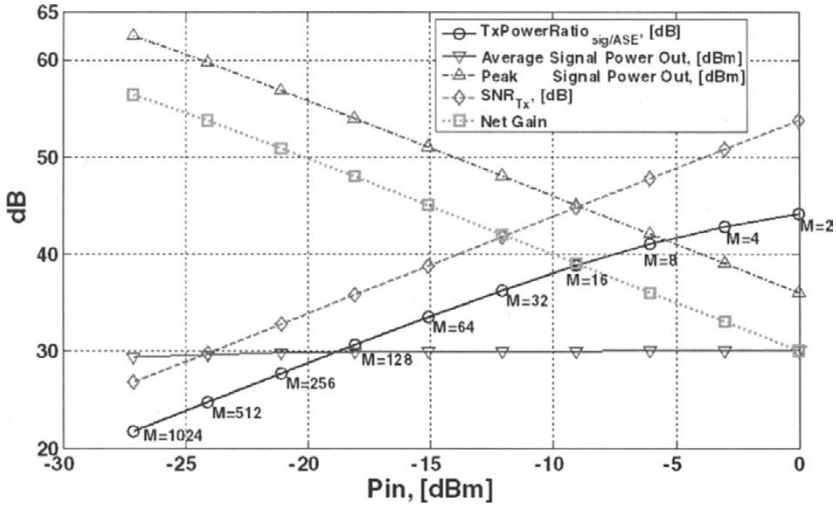


Fig. 41. Simulated amplifier signal to ASE power ratio ($\mathrm{TXPowerRatio_{sig/ASE}}$ in circles), average signal output power (down-pointing triangles), peak signal output power (up-pointing triangles), TX SNR (diamonds), and net amplifier gain (squares) as a function of input power.

P_{th} is about 250W to 1kW in commercially available EDFAs [21,198,199], although promising research-grade EDFAs have generated 100 kW-class peak power with little or no nonlinear effects [224]. Once beyond P_{th}, the nonlinear effects rapidly dominate. For the MOPA TX modeled above, the lowest duty cycle generated a peak power of $\sim$1.8 kW, which can lead to prohibitive nonlinear impairments in conventional power amplifiers. Such effects are illustrated in Fig. 42.

Note that the signal at $\sim$1557 nm is $>$ 33 dB larger than the amplifier ASE which peaks at $\sim$1567 nm for this amplifier. Therefore the net ASE power is less than 1% of the net transmitted power and can be neglected. As the peak power of the amplifier is increased (by lowering the duty cycle of the input), nonlinearities appear. The combined effects of dispersion and SPM can lead to *modulation instability* [111], which can generate spectral sidebands around the signal. These are often the first noticeable nonlinear effect. SPM on its own can broaden the amplified signal spectrum, although this is not as apparent when viewing the broad band optical spectrum. As the peak power continues to increase, the nonlinearly-induced side bands grow and take a larger percentage of the amplifier output power. For this amplifier, SRS tends to dominate at higher peak powers, shifting the optical spectrum to the longer wavelengths. As the peak power is increased further the SRS-generated spectral components can extend beyond 100 nm, stealing a significant portion of the net amplifier output power.

Another nonlinear effect is stimulated Brillouin scattering (SBS) which is a slow nonlinearity that scatters light in the backward direction. This attenuates the forward propagating signal, reducing the TX efficiency and can lead to catastrophic damage in the final amplifier stage due to SBS induced Q-switching. In standard single-mode fiber with an effective area of 50 $\mu\mathrm{m}^2$, an estimate of the SBS power-length threshold

Table 10. Design options for stable high-peak-power single-mode (spatial and temporal) wave guide amplifiers.

#	Option	Benefit	Limitation	Drawback
1	Larger core fiber	~2 – 10 x	Multimode operation	• Poor fiber coupling • Reduced NA • Coupling loss • Sensitive to bend loss
2	Minimize fiber lengths through higher doping concentration	~2 – 4 x	Maximum doping concentrations in fused silica	• Dopants may lead to enhanced radiation sensitivity for space-based applications • Reduced power efficiency
3	Coiled multimode [224, 384]	~100 x	Efficiency of fundamental mode	• Induced polarization dependence • Reduced power efficiency • Reliability of coiled fiber • Research grade only
4	Self-imaging waveguide [385]	~100 x	Technical, doping levels	• Coupling efficiency and stability • Thermal and modal stability • Research grade only
5	Post amplified chirped pulse compression	~2 – 4 x	Maximum chirp and compensation	• Reduced power efficiency • Potential for waveform distortion
6	WDM distribution [22]	~100 x	• Complexity and cost (SWAP) • Available optical bandwidth • FWM efficiency reduction	• Increased complexity, cost, and SWAP over single channel • FWM nonlinearities may lead to data-dependent power efficiency reduction, which can lead to an error floor without FEC

is ~20 W m within the ~50 MHz Brillouin linewidth. However, in contrast with the ultra-fast SPM and SRS nonlinearities, SBS has a narrow gain bandwidth ($\Delta\nu_b$ <~50 MHz), and therefore, can be mitigated by broadening the signal spectrum. This can be achieved with the use of short duration signaling waveforms or inducing additional phase modulation [110,111]. For example, by using short 100 ps pulsed waveforms with ~10 GHz class bandwidth, this threshold can be increased by a factor of 200, to ~4 kW m. In addition to the nonlinear mitigation techniques highlighted in Table 10, SBS can be suppressed through temperature gradients in the fiber [374]. As a result of the temperature dependent Brillouin frequency shift, temperature segmentation of the high-power-handling fiber can be used to disrupt the SBS phase matching, yielding up to 7 dB suppression [386].

The use of WDM (Fig. 43) to reduce peak power by distributing the amplified signal energy over many wavelengths is particularly attractive for use in communication systems employing strong FEC. This approach directly reduces slow SBS effects by lowering the peak power within the SBS bandwidth. However, fast SPM, FWM, and SRS nonlinearities have many THz of bandwidth, so the benefits of this approach occur only when the WDM signals do not overlap in time for relatively short, ~10 m amplifier lengths. Fortunately, as the duty cycles get lower and peak power impairments grow, the probability of pulse overlap is reduced. Nonlinear impairments that occur due to

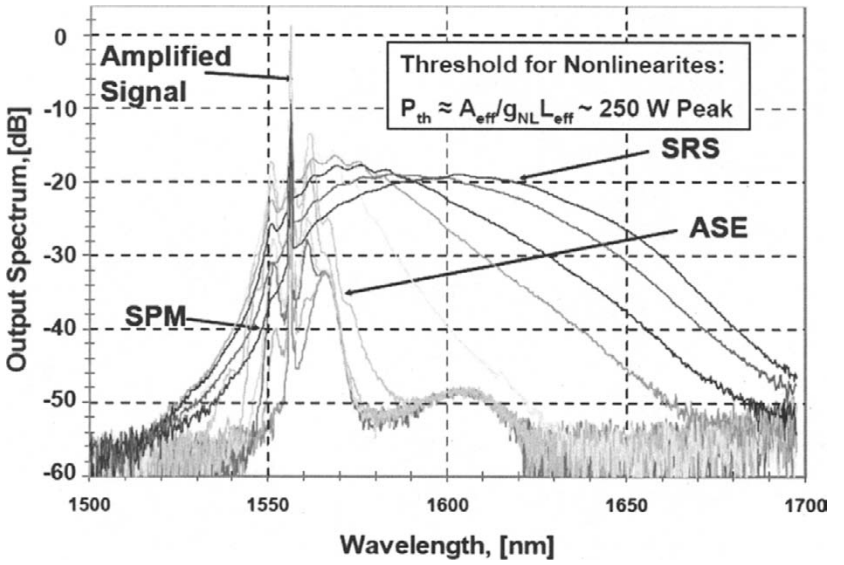


Fig. 42. Nonlinearly-induced spectrum for several peak output powers in an Er:Yb doped cladding pump amplifier. The bottom trace shows the amplified signal spectrum from a 4-W EDFA without any nonlinear effects. Fiber nonlinearities can generate new spectral components which shift usable signal power out of band, effectively making an average-power-limited amplifier look peak-power-limited [21].

sporadic pulse overlap can cause a pattern-dependent performance degradation. But as long as these events occur infrequently enough, they can be compensated for by FEC.

Consider, for example, a 10-W MOPA-TX with $P_{\rm th} = \sim 1$ kW using square 1024-PPM waveforms, and a receiver employing FEC with an error-free cutoff at a 10^{-3} error rate. Single-channel peak powers would be ~ 10 kW. By distributing the TX over 32 wavelengths, the peak-power per channel is reduced to 320 W. But the aggregate peak power can be higher due to multi-channel pulse overlap. Assuming equal probability for each PPM symbol and time-alignment between the independent channels, an upper-bound for the probability distribution of peak aggregate power is shown in Fig. 44 for 8, 16, 32, and 64 WDM channels.

For the 32-channel configuration, the overlap of 3-pulses yields a peak power of 960kW, below $P_{\rm th}$. The overlap of 4-pulses yields a peak power of 1.28 kW, above $P_{\rm th}$, but with a probability $< 3 \times 10^{-5}$, well below the FEC error cutoff. Therefore, the WDM signal distribution over 32 channels effectively suppresses broadband nonlinear impairments by about a factor of ten. This corresponds to the ratio of the number of WDM channels, w, to the effective number of pulses that overlap with a probability greater than the FEC threshold rate, $\kappa_{\rm eff}$, which is 3 in the example above. The effective peak power then becomes

$$P_{\rm pk}^{\rm eff} = \left(\frac{\kappa_{\rm eff}}{w}\right) \frac{\bar{P}\cdot}{DC_{\rm eff}}. \tag{90}$$

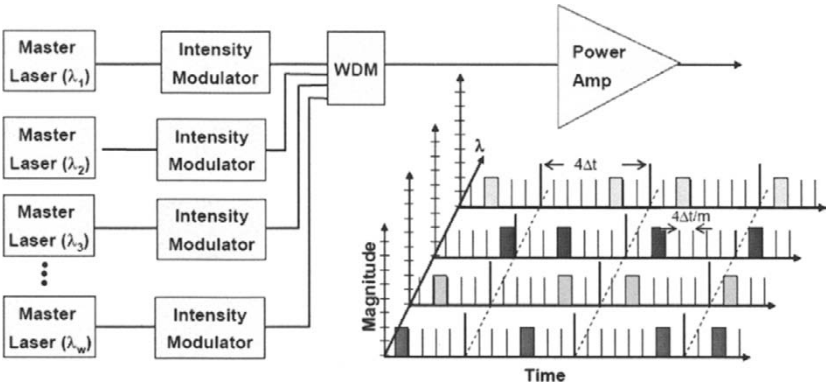


Fig. 43. M-PPM/WDM waveforms illustrating the WDM method of diminishing peak signal power by distributing it over many wavelengths Relative to a single channel TX of same duty cycle, average power, and symbol time Δt, a w-channel WDM transmitter reduces the peak power and bandwidth of each channel by a factor of w [22]

This benefit could be furthered by intentionally staggering the time alignment of the WDM channels to reduce κ_{eff}.

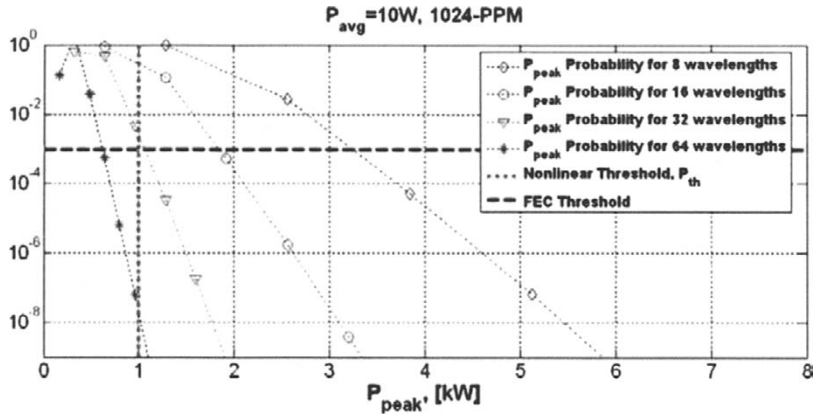


Fig. 44. Peak power probability distribution due to pulse overlap in a 10-W average power 1024-PPM multi-channel WDM MOPA TX. The pulses are assumed to be synchronized, so that the overlap, when it occurs is complete. Note that the resulting peak powers occur at discrete levels, depending on the integer number of pulses that overlap. Also shown are representative thresholds for the onset of nonlinear effects (vertical dotted line) and FEC cutoff (horizontal dashed line) [22,479] that break the chart into 4-quadrants. BER performance will be degraded for curves with points contained within the upper-right quadrant.

The use of WDM has additional advantages. First, it increases the average input to the TX power amp, which reduces the amplifier gain and output ASE, and improves the output signal-power to ASE ratio. It is also a means of overcoming electronic bandwidth and performance limitations, enabling scaleable data-rates that can leverage the available THz optical spectrum.

4. Receiver Technologies

Along with photon and power efficiency, important receiver metrics include reliability, especially in the space environment, and performance in the presence of background noise and atmospheric channel effects. In this section we will present the relative performance of photon-counting, coherent-homodyne, and preamplified receivers. Systems based on these technologies show the most promise for next generation solutions for ultra-long distance communication.

For a quantum-limited optical receiver, the SNR of received optical signals is ultimately limited by the number of photons/bit, which is independent of the shape the waveform. This leads to flexibility in the receiver to accept waveforms of arbitrary shape. The average-power-limited properties of optical power amplifiers discussed in section 3 allow similar freedom in waveform generation. This combination enables the use of aggressive pulse shaping to efficiently match the TX signal waveform to realizable RX filters. Furthermore, there is abundant optical spectrum available which eliminates channel bandwidth limitations and enables excess spectrum to be traded for improved sensitivity through modulation and coding.

Note that there is an important distinction between the quantum-limited SNR defined in Eq. (32) and the optical SNR (OSNR), which is a commonly used metric for accessing communication performance in fiber-based WDM systems. As discussed in section 2.3.5, the quantum-limited SNR is an absolute measure of SNR relative to quantum-noise ($h\nu$), equivalent to the incident number of photons/bit (for direct-detection receivers) regardless of signal shape. This is not necessarily the case for relative measurements of OSNR, which are often determined by measuring the signal and out-of-band noise power within a specified bandwidth, typically 0.1 nm. The signal power estimated using this technique can vary with pulse-shape, which can lead to dB-class waveform-dependent OSNR errors without proper calibration. Moreover, since the noise power can be much larger than the quantum noise, OSNR measurements of this type are generally a poor measure of absolute receiver sensitivity—the relevant metric for most free-space applications.

4.1. Direct Detection–PIN

The most common receiver for high-speed optical communications is the PIN-photodiode-based detector[10] shown in Fig. 45. PIN detectors can be extremely wide-band, with DC to 50+ GHz response commercially available. While PIN-based receivers are relatively simple, they are the least sensitive and require careful front-end electronic design to achieve good high-rate performance [20,39,387].

[10] The PIN or p-i-n photodiode is a semiconductor-based detector structure with an intrinsic (i) region in between n- and p-doped regions. Relative to p-n photodiodes, PIN detectors have a thicker depletion region,

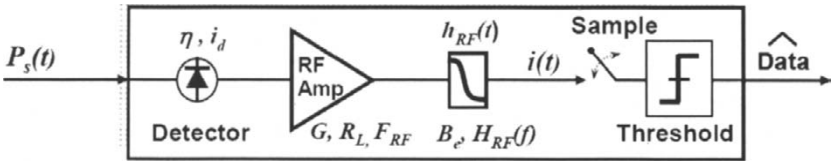


Fig. 45. Schematic of PIN-based direct-detection receiver with input waveform $s(t)$, electrical amplification with load resistance R_L, gain G and noise figure F_{RF}, postdetction filtering with impulse response $h_{\mathrm{RF}}(t)$ and effective electrical bandwidth B_e [see Eq. (39)], prior to sample and thresholding circuitry.

In section 2.3.5, we introduced the subject of shot-noise in the ideal detection process [(see Eqs. (32) and (35)–(40)]. In practice other noise sources such as dark current i_d and thermal noise contribute to the SNR, with

$$\sigma_{\mathrm{th}}^2 = (4k_BT/R_L)F_{\mathrm{RF}}B_e \quad [\mathrm{amps}^2], \tag{91}$$

$$\sigma_{\mathrm{sh}}^2 = 2q(i_s + i_d)B_e, \quad [\mathrm{amps}^2], \tag{92}$$

where σ_{sh}^2 is the shot noise resulting from signal and dark current, and σ_{th}^2 is the thermal noise generated in the load resister R_L [216] scaled by the amplifier noise figure F_{RF} [20, 111]. Incorporating these noise sources into the direct-detection SNR expression in Eq. (40) we obtain

$$\mathrm{SNR}_{\mathrm{DD}} = \frac{\langle i_s \rangle^2}{\sigma_{\mathrm{sh}}^2 + \sigma_{\mathrm{th}}^2} = \frac{\left(R_{\mathrm{A/W}}P_s\right)^2}{2q(R_{\mathrm{A/W}}P_s + i_d)B_e + (4k_BT/R_L)F_{\mathrm{RF}}B_e}, \tag{93}$$

where P_s is the incident signal power, and $i_s = R_{\mathrm{A/W}}P_s$ is the signal photocurrent. The detector responsivity $R_{\mathrm{A/W}} = \eta q/hv = 1$ mA/mW at 1.55 μm, for detection efficiency $\eta = \sim 0.8$, causing a $\sim$1 dB reduction in detected power. When $\sigma_{\mathrm{sh}}^2 \gg \sigma_{\mathrm{th}}^2$, Eq. (93) converges to the quantum-limited SNR in Eq. (40), as expected.

Amplifier noise is often given in terms of thermal noise current density i_{th} [pA/Hz$^{1/2}$], which incorporates both thermal noise and amplifier excess noise. This results in a current noise variance $\sigma_{\mathrm{th}}^2 = i_{\mathrm{th}}^2 B$ [amps^2], and corresponds to an effective resistance $R_{\mathrm{eff}} = 4k_BT/i_{\mathrm{th}}^2$. While thermal noise is a white noise source, RF amplifiers can add frequency dependent $1/f$ noise, which further degrades the SNR at frequencies below $\sim$100 MHz [20]. Relative to the dominant thermal noise, the few nanoamp dark current typical of detectors at 1.55 μm can be neglected in Eq. (93).

To relate the signal and noise terms to OOK communication performance, we introduce the Q-factor, which can be expressed as

$$Q = \frac{i_1 - i_0}{\sigma_1 + \sigma_0} = \frac{R_{\mathrm{A/W}}P_s\,(1 - \mathrm{ER})}{\sqrt{2q(R_{\mathrm{A/W}}P_s(1 + \mathrm{ER}) + i_d)B_e} + 2\sqrt{(4k_BT/R_L)F_{\mathrm{RF}}B_e}}, \tag{94}$$

which improves detection efficiency and reduces capacitance, resulting in wider bandwidth—making them well suited for communications applications.

where $i_{1,0}$ and $\sigma_{1,0}$ are mean signal photocurrent and noise terms for the '1' and '0' symbols, respectively, and ER $= i_0/i_1$ is the modulation extinction ratio (see section 3.5.2), ideally equal to 0. With optimized thresholding, the bit-error-rate (BER) can be estimated by the Q-function given in Eqs. (8) and (9), with the Q-factor above as the argument. A 10^{-9} BER is achieved for $Q = 6$.

To reduce the impact of thermal noise, the load resistance can be increased using transimpedance amplifiers (TIAs). However, due to residual capacitance this limits the bandwidth of the system to about $(2\pi \mathrm{RC})^{-1}$ [20]. Commercial 1-kΩ TIAs are available with B_e = 7 GHz suitable for 10 Gbps communications. Typical thermal noise current for this class of devices is $i_{\mathrm{th}} = \sim 7$ [pA/Hz$^{1/2}$], which corresponds to R_{eff} = 320 Ω, or equivalently $F_{\mathrm{RF}} = \sim 5$ dB. For these parameters, the 10^{-9} BER can be achieved with –21dBm power incident on the detector, which corresponds to ~6200 PPB.

Another means of improving the performance of direct-detection RXs limited by thermal noise is to use low-duty-cycle return-to-zero (RZ) impulsive coding. As noted by Personick [39] in 1973, and more recently, demonstrated by Boivin et al. [343], such techniques can lead to several dB enhancement of RX sensitivity. Discussion of the transmitter considerations of using such waveforms is given in section 3.5.5, and we elaborate on the use of this class of waveforms in section 5.1 for preamplified receivers. For further details on optimizing the direct-detection RX, we refer the reader to thorough analysis by Winzer and Calmar [345].

We note that while PIN-based RX performance is far from ideal, it provides a basis for understanding the benefits of techniques used to achieve near quantum-limited RX sensitivities discussed in the following sections.

4.2. *Direct Detection Avalanche-Photodiode (APD)*

APDs can improve performance over PIN detectors by providing internal gain within the detection process, which multiplies the photocurrent by an average factor of M. However as with any amplification process it is not noiseless. The resulting direct-detection SNR expression becomes [20,121,236]

$$\mathrm{SNR_{DD}} = \frac{i_{\mathrm{avg}}^2}{\sigma_i^2 + \sigma_{\mathrm{th}}^2} = \frac{\left(R_{\mathrm{A/W}} P\right)^2}{2qF_{\mathrm{APD}}(i_{\mathrm{avg}} + i_d)B_e + (4k_B T/R_L)F_{\mathrm{RF}} B_e/M^2}, \tag{95}$$

where F_{APD} is the excess noise factor associated with variations in M. The thermal noise term is reduced by M^2 relative to the PIN detector, whereas the shot-noise term is increased by F_{APD}, which generally grows with M. Typical values of M range from ~3 to 100, corresponding to F_{APD} values of 2 to 10, although these values are vary considerably depending on wavelength, data rate, and technology. At rates of 10 Gbps, $M = \sim 3$ to 10, and $F_{\mathrm{APD}} = \sim 4$, providing about a 10-dB benefit over PIN-receivers, with 10^{-9} BER achieved for –29 dBm or ~1000 PPB receiver sensitivity (see, e.g., [388,389]). At lower rates—less than 1 Gbps, APDs have demonstrated performance in the 100 to 200 PPB regime [390–392]. However, PIN-based RXs offers superior bandwidth and dynamic range of operation, and are less temperamental, offering robust performance over a wide range of environmental conditions.

4.3. Direct Detection–Photon Counting

Qualitatively, photon-counting detectors can be viewed as an extension of APD-type detectors with infinite gain in which, a digital output signal is generated for each detected photon. Due to the binary nature of the detection output, noise in the detection process appears in the form of dark counts or varying detection efficiency, and the detector is limited to counting at most one photon per reset time. As a result, the detector is blinded after a detection event, missing all incident photons until reset, an effect which is referred to as blocking loss [34,35]. Detector arrays can be used to extend the dynamic range of detected signal to multiple counts per-interval and mitigate blocking loss, although this comes at the expense of increased dark count rate.

Photon-counting receiver architectures have been proposed [1,26,43,169,170,393] and realized at Mbit/s [44] and ~Gbit/s [33,34] data rates with the best demonstrated coded RX sensitivities near 1 photon/bit (PPB), with potential for improvement to multiple-bits/photon sensitivities and greatly simplified processing due to the digital nature of the counting process.

However, these photon-counting RXs are presently limited to power-starved links with little or no background noise and <~Gbit/sec rates due to dark-count and reset-time constraints, which preclude their use with the sun in the field of view. Rates are also limited by the need of near-Shannon-limited-capacity coding [42,394] needed to avoid performance degradation due to dark counts, blocking losses, and background noise. These powerful FEC codes place additional burden on the RX electronics in terms of SWAP and bandwidth constraints[45]. Also, the suitability of the key detector technologies, Geiger-mode avalanche photodidode (GM-APD) arrays [32] and superconducting single photon counting detectors SSPDs [222] has not yet been established in the space environment. Si-based GM-APDs, for example, have shown sensitivity to radiation [395], and SSPDs require cooling to cryogenic temperatures, requiring a significant overhead in SWAP.

For ground-based receivers, however, where reliability and receiver SWAP are not a driving limitation, these technologies offer significant potential. Unlike coherent and preamplified RXs, these photon-counting detectors can receive multiple spatial modes and can efficiently collect signals distorted by the atmospheric channel without the need for wavefront correction. Furthermore, the net detection area can be efficiently scaled in distributed telescope arrays without the need to build large and costly telescopes [396].

The scalable-multimode-detection capability of photon-counting RXs comes with the expense of increased sensitivity to background noise, which is enhanced by the broad, several-hundred-nanometer spectral response of the detectors. However, the background can be mitigated with the use of single-polarization narrow-band low-loss high-rejection spectral filtering. Such filtering can be implemented using a cascade of loss-loss volume-Bragg-grating [397] and etalon filters with ~0.5 –1 dB insertion loss and noise-equivalent bandwidth of 0.01 to 0.1 nm over a spectral range exceeding 500 nm .[398]. Thus, many of the key elements needed for high-performance photon-counting receivers have been reduced to practice [44]. As the detector technologies mature, they may become the technology of choice for future ground and space-based receiver solutions.

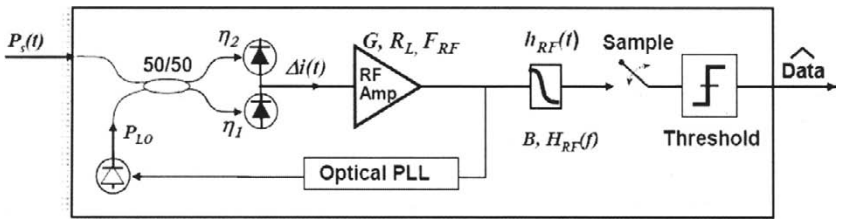


Fig. 46. Schematic of balanced coherent homodyne PSK receiver.

4.4. Coherent Homodyne Receivers

Coherent receivers are another means of boosting the shot noise relative to the thermal noise though coherent gain in the detection process. Homodyne detection of binary-PSK provides among the best theoretical RX sensitivity and is spectrally efficient. Coherent mode-matching requirements combined with matched electrical filtering also provide superior rejection of background noise over multi-mode receivers, which could enable, for example, improve communication performance with the sun in the field of view as noted in section 2.3.7. Information is carried in the phase $\phi_m(t)$ of the optical carrier and demodulated using a balanced receiver shown in Fig. 46.

The signal is received in one port of a 50/50 coupler and a local oscillator (LO) laser of the same frequency is injected into the other input port. The signal and LO mix at the two balanced square-law detectors, and generate a difference photocurrent that is given by [20,156,236]

$$\Delta i(t) = \left((\eta_1 - \eta_2) \left(\frac{P_{\rm LO}}{2} + \frac{P_s}{2} \right) + 2\eta_{\rm avg} \sqrt{P_s P_{\rm LO}} \cos\left(2\pi \Delta f t + \Delta\phi + \phi_m(t)\right) \right) \frac{q}{h\nu}$$

$$\xrightarrow[\substack{\eta_1=\eta_2=\eta \\ f_s=f_{\rm LO}}]{} \quad \frac{2\eta_{\rm avg}}{h\nu} \sqrt{P_s P_{\rm LO}} \cos\left(\Delta\phi_m(t)\right) \quad \text{[amps]}, \tag{96}$$

where $\eta_{\rm avg} = (\eta_1 + \eta_2)/2$ is the average detection (ideally equal to one), Δf and $\Delta\phi$ are the frequency and phase differences between LO and signal, that are set to zero with feedback from an optical phase-locked loop [153,156,166, 99], and $\phi_m(t)$ contains the (0 or π) phase modulation. The resulting SNR is given by [20, 156, 236]

$$\begin{aligned} {\rm SNR}_{\rm PSK} &= \frac{i_{\rm avg}^2}{\sigma_{\rm sh}^2 + \sigma_{\rm th}^2 + \sigma_{\rm LO}^2} \\ &= \frac{4R_{\rm avg}^2 P_s P_{\rm LO}}{2q(R_{\rm avg} P_{\rm LO})B + \sigma_{\rm th}^2 + (\eta_1 - \eta_2)\, P_{\rm LO}({\rm RIN})B}, \end{aligned} \tag{97}$$

where the shot noise term is due to the LO power and $R_{\rm avg} = \eta_{\rm avg} q/h\nu$. The local oscillator noise term $\sigma_{\rm LO}^2$ is the due to excess relative intensity noise (RIN) in the LO, which vanishes when the two detectors have the same detection efficiency [400,401]. In this case, the shot noise power can become the dominant noise source by increasing $P_{\rm LO}$. In the limit when $\sigma_{\rm sh}^2 \gg \sigma_{\rm th}^2$:

$$\mathrm{SNR_{PSK}} = 2\frac{\eta_{\mathrm{avg}} q P_s}{h\nu B} = 4\mathrm{SNR_{DD}}, \tag{98}$$

four times the SNR of ideal direct-detection. Even in the shot-noise limit, the SNR is proportional to the detection efficiency η_{avg}, meaning any detection losses, including spatial or polarization mode-mismatch, directly impact RX sensitivity.

However, in practice the high-sensitivity potential of optical PSK has not yet been realized, in part due to challenging component and laser linewidth requirements, the need for high-detection efficiencies, signal-LO polarization mismatch, and difficulties associated with phase-locking the local-oscillator [156,399]. In the Gbit/s regime the best reported uncoded PSK performance is ~35 PPB at ~6 Gbit/s and ~80 PPB at ~8 Gbps [158], providing little performance benefit over optically preamplified DPSK [24,25,133], which is WDM scalable and easier to implement (see Differential Phase Shift Keying (DPSK) in section 5.2). At lower rates < 1 Gbps, PSK performance is somewhat better, with uncoded RX sensitivities of 16 and 20 PPB at 4 and 565 Mbps, respectively [157,195].

Some of the reduction in PSK sensitivity at high data rates can be understood through the expression for SNR degradation in Eqs. (99) and (100). To the extent that the shot noise does not overwhelm all other sources of noise in the receiver, the SNR is degraded by

$$\Delta\gamma = 10\log_{10}\left(1 + \frac{\sigma_{\mathrm{th}}^2}{\sigma_{\mathrm{sh}}^2}\right) = 10\log_{10}\left(1 + \frac{i_{\mathrm{eq}}}{i_{\mathrm{avg}}}\right), \tag{99}$$

$$i_{\mathrm{eq}} = \frac{2k_B T}{qR_L} \quad [\mathrm{amps}], \tag{100}$$

where i_{eq} is the equivalent photocurrent needed to make $\sigma_{\mathrm{sh}}^2 = \sigma_{\mathrm{th}}^2$. For $\mathrm{R}_L = 50\Omega$ at 290K (room temperature), i_{eq} = 1 mA which corresponds to ~1 mW.

In order for shot noise to dominate thermal noise, large photocurrent or load resistance is required. For a maximum photocurrent of ~1 mA and $R_L = 50\Omega$ the shot noise and thermal noise are equal and the SNR is halved. This penalty can be reduced by boosting R_L, at the expensive of limiting the bandwidth due to residual capacitance in the TIA. Furthermore, as noted in the Direct detection—PIN section, the effective resistance in wide band TIAs is often limited by excess noise to ~300 Ω, which for a ~1 mA current reduces the RX sensitivity by > 1/ 2 dB.

Thus, the design of the front-end receiver electronics in coherent receivers remain critical [402]. At high data rates, achieving near-quantum-limited performance with coherent-homodyne PSK receivers is challenged by a combination of detection-efficiency, photocurrent, and transimpedance-gain limitations, whereas at lower data rates, laser-linewidth, RIN, and $1/f$ noise impose additional constraints [20]. In the following sections, we discuss the use of low-noise high-gain wide-bandwidth preamplification as a means of overcoming thermal noise and bandwidth limitations in the detection process, and reducing the impact of the front-end design on receiver performance.

4.5. Optically Preamplified Direct Detection

Preamplified RXs can directly leverage the field-tested heritage of telecom-type 1.55 μm technologies that are compatible with operation in the space environment and

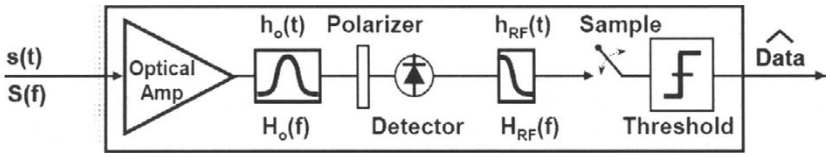


Fig. 47. Schematic of an optically preamplified receiver with input signal $s(t)$, optical and RF postdetction filters $h_o(t)$ and $h_{\mathrm{RF}}(t)$, respectively, optional polarizer, and sample and thresholding circuitry.

have demonstrated the best sensitivities at high data-rates (>∼Gbit/s) of 25–30 PPB for uncoded DPSK and 7–10 PPB with coding [24,25,133]. With the use of M-ary orthogonal modulation formats, RX sensitivities can approach the 1–2 PPB regime albeit with substantial bandwidth expansion [15,22]. In this case, the use of hybrid modulation formats including frequency, position, and polarization modulation along with WDM rate scaling can be used to access the many THz available in EDFAs at 1.55 μm and/or YDFAs at 1.06 μm, and overcome electronic bandwidth limitations. We address these modulation possibilities further in section 5.3.

In sections 3.5.3 and 3.5.5, we introduced some of the high-level characteristics of EDFA performance from the transmitter perspective. Here we provide additional detail on the impact of amplifier noise on SNR and receiver sensitivity. The noise processes in EDFAs and impact on receiver performance has been discussed extensively in the literature e.g., [20,36–38,76,77,126,336,340,403–406]. The basic optically preamplified receiver setup is shown in Fig. 47.

The signal field $s(t)$ with power $P_s(t) = |s(t)|^2$ is amplified by a factor of G which adds $\mathrm{GN_{sp}}$ noise photons per mode with noise power spectral density $N_{\mathrm{ASE}} = mh\nu GN_{\mathrm{sp}}$ [W/Hz]. Here, m is the number of polarization modes, and N_{sp} is the spontaneous emission factor of the amplifier (NF = $2N_{\mathrm{sp}}$). Out-of-band amplifier noise (ASE) is removed by optical filter $H_o(f)$, with field impulse response $h_o(t)$. The mean detected signal current is given by

$$i_{\mathrm{avg}}(t) = R_{\mathrm{A/W}} G\,|y(t)|^2 \underset{B_o \gg B_s}{\approx} R_{\mathrm{A/W}} G P_s(t), \tag{101}$$

where B_o and B_s are the noise equivalent bandwidths of the filter $H_o(f)$ and signal $S(f)$ spectra respectively, and

$$y(t) = s(t) * h_o(t) = \int_0^t s(t-\tau) h_o(\tau) d\tau \tag{102}$$

is the filtered signal field incident on the detector. Due to the square-law detection process, like-polarized signal and noise components mix, generating S × ASE and ASE × ASE noise terms in addition to shot-noise σ^2_{sh} and thermal noise σ^2_{th}. After the postdetection filter noise-terms are given by [344,345]

$$\sigma^2_{\mathrm{ASE}\times\mathrm{ASE}} = \int_{-\infty}^{\infty} d\tau_1 h_{\mathrm{RF}}(t-\tau_1) \int_{-\infty}^{\infty} d\tau_2 h_{\mathrm{RF}}(t-\tau_2) 2N^2_{\mathrm{ASE}} \left|R_o(\tau_1-\tau_2)\right|^2, \quad (103)$$

$$\sigma^2_{S\times\mathrm{ASE}}(t) = \mathrm{Re}\left\{ \int_{-\infty}^{\infty} d\tau_1 h_{\mathrm{RF}}(t-\tau_1) \times \int_{-\infty}^{\infty} d\tau_{21} h_{\mathrm{RF}}(t-\tau_2) 2N_{\mathrm{ASE}} y^*(\tau_1) R_o(\tau_1-\tau_2)\, y(\tau_2) \right\}, \quad (104)$$

$$\sigma^2_{\mathrm{sh}} = 2q(R_{\mathrm{A/W}} G P_s + mGN_{\mathrm{sp}} B_o + i_d) B_e, \quad (105)$$

$$\sigma^2_{\mathrm{th}} = (4k_B T/R_L) F_{\mathrm{RF}} B_e, \quad (106)$$

where $R_o(t) = h_o(t) * h_o(t)$ is the filter autocorrelation, $h_{\mathrm{RF}}(t)$ is the RF (electrical) filter impulse response, and B_e is the noise equivalent bandwidth of RF filter given in Eq. (39). For square optical filter $H_o(f)$ and $P_s(t)$ constant, Eqs. (103) and (104) reduce to [20, 37, 126]

$$\sigma^2_{S\times\mathrm{ASE}} \approx 4R^2_{\mathrm{A/W}} G P_s N_{\mathrm{ASE}} B_o = 4q^2 N_{\mathrm{sp}} \eta^2 G^2 \frac{P_s}{h\nu} B_o, \quad (107)$$

$$\sigma^2_{\mathrm{ASE}\times\mathrm{ASE}} \approx mR^2_{\mathrm{A/W}} N^2_{\mathrm{ASE}} B_e (2B_o - B_e) \approx m \left(N_{\mathrm{sp}} \eta q G B_o\right)^2, \quad (108)$$

and the expression for SNR becomes

$$\mathrm{SNR}_{\mathrm{Preamp}} = \frac{\left(P_s R_{\mathrm{A/W}} G\right)^2}{\sigma^2_{\mathrm{sh}} + \sigma^2_{\mathrm{th}} + \sigma^2_{S\times\mathrm{ASE}} + \sigma^2_{\mathrm{ASE}\times\mathrm{ASE}}} = \frac{\left(P_s \frac{\eta q}{h\nu}\right)^2}{\frac{\sigma^2_{\mathrm{sh}}}{G^2} + \frac{\sigma^2_{\mathrm{th}}}{G^2} + m\left(N_{\mathrm{sp}} \eta q B_o\right)^2 + 4q^2 N_{\mathrm{sp}} \eta^2 \frac{P_s}{h\nu} \frac{B_e}{B_o}}, \quad (109)$$

where we have assumed $B_o = \sim B_e$, a subject discussed further in section 5.1.2.

For large gain, optical bandwidth, and bit rate (e.g., $G > \sim$40 dB, $B_0 > 1$ GHz, and $R_b >$ 1Gbps), the σ^2_{sh} and σ^2_{th} terms are overwhelmed by the $\sigma^2_{S\times\mathrm{ASE}}$ and $\sigma^2_{\mathrm{ASE}\times\mathrm{ASE}}$ terms. For $m = 1$ and N_{sp}= 1, when

$$\frac{P_s}{h\nu} = \frac{B_o}{4}, \quad (110)$$

$\sigma^2_{S\times\mathrm{ASE}}$ and $\sigma^2_{\mathrm{ASE}\times\mathrm{ASE}}$ terms become equal [20]. Normalizing Eq. (110) by the bit rate this condition becomes

$$\frac{P_s}{h\nu R_b} = \mathrm{PPB} = \frac{B_o}{4R_b} \approx \frac{1}{4}. \quad (111)$$

For these filtering conditions ($B_o \approx B_e \approx R_b$), this implies that whenever the number of photons/bit $\gg$ 1/4, $\sigma^2_{S\times\mathrm{ASE}} \gg \sigma^2_{\mathrm{ASE}\times\mathrm{ASE}}$. In this case, Eq. (109) becomes

$$\mathrm{SNR}_{\mathrm{Preamp}} = \frac{\left(P_s R_{\mathrm{A/W}} G\right)^2}{\sigma^2_{S\times\mathrm{ASE}}} = \frac{P_s}{4h\nu B_e} = \frac{\mathrm{SNR}_{\mathrm{DD}}}{2}, \quad (112)$$

and we approach $\frac{1}{2}$ shot-noise limited direct-detection performance as expected due to the 3-dB noise figure of the optical amplifier. Note the performance in this limit

is not impacted by thermal noise or detection efficiency. For high-sensitivity applications, G $>\sim$40 dB is often necessary depending on the data rate and amplifier design, sometimes requiring more than two gain-stages and photocurrent $> \sim 1$ mA to overcome all other sources of noise. This makes high-current-capable high-speed 50 Ω photodetectors attractive [407,408]. Wideband transimpedance amplifiers can reduce this power requirement somewhat, but scaling the EDFA output power to the necessary levels is relatively straightforward and it is often challenging to obtain the combination of bandwidth, waveform fidelity, and dynamic-range of operation of the filtered EDFA-PIN photodetector combination with the addition of RF amplifiers. The saturating characteristics of the EDFA combined with adjustable pump power in the final gain stage can be useful for performing automatic gain control when operating over a wide range of power-levels or data rates.

The addition of a polarizer eliminates half the ASE (and other orthogonally polarized background [47,223]), leading to theoretical sensitivity improvements, which are typically about 0.5 dB in practice [41]. This requires some form of polarization control in the TX, RX, or both. Unlike coherent RXs, for preamplified RXs this is an option, and we note that control of polarization can increase sensitivity and extend modulation options as discussed in section 5.3.

Like coherent receivers, single-polarization optically preamplified receivers when implemented with matched optical filtering provide true-single-mode background rejection (see section 2.3.7), and therefore can perform well with the sun in the field of view (as long as coupling and tracking elements continue to function). In fact, since preamplified RX SNR is already degraded by ASE noise, the change in SNR due to solar background is less than that for coherent receivers.

The combination of low-noise, high-gain and bandwidth available in preamplified RXs relaxes many of the downstream component requirements in the receiver. Since EDFA gain and power in the preamplifier are inexpensive, they are a cost-effective means of achieving near quantum-limited receiver sensitivity over a wide range of data rates and modulation formats. Furthermore, rare-earth-doped EDFA and YDFA technologies have been shown to be robust in the space environment [198,204,380,409], making it the present technology of choice for future space-based applications. Methods of optimizing preamplified RX performance are discussed in the next section.

5. Performance and Implementation Considerations

For average-power-limited optical transmitters operating over the free-space channel, without dispersion or nonlinear impairments, the signaling waveform can be chosen arbitrarily without sacrificing transmitted power (subject to constraints discussed in section 3.5.5). This enables the selection of robust waveforms that are well suited for optimum communications performance and facilitates all-optical matching of the transmitter to the receiver [10]. With an abundance of available optical spectrum, waveforms can be chosen to provide nearly matched performance at multiple rates, diminished ISI, with little or no power penalty at the transmitter end. This approach is a paradigm shift from conventional techniques based on peak-power and bandwidth-limited RF designs, where there is always a significant trade between the choice of waveforms, filtering, and the power transmitted.

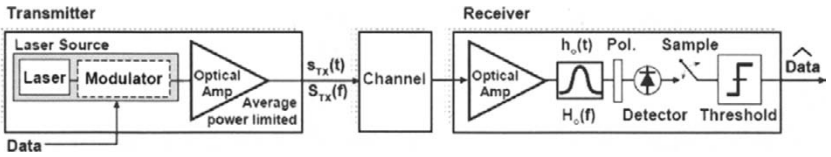


Fig. 48. Schematic of an average-power-limited MOPA TX with output field waveform $s_{\mathrm{TX}}(t)$ separated from a preamplified RX by an attenuating channel. A band-pass optical filter is followed by a polarizer, square-law detector, and postdetection sampling. The electrical response of the receiver is assumed to be relatively wideband, so that RF filtering is not explicitly included.

In the remainder of this section, we will combine well known concepts of matched filtering with the use of pulse shaping in the transmitter and optical filtering in the receiver to achieve robust high-sensitivity performance in optically preamplified receivers.

5.1. Waveform and Filtering Considerations

For high-rate and high-sensitivity applications, the MOPA TX and optically preamplified RX setup shown in Fig. 48 have demonstrated the best performance. It is well known that matched optical filtering yields the best performance for a signal corrupted by AWGN noise [156,410]. The SNR relative to the optimal matched filter condition is given by [18,120,411]

$$\Delta\gamma = \max_{t} \left\{ \frac{\left| \int_0^t s_{\mathrm{TX}}(\tau) h_o(t-\tau) d\tau \right|^2}{\int_{-\infty}^{\infty} |s_{\mathrm{TX}}(\tau)|^2 d\tau \int_{-\infty}^{\infty} |h_o(\tau)|^2 d\tau} \right\} \leq 1, \tag{113}$$

where $s_{\mathrm{TX}}(t)$ is the optical signal (field) waveform and $h_o(t)$ is the optical receiver filter (field) impulse response. The resulting SNR ($\Delta\gamma$) is maximized whenever

$$s(t) = h(-t) \quad \overset{\mathrm{FT}}{\longleftrightarrow} \quad S(f) = H^*(f). \tag{114}$$

The ratio in Eq. (113) represents the filtered signal waveform (power incident on the detector) as a function of time, normalized by the product of signal and filter waveform energies. The best SNR is achieved by sampling this waveform at its maximum.

The matching condition does not specify the waveform; rather it specifies the optimum relationship between the signal $s_{\mathrm{TX}}(t)$ and filter $h_o(t)$, that occurs when the filter impulse response is the time-reverse of the signal waveform. For waveforms that are symmetric in time, i.e., $s(t) = s(-t)$, the matching conditions reduce to

$$s_{\mathrm{TX}}(t) = h_o(t) \quad \overset{\mathrm{FT}}{\longleftrightarrow} \quad S_{\mathrm{TX}}(f) = H_o(f), \tag{115}$$

meaning the signal and filter impulse response have the same shape in both the time and frequency domains.

Early efforts towards optimizing preamplified RX sensitivity for high-rate optical systems focused on matching square non-return-to-zero (NRZ) signal waveforms[410,412] or optimizing performance with available Fabry–Perot (FP) optical filters .[128,404,413,414]. The NRZ waveform is particularly difficult to match as we show later in this section, and is prone to inter-symbol interference (ISI) penalties since the NRZ waveform energy is uniformly distributed between the symbol boundaries. This is especially problematic when received with the asymmetric exponentially decaying response of the FP filter, which requires a large optical bandwidth, $B_o = 3.7R_b$ and tight postdetection filtering in order to avoid significant ISI penalties [413]. The optimized FP bandwidth in this case is more than seven times larger than the best FP filter without ISI (see Table 11).

Matched optical filtering is desirable since it minimizes ASE or other out-of-band background optical noise, such as black-body radiation (see section 2.3.7) or inter-channel-interference (ICI) from adjacent WDM wavelengths, prior to conversion to the electrical domain. In this regard, all-optical matched filtering is desirable, from the standpoint of achieving optimum sensitivity [36,156] and a secondary benefit of improved spectral efficiency, since it enables tighter channel spacing [13].

The use of pulse-shaping in average-power-limited optical transmitters [10] improves RX sensitivity for two primary reasons. First it can reduce ISI penalties and second, it can facilitate the matching of the transmit waveform to realizable optical filters. This has led to high-sensitivity demonstrations (~2 dB from theory) of multi-rate variable-duty-cycle PPM [11,12], and RZ-OOK and RZ-DPSK [129] using FP optical filters.

Fortunately, there are many good waveforms that are easily generated (see section 3.5) and a wide selection of available filters such as single and multi-pass Fabry-Perot and dielectric filters, arrayed-waveguide (AWG), diffraction [415], and fiber Bragg (FBG) grating filters [133,416], and integrated waveguide filters [417]. Many of these can be customized to achieve a desired transfer function [418].

For the TX and RX in Fig. 48, the optimum signal-to-filter pulse-width-ratio β, defined as the ratio of signal FWHM to filter-impulse-response FWHM power waveforms, is calculated for a variety of waveform and filter type combinations summarized in the Matching Matrix in Table 11. For each waveform, the time-bandwidth product $(\Delta t_I \Delta f_I)$ is given for 3, 10, and 20 dB signal power levels. The SNR loss $(\Delta\gamma)$ is given in dB for the optimized β, without post-detection filtering. Note that ISI effects are assumed to be negligible, which is valid if these pulsed waveforms are of sufficiently low duty-cycle. No electrical filtering is included in this optimization, implying that the receiving electronics is relatively wide-band relative to the optical filter.

Also listed is the $\frac{1}{2}$ dB β-margin, which is the percent change from the optimum β that can be tolerated with less than $\frac{1}{2}$ dB reduction in SNR. This is an important measure of the SNR or matching robustness for the signal-filter pair and the sensitivity to deviations from the optimum. In this regard, the matched square waveform is a poor performer, a subject discussed further in section 5.1.2.

The example below illustrates how to use the matching matrix in Table 11.

Determine the optimum RX filter bandwidth for a 10-Gbps data stream formed by RZ-33% pulse carver.

From its name and as shown in Fig. 29, the FWHM of the RZ-33% is 33% of the period. For 10 Gbps, this would correspond to a 33-ps FWHM. Using a

Table 11. Matching matrix of 14 signaling waveform types (power pulse shapes) and a selection of four receiver filters. The SNR loss, $\Delta\gamma$ is calculated for the optimum $\beta \equiv \tau_o/\tau_s$, defined as the ratio of signal$_{\mathrm{FWHM}}$ to filter-impulse-response$_{\mathrm{FWHM}}$ (power) waveforms. Also listed for each filter type is the $\frac{1}{2}$ dB β margin, which is the percent deviation in β with $< \frac{1}{2}$ dB SNR change.

Time Domain Signal Power, $s_i(t)$ / Filter type: $h_i(t)$, Power Impulse Response				Square $h_i(t) = \mathrm{rect}(t/\tau)$			FP $h_i(t) = \exp(-2t/\tau)$			FP^2 $h_i(t) = (t)^2\exp(-2t/\tau)$			Gaussian $h_i(t) = \exp(-t^2/\sigma^2)$		
Waveform Type	$\Delta t_i\Delta f_i$ 3dB	10dB	20dB	$\Delta\gamma$ [dB]	Optim. β	1/2 dB Margin %	$\Delta\gamma$ [dB]	Optim. β	1/2 dB Margin %	$\Delta\gamma$ [dB]	Optim. β	1/2 dB Margin %	$\Delta\gamma$ [dB]	Optim. β	1/2 dB Margin %
Gaussian, $\exp(-t^2/\sigma^2)$	0.44	1.47	2.9	0.5	0.60	50	1	2.3	95	0.3	1.0	70	0.0	1	65
RZ-33% (~Gaussian)	0.49	1.54	2.7	0.5	0.60	55	1	2.2	100	0.3	1.1	70	0.0	1.1	60
RZ-40% (~Gaussian)	0.55	1.57	2.6	0.4	0.70	45	1	2.7	80	0.3	1.1	70	0.0	1.1	65
RZ-50%	0.66	1.62	4.0	0.3	0.80	35	1	2.9	90	0.3	1.3	70	0.1	1.4	50
RZ-53% (Raised Cos)	0.63	1.62	2.4	0.3	0.70	45	1	2.5	110	0.3	1.2	75	0.1	1.3	50
RZ-66⅔ (CSRZ)	0.73	1.62	2.9	0.2	0.80	40	1	3.1	85	0.3	1.5	55	0.2	1.3	66
$\mathrm{Sech}^2(t/\tau)$ (Soliton)	0.31	1.34	3.6	0.7	0.60	50	1	2	100	0.3	0.9	75	0.0	0.9	60
Gaussian^2 $\exp(-t^2/s^2)^2$	0.68	1.58	3.4	0.3	0.80	35	1	2.9	90	0.3	1.5	55	0.1	1.4	50
Gaussian^2 $\exp(-t^2/s^2)^4$	0.80	1.56	4.0	0.1	0.90	25	0.9	3.5	70	0.4	1.7	50	0.2	1.7	45
Square $\mathrm{rect}(t/\tau)$	0.89	1.49	3.6	0	1	10	0.9	3.5	80	0.6	1.8	45	0.5	1.7	50
FP^4 $(t^3/6)^2\exp(-2t/\tau)$	0.40	1.51	2.7	0.5	0.59	50	1.5	2.5	90	0.7	1.0	80	0.1	1	65
FP^3 $(t^2/2)^2\exp(-2t/\tau)$	0.39	1.51	1.5	0.5	0.25	50	1.6	2.4	100	0.8	1.0	80	0.2	1	70
FP^2 $(t)^2\exp(-2t/\tau)$	0.35	1.51	4.6	0.6	8.33	50	1.8	2.2	120	1.0	1.0	80	0.3	1	70
FP $\exp(-2t/\tau)$	0.11	1.11	7.3	0.9	0.01	80	2.7	1	130	1.8	0.4	105	1.0	0.4	105

matched Gaussian filter, the optimum filter impulse response ($\beta = 1.1$) has a FWHM of ~30 ps. From the time-bandwidth product, this corresponds to a 14.7 GHz filter, which provides nearly optimum SNR performance (~0 dB SNR loss) and can accommodate up to a ~60% bandwidth mismatch with only 0.5 dB added penalty.

The relaxed tolerances in this example can be used to assist in manufacturing yields and reduce component cost. The flexibility in filter bandwidth permits, for example, widening the filter to mitigate ISI or narrowing the filter to diminish ICI with little or no SNR penalty.

5.1.1. Symmetric Filtering

Pulse-shaping at the transmitter can be exploited to implement near optimum communications waveforms with non-ideal components (modulators, drivers, and receiver filters). Consider the TX in Fig. 37 with a wide-band optical source generating short (impulse-like) pulses in time followed by a high-gain APL amplifier with optical filtering. Passive optical filtering can be used to perform TX pulse shaping such that $S_{\mathrm{TX}}(f) \approx H_{\mathrm{net}}(f)$. A *symmetrically filtered* RX has the same net optical filtering,

$H_{\text{net}}(f)$, and if the filter time-response is also symmetric, match filtered performance can be achieved. For filters with an asymmetric response, such as the Fabry-Perot, with multi-pass or cascaded filtering, the response can be made progressively more symmetric (and Gaussian-like). In this manner, near optimum communications performance can achieved by cascading readily available filters which would otherwise yield sub-optimal performance when used individually [10]. Mathematically, this is equivalent to convolving a function with itself, which qualitatively tends to spread a function making it more symmetric.

The performance benefits of using symmetric TX and RX filtering for cascaded asymmetrical filters can be observed for the case of the commonly used Fabry–Perot filter (FP^n) listed in Table 11. Without additional filtering, a first-order FP incurs a 2.7-dB SNR loss. For the second order (FP^2), which can be implemented by a double-pass through a single FP, the SNR falls about 1 dB from optimum. A three-pass configuration yields a mere ~0.6 dB deviation, and as expected, departure from the optimum continues to diminish as additional iterative passes are made, and the time domain signal becomes increasingly symmetric. Note that along with increased symmetry, the time and frequency domain characteristics of FP^n start to resemble those of a Gaussian, the subject of the next section.

5.1.2. Gaussian Waveforms and Matched Optical Filtering

Combining *symmetric filtering* with the use of matched Gaussian-like waveforms [14], nearly optimal performance can be achieved with relaxed filter tolerances [13,133]. Gaussian waveforms are relatively easy to generate (see Pulsed Waveform Generation in section 3.5.2) and Gaussian filters can be realized with filter technologies such as diffraction gratings [13,415], AWGs [419,420], apodized FBGs [133,421], and cascaded filters discussed in the previous section. The Gaussian probability distribution function is known as a minimum-uncertainty-packet in quantum-mechanics, providing the minimum variance in two conjugate variables (e.g., position and momentum, or arrival time and energy of a particle). As a signaling waveform, Gaussian waveforms have a narrow time-bandwidth product, $\Delta f_I \Delta t_I = 0.44$ (at the 3-dB power point), making them well suited for both TDM and WDM communications. The Gaussian function is part of a class of functions known as "Self-characteristic functions" [422], which have a Fourier transform of the same form, i.e., a Gaussian time-domain waveform also has a Gaussian spectrum. Another well known optical waveform which shares this property is the soliton, which is Gaussian-like and also has a narrow time-bandwidth product (see Table 11). The soliton waveform is considered for use in ultra-high-speed fiber-optic links since the waveform maintains its shape due to a balance of dispersion and nonlinearities [111,236,423,424]. In this regard, fiber-based communications using solitons is similar to communication over the dispersionless free-space channel.

5.1.2.1. Relaxed Filter Tolerances

In contrast with *self-characteristic minimum-uncertainty waveforms*, the square-NRZ signaling waveform, which is completely contained in the time-domain, has a frequency domain spectrum with infinite extent, making it a poor choice for use in dense WDM systems. Furthermore, as noted earlier, the square waveform is susceptible to inter-symbol interference (ISI) penalties and requires tight filter tolerances to avoid

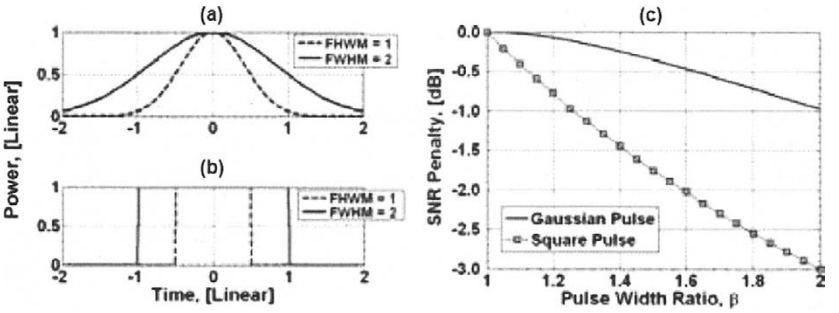


Fig. 49. Signal $s_i(t) = |s(t)|^2$ in solid line and filter impulse response $h_i(t) = |h_f(t)|^2$ in dashed line for Gaussian and square waveforms in (a) and (b) above. The pulse width ratio β is defined as the signal FWHM to filter-impulse-response FWHM. For the waveforms in (a) and (b), β =2. (c) Shows the reduction in SNR as a function of β for matched Gaussian and square filters.

SNR penalties as illustrated for the case of Gaussian and square matched filters in Fig. 49.

In Fig. 49(a), a wider Gaussian signal waveform (solid) and filter impulse response (dashed) are shown. The full-width-half-max pulse width ratio for the signal and filter response is defined as β, where the matched filter occurs for $\beta = 1$. Similarly, a wide square-signal-waveform and filter-response are shown in Fig. 49(b). For both waveform types illustrated, $\beta = 2$. Excluding ISI effects, the SNR penalty ($\Delta\gamma$) is calculated using Eq. (113) and plotted for Gaussian and Square waveform types in Fig. 49(c). While both Gaussian and square waveforms have no penalty for the matched condition (β = 0), the square degrades much more quickly from waveform-filter mismatch. For a 60% mismatch (β = 1.6), the Gaussian incurs ~0.5 dB, whereas the square waveform degrades by ~2 dB. For SNR loss $< \frac{1}{2}$ dB the square filter bandwidth must be within 10% of the signal bandwidth, effectively six $\times$ more or sensitive to mismatch than the Gaussian at the $\frac{1}{2}$ dB point. For β = 2, the square waveform SNR degrades by 2 dB more than the Gaussian, illustrating that the Gaussian waveform is more robust and easier to match than the square.

In an attempt to capture both SNR magnitude and its matching tolerance for a particular waveform combination in a single term, we introduce $\Delta\gamma_{\text{eff}}$, defined as the average SNR for $\beta_{\text{optimum}} < \beta < 1.6$, i.e.,

$$\Delta\gamma_{\text{eff}} = \frac{\int_{\beta_{\text{optimum}}}^{1.6\beta_{\text{optimum}}} \Delta\gamma\,(\beta)\,d\beta}{0.6\beta_{\text{optimum}}}. \tag{116}$$

$\Delta\gamma$ is listed in Table 12 and displayed in Fig. 50 for several signaling waveforms and filter types, sorted by $\Delta\gamma_{\text{eff}}$, and shows Gaussian-like properties and better effective SNR.

Table 12. Effective Matching-matrix. Effective SNR loss, $\Delta\gamma_{\text{eff}}$ resulting from the overlap of several signaling waveform types (power pulse shapes) with a selection of RX optical filters, sorted by $\Delta\gamma_{\text{eff}}$ from Eq. (116). ISI effects are neglected.

Signal	Receiver Filter, $h_i(t)$							
Waveform, $s_i(t)$	Gauss	FP^4	FP^3	$Gaus^{(2)}$	$FP^{(2)}$	$Gaus^{(4)}$	Square	FP
Gaussian	0.2	0.3	0.3	0.3	0.4	0.5	0.8	1.1
RZ-33%	0.2	0.4	0.4	0.2	0.5	0.4	0.7	1.1
RZ-40%	0.2	0.4	0.4	0.4	0.5	0.4	0.7	1.1
Soliton	0.2	0.4	0.4	0.4	0.5	0.5	0.8	1
RZ-50%	0.3	0.4	0.4	0.4	0.5	0.5	0.7	1.1
RZ-53%	0.3	0.4	0.4	0.2	0.5	0.4	0.6	1.1
RZ-66%	0.3	0.6	0.6	0.4	0.6	0.5	0.6	1.1
$SuperGauss^{(2)}$	0.3	0.4	0.4	0.3	0.5	0.5	0.7	1
$FP^{(4)}$	0.3	0.7	0.7	0.4	0.8	0.5	0.8	1.6
$FP^{(3)}$	0.3	0.7	0.7	0.4	0.9	0.6	0.8	1.7
$FP^{(2)}$	0.4	0.9	0.9	0.5	1.1	0.6	0.7	1.8
$SuperGauss^{(4)}$	0.5	0.5	0.5	0.4	0.7	0.5	0.7	1.2
Square	0.7	0.7	0.7	0.7	0.9	1.2	1	1
FP	1.0	1.6	1.6	1.0	1.9	1.0	1	2.7

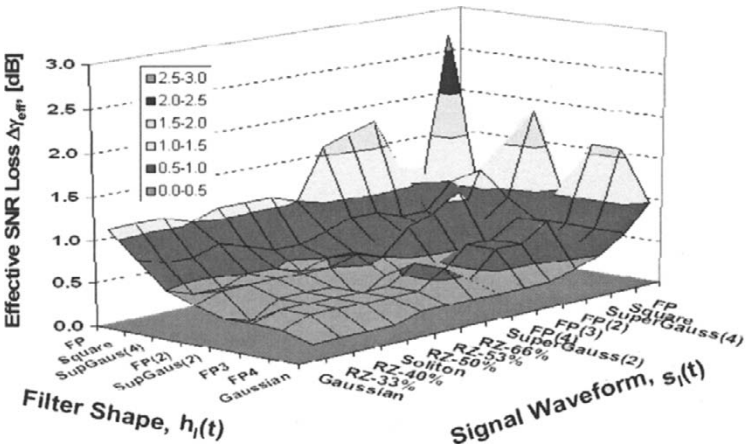


Fig. 50. Effective SNR loss for a variety of signal and filter shapes sorted $\Delta\gamma_{\text{eff}}$. The more symmetric, Gaussian-like waveforms yield better SNR performance, whereas the sharp-edged square and asymmetric FP filters incur the biggest losses.

5.1.2.2. Reduced Sensitivity to Timing Jitter

The SNR analysis in section 5.1 assumes that the received waveform is sampled at its maximum. However, deviations from the optimum sample point due to noise in the clock-recovery processes [236,419,425] or other sources of timing-jitter [426] reduce

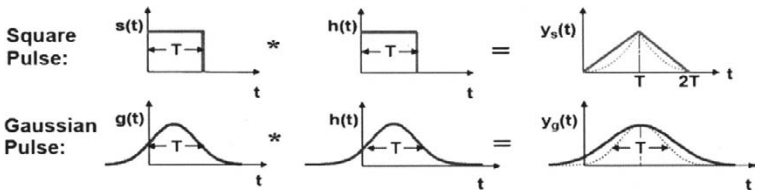


Fig. 51. Convolution of signal and matched filter field-impulse-response $h(t)$ for square (top) and Gaussian waveforms (bottom) of the same pulse-width. The filtered signal field $y(t)$ is the solid line and signal power (e.g., photocurrent after square law detection) is in the dotted line [13].

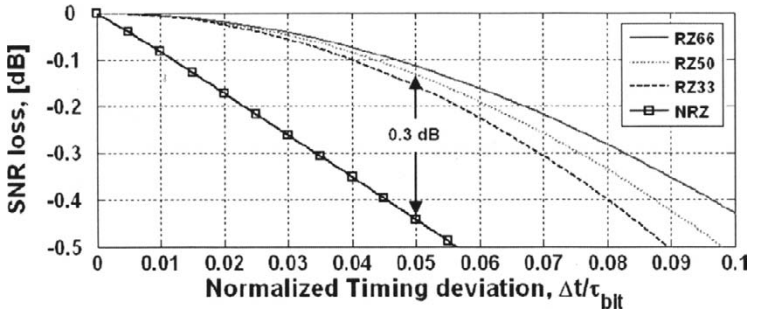


Fig. 52. SNR sensitivity to timing jitter for matched Gaussian-like and square-NRZ waveforms (excluding ISI) [13].

the SNR at the thresholding circuitry which degrades RX performance. Such effects are dependent on waveform shape as shown in Figs. 51 and 52, where, square $s(t)$ and Gaussian $g(t)$ field waveforms are received by their respective matched filters $h(t)$. The output from the matched optical filter $y(t) = s(t) * h(t)$, with detected photocurrent $i(t) = \eta|y(t)|^2$, indicated by the dotted-lines in Fig. 51.

The convolution of the square waveform and its matched filter yields the triangular waveform, which is subsequently narrowed by the square-law detection processes. The best SNR is achieved by sampling photocurrent at the peak. In this case, the peak is a relatively narrow target, which tends to magnify the any SNR penalties due to fluctuations in the sample time.

For the Gaussian waveform, the matched-filtered output is broadened 41% by the convolution process, but the waveform remains Gaussian. After square-law detection, the received photocurrent still has a Gaussian shape, but the pulse width is reduced by $\sqrt{2}$, so that it has the same pulse width and shape as the incident signal field. Note that the peak sampling point for the Gaussian is a much broader target than that of the square. The resulting SNR degradation ($\Delta\gamma$) as a function of time deviation from the optimum sampling point is shown in Fig. 52.

The sensitivity of matched-Gaussian-like waveforms to timing errors and sampling duration is significantly better than that of matched-square waveforms as shown above. This can reduce the impact of timing-errors from a variety of sources (e.g., Gordon-Haus timing-jitter [426]), and relax tolerances for sampling speed, accuracy, and stability of clock-recovery and detection hardware. This is especially important for M-ary PPM and high-rate communications which are more sensitive to the effects of timing jitter. For example, at 40 Gbps the timing precision required for less than 0.1 dB penalty is ~1 ps for 33% RZ waveforms. Despite being three-times wider, the square-NRZ waveform is over three-times more sensitive than 33% RZ to timing jitter, requiring $\Delta t <$ ~300 fs for < 0.1 dB penalty. For only $\Delta t =$ ~1 ps SNR is reduced by ~0.35 dB, excluding additional ISI penalties.

5.1.2.3. Combined Optical and Postdetection Filtering

For systems where electronic bandwidth is a limitation, both optical and RF postdetection filtering can be adjusted for better overall RX performance. Practically, obtaining RF bandwidths wide relative to the data rate is not always an option, but SNR loss ($\Delta\gamma$) can be avoided to a large extent, by increasing the optical filter bandwidth accordingly. As noted earlier, a similar trade between optical and RF bandwidths can be used to compensate when suitable narrow (matched) optical filtering is unavailable [345,346,404].

The SNR penalty as a function of optical and electrical bandwidths [344,345,404–406,427,428] has been evaluated for a variety of waveforms and filter shapes, often numerically and for the case of optical filters with bandwidths wide relative to the signal spectrum. Below, we evaluate Gaussian-like signal-waveforms and Gaussian-optical- and postdetection-RX filters for which the impact of deviations from the matched condition on SNR can be estimated analytically [429].

The SNR trade space for optical- and RF- filter time-response pulse widths (τ_o and $\tau_{\rm RF}$) relative to the optical signal pulse width (τ_s) is shown in Fig. 53. We use the signal-pulsewidth for normalization rather than the bit-duration ($\tau_{\rm bit}$) since matching is independent of data rate. The bit period, for instance, can be adjusted to vary the data rate or lengthened to reduce the effects of ISI if needed.

As expected, for wide RF filtering ($\tau_{\rm RF}/\tau_s \ll 1$), optimum performance is achieved for the matched condition, when $\beta = \tau_o/\tau_s = 1$, but extends over a large sweetspot surrounding the dashed *centerline* ($\tau_{\rm RF} \approx 3(\tau_s - \tau_o)$). As noted earlier, the Gaussian waveform relaxes the optical filter tolerances needed for nearly-matched performance, i.e., <~0.1 to 0.2 dB filtering penalty. Similar benefits are seen for Gaussian electrical filters.

If the $\tau_s/\tau_{\rm bit}$ ratio is sufficiently short, ISI penalties can be made negligibly small, and the analysis in Fig. 52 is broadly applicable. Otherwise, uncorrected ISI penalties start to degrade performance for longer τ_o and $\tau_{\rm RF}$ (upper right quadrant), compressing the SNR contours towards shorter pulse widths (lower left quadrant) as $\tau_s/\tau_{\rm bit}$ increases. Staying to the lower-left of the *centerline,* avoids ISI penalties as shown for the displayed RZ33, RZ40, and RZ50 data points. The SNR-optimized pulsewidth coordinates for Gaussian filter and RZ signals shown in Fig. 53 are linked to the data rate and bit duration is in Table 13.

For all three waveform types, optical and electrical bandwidths can be chosen for near-quantum-limited performance well within the 0.1-dB SNR sweetspot in Fig. 53,

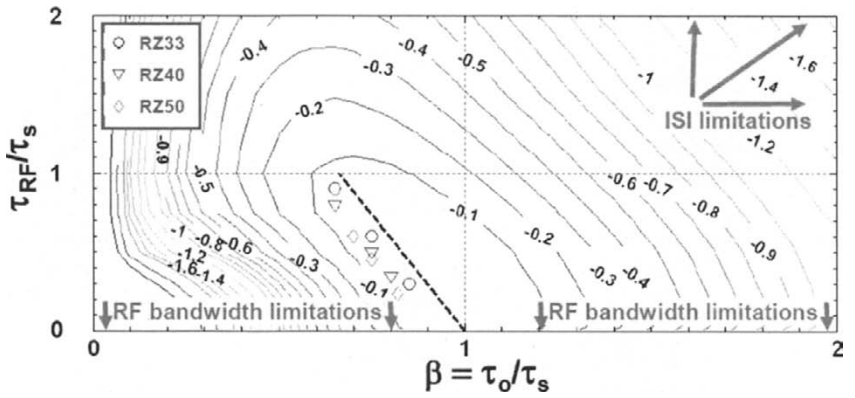


Fig. 53. Estimated SNR impact ($\Delta\gamma$) on a preamplified receiver with combined optical and RF Gaussian filtering and low photon/bit Gaussian signals. Contours represent lines of constant SNR (in dB) as a function of both RF and optical filter-response-pulse-widths relative to the signal power pulse width τ_s. ISI effects are not included in this calculation since they are dependent on both waveform shape and duty-cycle. However, ISI can be neglected for small τ_o and $\tau_{\rm RF}$ (the lower-right quadrant of the filter space) and for low-duty-cycle waveforms, i.e., $\tau_s \ll \tau_{\rm bit}$. Detailed parameters for RZ33, RZ40, and RZ50 points are given in Table 13. Simulations based on SNR analysis by J. Shapiro [429].

Table 13. Relative pulse width and bandwidth for data points shown in the optimized SNR region of Fig. 53. The bit duration $\tau_b = 1/R = 1$ for all entries. Optical parameters are in terms of FWHM intensity.

	Optical Signal			Optical Filter				RF Filter			
Signal Type	τ_s/τ_{bit}	$\tau_s\Delta f_s$	BW $\Delta f_s/R$	τ_o/τ_{bit}	$\tau_o/\Delta t_s$	$\tau_o\Delta f_o$	BW τ_o/R	τ_{RF}/τ_{bit}	$\tau_{RF}/\Delta t_s$	$\tau_{RF}\Delta f_{RF}$	BW $\Delta f_c/R$
RZ33	0.33	0.49	1.5	0.21	0.65	0.44	2.1	0.30	0.90	0.22	0.7
RZ33	0.33	0.49	1.5	0.25	0.75	0.44	1.8	0.20	0.60	0.22	1.1
RZ33	0.33	0.49	1.5	0.28	0.85	0.44	1.6	0.10	0.30	0.22	2.2
RZ40	0.4	0.55	1.4	0.26	0.65	0.44	1.7	0.32	0.80	0.22	0.7
RZ40	0.4	0.55	1.4	0.30	0.75	0.44	1.5	0.20	0.50	0.22	1.1
RZ40	0.4	0.55	1.4	0.32	0.80	0.44	1.4	0.14	0.35	0.22	1.6
RZ50	0.5	0.66	1.3	0.35	0.70	0.44	1.3	0.30	0.60	0.22	0.7
RZ50	0.5	0.66	1.3	0.38	0.75	0.44	1.2	0.23	0.45	0.22	1.0
RZ50	0.5	0.66	1.3	0.41	0.82	0.44	1.1	0.12	0.24	0.22	1.8

with *realizable* optical filter bandwidths $\sim$1 to 2 times the data rate (R) and RF filter bandwidths 0.7 to 2.2 R. When limited RF bandwidth drives the design, the optical filter can be chosen to reduce the RF bandwidth to $< 0.7R$ with little or no SNR penalty. The breadth of matching tolerance is also noticeable. For an optical filter bandwidth of about $1.3R$, nearly optimal performance can be achieved for all three waveforms with RF bandwidth between R and $2R$.

We note that with nearly matched optical filters, a single-polarization preamplified RX achieves true single-mode spatio-temporal filtering, which maximizes rejection to background noise, but also modifies the ASE noise statistics [122–125]. While accurate analysis of single-mode Bose-Einstein and noncentral-negative-binomial distributions for '0' and '1' ASE statistics, respectively, may influence theoretical predictions of optimum threshold and receiver sensitivity, they are unlikely to impact the results above. Near-quantum-limited demonstrations using this class of Gaussian waveforms and filters for OOK [13] and DPSK [133] modulation lend additional support to the analysis above.

5.1.3. Optimized Multi-Rate Transceivers

In many free-space applications, the flexibility to operate over a wide range of data rates is desirable since it provides a straightforward means of adjusting the transmitted signal-to-noise ratio, providing increased margin as needed. This can be used to accommodate varying link losses, due to changes in link distance or atmospheric channel effects, or compensate for performance degradation and extend the useful lifetime of the system. Especially in space-based systems, such capabilities should not incur increased complexity, SWAP, or reduction in RX sensitivity.

For systems that employ APL amplifiers (see section 3.5.3), multi-rate communications with nearly matched filtering can be efficiently implemented[11,15]. As discussed in the sections 5.1.1 and 5.1.2, the transmitter pulse shape can be adjusted so that it is well matched to the receiver without sacrificing transmitted energy-per-bit. This can be extended to multiple bit-rates simply by lowering the duty-cycle or average repetition rate [10–12,15], causing the peak output power to increase while the average saturated output power remains constant (see Figs. 35 and 36). Lowering the bit-rate in this manner while maintaining the same pulse shape and width (i.e., reducing the duty cycle), increasing the transmitted energy-per-bit thereby improving the error rate while maintaining optimum performance.

This *variable-duty-cycle* multi-rate approach has been demonstrated for binary-PPM at 51 to 1244 Mbps [11,12] with performance ~2 dB from quantum-limited theory at all rates. Subsequent M-PPM demonstrations at 2.5 Gslots/s [15] similarly yielded near-quantum-limited multi-rate performance but with improved sensitivities (see Fig. 61), spanning a range exceeding 20 dB with only a 12-dB change in data rate. Such features are particularly useful for providing bandwidth on demand or fall-back modes for communications over a noisy or uncertain channel since SNR can be improved by simply lowering the duty cycle (bit-rate) without additional penalties or hardware. It is particularly well suited for free space laser communications since it allows for graceful degradation without significant cost or complexity. Effects that can limit the range of practical duty cycles are discussed in section 3.5.5.

Variable duty cycle techniques can also be used to simplify multi-rate DPSK receivers [10,134] discussed in the next section. With a reconfigurable DI (see section 5.2.4) that can adjust the delay to accommodate the bit rate, a single filter design can be used to achieve nearly-quantum-limited receiver sensitivity at all rates. Alternatively, for harmonically related data rates, a fixed DI sized to the lowest rate can be used [134] along with appropriate differential precoding. In this case, the same DI can be used to simultaneously demodulate multiple-rate WDM-DPSK signals (see section 5.2.3), providing both rate-flexibility and WDM scalability [136].

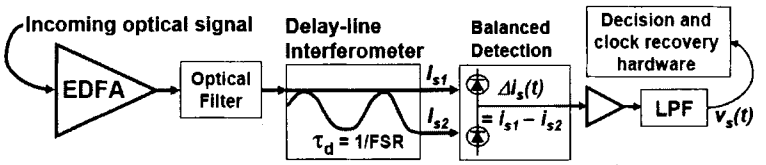


Fig. 54. Typical optically preamplified DPSK receiver.

5.2. *Differential Phase Shift Keying (DPSK)*

Optical differential-phase-shift-keying (DPSK) modulation was introduced in section 2.2.1 as a means of improving sensitivity over commonly used OOK. DPSK theoretically can offer among the best combined spectral and photon efficiency without requiring a coherent local-oscillator-based receiver. This has generated considerable attention by both the FSO community and the telecom industry, and led to the most sensitive high-rate demonstrations [24,25,41,128–131,133,172]. However the sensitivity benefits of DPSK come at the cost of increased complexity over OOK, requiring a phase modulator and differential precoding in the TX, and an optical delay-line interferometer (DI) and balanced detection in the RX as shown Fig. 54. The remaining DPSK RX elements including subsequent electrical amplification and filtering are similar to OOK RXs.

In practice, the complexities associated with DPSK make it challenging to achieve the potential 3-dB sensitivity advantage over OOK. The most significant performance penalties result from residual chirp in the transmitter waveforms [132,142] discussed section 3.5.2, and alignment of DI and the signal wavelength [144,146,430], discussed in the section 5.2.2. Smaller degradations in DPSK performance arise from non-ideal interferometer delay [135,146,31], extinction ratio, and detector imbalance in amplitude and timing [132,145]. The reader is directed to an excellent tutorial on long-haul fiber-based DPSK systems by Gnauck and Winzer in [132] and other references in this section for further details and analysis of DPSK performance.

Despite these possible degradations, DPSK receiver sensitivities closely approaching quantum-limited theory can be achieved [133]. However, it is important to realize that while implementation penalties may individually be considered negligible, the combined impact of several of these effects could be prohibitive. Unless the improvement in RX sensitivity can be realized with long-term reliability, DPSK provides little value for free-space applications, in contrast with long-haul fiber-based DPSK applications, which benefit from the reduction of nonlinear impairments that result from lower peak power.

In the DPSK receiver, the DI enables a comparison of the optical phase by splitting the optical signal field and recombining the two polarization-aligned components with a relative time-delay τ_d. At the output coupler, the two overlapping field components interfere constructively or destructively depending on the relative phase of the differentially encoded bits, with the resulting output power proportional to the product of the two fields, being directed to the first or second output port. For complete interference, the incident fields need to have the same amplitude, and a delay that is a positive integer multiple n of the bit-period τ_{bit} [134,135], i.e.,

$$\tau_d \equiv \frac{1}{\mathrm{FSR}} = n\tau_{\mathrm{bit}}, \tag{117}$$

where FSR is the DI free spectral range. Since the data rate $R_b = 1/\tau_{\mathrm{bit}}$, it is ideally equal to an integer multiple of the FSR, resulting in a delay-error $\Delta\tau = |\tau_d - \tau_{\mathrm{bit}}| = 0$. The associated delay-error or rate-to-FSR mismatch penalty is limited to less ~0.3 dB for $\Delta\tau < 0.05\tau_{\mathrm{bit}}$[146,431]. However, for larger errors, the SNR degradation is waveform and filter dependent [135]. For receivers optimized for 33%-RZ waveforms, 10% and 14% delay-errors lead to ~0.5 dB and ~1 dB SNR penalties, respectively. While manufacturers typically specify a ~1% tolerance for the DI delay accuracy, in practice, larger delay-errors can result from channel-rate changes in deployed systems to accommodate different FEC overhead [132] or to simplify WDM-DPSK receiver implementation while maintaining compliance with existing wavelength and rate standards [149,151,432,433].

For stability, size, and performance reasons, τ_d is typically chosen to equal only one bit period. This relaxes signal coherence requirements [264], minimizes frequency-alignment penalties, and simplifies TX precoding, which can be implemented with an OR gate to logically combine DATA and CLOCK inputs to drive a T-flip-flop [140,434–436], causing the transmitted phase to change whenever the DATA is a '1'. However, for multi-rate and multi-channel DPSK applications discussed later, it can be desirable to use a multiple-bit delay to accommodate simplified implementations that provide rate-flexibility and scalability with penalty-free performance that is compatible with existing channel-rate and channel-spacing standards [135, 136].

The DI outputs are received with balanced detection, which generates a signal $\Delta i(t)$ that is proportional to the power difference between the two output arms of the DI. This can be implemented by directly subtracting the photocurrent output from discrete [128] or integrated balanced photodetectors [25,408,437], or using subsequent differential electronics to generate the difference signal [25,130,159], all of which are commercially available with bandwidths exceeding 40 GHz. Balanced detection provides an implementation benefit over single detector RXs since the resulting photocurrent deviation between "0" and "1" symbols is twice that of single-detector receivers, making it easier to directly overcome the noise threshold in the subsequent decision circuitry without additional RF amplifiers.

5.2.1. DPSK Wavelength Alignment Considerations

Of the additional RX elements needed for DPSK, the delay-line interferometer often imposes the most demanding requirements. For good performance the DI is preferably polarization insensitive and the two arms must be stable to a small fraction of a wavelength. This requires careful, thermo-mechanical packaging and/or active stabilization [432,438–440], adding to DI size, weight, power, and cost. The DI is most commonly implemented with an asymmetric Mach–Zehnder design (see section 3.5.2 for detailed description), although Michelson[11] [432,441,442] and Sagnac-interferometer-based

[11] Albert Abraham Michelson, 1852–1931, invented and built what has become known as the Michelson-interferometer, used in accurate measurements of length and the speed of light. Conducted the Michelson-Morley experiment (1887) with E. W. Morley, which attempted to detect the expected difference in the speed of light caused by the motion of the earth. Michelson was awarded the Nobel Prize in physics in 1907.

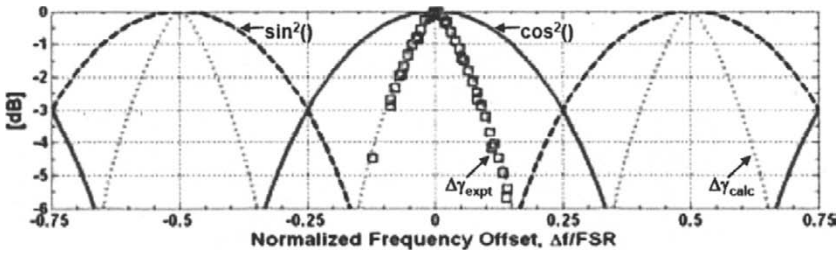


Fig. 55. Periodic transfer function (fringes) of a DI: $\sin^2()$ and $\cos^2()$ outputs; and calculated ($\Delta\gamma_{\mathrm{calc}}$) and measured ($\Delta\gamma_{\mathrm{expt}}$) SNR penalties at 10^{-9} BER as a function of carrier frequency offset Δf, normalized by the DI FSR (assuming polarity correction) [443].

designs [440] with an equivalent transfer function [see Eqs. (57)–(60)] may be used as well (see section 5.2.4 on Reconfigurable DPSK Demodulators below).

In contrast to the Mach–Zehnder interferometer-based modulators discussed in section 3.5.2, where a short sub-picosecond static time delay is desirable, the relatively long delay of at least one bit-period needed for DPSK demodulation causes an asymmetry that makes stabilization more difficult. Small changes in DI temperature, for example, even if uniform, generally lead to a phase difference between the arms that shift the spectral alignment of the interferometer, resulting in significant SNR degradation.

The ideal intensity transfer functions of the two interferometer output arms are periodic and complementary, given by

$$\begin{aligned} I_1 &= |E_1|^2 = \cos^2\left(\pi \tfrac{\Delta f}{\mathrm{FSR}} + \tfrac{\Delta\phi}{2}\right), \\ I_2 &= |E_2|^2 = \sin^2\left(\pi \tfrac{\Delta f}{\mathrm{FSR}} + \tfrac{\Delta\phi}{2}\right), \end{aligned} \tag{118}$$

where $\Delta\phi$ is a measure of the relative optical phase between the internal interferometer arms and is typically controlled to maximize the signal interference in one or both of the output arms. When the bias $\Delta\phi = 0$, the Δf term represents the frequency deviation from optimum alignment of the incoming signal to the interferometer. The SNR degradation associated with Δf (frequency offset error) forms periodic SNR fringes at even multiples of FSR/2 as shown in Fig. 55. The received signal is reduced by an amount proportional to the difference of signal intensities at the output arms. The received signal is maximum for $\Delta f = 0$ and goes to zero as Δf approaches odd multiples of FSR/4, when the outputs of the two arms are equal. Beyond this point, the received data are inverted, an effect that can be compensated with polarity correction, until the next maximum is reached which occur an even multiple FSR/2.

The net DPSK wavelength offset SNR penalty includes both signal reduction and increased noise that is numerically analyzed in [132,144,430,431]. A simplified estimate of the offset penalty can be obtained in closed-form following the approach in [116,120,404]. Using the Marcum Q-function defined as

$$Q_{\text{Marcum}}(a,b) = \int_b^\infty x \exp\left[-\tfrac{1}{2}\left(x^2+a^2\right)\right] I_0(ax)\,dx, \tag{119}$$

where I_0 is the modified Bessel function of the first kind of zero order, the BER and SNR estimate (γ_{calc}) are given by [135]

$$\text{BER}_{\text{calc}} \cong 0.5 \left[\begin{array}{l} 1 - Q_{\text{Marcum}}\left(\sqrt{2\gamma_{\text{th}}}\cos(\frac{\pi\Delta f}{\text{FSR}}), \sqrt{2\gamma_{\text{th}}}\sin(\frac{\pi\Delta f}{\text{FSR}})\right) \\ +Q_{\text{Marcum}}\left(\sqrt{2\gamma_{\text{th}}}\sin(\frac{\pi\Delta f}{\text{FSR}}), \sqrt{2\gamma_{\text{th}}}\cos(\frac{\pi\Delta f}{\text{FSR}})\right) \end{array}\right], \tag{120}$$

$$\gamma_{\text{calc}} \cong -\ln(2\text{BER}_{\text{calc}}), \tag{121}$$

$$\Delta\gamma \cong 10\log_{10}\left(\frac{\gamma_{\text{calc}}}{\gamma_{\text{th}}}\right), \tag{122}$$

where γ_{th} is the ideal SNR (or PPB) at the input to the preamplified RX (without frequency offset) and $\Delta\gamma$ is the effective SNR penalty in dB. This straightforward estimate for the offset SNR penalty provides excellent agreement with the measured data at 40 Gbit/s shown in Fig. 55 and in [135] at 2.5 Gbit/s, and is consistent with measurements and calcuations in [144,146,431]. For $|\Delta f/\text{FSR}| <\sim 4\%$, $\Delta\gamma$ is less than $\frac{1}{2}$ dB at the 10^{-3} BER, which is an error-free threshold point for commonly used FEC [24,25,41,115,444]. Note that for a fixed frequency offset, this alignment penalty is reduced as the FSR gets larger, an effect that makes it more challenging to avoid penalties at lower data rates. For 40 Gbit/s DPSK channels received with a 40-GHz FSR interferometer, wavelength alignment ($|\Delta f|$) needs to be within ± 1.6 GHz in order to limit $\Delta\gamma < \frac{1}{2}$ dB. At 2.5 Gbit/s, the $\frac{1}{2}$ dB wavelength tolerance is reduced to ± 100 MHz. But this can be achieved with commercial sources, as noted in section 3.3. With appropriate temperature and wavelength control, laser DFB laser wavelength can be stable for extended periods with less than 15 MHz standard deviation, consistent with 2.5 Gbit/s DPSK communication performance with < 0.1 dB SNR deviation [135].

5.2.2. Interferometer Stabilization

Given the significant SNR penalties for wavelength offset errors, the need for stable signal and DI wavelength alignment is clear. Passive athermal DI designs have been constructed to align to the 50 GHz ITU grid with Δf less than ± 0.8 GHz [432,433] over a 0–70C° temperature range. Assuming an otherwise perfectly aligned incoming signal wavelength, this would lead to only $\sim$0.1 dB temperature induced penalty at 40 Gbit/s. However, at 10 Gbit/s or lower data rates, even this level of stability would lead to prohibitive SNR penalties. Moreover, the TX signal wavelengths are not necessarily stable in free-space applications, an effect that discourages the use of passively-stabilized DIs. For example, space-based links between Geosynchronous and Low-Earth orbits (GEO-LEO) can experience Doppler shifts up to $\sim$9 GHz (for 1.55 μm signal wavelengths), that would preclude the use of an athermal interferometer without some form of TX signal compensation [135] for Doppler or other long-term wavelength shifts.

For single channel RXs, the preferred means of wavelength-offset compensation is local tracking at the RX for reasons of simplicity, speed, and stability of control. This

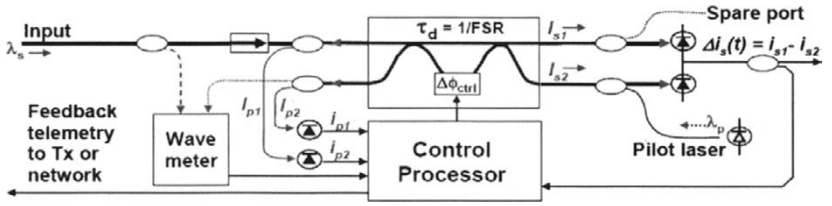


Fig. 56. Example setup for a pilot-stabilized DPSK receiver.

is typically implemented using synchronous phase-locking techniques that introduce a small dither on Δi_s in order to determine the sign of the phase error, and maximizing the peak baseband (RF) signal power accordingly [399]. The dither may be applied to one of the DI arms [438] or TX signal wavelength [439,445,446] [$\Delta\phi$ and Δf, respectively in Eq. (118)]. The amplitude of the dither can be reduced to sufficiently low levels so that active control can be achieved with little or no RX sensitivity penalty [133].

An alternative approach to interferometer stabilization makes use of pilot tones to stabilize the DI [135], which can provide dither-free open-loop DI control without an RX signal, and the flexibility to tune the DI to optimize performance when an RX signal is present [267,268]. Such capabilities are particularly useful during the spatial acquisition phase in a free-space link. With absolute DI wavelength control, for instance, the DI can be tuned to the proper wavelength in advance of the RX signal, thereby expediting the acquisition process and eliminating the need to re-align the DI after dropouts or fades. This also removes the polarity ambiguity, in which the sign of the data is unknown when using the standard phase-locking techniques.

An example of a pilot-based stabilization setup is shown in Fig. 56. The pilot signal, which could be generated by a low-power DFB laser, is injected into the DI through an optical tap in the reverse direction. As discussed in section 3.3 and earlier in this section, the pilot wavelength λ_p can be calibrated via temperature and current settings with a (short term) stability $<\sim$30 MHz [135], or through other methods, such as a feedback from a built in temperature controlled etalon [231] or an external wavelength reference such as a wavemeter, with sub-GHz long-term stability.

The pilot tone outputs from the DI (I_{p1} and I_{p2}) can be detected by low-speed photodetectors and the resulting photocurrents are processed to measure the normalized Δi_{pilot} or *contrast ratio* given by

$$D = \frac{I_{p1} - I_{p2}}{I_{p1} + I_{p2}} = \frac{\cos^2(\theta) - \sin^2(\theta)}{\cos^2(\theta) + \sin^2(\theta)} = \cos(2\theta). \tag{123}$$

Here $\theta = \pi\Delta f_p$ /FSR + $\Delta\phi_{\text{ctrl}}$ +$\Delta\phi_{\text{error}}$ is the net phase argument, Δf_p is the relative pilot carrier frequency, which is either known or measured, $\Delta\phi_{\text{ctrl}}$ is an interferometer phase control term, and $\Delta\phi_{\text{error}}$ is a random interferometer phase term, which can be constant or have time and temperature dependencies that need to be compensated for. Fig. 57 shows a plot of D and normalized I_{p1} and I_{p2} as a function of the Δf.

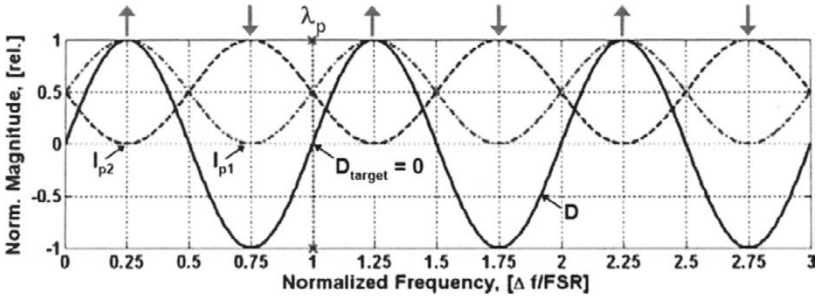


Fig. 57. The *contrast ratio* D and normalized pilot outputs I_{p1} and I_{p2} as a function of Δf normalized by the DI FSR. Arrows above the I_{p1} and I_{p2} fringe peaks indicate the location and polarity of signal wavelengths that the DI can accept without penalty when the DI is locked to the pilot wavelength at λ_p at $\Delta f/\mathrm{FSR} = 1$.

The contrast ratio ranges from +1 to −1 and has a period equal to the DI FSR. Thus, with the pilot capabilities shown in Fig. 56, the FSR and other important DI parameters such as extinction ratio can be measured by sweeping the pilot carrier frequency with fixed phase terms $\Delta\phi_{\mathrm{ctrl}}$ and $\Delta\phi_{\mathrm{error}}$, a capability that would also be useful for performing automated diagnostics in the field. Such capabilities could be enhanced with the use of multiple pilot tones (distinguishable by dither frequency and or wavelength), in which one pilot is used for stabilization, and the other for characterization.

By choosing a target contrast ratio, e.g., $D_{\mathrm{target}} = 0$, $\Delta\phi_{\mathrm{ctrl}}$ can be adjusted to make the error between the measured and target contrast equal to zero (i.e., $D_{\mathrm{error}} = D_{\mathrm{measured}} - -D_{\mathrm{target}}$) using standard feedback control techniques. This locks the DI to the pilot wavelength so that it is aligned to accept signal wavelengths at

$$\lambda_s(m) = \lambda_p \pm (2m+1)\frac{\mathrm{FSR}}{4}, \tag{124}$$

where m is an integer and polarity correction is needed to invert the data when m is odd. This provides flexibility in placement of the pilot wavelength. For instance, it could be placed outside of the communication band to conserve spectrum and prevent potential in-band interference. While the pilot tone in the example above can be used as a CW single wavelength source, dithers in either wavelength and/or amplitude can be introduced to allow synchronous control/detection methods.

In practice, a pilot-stabilized DPSK RX can be used in the following manner. Prior to acquisition, the pilot tone can be used to align the interferometer to the anticipated RX wavelength. This eliminates the need for scanning the interferometer to search for an incoming wavelength as long as the anticipated and received wavelength are relatively close, e.g., correct to within about ±10% of an FSR. Once the signal is received, with incoming power above a predefined threshold $\mathrm{P_{min}}$, the interferometer control can switch to tracking the signal using standard approaches. Therefore, the control system only needs to track on the signal, if at all, when the signal is relatively strong, which reduces the dynamic range of operation required for signal tracking.

While the signal is strong, any misalignment between the pilot wavelength needed for correct interferometer alignment to the signal can be eliminated by having the pilot tone align to the interferometer. In this manner, a locally resident pilot tone can effectively track an incoming signal that can be slowly drifting due to poor control, aging, or Doppler shifts in moving platforms. Should the signal drop out, the interferometer control can switch back to the well aligned pilot control, which can remain stably aligned to the correct wavelength until the incoming signal exceeds P_{min}. This provides seamless interferometer performance during fades, and reduces any interferometer induced acquisition delay [268].

5.2.3. Multi-Wavelength DPSK Receiver Options

For high-rate 100+ channel WDM-DPSK systems (e.g., [138–140,167,447]), the size, weight, power, and costs associated with reliably maintaining a stable delay-line interferometer for each channel can be substantial. Miyamoto and coauthors [149,151] demonstrated concurrent PSK to ASK conversion of 43 Gbit/s WDM-DPSK channels to WDM-duobinary channels on the 100 GHz ITU grid using a 50-GHz free-spectral-range (FSR) DI. While well suited for applications where spectral efficiency is required, this simplified DPSK RX incurs sensitivity penalties of 3 dB due to single-ended DPSK reception and another ~0.5 dB due to the ~14% mismatch between the 50-GHz DI FSR and the 42.8 GHz data rate when used with 66% RZ waveforms. As noted earlier, the mismatch penalty is waveform dependent; with ~14% delay error, the penalty increases to ~1 dB for 33% RZ waveforms further reducing the sensitivity benefit of DPSK.

Simplified multi-channel 'DPSK' receivers have also been implemented with periodic narrow band optical filtering and similar duobinary signals over dispersive channels [148,150,416,448]. While dispersion tolerant, these single-ended demodulators also incur sensitivity penalties in excess of 3 dB when compared to balanced DPSK receivers.

However, multi-channel WDM-DPSK reception can be achieved using a single DI with near-quantum-limited performance and be compatible with existing standards without waveform-dependent penalties [135].

By leveraging the DI's periodic transfer function shown in Fig. 55, and constraining the received wavelength spacing ($\Delta\nu_{ch}$), multiple WDM-DPSK channels can be simultaneously demodulated using a single interferometer. Unlike single-channel DPSK receivers in which the DI can track the incoming signal, in the multi-channel configuration, independent channels must have the same wavelength periodicity as the DI in order to avoid significant SNR penalties. For WDM-PDSK signals coming from the same TX, this periodicity can be configured locally at the TX. In this case, the RX DI can also track locally on any of the received signals in order to align to all of the incoming WDM channels. However, local tracking at the RX is precluded for multi-access applications, where independent TXs send WDM-DPSK signals to a common RX. In this case, proper wavelength alignment can be achieved with either absolute λ-control at the TX and RX, or with (slow) feedback from the RX via network-level maintenance and control of the TXs. Transmitter-centric control could also be used to compensate for aperiodicities that can occur in the DI [449]. As noted in the previous section, absolute λ-control of the DI can be achieved via stable thermal design [432] or with pilot tone stabilization, which could be a locally-resident calibrated laser. Another

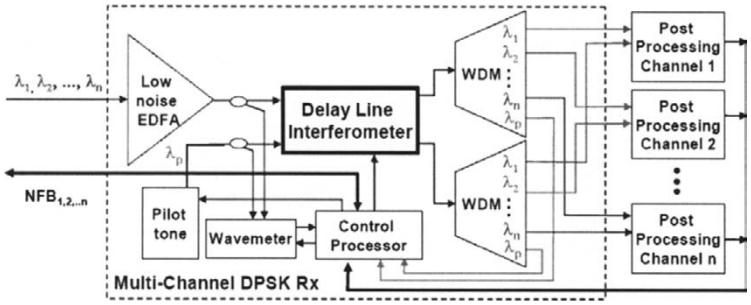


Fig. 58. A Multi-channel DPSK RX with feedback and feedforward alignment capabilities that can include a pilot tone (λ_p) or master-channel, wavemeter, channel power and bit-error rate (BER), and interactive communication with the network (NFB) [266].

option is to use a master-reference channel that the DI, and all other signal and pilot λs can align to with feedback [268], providing the capability for sub-GHz (perhaps sub 100 MHz) wavelength alignment capability across the network.

WDM-DPSK channel separation can be achieved via optical demultiplexing after the DI [149,151,266,450,451] as shown in Fig. 58. Since balanced detection is needed to achieve high-sensitivity, this WDM-DPSK RX requires one extra WDM filter, but the WDM cost is small relative to the potential cost of tens of interferometers and associated stabilization hardware.

In order to avoid SNR degradation when using a multi-channel DPSK receiver, three conditions must be satisfied. First, the channel rate (R) must be nearly equal to an integer multiple n of the FSR to satisfy Eq. (117) and minimize delay-error. Second, $\Delta\nu_{\text{ch}}$ must be within a small frequency offset (Δf) from an integer multiple m, of FSR/2 (with polarity correction), i.e.,

$$\Delta\nu_{\text{ch}} = m\frac{\text{FSR}}{2} \pm \Delta f = m\frac{R}{2n} \pm \Delta f. \tag{125}$$

The third condition requires that transmitted channels each be aligned to target SNR fringe peaks so that $\Delta f \to 0$.

For wavelengths on the ITU grid (e.g., 100 GHz channel separation), standard SONET rates of 2.5, 10, and 40 Gbps for instance, are compatible with these conditions since the channel rates (or half-channel rates) are integer submultiples ($1/m$) of the ITU channel spacing.

However, when commonly used 7%-overhead G.709 compliant FEC is used with standard SONET rates, bringing a 10-Gbps SONET data rate to a 10.7-Gbps coded channel rate, it is impossible for all of the ITU grid-based WDM channels to align with the periodicity of the interferometer and satisfy the conditions in Eq. (125) above. In this case, SNR penalties can be avoided by either: a) adjusting the channel spacing to be a multiple of half the channel rate, abandoning the ITU grid if necessary, or b) adjusting $R/2$ to be an even factor of the channel spacing, abandoning, for example, SONET or G.709 standards if necessary.

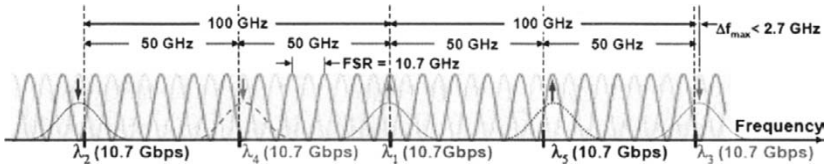


Fig. 59. WDM-DPSK channel spacing for 10.7 Gbps rate signals aligned near the 50/100 GHz ITU grid and to a 10.7-GHz FSR delay-line interferometer. Arrows centered on each wavelength indicate location and polarity of the received data.

If conforming with existing standards is the priority, Eq. (125) may not be satisfied exactly, but the associated performance penalties and deviation from the standards can be constrained to acceptable levels. For example, 10.7 Gbit/s channels on the 100-GHz ITU grid can be demodulated without performance penalty by a 10.7-GHz FSR DI, which can accept optical center frequencies every ~5.3 GHz with polarity correction. While most of the DI fringes will not align exactly to the 100-GHz ITU grid, none will misalign by more than 5.3/2 or ~2.7 GHz (FSR/4), with the average deviation of only ~1.4 GHz (FSR/8) as shown in Fig. 59. This deviation can be further reduced by a factor of n by demodulating with an n-bit DI, albeit with tighter wavelength alignment requirements as shown in Fig. 55 and bounded in Eq. (125) [266].

It is important to note that there are additional cost and performance benefits associated with the increase in alignment requirements needed for multi-channel WDM-DSPK RXs. For instance, spectral efficiency and RX sensitivity could benefit from reduced wavelength uncertainty throughout the network, which in turn, may allow reduced filter bandwidth margins (needed to accommodate wavelength uncertainties) and improved performance. Moreover, as noted in section 3.3, the additional cost of providing infrequent feedback to optimize wavelength alignment and communication performance of relatively stable tunable devices can be smaller than the expense of developing fieldable hardware with 20+ year stability.

Expanding on the example above, with proper differential precoding, the same 10.7-GHz DI could also be used to simultaneously demodulate 43 Gbit/s channels with negligible delay error. Thus, for near-harmonically related data rates, a single interferometer can be used to simultaneously demodulate multiple-rate WDM-DPSK signals [136]. To further diminish deviation from the ITU grid and to provide additional rate/alignment flexibility, a 2.68 GHz DI could be used to simultaneously demodulate 2.5, 2.67, 10.7, 40, 42.7, and 43.02 Gbps SONET and G.709-compliant WDM-DPSK signals within 700MHz of the ITU grid and with less than 7% delay-error.

For high-channel count WDM-DPSK links where spectral efficiency is more important, tight channel spacing is desirable but often comes with a decrease in RX sensitivity due to inter-channel interference resulting from limited WDM channel isolation. For single polarization RXs, this cross-talk penalty can be reduced by polarization multiplexing odd and even channels [452]. Another method of suppressing cross-talk can be seen in Fig. 55, for WDM channels separated by exactly an odd multiple of FSR/4. For this spacing, the signal intensities on the DI output arms are equal regardless of the data, and the resulting signal photocurrent $\Delta i(t)$ goes to zero. This property could be

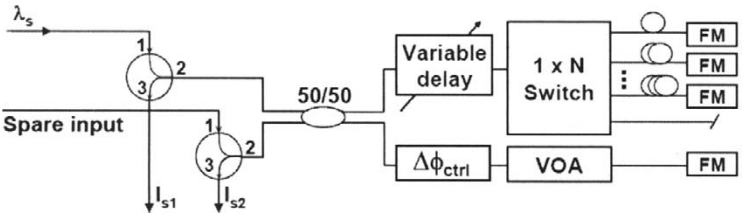


Fig. 60. A reconfigurable polarization-independent delay-line interferometer, based on a double-pass Michelson interferometer design using Faraday-rotator mirrors (FMs) to compensate for residual polarization rotation in the arms of the interferometer. Variable-delays and switching elements can be included to adjust the delay or FSR, and variable-optical-attenuators (VOAs) can adjust the relative power in the two arms to optimize the DI-extinction [267,268].

used to benefit dense WDM systems, by staggering the channel spacing so that delay-line interferometer improves rejection and reduces potential cross talk penalties. This is especially helpful for adjacent channels where WDM rejection is the poorest. For the multi-channel receivers described above, this could be implemented by sending odd and even WDM-DPSK channels to two separate WDM-DPSK RXs, staggered by FSR/4. Moreover, when n-bit DIs are used for demodulation (yielding FSR $= R_b/n$), the channel spacing can be adjusted by integer submultiples of the data rate, providing n-times finer wavelength adjustment for optimizing the trade between photon and spectral efficiency of the system.

5.2.4. Reconfigurable DPSK Demodulators

The ability to simultaneously demodulate many WDM-DPSK signals with a single DI, becomes even more useful if the DI can be reconfigured to optimize performance or accommodate other data rates. For example, a DPSK receiver could be adjusted to accept different (non-harmonically related) rates or FEC upgrades without delay-error penalties. Even in single-channel DPSK RX designs, it may be preferable to have one reconfigurable DI and associated control that can operate at several rates, rather than a fixed DI for each rate.

A polarization independent extension of the double-pass amplifier design [354] discussed in section 3.5.3 is a Michelson-based interferometer that incorporates Faraday rotator mirrors in both arms of the interferometer as shown in Fig. 60.

This geometry compensates for residual polarization rotation over a broad-range of wavelengths, making it well suited for WDM-DPSK applications, and the lack of polarization dependence enables useful non-polarization-maintaining elements to be incorporated into the arms of the interferometer without introducing polarization dependent losses (PDL) or degradation in DI performance. For example, a standard single-mode variable-delay-line with $>\sim$200 ps delay could be used to continuously tune a reconfigurable DI to accept DPSK signals at any rate greater than $\sim$2.5 Gbit/s, and provides a built-in means of adjusting the FSR to compensate for rate-FSR mismatch. The tuning and phase-control range in this case benefits by a factor of two from the double-pass geometry. Alternatively, $1 \times N$ switching elements can be used

to vary the DI-delay by discrete steps to extend the DI tuning range further. When used in combination with a variable-delay, the continuous tuning provides additional rate flexibility and relaxes path-length tolerances.

Other elements, such as a variable-optical-attenuator (VOA) can be used to compensate for relative changes in insertion loss between the two paths, which, in combination with pilot-based characterization, can be used to measure and optimize DI-extinction ratio. By including the ability to terminate the reflection from the arms of the interferometer, either by adjusting the VOA or switching to an angle-terminated port, the splitting ratio of the 50/50 coupler and the output intensities I_{s1} and I_{s2} and corresponding photocurrents can be measured and balanced.

Pilot tones can be injected at the spare input port in this geometry to stabilize and characterize the DI (see Fig. 58 and section 5.2.2. This, in combination with variable-delay adjustment enables autocalibration and *in-situ* feedback and optimization capabilities. For example, the DI-delay can be adjusted to minimize received BER, and pilot-based stabilization/characterization can provide independent measure of the DI-FSR (channel-rate) and the incoming signal wavelength. This information can be used to track/manage wavelengths and relative Doppler shifts throughout the network.

5.3. Hybrid Modulation Formats

In order to optimize overall system design, performance, and cost, we can consider hybrids of the orthogonal modulation formats described in section 2.2.1, which can be used to access the many THz of available optical spectrum. Selection of f-FSK frequencies, p-PPM positions, and L-polarization states distributed over w-WDM channels can be used to overcome practical limitations such as electrical bandwidth, nonlinear impairments, and available channel bandwidth as highlighted in Eqs. (126)–(129). The number of bits/symbol in such a multi-dimensional hybrid format is given by

$$k = \log_2(pfL) = \log_2(M_{\text{eff}}), \tag{126}$$

where M_{eff} is the effective number of orthogonal symbols in the constellation. In order to maintain orthogonality, the maximum number of polarization states $L_{\max} = 2$. In the Shannon-sense, the bandwidth expansion factor (as illustrated in Fig. 1) is given by

$$F = \frac{1}{r} = \frac{\log_2(pfL)}{pfL\,(1+\eta_{\text{FEC}})} = \frac{k}{M_{\text{eff}}\,(1+\eta_{FEC})} \quad \left[\frac{\text{bit/s}}{\text{Hz}}\right]^{-1}, \tag{127}$$

which can be used to improve RX sensitivity and photon efficiency. Here r is spectral efficiency, and the $(1+\eta_{\text{FEC}})$ term accounts for additional FEC overhead such that the channel rate, R_b $(1+\eta_{\text{FEC}})$. For a given data bit rate R_b, the combination of the number of wavelengths and symbol constellation size can be used to reduce the required electrical bandwidth per channel:

$$B_E = \frac{R_b\,(1+\eta_{\text{FEC}})\,p}{wk}, \tag{128}$$

which benefits from the use of orthogonal bases (polarization and frequency) and independent WDM channels. The net optical bandwidth required is

$$B_O = \frac{2R_b\,(1+\eta_{\mathrm{FEC}})\,pf}{k}, \tag{129}$$

where the factor of 2 is a conservative estimate for the minimum wavelength spacing. The corresponding effective peak TX power is given by

$$P_{\mathrm{pk}}^{\mathrm{eff}} = \left(\frac{\kappa_{\mathrm{eff}}}{w}\right)\frac{\bar{P}p}{\mathrm{DC}_{\mathrm{pulse}}}, \tag{130}$$

where κ_{eff} is the effective number of overlapping symbols defined Eq. (90) in section 3.5.5A.

From these equations, the impact of adjusting the hybrid-modulation parameters (p, f, L, and w) can readily evaluated, to optimize the lasercom system. For example, given a maximum electronic bandwidth limitation $B_E^{\max}$, the maximum achievable data rate can be determined from Eq. (128),

$$R_b^{\max} = \frac{B_E^{\max} wk}{(1+\eta_{\mathrm{FEC}})\,p} = \frac{Swk}{(1+\eta_{\mathrm{FEC}})\,p}, \tag{131}$$

where S is the bandwidth-limited p-PPM slot rate. From Eq. (131) it is easy to see that $R_b^{\max}$ can be increased directly through the number of WDM channels or by increasing the net bits per symbol k, by expanding the constellation size with supplementary frequency or polarization bases. The former linearly increases $R_b^{\max}$ while maintaining the same spectral efficiency and sensitivity per channel, whereas the latter can be used to augment $R_b^{\max}$ while improving RX sensitivity [see Eqs. (126) and (17)–(20)] at the expense of spectral efficiency.

5.4. Demonstrated Communication Performance

Throughout the paper, we have discussed methods of optimizing receiver sensitivity through use of various TX and RX designs, modulation and coding. An extension of Alexander's compilation of representative high-sensitivity lasercom demonstrations is given in Table 14 and Fig. 61 [20]. Receiver sensitivity at 10^{-9} BER is listed as a function of data-rate for various modulation formats, waveforms, and RX designs. These demonstrations bridge the gap between theory and what can be achieved given realistic design and technology constraints, and support many of the conclusions and design guidelines presented here. As laser communication technologies and TX/RX designs have matured, data rates have increased and sensitivities have improved, moving towards the lower-right quadrant of the Fig. 61, with the current state-of-the-art rapidly approaching the fundamental quantum and Shannon limits of communication performance.

As we can see in Fig. 61, at high rates, optically preamplified receivers have demonstrated the best sensitivities, despite inferior theoretical sensitivities relative to coherent-homodyne and photon-counting RXs. This is largely due to the nearly ideal characteristics of wideband high-gain average-power-limited optical amplifiers discussed throughout this paper that enable the preamplified RX structure to leverage the strengths and mitigate the deficiencies in essential RX electronics. As a result, preamplified systems can often leverage ongoing developments in wideband telecom optical and high-speed electronic technologies, proving a cost-effective means of advancing the state-of-the-art. When combined with the maturity and sophistication

Table 14. List of high-sensitivity optical receiver demonstrations shown in Fig. 61.

#	Rate, [Mbit/s]	Sensitivity, [avg. PPB]	Modulation/ Waveform	RX	Coding, O/H	Ref.
1	65*	1	32-PPM	Phot. Cnt. SSPD	SCPPM Turbo, 100%	[34]
2	14	1	64-PPM	Phot. Cnt. GM-APD	SCPPM Turbo, 100%	[44]
3	781	2	32-PPM	SSPD Phot. Cnt.	SCPPM Turbo, 100%	[34]
4	73	4	256-PPM	Opt. Preamp	RS255/239, 7%	[15]
5	2	5	PSK	Homodyne	Conv. Cd.	[157]
6	10000	7	DSPK-RZ	Opt. Preamp	Turbo Block, 25%	[24]
7	781	8	32-PPM	Phot. Cnt. SSPD	SCPPM Turbo, 100%	[33]
8	40000	8.7	DSPK-RZ	Opt. Preamp	E-RS255/239, 7%	[25]
9	78	9	256-PPM	Opt. Preamp	no	[15]
10	136	10	128-PPM	Opt. Preamp	no	[15]
11	10000	11	DSPK-RZ	Opt. Preamp	E-RS255/239, 7%	[41]
12	233	10	64-PPM	Opt. Preamp	no	[15]
13	389	12	32-PPM	Opt. Preamp	no	[15]
14	4	16	PSK	Homodyne	no	[157]
15	622	16	16-PPM	Opt. Preamp	no	[15]
16	10000	16	OOK-RZ	Opt. Preamp	E-RS255/239, 7%	[41]
17	565	20	PSK	Homodyne	no	[195]
18	781	20	32-PPM	Phot. Cnt. SSPD	SCPPM Turbo, 100%	[33]
19	933	20	8-PPM	Opt. Preamp	no	[15]
20	140	25	PSK	Homodyne	no	[453]
21	2500	25	DPSK-RZ	Opt. Preamp	no	[133]
22	10000	30	DSPK-RZ	Opt. Preamp	no	[129]
23	1244	31	4-PPM	Opt. Preamp	no	[15]
24	5600	35	PSK	Homodyne	no	[158]
25	10000	36	DSPK	Opt. Preamp	no	[129]
26	400	36	PSK	Homodyne	no	[155]
27	110	37	4-FSK	Heterodyne	no	[160]
28	1000	37	OOK	Opt. Preamp	4/5 Conv. Cd., 25%	[454]
29	10000	38	DSPK	Opt. Preamp	no	[128]
30	42700	38	DPSK	Opt. Preamp	no	[159]
31	5000	43	OOK-RZ	Opt. Preamp	no	[13]
32	1250	45	OOK-RZ	Opt. Preamp	no	[455]
33	400	45	DPSK	Heterodyne	no	[456]
34	12500	45	DSPK-RZ	Opt. Preamp	no	[24]
35	42700	45	DSPK-RZ	Opt. Preamp	no	[130]
36	1250	46	FSK-RZ	Opt. Preamp	no	[455]
37	1000	46	PSK	Homodyne	no	[164]
38	10000	57	OOK-RZ	Opt. Preamp	no	[406]
39	10000	60	OOK-RZ	Opt. Preamp	no	[129]
40	1244	61	2-PPM	Opt. Preamp	no	[15]
41	3000	62	DPSK	Opt. Preamp	no	[438]

*Note that #1 is the #3 32-PPM demonstration at 10 Gslots/s with a 12× repeat code. This lowers photon flux and reduces blocking losses by 3 dB, and lowers the data rate to 65 Mbps. It also emulates the performance of a 12-detector array operating at 781 Mbps.

Table 14. (Continued) List of high-sensitivity optical receiver demonstrations shown in Fig. 61.

#	Rate, [Mbit/s]	Sensitivity, [avg. PPB]	Modulation/ Waveform	RX	Coding, O/H	Ref.
42	51	65	2-PPM	Opt. Preamp	no	[11,12]
43	155	65	2-PPM	Opt. Preamp	no	[11,12]
44	311	65	2-PPM	Opt. Preamp	no	[11,12]
45	622	65	2-PPM	Opt. Preamp	no	[11,12]
46	1244	65	2-PPM	Opt. Preamp	no	[11,12]
47	2500	66	CPFSK	Heterodyne	no	[194,457]
48	60	68	4-PPM	APD-FET	no	[391]
49	120	71	4-PPM	APD-FET	no	[391]
50	4000	72	PSK	Homodyne	no	[458]
51	200	74	FSK	Heterodyne	no	[179]
52	560	78	DPSK	Heterodyne	no	[459]
53	30	80	4-PPM	APD-FET	no	[391]
54	7800	81	PSK	Homodyne	no	[158]
55	10000	81	OOK	Opt. Preamp	no	[129]
56	140	84	FSK	Heterodyne	no	[460]
57	10000	88	Duobinary	Opt. Preamp	no	[152]
58	6000	89	CPFSK	Heterodyne	no	[461]
59	5000	93	PSK	Homodyne	no	[166]
60	10000	100	PSK	Homodyne	no	[462]
61	15	108	4-PPM	APD-FET	no	[391]
62	325	110	OOK	APD-trans-Z	no	[390]
63	10000	112	OOK	Opt. Preamp	no	[463]
64	565	116	DPSK	Heterodyne	no	[464]
65	3000	116	DPSK	Opt. Preamp	no	[465]
66	140	119	FSK	Heterodyne	no	[162]
67	5000	135	OOK	Opt. Preamp	no	[466]
68	2500	137	OOK	Opt. Preamp	no	[316]
69	10000	147	OOK	Opt. Preamp	no	[467]
70	622	152	OOK	Opt. Preamp	no	[468]
71	622	155	OOK	APD-HEMPT	no	[469]
72	1244	156	OOK	Opt. Preamp	no	[470]
73	4000	175	OOK	Heterodyne	no	[471]
74	4000	191	FSK	Heterodyne	no	[471]
75	10000	193	OOK	Opt. Preamp	no	[316]
76	4000	209	DPSK	Heterodyne	no	[471]
77	20000	270	OOK	Opt. Preamp	no	[472]
78	10000	297	PSK	Homodyne	no	[165]
79	2000	705	OOK	APD-FET	no	[473]
80	1000	1162	OOK	APD-FET	no	[474]
81	20000	1600	OOK	Opt. Preamp	no	[475]
82	10000	3000	OOK-RZ	APD-FET	no	[388]

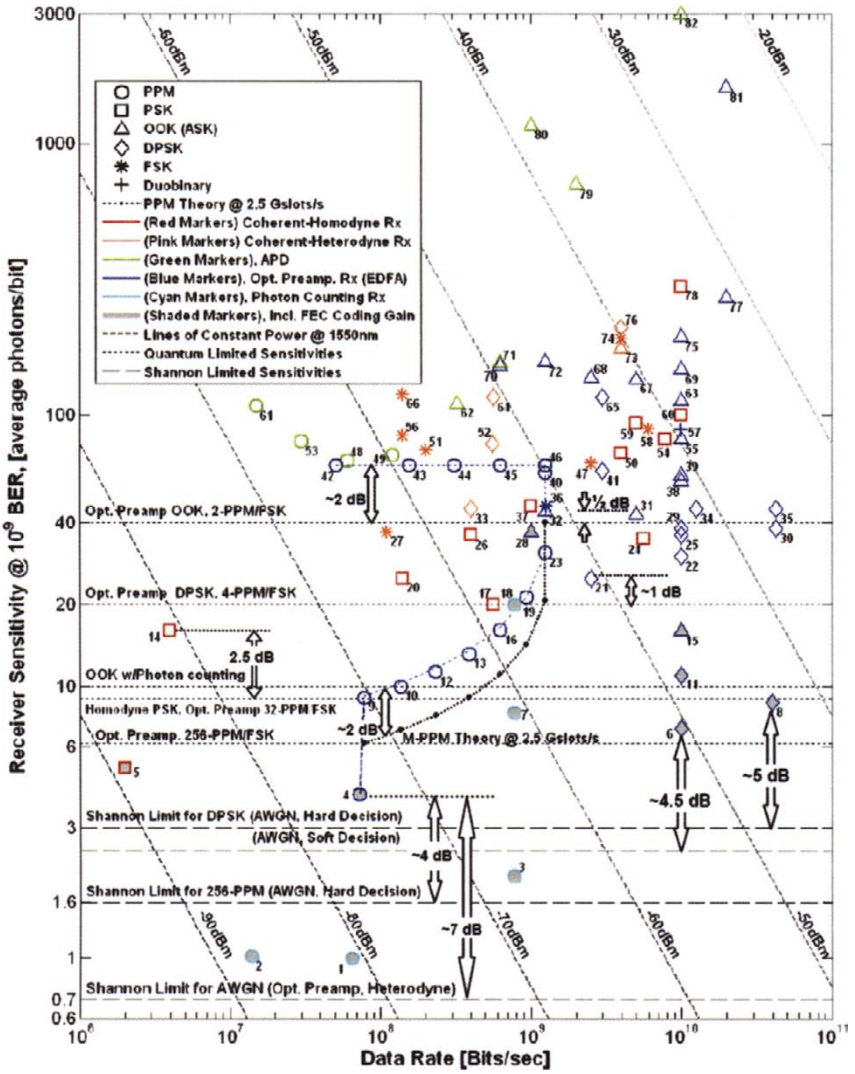


Fig. 61. Summary of high-sensitivity optical communication demonstrations showing reported sensitivities as a function of data rate. Modulation type is indicated by marker type, for instance PPM is represented by circles. Receiver type is indicated by color: Red indicates coherent (dark red = homodyne and pink = heterodyne), dark blue indicates an optically preamplified, and light-blue indicates a photon-counting RX. Coded demonstration results are shaded. Also shown are the quantum-limited sensitivities for the various modulation formats (uncoded), M-PPM theory for 2.5 Gslots/s, Shannon-limited sensitivities, and lines of constant power (dashed diagonal). See also Robinson et al. [214], 1.5 PPB sensitivity at 100 kbps. *Adapted from S. Alexander* [20].

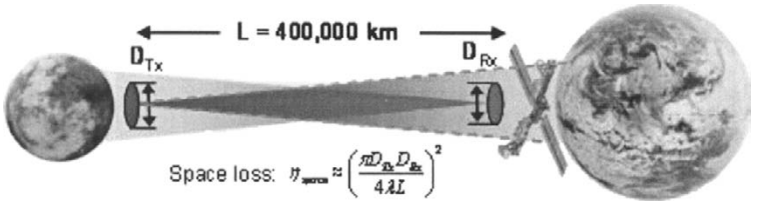


Fig. 62. Illustration of bidirectional Earth–Moon lasercom link [22].

of these technologies, performance is more robust and much closer to the theoretical limits over a wide range of data rates from the Mbit/s regime to tens of Gbit/s [13,15,24,25,41,130,159]. Furthermore, the data rate for high-speed preamplified receivers can be scaled using standard WDM techniques without additional performance penalty, providing a significant performance advantage over coherent- and photon-counting-based receiver designs.

Recent advancement in photon-counting detector technologies combined with powerful FEC has led to the most sensitive demonstrations in Fig. 61. These demonstrations include detector characteristics such as detection-efficiency and internal detector noise sources such as dark-counts, after-pulsing, and blocking loss (due to limited detector bandwidth). Note that these demonstrations do not include external background noise, which is an important consideration that can vary widely with channel conditions [108,109,476]. The multi-mode detection area of these detectors that is beneficial for mitigating channel turbulence and simplifying spatial acquisition, tracking, and photon-collection, combined with a broad spectral acceptance, make background noise a practical limitation in many photon-counting applications. Thus, for photon-counting receivers limited by background noise, optical filtering along with associated insertion losses must be included in the evaluation of RX performance. However after incorporating the additional ~0.5 to 1 dB losses for narrow-band optical filtering [398,477,478], photon-counting receivers sill show great promise for use in power-starved links, especially ground-based receivers, where atmospheric effects may be considerable and size, weight, power, and reliability are not necessarily critical design drivers.

5.5. Applications: to the Moon and Beyond

We conclude this paper with discussion of a long-haul application, which illustrates many of the benefits of FSO. Renewed interest in manned missions to the moon and beyond would require robust bidirectional high-rate links to support the human infrastructure, telemetry, science data, diagnostics, remote monitoring and control, and web-based connectivity. Here we present four notional lasercom link budgets for the ~400,000 km Earth–Moon link (Fig. 62), based on mature technologies that are readily available today and incorporate many of the concepts discussed earlier in the paper.

The budgets in Table 15 use the following communication parameters:

Modulation format M-PPM

Slot Rate: $S = 10$ Gslot/s
Coding gain: 5 dB
Waveform: 50% RZ
Wavelength: $\lambda = 1550$ nm
Aperture size: $D_{\mathrm{TX}} = D_{\mathrm{TX}} = 0.2$ m $\rightarrow \eta_{\mathrm{space}} = 86$ dB.

The budgets also include the following estimates of implementation losses:

Transmitter optics: $\eta_{\mathrm{TX}} = -3$ dB
Receiver optics: $\eta_{\mathrm{RX}} = -3$ dB
Pointing and tracking: $\eta_{\mathrm{PAT}} = -2$ dB
Receiver dB from theoretical: $\eta_{\mathrm{QL}} = -2$ dB.

Raybon and co-authors presented an impressive all optical 400 km regenerating optical loop experiment at 40 Gbit/s, with each loop consisting of four EDFA- and Raman-amplified 100 km spans [138,303]. They postulated that such technologies could be used to span a distance equivalent to the Earth– Moon link, but would require ~1000 regenerators to overcome the ~100,000 dB fiber loss. Long-haul FSO has clear benefits over such guided-wave links, since the net space-loss for 0.2 m TX and RX apertures is ~86 dB, a savings of over 99,000 dB. This advantage combined with more energy-efficient modulation and coding, and reduced data-rates enables the link to be closed with Watt-class transmitters.

Table 15. Lunar lasercom link budgets for various configurations of PPM constellation size (M), number of channels (w), and aperture diameter. Common to each of the four configurations is the slot rate $S = 10$ Gslot/s, coding overhead, wavelength ($\lambda = 1550$ nm), link distance, and assumptions for implementation penalties. These include losses in the TX ($\eta_{\mathrm{TXOpt}} = -3$ dB) and RX ($\eta_{\mathrm{RXOpt}} = -3$ dB) optics, pointing and tracking ($\eta_{\mathrm{PAT}} = -2$ dB) and RX deviation from theoretical performance ($\eta_{\mathrm{QL}} = -2$ dB). Adapted from [22].

Key Parameters for Earth ↔ Moon link	Units	#1	#2	#3	#4
Communication System Parameters:					
1 Constellation size (M)		1024	128	1024	256
2 Number of WDM channels (w)		1	1	8	120
3 Data Rate/channel, $R_{ch} = S \cdot \log_2(M)/[M \cdot (1+\eta_{FEC})]$	GBit/sec	0.065	0.365	0.065	0.208
Transmitter:					
4 Average Tx Power/channel (P_{Tx})	dBW	-3	-3	-12.0	-7.8
5 Peak Tx Power/channel = $P_{Tx} \cdot M/(w \cdot \mathrm{Pulse}_{DC})$	dBW	30	21	21	19
6 Net Tx Power = $w \cdot P_{Tx}$	dBW	-3	-3	-3	13
Receiver:					
7 Approx. Uncoded Rx sensitivity M-PPM, (PPB_{QL})	dB-PPB	6.0	7.6	6.0	7.0
8 Net Rx sensitivity @ Data Rate (S_{Rx})	dBW	-107.8	-98.7	-107.8	-101.7
9 Received Power/channel at EDFA input (P_{Rx})	dBW	-96.9	-96.9	-105.9	-101.7
Performance					
10 Implemented Receiver Sensitivity	dB-PPB	3.0	4.6	3.0	4.0
11 Net Bandwidth Required = $2 \cdot S \cdot w \cdot (1+\eta_{FEC})/125E9$	nm	0.24	0.24	1.92	28.8
12 Net Rate: Rnet = $R_{ch} \cdot w$	GBit/sec	0.065	0.365	0.521	25.0
13 Margin	dB	10.9	1.8	1.8	0.0

A representative TX and RX pair suitable for long-term use in space [21,198,380,409] that can achieve the performance in the budget above is a Master-Oscillator-Power Amplifier (MOPA) transmitter with an optically preamplified receiver [15]. Quasi-Gaussian 50%-RZ pulse-shaping in the transmitter is used to ensure a robust match between the TX-waveforms and the Gaussian optical filter in the RX and minimize

the impact of timing jitter [14], with demonstrated performance within 1 dB of the quantum limit [13,133]. Furthermore, as discussed in sections 3.5.5 and 5.1.3, fixing the pulse shape for all M enables flexible multi-rate capability with a single receiver optimized for all rates [11,12,15].

In Table 15, configuration #1 highlights a low-rate 65-Mbit/s link achieved with ~11 dB margin using a 0.5 W transmitter, 0.2m TX and RX apertures, and 1024-PPM. Of course, the margin could be traded to reduce aperture size or net output power. Alternatively, by simply changing the PPM alphabet size to M = 512, 256, or 128 as in configuration #2, the excess margin could be used to increase the data rate up to 365 Mbit/s with ~2 dB margin, providing the option for bandwidth-on-demand or fall-back modes of operations as needed.

Another option would be to include additional WDM channels with 0.2nm spacing as in configuration #3, which can provide up to 521 Mbit/s with ~2 dB margin. Relative to #2, the increase in net data rate is enabled by the 1.6 dB improvement in RX sensitivity due to the larger alphabet size. Although more hardware is required to support the additional WDM channels, #3 provides similar flexibility in net data throughput as #1 and #2 with additional redundancy. In addition, having multiple channels share a common power amplifier effectively reduces the peak TX power avoiding the nonlinear impairments that can limit M [21]. Such effects start to impact communication performance at 24 to 30 dBW peak levels as discussed in section 3.5.5C, and could necessitate a change of design parameters in #1. For example, hybrid-orthogonal constellations including polarization and frequency bases can be used to maintain M_{eff} and reduce the peak power.

By increasing average TX power to 13 dBW (20 W), and WDM channel count to 120 in configuration #4, a link with up to ~25 Gbit/s throughput can be achieved. This requires ~29nm of bandwidth, which easily falls within the 35-nm EDFA C-band, and corresponds to only 0.007 bit/s/Hz spectral efficiency. However, the resultant improvement in receiver sensitivity to 2.5 PPB enables a respectable distance-rate product of $\sim 10^{16}$km-bit/s—comparable to the best fiber-based demonstrations [93,138,139,141], without the need for ~1000 regenerators, emphasizing the difference between square- and exponentially-dependent channel loss. For #4, the overall link efficiency (using $\eta_{\text{E/O}}$ = 13% [199]) is

$$\eta_{\text{link}} = \frac{P_{\text{TX}}^{\text{net}}}{\eta_{\text{E/O}} R_{\text{net}}} = 6 \quad [\text{nJ/bit}] \tag{132}$$

for the 400,000 km lunar link.

We conclude that established technologies developed for the telecom industry can be leveraged to implement scalable photon-efficient optical communication capabilities today. These example lunar link budgets, based on an average-power-limited MOPA TX and a near-quantum-limited M-orthogonal optically-preamplified RX, show the potential for high-capacity long-haul free-space optical links at variety of data rates from the Mbit/s regime to tens of Gbit/s, rates that cannot be practically achieved using RF technologies. And, as photon-counting technologies mature, the sensitivity of optical receivers could potentially improve another 5–10 dB. These capabilities are likely to become an essential part of the space-based communications infrastructure, providing cost-effective support for deep-space exploration and future manned missions to the Moon, Mars, and beyond.

Acknowledgments

For most of the past decade, I've had the pleasure to work in the challenging and innovative atmosphere at MIT Lincoln Laboratory, perhaps best described by Don McClellan, the former head of the Communications Division who said, "If the laws of physics will permit it, we can build it."

Many thanks to Vincent Chan, Roy Bondurant, Kristin Rauschenbach, and Eric Swanson for providing me with the initial opportunity to work in this stimulating environment, where I have been able to contribute to the field of optical communications and bring many new ideas to reality, alongside an extremely talented group of present and former colleagues from the MIT community. These include Walid Atia, Richard Barron, David Bold, Don Boroson, Larry Candell, Steve Constantine, Jerry Chen, Eric Dauler, Bill Freking, Jeff Gottschalk, Matt Grein, Farhad Hakimi, Katie Hall, Scott Hamilton, Scott Henion, Paul Juodawlkis, Alan Kachelmyer, John Kaufmann, Andrew Kerman, Sumanth Kaushik, Bill Keicher, Mark Kuzenetzov, Farzana Khatri, Emily Kintzer, Sarah Klein, Jeff Livas, Jeff Mendenhall, John Moores, Bob Murphy, Dan Murphy, Ron Parenti, Wayne Phoel, Al Pillsbury, Jinendra Ranka, Jeff Shapiro, Neal Spellmeyer, Peter Schulz, Todd Ulmer, Simon Verghese, Matt Willis, Tim Yarnell, and many others in the community.

It has been a privilege to work closely with people of the caliber of John Carney, Bryan Robinson, Fred Walther, and especially Mark Stevens. As friends and colleagues with tremendous breadth and depth of experience with optics, communications, and making things happen, they have been a catalyst for many of the new ideas that are reflected in this paper. Along with other members of the Optical Communication Technology Group, they have contributed significantly to the content and initial demonstration of many of the concepts presented.

To my wife Rose and my family I owe the most gratitude. Without their continuing support and encouragement, this would not have been possible.

Symbols

α	Chirp parameter
$\Delta\gamma$	SNR penalty
κ_{eff}	Effective number of overlapping WDM pulses
λ	Wavelength $= c/\nu$, [m]
η_{atm}	Atmospheric loss
η_{ch}	Net channel loss = $\eta_{\text{TXOpt}}\eta_{\text{atm}}\eta_{\text{PAT}}g\eta_{\text{space}}\eta_{\text{RXOpt}}$
η_{link}	Overall link efficiency = $\eta_{\text{TX}}\eta_{\text{ch}}\eta_{\text{RX}}$TX-Joules/RX-bit, [J/bit]
η_{PAT}	Pointing and tracking losses
η_{RX}	Net Receiver efficiency= $\text{PPB}_{\text{QL}}G_{\text{code}}$[Photons/bit]
η_{RXOpt}	Receiving optics (telescope) losses
η_{space}	Space diffraction loss = $[\pi D_T D_R/(4\lambda L)]^2$
η_{TX}	Net Transmitter efficiency (E/O) [TX-Joules/Photon]
η_{TXER}	Transmitter ER penalty
η_{TXOpt}	Transmitting optics (telescope) losses
ν	Optical frequency = c/λ [Hz]
τ_{bit}	Bit or symbol duration [s]
τ_o	3 dB pulsewidth of optical filter impulse resp. [s]
τ_s	3 dB pulsewidth of signal (power) [s]
B	Bandwidth [Hz]
B_o	Optical bandwidth [Hz]
c	Speed of light in vacuum $c = \lambda\nu = 3.0 \times 10^8$ [m/s]
D_{TX}	Diameter of transmit aperture [m]
D_{RX}	Diameter of receive aperture [m]
DC_{eff}	Effective duty cyle $\text{DC}_{\text{eff}} = \text{DC}_{\text{mod}}\ \text{DC}_{\text{pulse}}$
DC_{mod}	Modulation duty cycle
DC_{pulse}	Pulse shape duty cycle
e	Electron charge 1.601×10^{-19} [C]
f_c	RF filter 3-dB bandwidth (cutoff frequency) [Hz]
f-FSK	f= # of FSK frequencies in a hybrid symbol
g_o	Optical amplifier small-signal gain
G	Gain
G_{code}	FEC coding gain
h	Planck's constant 6.626×10^{-34} [J/s]
k_B	Boltzman's constant 1.379×10^{-23} [W/°K-Hz]
k	number of bits-per M-ary symbol $k = \log_2(M)$
k_{eff}	Effective number of bits-per M-ary symbol
L-PolSK	L = # of PoSLK polarizations in a hybrid symbol
M	Symbol constellation size, as in M-ary PPM
$M =$	$2^k, k = \log_2(M)$ = # of bits/symbol
N_{avg}	Average photon count per interval [photons]
N_b	Background-noise photon count [photons]
$N_o/2$	Noise power spectral density [W/Hz]
$N_{\text{RX}-\text{min}}$	Minimum required RX SNR
N_{shot}	Shot noise power spectral density $= h\nu$ [W/Hz]
N_{sp}	Spontaneous emission factor [photons/mode]
N_{th}	Thermal noise power spectral density $= k_B T$ [W/Hz]

$P_{\text{in-sat}}$	Amplifier input saturation power $P_{\text{in-sat}} = P_{\text{sat}}/g_o$
P_{sat}	Amplifier saturated output power
P_{shot}	Shot noise power spectral density $= h\nu B$ [W]
P_{th}	Thermal noise power $= k_B T B$ [W]
PPB_{QL}	Uncoded quantum-limited RX sensitivity [photons/bit]
p-PPM	p = # of PPM positions in a hybrid symbol
r	Spectral efficiency [(bits/s)/Hz]
R_{ph}	Photon flux rate $P/(h\nu)$ [photons/s]
R_b	Bit rate, $R_b = 1/\tau_{\text{bit}}$[bits/s]
$R_{\text{A/W}}$	Detector responsivity, $R_{\text{A/W}} = \eta q/hv$ [mA/mW]
V_π	Switching voltage
w	Number of WDM channels
Z_o	Impedance Ω = [V/A]

Acronyms and Abbreviations

ADC	Analog-to-digital converter
APL	Average power limited
ARTEMIS	Advanced Relay Technology Mission Satellite
ASK	Amplitude shift keying, 2-ASK the same as OOK
AWGN	Additive white Gaussian noise
BER	Bit-error rate
CML	Chirp-managed laser
CW	Continuous wave
COTS	Commercial off-the-self
DAC	Digital-to-analog converter
DC	Duty cycle
dB	Decibel = 10*$\log_{10}$(Power/P_{ref})
dBm	10 $\log_{10}$(Power/1 mW)
dBW	10 $\log_{10}$(Power/1 W)
DI	Delay-line interferometer
DPSK	Differential phase shift keying
EDFA	Erbium-doped fiber amplifier
EYDFA	Erbium-Ytterbium (codped) fiber amplifier
EM	Electromagnetic
EMI	Electromagnetic interference
ER	Extinction ratio = $\text{Power}_{\text{off}}/\text{Power}_{\text{on}}$
FBG	Fiber Bragg grating
FEC	Forward error correction (coding)
FIT	Failure in time (#failures per billion hours of operation)
FM	Frequency modulation
FP	Fabry-Perot
FP^n	nth order Fabry-Perot
FSK	Frequency-shift keying
FSO	Free-space optical
FWHM	Full width half max
FWM	Four-wave mixing

Gbps	Gigabit per second
GEO	Geostationary earth orbit
GeoLITE	Geosynchronous Lightweight Technology Experiment
GOLD	Ground-to-Orbit Lasercom Demonstration
IM	Intensity modulation
IM-DD	Intensity-modulated direct-detection
ICI	Inter-channel-interference
ISI	Inter-symbol-interference
ITU	International Telecommunications Union
ITU-R	Radiocommunication Standard
ITU-T	Telecommunication Standard
ITU-grid	Standard 100 GHz WDM channel spacing
LEO	Low earth orbit
Mbps	Megabit per second
MLCD	Mars Laser Communication Demonstration
MZI	Mach-Zehnder interferometer
MZM	Mach-Zehnder Modulator
MTTF	Mean time to failure
NA	Numerical aperture, NA = $n\sin(\theta)$
NF	Noise Figure
OICETS	Optical Inter-orbit Communications Engineering Test Satellite
OOK	On-off keying (binary-ASK)
PAT	Pointing acquisition and tracking
PM	Polarization maintaining
PPB	Number of photons/bit
PPM	Pulse position modulation
PPL	Peak power limited
PRBS	Pseudo-random bit sequence
PSD	Power spectral density, [dBm/Hz]
P[$a|b$]	Probability of a given b
PER	Polarization extinction ratio
RF	Radio frequency, generic term for electrical
RX	Receiver
SCOWA	Slab-coupled optical waveguide amplifier
SCM	Subcarrier multiplexing
SBS	Stimulated Brillouin scattering
SILEX	Semiconductor Intersatellite Link Experiment
SPM	Self-phase modulation
SPOT-4	System Probatoire/Satellite Pour d'Observation de la Terra 4 satellite
SNR	Signal-to-noise-ratio
SRS	Stimulated Raman scattering
TDM	Time division multiplexing
TEC	Thermo-electric cooler
TIA	Transimpedance amplifier
TX	Transmitter
VDC	Variable duty cycle
VBG	Volume Bragg grating
WDM	Wavelength division multiplexing or multiplexer
YDFA	Ytterbium-doped fiber amplifier

References

1. B.L. Edwards, et. al., "Overview of the Mars laser communications demonstration project," in American Institute of Aeronautics and Astronautics, Space 2003 Conference & Exposition, 2003.
2. S.A. Townes, et. al., "The Mars laser communication demonstration," presented at IEEE Aerospace Conf., 2004.
3. E.A. Swanson and R.S. Bondurant, "Fiber-based free-space optical system," in US Pat. 5,062,150, 1991.
4. "Reliability assurance practices for optoelectronic devices in interoffice applications," Bellcore TR-NWT-000468 Issue 1, Dec. 1991.
5. "Generic reliability assurance requirements for optoelectronic devices used in telecommunications equipment," Telcordia GR-468-CORE no. 2, Dec. 2002.
6. "Generic requirements for fiber optic branching components," Bellcore GR-1209-CORE, Issue 1, Nov. 1994.
7. "Generic requirements for optical fiber amplifiers," GR-1312-CORE, Issue 2, Dec. 1996.
8. "Generic reliability assurance requirements for fiber optic branching components," GR-1221-CORE, Issue 1, Dec. 1994.
9. "Test methods and procedures for microelectronics," Military Standard MIL-STD-883C, Aug. 1983.
10. D.O. Caplan, M.L. Stevens, and D.M. Boroson, "Variable-rate communication system with optimal filtering," in US Pat. 6,694,104, 2004, (filed 1998).
11. D.O. Caplan, M.L. Stevens, D.M. Boroson, and J.E. Kaufmann, "A multi-rate optical communications architecture with high sensitivity," in LEOS, 1999.
12. M.L. Stevens, D.M. Boroson, and D.O. Caplan, "A novel variable-rate pulse-position modulation system with near quantum limited performance," in LEOS, 1999.
13. D.O. Caplan and W.A. Atia, "A quantum-limited optically-matched communication link," in Optical Fiber Conference (OFC), 2001.
14. D.O. Caplan and W.A. Atia, "Methods of achieving optimal communications performance," in US Pat. 7,181,097, 2007.
15. D.O. Caplan, B.S. Robinson, R.J. Murphy, and M.L. Stevens, "Demonstration of 2.5-Gslot/s optically-preamplified M-PPM with 4 photons/bit receiver sensitivity," in Optical Fiber Conference (OFC): Paper PDP23, 2005.
16. C.E. Shannon, "A mathematical theory of communication," Bell Syst. Technol. J. **27**, 379-423, 623-656, 1948.
17. S. Haykin, *Digital Communications* (John Wiley & Sons, Inc., 1988).
18. J.G. Proakis and M. Salehi, *Communication Systems Engineering* (Prentice-Hall, Inc., 1994).
19. Peebles, Peyton Z. Jr., *Digital Communications Systems* (Englewood Cliffs, NJ: Prentice-Hall, 1987).
20. S.B. Alexander, *Optical communication receiver design* (Bellingham, Washington, USA: SPIE Optical Engineering Press, 1997).
21. D.O. Caplan, "High-performance free-space laser communications and future trends," in Optical Amplifiers and Their Applications (OAA'05) Topical Meeting, Budapest, Hungary, 2005.

22. D.O. Caplan, B.S. Robinson, M.L. Stevens, D.M. Boroson, and S.A. Hamilton, "High-Rate Photon-Efficient Laser Communications with Near Single Photon/bit Receiver Sensitivities," in Optical Fiber Conference (OFC), 2006.

23. D.M. Boroson, "Optical Communications, A Compendium of Signal Formats, Receiver Architectures, Analysis Mathematics, and Performance Comparisons," 2005.

24. T. Mizuochi, et. al., "Forward error correction based on block turbo code with 3-bit soft decision for 10-Gb/s optical communication systems," IEEE Sel. Top. Quantum Electron. **10**, 376–386, 2004.

25. N.W. Spellmeyer, J.C. Gottschalk, D.O. Caplan, and M.L. Stevens, "High-sensitivity 40-Gb/s RZ-DPSK with forward error correction," IEEE Photon. Technol. Lett. **16**, 1579–1581 (2004).

26. J.R. Pierce, "Optical Channels: Practical Limits with Photon Counting," IEEE Trans. Commun. **COM-26**, 1819-1821, 1978.

27. Y. Yamamoto and H.A. Haus, "Preparation, measurement and information capacity of optical quantum states," Rev. Mod. Phys. **58**, 1001-1020, 1986.

28. H.A. Haus, "Limits on communication using photons," in MIT-EECS Colloquium Series, 1997.

29. G.N. Gol'tsman, O. Okunev, G. Chulkova, A.L.A. Semenov, K. Smirnov, B. Voronov, A. Dsardanov, C. Williams, and R. Sobolewski, "Picosecond superconducting single-photon optical detector," Appl. Phys. Lett. **79**, 705–707, 2001.

30. K.A. McIntosh, et. al., "InGaAsP/InP avalanche photodiodes for photon counting at 1.06 μm," Appl. Phys. Lett. **81** (2002).

31. J. Zhang, W. Slysz, A. Verevkin, O. Okunev, G. Chulkova, A. Korneev, A. Lipatov, G.N. Gol'tsman, and R. Sobolewski, "Response time characterization of NbN superconducting single-photon detectors," IEEE Trans. App. Supercond. **13**, 180–183, 2003.

32. K.A. McIntosh, et. al., "Arrays of III-V semiconductor Geiger-mode avalanche photodiodes," in LEOS, 2003.

33. B.S. Robinson, A.J. Kerman, E.A. Dauler, R.J. Barron, D.O. Caplan, M.L. Stevens, J.J. Carney, S.A. Hamilton, J.K.W. Yang, and K.K. Berggren, "781-Mbit/s Photon-Counting Optical Communications Using Superconducting NbN-Nanowire Detectors," Opt. Lett. **31**, 444– 446 (2006).

34. B.S. Robinson, A.J. Kerman, E.A. Dauler, R.J. Barron, D.O. Caplan, M.L. Stevens, J.J. Carney, S.A. Hamilton, J.K.W. Yang, and K.K. Berggren, "High-Data-Rate Photon-Counting Optical Communications Using a NbN Nanowire Superconducting Detector," in Conference on Lasers and Electro-Optics (CLEO), 2006.

35. E.A. Dauler, B.S. Robinson , A.J. Kerman, V. Anant, R.J. Barron, K. Berggren, D.O. Caplan, J.J. Carney, S.A. Hamilton, K.M. Rosfjord, M.L. Stevens, and J.K.W. Yang, "1.25-Gbit/s photon-counting optical communications using a twoelement superconducting nanowire single photon detector," in Proc. SPIE, (Advanced Photon Counting Techniques), 2006.

36. P.S. Henry, "Error-rate performance of optical amplifiers," in Optical Fiber Conference (OFC), 1989.

37. N.A. Olsson, "Lightwave Systems With Optical Amplifiers," J. Lightwave Technol. **7**, 1071–082 (1989).

38. P.A. Humblet and M. Azizoglu, "On bit error rate of lightwave systems with optical amplifiers," J. Lightwave Technol. **9** (1991).

39. S.D. Personick, "Receiver design for digital fiber optic communication systems, I & II," Bell Syst. Tech. J. **52**, 843-886, 1973.

40. S. Vanstone and P.C. v. Oorschot, *An introduction to Error Correcting Codes with Applications* (Kluwer Academic Publishers, 1989).

41. D.O. Caplan, J.C. Gottschalk, R.J. Murphy, N.W. Spellmeyer, M.L. Stevens, and A.M.D. Beling, "Performance of high-rate high-sensitivity optical communications with forward error correction coding," in Conference on Lasers and Electro-Optics (CLEO): Paper CPDD9, 2004.

42. B.E. Moision and J. Hamkins, "Coded modulation for the deep-space optical channel: serially concatenated pulse-position modulation," 42-161, 2005.

43. D.M. Boroson, C.C. Chen, and B.L. Edwards, "The Mars laser communication demonstration project: truly ultralong-haul optical transport," in Optical Fiber Conference (OFC), 2005.

44. P.I. Hopman, P.W. Boettcher, L.M. Candell, J.B. Glettler, R. Shoup, and G. Zogbi, "An End-to-End Demonstration of a Receiver Array Based Free-Space Photon Counting Communications Link," in SPIE, (Free-Space Laser Communications VI), vol. 6304, 2006.

45. R. Shoup, "Hardware implementation of a high-throughput 64-PPM serial concatenated turbo decoder," in SPIE, (Optical Information Systems IV), vol. 6311, 2006.

46. E.A. Swanson and R.S. Bondurant, "Using fiber optics to simplify free-space lasercom systems," in Proc. SPIE, (Free-Space Laser Communication Technologies II), **1218**, 70–82 (1990).

47. R.M. Gagliardi and S. Karp, *Optical Communication*, 2nd Ed. (New York,: John Wiley & Sons, Inc., 1995).

48. F. G. Walther, J.M. Roth, W.E. Keicher, A.E. DeCew, and "Wavelength division and polarization division multiple access free space optical terminal using a single aperture," in US Pat. Appl. 20040081466, 2004.

49. A.G. Bell and S. Tainter, "Photophone transmitter," in US Pat. 235,496, 1880.

50. A.G. Bell, "On the Production and Reproduction of Sound by Light," Am. J. Sci., 3rd Series, **XX**, 305–324 (1880).

51. A.G. Bell and S. Tainter, "Photophonic receiver," in US Pat. 241,909, 1881.

52. A.G. Bell, C.A. Bell, and S. Tainter, "Transmitting and recording sounds by radiant energy," in US Pat. 324,213, 1886.

53. D. Killinger, "Free space optics for laser communication through the air," in *Optics & Photonics News*, 2002.

54. Hertz, "Heinrich Rudolf Hertz discovered GHz radio waves in 1887," 1887.

55. A.G. Bell, "Improvement in Telegraphy," in US Pat. 174,465, 1876.

56. T.A. Edison, "Telephones or Speaking Telegraphs," in US Pat. 203,018, 1878.

57. G. Marconi, "A system of telegraphy using Hertzian waves," in British Pat. 12039, (filed June 2, 1896, demonstrated Sept. 2, 1896), 1896.

58. M. Loomis, "Improvement in Telegraphing," in US Pat. 129,971, 1872 (first US wireless telegraphy patent).

59. R.A. Fessenden, "Wireless Telegraphy," in US Pat. 706,737, 1902, (filed 1901).

60. A.L. Schawlow and C.H. Townes., "Infrared and Optical Masers," Phys. Rev.**112**, 1940–1949 (1958).

61. T.H. Maiman, "Stimulated optical radiation in ruby," Nature **187**, 493 (1960).

62. T.H. Maiman, "Ruby Laser Systems," in US Pat. 3,353,115, 1967, (filed 1963).

63. A. Javan, Bennett, W.R. Jr., and D.R. Herriott, "Population Inversion and Continuous Optical Maser Oscillation in a Gas Discharge Containing a He-Ne Mixture," Phys. Rev. Lett. **6**, 106–110, 1961.

64. "Research groups at GE, IBM, and Lincoln Laboratory at MIT demonstrated semiconductor lasers using gallium arsenide (GaAs) in 1962. (Courtesy W. Keicher)," 1962.

65. C.H. Gooch, *Gallium Arsenide Lasers* (New York: John Wiley & Sons, 1969).

66. C.J. Koester and E. Snitzer, "Amplification in a Fiber Laser," Appl. Opt. **3**, 1182–86 (1964).

67. E. Snitzer, "Means for producing and amplifying optical energy," in US Pat. 3,729,690, (filed 1969, 1961), 1973.

68. Z.I. Alferov, V.M. Andreev, D.Z. Garbuzov, Y.V. Zhilyaev, E.P. Morozov, E.L. Portnoi, and V.G. Trofim, "Investigation of the influence of the AlAs-GaAs heterostructure parameters on the laser threshold current and the realization of continuous emission at room temperature," Sov. Phys. Semicond. **4**, 1573–1575 (1971).

69. C.K. Kao and G.A. Hockham, "Dielectric-Fiber Surface Waveguides for Optical Frequencies," Proc. IEE **133**, 1151–58 (1966).

70. R.D. Maurer and P.C. Schultz, "Fused silica optical waveguide," in US Pat. 3,659,915, 1972, (filed 1970).

71. D.B. Keck and P.C. Schultz, "Method of producing optical waveguide fibers," in US Pat. 3,711,262, 1973, (filed 1970).

72. C.K. Kao and T.W. Davies, "Spectroscopic studies of ultra low loss optical glasses," J. Sci. Instrum. (1968).

73. F.P. Kapron, D.B. Keck, and R.D. Mauer, "Radiation losses in glass optical waveguides," Appl. Phys. Lett. **17**, (1970).

74. E. Snitzer, "Cylindrical dielectric waveguide modes," J. Opt. Soc. Am. (1961).

75. E. Snitzer, H. Po, R.P. Tumminelli, and F. Hakimi, "Optical fiber lasers and amplifiers," in US Pat. 4,815,079, (filed 1987), 1989.

76. R.J. Mears, L. Reekie, I.M. Jauncey, and D.N. Payne, "Low-noise erbim-doped fibre amplifier operating at 1.54 μm," Electron. Lett. **23**, 1026 (1987).

77. E. Desurvire, J.R. Simpson, and P.C. Becker, "High-gain erbium-doped traveling wave fiber amplifier," Opt. Lett. **12**, 888 (1987).

78. "TAT-8, the first trans-Atlantic fiber cable (2 $\times$ 280 Mb/s, 1.3 μm), was AT&T's 8th transatlantic telephone cable, in operation from 1988, initially carrying 40,000 telephone circuits between USA and France.."

79. "TAT-10, AT&T's 10th transatlantic telephone cable, in operation from 1992, initially carrying 2 $\times$ 565Mb/s between USA and Germany, using 1.5 μm technology."

80. Y. Arimoto, M. Toyoshima, M. Toyoda, T. Takahashi, M. Shikatani, and K. Araki, "Preliminary result on laser communication experiment using Engineering Test Satellite-VI (ETS-VI)," in Proc. SPIE 2381, 1995.

81. K.E. Wilson, J.R. Lesh, K. Araki, and Y. Arimoto, "Preliminary results of the Ground/Orbiter Lasercom Demonstration experiment between Table Mountain and the ETS-VI satellite," in Proc. SPIE (Free-Space Laser Communication Technologies VIII), 1996.

82. K.E. Wilson, et. al., "Results from Phase-1 and Phase-2 GOLD experiments," Feb. 15 1997.

83. K.E. Wilson and J.R. Lesh, "Overview of the Ground-to-Orbit Lasercom Demonstration (GOLD)," in Proc. SPIE (Free-Space Laser Communication Technologies IX), 1997.

84. "TAT-12/13 (1996) first used optical amplification (EDFA) and ring topology."

85. A.R. Chraplyvy, A.H. Gnauck, R.W. Tkach, J.L. Zyskind, J.W. Sulhoff, A.J. Lucero, Y. Sun, R.M. Jopson, F. Forghieri, R.M. Derosier, C. Wolf, and A.R. McCormick, "1-Tb/s transmission experiment," Photonics Technol. Lett. **8**, 1264–1266 (1996).

86. A.H. Gnauck, et. al., "One Terabit/s Transmission Experiment," presented at Optical Fiber Conference (OFC), 1996.

87. T. Morioka, et. al., "100 Gb/s × 10 Channel OTDM/WDM Transmission Using a Single Supercontinuum WDM Source," presented at Optical Fiber Conference (OFC), 1996.

88. H. Onaka, et. al., "1.1 Tb/s WDM Transmission over a 150 km 1.3 mm Zero-Dispersion Single-Mode Fiber," presented at Optical Fiber Conference (OFC), 1996.

89. KMI Corporation, Nashua, NH, USA.

90. S. Bigo, A. Bertaina, Y. Frignac, S. Borne, L. Lorcy, D. Hamoir, D. Bayart, J.-P. Hamaide, W. Idler, E. Lach, B. Franz, G. Veith, P. Sillard, L. Fleury, P. Guénot, and P. Nouchi, "5.12 Tbit/s (128 × 40 Gbit/s WDM) transmission over 3 × 100 km of TeraLight fiber," in Proc. Eur. Conf. Optical Communications (ECOC), paper PD1.2, Munich, Germany, 2000.

91. W. Idler, S. Bigo, Y. Frignac, B. Franz, and G. Veith, "Vestigial side-band demultiplexing for ultra-high capacity (0.64 bit/s/Hz) of 128 × 40 Gbit/s channels," in Optical Fiber Conference (OFC), 2001.

92. K. Fukuchi, T. Kasamatsu, M. Morie, R. Ohhira, T. Ito, K. Sekiya, D. Ogasahara, and T. Ono, "10,92 Tb/s (273 × 40 Gb/s) triple band/ultra dense WDM optical repeatered transmission experiment," presented at Optical Fiber Conference (OFC), 2001.

93. S. Bigo, Y. Frignac, G. Charlet, W. Idler, S. Borne, H. Gross, R. Dischler, W. Poehlmann, P. Tran, C. Simonneau, D. Bayart, G. Veith, A. Jourdan, and J.-P. Hamaide, "10.2 Tbit/s (256 × 42.7 Gbit/s PDM/WDM) transmission over 100 km TeraLight fiber with 1.28 bit/s/Hz spectral efficiency," in Optical Fiber Conference (OFC), paper PD25, 2001.

94. Y. Frignac, G. Charlet, W. Idler, R. Dischler, P. Tran, S. Lanne, S. Borne, C. Martinelli, G. Veith, A. Jourdan, J.-P. Hamaide, and S. Bigo, "Transmission of 256 wavelength-division and polarization-division multiplexed channels at 42.7 Gb/s (10.2 Tb/s capacity) over 3 × 100 km of TeraLight fiber," in Optical Fiber Conference (OFC), 2002.

95. "Press Release: NRO GeoLITE Satellite Successfully Launched". http://cartome.org/geolite .htm, 2001.

96. "Press Release: Delta Launches GeoLITE Satellite for U.S. NRO." http://www.spaceandtech.com/digest/flash2001/flash2001-038.shtml Andrews Space & Technology, 2001.

97. "Press Release: NRO Awarded the David Packard Excellence in Acquisition Award". http://www.nro.gov/PressReleases/prs_rel62.html, 2002.

98. T.T. Nielsen and G. Oppenhaeuser, "In-orbit test result of an operational optical intersatellite link between ARTEMIS and SPOT4, SILEX," in Proc. SPIE, (Free-Space Laser Communication Technologies XIV), 2002.

99. T.T. Nielsen, G. Oppenhaeuser, B. Laurent, and G. Planche, "In-orbit test results of the optical intersatellite link, SILEX. A milestone in satellite communication," in 53rd International Astronautical Congress, IAC-02-M.2.01, 2002.

100. H.P. Lutz, "Optical Communications in Space—Twenty Years of ESA Effort," in *ESA Bulletin (* http://esapub.esrin.esa.it/bulletin/bullet91/b91lutz.htm), **1**, 25–31 (1997).

101. B.I. Edelson, J.N. Pelton, C.W. Bostian, W.T. Brandon, V.W.S. Chan, E.P. Hager, N.R. Helm, R.D. Jennings, R.K. Kwan, C.E. Mahle, E. F. Miller, and L. Riley, "Satellite Communications Systems And Technology," in http://www.wtec.org/loyola/satcom/ ,Loyola College in Maryland, 1993.

102. M. Reyes, S. Chueca, A. Alonso, T. Viera, and Z. Sodnik, "Analysis of the preliminary optical links between ARTEMIS and the Optical Ground Station," in Proc. SPIE 4821, 2003.

103. A. Alonso, M. Reyes, and Z. Sodnik, "Performance of satellite-to-ground communications link between ARTEMIS and the Optical Ground Station," in Proc. SPIE 5572, 2004.

104. D.M. Boroson, C.C. Chen, and B.L. Edwards, "The Mars laser communications demonstration project: truly ultralong-haul optical transport," in Optical Fiber Conference (OFC), 2005.

105. M. Toyoshima, "Special Report: Trends of research and development of optical space communications technology," Space Japan Review 12 - 1, No. 44, 2005.

106. E. Hecht, *Optics, 2nd Ed.* (Addison Wesley, 1987).

107. A. Biswas, K.E. Wilson, S. Piazzolla, J.P. Wu, and W.H. Farr, "Deep-space optical communications link availability and data volume," in Proc. SPIE **5338**, 175 (2004).

108. F.I. Khatri, D.M. Boroson, D.V. Murphy, and J. Sharma, "Link analysis of Mars-Earth optical communications system," in Proc. SPIE **5338**, 143 (2004).

109. F.I. Khatri and A. Biswas, "Signal and Background Levels for the Mars Lasers Communications Demonstration (MLCD)," in IEEE LEOS Summer Topical Meetings, 2005.

110. R.W. Boyd, *Nonlinear Optics* (New York: Academic Press, Inc., 2003).

111. G.P. Agrawal, *Nonlinear Fiber Optics*, 2^{nd} ed. (New York: Academic Press, Inc., 1995).

112. W.E. Webb and J.T. Marino, Jr., "Threshold detection in an on-off binary communications channel with atmospheric scintillation," Appl. Opt. **14**, 1413–1417 (1975).

113. W.C. Brown, "Optimum Thresholds for Optical On-Off Keying Receivers Operating in the Turbulent Atmosphere," in Proc. SPIE, (Free-Space Laser Communication Technologies IX), vol. 2290, 254–261 (1997).

114. H. Haunstein, R. Schlenk, and K. Sticht, "Control of Combined Electrical Feed-Forward and Decision Feedback Equalization by Conditional Error Counts from FEC in the Presence of PMD," in Optical Fiber Conference (OFC), 2003.

115. D.M. Castagnozzi, "Digital signal processing and electronic equalization (EE) of ISI," in Optical Fiber Conference (OFC), 2004.

116. M. Schwartz, W.R. Bennett, and S. Stein, *Communication Systems and Techniques* (New York: IEEE Press, 1996).

117. S.D. Personick, P. Balaban, J. Bobsin, and P. Kumar, "A Detailed Comparison of Four Approaches to the Calculation of the Sensitivity of Optical Fiber System Receivers," IEEE Trans. Comm. **25**, 541–548 (1977).

118. R.G. Smith and S.D. Personick, *Semiconductor devices for optical communication* (New York: Springer-Verlag, 1982), vol. 39.

119. S.O. Rice, "Mathematical analysis of random noise," Bell Syst. Technol. J. **24**, 46–156 (1945).

120. S. Stein and J.J. Jones, *Modern communication principles* (New York: McGraw-Hill, 1967).

121. Saleh, B.E.A. and M.C. Teich, *Fundamentals of Photonics* (New York: Wiley, 1991).

122. T. Li and M.C. Teich, "Bit-Error Rate For A Lightwave Communication System Fibre Amplifier Incorporating An Erbium-doped Fibre Amplifier," Electron. Lett. **27**, 598–599 (1991).

123. T. Li and M.C. Teich, "Photon Point Process for Traveling-Wave Laser Amplifiers," IEEE J. Quantum Electron. **29**, 2568–2578 (1993).

124. W.S. Wong, H.A. Haus, L.A. Jiang, P.B. Hansen, and M. Margalit, "Photon statistics of amplified spontaneous emission noise in a 10-Gbitys optically preamplified direct-detection receiver," Opt. Lett. **23**, 1832-834 (1998).

125. W.S. Wong, J.D. Moores, J. Korn, and H.A. Haus, "Photon statistics of NRZ signals in high-bit-rate optically pre-amplified direct detection receiver," Optical Fiber Conference (OFC), 1999.

126. E. Desurvire, *Erbium-doped fiber amplifiers* (New York: John Wiley & Sons, 1994).

127. H.A. Haus, *Electromagnetic Noise and Quantum Optical Measurements* (Springer-Verlag, 2000).

128. J.C. Livas, "High sensitivity optically preamplified 10 Gb/s receivers," in Optical Fiber Conference (OFC), Paper PD4, 1996.

129. W.A. Atia and R.S. Bondurant, "Demonstration of return-to-zero signaling in both OOK and DPSK formats to improve receiver sensitivity in an optically preamplified receiver," in LEOS, 1999.

130. A.H. Gnauck, S. Chandrasekhar, J. Leuthold, and L. Stulz, "Demonstration of 42.7-Gb/s DPSK receiver with 45 photons/bit sensitivity," IEEE Photon. Technol. Lett. **15**, 99–101 (2003).

131. J.H. Sinsky, A. Adamiecki, A.H. Gnauck, C.A. Burrus, J. Leuthold, O. Wohlgemuth, and A. Umbach, "A 42.7-Gb/s Integrated Balanced Optical Front End with Record Sensitivity," in Optical Fiber Conference (OFC): Paper PD39-1, 2003.

132. A.H. Gnauck and P.J. Winzer, "Optical phase-shift-keyed transmission," J. Lightwave Technol. **23**, 115–130 (2005).

133. D.O. Caplan, M.L. Stevens, J.J. Carney, and R.J. Murphy, "Demonstration of Optical DPSK Communication with 25 Photons/Bit Sensitivity," in Conference on Lasers and Electro-Optics (CLEO), 2006.

134. J.R. Minch, D. J. Townsend, and D.R. Gervais, "Rate Adjustable NRZ-DPSK Modulation Scheme with a Fixed Interferometer," in IEEE LEOS, 2005.

135. D.O. Caplan, M.L. Stevens, and J.J. Carney, "A High-Sensitivity Multi-Channel Single-Interferometer DPSK Receiver," Opt. Express **14**, 10984–10989 (2006).

136. D.O. Caplan, M.L. Stevens, and J.J. Carney, "High-Sensitivity Demodulation of Multiple-Data-Rate WDM-DPSK Signals using a Single Interferometer," in Optical Fiber Conference (OFC), 2007.

137. K. Yonenaga and K. Hagimoto, "10-Git/s × four-channel WDM transmission experiment over 2400-km DSF using optical DPSK direct detection scheme," in Optical Fiber Conference (OFC), 1997.

138. A.H. Gnauck, G. Raybon, S. Chandrasekhar, J. Leuthold, C. Doerr, L. Stulz, A. Agarwal, S. Banerjee, D. Grosz, S. Hunsche, A. Kung, A. Marhelyuk, D. Maywar, M. Movassaghi, X. Liu, C. Xu, X. Wei, and D.M. Gill, "2.5 Tb/s (64 × 42.7 Gb/s) transmission over 40 × 100 km NZDSF using RZ-DPSK format and all-Raman-amplified spans," in Optical Fiber Conference (OFC), 2002.

139. J.-X. Cai, D.G. Foursa, C.R. Davidson, Y. Cai, G. Domagala, H. Li, L. Liu, W. Patterson, A. Pilipetskii, M. Nissov, and N. Bergano, "A DWDM Demonstration of 3.73 Tb/s over 11,000 km using 373 RZ-DPSK Channels at 10 Gb/s," in Optical Fiber Conference (OFC): Paper PD22-1, 2003.

140. C. Rasmussen, T. Fjelde, J. Bennike, F. Liu, S. Dey, B. Mikkelsen, P. Mamyshev, P. Serbe, P. van der Wagt, Y. Akasaka, D. Harris, D. Gapontsev, V. Ivshin, and P. Reeves-Hall, "DWDM

40G transmission over trans-Pacific distance (10 000 km) using CSRZ-DPSK, enhanced FEC and all-Raman amplified 100 km Ultrawave fiber spans," in Optical Fiber Conference (OFC): Paper PD18-1, 2003.

141. B. Zhu, L. E. Nelson, S. Stulz, A.H. Gnauck, C. Doerr, J. Leuthold, L. Gruner-Nielsen, M.O. Pedersen, J. Kim, R.L. Lingle, Jr., Y. Emori, Y. Ohki, N. Tsukiji, A. Oguri, and S. Namiki, "6.4 Tbit/s (160×42.7 Gb/s) transmission with 0.8 bit/s/Hz spectral efficiency over 32×100 km of fiber using CSRZ-DPSK format," in Optical Fiber Conference (OFC): Paper PD19, 2003.
142. M.L. Stevens, "Transmitter pulse issues," private communication, 2004.
143. R.A. Griffin, "Integrated DQPSK Transmitters," in Optical Fiber Conference (OFC), 2005.
144. H. Kim and P.J. Winzer, "Robustness to laser frequency offset in direct-detection DPSK and DQPSK systems," J. Lightwave Technol. **21**, 1887–1891 (2003).
145. G. Bosco and P. Poggiolini, "On the joint effect of receiver impairments on direct-detection DQPSK systems," J. Lightwave Technol. **24**, 1323–1333 (2006).
146. P.J. Winzer and H. Kim, "Degradations in balanced DPSK receivers," IEEE Photon. Technol. Lett. **15**, 1282–1284 (2003).
147. I. Morita and N. Yoshikane, "Merits of DQPSK for Ultrahigh Capacity Transmission," in LEOS, 2005.
148. A. Royset and D.R. Hjelme, "Novel dispersion tolerant optical duobinary transmitter using phase modulator and Bragg grating filter," in ECOC, 1998.
149. Y. Miyamoto, H. Masuda, A. Hirano, S. Kuwahara, Y. Kisaka, H. Kawakami, M. Tomizawa, Y. Tada, and S. Aozasa, "S-band WDM coherent transmission of 40 × 43-Gbit/s CS-RZ DPSK signals over 400 km DSF using hybrid GS-TDFAs/Raman amplifiers," Electron. Lett. **39**, 1569–1570 (2002).
150. D. Penninckx, H. Bissessur, P. Brindel, E. Gohin, and F. Bakhti, "Optical differential phase shift keying (DPSK) direct detection considered as a duobinary signal," in ECOC, 2001.
151. Y. Miyamoto, A. Hirano, S. Kuwahara, M. Tomizawa, and Y. Tada, "Novel modulation and detection for bandwidth-reduced RZ formats using duobinary-mode splitting in wideband PSK/ASK conversion," J. Lightwave Technol. **20**, 2067–2078 (2002).
152. L. Moller, C. Xie, R. Ryf, L. Xiang, and X. Wei, "10 Gb/s duobinary receiver with a record sensitivity of 88 photons per bit," presented at Optical Fiber Conference (OFC), 2004.
153. L.G. Kazovsky and D.A. Atlas., "PSK synchronous heterodyne and homodyne experiments using optical phase-locked loops," presented at Optical Fiber Conference (OFC), 1990.
154. L.G. Kazovsky and D.A. Atlas, "A 1320-nm Experimental Optical Phase-Locked Loop: Performance Investigation and PSK Homodyne Experiments at 140 Mb/s and 2 Gb/s," J. Lightwave Technol. **8**, 1414–1425 (1990).
155. F.T. Herzog, "An optical phase locked loop for coherent space communications," Swiss Federal Institute of Technology, 2006.
156. L.G. Kazovsky, S. Benedetto, and A. Willner, *Optical Fiber Communication Systems* (Norwood, MA: Artech House, Inc., 1996).
157. B. Wandernoth, "5 photon/bit low complexity 2 Mbit/s PSK transmission breadboard experiment with homodyne receiver applying synchronization bits and convolutional coding," Proc. Eur. Conf. Opt. Commun. (ECOC), **1**, 59–62 (1994).
158. R. Lange and B. Smutny, "Highly-Coherent Optical Terminal Design Status and outlook," in LEOS, 2005.

159. J.H. Sinsky, A. Adamiecki, A.H. Gnauck, C.A. Burrus, J. Leuthold, and O. Wohlgemuth, "RZ-DPSK Transmission Using a 42.7-Gb/s Integrated Balanced Optical Front End With Record Sensitivity," J. Lightwave Technol. **22**, (2004).

160. S.B. Alexander, R. Barry, D.M. Castagnozzi, V.W.S. Chan, D.M. Hodsdon, L.L. Jeromin, J.E. Kaufmann, D.M. Materna, R.J. Parr, M.L. Stevens, and D.W. White, "4-ary FSK coherent optical communication system," Electron. Lett. **26**, 1346–1348 (1990).

161. M.L. Stevens, D.M. Boroson, and J.E. Kaufmann, "A near-optimum discriminator demodulator for binary FSK with wide tone spacing," IEEE Microwave and Guided Wave Lett. **3**, 227–229 (1993).

162. R. Noe, et. al., "Optical FSK transmission with pattern independent 119 photoelectrons/bit receiver sensitivity with endless polarization control," Electron. Lett. **25**, 757–758 (1989).

163. B.S. Robinson "Semiconductor-based all-optical switching for optical time-division multiplexed networks," Thesis, MIT, 2003.

164. J.M. Kahn, "1 Gbit/s PSK homodyne transmission system using phase-locked semiconductor lasers," IEEE Photon. Technol. Lett. **1**, 340–342 (1989).

165. S. Norimatsu, K. Iwashita, and K. Noguchi, "10 Gbit/s optical PSK homodyne transmission experiments using external cavity DFB LDs," Electron. Lett. **26**, 648–649 (1990).

166. S. Norimatsu, K. Iwashita, and K. Sato, "PSK optical homodyne detection using external cavity laser diodes in Costas loop," IEEE Photon. Technol. Lett. **2**, 374–376 (1990).

167. B. Zhu, L. Leng, L. E. Nelson, L. Gruner-Nielsen, Y. Qian, J. Bromage, S. Stulz, S. Kado, Y. Emori, S. Namiki, P. Gaarde, A. Judy, B. Palsdottir, and R.L. Lingle, Jr., "3.2 Tb/s (80 × 42.7 Gb/s) transmission over 20 × 100km of nonzero dispersion fiber with simultaneous C + L-band dispersion compensation," in Optical Fiber Conference (OFC), 2002.

168. J.M. Ross, S. I. Green, and J. Brand, "Short-pulse optical communication experiments," Proc. IEEE **58**, (1970).

169. J.R. Lesh, J. Katz, H.H. Tan, and D. Zwillinger, "2.5 bit/detected photon demonstration program: description, analysis, and phase 1 results," Jet Propulsion Laboratory, Pasadena, CA 42-66, Dec. 1981.

170. J.R. Lesh, "Capacity Limit of the Noiseless, Energy-Efficient Optical PPM Channel," IEEE Trans. Comm., **31**, 546–548 (1983).

171. A.J. Phillips, R.A. Cryan, and J.M. Senior, "An optically preamplified intersatellite PPM receiver employing maximum likelihood detection," IEEE Photon. Technol. Lett. **8**, 691–693, 1996.

172. D.O. Caplan, P.W. Juodawlkis, J.J. Plant, and M.L. Stevens, "Performance of high-sensitivity OOK, PPM, and DPSK communications using high-power slab-coupled optical waveguide amplifier (SCOWA) based transmitters," in Optical Fiber Conference (OFC), 2006.

173. C. Chen and C.S. Gardner, "Performance of PLL Synchronized Optical PPM Communication Systems," IEEE Trans. Comm. **COM-34**, 988–994 (1986).

174. F. Nekoogar, *Ultra-Wideband Communications: Fundamentals and Applications* (Upper Saddle River: Prentice Hall, 2005).

175. M.L. Stevens, "Estimation of M-PPM Spectra for Pseudo-Random Bit Sequences." private communication, 2006.

176. B.S. Robinson, S.A. Hamilton, and E.P. Ippen, "Demultiplexing of 80 Gbit/s pulse-position modulated data with an ultrafast nonlinear interferometer," IEEE Photon. Technol. Lett. **14**, 2002.

177. R.S. Vodhanel, J.L. Gimlett, N.K. Cheung, and S. Tsuji, "FSK Heterodyne Transmission Experiments at 560 Mbit/s and 1 Gbit/s," J. Lightwave Technol. **LT-5**, 461–468 (1987).

178. A.R. Chraplyvy, R.W. Tkach, A.H. Gnauck, and R.M. Derosier, "8Gbit/s FSK modulation of DFB lasers with optical demodulation," Electron. Lett., **25**, 319–321 (1989).

179. B. Glance, et. al., "Densely spaced FDM optical coherent system with near quantum-limited sensitivity and computer controlled random access channel selection," in Optical Fiber Conference (OFC): Paper PD11, 1989.

180. M.L. Stevens, B. R. Hemenway, and S.B. Alexander, "Simultaneous TDM/FDM digital optical transmission with polarization-diversity heterodyne detection," Microwave Symposium Digest, **1**, 171–173 (1994).

181. R. Gross and R. Olshansky, "Multichannel Coherent FSK Experiments Using Subcarrier Multiplexing Techniques," J. Lightwave Technol. **8**, 406–415 (1990).

182. H. Gross, M. Schmidt, R. Olshansky, and V. Lanzisera, "Coherent Transmission of 60 FM-SCM Video Channels," Photonics Technol. Lett. **2**, 288–290 (1990).

183. S. Bendetto, R. Gaudino, and P. Poggiolini, "Direct detection of optical digital transmission based on polarization shift keying modulation," IEEE Sel. Areas Commun., **13**, 531–542 (1995).

184. E. Hu, K. Wong, M. Marhic, L.G. Kazovsky, K. Shimizu, and N. Nikuchi, "4-Level Direct-Detection Polarization Shift-Keying (DD-PolSK) System with Phase Modulators," in Optical Fiber Conference (OFC), 2003.

185. M.M. Matalgah and R.M. Radaydeh, "Hybrid Frequency-Polarization Shift-Keying Modulation for Optical Transmission," J. Lightwave Technol. **23**, 1152–1162 (2005).

186. M. Nazarathy and E. Simony, "Generalized Stokes Parameters-Shift Keying: A New Perspective on Optimal Detection Over Electrical and Optical Vector Incoherent Channels," IEEE Trans. Comm. **54**, 499–509 (2006).

187. M. Nazarathy and E. Simony, "Stokes Space Optimal Detection of Multidifferential Phase and Polarization Shift Keying Modulation," J. Lightwave Technol. **24**, 1978–1988 (2006).

188. R. Zhang and G.S. La Rue, "Clock and data recovery circuits with fast acquisition and low jitter," presented at IEEE Workshop on Microelectronics and Electron Devices, 2004.

189. N.W. Spellmeyer, "Communications performance of a multimode EDFA," IEEE Photon. Technol. Lett. **12**, 1337–1339 (2000).

190. R.H. Kingston, *Detection of Optical and Infrared Radiation* (New York: Springer-Verlag, 1978).

191. B. Glance, "Polarization independent coherent optical receiver," J. Lightwave Technol. **5**, 274–276 (1987).

192. B.S. Glance, K. Pollock, C.A. Burrus, B.L. Kasper, G. Einstein, and L.W. Stulz, "WDM coherent optical star network," J. Lightwave Technol. **6**, 67–72 (1988).

193. B.S. Glance and M. Kavehrad, "Polarization-insensitive frequency-shift-keying optical heterodyne receiver using discriminator demodulation," J. Lightwave Technol. **6**, 1386–1394 (1988).

194. N. Ohkawa, T. Sugie, and Y. Hayashi, "A highly sensitive balanced receiver for 2.5 Gb/s het-erodyne detection systems," IEEE Photon. Technol. Lett. **3**, 375–377 (1991).

195. B. Wandernoth, "20 photon/bit 565 Mbit/s PSK homodyne receiver using synchronization bits," Electron. Lett. **28**, (1992).

196. F.H. Raab, P. Asbeck, S. Cripps, P.B. Kenington, Z.B. Popovic, N. Pothecary, J.F. Sevic, and N.O. Sokal, "RF and Microwave Power Amplifier and Transmitter Technologies—Part 1," High Frequency Electronics, 22–36 (2003).

197. P.W. Juodawlkis, J.J. Plant, R.K. Huang, L.J. Missaggia, and J.P. Donnelly, "High-power 1.5-/spl mu/m InGaAsP-InP slab-coupled optical waveguide amplifier," IEEE Photon. Technol. Lett. **17**, (2005).

198. N.W. Spellmeyer, D.O. Caplan, and M.L. Stevens, "Design of a 5-Watt PPM transmitter for the Mars Laser Communications Demonstration," in LEOS, 2005.

199. P. Wysocki, T. Wood, A. Grant, D. Holcomb, K. Chang, M. Santo, L. Braun, and G. Johnson, "High Reliability 49 dB Gain, 13W PM Fiber Amplifier at 1550 nm with 30 dB PER and Record Efficiency," in Optical Fiber Conference (OFC), paper PDP17, 2006.

200. L. Goldberg, J.P. Koplow, and D.A.V. Kliner, "Highly efficient 4-W Yb-doped fiber amplifier pumped by a broad-stripe laser diode," Opt. Lett. **24**, 673–675 (1999).

201. A.N. Curren, J.A. Dayton, Jr., R.W. Palmer, K.J. Long, D.A. Force, C.E. Weeder, Z.A. Zachar, and W.L. Harvey, "The Cassini mission Ka-band TWT," in International Electron Devices Meeting, 1994.

202. D. Morabito, S. Butman, and S. Shambayati, "The Mars Global Surveyor Ka-Band Link Experiment (MGS/KaBLE-II)," in Telecommunications and Mission Operations Progress Report 42-137, Jet Propulsion Laboratory, 1999.

203. I. Haque, "Ka-Band Traveling Wave Tube Amplifier," in IND Technology and Sciency News, Jet Propulsion Laboratory, 2002, pp. 11–14: http://tmot.jpl.nasa.gov/Program_Overview_Information/IND_Program_News/Issue15.pdf.

204. N.W. Spellmeyer, D.O. Caplan, B.S. Robinson , D. Sandberg, M.L. Stevens, M.M. Willis, D.V. Gapontsev, N.S. Platonov, and A. Yusim, "A High-Efficiency Ytterbium-Doped Fiber Amplifier Designed for Interplanetary Laser Communications," in Optical Fiber Conference (OFC), 2007.

205. R. Loudon, *The Quantum Theory of Light* (New York: Oxford University Press, Inc., 2000).

206. C.W. Gardiner and P. Zoller, *Quantum Noise* (New York: Springer Verlag, 2000).

207. J.W. Goodman, *Statistical Optics* (New York: John Wiley & Sons, Inc., 2000).

208. A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (New York: McGraw-Hill, Inc., 1984).

209. R. Sobolewski, et. al., "Ultrafast superconducting single-photon optical detectors and their applications," IEEE Trans. Appl Superconduct **13**, 1151–1157 (2003).

210. J.K.W. Yang, et. al., "Fabrication development for nanowire GHz-counting-rate single-photon detectors," IEEE Trans. Appl. Superconduct., 2005.

211. B.S. Robinson, "Private communication," 2005.

212. J.H. Shapiro, "Imaging and Optical Communication through Atmospheric Turbulence," in *Laser Beam Propagation in the Atmosphere*, J.W. Strohbehn, Ed. (Berlin: Springer-Verlag, 1978).

213. R.J. Barron, "Binary shaping for low-duty-cycle communications," in International Symposium on Information Theory (ISIT), 2004.

214. B.S. Robinson, D.O. Caplan, M.L. Stevens, R.J. Barron, E.A. Dauler, and S.A. Hamilton, "1.5-photons/bit Photon-Counting Optical Communications Using Geiger-Mode Avalanche Photodiodes," in IEEE LEOS Summer Topical Meetings, 2005.

215. J.B. Johnson, "Thermal agitation of electricity in conductors," Phys. Rev. **32**, 97–109 (1928).

216. H. Nyquist, "Thermal agitation of electric charge in conductors," Phys. Rev. **32**, 110–113 (1928).

217. B. M. Oliver, "Thermal and quantum noise," presented at Proc. IEEE, 1965.

218. A. Yariv, *Optical Electronics in Modern Communications, 5th Ed.* (New York: Oxford Un. Press, 1997).

219. S.G. Lambert and W.L. Casey, *Laser Communications in Space* (Boston: Artech House, 1995).

220. ITU-R Recommendation PI.372-6: Radio Noise, 1994.

221. K. P. Phillips, "An overview of propagation factors influencing the design of mobile satellite communication systems," Electron. & Commun. Eng. J., (1997).

222. K. Rosfjord, J. Yang, E. Dauler, A. Kerman, V. Anant, B. Voronov, G. Gol'tsman, and K. Berggren, "Nanowire single-photon detector with an integrated optical cavity and anti-reflection coating," Opt. Express, **14**, 527–534 (2006).

223. W.R. Leeb, "Degradation of signal to noise ratio in optical free space data links due to background illumination," Appl. Opt. **28**, 3443–3449 (1989).

224. F.D. Teodoro, J.P. Koplow, and S.W. Moore, "Diffraction limited, 300-kW peak power pulses from a coiled multimode fiber amplifier," Opt. Lett., (2002).

225. V.W.S. Chan, "Space coherent optical communication systems–An introduction," J. Lightwave Technol. **5**, 633–637 (1987).

226. S.B. Alexander, E.S. Kintzer, and J.C. Livas, "A Gbps, 1 Watt Free-space Coherent Optical Communication System," in LEOS, 1992.

227. A. Biswas, H. Hemmati, and J.R. Lesh, "High data-rate laser transmitters for free space laser communications," in Photonics West, 1999.

228. E. Rochat, R. Dändliker, K. Haroud, R.H. Czichy, U. Roth, D. Costantini, and R. Holzner, "Fiber Amplifiers for Coherent Space Communication," IEEE Sel. Top. Quantum Electron., **7**, 64–81 (2001).

229. R. Lange and B. Smutny, "BPSK Laser Communication Terminals to be verified in space," in Milcom, 2004, pp. 441–444.

230. R. Lange and B. Smutny, "Optical inter-satellite links based on homodyne BPSK modulation: Heritage, status and outlook," in Proc. SPIE, (Free-Space Laser Communication Technologies XVII), 2005.

231. T. Shinagawa, "Detailed investigation on reliability of wavelength-monitor-integrated fixed and tunable DFB laser diode modules," J. Lightwave Technol. **23**, 1126–1136 (2005).

232. T. Ikegami and Y. Suematsu, "Resonance-like characteristics of the direct modulation of a junction laser," presented at Proc. IEEE, 1967.

233. L.A. Coldren and S.W. Corzine, *Diode Lasers and Photonic Integrated Circuits* (John Wiley & Sons, Inc., 1995).

234. D.A. Ackerman, "Laser diodes for CATV," in IEEE LEOS, 1995.

235. M.R. Phillips, et. al., "112 channel split-band WDM lightwave CATV system," IEEE Photon. Technol. Lett. **4**, 790–792 (1992).

236. G.P. Agrawal, *Fiber-Optic Communication Systems* (New York: John Wiley & Sons, 1992).

237. J. Piprek and J.E. Bowers, "Analog modulation of semiconductor lasers," in *RF Photonic Technology in Optical Fiber Links*, W.S. C. Chang, Ed. (Cambridge Univ. Press, 2002).

238. R.S. Vodhanel, A. F. Elrefaie, M. Z. Iqbal, R. E. Wagner, J.L. Gimlett, and S. Tsuji, "Performance of directly modulated DFB lasers in 10-Gb/s ASK, FSK, and DPSK lightwave systems," J. Lightwave Technol. **8**, 1379–1386 (1990).

239. J. Zhang, N. Chi, P. V. Holm-Nielsen, C. Pueucheret, and P. Jeppesen, "An Optical FSK Transmitter Based on an Integrated DFB Laser–EA Modulator and Its Application in Optical Labeling," IEEE Photon. Technol. Lett. **15**, 984–986 (2003).

240. F. Koyama and K. Iga, "Frequency chirping in external modulators," J. Lightwave Technol. **6**, 87–93 (1988).

241. M. Kuznetsov, J. Stone, and L.W. Stulz, "Time- and frequency-resolved measurements of frequency-modulation and switching of a tunable semiconductor laser," Appl. Phys. Lett. **59**, 2492–2494 (1991).

242. M. Kuznetsov and D.O. Caplan, "Time-frequency analysis of optical communication signals and the effects of second and third order dispersion," in Conference on Lasers and Electro-Optics (CLEO), 2000.

243. R.A. Linke, "Modulation Induced Transient Chirping in Single Frequency Lasers," IEEE J. Quantum Electron. **QE-21**, 593–597 (1985).

244. P.A. Morton, G.E. Shtengel, L.D. Tzeng, R.D. Yadvish, T. Tanbun-Ek, and R.A. Logan, "38.5 km error free transmission at 10 Gbit/s in standard fibre using a low chirp, spectrally filtered, directly modulated 1.55 μm DFB laser," Electron. Lett. **33**, 310–311 (1997).

245. D.R. Hjelme and A. Royset, "RZ versus NRZ in space communication system using direct current modulated transmitter and optically pre-amplifed receiver with ultra-narrow optical filter," in LEOS, San Francisco, 1999.

246. M.M. Strasser, P.J. Winzer, M. Pfennigbauer, and W.R. Leeb, "Significance of Chirp-Parameter for Direct Detection Free-Space Laser Communication," in SPIE, 2001.

247. P. Corvini and T. Koch, "Computer simulation of high-bit-rate optical fiber transmission using single-frequency lasers," J. Lightwave Technol. **5**, 1591–1595 (1987).

248. D. Mahgerefteh, P.S. Cho, J. Goldhar, and H. I. Mandelberg, "Penalty-free propagation over 600 km of nondispersionshifted fiber at 2.5 Gb/s using a directly laser modulated transmitter," in Conference on Lasers and Electro-Optics (CLEO), 1999.

249. Y. Matsui, D. Mahgerefteh, X. Zheng, C. Liao, Z.F. Fan, K. McCallion, and P. Tayebati, "Chirp-Managed Directly Modulated Laser (CML)," Photon. Technol. Lett. **18**, 385–386 (2006).

250. S. Chandrasekbar, C.R. Doerr, L.L. Buhl, Y. Matsui, D. Mahgerefteh, X. Zheng, K. McCallion, Z. Fan, and P. Tayebati, "Repeaterless Transmission With Negative Penalty Over 285 km at 10 Gb/s Using a Chirp Managed Laser," Photon. Technol. Lett. **17**, 2454–2457 (2005).

251. S. Chandrasekbar, A.H. Gnauck, G. Raybon, L.L. Buhl, D. Mahgerefteh, X. Zheng, Y. Matsui, K. McCallion, Z. Fan, and P. Tayebati, "Chirp-Managed Laser and MLSE-RX Enables Transmission Over 1200 km at 1550 nm in a DWDM Environment in NZDSF at 10 Gb/s Without Any Optical Dispersion Compensation " Photonics Technol. Lett. **18**, 1560–1562 (2006).

252. M. Ito and T. Kimura, "Stationary and transient thermal properties of semiconductor laser diodes," IEEE J. Quantum Electron. **17**, 787–795 (1981).

253. H. Shalom, A. Zadok, M. Tur, P.J. Legg, W.D. Cornwell, and I. Andonovic, "On the Various Time Constants of Wavelength Changes of a DFB Laser Under Direct Modulation," IEEE J. Quantum Electron. **34**, 1816–1822 (1998).

254. A. Ma, J.C. Cartledge, and H.E. Lassen, "Performance implications of the thermal-induced frequency drift in fast wavelength switched systems with heterodyne detection," J. Lightwave Technol. **14**, 1090–1096 (1996).

255. D.O. Caplan, G.S. Kanter, and P. Kumar, "Characterization of dynamic optical nonlinearities by continuous time-resolved Z-Scan," Opt. Lett, **21**, 1342–1344 (1996).

256. A. A. Saavedra, R. Passy, and J.P. von der Weid, "Thermal drift in wavelength-switching DFB and DBR lasers," Electron. Lett. **33**, 780–781 (1997).

257. C.R. Giles, T. Erdogan, and V. Mizrahi, "Simultaneous wavelength-stabilization of 980-nm pump lasers," IEEE Photonics Technol. Lett. **6**, (1994).

258. S. Mohrdiek, T. Plisk, and C. Harder, "Coolerless operation of 980 nm pump modules," in Optical Fiber Conference (OFC), 2001.

259. J.-L. Archambault and S.G. Grubb, "Fiber Gratings in Lasers and Amplifiers," J. Lightwave Technol. **15**, 1378–1390 (1997).

260. B. R. Hemenway and M.L. Stevens, "Simultaneous TDM/FDM Using Rapidly-tunable Transmitters and Receivers For Multi-access Optical Networks," presented at IEEE/LEOS Integrated Optoelectronics Proceedings, 1994.

261. C.H. Henry, "Theory of the linewidth of semiconductor lasers," IEEE J. Quantum Electron. **QE-18**, 259–264 (1982).

262. K. Kojima, K. Kyuma, and T. Nakayama, "Analysis of the spectral linewidth of distributed feedback laser diodes," J. Lightwave Technol. **3**, 1048–1055 (1985).

263. H. Nakano, S. Sasaki, S. Tsuji, N. Chinone, and M. Maeda, "Comparison of optical reflection tolerance between conventional and ?/4-shifted DFB lasers in a 2.4 Gbit/s system," Electron. Lett. **4**, 1049–1051 (1988).

264. G. Jacobsen, "Performance of DPSK and CPFSK Systems with Significant Post-Detection Filtering," J. Lightwave Technol. **11**, (1993).

265. H. Nasu, T. Mukaihara, T. Takagi, M. Oike, T. Nomura, and A. Kasukawa, "25-GHz-spacing wavelength-monitor integrated DFB laser module for DWDM applications," Photon. Technol. Lett. **15**, 293–295 (2003).

266. D.O. Caplan, "Multi-channel DPSK Receiver," in US Pat. Appl. 11/022,344, 2004.

267. D.O. Caplan, "Polarization independent optical interferometers," in US Pat. Appl., 2004.

268. D.O. Caplan, "Reconfigurable Polarization Independent Interferometers and Methods of Stabilization," in US Pat. Appl. 11/318,255, 2005.

269. J.B. Abshire, et. al., "The Geoscience Laser Altimeter System (GLAS) for the ICESat mission," presented at Conference on Lasers and Electro-Optics (CLEO), 2000.

270. M. Albota, et. al. , "Three-dimensional imaging laser radars with Geiger-mode avalanche photodiode arrays," MIT Lincoln Lab. J. **13**, 351–370 (2002).

271. W.T. Roberts, "Cavity-dumped communication laser design," 42–152, Feb. 2003.

272. X. Sun, et. al., "Cloud and aerosol lidar channel design and performance of the Geoscience Laser Altimeter System on the ICESat mission," presented at Conference on Lasers and Electro-Optics (CLEO), 2004.

273. X. Sun, et. al., "Design and performance measurement of the mercury laser altimeter," in Conference on Lasers and Electro-Optics (CLEO), 2004.

274. M. Aoki, M. Suzuki, H. Sano, T. Kawano, T. Ido, T. Taniwatari, K. Uomi, and A. Takai, "InGaAs/InGaAsP MQW electroabsorption modulator integrated with a DFB laser fabricated by band-gap energy control selective area MOCVD," IEEE J. Quantum Electron. **29**, 2088–2096 (1993).

275. W.S.C. Chang, "Multiple quantum well electroabsorption modulators for RF photonic links," in *RF Photonic Technology in Optical Fiber Links*, W.S.C. Chang, Ed. (Cambridge Univ. Press, 2002).

276. G. Raybon, U. Koren, M.G. Young, B.I. Miller, M. Chien, T.H. Wood, and H. M. Presby, "Low chirp transmission at 5.0 Gbit/s using an integrated DBR laser-modulator transmitter," Electron. Lett. **30**, 1330–1331 (1994).

277. Y. Akage, K. Kawano, S. Oku, R. Iga, H. Okamoto, Y. Miyamoto, and H. Takeuchi, "Wide bandwidth of over 50 GHz travelling-wave electrode electroabsorption modulator integrated DFB lasers," Electron. Lett. **37**, 299–300 (2001).

278. A.H. Gnauck, S.K. Korotky, J.J. Veselka, J. Nagel, C.T. Kemmerer, W. J. Minford, and D. T. Moser, "Dispersion penalty reduction using an optical modulator with adjustable chirp," IEEE Photon. Technol. Lett. **3**, 916–918 (1991).

279. A.H. Gnauck, S.K. Korotky, and J.E. Zucker, "Tunable chirp, lightwave modulator for dispersion compensation," in US Pat. 5,303,079,. USA, 1992.

280. R.M. Jopson and A.H. Gnauck, "Dispersion compensation for optical fiber systems," IEEE Communi. Mag., **33**, 96–102 (1995).

281. L. Zehnder, "Ein neuer Interferenzrefractor," Z. Instrkde **11**, 275–285 (1891).

282. L. Mach, "Uber einer Interferenzrefractor," Z. Instrkde **12**, 89–93 (1892).

283. R.C. Alferness, "Waveguide Electrooptic Modulators,"IEEE Trans. Microwave Theory Technol., **82**, 1121–1137 (1982).

284. R.C. Alferness, "Corrections to "Waveguide Electrooptic Modulators"," IEEE Trans. Microwave Theory Technol. **83**, 315 (1983).

285. R.C. Alferness, L.L. Buhl, J.L. Jackel, S. P. Lyman, and V. Ramaswamy, " Fabrication method for $LiNbO_3$ and $LiTaO_3$ integrated optics devices," in U.S. Pat. 4,439,265, USA, 1984.

286. L. Thylen, "Integrated optics in LiNbO3: recent developments in devices for telecommunication," J. Lightwave Technol. **6**, 847– 861 (1988).

287. K. Noguchi, O. Mitomi, and H. Miyazawa, "Low-voltage and broadband Ti:$LiNbO_3$ modulators operating in the millimeter wavelength region," in Optical Fiber Conference (OFC), 1996.

288. K. Noguchi, O. Mitomi, and H. Miyazawa, "Millimeter-wave Ti:$LiNbO_3$ optical modulators," J. Lightwave Technol. **16**, 615–619 (1998).

289. G.E. Betts, "$LiNbO_3$ external modulators and their use in high performance analog links," in *RF Photonic Technology in Optical Fiber Links*, W.S. C. Chang, Ed. (Cambridge Univ. Press, 2002).

290. M.M. Howerton and W.K. Burns, "Broadband traveling wave modulators in $LiNbO_3$," in *RF Photonic Technology in Optical Fiber Links*, W.S. C. Chang, Ed. (Cambridge Univ. Press, 2002).

291. O. Leclerc, et. al., "40 Gbit/s polarization-independent, push-pull InP Mach-Zehnder modulator for all-optical regeneration," presented at Optical Fiber Conference (OFC), PD35, 1999.

292. O. Leclerc, B. Dany, D. Rouvillain, P. Brindel, E. Desurvire, C. Duchet, A. Shen, F. Devaux, E. Coquelin, M. Goix, S. Bouchoule, L. Fleury, and P. Nouchi, "Simultaneously regenerated 4×40 Gbit/s dense WDM transmission over 10000 km using single 40 GHz InP Mach-Zehnder modulator," Electron. Lett. **36**, 1574–1575 (2000).

293. B. Li, et. al. , "SiGe/Si Mach-Zehnder Interferometer Modulator based on the Plasma Dispersion Effect," Appl. Phys. Lett. **74**, (1999).

294. L. Liao, D. Samara-Rubio, M. Morse, A. Liu, and D. Hodge, "High speed silicon Mach-Zehnder modulator," Opt. Express, **13**, (2005).

295. G. L. Li and P.K.L. Yu, "Optical intensity modulators for digital and analog applications," J. Lightwave Technol. **21**, 2010–2030 (2003).

296. J.P. Sokoloff, P.R. Prucnal, I. Glesk, and M. Kane, "A terahertz optical asymmetric demultiplexer (TOAD)," IEEE Photon. Technol. Lett.**5**, (1993).

297. T. Durhuus, C. Joergensen, B. Mikkelsen, and K.E. Stubkjaer, "Penalty free all-optical wavelength conversion by SOA's in Mach–Zehnder configuration," in ECOC'93. Montreux, 1993.

298. B. Mikkelsen, et. al., "20 Gbit/s polarisation insensitive wavelength conversion in semiconductor optical amplifiers," in ECOC'93. Montreux, 1993.

299. N. Patel, et. al., "40-Gb/s demultiplexing using an ultrafast nonlinear interferometer (UNI)," IEEE Photon. Technol. Lett. **8**, 1695–1697 (1996).

300. N.S. Patel, K. L. Hall, and K.A. Rauschenbach, "Interferometric all optical switches for ultrafast signal processing," Appl. Opt. **37**, 2831–2842 (1998).

301. C. Janz, F. Poingt, F. Pommereau, W. Grieshaber, F. Gaborit, D. Leclerc, I. Guillemot, and M. Renaud, "All-active Dual Order Mode (DOMO) Mach-Zehnder wavelength converter for 10 Gb/s operation," Electron. Lett. **35**, 1862 (1999).

302. B.S. Robinson, S.A. Hamilton, and E.P. Ippen, "Demultiplexing of 80 Gbit/s pulse-position modulated data with an ultrafast nonlinear interferometer," IEEE Photon. Technol. Lett. **14**, 2002.

303. G. Raybon, Y. Su, J. Leuthold, R.-J. Essiambre, T. Her, C. Joergensen, P. Steinvurzel, and K. D. K. Feder, "40 Gbit/s pseudo-linear transmission over one million kilometers," in Optical Fiber Conference (OFC): Paper FD10-1, 2002.

304. J. Leuthold, G. Raybon, Y. Su, R. Essiambre, S. Cabot, J. Jaques, and M. Kauer, "40 Gbit/s transmission and cascaded all-optical wavelength conversion over 1000000 km," Electron. Lett. **38**, 890–892 (2002).

305. Y. Su, G. Raybon, R.-J. Essiambre, and T.-H. Her, "All-optical 2R regeneration of 40-Gb/s signal impaired by intrachannel four-wave mixing," Photonics Technol. Lett. **15**, 350–352 (2003).

306. J. Nayyer and H. Nagata, "Suppression of thermal drifts of high speed Ti:$LiNbO_3$ optical modulators," IEEE Photon. Technol. Lett. **6**, 952–955 (1994).

307. S.K. Korotky and J.J. Veselka, "An RC network analysis of long term Ti:LiNbO3 bias stability," J. Lightwave Technol. **14**, 2687–2697 (1996).

308. A. Waksberg and J. Wood, "An Automatic Optical Bias Control for Laser Modulators," Rev. Sci. Instrum., **43**, 1271–1273 (1972).

309. C.T. Mueller and J.G. Coffer, "Temperature-dependent bias drift in proton-exchanged lithium niobate Mach-Zehnder modulators," in Conference on Lasers and Electro-Optics (CLEO), 1999.

310. H. Nagata, "DC drift failure rate estimated on 10 Gb/s x-cut lithium niobate modulators," IEEE Photon. Technol. Lett. **12**, 1477–1479 (2000).

311. H. Nagata, N. Papasavvas, and D.R. Maack, "Bias stability of OC48 x-cut lithium-niobate optical modulators: four years of biased aging test results," Photon. Technol. Lett. **15**, 42–44 (2003).

312. H. Nagata, G. D. Feke, Y. Li, and W.R. Bosenberg, "DC drift of Z-cut $LiNbO_3$ modulators," IEEE Photon. Technol. Lett. **16**, 1655–1657 (2004).

313. H. Nagata, Y. Li, D.R. Maack, and W.R. Bosenberg, "Reliability Estimation From Zero-Failure $LiNbO_3$ Modulator Bias Drift Data," IEEE Photon. Technol. Lett. **16**, 1477–1479 (2004).

314. D.O. Caplan, "A technique for measuring and optimizing modulator extinction ratio," in Conference on Lasers and Electro-Optics (CLEO), 2000.

315. N. Kuwata, H. Nishimoto, T. Horimatsu, and T. Touge, "Automatic bias control circuit for Mach-Zehnder modulator," presented at Nat. Meet. Inst. Electon. Comm. Eng., Japan, 1990.

316. A.H. Gnauck and C.R. Giles, "2.5 and 10 Gb/s transmission experiments using a 137 photon/bit erbium-fiber preamplifier receiver," IEEE Photon. Technol. Lett. **4**, 80–82 (1992).

317. Q. Jiang and M. Kavehrad, "A Subcarrier-Multiplexed Coherent FSK System Using a Mach-Zehnder Modulator with Automatic Bias Control," IEEE Photon. Technol. Lett. **5**, 941–943 (1993).

318. H. Nagata, Y. Li, K. R. Voisine, and W.R. Bosenberg, "Reliability of Nonhermetic Bias-Free $LiNbO_3$ Modulators," IEEE Photon. Technol. Lett. **16**, (2004).

319. L. Fenghai, C.J. Rasmussen, and R.J.S. Pedersen, "Experimental verification of a new model describing the influence of incomplete signal extinction ratio on the sensitivity degradation due to multiple interferometric crosstalk," IEEE Photon. Technol. Lett. **11**, 137 (1999).

320. Z. Li, Y. He, B.F. Jorgensen, and R.J. Pedersen, "Extinction ratio effect for high-speed optical fiber transmissions," presented at Int. Conf. on Comm. Tech. Proc. (ICCT '98), 1998.

321. M. Pauer and P.J. Winzer, "Impact of Extinction Ratio on Return-to-Zero Coding Gain in Optical Noise Limited Receivers," IEEE Photon. Technol. Lett. **15**, 879 - 881 (2003).

322. H. Kim and A.H. Gnauck, "Chirp characteristics of dual-drive mach–zehnder modulator with a finite dc extinction ratio," IEEE Photonics Tech. Lett., **14**, (2002).

323. T. Kawanishi, T. Sakamoto, M. Tsuchiya, and M. Izutsu, "70 dB extinction-ratio $LiNbO_3$ optical intensity modulator for two-tone lightwave generation," in Optical Fiber Conference (OFC), 2006.

324. "Measuring extinction ratio of optical transmitters," Hewlett Packard App. Note, 1550-8, 1998.

325. P.O. Andersson and K. Akermark, "Accurate-optical extinction ratio measurement," IEEE Photon. Technol. Lett. **6**, 1356–1358 (1994).

326. C.R. Yang, W.-Y. Hwang, H. Park, H.H. Hong, and S.G. Han, "Off-level sampling method for bias stabilisation of electro-optic Mach-Zehnder modulator," Electron. Lett. **35**, 590–591 (1999).

327. R.C. Alferness, S.K. Korotky, and E. Marcatili, "Velocity-matching techniques for integrated optic traveling wave switch/modulators," IEEE J. Quantum Electron. **20**, 301–309 (1984).

328. R.C. Alferness, " Traveling wave, electrooptic devices with effective velocity matching " in US Pat. 4,448,479, 1984.

329. O. Mitomi, K. Noguchi, and H. Miyazawa, "Broadband and low driving-voltage $LiNbO_3$, optical modulators," IEE Proc. Optoelectron. **145**, 360–364 (1998).

330. R.C. Alferness, "Guided-wave devices for optical communication," IEEE J. Quantum Electron. **17**, 946–959 (1981).

331. N. Henmi, T. Saito, and T. Ishida, "N. Henmi, T. Saito, and T. Ishida, "Prechirp technique as a linear dispersion compensation for ultrahigh-speed long-span intensity modulation direct detection optical communication systems," J. Lightwave Technol. **12**, 1706–1719 (1994).

332. S.K. Kim, O. Mizuhara, Y.K. Park, L.A. Tzeng, Y.S. Kim, and J. Jeong, "Theoretical and experimental studies of 10 Gb/s transmission performance using 1.55 μm $LiNbO_3$ – based transmitters using adjustable extinction ratio and chirp," J. Lightwave Technol. **17**, 1320–1325 (1999).

333. J.J. Veselka, S.K. Korotky, P. V. Mamyshev, A.H. Gnauck, G. Raybon, and N. M. Froberg, "A Soliton Transmitter Using a CW Laser and an NRZ Driven Mach-Zehnder Modulator," IEEE Photon. Technol. Lett. **8**, 950–952 (1996).

334. N.M. Froberg, A.H. Gnauck, G. Raybon, and J.J. Veselka, "Method and Apparatus for Generating Data Encoded Pulses in Return-to-Zero Format," in US Pat. 5,625,722, 1997.

335. B. Zhu, L. Leng, A.H. Gnauck, M.O. Pedersen, D. Peckham, L.E. Nelson, S. Stulz, S. Kado, L. Gruner-Nielsen, R.L. Lingle, Jr., S. Knudsen, J. Leuthold, C. Doerr, S. Chandrasekhar, G. Baynham, P. Gaarde, Y. Emori, and S. Namiki, "Transmission of 3.2 Tb/s (80 $\times$ 42.7 Gb/s) over 5200 km of UltraWave^{TM} fiber with 100-km dispersion-managed spans using RZ-DPSK format," in ECOC: Paper PD4.2, 2002.

336. P.J. Winzer, "Optical transmitters, receivers, and noise," in *Wiley Encyclopedia of Telecommunications*, J.G. Proakis, Ed. (New York: Wiley, 2002), pp. 1824–1840.

337. M.L. Dennis, W. I. Kaechele, W.K. Burns, T. F. Carruthers, and I.N. Duling, "Photonic Serial–Parallel Conversion of High-Speed OTDM Data," IEEE Photon. Technol. Lett. **12**, 1561–1563 (2000).

338. R.C. Williamson, J.L. Wasserman, G.E. Betts, and J.C. Twichell, "Sinusoidal Drives for Optical Time Demultiplexers," IEEE Trans. Microwave Theory and Technol. **49**, 1945–1949 (2001).

339. V.W.S. Chan, "Optical satellite networks," J. Lightwave Technol. **21**, 2811–2827 (2003).

340. M.J.F. Digonnet, *Rare Earth Doped Fiber Lasers and Amplifiers* (New York: Marcel Dekker, 1993).

341. C.R. Giles and T. Li, "Optical amplifiers transform long-distance lightwave telecommunications," Proc. IEEE, **84**, 870–883 (1996).

342. J.-M.P. Delavaux and J.A. Nagel, "Multi-Stage Erbium-Doped Fiber Amplifier Designs," J. Lightwave Technol. **13**, 703–720 (1995).

343. L. Boivin, M.C. Nuss, J. Shah, D.A.B. Miller, and H.A. Haus, "Receiver sensitivity improvement by impulsive coding," Photon. Technol. Lett. **9**, 684–686 (1997).

344. L. Boivin and G.J. Pendock, "Receiver sensitivity for optically amplified RZ signals with arbitrary duty cycle," presented at Optic. Amplifiers and their Applications (OAA'99), 1999.

345. P.J. Winzer and A. Kalmar, "Sensitivity enhancement of optical receivers by impulsive coding," J. Lightwave Technol. **8**, 171–177 (1999).

346. W.R. Leeb, P.J. Winzer, and M. Pauer, "The potential of return-to-zero coding in optically amplified lasercom systems," in LEOS, 1999, pp. 224–225.

347. M. Pauer, P.J. Winzer, and W.R. Leeb, "Booster EDFAs in RZ-coded links: Are they average-power limited?," in Proc. SPIE, (Free-Space Laser Communication Technologies XIII), vol. 4272, 118–127, San Jose, USA, 2001.

348. P.J. Winzer, A. Kalmar, and W.R. Leeb, "Role of amplied spontaneous emission in optical free-space communication links with optical amplication - impact on isolation and data transmission; utilization for pointing, acquisition, and tracking," in Proc. SPIE (Free-Space Laser Communication Technologies XI), vol. 3615, 104–114, San Jose, CA, USA, 1999.

349. D.J. DiGiovanni and C.R. Giles, "Multistage optical amplifiers " in US Pat. 5115338, 1992.

350. R.I. Laming, M.N. Zervas, and D.N. Payne, "Erbium-Doped Fiber Amplifier with 54 dB Gain and 3.1 dB Noise Figure," IEEE Photon. Technol. Lett. **4**, 1345–1347 (1992).

351. T. Pliska, S. Mohrdiek, and C. Harder, "Power stabilisation of uncooled 980 nm pump laser modules from 10to 100°C," Electron. Lett. **37**, 33–34 (2001).

352. B. Schmidt, S. Pawlik, N. Matuschek, J. Muller, T. Pliska, J. Troger, N. Lichtenstein, A. Wittmann, S. Mohrdiek, B. Sverdlov, and C. Harder, "980 nm single mode modules yielding 700 mW fiber coupled pump power," in Optical Fiber Conference (OFC), 2002.

353. F. Hakimi, E.S. Kintzer, and R.S. Bondurant, "High-power single-polarization EDFA with wavelength mulitplexed pumps," in Conference on Lasers and Electro-Optics (CLEO), 1998.

354. D.O. Caplan and F. Hakimi, "A high-power high-gain single-polarization EDFA," in Conference on Lasers and Electro-Optics (CLEO), 2000.

355. C. Lester, A. Bjarklev, T. Rasmussen, and P. G. Dinesen, "Modeling of Yb^{3+}-Sensitized Er^{3+}-Doped Silica Waveguide Amplifiers," J. Lightwave Technol. **13**, 740–743 (1995).

356. Z. J. Chen, J.D. Minelly, and Y. Gu, "Compact low cost Er^{3+}/Yb^{3+} c o-doped fibre amplifiers pumped by 827 nm laser diode," Electron. Lett. **32**, 1812–1813 (1996).

357. M. Karasek, "Optimum Design of Er^{3+} –Yb^{3+} Codoped Fibers for Large-Signal High-Pump-Power Applications," IEEE J. Quantum Electron. **33**, 1699–1705 (1997).

358. J. Nilsson, P. Scheer, and B. Jaskorzynska, "Modeling and Optimization of Short Yb^{3+}-Sensitized Er^{3+}-Doped Fiber Amplifiers," Photon. Technol. Lett. **6**, 383–385 (1994).

359. A. Galvanauskas and B. Samson, "High Fiber," in SPIE's oemagazine, 2004, pp. 15–17.

360. I. Berishev, A. Komissarov, N. Moshegov, P. Trubenko, L. Wright, A. Berezin, S. Todorov, and A. Ovtchinnikov, "AlGaInAs/GaAs record high-power conversion efficiency and record high-brightness coolerless 915-nm multimode pumps," in SPIE, 2005.

361. V. Gapontsev, I. Berishev, G. Ellis, A. Komissarov, N. Moshegov, O. Raisky, P. Trubenko, V. Ackermann, E. Shcherbakov, J. Steineke, and A. Ovtchinnikov, "High-efficiency 970-nm multimode pumps " in SPIE, 2005.

362. V. Gapontsev, I. Berishev, G. Ellis, A. Komissarov, N. Moshegov, A. Ovtchinnikov, O. Raisky, P. Trubenko, V. Ackermann, and E. Shcherbakov, "9xx nm single emitter pumps for multi-kW systems," in SPIE, 2006.

363. D.J. Ripin and L. Goldberg, "High efficiency side-coupling of light into optical fibres using imbedded v-grooves," Electron. Lett. **31**, 2204–2205 (1995).

364. L. Goldberg, B. Cole, and E. Snitzer, "V-groove side-pumped 1.5 μm fibre amplifier," Electron. Lett. **33**, 2127–2129 (1997).

365. L. Goldberg and J. Koplow, "Compact, side-pumped 25 dBm Er/Yb co-doped double cladding fibre amplifier," Electron. Lett. **34**, 2027–2028 (1998).

366. D.J. DiGiovanni and A.M. Vengsarkar, "Article comprising a cladding-pumped optical fiber laser," in US Pat. 5,708,669, 1998.

367. D.J. DiGiovanni and A.J. Stentz, "Tapered fiber bundles for coupling light into and out of cladding-pumped fiber devices," in US Pat. 5,864,644, 1999.

368. V.P. Gapontsev and I. Samartsev, "Coupling arrangement between a multi-mode light source and an optical fiber through an intermediate optical fiber length," in US Pat. 5,999,673, 1999.

369. R.P. Espindola, I. Ryazansky, A.J. Stentz, K. L. Walker, and P.F. Wysocki, "Multi-stage optical fiber amplifier having high conversion efficiency," in US Pat. 6,104,733, 2000.

370. F. Hakimi and H. Hakimi, "New side coupling method for double-clad fiber amplifiers," in Conference on Lasers and Electro-Optics (CLEO), 2001.

371. L. Goldberg and M. Le Flohic, "Optical fiber amplifiers and lasers and optical pumping device therefor " in US Pat. 6,608,951, 2003.

372. A.B. Grudinin, D.N. Payne, W. Paul, L.J.A. Nilsson, M.N. Zervas, M. Ibsen, and M.K. Durkin, " Multi-fibre arrangements for high power fibre lasers and amplifiers " in US Pat. 6,826,335, 2004.

373. Y. Jeong, J. Sahu, D.B.S. Soh, C.A. Codemark, and J. Nilsson, "High-power, tunable, single-frequency, single-mode erbium:ytterbium codoped large core fiber master-oscillator power amplifier source," Opt. Lett. **30**, 2997 (2005).

374. Y. Jeong, J. Nilsson, J. Sahu, D.B.S. Soh, C. Alegria, P. Dupriez, C.A. Codemark, and D.N. Payne, "Single-frequency, single-mode, plane-polarized ytterbium-doped fiber master oscillator power amplifier source with 264 W of output power," Opt. Lett. **30**, 459–461 (2005).

375. N.G. Walker and G.R. Walker, "Polarization Control for Coherent Communications," J. Lightwave Technol. **8**, 438–458 (1990).

376. F. Heismann and M.S. Whalen, "Fast Automatic Polarization Control System," Photonics Technol. Lett. **4**, 503–505 (1992).

377. P. Oswald and C.K. Madsen, "Deterministic Analysis of Endless Tuning of Polarization Controllers," J. Lightwave Technol. **24**, 2932–2939 (2006).

378. Duling, I.N. III and R.D. Esman, "Single-polarisation fibre amplifier," Electron. Lett. **28**, 1126–1128 (1992).

379. D.O. Caplan, "Method and apparatus for stabilizing a high-gain, high-power single polarization EDFA," in US Pat. 6,831,779, 2004.

380. F. Hakimi, D.O. Caplan, H. Hakimi, and A.L. Tuffli, "Radiation effects on a two-stage double-pass single-polarization erbium fiber amplifier," in Conference on Lasers and Electro-Optics (CLEO), 2002.

381. K. Morito and S. Tanaka, "Record High Saturation Power (+22 dBm) and Low Noise Figure (5.7 dB) Polarization-Insensitive SOA Module," in Optical Amplifiers and Their Applications (OAA) Topical Meeting, paper TuC2, 2005.

382. Z. Jiang, D.E. Leaird, and A.M. Weiner, "Optical Arbitrary Waveform Generation and Characterization Using Spectral Line-by-Line Control," J. Lightwave Technol. **24**, 2487–2494 (2006).

383. J. Nilsson, Y. Jeong, C. Alegria, R. Selvas, J. Sahu, R. Williams, K. Furusawa, W. Clarkson, D. Hanna, D. Richardson, T. Monro, D. Payne, K. Yla-Jarkko, S. Alam, and A. Grudinin, "Beyond 1 kW with Fiber Lasers and Amplifiers," in Optical Fiber Conference (OFC), 2003.

384. J. Limpert, A. Liem, H. Zellmer, and A. Tuennermann, "Continuous wave ultrahigh brightness fiber laser systems," in IEEE Photonics West, 2003.

385. I.T. McKinnie, J.E. Koroshetz, W.S. Pelouch, D.D. Smith, J.R. Unternahrer, S.W. Henderson, and M. Wright, "Self-imaging waveguide Nd:YAG laser with 58% slope efficiency," in Conference on Lasers and Electro-Optics (CLEO), 2002.

386. M.D. Mermelstein, A.D. Yablon, and C. Headley, "Suppression of Stimulated Brillouin Scattering in an Er-Yb Fiber Amplifier Utilizing Temperature-Segmentation," in OAA, Paper TuD3, Budapest, Hungary, 2005.

387. R.G. Smith and S.D. Personick, "Receiver Design," in *Semiconductor Devices for Optical Communication*, H. Kressel, Ed. (New York: Springer-Verlag, 1980).

388. Y. Miyamoto, Y. Hagimoto, and T. Kagawa, "A 10 Gb/s high sensitivity optical receiver using an InGaAs-InAlAs superlattice APD at 1.3 μm/1.5 μm," IEEE Photon. Technol. Lett. **3**, 372–374 (1991).

389. T.Y. Yun, M.S. Park, J.H. Han, I. Watanabe, and K. Makita, "IO-Gigabit-per-Second High-Sensitivity and Wide-Dynamic-Range APD-HEMT Optical Receiver," Photonics Technol. Lett. **8**, 1232–1234 (1996).

390. T.V. Muoi, "Extremely sensitive direct detection receiver for laser communications," in Conference on Lasers and Electro-Optics (CLEO), 1987.

391. A. MacGregor and B. Dion, "39 Photons/bit direct detection receiver at 810 nm, BER=$1E^{-6}$, 60 Mb/s, QPPM," Proc. SPIE (Free-Space Laser Communication Technologies III), **1417**, (1991).

392. H. Matsuda, A. Miura, H. Irie, S. Tanaka, K. Ito, S. Fujisaki, T. Toyonaka, H. Takahashi, H. Chiba, S. Irikura, R. Takeyari, and T. Harada, "High-sensitivity and wide-dynamic-range 10 Gbit/s APD/preamplifier optical receiver module," Electron. Lett. **38**, 650–651 (2002).

393. J.R. Lesh, "Power Efficient Communications for Space Applications," in International Telemetering Conference, 1982.

394. B.E. Moision and J. Hamkins, "Deep-Space Optical Communications Downlink Budget: Modulation and Coding," JPL IPN Progress Report 42-154, 2003.

395. X. Sun, et. al., "Space-qualified silicon avalanche-photodiode single-photon-counting modules," J. Mod. Optics **51**, 1333–1350 (2004).

396. D.M. Boroson, R.S. Bondurant, and D.V. Murphy, "LDORA: A Novel Laser Communication Receiver Array Architecture," Proc. SPIE **5338**, 16–28 (2004).

397. O.M. Efimov, L.B. Glebov, L.N. Glebova, K.C. Richardson, and V.I. Smirnov, "High-Efficiency Bragg gratings in photothemorefractive glass," Appl. Opt. **38**, 619–627 (1999).

398. D.O. Caplan, "Spectral Filtering," in *Mars Laser Communication Demonstration (MLCD) Receiver Final Report* (MIT Lincoln Laboratory, 2005).

399. F. Herzog, K. Kudielka, D. Erni, and W. Bachtold, "Optical Phase Locking by Local Oscillator Phase Dithering," IEEE J. Quantum Electron. **42**, 973–985 (2006).

400. H.P. Yuen and V.W.S. Chan, "Noise in Homodyne and Heterodyne Detection," Opt. Lett. **8**, 177–179 (1983).

401. G.L. Abbas, V.W.S. Chan, and T.K. Yee, "A Dual-Detector Optical Heterodyne Receiver for Local Oscillator Noise Suppression," IEEE J. Lightwave Technol. **LT-3** (5), 1110–1122 (October 1985).

402. S.B. Alexander, "Design of Wide-Band Optical Heterodyne Balanced Mixer Receivers," J. Lightwave Technol. **LT-5**, 523–537 (1987).

403. L.G. Kazovsky and O.K. Tonguz, "Sensitivity of Direct-Detection Lightwave Receivers Using Optical Preamplifiers," Photon. Technol. Lett. **3**, 53–55 (1991).

404. S.R. Chinn, D.M. Boroson, and J.C. Livas, "Sensitivity of optically preamplified DPSK receivers with Fabry-Perot filters," J. Lightwave Technol. **14**, (1996).

405. P.J. Winzer, M. Pfennigbauer, M.M. Strasser, and W.R. Leeb, "Optimum filter bandwidths for optically preamplified NRZ receivers," J. Lightwave Technol. **19**, 1263–1273 (2001).

406. M. Pfennigbauer, M.M. Strasser, M. Pauer, and P.J. Winzer, "Dependence of Optically Preamplified Receiver Sensitivity on Optical and Electrical Filter Bandwidths—Measurement and Simulation," Photon. Technol. Lett. **14**, 831–833 (2002).

407. L.Y. Lin, M.C. Wu, and T. Itoh, "Figure of merit for high-power, high-speed photodetectors," in Optical Fiber Conference (OFC), 1997.

408. A.M.D. Beling, D. Schmidt, H.-G. Bach, G.G. Mekonnen, R. Ziegler, V. Eisner, M. Stollberg, G. Jacumeit, E. Gottwald, and J.-J. Weiske, "High power 1550 nm twin-photodetector modules with 45 GHz bandwidth based on InP," in Optical Fiber Conference (OFC), 2002.

409. T.S. Rose, D. Gunn, and G.C. Valley, "Gamma and proton radiation effects in erbium-doped fiber amplifiers: active and passive measurements," J. Lightwave Technol. **19**, 1918–1923 (2001).

410. P.A. Humblet, "Design of optical matched filters," in Globecom '91, 1991.

411. H.L. Van Trees, *Detection, estimation, and modulation theory, Part I* (New York: Wiley, 1968).

412. H. Geiger, M. Ibsen, and R.I. Laming, "Optimum receivers with fiber gratings," in Optical Fiber Conference (OFC), 1998.

413. S.R. Chinn, "Error-rate performance of optical amplifiers with Fabry-Perot filters," Electron. Lett. **31**, 756–757 (1995).

414. R.C. Steele and G.R. Walker, "High-sensitivity FSK signal detection with an erbium-doped fiber preamplifier and Fabry-Perot etalon demodulation," IEEE Photon. Technol. Lett. **2**, 753–755 (1990).

415. J.D. Berger, F. Ilkov, D. King, A. Tselikov, and D. Anthon, "Widely tunable, narrow optical bandpass Gaussian filter using a silicon microactuator," in Optical Fiber Conference (OFC), 2003.

416. A. D'Errico, R. Proietti, N. Calabretta, L. Giorgi, G. Contestabile, and E. Ciaramella, "WDM-DPSK Detection by Means of Frequency-Periodic Gaussian Narrow Filtering " in Optical Fiber Conference (OFC), 2006.

417. B.E. Little, et. al., "Very high-order microring resonator filters for WDM applications," IEEE Photon. Tech. Lett. **16**, 2263–2265 (2004).

418. C.K. Madsen and J.H. Zhao, *Optical filter design and analysis* (New York: John Wiley & Sons, Inc., 1999).

419. H. Takahashi, K. Oda, H. Toba, and Y. Inoue, "Transmission characteristics of arrayed-waveguide N × N wavelength multiplexer," J. Lightwave Technol. **13**, 447–455 (1995).

420. K. Takada, M. Abe, T. Shibata, and K. Okamoto, "A 25-GHz-Spaced 1080-Channel Tandem Multi/Demultiplexer Covering the S-, C-, and L-Bands Using an Arrayed-Waveguide Grating With Gaussian Passbands as a Primary Filter," Photon. Technol. Lett. **14**, 648–650 (2002).

421. I. Littler, M. Rochette, and B. Eggleton, "Adjustable bandwidth dispersionless bandpass FBG optical filter," Opt. Express **13**, 3397–3407 (2005).

422. A. Nosratinia, "Self-characteristic Distributions," J. Franklin Institute **36**, 1219–1224 (1999).

423. A. Hasegawa and F. Tappert, "Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers. I. Anomalous dispersion," Appl. Phys. Lett. **23**, 142–144 (1973).

424. L.F. Mollenauer, "Ultra-long distance soliton transmission: putting fiber nonlinearity to work," in LEOS, 1993.

425. J. O'Reilly, J. da Rocha, and K. Schumacher, "Optical Fiber Direct Detection Receivers Optimally Tolerant to Jitter," IEEE Trans. Commun. **34**, 1141–1147 (1986).

426. J.P. Gordon and H.A. Haus, "Random walk of coherently amplified solitons in optical fiber transmission," Opt. Lett. **11**, 665–667 (1986).

427. G. Bosco, R. Gaudino, and P. Poggiolini, "An exact analysis of RZ versus NRZ sensitivity in ASE noise limited optical systems," in ECOC, 2001.

428. G. Bosco, A. Carena, V. Curri, R. Gaudino, and P. Poggiolini, "On the Use of NRZ, RZ, and CSRZ Modulation at 40 Gb/s With Narrow DWDM Channel Spacing," J. Lightwave Technol. **20**, 1694–1704 (2002).

429. J.H. Shapiro, "Signal-to-Noise Ratio Analysis for a Preamplified Direct-Detection Receiver with Pre- and Post-Detection Matched Filters," private communication, 2005.

430. G. Bosco and P. Poggiolini, "The Effect of Receiver Imperfections on the Performance of Direct-Detection Optical Systems using DPSK Modulation," in Optical Fiber Conference (OFC), 2003.

431. G. Bosco and P. Poggiolini, "The Impact of Receiver Imperfections on the Performance of Optical Direct-Detection DPSK," J. Lightwave Technol. **23**, 842–848 (2005).

432. J. Hsieh, A. Chiayu, V. Chien, X. Liu, A. Gnauck, and X. Wei, "Athermal Demodulator for 42.7-Gb/s DPSK Signals," in ECOC, 2005.

433. X. Liu, A.H. Gnauck, X. Wei, J. Hsieh, C. Ai, and V. Chien, "Athermal optical demodulator for OC-768 DPSK and RZ-DPSK signals," Photon. Technol. Lett. **17**, 2610–2612 (2005).

434. M.L. Stevens, "A High-speed DPSK encoder," private communication, 1998.

435. W. Kaiser, T. Wuth, M. Wichers, and W. Rosenkranz, " Reduced complexity optical duobinary 10-Gb/s transmitter setup resulting in an increased transmission distance," Photon. Technol. Lett. **13**, 884–886 (2001).

436. I. Kang, C. Xie, C. Dorrer, and A. Gnauck, "Implementations of alternate-polarization differential-phase-shift-keying transmission," Electron. Lett. **40**, 333–335 (2004).

437. C. Schramm, H.-G. Bach, A.M.D. Beling, G. Jacumeit, S. Ferber, R. Ludwig, R. Ziegler, G. G. Mekonnen, R. Kunkel, D. Schmidt, W. Schlaak, and G. Unterborsch, "High-bandwidth balanced photoreceiver suitable for 40-gb/s RZ-DPSK modulation formats," IEEE Sel. Topics Quantum Electron. **11**, 127–134 (2005).

438. E.A. Swanson, J.C. Livas, and R.S. Bondurant, "High sensitivity optically preamplified direct detection DPSK receiver with active delay-line stabilization," IEEE Photon. Technol. Lett. **6**, 263–265 (1994).

439. D.G. Heflinger , J.S. Bauch, and T. E. Humes, "Apparatus and method for tuning an optical interferometer," in US Pat. 6,396,605, 2002.

440. F. Séguin and F. Gonthier, "Tuneable All-Fiber® Delay-Line Interferometer for DPSK Demodulation," in Optical Fiber Conference (OFC), 2005.

441. D.O. Caplan, "Polarization Independent Interferometer designs," MIT Lincoln Laboratory, private correspondence, 1999.

442. D.A. Rockwell, D. H. Matsuoka, and C. L. Schulz, "Differential Phase Shift Keyed Demodulator System," in US Pat. 6,834,146 B2, 2004.

443. M.L. Stevens, S. Constantine, and D.O. Caplan, "Measured and calculated DPSK SNR change due to frequency offset error at 40 Gbit/s," private communication, 2006.

444. P.J. Winzer, F. Fidler, M.J. Matthews, L.E. Nelson, H.J. Thiele, J.H. Sinsky , S. Chandrasekbar, M. Winter, D.M. Castagnozzi, L.W. Stulz, and L.L. Buhl, "10-Gb/s Upgrade of Bidirectional CWDM Systems Using Electronic Equalization and FEC," J. Lightwave Technol. **23**, 203–210 (2005).

445. S.K. Nielsen, B.F. Skipper, and J.P. Vailladsen, "Universal AFC for use in optical DPSK systems," Electron. Lett. **29**, 1445–1446 (1993).

446. K. Kudielka and W. Klaus, "Optical homodyne PSK receiver: Phase synchronization by maximizing base-band signal power," in LEOS, 1999.

447. A.H. Gnauck, G. Raybon, S. Chandrasekhar, J. Leuthold, C. Doerr, L. Stulz, and E. Burrows, "25 × 40-Gb/s copolarized DPSK transmission over 12 × 100-km NZDF with 50-GHz channel spacing," IEEE Photon. Technol. Lett. **15**, 467–469 (2003).

448. H. Bissessur, G. Charlet, E. Gohin, C. Simonneau, L. Pierre, and W. Idler, "1.6 Tbit/s (40 × 40 Gbit/s) DPSK transmission over 3 × 100 km of TeraLight fibre with direct detection," Electron. Lett. **39**, 192–193 (2003).

449. T. Mizuno, M. Oguma, T. Kitoh, Y. Inoue, and T. Takahashi, "Mach–Zehnder Interferometer Exactly Aligned With ITU Grid Frequencies," Photon. Technol. Lett. **18**, 325–327 (2006).

450. T. Hoshida and H. Onaka, "Method and system for demultiplexing non-intensity modulated wavelength division multiplexed (WDM) signals," in U.S. Pat. 7,035,543, 2006 (filed 2001).

451. P.J. Winzer, "Optical Receiver for Wavelength-Division-Multiplexed Signals," in US Pat. Appl. 2004/0258423 A1, 2004.

452. B. Zhu, L. E. Nelson, S. Stultz, A.H. Gnauck, C. Doerr, J. Leuthold, L. Gruner-Nielsen, M.O. Pedersen, and R.L. Lingle, Jr., "High Spectral Density Long-Haul 40-Gb/s Transmission Using CSRZ-DPSK Format," J. Lightwave Technol. **22**, 208–214 (2004).

453. D.A. Atlas and L.G. Kazovsky, "An optical PSK homodyne transmission experiment using 1320 nm diode-pumped Nd:YAG lasers," IEEE Photon. Technol. Lett. **2**, 367–370 (1990).

454. D.M. Castagnozzi, J.C. Livas, E.A. Bucher, L.L. Jeromin, and J.W. Miller, "Performance of a 1 Gbit/s optically preamplified communication system with error correcting coding," Electron. Lett. **30**, 65–66 (1994).

455. D.O. Caplan, S. Constantine, and M.L. Stevens, "Near-quantum-limited OOK and binary-FSK receiver sensitivity using a Filtered Direct-Drive pulse-carved MOPA transmiter and an optically preamplified receiver at 1.25 Gbit/s," private communication, 2006.

456. R.A. Linke, et. al., "Coherent lightwave transmission over 150km fiber lengths at 400 Mbit/s and 1 Gbit/s data rates using phase modulation," Electron. Lett. **22**, 30–31 (1985).

457. T. Imai, T. N. Ohkawa, Y. Ichihashi, T. Sugie, and T. Ito, "Over 300 km CPFSK transmission experiment using 67 photon/bit sensitivity receiver at 2.5 Gbit/s," Electron. Lett. **26**, 357–358 (1990).

458. J.M. Kahn, A.H. Gnauck, J.J. Veselka, S.K. Korotky, and B.L. Kasper, "4-Gb/s PSK homodyne transmission system using phase-locked semiconductor lasers," IEEE Photon. Technol. Lett. **2**, 285–287 (1990).

459. T. Chikama, et. al., "Modulation and demodulation techniques in optical heterodyne PSK transmission systems," J. Lightwave Technol. **8**, 309–321 (1990).

460. I. Hardcastle, T. Large, F. Davis, and A. Hadjifotiou, "High performance 140 Mbit/s FSK coherent system," Electron. Lett. **26**, 1523–1525 (1990).

461. T. Naito, T. Chikama, and G. Ishikawa, "Optimum system parameters for multigigabit CPFSK optical heterodyne detection systems," J. Lightwave Technol. **12**, 1835–1841 (1994).

462. S. Norimatsu, H. Mawatari, Y. Yoshikuni, O. Ishida, and K. Iwashita, "10 Gbit/s optical BPSK homodyne detection experiment with solitary DFB laser diodes," Electron. Lett. **31**, 125–127 (1995).

463. R.I. Laming, A.H. Gnauck, C.R. Giles, M.N. Zervas, and D.N. Payne, "High-sensitivity two-stage erbium-doped fiber preamplifier at 10 Gb/s," IEEE Photon. Technol. Lett. **4**, 1348–1350 (1992).

464. E. Meissner, "116 photons/bit in a 565 Mbit/s optical DPSK heterodyne transmission experiment," Electron. Lett. **25**, 281–282 (1989).

465. T.J. Paul, E.A. Swanson, J.C. Livas, R.S. Bondurant, and R.J. Magliocco, "3 Gbit/s optically preamplified direct detection DPSK receiver with 116 photon/bit sensitivity," Electron. Lett. **29**, 614–615 (1993).

466. Y.K. Park, J.-M. P. Delavaux, Mizuhara, L.D. Tzend, T.V. Nguyen, M.L. Kao, P.D. Yeates, S.W. Granlund, and S.J., "5 Gbit/s Optical Preamplifier Receiver with 135 Photons/bit Usable Receiver Sensitivity," in Optical Fiber Conference (OFC), paper TuD4, 1993.

467. T. Saito, Y. Sunohara, K. Fukagai, S. Ishikawa, N. Henmi, S. Fujita, and Y. Aoki, "High Receiver Sensitivity at 10 Gb/s Using an Er-Doped Fiber Preamplifier Pumped with a 0.98 μm Laser Diode " IEEE Photon. Technol. Lett. **3**, 551–553 (1991).

468. P.P. Smyth, R. Wyatt, A. Fidler, P. Eardley, A. Sayles, and S. Craig-Ryan, "152 photons per bit detection at 622 Mbit/s to 2.5 Gbit/s using an erbium fibre preamplifier," Electron. Lett. **26**, 1604–1605 (1990).

469. L.D. Tzeng, R.E. Frahm, and W. Asous, "A high-performance optical receiver for 622 Mb/s direct-detection systems," Photon. Technol. Lett. **2**, 759–761 (1990).

470. K. Kannan, et. al., "High-sensitivity receiver optical preamplifiers," IEEE Photon. Technol. Lett. **4**, 272–275 (1992).

471. A.H. Gnauck, K.C. Reichmann, J.M. Kahn, S.K. Korotky, J.J. Veselka, and T.L. Koch, "4-Gb/s heterodyne transmission experiments using ASK, FSK and DPSK modulation," IEEE Photon. Technol. Lett. **2**, 908–910 (1990).

472. T. Kataoka, Y. Miyamoto, K. Hagimoto, and K. Noguchi, "20 Gbit/s long distance transmission using a 270 photon/bit optical preamplifier receiver," Electron. Lett. **30**, 716–716 (1994).

473. M. Shikada, et. al., "1.5 m high bit rate long span transmission experiments employing a high power DFB-DC-PBH laser diode," in *European Conference on Optical Communication* (Istituto Internazionale delle Communicazioni: Genva, 1985).

474. J.C. Campbell, et. al., "High performance avalanche photodiode with separate absorption grading and multiplication regions," Electron. Lett. **19**, 818–820 (1983).

475. K. Hagimoto, et. al., "Twenty-Gbit/s signal transmission using a simple high-sensitivity optical receiver," in Optical Fiber Conference (OFC), 1992, paper Tul3.

476. V. Vilnrotter, C.-W. Lau, M. Srinivasan, K. Andrews, and R. Mukai, "Optical Array Receiver for Communication Through Atmospheric Turbulence," J. Lightwave Technol. **23**, 1664–1675 (2005)..

477. D.O. Caplan and J.J. Carney, "Angstrom class narrow band filters at 1060–1080 nm," MIT Lincoln Laboratory, internal memorandum, 2003.

478. D.O. Caplan, "Spectral Filtering " in *Mars Laser Communication Demonstration (MLCD) Preliminary Design Review (PDR)* (MIT Lincoln Laboratory and NASA, 2005).

479. D.O. Caplan and B.S. Robinson, "WDM Mitigation of Nonlinear Impairments in Low-Duty-Cycle M-PPM Free-Space Optical Transmitters," submitted to OFC2008.

DOI: 10.1007/s10297-005-0033-5
Originally published in J. Opt. Fiber. Commun. Rep. 1, 355–379 (2004)

Free-space laser communications with adaptive optics: Atmospheric compensation experiments

Thomas Weyrauch[1] and Mikhail A. Vorontsov[1,2]

[1] Intelligent Optics Laboratory, Institute for Systems Research, University of Maryland, A.V. Williams Bldg., Mail Stop 1103, College Park, Maryland 20742
Email: weyrauch@umd.edu

[2] Computational and Information Sciences Directorate, U.S. Army Research Laboratory, 2800 Powder Mill Road, Adelphi, Maryland 20783
Email: mvorontsov@arl.army.mil

Abstract

Refractive index inhomogeneities of the turbulent air cause wave-front distortions of optical waves propagating through the atmosphere, leading to such effects as beam spreading, beam wander, and intensity fluctuations (scintillations). These distortions are responsible for severe signal fading in free-space optical communications systems and therefore compromise link reliability. Wave-front distortions can be mitigated, in principle, with adaptive optics, i.e., real-time wave-front control, reducing the likeliness of signal fading. However, adaptive optics technology, currently primarily used in astronomical imaging, needs to be adapted to the requirements of free-space optical communication systems and their specific challenges.

In this chapter we discuss a non-conventional adaptive optics approach that has certain advantages with respect to its incorporation into free-space optical communication terminals. The technique does not require wave-front measurements, which are difficult under the strong scintillation conditions typical for communication scenarios, but is based on the direct optimization of a performance quality metric, e.g., the communication signal strength, with a stochastic parallel gradient descent (SPGD) algorithm.

We describe an experimental adaptive optics system that consists of a beam-steering and a higher-resolution wave-front correction unit with a 132-actuator MEMS piston-type deformable mirror controlled by a VLSI system implementing the SPGD algorithm. The system optimizes the optical signal that could be coupled into a single-mode fiber after propagating along a 2.3-km near-horizontal atmospheric path. We

investigate characteristics of the performance metric under different atmospheric conditions and evaluate the effect of the adaptive system. Experiments performed under strong scintillation conditions with beam-steering only as well as with higher-resolution wave-front control demonstrate the mitigation of wave-front distortions and the reduction of signal fading.

1. Introduction

Although the term "free-space optical wave propagation" in its strict sense is associated with propagation in vacuum, in most applications of the so-called "free-space optical (FSO) communication" technology wave propagation occurs in the atmosphere, which constitutes a media with random optical inhomogeneities caused by turbulence. The wave-front distortions arising from these inhomogeneities may severely impact the performance of FSO communication systems resulting in communication link deterioration, i.e., an increase of the number of errors in the received signal. The bit error rate (BER) – the major characteristic of communication system performance – depends on both short-term errors resulting from electronic-circuit related noise and the turbulence-induced long-term random breaks (up to tens of milliseconds) in communication data traffic known as atmospheric signal fading. The short-term data losses can be recovered (at least partially) using various data coding techniques developed for wire and fiber-based communication systems [1]. In contrast, atmospheric-turbulence-induced deep signal fading represents a unique and significantly more challenging problem, which cannot be solved using conventional data coding techniques without sacrificing communication system efficiency, i.e., data throughput rate.

The major incentive for incorporation of adaptive optics (AO) technology into atmospheric laser communication systems is the active prevention of long-term data loss rather than data recovery. This could be accomplished by introducing aberrations to the communication system's optical antenna that compensate or at least mitigate atmospheric-turbulence-induced wave-front phase distortions. The challenge is the real-time control of the active/adaptive optical elements (wave-front correctors) incorporated in the receiver or transmitter optics in order to make communication signal fading less likely. Atmospheric turbulence mitigation with adaptive optics technique will hopefully result in a communication system performance comparable to that of true free-space propagation. In this sense, the term "free-space laser communication" being applied to atmospheric systems represents an advanced (optimistic) view of future developments in atmospheric laser communication technology.

In order to achieve this goal, adaptive optics has to be adapted for the specific requirements of free-space laser communications. AO technology was developed initially for atmospheric compensation in astronomical observatory sites, whose location are selected specifically in view of low turbulence-induced distortions. Novel approaches for AO may be more successful in meeting the challenges of free-space laser communication systems in near-ground scenarios characterized by stronger turbulence effects. This chapter is devoted to the experimental evaluation of such a non-conventional adaptive optics technique and is organized as follows. In Sec. 2 we discuss architectures for the incorporation of adaptive optics in FSO communication systems from a more general point of view. In Sec. 3 an experimental setup for evaluation of adaptive optics performance for FSO communications is described and results of the characterization

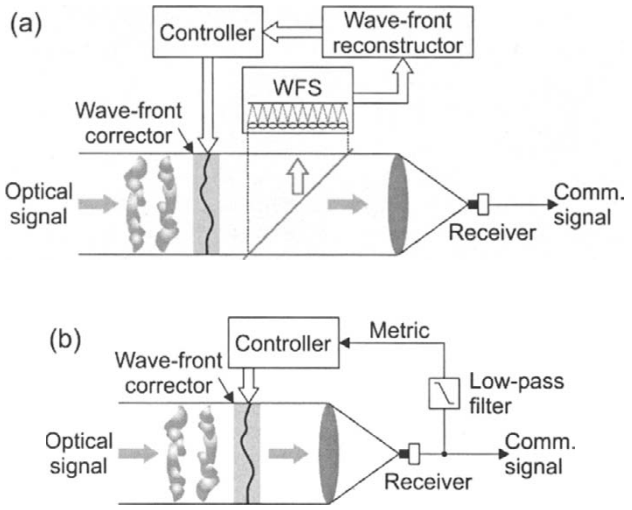


Fig. 1. Schematics of adaptive optics approaches for FSO communication. (**a**) Conventional adaptive optics uses a wave-front sensor (WFS) and wave-front reconstruction for control of the wave-front corrector. (**b**) Wave-front distortion compensation based on blind optimization of a system performance metric may use the received signal strength (determined from the communication signal after low-pass filtering) for feedback

of the atmospheric propagation path used in the experiments are presented. In Sec. 4 low-order wave-front control with a beam-steering system is analyzed. The high-order wave-front control system as well as results from atmospheric compensation experiments are described in Sec. 5.

2. Adaptive Optics Architectures for Free-Space Optical Communication Systems

The conventional adaptive optics approach for atmospheric compensation is based on wave-front sensing and reconstruction [2, 3]. In order to apply this technique for FSO communication systems, a part of the received beam has to be directed to a wave-front sensor. The wave-front is reconstructed from the measured data and used to calculate the control signals for the actuators of the wave-front corrector (Fig. 1a). While conventional AO has been successfully implemented in a number of (mostly astronomical) systems, in these applications atmospheric turbulence induced wave-front distortions occur in the near field of the (receiver) telescope's aperture, and intensity fluctuations (scintillations) can be considered as relatively weak [2, 3]. The situation is different for laser communication applications, where distributed turbulence is affecting the beam along the entire propagation path [4, 5]. Wave diffraction from continuously distributed refractive index inhomogeneities results in corruption of both phase and intensity. Strong intensity scintillations in the receiver aperture make wave-front measurements difficult, if not impossible, mostly because of the occurrence of branch points in the optical field's phase that represents a challenge for phase reconstruction techniques [6, 7].

In order to avoid wave-front measurements that are undesired in strong scintillation conditions, control of the wave-front correctors in AO systems can be performed by blind (or model-free) optimization of a system performance metric, e.g., the Strehl ratio or the received signal strength (Fig. 1b). Although this concept of wave-front control without wave-front measurements has been known since the early stages of AO technology development [8–10], it was largely disregarded because estimation techniques for the gradient of the metric (cost function) in respect to the control signals imposed limitations for the control bandwidth [11, 12].

This situation has changed, mainly because of the development of several novel technologies: new efficient control algorithms such as stochastic parallel gradient descent (SPGD) [13–15], their implementation using parallel processing hardware based on VLSI micro-electronics [16], and the emergence of high-bandwidth wave-front phase controllers, e.g., deformable mirrors based on micro-electromechanical systems (MEMS) [17]. For free-space optical communication there can be additional advantage of using wavefront control by optimization of a performance metric. Information about the received power level is often available in communication terminals (e.g. it can be obtained by low-pass filtering of the communication signal) and may be used as metric for the AO controller.

Adaptive optics based on model-free optimization can be used in different ways in FSO communication terminals. In general, one can distinguish between adaptive receiver, adaptive transmitter and adaptive transceiver architectures. The most straightforward approach is the adaptive receiver (Fig. 2a). Here, the distortions of the received beam are compensated by the adaptive optics (AO) allowing for a better focusing of the optical signal onto a small receiver area. The received signal strength is determined by the receiver sensor (Rx) and provided as metric signal, J, to the controller, which evaluates this feedback signal and applies the control signals, $\{u_j\}$, to the adaptive optics.

Compensating only the lowest-order aberrations, i.e., wave-front tip and tilt, which stabilizes the beam centroid location in the focal plane, may be sufficient if the beam is focused directly on a detector with a large enough photosensitive area. However, many FSO communication scenarios for enhanced signal throughput require the coupling of the received beam into an optical fiber (often a single-mode fiber) and use of wavelength-division multiplexed (WDM) beams and standard fiber-optic WDM technology for combining or separating different wavelengths. Since wave-front aberrations beyond tip and tilt errors sincerely impact the fiber-coupling efficiency, high-resolution phase compensation is required in addition to beam steering in order to avoid (mitigate) signal fading.

The potential of the adaptive receiver configuration for FSO communication systems is limited, because it allows only to compensate distortions of light waves that already entered the receiver aperture. In order to maximize the irradiance at a receiver aperture, an adaptive transmitter system at the other end of the propagation path is required, which mitigates atmospheric turbulence-induced beam spreading by pre-compensation of the transmitted wave.

An adaptive transmitter may be realized in different ways. The implementation shown in Fig. 2b is specifically designed for model-free-optimization AO [18]. The adaptive transmitter uses feedback from its FSO communication counterpart, i.e., the signal strength, J_2, received by the remote receiver (Rx_2). The disadvantage of this approach is the need for a means to transmit metric data from Rx_2 to the adaptive trans-

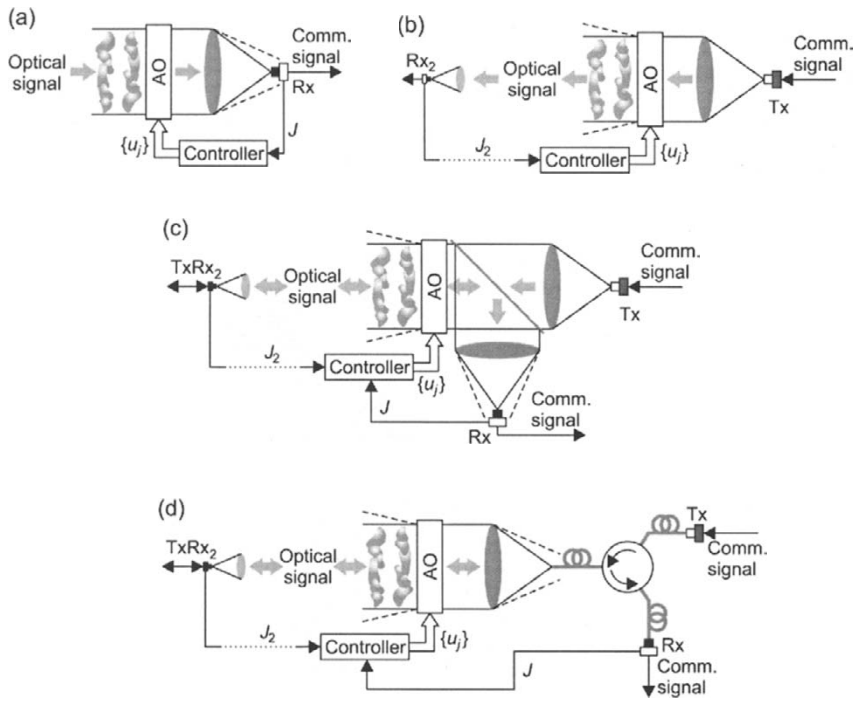


Fig. 2. Simplified schematics of SPGD adaptive optics architectures in FSO communication. (**a**) An adaptive receiver improves focusing the received beam on the receiver (Rx) using the received signal strength (J) as feedback signal and reduces thus the risk of data loss in the communication signal (dashed lines indicate waves without adaptive optics). (**b**) An adaptive transmitter reduces turbulence-induced beam spreading of the beam originating from the transmitter (Tx). It uses the signal strength received by the remote receiver (Rx_2) as feedback signal (J_2). (**c**) Received and transmitted optical signals share the same adaptive optical elements in an adaptive transceiver. Feedback is provided from the local receiver (Rx) and/or the remote transceiver ($TxRx_2$). (**d**) The use of fiber optics components reduces the complexity of the free-space optical setup in an adaptive transceiver.

mitter's AO controller. Metric data could be sent multiplexed (in time or wavelength division) with other information over the optical transmission channel or transmitted by a radio frequency signal [19]. The feedback signal's inherent latency due to the propagation time constitutes a basic limitation for this approach, and very long distances between transmitter and receiver are not feasible with a reasonable update rate for the control voltages.

The second approach for realization of an adaptive transmitter with model-free optimization strategy uses the beam from a beacon at the remote FSO communication terminal to obtain information about the wave-front distortions during propagation between the terminals. Control of the adaptive optics is performed by maximization of the beam quality (e.g., the Strehl ratio) of the received light from the beacon, and the transmitted laser beam is sent in the opposite direction through the same adaptive optics elements. This control scheme, similar to a conventional AO transmitter, relies on the assumption that wave-front correction for the received beam is simultaneously

the optimal pre-compensation for the outgoing beam. This may not necessarily be the case considering the effects of long, horizontal atmospheric propagation and the related scintillation effects typical for many laser communication applications.

In FSO communication systems, the optical signal from the remote terminal may serve as beacon, so that the configuration corresponds actually to an adaptive transceiver that uses the received signal strength, J, as performance metric and where received and transmitted optical communication signals (beams) are counter-propagating through the same adaptive optical system (Fig. 2c). Feedback for the adaptive optics controller could be provided by the signal strength (metric J) from the local receiver, from the remote receiver (metric J_2), or any combination thereof (with the need for experimental determination of the rule for the optimal combination of the performance metrics [19]).

Mutual misalignment of transmitter and receiver is a particular problem related with the adaptive transceivers setup shown in Fig. 2c. It can be avoided if the FSO communication terminals link the counter-propagating optical signals directly from fiber to fiber. Separation of the counter-propagating received and transmitted waves can then be performed in the fiber domain, e.g., using optical circulators or WDM technology. A schematic of the corresponding setup is shown in Fig. 2d.

3. Experimental System Arrangement and Components

Experimental evaluation of the performance of an adaptive optics system based on blind optimization of received signal strength was performed on a 2.3-km atmospheric propagation path that linked free-space optical communication transceivers located on the top of a 73-m-high water tower and an AO transceiver system in the laboratory. A detailed description of the propagation path as well as atmospheric characterization capabilities of the used atmospheric laser optics testbed can be found in [20].

A schematic of the wave propagation geometry is shown in Fig. 3. The first laser communication transceiver (T_1) on the tower uses a 15-mW laser diode at 785 nm and a 5-cm-diameter collimating lens. A diffuser placed between the laser diode and the collimating lens provides a beam divergence of about 2 mrad. The second transmitter (T_2) is collimating light emerging from a multi-mode fiber with 200 μm core diameter using a 60 mm-diameter lens. The light from a 100-mW laser diode (808 nm) is coupled into the other end of the 100-m-long fiber. The laser beam divergence is 2.8 mrad that results in a 6.5-m beam footprint at the receiver plane.

At the laboratory site, a 45° mirror mounted on the laboratory's roof directs the light from the transceivers toward an optical table, where the adaptive optics system is installed. A telescope (F/8 refractor, focal length $F_1 = 1.2$ m) was used to reduce the received beam diameter from 15 cm at the pupil $\mathcal{P}$ to 12.5 mm. For mitigation of atmospheric turbulence effects, we used an AO system comprising two subsystems with three active mirrors (depicted as transmissive elements in Fig. 3). The beam tracking system uses the slow tracking mirror, *TM*, to correct for large-scale pointing errors induced, e.g., by thermal expansion of the laboratory building and the water tower. The fast beam steering mirror, *FSM*, compensates tip and tilt errors induced by atmospheric turbulence and jitter from building vibrations. The MEMS piston-type deformable mirror, μDM, in the high-resolution adaptive optics (μAO) system mitigates the higher-order wave-front phase distortions from atmospheric turbulence.

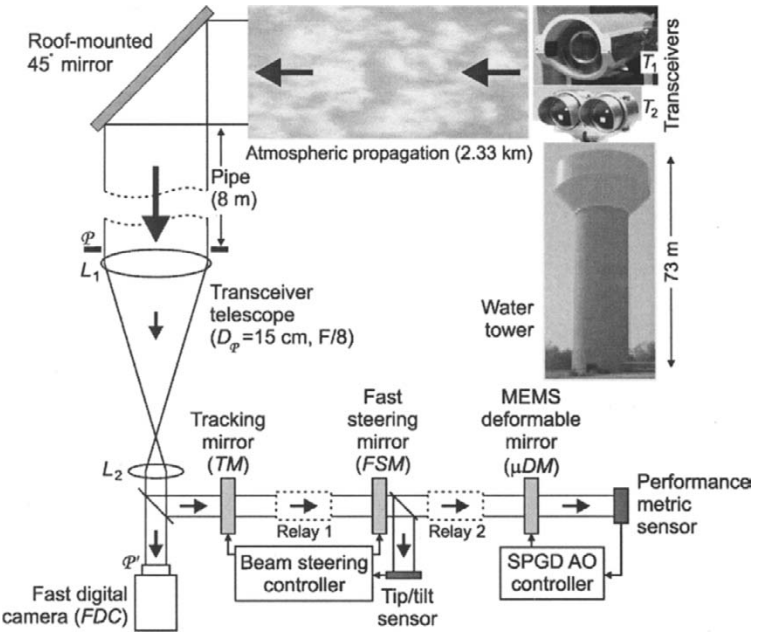


Fig. 3. Schematic of the experimental setup with lasercom transceiver modules (T_1, T_2) on the water tower, the transceiver telescope with $D_{\mathcal{P}} = 15$ cm aperture, and the beam steering system with the tracking mirror *TM* and the fast steering mirror *FSM*. Higher-order wave-front distortions were compensated with the MEMS deformable mirror μDM, driven by the SPGD controller. (The active mirrors of the AO subsystems are depicted as transmissive elements to simplify the schematic.) The fast digital camera *FDC* was used for recording irradiance fluctuations (scintillations) in the image plane $\mathcal{P}'$ of the entrance pupil

The tracking mirror is placed in an image plane of the transceivers telescope's pupil $\mathcal{P}$. With subsequent re-imaging onto the fast beam steering mirror and the micro-machined mirror by two optical relays, all active mirrors are in planes conjugate to the pupil $\mathcal{P}$.

To evaluate the irradiance fluctuations (scintillations) for waves from the beacon T_2, we imaged the pupil plane of the receiver telescope using a fast framing camera and recorded the temporal evolution of the irradiance distribution at about 800 frames/s. Then the probability distribution of irradiance values, the scintillation index as well as the irradiance fluctuations for individual pixels and their temporal spectra were calculated from the recorded data (929 frames). Figure 4a presents the probability density function, $p(I)$, for intensity scintillations for three different turbulence strengths corresponding to Rytov variances, σ_R^2, of 0.2, 1, and 10, respectively. The σ_R^2 values were calculated from the path-averaged refractive index structure constant, C_n^2, measured during the experiments, according to $\sigma_R^2 = 1.23 C_n^2 k^{7/6} L^{11/6}$ (k = wave number, L = propagation distance). These values cover the range from relatively weak turbulence ($\sigma_R^2 = 0.2$) to the scintillation saturation regime [21] ($\sigma_R^2 = 10$). The irradiance values, I, were normalized by their respective average, $\langle I \rangle$, in order to allow for an easier comparison of the curves. The insets show single frames, $I(\mathbf{r})$, for each condition ($\mathbf{r} = (x, y)$ is a pixel location vector in the pupil plane). The scintillation indices,

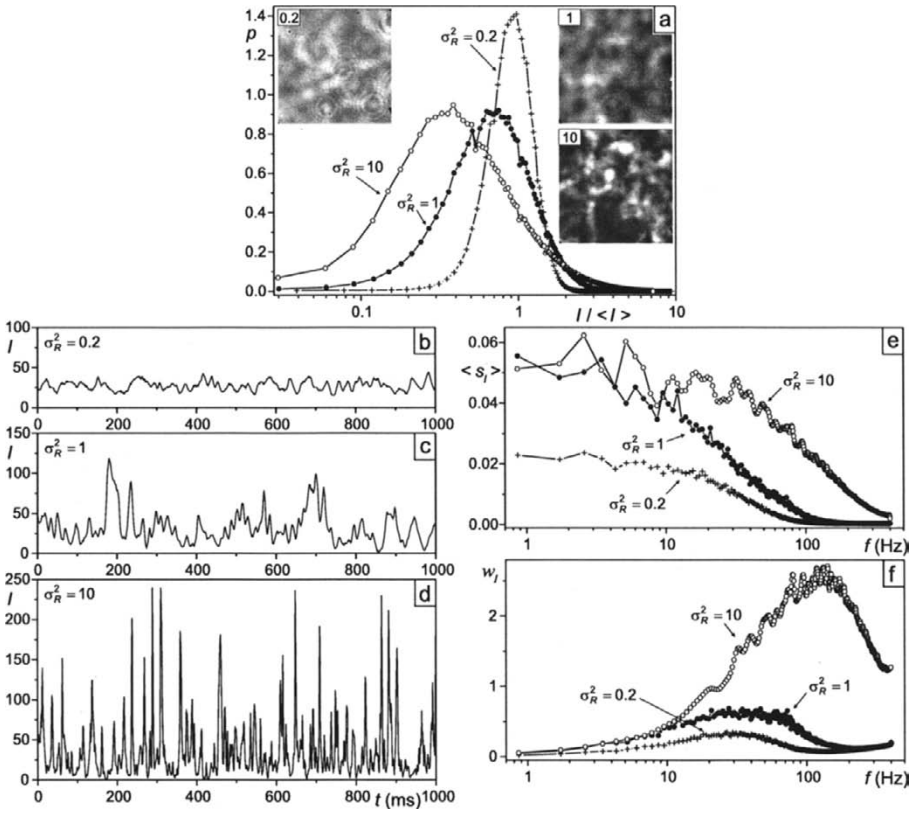


Fig. 4. Irradiance distributions in the receiver's pupil plane ($\mathcal{P}$): (**a**) Probability density function $p(I)$ of measured irradiance values I for three different atmospheric conditions ($C_n^2 = 1 \times 10^{-15}\,\mathrm{m}^{-2/3}, 5\times10^{-15}\,\mathrm{m}^{-2/3}$, and $5\times10^{-14}\,\mathrm{m}^{-2/3}$), which correspond to Rytov variances $\sigma_R^2 = 0.2$, 1, and 10, respectively. Insets: sample short-exposure irradiance distributions for each condition (labelled by the σ_R^2 value; photo area corresponds to $11\times11\,\mathrm{cm}^2$ in $\mathcal{P}$). (**b-d**) Temporal evolution of the irradiance at a sample pixel for $\sigma_R^2 = 0.2$, 1, and 10 (b,c, and d, respectively). (**e**) Average temporal spectra $\langle s_I(f)\rangle$ of the irradiance fluctuations in individual pixels. (**f**) The contents spectra $w_I(f) = f\langle s_I(f)\rangle$ reveal the main spectral contributions on the logarithmic frequency scale

$\sigma_I^2 = \langle I^2\rangle/\langle I\rangle^2 - 1$, were 0.09, 0.31, and 1.06 for this particular set of measurements. These values are smaller than the respective Rytov variance values ($\sigma_R^2 = 0.2$, 1, and 10) even for the weak fluctuation regime, which might be explained by the partial loss of coherence of the light at the transceiver optical antenna (T_2) after 100-m propagation in a multi-mode fiber.

The temporal evolution curves, $I(t)$, of the irradiance in an individual sample pixel are shown in Figs. 4b-d for $\sigma_R^2 = 0.2$, 1, and 10, respectively. To obtain the temporal spectra of the irradiance fluctuations we used Fourier transformations of the evolution curve $I(t)$ for all pixel locations $\mathbf{r}$ within the receiver pupil and calculated the average spectra $\langle s_I(f)\rangle$; the results are shown in Fig. 4e. The curves were normalized by their

d.c. spectral component $s_I(f = 0)$. The main spectral contributions, which can be identified in the contents spectra, $w_I(f) = f\langle s_I(f)\rangle$, shown in Fig. 4f, are in the range of a few Hz to more than 400 Hz (the maximum frequency detectable with the used camera). The crosswind speed v measured at the laboratory site was less than 1 m/s for the measurements corresponding to $\sigma_R^2 = 0.2$ and 1, while for $\sigma_R^2 = 10$ crosswind was considerably stronger with $v \approx 6$ m/s, which contributes to the shift of the maximum in the contents spectrum, $w_I(f)$.

4. Compensation of Low-Order Distortions

Conventional (nonadaptive) laser communication systems typically use a divergent beam for transmission so that a sufficiently large beam footprint at the receiver end of the communication link eliminates the need for precision tracking. However, only a small portion of the light emitted by the transmitter is actually entering the communication receiver aperture in this case. The available laser power could be utilized much more efficiently with a narrow beam using active beam steering in a FSO communication terminal in adaptive transmitter configuration. To evaluate the potential, we performed atmospheric compensation experiments with low-order (tip/tilt) adaptive optics compensation, i.e., using only the tracking and fast beam steering subsystem (BS) of the setup.

4.1. Tracking and Fast Beam Steering System

The beam tracking and steering controllers share a single sensor for driving both mirrors as shown in Fig. 5 (dashed box). The received wave is reflected sequentially from the tracking and beam steering mirrors (*TM* and *FSM*, respectively), which are optically interfaced by an optical relay (lenses L_3 and L_4). To provide a feedback signal, the wave is split after reflection from the fast steering mirror by beam splitter BS_1, and a portion of the beam is focused by lens L_5 onto the 2-dimensional linear position sensitive detector (*PSD*); its output voltages, V_x and V_y, correspond to the horizontal and vertical beam-centroid displacement from the sensor's center point, respectively. The fast steering mirror, *FSM*, is actuated using a piezoelectric tip/tilt platform; its angular control range corresponds to $\pm 50\,\mu$rad in the pupil plane $\mathcal{P}$. Control of both x and y actuators of *TM* and *FSM* are performed independently; in the following only the x control channel is described. The error signal $e_x^{(n)} = V_x^{(n)} - V_{x,0}$, i.e., the deviation of the *PSD* output at iteration n from the desired value $V_{x,0}$, is used to calculate the control signal $U_x^{(n)}$ using the discretized proportional-integral-derivative (PID) algorithm

$$U_x^{(n)} = U_x^{(n-1)} + k_{P,x} e_x^{(n)} + k_{I,x} S_x^{(n)} + k_{D,x} \left[e_x^{(n)} - e_x^{(n-1)} \right] . \qquad (1)$$

The error integral $S_x^{(n)}$ is calculated iteratively as $S_x^{(n)} = \alpha e_x^{(n)} + (1 - \alpha) S_x^{(n-1)}$. The coefficients k_P, k_I, and k_D, which determine the magnitude of the proportional, integral and derivative terms in (1), respectively, as well as α were determined by optimizing the response to a step-wise input signal. To suppress mirror ringing due to the fast steering mirror's mechanical resonance at about 2.4 kHz, the output voltages

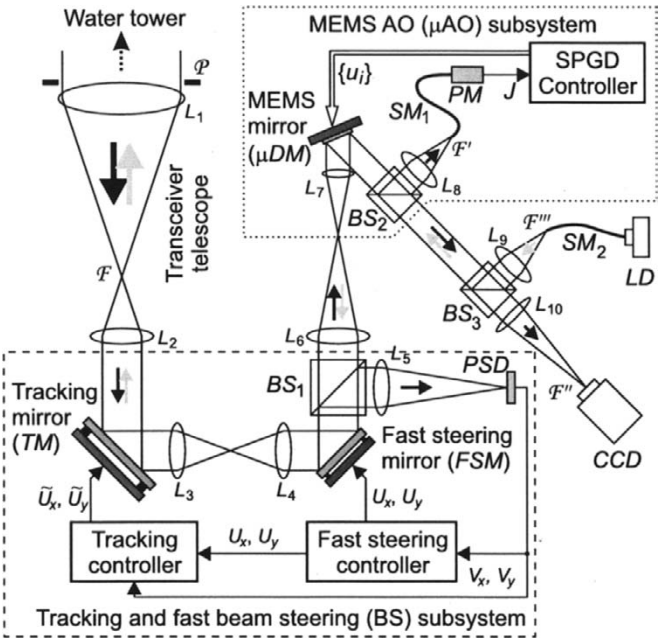


Fig. 5. Schematics of the AO setup. BS (*dashed box*): beam steering subsystem with active mirrors for slow tracking and fast beam steering. μAO(*dotted box*): higher resolution wave-front correction subsystem with MEMS mirror and SPGD controller. The pupil plane $\mathcal{P}$ is imaged by the transceiver telescope (L_1, L_2) onto the tracking mirror (*TM*), and subsequently by an optical relay (L_3, L_4) onto the fast steering mirror (*FSM*). Part of the light is coupled out by beam splitter BS_1 and focused by lens L_5 onto the position sensitive detector, *PSD*, to provide feedback signals V_x, V_y to the fast steering controller. The tracking controller monitors the *FSM* control voltages (U_x, U_y) and applies correction signals ($\tilde{U}_x$, $\tilde{U}_y$) to *TM* if required. The optical relay (L_6, L_7) connects the subsystems BS and μAO imaging the pupil onto the MEMS deformable mirror μDM. After reflection from μDM and beam splitter BS_2, a part of the received light is coupled into the single-mode fiber SM_1 in focal plane $\mathcal{F}'$. The optical power in the fiber, measured by the photomultiplier (*PM*), is the feedback signal J for the SPGD controller. Another part of the received light is transmitted by BS_2 and BS_3 and focused by onto the camera *CCD* in focal plane $\mathcal{F}'''$ (far field intensity distribution). For adaptive transmission, the light emerging from SM_2 is collimated by L_9 and propagates after reflection from beam splitter BS_3 in opposite direction through the AO system's optical train

U_x and U_y were low-pass filtered with 1-kHz cutoff frequency. The iteration rate for the discretized PID controller was up to 12 kHz.

The tracking controller continuously monitors the voltage U_x applied to the fast steering mirror, *FSM*, and actuates the tracking mirror,*TM*, with control signal $\tilde{U}_x$ when the control signal value U_x is outside of a tolerance range $U_{x,0} \pm \Delta U_C$ around the center of the dynamic range, $U_{x,0}$. The angular tracking range of *TM* corresponds to $\sim$ 10 mrad, and the maximum angular speed to $\sim$ 50 μrad/s in the pupil plane $\mathcal{P}$.

The closed-loop performance of the fast steering system was determined in experiments without atmospheric propagation but laboratory-generated tilt aberrations with sinusoidal oscillating amplitude. A comparison with open-loop behavior revealed

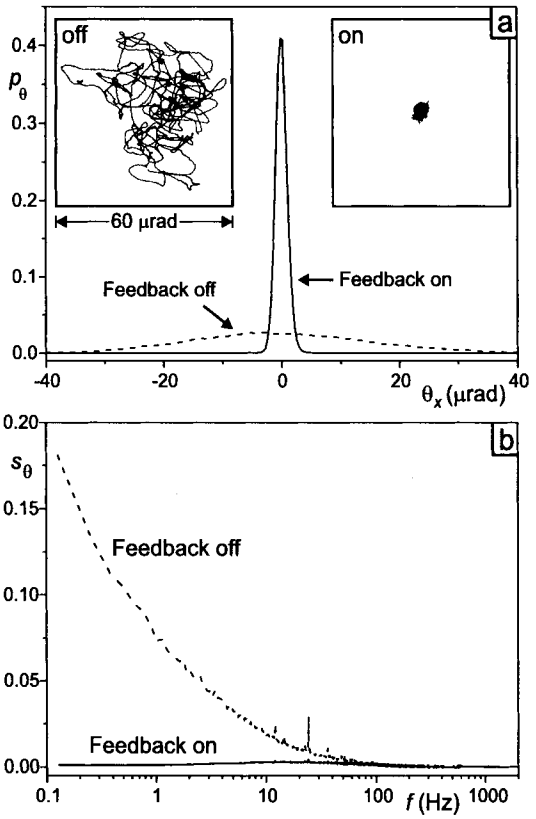


Fig. 6. (**a**) Probability density function p_θ for the tilt angle θ_x in the receiver pupil plane $\mathcal{P}$ after 2.3 km atmospheric propagation with and without tip/tilt correction (feedback on and off, respectively). The left inset shows the trajectory of tip/tilt angles acquired during two seconds under open-loop conditions; the right inset is the corresponding measurement with closed feedback loop. (**b**) Temporal spectra $s_\theta(f)$ for the tilt angle θ_x without tip/tilt correction (feedback off) and for the residual tilt angle with beam-steering system operation (feedback on)

better than 90% suppression of the tilt aberration amplitude at frequencies $f \leq 60$ Hz and better than 50% for $f \leq 330$ Hz.

4.2. *Compensation of Atmospheric Wave-Front Tilt Distortions*

For evaluation of the fast beam steering system performance in atmospheric conditions, we compared the statistics of the measured tip and tilt wave-front angles after 2.3 km propagation with the feedback off and on. Figure 6a shows the probability density functions p_θ for the tilt angle θ_x for both cases. Experiments demonstrated the reduction of the root-mean-square value $\theta_{\rm rms}$ of the total deviation angle $\theta = (\theta_x^2 + \theta_y^2)^{1/2}$ from $\theta_{\rm rms}^{\rm off} = 28$ μrad (beam steering system off) to $\theta_{\rm rms}^{\rm on} = 2$ μrad with closed-loop control (these values refer to the angle in the pupil plane $\mathcal{P}$, see Fig. 5). The left inset in Fig. 6a shows the trajectory of the beam centroid's angular coordinates (θ_x, θ_y) during a period of two seconds while the tip/tilt control was off (the side length of the box corresponds to 60 μrad in the pupil plane $\mathcal{P}$). For comparison, the trajectory of (θ_x, θ_y), recorded

during two seconds with closed tip/tilt control loop, is presented on the same scale in the right inset of Fig. 6a.

Figure 6b shows the temporal spectra of the wave-front tilt angle for both cases. Distortions of the tilt angle were significant only below 100 Hz; the fast beam steering system compensated these distortions effectively. The spectrum for the wave-front tilts in the absence of feedback control also reveals distortions at discrete frequencies (e.g. at 23 Hz) that were caused by vibrations of the laboratory building. As seen from the results in Fig. 6b, this jitter was efficiently mitigated by the tip/tilt compensation system.

4.3. Laser Communication with Tip/Tilt Control

For exploration of the influence of low-order (tip/tilt) adaptive beam steering on the performance of a laser communication system, we built a communication setup combining nonadaptive transceivers with an adaptive communication transmitter as shown in Fig. 7a. The bit pattern generator of a bit error rate (BER) performance analyzer modulated the laser source of the adaptive transmitter A, which transmitted the optical signal to the communication transceiver B on the water tower. The communication signal was retransmitted by transceiver B toward the laboratory building and was received by the communication receiver C, which was connected to the input of the BER analyzer to determine the overall BER of the closed-loop communication link.

Transceiver B was also used as the beacon for the adaptive transmitter A. It has concentric transmitter and receiver apertures (diameters of 5 cm and 10 cm, respectively; see T_1 in Fig. 3). Note that a diffuser was placed between the laser diode and the transmitter's collimating lens to achieve a beam divergence of about 2 mrad. The optical field from the beacon (dashed arrow in Fig. 7a) entered the AO system telescope at the laboratory (adaptive transmitter A in Fig. 7a). The power received by the AO telescope was in the order of 0.5 μW.

The adaptive transmitter A used the received optical wave to determine and compensate the wave-front phase tilts introduced by atmospheric turbulence and beam jitter. A laser beam was transmitted through the compensated beam propagation train as shown in Fig. 7b. For this the light from a laser diode, *LD*, was coupled into a single mode fiber, *SMF*, and the end of this fiber was placed in a plane conjugate to the beam positioning sensor, *PSD*. Thus the transmitted wave-front was pre-compensated by removing the wave-front tilts that resulted from the beam propagation from the beacon to the optical receiver system. The transmitter laser diode, *LD*, was modulated with an 125 Mbit/s bit stream.

The large footprint of the divergent beam received from transceiver B allowed the placement of the communication receiver C at a distance $h \approx 2$ m from the AO transmitter A, thus providing weak correlation of atmospheric phase aberrations between transmitted and received waves (space diversity) and avoiding backscatter enhancement effects that accompany counter-propagating waves [22,23]. Note that in the adaptive communication transmitter system no diffuser was used, and the transmitted power was intentionally reduced to only 70 μW in order to obtain a variation of the atmospheric-turbulence-induced BER that was large enough for a reasonable measurement duration. The beam footprint at the water tower site was less than 20 cm.

The influence of beam steering on the BER was investigated by sequentially turning on and off the fast beam steering feedback control in intervals of 60 seconds and

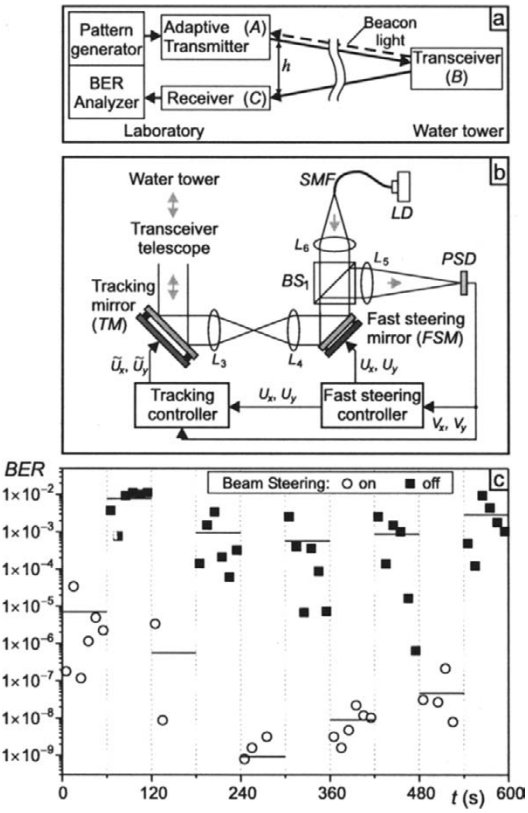


Fig. 7. (**a**) Schematic of the setup for bit error rate (BER) measurements for a closed-loop free-space optical communication link with a single adaptive transmitter. The bit pattern generator modulates the light transmitted by the adaptive transmitter A. The transceiver B on the water tower receives and retransmits the optical bit stream, which is received by the nonadaptive receiver C and sent to the bit error rate analyzer. The telescopes for A and C are separated by distance $h = 2$ m to introduce spatial diversity. (**b**) Schematic of the adaptive transmitter with the tracking mirror TM and the fast beam steering mirror FS. It uses the light from transceiver B as beacon for tip/tilt correction. (**c**) Bit error rates measured during a 600 s period with the beam steering turned on and off alternatively in intervals of 60 s each. The horizontal lines indicate the average BER for each 60 s interval

simultaneously monitoring the BER analyzer. Figure 7c shows the rates for bit errors accumulated during intervals of ten seconds within a total measurement time of ten minutes. No errors were observed in the cases where no symbol is shown (the smallest possible non-zero value for the BER is 8×10^{-10}, which corresponds to a single error within ten seconds for 125 MBit/s). The horizontal lines show the average values for the 60-second interval. As can be seen from Fig. 7c, the BER in an interval without wave-front tilts compensation was at least three orders of magnitude higher than in the respective preceding period with active beam steering. Since both free-space optical links (the commercial configuration with diffuser [link path $A \rightarrow C$] and the configuration with the wave-front tip/tilt compensation [link path $B \rightarrow A$]) contribute to the

BER, but compensation was only performed over the propagation path $B \rightarrow A$, one can expect an even higher improvement if compensation occurred in both directions.

5. SPGD High-Resolution Wave-Front Control

The high-resolution adaptive system used for the atmospheric compensation experiments is based on analog VLSI controller chips that implement a stochastic parallel gradient descent (SPGD) algorithm for optimization of a system performance metric in combination with a PC as supervisory controller and a MEMS deformable mirror as wave-front corrector. Details about the algorithm and the VLSI architecture can be found in [16, 24]. An analysis of the compatibility of the SPGD approach with various MEMS deformable mirrors and an experimental evaluation of the efficiency of SPGD control with a low-resolution AO system were presented in [25]. First experiments for wave-front distortion compensation using an AO system similar to the one described here were performed using artificially generated turbulence [26]. In the following, we discuss wavefront control using the SPGD adaptive system operating on a 2.3 km atmospheric propagation path with respect to free-space laser communication applications.

5.1. *SPGD Adaptive Optics System*

A 140-element piston-type micro-machined mirror (μDM) [17, 27] is used for the MEMS adaptive optics (μAO) system. The mirror is placed in an image plane of the AO receiver system pupil $\mathcal{P}$ (see Fig. 5). An optical relay system (lenses L_6 and L_7) reduces the beam size to match the 3.3 mm side length of the MEMS mirror. The control of the μDM is based on the optimization of a signal proportional to the optical power received through the single-mode fiber SM_1 (system performance metric J). The VLSI controller chips generate at each iteration step, n, a set of statistically independent control parameter perturbations, $\{\delta u_j^{(n)}\}$, in parallel for all N control voltages, $u_j^{(n)}$ ($j = 1, \ldots, N$; $N = 132$ – only the 132 mirror elements that were within the circular aperture of the beam were used). The perturbations have all the same absolute values ε but pseudo-random signs (Bernoulli distribution, $\delta u = \pm\varepsilon$) with a probability of 0.5 for either value. After application of the control voltage perturbations to the μDM, the system measures a perturbed metric value $J_+^{(n)}$, then the signs of all perturbation voltages are inverted and the corresponding value $J_-^{(n)}$ is measured, i.e.,

$$J_\pm^{(n)} = J(u_1^{(n)} \pm \delta u_1^{(n)}, \ldots, u_j^{(n)} \pm \delta u_j^{(n)}, \ldots, u_N^{(n)} \pm \delta u_N^{(n)}) \,. \tag{2}$$

The control voltages are then updated according to

$$u_j^{(n+1)} = u_j^{(n)} + \gamma \left[J_+^{(n)} - J_-^{(n)} \right] \operatorname{sign}(\delta u_j^{(n)}) \,. \tag{3}$$

The update coefficient, γ, was controlled by a supervisory control loop [25] using $\gamma = \gamma_0/(J + C)$ in order to allow for faster corrections when the metric value J is low (the constant $C > 0$ limits γ for very low J values). The response bandwidth of the MEMS deformable mirror limited the maximum achievable iteration rate, which was about 7,000 s^{-1}.

5.2. *Temporal Behavior of the Received Power in the Tip/Tilt-Compensated Receiver System*

Since the power of the received optical signal is used by the SPGD AO system as feedback signal J, the study of its characteristics after compensation of tip and tilt errors under real atmospheric propagation conditions is important for the understanding of the system's challenges and performance. In experiments, we recorded the temporal dependence of the metric, J, under different atmospheric conditions and measured simultaneously the path-averaged refractive index structure constant, C_n^2. As an example, Figure 8a shows two different time series of the received power signal $J(t)$, which were recorded while no feedback control was applied to the mirror *μDM*. Both beam tracking and fast steering control systems were operating during the experiments. The data presented in Fig. 8a were acquired on different days, but the average refractive index structure constant $C_n^2 = 4 \times 10^{-14}\,\mathrm{m}^{-2/3}$ was identical during both experiments. A weather station at the laboratory site recorded no wind during measurement of curve 1 while the crosswind speed was $v = 3.5$ m/s during measurement of curve 2. Both cases are characterized by strong intensity scintillations (scintillation saturation regime with the Rytov variance $\sigma_R^2 = 1.23 C_n^2 k^{7/6} L^{11/6} = 8$ for both cases). Despite similar C_n^2 values the normalized variances of metric values $\sigma_J^2 = \langle J^2\rangle / \langle J\rangle^2 - 1$ were different: $\sigma_{J,1}^2 = 0.92$ and $\sigma_{J,2}^2 = 0.57$ for curves 1 and 2, respectively.

The distinct difference in the temporal behavior is also visible in the average temporal power spectra, $\langle s_J(f)\rangle$, shown in Fig. 8b. To obtain the $\langle s_J(f)\rangle$ curves, we calculated the power spectra, $s_J(f)$, for 1000 recorded time series, $J(t)$, with about 1 s duration each by fast Fourier transformation and averaged these spectra. The average spectra are normalized to their d.c. component, $\langle s_J(0)\rangle$. Some spectrum features are better visible in the contents spectra, $w_J(f) = f\langle s_J(f)\rangle$, shown in Fig. 8c. Apparent parasitic signals at $f \approx 0.8$ kHz and $f \approx 2.4$ kHz were the result of incomplete suppression of the ringing of the fast steering mirror, *FSM*. Atmospheric distortions contribute to the signal mostly in the frequency range < 500 Hz, which is very similar to the scintillation spectra presented in Figs. 4e,f. The main contributions to the distortions were below 300 Hz.

An important characteristic of the metric J, i.e., the received power, in respect to free-space laser communication is the probability density distribution (PDF) for the metric values, $p(J)$, because the probability of low power values determines the likelihood of fading and thus the bit error ratio [22, 28]. Hence, we analyzed the tip/tilt-compensated PDFs of the metric for a broad range of turbulence conditions. Fig. 9 shows three representative examples corresponding to different metric variances $\sigma_J^2 = \langle J^2\rangle / \langle J\rangle^2 - 1$. In general, the variance of the received power observed in our experiments was between $\sigma_J^2 = 0.15$ and $\sigma_J^2 = 1.1$. In Fig. 9, the J values were normalized by their respective average value, $\langle J\rangle$, in order to allow for a comparison of the curve shapes. The curves $p(J)$ for $\sigma_J^2 = 1.03$ and 0.48 have apparently a Gaussian distribution-like shape on the logarithmic J-scale, i.e., they are similar to a log-normal distribution, while $p(J)$ for $\sigma_J^2 = 0.16$ is significantly different. The curves in Fig. 9 show that even for weak fluctuations (as for $\sigma_J^2 = 0.16$) there is a considerable probability for the occurrence of low metric values, i.e., deep fading in terms of laser communication, which could be (at least partially) prevented by higher-resolution adaptive optics.

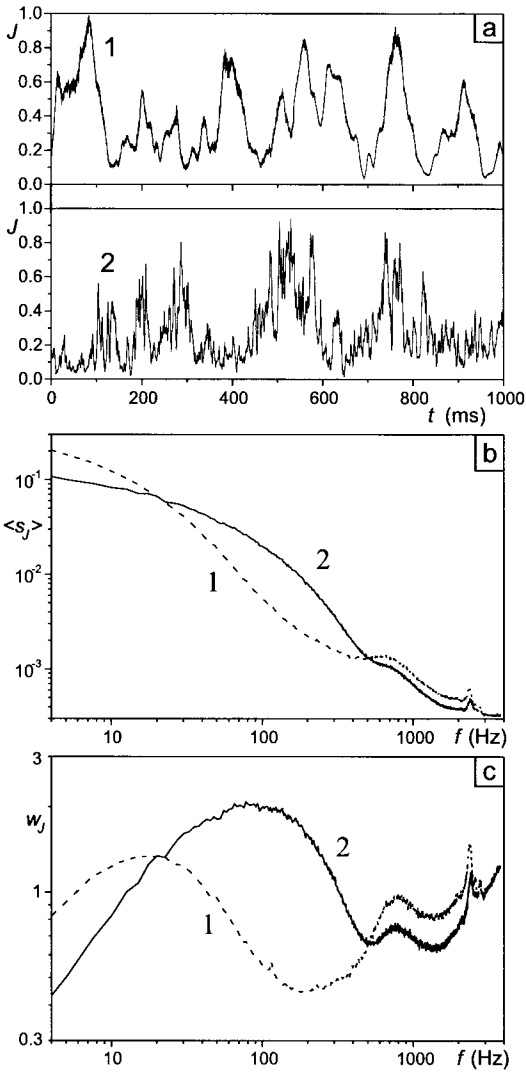


Fig. 8. (**a**) Two samples of the metric (received power) evolution $J(t)$ after 2.3 km atmospheric propagation; C_n^2 values were identical for the measurements but the wind speed was considerably higher for measurement 2 (beam steering was on in both cases). (**b**) The corresponding (averaged) power spectra $\langle s_J(f)\rangle$ and (**c**) the contents spectra $w_J(f) = f\langle s_J(f)\rangle$ reveal the difference in the the main spectral components (f = frequency). Note the frequency shift of the maxima of $w_J(f)$ by a factor of 5

5.3. Wave-Front Control with the SPGD AO System

For evaluation of the SPGD AO system performance we compared the average values of the metric (received power) with and without operation of the system. Since the AO system was not able to cancel out the strong fluctuations (cf. Fig. 8a) of the received

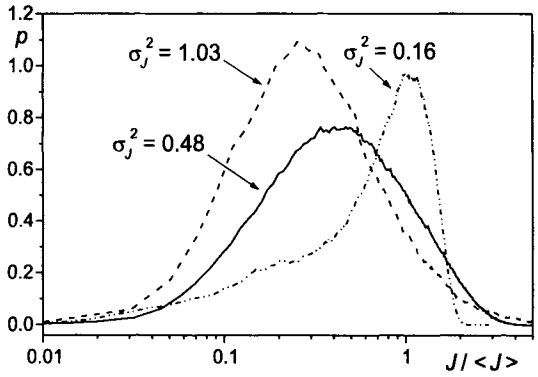


Fig. 9. Three sample probability density functions $p(J)$ for the received power J demonstrating the range of the variance $\sigma_J^2 = \langle I^2 \rangle / \langle I \rangle^2 - 1$ observed under various atmospheric conditions. J values were normalized by their corresponding mean value $\langle J \rangle$

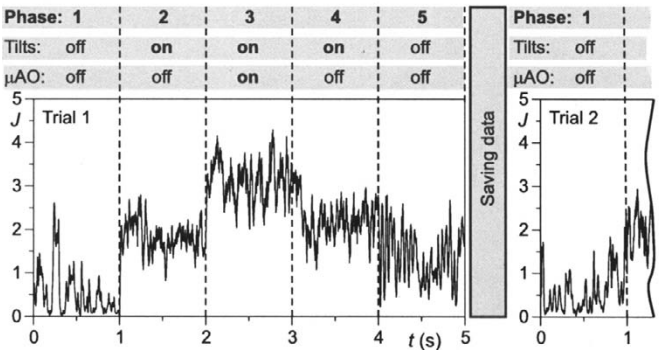


Fig. 10. Method for evaluation of the performance of the AO system: Metric (received power) values J were recorded for a trial with five phases of different operation conditions for the adaptive systems (as indicated) and subsequently saved before the next trial was started. Data evaluation (e.g. calculation of average metric evolution curves) was performed on data sets of 1000 or more recorded trials

power completely (due to the received wave's intensity scintillations and residual wave-front errors), a rather long integration time was required in order to obtain an acceptable error figure for the average power, $\langle J \rangle$. Monitoring the atmospheric conditions of our propagation path with a scintillometer showed that atmospheric conditions could change quite rapidly and correlation could be lost within minutes. Thus, it was not feasible to record $J(t)$ for a longer time when the AO system was off for obtaining the uncompensated power average, $\langle J_{\text{off}} \rangle$, and compare it with the value $\langle J_{\text{on}} \rangle$, recorded earlier or later with the AO system working. A separation between the effects from the AO system and from the changes in atmospheric conditions would not be possible.

Since correlation between turbulence conditions is higher within short temporal distances, we performed experiments where we switched among five different adaptive system operation conditions in short identical time intervals (usually one second or less) and repeated these trials many (typically 1000) times. The procedure is depicted in Fig. 10. In phase 1 we observed the metric J with the feedback control off for

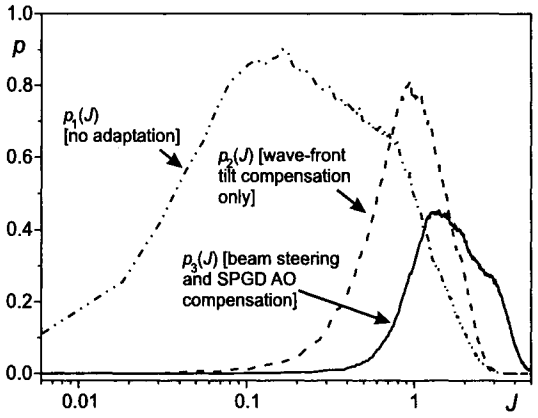


Fig. 11. Probability density functions $p_i(J)$ of received power values J for three different operation conditions of the AO system: no adaptation ($i = 1$, BS and μAO off), beam steering only ($i = 2$, BS on, μAO off), and beam steering together with high-resolution SPGD AO control ($i = 3$, BS and μAO on). Deep fading (occurrence of low J values) is strongly reduced with wave-front correction

both the fast beam steering and the SPGD AO systems; in phase 2 the beam steering (wave-front tilts compensation) system was on while the SPGD AO control remained off. In phase 3 both adaptive subsystems were operating. In phase 4 the μAO system was off, and finally in phase 5 no adaptive correction was used (both subsystems were off).

In each phase, a fixed number of iterations (typically several thousand) were performed. The values of the metric J measured for all iteration steps were saved after one trial (phases 1 through 5) was completed and before the next trial was started. By implementing this measurement scheme we were able to compare only data that were measured during close time intervals and thus minimized the influence of long-term fluctuations of atmospheric conditions, while short-term fluctuation effects were mitigated by averaging a large number of trials.

The influence of both the beam steering system (BS) and the SPGD AO system (μAO) on the probability density function, $p(J)$, of the received power values is shown in Fig. 11. Curves $p_i(J)$ were calculated for each of the phases $i = 1$ to 3 by processing the recorded metric data after the whole set of trial measurements was finished. Data points within the first 0.1 seconds of each trial phase were neglected so as to exclude the influence of the transition processes. The PDF curve $p_2(J)$ corresponding to wave-front tilt compensation is significantly narrower than the PDF $p_1(J)$ without compensation and its maximum value is shifted to larger values of the metric. Just wave-front tip/tilt compensation alone allowed reduction of the received power variance $\sigma_J^2 = \langle J^2 \rangle / \langle J \rangle^2 - 1$ from 0.53 (no compensation) to 0.19 (wave-front tip/tilt compensation). When the SPGD AO system was turned on in addition to the beam steering system (PDF $p_3(J)$), the probability of low values of the received signal, i.e., the fading responsible for bit errors in laser communication systems, was further reduced and the average $\langle J \rangle$ almost doubled. The average C_n^2 value during this experiment was 3.6×10^{-15} m$^{-1/3}$ corresponding to $\sigma_R^2 = 0.73$.

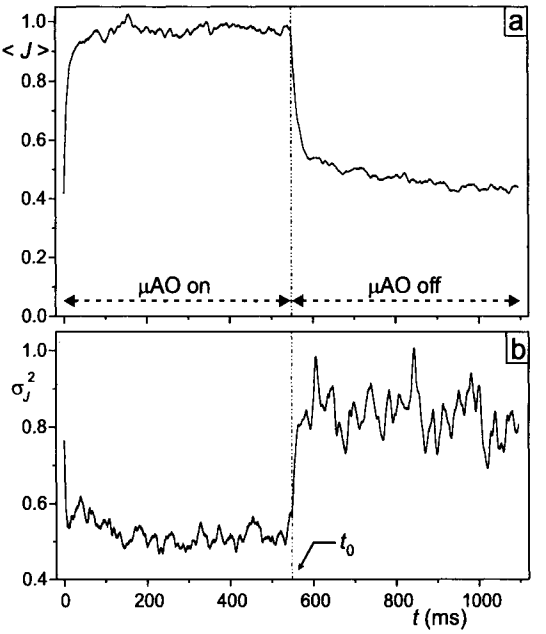


Fig. 12. Mitigation of atmospheric-turbulence induced wave-front distortions with the SPGD AO system (μAO). (**a**) The evolution curve demonstrates the improvement of the average received power $\langle J(t)\rangle$ after μAO is switched on at $t = 0$ and its decrease after feedback is switched off at $t \approx 550$ ms. The beam steering system is always on during the experiment. (**b**) The curve shows the corresponding improvement of the normalized metric variance $\sigma_J^2(t) = \langle J^2\rangle/\langle J\rangle^2 - 1$

Figure 12a shows the averaged metric evolution curve, $\langle J(t)\rangle$, for an experiment with $C_n^2 = 2.1 \times 10^{-15}\,\mathrm{m}^{-2/3}$ (Rytov variance $\sigma_R^2 = 0.42$). Since we were interested particularly in the performance of the SPGD AO system, a shorter trial scheme comprising only phases 3 and 4 of the measurement procedure shown in Fig. 10 was used, i.e., the beam steering system was continuously operating. The SPGD AO system (μAO) was on for 550 ms and then turned off during the consecutive 550 ms. The average curve $\langle J(t)\rangle$ was calculated from 1000 trials. As can be seen from Fig. 12a, operation of the SPGD AO system improved the averaged metric value $\langle J\rangle$ by a factor of 2.3. Metric fluctuations were significantly lowered by the μAO system as shown in Fig. 12b, where the normalized metric variance $\sigma_J^2(t) = \langle J(t)^2\rangle/\langle J(t)\rangle^2 - 1$ is plotted.

The dynamics of the average metric $\langle J(t)\rangle$ during the optimization process after switching the SPGD AO system on at $t = 0$ characterize the adaptation speed (and hence the closed-loop compensation bandwidth) of the system for compensation of distortions from atmospheric turbulence. In contrast, the decay of $\langle J(t)\rangle$ after μAO is switched off at $t_0 = 550$ ms reveals information about the temporal correlation and hence the dynamics of atmospheric-turbulence-induced phase distortions. Wave-front correction for the data shown in Fig. 12a was possible because the SPGD adaptation process convergence was faster than the characteristic time scale of turbulence induced distortions changes (80% of the metric improvement during the optimization process

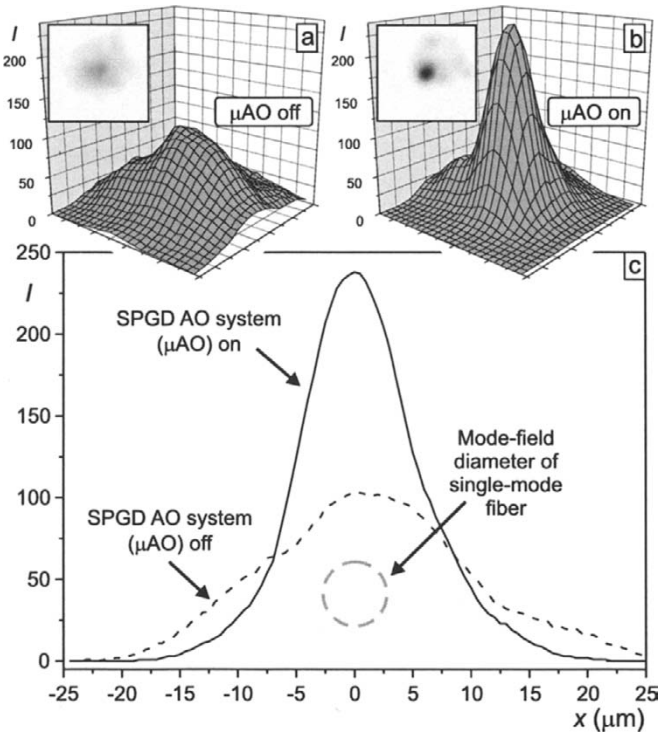


Fig. 13. (**a**, **b**) Long-exposure distributions of the irradiance $I(\mathrm{r})$ in the focal plane $\mathcal{F}''$ (in Fig. 5) and the corresponding 3-D plots recorded while the SPGD AO system (μAO) was off and on (a and b, respectively). Wave-front tilts control with the beam steering system (BS) was on in both cases. (**c**) The corresponding 2-D beam profiles were re-scaled to the dimensions in the focal plane $\mathcal{F}'$ of the receiver fiber SM_1. The mode field diameter of SM_1 (5.5 μm) is indicated as circle

was accomplished within 20 ms, while it took 50 ms to loose 80% of the improvement after the μAO system was turned off).

The long-exposure irradiance distribution $\langle I_{\mathcal{F}''}(\mathbf{r})\rangle$ in the focal plane $\mathcal{F}''$ recorded under the same atmospheric conditions with the camera *CCD* (see Fig. 5) for both operation conditions, SPGD AO system off and on, are shown in Fig. 13a,b along with the respective 3-dimensional plots of the focal plane irradiance. Beam steering was on for both conditions; we discuss here only the improvement of the beam in addition to beam steering. The irradiance pattern $\langle I_{\mathcal{F}''}(\mathbf{r})\rangle$ was obtained by averaging 400 frames recorded during one minute. The corresponding beam profiles are depicted in Fig. 13c. Similar to the metric values in Fig. 12a, the maximum intensity was about 2.3 times higher with the SPGD AO system operating than with beam steering only. However, the intensity profile curve obtained even with higher-order wave-front control (i.e., during μAO operation) was still about two times wider than the diffraction limit, which corresponds approximately to the mode-field diameter (5.5 μm) of the single-mode fiber SM_1 (indicated in Fig. 13c as dashed circle). Tests showed that this can not

be attributed to the optical system's aberrations, but is due to residual distortions from atmospheric turbulence.

5.4. *SPGD Adaptive Transceiver System*

In the atmospheric compensation experiments described in Sec. 5.3, the SPGD AO system was used in an adaptive receiver configuration (cf. Fig. 2a). However, the experimental setup allowed also to investigate an adaptive transceiver as shown in Fig. 2c. In order to transmit light from the laboratory to the water tower through the adaptive optics system, we coupled the light from laser diode *LD* (see Fig. 5) into single-mode fiber SM_2 and positioned the fiber end within focal plane $\mathcal{F}'''$ in a position conjugate to the receiver fiber SM_1 in plane $\mathcal{F}'$. After collimating the beam with lens L_9, the beam was coupled by beam splitter BS_3 with opposite propagation direction into the light path of the received beam.

The transmitter of the optical antenna T_2 (Fig. 1) was used as beacon, and its receiver unit determined the magnitude of the optical signal received on the water tower, i.e., the remote metric, J_2. Customized transceiver electronics digitized J_2 data so that they could be transmitted optically to the laboratory site, where they were received by a second transceiver unit, which provided the J_2 values to the supervisory computer of the SPGD system. Although data transmission was performed at a rate of nearly $10{,}000\,\mathrm{s}^{-1}$, the latency from the signal coding electronics was too large to use J_2 data as feedback signal for the SPGD controller. Hence, it was not possible to use this setup as an adaptive transceiver that uses a combination of the local and the remote metric (J and J_2, respectively), or even as an adaptive transmitter (Fig. 2b), which relies solely on feedback from J_2. Experiments were thus performed maximizing the local metric J while recording the simultaneously measured data for J_2.

Results for the averaged evolution curves of both metrics J and J_2 for an experiment with $C_n^2 = 2.1 \times 10^{-15}\,\mathrm{m}^{-2/3}$ ($\sigma_R^2 = 0.42$) are shown in Fig. 14. Again the trial consisted only of phases 3 and 4 (cf. Fig. 10), i.e., the SPGD AO system was periodically turned on and off. Indeed both metrics $\langle J(t)\rangle$ and $\langle J_2(t)\rangle$ were improved when the μAO system was working, i.e., the SPGD adaptive transmitter configuration was capable of mitigating beam spreading and increasing the irradiance at the receiver aperture on the water tower. Note that this improvement in J_2 was in addition to beam steering; the tilt compensation subsystem was always on).

Although there is a simultaneous improvement of metrics J and J_2 when the μAO system is working, the system's influence on J_2 is much less than on J. This can be explained by the anisoplanatism of the setup. Transmitter and receiver apertures of the lasercom antenna T_2 on the water tower (both 6 cm in diameter) are mounted with a distance (between centers) of about 8 cm corresponding to an angle of 34 μrad as seen from the AO system in the laboratory. This is larger than the isoplanatic angle, θ_{iso}, which can be estimated to $\theta_{\mathrm{iso}} = [1.09k^2C_n^2L^{8/3}]^{-3/5} \approx 13\,\mu\mathrm{rad}$ for the atmospheric conditions during the experiments (assuming a constant C_n^2 value along the propagation path). Since θ_{iso} corresponds to a displacement of 3 cm at the water tower, wave-front distortion compensation for the transmitted beam can not be as successful as under isoplanatic conditions. This example illustrates that a transceiver with an aperture shared by transmitter and receiver is important for the efficiency of atmospheric compensation with adaptive optics in free-space laser communications.

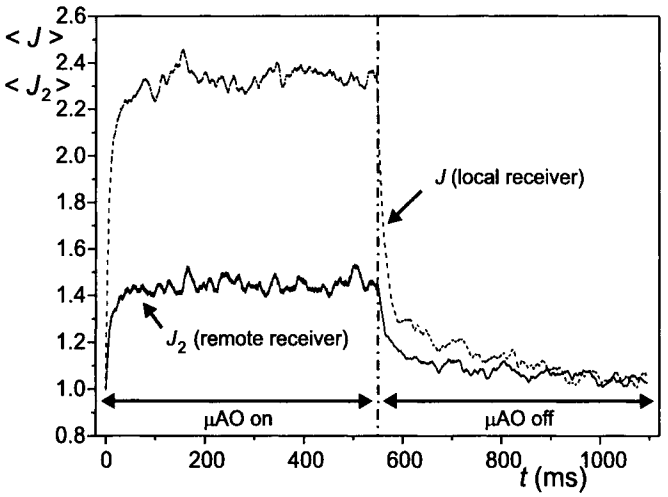


Fig. 14. Results from an experiment using an adaptive transceiver configuration: $\langle J(t) \rangle$ describes the evolution of the local metric, J (power of the beam received at the laboratory with the AO setup), which served as feedback signal for the SPGD controller. The SPGD AO system (μAO) was turned on at $t = 0$ and off at $t = 550$ ms while beam steering (BSS) was continuously working. The curve of the remote receiver metric, J_2, i.e., the signal strength of the beam received by transceiver T_2 at the water tower, demonstrates simultaneous improvement of incoming and outgoing beam.

6. Summary and Conclusion

Wave-front distortions induced by atmospheric turbulence can severely impact performance of free-space optical communications causing signal fading on a millisecond time scale. In principle, these aberrations can be mitigated by adaptive optics (AO), but conventional techniques with wave-front sensing and reconstruction face very challenging problems and perform poorly under conditions of strong scintillations caused by distributed turbulence along the propagation path. Using a laboratory-based non-conventional, SPGD based adaptive optics system, we demonstrated wave-front control for free-space optical communication scenarios under adverse conditions (scintillation index $\sigma_I^2 \simeq 1$) for optical wave propagation over a 2.3 km near-horizontal atmospheric path. The AO setup consisted of a beam steering system for tip/tilt control and the higher-resolution AO system that omitted wave-front measurements by using a stochastic parallel gradient descent (SPGD) algorithm for optimization of the received signal strength used as performance metric. Thus, SPGD AO offers an attractive alternative that can operate under strong scintillation conditions. Because of its system architecture, incorporation of SPGD adaptive optics in free-space laser communication terminals is relatively simple.

In the future, there are several ways to improve the performance of SPGD AO systems, i.e., primarily to increase convergence speed. First, there is the prospect of faster wave-front correctors [17]. Second, the SPGD system performance can be increased by the introduction of an additional low-resolution deformable mirror with an independent control loop for compensation of low-order but high-amplitude distortions (as defocus or astigmatism) in the AO system [29]. Third, advanced SPGD

controllers may implement "turbulence-friendly" perturbation statistics, which provide much faster convergence than Bernoulli statistics [14]. New SPGD control hardware must be developed before the advantage of such improved perturbation statistics can be demonstrated on an atmospheric propagation path.

Acknowledgment

The authors are grateful to their colleagues Gary Carhart and Matthew Banta at the Army Research Laboratory's Intelligent Optics Laboratory for their important contributions to the development of the atmospheric laser optics testbed used in our experiments. We also thank Thomas Bifano for providing the MEMS deformable mirror and Marc Cohen for assistance with the AdOpt VLSI controller. Part of the experimental work was performed while T. Weyrauch held a National Research Council Research Associateship at the Army Research Laboratory. This research was funded by the U.S. Joint Technology Office under contract JTO-02-602-18, the Battelle Memorial Institute under contract DAAD0599D7014, and the National Science Foundation under contract ECS-0010026.

References

1. I. S. Reed and X. Chen, *Error-Control Coding for Data Networks*, Vol. 508 of *The Kluwer International Series in Engineering and Computer Science* (Kluwer Academic, Boston, 1999).
2. J. W. Hardy, *Adaptive Optics for Astronomical Telescopes*, Vol. 16 of *Oxford Series in Optical and Imaging Sciences* (Oxford University Press, UK, 1998).
3. F. Roddier, *Adaptive Optics in Astronomy* (Cambridge University Press, UK, 1999).
4. C. A. Primmerman, T. R. Price, R. A. Humphreys, B. G. Zollars, H. T. Barclay, and J. Herrmann, "Atmospheric-compensation experiments in strong-scintillation conditions," Appl. Opt. **34,** 2081–2088 (1995).
5. B. M. Levine, E. A. Martinsen, A. Wirth, A. Jankevics, M. Toledo-Quinones, F. Landers, and T. L. Bruno, "Horizontal line-of-sight turbulence over near-ground paths and implications for adaptive optics correction in laser communications," Appl. Opt. **37,** 4553–4560 (1998).
6. N. B. Baranova, A. V. Mamaev, N. F. Pilipetsky, V. V. Shkunov, and B. Y. Zel'dovich, "Wavefront dislocations: topological limitations for adaptive systems with phase conjugation," J. Opt. Soc. Am. **73,** 525–528 (1983).
7. D. L. Fried, "Branch point problem in adaptive optics," J. Opt. Soc. Am. A **15,** 2759–2768 (1998).
8. A. Buffington, F. S. Crawford, R. A. Muller, A. J. Schwemin, and R. G. Smits, "Correction of atmospheric distortion with an image-sharpening telescope," J. Opt. Soc. Am. **67,** 298–303 (1977).
9. S. L. McCall, T. R. Brown, and A. Passner, "Improved optical stellar image using a real-time phase-correction system: initial results," Astrophysical Journal **211,** 463–468 (1977).
10. J. W. Hardy, "Active optics: a new technology for the control of light," Proc. IEEE **66,** 651–697 (1978).
11. T. R. O'Meara, "The multi-dither principle in adaptive optics," J. Opt. Soc. Am. **67,** 306–315 (1977).

12. J. E. Pearson and S. Hansen, "Experimental studies of a deformable-mirror adaptive optical system," J. Opt. Soc. Am. **67,** 325–333 (1977).

13. M. A. Vorontsov, G. W. Carhart, and J. C. Ricklin, "Adaptive phase-distortion correction based on parallel gradient-descent optimization," Opt. Lett. **22,** 907–909 (1997).

14. M. A. Vorontsov and V. P. Sivokon, "Stochastic parallel-gradient-descent technique for high-resolution wave-front phase-distortion correction," J. Opt. Soc. Am. A **15,** 2745–2758 (1998).

15. J. C. Spall, *Introduction to Stochastic Search and Optimization* (Wiley, Hoboken, NJ, 2003).

16. R. T. Edward, M. Cohen, G. Cauwenberghs, M. A. Vorontsov, and G. W. Carhart, "Analog VLSI parallel stochastic optimization for adaptive optics," in *Learning on Silicon*, G. Cauwenberghs and M. A. Bayoumi, eds., (Kluwer Academic, Boston, 1999), Chap. 16, pp. 359–382.

17. T. G. Bifano, J. A. Perreault, P. A. Bierden, and C. E. Dimas, "Micromachined deformable mirrors for adaptive optics," In *High-Resolution Wavefront Control: Methods, Devices, and Applications IV*, J. D. Gonglewski, M. A. Vorontsov, M. T. Gruneisen, S. R. Restaino, and R. K. Tyson, eds., Proc. SPIE **4825,** 10–13 (2002).

18. M. A. Vorontsov, G. W. Carhart, J. W. Gowens II, and J. C. Ricklin, "Adaptive correction of wave-front phase distortions in a free-space laser communication: system and method," Patent pending.

19. M. A. Vorontsov, G. W. Carhart, L. A. Beresnev, and J. W. Gowens II, "Adaptive WDM free-space laser communication system: low-order aberration compensation experiments," presented at the *International Symposium on Optical Science and Technology*, Denver, Colorado, Conference 5550 "Free-Space Laser Communications IV," 2–6 August 2004.

20. M. A. Vorontsov, G. W. Carhart, M. Banta, T. Weyrauch, J. Gowens II, and J. C. Carrano, "Atmospheric laser optics testbed (A_LOT): atmospheric propagation characterization, beam control and imaging results," In *Advanced Wavefront Control: Methods, Devices, and Applications*, J. D. Gonglewski, M. A. Vorontsov, and M. T. Gruneisen, eds., Proc. SPIE **5162,** 37–48 (2003).

21. M. E. Gravecha, A. S. Gurvich, S. S. Kashkarov, and V. L. V. Pokasov, "Similarity relations and their experimental verification for strong intensity fluctuations of laser radiation," in *Laser Beam Propagation in the Atmosphere*, J. Strohbehn, ed., (Springer, New York, 1978).

22. L. C. Andrews, R. L. Phillips, and C. Y. Hopen, *Laser Beam Scintillation with Applications* (SPIE Press, Bellingham, WA, 2001).

23. Y. A. Kravtsov, "New effects in wave propagation and scattering in random media (a mini review)," Appl. Opt. **32,** 2681–2691 (1993).

24. M. A. Vorontsov, G. W. Carhart, M. Cohen, and G. Cauwenberghs, "Adaptive optics based on analog parallel stochastic optimization: analysis and experimental demonstration," J. Opt. Soc. Am. A **17,** 1440–1453 (2000).

25. T. Weyrauch, M. A. Vorontsov, T. G. Bifano, J. A. Hammer, M. Cohen, and G. Cauwenberghs, "Microscale adaptive optics: wave-front control with a μ-mirror array and a VLSI stochastic gradient descent controller," Appl. Opt. **40,** 4243–4253 (2001).

26. T. Weyrauch and M. A. Vorontsov, "Dynamic wave-front distortion compensation with a 134-control-channel submillisecond adaptive system," Opt. Lett. **27,** 751–753 (2002).

27. T. G. Bifano, J. Perrault, R. Krishnamoorthy Mali, and M. N. Horenstein, "Microelectromechanical deformable mirrors," IEEE J. Sel. Top. Quantum Electron. **5,** 83–89 (1999).

28. J. C. Ricklin and F. M. Davidson, "Atmospheric optical communication with a Gaussian Schell beam," J. Opt. Soc. Am. A **20,** 856–866 (2003).

29. T. Weyrauch and M. A. Vorontsov, "Atmospheric compensation with a multiconjugate (piston-MEMS/modal DM) adaptive system," In *Target-in-the-Loop: Atmospheric Tracking, Imaging and Compensation*, M. T. Valley and M. A. Vorontsov, eds., Proc. SPIE **5552,** 73–84 (2004).

DOI: 10.1007/s10297-004-0025-x
Originally published in J. Opt. Fiber Commun. Rep. 2, 56–85 (2005)

Optical networks, last mile access and applications

E. Leitgeb, M. Gebhart, and U. Birnbacher

Department of Communications and Wave Propagation
Graz Technical University, A-8010 Graz, Inffeldgasse 12, Austria
Email: leitgeb@inw.tu-graz.ac.at

Abstract. Free Space Optical (FSO) links can be used to setup FSO communication networks or to supplement radio and optical fiber networks. Hence, it is the broadband wireless solution for closing the "last mile" connectivity gap throughout metropolitan networks. Optical wireless fits well into dense urban areas and is ideally suited for urban applications. This paper gives an overview of free-space laser communications. Different network architectures will be described and investigated regarding reliability. The usage of "Optical Repeaters", Point-to-Point and Point-to-Multipoint solutions will be explained for setting up different network architectures. After having explained the different networking topologies and technologies, FSO applications will be discussed in section 2, including terrestrial applications for short and long ranges, and space applications. Terrestrial applications for short ranges cover the links between buildings on campus or different buildings of a company, which can be established with low-cost technology. For using FSO for long-range applications, more sophisticated systems have to be used. Hence, different techniques regarding emitted optical power, beam divergence, number of beams and tracking will be examined. Space applications have to be divided into FSO links through the troposphere, for example up- and downlinks between the Earth and satellites, and FSO links above the troposphere (e.g., optical inter-satellite links). The difference is that links through the troposphere are mainly influenced by weather conditions similar but not equal to terrestrial FSO links. Satellite orbits are above the atmosphere and therefore, optical inter-satellite links are not influenced by weather conditions. In section 3 the use of optical wireless for the last mile will be investigated and described in more detail. Therefore important design criteria for connecting the user to the "backbone" by FSO techniques will be covered, e.g., line of sight, network topology, reliability and availability. The advantages and disadvantages of different FSO technologies, as well as the backbone technology are discussed in this respect. Furthermore, the last mile access using FSO will be investigated for different environment areas (e.g., urban, rural, mountain) and climate zones. The availability of the FSO link is mainly determined by the local atmospheric condi-

tions and distance and will be examined for the last mile. Results of various studies will complete these investigations. Finally, an example for realizing a FSO network for the last mile will be shown. In this example FSO transmitters with light emitting diodes (LED) instead of laser diodes will be described. By using LEDs, problems with laser- and eye safety are minimized. Some multimedia applications (like video-conferences, live TV-transmissions, etc.) will illustrate the range of applications for FSO last mile networks.

1. Optical Networks

Free Space Optics (FSO) links can be used to setup a complete FSO network or as a supplement to conventional radio links and fiber optics.

Most FSO links are point-to-point links between one transmitting and one receiving station. However, current research [1,6,9] is also investigating point-to-multipoint configurations. By using light sources like LEDs with a wide beam angle and/or multiple transmitter optics, the output power within the same laser class can be increased multiple times, and several terminal stations can be connected to central station.

1.1. Types of FSO Systems for Different Network Architectures

Different FSO systems for various applications have been developed by research groups and industry. In order to describe the different technologies implemented in available systems, we simplify the concepts to some important elements. Such a simplified scheme for a typical FSO unit for data transmission is shown in Fig. 1. On the basis of this concept, three types of system design can be distinguished, each as a compromise regarding reliable operation and installation costs for certain applications and distances.

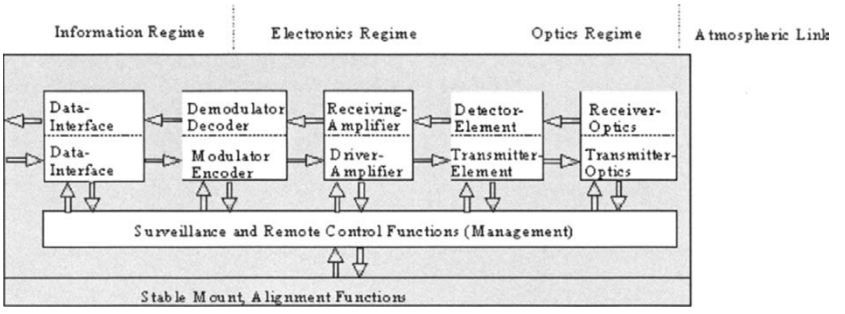


Fig. 1. Simplified scheme of the elements of a FSO unit.

For the first type of systems, available standard components are combined to build cost-effective solutions. Generally, the value of the beam divergence is high (up to 60 mrad) and allows an easy alignment for the user without the need for a telescope. All elements including the optics do not require as high precision as compared to the other two approaches, allowing higher tolerances at fabrication, and allowing the use of simple mounts at installation. Owing to large divergence the requirements for a stable underground are not very high, allowing a quick installation. Suitable distances for high availability operation are limited up to 300 m because of the wide beam angle (and depending on local climate). Owing to laser safety regulations the use of light

sources with higher beam divergence allow more output power in the same safety class, which improves the link budget for short distance applications.

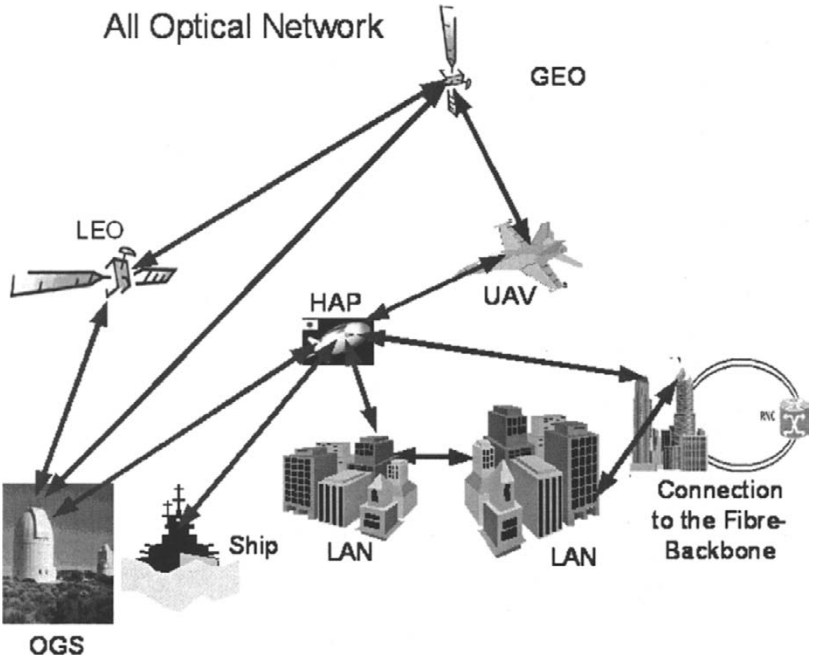


Fig. 2. All Optical Network.

Systems of the second type take advantage of a collimated beam as much as possible without active alignment, which leads to a beam divergence of about 2 to 6 mrad. The implementation of fiber-optic technology for coupling directly from fiber to FSO is a suitable concept. To achieve the same intensity at the receiver in longer distance, the concept requires precise optical components being carefully adjusted, and telescopes or adjustment procedures are needed for the installation. A stable mount on stable ground is required. These first two types of FSO systems are mainly used for short range terrestrial applications and especially for last mile access. For increasing the intensity at the receiver and to overcome atmospheric turbulences (scintillation) caused by variation of the refraction index, more than one transmitter and/ or receiverunits are used in one housing. Systems with multiple transmitter/receiver units are called Multibeam- or Multlink-FSO systems and are found in both system types.

The third category contains the sophisticated solutions, including strategies to extend the distance for reliable operation as much as possible. The beam divergence reaches values less than 1 mrad leading to low geometrical losses, concentrating all incoming light on a small area around the receiver. To compensate building sway and deviation, automatic tracking of the beam is implemented, adaptive optics may help to compensate atmospheric fluctuations or allow higher optical output over a large area. These systems are in use for long distances and the most accurate solutions with acquisition and tracking for space applications. For space applications not only direct detection for the receiver is used, but also coherent detection.

Using FSO systems from all three categories, various communication links or networks can be realized, including terrestrial FSO (between buildings, hospitals, campus), horizontal and slant paths, air-borne and space communications (between aircraft, unmanned aerial vehicles (UAV), high altitude platforms (HAP) and satellites. By connecting FSO links and networks to the Backbone realized with optical fibers, the FSO networks and links can be seen as parts of a global all optical network around the world (Fig. 2). A very important issue is the reliability and the availability of FSO links and FSO networks. In order to estimate the reliability of a network, the reliability of each single component has to be taken into account. For exampled, the reliability of one FSO terminal depends on the reliability of all electrical and optical components built into the terminal, including the connection and interface to the network (Fig. 3). The reliability of the FSO link is determined by the reliability of the transmitting FSO terminal (unit 1 at location A) and the receiving FSO terminal (unit 2 at location B) and the quality of the optical path in-between. The reliability and availability of the optical path is mainly influenced by the local weather conditions, with fog being the most limiting factor. If not only a single FSO link but a network of optical wireless connections is used, the overall reliability is calculated by taking into account all the links in the network and their specific arrangement to each other (parallel or serial). Hence, different network architectures have different availabilities and reliabilities.

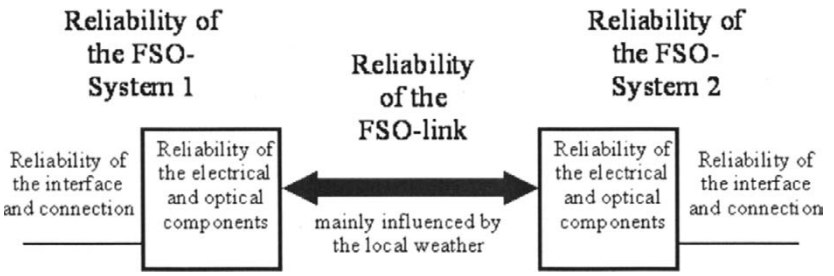


Fig. 3. Reliability of FSO links.

1.2. Architectures of FSO Networks (Point-to-Point and Point-to-Multipoint Configurations)

In the following section different architectures are described and compared regarding the overall reliability.

1.2.1. Optical Wireless in Ring Architecture

In Fig. 4 a FSO network in ring architecture is shown. In the given example, the distances between the buildings are up to 500 m. In the minimum configuration two optical receiver/transmitter units are installed on the top of each building. Optical Repeaters have to be used, if there is no line of sight between transmitter and receiver.

In the event of a broken link or for example a link failure between buildings 1 and 2, the indirect connection can be used. Thereby, the information is sent in the other

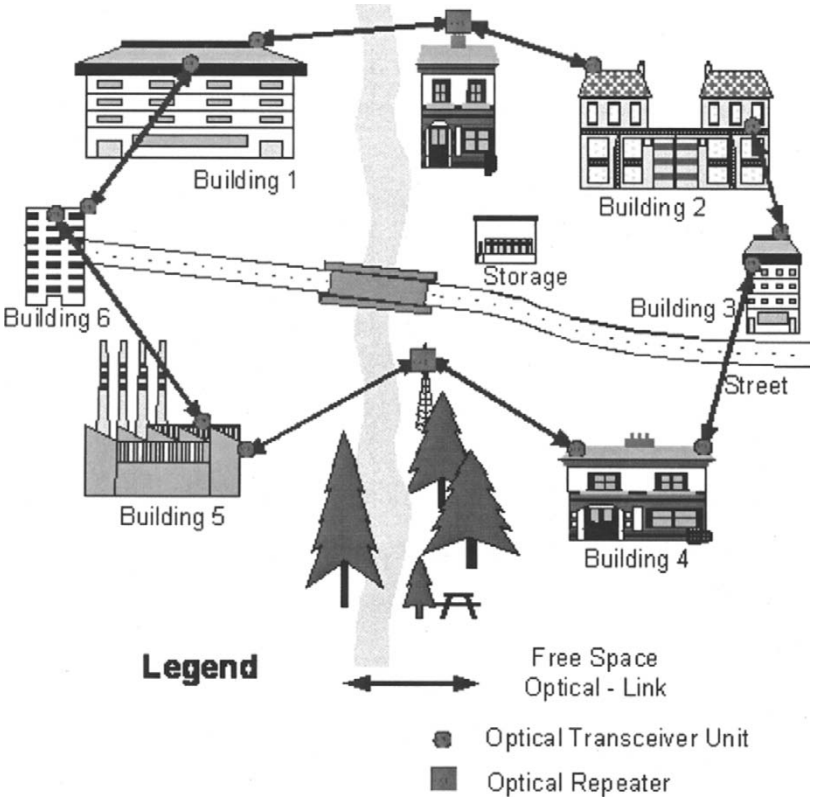


Fig. 4. FSO network with ring architecture.

direction of the ring network passing buildings 1, 6, 5, 4, 3, and 2. Hence, using a ring architecture, a partial security against failure can be achieved. The installation of additional, redundant links increases the availability and the security against failure.

1.2.2. Optical Wireless in Star Architecture

In Fig. 5 a test installation at the University of Technology Graz in Austria (TU Graz) is shown, using a star architecture. The coverage area of this FSO network is about 300 m in diameter. An optical multipoint unit is mounted at the roof of a building. Five user terminals are permanently connected by their optical transceiver units to the optical multipoint hub station. The five user terminals are located at the surrounding buildings and offer a connection via the FSO link to the optical multipoint unit. In this configuration the optical multipoint unit is realized with five FSO Point-to-Point units, each of them directed to one user FSO terminal. The optical multipoint unit is interconnected by switches with the backbone network of TU Graz.

The advantage of this configuration is the shorter distance between any two FSO units. In general, the optical multipoint unit is located in the center of the area, mounted on the tallest building. But this architecture has the disadvantage of a single point of

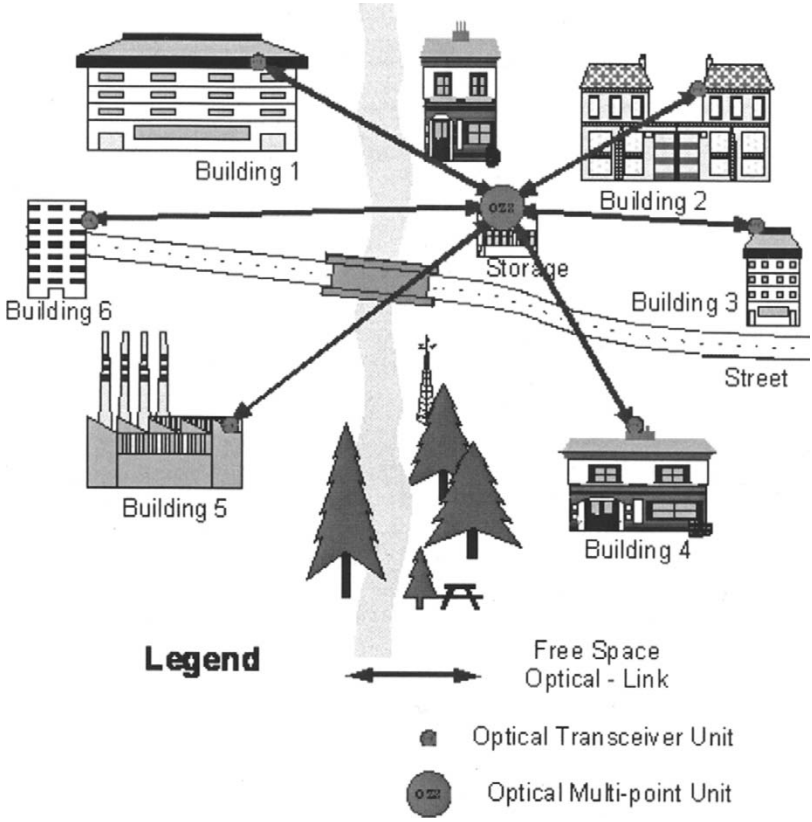


Fig. 5. FSO network with star architecture.

failure. If the optical multipoint unit fails, a system breakdown of the whole installation is caused. To improve the reliability of this architecture, a redundant Multipoint Unit would have to be installed. The second optical multipoint unit can also be mounted on moveable platforms, e.g., a van, for increasing the flexibility and decreasing the setup time. However, for the installation of multipoint units on cars, FSO systems with "auto tracking" are preferable.

1.2.3. Optical Wireless in Meshed Architecture

For high reliability, the optimum network architecture is a meshed network. Meshed networks combine the benefits of the above described architectures, because different connections are possible. An example of a meshed FSO network is given in Figs. 19 and 20 in section 3.

1.3. Connecting to the Backbone

The access network resides between the individual subscribers and the network operator's backbone network. fiber often does not reach paying customers. Only 3% of

worldwide businesses are on fiber, and 75% are within a mile from fiber. Nowadays the backbone is realized with fiber and so the connection with FSO offers many advantages. The FSO systems work protocol transparent like a fiber link. Hence, the same networking technologies used in fibers can be used over FSO. It is possible to couple from FSO systems directly into the optical domain or with optical/electrical conversion and regeneration to the network-fiber. Also coupling into a Wavelength Division Multiplex (WDM) is possible, by connecting different optical wireless systems with a wavelength division multiplexing unit to the WDM-fiber network.

In different architectures (section 1.2) and configurations, the FSO unit can be connected to satellites, directional radio links, (mobile) telephone networks, or fiber optics. In Fig. 6 a connection to a fiber- based backbone realized with a Point-to-Multipoint architecture is shown. The optical multipoint unit is connected with a switch or router to the backbone network (thick solid line in Fig. 6). The users in the buildings 1, 2, 3, and 4 are linked with their FSO terminal units to the central optical multipoint unit.

2. FSO Applications

This section covers terrestrial applications for short and long ranges and space applications. Terrestrial applications and FSO links trough the troposphere are mainly influenced by weather conditions. Therefore, some important characteristics of the atmosphere have to be discussed before describing the FSO applications in more detail.

The lowest part of the atmosphere up to 10 km above the Earth's surface is called the troposphere or the weather sphere. It has a varying refraction index which is dependent on the height above the Earth's surface. Normally the refraction index decrease with the height, but at inversion situations there is different relationship.

Atmospheric conditions degrade laser communications through the atmosphere in two ways. First, the atmosphere acts as a variable attenuator between the transmitting and receiving terminals. Second, a free space laser link is subjected to scintillations.

Attenuation is caused by the weather conditions along the transmission path. Generally, there is low atmospheric attenuation during clear days and high attenuation during foggy days. Rain does not influence optical transmissions heavily, because raindrops have the size of a few millimeters and are large compared to laser wavelengths (1.5 microns) and thus cause minimal scattering of the laser energy. Furthermore, water has minimal absorption at a 1550-nm laser wavelength. Therefore, it is not surprising that the optical transmission is not heavily impacted by rain (only about 3 dB/km). Similarly it is not surprising to find out that optical transmission is impacted dramatically by heavy fog (30 dB/km). This is because the fog aerosols have a comparable size as the used wavelengths, causing much scattering of the laser energy as the fog gets thicker. Absorption effects can be subdivided into absorption and scattering effects. Absorption is caused by many different species of gas in the atmosphere, the dominant one being water vapor, which is in the wavelength region used for wireless optical links. There are two types of scattering, Rayleigh and Mie scattering (see section 3). The amount of scattering depends on the particle size distribution and the density of particles. Generally, any optical wavelength could be used for FSO, but because of the atmospheric conditions and due to the laser safety regulations, 1550 nm is the best

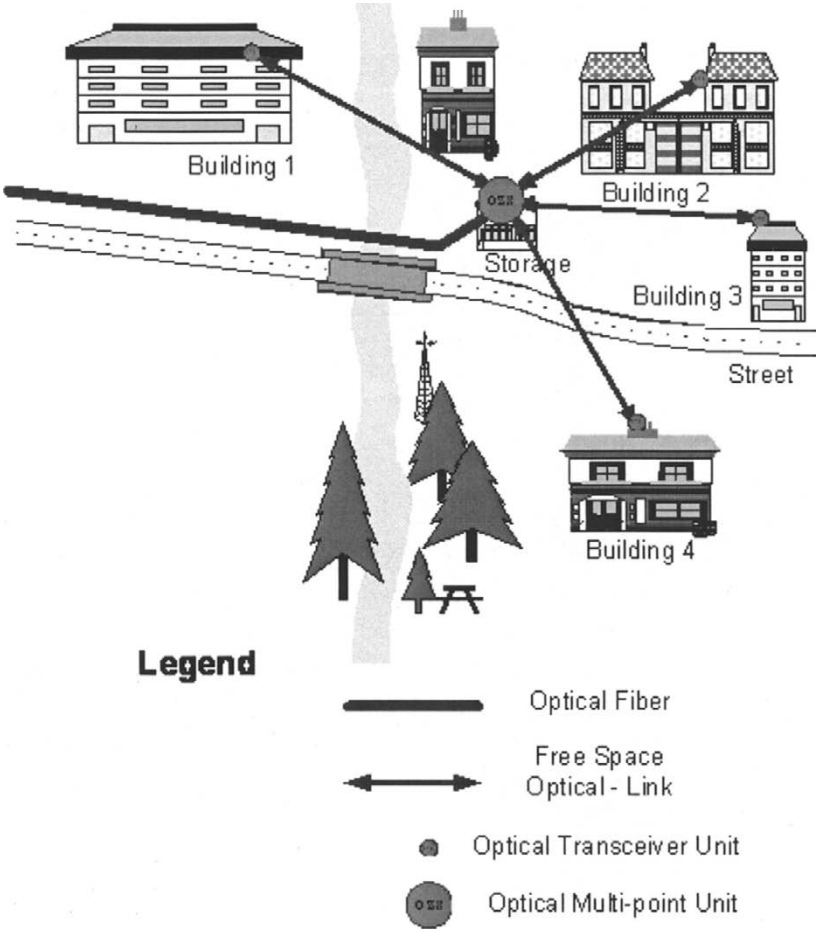


Fig. 6. Optical Wireless Access with Point-to-Multipoint Architecture.

suitable wavelength. The losses due to Mie-scattering in haze or light fog are smaller at longer wavelengths (1550 nm) than at shorter ones (850 nm).

The second major influence on FSO transmission is scintillation, which is caused by small-scale fluctuations in the refraction index of the atmosphere. Its primary effect is signal fading due to phase changes in the wave front arriving at the receiver. Unless the receiver has a very high dynamic range or the aperture is large enough to average out the scintillation spots, this can have an extremely detrimental effect on the signal. As described above, a foggy day is very bad for free space communication due to the high attenuation. On the other hand, it can be observed that scintillations are very low on a foggy day, because of low fluctuations in the atmosphere.

2.1. *Short Range Aapplications and Last Mile Access*

FSO short-range applications will be used up to 1 km (maximum 2 km) for connecting buildings or establishing a connection to the backbone.

Within the range of up to 2 km distance, it makes sense to use FSO systems of type 1 or type 2 (see section 1.1). These systems will be cost-efficient and they have a high availability and reliability. For short ranges, "auto tracking" is not necessary. In [1] the ranges for developed type 1 FSO systems are 100, 300, and 800 m (Figs. 7A and B). In Fig. 8 an example for a type 2 system is shown.

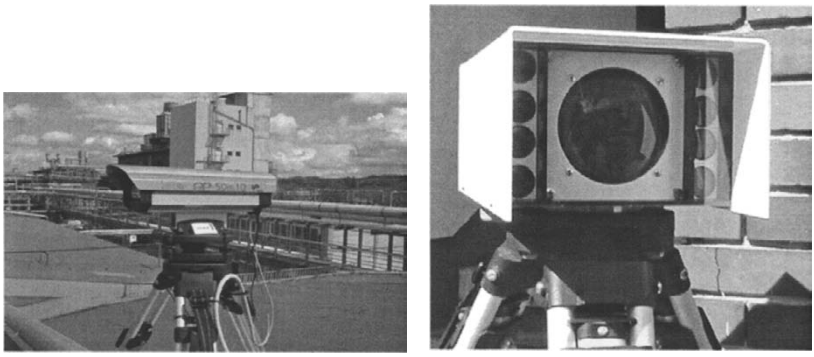


Fig. 7. Point-to-Point system of type 1 for ranges of A) 100 and B) 800 m (with 8 transmitter units).

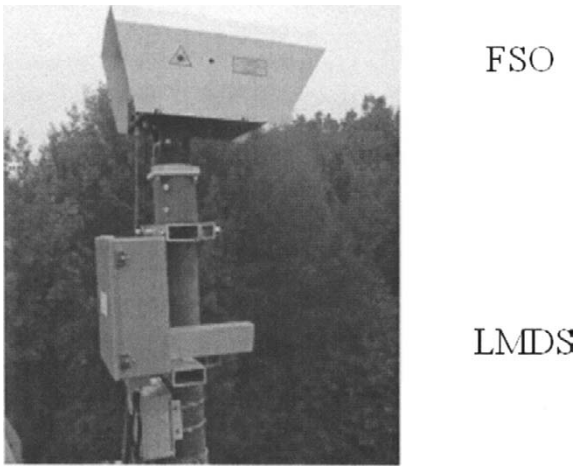


Fig. 8. Installation of a GoC Multilink 155F in combination with LMDS.

In [2,6] various FSO systems have been evaluated in regard to different weather conditions. For terrestrial applications it is important to know the influence of the path through the atmosphere, which can be either simulated and modelled or measured over long time periods. The atmospheric path for transmission of a collimated beam of light may be seen as an information channel, to which information is added and on which information gets lost (Fig. 9). A variety of effects can be described and several theoretical models are known, but not all of them are relevant to this technology. Once more it is a question of technology used, distance and application to find out relevant impacts on FSO data transmission. Before coming to quantitative results from experiments, a qualitative introduction to some relevant problems is given in this section.

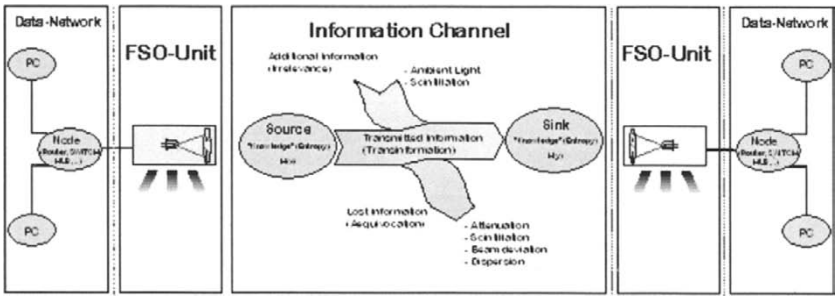


Fig. 9. Scheme of data transmission and impacts on the channel.

Ambient light that passes the optics and reaches the detector element, which usually is a PIN-photodiode or an avalanche-photodiode (APD), causes additional current that leads to additional white Gaussian noise in the receiver. Depending on the semiconductor materials and construction principle of the detector there are additional effects coming up. Flicker noise produces irregular disturbances which, if we consider data transmission, affects throughput of longer data packets most. At the same time the noise level raises, the impedance of the detector element decreases, which in most receiver principles leads to a level reduction of the received information signal, sometimes also referred as "burn out" effect. The dominant source of ambient light for outdoor FSO systems is the sun. Therefore the sunlight and the possible positions of the sun have to be considered for practical applications.

An intensity of about 1340 W/m^2 reaches the Earth, also known as the solar constant, which is reduced to about 1000 W/m^2 global radiation on ground under clear sky conditions due to wavelength dependent absorption in the atmosphere. Over 90% comes as direct sunlight, the rest looses directivity and is scattered over the sky contributing to background illumination. The sun is seen from the Earth under an angle of about 9.31 mrad (diameter) with a relative movement of 0.0728 mrad/s. The so-called windows of transmittance of the atmosphere include the two most important wavelength regions, leading to specific intensities of about 0.5 to 0.9 W/m^2 and per nanometer (W/m^{2}nm) in the 850-nm spectral region and about 0.17 W/m^{2}nm in the 1550-nm region for direct sunlight on ground. All these figures are estimations for clear sky conditions being subject to atmospheric influences, leading to variations. In general, the impact of sunlight on FSO depends on the angular and the spectral

sensitivity of the receiver and the possibility of direct sunlight at the receiver, referring to orientation and mount of the installed system.

Attenuation is the most critical factor for longer FSO links. The contribution of the free atmosphere is comparatively small at the most commonly used wavelengths around 850 and 1550 nm. Values for specific attenuation between 0.2 dB/km under very clear atmospheric conditions up to about 10 dB/km due to dust in urban regions are reasonable.

Far more critical is the impact of the weather situation. It has been reported, that specific attenuation may temporarily raise to more than 300 dB/km in heavy fog, even though these are very rare occasions depending on local climate (and within 3 years of measurement have never been observed in Graz). The situation can be explained by a theory describing the interaction of electromagnetic waves and particles, which can be simplified depending on the ratio of particle size and wavelength, to geometrical optics for interaction with comparatively large particles causing wavelength-independent absorption, and to Rayleigh-scattering for comparatively small particles. The most critical condition appears when both are approximately in the same order, such as optical wavelengths in the order of 1 μm and haze or fog that consists of water droplets at diameters from about 1 to 15 μm. This can be characterized by Mie's theory of scattering.

However, many factors depending on the properties of the particles are not known, therefore a different approach leads to a well known deterministic formula (1) based on visibility, which can be used in practice. Attenuation caused by scattering can be estimated by

$$a_S \cong \frac{17}{S}\left(\frac{555}{\lambda}\right)^{0.195-S}. \tag{1}$$

In this equation a_S is the specific attenuation in decibel per kilometer, S is the visibility for human eyes (sight) in kilometers and λ is the wavelength of transmitted light in nanometers. Usually, the range of visibility is defined as a path of 2% transmission in air (corresponding to about 17 dB attenuation at 555 nm wavelength), the contrast resolution of the human eye. Records of visibility are available at airports or meteorological stations and may be used as one reference to estimate the probability of fog for the local climate of a certain location. The impact of rain is less critical because of the larger particle size in the order of 0.1 to about 5 mm diameter, which has more effect on longer wavelengths such as millimeter waves at several tens of GHz. Falling snow simply absorbs the light by the irregular shapes of particles in the size of about 2 up to 25 mm, leading to a varying attenuation depending on the relation of particle and receiver optics area. For the same visibility range the conditions for FSO are worse than fog, although site diversity realized by using multiple transmitter optics can improve the quality of the link.

Light transmission in media follows the principle of Fermat, according to which the way of light from one point to another follows the shortest optical path length, which depends on geometrical distance and optical properties of the medium, given by the fraction index n. Consequently the fraction index has an impact on light propagation. For free air at sea level n is in the order of 1.0003, but depends on many factors, e.g., the wavelength used for transmission, the temperature, the atmospheric pressure and the humidity of the air.

Different sheets of air being crossed by light lead to beam deviation (refraction of light) due to a temperature or pressure gradient, an effect which leads to results similar

to misalignment or building sway, causing a reduction of received power which can get critical for very small divergence.

Short-time fluctuations of n, so-called turbulence cells, lead to irregular changing intensity in the beam, wave-front distortion (changing the angle of received light and causing a spot of light dancing around the focal point) and changing deviation, causing a varying input power at the receiver, also known as scintillation. The largest gradients and therefore strongest effects are observed near hot surfaces like streets, rooftops or side walls, or over outlets of air conditions or in winter heaters.

The influence of the turbulence cells is not interesting for short ranges, but very important for long distances (section 2.2). Type 1 and 2 systems will be cost effective and they will have a high availability and reliability at the last mile area.

2.1.1. FSO in Combination with Satellite and Wireless LAN

As an example for short-range applications, the use of FSO for connecting a satellite-Earth station to the backbone or to a local area network is described. This is a example for a temporary installation and shows the flexibility of FSO systems.

The satellite videoconferencing system was used at a tele-medicine conference in Vienna (Fig. 10), but it is also well suited for any conference, for tele-teaching and for military issues. In Fig. 11 we can see the video presentation of a doctor in Graz and experts at the conference in Vienna. In the lecture room in Vienna Wireless LAN was available. In order to connect the Wireless LAN with the mobile satellite-Earth station outside the UNO-City building, a FSO link was used.

2.2. Long Range Applications

For communication links, which run over distances above 2 km, it is important to use systems of type 3 technology. An example of a commercially available system is shown in Fig. 12.

Another important installation is the Wallberg experiment [4] in Germany at the German Aerospace Centre (DLR). DLR has performed an optical free-space data transmission experiment along with the European Aeronautic Defence and Space Company (EADS), Germany and Contraves Space AG, Switzerland. The scope of this experiment is to verify the tracking capabilities of the OPTEL 02, a space-qualified optical terminal, and to demonstrate optical high data rate transmission through the atmosphere. Two laser diode transmitters at 980 nm, each mounted on a tripod for static pointing and laterally separated by about 4 m, were placed on a mountain top in the German Alps at a height of 1620 m. Either a pseudo-noise pattern or video data can be transmitted. The OPTEL 02, performing acquisition and tracking, is situated at the DLR site in Oberpfaffenhofen near Munich at 620 m. An APD receiver front end is connected to an additional 75-mm telescope in order to receive data up to 270 Mbps. The optical path length between both terminals is 61 km.

At Wallberg a building with two rooms with one window each, physically separated by a distance of four meters, is at their disposal. Despite facing towards a popular foot path, there are no problems with laser safety. The two separated windows allow the use of a two-transmitter concept to reduce scintillation effects. Figure 13 shows a map of the southern Munich area showing Wallberg and Oberpfaffenhofen including an altitude profile for the optical link.

Fig. 10. Mobile satellite-Earth station with FSO Connection demonstrated at the UNO-City Vienna.

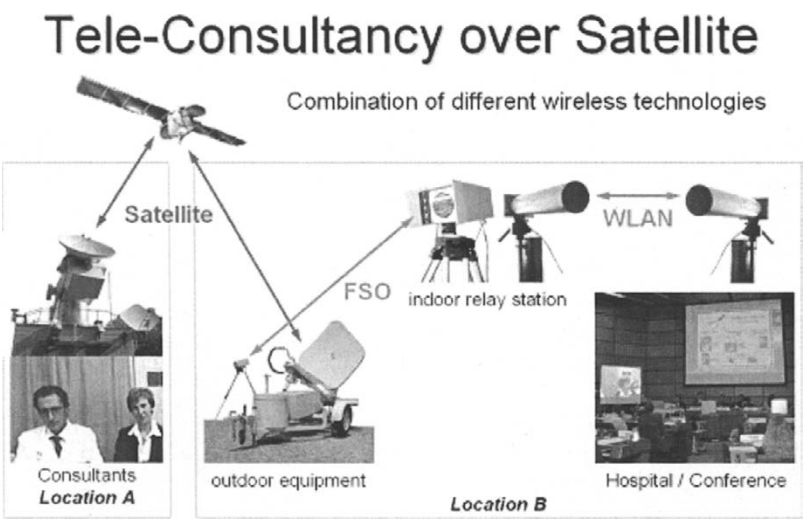


Fig. 11. Video conference connection between Graz (Location A) and Vienna (Location B) for tele-medicine.

Fig. 12. DT-50/DT-IF156 CanoBeam.

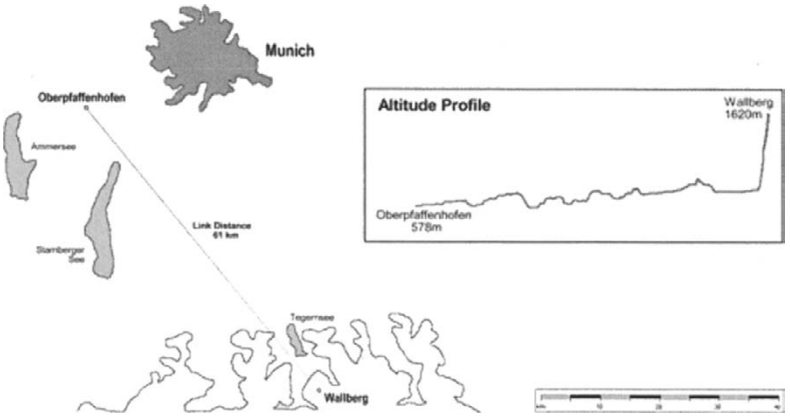


Fig. 13. Location and scenario of the Wallberg experiment.

As can easily be seen in the profile, there is a steep slope from the Wallberg mountain top to Lake Tegernsee on the plain. This gives a good reason to place the transmitter on top of the mountain, as little atmospheric disturbance can be expected there. The last few kilometers of the link are close to ground. North of Starnberger lake, there is a marshland, which might cause severe beam disturbance.

The described link is a terrestrial FSO path for very long distance. Hence, the beam divergence must be smaller than 1 mrad, resulting in low geometrical losses and concentrating all incoming light on a small area around the receiver. To compensate building sway and deviation, automatic tracking of the beam is necessary. Adaptive optic helps to compensate atmospheric fluctuations. The path characteristic of the Wallberg experiment is similar to space applications (section 2.3).

2.3. *Space Applications (Aircraft and Satellites)*

The main advantages of FSO links compared to microwave links are small high gain antennas (telescopes), light terminals, highest possible data rates at low signal power, no interference with other transmission systems and tap-proof operation using coherent methods. Possible applications for optical free-space links are inter-satellite links (ISL) in satellite networks, links for deep space missions, links between unmanned aerial vehicles (UAV) and high altitude platforms (HAP) and data links from geostationary satellites (GEO) to earth ground stations.

Space applications are divided into FSO links in the troposphere (for example, up- and down-links between the Earth and satellite) and FSO links above the troposphere (e.g., optical inter satellite links). Links through the troposphere are mainly influenced by the weather conditions similar but not equal to terrestrial FSO links. Optical inter satellite links are not influenced by weather conditions, because satellite orbits are above the atmosphere.

While for terrestrial optical wireless the most effort is given to compensate for the local weather conditions, for space applications, the acquisition and tracking [7] is the greatest challenge, because satellites and aerial vehicles have relative movements. Hence, we can see that for these applications only the most expensive systems of type 3 can be used.

Regarding FSO links through the troposphere, there is furthermore a difference depending on the communication direction. On the up-link, the FSO beam from the Earth to the satellite, the divergence of the optical beam is increasing quickly, because the refraction-index is decreasing with height and therefore the optical ray will bend away from the normal. Hence, the light beam gets wider on its first part of the path and is affected by the troposphere.

On the down-link, the FSO beam from the satellite to the Earth, the narrow beam transmitted from the FSO unit onboard the satellite is influenced by the troposphere on the last 10 to 15 km of its path. Because of the character and distribution of the refraction index in the troposphere, the light is coming from the medium with lower index and is going to the medium with higher index. Hence, the ray bends towards the normal. These effects result in a larger beam diameter for the received up-link beam compared to the received beam of the down-link, which has a smaller diameter.

An interesting European experiment is shown in Fig. 14, the world first data transmission between satellites using laser light which took place at 22nd November 2001. A data link between satellites was established using a laser beam as signal carrier. On board ESA's Artemis satellite is the SILEX [14,15] system. This system provides an optical data transmission link with the CNES Earth observation satellite SPOT 4, which is orbiting the earth at an altitude of 832 km, while Artemis was temporarily in a parking orbit at 31,000 km. Through the laser data link, images taken by SPOT 4 can be transmitted in real-time to the image processing center in Toulouse, France, via Artemis, thus drastically reducing the time between taking the picture and its delivery to the Earth. This is possible whenever the two satellites are in line of sight. Without the Artemis relay the images are stored on board in SPOT 4's memory and dumped to the ground stations at a later time.

The experiment performed consisted in establishing the link four times: in the course of four successive SPOT 4 orbits, the SILEX terminal on board Artemis activated its optical beacon to scan the area where SPOT was expected to be. When contact was

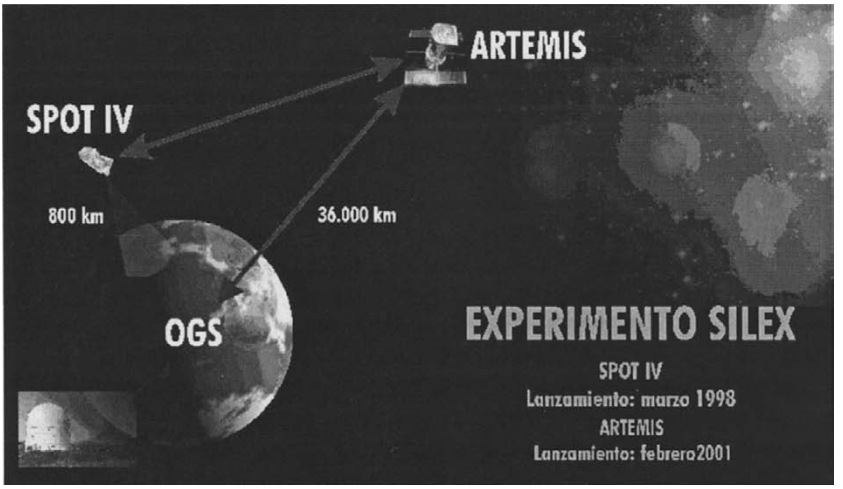


Fig. 14. SILEX Experiment.

made, SPOT 4 responded by sending its own laser beam to Artemis. On receiving the SPOT 4 beam, Artemis stopped scanning and the optical link was maintained for a pre-programmed period lasting from 4 to 20 minutes.

During the period when the two satellites were "communicating", test data were transmitted from SPOT 4 to the ground via Artemis at a rate of 50 Mbps. The extremely high accuracy of the data stream was confirmed at ESA's test station in Redu (Belgium) and the SPOT 4 receive station in Toulouse.

The main challenge in establishing an optical link between satellites is to point a very narrow beam with extreme accuracy to illuminate the partner spacecraft flying at a speed of 7000 m/s. This experiment was performed under worst-case conditions since Artemis was not in its nominal geostationary position but in a lower parking orbit, circling the Earth every 19 hours. The experiment was preceded by a series of tests a week earlier, during which a link was established between Artemis and ESA's optical ground station in Tenerife. Those tests demonstrated the correct operation of the SILEX terminal and paved the way for the subsequent steps.

After moving Artemis to its final geostationary orbit at 36,000 km the operational phase has started and the link between the two satellites is established at least 5 times a day. The SILEX system consists of two terminals: one on board Artemis, the other on SPOT 4.

For optical wireless space applications there have been developed coherent receiver instead of the direct detection. The optical free-space links have to cover large distances in space applications (inter-satellite links, inter orbit links, connections to deep space probes), the main criterion for selecting an appropriate transmission scheme is power efficiency or the sensitivity of the receiver. Hereby coherent methods are superior to direct detection methods, which are commonly used in terrestrial communications. In particular the PSK homodyne receiver shows a very high sensitivity. The advantages of a homodyne receiver against a direct detection concept are a higher sensitivity and less problems with background light. One key element of the homodyne receiver is the synchronisation of the local oscillator laser with the incoming signal using a

rather complex optical phase locked loop (OPLL). If the optical link goes through the atmosphere the atmosphere will cause wavefront distortions and fading of the incoming light. These effects make the realisation of an OPLL very complicated and also reduce the error performance in an direct detection system. So special techniques like array detection, channel coding or diversity are studied [4].

3. Last Mile

Now the use of optical wireless for the last mile will be investigated and described in more detail. Therefore line of sight, reliability and availability, different FSO techniques for the last mile and multimedia applications are described. Results of various studies will complete these investigations. Finally, an example for realizing a FSO network for the last mile will be shown.

3.1. Line of Sight

Optical wave propagation through the atmosphere requires a free line of sight from the location of the transmitting terminal to the receiving terminal. Even though receivers for infrared transmission are more sensitive compared to the human eye in the visible light region, the influence on the radiation is very similar. Especially for networking applications, where high availability is essential, the weather influence is a key factor. Visibility data, collected over several years either by free eye estimation or, more accurately, with a transmissionmeter instrument at meteorological stations or at airports as runway visibility range (RVR) can be used to calculate availability in free-space optic network planning. Atmospheric transmission can be described by the Beer-Lambert law:

$$\tau = \frac{P_d}{P_0} = e^{-\gamma d}. \tag{2}$$

The transmission τ for an optical wavelength means the relation of the optical power P_d at the end of an atmospheric path of distance d to the optical power P_0 which was originally sent (Fig. 15). According to the Beer-Lambert law transmission can be expressed by the extinction coefficient γ and the atmospheric path distance d. The extinction coefficient γ is constituted by processes of absorption α and scattering β by particles in the atmosphere. In detail, these particles can be the molecules of the atmospheric gas constituents and the larger aerosol particles:

$$\begin{aligned} \gamma(\lambda) &= \alpha(\lambda) + \beta(\lambda) \\ &= \alpha_m(\lambda) + \alpha_a(\lambda) + \beta_m(\lambda) + \beta_a(\lambda). \end{aligned} \tag{3}$$

For the wavelengths used in FSO, especially 850 and 1550 nm, atmospheric absorption can be neglected. Scattering by molecules $\beta_m(\lambda)$, which can be described by the Rayleigh scattering model, only get important for optical wavelengths smaller than 400 nm. The main important process that remains is the scattering of light by aerosol particles $\beta_a(\lambda)$ in fog. It can be described by the Mie scattering model for single interactions between particles and waves.

By definition the visibility range V is the atmospheric path distance for a transmission of 2%. This is expressed in the Koschmieder equation:

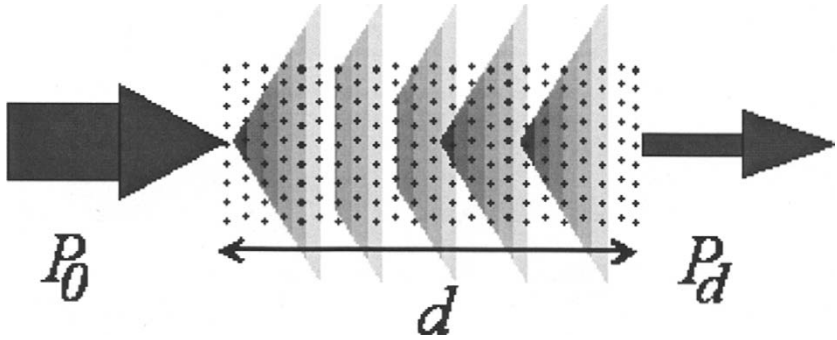


Fig. 15. Transmission according to the Beer-Lambert law.

$$V = \frac{\ln\left(\frac{1}{\tau}\right)}{\gamma_{550\text{ nm}}} = \frac{\ln\left(\frac{1}{0.02}\right)}{\gamma_{550\text{ nm}}} = \frac{3.912}{\gamma_{550\text{ nm}}}. \tag{4}$$

Sometimes a slightly different definition is used in the literature, defining the visibility range at 5% transmission. This leads to a factor of 2.996 instead of 3.912. Reforming this equation allows to calculate the extinction coefficient for 550 nm, at the center of the visible light range.

To adapt visibility data for infrared wavelength transmission, the wavelength dependency of the scattering process has to be taken into account. A first relation based on empirical measurement data was proposed by Kruse [5]:

$$\gamma(\lambda) \cong \beta_a(\lambda) \cong \frac{3.912}{V}\left(\frac{\lambda}{550\text{ nm}}\right)^{-q}. \tag{5}$$

The exponent q in Eq. (5) depends on the visibility distance range. To include also low visibilities in dense fog, the original Kruse equation was modified by Kim [3, 5] to

$$q = \begin{cases} 1.6 & \text{for } V > 50\text{ km} \\ 1.3 & \text{for } 6\text{ km} < V < 50\text{ km} \\ 0.16V + 0.34 & \text{for } 1\text{ km} < V < 6\text{ km} \\ V - 0.5 & \text{for } 0.5\text{ km} < V < 1\text{ km} \\ 0 & \text{for } V < 0.5\text{ km}. \end{cases} \tag{6}$$

In general it is problematic to find a relation which allows the recalculation of visibility data from 550 nm to longer wavelengths. Different types of fog can cause different attenuation for longer wavelengths at the same visibility range, actually depending on the particle size distribution and the particle density according to the Mie scattering model. But visibility data does not include information about the particles, even rain and snow are included. Measurement data indicates that longer wavelengths are less attenuated in haze and light fog, while there is no wavelength dependency in dense fog. New theoretical calculations from Alnaboulsi for advection fog and convection fog indicate that shorter wavelengths are less attenuated at very short visibilities, which is the most critical case for high available FSO.

Hence, even if visibility data should not be over-interpreted, it is a useful information for availability prediction. Probably the best compromise is to assume no

wavelength dependency. Especially for short-range high availability applications this seems to be the best approach, except for 10 μ technology. It is convenient for technical applications to calculate in decibels, so the extinction coefficient can be recalculated to attenuation due to scattering:

$$a_{\mathrm{SCAT}} = \frac{10}{\ln(10)}\gamma(\lambda)d. \tag{7}$$

Equation (7) considers attenuation a_{SCAT} (in dB) over the full atmospheric path distance d, while Eq. (8) considers the specific attenuation $a_{\mathrm{SCAT,SPEC}}$ (in dB/km) for a transmission threshold of 2%:

$$a_{\mathrm{SCAT,SPEC}} = \frac{10}{\ln(10)}\gamma(\lambda) = \frac{10\log\left(\frac{1}{\tau}\right)}{V} \cong \frac{17\,\mathrm{dB}}{V}. \tag{8}$$

3.2. *Reliability and Availability*

To introduce this topic, which is of high relevance for the practical use of FSO systems in communication networks, the meaning of the expressions shall be explained.

- System reliability $R(T)$ is the probability that the system works correctly during the time period T under defined environmental conditions.
- System availability $A(t)$ is the probability that the system works correctly at the time t.

Because the FSO link just offers the physical layer for data transmission, the conditions for correct operation of the system usually are defined by a maximum tolerable bit error rate (BER) for the specified data rate (e.g., BER = 10^{-9} for 100 Mbps). The bit error rate increases, if too much or not enough optical power is received. In this sense, FSO systems usually are very secure by construction, if they can avoid overload and if they simply switch off the connection to the network, when the received optical power gets too low because of atmospheric attenuation, before too many bit errors are produced. This threshold is also referred as the receiver sensitivity limit P_{RS}. Especially short-range systems can be described in this way and behave like a wired connection, offering the specified performance over the full distance. For long distance FSO connections using more collimated and more coherent light, fluctuations in the received power get increasingly important. This makes it difficult to set the right sensitivity limit and could also cause performance degradation by burst errors, which can be avoided by channel coding and error correction techniques.

Related are safety issues, which do not actually refer to data transmission, but in general to conditions for operation. To satisfy laser safety regulations according to IEC/EN 60 825, especially to meet laser class 1 which certifies safety under all possible conditions over long time, or the classes 1M, 2 (visible light) or 3A is important to allow the operation without further restrictions. The main concern of laser safety is the human eye, which focuses light coming through the iris aperture of max. 7 mm diameter to the retina, so light intensity is actually measured. In the same eye safety laser class, more power is allowed for longer wavelengths up to 1400 nm because of transmittance and focus characteristics of the eye. Hence, 1550 nm technology has an advantage against 850 nm here. More power is also allowed for light sources with a higher divergence angle compared to point sources, because only point sources can be

concentrated in a very small spot on the retina. For this reason, light sources like LEDs, optics using diffusers, or multiple transmitter optics have an advantage. In addition to laser safety also the regulations for mounting the systems must be met, too.

The availability of an installed FSO link mainly depends on the power link budget and the local climate conditions, causing increased attenuation over periods of time. The properties of an FSO system excluding the effects of the atmospheric path and distance can be summarized in the system power factor P_{SYS} (in dBm) to

$$\begin{aligned} P_{\mathrm{SYS}} &= P_{\mathrm{TX}} + G_{\mathrm{TX}} + A_{\mathrm{RX}} - \sum(a_{\mathrm{TX}} + a_{\mathrm{RX}}) \\ &= P_{\mathrm{TX}} + 10\log\left(\frac{4\pi}{2\pi(1-\cos a/2)}\right) + 20\log(R_A) - \sum(a_{\mathrm{TX}} + a_{\mathrm{RX}}). \end{aligned} \tag{9}$$

In Eq. (9) the optical output power in decibels over one mill watts (dBm) is represented by P_{TX}, the geometrical transmitter gain in decibels is G_{TX} (depending on the full beam divergence angle α), the geometrical gain at the receiver due to optics aperture radius R_A is A_{RX} and losses in the optics of transmitter and receiver are represented as a_{TX} and a_{RX} (dB). This allows calculating the power received P_{RX} in distance d by

$$\begin{aligned} P_{\mathrm{RX}} &= P_{\mathrm{SYS}} - D_L \\ &= P_{\mathrm{SYS}} - 20\log\left(\frac{2d}{1\ \mathrm{m}}\right). \end{aligned} \tag{10}$$

Equation (10) includes the decibel value D_L of the geometrical distance d in meters and gives the received power P_{RX} without any atmospheric influence or pointing losses.

The specific link margin M_{SPEC} for an installation can be found with the receiver sensitivity limit P_{RS} for proper operation

$$\begin{aligned} M_{\mathrm{SPEC}} &= \frac{1000\ \mathrm{m}}{d} M \\ &= \frac{1000\ \mathrm{m}}{d}[P_{\mathrm{SYS}} - D_L - P_{\mathrm{RS}}]. \end{aligned} \tag{11}$$

The basic condition for system availability A is that the specific link margin exceeds the specific attenuation, which can be expressed by

$$A = \begin{cases} 1 & \text{for } M_{\mathrm{SPEC}} \geq \sum a_{\mathrm{SCAT,SPEC}} \\ 0 & \text{for } M_{\mathrm{SPEC}} < \sum a_{\mathrm{SCAT,SPEC}} \end{cases} \tag{12}$$

at the time t.

The appearance of fog can only be statistically described, usually as the steady-state availability, which is approximated by measurements over long periods of time. As an example, a one-year runway visibility range data set from Graz, Austria, is shown in Fig. 16. In this data, point 3 represents the probability for a visibility of less than 1500 m, equal to a specific attenuation of 11.4 dB/km, which was approximately 5% of one year's time. The probability for visibilities less than 200 m indicated in point 1 was 0.01%, which means that a FSO installation would reach 99.99% average availability during this year if it has more than 85 dB/km specific margin. Such link availabilities and even more are possible if only the link distance does not exaggerate a certain value.

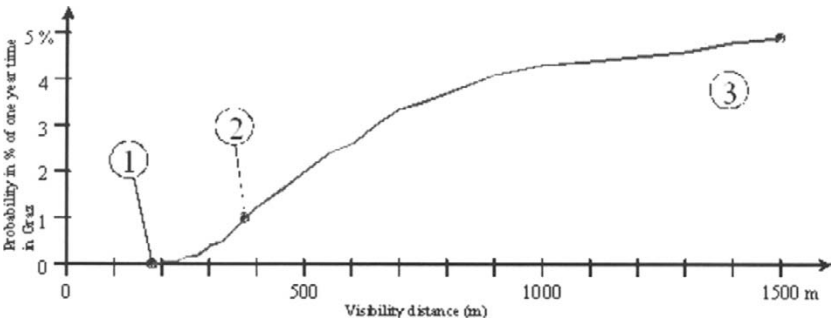


Fig. 16. Full year visibility data profile for the city of Graz, Austria.

Based upon this climate example, it is possible to calculate availability contours for specific FSO systems characterized by their system power factor. Figure 17 shows this for two types of systems. Type one is a cost efficient general purpose type with a comparatively large beam divergence and receiver acceptance angle, allowing quick installation and easy mount. Type three is a sophisticated system with small beam divergence and automatic beam tracking specialized for longer distances. Typical system properties may be between these two cases.

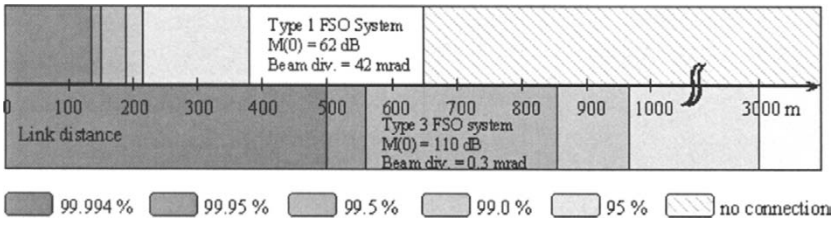


Fig. 17. Calculated availability contours for two FSO systems for the visibility data set of Fig. 16.

A practical measurement with a commercial FSO system was performed in [2,6] over the same period of time as used in Fig. 16. The system used 850 nm wavelength and a beam divergence of 2 mrad for an installed link distance of 2.7 km. Multiple beam technology was implemented to reduce scintillation effects in the optical domain and to increase system reliability. The system power figure was 96 dBm and the specific link margin approximately 7 dB/km. Figure 18 shows the seasonal and diurnal dependencies of the unavailability of the link.

A clear dependence on the season can be seen, as it is typical for Central Europe and North America, for example. The major cause of unavailability was fog, especially convection or radiation fog, which appears at the end of the day and during night, caused by ground cooling due to radiation, and is resolved by sunlight during the day. This is the cause for the daytime dependency.

The performance of the system was very good; at 155 Mbps data rate the monthly averaged bit error rates (BER) reached values between 10^{-11} and 10^{-13}. A one-year follow-up trial [11] of the same optical link combined with 40 GHz technology with approximately 2.5 dB/km specific margin did result in 99.926% reliability for the

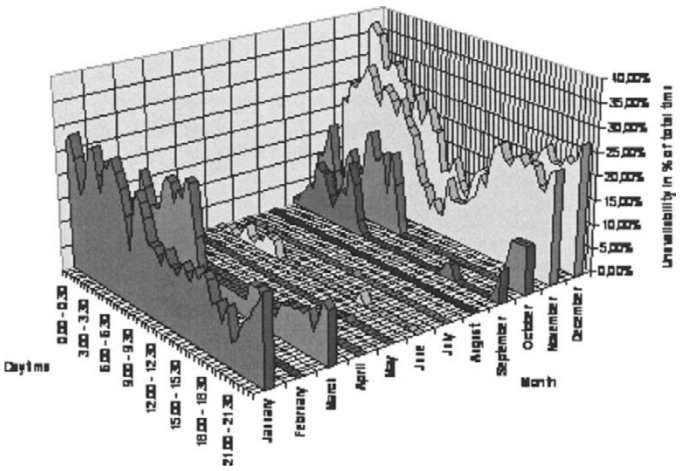


Fig. 18. Seasonal and diurnal dependency of measured unavailability of a 850-nm FSO system at 7 dB/km specific margin in Graz, Austria.

combination of optical and millimeter wave, a so-called hybrid wireless link, while the availability of the millimeter wave link was 97.989% and for the FSO system 96.813% during this year. A short range FSO system with 70 dB/km specific margin did show an availability higher than 99.96% in the winter season during this time.

For practical applications this means that broadband and high available wireless links can either be reached over short distances with optical links only, or over longer distances by the combination of very high frequency and optical links. If the application is not time critical or does not require high availability (like, for example, remote data storage), additional FSO links over several kilometer distances can increase the network throughput dramatically.

Reliability refers to the inner properties of the system in first place, which can be expressed by the mean time-to-failure (MTTF), or the FIT rate actually. 1 FIT means the probability for a device to fail in 10^9 hours of operation. In FSO systems especially the transmitting elements have a limits of lifetime, depending on the conditions under which they are operated. In general, the output power reduces gradually, and finally the device will fail. Major manufacturers qualify their products and provide statistical data for the performance over time. VCSELs and LEDs have a typical lifetime of 10^5 to 10^8 hours of operation. Providers of fiber optic networks state that typically 1–2 of 1000 media converters can fail within a year. Operating the devices below the output power limit and avoiding extreme temperature variations can improve the lifetime a lot.

3.3. Different FSO Techniques for the Last Mile

In sections 1 and 2 we have seen the different types of FSO systems, their characteristics and their advantages. For the last mile FSO systems with light emitting diodes (LED) as source are a well suited solution, because they are cheap and data rates in the range

of tens of Mbps are sufficient. Using LEDs has the advantage that problems with laser- and eye safety are minimized. Low-cost FSO systems using LEDs instead of laser diodes have been developed by [1,9] for data rates of 10 and 100 Mbps.

The developed systems combine available standard components to realize cost-effective solutions. The beam divergence usually covers values from about 8 to 60 mrad, allowing an easy alignment for the user without the need of a telescope. All elements including the optics do not require a precision as high as needed for other approaches, allowing higher tolerances at production, and the use of simple mounts. Owing to large divergence, the requirements for a stable underground are not very high allowing a quick installation. Suitable distances for high availability operation are limited by the wide beam angle and depending on local climate to up to 300 m. Owing to Laser Safety Regulations the use of sources with larger emitting area allows more output power in the same safety class, which improves the link budget.

3.3.1. Small FSO System for 100 m

The developed FSO system for a range of 100 m consists of two main parts, one transmitter (LED) and one receiver (photo-PIN-diode). The LED and the photo-PIN-diode are mounted in the focus of a cheap plastic lens encased by a tube of aluminium alloy. The transmitter and receiver units are mounted on a printed circuit board in a housing of a video camera (Fig. 7A). On a single printed circuit board the electronic part of the FSO system is located. The receiver electronic includes the photo-PIN-diode, the amplifier, and the data-interface. For the transmitter electronic, the data-interface, the driver amplifier and the LED are necessary. The beam divergence of this system is about two degrees.

Currently, solutions are available for 10 and 100 Mbps. The data-interface converts the signal from the electrical domain into fiber or RJ-45 (Ethernet or Fast-Ethernet).

By using VCSELs instead of LEDs and three transmitting units instead of one, the range of the system can be increased. The measured voltage of the receiver unit is shown on an analogue voltage display (integrated at the back of the system). The installation of this FSO system is very easy, because of the beam divergence of two degrees and the above mentioned voltage display.

3.3.2. FSO System for 300 m

The developed FSO system for a range of 300 m is mounted in a plastic housing (Fig. 7B). The FSO unit consists of 8 transmitters (LED) and one receiver (photo-PIN-diode). The LEDs are also mounted in the focus of a cheap plastic lens encased by a tube of Aluminium alloy.

The received light is focused by a large lens to the photo-PIN-diode. The electronic part on the printed circuit board is similar to the small system for the 100-m range. The data-interface converts the signal from the electrical domain into fiber or RJ-45 (Ethernet). By using VCSELs instead of LEDs, the range of the system can be increased up to 800 m.

3.4. FSO network for a Small City and Multimedia Applications

The use of FSO is shown for selected scenarios and some multimedia applications (like video-conferences, live TV-transmissions, etc.) will illustrate the range of applications for FSO last mile networks.

3.4.1. Multimedia Applications

In this section some examples for temporary and nomadic use of FSO are described.

Nomadic Use of FSO at the Folk-Festival "Aufsteirern" A real service demonstration for nomadic use of optical wireless networks took place at the folk-festival "Aufsteirern 2002" in the historical town of Graz. Folk dancers and brass bands of the whole country of Styria came together, playing at at different places and streets. Free Space Optic links were used to transmit live video pictures (with a data rate of 10 Mbps) from two locations ("Landhaushof" and "Tummelplatz") to the central enquiry kiosk located at "Herrengasse". Thereby, the visitors of the festival could inform themselves what was going on at the other locations. In Fig. 19 the setup for this demonstration is shown in the city map of Graz. The transmission from the location Tummelplatz was realized with the use of an optical repeater, because there was no line of sight between Tummelplatz and the enquiry kiosk. For this purpose two systems (for 300 m range) have been interconnected. The transmission from the Landhaushof was implemented with a small FSO (100 m) system.

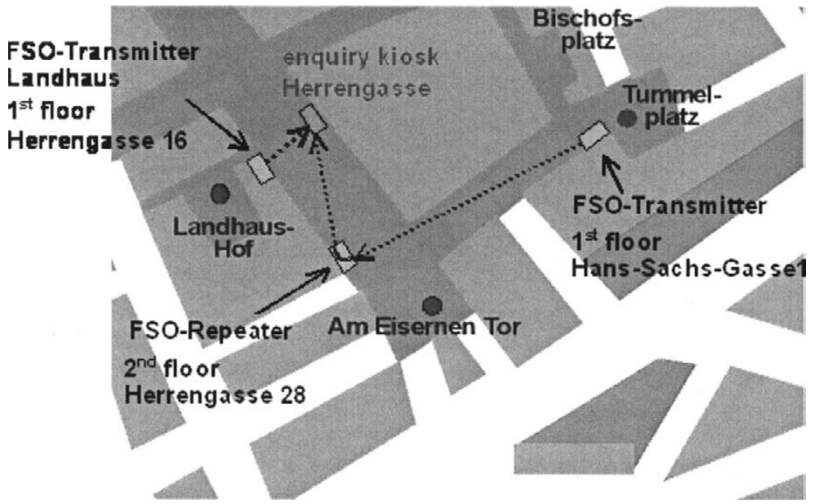


Fig. 19. Locations Landhaushof and Tummelplatz shown in the city map of Graz (Event Aufsteirern).

Hybrid Wireless Networks—High Availability with Combined Optical/Microwave Links The availability of FSO links is limited by weather patterns like fog and heavy snowfall. Microwave-based communication links operating at high frequencies (40–43

GHz) offer comparable data rates and need line-of-sight as well. Link availability for microwave systems is limited by heavy rain. Combining FSO links with microwave links within a hybrid FSO/microwave communication network has the advantage of added redundancy and higher link availability. Measurements over a period of one year show a combined availability of 99,93

Discussing the individual advantages and disadvantages of FSO and radio communication technologies, the idea was born to combine optical and microwave links for introducing redundancy and to achieve higher availability. Therefore a measurement scenario has been set up in Graz, Austria to compare the link availability of a microwave and a FSO link between the same locations influenced by the same weather conditions. The distance of this redundant link is 2.7 km and the two locations are in line of sight. Graz is located in an area where various weather conditions appear and the climatic zone is very problematic for optical and microwave communication systems. Microwave and FSO communication links (Fig. 8) have roughly the same properties regarding offered data rates and flexibility of setup, but operate under different conditions, with their benefits and challenges.

The measurement result is drawn in Fig. 20 for each month. The dotted line indicates the optical link. Obviously the optics is highly influenced by fog and snow fall, which can be seen in December. In contrast the microwave link does not have excessive losses during this period, and the hybrid reaches nearly 100%. A similar situation appears in January and February. During this time snow fall, fog and rain occur, which can be seen in the low availability of both links. Nevertheless, the hybrid has very high availability. This can be seen as proof, that the systems have nearly orthogonal behavior. From March to July both systems have approximately the average availability. The next interesting section is the summer with heavy thunderstorms. In the observed year this happened in August, where both systems do not have excessive bad availability, but the hybrids availability decreases to 99.28%. The reason for this behavior is the loss of the optical link during heavy thunderstorms. A month later nearly the same single availability is obtained, but the hybrid has again a high availability.

The big advantage of the hybrid is the fact, that systems with a single system availability of 97% can be improved to 99.926%, which is identical with a gain enlargement of transmitting power of 12 dB in the microwave link case.

3.4.2. FSO Network for a Small City

Using FSO technology a small network, offering various applications can be setup. The architecture and possible applications are described in the following.

For high reliability, the optimum network architecture is a meshed network, because it combines the advantages of the above described architectures. Different connections are possible, of which two examples are described. In Fig. 21 an expanded version of a ring architecture is shown. The installation of additionally links (the diagonal dashed lines in Fig. 21) increases the reliability and the security against failure. For example, a connection failure in building 2 (in the worst case a total system breakdown of all FSO links of building 2) does not cause a disconnection of the other buildings. An information flow from building 1 to building 3 can be realized in different forms. The information can be sent in the other direction (counter-clockwise) of the ring network (buildings 1, 6, 5, 4, 3), as shown in Fig. 3. A further routing possibility is from

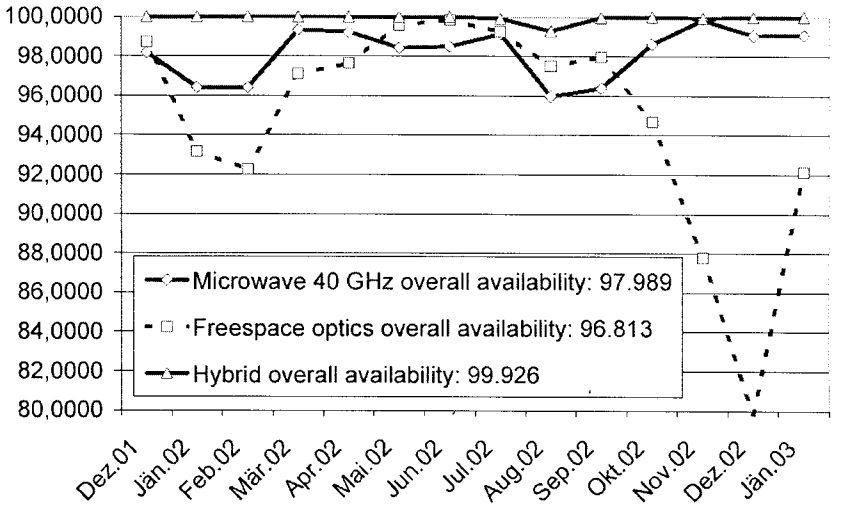


Fig. 20. Availability of the separated links and the hybrid.

building 1 over g 4 to 3 or from building 1 over 6 to 3. Hence, more security against failure can be achieved and a network with high reliability can be realized.

In Fig. 22 a mixture of a ring and a star network is shown. In regard to reliability, shorter distances are advantageous, because the optical multipoint unit is located in the center of the area. In comparison to Fig. 21 the optical multipoint unit is used as additional active optical repeater. If there is more than one optical multipoint unit connected to the backbone a partial security against failure can be achieved.

Both examples of meshed FSO networks have shown advantages in regard to ring or star configurations. Meshed networks combine the benefits of the above described architectures.

In each configuration, the central FSO unit can be connected to satellites, directional radio links, (mobile) telephone networks, or fiber optics. In Fig. 5 a connection to a fiber-based backbone realized with a Point-to-Multipoint architecture is shown. The optical multipoint unit is connected with a switch or router to the backbone network. The clients in the buildings 1, 2, 3 and 4 are linked with their FSO units to the Optical Multipoint. This solution is similar to the FSO network installed at the Department of Communications and Wave Propagation at the TU Graz [1].

A wireless optical access technology allows quick deployment like a RF cellular network, combined with the good performance of an optical technology. The most critical aspect is the weather dependency and the reliability of the links, but as shown in the previous section, this is not a problem if only the point-to-point distance is not too long. FSO technology can offer high data rates very cost effective, and allows implementing the same networking protocols as fiber optic links.

If the network is not just offering transmission capabilities at the physical layer, but is used by a network operator to offer services to the customers, higher layer protocols have to be considered. Nowadays, most applications can already be offered of the Internet, hence all-IP networks seem to be the choice of future access networks.

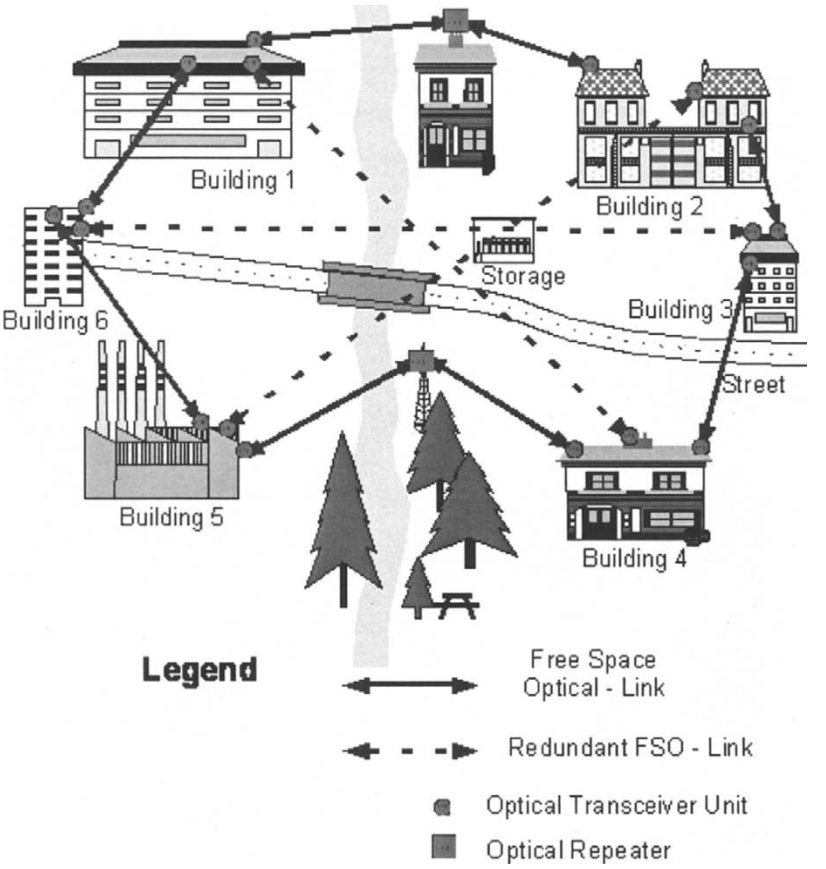


Fig. 21. Optical wireless (meshed architecture, solution A).

In general, a typical access network consists of a central office, the network infrastructure and the customer premise's equipment. At the central office, servers are needed to offer the services to be provided, for example, a mail server, servers for Web-service or external data storage. Access to the Internet requires an Internet router and a firewall. Furthermore, a router/switch is needed in order to connect the incoming links from the users to the backbone network. Additionally, software is needed to manage and control the network, for example software tools for user administration and billing. A typical structure for the central station is shown in Fig. 23.

Regarding the data rates of the network links to the end users, the applications and services have to be considered and will be in the range of 1 to 100 Mbps. If for example, TV-broadcasting services are offered, the link bandwidth has to allow for the simultaneous transmission of up to three TV-program streams, this means three data streams between 2 and 7 Mbps, depending on the image content and compression format.

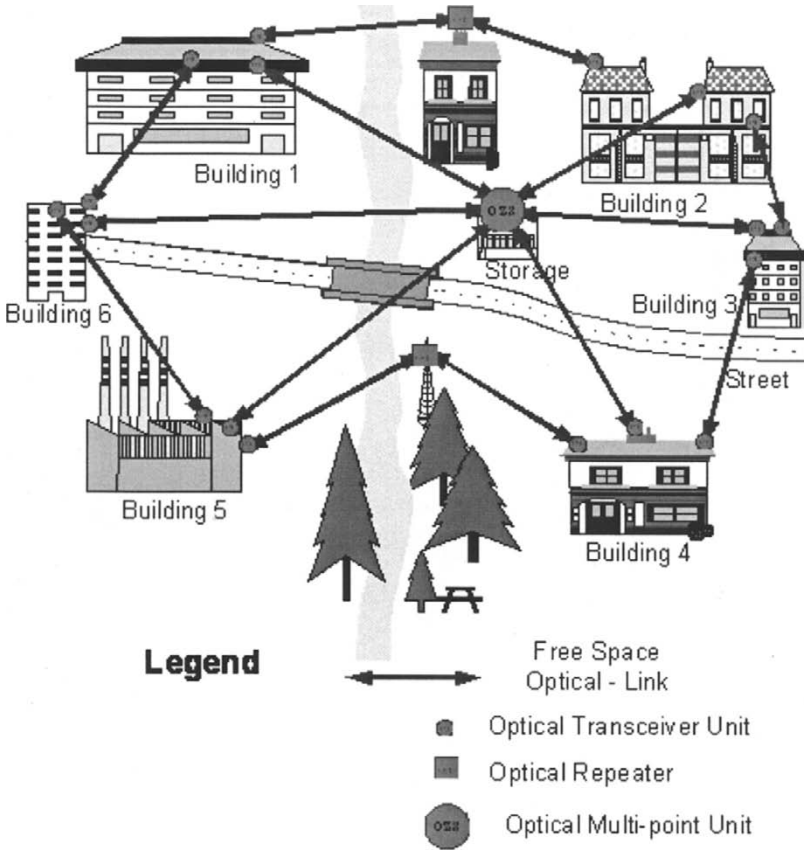


Fig. 22. Optical wireless (meshed architecture, solution B).

For the deployment network the implementation of a broadband technology is essential. To make the network attractive, it should not only satisfy today's needs, but leave room for future expansion. On the other hand, the affordable costs per user are very limited, for TV and 10 Mbps networking services the break-even point which customers were willing to pay was around 70 Euro per month in 2002.

For best performance, the solution of choice for an access network is optical technology, especially if the network provider is an emerging company and does not yet have infrastructure like the Telecom companies. Thereby networks can be built using FSO links only, or fiber and FSO links can be combined in order to satisfy the bandwidth demand of broadband applications like TV broadcasting.

4. Summary

Optical wireless is an excellent broadband solution for connecting end users to the backbone (last mile access). This technology should be seen as supplement to conven-

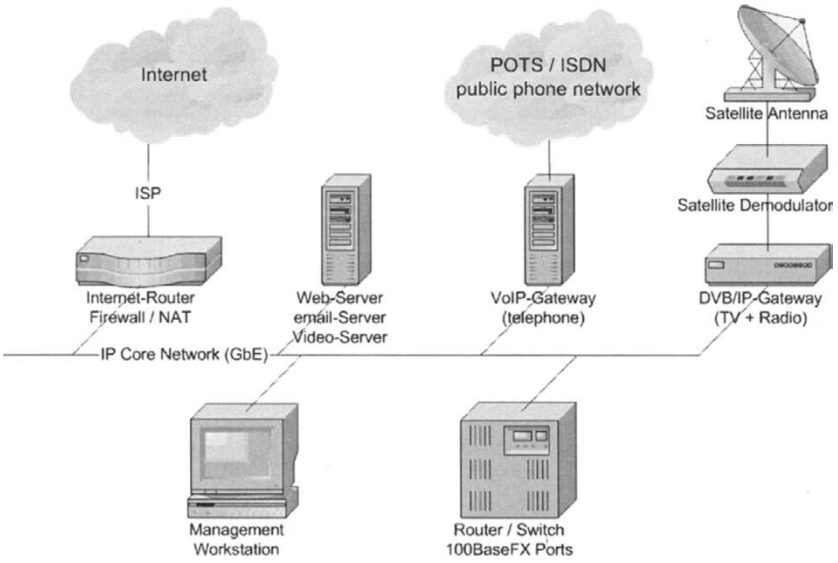


Fig. 23. Central office network structure.

tional radio links and fiber Optics. The use of low-cost FSO systems for short distances makes this technology interesting for private users.

At the moment the main work in this field is to increase reliability and availability. Those two parameters of the FSO link are mainly determined by the local atmospheric conditions. Good reliability and availability can be achieved by using the FSO for short distances, by calculating enough link-budget and by using the optimal network architecture for each FSO application.

The optimal solution for FSO configurations is a meshed architecture. This network architecture combines shorter distances and high reliability, because of the location of the optical multipoint unit in the center of the network. For increasing the reliability and availability it is also necessary to perform field-tests with FSO systems regarding the local atmospheric conditions. Models for propagation and for predicting link availability in different climate zones could improve the installation of this technology.

The combination of FSO and microwave links is also a further possibility for increasing reliability and availability, because terrestrial FSO is most effected by fog, whereas the microwave propagation is mainly influenced by rain. Within parallel studies, wireless hybrid (optical/microwave) links have been evaluated at the Department of Communications and Wave Propagation [1,11]. First results show a reliability of 99.9991% for hybrid systems.

References

1. E. Leitgeb, J. Bregenzer, M. Gebhart, P. Fasser, A. Merdonig, "Free-space optics: Broadband wireless supplement to fiber networks", Proceedings SPIE, Vol. 4975-07 (2003), Jänner 2003, San Jose, USA

2. E. Leitgeb, M. Gebhart, P. Fasser, J. Bregenzer, J. Tanczos, "Impact of atmospheric effects in free-space optics transmission sytems", Proceedings SPIE, Vol. 4976-28 (2003), Jänner 2003, San Jose, USA
3. I. Kim, B. McArthur, E. Korevaar, "Comparison of laser beam prop. at 785 nm and 1550 nm in fog and haze for optical wireless communications", Proceedings SPIE, **4214**, 26–37 (2001)
4. F. David, D. Giggenbach et al., "Preliminary results of a 61 km Ground-to-Ground Optical IM/DD Data Tramsmission Experiment", Proceedings SPIE, Vol. 4635 (2002)
5. H. Sizun, M. Alnaboulsi, O. Veyrunes, O. Bouchet, "Free-space optical communication links, Bibliographical study and experimentation", European workshop on integrated radio-communication systems, Angers, France (5/2002)
6. E. Leitgeb, M. Gebhart, P. Fasser, "Reliability of Free Space Laser Communications—Investigations at the TU Graz", Proceedings of the 8th Annual WCA Technical Symposium, 14–16 Jan. 2002, San Jose, CA, USA
7. P. J. Winzer, W. R. Leeb, "Space-borne optical communications—a challenging reality", Proceedings of the 15th Annual IEEE/LEOS-Meeting Nov. 2002, Glasgow
8. G. S. Mecherle, "Active Pointing for Terrestrial FSO", Proceedings of the 15th Annual IEEE/LEOS-Meeting Nov. 2002, Glasgow
9. E. Leitgeb, J. Bregenzer, P. Fasser, M. Gebhart, "FSO –Extension to fiber-Networks for the 'Last Mile' ", Proceedings of the 15th Annual IEEE/LEOS-Meeting Nov. 2002, Glasgow
10. H. Willebrand, B. S. Ghuman, *FSO*, ISBN 0-672-32248-X (SAMS, 2001)
11. W. Kogler, P. Schrotter, U. Birnbacher, E. Leitgeb, O. Koudelka, "Hybrid Wireless Networks—High Availability with combined Optical/Microwave links", Proceedings of the Conference of Telecommunications and Mobile Computing (TCMC03) April 2002, Graz
12. E. Kube, "Nachrichtenübertragung mit Lichtstrahlen in der Atmosphäre", VEB Verlag Technik, Nachrichtentechnik Heft 19, Seite 201–207 (1969)
13. J. R. Barry, "Wireless Infrared Communications", ISBN 0-7923-9476-3 (Kluwer Academic Publishers, 1994)
14. Webpage ESA (Dec. 22nd 2003) http://www.esa.int/export/esaCP/ESASGBZ84UC_index_0.html
15. Webpage OGS (Dec. 22nd 2003) http://www.iac.es/gabinete/oteide/ogs/2ogs.html and http://www.iac.es/gabinete/iacnoticias/1-2001/47.pdf

DOI: 10.1007/s10297-006-0077-1
Originally published in J. Opt. Fiber. Commun. Rep. 4, 363–405 (2007)

Communication techniques and coding for atmospheric turbulence channels

Xiaoming Zhu[1] and Joseph M. Kahn[2]

[1] QUALCOMM Inc., 675 Campbell Technology Parkway
Suite 200, Campbell, CA 95008
E-mail: xzhu@qualcomm.com
[2] Department of Electrical Engineering, Stanford University
372 Packard Building, 350 Serra Mall
Stanford, CA 94305-9515
E-mail: jmk@ee.stanford.edu

Abstract. In free-space optical communication links, atmospheric turbulence causes fluctuations in both the intensity and the phase of the received light signal, impairing link performance. In this paper, we describe several communication techniques to mitigate turbulence-induced intensity fluctuations, i.e., signal fading. These techniques are applicable in the regime in which the receiver aperture is smaller than the correlation length of the fading, and the observation interval is shorter than the correlation time of the fading. We assume that the receiver has no knowledge of the instantaneous fading state. The techniques we consider are based on the statistical properties of fading, as functions of both temporal and spatial coordinates. Our approaches can be divided into two categories: temporal domain techniques and spatial domain techniques.

In the spatial domain techniques, one must employ at least two receivers to collect the signal light at different positions or from different spatial angles. Spatial diversity reception with multiple receivers can be used to overcome turbulence-induced fading. When it is not possible to place the receivers sufficiently far apart, the fading at different receivers is correlated, reducing the diversity gain. We describe a ML detection technique to reduce the diversity gain penalty caused by such fading correlation.

In the temporal domain techniques, one employs a single receiver. When the receiver knows only the marginal statistics of the fading, a symbol-by-symbol ML detector can be used to optimize performance. When the receiver also knows the temporal correlation of the fading, maximum-likelihood sequence detection (MLSD) can be employed, yielding a further performance improvement, but at the cost of very high complexity. We describe two reduced-complexity implementations of the MLSD, which make use of a single-step Markov chain model for the fading correlation in conjunc-

tion with per-survivor processing. Next, we also investigate the performance of using error-control coding and pilot symbol-assisted detection schemes through atmospheric turbulence channels.

1. Introduction

Recently, free-space optical communication has attracted considerable attention for a variety of applications [1–8]. Because of the complexity associated with phase or frequency modulation, current free-space optical communication systems typically use intensity modulation with direct detection (IM/DD). However, in practice, the performance of free-space optical communication systems can be degraded by many effects, such as fog, obstruction of the line-of-sight path, atmospheric turbulence and the nonideal characteristics of optical transmitters and receivers. In this chapter, we focus on communication techniques and coding schemes to counter the degradation caused by atmospheric turbulence in IM/DD links.

Atmospheric turbulence can degrade the performance of free-space optical links, particularly over ranges of the order of 1 km or longer. Inhomogeneities in the temperature and pressure of the atmosphere lead to variations of the refractive index along the transmission path. These index inhomogeneities can deteriorate the quality of the image formed at the focal plane of an imaging receiver, and can cause fluctuations in both the intensity and the phase of the received signal. These fluctuations can lead to an increase in the link error probability, limiting the performance of communication systems. Aerosol scattering effects caused by rain, snow and fog can also degrade the performance of free-space optical communication systems [9,10], but are not treated in this paper.

Atmospheric turbulence has been studied extensively, and various theoretical models have been proposed to describe turbulence-induced image degradation and intensity fluctuations (i.e., signal fading) [11–17]. Two useful parameters describing turbulence-induced fading are d_0, the correlation length of intensity fluctuations and τ_0, the correlation time of intensity fluctuations. When the receiver aperture D_0 can be made larger than the correlation length d_0, then turbulence-induced fading can be reduced substantially by aperture averaging [17]. However it is not always possible to satisfy the condition $D_0 > d_0$. In this paper, we will introduce communication techniques to mitigate intensity fading in IM/DD links, in the regime where $D_0 < d_0$. At the bit rates of interest in most free-space optical systems, the receiver observation interval T_0 during each bit interval is smaller than the turbulence correlation time τ_0. Throughout this paper, we will assume that $D_0 < d_0$ and $T_0 < \tau_0$.

The techniques we consider in this paper are based on the statistical properties of turbulence-induced signal intensity fading, as functions of both temporal and spatial coordinates. Our approaches can be divided into two categories: temporal domain techniques and spatial domain techniques.

In the spatial domain techniques, one must employ at least two receivers to collect the signal light at different positions or from different spatial angles [6]. To maximize the diversity reception gain, the multiple receivers should be placed as far apart as possible, so that the turbulence-induced fading is uncorrelated at the various receivers. In practice, however, it may not always be possible to place the receivers sufficiently far apart. Hence, in this chapter, making use of the spatial correlation of turbulence-

induced fading, we derive the optimal ML detection scheme for correlated spatial diversity reception [6].

In the temporal domain techniques, one employs a single receiver. If the receiver has knowledge of the marginal fading distribution, but knows neither the temporal fading correlation nor the instantaneous fading state, a maximum-likelihood (ML) symbol-by-symbol detection technique can be used. If the receiver further knows the joint temporal fading distribution, but not the instantaneous fading state, the receiver can use a ML sequence detection (MLSD) technique [7][8]. Error-control coding [18] and pilot-symbol assisted detection schemes [19] can also be applied in free-space communication systems to mitigate turbulence-induced fading.

We note that the statistical properties of turbulence-induced signal intensity fading, as functions of both temporal and spatial coordinates, can also be applied to study space-time codes [19]. In this paper, however, we will not cover the design of space-time codes and coded modulation schemes for free-space optical communications.

The remainder of this paper is organized as follows. In section 2, we review the theories used to model atmospheric turbulence, and use them to derive the spatial and temporal coherence properties of the optical field in weak turbulence channels. We then use the coherence properties obtained to derive the joint probability distribution of the light intensity received on turbulence channels, characterizing the correlation of signal fading over both space and time. In section 3 we first use the marginal distribution of fading to derive a ML symbol-by-symbol detector for systems using on-off keying (OOK). We then use the joint temporal distribution of fading to derive a MLSD for OOK. In section 4, we introduce the spatial-diversity reception scheme, where the instantaneous fading state at each receiver is unknown. With knowledge of the joint spatial distribution of fading, ML detection can help to mitigate the diversity reception gain penalty caused by correlation among multiple receivers. We also numerically demonstrate that in the dual-receiver case, this ML detection scheme has better performance than conventional equal-gain combining (EGC). Then, we describe the temporal domain techniques in section 5. As shown in section 2, the high computational complexity of the MLSD makes it impractical for most applications. Hence, in section 5, we first describe two reduced-complexity implementations of the MLSD, which make use of a single-step Markov chain model for the fading correlation in conjunction with per-survivor processing. We then introduce the pilot-symbol assisted detection (PSAD) scheme. We experimentally demonstrate the effectiveness of these two temporal domain techniques in a 500-m terrestrial link using OOK, where MLSD and PSAD yield signal-to-noise ratio gains of 2.4 and 1.9 dB, respectively. In section 6 we study coding for free-space optical communication systems and derive an error-probability bound for coded OOK free-space optical communication through atmospheric turbulence channels. In section 7, we summarize the conclusions of this paper.

2. Modeling of Optical Communication through Atmospheric Turbulence

In this section, we first review the theories used to model atmospheric turbulence. We then use these to derive the spatial and temporal coherence properties of the optical signal in weak-turbulence channels. Finally, we present the joint spatial and temporal distribution of turbulence-induced fading.

2.1. *Modeling of Atmospheric Turbulence*

Atmospheric turbulence can be physically described by Kolmogorov theory [11–14]. The energy of large eddies is redistributed without loss to eddies of decreasing size until finally dissipated by viscosity. The size of turbulence eddies normally ranges from a few millimeters to a few meters, denoted as the inner scale l_0 and the outer scale L_0, respectively. We can express the refractive index as $n(\mathbf{r}, t) = n_0 + n_1(\mathbf{r}, t)$, where n_0 is the average index and n_1 is the fluctuation component induced by spatial variations of temperature and pressure in the atmosphere. The correlation function of n_1 is defined as

$$\Gamma_{n_1}(\mathbf{r}_1, t_1; \mathbf{r}_2, t_2) = E[n_1(\mathbf{r}_1, t_1) n_1(\mathbf{r}_2, t_2)]. \tag{1}$$

Setting $t_1 = t_2$ in Eq. (1), we obtain $\Gamma_{n_1}(\mathbf{r}_1, \mathbf{r}_2)$, which describes the spatial coherence of the refractive index. To study the spatial coherence of the refractive index, many models have been proposed, which assume exponential, Gaussian or other solvable function forms for $\Gamma_{n_1}(\mathbf{r}_1, \mathbf{r}_2)$. We define the wavenumber spectrum $\Phi_n(k)$ to be the spatial Fourier transform of $\Gamma_{n_1}(\mathbf{r}_1, \mathbf{r}_2)$. A widely used model with good accuracy was proposed by Kolmogorov, which assumes the wavenumber spectrum to be

$$\Phi_n(\mathbf{k}) = 0.033 C_n^2 k^{-11/3}. \tag{2}$$

Here, C_n is the wavenumber spectrum structure parameter, which is altitude-dependent. Hufnagel and Stanley gave a simple model for C_n [13]:

$$C_n^2(z) = K_0 z^{-1/3} \exp(-z/z_0), \tag{3}$$

where K_0 is parameter describing the strength of the turbulence and z_0 is effective height of the turbulent atmosphere. For atmospheric channels near the ground ($z <$ 18.5 m), C_n can vary from 10^{-13} $\mathrm{m}^{-2/3}$ for strong turbulence to 10^{-17} $\mathrm{m}^{-2/3}$ for weak turbulence, with 10^{-15} $\mathrm{m}^{-2/3}$ often quoted as a typical "average" value [12]. Other models and recent measurements of the vertical profile of C_n can be found in [14,15].

2.2. *Spatial and Temporal Coherence of Optical Signals through Turbulence*

To describe spatial coherence of optical waves, the so-called mutual coherence function (MCF) is widely used [12]:

$$\Gamma(P_1, t_1; P_2, t_2) = E[u(P_1, t_1) u^*(P_2, t_2)], \tag{4}$$

where $u(P, t)$ is the complex optical field and $\mathbf{r}$ is the space vector at P. Setting $t_1 = t_2$ in Eq. (4), we obtain the spatial MCF $\Gamma(\mathbf{r}_1, \mathbf{r}_2)$. The Rytov method is frequently used to expand the optical field $u(\mathbf{r})$:

$$u_0(\mathbf{r}) = A(\mathbf{r}) \exp[i\phi(\mathbf{r})] = u_0(\mathbf{r}) \exp(\Phi), \tag{5}$$

where $u_0(\mathbf{r})$ is the field amplitude without air turbulence:

$$u_0(\mathbf{r}) = A_0(\mathbf{r}) \exp[i\phi(\mathbf{r})]. \tag{6}$$

The exponent of the perturbation factor is

$$\Phi_1 = \log[A(\mathbf{r})/A_0(\mathbf{r})] + i[\phi(\mathbf{r}) - \phi_0(\mathbf{r})] = X + iS, \tag{7}$$

where X is the log-amplitude fluctuation and S is the phase fluctuation. We assume X and S be homogeneous, isotropic and independent Gaussian random variables. This assumption is valid for long propagation distances through turbulence.

In order to characterize turbulence-induced fluctuations of the log-amplitude X, we use the log-amplitude covariance function

$$B_X(P_1; P_2) = E[X(P_1)X(P_2)] - E[X(P_1)]E[X(P_2)]. \tag{8}$$

Since the random disturbance is Gaussian-distributed under the assumption of weak turbulence, we can use the Rytov method to derive the normalized log-amplitude covariance function for two positions in a receiving plane perpendicular to the direction of propagation [17]:

$$b_X(d_{12}) = \frac{B_X(P_1, P_2)}{B_X(P_1, P_1)}, \tag{9}$$

where d_{12} is the distance between P_1 and P_2. We define the correlation length of intensity fluctuations, d_0, such that $b_X(d_0) = e^{-2}$. When the propagation path length L satisfies the condition $l_0 < \sqrt{\lambda L} < L_0$, where λ is the wavelength and l_0 and L_0 are inner and outer length scales, respectively, d_0 can be approximated by [13]:

$$d_0 \approx \sqrt{\lambda L}. \tag{10}$$

In most free-space optical communication systems with visible or infrared lasers and with propagation distance of a few hundred meters to a few kilometers, Eq. (10) is valid. We note that aerosol forward scattering can further degrade the coherence of the optical field and thus affect the correlation length. In this paper, however, we focus only on atmospheric turbulence effects.

Atmosphere turbulence also varies with time and leads to intensity fluctuations that are temporally correlated. Modeling the movement of atmospheric eddies is extremely difficult, and a simplified "frozen air" model is normally employed, which assumes that a collection of eddies will remain frozen in relation to one another, while the entire collection is translated along some direction by the wind. Taylor's frozen-in hypothesis can be expressed as [14]:

$$n_1(\mathbf{r}, t) = n_1(\mathbf{r} - \mathbf{v}t, 0), \tag{11}$$

where $\mathbf{v}$ is the velocity of the wind, which has an average $\mathbf{u}$ and a fluctuation $\mathbf{v}_f$. If $\mathbf{v}_f$ is negligible and $\mathbf{u}$ is transverse to the direction of light propagation, then temporal correlation becomes analogous to spatial correlation; in particular, the correlation time is $\tau_0 = d_0/u$. Assuming a narrow beam propagating over a long distance, the refractive index fluctuations along the direction of propagation will be well-averaged, and will be weaker than those along the direction transverse to propagation. Therefore we need only consider the component of the wind velocity vector perpendicular to the propagation direction $u_\perp$. The turbulence correlation time is therefore

$$\tau_0 = \frac{d_0}{u_\perp}. \tag{12}$$

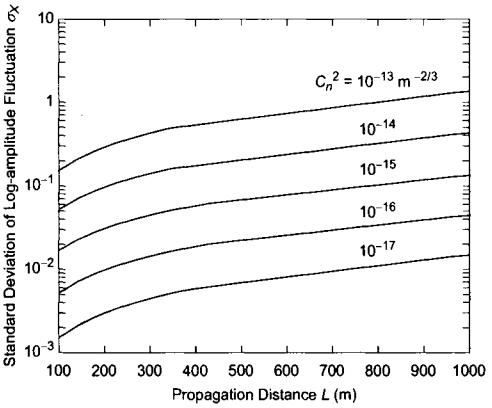


Fig. 1. Standard deviation of the log-amplitude fluctuation versus propagation distance for a plane wave.

2.3. *Probability Distributions of Turbulence-Induced Intensity Fading*

As discussed previously, when the propagation distance is long, log-amplitude fluctuations can become significant. In this section, we will derive the statistical properties of the log-amplitude fluctuations, which we refer to as "intensity fading" or simply "fading". The marginal distribution of fading is derived in section 2.3.1, while the joint spatial and temporal distribution of fading are derived in section 2.3.2.

2.3.1. Marginal Distribution of Fading

In this section, we derive the marginal distribution of fading at a single point in space at a single instant in time. The marginal distribution is used in symbol-by-symbol ML detection, which is discussed in section 3.1.

For propagation distances less than a few kilometers, variations of the log-amplitude are typically much smaller than variations of the phase. Over longer propagation distances, where turbulence becomes more severe, the variation of the log-amplitude can become comparable to that of the phase. Based on the atmosphere turbulence model adopted here and assuming weak turbulence, we can obtain approximate analytic expressions for the covariance of the log-amplitude fluctuation X of plane and spherical waves [14]:

$$\sigma_X^2|_{\mathbf{plane}} = 0.56\left(\frac{2\pi}{\lambda}\right)^{7/6}\int_0^L C_n^2(x)(L-x)^{5/6}dx, \tag{13}$$

$$\sigma_X^2|_{\mathbf{spherical}} = 0.56\left(\frac{2\pi}{\lambda}\right)^{7/6}\int_0^L C_n^2(x)\left(\frac{x}{L}\right)^{5/6}(L-x)^{5/6}dx. \tag{14}$$

Figure 1 shows the standard deviation of the log-amplitude fluctuation σ_X for a plane wave, computed using Eq. (13), as a function of the propagation distance L. In Fig. 1, we again assume a wavelength of 529 nm, and assume $C_n^2(z)$ to be constant.

Fig. 1 shows that for propagation distances of a kilometer, σ_X varies from 10^{-2} to 1 for different values of C_n^2.

Consider the propagation of light through a large number of elements of the atmosphere, each causing an independent, identically distributed phase delay and scattering. By the Central Limit Theorem, the marginal distribution of the log-amplitude is Gaussian:

$$f_X(X) = \frac{1}{(2\pi\sigma_X^2)^{1/2}} \exp\left\{-\frac{(X - E[X]^2}{2\sigma_X^2}\right\}. \tag{15}$$

The light intensity I is related to the log-amplitude X by

$$I = I_0 \exp(2X - 2E[X]), \tag{16}$$

where $E[X]$ is the ensemble average of log-amplitude X.

From Eqs. (15) and (16), the average light intensity is

$$E[I] = E[I_0 \exp(2X - 2E[X])] = I_0 \exp(2\sigma_X^2). \tag{17}$$

Hence, the marginal distribution of light intensity fading induced by turbulence is log-normal:

$$f_I(I) = \frac{1}{2I} \frac{1}{(2\pi\sigma_X^2)^{1/2}} \exp\left\{-\frac{[\ln(I) - \ln(I_0)]^2}{8\sigma_X^2}\right\}. \tag{18}$$

2.3.2. Joint Spatial and Temporal Distributions of Fading

In this section, we derive the joint spatial and temporal distributions of fading. The joint spatial distribution describes the fading at multiple points in space at a single instant of time, and is used in section 4 in evaluating the performance of spatial diversity reception. The joint temporal distribution describes the fading at a single point in space at multiple instants of time. This distribution is the basis for the MLSD introduced in section 3.2.

We assume that the log-amplitude at n receivers is described by a joint Gaussian distribution. From Eq. (9), the auto-covariance matrix of the log-amplitude at n receivers in a plane transverse to the direction of propagation is given by

$$C_X = \begin{bmatrix} \sigma_X^2 & \sigma_X^2 b_X(d_{12}) & \dots & \sigma_X^2 b_X(d_{1n}) \\ \sigma_X^2 b_X(d_{21}) & \sigma_X^2 & \dots & \sigma_X^2 b_X(d_{2n}) \\ \dots & \dots & \dots & \dots \\ \sigma_X^2 b_X(d_{n1}) & \sigma_X^2 b_X(d_{n2}) & \dots & \sigma_X^2 \end{bmatrix}_{n \times n}, \tag{19}$$

where d_{ij} is the distance between points i and j in the receiver plane. Based on "frozen-in" model, and ignoring wind velocity fluctuations, Eq. (19) can also be modified to describe temporal fading correlation by making the substitution

$$d_{ij} = |i - j| T u_\perp, \tag{20}$$

where T is the time interval between observations. In all that follows, we assume a communication system using OOK, in which case, T is the bit interval. We denote the covariance matrix of a string of n bits as

$$C_X^T = \begin{bmatrix} \sigma_X^2 & \sigma_X^2 b_X\left(\frac{T}{\tau_0}d_0\right) & \dots & \sigma_X^2 b_X\left[\frac{(n-1)T}{\tau_0}d_0\right] \\ \sigma_X^2 b_X\left(\frac{T}{\tau_0}d_0\right) & \sigma_X^2 & \dots & \sigma_X^2 b_X\left[\frac{(n-2)T}{\tau_0}d_0\right] \\ \dots & \dots & \dots & \dots \\ \sigma_X^2 b_X\left[\frac{(n-1)T}{\tau_0}d_0\right] & \sigma_X^2 b_X\left[\frac{(n-2)T}{\tau_0}d_0\right] & \dots & \sigma_X^2 \end{bmatrix}_{n\times n}. \tag{21}$$

From the previous discussion about spatial and temporal correlation (see section 3.2), the correlation time τ_0 equals $d_0/u_\perp$.

The joint distribution of intensity for a sequence of transmitted On bits is

$$f_I(I_1, I_2, \dots, I_n) = \frac{1}{2^n \prod_{i=1}^n I_i} \frac{1}{(2\pi)^{n/2}|C_X^T|^{1/2}} \times \exp\left\{-\frac{1}{8}\left[\ln\left(\frac{I_1}{I_0}\right)\cdots\left(\frac{I_n}{I_0}\right)\right](C_X^T)^{-1}\begin{bmatrix}\ln\left(\frac{I_1}{I_0}\right)\\ \dots \\ \ln\left(\frac{I_n}{I_0}\right)\end{bmatrix}\right\}. \tag{22}$$

Similarly, when an On bit is transmitted, the joint distribution of intensity at nreceivers is given by

$$f_I(I_1, I_2, \dots, I_n) = \frac{1}{2^n \prod_{i=1}^n I_i} \frac{1}{(2\pi)^{n/2}|C_X|^{1/2}} \times \exp\left\{-\frac{1}{8}\left[\ln\left(\frac{I_1}{I_0}\right)\cdots\left(\frac{I_n}{I_0}\right)\right]C_X^{-1}\begin{bmatrix}\ln\left(\frac{I_1}{I_0}\right)\\ \dots \\ \ln\left(\frac{I_n}{I_0}\right)\end{bmatrix}\right\}. \tag{23}$$

3. Maximum-Likelihood Detection of On-Off Keying in Turbulence Channels

In most practical IM/DD free-space optical communication systems using OOK, the receiver signal-tonoise ratio (SNR) is limited by shot noise caused by ambient light much stronger than the desired signal and/or by thermal noise in the electronics following the photodetector. In either case, the noise can usually be modeled to high accuracy as additive, white Gaussian noise that is statistically independent of the desired signal. Let T denote the bit interval of the OOK system, and assume that the receiver integrates the received photocurrent for an interval $T_0 \leq T$ during each bit interval. At the end of the integration interval, the resulting electrical signal can be expressed as

$$r_e = \eta(I_s + I_b) + n, \tag{24}$$

where I_s is the received signal light intensity and I_b is the ambient light intensity. Both of these quantities can be assumed to be constant during the integration time. The optical-to-electrical conversion coefficient is given by

$$\eta = \gamma T_0 \frac{e\lambda}{hc}, \tag{25}$$

where γ is the quantum efficiency of the photodetector, e is the electron charge, λ is the signal wavelength, his Plank's constant, c is the speed of light. The additive noise n is

white and Gaussian, and has zero mean and covariance $N/2$, independent of whether the received bit is Off or On.

In this section, we will describe symbol-by-symbol ML detection and compute its error probability in the absence of error-correction coding. We will then describe the MLSD.

3.1. Symbol-by-Symbol Maximum-Likelihood Detection

We assume that the receiver has knowledge of the marginal distribution of the turbulence-induced fading, but has no knowledge of the channel's instantaneous fading. After subtraction of the ambient light bias ηI_b, the signal $r = r_e - \eta I_b$ is described by the following conditional densities when the transmitted bit is Off or On, respectively:

$$P(r|\text{Off}) = \frac{1}{\sqrt{\pi N}} \exp\left(-\frac{r^2}{N}\right), \tag{26}$$

$$\begin{aligned} P(r|\text{On}) &= \int_{-\infty}^{\infty} P(r|\text{On}, X) f_X(X) dX \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N}} f_X(X) \exp\left[-\frac{(r - \eta I_0 e^{2X - 2E[X]})^2}{N}\right] dX, \end{aligned} \tag{27}$$

The optimal maximum a posteriori (MAP) symbol-by-symbol detector decodes the bit $\hat{s}$ as [21,22]:

$$\hat{s} = \underset{s}{\text{argmax}} P(r|s) P(s), \tag{28}$$

where $P(s)$ is the probability that a On bit or Off bit is transmitted. $P(r|s)$ is the conditional distribution that if a bit s(On or Off) is transmitted, a signal level r will be received. If On and Off bits are equally probable, or if their a priori probabilities are unknown, the symbol-by-symbol ML detector decodes the bit $\hat{s}$ as

$$\hat{s} = \underset{s}{\text{argmax}} P(r|s). \tag{29}$$

The likelihood function is

$$\lambda(r) = \frac{P(r|\text{On})}{P(r|\text{Off})} = \int_{-\infty}^{\infty} f_X(X) \exp\left[\frac{(r - \eta I_0 e^{2X - 2E[X]})^2 - r^2}{N}\right] dX. \tag{30}$$

In Fig. 2, we see that the likelihood ratio increases monotonically with r for $0 \leq r \leq 1$, so that ML detection can be implemented by simply thresholding the received signal. In Fig. 3, we plot the optimal ML threshold for $0 \leq r \leq 1$ versus the log-amplitude standard deviation σ_X. We see that as σ_X increases, the ML threshold decreases toward zero, because turbulence-induced fading increases the fluctuation of the On-state signal level, while leaving the fluctuation of the Off-state signal level unchanged. In Fig. 3, we also see that as the additive Gaussian noise covariance $N/2$ increases, the fluctuations of the Off and On state becomes more closely equal, and the threshold increases toward 1/2.

The bit-error ratio (BER) of OOK can be computed as

$$P_b = P(\text{Off}) P(\text{Bit Error}|\text{Off}) + P(\text{On}) P(\text{Bit Error}|\text{On}), \tag{31}$$

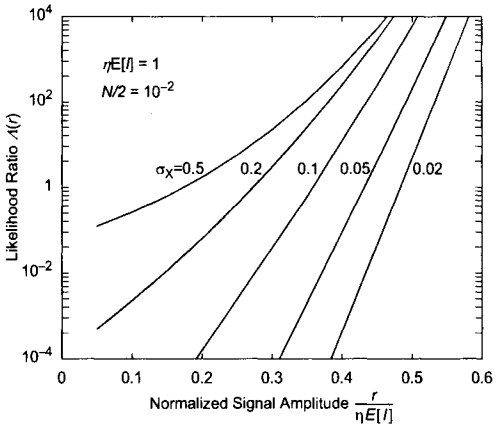


Fig. 2. Likelihood ratio versus normalized received signal amplitude for different values of the log-amplitude standard deviation.

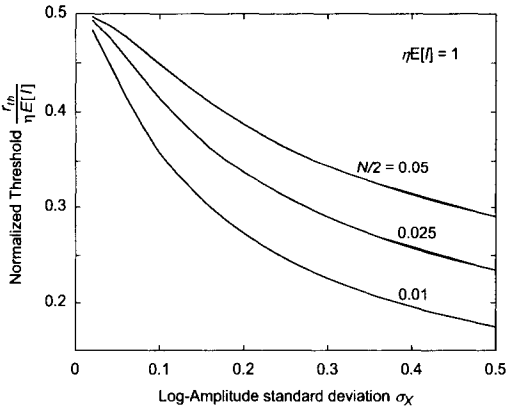


Fig. 3. Normalized threshold for maximum-likelihood symbol-by-symbol detection versus log-amplitude standard deviation for different values of the noise covariance.

where $P(\text{Bit Error}|\text{Off})$ and $P(\text{Bit Error}|\text{On})$ denote the bit-error probabilities when the transmitted bit is Off and On, respectively. Without considering intersymbol interference, which can be ignored when the bit rate is not high and multipath effects are not pronounced, we have

$$P(\text{Bit Error}|\text{Off}) = \int_{\Lambda(r)>1} p(r|\text{Off})dr, \tag{32}$$

$$P(\text{Bit Error}|\text{On}) = \int_{\Lambda(r)<1} p(r|\text{On})dr. \tag{33}$$

3.2. *Maximum-Likelihood Sequence Detection*

The MLSD exploits the temporal correlation of turbulence-induced fading, and is thus expected to outperform the symbol-by-symbol ML detector. For a sequence of ntransmitted bits, the MLSD computes the likelihood ratio of each of the 2^npossible bit sequences $\hat{s} = [s_1 s_2 \ldots s_n]$ and chooses

$$\begin{aligned}\hat{s} &= \underset{\hat{s}}{\operatorname{argmax}}\, P(\mathbf{r}|\mathbf{s}) \\ &= \underset{\hat{s}}{\operatorname{argmax}} \int_{\mathbf{X}} f_{\mathbf{X}}(\mathbf{X}) \exp\left[-\sum_{i=1}^{n} \frac{(r_i - \eta s_i I_0 e^{X_i - 2E[X_i]})^2}{N_i}\right] d\mathbf{X}. \end{aligned} \quad (34)$$

Here, each s_i can take the value Off or On, so that $s \in \{0, 1\}$:

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{X}) = \frac{1}{(2\pi)^{n/2}|C_X^T|^{1/2}} \exp\Bigg\{ &-\tfrac{1}{2}[(X_1 - E[X_1]) \cdots (X_n - E[X_n])] \\ &\times (C_X^T)^{-1} \begin{bmatrix} (X_1 - E[X_1]) \\ \cdots \\ (X_n - E[X_n]) \end{bmatrix} \Bigg\}. \end{aligned} \quad (35)$$

The complexity of MLSD is proportional to $n2^n$, because it requires computing an n-dimensional integral for each of 2^n bit sequences. This complexity is excessive for most applications. In section 5.1, we describe two reduced-complexity implementations of the MLSD, which make use of a single-step Markov chain model for the fading correlation in conjunction with per-survivor processing.

4. Spatial Diversity Reception

Spatial diversity reception, which has been well-studied for application at radio and microwave frequencies, has the potential to mitigate the degradation caused by atmospheric turbulence [6,21–25]. Spatial diversity reception in free-space optical communication has been proposed and studied for both near-earth links [6,23] and interplanetary links [25]. Ibrahim [24] has studied the performance of spatial-diversity optical reception on turbulence channels, assuming that turbulence-induced fading is uncorrelated at each of the optical receivers. In order for this assumption to hold true, the spacing between receivers should exceed the fading correlation length in the plane of the receivers. It may be difficult to satisfy this assumption in practice, for various reasons. Available space may not permit sufficient receiver spacing. In power-limited links, which often employ well-collimated beams, the receiver spacing required for uncorrelated fading may exceed the beam diameter.

4.1. *Maximum-Likelihood Diversity Detection on Turbulence Channels*

In the case of spatial diversity reception with nreceivers, the received signal is described by an n-component vector $\mathbf{r}$ Taking account of correlation between the receivers, Eqs. (26) and (27) are modified as

$$P(\mathbf{r}|\text{Off}) = \exp\left[-\sum_{i=1}^{n}\frac{r_i^2}{N_i}\right]\prod_{1}^{n}\frac{1}{\sqrt{\pi N_i}}, \tag{36}$$

$$P(\mathbf{r}|\text{On}) = \int_{\mathbf{X}} f_{\mathbf{X}}(\mathbf{X})\exp\left[-\sum_{i=1}^{n}\frac{(r_i-\eta I_0 e^{2X_i-2E[X_i]})^2}{N_i}\right]\prod_{1}^{n}\frac{1}{\sqrt{\pi N_i}}dX, \tag{37}$$

where $N_i/2$ is the noise covariance of the ith receiver and

$$f_{\mathbf{X}}(\mathbf{X}) = \frac{1}{(2\pi)^{n/2}|C_X|^{1/2}}\exp\left\{-\frac{1}{2}\left[(X_1-E[X_1])\right.\right.$$
$$\left.\left.\times\cdots(X_n-E[X_n])\right]C_X^{-1}\begin{bmatrix}(X_1-E[X_1])\\ \cdots\\ (X_n-E[X_n])\end{bmatrix}\right\}. \tag{38}$$

Here, C_X is the covariance matrix of the log-amplitudes in the n receivers, as in Eq. (19). The likelihood function is

$$\Lambda(\mathbf{r}) = \frac{P(\mathbf{r}|\text{On})}{P(\mathbf{r}|\text{Off})} = \int_{\mathbf{X}} f_{\mathbf{X}}(\mathbf{X})\exp\left[-\sum_{i=1}^{n}\frac{(r_i-\eta I_0 e^{2X_i-2E[X_i]})^2}{N_i}\right]dX. \tag{39}$$

On the ML detector employs the decision rule $\Lambda(\mathbf{r}) \underset{\text{Off}}{\overset{\text{On}}{\gtrless}} 1$. Since the log-amplitude follows a joint log-normal distribution, calculation of the likelihood function in Eq. (39) involves multi-dimensional integration. We emphasize that this decision rule has been derived under the assumption that the receiving party knows the fading correlation but not the instantaneous fading state.

The BER of the ML receiver is given by

$$P_b = P(\text{Off})P(\text{Bit Error Off}) + P(\text{On})P(\text{Bit Error On}), \tag{40}$$

where $P(\text{Bit Error}|\text{Off})$ and $P(\text{Bit Error}|\text{On})$ denote the bit-error probabilities when the transmitted bit is Off and On, respectively. Without considering intersymbol interference, which can be ignored when the bit rate is not high and multipath effects are not pronounced, we have

$$P(\text{Bit Error Off}) = \int_{\Lambda(\mathbf{r}>1} p(\mathbf{r}|\text{Off})d\mathbf{r} \tag{41}$$

and

$$P(\text{Bit Error On}) = \int_{\Lambda(\mathbf{r}<1} p(\mathbf{r}|\text{On})d\mathbf{r}. \tag{42}$$

To evaluate the optimal ML diversity detection scheme, we compare it with the conventional equal-gain combining (EGC) scheme [21]. In the EGC scheme, we assume that the receiving party has knowledge of the marginal distribution of the fading at each receiver, but has no knowledge of the fading correlation or the instantaneous fading state. For each individual receiver output, we can find an optimum threshold τ_i. The EGC detector then adds together the n receiver outputs with equal gains and compares the sum to the threshold:

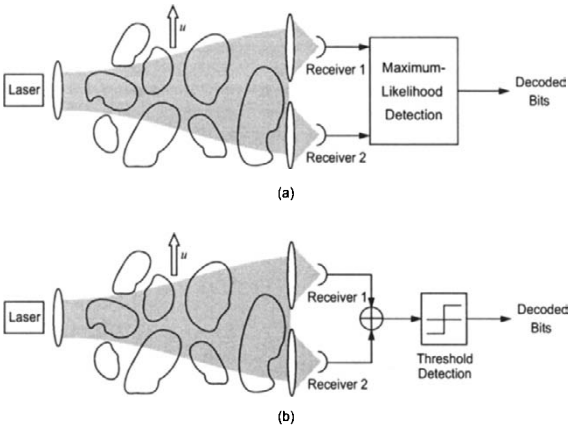


Fig. 4. Dual-branch reception on atmospheric turbulence channels with correlated turbulence-induced fading: (a) maximum-likelihood detection, (b) equal-gain combining with threshold detection.

$$T_{\text{th}} = \sum_{i=1}^{n} \tau_i. \tag{43}$$

The error probability of EGC is

$$P_b = \int_{\mathbf{X}} f_{\mathbf{X}}(\mathbf{X}) P(\text{Bit Error}|\mathbf{X}) d\mathbf{X}, \tag{44}$$

where

$$P(\text{Bit Error}|\mathbf{X}) = P(\text{Off}) Q\left(\frac{T_{\text{th}}}{\sqrt{2N}}\right) + P(\text{On}) Q\left[\frac{\eta I_0 (e^{2X_1 - 2E[X_1]} + e^{2X_2 - 2E[X_2]}) - T_{\text{th}}}{\sqrt{2N}}\right] \tag{45}$$

4.2. *Numerical Simulation for Dual Receivers*

In this section, we present numerical simulations of the performance of spatial-diversity detection for the dual-receiver case, which is illustrated in Fig. 4. As described above, the ML receiver [Fig. 4(a)] has full knowledge of the turbulence-induced fading correlation matrix C_X, while the EGC receiver [Fig. 4(b)] has knowledge only of the marginal fading distributions at the individual receivers. We assume that $E[I_1] = E[I_2] = E[I]$ and $N_1 = N_2 = N$, and define the electrical signal-to-noise ratio SNR $= (\eta E[I])^2/N$. We have used the expressions given in the previous section to numerically compute the average bit-error probabilities.

In Fig. 5(a) we plot the simulation results, assuming $E[X] = 0$ and $\sigma_X = 0.1$, varying the normalized correlation $\rho_d = b_X(d_{12})$ from 0 to 0.9. In Fig. 5(b), we present corresponding results assuming $\sigma_X = 0.25$. Comparing Figs. 5(a) and (b), we

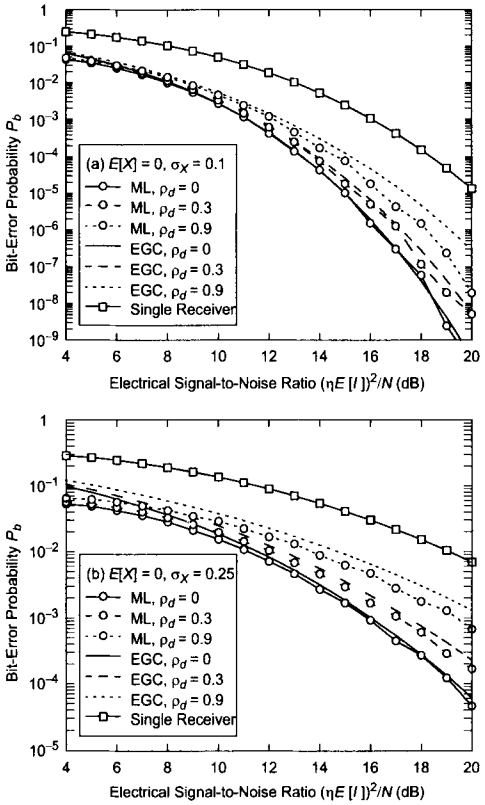


Fig. 5. BER of dual-branch receiver versus average electrical signal-to-noise ratio using maximum-likelihood detection (lines with circles) and equal-gain combining (lines without symbols) for different values of ρ_d, the normalized correlation between the two receivers. The line with squares represents the BER using a single receiver. The turbulence-induced fading has standard deviation $\sigma_x = 0.1$ in (a) and $\sigma_x = 0.25$ in (b). In both (a) and (b), this fading has mean $E[X] = 0$.

see that turbulence-induced fading causes a greater degradation of the BER when the standard deviation σ_X is larger. Diversity reception with two receivers can improve the performance as compared to a single receiver. With two receivers, ML detection achieves better performance than EGC for a given SNR. The advantage of ML over EGC is more pronounced when the correlation ρ_d between the two receivers is high. It is also more pronounced when the SNR is high, so that errors are caused mainly by turbulence-induced fading, as opposed to noise.

4.3. Summary

For IM/DD free-space optical communication in the presence of turbulence-induced fading, spatial diversity with multiple receivers can mitigate this fading, improving

system performance. When the instantaneous fading state is not known, the conventional spatial-diversity detector must employ EGC and simple threshold detection, but the correlation of fading at different receivers can reduce the diversity gain achieved by EGC. In this section, we have derived the ML decision rule under the assumption that the fading correlation properties are known, but the instantaneous fading state is not known. We have performed numerical simulations for the dual-receiver case. Our results show that the ML scheme provides better performance than EGC, particularly when the fading at different receivers is highly correlated and/or when the average SNR is high.

5. Temporal Domain Techniques

5.1. Markov Chain Model in Maximum-Likelihood Sequence Detection through Turbulence

The optimal maximum-likelihood sequence detection (MLSD) requires complicated multidimensional integration, and its computational complexity is exponential in the length of the transmitted bit sequence. To simplify implementation of the MLSD, we propose a single-step Markov chain (SMC) model for the fading temporal correlation, and we use the SMC model to derive an approximate higher-order distribution of bit errors, as well as two reduced-complexity MLSD algorithms based on sub-optimal per-survivor processing (PSP). We use simulations to investigate the accuracy of the SMC model and the effectiveness of the suboptimal MLSD techniques.

5.1.1. Joint Temporal Distribution for Turbulence Induced Fading

In a free-space optical communication system using OOK, we assume a n-bit sequence $\mathbf{s} = [s_1 s_2 \cdots s_n]$ is transmitted. We define the index subset of On-state symbols $S_{\mathrm{On}} = \{n_i \in \{1, 2, \ldots, n\}, s_{n_i} = 1\}_{i=1}^{m}$. We also have the index subset of Off-state symbols $S_{\mathrm{Off}} = \{l_j \in \{1, 2, \ldots, n\}, s_{l_j} = 1\}_{j=1}^{n-m}$. Ignoring intersymbol interference (ISI), the receiver would only receive signal light when the On-state is transmitted. The joint distribution of the signal intensity of On-state symbols is

$$f(I_{n_1}, I_{n_2}, \ldots, I_{n_m}) = \frac{1}{2^m \prod_{i=1}^{m} I_{n_i}} \frac{1}{(2\pi)^{m/2} |C_X^{\mathrm{On}}|^{1/2}}$$

$$\times \exp\left\{ -\frac{1}{8} \left[\ln\left(\frac{I_{n_1}}{I_0}\right) \cdots \ln\left(\frac{I_{n_m}}{I_0}\right) \right] \left(C_X^{\mathrm{On}}\right)^{-1} \begin{bmatrix} \ln\left(\frac{I_{n_1}}{I_0}\right) \\ \cdots \\ \ln\left(\frac{I_{n_m}}{I_0}\right) \end{bmatrix} \right\}, \tag{46}$$

where the ith On-state symbol intensity

$$I_{n_i} = I_0 \exp(2X_{n_i} - 2E[X]) = I_0 \exp(2x_{n_i}). \tag{47}$$

Here, $x_{n_i} = X_{n_i} - E[X]$ can be modeled as a Gaussian random variable with zero mean and covariance σ_X^2. For a string of bits, the covariance matrix of On-state bits is

$$C_X^{\mathrm{On}} =$$

$$\begin{bmatrix} \sigma_X^2 & \sigma_X^2 b_X \left[\frac{|n_1-n_2|T}{\tau_0}\right] d_0 & \dots & \sigma_X^2 b_X \left[\frac{|n_1-n_m|T}{\tau_0}\right] d_0 \\ \sigma_X^2 b_X \left[\frac{|n_2-n_1|T}{\tau_0}\right] d_0 & \sigma_X^2 & \dots & \sigma_X^2 b_X \left[\frac{|n_2-n_m|T}{\tau_0}\right] d_0 \\ \dots & \dots & \dots & \dots \\ \sigma_X^2 b_X \left[\frac{|n_m-n_1|T}{\tau_0}\right] d_0 & \sigma_X^2 b_X \left[\frac{|n_m-2_m|T}{\tau_0}\right] d_0 & \dots & \sigma_X^2 \end{bmatrix}_{n\times n}, \tag{48}$$

where T is the bit interval.

One can also show that the joint distribution of $\mathbf{x} = [x_{n_1}, x_{n_2}, \dots, x_{n_m}$ [7]:

$$f() = \frac{1}{(2\pi)^{m/2}|C_X^{\mathrm{On}}|} \exp\left[-\frac{1}{2}x(C_X^{\mathrm{On}})^{-1}x^T\right]. \tag{49}$$

5.1.2. Single-Step MC Model for Fading Correlations

Here, we consider the SMC model describing the correlation of fading at a sequence of equally spaced times. It is straightforward to extend the treatment to spatial correlation provided that the receivers are equally spaced along a line perpendicular to the direction of propagation. Let x_n denote the log-amplitude at time n, and let x_1^{n-1} denote $[x_1, \dots, x_{n-1}]$. Assuming the turbulence-induced fading is a single-step MC, we have

$$P(x_n|x_1^{n-1}) = P(x_n|x_{n-1}). \tag{50}$$

Define $\rho_T = b_X(d_0T/\tau_0)$. If x_1^n follows a joint Gaussian distribution, the conditional distribution of x_n given x_1 is

$$f(x_n|x_1) = \frac{1}{[2\pi(1-\rho_T^{2n-2})\sigma_X^2]^{1/2}} \exp\left[-\frac{(x_n-\rho_T^{n-1}x_1)^2}{2(1-\rho_T^{2n-2})\sigma_X^2}\right]. \tag{51}$$

From Eq. (51), we see that the exponent of correlation has dropped to 1, implying that the correlation in the SMC model is stronger than in the exact correlation model.

Since the SMC model only takes into consideration the probability distribution of the most adjacent On-bit that contains information on the instantaneous fading state, it is applicable to modeling spatial correlation only in the special case that the multiple receivers are aligned along a line perpendicular to the direction of propagation. However since the number of receivers is typically small, the computational complexity of ML detection with diversity reception will typically be reasonable, as we have shown in [6,7].

5.1.3. Burst Error Distribution for Symbol-by-Symbol Detection

To test the validity of the SMC model, we compute the burst error probability for symbol-by-symbol detection of OOK. Assume that a sequence of On bits is transmitted. The probability of having m consecutive erasures is

$$P_{\mathrm{On}}^m = \mathrm{Prob}\{r_0 > r_{\mathrm{th}}; r_{m+1} > r_{\mathrm{th}}; r_i < r_{\mathrm{th}} \quad \forall i, 1 \le i \le m\}, \tag{52}$$

where r_{th} denotes the receiver decision threshold. Using the chain rule, Eq. (52) can be written as

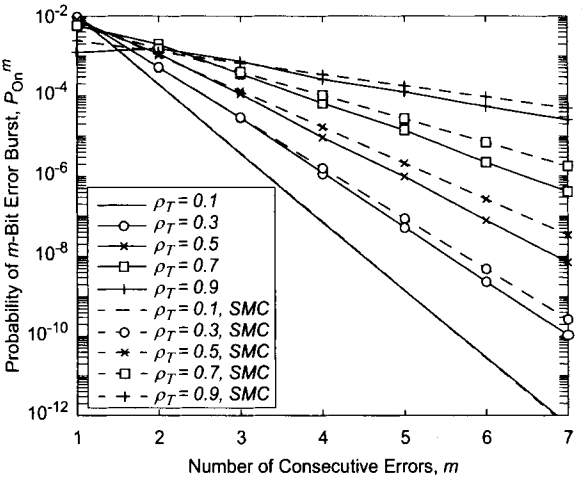


Fig. 6. Distribution of consecutive bit errors and its upper bound derived using the single-step Markov chain model. For simplicity, additive white Gaussian noise is ignored here.

$$P_{\mathrm{On}}^{m} = P_{\mathrm{On}}(r_0 > r_{\mathrm{th}})P_{\mathrm{On}}(r_1 < r_{\mathrm{th}}|r_0 > r_{\mathrm{th}}) \times \cdots P_{\mathrm{On}}(r_{m+1}|r_0 > r_{\mathrm{th}}; \quad r_i < r_{\mathrm{th}}, \quad \forall i, 1 \le i \le n), \tag{53}$$

which is upper-bounded by

$$\begin{aligned} P_{\mathrm{On}}^{m} &= P_{\mathrm{On}}(r_0 > r_{\mathrm{th}}, r_1 < r_{\mathrm{th}})P_{\mathrm{On}}(r_2 < r_{\mathrm{th}}|r_1 < r_{\mathrm{th}} \\ &\times \cdots P_{\mathrm{On}}(r_{m+1} > r_{\mathrm{th}}|r_m < r_{\mathrm{th}} \\ &= P_{\mathrm{On}}(r_0 > r_{\mathrm{th}}, r_1 < r_{\mathrm{th}})[P_{\mathrm{On}}(r_2 < r_{\mathrm{th}}|r_1 < r_{\mathrm{th}}]^{m-1} \\ &\times P_{\mathrm{On}}(r_{m+1} > r_{\mathrm{th}}|r_m < r_{\mathrm{th}}). \end{aligned} \tag{54}$$

To simplify the calculation, let us first ignore AWGN and focus on errors caused by turbulence-induced fading. In the absence of AWGN, we have

$$\begin{aligned} P_{\mathrm{On}}(r_n < r_{\mathrm{th}}|r_1 < r_{\mathrm{th}}) &= \frac{P_{\mathrm{On}}(r_1 < r_{\mathrm{th}}, r_n < r_{\mathrm{th}})}{P_{\mathrm{On}}(r_1 < r_{\mathrm{th}})} \\ &= \frac{\int_{-\infty}^{x_{\mathrm{th}}} f(x_1) \int_{-\infty}^{x_{\mathrm{th}}} f(x_n||x_1) dx_n dx_1}{\int_{-\infty}^{x_{\mathrm{th}}} f(x_1) dx_1}, \end{aligned} \tag{55}$$

where x_{th} can be calculated by r_{th}:

$$r_{\mathrm{th}} = \eta I_0 \exp(2x_{\mathrm{th}}). \tag{56}$$

Note that to compute eqs. (54) and (55), we need only perform two-dimensional integration, independent of the number of bits in the sequence.

In Fig. 6, we present the distribution of consecutive bit errors for several different correlation parameters, comparing the SMC model to the exact correlation model.

Figure 6 has been computed assuming $\eta I_0 = 1$, $E[X] = 0$, $\sigma_X = 1$, and a decision threshold

$$r_{\text{th}} = \frac{1}{10}\exp(2E[X]) = 0.1.$$

As we would expect, the probability of a burst of length $m > 1$ increases as we increase the correlation parameter ρ_T. We see that the SMC model yields a fairly tight upper bound on the exact burst-error probability, validating its use in modeling turbulence-induced fading correlation. Note that as the correlation ρ_T increases, the probability of a single bit error ($m = 1$) decreases, because it becomes more likely that a burst of more than one bit error will occur.

Considering AWGN, (55) is modified to

$$P_{\text{On}}(r < r_{\text{th}}) = \int_{-\infty}^{r_{\text{th}}} P(r|\text{On})dr, \tag{57}$$

$$P_{\text{On}}(r_1 < r_{\text{th}}, r_n < r_{\text{th}}) = \int_{-\infty}^{r_{\text{th}}}\int_{-\infty}^{r_{\text{th}}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1)f(x_n|x_1) \times \frac{1}{\pi\sqrt{N_1 N_n}}\Phi(r_1, r_n, x_1, x_n)dx_n dx_1 dr_n dr_1, \tag{58}$$

where

$$\Phi(r_i, r_j, x_i, x_j) = \exp\left[-\frac{(r_i - \eta I_0 e^{2x_i})^2}{N_i} - \frac{(r_j - \eta I_0 e^{2x_j})^2}{N_j}\right]. \tag{59}$$

Note that a four-dimensional integration is required to compute Eq. (58), independent of the length of the sequence.

In the above, we have focused on missed detection of On bits. When a sequence of Off bits is transmitted, the probability of m consecutive false alarms (falsely detecting a sequence of m On bits) is simply

$$P_{\text{Off}}^m = (1 - P_{\text{Off}})^2 (P_{\text{Off}})^m, \tag{60}$$

where

$$P_{\text{Off}} = \int_{r_{\text{th}}}^{\infty} P(r|\text{Off})dr. \tag{61}$$

Because the correlation of atmosphere turbulence only affects the detection of On bits, a burst of m missed On bits is much more likely than a burst of m false alarms.

It is obvious that

$$P_{\text{On}}(r > r_{\text{th}}) = 1 - P_{\text{On}}(r < r_{\text{th}}), \tag{62}$$

$$P_{\text{On}}(r_1 < r_{\text{th}}, r_n > r_{\text{th}}) = P_{\text{On}}(r_1 < r_{\text{th}}) - P_{\text{On}}(r_1 < r_{\text{th}}, r_n < r_{\text{th}}). \tag{63}$$

We can modify Eqs. (54)–(61) to derive the higher-order BER distribution for a sequence of bits with symbol-bysymbol detection. Since only On-state bits are affected by atmospheric turbulence, based on the SMC model, the On-state BER would depend on the joint probability distribution of the most recent preceding On-state bit and the current On-bit.

5.1.4. Sub-Optimal Per-Survivor Processing for MLSD

The MLSD, as expressed in Eq. (51), is optimal for detecting a sequence of bits that is i.i.d. and uniform on the set {Off, On}. Detecting a sequence of nbits requires a complexity of order $n2^n$, because it requires computing an n-dimensional integral for each of 2^n bit sequences. To reduce the complexity, we consider PSP, which was proposed by Polydoros to extend the Viterbi algorithm to uncertain environments [26]. The key idea is to use the received intensity of recently detected On bits to reduce uncertainty about the state of the turbulence-induced fading. If we consider AWGN and use the SMC model for the fading temporal correlation, then knowing the correlation between two consecutive On bits is sufficient to perform MLSD. Even under the SMC assumption, however, the likelihood function in Eq. (34) cannot be decoupled into a sum of per-branch metrics, which is required for a reduced-complexity implementation. We will modify the metric function to allow us to implement a sub-optimal MLSD using PSP. This sub-optimal MLSD can decode an n-bit sequence with a complexity of order n^2, as compared to the optimal MLSD described in Section 3.2, which has a complexity of order $n2^n$.

Assume a transmitted n-bit sequence $\mathbf{s} = [s_1 s_2 \cdots s_n]$. Define the index subset of On-state symbols $S_{\text{On}} = \{n_i \in \{1, 2, \ldots, n\}, s_{n_i} = 1\}_{i=1}^{m}$ with $n_1 = 1$, i.e., assume the first bit is On. We also have the index subset of Off-state symbols $S_{\text{Off}} = \{l_j \in \{1, 2, \ldots, n\}, s_{l_j} = 0\}_{j=1}^{n-m}$. The exact likelihood function in Eq. (34) is

$$
\begin{aligned}
L(\mathbf{s}) &= \int_{\mathbf{x}} f(\mathbf{x}) \exp\left[-\sum_{i=1}^{n} \frac{(r_i - \eta s_i I_0 e^{2x_i})^2}{N_i}\right] d\mathbf{x} \\
&= \exp\left[-\sum_{\substack{i=1 \\ l_i \in S_{\text{Off}}}}^{n-m} \frac{r_{l_i}^2}{N_{l_i}}\right] \\
&\quad \times \int_{\mathbf{x}_{\text{On}}} f(\mathbf{x}_{\text{On}}) \exp\left[-\sum_{\substack{i=1 \\ n_i \in S_{\text{On}}}}^{m} \frac{(r_{n_i} - \eta I_0 e^{2x_{n_i}})^2}{N_{n_i}}\right] d\mathbf{x}_{\text{On}}, \qquad (64)
\end{aligned}
$$

where $\mathbf{x}_{\text{On}} = \{x_{n_i}, n_i \in S_{\text{On}}\}_{i=1}^{m}$. Based on the SMC model, we have

$$
\begin{aligned}
L(\mathbf{s}) \cong \exp&\left[-\sum_{\substack{i=1 \\ l_i \in S_{\text{Off}}}}^{n-m} \frac{r_{l_i}^2}{N_{l_i}}\right] \int_{\mathbf{x}_{\text{On}}} f(x_{n_1}) f(x_{n_2}|x_{n_1}) \cdots f(x_{n_m}|x_{n_{m-1}}) \\
&\times \exp\left[-\sum_{\substack{i=1 \\ n_i \in S_{\text{On}}}}^{m} \frac{(r_{n_i} - \eta I_0 e^{2x_{n_i}})^2}{N_{n_i}}\right] d\mathbf{x}_{\text{On}}. \qquad (65)
\end{aligned}
$$

In order to decouple Eq. (65), we modify the weighted integration of Eq. (65) by

$$\int_{\mathbf{x}_{\mathrm{On}}} f(x_{n_1}) \exp\left[-\frac{(r_1 - \eta I_0 e^{2x_1})^2}{N_1}\right] \times \prod_{\substack{i=1 \\ n_i \in S_{\mathrm{On}}}}^{m-1} \frac{f(x_{n_i}, x_{n_{i+1}}) \Phi(r_{n_i}, r_{n_{i+1}}, x_{n_i}, x_{n_{i+1}})}{f(x_{n_i}) \exp\left[-\frac{(r_{n_i} - \eta I_0 e^{2x_{n_i}})^2}{N_{n_i}}\right]} d\mathbf{x}_{\mathrm{On}}. \tag{66}$$

We modify Eq. (66) by decoupling the integration as

$$\frac{\left\{\int_{\mathbf{x}_{\mathrm{On}}} f(x_{n_1}) \exp\left[-\frac{(r_1 - \eta I_0 e^{2x_1})^2}{N_1}\right] dx_{n_i}\right\} \prod_{\substack{i=1 \\ n_i \in S_{\mathrm{On}}}}^{m-1} \int f(x_{n_i}, x_{n_{i+1}}) \Phi(r_{n_i}, r_{n_{i+1}}, x_{n_i}, x_{n_{i+1}}) dx_{n_i} dx_{n_{i+1}}}{\prod_{\substack{i=1 \\ n_i \in S_{\mathrm{On}}}}^{m-1} \int_{-\infty}^{\infty} f(x_{n_i}) \exp\left[-\frac{(r_{n_i} - \eta I_0 e^{2x_{n_i}})^2}{N_{n_i}}\right] dx_{n_i}} \tag{67}$$

since the first term in the denominator is identical in the likelihood function for all codewords. We can write the modified likelihood function of $\mathbf{s}$ as

$$L(\mathbf{s}) = \frac{\prod_{\substack{i=1 \\ n_i \in S_{\mathrm{On}}}}^{m-1} \int f(x_{n_i}, x_{n_{i+1}}) \Phi(r_{n_i}, r_{n_{i+1}}, x_{n_i}, x_{n_{i+1}}) dx_{n_i} dx_{n_{i+1}}}{\prod_{\substack{i=1 \\ n_i \in S_{\mathrm{On}}}}^{m-1} \int_{-\infty}^{\infty} f(x_{n_i}) \exp\left[-\frac{(r_{n_i} - \eta I_0 e^{2x_{n_i}})^2}{N_{n_i}}\right] dx_{n_i}} \times \exp\left[-\sum_{\substack{i=1 \\ l_i \in S_{\mathrm{Off}}}}^{n-m} \frac{r_{l_i}^2}{N_{l_i}}\right], \tag{68}$$

From Eq. (68), we can define the metric function of the kth branch ($k > 1$):

$$BM_k(\mathrm{On}) = \frac{\int f(x_{n_i}, x_{n_{i-1}}) \Phi(r_{n_i}, r_{n_{i-1}}, x_{n_{i-1}}, x_{n_i}) dx_{n_i} dx_{n_{i-1}}}{\int f(x_{n_i}) \exp\left[-\frac{(r_{n_i} - \eta I_0 e^{2x_{n_i}})^2}{N_{n_i}}\right] dx_{n_i}}, \tag{69}$$

$$BM_k(\mathrm{Off}) = \exp\left(-\frac{r_k^2}{N_k}\right), \tag{70}$$

where $k = n_i$ and n_{i-1} denotes the position of the most recent On-state bit.

In terms of the branch metrics, the MLSD can be expressed as

$$\mathbf{s} = \underset{\mathbf{s}}{\mathrm{argmax}} \prod_{\substack{i=2 \\ n_i \in S_{\mathrm{On}}}}^{m} BM_{n_i}(\mathrm{On}) \prod_{\substack{i=1 \\ l_i \in S_{\mathrm{Off}}}}^{n-m} BM_{l_i}(\mathrm{Off}). \tag{71}$$

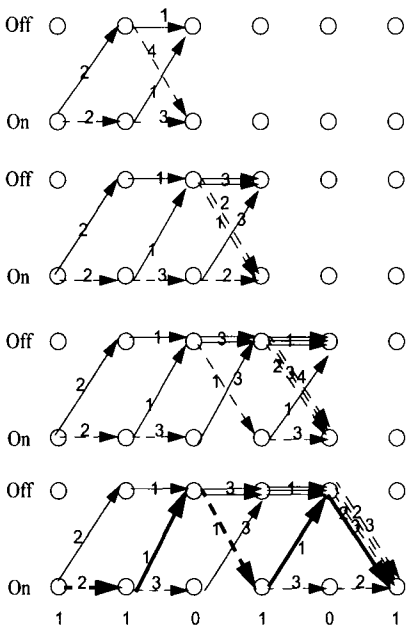


Fig. 7. Viterbi algorithm for maximum-likelihood sequence detection with turbulence-induced fading. Solid lines denote Off bits and dashed lines denote On bits. The branch metric is marked on each branch. The decoded bit sequence is indicated at the bottom.

Since the MLSD considers path metrics that are the product of branch metrics, it can be implemented using the Viterbi algorithm, with a complexity of the order of $n^2/2$. Note that computation of each branch metric of the form (69) requires only a two-dimensional integration, independent of n. In Eq. (69), we see that computing the branch metric for an On bit requires information obtained during the most recently transmitted On bit, so we can only choose a survivor path when the previous bit is known to be On; otherwise, we must keep track of the amplitude of the most recently received On bit, and must also keep track of all survivor paths whose last bit is Off. An example of the asymmetric PSP is shown in Fig. 7. The number on each branch is the branch metric computed using Eqs. (69) and (70). The first and sixth bits correspond to the On state. We see that in this asymmetric PSP, we can only eliminate non-survivor paths when the most recent bit corresponds to the On state. In order to reduce the complexity of this algorithm, we can add an On bit at the beginning and ending of each n-bit sequence, as we have done in this example. We can simply employ the starting On bit of the next sequence as the ending On bit of the previous sequence. The complexity of such an algorithm would be of order $n^2/2$, and only two-dimensional is required. However, we need extra memory to keep track of the survivor path information. Also, some bit overhead is required for implementation of this algorithm.

The algorithm described above still requires a large computational load to perform the two-dimensional integration. To reduce this complexity, we can estimate x_n using r_{n-1}. Since

$$\mathbf{s} = \underset{\mathbf{s}}{\operatorname{argmax}} \int_{\mathbf{x}} f(\mathbf{X}) \exp\left[-\sum_{i=1}^{n} \frac{(r_i - \eta s_i I_0 e^{2x_i})^2}{N_i}\right] d\mathbf{x}$$

$$= \underset{\mathbf{s}}{\operatorname{argmax}} \int_{\mathbf{x}} f(x_1)\phi(s_1, r_1, x_1) \prod_{i=2}^{n} [f(x_i|x_{i-1})\phi(s_i, r_i, x_i)]\, d\mathbf{x}, \quad (72)$$

$$\phi(s_i, r_i, x_i) = \exp\left[-\frac{(r_i - \eta s_i I_0 e^{2x_i})^2}{N_i}\right]. \quad (73)$$

Replacing $f(x_i|x_{i-1})$ with $f(x_i r_{i-1})$, we can define the metric function for branch i:

$$BM_i(\text{On}) = \begin{cases} \int_{x_1} f(x_1)\phi(s_1 = 1, r_1, x_1)dx_1, & i = 1 \\ \int_{x_i} f(x_i|r_{i-1})\phi(s_i = 1, r_i, x_i)dx_i, & i > 1 \end{cases} \quad (74)$$

$$BM_i(\text{Off}) = \int_{x_i} f(x_i|r_{i-1})\phi(s_i = 0, r_i, x_i)dx_i. \quad (75)$$

In terms of the branch metrics, the sub-optimal MLSD can be expressed similar to Eq. (74) with branch mteric functions in Eqs. (74) and (75).

To demonstrate the effectiveness of the sub-optimal PSP algorithm, we present the simulation results in Fig. 8. In this simulation, we assume there is at most one error between two correctly decoded On bits, and we assume $\sigma_X = 0.25$, $I_0 = 1$. In Fig. 8(a), we assume $E[X] = 0$ and $\sigma_X = 0.1$. We plot the BER versus average electrical SNR, given by SNR $= \eta E[I])^2/N$. We consider the two choices of branch metric discussed above; Method 1 uses Eqs. (69) and (70), while Method 2 uses Eqs. (74) and (75). The temporal correlation coefficient ρ_T is chosen to be 0.15 and 0.95. In Fig. 8, we see that both Methods 1 and 2 can achieve much better bit-error performance than the symbol-by-symbol decoding scheme. Method 2 is subject to a penalty of a few dB compared to Method 1, but avoids the two-dimensional integration required by Method 1.

From the discussions above, we see the SMC model can help to greatly simplify the implementation of MLSD with the sub-optimal PSP algorithm, leading to a significant improvement in bit-error performance.

5.2. *Pilot-Symbol Assisted Detection for Correlated Turbulent Free-Space Optical Channels*

Pilot-symbol assisted detection (PSAD) has been employed in applications at radio and microwave frequencies to mitigate the effects of channel fading in situations where channel side information (SI), i.e., the instantaneous fading state, is not known to the receiver [27,28]. In PSAD, the transmitter periodically inserts known symbols, which comprises a reference for receiver to derive the correlated channel fading. PSAD can also be incorporated with maximum likelihood sequence detection (MLSD) to truncate the survivor path of information bit string [8].

In this section, we first introduce the pilot-symbol assisted maximum likelihood (PSA-ML) detection as well as another simpler PS-assisted detection scheme with variable threshold (PSA-VT). We then present the numerical simulations for these detection techniques and compare them with the normal ML symbol-by-symbol detection.

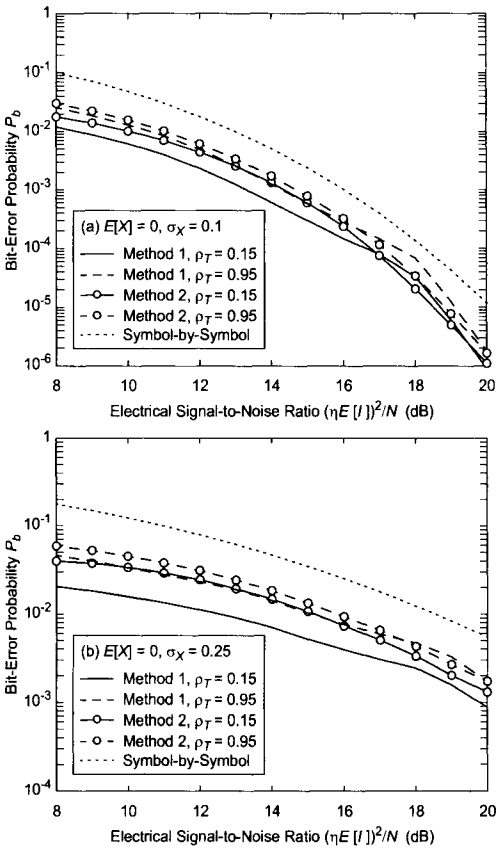


Fig. 8. BER of different decoding schemes versus average electrical signal-to-noise ratio with turbulence-induced fading. The dotted line represents the BER using a symbol-by-symbol decoding scheme. The solid lines consider a PSP algorithm based on the SMC model using branch metric functions (69) and (70), while the dashed lines represent a similar algorithm using Eqs. (74) and (75).

5.2.1. Pilot-Symbol Assisted Maximum-Likelihood Detection

In PSAD, we periodically insert On-state symbols in the information bit string. The composite symbols are transmitted in the usual way over the channel. The resulting frame structure is shown in Fig. 9. At the receiver, we can decode the information bit by considering the joint distribution of the turbulence induced fading at the information bit and the adjacent pilot symbols. Although PSAD will lead to delay in the receiver because of the need to store the whole frame before decoding, the pilot symbols (PS) provide the receiver with information on the state of turbulence-induced fading, and can thus help mitigate the effects of fading. In terms of the branch metrics, the sub-optimal MLSD can be expressed similar to Eq. (71) with branch metric functions in Eqs. (74) and (75).

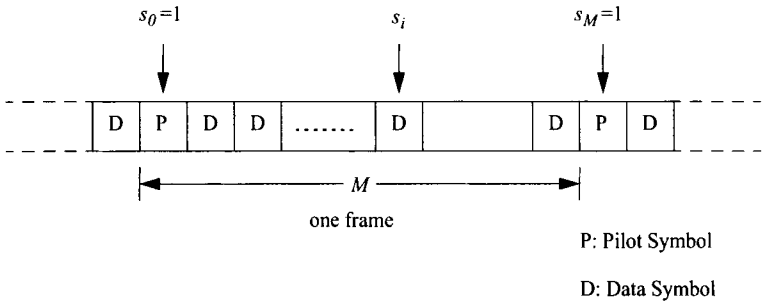


Fig. 9. Pilot-symbol assisted detection frame structure.

In Fig. 9, the composite frame is of length M, the first symbol of each frame is the On-state PS followed by $M-1$ information symbols. In detecting the information bits in one frame, we can refer to the PS in the current frame and in the next frame. Assume r_i is the received photo-current signal of the ith information bit in the frame, r_0 and r_M are the received signal of the PS of the current frame and the next frame. The joint probability distribution of $\mathbf{r} = [r_0, r_i, r_M]$ conditioned on the ith $(1 \leq i \leq M-1)$ information bit s_i are

$$p(\mathbf{r}|s_i = 0) = \frac{1}{(\pi N)^{3/2}} \exp\left[-\frac{r_i^2}{N}\right] \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\mathbf{X}}(X_0, X_M) \times \exp\left[-\sum_{j=0,M} \frac{(r_j - \eta I_0 e^{2X_j - 2E[X_j]})^2}{N}\right] dX_0 dX_M, \quad (76)$$

$$p(\mathbf{r}|s_i = 1) = \frac{1}{(\pi N)^{3/2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\mathbf{X}}(X_0, X_i, X_M) \times \exp\left[-\sum_{j=0,i,M} \frac{(r_j - \eta I_0 e^{2X_j - 2E[X_j]})^2}{N}\right] dX_0 dX_i dX_M. \quad (77)$$

The PDFs of log-amplitude $f(X_0, X_M)$ and $f(X_0, X_i, X_M)$ are shown as in Eq. (35). The likelihood ratio is

$$\Lambda(\mathbf{r}) = \frac{p(\mathbf{r}|s_i=1)}{p(\mathbf{r}|s_i=0)} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\mathbf{X}}(X_0,X_i,X_M) \exp\left[-\sum_{j=0,i,M} \frac{(r_j - \eta I_0 e^{2X_j - 2E[X_j]})^2}{N}\right] dX_0 dX_i dX_M}{\exp\left[-\frac{r_i^2}{N}\right] \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\mathbf{X}}(X_0,X_M) \exp\left[-\sum_{j=0,M} \frac{(r_j - \eta I_0 e^{2X_j - 2E[X_j]})^2}{N}\right] dX_0 dX_M} \quad (78)$$

The ML detector employs the decision rule $\Lambda(\mathbf{r}) \underset{0}{\overset{1}{\gtrless}} 1$. We emphasize that this decision rule has been derived under the assumption that the receiver knows the fading correlation but not the instantaneous fading state.

The BER of the ith information bit in the M-bit frame is given by

$$P_i = P(s_i = 0)P(\text{BitError}|s_i = 0) + P(s_i = 1)P(\text{BitError}|s_i = 1), \quad (79)$$

where $P(\text{Bit Error } s_i = 0)$ and $P(\text{Bit Error } s_i = 1)$ denote the bit-error probabilities when the transmitted ith information bit is 0 and 1, respectively. Without considering intersymbol interference, which can be ignored when the bit rate is not high and multipath effects are not pronounced, we have

$$P(\text{BitError}|s_i = 0) = \int_{\Lambda(\mathbf{r}>1} p(\mathbf{r}|s_i = 0)d\mathbf{r} \quad (80)$$

and

$$P(\text{BitError}|s_i = 1) = \int_{\Lambda(\mathbf{r}<1} p(\mathbf{r}|s_i = 1)d\mathbf{r}. \quad (81)$$

The average BER of a frame is

$$P_b = \frac{1}{M-1}\sum_{i=1}^{M-1} P_i. \quad (82)$$

5.2.2. Pilot-Symbol Assisted Symbol-by-Symbol Detection with Variable Threshold

The PSA-ML detection scheme described above employs the joint PDF of turbulence-induced fading. This scheme requires very lengthy computations to perform multi-dimensional integration. Here, we propose a simpler PS-assisted detection scheme with variable threshold set with the aid of received pilot symbols. We refer to this scheme as PS-assisted variable-threshold (PSA-VT) detection.

Since $[X_0, X_i, X_M]$ is jointly Gaussian-distributed, $E[X_i|X_0, X_M]$ is an affine function of the log-amplitude at two adjacent PS positions, X_0 and X_M:

$$E[X_i|X_0, X_M] = a_i^0 X_0 + a_i^M X_M, \quad (83)$$

$$\begin{bmatrix} a_i^0 \\ a_i^M \end{bmatrix} = \begin{bmatrix} 1 & b_X\left(\frac{MT}{\tau_0}d_0\right) \\ b_X\left(\frac{MT}{\tau_0}d_0\right) & 1 \end{bmatrix}^{-1} \begin{bmatrix} b_X\left(\frac{iT}{\tau_0}d_0\right) \\ b_X\left[\frac{(M-i)T}{\tau_0}d_0\right] \end{bmatrix}. \quad (84)$$

We simply assume that the received PS signals are noise free. Therefore we can set the variable threshold for the ith $(1 \le i \le M-1)$ information bit in the frame to be

$$\tau_i = \frac{(r_0)^{a_i^0}(r_M)^{a_i^M}}{2}. \quad (85)$$

Using PSA-VT, the BER of the ith information bit in an M-bit frame is given by

$$P_i = P(s_i = 0)P(\text{BitError}|s_i = 0) + P(s_i = 1)P(\text{BitError}|s_i = 1), \quad (86)$$

$$P(\text{BitError}|s_i = 0) = \int_{\mathbf{r}, r_i > \tau_i} p(\mathbf{r}|s_i = 0)d\mathbf{r}, \quad (87)$$

$$P(\text{BitError}|s_i = 1) = \int_{\mathbf{r}, r_i < \tau_i} p(\mathbf{r}|s_i = 1)d\mathbf{r}. \quad (88)$$

The expression for the average BER is the same as Eq. (82).

5.2.3. Numerical Simulation

In this section, we present numerical simulations of the detection schemes we proposed in sections 5.2.1 and 5.2.2. We assume the turbulence-induced log-amplitude fluctuation has zero mean, covariance $\sigma_X = 0.1$ and $T/\tau_0 = 0.001$. The BER of ML symbol-by-symbol detection versus the average electrical signal-to-noise ratio $\mathrm{SNR} = (\eta I_0)^2/N$ is shown in Fig. 10 by a dotted line. Recall that the ML symbol-bysymbol detector has knowledge only of the marginal distribution of fading, but not the instantaneous fading. If the ML symbol-by-symbol detector had knowledge of the instantaneous fading, it could achieve better performance; this idealized case is plotted in Fig. 10 using a dash-dot line. SI on the instantaneous fading state is typically not available to the receiver. If we assume the receiver has full knowledge of the joint temporal distribution of the correlated fading, the PS-assisted detection schemes discussed in Sections 5.2.1 and 5.2.2 can be applied. In our simulation, we use the PSAD frame structure shown in Fig. 9. In Fig. 10, we plot the BER of the ith $(i = M/2)$ symbol in an M-bit frame with PSA-ML detection and PS-assisted variable threshold (PSA-VT) detection, respectively. In Fig. 10(a), we plot the BER versus the average electrical SNR of each channel bit and in Fig. 10(b), we plot the BER versus the average electrical SNR of each information bit. We assume the frame size M to be 10, 100, and 1000, respectively. In Fig. 11(a) and (b), we plot the average BER of an M-bit frame with PSA-VT detection versus the average electrical SNR of each channel bit and of each information bit respectively, varying the frame size M to be 10, 100, and 1000.

In Figs. 10 and 11, we observe that when the receiver has knowledge of the temporal correlation of fading, but does not have SI on the instantaneous fading state, PS-assisted detection techniques can give better BER performance than ML symbol-by-symbol detection. However with larger frame sizes, the correlation between the PS and information symbols will become weaker. This will make the PS-assisted detection less effective. In our simulation, we observe that in order to ensure that PS-assisted detection achieves a gain of at least 0.5 dB over ML symbol-by-symbol detection, we require $M < d_T/T$.

In peak power-limited systems, for a fixed transmission bit rate over the channel, using smaller M will improve BER performance, as shown in Figs. 10(a) and 11(a). Another benefit of choosing smaller M is that it will reduce the delay incurred during detection. Choosing smaller M, however, does entail a loss of information throughput.

In average power-limited systems, the inserted PS consume additional power and decrease the SNR of bits transmitted on the channel. As we can see in Figs. 10(b) and 11(b), the BER performance versus average SNR per information bit is very close for $M = 10$ and $M = 100$. If we choose M even smaller, however, the penalty due to the additional power required to transmit the PS will surpass the PSAD gain. Therefore, in average power-limited systems, we should choose M to be too small.

We also note that PSA-ML can achieve better BER performance than PSA-VT, as seen in Fig. 10. PSA-VT does not require multi-dimensional integration and is much faster and simpler to implement than PSA-ML detection, however.

5.3. Experimental Demonstration on Temporal Domain Techniques

We have performed transmission experiments to demonstrate the effectiveness of the MLSD and PSAD techniques in mitigating scintillation noise [29]. Because our proof-

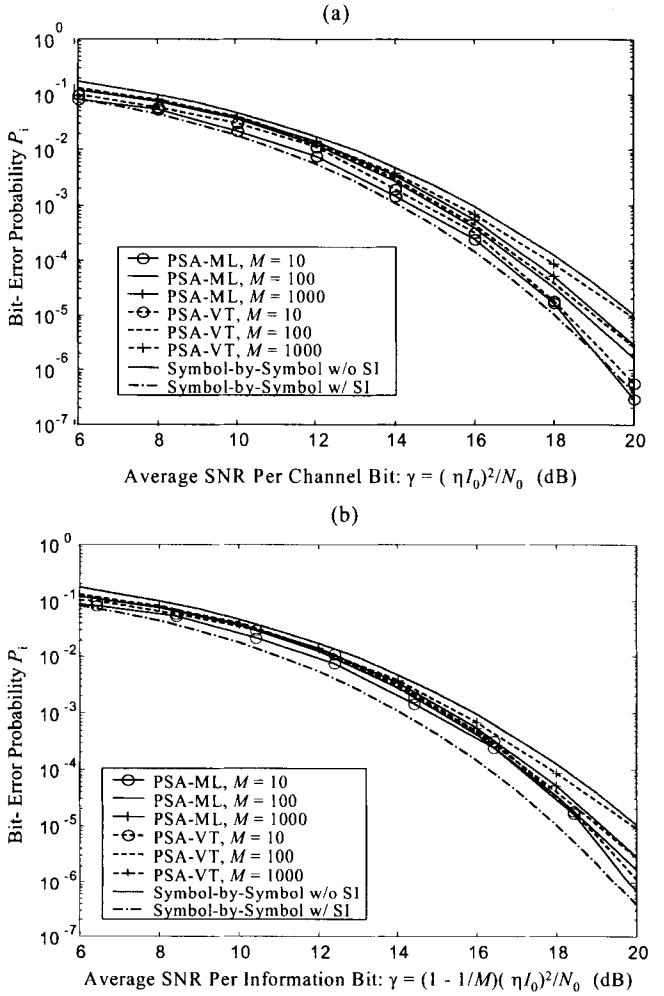


Fig. 10. BER of ith symbol in an M-bit frame versus (a) average electrical SNR of each channel bit and (b) average electrical SNR of eachinformation bit using PSA-ML and PSA variable threshold (VT). Wechoose different values of frame size M (10, 100, and 1000, respectively). Here $i = M/2$. The covariance oflog-amplitude due to turbulence-induced fading is $\sigma_X = 0.1$. The ratio of bit-interval T versus fading coherencetime d_T is $T/d_T = 0.001$. The BER of ML symbol-by-symbol detection with or without instantaneous channel fading side-information (SI) are also shown in the figure for comparison.

of-concept system employs a personal computer (PC) for data acquisition and decoding, the bit rate is limited to 3 kb/s. Using appropriate special-purpose hardware, these detection techniques can be implemented at the bit rates of interest in most applications (Mb/s to Gb/s).

Our experimental system is shown in Fig. 12. At the transmitter, we employ a 675 nm, 0.95 mW (Class II) laser diode module with an output beam divergence of

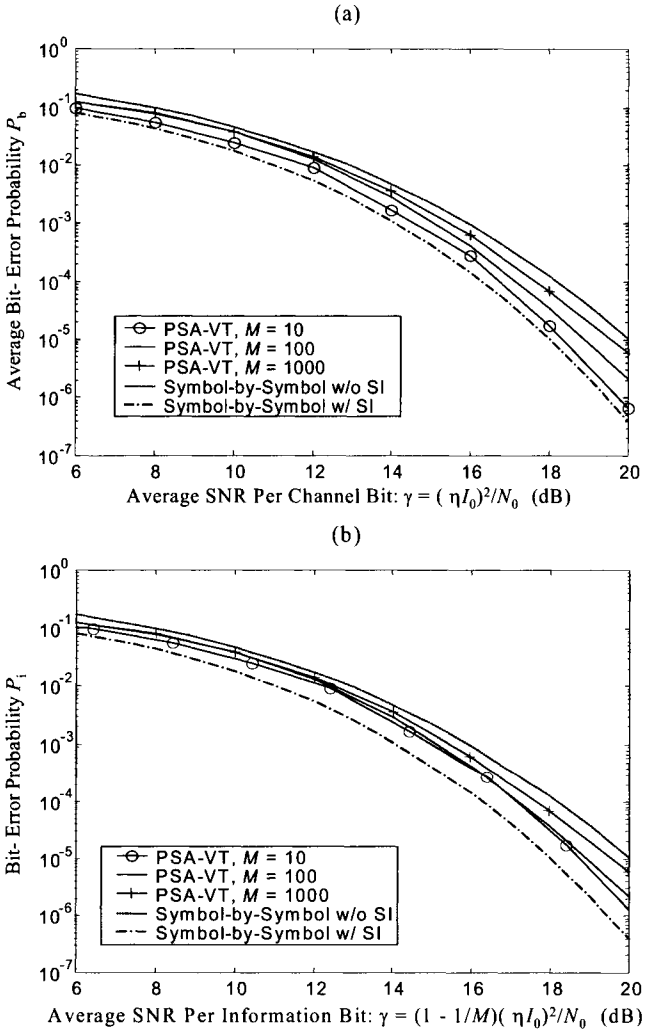


Fig. 11. Average BER of an M-bit frame versus (a) average electrical SNR of each channel bit and (b) average electrical SNR of each information bit using PSA-VT detection. We choose the frame size M to be 10, 100, and 1000, respectively. The covariance of log-amplitude fluctuation due to turbulence-induced fading is $\sigma_X = 0.1$. The ratio of bit-interval T versus fading coherence time d_T is $T/d_T = 0.001$. The BER of ML symbol-by-symbol detection with or without instantaneous channel fading side-information (SI) are also shown in the figure for comparison.

0.35 mrad. We modulate the injection current to achieve OOK with non-return-to-zero pulses. The receiver uses a telescope with an entrance aperture diameter of $D = 8$ cm. An optical bandpass filter with 10 nm bandwidth (–3 dB) is used to minimize ambient light noise. A 1.1-mm^2 p-i-n photodiode is coupled to a transimpedance preamplifier and a second-order Bessel lowpass filter having a 2-kHz bandwidth (–3 dB). The back-to-back receiver sensitivity is –69 dBm at a BER of 10^{-4}. Received electrical signals

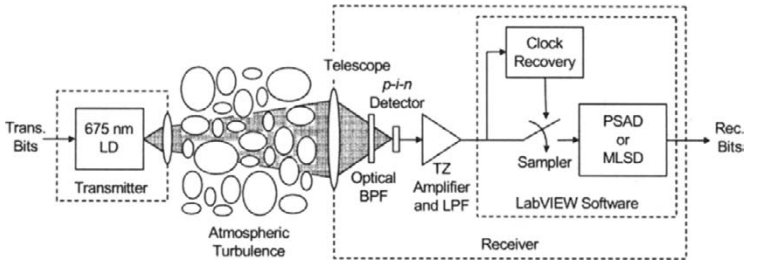


Fig. 12. Experimental 3 kb/s free-space optical link using pilot-symbolassisted detection (PSAD) or maximum-likelihood sequence detection (MLSD) to mitigate turbulence-induced scintillation noise.

are sampled using a PC-interfaced data acquisition card, and the MLSD and PSAD algorithms are implemented in LabVIEW software.

Optical signals are transmitted over a 500-m outdoor path between Cory Hall and Doe Library on the U.C. Berkeley Campus. Because of the relatively large photodiode size, turbulence-induced image degradation has negligible impact over the transmission range employed. The intensity correlation length can be estimated as λL, where λ is the transmission wavelength and L is the transmission range [17], yielding $d_0 \approx 1.8$ cm. Although the receiver aperture, $D = 8$ cm, is somewhat larger than this estimate of d_0, aperture averaging does not completely eliminate scintillation noise, and the observed standard deviation of the log-amplitude is $\sigma_X \approx 1.6$. Based on experimental measurements, we estimate the intensity correlation time to be $\tau_0 \approx 35$ ms, which is much longer than the receiver observation interval $T_0 \approx 0.01$ ms. Hence, both MLSD and PSAD are expected to be effective in mitigating scintillation noise. We have implemented MLSD using the SMC model and the Method 2 described in [8], and have implemented PSAD using PSA-VT with a frame size of $M = 12$ bits.

Our experimental results are shown in Fig. 13, which shows the BER for On-state bits versus the average received electrical signal-to-noise ratio (SNR) for various detection techniques. The measured results for ML symbol-by-symbol detection, PSAD and MLSD are indicated by squares, triangles and circles, respectively. Using measured values of σ_X and τ_0, we have used the theories described in [7,8,19] to compute the dashed, dot-dashed, and solid lines, respectively. We obtain excellent agreement between experimental and theoretical results with no adjustable parameters. As compared to symbol-by-symbol detection, we observe experimentally that PSAD gives an SNR gain of about 1.9 dB, while MLSD yields an SNR gain of about 2.4 dB (both at 10^{-3} error probability). Even larger gains are expected for longer propagation paths with larger values of the log-amplitude standard deviation σ_X.

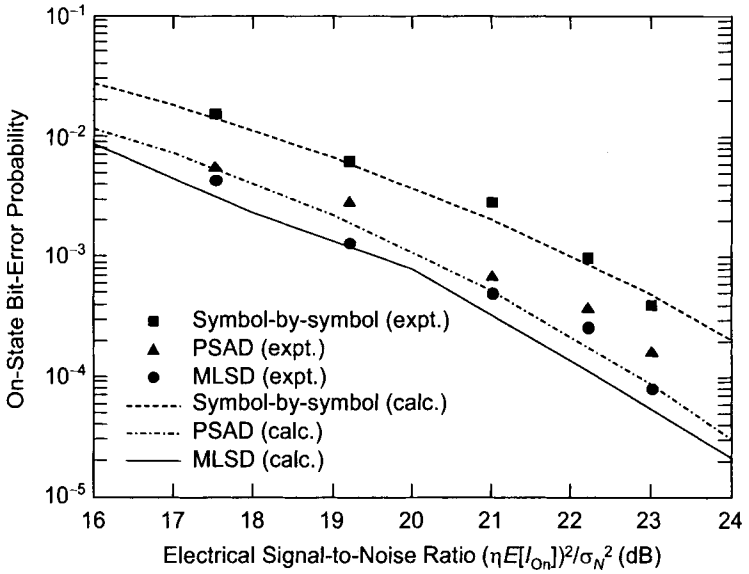


Fig. 13. BER for On-state bits versus average received electrical signal-tonoise ratio in a 500-m outdoor transmission experiment. PSAD: pilot symbol-assisted detection. MLSD: maximum-likelihood sequence detection.

6. Performance Bounds for Coded Free-Space Optical Communication Through Atmospheric Turbulence

Error-control coding can be applied to improve the error performance of free-space optical communication links over atmospheric turbulence channels [18]. The theoretical error performance of coded systems over time-varying channels has been under research for many years [22,30–32]. In most wireless communication systems, the channel is not memoryless. The error performance of such continuous fading channels with memory often requires lengthy computer simulations. Analytical performance bounds have been widely used to study the error performance of communication systems. In this section, we find an expression for the pair-wise codeword-error probability and upper-bound the codeword-error probability P_{block} by

$$P_{\text{block}} \leq \sum_{j, C_j \in S_C} P(C_j) \left[\sum_{\substack{k, C_k \in S_C \\ j \neq k}} P_e(C_j, C_k) \right], \tag{89}$$

where S_C is the set of all codewords and PC_j is the probability that codeword C_j is transmitted. The pairwise error probability $P(C_j, C_k)$ is the probability that when codeword C_j is transmitted, the decoder favors selection of an incorrect codeword C_k over C_j. With the knowledge of the weight enumerating function (WEF), we can further simplify the calculation of Eq. (89) and extend it to accurately estimate the

error bounds of constituent codes where the number of codewords is infinite, such as convolutional codes and turbo codes [32,33]. Many error bounds have been introduced for radio-frequency channels, which often can be well-modeled as Rayleigh or Rician fading channels [30,31].

In this section we will derive an error performance bound for coded OOK free-space optical communication through atmospheric turbulence channels, where the fading is described by a joint log-normal distribution.

6.1. Pairwise Codeword-Error Probability Bound

In this analysis we assume the turbulence-induced fading to be piecewise-constant during each bit interval and known to the receiver, i.e., receiver has perfect state information (SI) about the channel. The receiver utilizes ML soft decoding.

Consider two n-bit codewords $C_j = [c_1^j, c_2^j, \ldots, c_n^j]$ and $C_k = [c_1^k, c_2^k, \ldots, c_n^k]$. Define the symmetric difference of their On-state symbol index subsets, i.e., the set consists of all those points that belong to one or the other of the two sets but not to both:

$$S_d^{C_j,C_k} = \left\{ n_i \in \{1,2,\ldots,n\}, c_{n_i}^j \neq c_{n_i}^k \right\}_{i=1}^{m_{j,k}}, \tag{90}$$

and define its complement set:

$$S_c^{C_j,C_k} = \left\{ n_i \in \{1,2,\ldots,n\}, c_{n_i}^j = c_{n_i}^k \right\}_{i=1}^{n-m_{j,k}}. \tag{91}$$

The energy over the symmetric difference set of two codewords can be defined as

$$\epsilon_{C_j,C_k} = \sum_{\substack{i=1 \\ n_i \in S_d^{C_j,C_k}}}^{M_{j,k}} (\eta I_{n_i})^2. \tag{92}$$

Owing to the linearity of the code, the pairwise error probability between C_j and C_k can be approximated by the pairwise error probability between a codeword whose index subset of On-state bits is the same as the symmetric difference set $S_d^{Cj,Ck}$ and the all-zero codeword:

$$\begin{aligned} E[P_e(C_j,C_k)] &\approx E\left[Q\left(\sqrt{\frac{\epsilon_{C_j,C_k}}{2N_0}}\right)\right] \\ &= \int_{\mathbf{X}} f(\mathbf{X}) Q\left(\sqrt{\frac{\epsilon_{C_j,C_k}}{2N_0}}\right) d\mathbf{X}. \end{aligned} \tag{93}$$

Since $Q(\sqrt{x}) \leq \frac{1}{2}\exp\left(\frac{-x}{2}\right)$, an upper bound on the pairwise error probability is

$$E[P_e(C_j,C_k)] \leq \frac{1}{2} E\left[\left(\frac{\epsilon_{C_j,C_k}}{4N_0}\right)\right]. \tag{94}$$

Defining the average signal to noise ratio (SNR) $\gamma = \frac{(\eta I_0)^2}{N_0}$ we can express Eq. (94) as

$$E[P_e(C_j, C_k)] \leq \frac{1}{2}\int_{\mathbf{X}} f(\mathbf{X}) \exp\left\{-\frac{\gamma}{4} \sum_{\substack{i=1 \\ n_i \in S_d^{C_j, C_k}}} [\exp(2X_{n_i})]^2\right\} d\mathbf{X}$$

$$\cong \frac{1}{2}\int_{\mathbf{X}} f(\mathbf{X}) \exp\left\{-\frac{\gamma}{4} \sum_{\substack{i=1 \\ n_i \in S_d^{C_j, C_k}}} \left[\sum_{l=0}^{\infty} \frac{(4X_{n_i})^l}{l!}\right]\right\} d\mathbf{X}. \quad (95)$$

Under the assumption of weak turbulence ($\sigma_X \ll 1$), we can approximate the upper bound (95) as

$$E[P_e(C_j, C_0)] \lesssim \frac{1}{2}\int_{\mathbf{X}} f(\mathbf{X}) \exp\left[-\frac{\gamma}{4} \sum_{l=0}^{2} \sum_{\substack{i=1 \\ n_i \in S_d^{C_j, C_k}}}^{m_{j,k}} \frac{(4X_{n_i})^l}{l!}\right] d\mathbf{X} \quad (96)$$

and

$$E[P_e(C_j, C_0)] \lesssim \frac{1}{2} \exp\left(-\frac{\gamma m_{j,k}}{4}\right) \prod_{i=1}^{m_{j,k}} \frac{1}{\sqrt{1+4\gamma\lambda_i}} \exp\left[\sum_{i=1}^{m_{j,k}} \frac{\gamma^2 u_i^2 \lambda_i}{2(1+4\gamma\lambda_i)}\right], \quad (97)$$

where λ_i is the ith eigenvalue of the covariance matrix $C_X^{j,k}$ and u_i is the sum of the elements in the corresponding eigenvector.

To verify the accuracy of the approximate upper bound (97), we calculate the pairwise error probability of two codewords whose difference set is $S_d^{C_j, C_k} = \{1, 3, 4, 6\}$ versus the average SNR, choosing $T/\tau_0 = 0.04$. The approximate upper bounds for log-amplitude variances $\sigma_X = 0.05$, 0.15, and 0.25 are indicated by the solid lines in Fig. 14. The pairwise codeword-error probability calculated using Eq. (93) is indicated by the dashed lines in Fig. 14 for comparison. We can see that the upper bound (97) is accurate under the assumption of small σ_X in the weak-turbulence regime.

To demonstrate the limits of applicability of Eq. (97), we also consider larger values of σ_X in Fig. 15. In Fig. 15, we see when σ_X increases from 0.25 to 0.65, which should obviously result in higher error probability, the approximate pairwise error probability computed using Eq. (97) decreases instead. This is because terms of higher order in the log-amplitude, which we ignored in deriving Eq. (96), are no longer negligible when σ_X is large, even though the weak-turbulence approximation made in Eq. (5) is still valid. Therefore, Eq. (97) is no longer valid for large σ_X.

Since many free-space optical communication systems operate in the weak-turbulence regime, the approximate upper bound (97) should be widely applicable to estimate the pairwise error probability for long block codes and constituent codes, such as convolutional codes and turbo codes, which can be useful in optimizing the design of codes for free-space optical turbulence channels.

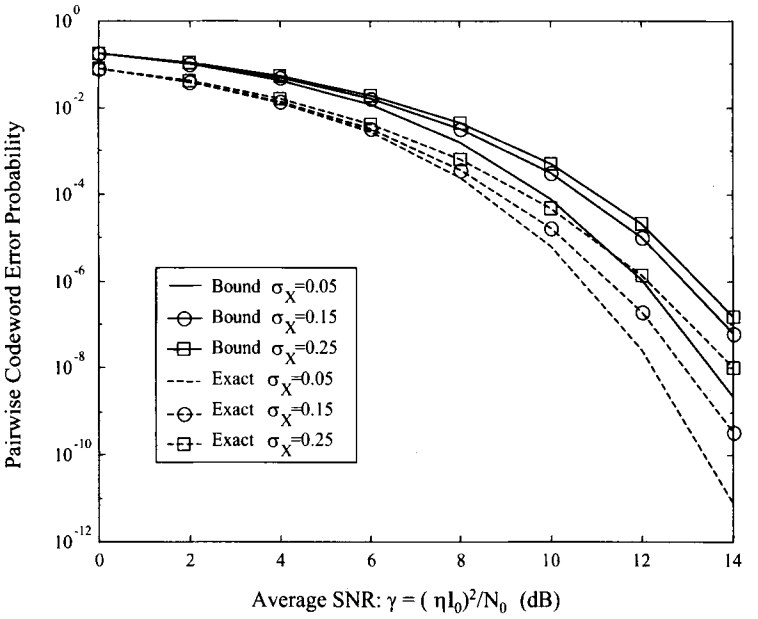


Fig. 14. Pairwise codeword-error probability versus average SNR using exact integration (dashed lines) and using approximate upper bounds (solid lines) for various log-amplitude variances ($\sigma_X = 0.05, 0.15, 0.25$).

6.2. *Error-Probability Bounds for Various Coding Schemes*

Using the approximate pairwise codeword-error upper bound that we have derived, we can compute upper bounds on the error probability for various coding schemes. In this section, we restrict ourselves to considering linear codes, and assume that all codewords are selected with equal probability. We derive upper bounds on the BER of block codes, convolutional codes and turbo codes.

6.2.1. Block Codes

For a binary linear (n, k) block code with a set of codewords $\{C_0, C_1, \ldots, C_{2^k-1}\}$, where C_0 denotes the all-zero codeword, the average block-error probability with maximum likelihood (ML) decoding is:

$$E[P_{\text{block}}] \leq \sum_{i=1}^{2^k-1} E[P_e(C_j, C_0)], \tag{98}$$

and the average BER upper bound is

$$E[P_b] \leq \frac{1}{k} \sum_{i=1}^{2^k-1} B_j E[P_e(C_j, C_0)], \tag{99}$$

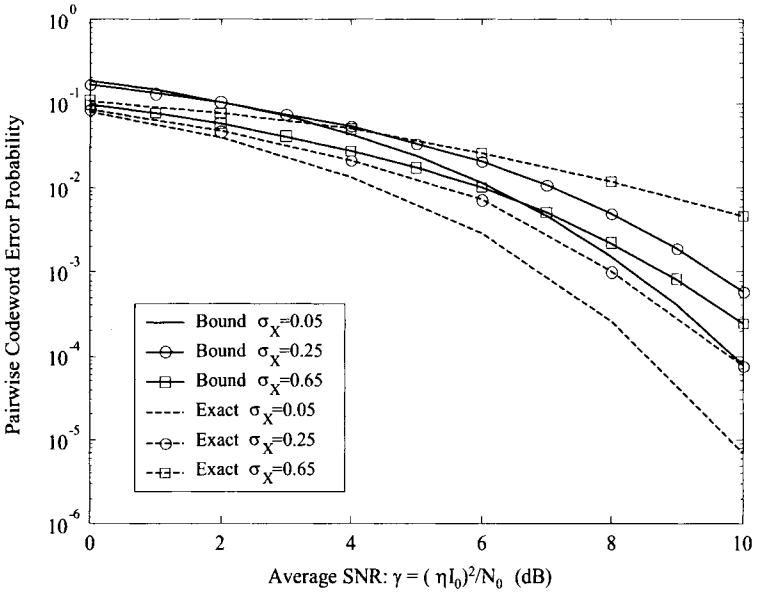


Fig. 15. Pairwise codeword-error probability versus average SNR using exact integration (dashed lines) and using approximate upper bounds (solid lines) for larger values of the log-amplitude variances ($\sigma_X = 0.05, 0.25, 0.65$).

where B_j is the Hamming weight of the information sequence corresponding to codeword C_j.

It is straightforward to apply Eq. (97) in Eqs. (98) and (99) to obtain performance bounds for block codes.

6.3. Convolutional Codes

Consider a rate $R = k/n$ linear convolutional code, where C_0 is the all-zero codeword and $\{C_j, j > 0\}$ is the set of nonzero codewords whose initial state is the all-zero state which first re-merge with the all-zero state at their final state. With ML decoding, the average BER can be upper-bounded by

$$E[P_b] \leq \frac{1}{k} \sum_{j=1}^{\infty} B_j E[P_e(C_j, C_0)]. \tag{100}$$

To estimate Eq. (100), we have to reorder the pairwise error elements by sorting the codewords according to their Hamming weights. Define $w_H(C_i, C_j)$ to be the Hamming distance between two codewords. Let S_w be the subset of codewords with Hamming weight w: $S_w = \{C_{j_k} | w_H(C_{j_k}, C_0) = w\}_{k=1}^{n_w}$, where n_w is the number of codewords in S_w. The upper bound (100) can be expressed as

$$E[P_b] \leq \frac{1}{k} \sum_{w=w_{\text{free}}}^{\infty} \left\{ \sum_{\substack{l=1 \\ C_{j_l} \in S_w}}^{n_w} B_j E[P_e(C_{j_l}, C_0)] \right\}, \tag{101}$$

where w_{free} is the free distance of the code, which is the minimum Hamming weight of any codeword except C_0. We can truncate the sum in Eq. (101) by ignoring negligible terms having large Hamming weights. When the SNR is high, such an approximation is quite accurate and we can approximate the average BER by

$$E[P_b] \cong \frac{1}{k} \sum_{\substack{l=1 \\ C_{j_l} \in S_{w_{\text{free}}}}}^{n_{w_{\text{free}}}} B_j E[P_e(C_{j_l}, C_0)]. \tag{102}$$

Applying Eq. (97) to Eq. (102), we obtain the approximate BER:

$$E[P_b] \cong \frac{1}{2k} \exp\left(-\frac{\gamma w_{\text{free}}}{4}\right) \times \sum_{\substack{l=1 \\ C_{j_l} \in S_{w_{\text{free}}}}}^{n_{w_{\text{free}}}} \left\{ B_{j_l} \prod_{i=1}^{w_{\text{free}}} \frac{1}{\sqrt{1+4\gamma\lambda_i^{j_l}}} \exp\left[\sum_{i=1}^{w_{\text{free}}} \frac{\lambda_i^{j_l} (\gamma u_i^{j_l})^2}{2(1+4\gamma\lambda_i^{j_l})} \right] \right\}. \tag{103}$$

We can also apply Eq. (97) to Eq. (101) to find an upper bound to the BER:

$$E[P_b] \lesssim \frac{1}{2k} \sum_{w=w_{\text{free}}}^{\infty} \left\{ \exp\left(-\frac{\gamma w_{\text{free}}}{4}\right) \times \sum_{\substack{l=1 \\ C_{j_l} \in S_w}}^{n_w} \left[B_{j_l} \prod_{i=1}^{w} \frac{1}{\sqrt{1+4\gamma\lambda_i^{j_l}}} \exp\left(\sum_{i=1}^{w} \frac{\lambda_i^{j_l} (\gamma u_i^{j_l})^2}{2(1+4\gamma\lambda_i^{j_l})} \right) \right] \right\}. \tag{104}$$

In Eqs. (103) and (104), $\lambda_i^{j_l}$ is the ith eigenvalue of the covariance matrix C_X and u_i^l is the sum of the elements in the corresponding eigenvector. To simplify Eq. (104), we can define the fading-induced degradation factors:

$$\alpha_w = \max_{C_{j_l} \in S_w, 1 \leq l \leq n_w} \prod_{i=1}^{w} \frac{1}{\sqrt{1+4\gamma\lambda_i^{j_l}}} \tag{105}$$

and

$$\beta_w = \max_{C_{j_l} \in S_w, 1 \leq l \leq n_w} \sum_{i=1}^{w} \frac{\lambda_i^{j_l} (\gamma u_i^{j_l})^2}{2(1+4\gamma\lambda_i^{j_l})}. \tag{106}$$

Then the upper bound (104) can be more simply expressed using Eqs. (105) and (106):

$$E[P_b] \lesssim \frac{1}{2k} \sum_{w=w_{\text{free}}}^{\infty} B^w \alpha_w \exp\left(-\frac{\gamma w}{4} + \beta_w \gamma^2\right), \tag{107}$$

where B^w is the sum of all B_{j_l} for which codeword C_{j_l} has Hamming weight w. Note that B^w can be obtained using the transfer function of the code [32,33]. The upper bound (107) can also be applied to long block codes for which Eq. (99) would involve a prohibitive amount of computation.

6.3.1. Turbo Codes

Turbo codes offer excellent performance in a variety of applications, including optical communications, and have aroused significant interest in the coding community. Turbo codes are parallel concatenated convolutional codes (PCCC) in which the encoder is formed by two or more constituent systematic recursive convolutional encoders joined through an interleaver [32–35]. The information sequence is divided into blocks of length equal to the interleaver length K. The input to the first encoder is the original information bit sequence, and the inputs to the other encoders are interleaved versions of the information block. The encoded sequence consists of the information sequence and the parity check bits from all encoders. Many decoding schemes and error-performance analyses for turbo codes have been documented in the literature. An abstract uniform interleaver approach has been widely used to derive the average of the upper bounds obtainable for the whole class of deterministic interleavers [34].

We start with the definition of the input-redundancy weight enumerating function (IRWEF) for systematic convolutional code:

$$A_{i,j}(W,Z) \equiv \sum_{w,z} A_{i,j}^{w,z} W^w Z^z, \tag{108}$$

where i and j denote the initial and final states of the codewords. $A_{i,j}^{w,z}$ denotes the number of codewords generated by an input information word having Hamming weight w and having parity check bits of weight z, so that the overall Hamming weight of the systematic codeword is $w+z$. We can also define the conditional WEF:

$$A_{i,j}^{w}(w,Z) \equiv W^w A_{i,j}^{C}(Z) = W^w \sum_{z} A_{i,j}^{w,z} Z^z. \tag{109}$$

Making use of the properties of a uniform interleaver [34], we obtain the average conditional WEF of all possible turbo codes with respect to the whole class of interleavers:

$$A_{ip,jq}^{w}(w,Z) = W^w \frac{A_{i,j}^{C_1}(Z) A_{p,q}^{C_2}(Z)}{\binom{K}{w}}, \tag{110}$$

where $A_{i,j}^{C_1}(Z)$ and $A_{p,q}^{C_2}(Z)$ are the conditional WEFs of two encoders, respectively, and K is the interleaver length. It has been shown when K is sufficiently large, we can accurately approximate the error performance with the paths which diverge from the Off-states of both constituent encoders and re-merge into Off-states after K steps [34], which have WEF $A_{00,00}(WZ) = \sum_{w,z} A_{00,00}^{w,z,} W^w Z^z$. A tight bound for the pairwise codeword error in correlated turbulence-induced fading channel requires knowledge of the positions of differing symbols, as we discussed above. For simplicity, we can loosen the bound by making the pessimistic assumption that all $d = w + z$ differing symbols are adjacent. Therefore the average BER is upper-bounded as

$$E[P_b] \leq \sum_{w,z} \frac{w}{k} A_{00,00}^{w,z} P_e(w+z) = \sum_{w=1}^{K} \frac{w}{k} \left[\sum_{d \geq w} A_{00,00}^{w,d-w} P_e(d) \right], \tag{111}$$

where $P_e(d)$ is the pairwise codeword-error bound of the all zero codeword and the codeword with d adjacent On-state bits. Let us define

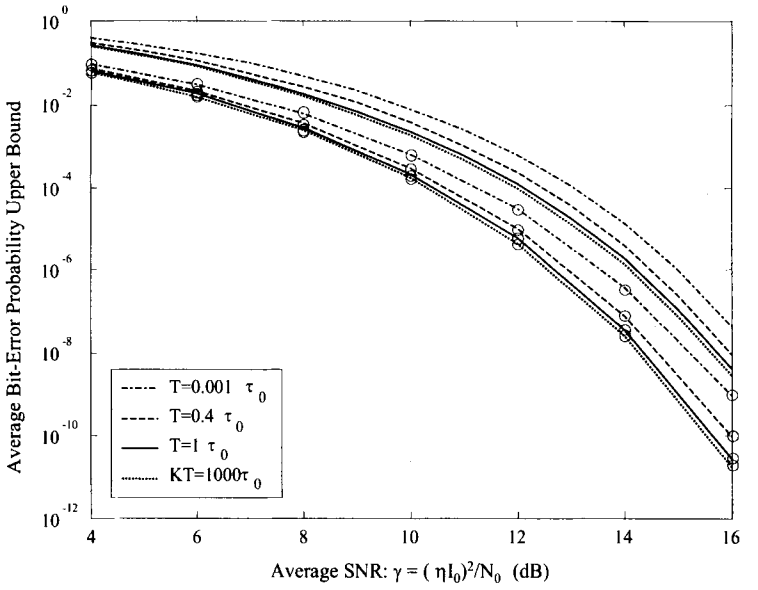


Fig. 16. The simulated BER (lines with circles) and the approximate upper bound (lines without circles) for (7, 4) Hamming codes versus average SNR. The log-amplitude variance $\sigma_X = 0.2$ and the ratio of the adjacent codeword bit interval T to the coherence time of turbulence-induced fading (T/τ_0) is chosen to be 0.001, 0.4, and 1, respectively. We also show the simulation results where a block interleaver is used with $KT/\tau_0 = 1000$.

$$\left[\frac{W}{K}\frac{\partial A_{00,00}(W,Z)}{\partial W}\right]\Bigg|_{W=Z=D} = \sum_d A_d D^d, \tag{112}$$

which allows us to express Eq. (111) as

$$E[P_b] \leq \sum_d A_d P_e(d). \tag{113}$$

We can simplify Eq. (113) by ignoring the negligible higher-order terms with larger Hamming weights as discussed in section 6.2.2.

6.4. Numerical Simulation Results

In this section, we use the approximate error-probability upper bounds derived above to numerically evaluate the performance of some practical codes.

We first study a Hamming (7, 4) code. In Fig. 16, we plot the approximate BER upper bound (lines without circles) versus average SNR with log-amplitude variance $\sigma_X = 0.2$. T/τ_0 takes on the values 0.001, 0.4, 1, and 1000, respectively. We also present the BER estimated using Monte-Carlo simulation (lines with circles) for comparison. Fig. 16 shows that when T/τ_0 increases, the average BER for (7, 4) Hamming

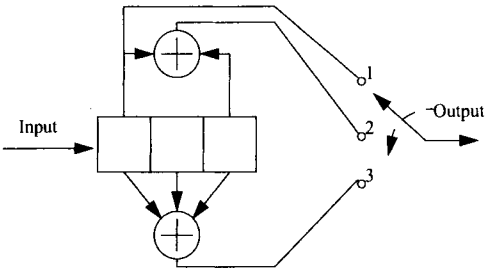


Fig. 17. Rate1/3 convolutional encoder with 3-stage shift registers.

codes will decrease. Therefore, we can implement interleaving to compensate for the coding-gain penalty due to the memory of channel. Consider the normal block interleaver of degree K [22], where the codewords are interleaved so that the adjacent bits of coded bit sequence are transmitted at intervals of KT. Figure 16 shows that interleaving will improve the bit-error performance when KT/τ_0 is large. However when $KT > \tau_0$, further increase of the interleaver depth will not significantly improve the bit-error performance. As shown in Fig. 16, when T/τ_0 increases from 1 to 1000, the corresponding average upper bounds to the BER are very close. Therefore it is not necessary to further increase the interleaver depth when $KT > \tau_0$.

For convolutional codes, we consider the example of a rate-1/3 code, whose encoder diagram is shown in Fig. 17. We choose the log-amplitude variance $\sigma_X = 0.2$. T/τ_0 takes on the values 0.001, 1, and 1000, respectively. In Fig. 18, the approximate BER versus average SNR, computed using Eqs. (103) (lines with squares), (104) (pure lines), and (107) (lines with crosses) respectively. The simulated BER is also shown (lines with circles) in Fig. 18 for comparison. We see that Eq. (107) yields a very accurate estimate of the approximate upper bound (104). Also, when the SNR is high, the higher Hamming weight terms in Eq. (104) are negligible, and the approximate BER (103) merges with the approximate upper bound (104). Interleaving can help to improve the system performance for convolutional codes as well. Similar to the block coding case, there is not much difference in error performance when T/τ_0 increases from 1 to 1000. as shown in Fig. 18. We conclude that for convolutional codes, the interleaver depth K is sufficient when it satisfies $KT > \tau_0$.

Finally, we present simulations for a rate-1/3 turbo-coded system whose encoder structure is shown in Fig. 19. The approximate BER upper bound is calculated using Eq. (113) with log-amplitude variance $\sigma_X = 0.2$ and $T/\tau_0 = 0.001$. The interleaver length K of the uniform interleaver takes on the values 100, 1000 and 10000, respectively. In Fig. 20, we plot the approximate BER upper bound (lines without circles) and the simulated BER (lines with circles) versus the average SNR for comparison. Comparing the simulation results shown in Fig. 20 to those of the rate-1/3 convolutional code in Fig. 18, we see turbo coding can achieve better BER performance through atmospheric turbulence-induced fading channels if the interleaver length is sufficiently long. In Fig. 20, we see that the error-probability performance of turbo codes continues to improve with increasing interleaver length even when $KT > \tau_0$, especially at low average

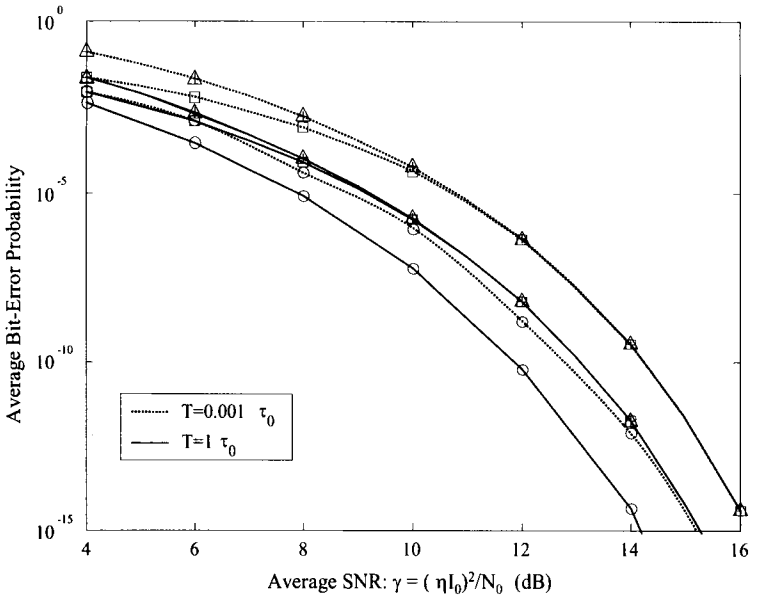


Fig. 18. Average BER for a three-stage, rate-1/3 convolutional code versus average SNR with log-amplitude variance $\sigma_X = 0.2$ and for $T/\tau_0 = 0.001$ and 1. The estimates are provided by Eqs. (103) (lines with squares), (104) (lines with triangles), and (107) (lines with crosses) respectively. The simulated average BER (lines with circles) is also shown for comparison.

SNR [34]. This favors the use of longer interleavers with turbo codes. However, system complexity and delays in coding and decoding will limit the length of interleavers in practical systems using turbo codes.

6.5. Summary

Error-control codes can help to mitigate turbulence-induced signal fading in wireless optical communication through atmosphere turbulence. To study the efficiency of various coding schemes, performance bound analysis has been used for its simplicity.

In this section, we first derived an upper bound on the pairwise codeword-error probability for correlated atmospheric turbulence channels. To avoid complicated multidimensional integration, we have also derived an approximation for this upper bound under the assumption of weak turbulence. The accuracy and the limits of applicability of this approximation have been demonstrated using numerical simulations. We then applied this approximate upper bound to derive the error performance bounds and their approximations for various coding schemes through atmospheric turbulence channels, including block coding, convolutional coding and turbo coding. The analytical upper bounds were then applied to compare the performance of a few specific example codes. The effect of varying the interleaver depth was studied through numerical evaluation of the performance bounds.

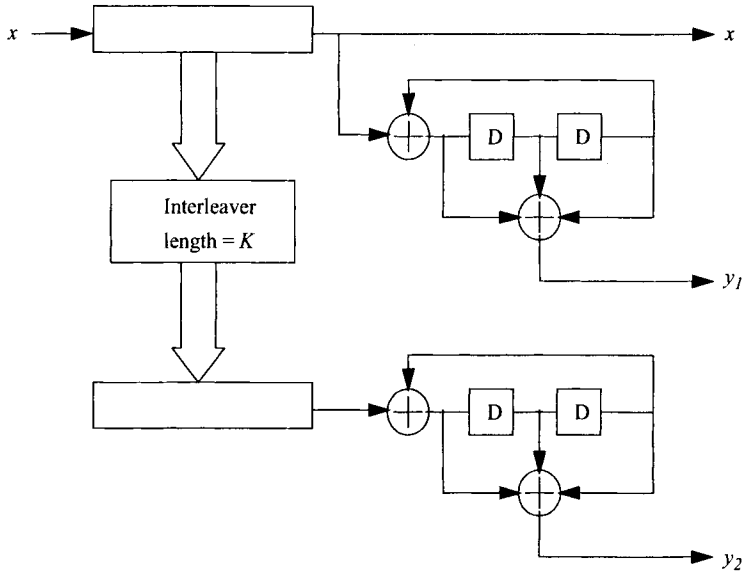


Fig. 19. Encoder structure of a rate-1/3 turbo code with uniform interleaver of length K.

Conclusions

In this paper, we have discussed free-space optical links using intensity modulation with direct detection (IM/DD). We have described several communication techniques to mitigate turbulence-induced log-amplitude fluctuations (i.e., signal fading) in the regime in which the imaging receiver's diameter D_0 is smaller than the turbulence coherence length d_0, and the observation interval T_0 is smaller than the correlation time τ_0. We assume that the receiver has no knowledge of the instantaneous fading state. We have described several classes of communication techniques:

- When a single receiver has knowledge only of the marginal statistics of fading, a symbol-by-symbol ML detector can be used to improve detection performance.
- When a single receiver has knowledge of the joint temporal statistics of the fading, maximum-likelihood sequence detection (MLSD) can be employed, yielding a further performance improvement, at the cost of very high complexity. We introduced a single-step Markov chain model for the fading correlation, and used it to derive a low-complexity, sub-optimal MLSD based on per-survivor processing.
- Spatial diversity reception with multiple receivers can also help to mitigate turbulence-induced fadings. Using the joint spatial statistics of fading at multiple receivers, we have derived the maximum-likelihood (ML) decision rule for spatial diversity reception. ML spatial diversity is applicable when the fading correlation properties are known, but the instantaneous fading state is known.
- When a single receiver is employed, pilot symbol-assisted detection and error-control codes can also be used to mitigate the impact of turbulence-induced signal fading.

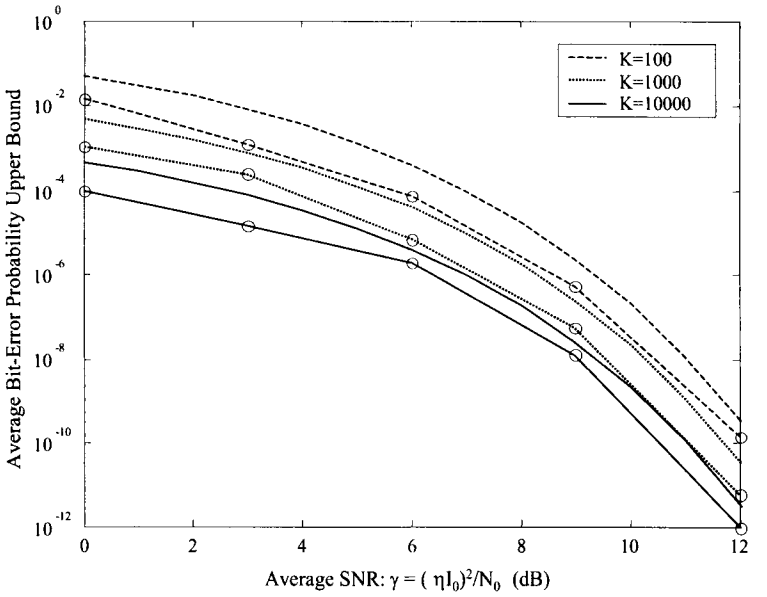


Fig. 20. The simulated BER (lines with circles) and the approximate upper bound (lines without circles) of a rate-1/3 turbo code versus average SNR with log-amplitude variance $\sigma_X = 0.2$ and $T/\tau_0 = 0.001$. A uniform interleaver with different interleaver depths ($K = 100$, 1000, 1000) is considered.

Acknowledgment

This research has been supported by the DARPA STAB Program under contract number DAAH01-00-C0089 and the DARPA MTO MEMS Program under Contract Number DABT63-98-1-0018.

References

1. P.F. Szajowski, G. Nykolak, J.J. Auborn, H.M. Presby and G.E. Tourgee, "High-Power Optical Amplifier-Enabled 1550-nm Terrestrial Free-Space Optical Data-Link Operating at 10 Gb/s", Proc. of IEEE Mil. Commun. Conf., Atlantic City, NJ, Oct. 31–Nov. 3, 1999, pp. 687–689.
2. V.S. Hsu, "MEMS Corner Cube Retroreflectors for Free-space Optical Communications", Masters Thesis, EECS Department, U.C. Berkeley, October 1999.
3. J.M. Kahn, R.H. Katz and K S.J. Pister, "Mobile Networking for Smart Dust", Proc. of ACM/IEEE Intl. Conf. on Mobile Computing and Networking (MobiCom 99), Seattle, WA, August 17–19, 1999.
4. Chu, P.B., Lo, N.R., Berg, E., Pister, K.S.J, "Optical Communication Using Micro Corner Cuber Reflectors", Proc. of IEEE MEMS Workshop, Nagoya, Japan, 26–30 Jan 1997, pp. 350–5.

5. L. Zhou, K.S.J. Pister and J.M. Kahn, "Assembled Corner-Cube Retroreflector Quadruplet", to be presented at 15th IEEE Intl. Conf. on MicroElectroMechanical Systems, Las Vegas, NV, January 2024, 2002.
6. Xiaoming Zhu, J.M. Kahn, "Maximum-Likelihood Spatial-Diversity Reception on Correlated Turbulent Free-Space Optical Channels", IEEE Conf. on Global Commun., San Francisco, CA, November 27–December 1, 2000.
7. X. Zhu and J.M. Kahn, "Free-Space Optical Communication through Atmospheric Turbulence Channels", IEEE Trans. Commun. **50** (8), 1293–1300 (August 2002).
8. X. Zhu and J.M. Kahn, "Markov Chain Model in Maximum-Likelihood Sequence Detection for Free-Space Optical Communication through Atmospheric Turbulence Channels", IEEE Trans. Commun. **51** (3), 509–516 (March 2003).
9. D. Sadot, A. Melamed and N.S. Kopeika, "Effects of Aerosol Forward Scatter on the Long- and Short-Exposure Atmospheric Coherence Diameter", Waves in Random Media **4** (4), 487–98 (Oct. 1994).
10. S. Arnon and N.S. Kopeika, "Effect of Particulates on Performance of Optical Communication in Space and an Adaptive Method to Minimize Such Effects", App. Opt. **33** (21), 4930–37 (20 July 1994).
11. Max Born and Emil Wolf, *Principles of Optics*, Sixth edition (Pergamon Press, Canada, 1980).
12. Joseph W. Goodman, *Statistical Optics* (John Wiley & Sons, Inc., 1985).
13. Akira Ishimaru, *Wave Propagation and Scattering in Random Media*, Vol. 1-2 (Academic Press, Inc., 1978).
14. S. Karp, R. Gagliardi S.E. Moran and L.B. Stotts, *Optical Channels* (Plenum Press, 1988).
15. N.S. Kopeika, A. Zilberman and Y. Sorani, "Measured Profiles of Aerosols and Turbulence for Elevations of 2-20 km and Consequences on Widening of Laser Beams", Optical Pulse and Beam Propagation III, Proc. SPIE, Vol. 4271, 2001.
16. A. Zilberman, N.S. Kopeida and Y. Sorani, "Laser Bean Widening as a Function of Elevation in the Atmosphere for Horizontal Propagation", Laser Weapons Technology II, Proc. SPIE, Vol. 4376, 2001.
17. L.C. Andrews and R.L. Phillips, "Laser Beam Propagation Through Random Media", SPIE Optical Engineering Press, 1998.
18. X. Zhu and J.M. Kahn, "Performance Bounds for Coded Free-Space Optical Communications through Atmospheric Turbulence Channels", IEEE Trans. Commun.**51** (8), 1233–1239 (August 2003).
19. Shane M. Haas and Jeffrey H. Shapiro, "Capacity of Wireless Optical Communications", IEEE J. Sel. Areas Commun. **21** (8), 1346–1357 (October 2003).
20. X. Zhu and J.M. Kahn, "Pilot-Symbol Assisted Modulation for Correlated Turbulent Free-Space Optical Channels", SPIE's 46th Annual Meeting, The International Symposium on Optical Science and Technology, San Diego, California, July 29–August 3, 2001.
21. Gordon L. Stuber, *Principles of Mobile Communication* (Kluwer Academic Publishers, 1996).
22. John G. Proakis, *Digital Communication*, Third Edition (McGraw-Hill, Inc., 1995).
23. M. Srinivasan and V. Vilnrotter, "Avalanche Photodiode Arrays for Optical Communication Receivers", NASA TMO Progress Report 42-144, Feb. 15, 2001.
24. Ibrahim, M.M., Ibrahim, A.M., "Performance analysis of optical receivers with space diversity reception", IEE Proc. Commun., **143** (6), 369–372 (Dec. 1996).

25. D.M. Boroson, "Interplanetary Deep-Space Optical Communication Systems and Techniques", Proc. of IEEE Lasers and Electro-Optics Society Annual Meeting, Tucson, AZ, October 26–30, 2003.

26. R. Raheli, A. Polydoros and C-K. Tzou, "Per-survivor processing: A general approach to MLSE in uncertain enviroments", IEEE Trans. Commun. **43**, 354–364 (Feb./Mar./Apr. 1995).

27. J.K. Cavers, "An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels", IEEE Trans. Commun. **40** (4), 686–93 (Nov. 1991).

28. S.B. Bulumulla, S.A. Kassam, S.S. Venkatesh, "Pilot symbol assisted diversity reception for a fading channel", Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, Seattle, WA, USA, Vol. 6, pp.3417–20, May 1998.

29. X. Zhu, J. M. Kahn and J. Wang, "Mitigation of Turbulence-Induced Scintillation Noise in Free-Space Optical Links using Temporal-Domain Detection Techniques", IEEE Photon. Technol. Lett. **15** (4), 623–625 (April 2003)

30. Francois Gagnon and David Haccoun, "Bounds on the Error Performance of Coding for Nonindependent Rician-Fading Channels", IEEE Trans. Commun. **40** (2), 351–360 (Feb. 1992).

31. Gideon Kaplan and Shlomo Shamai, "Achievable Performance over the Correlated Rician Channel", IEEE Trans. Commun. **42** (11), 2967–78 (Nov. 1994).

32. Tolga M. Duman and Masoud Salehi, "Performance Bounds for Turbo-Coded Modulation Systems", IEEE Trans. Commun. **47** (4), 511–21 (Apr. 1999).

33. Kamran Kiasaleh, "Turbo-Coded Optical PPM Communication Systems", J. Lightwave Technol. **16** (1), 18–26 (Jan. 1998).

34. Sergio Benedetto and Guido Montorsi, "Unveiling Turbo Codes: Some Results on Parallel Concatenated Coding Schemes", IEEE Trans. Commun. **42** (2), 409–428 (March 1996)

35. Eric K. Hall and Stephen G. Wilson, "Design and Analysis of Turbo Codes on Rayleigh Fading Channels", IEEE J. Sel. Areas Commun. **16** (2), 160–74 (Feb. 1998).

DOI: 10.1007/s10297-005-0057-x
Originally published in J. Opt. Fiber. Commun. Rep. 2, 558–602 (2005)

Optical communications in the mid-wave IR spectral band

Narasimha S. Prasad

NASA Langley Research Center, 468/LEOB
5 North Dryden Street, B1202
Hampton, VA 23681-9403
Email: n.s.prasad@larc.nasa.gov

Abstract. The mid-wave IR (MWIR) spectral band extending from 3 to 5 microns is considered to be a low loss atmospheric window. The MWIR wavelengths are eye safe and are attractive for several free-space applications including remote sensing of chemical and biological species, hard target imaging, range finding, target illumination, and free-space communications. Due to the nature of light-matter interaction characteristics, MWIR wavelength based systems can provide unique advantages over other spectral bands for these applications. The MWIR wavelengths are found to effectively penetrate natural and anthropogenic obscurants. Consequently, MWIR systems offer increased range performance at reduced power levels. Free-space, line-of-sight optical communication links for terrestrial as well as space based platforms using MWIR wavelengths can be designed to operate under low visibility conditions. Combined with high-bandwidth, eye-safe, covert and jam proof features, a MWIR wavelength based optical communication link could play a vital role in hostile environments.

A free-space optical communication link basically consists of a transmitter, a receiver and a scheme for directing the beam towards a target. Coherent radiation in the MWIR spectral band can be generated using various types of lasers and nonlinear optical devices. Traditional modulation techniques are applicable to these optical sources. Novel detector and other subcomponent technologies with enhanced characteristics for a MWIR based system are advancing. Depending on the transmitter beam characteristics, atmospheric conditions may adversely influence the beam propagation and thereby increasing the bit error rate. For satisfactory transmission over a given range, the influence of atmosphere on beam propagation has to be analyzed. In this chapter, salient features of atmospheric modeling required for wavelength selection and performance prediction is presented. Potential optical sources and detectors for building a practical MWIR communication link are surveyed. As an illustration, the design configuration and experimental results of a recently demonstrated free-space, obscurant penetrating optical data communication link suitable for battlefield appli-

cations is discussed. In this case, the MWIR wavelength was derived using an all solid-state, compact, optical parametric oscillator device. With this device, weapon codes pertaining to small and large weapon platforms were transmitted over a range of 5 km. Furthermore, image transmission through light fog, accomplished using this hardware, is also presented.

Advances in source and detector technologies are contributing to the development of cost effective systems compatible with various platforms requirements. In coming years, MWIR wavelengths are anticipated to play a vital role in various human endeavors.

1. Introduction

The mid-wave IR (MWIR) spectral band is generally defined over 3–5 μm. In this spectral band, extraneous thermal and solar background emissions are relatively low. Furthermore, there are several spectral sub-bands with relatively low atmospheric attenuation in the MWIR region. Hence, the MWIR spectral band is considered to be a low loss atmospheric window and hence is very attractive for several free-space applications. MWIR wavelengths can be generated using lasers and nonlinear optical devices. The MWIR wavelengths are useful in probing various states of matter by quantifying the effects on wavelength, amplitude, phase and polarization due to various optical phenomenon including reflection, refraction, diffraction and interference. The MWIR window plays a significant role in applications involving transmission of optical wavelengths through atmosphere over long ranges. Such applications include but not limited to remote sensing of chemical and biological species, hard target imaging, range finding, illumination, and free-space communications.

In general, for a given application, the characteristics of the atmospheric medium, optical transmitter, remote target and receiver characteristics have to be evaluated. In free-space based applications, the effects of atmospheric components on propagating optical radiation parameters have to be ascertained to choose operational wavelengths. Unwanted background emissions due to various sources and reduction in transmitted beam intensity have to be minimized. In many practical applications, the influence on optical amplitude or intensity over a given range is of primary importance. The optical radiation undergoes spectrally dependent attenuation due to various atmospheric processes. The extent of attenuation depends upon a variety of atmospheric weather conditions.

The use of an appropriate wavelength or wavelengths for obtaining maximum range performance accompanied by detection schemes that permit precision amplitude measurements provide valuable information. The spectral attenuation characteristics are exploited in the atmospheric measurements of chemical or biological species. For, e.g., in Light Detection And Ranging (lidar) measurements, depending on the operational scheme, the relative attenuation of judiciously selected two or more wavelengths with respect to the molecular resonance features provide useful information about the concentration of desired species. In free-space communication applications, wavelengths that undergo minimum attenuation under desired weather conditions to provide maximum range performance with sufficient signal-to-noise ratio (SNR) are of importance. Target characteristics in addition to spectral attenuation have to be considered for hard target imaging and range finding applications. In any case, the atmospheric modeling

has to be carried out to achieve optimum system performance and is discussed in the following section.

2. Atmospheric Modeling

Absorption and scattering processes attenuates the propagating optical radiation. Molecular absorptions are attributed to electronic, vibrational, or rotational transitions, which manifest as sharp lines that are localized in wavelength. Water vapor and carbon dioxide are two prominent contributors for atmospheric absorption. On the contrast, broad featureless absorption bands are referred to as continuum absorptions and are caused by the wings of vibrational and vibrational rotational absorption bands or to collision-induced absorption bands. Backscattering off airborne particulates such as dust is often referred to as aerosol scattering, which also contributes to atmospheric attenuation. The primary source of dust is the terrestrial soil carried by wind. However, the aerosol particulates could be hygroscopic in nature that generally exists in the vicinity of water. Hygroscopic aerosols are typically tars and resins created by the oxidation of vegetation extruded aromatic hydrocarbons or sea salts that are injected into the atmosphere at ocean surfaces. The mean size of hygroscopic particles increases with increasing humidity. Scattering efficiency increases with particulate size. The atmospheric attenuation due to aerosol scattering is inversely proportional to the wavelength or in other words it becomes weaker as wavelengths increase.

Atmospheric modeling allows the selection of appropriate wavelengths for a given application. Atmospheric transmission can be modeled using FASCOD3 (Fast Atmospheric Signature CODe) and FASCOD3p. Both simulation codes were developed under the sponsorship of the U.S. Air Force Geophysics Laboratory (AFGL) (Hanscom AFB) and represent the culmination of several decades of research. FASCOD3 is the newest version of the transmission code and was released in 1996. The codes provide an "exact" or line-by-line radiance and transmission model that calculates the effect of the Earth's atmosphere on the emission and propagation of electromagnetic radiation. The codes are anchored by atmospheric measurements to the extent allowed by available databases. Transmission losses due to molecular absorptions, continuum absorptions, and aerosol scattering can be calculated either separately or aggregately by FASCOD [1]. Either version of FASCOD uses HITRAN molecular spectroscopic databases to generate high-resolution spectral transmittance data. Each version of HITRAN contains line-by-line absorption data, including absorption cross-sections, for 32 (HITRAN'92) or 38 (HITRAN'96) most prevalent molecular species. The databases also include aerosol data. FASCOD3p interrogates the HITRAN database and then applies empirical line-broadening algorithms upon the HITRAN molecular absorption bands. These line-broadening algorithms take into account many simulation parameters such as, the simulation path, the model atmosphere, visibility, rain, aerosol models etc. The resolution of FASCOD is many orders of magnitude lower than LOWTRAN and is predominately a function of the resolution of the HITRAN database.

The atmospheric transmission can be calculated using Beer's Law and is given by

$$T(R) = \exp\left(-\int_0^R \alpha(r)\,dr\right), \tag{1}$$

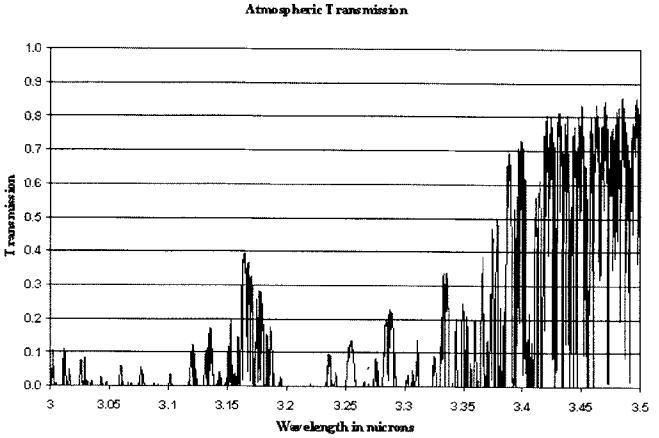


Fig. 1. Atmospheric transmission from 3 to 3.5 μm.

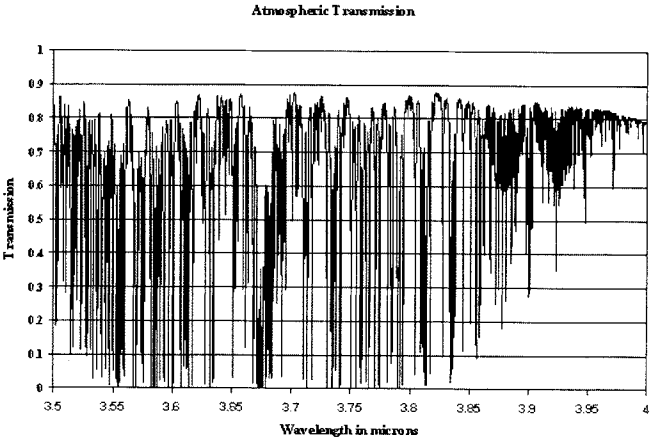


Fig. 2. Atmospheric transmission from 3.5 to 4 μm.

where $\alpha(r)$ is the extinction coefficient profile along the path [0,R] and R is the range from the transmitter to the receiver. The extinction coefficient consists of a combination of aerosol and molecular extinctions and can be modeled using AFGL's FASCODE coupled with the HITRAN data base. Figures 1 to 4 illustrate atmospheric modeling using FASCOD3p in the MWIR spectral band. It is seen that several sub-bands with relatively high atmospheric transmission over 3.4 to 4 μm exists.

Besides FASCODE, Line-by-Line Radiative Transfer Model or LBLRTM is another code available to compute atmospheric transmissions. The information on LBLRTM can be accessed from `www.rtweb.aer.com`. This code was developed

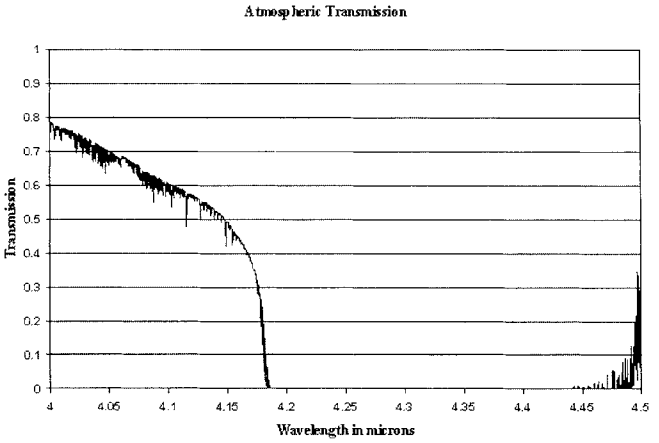


Fig. 3. Atmospheric transmission from 4 to 4.5 μm.

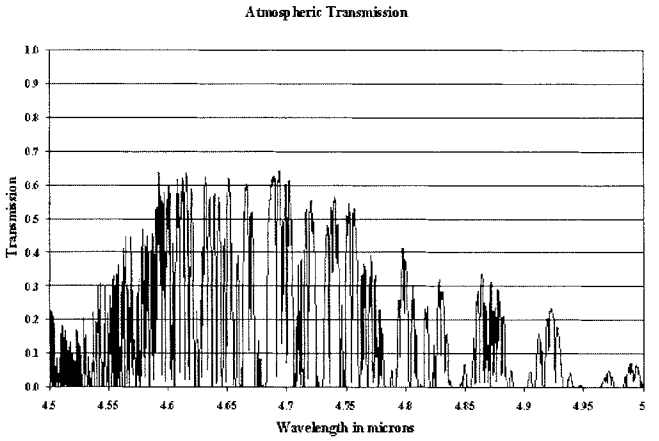


Fig. 4. Atmospheric transmission from 4.5 to 5 μm.

by Atmospheric and Environmental Research Inc. (AER) with the support of Department of Energy Developed on the legacy of FASCOD. The LBLRTM uses the HITRAN database and is considered to be efficient and highly flexible than FASCOD.

Plots of the one-way atmospheric transmission at a representative wavelength of 3.8 μm wavelength versus range, calculated using FASCOD3p using a Mid-Latitude Summer model with rural haze (5 and 23 km) visibility, no rain, and at sea level is shown in Fig. 5. The extinction profiles (extinction coefficient versus height) for 3, 4, and 5 μm wavelengths are shown in Fig. 6.

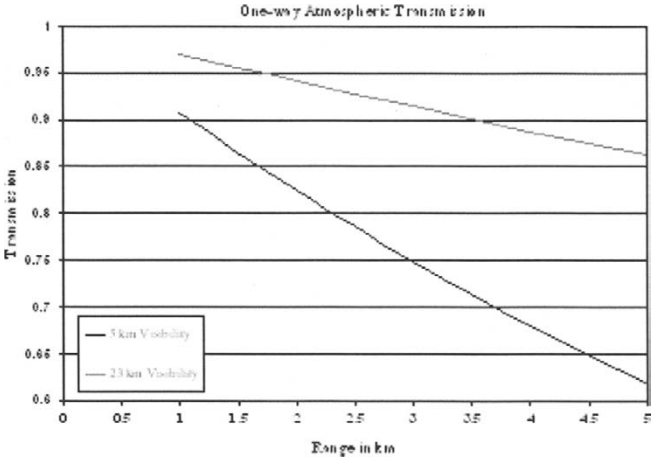


Fig. 5. One-way transmission at 3.8 μm for 5 and 23 km visibilities.

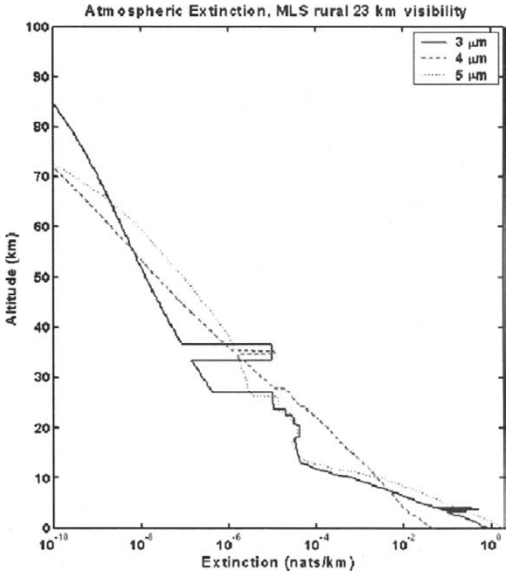


Fig. 6. Extinction profiles at 3, 4, and 5 μm.

2.1. *Atmospheric Turbulence*

Laser based optical sources generate Gaussian intensity beams. The TEM_{00} spatial mode is normally selected for applications involving beam propagation in free-space.

However, the Gaussian beam spreads and breaks up into beamlets as it propagates through the atmosphere due to turbulence. Atmospheric turbulence is the result of erratic air movements arising due to wind and convection processes driven by temperature differences. The effect of atmospheric turbulence has to be carefully considered when estimating the system performance. [2,3] The two parameters that characterize the effects of turbulence are Rytov variance, which describes scintillation, and coherence length, which affects beam spreading and determines the size of hot spots in the detector plane. Salient features of these parameters are discussed in the following sections.

Simulation of beam propagation through turbulence can be done numerically using the Fresnel approximation to the Huygens-Fresnel integral. [3–5] Two-dimensional fast Fourier transforms (FFTs) allows efficient implementation of Fresnel approximation through numerical means. To incorporate the effects of turbulence into the simulation, the propagation is done in multiple steps and turbulence is added after each step. The turbulence is simulated using phase screens. Phase screens are random realizations of phase noise that the beam encounters as it propagates. The phase screens are generated by assuming that the phase noise follows a specified atmospheric power spectral density (PSD). Spatially correlated Gaussian noise with the correct statistics must be created in order to properly generate phase screens to simulate propagation through turbulence. Simulations can be carried out using several different types of phase screens, namely, Kolmogorov, Tatarskii, Von Karman, and Hill phase screens.

In order to simulate the propagation of beam through atmospheric turbulence, a turbulence spectrum is selected to model the atmosphere. Then, random phase screens are generated and the phase values of the numerically sampled beam are changed by those values to simulate the turbulence. It can be shown that these phase screens distort the beam significantly over a path of few kilometers. Both truncated as well as untruncated Gaussian beams can be used for modeling. However, the numeric propagation of an untruncated Gaussian beam is very accurate, but the numeric propagation of a truncated Gaussian beam introduces slight errors in the field magnitude.

The beam distortion due to atmospheric turbulence simulated using Hill's phase screens for collimated untruncated Gaussian beams is illustrated in Fig. 7. For this simulation, a wavelength of 3.5 μm, transmission range of 2 km, and a C_n^2 value of $1E - 13$ m$^{-2/3}$ are assumed. It can be shown that beam distortion increases with turbulence.

2.1.1. Rytov Variance

The Rytov variance is a measure of the strength of scintillation. The scintillation is considered weak when the Rytov amplitude variance is less than 0.3. In the weak regime, where log-normal statistics apply, the normalized intensity variance [var/mean^2] or scintillation index is approximately four times the log-amplitude variance ($\sigma_I^2 \sim 4\ \sigma_\chi^2$), where $A = \exp[\chi]$ is the amplitude, χ is the log-amplitude and $I = A^2$ is the intensity. For a Kolmogorov spectrum and a spherical wave model, the log-amplitude variance is given by [1]

$$\sigma_\chi^2 = 0.56k^{7/6} \int_0^R C_n^2(r)\, w(r)\, dr, \tag{2}$$

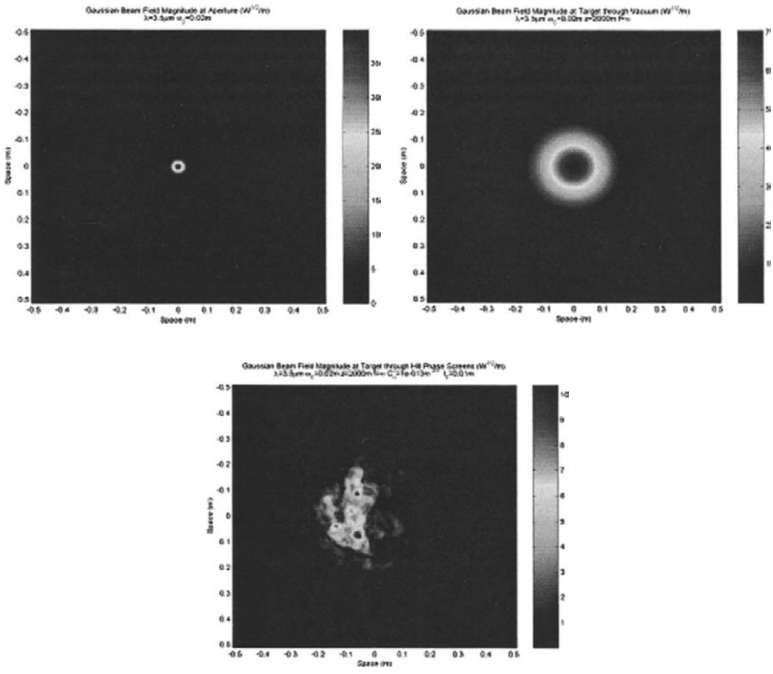


Fig. 7. The effect of atmospheric turbulence on beam propagation. Top left: The spatial beam profile of size 2 cm diameter at the transmitter aperture. Top right: The anticipated beam profile at a range of 2 km in vacuum (under no turbulence). Bottom: The beam distortion due to turbulence.

where $C_n^2(r)$ is the refractive index structure constant profile along the path, r and k= $2\pi/\lambda$. The weighting function, $w(r)$, is, for a spherical wave, independent of the path direction because it is symmetric about the middle of the path and is given by

$$w(r) = [r(1 - r/R)]^{5/6}. \tag{3}$$

For a constant C_n^2 and a spherical wave, Eq. (3) becomes

$$\sigma_\chi^2 = 0.1235k^{7/6}C_n^2R^{11/6}. \tag{4}$$

Figure 8 shows plots of the Rytov variance versus range parametric in C_n^2. Values of Rytov variance below 0.3 is usually indicate weak turbulence effects.

2.1.2. Coherence Length

The coherence length is a measure of the transverse dimension over which the electric field is coherent. It is the exp(-1) point of the electric field autocorrelation function. For a spherical wave and a Kolmogorov spectrum, the coherence length is given by

$$\rho_o = \left(2.91k^2 \int_0^R C_n^2(r)w(r)\,dr\right)^{-3/5}. \tag{5}$$

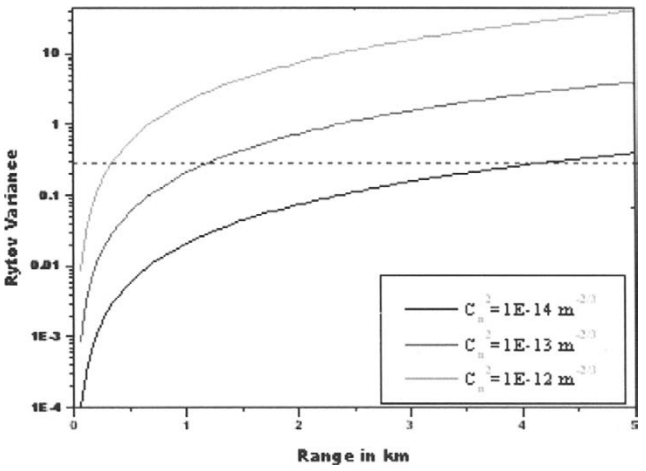


Fig. 8. Plots of Rytov variance as a function of range for different values of constant C_n^2. When the Rytov variance is below the dashed line at 0.3, the turbulence is considered weak.

The weighting function depends upon the path direction. For transmitter to receiver paths a reverse weighting function, $w_r(r)$ is used:

$$w_r(r) = (1 - r/R)^{5/3}. \tag{6}$$

This weighting function weights C_n^2 near the receiver more heavily than near the transmitter. The reverse path weighting is used to estimate the coherence length affecting the transmitted beam diameter at the receiver plane. A forward path weighting that weights C_n^2 near the transmitter more heavily than near a target is given by

$$w_r(r) = (r/R)^{5/3}. \tag{7}$$

Equation (7) is used to estimate the coherence diameter in a target plane. For a free space communication system (transmitter to receiver paths), the reverse weighting function is appropriate. For a constant C_n^2 and a spherical wave, however, Eq. (5) simplifies to

$$\rho_0 = \left[2.91k^2 C_n^2 R \frac{3}{8}\right]^{-3/5}. \tag{8}$$

Figure 9 shows plots of the coherence diameter (twice the coherence length or $2\times\rho_0$) versus range for difference values of constant C_n^2.

2.2. *Beam Diameter*

Taking into account near and far field effects as well as transmitter focus (F), and refractive turbulence, a Gaussian beam with initial beam diameter $D_b(0)$ has a beam diameter at range R given by

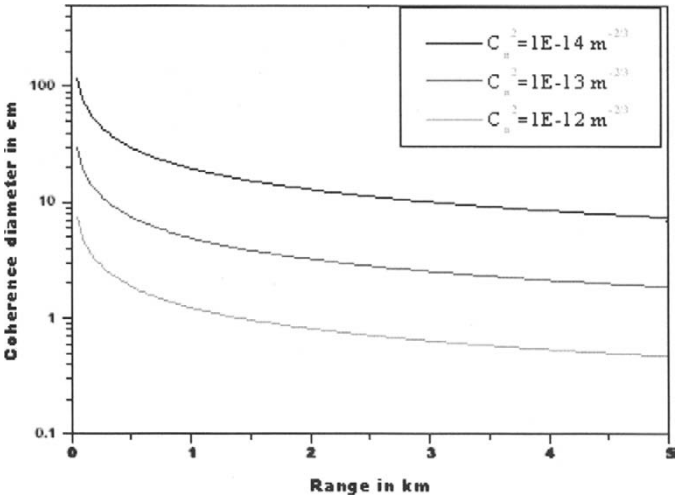


Fig. 9. Plots of the coherence diameter ($2\times\rho_0$) as a function of range for different values of constant C_n^2.

$$D_b(R) = 2\left[\left(\frac{D_b(0)}{2}\right)^2\left(1-\frac{R}{F}\right)^2 + \frac{R^2\lambda^2}{\pi^2}\left(\frac{4}{D_b^2(0)}+\frac{1}{\rho_0^2}\right)\right]^{1/2}. \quad (9)$$

The first term in the square brackets of Eq. (9) determines the near field effects and effect of focus on the beam diameter. The next term describes the far field effects and the last term describes the effects of coherence length (ρ_0) on the beam diameter. Thus, for a collimated ($F=\infty$) Gaussian beam, Eq. (9) becomes

$$D_b(R) = 2\left[\left(\frac{D_b(0)}{2}\right)^2 + \frac{R^2\lambda^2}{\pi^2}\left(\frac{4}{D_b^2(0)}+\frac{1}{\rho_0^2}\right)\right]^{1/2}. \quad (10)$$

When the coherence diameter ($2\rho_0$) becomes similar in size to the initial beam diameter, the beam diameter starts to increase more quickly than it would for a diffraction limited system where $\rho_0 = 0$.

Figure 10 shows plots of beam diameter versus range for collimated ($F=\infty$) Gaussian beams with 1 cm initial beam diameter propagating through various amounts of refractive turbulence. The beam diameter versus range for a diffraction limited collimated Gaussian beam is also plotted in each panel. Note that, for a refractive index structure constant of $C_n^2 = 1\times10^{-1}$, when the coherence diameter ($2\rho_0$) becomes larger than the initial beam diameter of 1 cm at a range of approximately 4 km the area of the beam is approximately doubled (i.e., the beam diameter is $\sqrt{2}$ larger) compared to that of the diffraction limited system. If the transmitted beam is focused at each range, the equation for the beam diameter becomes

$$D_b(R) = \frac{2R\lambda}{\pi}\sqrt{\left(\frac{2}{D_b(0)}\right)^2+\left(\frac{1}{\rho_o}\right)^2}. \quad (11)$$

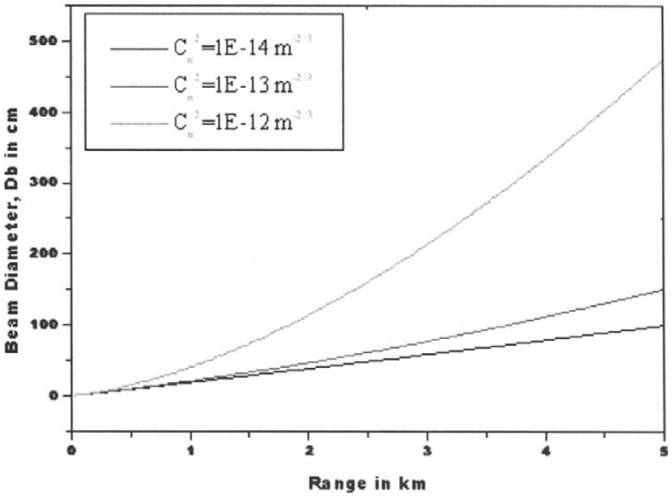


Fig. 10. Beam diameter versus range for collimated Gaussian beam under turbulence.

Note there is very little difference between a transmitted beam that is focused at each range and a collimated transmitted beam (focused at infinity). This is due to the fact that the Rayleigh range is around 50 m for a 1-cm beam diameter and around 0.5 m for a 1-mm beam diameter. Thus, most ranges in these cases are in the far field where the beam diameter will be affected more by diffraction and refractive turbulence than by focusing at the aperture.

2.3. Direct Detection SNR

Estimates of signal to noise ratio (SNR) for each range can be found by using noise models for a particular detector (e.g. an avalanche photodiode) and combining these with the calculated received power $P_r(R)$ in the form

$$\mathrm{SNR}\,(R) = \frac{[\Re P_r\,(R)\,M]^2}{[(\Re P_r\,(R) + \Re P_{bg} + I_{db})\,M^2 F + I_{ds}]\,2qB_n + N_{\mathrm{SD}} B_n}, \tag{12}$$

where $\Re$ is the detector responsivity given by

$$\Re = \frac{\eta_q q}{h\nu} \tag{13}$$

and $q = 1.602 \times 10^{-19}$ J is the electron charge, $h = 6.626 \times 10^{-34}$ Js is Planck's constant, and ν is the frequency of radiation.

The received power at range R, incident on an aperture of diameter D_r can be estimated by

$$P(R)_r = P_x \eta T\,(R)\,\eta_O(R), \tag{14}$$

where P_x is the power of the transmitted beam, $T(R)$ is the atmospheric transmission at range R, η is the combined transmitter/receiver optical efficiency, and η_O is a range dependent overlap (between detector area and beam area) efficiency given by

$$\eta_O = 1 - \exp\left[-2\left(\frac{D_r}{D_b(R)}\right)^2\right]. \tag{15}$$

When the detector diameter, D_r, is significantly smaller than the beam diameter at range $R(D_r \ll D_b)$, the argument of the exponential in Eq. (15) becomes much smaller than unity. Thus the approximation

$$\exp(-x) \approx 1 - x \quad \text{for} \quad x \ll 1 \tag{16}$$

can be used to find an approximation for received power incident on the detector:

$$P_r \approx P_x \frac{\eta T(R)}{R^2} \frac{\pi D_r^2}{4\Omega_x}, \tag{17}$$

where

$$R^2 \Omega_x = A_x r = \frac{\pi D_b^2(R)}{4} \tag{18}$$

is the area defined by the diameter of the transmitted beam at range R.

2.4. *Large Scale and Small Scale Turbulence—Spherical Waves*

The scintillation index, defined by (pp. 236–237 of [1]):

$$\sigma_I^2(r,R) = \frac{\langle I^2(r,R)\rangle}{\langle I(r,R)\rangle^2} - 1 = \frac{\Gamma_4(r,r,r,r,R)}{\Gamma_2(r,r,R)\,\Gamma_2(r,r,R)} - 1, \tag{19}$$

where $I(r,R)$ is the irradiance at point r in the field at range R, $\Gamma_4(r_1,r_2,r_3,r_4,R)$ is the fourth-order coherence function of the field at R, and $\Gamma_2(r_1,r_2,R)$ is the mutual coherence function of that field. The irradiance at a given range can be modeled as a product of large scale (x) and small scale (y) irradiance fluctuations written as

$$I = xy, \tag{20}$$

where the large scale and small scale irradiance fluctuations have gamma distributions given by

$$\begin{aligned} p_x(x) &= \frac{\alpha(\alpha x)^{(\alpha-1)}}{\Gamma(\alpha)} \exp(-\alpha x) \quad x > 0 \\ p_y(y) &= \frac{\beta(\beta y)^{(\beta-1)}}{\Gamma(\beta)} \exp(-\beta y) \quad y > 0 \end{aligned}. \tag{21}$$

The overall irradiance will therefore have a gamma-gamma PDF, written as

$$p_I(s) = \frac{2(\alpha\beta)^{\frac{(\alpha+\beta)}{2}}}{\Gamma(\alpha)\,\Gamma(\beta)\,\langle i_s\rangle}\left[\frac{s}{\langle i_s\rangle}\right]^{\frac{(\alpha+\beta)}{2}-1} K_{\alpha-\beta}\left(2\sqrt{\frac{\alpha\beta s}{\langle i_s\rangle}}\right) \quad s > 0. \tag{22}$$

The α and β parameters in the distribution(s) can be identified with large-scale and small-scale scintillation as

$$\sigma_I^2 = \frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\alpha\beta} = \sigma_x^2 + \sigma_y^2 + \sigma_x^2\sigma_y^2, \tag{23}$$

where σ_x^2 describes the large scale scintillation and σ_y^2 describes the small-scale scintillation. Reference [2] presents the process for finding a form for the scintillation index dependent only on the Rytov variance. The result is

$$\sigma_I^2 = \exp\left[\frac{0.49B_0^2}{\left(1+0.56B_0^{12/5}\right)^{7/6}} + \frac{0.51B_0^2}{\left(1+0.69B_0^{12/5}\right)^{5/6}}\right] - 1, \tag{24}$$

where

$$B_0^2 = 0.494k^{7/6}C_n^2R^{11/6} \tag{25}$$

is a strength of turbulence parameter approximately equal to 4 times the Rytov variance defined in Eq. (2).

Contributions to scintillation from small-scale fluctuations are related to turbulent cells that have a size less than the smaller of either the first Fresnel zone, $L_{\mathrm{Fz}} = \sqrt{R/k}$, or the coherence radius, ρ_0. Contributions to scintillation from large-scale fluctuations, however, are related to turbulent cells that have a size greater than the larger of either the Fresnel zone or the "scattering disk", where the scattering disk is defined as

$$L_{\mathrm{SD}} = \frac{R}{k\rho_0}. \tag{26}$$

The plots of all three of these parameters (coherence length, scattering disk, and Fresnel zone). help determine the contributions of different types of cells to either large scale or small scale fluctuations for a given refractive turbulence parameter. Thus, turbulent cells that are smaller than the coherence length will contribute most to the small-scale scintillation at these ranges. If the scattering disk is greater than the Fresnel zone, it indicates that contributions to large-scale scintillations will come from cells that are larger than the scattering disk. The Fresnel zone is the limiting factor in determining which turbulence cells contribute to large-scale scattering. One can see that large-scale scintillation has a greater influence (more cell sizes will contribute) in situations with smaller refractive turbulence structure constants. We see these effects when the mean SNR for small aperture diameters is smaller for low refractive turbulence than it is for higher refractive turbulence.

2.5. Bit Error Rate (BER) and Minimum SNR

The Bit Error Rate (BER) is a parameter that is often used to characterize the performance of a communication system. In this section, the relationship between BER and receiver SNR for different receiver aperture diameters and different refractive index structures, C_n^2 is discussed. The main conclusion to draw from this section is that the mean SNR at a given range and given C_n^2 is highly dependent on the receiver aperture diameter and the scintillation index. The scintillation index for each range in turn depends on C_n^2 and receiver aperture diameter. Hence, care must be taken in choosing the correct combination for the particular range(s) of interest. In general, however, increasing the receiver aperture diameter will improve mean SNR unless the refractive turbulence is high (e.g., $C_n^2 = 1 \times 10^{-12}$) in which case the receiver SNR is low at most ranges and so the mean SNR saturates with increasing receiver aperture diameter.

The bit error rate (BER) or probability of error in an on-off keying diffraction limited communication system is given by

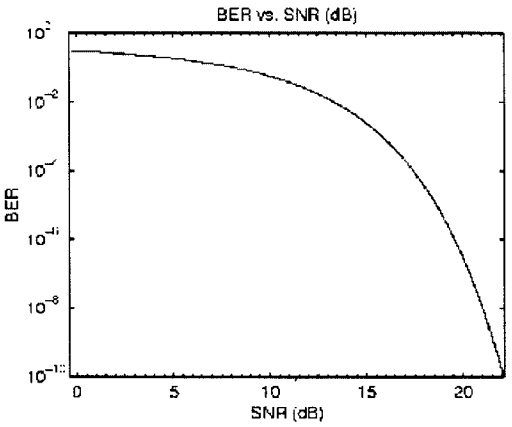


Fig. 11. Plot of theoretical bit error rate (BER) vs. SNR.

$$\mathrm{BER} = \frac{1}{2}\mathrm{erfc}\left(\frac{\mathrm{SNR}\,(R)}{2\sqrt{2}}\right). \tag{27}$$

For a maximum required BER of 1×10^{-9}, the minimum required SNR is about 21 dB as can be seen in Fig. 11.

When refractive turbulence is present in the communication system, the probability of error must be averaged over the probability distribution function of the signal, s. Thus the BER becomes (pp. 236–237 of [2]):

$$\mathrm{BER} = \frac{1}{2}\int_0^\infty p_I\,(s)\mathrm{erfc}\left(\frac{\langle \mathrm{SNR}\rangle^{1/2}\, s}{2\sqrt{2}\,\langle i_s\rangle}\right) ds. \tag{28}$$

where the probability density function, $p_I(s)$, is given by a gamma-gamma distribution:

$$p_I\,(s) = \frac{2\,(\alpha\beta)^{\frac{(\alpha+\beta)}{2}}}{\Gamma\,(\alpha)\,\Gamma\,(\beta)\,\langle i_s\rangle}\left[\frac{s}{\langle i_s\rangle}\right]^{\frac{(\alpha+\beta)}{2}-1} K_{\alpha-\beta}\left(2\sqrt{\frac{\alpha\beta s}{\langle i_s\rangle}}\right) \quad s > 0. \tag{29}$$

and $\mathrm{K}_\nu()$ is a modified Bessel function of the second kind. The α and β parameters for the gamma distributions are found using:

$$\begin{aligned} \frac{1}{\alpha} &= \exp\left[\frac{0.49\beta_0^2}{\left(1+0.18d^2+0.56\beta_0^{12/5}\right)^{7/6}}\right] \\ \frac{1}{\beta} &= \exp\left[\frac{0.51\beta_0^2\left(1+0.69\beta_0^{12/5}\right)^{-5/6}}{\left(1+0.90d^2+0.62d^2\beta_0^{12/5}\right)}\right], \end{aligned} \tag{30}$$

where

$$B_0^2 = 0.494k^{7/6}C_n^2R^{11/6} \tag{31}$$

is a strength of turbulence parameter approximately equal to 4 times the Rytov variance, and the term d is dependent on the aperture diameter D_r and range, R, according to

$$d = \sqrt{\frac{kD_r^2}{4R}}. \tag{32}$$

The parameters α and β can be thought of as representing the effective number of large-scale cells and small-scale cells, respectively. The mean SNR, $\langle \mathrm{SNR} \rangle$, is determined by

$$\langle \mathrm{SNR} \rangle = \frac{\mathrm{SNR}}{\frac{P_{\mathrm{rd}}}{P_{\mathrm{sr}}} + \sigma_I^2 (D_r)\,\mathrm{SNR}}, \tag{33}$$

where SNR is the receiver SNR defined previously and P_{rd}/P_r is a power reduction factor consisting of the ratio between the received power in a diffraction limited system and the received power when optical turbulence is present. This ratio is given by

$$\frac{P_{\mathrm{rd}}}{P_{\mathrm{sr}}} = \left(1 + 1.63\sigma_x^{12/5} \frac{2R}{k\omega^2 (R)}\right), \tag{34}$$

where

$$\omega^2 (R) = \omega_0^2 + \frac{R^2\lambda^2}{\pi^2}\left(\frac{1}{\omega_0^2} + \frac{1}{\rho_0^2}\right) \tag{35}$$

is square of the beam radius at the receiver. The scintillation index σ_I^2, in Eq. (33) is given, for a spherical wave, by

$$\sigma_I^2(D) = \exp\left[\frac{0.49\beta_0^2}{\left(1 + 0.18d^2 + 0.49\beta_0^{12/5}\right)^{7/6}} + \frac{0.51\beta_0^2\left(1 + 0.69\beta_0^{12/5}\right)^{-5/6}}{1 + 0.90d^2 + 0.47d^2\beta_0^{12/5}}\right] - 1. \tag{36}$$

In summary, the $\langle \mathrm{SNR} \rangle$ of a received signal is dependent on the scintillation index of that signal which in turn is dependent on the refractive index structure constant, C_n^2, the receiver aperture diameter, and the range to the receiver. A smaller refractive index structure constant ($C_n^2 = 1 \times 10^{-14}$) along the path will result in breakup of the beam into larger sections that may miss a small receiver aperture altogether. A larger refractive index structure constant ($C_n^2 = 1 \times 10^{-12}$), however, will break the beam up into smaller sections, increasing the probability that there will be signal on the detector at a given time, thus increasing the $\langle \mathrm{SNR} \rangle$ and decreasing the BER. If the aperture diameter is increased, however, the probability that the large beam sections (from $C_n^2 = 1 \times 10^{-14}$ along the path) increases and the $\langle \mathrm{SNR} \rangle$ becomes larger than that for a beam that is more broken up (from $C_n^2 = 1 \times 10^{-12}$ along the path). Another way to look at this is to notice that the scintillation index for a beam propagating through strong refractive turbulence ($C_n^2 = 1 \times 10^{-12}$) is fairly insensitive to an increase in receiver aperture diameter whereas the scintillation index for a beam propagating through weaker refractive turbulence ($C_n^2 = 1 \times 10^{-14}$) is highly sensitive to increases in receiver aperture diameter.

The C_n^2 and hence the scintillation index can be measured using a scintillometer. In a typical arrangement, the scintillometer consists of a transmitter and a receiver. These two units are placed at a convenient distance say 1 km at a site where the C_n^2 has to be measured. Figure 12 shows a typical measurement in which the refractive index structure constant, $C_n^2$was measured using a scintillometer at 1 Hz rate. In this figure, the C_n^2 level remains between 1×10^{13} and 2×10^{13} $\mathrm{m}^{-2/3}$ during the times shown.

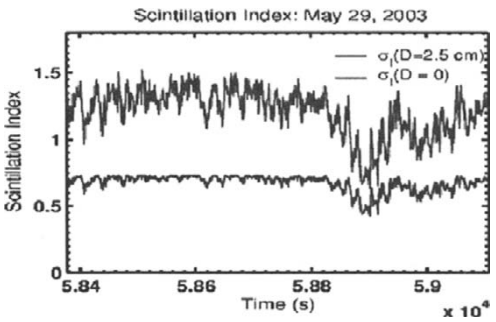


Fig. 12. The C_n^2 measurements using a scintillometer is shown. The computed scintillation index using Eq. (19) for a time interval (ΔT) of 8 seconds is also plotted.

3. The MWIR Optical Sources

Following the selection of an appropriate candidate wavelength or wavelength windows, the hardware availability has to be carefully considered. The MWIR wavelengths are considered as eyesafe. [7] The source characteristics have to be blended into the choice of an optimum operational wavelength. Often, this could be an iterative process. An optical transmitter generally consists of an optical source, drive electronics, and an output telescope or beam shaping optics. Sometimes scanners could be a part of the transmitter unit. Lasers and nonlinear optical devices such as optical parametric oscillators can be utilized for the generation of MWIR wavelengths. Due to the physical nature of operation, lasers generate discrete wavelengths in the MWIR region with very limited tunability. Some of the considerations that go into selection process of an optical source include overall output power (average or peak), CW or pulsed mode of operation, spectral width, reliability, size, weight, operational elegance, environmental compatibility and cost. In the case of pulsed operation, PRF, pulsewidth and pulse energy parameters have to be looked into. In this section, a brief description of different types of MWIR sources is presented.

The most common MWIR source is the Helium Neon laser. The Helium Neon gas laser operating at 3.39 μm in CW mode is commercially available from several vendors. However, Helium Neon gas lasers can only provide limited output powers (tens of milliwatts) and furthermore they are bulky and less versatile. These lasers can be externally modulated using acousto-optic or electro-optic modulators. For pulsed operational requirements, theses lasers are of limited use. Hence these lasers are generally used in laboratory experiments. Another laser type is XeHe that generates wavelengths over 2 to 4 μm band with a prominent emission at 3.506 μm. However, so far XeHe lasers have not been practical as far as applications are considered.

3.1. Semiconductor Based Lasers

Semiconductor-based lasers are also utilized to generate the MWIR wavelengths. In this category, there are three classes of devices namely, lead salt lasers, III–V strained

quantum-well lasers, and quantum cascade lasers. Small size, wavelength diversity and ease of modulation are primary advantages. Disadvantages of these devices typically include cryogenic operation, low duty cycles, and low efficiencies. However, the technology of these devices is slowly maturing that minimize these disadvantages. Each of these classes of devices is discussed in greater detail in the following sections.

3.1.1. Lead-Salt Lasers

Lead salt lasers are based on IV–VI semiconductor materials. They are so called because majority of these compounds contain lead. These lasers operate at wavelengths ranging from 3.3 to 29 μm. Devices operate exclusively at sub-ambient temperatures. Typical operation is at cryogenic temperatures with average output powers of tens of milliwatts. [8–10] Limited pulsed operation of a MWIR device has been observed at room temperature. However, robust operation of devices in the 3–10 μm wavelength range is restricted to liquid nitrogen temperatures. Average output powers can be as high as tens of milliwatts. However, device efficiencies are well under 1%. Efficiencies are limited by low internal quantum efficiencies and high material losses. Room temperature lead salt laser research is in the exploratory stage.

3.1.2. III–V Strained Quantum-Well Lasers

Quantum well lasers (QWL) are commonly used in the Short-Wave InfraRed (SWIR) telecommunications windows at 1.31 and 1.55 μm. They are less developed and less robust at MWIR wavelengths with longer wavelength CW operation limited to cryogenic temperatures. The longest reported emission wavelength of a III–V QWL is 4.5 μm. [11] The device lased in pulsed mode at temperatures up to 144 K and in CW mode at temperatures up to 105 K. With electrical pumping, cw output powers of 30 mW were produced at 70 K. Electrical-to-optical efficiencies were less than 1%. The development of MWIR QWLs is difficult because the valence-band offset and refractive index separation between the active and barrier layers must be made smaller for operation at longer wavelengths. Growing devices with appropriate characteristics is difficult. This limitation may be due to non-uniform carrier distribution in the quantum wells at elevated temperatures.

3.1.3. Quantum Cascade Lasers

Quantum cascade lasers (QCLs) are fundamentally different from diode lasers. [12–13] Diode lasers generate light when conduction band electrons and valance band holes radiatively recombine across a band gap that exists because of the presence of discrete energy bands in the semiconductor. QCLs generate light when electrons (only) make transitions between conduction band states (no valance band) arising from size quantization in semiconductor heterostructures. QCL devices are created using molecular beam epitaxy and band structure engineering. QCL are waveguide devices with core and cladding regions. Continuous-wave (CW) operation of QCLs is limited to cryogenic temperatures. Pulsed devices have been demonstrated at room temperature. These devices output nsec duration pulses at kHz to MHz repetition frequencies. QCLs operate better at lower temperatures, since (1) laser gain decreases

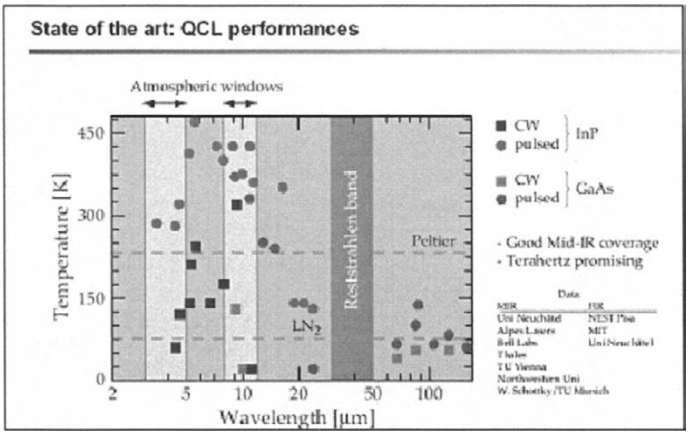


Fig. 13. The state-of-art performance of QCL devices (Courtesy of Alpes lasers).

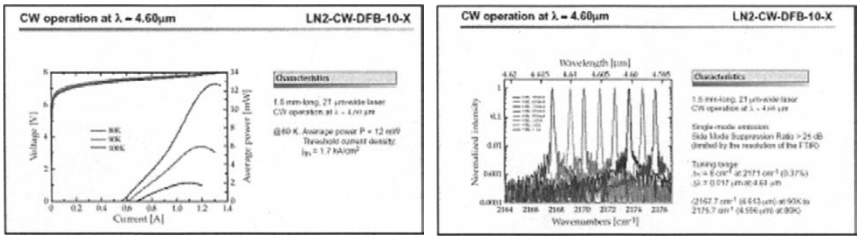


Fig. 14. The MWIR QCL laser marketed by Alpes lasers.

with increasing temperature due to an increase in the width of the gain spectrum, (2) the laser's upper state lifetime also decreases with increasing temperature due to the presence of larger numbers of optical phonons at elevated temperatures, and (3) the thermal population of the laser's ground state increases with increasing temperature. QCLs tend to have high thresholds (compared to diode lasers) due to high waveguide losses. The origin of these losses is not well understood. Typical electrical-to-optical efficiencies for QCLs are approximately 0.1%.

Currently, Alpes Lasers from Switzerland (`www.alpeslasers.ch`) is a leading manufacturer of QCL lasers. Pulsed DFB QCL on Peltier cooler have been commercialized for wavelengths in the range 4.3 to 16.5 μm as illustrated in Figs. 13 and 14.

3.1.4. Solid-State Lasers

There are a variety of solid-state laser options in the MWIR spectral region. Available and potential sources in both the MWIR spectral region are reviewed below.

There are few demonstrated solid-state lasers in the MWIR. None are tunable. Reported systems include a cryogenic Er:YLF laser at 3.41 μm, [14] a room temperature Er:ZBLAN fiber laser at 3.5μm,[15] a Ho:ZBLAN fiber laser at 3.9 μm [16] and cascade laser emission in Ho:YLF at 2.9, 3.4, and 3.9 μm [17] These systems and several potential alternatives are reviewed below.

The Er:YLF work was done at 77 K and the laser had a 2% slope efficiency. The Er:ZBLAN laser was demonstrated at both reduced and room temperatures. Room temperature output powers were a couple of milliwatts for pump powers, at 488 nm, approaching 1 W. The probability of adapting this fiber laser in a hardened package is low. Problems include the lack of a convenient demonstrated pump source; the fragility of ZBLAN fiber; and difficulties in stripping, cleaving, and generally working with ZBLAN fiber. Additionally, the Er:ZBLAN laser is expected to offer lower overall optical efficiency than some of the nonlinear frequency conversion techniques discussed in the next section.

The Ho lasers are cascade devices that function by lasing multiple transitions simultaneously. The YLF work reported stimulated emission but laser energies were too low to measure. The fiber work is more intriguing but has only been demonstrated at cryogenic temperatures (below 150 K). Laser power at 77 K was limited to 11 mW for 880 mW of launched pump power. The laser was pumped by either a dye laser at 640 nm or a Ti:Sapphire laser at 890 nm. In our opinion, these lasers have limited use at room temperature due to low fluorescence quantum efficiencies.

A promising possibility in this wavelength range is to lase a Co doped isomorph of Cr:ZnSe. Two interesting possibilities are Co:ZnSe and Co:ZnTe. The former has a broad emission extending from 2.5 to 4.0 μm and the latter from 2.7 to beyond 4.2 μm. The materials also possess broad absorption bands in the 1.4–2.0 μm wavelength range. Room temperature fluorescence quantum efficiencies are approximately 25% in Co:ZnSe and 8% in Co:ZnTe. These efficiencies substantially increase as temperatures are reduced to 200 K.

3.1.5. Chemical Lasers

In the chemical laser category, DF and HF based lasers are available. [9] The DF laser provides wavelengths over a range of 3.5 to 4.2 whereas the HF laser provides wavelengths in from 2.6 to 3.3 μm. Large size and toxic nature are major limitations of these laser systems.

3.2. Nonlinear Frequency Converters

Three-wave mixing (TWM), also known as the optical parametric interaction, is an alternate process to generate MWIR wavelengths through light-matter interaction. [18–20] The TWM process basically requires an intense laser pump source and a nonlinear medium. In this process, two propagating monochromatic waves with frequencies ω_1 and ω_2 mutually interact to generate new frequency through sum (sum frequency generation or SFG) or difference (difference frequency generation or DFG) processes. The relation between polarization density and the electric field vectors describes light and medium interaction. Three wave mixing (TWM) is possible through the second order nonlinearity, $\mathbf{P}_{\mathrm{NL}} = 2d\mathbf{E}^2$; d = 2nd order nonlinear coefficient and $\mathbf{E}$ is the electric

field vector. Optical parametric amplification (OPA), Optical parametric oscillation (OPO) and frequency up-conversion are examples of TWM. In principle, all three processes can be used in the generation of MWIR wavelengths. The second harmonic generation is a special case of SFG. In the case of second harmonic generation, two pump photons of frequency ω_o interact inside a suitable nonlinear medium to generate a photon of frequency $2\omega_o$. In the wavelength domain, two longer wavelength photons are combined to generate a shorter wavelength photon. However, second harmonic generation devices are not attractive for generating MWIR because of the lack of suitable small and compact Long-Wave InfraRed (LWIR) laser pump sources. To generate 4 μm wavelength we need a laser pump source operating at 8 μm. One possibility is to use CO_2 gas laser generating wavelengths in the vicinity of 10.6 μm. However, these lasers are bulky and impractical for many applications. In the case of frequency mixing, two pump beams usually of different frequency have to interact inside a nonlinear medium to obtain either the sum or difference of the input frequencies. Again, similar to second harmonic generation sum frequency is not practically viable due to source limitations. However, difference frequency mixing (DFM) is more viable using readily available short wavelength sources. The main disadvantage is that this technique requires two pump sources and hence becomes less practical. The optical parametric oscillator (OPO) is the opposite of SFG and is considered most versatile and practical means to generate MWIR wavelengths among the above three types of nonlinear interaction and hence is described in greater detail in the following section.

An OPO generates broadly tunable coherent optical radiation over a wide wavelength range via second order nonlinear process called three-wave-mixing in certain non-centrosymmetric crystals. Near-UV to LWIR wavelengths can be generated using OPOs. In an OPO, the intense pump photon generates two higher wavelength photons when frequency and phase matching conditions are simultaneously satisfied. These conditions are $\omega_3 = \omega_1 + \omega_2$ and $\mathbf{k}_3 = \mathbf{k}_1 + \mathbf{k}_2$; ω = frequency, $\mathbf{k}$ = momentum vector, and 1,2 and 3 correspond to pump, signal and idler wavelengths, and can be satisfied in certain anisotropic crystals. Figure 15 shows the schematic of an OPO. For nonlinear interaction to occur, an OPO requires a strong pump source whose intensity is above a certain oscillation threshold value. Intense pump radiation can be achieved from a suitable Q-switched laser.

The OPOs can be normally pumped using pulsed solid-state laser sources. Generally, the Q-switched pulsed lasers provide sufficient intensities to initiate parametric interaction in many OPO materials to overcome the oscillation threshold. Besides pulsed laser pumping, OPOs can also be pumped using lasers operating in the CW mode. [21] OPOs can also be directly diode pumped provided high power diodes and/or materials with high nonlinear figure of merit (D_{eff}) are available. [22] Normally, indirect diode pumping of OPO are adopted due to limitations of pump laser diode power in single emitter configuration, power supply constraints, overall size, and cost considerations. With the advancement in laser diode technology, direct pumping using laser diodes could become attractive. The advantage of diode pumping is that the diode can be modulated and the OPO output will be similarly modulated. Additionally, diode pumping enables the construction of small and compact transmitters.

Above threshold, the pump radiation generates two wavelengths called signal and idler when phasematching conditions are satisfied. The signal wavelength is less than or equal to idler wavelength. The phase matching conditions are derived from fundamental energy and momentum conservation laws. For phasematching to occur, the crystal

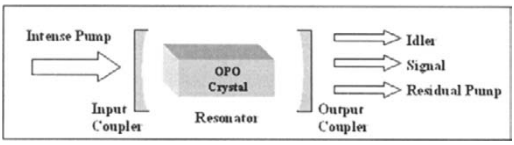


Fig. 15. The schematic of an OPO.

has to be cut at a phasematching angle to obtain desired wavelengths under normal operating conditions. Without proper phasematching a laser beam propagating through a nonlinear crystal will continue to transfer energy back and forth between the pump beam and the signal/idler beam pair. The result is that no net energy conversion takes place.

The OPOs are intensity driven devices. Although the OPO output is at two wavelengths, it offers zero quantum defect. Furthermore, the device gain is independent of tuning. The pump intensity has to exceed the oscillation threshold. Large $D_{\text{effective}}$, and doubly resonant oscillation (DRO) reduces the oscillation threshold and increases conversion efficiency. The crystal selection depends on its transparency range, nonlinear coefficient, (higher the better), phasematching capability, availability in appreciable sizes, optical quality, damage threshold (higher the better), commercial availability and cost. The input and output couplers are achieved by having dielectric or trichroic coatings of desired bandwidth and transmittance on suitable mirror substrates. Broad tunability allows an arbitrary wavelength to be accessed.

In bulk crystals, a unique orientation of a birefringent crystal can be chosen for phasematching purposes. The birefringent phasematching can be further sub-classified as either Critical or Non-Critical Phase Matching (CPM or NCPM). The CPM technique offers large wavelength tunability through angle rotation. However, for achieving optimal OPO efficiency, the limitations imposed by low angular acceptance and large walkoff parameters have to be carefully considered. The NPCM offers high angular acceptance and very low walkoff with limited wavelength tunability [23]. Phase matching can be either Type I or Type II depending on the polarization of three beams. With reference to the frequency matching condition, $\omega_p = \omega_s + \omega_i$, a particular combination of the polarizations of pump, signal and idler wavelengths are allowed. In the case of a negative uniaxial crystal, Type I phasematching refers to $e = o + o$ and Type II phase matching refers to $e = e + o$ or $o + e$, where "e" is the extra-ordinary ray and "o" is the ordinary ray. This is reversed in the case of a positive uniaxial crystal. Wavelength tuning is achieved by varying refractive index via crystal rotation (angle tuning), temperature, or electro-optic means.

Alternatively, a material can be engineered so that light is re-phased every coherence length in the crystal. The former technique is referred to as birefringent phasematching and the later technique is known as Quasi-Phase Matching (QPM). [24,25] QPM based materials were selected instead of bulk materials since they can provide higher nonlinear coefficient, D_{eff} than corresponding bulk materials. Other advantages of QPM materials include zero walkoff, and high angular acceptance. The QPM material is fabricated from wafer of a bulk material by "periodically poling" the ferroelectric domains. Effectively, this changes the sign of the nonlinear coefficient, D_{eff},

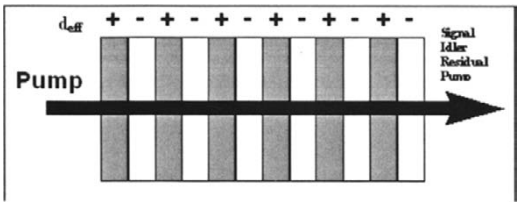


Fig. 16. Schematic of a QPM material.

for every coherence length and is illustrated in Fig. 16. The alternating dark and light regions indicate a change in the sign of the nonlinear coefficient. Pump light propagates through the crystal perpendicular to the poled regions. The domain (also known as grating) width determines the idler and signal wavelength for a given pump wavelength. The QPM materials are engineered to provide relatively high D_{eff} convenient for TWM operation which otherwise is difficult to achieve in bulk materials. The output wavelengths can be conveniently tuned either by varying the OPO crystal temperature or by mechanical translation of the crystal over varying grating periods. The former one provides fine tuning of wavelengths where as the latter allows for course change in wavelengths. A commonly used QPM material to generate near IR and MWIR wavelengths is the periodically poled lithium niobate (PPLN) material. Other recently developed materials include periodically poled KTA and periodically poled RTA.

The anisotropic OPO crystal is located inside a resonator. For OPO operation, idler and/or signal and/or pump can be resonated. An OPO can be operated as a Singly Resonant Oscillator (SRO) or a Doubly Resonant Oscillator (DRO. In a SRO, either signal or idler is resonant. In a DRO both the signal and idler wavelengths are resonant. In general, the output coupling could vary from 0% to close to 100%. A DRO with 100% output coupling at either wavelength is generally called an SRO. The optimal oscillator design is determined by the application. SROs have higher spectral stability than DROs because DROs suffer from mode "cluster hopping". This phenomenon is manifested by frequency hops from one pair of resonant (signal and idler) frequencies to another spectrally close pair. The underlying cause is that a cavity length that is optimized for the signal frequency is unlikely to be simultaneously optimized at the idler frequency. The result is small frequency hops as the OPO tries to find the optimum (highest gain) axial mode pairs to operate at. For the same reason, SROs are easier to wavelength tune than DROs and require mirrors with less coating constraints. As such, SRO mirrors cost less. The main advantage of a DRO is that it will have a lower power threshold than an SRO. The pump radiation itself can be resonated in both cases. If all three wavelengths are resonated, the OPO is known as triply resonant oscillator (TRO). TRO configuration, in general, offers lower oscillation threshold but present practical difficulties in the fabrication of resonator mirrors. Besides having an OPO resonator arrangement external to the laser resonator, it can be operated in an intracavity configuration by placing the OPO crystal inside the laser resonator. [26]

OPOs operating in the MWIR spectral region have been widely demonstrated. [27–28] The OPOs with various cavity configurations can be developed to generate CW or pulsed output. In one configuration, the lasing and OPO interaction was accomplished

inside a single nonlinear medium namely Nd:MgO:$LiNbO_3$. [29] Additionally, a variety of nonlinear gain media may be used depending on the pump wavelength chosen and desired device attributes. Practical pump sources include Nd:YAG laser (1.064 μm), Er:YAG laser (1.54 μm), and Tm:Ho:YLF (2.06 μm) and their designs are discussed in [30]. The potential nonlinear media for generating mid-wave IR (MWIR), 3-5 μm wavelengths are $LiNbO_3$, periodically poled $LiNbO_3$ (PPLN), RTA and KTA. For output at 4 μm OPOs are typically pumped around 1 μm due to the availability of commercial Nd:YAG laser pump sources. In pulsed mode, these systems typically output nanosecond duration pulses due to the short pulse duration of the Nd:YAG pump lasers. Finally, OPOs can outperform state-of-the-art QCL devices in peak power, room temperature operation, wavelength agility, size and cost.

3.2.1. Performance Modeling of OPOs

To start with, the performance of an OPO has to be modeled. The mathematical relations for the estimation of various OPO parameters are described in the literature [31]. The performance modeling is an iterative process based on the selection of a suitable pump source wavelength and the OPO material characteristics. For performance modeling, the OPO material, pump wavelength, OPO material characteristics including length and dispersion characteristics and resonator parameters have to be determined. In the beginning, phase matching curves (or tuning curves) are generated for a given pump wavelength by using the phase matching conditions, Snell's law and corresponding Sellmeyer's equations for the refractive indices for a given OPO material. Based on the D_{eff}, Type 1 or Type II operation is selected. Note that larger the D_{eff}, lower the oscillation threshold and hence the pump intensity requirement. The crystal cut angle is then determined for nominal output wavelength operation. Based on walkoff considerations, optimum OPO crystal length is determined. Most of the time, the selection of an OPO crystal and its dimensions are dictated by its commercial availability, procurement time and cost. The pump beam is modematched to the OPO resonator mode. The mirror closest to the laser pump is designated as the input coupler and the farthest mirror as the output coupler. The OPO resonator mirrors are designed based on the SRO or DRO configuration. In the case of a SRO OPO, the input coupler has to be AR coated at the pump wavelength and highly reflective at the signal and idler wavelengths. Also, the output coupler has to be highly reflective at the resonant wavelength. Based on the source and OPO material characteristics and the transmission characteristics of resonator mirrors, the oscillation threshold intensity is computed. The total conversion efficiency then depends on the pump intensity above threshold. The OPO design details can be found in [20] and [32]. In the following subsection, salient features of PPLN OPO for MWIR wavelength generation design are discussed.

3.2.2. PPLN OPOs

PPLN is a commonly used material for MWIR OPO operation. The major advantages of this material include: (1) offers good optical transmission from approximately 0.35 to 4.2 μm wavelengths, (2) provides a relatively high nonlinear figure-of-merit (D_{eff} ~17 pm/V), (3) offers low optical loss at most wavelengths shorter than 4 μm, and (4) allows phasematching with visible to near-IR wavelengths where robust laser sources

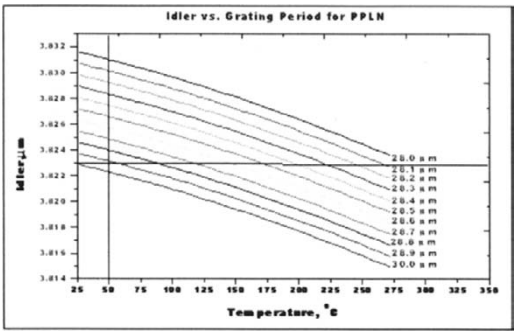


Fig. 17. The dependence of grating period on temperature at pump wavelength of 1064 nm. The grating period selected for operation at 3.82 μm is shown.

are available. As discussed above, the PPLN material is fabricated by "periodically poling" the ferroelectric domains of a bulk wafer of $LiNbO_3$. Up to 1 mm thick wafers and crystal lengths up to 5 cm are commercially available. Several examples of PPLN OPOs being used to generate 4 μm light have been reported in the literature. [21,24,29] In the following section, PPLN OPO for generating wavelengths around 3.82 μm is discussed. Note that 3.82 μm is one of several low loss sub-bands.

PPLN OPO Design Considerations. To obtain a given MWIR wavelength, proper grating width for a given operational temperature is required. It is determined using the following equation.

$$\frac{n_p}{\lambda_p} = \frac{n_s}{\lambda_s} + \frac{n_i}{\lambda_i} + \frac{1}{\Lambda}, \tag{37}$$

where n_p, n_s, and n_i are the refractive indices corresponding to pump wavelength, λ_p, signal wavelength, λ_s, and idler wavelength, λ_i. Figure 17 illustrates the dependence of grating widths on wavelength and temperature. It can be seen that for obtaining 3823 nm at 50°C (323 K) PPLN with 28.9 μm is required. If a slightly different grating period is used then operational temperature has to be adjusted accordingly. Figure 18 shows the tuning of idler and signal wavelengths as the temperature is varied for the PPLN crystal of grating spacing equal to 28.9 μm. Due to weak dependence of idler and signal wavelengths on temperature, the operational wavelength can be kept stable within the tolerance range of 9 nm or 0.009 μm using temperature controllers.

3.2.3. Oscillation Threshold Calculations

The oscillation threshold calculations are widely described in literature. [33–35] Two cases, unfocused (or unoptimized) and focused Gaussian pump beams are considered for modeling. For un-optimized pump beam, the oscillation threshold was computed to be 757 kW/cm^2. The input parameters used were a PPLN crystal of length 5 cm, D_{eff}= 17 pm/V and a pump radius of 100 μm with TEM mode profile. This is for the case of single pass pump beam. If the pump beam is double passed, the threshold decreases to 489 KW/cm^2. [30] When the pump beam is optimized with a focusing parameter of

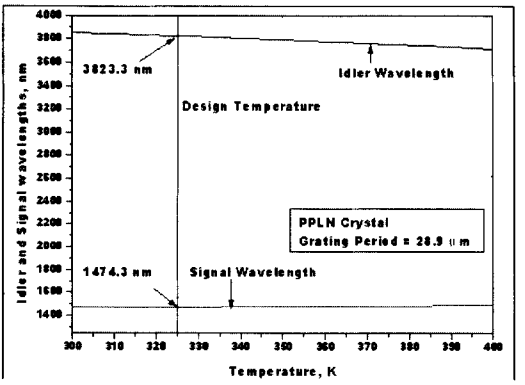


Fig. 18. Temperature tuning of PPLN crystal of 28.9 mm grating period.

1.26, the threshold drops to about 33 kW/cm^2. The threshold calculations give us an approximate pump intensity required to start seeing an OPO output. The peak intensity of a 15 μJ/pulse at 10 ns when focused to a 100-μm spot radius is 9.55 MW/cm^2. This peak intensity is nearly 12.5 times the single pass, un-optimized pump beam. When the pump beam is optimally focused, then this figure goes up even further. The rule of thumb is that the conversion efficiency can be maximized if the threshold is more than five times the threshold.

The interaction length is proportional to the crystal length. Since PPLN crystals have no walkoff, longer the interaction length, higher is the conversion efficiency. However, OPO parameters have to be optimized to minimize back conversion. To achieve maximum conversion efficiency, the pump beam with a TEM_{00} mode profile has to be properly injected into the OPO cavity. The pump beam spot has to be located in the center of PPLN crystal. Furthermore, the signal and idler radiations have to be contained inside the crystal. It is also clear that power threshold increases as the pump beam waist diameter increases. However, smaller beam waists result in greater beam divergence in the PPLN gain medium. The pump beam optimization has to be carried out to avoid beam clipping by the PPLN crystal. In this case, the Rayleigh range has to be greater than the PPLN crystal thickness. The pump beam spot radius of about 100 μm at the center of the PPLN crystal of length 5 cm will be appropriate. At 10 μJ/pulse, a 100-μm spot radius has intensity greater than 2 MW/cm^2. Greater than 100 μm spot radius at the output coupler of the pump laser can be obtained for the plano-concave cavity configuration of length 25 mm and output coupler radius of curvature of 50 mm. A 2-W diode pumped, passively Q-switched Nd:YAG laser with 7.5 % total optical slope efficiency operating at 10 kHz PRF would provide 15 μJ/pulse. Assuming 25% pump depletion, >1 μJ/pulse at idler wavelengths can be obtained. These are very conservative figures and a better performance can be anticipated by minimizing coupling losses.

The output beam from the laser resonator can be mode-matched to the PPLN crystal using suitable beam shaping arrangement. The PPLN crystal has to be enclosed inside an oven for temperature control. For partial DRO operation, the specifications of the OPO resonator can be as follows: A concave-concave type resonator with 50

mm radius of curvature mirrors and a 5-cm length × 1-cm width × 500-μm thick PPLN crystal. The input coupler mirror will have an AR coating at pump wavelength and high reflectance coating at signal and idler wavelengths. The output coupler will totally reflect the pump and signal wavelengths, and completely transmit the idler (3.82 μm) beam. The partial DRO configuration can be chosen to enhance the OPO conversion efficiency. For this OPO cavity, the signal beam that resonates inside will also be completely contained within the 500-μm-thick PPLN crystal.

4. The MWIR Detectors

Pyroelectric detectors and PbSe detectors are available for use from 3 to 5 μm. These detectors allow room temperature operation. [36] Pyroelectric detectors generate charge, and, hence, a signal, in response to heating. They can be operated in either voltage or current mode. Advantages include (1) extremely broad spectral response (0.0001–1000 μm), (3) wide operating temperature ranges (−55 to 125°C). [37] PbSe detectors have peak sensitivity at 3.8 μm. They are photoconductive devices and operate by exciting photoelectrons across a semiconductor band gap into a conduction band. Advantages of this type of detector include (1) slightly higher specific detectivity, or $D*$ values, than the pyroelectric detector, (2) shock resistance since they are photoconductive devices, (3) wide operating temperature ranges (−20 to 50°C). Both detectors can be equipped with a wide array of optical bandpass filters. [38] Both the PbSe and pyroelectric detectors are inexpensive, can be ruggedized. Typical specifications of PbSe and pyroelectric detectors are given in Table 1. New detector technologies in the MWIR spectral region, however, offer substantially better performance. Examples of these new detectors include (1) "Dember effect" or "non-equilibrium" detectors, (2) Quantum Well Infrared Photodetectors or QWIPS, (3) microbolometers, (4) resonant enhanced cavity detectors, and (5) Schottky barrier PdSi, PtSi, and $PtGe_xSi_{1-x}$ detectors. A disadvantage of some of these technologies, however, is that performance specifications are classified and/or not commercially available.

Table 1. Atmospheric transmission from 3.5 to 4 μm.

Type Model Number	Active Area (mm)	Peak Sensitivity Wavelength (μm)	Sensitivity	Detectivity (D*) ($cmHz^{1/2}/W$)	Rise Time
PbSe P791-02	3 × 3	3.8	70 μV	1×10^9	2 to 5 μsec
Pyroelectric P3782	$\phi 2$	flat	1500 V/W	1.7×10^8	100 msec

4.1. Dember Effect Detectors

These devices are based on Zn-doped HgCdTe semiconductor materials with free carrier concentrations below their (nominal) equilibrium values (non-equilibrium mode). [39] Typical $D*$ value of these devices at room temperature is about 1×10^{11} cm $Hz^{1/2}/W$ in the MWIR spectral region. This is a significant performance enhancement over the PbSe and pyroelectric detectors previously discussed, and use of Dember effect detectors would substantially reduce the required laser transmitter power. The Dember

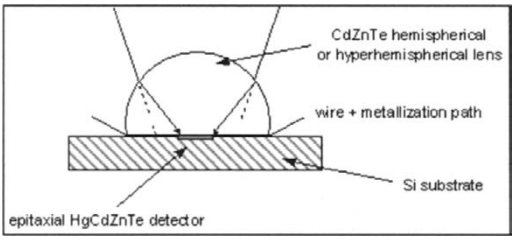


Fig. 19. Schematic diagram of a commercially available HgCdTe non-equilibrium or "Dember effect" detector.

Parameter	Specifications
Peak Detectivity, cm $Hz^{0.5}/W$	$>3*10^9$
Responsivity V/W	24
Response Time, ns	< 10
Resistance, Ω	35
Active Area, mm^2	1
Noise Density, $nV/Hz^{0.5}$	0.8
FOV, degrees	42

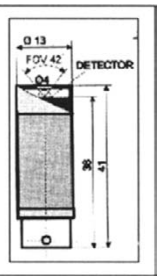


Fig. 20. The detector specifications supplied by the manufacturer and its dimensions.

effect detectors are attractive due to several reasons. They include (a) ambient temperature operation, (b) no bias requirement, (c) short response time, (d) no flicker noise, (e) wide dynamic range and (f) commercial availability at relatively low cost. These detectors are available for operation in the spectral range of 2–12 μm and individual detectors are designed to exhibit peak performance at a specific wavelength. Dember effect detectors are commercially available from VIGO Systems Ltd. (`www.vigo.com`) from Poland and are distributed by Boston Electronics, Inc. in USA. Figure 19 shows a schematic diagram of a Dember effect detector. The HgCdZnTe photodetector is seated on a Si substrate. A CdZnTe hyperhemispherical lens is formed over the detector using ultrasonic and classical mechanical manufacturing techniques. The monolithic design permits the construction of rugged and mechanically stable devices. Additionally, by properly coating the spherical lens surface the device can be (1) shielded from unwanted background radiation and (2) antireflection coated for operation at the desired wavelengths.

For MWIR wavelength detection, the PDI-5 detector and a matching transimpedance amplifier, Model 481-20PEM are available from Vigo Systems.

Figure 20 shows the vendor measured specifications and the detector dimensions. The procured detector has a specific detectivity, D^* of $> 3E9$ cmHz $^{0.5}$/W. The PDL-5 is a photovoltaic heterostructure based on epitaxial variable band gap $Hg_{1-x}CD_xTe$ structures grown using Isothermal Vapor Phase Epitaxy. The photovoltage signal is due to both bulk photodiffusion effect (the Dember effect) and photovoltage at hetero-

junctions. Actually, in the short wavelength (≈3.4μm) devices, the response is mainly due to the photovoltaic response at heterojunctions, the contribution of the volume photovoltage is small. The devices are monolithically optically immersed to CdZnTe lens ($n = 2.7$, radius of curvature 0.8 mm, position of active layer-aplanatic plane ($r + r/n$ distance from surface of the lens). Coating of the devices can increase the quantum efficiency by ≈20%. Due to optical immersion the physical active area is n^4 times (≈50) times smaller than the apparent optical area. In consequence, this reduces electrical capacitance by the same factor in comparison with conventional device of the same optical area. This aspect is very important for frequency response! The main issue is the low resistance of active device, comparable or even lower than the parasitic series resistance of the device. The large series resistance is due to resistance of the p-type layer. Actually, the device has distributed parameters and is complex to analyze its performance. The result of the parasitic series resistance is a low current responsivity and apparent quantum efficiency. Simply, the series resistance reduces current responsivity, even if the devices would have 100% percent of quantum efficiency. This is specific feature of photovoltaic devices operating at long wavelength and high temperatures. One can bias this detector device. The device exhibits soft breakdown due to heating. One has to be careful to prevent excess thermal breakdown. Starting with a low bias current (few μA) and increasing it to the highest photoresponse one will notice an increase of photocurrent and, more pronounced, photovoltage. The decrease of photoresponse means heating of the device, so it is better to work with bias current below the current of the maximum response. One can expect significant drop of photovoltages on parasitic impedances, which will reduce possible gains. The noise of the non biased photovoltaic device with no or weak illumination is the Johnson-Nyquist noise [$Vn = (4kTR\Delta f)^{1/2}$]. When biased one can see increased current and low frequency noise. In the future, detectors providing a quantum efficiency of about 80–90% and frequency response up to several hundreds of MHz are possible to fabricate with the following techniques: (a) reduction of series resistance by the use of heterostructure with heavily doped n^+ layer as the first deposited layer on substrate. This will significantly improve the measured quantum efficiency and response time (b) use of AR coating on immersion lens.

Current Dember effect detectors are within a factor of 6 of the theoretical Background LImited Performance (BLIP) for 2π steradian field-of-view and 300 K sky temperature. It is expected that BLIP limited performance at 4 μm will be achieved within the next few years with improvements in semiconductor materials processing. It is anticipated that Dember Effect Detectors with D^*= 1 × 10^{11} cm $Hz^{0.5}$/W, within a factor of 100 of those available at 1 μm, will be available in the near future. Higher sensitivity detectors mean that lower transmitter powers are required.

5. Data Communications in the Mid-IR

Using the emerging QCL device technology, data communication through direct modulation can be carried out. [40] Due to limited wavelengths and output power (less than 25 mW) available, the QCL lasers are suitable for very limited free-space range applications. These diodes would be attractive provided low loss optical fibers in the MWIR become available. However, another approach is to use the OPO devices. These OPO devices are capable of providing very high peak powers in the pulsed mode. In

the following section, the performance of a compact, eyesafe, all-solid-state, mid-wave IR (MWIR) transceiver for data communication through low visibility conditions is discussed.

The transceiver was developed for Multiple Integrated Laser Engagement System (MILES) application. [41–42] The MWIR wavelengths are derived using a passively Q-switched Nd:YAG laser pumped periodically poled lithium niobate based OPO. MILES weapon code transmission for small and heavy weapon platforms have been demonstrated through dense theatrical fog. With less than 3 μJ/pulse at $\sim$4 μm and a room temperature Dember effect IR detector, greater than 5 km range has been successfully demonstrated [43] and is discussed in the following subsection.

5.1. Weapon Code Transmission in MWIR

Multiple Integrated Laser Engagement System (MILES) is a tactical engagement simulator that facilitates force-on-force training in realistic battlefield scenarios. MILES is a free-space communication transceiver that generates weapon codes that is characteristic of the weapon platform, ammunition features and user. It transmits an encoded message simulating weapon firing characteristics, round dispersion patterns, hit or miss probability with range information. The receiver decodes the message following which further action is initiated. MILES is a laser communications system. It operates by transmitting and receiving encoded "words" that identify the shooter and the amount of damage inflicted on the receiver (hit or miss). MILES transmitters are designed to mimic the actual firing rates and effective ranges of the weapons they simulate. The MILES has been proven to increase the survivability, combat readiness, and fighting effectiveness of military forces. The current MILES, however, suffers range limitations under certain reduced visibility conditions. These conditions include inclement weather (rain, snow, fog), dust clouds, and camouflage smokes. This range limitation reduces training effectiveness, can lead to training postponement and hence higher training costs. For the warrior to effectively train to the capability of his weapon it will be necessary for future MILES to penetrate both natural and anthropogenic obscurants as effectively as a FLIR.

The current MILES and MILES 2000 trainers operate at 904 nm. The 904-nm MILES that is currently being widely deployed provides exceptional performance in clear atmospheric conditions only. The weapon code sequence in a 904 nm wavelength MILES is generated using a GaAs laser diode. A Si photodiode based receiver decodes the received code sequence. Besides eyesafety concerns, the device does not operate in the presence of obscurants. Advantages of the system include small size and weight, environmental ruggedness, demonstrated effectiveness, and low component cost. The principal disadvantage is worse transmission than a thermal imager through some weather conditions and obscurant clouds. This limitation can result in reduced training effectiveness, canceled training activities, misses and false alarms, and troop frustration. Although in the past, MILES operating at 1.5-, 2.1- and 10.6-μm wavelengths have been developed to address the transmission shortcomings, they do not provide optimal performance. The 1.5- and 2.1-μm wavelengths do not penetrate obscurants. The 10.6-μm system paired with relatively inefficient pyroelectric detectors are bulky and requires high prime power.

MILES operating in the mid-wave IR (MWIR) spectral region for small and heavy weapon platforms has been investigated by Coherent Technologies, Inc. (CTI). It was

shown that the MWIR MILES has the potential to provide eyesafe and optimal performance in the presence of battlefield and natural obscurants. The MWIR MILES obscurant penetration capability has been validated using a theatrical fog generator at the Table Mountain test site located near Boulder. The MWIR wavelength is generated using an all-solid-state device technology using a PPLN OPO. Field tests using compact MWIR MILES have produced range performance greater than 5 km. Although specifically developed for MILES, the proposed approach is, in general, applicable for data communications in the MWIR region. In this section, wavelength selection and obscurant penetration analysis, transceiver design approach, and field performance of a MWIR MILES transceiver in the presence of obscurants are discussed.

5.1.1. Wavelength Selection

A variety of atmospheric conditions including clear days, rain, snow, and fog were modeled using FASCOD3p. Furthermore, fog oil based battlefield obscurants consisting of airborne particulates that reduce visibility were also considered for modeling. The FASCOD3p parameters were chosen to mimic conditions in locales where MILES training is conducted. The data required for simulation was collected from respective locations through STRICOM/US Army. The DoD obscurant handbook [44] describes typical CL products for battlefield use of fog oil smokes. Manmade battlefield obscurants behave differently because they have been engineered to greatly reduce visibility at visible wavelengths. Typically, the particles that compose manmade obscurants have wavelength dependent absorptions associated with them. An example of such an obscurant - that is commonly used in training is fog oil. The DoD obscurant handbook describes typical CL products for battlefield use of fog oil smokes. With enough power any wavelength will penetrate the atmosphere regardless of weather or obscurant. However, from eyesafety, power consumption, and platform compatibility considerations, MWIR wavelength offers significant advantages. The 4-μm wavelength offers several significant advantages over the other wavelengths considered. Based on modeling, the 4-μm window was selected because it offers high transmission and can be readily accessed using "practical" all solid-state laser hardware. Furthermore, longer wavelengths offer relatively reduced scintillation effects. Additionally, the 4-μm wavelength certain atmospheric conditions this wavelength offers advantages over either the 900-nm or 1.55-μm wavelengths. The FLIR sight is unaffected by the smoke. The 905-nm and 1.55-μm wavelengths are severely attenuated by the smoke and the 3.8-μm wavelength is somewhat attenuated. Despite the reduction in visibility at 3.8 μm due to the smoke, the 3.8-μm and the 10-μm wavelengths are attractive for transmission through obscurants.

5.1.2. Transceiver Design Approach

The critical specifications for small weapon platform such as M16 rifle include range equal to 500 m, beam divergences of 0.7 and 1.5 mrad for hit and near miss cases, and a single 9V battery operation satisfying 200 rounds. For heavy weapon platforms such as M1A2 tanks, the range is 3.75 km, and beam divergences for hit and near miss cases are 1.2 and 1.7 mrad, respectively. Ample power is available in tank platforms. One attractive spectral region is around 3.8 μm wavelength. There are several low atmospheric loss sub-bands available for use and are illustrated in Fig. 21. The 3.8-μm

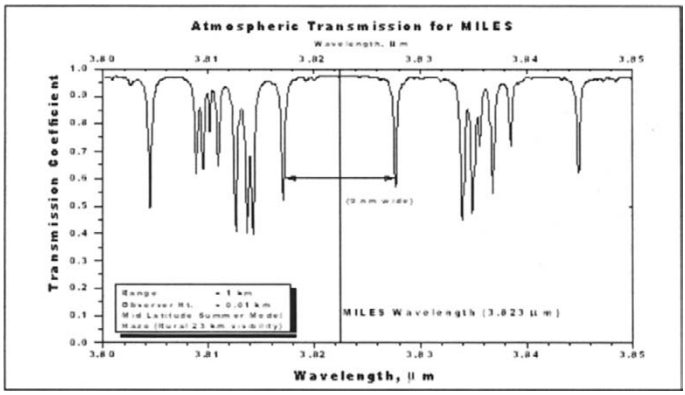


Fig. 21. One-way transmission in the 4-mm region expanded around 3.82 mm wavelength.

wavelength region is farther away from C-H absorption bands. From obscurant penetration analysis based on atmospheric modeling, fog oil characteristics and source availability, operational wavelength selected for MWIR MILES was 3.82 μm with a tolerance width of 9 nm. To obtain 3.82 μm radiation, the optical parameter oscillator (OPO) based approach was selected. OPOs are viable candidates for a next-generation MILES transmitter due to their broad tunability in the MWIR and potential high efficiency. As mentioned earlier, OPOs can be directly or indirectly laser diode-pumped. The advantage of either pumping scheme is that the laser diode can be modulated to output the appropriate MILES message code and the OPO output will be similarly modulated. Additionally, diode pumping enables the construction of small and compact transmitters. The indirect diode pumping based on Nd:YAG pump laser approach provides short pulse, high peak power, and extended range performance that is not practical with direct diode pumping.

The all-solid-state MWIR MILES transmitter consisted of a passively Q-switched, Nd:YAG pump laser coupled to a periodically poled lithium niobate (PPLN) OPO. The passive Q-switching was achieved using the Cr^{4+}:YAG material. PPLN is an engineered quasi-phase-matched (QPM) nonlinear optical material that can efficiently generate mid-IR wavelengths. The advances in the QPM technology have permitted smaller and more efficient devices to be built. A 5-cm-long PPLN crystal was located inside a temperature-controlled oven. The PPLN OPO resonator was a concave-concave type and was operated in a singly resonant oscillator mode with signal resonance. The PPLN OPO was temperature tuned from 3.6 to 4 μm. Initially, photorefractive damage of the PPLN material was a concern. The effects of photorefractive damage were evaluated by operating the PPLN OPO at 50°C for 10 days with 8 hours of continuous operation each day. However, no visible damage to the PPLN crystal was observed. A miniature telescope was used for output beam shaping of the idler wavelength.

The transmitter drive electronics consisted of laser diode driver, encoder and signal conditioning electronics. The pump laser cavity was optimized for a free-running pulse repetition frequency of 10 kHz. Typical laser pulsewidths were around 6.5 ns. Short pump pulsewidths are advantageous since an OPO requires intense pump pulses whose

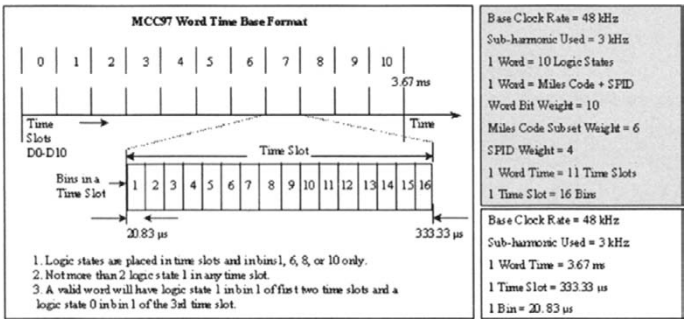


Fig. 22. A summary of MCC97 MILES code format (MCC97 manual obtained from STRICOM website).

intensity has to be above the oscillation threshold. The laser diode was operated in the pulsed mode. The laser drive current pulsewidth was adjusted such that it generates only one laser pulse. Typical current pulsewidths were less than 120 μs. A laser pulse was generated within 100 μs after the onset of laser drive current pulse and this interval can be adjusted by varying the laser cavity parameters.

The encoder pulses trigger the laser diode driver current. Basically, the encoder drives the laser diode electronics to generate a desired MILES code. The MILES code format had to be compatible with MCC97 regulations. For details of MCC97 code, readers are referred to the STRICOM website. A summary of MCC97 is given in Fig. 22. For generating MCC 97 compatible MILES weapon codes, an encoder developed by Lockheed Martin Corporation was used. Figure 23 shows the internal view of a battery operated MILES encoder device. The encoder unit can be set to function as a decoder and vice versa. The encoder device was compliant with MCC97 requirements. A toggle switch is used to change or set the MILES code parameters while a fire control button to is used to transmit the desired code. The LCD panel displays the MILES code parameters. The encoder generates $\sim$ 1 μs width pulses corresponding to a 1 bit that acts as the trigger pulse to the laser diode current driver. Similarly, the decoder requires pulses of 1 μs width corresponding to 1 bit for decoding the code. The received sequence from the preamplifier is inverted TTL logic to be compatible with the decoder. The receiver electronics detects and conditions the received code for decoding purposes. The receiver electronics consisted of a MWIR detector, signal conditioning electronics, and a decoder.

A room temperature "Dember" effect based IR detector with a peak response at 4 μm was used for the MWIR MILES experiments. These IR detectors, as discussed in Section 4.1, are fabricated using Zn:HgCdTe based compound whose principle of operation is based on volume photovoltaic effect. The detector specifications for our experiments include: active area of 1 mm^2, specific detectivity of 3×10^9 cm Hz$^{0.5}$/W, voltage responsivity of 24 V/W and response time less than 10 ns. A CdZnTe lens above the active area provided a 42° field-of-view. Measured characteristics of this detector include: dark current of I_d $\sim$0.4 mA, noise equal to -70.82 dBm ($8.28E-11$ W), and a quantum efficiency of $\sim$20% at 3.39 μm. With AR coatings, the quantum efficiency of the detector can be substantially enhanced.

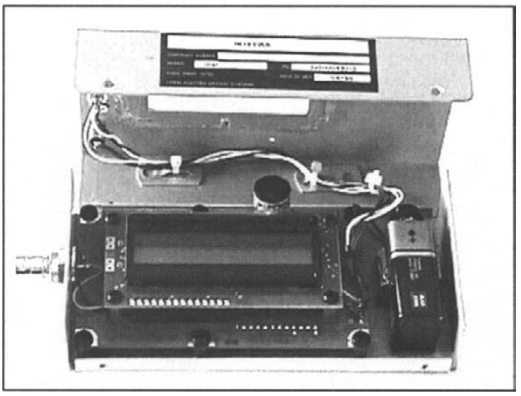


Fig. 23. The MILES encoder and decoder unit.

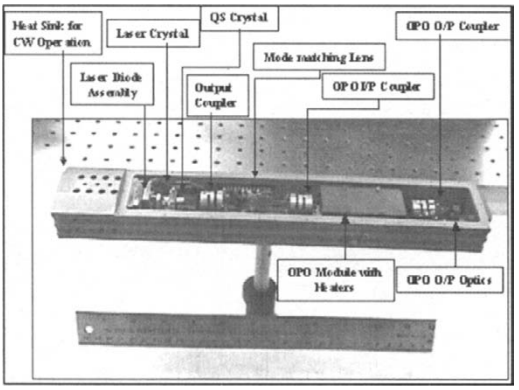


Fig. 24. The MWIR MILES brassboard transmitter unit.

The brassboard unit used to transmit heavy weapon codes is shown in Fig. 24. The breadboard transmitter satisfied the range requirements of a small as well as heavy weapon platforms and validates obscurant penetrating property of the selected MWIR wavelength. Table 1 illustrates the performance of the MWIR MILES transmitter.

The transmitter electronics shown in Fig. 25 comprises of an encoder interface, laser diode driver, OPO heater controller, TEC controller and display interfaces. The battery operated unit consists of a low noise, high bandwidth and high gain preamplifier, automatic noise thresholding circuit, an one shot pulse digitizer circuit, and a pulse stretcher. The pulse stretcher circuit stretches received ~6.5 ns idler pulses up to 1 μsec duration to be compatible with the MILES decoder. The output is inverting and follows TTL logic. The 3 dB bandwidth was 26.4 MHz. The receiver circuit shown in Fig. 25 can be reduced to less than 1 square inch.

Table 2. Demonstrated performance characterisitics of the compact, MWIR MILES receiver.

Laser Diode Power	= 2 W
Max. CW Laser Power	= 840 mW
Passive Absorber	= Cr:YAG, T = 90%
Max. Q-switch Average Power	= 320 mW
PRF Range	= 4–20 ns
Pulsewidth Range	= 4–66 kHz
Max. Laser Pulse Energy	= 16 μJ/pulse
Spatial Mode	= TEM_{00}
Max. Idler Pulse Energy	= 2 μJ/pulse
Max. PRF Jitter	= $\pm$ 2.5 μs
Max. Amplitude Jitter	= $<$ 5%
Max. Pulsewidth Jitter	$\sim$ 2 ns
Pulse Mode Operation	= Highly stable
Size	= 12 $\times$ 2 $\times$ 1.31 inches

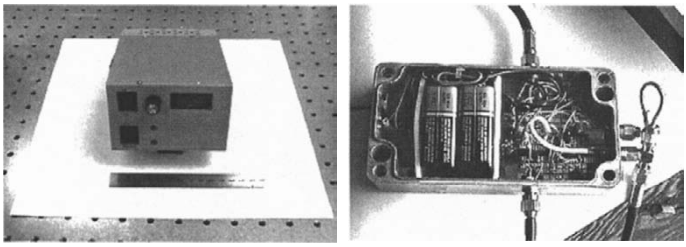


Fig. 25. Transmitter electronics (left) and receiver electronics (right).

The fog was generated using a fog oil based commercially available theatrical smoke generator. The fog fluid was completely water soluble at 70°F, with a vapor pressure of 8.4 mmHg. The unit generated concentrated heavy smoke output with pleasant mild odor.

Finally, with MILES devices, eye safety is an issue. Based on the calculations that were carried out using ANSI Z136-1 1993, it was apparent that any of the longer wavelength options remains eye safe at transmit power levels substantially higher than those permissible at 905 nm. For instance, the MPE/pulse increases from 2.5 ergs/cm^2 to 3.5 $\times$ 10^4 ergs/cm^2 when the wavelength is increased from 905 nm to 3.82 μm. Hence, the MWIR MILES was eyesafe for all practical distances.

5.1.3. Experimental Results

The output characteristics of the PPLN OPO is greatly influenced by the Nd:YAG pump laser characteristics. Figure 26 shows a typical pulse profile and PRF characteristics of the passively Q-switched Nd:YAG pump laser. In this case, the pulsewidth is 6.5 ns and the PRF is around 13 kHz. While the PRF MWIR OPO remains the same, the pulsewidth of the OPO will be lesser than that of the pump laser due to parametric

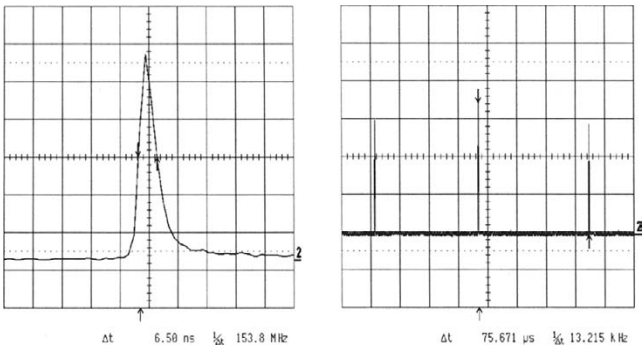


Fig. 26. Left: Typical pump laser pulse profile. Right: PRF characteristics of the pump laser.

conversion process and is illustrated in Fig. 27. Also shown in Fig. 27 is the jitter characteristics of the passively Q-switched Nd:YAG laser. The typical amplitude and PRF jitter characteristics of the pump laser is also shown in Fig. 27. For data communication application under consideration, the PRF jitter has to be substantially smaller than the duration of bin in each time slot as illustrated in Fig. 27. The PRF jitter of a passively Q-switched laser is larger than that in the case of an actively Q-switched laser. The timing jitter has to be controlled for data communication applications since it will effectively limits the data rate. Although in this discussion, the application of a passively Q-switched laser for communications is illustrated, an active Q-switch such as an electro-optic Q-switch based laser can be utilized to provide PRF jitter of less than 100 ns for large data rate applications.

As mentioned earlier, the MILES transmitter operational wavelength was set by fine tuning the temperature of the PPLN OPO. Figure 28 illustrates a sequence of output wavelengths measured using a monochromator as the PPLN crystal wavelength is set at four different temperature values. Following several measurements, the MWIR wavelength tuning rate of the PPLN OPO was determined to be $\sim$1.37 nm/°C. To obtain 3.82 μm, the operating temperature was set at $\sim$50°C. The measured FWHM of the idler pulse was $\pm$ 2 nm. The heater controller kept the OPO temperature within $\pm$ 0.1°C.

Up to 30 mW of average power with good spatial quality was achieved at idler wavelengths. At 10 kHz, this average power corresponds to 3 μJ/pulse. Using this device, a total OPO conversion efficiency greater than 53% has been demonstrated. The field experiments using MWIR MILES were conducted at a laser test range located on the Table Mountain near Boulder, Colorado. The free-space range available for testing was 2.5 km. Typical idler pulse energy used for our small weapon platform related experiments was 0.5 μJ/pulse. The MILES transmitter and receiver were located inside a 30-ft-long shelter. A front surface coated mirror was placed at a distance of 250 m from the MILES transmitter to reflect the idler beam back into the shelter.

In order to transmit the MILES weapon code, pulses corresponding to a "1" bit were generated. An electronics interface to the encoder was constructed to generate a TTL trigger pulse for each "1" bit. The trigger pulse starts the laser diode current. Based

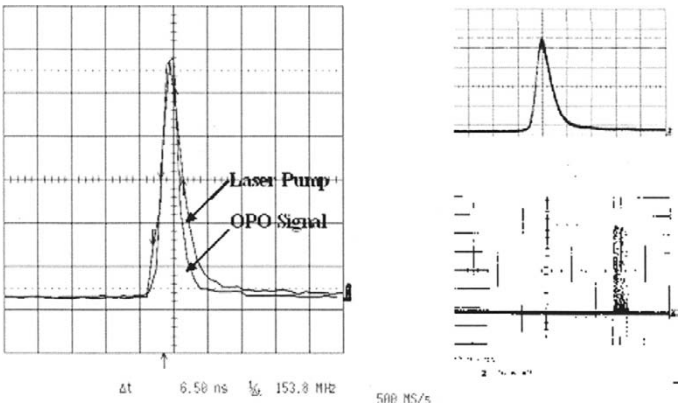


Fig. 27. Right: Pulse compression due to OPO interaction. Pulse profiles before and after OPO interaction are shown. Left: Amplitude and PRF jitter characteristics of the pump laser. Pulse amplitude jitter is less than $\pm$ 5% (peak-to-valley). The PRF jitter of the pump laser was around $\pm$ 2.6 μs.

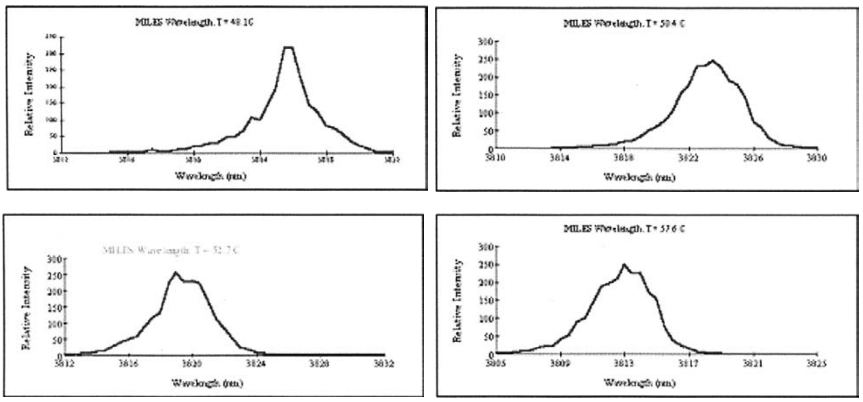


Fig. 28. The MWIR wavelength tuning by varying the PPLN crystal temperature.

on the resonator dynamics, the time required to generate a laser pulse was determined. As soon as the laser pulse is generated, the laser diode current was shut off. In this way, Fig. 29 illustrates the generation of a "1" bit. Whenever the trigger pulse corresponding to a "1" bit occurs, the laser diode current is turned on for a set period to allow the generation of laser pulse. Due to 4 μm detector characteristics, OPO pulse is broadened by several ns and is also shown in Fig. 29. However, the OPO pulse broadening by several nanoseconds had no significant effect on decoding process.

The required transmission range was confirmed by calculating the time elapsed between the instant the OPO pulse leaves the transmitter aperture and the moment it returns back at the receiver using an oscilloscope. A retro-reflector was located at

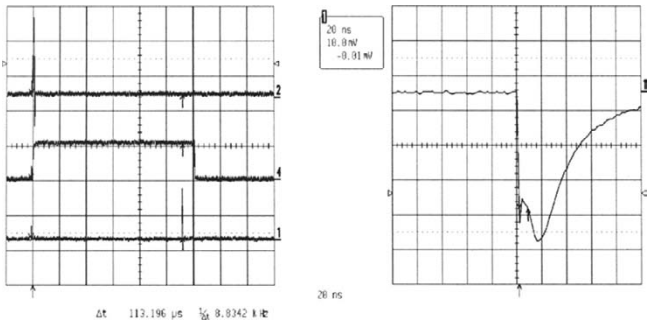


Fig. 29. Left: A weapon code bit generation. Top trace: Trigger pulse to the laser diode driver from an encoder interface. Middle trace: Current pulse to the laser diode. Bottom trace: The OPO pulse. The onset of OPO pulse can be adjusted by varying laser cavity parameters. Right: The detected OPO pulse is broadened to $\sim$ 25–30 ns due to the 4 μm detector characteristics.

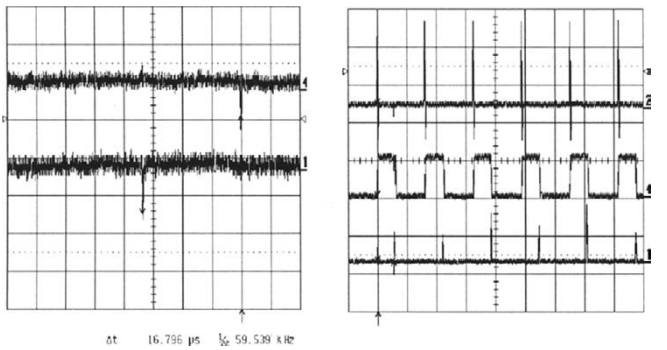


Fig. 30. Left: Range measurements. Top trace: Return OPO pulse. Bottom trace: Transmitted OPO pulse. Right. Checking for bit drop outs. Top Trace: Trigger pulses from the encoder. Center Trace: Drive current pulses to the laser diode. Bottom Trace: The OPO pulse train.

about 2.5 km away from the transmitter. As shown in Fig. 30, the elapsed time was found to be Δt = 16.796 ms. Using the relation, Range = $c\Delta t$, where c, the speed of light, equal to 3×10^8 m/s, the transmission range was 5.039 km. Following range finding measurements, bit drop outs measurements under clear weather conditions and mild atmospheric turbulence. These measurements were done to check whether the transmitter, retro-reflector, and receiver optics were aligned by sending a continuous train of pulses that correspond to a "1" bit. In this case, continuous trigger pulses were transmitted to make sure that for each encoder trigger pulse, a corresponding OPO pulse is emitted.

Following free space transmission experiments, the shelter was filled with dense fog. The visibility was less than 1 ft. The path of the idler beam inside the fog filled

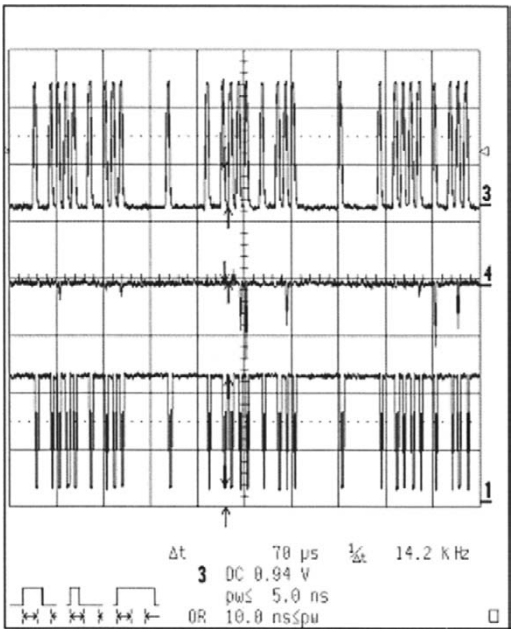


Fig. 31. Oscilloscope traces of the M16 weapon code sequence transmission. Top trace is the transmitted code sequence, middle trace represents OPO pulses, and the bottom trace is the decoded code sequence.

shelter was 60 ft. Under dense fog conditions, successful M16 weapon code transmission was accomplished. The transmission results captured on an oscilloscope are illustrated in Fig. 31. The top trace is the input weapon code sequence. The middle trace corresponds to the output idler pulses. The pulses are very short and are not properly captured due to inadequate sampling. The bottom trace corresponds to the received code sequence. The MWIR MILES performance was compared with that of a 904-nm MILES. The 904-nm MILES totally stopped working at a distance of 30 ft range in the presence of very light fog.

Figure 32 shows the experimental arrangement used for testing heavy weapon codes. The code corresponding to M105 weapon was selected for our tests. Since the one-way range available at the test site was 2.5 km, a set of gold-coated retroreflectors was placed at a distance of 2.5 km from the transmitter location. The smoke was released along its path.

Figure 33 illustrates the results of heavy weapon code transmission. The received sequence was reconditioned to be inverted TTL logic such that it was compatible with the decoder requirements

In summary, data communication at MWIR wavelengths for obscurant penetration is discussed in this section. The MWIR transceiver, that provided high peak power and short pulses, was designed for MILES applications. Weapon codes corresponding to small and heavy weapon platforms were successfully transmitted though fog. The MWIR MILES transmitter concept, in general, is applicable to all types of MILES

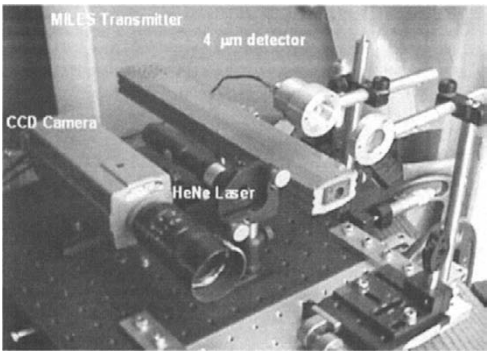


Fig. 32. The MWIR MILES transmitter setup. The MWIR detector and the collecting optics are located on the top of transmitter. A visible CCD camera on the right is used for bore sighting.

platforms. In the following subsection, a bit map image transmission at MWIR wavelengths using this device is presented. Test images consisting of 50 × 40 pixels and 100 × 80 pixels were successfully transmitted through free space.

5.2. *Image Transmission in MWIR using an OPO*

There is a need for free-space, short range, compact, secure, eye-safe, obscurant penetrating and cost effective image transmission systems for defense applications. Such systems can play a significant role in battlefield communications. The MWIR MILES transmitter scheme described in the previous section is capable of providing multiple functions. Besides weapon code transmission, other functions demonstrated using MWIR MILES include: Direct Detection LADAR, Target illumination, aimpointing, Monostatic Reflectance Distribution Function (MRDF) measurements, Bistatic Reflectance Distribution Function (BRDF) measurements and Image Transmission through obscurants. The MWIR MILES device can be easily adapted for Combat ID applications and flying-spot active imaging through obscurants. In this section, the feasibility of image transmission using the MWIR MILES transceiver through low visibility conditions is discussed. [45]

The MWIR MILES transceiver, consisting of a passively Q-switched, short-pulsed Nd:YAG laser pumping a PPLN based OPO and a Dember effect detector was utilized for image transmission experiments. The MWIR MILES transceiver generates pulse position waveforms. The optical data link consisting of transmitter drive electronics, pulse conditioning electronics and a computer generating pulses compatible with the 2400-baud rate RS232 receiver was utilized. Data formatting and RS232 transmission and reception were achieved using a computer. Data formatting transformed an arbitrary image file format compatible with the basic operation of pump laser. Images were transmitted at a date rate of 2400 kbits/sec with 16 bits/pixel. Test images consisting of 50 × 40 pixels and 100 × 80 pixels were transmitted through free-space filled with light fog (visibility ~5 ft) up to 120 ft. Fog was generated using a theatrical fog generator.

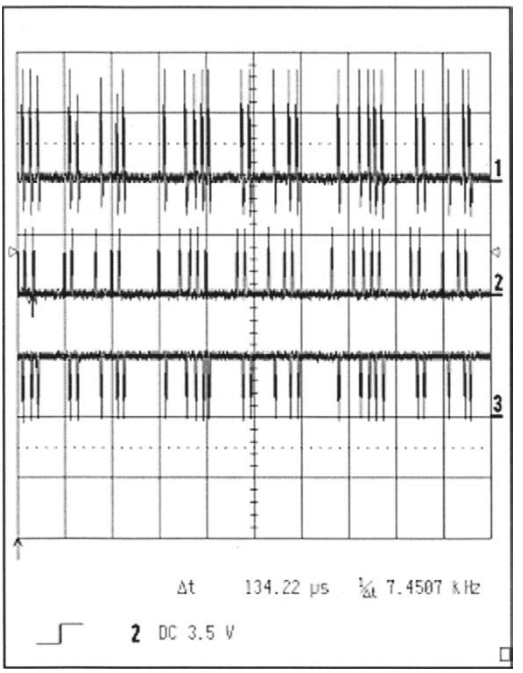


Fig. 33. Oscilloscope traces of the M105 weapon code sequence transmission. Top trace is the received code sequence (before decoding), middle trace is the transmitted code sequence, and the bottom trace is the inverted decoded code sequence.

Besides OPOs, the proposed concept can be extended to optical parametric amplifiers, Raman lasers and other nonlinear optical devices to achieve multi-functionality.

The elements of the data link are illustrated in Fig. 34. The data link consisted of a computer, a level shifter and pulse generator, the laser drive electronics, the MILES transmitter, a Dember effect detector, the pre-amplifier and threshold detection circuit, and a pulse generator and level shifter. The Q-switch timing of the passively Q-switched pump laser was electronically adjustable. The externally generated trigger pulse to the laser diode driver starts the laser diode current pulse. The diode current pulsewidth is variable and can be set to a predetermined value. As soon as the Q-switched laser pulse is emitted, the diode current pulse is shut off to eliminate spurious pulses. In our case, the laser pulse emission was typically adjusted to occur after 80 μs from the start of laser diode pulse. The Q-switch laser pulse onset was adjusted by varying the laser cavity parameters and the Q-switch crystal transmission.

A computer was used for data formatting and RS232 transmission and reception. The level shifter and pulse generator between the computer and the laser drive electronics was required to generate pulses compatible with the laser drive electronics and the basic laser operation. The level shifter and pulse generator between the threshold detection circuit and the computer was required to "stretch" the detected pulses to a length compatible with the 2400 baud rate RS232 receiver. Data formatting was

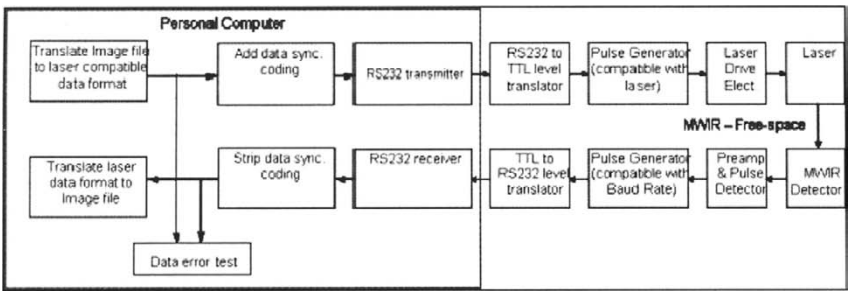


Fig. 34. The schematic of imaging scheme.

required to make the arbitrary image file format compatible with the basic operation of the laser. For RS232 computer serial data transmission the output will remain at a high voltage level for a "1" data bit level and remain at that level until a "0" data bit is encountered or until the data byte is transmitted. The MILES transmitter can only respond to distinct drive pulses. Driving the laser electronics with a high level for an extended period will only result in multiple erroneous pulses being sent out of the laser (after the first pulse the laser emits pulses at a high PRF). This is the result of the build-up time of the laser varying in the short term such that a specific drive level to create exact correspondence of output pulses to high level drive voltage duration. Subsequently, to prevent the occurrence of concatenated "1"s (high level laser drive voltage) in the data output, "0"s are inserted in the data for every other bit. For RS232 data format there is a 1 transmitted as a "Stop" bit with every word transmitted, which is appended to the eight bit of data. To accommodate the stop bit the "0"s are inserted in the even bit location, the LSB being bit 1.

For data framing error mitigation, a synchronization bit pattern is inserted in the data and the output data stream is reformatted to conform to the synchronization rule. The error sources are from undetected pulses or false alarms in the data that may occur during transmission. For data reception the program detects this synchronization pattern and attempt to restructure the data to maintain the synchronization logic. After synchronizing the formatting is translated and the image is returned to its original format, including any residual transmission errors.

Figure 35 illustrates transmitted and received images through light fog over 120 ft at 3.82 μm. For a typical case of operation, the transmitter PRF and pulse length were 10 kHz and 10 ns. The images consisted of low resolution (50 $\times$ 40 pixels) and high resolution (100 $\times$ 80 pixels). The data rate selected was 2400 bits/sec with 16 bits/pixel. No bit errors were detected in this demonstration. The dimensions of the room used to contain the fog limited the transmission range. The interface electronics and the laser PRF limited the data rate.

However, energy scaling and increasing the transmitter PRF combined with improvements in electronics hardware can alleviate these deficiencies to a certain extent. For improved timing control and higher PRF, the passive Q-switch can be replaced by an electro-optic or an acousto-optic Q-switch at the expense of higher power consumption and relatively complex electronic drivers.

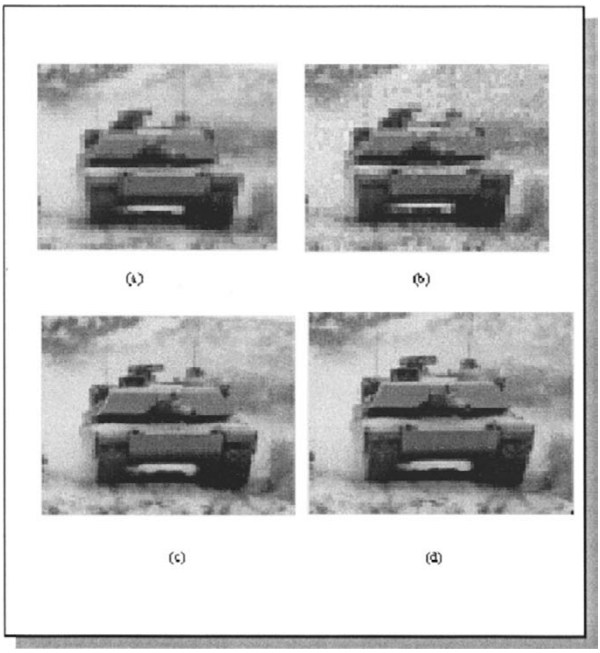


Fig. 35. Image transmission through light fog (visibility $\sim$ 5 ft) using the MILES unit. Data rate is 2400 kbits/sec with 16 bits/pixel. (a) Input low-resolution image (50 $\times$ 40 pixels), (b) Received low-resolution image, (c) Input high-resolution image (100 $\times$ 80 pixels), (d) Received high-resolution image.

Data link was designed to be highly compatible with common computer interface. Testing demonstrated transmission of image files of over 4 kilobytes with no errors in a moderately smoke filled path up to 120 ft at a 2400-Hz baud rate. The system proved the capability to transfer bit-mapped images through moderate smoke/fog and could have been modified for other image or data formats easily. Initially, MWIR MILES was designed to be battery operated and hence passive Q-switch scheme was selected. Such a unit with image transmission capability will be compatible with small weapon platforms.

In summary, Image transmission using passively Q-switched laser pumped OPOs is feasible. Increasing the transmitter PRF can increase the data rate. As such, this concept can be extended to any optical wavelength and to high-power diode-pumped pulsed laser based transmitters. Such transmitters can be utilized to pump OPOs, optical parametric amplifiers, frequency up-converters, Raman lasers, and other nonlinear optical devices to achieve multi-functionality.

6. Summary and Conclusions

The MWIR spectral region is considered to be a low loss atmospheric window. As such, it is playing a considerable role in lidar for hydrocarbon detection [46,47] and ladar applications. With the emergence of novel source and detector technologies, it will be contributing significantly to the development of systems whose functionality depends upon atmospheric propagation. In the above discussion, salient features of atmospheric effects on beam propagation, atmospheric modeling, MWIR sources with emphasis to OPOs, and MWIR detectors are presented. Furthermore, the free-space data communication and image transmission in the MWIR region using OPOs are discussed. In conclusion, the MWIR spectral region offers significant advantages for lidar, ladar, free-space optical communication, illumination, and hard target imaging applications for military and civilian use.

References

1. M.E. Thomas, D.D. Duncan, Atmospheric transmission, in Atmospheric Propagation of Radiation, vol. 2 of The Infrared & Electro-Optical Systems Handbook, edited by F. G. Smith, pp. 1–156, 1993.
2. L. Andrews, R. Phillips, *Laser Beam Propagation through Random Media* (SPIE Optical Engineering Press, Bellingham, WA, 1998).
3. L. Andrews et. al., *Laser beam scintillation with applications* (SPIE Optical Engineering Press, Bellingham, WA, 2001).
4. J. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1996).
5. W. Coles et al., Simulation of wave propagation in three-dimensional random media, Appl. Opt., **34** (12), 2089–2101 (1995).
6. R. Frehlich, Simulation of laser propagation in a turbulent atmosphere, Appl. Opt., **39** (3), 393–397 (2000).
7. American National Standards for Safe Use of Lasers, ANSI Z136.1-2000, published by American National Standards Institute, New York, 2000.
8. M. Tacke, New developments and applications of tunable IR lead salt lasers, Infrared Phys. Technol., **36**, 447–463 (1995).
9. J. Hecht, *The Laser Handbook* (McGraw-Hill, New York, 1992).
10. Improved lead-salt lasers set power record, in Laser Focus World, vol. 33, p. 9, January 1997.
11. H.K. Choi, G.W. Turner, H.Q. Le, InAsSb/InAlAs strained quantum-well lasers emitting at 4.5 μm, Appl. Phys. Lett., **66**, 3543–3545 (1995).
12. J. Faist, F. Capasso, D.L. Sivco, C. Sirtori, A.L. Hutchinson, A.Y. Cho, Quantum cascade laser, Science, **264**, 553–556 (1994).
13. J. Faist, F. Capasso, C. Sirtori, D.L. Sivco, J.N. Baillargeon, A.L. Hutchinson, S.N.G. Chu, A.Y. Cho, High power mid-infrared ($\lambda \sim 5$ μm) quantum cascade lasers operating above room temperature, Appl. Phys. Lett., **68**, 3680–3682 (1996).
14. J.F. Pinto, G.H. Rosenblatt, L. Esterowitz, Continuous-wave laser action in Er^{3+}:YLF at 3.41 μm, Electron. Lett., **30**, 1596–1598 (1994).

15. H. Többen, Room temperature cw fibre laser at 3.5μm in Er^{3+}-doped ZBLAN glass, Electron. Lett., **28**, 1361–1362 (1992).

16. J. Schneider, C. Carbonnier, U.B. Unrau, Continuous wave fiber laser operation at a wavelength of 3.9 micrometers, OSA TOPS on Advanced Solid-State Lasers, **1**, 333–334 (1996).

17. R.C. Eckardt, L. Esterowitz, I. D. Abella, Multiwavelength mid-IR laser emission in Ho:YLF, paper FM5, Digest for Conference on Lasers and Electro-Optics, p. 160, 1982.

18. R.L. Byer, Optical parametric oscillators, in Quantum Electronics: A Treatise, edited by H. Rabin and C.L. Tang, pp. 587–701, 1975.

19. A. Yariv and P. Yeh, *Optical Waves in Crystals* (John Wiley & Sons, New York, 1984).

20. V.G. Dmitriev et. al., *Handbook of Nonlinear Optical Crystals* (Springer-Verlag, New York, 1991).

21. W.R. Bosenberg, A. Drobshoff, J.I. Alexander, L.E. Myers, R. L. Byer, Continuous-wave singly resonant optical parametric oscillator based on periodically poled $LiNbO_3$, Opt. Lett., **21**, 713–715 (1996).

22. M. Scheidt, B. Beier, R. Knappe, K.-J. Boller, R. Wallenstein, Diode-laser-pumped continuous-wave KTP optical parametric oscillator, J. Opt. Soc. Am. B, **12**, 2087–2094 (1995).

23. H. Komine, J.M. Fukumoto, W.H. Long, Jr., E.A. Stappaerts, Noncritically phase matched mid-infrared generation in $AgGaSe_2$, IEEE J. Select. Top. Quantum Electron., **1**, 44–49 (1995).

24. L.E. Myers, R.C. Eckardt, M.M. Fejer, R.L. Byer, W.R. Bosenberg, J.W. Pierce, Quasi-phase-matched optical parametric oscillators in periodically poled $LiNbO_3$, J. Opt. Soc. Am. B,. **12**, 2102–2116 (1995).

25. L.E. Myers, R.C. Eckardt, M.M. Fejer, R.L. Byer, W.R. Bosenberg, Multigrating quasi-phase-matched optical parametric oscillator in periodically poled $LiNbO_3$, Opt. Lett., **21**, 591–593 (1996).

26. R. Lavi, A. Englander, R. Lallouz, Highly efficient low-threshold tunable all-solid-state intracavity optical parametric oscillator in the mid infrared, Opt. Lett., **21**, 800–802 (1996).

27. H. Plaessmann, A. Drobshoff, W.R. Bosenberg, Long-pulse, amplitude-modulated optical parametric oscillator, Appl. Opt., **35**, 5964–5966 (1996).

28. M. Scheidt, B. Beier, R. Knappe, K.-J. Boller, R. Wallenstein, Diode-laser-pumped continuous-wave KTP optical parametric oscillator, J. Opt. Soc. Am. B, **12**, 2087–2094 (1995).

29. A.R. Geiger, H. Hemmati, W.H. Farr, N.S. Prasad, A Directly Diode Pumped Optical Parametric Oscillator, Opt. Lett., **21**, 3 (1996).

30. W. Koechner, *Solid-state Laser Engineering*, 4th ed. (Springer, New York, 1996).

31. Selected papers on Optical Parametric Oscillators and Amplifiers and Their Applications, edited by Jeffrey H. Hunt, SPIE Milestone Series, Vol. MS 140, Bellingham, WA, 1997.

32. SNLO nonlinear optics code available from A.V. Smith Sandia National Laboratories, Albuquerque, NM 87185-1423.

33. S.L. Brosnan and R.L. Byer, Optical parametric oscillator threshold and linewidth studies, IEEE J. Quantum Electron., **QE-15**, 415–431 (1979).

34. S.D. Boyd and D.A. Kleiman, Parametric interaction of focused Gaussian Light beams, J. Appl. Phys., **19** (8), 3597–3639 (1968).

35. S. Guha et.al., The effects of focusing on Parametric Oscillation, IEEE J. Quantum Electron., **QE-18** (5), 907–912 (1982).

36. K. Kaufmann, Detectors cover the spectrum of instrument applications, in *Laser Focus World*, vol. 30, pp. 99–105, 1994.

37. Eltec Instruments, Inc., Passive Infrared Technology, High Megohm Resistors, Hybrid Electronics Product Catalog, Daytona Beach, FL.

38. Hamamatsu Corp., Optosemiconductors Condensed Catalog, Bridgewater, NJ.

39. J. Piotrowski, W. Gawron, Z. Djuric, New generation of near-room-temperature photodetectors, Opt. Eng., **33**, 1413–1421 (1994).

40. S. Blaser, D. Hofstetter, M. Beck, J. Faist, Free-space optical data link using Peltier-cooled quantum cascade laser, Electron. Lett., **37**, 12 (2001).

41. Eye-Safe Multiple Integrated Laser Engagement System, Phase I Final Report CTI-TR-9708.

42. All Solid-State Mid-Wave Infrared Multiple Integrated Laser Engagement System (MILES), Phase II Final report CTI-TR-2001-27.

43. Narasimha S. Prasad, Duane D. Smith, James R. Magee, Data communication in mid-IR using a solid-state laser-pumped optical parametric oscillator, Proc. SPIE, **4821**, 214–224 (Dec. 2002).

44. Quantitative Description of Obscuration Factors for Electro-Optical and Millimeter Wave Systems, DoD-HDBK-178 (ER), 1986.

45. Narasimha S. Prasad, Pat Kratovil, James R. Magee, Image transmission in mid-IR using a solid-state laser pumped optical parametric oscillator, Proc. SPIE, **4635**, 272–277 (April 2002).

46. R.M. Measures, *Lasing Remote Sensing* (John Wiley & Sons, New York, 1984).

47. Narasimha S. Prasad, Allen R. Geiger, Remote Sensing of Propane and Methane Using a Differential Absorption Lidar by topographic reflection, Opt. Eng., **35**, 4 (1996).

DOI: 10.1007/s10297-005-0052-2
Originally published in J. Opt. Fiber. Commun. Rep. 2, 279–292 (2005)

Quantum cascade laser-based free space optical communications

R. Martini and E.A. Whittaker

Department of Physics and Engineering Physics
Stevens Institute of Technology
Hoboken, NJ 07030
Email: rmartini@stevens.edu
and ewhittak@stevens.edu

Abstract. The advent of the quantum cascade laser (QCL) with emission wavelengths available in the infrared range from 3 μm through more than 100 μm opens up the possibility of exploiting infrared atmospheric transparency windows for free space optical communications. In an effort to establish the efficacy of using directly modulated QCLs for free space communications we have conducted a series of investigations that demonstrate the potential advantages of this technology. In these experiments we first establish that the QCL has very high modulation bandwidth. We then implement a practical free space communications link that under conditions of atmospheric fog, dust and other obscurants offers significant transparency advantage when compared with near infrared wavelength sources presently used in commercial free space optical communications links.

1. Introduction

Free space optical communications links have recently received attention in the research and commercial development communities because of their potential to serve applications requiring high bandwidth with relatively high security to eavesdropping but without the complexity of installing optical fiber. Commercially available systems have largely been based on the adaptation of telecommunications components in the near-infrared spectral region. However, the free space (atmospheric) optical channel has characteristics that are quite different from fiber-based channels. Molecular absorption, light scattering from suspended particulates, fog, rain and even snow all conspire to make a more complex and less controlled transmission medium for free space optical transmission systems. The choice of source wavelength is thus a critical parameter

in the design of such systems and it is natural to explore whether a significant advantage may be gained by employing longer wavelength systems. However, existing long wavelength laser systems do not enjoy the advantages of the typical semiconductor system employed in telecommunication system such as small size, potential for high bandwidth direct current modulation and high quantum efficiency. A possible solution for this niche has recently become available in the form of the quantum cascade laser (QCL) emitting in the middle infrared (8–10 μm). QCL's have established themselves as versatile semiconductor light sources for the mid-infrared (IR) spectrum and beyond ($\lambda \sim 3.5 - 24$ μm) and are already widely used for spectroscopic applications [1].

We will discuss first the high-speed modulation properties of the QCL, address then a comparison of FSO experiments between a QCL and NIR laser based link in controlled low visibility experiments and finally present a recent FSO application of a QCL based link used for transmission of satellite TV data streams.

2. High Frequency Analog and Digital Modulation

Based on a unipolar lasing mechanism, the QCL possesses unique high-frequency characteristics with theoretical bandwidths in excess of several hundred Gigahertz [2]. Several publications address this issue for applications like gain-switching [3], mode-locking [4], and high-speed modulation without relaxation oscillations [5]. We report in the following from measurements on the high-speed bandwidth of QCLs as well as successful high-speed data transmission as reported in [6]. These experiments were performed using experimental QCLs provided by Lucent Technologies, grown by molecular beam epitaxy in the GaInAs/AlInAs material system, and based on the "3-well vertical" design of the active region [1]. The results presented herein refer to 1.25-mm long, 4.5-μm-wide deep-etched ridge lasers with an emission wavelength near 8.1 μm. To optimize their high-frequency characteristics, the lasers were packaged and processed using a chalcogenide lateral waveguide as described in [3].

The inset of Fig. 1 shows the experimental setup used to measure the high-frequency modulation response of the specified QCL. The device was biased above its cw threshold, and modulated by the output of a synthesized signal generator with frequencies ranging from 0.1 to 10 GHz at a level of 0 dBm. The light output was transmitted over a distance of 1 m and detected with a GaAs/AlGaAs quantum-well infrared photodetector (QWIP) [9], packaged for high-speed operation. The resulting photocurrent was amplified and fed into a microwave spectrum analyzer, where the modulation amplitude was measured. The main part of Fig. 1 shows a typical frequency response of this device for a dc current of 300 mA and a laser temperature of 20 K. The data were normalized to the frequency response of the detector [9] to reflect the modulation response of the QCLs. Aside from a low-frequency shoulder around 2 GHz, which we ascribe to residual parasitic effects, the modulation response remains relatively flat up to roughly 7 GHz, which is adequate for high speed broadband transmission. As discussed in [4], the observed modulation response is still fully parasitics-limited, and nearly independent of the dc current.

To demonstrate the high-speed data transmistion capability of the QCL, the device was modulated with a non-return-to-zero (NRZ) pseudo-random bit stream (PRBS) from a bit error rate (BER) test system. This test pattern is the most common test system to monitor the performance of standard network systems. The quality of the

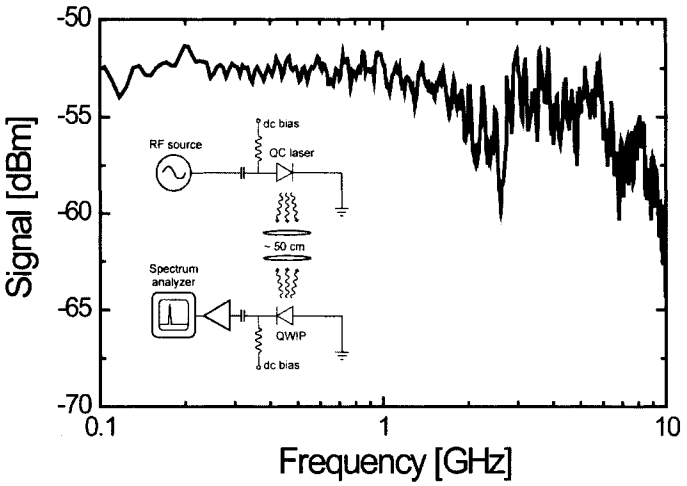


Fig. 1. High-frequency modulation response of the 1.25-mm-long quantum cascade laser at a temperature of 20 K. The inset shows schematically the experimental setup.

communication link is thereby measured using an eye-diagram test as well as in the bit error rate (BER) measurement. In an eye-diagram measurement the clarity of a decision process between a logical "0" and "1" is determined as "openness" of the eye, which is created by averaging over multiple transmitted bit patterns, with the recording oscilloscope synchronized with the transmission clock rate. Figure 2 shows a typical eye diagram observed using the same setup as mentioned above for a $2^{31}-1$ bit long pattern (amplitude 2 V, corresponding to 10 dBm rf power) at a data rate of 2.5 Gbit/s for a heat-sink temperature of 20 K and a dc current of 250 mA. In general, a clear and open eye is observed for heat sink temperatures up to 85 K, demonstrating that the receiving system can distinguish between a logical "0" or "1". However, some deviation in the average "1" and "0" levels can be observed, which we attribute to heating and cooling of the laser due to preceding long "on" and "off"-times. This effect can be prevented using a return-to-zero (RZ) modulation, which is especially suitable at higher data rates, and at operating temperatures of the laser close to its maximum cw temperature.

We also performed BER measurements on the QCL data transmission system. In these tests the average percentage of incorrectly transmitted bits ("0" identified as "1" and vice versa) within a standard transmission session is evaluated. Typical high performance network systems must achieve a minimum 10^{-9} BER, which allows for error-free data communications utilizing enhanced forward error correction algorithms. Naturally preferable are lower BER as they enhance the stability of the communication link. Our QCL transmission link showed error-free data transmission with a BER $< 10^{-12}$ observed for temperatures up to 85K and over the entire range of dc current tested (175 to 300 mA). To quantify the power margins for these error-free data transmission, a variable beam attenuator was included in the beam path. The threshold value for 10^{-9} BER transmission was reached around 3 dB attenuation, which corresponds to a received modulation power of 500 μW (−3 dBm). Similar results were obtained

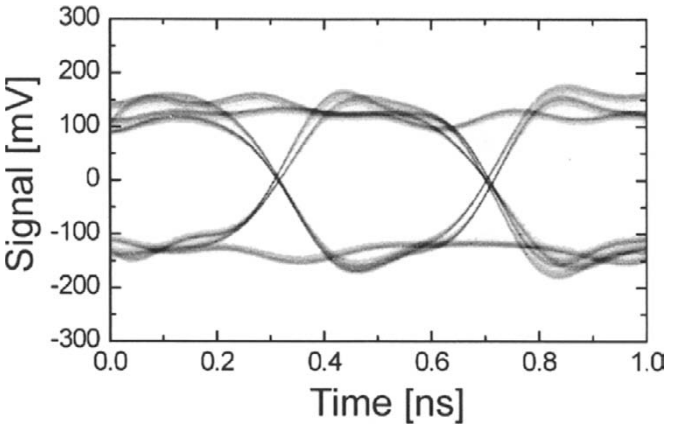


Fig. 2. Observed transmitter eye diagram of the QC laser at 20 K.

for attenuation of the electrical modulation amplitude. This relatively high value in comparison to fiber optic communication systems (1.3 or 1.55 μm) is attributed to the lower quantum efficiency and higher noise equivalent power (−20 dBm) of the QWIP detector as well as to the previously discussed deviation in detected digital level. Nonetheless, these measurements prove the ability of a QCL based FSO link to serve in high bandwidth digital networks.

2.1. Stability for NIR-Laser and a QCL-Based FSO Link Under Strong Scattering

Aside from the modulation speed, other parameters have to be taken into account for the design of a QCL-based FSO link. Design improvements in the basic technology are resulting in steady improvements in optical power, room temperature operation and maximum wavelength of QCLs but whether these improvements will make a QCL-based long-wavelength free space links a logical choice for commercial development is still not clear. Additional factors that must be considered include receiver performance as well as the specific application environment.

A first step in accessing these parameters was made in recent experiments [9], which analyzed the transmission characteristics of three commonly used telecommunications wavelengths and the QCL using standard commercial photodetectors without evaluating the performance of the receivers themselves.

3. Experimental Apparatus

To better understand the wavelength dependence of a free space link we conducted a systematic study to compare systems using three wavelengths used by commercially available free space links (0.85, 1.3 and 1.55 μm) with a quantum cascade laser based link operating at 8 μm. The measurements were performed within a wind tunnel facility

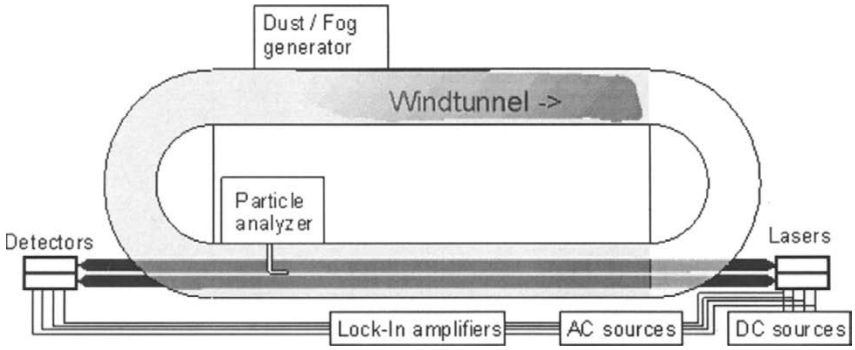


Fig. 3. Arrangement of laser source, detector and wind tunnel for the transmission experiments. The optical path was approximately 65 ft.

that can simulate specific weather conditions and allow for detailed comparison of the four wavelengths. The 76-foot-long Pacific Northwest National Laboratory's (PNNL) Aerosol Wind Tunnel Research Facility allows us to inject particles with well defined size distributions into the beam path and then to measure power transmission and the bit-error-rate for each of the four wavelength lasers.

Figure 3 shows schematically the setup used for these measurements. The PNNL wind tunnel was modified by adding optical windows at either end of the 76-foot-long test leg of the tunnel. The three NIR wavelengths passed through two BK7 glass windows and the QCL beam was directed through an Amtir-1 window. The four different optical links were sent through the 65-foot-long wind tunnel and detected by independent photodetectors (Thorlabs PDA-400) as well as a liquid-N_2 cooled MCT detector. All lasers were operated in cw-mode above threshold and were electrically modulated around 25 kHz. The modulation amplitude of the detected radiation was extracted using lock-in detection. All four transmitted signals were recorded using an electronic strip chart recorder implemented using LabView® on a personal computer. As the intensity varied over several orders of magnitude, the scale settings on the lock-in amplifiers were changed several times. Raw data were recorded and then processed by factoring in the scale changes and referencing transmitted intensity to initial intensity to produce the smoothly varying plots shown in this report. We also analyzed the transmitted intensities to generate an approximate measure of signal-to-noise ratio (SNR) by calculating the rms fluctuation in intensity divided into the rms magnitude of the intensity.

The wind tunnel was equipped with an injection apparatus for water fog, oil fog and dust particles. A sampling apparatus was used to monitor the density and in some cases, the particle mass/size distribution of scatterers in the tunnel. During a particular measurement, the air velocity would be set at a fixed value, typically a few meters per second. The particle injection apparatus would then be activated. The particle density then gradually increased with time over the course of 30–60 minutes and the laser intensities were recorded. The particle measuring apparatus generally showed a linear increase in particle density as a function of time and thus the strip chart records of laser

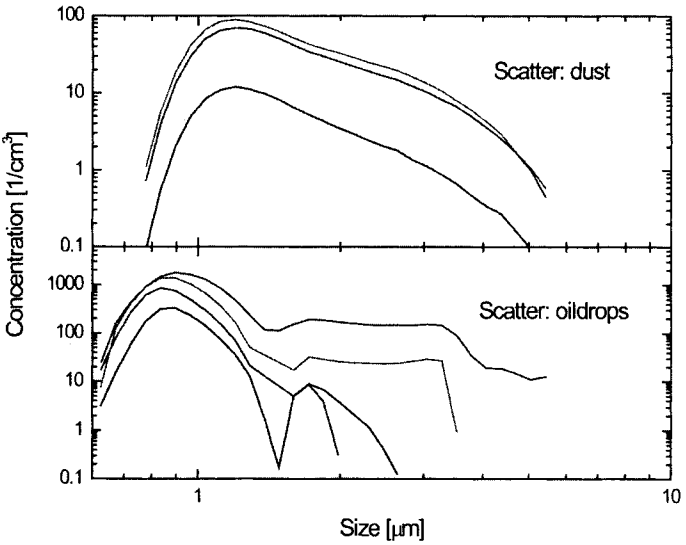


Fig. 4. Typical particle size distributions generated in the wind tunnel for oil (lower) and dust (upper) particles. The different curves represent different measuring time with increasing concentrations of scatter particles.

intensity as a function of time are also uncalibrated measurements of laser intensity as a function of particle density. In a typical data run, we would activate the particle injector and run the measurement until the transmitted intensity of the most severely attenuated beam dropped below our detection limits. We then turned off the particle injector and simultaneously activated a ventilator, allowing the particle density to diminish and the transmitted intensities to return to their unattenuated levels. Such a data run would last approximately one hour.

The particle size distribution depends on the type of particle and particle density. A laser scattering apparatus measured the size distribution of the oil vapor and dust particles. Some typical plots of size distributions for these two types of particles are shown in Fig. 4. For dust we obtained a broadband size distribution peaking at $\sim$1.1 μm and ranging from 800 nm to 5.5 μm, whereas the oil vapor distribution peaked around 850 nm with a range from 600 nm to beyond 6 μm. With increasing density the size distribution seemed to be constant for dust particle, whereas for oil vapor the smaller particle concentration rises first strongly, with a reverting effect for higher concentration.

In the next two figures, the data obtained for scattering from water fog and oil fog are shown. In the left-hand side of Fig. 5 the water fog results show a dramatic decrease in intensity for the three NIR wavelengths, whereas the right-hand side shows the extracted SNR ratio separately for each of the communication links separately. At the highest concentration (at 15:52), the 1.3- and 1.55-μm beams experienced over 40 dB losses and the 0.85 μm somewhat less.

However, the 8-μm beam experienced only 3 dB of attenuation. The corresponding SNR charts show a similar result with nearly constant SNR for the MIR link and drops

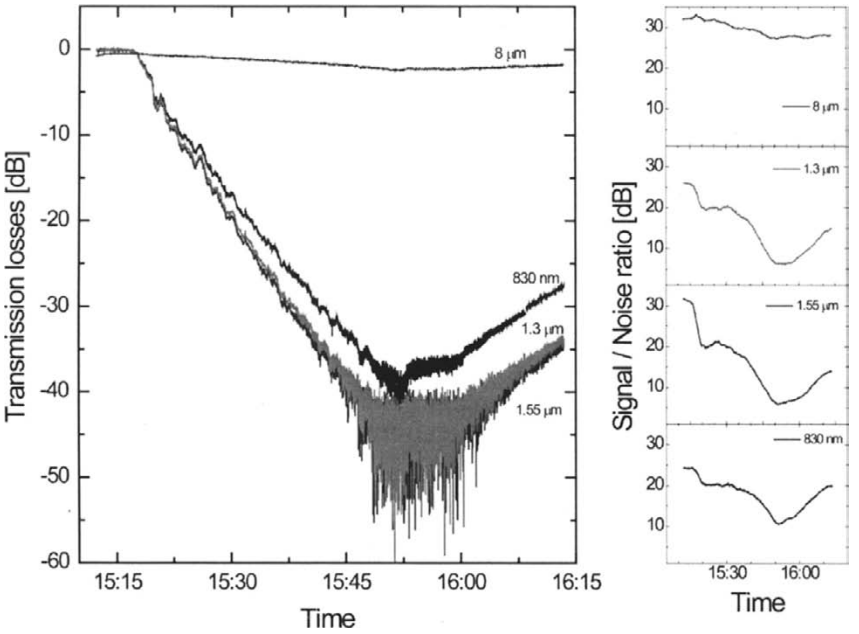


Fig. 5. Transmission losses for the four different wavelength links are shown on the right as a function time during which the water vapor concentration increased. Independent measurements show that the concentration of the particles increased linearly with time so the horizontal axis also can be interpreted as scattering particle concentration. The left side shows the resultant signal-to-noise ratio as a function of water vapor concentration.

in the signal down to (and below) 10 dB for the NIR links. Figure 6 shows the data set for oil fog. The oil fog differs from the water vapor in that the 0.85-μm beam was attenuated more than the 1.3- and 1.55-μm beams but as with the water vapor the 8-μm beam shows almost no loss under circumstances that completely attenuate the NIR beams. The difference between the water and oil fog most likely reflects differences in particle size distribution. Unfortunately, the laser scattering apparatus that we used to measure the particle distribution is unable to measure the water fog size distribution directly so a direct test of this assumption could not be made.

However, the slope of the transmission data obtained during the particle density increase phase allowed us to make a summary comparison of the attenuation efficacy of each of the scattering media as a function of the transmission wavelength. To determine the scattering efficiency for the different wavelengths the transmission losses along the length of the link were calculated assuming a scattering concentration of 1000 scatterers/cm^3. The results of this analysis are shown in Fig. 7. Although there are slight variations in scattering efficiency for the three NIR wavelengths, in all cases the 8 μm beam is almost unaffected by scattering conditions (0.06–0.4 dB), whereas the NIR beams experience much stronger attenuation ranging from ~2.5 dB for dust to 10 dB for water, having at least one order of magnitude stronger losses. The case for water droplets is thereby even more interesting, as it reveals again the different wave-

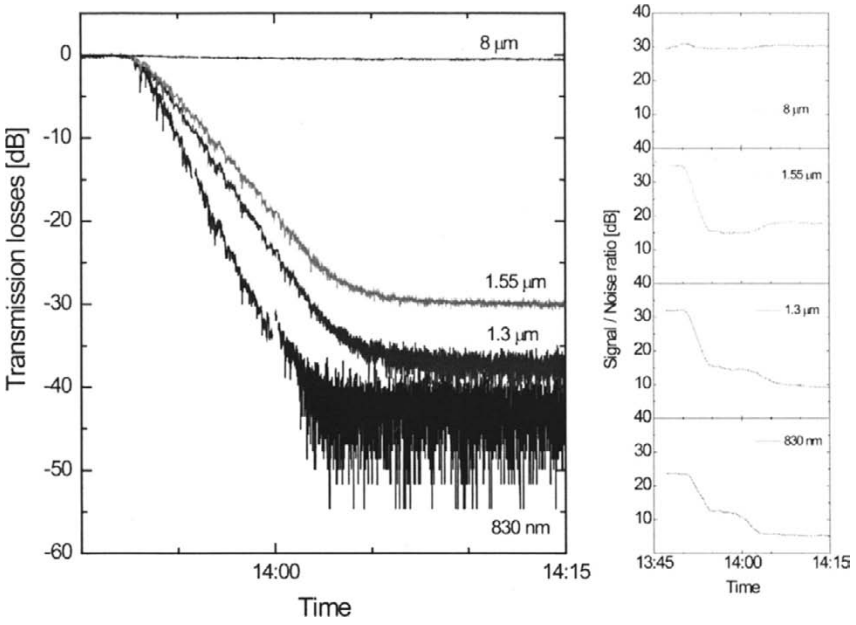


Fig. 6. Transmission losses for the four different wavelength links are shown on the right as a function of increasing oil vapor concentration. The left side shows the resultant signal-to-noise ratio as a function of oil vapor concentration.

length dependent behavior compared to dust or oil scattering: Here the link operating at 850 nm experienced less losses then the other NIR links—whereas for the other scattering system the losses seemed to be qualitatively stronger for longer wavelength. We attribute this behavior to a significantly different particle size distribution for the created water fog.

In an additional set of experiments we evaluated the influence of the scattering environment on digital communication properties by measuring bit error rates (BER) and eye diagrams for each of the four wavelengths while stepwise increasing the particle concentration. For this purpose the sinusoidal modulation was replaced by a PRBS data stream at a data rate of 34 MBit/s (corresponding to a DS-3 data transmission rate) and the detectors of the NIR wavelength links were replaced by Thorlabs D400FC detector with 1 GHz bandwidth (the MCT detector for the MIR link provided enough bandwidth for the communication measurements). The digital data were then amplified corresponding to automatic gain compensation in standardized communication systems and analyzed with a Hewlett-Packard 3784A BERT. For further evaluation corresponding eye diagrams were recorded in parallel using a Tektronix storage oscilloscope.

For the NIR communication links the signal quality decreased drastically with scatter particle density, resulting in noisier eye diagrams. Utilizing high gain (factor of 125) and based on the relatively high link margin (typically SNR > 30 dB in the start)

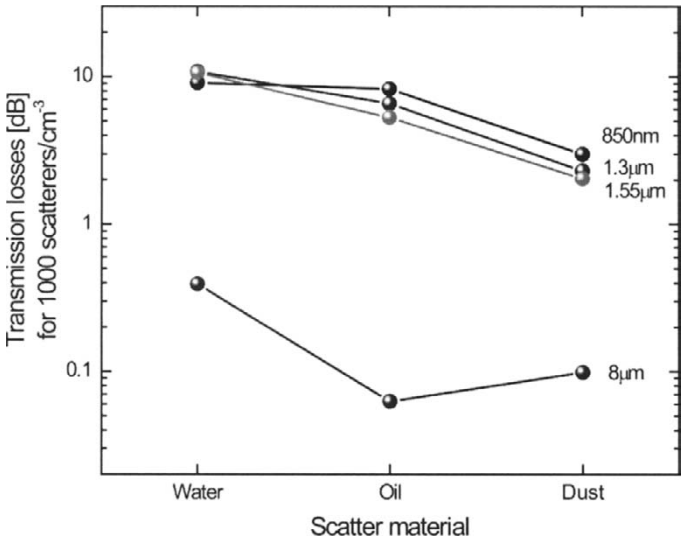


Fig. 7. Summary analysis of the data shown in Figs. 3–5. By extracting the slope of the attenuation curve during the scattering density increase phase we obtain a measure of scattering efficacy for each of the wavelengths and each of the scattering types.

error-free data transmission could still be ensured until higher scattering densities, but typically could not be sustained at the highest densities. The upper part of Fig. 8 shows typical eye diagrams obtained for the 1.3-μm link under conditions of low oil concentration (0.57 mg/cm^3) and roughly 20$\times$ higher oil concentration (11 mg/cm^3) with 125$\times$ amplification. The open eye at the low concentration has completely closed at the high concentration, and the signal has highly degraded from 1.5 V to 200 mV. On the other hand, data transmission in the MIR-link only slightly affected and no amplification was needed for the tested environments. The comparable eye diagrams for the 8 μm link, shown in the lower part of Fig. 8, shows almost no degradation and we still obtained error free transmission (BER $> 10^{-13}$) at the highest oil vapor concentration. We obtained similar results for each of the other scattering materials.

Within these experiments the performance of free space optical communication links operating at three near-infrared wavelengths (0.85, 1.3 and 1.55 μm) using commercially available telecommunications diode lasers and at 8 μm using an experimental quantum cascade laser produced at Lucent's Bell Laboratories were compared under low visibility conditions. To simulate a variety of atmospheric conditions, the four links were transmitted through the 65-foot-long measurement leg of the Pacific Northwest National Laboratory's (PNNL) Aerosol Wind Tunnel Research Facility. The wind tunnel was selectively injected with water vapor, oil vapor, and dust particles. We then measured the transmitted intensity of each of the links as the scattering density increased. We also recorded bit error rate and eye diagrams using a bit error rate tester operating at 34 MHz. The results show that each of the NIR wavelengths become strongly attenuated at the highest concentration of scattering particles whereas the

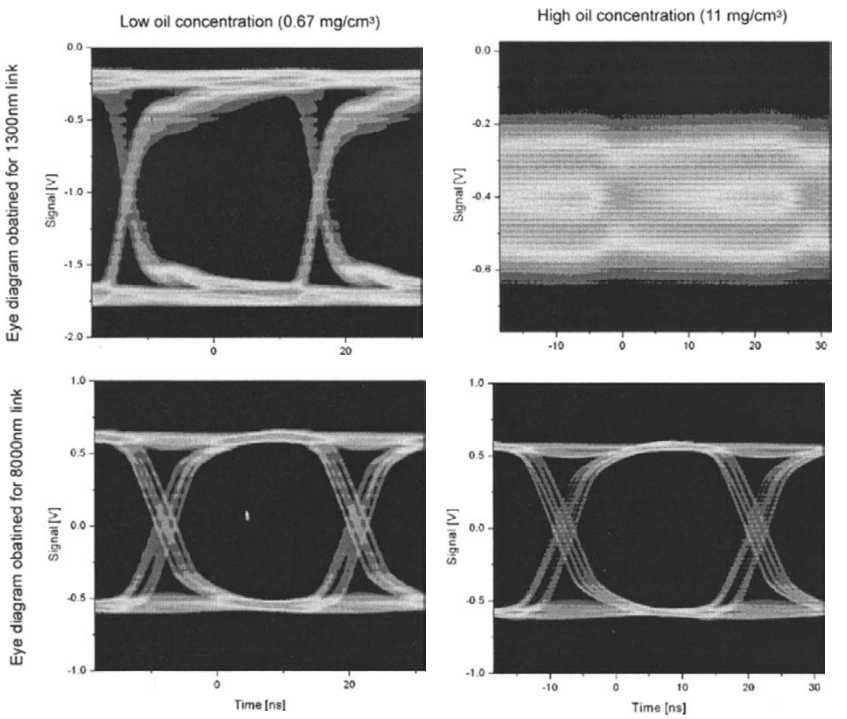


Fig. 8. The top row shows eye diagrams for the 1.33-μm laser link operating at a data rate of 34 MBit/s (DS-3) under conditions of low concentration (left) and high concentration. The bottom row shows eye diagrams for the 8 μm quantum cascade laser link operating at a data rate of 34 MBit/s (DS-3) under corresponding conditions of low concentration (left) and high concentration.

8-μm QCL does not experience significant attenuation. Likewise, the NIR links are unable to sustain error free transmission at the highest attenuation levels but the QCL link remains error free under the same particle density conditions.

3.1. Satellite TV Transmission Using a QCL-Based FSO Link

The first two parts of this chapter addressed the two dominant requirements for QCL based FSO links: high modulation bandwidth and high transmissivity in scattering environments. In this final section we report on the application of a QCL–FSO link for high-bandwidth satellite TV-data transmission as described in [12]. The upper part of Fig. 9 shows the optical setup of the transmission link. The same QCL's described in the modulation experiment were employed in this setup. The QCL beam was collimated using an f/3 ZnSe-lens and then transmitted over an open-air 100m path to a retroreflector, mounted on another building of Bell Laboratories in Murray Hill, New Jersey. The reflected light was collected using an $f/9$ telescope with an aperture of 76 mm and focused onto a high-speed liquid nitrogen cooled MCT detector (Sagem HgCdTe 011). To compare the effect of the longer wavelength on the link quality and

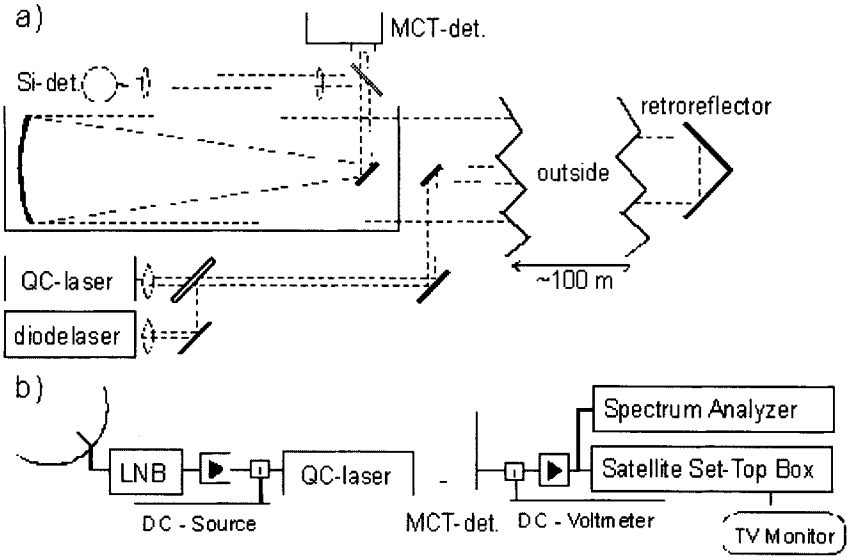


Fig. 9. Schematics of the optical and electrical setups of the transmission links.

stability, a second beam was included in the path, originating from a 0.85 μm diode laser (10 mW output power) and detected with a standard Si-detector. To ensure an identical beam path and easy adjustment the optics for the outgoing beam and for the detection were rigidly connected to the telescope.

The lower part of Fig. 9 shows the electrical setup of the QC-laser and of the detection system. The signal was received from a satellite dish using a LNB (low noise block down converter)-module. This high-frequency signal (750 MHz–1.45 GHz) was combined with a dc current to drive the QC-laser continuously above its threshold. The modulated laser radiation was transmitted over the total distance of ~200 m before it was detected. The dc component was split off with a bias-Tee and used as monitor of the received laser intensity. The high- frequency part was amplified and fed into a spectrum analyzer as well as into a standard satellite set-top-box connected to a TV monitor.

Under typical QC laser operating conditions (500 mA dc current at a temperature of 25 K) the link could be run continuously and stably for at least five hours. Owing to the beamsplitter and the multiple optical elements in the outgoing beam path only 7.5 mW of the initial 25-mW output power were actually used for the transmission. About 10% of the original intensity (0.75 mW) could still be detected under good weather conditions after the transmission and the collecting telescope optics. The simultaneously transmitted beam with a wavelength of 0.85 μm had comparable losses, which are therefore attributed to beam spreading and losses in the optical elements.

A typical example of the transmitted data stream is shown in the upper part of Fig. 10. The modulation in the frequency region from 900 MHz to 1.45 GHz contains the digitally encoded information (QPSK-code: quadrature phase shift keying) consisting of around 800 television channels and 100 radio channels. The dashed line in Fig. 10

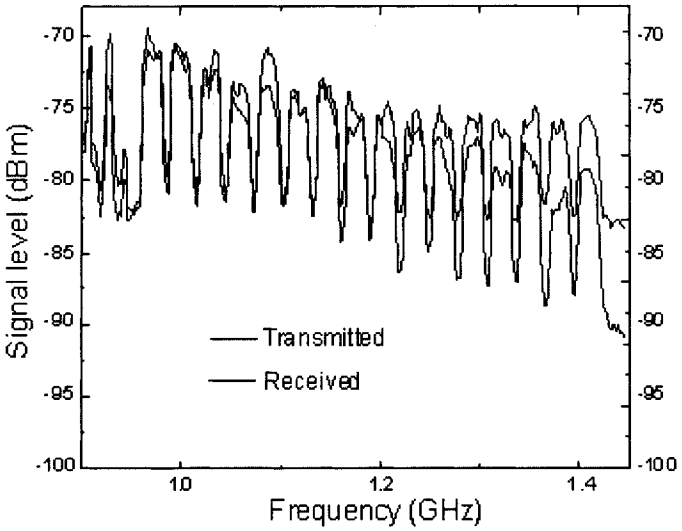


Fig. 10. Signal at the output of the LNB (upper curve, dashed) shown together with the signal after transmission over the free-space link (solid line).

represents the original signal received from the LNB, whereas the solid line shows a typical signal detected after the QC laser-link. Owing to the limited bandwidth of the detector, the channels in the higher frequency region were detected with a 10dB higher loss, reducing the number of actually decodable channels to around 650. The link power margin is 7 dB corresponding to a received power of 0.125 mW, below which the receiver becomes unstable.

To evaluate the advantage of the longer wavelength relative to the collinearly propagating near-infrared beam (0.85 μm) the intensity of the latter was monitored in parallel. For typical weather conditions including clear air, strong rain as well as a thunderstorm, no differences in sensitivity were observed. Nevertheless, a strongly pronounced deviation was observed during a dense fog situation, with nearly zero visibility. Figure 11 shows the temporal evolution of the detected dc intensities for both laser links, starting at a very dense fog situation in the early morning of 8/14/01. As the fog lifted slowly around 3:15 AM. the QC-laser link regained transmission much quicker than the near-IR link. The QC-laser link had reached transmission of nearly 70% of its optimal value, when the intensity of the near-IR link was still below the detection limit. As a result around 4:00 AM, the mid-IR television link became stable again, whereas the near-IR link was still unstable almost for another hour. The superior performance of the QC-laser link compared to the near-IR link can readily be understood from the wavelength dependence of Rayleigh- and Mie-scattering and relates well to the aforementioned windtunnel measurement.

Following this measurement the conclusion has to be drawn, that reliable free space optical communications should be possible using the QCL under conditions that would make communications impossible using shorter wavelengths. As detector and

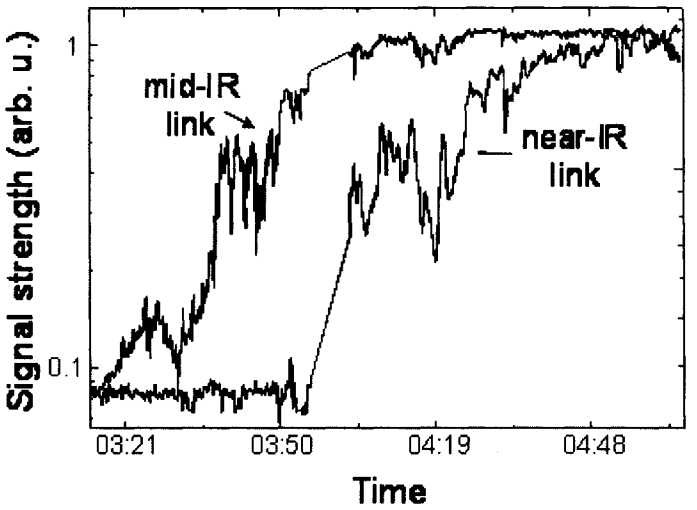


Fig. 11. Comparison of the received intensities of the mid-IR ($\lambda = 8.1$ μm) and the near-IR ($\lambda = 0.85$ μm) link as a function of time. Fog came up around 2 AM and progressively dissipated during the measurement.

laser technology continue to improve in the mid-infrared spectral range, this result is likely to be a driver for commercial development of such links.

Acknowledgments

The work reported in this chapter represents the results of collaborative efforts with the Bell Laboratories-Lucent Technologies Quantum Cascade Laser group (F. Capasso, C. Gmachl, R. Paiella, H.Y. Wang, D. L. Sivco, J.N. Baillargeon, A.Y. Cho), the remote sensing group at Pacific Northwest National Laboratory (W.W. Harper, Y-F. Su, J.F. Shultz), and H. C. Liu from National Research Council of Canada and we gratefully acknowledge the contributions of these collaborators. Portions of this work were supported by the United States Department of Energy under contract DE-FG03-99NV13656 and by the US Army/DARPA under contract DAAD19-00-C-0096. R. Martini acknowledges the support of Bell Laboratories-Lucent Technologies during the early portions of this research.

References

1. Killinger, D., Free Space Optics for Laser Communication Through the Air, Optics and Photonics News, October 2002, pp. 36–42.
2. Capasso, F., Gmachl, C., Paiella, R., Tredicucci, A., Hutchinson, A.L., Sivco, D.L., Baillargeon, J.N., Cho, A.Y., and Liu, H.C., New Frontiers in Quantum Cascade Lasers and Applications, IEEE J. Select. Top. Quant. Electron., **6**, 931–947 (2000) and references therein

3. Capasso, F. et al., IEEE J. Quant. Electron., **38** (6), 511–532 (2002).

4. Faist, J., Capasso, F., Sirtori, C., Sivco, D.L., and Cho, A.Y., Quantum Cascade Lasers, Chapter VIII, in *Intersubband Transition in Quantum Wells: Physics and Device Application II*, edited by Liu, H.C. and Capasso, F. (Academic Press, London, 2000), pp. 29–40.

5. Paiella, R., Capasso, F., Gmachl, C., Bethea, C.G., Sivco, D.L., Baillargeon, J.N., Hutchinson, A.L., Cho, A.Y., and Liu, H.C., Generation and Detection of High-Speed Pulses of Mid-Infrared Radiation with Intersubband Semiconductor Lasers and Detectors, IEEE Photon. Technol. Lett., **12**, 780– 782 (2000).

6. Paiella, R., Capasso, F., Gmachl, C., Sivco, D.L., Baillargeon, J.N., Hutchinson, A.L., Cho, A.Y., and Liu, H.C., Self-mode-locking in quantum cascade lasers with giant ultrafast optical nonlinearities, Science, **290**, 1739–1742 (2000).

7. Paiella, R., Martini, R., Capasso, F., Gmachl, Wang, H.Y., Sivco, D.L., Baillargeon, J.N., Cho, A.Y, Whittaker, E.A., and Liu, H.C., High-frequency modulation without the relaxation oscillation resonance in quantum cascade lasers, Appl. Phys. Lett., **19**, 2526–2528 (2001).

8. Szajowski, P.F., Nykolak, G., Auburn, J.J., Presby, H.M., Tourgee, G.E., Korevaar, E., Schuster, J., and Kim, I.I.,: 2.4 km Free-Space Optical Communication 1550 nm Transmission Link Operating at 2.5 Gb/s—Experimental Results in *Optical Wireless Communications*, 1998, edited by Korevaar, E.J., Proc. SPIE 3552, pp. 29–40.

9. Martini, R., Gmachl, C., Falciglia, J., Curti, F.G., Bethea C.G., Capasso, F., Whittaker, E.A., Paiella, R., Tredicucci, A., Sivco, D.L., and Cho, A.Y., High-speed modulation and free-space optical audio/video transmission using quantum cascade lasers, Electron. Lett., **37**, 111–112 (2001).

10. Blaser, S., Hofstetter, D., Beck, M., and Faist, J., Free-space optical data link using Peltier-cooled quantum cascade laser, Electron. Lett., **37**, 778–780 (2001).

11. Liu, H.C., Lin, J., Buchanan, M., and Wasilewski, Z.R.,: High-Frequency Quantum-Well Infrared Photodetectors Measured by Microwave-Rectification Technique, IEEE J. Quantum Electron., **32**, 1024–1028 (1996).

12. Martini, R., Gmachl, C., Falciglia, J., Curti, F.G., Bethea C.G., Capasso, F., Whittaker, E.A., Paiella, R., Tredicucci, A., Sivco, D.L., and Cho, A.Y., High-speed modulation and free-space optical audio/video transmission using quantum cascade lasers, Electron. Lett., **37**, 111–112 (2001).

13. Blaser, S., Hofstetter, D., Beck, M., and Faist, J., Free-space optical data link using Peltier-cooled quantum cascade laser, Electron. Lett., **37**, 778–780 (2001).

14. Martini, R., Gmachl, Paiella, R., Capasso, F., Whittaker, E.A., Liu, H.C., Hwang, H.Y., Sivco, D.L., Baillargeon, J.N., and Cho, A.Y., High-speed digital data transmission using mid-infrared Quantum Cascade lasers, Electron. Lett., **37**, 1290–1291 (2001).

DOI: 10.1007/s10297-006-0078-0
Originally published in J. Opt. Fiber. Commun. Rep. 4, 214–224 (2007)

All-weather long-wavelength infrared free space optical communications

D.P. Hutchinson and R.K. Richards

Engineering Science and Technology Division
Oak Ridge National Laboratory
Bethel Valley Road, Oak Ridge, TN 37831

Abstract. Long-wavelength infrared radiation possesses superior transmission through common atmospheric problems such as fog, clouds, and smoke than the shorter wavelength laser sources in use today. Recent improvements in LWIR laser and modulator design makes possible reliable optical replacements for radio and microwave communications links in many applications. This paper describes components and techniques developed for high-speed, full-duplex all-weather infrared communications systems.

1. Introduction

Long-wavelength infrared radiation possesses better all-weather transmission than the shorter wavelength laser sources in use today. The superior transmission through common atmospheric problems such as fog, clouds, and smoke coupled with improvements in LWIR laser and modulator design makes possible reliable optical replacements for radio and microwave communications links in many applications Another advantage of LWIR laser radiation is the inherent eye safety of this wavelength region. To take advantage of the transmission characteristics of infrared radiation in the atmosphere, we have developed components for a high-speed, full-duplex all-weather infrared communications link for ranges up to 12 km. In support of this program, we have designed and constructed a compact, RF-driven waveguide CO_2 laser and a dielectric-waveguide Stark modulator. The modulator is based on the Stark shift of NH_2D (deuterated ammonia). The laser is an RF-driven waveguide design.

2. Atmospheric Transmission: The Case for LWIR

Transmission of optical radiation through the atmosphere is reduced by molecular absorption and aerosol extinction (absorption + scattering). Absorption and scattering are highly wavelength dependent phenomenon. For this reason, selecting a laser wavelength may be the most important variable when considering a given communications application. This section on atmospheric transmission includes four subsections that we feel are the most important aspects to consider for an infrared communications and/or lidar application on the battlefield. The first subsection, molecular absorption, uses the well known HITRAN [1] database to show the atmospheric transmission superiority of the $^{13}C^{16}O_2$-isotope wavelengths ($\sim$10.8 μm) over the standard $^{12}C^{16}O_2$ laser wavelengths ($\sim$10.6 μm). The remainder of the subsections deal with aerosols. Aerosols, particles suspended in the atmosphere, broadly encompass everything from microscopic dust to macroscopic fog droplets. A baseline continental aerosol extinction model is presented, then smoke and fog are covered individually. Aerosol extinction models for CO_2 wavelengths are plotted in comparison to visible, and when possible, 1.55 μm wavelengths. The aerosol extinction coefficients available for the standard $^{12}CO_2$ wavelengths are assumed to be representative of the $^{13}CO_2$ isotope wavelengths as well. Owing to the numerous possible combinations of variables within each subsection one or two representative graphs will be provided and trends and exceptions will be noted in the paragraphs.

2.1. *Molecular Absorption Model*

The molecular absorption study was conducted with the computer program HITRAN-PC [2]. The transmission calculations, shown in Figs. 1 and 2, were preformed with the 1976 U.S. Standard for partial pressures of the constituent gases in air. The roundtrip path length used in the calculations is 12 km. The calculations used a standard pressure of 1 atm and temperature of 296°K. Figure 1 shows the atmospheric transmission around the standard $^{12}C^{16}O_2$ laser wavelengths. The most intense P (20) line at 10.59 μm only provides a 40% transmission. Other nearby laser lines are also shown in Fig. 1, and a similar reduction in radiation intensity is also experienced due to the natural abundance of $^{12}C^{16}O_2$ in the atmosphere. Figure 2 shows a very promising alternative, the $^{13}C^{16}O_2$-isotope. The three specific laser lines shown in Fig. 2 are the R(20), R(18), and R(16) at the wavelengths of 10.768, 10.784, and 10.80 μm, respectively. Each of these laser lines provides more than 80% transmission over a total distance of 12 km at the given conditions.

In addition, a few of the other R-branch transitions of this isotope also provide over 95% transmission.

2.2. *Smoke Extinction Model*

The smoke extinction model employed in this subsection is built around optical theory of small particles that group together to form larger particles. For the sake of clarity, the small particles will be referred to as spherules and the larger particles will be referred to as clusters. Each of these clusters has a fractal dimension, which describes how well packed the individual spherules are within the cluster. A minimum fractal dimension of

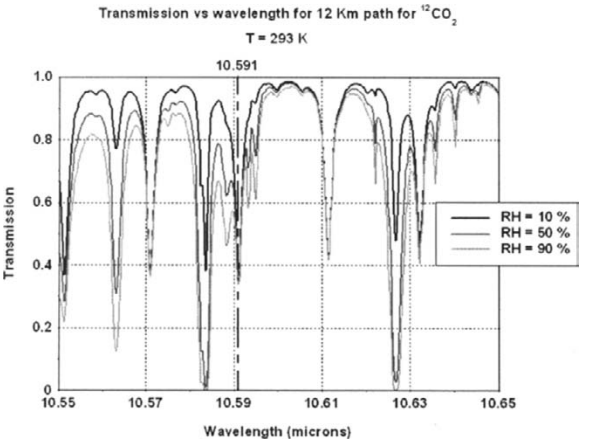


Fig. 1. Atmospheric transmission for a 12-km range.

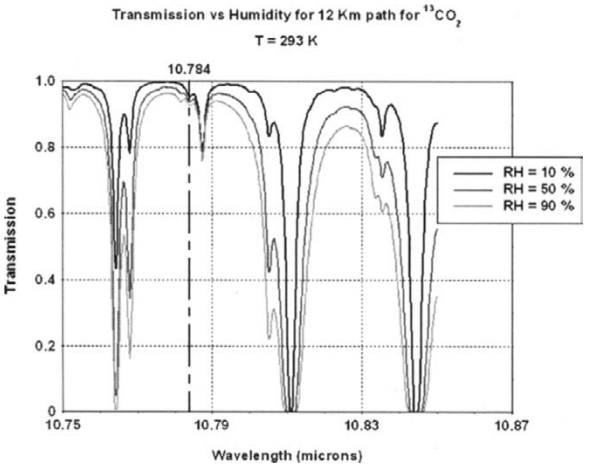


Fig. 2. Transmission of three lines for the Carbon-13 isotope of carbon dioxide.

1 refers to an array of just barely touching spherules, and a maximum fractal dimension of 3 refers to a totally compact cluster with no voids. A fractal dimension, D, of 1.78 has been accepted for smoke [3]. The average number of individual spherules in each cluster is given by [4]:

$$N = (R/a)D, \tag{1}$$

where R is the average cluster radius and a is the average spherule radius.

The absorption cross-sections were calculated for vector waves [4]:

$$\sigma_a = 4\pi N k a^3 \mathrm{Im}[(M^2 - 1)/(M^2 + 2)], \tag{2}$$

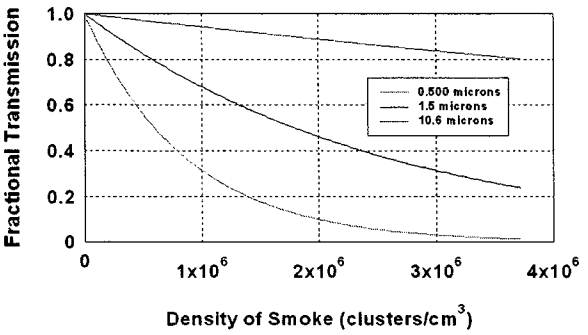


Fig. 3. Comparison of transmission through smoke.

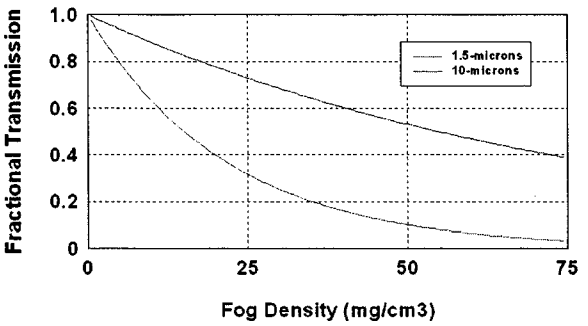


Fig. 4. Fractional transmission of 1.5 and 10 μm light through 100 m of fog.

where $k = 2\pi/\lambda$ and M is the complex index of refraction at wavelength λ.

The transmission of radiation is calculated:

$$T = e^{-\sigma\rho L}, \tag{3}$$

where σ is the extinction cross-section, ρ is the density, and L is the total path length. The average spherule radius is around 0.01 μm [4]. For smoke the absorption cross-section is much larger than the scattering cross-section, thus σ_a can be taken as a good estimate of the extinction cross-section σ.

Figure 3 shows the transmission of 10-μm, 1.5-μm, and 500-nm radiation versus smoke density. The total path length is 1000 m. All calculations are for a cluster size of R =0.1 μm. The indices of refraction used for smoke particles for 10-μm and 500-nm were $M_{10} = 3 + i$ and $M_{0.5} = 1.75 + .3i$, respectively [4]. Since Berry and Percival did not provide an index at 1.5 μm for soot particles one was calculated by using the same value as for visible light, i.e., $M_{1.5} \approx M_{0.5} = 1.75 + .3i$.

Other sources report smoke particle size within the range of 0.005–0.15 μm range [4]. This reduction in cluster size would correspondingly yield much longer path

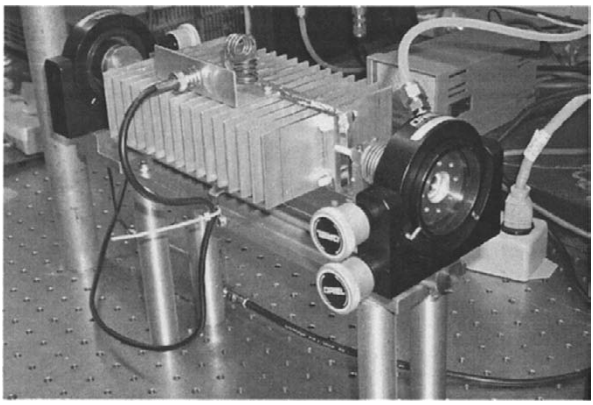


Fig. 5. Air-cooled, RF-driven CO_2 Laser.

lengths (for all wavelengths), but the wavelength trends shown in Fig. 3 will stay the same. With this in mind, Fig. 3 is best used as a relative comparison of the superiority of 10-μm to visible and 1.5-μm radiation.

2.3. Fog Extinction Model

The estimate for extinction of 10- and 1.5-μm radiation by fog is shown in Fig. 4. These extinction coefficients were obtained from graphical data presented in Rensch and Long [5]. The extinction coefficients for 10- and 1.5-μm radiation were 0.55 and 2.0 (dB/km)/(mg/m^3), respectively. The extinction coefficient chosen above from Rensch and Long for 10-μm seems to agree within 10% with the experimentally measured coefficients in Chimelis [3,6]. The average water droplet radius for the fog is 1 μm and the path length is 100 m.

3. Component Development for LWIR Communications

3.1. Compact RF-Driven Laser

One of the drawbacks to more widespread applications for long wavelength infrared communications systems is an inexpensive, compact source of 10-μm radiation. We have developed a compact CW RF-driven, air-cooled, sealed-off waveguide CO_2 laser featuring a power level of over 1 W. The hollow Al_2O_3 ceramic waveguide has an i.d. of 2.4-mm and provides a gain length of 20 cm. Our waveguide laser is based on a design reported by Walsh [7]. A photograph of the laser is shown in Fig. 5. This laser produces a power level of approximately 1.6 W using $^{12}CO_2$ and 0.8 W using $^{13}CO_2$, both in the EH_{11} dielectric waveguide mode. With a grating installed, our laser produces 0.4 W on the 10.59-μm line of $^{12}CO_2$. The output mode structure of the laser, measured with a thermally sensitive liquid crystal film, appears to be a TEM_{00}

Gaussian mode. Frequency measurements indicate that the laser operates in a single transverse and single longitudinal mode.

The gas mixture in the laser is He:CO_2:N_2:CO in the ratio of 65:18:15:2, respectively. This is a commercial mixture purchased for a pulsed CO_2 laser and the composition has not yet been optimized for our CW waveguide laser. Walsh used a mixture of He:CO_2:N_2:Xe in the mixture ratio 77:10:10:3, respectively.

One end of the laser cavity consists of a 0.5-in.-diameter, 3-m-radius concave ZnSe 95% reflective output coupler attached to the waveguide through a brass bellows. The ceramic waveguide is attached with epoxy to a brass fitting soldered to the bellows. Epoxy is also used to attach a ZnSe Brewster window is to the other end of the ceramic tube, which was ground to the Brewster's angle. A 150 l/mm flat master grating mounted on a piezoelectric actuator in an adjustable mirror mount forms the other end of the cavity. A 1-in.-diameter invar rod machined flat on opposing sides serve as a temperature stable mounting surface for the insulating support for the electrodes and the mirror mounts. The optimum operating pressure of the sealed-off laser is 60–65 torr. The laser is driven by a 58.5-MHz RF amplifier at a power level of approximately 50 W. Machined aluminum heatsinks are mechanically clamped to each side of the waveguide serve as RF electrodes and provide cooling for the tube. The two electrodes are shaped to conform to the round dielectric waveguide. A thin coating of heatsink compound applied to the electrodes during assembly improves thermal contact with the dielectric waveguide and enhances cooling. One side of the electrode is grounded and an air wound autotransformer couples RF power to the other electrode. The inductance of the 4-turn autotransformer resonates with the capacitance formed by the heatsinks attached to the waveguide at a frequency of 58.5 MHz. Another air-core inductor is used to connect the 50-Ω coaxial cable from the RF power supply to a tap (approximately 1-turn from the grounded end) on the autotransformer to efficiently couple power to the laser. The inductance of the coupling coil is adjusted by compressing or expanding the coil to optimize the impedance match to the 50-Ω cable. Both coils are constructed from #12 AWG copper wire.

3.2. *Stark-Effect Modulator*

Stark-effect modulation occurs when an electric field is applied to a gas molecule that has a substantial polarization. The applied electric field effectively changes the energy spacing of the molecular levels changing the optical frequency or wavelength that is absorbed by the gas. Also, the energy spacing is very small compared to the energy of the optical photon interacting with the gas. The modulator is filled with approximately 2 or more torr of partially deuterated ammonia (NH_2D), which has a molecular absorption resonance near the 10.59-μm wavelength line of emission of a CO_2 laser. The frequency difference between the absorption resonance and the laser line is reported to be approximately 2189-MHz [8] from the laser line. The dotted curve in Fig. 6 depicts the transmission of a 30-cm cell containing 2-torr of NH_2D in the absence of an applied electric field. The laser light (located at zero difference frequency on the scale) is not strongly absorbed by the cell. The solid curve in the left graph in Fig. 6 shows the transmission of the same cell in the presence of an applied electric field of 300 V/mm. The electric field causes this absorption line to split into nine Stark components classified by the designation $M = 0, \pm 1, \pm 2, \pm 3$, and ± 4. The $M = 0$ component is roughly in the center of the plot at a difference frequency of

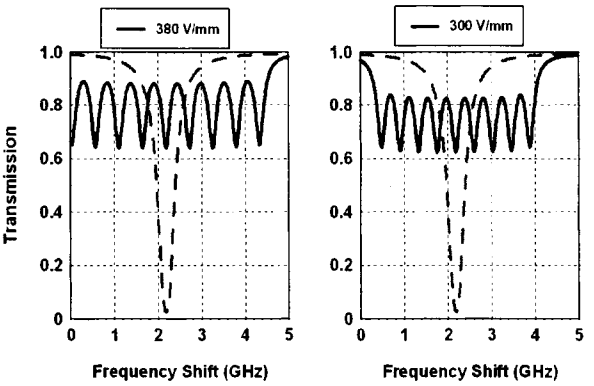


Fig. 6. The application of an electric field to deuterated ammonia causes the absorption peak to split into nine components.

2189 MHz and the $M = \pm 4$ components are at the extreme left and right ends of the structure. As the electric field is increased the $M = \pm 4$ absorption component moves closer to zero difference frequency until the peak of the absorption is in coincidence with the laser line corresponding to an electric field of approximately 380 V/mm. This situation is depicted in left graph shown in Fig. 6. As the applied electric field is varied from 350 to 380 V/mm, the transmission of the 30-cm-long cell varies from 90% to 75%. If the length of the cell is increased, the modulation amplitude can exceed 70%. If we impressed a steady state or dc electric field on the electrodes half-way between 350 and 380 V/mm or 365 V/mm and applied a sinusoidal alternating or ac voltage with an amplitude of 30-V peak-to-peak, a sinusoidal modulation of the laser beam would occur. Digital modulation can be achieved by setting the steady field at 350 V/mm and applying a 30-V digital signal to switch the transmission of the cell between the 75% and 90% states.

3.3. Dielectric Waveguide Modulator

Normally Stark modulators are constructed from two parallel electrodes separated by a distance much smaller than the width to minimize the variation of the electric field across the laser beam. The electrodes are placed inside of a dielectric tube, typically glass or ceramic, held at a pressure appropriate for the gas used for modulation, typically a few torr. A typical spacing would be on the order of 2-mm with a width of 20 or more mm and a length of more than 200 mm. One of the problems with this design is that normal expansion of the laser beam through such a structure causes a loss of laser light due to vignetting of the beam by the electrodes. To reduce this vignetting loss, we have designed a Stark modulator using a hollow glass dielectric tube to confine the optical beam with minimum loss. The laser beam is focused into the proper size by a lens to form a match to the EH_{11} waveguide mode of the dielectric tube. A lens following the waveguide collimates the beam for propagation to a detector or to other optical components. The electrodes are external to the dielectric tube that serves to

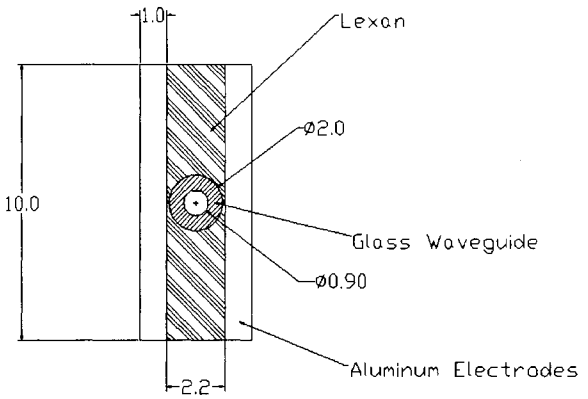


Fig. 7. End view of 0.9-mm i.d. glass modulator. All dimensions are in millimeters.

confine the optical beam and provide a means to operate the modulator at the reduced pressures required for operation.

To test this concept we constructed a 0.9-mm i.d. × 20-cm dielectric waveguide modulator. The waveguide material is borosilicate glass and is suspended between two 1-cm-wide by 20-cm-long flat aluminum electrodes. Lexan inserts are used to fill the gap between the electrodes outside of the waveguide. These Lexan inserts support the waveguide, align the electrodes and flatten and enhance the electric field in the waveguide. A sketch of the cross-section of the modulator is shown in Fig. 7. The ends of the waveguide we are sealed with ZnSe Brewster windows.

One problem that arises with this external electrode design is charge build-up from stray ions and electrons in the insulating waveguide. When an electric field is applied to the electrodes, the stray free charge in the waveguide migrates to the sides of waveguide nearest the electrodes with electrons attracted to the most positive electrode and ions attracted to the most negative electrode. This charge migration acts to cancel the electric field inside the waveguide thereby canceling modulation. We measured the time required for this canceling charge to build-up and found that it is on the order of or less than 1 sec. Because the charge build-up requires a finite time to occur, we have developed a biasing method that negates the effect of the stray canceling charge.

3.4. AC Biasing Method

The Stark components are created when an electric field is applied to the molecule. For the deuterated ammonia line that is in near coincidence with the 10.59-μm CO_2 laser line, the selection rules for absorption require that the electric field be applied perpendicular to the electric field of the optical beam. The polarity of the electric field is not important. For instance, if the electric field is applied with the positive electrode on the right and the negative electrode on the left or vice-versa, the absorption structure is not changed. The absorption for plus and minus M components are the same. The absorption is proportional only to the magnitude of the electric field, not direction, as long as the proper polarization relationship between the applied field and the field of

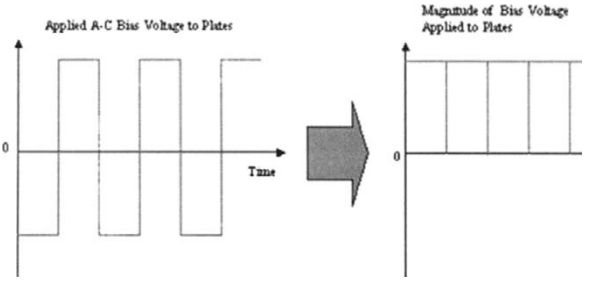


Fig. 8. The applied ac bias on the left produces the equivalent to a steady "dc" bias as shown on the right of the diagram.

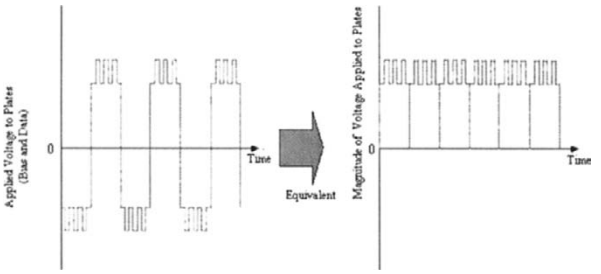


Fig. 9. The polarity of the data waveform must match the polarity of the bias waveform to properly modulate the laser beam.

the laser beam is maintained. Therefore a circuit has been developed that generates a low frequency (on the order of 100 to 1000 Hz) square wave that is applied to the electrodes. The waveform produced by the "ac bias" circuit is shown in Fig. 8. The polarity of the data waveform must match the polarity of the bias waveform to properly modulate the laser beam. This relationship is shown in Fig. 9. In addition the logic circuits buffer and temporally store the data and transmit the data waveform during the constant portion of the bias signal. High voltage amplifiers such as the Apex PA-94 have slew rates of approximately 1000 /μsec. Assuming a bias voltage level of 500 V and a bias frequency of 100 Hz, the switching transient will cause a duty cycle loss of only 0.05%. Using this type of biasing and data amplifiers, the waveform transmitted by the modulator should resemble the waveform shown in Fig. 10. The bandwidth of the modulator is limited only by the homogeneous linewidth of the deuterated ammonia gas. Using a filling pressure of 8-torr, a data rate of over 300-Mbits/sec is available.

4. Experimental Results

Using 0.9-mm i.d. $\times$ 20-cm dielectric waveguide modulator design discussed above and shown in Fig. 7, we measured the transmission of the modulator at pressures of 8, 11

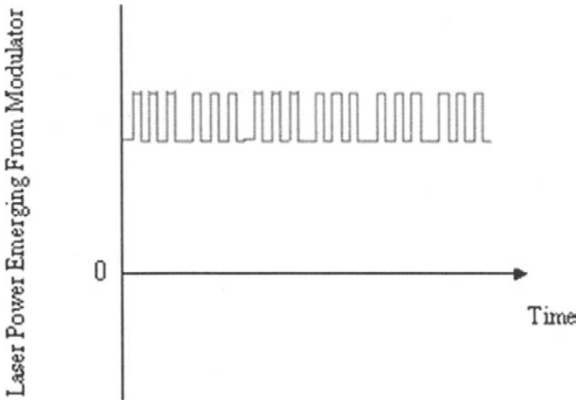


Fig. 10. The total waveform produced by the ac bias circuit combined with the data circuit appears identical to that produced using a dc bias voltage.

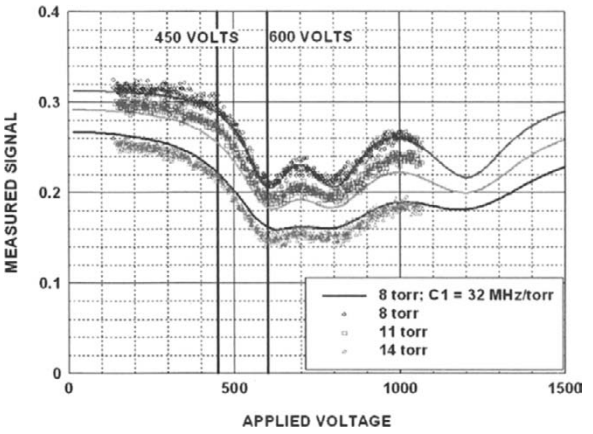


Fig. 11. The measured transmission of the 20-cm-long cell shows excellent agreement with theory.

and 14 torr as a function of the applied voltage. An "ac bias" waveform was generated by applying a square wave drive voltage from a signal generator to an iron-core step-up transformer (a 110 VAC to 6.3 VAC filament transformer operated in reverse). The bias was scanned from 0 to 1100 V peak and the transmitted signal was measured with a pyroelectric power meter. The "ac bias" frequency was 3000 Hz. The resulting modulator transmission as a function of applied voltage is shown in Fig. 11. The open circles are the measured data and the solid curves are the calculated transmission based on a theoretical model taking into account the cell length, pressure, e-field enhancement and deuterated ammonia fraction. The $M = 3$ and $M = 4$ absorption

peaks are clearly visible. The peak modulation index for this cell is approximately 30%. Using this information and verification of the model, we have designed and are constructing a 1-mm i.d. × 36-cm modulator for a prototype free-space optical link for high-speed digital transmission experiments.

References

1. 1996 HITRAN Molecular Line Database, U.S. Air Force.
2. HITRAN-PC, version 2.51, distributed by Ontar Corporation.
3. M.V. Berry and I.C. Percival, "Optics of fractal clusters such as smoke," Optica Acta **33**, 5 (1986).
4. P. Chylek, V. Ramaswamy, R. Cheng, and R.G. Pinnick, "Optical properties and mass concentration of carbonaceous smokes," Appl. Opt. **20**, 17 (1981).
5. D.B. Rensch and R.K. Long, "Comparative studies of extinction and backscattering by aerosols, fog, and rain at 10.6 μ and 0.63 μ," Appl. Opt. **9**, 7 (1970).
6. V. Chimelis, "Extinction of CO_2 laser radiation by fog and rain," Appl. Opt. **21**, 18 (1982).
7. C.J. Walsh, "An rf excited circular waveguide CO_2 laser," Rev. Sci. Instrum. **61**, 9 (Sept. 1990).
8. A.R. Johnston and R D.S. Melville, "Stark-Effect Modulation of a CO_2 Laser by NH_2D," Appl. Phys. Lett. **19**, 12 (1971).