TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology

Diplomarbeit

Channel Modeling for Free Space Optical Transmission of Satellite Data

ausgeführt zum Zwecke der Erlangung des akademischen Grades
Dipl.-Ing.

unter der Leitung von
Univ.Prof. Dipl.-Ing. Dr.-Ing. Norbert Görtz
Institute of Telecommunications

Univ.Prof. Dipl.-Ing. Dr.sc.techn. Georg Schitter Dipl.NDS ETHZ
Institut für Automatisierungs- und Regelungstechnik

Univ.Ass. Dipl.-Ing. Andreas Sinn, BSc
Institut für Automatisierungs- und Regelungstechnik

eingereicht an der Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

von
Johannes Steinbach
Matrikelnummer 01326094

Wien, im November 2019 _____

Abstract

Free space optical satellite communication systems offer numerous advantages to satisfy today's and future demands of satellite up- and downlink. In comparison to conventional RF communication they provide smaller wavelengths, low-complexity components and the absence of spectrum licensing. However, the optical frequency regime is more susceptible to atmospheric influences and therefore requires integral, well balanced system design and operation. This work is concerned with atmospheric optical links, where turbulence distorts the signal propagating through an inhomogeneous medium. Modeling this atmospheric influence is a key step when designing a communication system – in order to ensure its operation as well as to save costs by quantifying the effects of different design choices.

This thesis analyzes Gaussian and log-normally distributed fading coefficients in depth – the respective optimal symbol detectors for on-off keying are derived and the resulting channel capacity is calculated. Furthermore, models for the temporal correlation are discussed and detection strategies are compared.

An optical link of 600 m total length in urban environment forms the physical side of this work. The fading and error statistics of the channel are extracted and display evidence for the log-normal model. The observed scintillation indices are below 0.1, classifying the link as weakly turbulent. Several methods determining the capacity of the optical link are contrasted. A systematic approach for Gilbert burst error model fits based on measured error statistics is presented and the resulting capacity evaluated.

Zusammenfassung

Sowohl satellitengebundene als auch terrestrische Kommunikation wird über Funk- oder optische Kanäle geführt. Letztere haben die Vorteile kürzerer Wellenlängen, simpler Bauelemente und der lizenzfreier Frequenzbänder. Diese Diplomarbeit behandelt atmosphärische optische Kanäle, in denen Turbulenzen das sich durch ein inhomogenes Medium ausbreitende Signal verzerren. Die Modellierung dieses Einflusses ist ein wichtiger Schritt beim Entwurf eines Kommunikationssystems – um die ordnungsgemäße Funktion sicherzustellen wie auch durch Bezifferung von Effekten verschiedener Entscheidungen Kosten zu sparen.

Die vorliegende Arbeit analysiert normalverteilte sowie log-normalverteilte Übertragungskoeffizienten im Detail – die jeweiligen optimalen Symboldetektoren für Amplitudenumtastung (OOK) werden hergeleitet und die resultierende Kapazität wird berechnet. Weiters werden Modelle für die zeitliche Korrelation des Übertragungskoeffizienten betrachtet und Detektionsverfahren verglichen.

Ein optischer Kanal einer Gesamtlänge von 600 Metern in städtischer Umgebung bildet den physikalischen Aspekt dieser Diplomarbeit. Die Verteilung von Bitfehlern und des Übertragungskoeffizienten wird bestimmt und liefert Evidenz für das log-normalverteilte Modell. Die Messungen ergeben einen Scintillationsindex kleiner als 0.1, wodurch der Kanal als schwach turbulent klassifiziert wird. Verschiedene Ansätze, die Kapazität des Kanals zu bestimmen, werden gegenübergestellt und ein systematischer Zugang, die Parameter eines Gilbert-Bündelfehlerkanals anhand gemessener Fehlerverteilungen zu schätzen, wird angeführt und die daraus resultierende Kapazität bestimmt.

Acknowledgements

I would like to thank my patient advisors, Prof. Norbert Görtz, Prof. Georg Schitter and DI Andreas Sinn, for their time and guidance. I also want to express my gratitude for the numerous faculty members who go out of their way to help or teach interested students.

I also owe special thanks to my colleague Dr. Franz Stögerer for his kind nature and being a great example.

Moreover, I want to express my gratitude to my great teachers DI Bettina Burger, the person who probably made the biggest impact on my life; and DI Zsolt Nagy, who greatly lead his class despite having to put up with us.

I cannot overstate the positive effect spending time with student representation had on me. I want to thank all my companions at the Fachschaft Elektrotechnik, at the university-level representation (HTU) and within the Fachschaftslisten Österreichs.

My partner Claudia again and again motivated me to continue writing this thesis and I sincerely want to thank her for her time and support.

Finally, thanks to those who provide moments of smiles every day, those who make the world a better place and those who do not judge.

Contents

List of Figures

Nomenclature

AIC	Akaike Information Criterion
APD	Avalanche Photodiode
AWGN	Additive White Gaussian Noise
BAC	Binary Asymmetric Channel
BER	Bit Error Rate
BSC	Binary Symmetric Channel
CCDF	Complementary Cumulative Distribution Function
DAQ	Data Acquisition
ESA	European Space Agency
FSPL	Free-Space Path Loss
FWHM	Full-Width at Half Maximum
GLRT	Generalized Likelihood Ratio Test
HMM	Hidden Markov Model
iid	Independent and Identically Distributed (random variables)
IIR	Infinite Impulse Response
IM/DD	Intensity Modulation / Direct Detection

ISI	Inter-Symbol Interference
ISL	Inter-Satellite Link
KL	Kullback-Leibler
LLR	Log-Likelihood Ratio
LN	Log-Normal
M-PAM	M-ary Pulse Amplitude Modulation, Multilevel Pulse Amplitude Modulation
m-Sequence	Maximum Length Sequence (type of PRBS)
MCMC	Markov Chain Monte Carlo
MCS	Modulation and Coding Scheme
ML	Maximum Likelihood
MLSD	Maximum Likelihood Sequence Detector
MPPM	Multipulse Pulse Position Modulation
MSa	Megasample
ND	Neutral-Density
NP	Neyman-Pearson
OGS	Optical Ground Station
OOK	On-Off Keying
OWC	Optical Wireless Communication
PAPR	Peak-to-Average Power Ratio
pdf	Probability Density Function
PIM	Pulse Interval Modulation
PPM	Pulse Position Modulation
PRBS	Pseudo-Random Bit Sequence

PSD	Power Spectral Density
PWM	Pulse Width Modulation
RF	Radio Frequency
ROC	Receiver Operating Characteristic
SCPI	Standard Commands for Programmable Instruments
SFM	Spectral Flatness Measure
SILEX	Semi Conductor Inter Satellite Link Experiment
SMC	Single-Step Markov Chain
SNR	Signal-to-Noise Ratio
SPAD	Single-Photon Avalanche Photodiode
VISA	Virtual Instrument Software Architecture
wrt.	With Regards to

1

Introduction

1.1 Overview and Motivation

Wireless satellite communication links are primarily realized by radio frequency (RF) schemes. These communication systems should be lightweight, small and power-efficient while the data rate requirements are as high as 5.6 Gbps [1]. Using higher frequencies, and, consequentially, shorter wavelengths, enables higher antenna gains at the same aperture. This basic property of electromagnetic waves led not only to higher frequencies being used, but also to a push for using optical instead of RF wireless communication systems.

The constraints put upon wireless communication systems are numerous: The bit rate is to be maximized, bandwidth is a scarce resource, the power consumption should be as low as possible while the complexity needs to be modest enough to be not only realized but space-ready and the system should be robust against errors. These requirements all conflict with each other, rendering tradeoffs necessary. Overall progress is typically driven by hardware improvements or the development of low-complexity algorithms with good performance.

In the case of FSO communication, the advent of reliable photodiodes (PIN diodes and avalanche photodiodes) in the 1970s as well as semiconductor lasers in the 1980s set the path for this technology's maturation [2]. Due to delays and cancellations mostly unrelated to the optical technology itself, the first inter-satellite laser communication was not demonstrated until November 2001 [3]. The *Semi conductor Inter satellite Link EXperiment*, SILEX, showed exceptional performance of the optical link between the French earth observation satellite SPOT4 and the European Space Agency's (ESA's) telecommunication satellite ARTEMIS [3, 4].
While inter-satellite links (ISLs) have been the main focus of FSO or optical wireless communication (OWC) schemes, up- and downlink FSO communication has gained significant popularity over the last decade. This thesis also focuses on up- and downlink, where the laser beam propagates through the atmosphere and is distorted by turbulences, temperature and pressure inhomogeneities. Modeling this time-varying channel is the crucial step when designing the communication system.
Tests performed on the link between ARTEMIS and ESA's Optical Ground Station (OGS) showed good downlink performance, but substantially worse uplink operation [5, 6]. This is caused by the asymmetric link geometry introduced by the atmosphere: While the downlink is undisturbed for nearly the full link distance, the uplink experiences atmospheric perturbation at the very beginning, leading to divergence and hence increased loss over the whole link distance. Furthermore, The OGS receiver counters beam wandering and scintillation, hence reducing atmospheric influence on the communication channel [5].
As opposed to RF modulation and coding schemes (MCS) where coherent detectors, which allow amplitude, phase and frequency modulation, are widely used, FSO communication mainly uses intensity modulation / direct detection (IM/DD). The most common approach is on-off keying (OOK), which is used due to its high bandwidth efficiency and simplicity. Pulse position modulation (PPM) is the second IM/DD approach widely discussed in the literature, but was not further investigated since it provides less spectral efficiency (and hence low data rates) as well as no limiting constraints on the average power are imposed by the setup.

This thesis focuses on the evaluation of channel models for optical communication systems. Channel models with different amounts of complexity are analyzed and compared. This includes real-valued stochastic channels and their resulting bit error rates as well as discrete models as the highest-layer view. The binary Gilbert and Elliott burst error channels [7, 8] with hidden Markov model (HMM) structure are the prime discrete channels under scrutiny.

1.2 Challenges and Thesis Goals

The prime goal of this thesis is to provide optimal detectors for various models of optical links' behavior under OOK. Where these are not feasible, approximations shall be explored and compared. Following directly, the achievable capacity of said models is to be quantified and examined. A further objective is the analysis and evaluation of channel models for FSO communication wrt. the optical link in urban environment studied herein. Core topics addressed by this work include:

- Evaluation and comparison of different channel models for FSO links
- Derivation of optimal detectors for the analyzed models when OOK is used
  - Approximation of the optimal detector if it is prohibitively complex or impractical
- Analysis of the theoretical channel capacity of the discussed models
- Characterization of the physical link, evaluation of its performance and its accordance with the assessed models.

The main challenge of this work is balancing abstraction and detail when modeling FSO communication channels. Further challenges include the complexity of optimal sequence detectors or estimators of the instantaneous fading state, as well as the general difficulty determining the channel capacity.

1.3 State of the Art

After the SILEX mission 2001, numerous other FSO communication systems were designed or deployed, including ISLs (Optical Inter-orbit Communications Engineering Test Satellite), deep space missions (e.g. Mars Laser Communications Demonstration, designed only) and down- (KIrari's Optical Downlink to Oberpfaffenhofen, KIODO) as well as bidirectional (e.g. Engineering Test Satellite VI) links, summarized in [9]. Coherent detection as well as diode-pumped solid state (Nd:YAG) lasers were promoted by the ESA in order to approach the theoretic quantum limit [4]. Both coherent and direct detection are currently in use and under research.
Apart from OOK and PPM, IM/DD FSO schemes with binary laser states include [10]: multipulse PPM (MPPM), pulse width modulation (PWM), pulse interval modulation (PIM) and polarization modulation schmemes. The prime example of intensity modulation with more than two laser states is M-ary pulse amplitude modulation (M-PAM) which uses multiple values of the laser intensity. The modulation schemes differ in their spectral efficiency, synchronization, energy efficiency, peak-to-average power ratio (PAPR) and implementation complexity [10]. The differences are summarized in Table 1.1, for schemes without guard slots, i.e. adjacent on-pulses are allowed for all schemes that do not include them inherently (OOK and M-PAM).
OOK directly modulates the bit stream onto the transmit power: a logic "1" is represented by a turned-on laser during the bit interval, a logic "0" by the absence of laser current. This leads to the spectral efficiency of 1 bit per channel usage in a straightforward manner – as well as to the PAPR of 2 (for uniform data) since the laser is turned on 50% of the time. The synchronization is necessary on symbol duration, which is at the same time the smallest occurring time interval.
PPM uses a symbol duration of M chip durations ("chips" are the smallest time intervals on the channel), during one of these slots the laser is active. Thus M symbols can be distinguished, leading to $log_2(M)$ bits per symbol, and a spectral efficiency of $\frac{1}{M} log_2(M)$, rendering it unsuitable for bandwidth-limited channels with no constraint on the average transmit power. The PAPR of M is a direct consequence of the laser being turned on during

Table 1.1: Comparison between different IM/DD schemes without guard slots (M refers to the number of symbols) [9, 10].

Modulation Scheme	**Average Power**	**Spectral Efficiency**	**PAPR**	**Synchronization**	**Complexity**
OOK	medium	1	2	symbol	lowest
PPM	very low	$\frac{1}{M}\,log_2(M)$	M	symbol, chip	low
MPPM	low	$\frac{log_2(M)}{n\sqrt[2n]{2\pi M^2}}$ for $M \gg n$	$\sqrt[2n]{2\pi M^2}$ for $M \gg n$	symbol, chip	high
PWM	medium	$\frac{1}{M}\,log_2(M)$	$\frac{M+1}{2M}$	chip	low
PIM	low	$\frac{2}{M+1}\,log_2(M)$	$\frac{M+1}{2}$	chip	medium
M-PAM	high	$log_2(M)$	2	symbol	medium

one of M slots. The synchronization between transmitter and receiver must happen on symbol and chip duration – there is no additional information encoded that allows the receiver to determine the begin or end of a symbol. PPM is widely used due to its low average power consumption [11].
MPPM is akin to PPM, but with $n > 1$ on-pulses during one symbol duration L, which does not equal M. Evidently, $M = \binom{L}{n}$, leading to a noninteger number of bits encoded per symbol. The Stirling formula $N! \approx \sqrt{2\pi N}\left(\frac{n}{e}\right)^N$ leads to the approximation $M \approx \frac{1}{\sqrt{2\pi n}}\left(\frac{L}{n}\right)^n$ for $L \gg n$ (and $M \gg n$). The PAPR L/n can be written in terms of M and n with this approximation. Furthermore, the spectral efficiency can be approximated by $\frac{1}{L}\,log_2(M) \approx \frac{log_2(M)}{n\sqrt[2n]{2\pi M^2}}$. In conclusion, MPPM outperforms PPM on peak-power-limited channels while the opposite is true for average-power-constrained channels [10].

PWM begins each symbol interval with a rising edge of the transmit power and modulates the number of slots until its falling edge to encode information. For M slots, $log_2(M)$ bits can be encoded per symbol. An advantage over (M)PPM is given by the simple symbol-frame synchronization by the periodic rising edge of the transmit power, at the cost of increased average power consumption.

PIM is obtained by stripping PPM of the "idle" slots at the end of each symbol – basically modulating the duration between two pulses of the transmit power. Obviously, this discards the constant symbol duration of the other schemes. However, it also allows for easy symbol-synchronization since each pulse marks the end of a symbol. Again, $log_2(M)$ bits are encoded per symbol, but the *average* symbol duration equals $\frac{M+1}{2}$. The consequence is increased spectral efficiency at the cost of variable symbol duration and increased average transmit power due lower PAPR.

M-ary PAM can be described as OOK with more than two states, leading to $log_2(M)$ bits per symbol interval, which is already the smallest occurring time-unit. If the power levels corresponding to the various symbols are evenly spaced from zero to the peak power, the PAPR equals 2 as in the OOK case (generally, OOK can be seen as 2-PAM). Downsides of M-PAM are increased complexity due to modulating the power level (as opposed to the presence of a pulse) and the high power needed to attain suitable symbol distances. Furthermore, in the absence of perfect knowledge of the channel fading coefficient, its fluctuation can introduce additional symbol errors via its effect on the received amplitude.

In the case of (M)PPM, PWM and PIM, guard slots are often added at the end of each symbol. These are intervals of e.g. one chip duration with disabled transmitter. Their purpose is to prevent adjacent/consecutive on-pulses across symbols and help with symbol-level synchronization. Also, they decrease average transmit power and spectral efficiency while increasing PAPR.

The nomenclature for the modulation schemes explained above varies in the literature. Often, a preceding letter D denotes a digital modulation, sometimes it refers to "differential". Care has to be taken in order not to confuse multipulse PPM with M-ary PPM, at times referred to as M-PPM.

The three main propagation effects of the FSO channel on the laser beam are free-space path loss (FSPL), atmospheric extinction/damping and scintillation. While the former two affect the mean received power, scintillation also increases its variance. For in-orbit links, FSPL plays the only role as long as the channel remains outside of the atmosphere [12]. In commercial FSO communication systems, high link margins [13] or spatial diversity measures [12] are typically used to counter the adverse propagation effects.
For very low received intensities, (coherent) photon counting detectors are typically used. These usually employ photon multipliers followed by an integrator or edge detector whose output is able to resolve single photons [14]. While the principle of avalanche photodiodes (APDs) is very similar (carrier multiplication by the avalanche effect) and they can be operated in Geiger mode (single-photon avalanche photodiode, SPAD), their typical usage is in the continuous (i.e. large number of photons) regime. Photon counters lead to discrete probability distributions (Poisson) while continuous distributions are the consequence of approximating the number of photons by e.g. a real-valued charge.
The intensity sensor used in this thesis is of type APD, hence the models discussed feature continuous distributions. The most prominent member of this class of models is the additive white Gaussian noise (AWGN) channel. This model states that the (possibly scaled) input to the communication is disturbed by adding noise that is normally distributed (Gaussian) – as a result, all conditional density functions are Gaussian with the same variance. The approximation of the Poisson process of photons being absorbed by the sensor by a Gaussian distribution and subsequent additive Gaussian noise leads to data-dependent variance and a channel that has been termed optical intensity channel [15]. By conflating the stochastic attenuation of the channel with the Poisson approximation, the Gaussian fading coefficient hypothesis of this work is developed. Here, the conditional probability density functions are both Gaussian, but with different variance.
The most abundantly used model for FSO communications is the log-normal fading, where the laser intensity is multiplied by a positive random variable whose logarithm is normally distributed. The physical justification stems from many independent multiplicative effects on the beam – by the central

limit theorem, their influence in the log-domain converges to the normal distribution. While the log-normal model is recommended for weak scintillation [10], it is also appropriate when the receiver aperture is relatively large [16].
Another notable model is the Gamma-Gamma or ΓΓ channel, which also occupies a prominent position in the FSO literature. It is used to model weak, medium as well as strong scintillation and is in part based on the physics of large and small eddies or turbulence cells [10]. Numerous other models have been proposed, including the negative exponential, Rayleigh and K distribution, summarized e.g. in [9, 10].
The temporal correlation of the fading coefficient is also an active topic of research. While most publications discuss fading mitigation by use of multiple uncorrelated receivers (or the respective effect if they are not perfectly uncorrelated), there is literature on simulating temporally correlated optical channels, e.g. [17] and sub-optimal detectors [18], since optimal detection is computationally infeasible.
Since temporally correlated fading coefficients give rise to burst errors (in the presence of noise), abstractions that deal with the effects on the binary level have to be considered. The first and most known examples are the Gilbert [7] and Elliott [8] burst error channels. Both are based on binary Markov chains with a so-called "good" and "bad" state respectively. The states can not be observed directly, but their effects on the transmission: In the bad state, errors happen frequently, while they are absent (Gilbert) or happen less often (Elliott) in the good channel state. The transition probabilities between the states specify how long error-free (or low-error) periods typically persist. A timestep in the model's reference does not necessarily have to occur for each bit interval; packet-level considerations are e.g. found in [19].

1.4 Research Questions

In combination with the results of the state of the art analysis, following research questions emerge:

Research Question 1.
Is the studied 600 m link sufficiently well-characterized by the Gaussian and log-normal models for the channel coefficient, and which model yields a better fit to the measured values?

Research Question 2.
Are optimal detectors for OOK over channels with iid or correlated Gaussian or log-normal fading coefficient practically realizable?

Research Question 3.
Can the capacity of the studied link and the discussed models be determined?

1.5 Structure of this Thesis

Chapter 2 is concerned with the atmospheric link, including physical justifications for the different stochastic channel models used. The chapter concludes with considerations on temporally correlated channels and methods to simulate their behavior as well as detectors exploiting this correlations.
Chapter 3 summarizes the measurements that were performed on the optical link. The conformity between the theoretic models and the physical link is determined and key parameters of the channel are quantified.
Chapter 4 addresses the information-theoretic side of this work, discussing the channel capacity of the models derived in Chapter 2 and of the FSO link studied in this thesis.
The final Chapter 5 summarizes the conclusions of the thesis at hand and gives a short outlook on future research topics with thematic ties to this work.

2

Atmospheric Link

In this chapter, the physical properties of the communication channel and the stochastic real-valued channel models are discussed. A block diagram of the system is shown in Figure 2.1. At time instant n, a binary symbol u_n is multiplied by the positive amplitude h_n, yielding the noiseless received (but unobservable) symbol x_n. The additive receiver noise is added, producing the noisy received signal y_n:

$$\mathsf{y}_n = \mathsf{u}_n \mathsf{h}_n + \mathsf{v}_n \tag{2.1}$$

$$\mathbf{y} = \mathbf{u}\, diag\,(\mathbf{h}) + \mathbf{v} \tag{2.2}$$

The absence of transmitter intensity will be referred to as "L", 0 or low in this thesis, while the presence will be labeled "H", 1 or high. Equation 2.2 displays the vector-valued representation of the individual symbols, where $diag\,(\cdot)$ is a square matrix where the main diagonal is populated with the argument vector and all other components are zero. All vectors are represented as column vectors.

In the absence of additive noise, i.e. $\mathsf{y}_n = \mathsf{x}_n$, the sent symbol can always be recovered under the assumption of a positive, nonzero amplitude h_n. This assumption is valid as long as no total blockage of the communication channel is occurring.

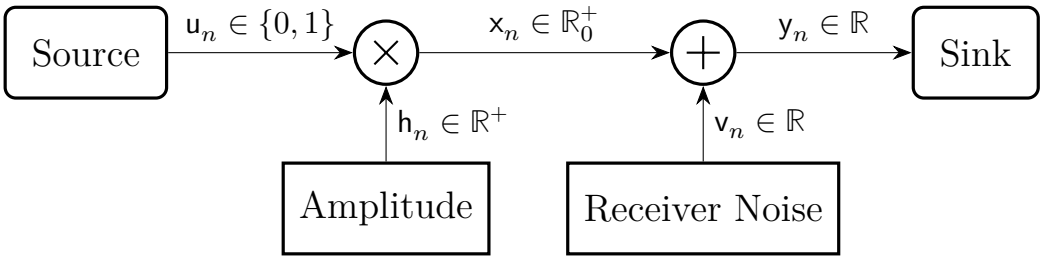


Figure 2.1: Channel block diagram at time index n. The main differences between the models arise from differently characterized h_n and v_n.

The restriction to positive channel coefficients is a direct consequence of the IM/DD scheme. Its received signal is directly proportional to the power at the aperture, which is nonnegative. Different models for the amplitude and the receiver noise as well as their respective temporal evolution lead to different channel models with the same basic structure (cf. Figure 2.1).
The receiver noise v_n consists of thermal and shot noise. The latter is caused by the discrete nature of photons as well as charge carriers and follows a Poisson distribution; however, given a sufficient mean number of photons, it can be approximated by a Gaussian distribution [10]. The thermal noise is caused by the electronic receiver circuit and is typically assumed to be independent of the received signal presence or strength. The shot noise depends on the background light intensity as well as on the signal strength, leading to different noise variances depending on the data. By using a stochastic model for the channel coefficient h_n, the data-dependent part of the shot noise can be shifted away from the receiver noise v_n, which can then be viewed as independent of the data. Both data-independent noise processes are uncorrelated and white – together they can be represented by a Gaussian distribution:

$$f_{\mathsf{v}_n}(v) = \frac{1}{\sqrt{2\pi}\sigma_{\mathsf{v}}} exp\left(-\frac{(v-\mu_{\mathsf{v}})^2}{2\sigma_{\mathsf{v}}^2}\right) \qquad \forall n \in \mathbb{N} \tag{2.3}$$

$$E\{\mathsf{v}_n\} = \mu_{\mathsf{v}}$$

$$E\left\{(\mathsf{v}_n - \mu_{\mathsf{v}})^2\right\} = \sigma_{\mathsf{v}}^2$$

The parameters μ_{v} and σ_{v} vary slowly due to ambient light and receiver temperature, but stay essentially constant over at least 10^7 symbols. Thus these parameters are treated as constant, but unknown. Together with a deterministic decision function $\hat{u}(\cdot)$ (to be specified in the following sections) we can now formally define the resulting channel:

Definition 2.1. Channel with Positive Fading Coefficient and Gaussian Noise

$$\left.\begin{aligned} \mathsf{y}_n &= \mathsf{u}_n \mathsf{h}_n + \mathsf{v}_n \\ \mathsf{u}_n &\in \{0,1\} \\ \mathsf{h}_n &\in \mathbb{R}^+ \\ \mathsf{h}_n &\sim f_{\mathsf{h}_n}(\cdot) \\ \mathsf{v}_n &\in \mathbb{R} \\ \mathsf{v}_n &\sim f_{\mathsf{v}}(\cdot) \end{aligned}\right\} \quad \forall n \in \mathbb{N}$$

$$f_{\mathsf{v}}(v) = \frac{1}{\sqrt{2\pi}\sigma_{\mathsf{v}}} exp\left(-\frac{(v-\mu_{\mathsf{v}})^2}{2\sigma_{\mathsf{v}}^2}\right)$$

$$\hat{\mathsf{u}}_n = \hat{u}(\mathsf{y}_n)$$

$$\sigma_{\mathsf{v}} > 0$$

The optimal detector $\hat{u}(\cdot)$ in terms of BER – $P\{\hat{\mathsf{u}}_n \neq \mathsf{u}_n\}$ (without knowledge of the prior probabilities of or for uniform distribution of the data) is the maximum likelihood (ML) detector. It is defined for individual symbols in Definition 2.2 and sequences of several received symbols in Definition 2.3.

Definition 2.2. Maximum Likelihood Symbol Detector

$$\hat{u}_n(y_n) = arg\,\underset{u}{max}\left\{f_{\mathsf{y}_n|\,\mathsf{u}_n}(y_n \mid u)\right\} \tag{2.4}$$

Definition 2.3. Maximum Likelihood Sequence Detector (MLSD)

$$\hat{\boldsymbol{u}}(\boldsymbol{y}) = arg\,\underset{\boldsymbol{u}}{max}\left\{f_{\mathbf{y}|\,\mathbf{u}}(\boldsymbol{y} \mid \boldsymbol{u})\right\} \tag{2.5}$$

2.1 Additive White Gaussian Channel Hypothesis

Within the coherence time of fading (which is usually in the range of 0.1 to 10 ms [9, 12], resulting in hundreds to several thousands of bits within the coherence time), the channel fading coefficient h stays approximately constant (reflected by the use of italics as opposed to the random variable h_n). This leads to the commonly known additive white Gaussian noise (AWGN) channel, which has been studied in detail. Moreover, closed-form expressions of key relations are known. The considerations in this section highlight the best receiver performance attainable when complete knowledge of the channel fading coefficient is available.
Two closely linked expressions for the comparison between the likelihood functions are the likelihood ratio $L(y)$ and log-likelihood ratio (LLR) $l(y)$. Their definition wrt. the AWGN detection problem of this section is found in Definition 2.4 [20]. Note that the ML comparison can be given by $L(y) > 1$ or, equivalently by $l(y) > 0$.

Definition 2.4. Likelihood Ratio and Log-Likelihood Ratio for the AWGN Channel

$$L(y) = \frac{f_{\mathsf{y}_n|\mathsf{u}_n}(y \mid 1)}{f_{\mathsf{y}_n|\mathsf{u}_n}(y \mid 0)} = \frac{f_{\mathsf{v}}(y-h)}{f_{\mathsf{v}}(y)} \tag{2.6}$$

$$l(y) = log\Big(L(y)\Big) \propto -\frac{(y-h-\mu_{\mathsf{v}})^2}{2\sigma_{\mathsf{v}}^2} + \frac{(y-\mu_{\mathsf{v}})^2}{2\sigma_{\mathsf{v}}^2} = \frac{h(y-h/2-\mu_{\mathsf{v}})}{\sigma_{\mathsf{v}}^2} \tag{2.7}$$

The ML symbol detector in the AWGN case reduces to a threshold comparison, deciding "L" for all received symbols below γ. The threshold for the system of Definition 2.1 is given by Equation 2.8 where μ_{v} is the noise DC component. By a simple coordinate transformation, this can always be set to zero.
The ML decision rule is given by Definition 2.5 with the indicator function of Definition 2.6. For continuous random variables, the difference between using $>$ and $\geq$ does not affect performance since $P\{\mathsf{y}_n = \gamma\} = 0$.

Definition 2.5. ML Symbol Detector for AWGN Channel with known fading coefficient h

$$\gamma = \mu_{\mathsf{v}} + \frac{h}{2} \tag{2.8}$$

$$\hat{u}\,(y) = I\,(y > \gamma) \tag{2.9}$$

Definition 2.6. Indicator function $I\,(\cdot)$

$$I\,(\text{true}) = 1$$
$$I\,(\text{false}) = 0$$

The performance of the ML decision rule can be explicitly derived: A Type I error or false alarm (erroneously assuming presence of laser intensity) happens in case "L" has been sent and the noise exceeds the threshold γ. A Type II error or miss (assigning "L" to a symbol where "H" has been sent) occurs when the noisy received signal y_n is below the decision threshold γ. Correctly detecting the present laser intensity is called detection while properly deciding in favor of "L" results in a correct rejection.
Due to Equation 2.8, the probabilities for false alarms (P_{FA}) and misses (P_{M}) as well as for detections (P_{D}) and correct rejections (P_{CR}) are very similar (cf. Equation 2.10), and, together with the symmetry of the Gaussian probability density function (pdf) and Definition 2.7 lead to the expression in Equation 2.11 for the respective probabilities.

$$P_{\text{D}} = P\{\hat{\mathsf{u}}_n = 1 \mid \mathsf{u}_n = 1\} = P\{\mathsf{v}_n > -\gamma\} \tag{2.10a}$$
$$P_{\text{FA}} = P\{\hat{\mathsf{u}}_n = 1 \mid \mathsf{u}_n = 0\} = P\{\mathsf{v}_n > \gamma\} \tag{2.10b}$$
$$P_{\text{M}} = P\{\hat{\mathsf{u}}_n = 0 \mid \mathsf{u}_n = 1\} = P\{\mathsf{v}_n < -\gamma\} \tag{2.10c}$$
$$P_{\text{CR}} = P\{\hat{\mathsf{u}}_n = 0 \mid \mathsf{u}_n = 0\} = P\{\mathsf{v}_n < \gamma\} \tag{2.10d}$$

Definition 2.7. The Q-function $Q(\cdot)$ is the complementary cumulative distribution function (CCDF) of the standard normal distribution and can be expressed by ($erfc$ is the complementary error function):

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty exp\left(-\frac{y^2}{2}\right) dy$$
$$Q(x) = \frac{1}{2} erfc\left(\frac{x}{\sqrt{2}}\right)$$

It follows from Equation 2.11b that an AWGN channel with ML decision leads to a binary symmetric channel (BSC). The term h/σ_{v} expresses the signal-to-noise ratio (SNR) in field quantities (as opposed to power). The error probability vs. the SNR is shown in Figure 2.2.

$$P_{\mathrm{D}} = P_{\mathrm{CR}} = 1 - Q\left(\frac{h}{2\sigma_{\mathsf{v}}}\right) \tag{2.11a}$$

$$P_{\mathrm{FA}} = P_{\mathrm{M}} = Q\left(\frac{h}{2\sigma_{\mathsf{v}}}\right) \tag{2.11b}$$

However, this performance is only attainable in the case of complete knowledge of the channel coefficient. In the case of time-variant coefficients, the ML decision threshold and the error probability also vary with time. Time-variant receiver noise does not affect the optimal threshold, but alters the performance by changing the SNR.
Since the threshold γ lies exactly between the expected values for "L" and "H", it can be approximated from the received sequence even in the absence of information about the channel as shown in [21]. The core assumptions made are that the fading coefficient h stays constant for the duration of the transmission and that the data is uniformly distributed.
Furthermore, time-variant fading coefficients make different detection strategies necessary. Those include approximation of the channel coefficient, e.g. the generalized likelihood ratio test (GLRT) [20], and Neyman-Pearson (NP) detectors, which depend solely on the statistics of the receiver noise v_n.

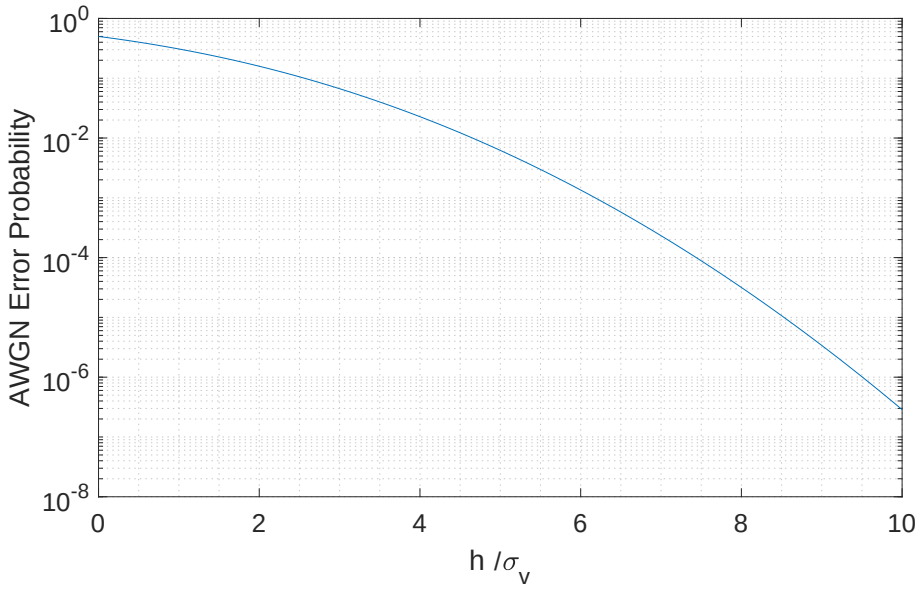


Figure 2.2: Error probability for the AWGN channel with ML decision vs. SNR in field quantities.

The following two sections will consider the GLRT strategy of Definition 2.8 in-depth. The NP approach was not further pursued in this thesis due to the following considerations: The pdf of the receiver noise is not known a priori, it depends on ambient light conditions and the receiver temperature. Even though those parameters do not vary fast enough to necessitate on-line tracking, they render estimation of the mean and variance of v_n necessary – in that case, estimating parameters of the signal with the laser turned on does not lead to a notable increase of complexity.

Definition 2.8. A Generalized Likelihood Ratio Test (GLRT) for a detection problem with unknown parameters $\boldsymbol{\theta}$ is a likelihood ratio test for the same problem where the ML estimates of the parameters $\hat{\boldsymbol{\theta}}$ are used instead of the unknown values.

$$\begin{aligned} H_0 &: \boldsymbol{\theta} \in \Theta_0 \\ H_1 &: \boldsymbol{\theta} \in \Theta_1 \\ L_g(\boldsymbol{y}) &= \frac{f_{\mathbf{y};\boldsymbol{\theta}}\left(\boldsymbol{y};\hat{\boldsymbol{\theta}}_1\right)}{f_{\mathbf{y};\boldsymbol{\theta}}\left(\boldsymbol{y};\hat{\boldsymbol{\theta}}_0\right)} \\ l_g(\boldsymbol{y}) &= log\left(L_g(\boldsymbol{y})\right) \\ \hat{\boldsymbol{\theta}}_0 &= arg \max_{\boldsymbol{\theta}\in\Theta_0} \left\{f_{\mathbf{y};\boldsymbol{\theta}}(\boldsymbol{y};\boldsymbol{\theta})\right\} \\ \hat{\boldsymbol{\theta}}_1 &= arg \max_{\boldsymbol{\theta}\in\Theta_1} \left\{f_{\mathbf{y};\boldsymbol{\theta}}(\boldsymbol{y};\boldsymbol{\theta})\right\} \end{aligned}$$

Definition 2.9. The Neyman-Pearson detector with level α is the detector which maximizes the detection probability P_{D} with a false alarm rate P_{FA} below or equal to α. For Definition 2.1 the NP detector is obtained by

$$\begin{aligned} \hat{\mathsf{u}}_n &= I(\mathsf{y}_n > \gamma) \\ \gamma &= \mu_{\mathsf{v}} + \sigma_{\mathsf{v}} Q^{-1}(\alpha), \end{aligned}$$

where $Q^{-1}(\cdot)$ is the inverse of the Q-function of Definition 2.7.

Since Definition 2.9 is asymmetric, the resulting binary channel will not be symmetric: $P_{\mathrm{FA}} = \alpha$ while $P_{\mathrm{M}} = Q\left(h/\sigma_{\mathsf{v}} - Q^{-1}(\alpha)\right)$. For linear channel codes however, only the total error rate matters – thus rendering the NP approach sub-optimal in general and choosing an appropriate level α difficult. Figure 2.3 shows the receiver operating characteristic (ROC) for the AWGN channel and the operating points of the ML detector. The dash-dotted line corresponds to the pure guess.

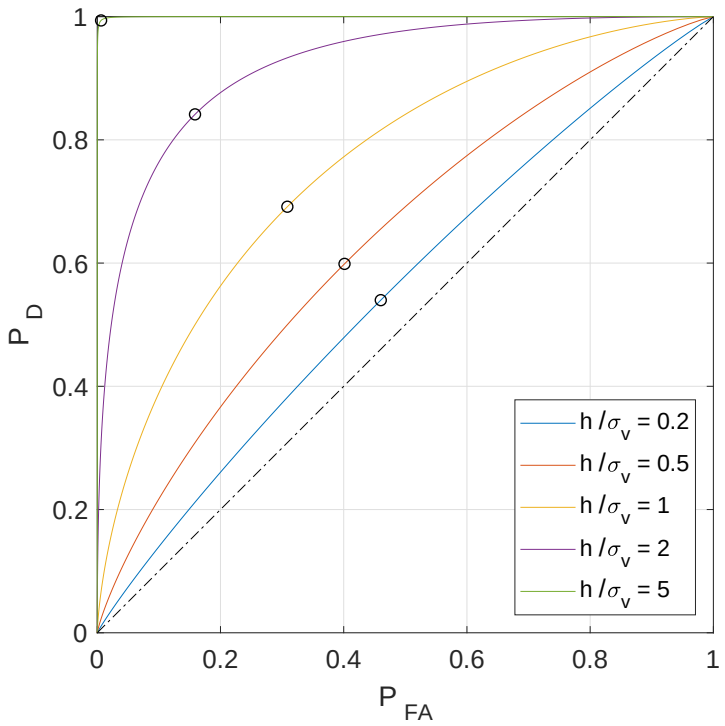


Figure 2.3: Receiver operating characteristic for AWGN channel, black circles indicate ML detector operating points. The dash-dotted line represents pure guess.

2.2 Fluctuating Fading Coefficient

In this section we regard the channel coefficients h_n independent and identically distributed (iid) according to the probability density function $f_{\mathsf{h}_n}(x) = f_{\mathsf{h}}(x)$ (note that this implies that the stochastic process to be strongly stationary). From $\mathsf{y}_n = \mathsf{x}_n + \mathsf{v}_n$, the pdf of y_n can be derived via convolution as shown in Equation 2.12.

$$f_{\mathsf{y}_n}(y) = \left(f_{\mathsf{x}_n} * f_{\mathsf{v}}\right)(y) \tag{2.12}$$

$$f_{\mathsf{x}_n|\mathsf{u}_n}(x \mid 0) = \delta(x) \tag{2.13a}$$

$$f_{\mathsf{x}_n|\mathsf{u}_n}(x \mid 1) = f_{\mathsf{h}}(x) \tag{2.13b}$$

$$f_{\mathsf{y}_n|\mathsf{u}_n}(y \mid 0) = f_{\mathsf{v}}(y) \tag{2.14a}$$

$$f_{\mathsf{y}_n|\mathsf{u}_n}(y \mid 1) = (f_{\mathsf{h}} * f_{\mathsf{v}})(y) \tag{2.14b}$$

Definition 2.10. Unit Impulse Function, Dirac-Delta Distribution $\delta(\cdot)$

$$\delta(x) = 0 \qquad x \neq 0$$

$$\int_a^b f(x)\,\delta(x)\,dx = \begin{cases} f(0) & 0 \in [a,b]\,, \\ 0 & 0 \notin [a,b]\,. \end{cases}$$

With a slight abuse of notation, the unit impulse function $\delta(x)$ with intuitive Definition 2.10 is used to represent the degenerate density function of the deterministic variable x_n given $\mathsf{u}_n = 0$. Since the unit impulse function is the identity element of the convolution, Equation 2.14a reduces to the pdf of the receiver noise. Together with Equation 2.14b, representations of the ML symbol detector can be derived:

Definition 2.11. ML Symbol Detector for iid Channel Coefficients h_n

$$\hat{\mathsf{u}}_n = I\left(f_{\mathsf{y}_n|\mathsf{u}_n}(\mathsf{y}_n \mid 1) > f_{\mathsf{y}_n|\mathsf{u}_n}(\mathsf{y}_n \mid 0)\right) \tag{2.15}$$

$$\gamma = min\left\{y : f_{\mathsf{y}_n|\mathsf{u}_n}(y \mid 1) > f_{\mathsf{y}_n|\mathsf{u}_n}(y \mid 0)\right\} \tag{2.16}$$

$$\hat{\mathsf{u}}_n = I(\mathsf{y}_n > \gamma) \tag{2.17}$$

From Equations 2.14a and 2.14b we can derive the (log-)likelihood ratio in a straightforward manner. Note that the thresholds obtained by using the likelihood ratio $L(y)$ and the LLR $l(y)$, expressed via $\gamma = min\{y : L(y) > 1\} = min\{y : l(y) > 0\}$ lead to the same threshold γ as in Equation 2.16.

Definition 2.12. Likelihood Ratio and Log-Likelihood Ratio for iid Fading Coefficients h_n

$$L(y) = \frac{(f_{\mathsf{h}} * f_{\mathsf{v}})(y)}{f_{\mathsf{v}}(y)} \tag{2.18}$$

$$l(y) = log\Big((f_{\mathsf{h}} * f_{\mathsf{v}})(y)\Big) - log\Big(f_{\mathsf{v}}(y)\Big) \tag{2.19}$$

Two different two-parameter distributions are used in this thesis for modeling the channel coefficient h: The normal or Gaussian distribution and the log-normal (LN) distribution. The former originates from the real-valued approximation of the Poisson processes and allows closed-form expressions of key quantities, but requires the restriction to positive channel coefficients to be dropped as its support includes negative values. The log-normal distribution naturally arises from multiple uncorrelated, multiplicative factors that influence a quantity and is widely used in optical communications.
In the FSO literature, the scintillation index σ_{SI}^2 (cf. Definition 2.13) and the Rytov variance σ_{R} (Definition 2.14) are commonly used to characterize the severeness of fading [9]. The three tiers commonly used are weak ($\sigma_{\mathrm{SI}}^2 < 0.3$ or $\sigma_{\mathrm{R}} \ll 1$), moderate ($0.3 < \sigma_{\mathrm{SI}}^2 < 5$ or $\sigma_{\mathrm{R}} \approx 1$) and strong turbulence or saturation regime ($\sigma_{\mathrm{SI}}^2 > 5$ or $\sigma_{\mathrm{R}} \gg 1$). The log-normal channel model is the preferred model for the weak turbulence domain, but numerous other stochastic models have been proposed for moderate and strong turbulence, e.g. the Gamma-Gamma distribution [9, 16, 22, 23]. However, for large apertures, log-normal and Gamma-Gamma channel models become practically equivalent [10, 24] and measurements confirmed that the log-normal model describes moderate to strong turbulence reasonably well [16, 17].

Definition 2.13. Scintillation Index σ_{SI}
(μ_{h} and σ^2_{h} refer to the parameters of the respective distributions)

$$\sigma^2_{\mathrm{SI}} = \frac{Var\{\mathsf{h}\}}{\left(E\{\mathsf{h}\}\right)^2} \qquad \text{(general)} \qquad (2.20)$$

$$\sigma^2_{\mathrm{SI}} = \frac{\sigma^2_{\mathsf{h}}}{\mu^2_{\mathsf{h}}} \qquad \text{for Gaussian } \mathsf{h}$$

$$\sigma^2_{\mathrm{SI}} = \frac{\left(exp\left(\sigma^2_{\mathsf{h}}\right) - 1\right) exp\left(2\mu_{\mathsf{h}} + \sigma^2_{\mathsf{h}}\right)}{exp\left(2\mu_{\mathsf{h}} + \sigma^2_{\mathsf{h}}\right)} = exp\left(\sigma^2_{\mathsf{h}}\right) - 1 \qquad \text{for log-normal } \mathsf{h}$$

Definition 2.14. Rytov Variance σ_{R}

$$\sigma^2_{\mathrm{R}} = 1.23\, C_n^2\, k^{7/6}\, r^{11/6}, \qquad (2.21)$$

where C_n^2 is the refractive-index structure constant, k is the wave number $\frac{2\pi}{\lambda}$ and r is the link length.

The noise DC value μ_{v} and variance σ^2_{v} can always be normalized, i.e. set to 0 and 1 respectively. The parameters of the transformed fading coefficient h' under the Gaussian and log-normal models change as follows (the scintillation index is not affected in both cases):

Gaussian: $\mu_{\mathsf{h}'} = \frac{\mu_{\mathsf{h}}}{\sigma_{\mathsf{v}}}$ $\sigma^2_{\mathsf{h}'} = \frac{\sigma^2_{\mathsf{h}}}{\sigma^2_{\mathsf{v}}}$ (2.22)

Log-Normal: $\mu_{\mathsf{h}'} = \mu_{\mathsf{h}} - ln\left(\sigma_{\mathsf{v}}\right)$ $\sigma^2_{\mathsf{h}'} = \sigma^2_{\mathsf{h}}$ (2.23)

2.2.1 Gaussian Channel Coefficient

The implications of using the Gaussian distribution will follow in-depth. Since its support includes the negative numbers, it does not comply to the channel Definition 2.1 and the ML detection does not reduce to comparing the data to a threshold γ. The resulting channel definition is given in Definition 2.15 below. The likelihood functions in Equation 2.26 are obtained either by convolution or simply using Gaussian random variables' property of being closed under addition together with the first two cumulants μ and σ^2.

Definition 2.15. Channel with Gaussian Fading Coefficient and Gaussian Noise, without Restriction on Positive Fading Coefficients

$$\left.\begin{aligned} \mathsf{y}_n &= \mathsf{u}_n \mathsf{h}_n + \mathsf{v}_n \\ \mathsf{u}_n &\in \{0,1\} \\ \mathsf{h}_n &\in \mathbb{R} \\ \mathsf{h}_n &\sim f_{\mathsf{h}}(\cdot) \\ \mathsf{v}_n &\in \mathbb{R} \\ \mathsf{v}_n &\sim f_{\mathsf{v}}(\cdot) \end{aligned}\right\} \quad \forall n \in \mathbb{N}$$

$$f_{\mathsf{v}}(v) = \frac{1}{\sqrt{2\pi}\sigma_{\mathsf{v}}} exp\left(-\frac{(v-\mu_{\mathsf{v}})^2}{2\sigma_{\mathsf{v}}^2}\right)$$

$$f_{\mathsf{h}}(h) = \frac{1}{\sqrt{2\pi}\sigma_{\mathsf{h}}} exp\left(-\frac{(h-\mu_{\mathsf{h}})^2}{2\sigma_{\mathsf{h}}^2}\right)$$

$$\sigma_{\mathsf{v}} > 0, \quad \mu_{\mathsf{h}} > 0, \quad \sigma_{\mathsf{h}} > 0$$

Definition 2.16. Relationships between Stochastic Channel Properties and Observed Quantities for the Gaussian Channel Coefficient

$$\mu_L = E\{\mathsf{y}_n \mid \mathsf{u}_n = 0\} = \mu_{\mathsf{v}} \tag{2.24a}$$

$$\sigma_L^2 = E\left\{(\mathsf{y}_n - \mu_L)^2 \mid \mathsf{u}_n = 0\right\} = \sigma_{\mathsf{v}}^2 \tag{2.24b}$$

$$\mu_H = E\{\mathsf{y}_n \mid \mathsf{u}_n = 1\} = \mu_{\mathsf{v}} + \mu_{\mathsf{h}} \tag{2.24c}$$

$$\sigma_H^2 = E\left\{(\mathsf{y}_n - \mu_H)^2 \mid \mathsf{u}_n = 1\right\} = \sigma_{\mathsf{v}}^2 + \sigma_{\mathsf{h}}^2 \tag{2.25a}$$

$$f_{\mathsf{y}_n|\mathsf{u}_n}(y \mid 0) = \frac{1}{\sqrt{2\pi}\sigma_L} exp\left(-\frac{(y-\mu_L)^2}{2\sigma_L^2}\right) \tag{2.25b}$$

$$f_{\mathsf{y}_n|\mathsf{u}_n}(y \mid 1) = \frac{1}{\sqrt{2\pi}\sigma_H} exp\left(-\frac{(y-\mu_H)^2}{2\sigma_H^2}\right) \tag{2.25c}$$

From Equation 2.24, the estimation of the parameters of h and v is an easy task, given a long enough training sequence since the sample mean and the adjusted sample variance are unbiased and consistent estimators of the respective true values. Physical origins and justifications for the parameters μ_L, σ_L, μ_H and σ_H can be found in [11].

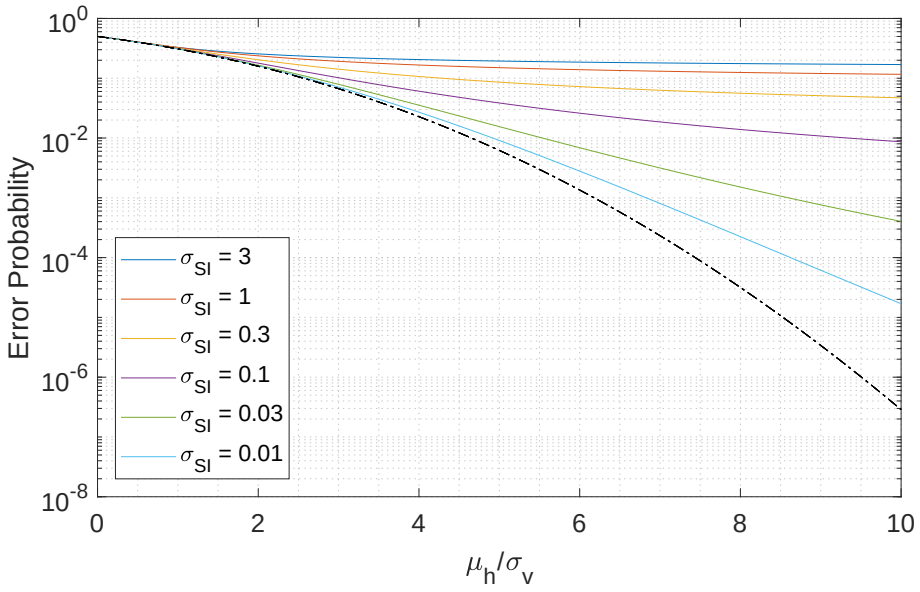


Figure 2.4: Error probability for Gaussian channel coefficient with threshold decision vs. SNR in field quantities (cf. Equation 2.31). The dash-dotted line corresponds to the AWGN channel.

Since the pdfs in question are continuous and unimodal, the values $y_i \in \mathbb{R} : l(y_i) = 0$ correspond to the boundaries of the ML decision scheme. Due to $\mathsf{y}_n \,|\, \mathsf{u}_n = 1$ having greater variance than $\mathsf{y}_n \,|\, \mathsf{u}_n = 0$ (reflected by positive, nonzero σ_h), a region where $y \ll \mu_\mathsf{v}$ leads to $\hat{\mathsf{u}}_{ML} = 1$, which is physically implausible and was hence omitted in this thesis. The decision rule was trimmed down to a simple threshold test with γ given by Equation 2.30a. In terms of the conditional means and variances of the received signal, the threshold is given by Equation 2.30b. The error probabilities are given by Equation 2.31ff and the average error rate for uniform data is shown for different noise constellations in Figure 2.4.

$$f_{\mathsf{y}_n|\,\mathsf{u}_n}(y \,|\, 0) = \frac{1}{\sqrt{2\pi}\sigma_\mathsf{v}} \, exp\left(-\frac{(y-\mu_\mathsf{v})^2}{2\sigma_\mathsf{v}^2}\right) \tag{2.26a}$$

$$f_{\mathsf{y}_n|\,\mathsf{u}_n}(y \,|\, 1) = \frac{1}{\sqrt{2\pi\left(\sigma_\mathsf{v}^2+\sigma_\mathsf{h}^2\right)}} \, exp\left(-\frac{(y-\mu_\mathsf{v}-\mu_\mathsf{h})^2}{2\left(\sigma_\mathsf{v}^2+\sigma_\mathsf{h}^2\right)}\right) \tag{2.26b}$$

$$l(y) \propto \frac{1}{2} ln\left(\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2}\right) - \frac{(y-\mu_{\mathsf{v}}-\mu_{\mathsf{h}})^2}{2(\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2)} + \frac{(y-\mu_{\mathsf{v}})^2}{2\sigma_{\mathsf{v}}^2} \tag{2.27}$$

The boundaries of the decision regions are found by solving the quadratic equation $l(y) = 0$:

$$l(y_i) = 0, \qquad i \in \{1,2\} \tag{2.28a}$$

$$\begin{aligned} 0 = {} & \sigma_{\mathsf{v}}^2(\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2) ln\left(\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2}\right) \\ & - \sigma_{\mathsf{v}}^2(y-\mu_{\mathsf{v}}-\mu_{\mathsf{h}})^2 + (\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2)(y-\mu_{\mathsf{v}})^2 \end{aligned} \tag{2.28b}$$

$$\begin{aligned} 0 = {} & -\sigma_{\mathsf{v}}^2\left(y^2 - 2y(\mu_{\mathsf{v}}+\mu_{\mathsf{h}}) + (\mu_{\mathsf{v}}+\mu_{\mathsf{h}})^2\right) \\ & + (\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2)(y^2 - 2y\mu_{\mathsf{v}} + \mu_{\mathsf{v}}^2) \\ & + \sigma_{\mathsf{v}}^2(\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2) ln\left(\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2}\right) \end{aligned} \tag{2.28c}$$

$$\begin{aligned} 0 = {} & y^2\sigma_{\mathsf{h}}^2 - 2y\left((\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2)\mu_{\mathsf{v}} - \sigma_{\mathsf{v}}^2(\mu_{\mathsf{v}}+\mu_{\mathsf{h}})\right) \\ & - \sigma_{\mathsf{v}}^2(\mu_{\mathsf{v}}+\mu_{\mathsf{h}})^2 + (\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2)\mu_{\mathsf{v}}^2 \\ & + \sigma_{\mathsf{v}}^2(\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2) ln\left(\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2}\right) \end{aligned} \tag{2.28d}$$

$$\begin{aligned} 0 = {} & y^2\sigma_{\mathsf{h}}^2 - 2y(\sigma_{\mathsf{h}}^2\mu_{\mathsf{v}} - \sigma_{\mathsf{v}}^2\mu_{\mathsf{h}}) \\ & + \sigma_{\mathsf{h}}^2\mu_{\mathsf{v}}^2 - 2\sigma_{\mathsf{v}}^2\mu_{\mathsf{v}}\mu_{\mathsf{h}} - \sigma_{\mathsf{v}}^2\mu_{\mathsf{h}}^2 \\ & + \sigma_{\mathsf{v}}^2(\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2) ln\left(\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2}\right) \end{aligned} \tag{2.28e}$$

$$\begin{aligned} 0 = {} & y^2 - 2y\left(\mu_{\mathsf{v}} - \frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{h}}^2}\mu_{\mathsf{h}}\right) \\ & + \mu_{\mathsf{v}}^2 - 2\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{h}}^2}\mu_{\mathsf{v}}\mu_{\mathsf{h}} - \frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{h}}^2}\mu_{\mathsf{h}}^2 \\ & + \sigma_{\mathsf{v}}^2\left(1+\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{h}}^2}\right) ln\left(\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2}\right) \end{aligned} \tag{2.28f}$$

$$y_i = \left(\mu_{\mathsf{v}} - \frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{h}}^2}\mu_{\mathsf{h}}\right) \pm \sqrt{\mu_{\mathsf{h}}^2\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{h}}^2}\left(1+\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{h}}^2}\right) - \sigma_{\mathsf{v}}^2\left(1+\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{h}}^2}\right) ln\left(\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2}\right)} \tag{2.29a}$$

$$y_i = \mu_{\mathsf{v}} - \frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{h}}^2}\mu_{\mathsf{h}} \pm \frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{h}}^2}\sqrt{\left(1+\frac{\sigma_{\mathsf{h}}^2}{\sigma_{\mathsf{v}}^2}\right)\left(\mu_{\mathsf{h}}^2 + \sigma_{\mathsf{h}}^2\, ln\left(1+\frac{\sigma_{\mathsf{h}}^2}{\sigma_{\mathsf{v}}^2}\right)\right)} \tag{2.29b}$$

The resulting threshold can be expressed in terms of the parameters of h and v or by the conditional means and variances:

$$\gamma = \mu_{\mathsf{v}} + \mu_{\mathsf{h}}\frac{\sigma_{\mathsf{v}}^2}{\sigma_{\mathsf{h}}^2}\left(\sqrt{\left(1+\frac{\sigma_{\mathsf{h}}^2}{\sigma_{\mathsf{v}}^2}\right)\left(1+\frac{\sigma_{\mathsf{h}}^2}{\mu_{\mathsf{h}}^2}\, ln\left(1+\frac{\sigma_{\mathsf{h}}^2}{\sigma_{\mathsf{v}}^2}\right)\right)} - 1\right) \tag{2.30a}$$

$$\gamma = \mu_L - \frac{\sigma_L^2}{\sigma_H^2-\sigma_L^2}(\mu_H-\mu_L) + \frac{\sigma_H\sigma_L}{\sigma_H^2-\sigma_L^2}\sqrt{(\mu_H-\mu_L)^2 + (\sigma_H^2-\sigma_L^2)\, ln\left(\frac{\sigma_H^2}{\sigma_L^2}\right)} \tag{2.30b}$$

$$P_{\mathrm{FA}} = Q\left(\frac{\gamma-\mu_L}{\sigma_L}\right) = Q\left(\frac{\gamma-\mu_{\mathsf{v}}}{\sigma_{\mathsf{v}}}\right) \tag{2.31a}$$

$$P_{\mathrm{M}} = Q\left(\frac{\mu_H-\gamma}{\sigma_H}\right) = Q\left(\frac{\mu_{\mathsf{v}}+\mu_{\mathsf{h}}-\gamma}{\sqrt{\sigma_{\mathsf{v}}^2+\sigma_{\mathsf{h}}^2}}\right) \tag{2.31b}$$

$$P_{\mathrm{E}} = \frac{P_{\mathrm{FA}}+P_{\mathrm{M}}}{2} \tag{2.32}$$

$$P_{\mathrm{FA}} = Q\left(\frac{\mu_{\mathsf{h}}\sigma_{\mathsf{v}}}{\sigma_{\mathsf{h}}^2}\left(\sqrt{\left(1+\frac{\sigma_{\mathsf{h}}^2}{\sigma_{\mathsf{v}}^2}\right)\left(1+\frac{\sigma_{\mathsf{h}}^2}{\mu_{\mathsf{h}}^2}\, ln\left(1+\frac{\sigma_{\mathsf{h}}^2}{\sigma_{\mathsf{v}}^2}\right)\right)} - 1\right)\right) \tag{2.33a}$$

$$P_{\mathrm{M}} = Q\left(\frac{\mu_{\mathsf{h}}\sigma_{\mathsf{v}}}{\sigma_{\mathsf{h}}^2}\left(\sqrt{1+\frac{\sigma_{\mathsf{h}}^2}{\sigma_{\mathsf{v}}^2}} - \sqrt{1+\frac{\sigma_{\mathsf{h}}^2}{\mu_{\mathsf{h}}^2}\, ln\left(1+\frac{\sigma_{\mathsf{h}}^2}{\sigma_{\mathsf{v}}^2}\right)}\right)\right) \tag{2.33b}$$

2.2.2 Log-Normal Channel Coefficient

Definition 2.17. Channel with Log-Normal Fading Coefficient and Gaussian Noise

$$\left.\begin{aligned} \mathsf{y}_n &= \mathsf{u}_n \mathsf{h}_n + \mathsf{v}_n \\ \mathsf{u}_n &\in \{0,1\} \\ \mathsf{h}_n &\in \mathbb{R}^+ \\ \mathsf{h}_n &\sim f_{\mathsf{h}}(\cdot) \\ \mathsf{v}_n &\in \mathbb{R} \\ \mathsf{v}_n &\sim f_{\mathsf{v}}(\cdot) \end{aligned}\right\} \quad \forall n \in \mathbb{N}$$

$$f_{\mathsf{v}}(v) = \frac{1}{\sqrt{2\pi}\sigma_{\mathsf{v}}} exp\left(-\frac{(v-\mu_{\mathsf{v}})^2}{2\sigma_{\mathsf{v}}^2}\right)$$

$$f_{\mathsf{h}}(h) = \frac{1}{\sqrt{2\pi}\sigma_{\mathsf{h}} h} exp\left(-\frac{(ln(h)-\mu_{\mathsf{h}})^2}{2\sigma_{\mathsf{h}}^2}\right)$$

$$\sigma_{\mathsf{v}} > 0,\ \sigma_{\mathsf{h}} > 0$$

Selecting the log-normal distribution for h_n leads to the channel defined in Definition 2.17 with the relations between the channel parameters and the observed mean and variance given by Definition 2.18. Note that the parameters μ_{h} and σ_{h}^2 are not the first two cumulants for the log-normal distribution, they are log-mean and log-variance, respectively.
The likelihood functions are given by Equation 2.34. A closed-form expression is not known, but all cumulants can be computed. To determine the parameters for the fading coefficients' log-normal distribution, the first two are sufficient (cf. Equation 2.35).
The threshold for the ML detector is obtained by the simple comparison in Equation 2.16 after numerical computation of the convolution integral. The resulting error probability vs. SNR is shown in Figure 2.5. In contrast to the Gaussian fading coefficient, the log-normal model's error probabilities can not be displayed in closed form.

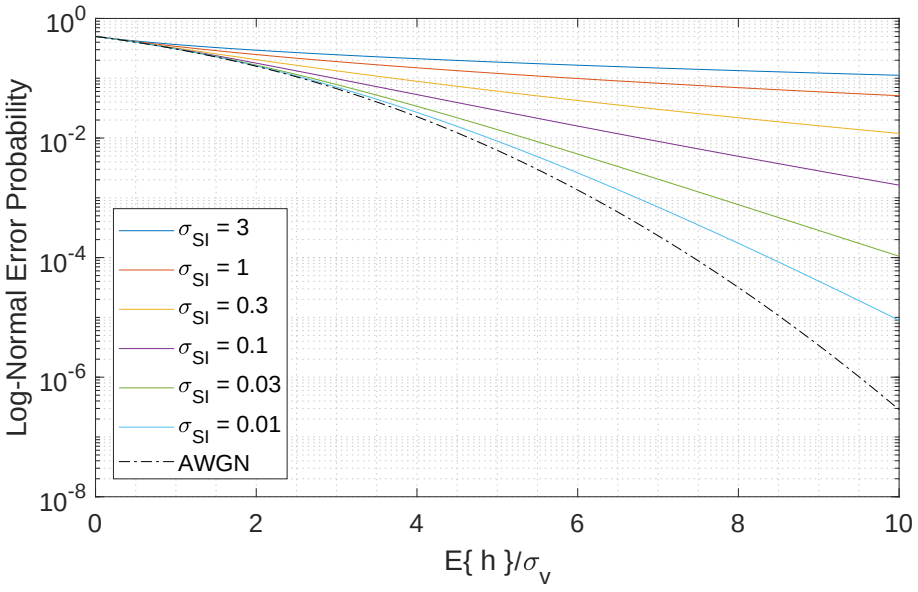


Figure 2.5: Error probability for log-normal channel coefficient with threshold decision vs. SNR in field quantities. The dash-dotted line corresponds to the AWGN channel.

Definition 2.18. Relationships between Stochastic Channel Properties and Observed Quantities for Log-Normal Channel Coefficient

$$
\begin{aligned}
\mu_L &= E\{\mathsf{y}_n \mid \mathsf{u}_n = 0\} && = \mu_{\mathsf{v}} \\
\sigma_L^2 &= E\left\{(\mathsf{y}_n - \mu_L)^2 \mid \mathsf{u}_n = 0\right\} && = \sigma_{\mathsf{v}}^2 \\
\mu_H &= E\{\mathsf{y}_n \mid \mathsf{u}_n = 1\} && = \mu_{\mathsf{v}} + exp\left(\mu_{\mathsf{h}} + \frac{\sigma_{\mathsf{h}}^2}{2}\right) \\
\sigma_H^2 &= E\left\{(\mathsf{y}_n - \mu_H)^2 \mid \mathsf{u}_n = 1\right\} && = \sigma_{\mathsf{v}}^2 + \left(exp\left(\sigma_{\mathsf{h}}^2\right) - 1\right) exp\left(2\mu_{\mathsf{h}} + \sigma_{\mathsf{h}}^2\right)
\end{aligned}
$$

$$f_{\mathsf{y}_n|\mathsf{u}_n}(y \mid 0) = \frac{1}{\sqrt{2\pi}\sigma_{\mathsf{v}}} exp\left(-\frac{(y-\mu_{\mathsf{v}})^2}{2\sigma_{\mathsf{v}}^2}\right) \tag{2.34a}$$

$$f_{\mathsf{y}_n|\mathsf{u}_n}(y \mid 1) = \int_0^\infty \frac{1}{2\pi\sigma_{\mathsf{v}}\sigma_{\mathsf{h}}x} exp\left(-\frac{(ln(x)-\mu_{\mathsf{h}})^2}{2\sigma_{\mathsf{h}}^2} - \frac{(y-x-\mu_{\mathsf{v}})^2}{2\sigma_{\mathsf{v}}^2}\right) dx \tag{2.34b}$$

$$\sigma_{\mathsf{h}}^2 = ln\left(1 + \frac{\sigma_H^2 - \sigma_L^2}{(\mu_H - \mu_L)^2}\right) \tag{2.35a}$$

$$\mu_{\mathsf{h}} = ln(\mu_H - \mu_L) - \frac{\sigma_{\mathsf{h}}^2}{2} \tag{2.35b}$$

2.3 Temporal and Spatial Correlation

Atmospheric turbulence is usually described by Kolmogorov theory – large eddies transfer energy approximately lossless to smaller-scale eddies which in turn redistribute the energy or dissipate through viscosity, leading to an inner scale l_0 and an outer scale L_0, typically ranging from millimeters to meters [23]. The correlation time is derived by assuming static spatial eddies that are moved by the wind (considering only the velocity normal to the beam since the influence of translation along the beam axis will be well-averaged). In case the optical link satisfies $l_0 < \sqrt{\lambda r} < L_0$ for propagation distance r and wavelength λ, the characteristic correlation time τ_c and distance d_c for which the log-amplitude covariance coefficient has decayed to e^{-2} can be approximated by [23]:

Definition 2.19. Characteristic Turbulence Correlation Distance and Time for "Frozen Air" Model

$$d_c \approx \sqrt{\lambda r} \tag{2.36}$$

$$\tau_c = \frac{d_c}{u_\perp} \tag{2.37}$$

for propagation distance r, wavelength λ and orthogonal wind speed $u_\perp$.

Which leads to $d_c \approx 20$ mm and $\tau_c \approx 5$ ms for $u_\perp \approx 4$ m/s. [17] compares the resulting low-pass behavior for several optical links to the estimated cutoff frequency given by Equation 2.38, which is about 500 Hz for the studied link and the assumptions made.

$$f_c = u_\perp \sqrt{\frac{2\pi}{\lambda r}} \tag{2.38}$$

2.3.1 Estimation/Detection Strategies

The absence of high-frequency components in the fading coefficient fluctuation can be exploited to reduce the error probability by estimating the current channel state. Using the ML estimate of the fading coefficient's temporal evolution h_n as "known" input to the ML symbol detector (Equation 2.8)

embodies a GLRT strategy, specified in Definition 2.20. $\hat{h}_{n,0}$ and $\hat{h}_{n,1}$ correspond to the ML estimates of h_n under the corresponding hypotheses. Since the channel output does not depend on the fading state in case "0" has been sent, $\hat{h}_{n,0}$ need not be estimated.

Definition 2.20. Generalized Likelihood Ratio Test for the Current Received Symbol y_n

$$L_g(y) = \frac{f_{\mathsf{y}_n|\mathsf{u}_n,\mathsf{h}_n}\left(y \mid 1, \hat{h}_{n,1}\right)}{f_{\mathsf{y}_n|\mathsf{u}_n,\mathsf{h}_n}\left(y \mid 0, \hat{h}_{n,0}\right)} \tag{2.39}$$

$$\hat{h}_{n,0} = arg\max_{h}\left\{f_{\mathsf{h}_n|\mathsf{u}_n,\mathbf{y}}(h \mid 0, \boldsymbol{y}_{1..n})\right\} \qquad \text{(no need to evaluate)}$$

$$\hat{h}_{n,1} = arg\max_{h}\left\{f_{\mathsf{h}_n|\mathsf{u}_n,\mathbf{y}}(h \mid 1, \boldsymbol{y}_{1..n})\right\} \tag{2.40}$$

$$L_g(y) = \frac{f_{\mathsf{y}_n|\mathsf{u}_n,\mathsf{h}_n}\left(y \mid 1, \hat{h}_{n,1}\right)}{f_{\mathsf{y}_n|\mathsf{u}_n}(y \mid 0)} = \frac{f_{\mathsf{v}}\left(y - \hat{h}_{n,1}\right)}{f_{\mathsf{v}}(y)} \tag{2.41}$$

$$l_g(y) \propto -\frac{\left(y - \hat{h}_{n,1} - \mu_{\mathsf{v}}\right)^2}{2\sigma_{\mathsf{v}}^2} + \frac{(y-\mu_{\mathsf{v}})^2}{2\sigma_{\mathsf{v}}^2} = \frac{\hat{h}_{n,1}\left(y - \hat{h}_{n,1}/2 - \mu_{\mathsf{v}}\right)}{\sigma_{\mathsf{v}}^2} \tag{2.42}$$

From Equation 2.42 it becomes obvious that simply replacing the channel coefficient of the AWGN ML detector by its ML estimate is sufficient to constitute a GLRT ML detector. The three main problems wrt. the optical channel are the correlation model (only approximations are known), the absence of information about h_n given $\mathsf{u}_n = 0$ and the numerical complexity of the estimation itself.

The second notable method uses the joint distribution of a sequence of channel coefficients to determine which series of data most likely caused the received data. This is called the maximum likelihood sequence detector (MLSD) as defined in Definition 2.3. The likelihood $f_{\mathbf{y}|\mathbf{u}}(\boldsymbol{y} \mid \boldsymbol{u})$ is obtained (e.g. in [18]) via marginalization of $f_{\mathbf{y},\mathbf{h}|\mathbf{u}}$ over the support $R_{\mathbf{h}}$ of $\mathbf{h}$:

$$\hat{\boldsymbol{u}}(\boldsymbol{y}) = arg\max_{\boldsymbol{u}}\left\{f_{\mathbf{y}|\mathbf{u}}(\boldsymbol{y} \mid \boldsymbol{u})\right\} \tag{2.43a}$$

$$\hat{\boldsymbol{u}}(\boldsymbol{y}) = arg\max_{\boldsymbol{u}}\left\{\int_{R_{\mathbf{h}}} f_{\mathbf{h}}(\boldsymbol{h})\, f_{\mathbf{y}|\mathbf{u},\mathbf{h}}(\boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{h})\, d\boldsymbol{h}\right\} \tag{2.43b}$$

Since $\mathbf{y} \big| \mathbf{u}, \mathbf{h}$ depends only on the (white) noise $\mathbf{v}$, the last density decomposes to a product of the N independent terms:

$$\begin{aligned} f_{\mathbf{y}\,|\,\mathbf{u},\mathbf{h}}\left(\boldsymbol{y} \,\middle|\, \boldsymbol{u}, \boldsymbol{h}\right) &= \prod_{i=1}^{N} f_{\mathsf{y}_i\,|\,\mathsf{u}_i,\mathsf{h}_i}\left(y_i \mid u_i, h_i\right) \\ &= \prod_{i=1}^{N} f_{\mathsf{v}}\left(y_i - u_i h_i\right) \end{aligned} \tag{2.44}$$

$$\hat{\boldsymbol{u}}\left(\boldsymbol{y}\right) = arg\,\underset{\boldsymbol{u}}{max} \left\{ \int_{R_{\mathbf{h}}} f_{\mathbf{h}}\left(\boldsymbol{h}\right) exp\left(-\sum_{i=1}^{N} \left(y_i - u_i h_i - \mu_{\mathsf{v}}\right)^2 \right) d\boldsymbol{h} \right\} \tag{2.45}$$

The problems of the MLSD mirror are twofold: The temporal correlation has to be encoded in $f_{\mathbf{h}}\left(\boldsymbol{h}\right)$ and the complexity of the MLSD in Equation 2.45 increases with $N \cdot 2^N$ for sequence length N (evaluating an N-dimensional integral for each of the 2^N possible sequences), rendering it impractical for most applications, but suboptimal approximations exist [18].
While the GLRT ML detector can be approximated on-line, the MLSD operates on (long) sequences of received bits, implying a certain delay added by the blockwise nature of the detector.
A (suboptimal) GLRT ML strategy still leads to temporally correlated errors – even with perfect knowledge of the fading coefficient over time, cf. Equation 2.11b. The implications will be discussed in Chapter 4.
An optimal GLRT detector needs a generative model for the fading state evolution, e.g. the single-step Markov chain (SMC) approximation in the form of $f_{\mathsf{h}_n\,|\,\mathsf{h}_{n-1}}\left(h_n \,\middle|\, h_{n-1}\right)$, which also allows sequential Bayesian estimation (Definition 2.21). Note that in the general, nonlinear case, this solution is prohibitively complex – typically Markov Chain Monte Carlo (MCMC) approaches like the particle filter are employed to approximate the optimal solution. In the linear-Gaussian case, the Kalman filter is the exact representation of the sequential estimation in Definition 2.21.

Definition 2.21. Sequential Bayesian Estimation [25, 26]

$$f_{\boldsymbol{\theta}_n|\mathbf{y}}(\boldsymbol{\theta}_n \mid \boldsymbol{y}_{1..n-1}) = \int_{R_{\boldsymbol{\theta}_{n-1}}} f_{\boldsymbol{\theta}_n|\boldsymbol{\theta}_{n-1}}(\boldsymbol{\theta}_n \mid \boldsymbol{\theta}_{n-1})\, f_{\boldsymbol{\theta}_{n-1}|\mathbf{y}}(\boldsymbol{\theta}_{n-1} \mid \boldsymbol{y}_{1..n-1})\, d\boldsymbol{\theta}_{n-1} \tag{2.46}$$

$$f_{\boldsymbol{\theta}_n|\mathbf{y}}(\boldsymbol{\theta}_n \mid \boldsymbol{y}_{1..n}) = \frac{f_{\mathbf{y}_n|\boldsymbol{\theta}_n}(\boldsymbol{y}_n \mid \boldsymbol{\theta}_n)\, f_{\boldsymbol{\theta}_n|\mathbf{y}}(\boldsymbol{\theta}_n \mid \boldsymbol{y}_{1..n-1})}{f_{\mathbf{y}_n|\mathbf{y}}(\boldsymbol{y}_n \mid \boldsymbol{y}_{1..n-1})} \tag{2.47}$$

$$f_{\mathbf{y}_n|\mathbf{y}}(\boldsymbol{y}_n \mid \boldsymbol{y}_{1..n-1}) = \int_{R_{\boldsymbol{\theta}_n}} f_{\mathbf{y}_n|\boldsymbol{\theta}_n}(\boldsymbol{y}_n \mid \boldsymbol{\theta}_n)\, f_{\boldsymbol{\theta}_n|\mathbf{y}}(\boldsymbol{\theta}_n \mid \boldsymbol{y}_{1..n-1})\, d\boldsymbol{\theta}_n \tag{2.48}$$

With initialization $f_{\boldsymbol{\theta}_n|\mathbf{y}}(\boldsymbol{\theta}_n \mid \boldsymbol{y}_{1..n})\big|_{n=0} = f_{\boldsymbol{\theta}_0}(\boldsymbol{\theta}_0)$ and repeated application of prediction (2.46) and update (2.47) steps, the posterior pdf $f_{\boldsymbol{\theta}_n|\mathbf{y}}(\boldsymbol{\theta}_n \mid \boldsymbol{y}_{1..n})$ is calculated for each data vector $\boldsymbol{y}_n$.
In the linear-Gaussian case, i.e. if $\boldsymbol{\theta}_0$, $\boldsymbol{\theta}_n \mid \boldsymbol{\theta}_{n-1}$ and $\mathbf{y}_n \mid \boldsymbol{\theta}_n$ are Gaussian, Equations 2.46 to 2.48 reduce to the Kalman filter equations.

A special case is the linear/affine state space model (Equation 2.49) with correlation a and Gaussian driving noise/innovation w (zero-mean and variance σ_{w}) that leads to the state transition pdf given in Equation 2.50. The Kalman filter for Equation 2.49 is given by Equations 2.51 through 2.52, where the indices 0 and 1 indicate the sent symbol. Note that this knowledge is not required for the GLRT detector a priori, but for the following Kalman filter step. It is proposed to use the (GLRT-ML) detected symbol as given by Equation 2.53, leading to the possibility of error propagation.

$$\mathsf{h}_n = a\mathsf{h}_{n-1} + (1-a)\,\mu_{\mathsf{h}} + \mathsf{w} \tag{2.49}$$

$$f_{\mathsf{h}_n|\mathsf{h}_{n-1}}(h_n \mid h_{n-1}) = \frac{1}{\sqrt{2\pi}\sigma_{\mathsf{w}}}\, exp\left(-\frac{(h_n - ah_{n-1} - (1-a)\,\mu_{\mathsf{h}})^2}{2\sigma_{\mathsf{w}}^2}\right) \tag{2.50}$$

$$\hat{h}_{n,0} = a\hat{h}_{n-1} + (1-a)\,\mu_{\mathsf{h}} \tag{2.51a}$$
$$\sigma^2_{e,n,0} = a^2\sigma^2_{e,n-1} + \sigma^2_{\mathsf{w}} \tag{2.51b}$$

$$\hat{h}_{n,1} = a\hat{h}_{n-1} + (1-a)\,\mu_{\mathsf{h}} + \frac{a^2\sigma^2_{e,n-1} + \sigma^2_{\mathsf{w}}}{a^2\sigma^2_{e,n-1} + \sigma^2_{\mathsf{w}} + \sigma^2_{\mathsf{v}}}\left(y_n - a\hat{h}_{n-1} - (1-a)\,\mu_{\mathsf{h}}\right) \tag{2.52a}$$
$$\sigma^2_{e,n,1} = \frac{\left(a^2\sigma^2_{e,n-1} + \sigma^2_{\mathsf{w}}\right)\sigma^2_{\mathsf{v}}}{a^2\sigma^2_{e,n-1} + \sigma^2_{\mathsf{w}} + \sigma^2_{\mathsf{v}}} \tag{2.52b}$$

$$\hat{h}_n = \begin{cases} \hat{h}_{n,0} & \mathsf{y} \le \mu_{\mathsf{v}} + \dfrac{\hat{h}_{n,1}}{2}, \\ \hat{h}_{n,1} & \mathsf{y} > \mu_{\mathsf{v}} + \dfrac{\hat{h}_{n,1}}{2}. \end{cases} \tag{2.53a}$$

$$\sigma_{e,n} = \begin{cases} \sigma_{e,n,0} & \mathsf{y} \le \mu_{\mathsf{v}} + \dfrac{\hat{h}_{n,1}}{2}, \\ \sigma_{e,n,1} & \mathsf{y} > \mu_{\mathsf{v}} + \dfrac{\hat{h}_{n,1}}{2}. \end{cases} \tag{2.53b}$$

Since the Kalman filter is likely still too computationally intensive for online operation at FSO communication frequencies, the results are mostly of theoretical nature. A more general, yet also theoretic result is the spectral flatness measure Ξ_{h} [27] as given by Definition 2.22.

Definition 2.22. Spectral Flatness Measure (SFM) Ξ_{h} for Gauss-Markov h with power spectral density (PSD) $S_{\mathsf{hh}}(\omega)$

$$\Xi_{\mathsf{h}} = exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} ln\left(S_{\mathsf{hh}}(\omega)\right)d\omega\right), \qquad 0 < \Xi_{\mathsf{h}} \le 1 \qquad \text{(general)}$$
$$\Xi_{\mathsf{h}} = \frac{\sigma^2_{\mathsf{w}}}{\sigma^2_{\mathsf{h}}} = 1 - a^2 \qquad \text{(for Equation 2.49)}$$

The gain of exploiting correlations within the sequence h_n is upper bounded by the SFM; i.e. the minimum achievable variance of the predicted fading coefficient equals the added "innovation" variance σ^2_{w} – a reduction by the factor $1-a^2$ for the SMC case.

2.3.2 Simulated Correlated Channels

Two reduced-complexity estimators are simulated in this thesis. Both do not depend on knowledge of transmitted data or detected bits – they operate on the assumption of uniform data (i.e. 50% high and low bits each) without extremely long sequences of ones or zeros. Most importantly, they can be implemented in simple electronic hardware in the form of low-order filters that operate directly on the received symbols y_n.
The first approach is designing a stationary Kalman filter to estimate $\mu_\mathsf{v}+\mathsf{h}/2$ for a plant with no observed inputs (the colored noise model) with additional observation variance $E\{\mathsf{h}^2\}/4$ based on Equation 2.54, which follows from the pairwise independence of h, v and u; as well as u taking only values 0 and 1. The last expression holds for log-normal h.
In order to design a Kalman filter for the nonlinear transformed random value, the exponentiation was linearized around the expected value of h, i.e. the Kalman filter was designed for the correlation model scaled by the constant $exp\left(E\{\mathsf{h}\}\right) = exp\left(\mu_\mathsf{h} + \frac{\sigma_\mathsf{h}^2}{2}\right)$.

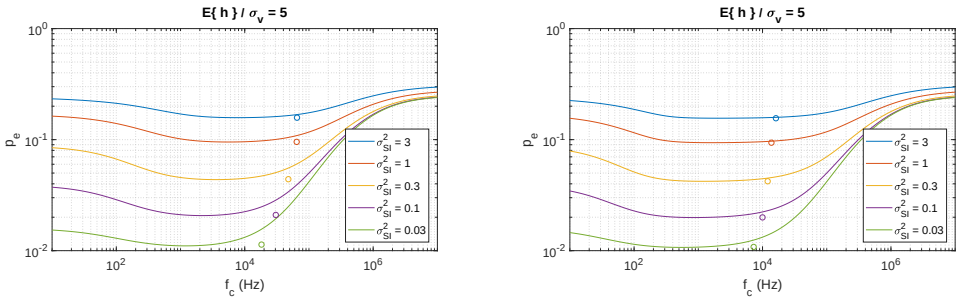
$$Var\{\mathsf{y}-\gamma_{\text{opt}}\} = E\left\{\left(\mathsf{uh}+\mathsf{v}-\left(\mu_\mathsf{v}+\frac{\mathsf{h}}{2}\right)\right)^2\right\} \tag{2.54a}$$

$$= \sigma_\mathsf{v}^2 + E\left\{\left(\mathsf{h}\frac{2\mathsf{u}-1}{2}\right)^2\right\} \tag{2.54b}$$

$$= \sigma_\mathsf{v}^2 + \frac{1}{4}E\{\mathsf{h}^2\} \tag{2.54c}$$

$$= \sigma_\mathsf{v}^2 + \frac{1}{4}\,exp\left(2\mu_\mathsf{h}+2\sigma_\mathsf{h}^2\right) \tag{2.54d}$$

The second approach consists of a simple first-order infinite impulse response (IIR) low-pass filter. The performance for different cutoff frequencies ranging from below the characteristic coherence frequency of the channel coefficient (Equation 2.38) to near the communication frequency.



(a) (b)

Figure 2.6: Error probability of a temporally correlated log-normal channel coefficient vs. filter cutoff frequency for a SNR of 5, for (a) SMC (first-order) and (b) Bessel filter model. Circles indicate Kalman-based operating points, lines refer to the threshold derived by the first-order lowpass.

Simulations have been performed for log-normal channel coefficients by using the method presented by [17]: Low-pass filtering (cutoff frequency of 100 Hz have been assumed) of Gaussian iid random variables, applying affine transformations to get required log-mean and log-variance, exponentiation. The resulting random sequence does obviously not comply to Gaussian assumptions – a clear gain over the ML symbol detector manifests and thus justifies these low-complexity, suboptimal estimators.

Results are shown in Figures 2.6, 2.7 and 2.8 for SNR of 5, 20 and 80 respectively (in field quantities) – the corresponding $\mathrm{SNR}_{\mathrm{dB}}$ values are approximately 14, 26 and 38 dB. Images (a) show the performance for the SMC model for h while the second set is concerned with a second-order Bessel filter as generative model (based on [17]). The operating points of the Kalman filter for each scenario is indicated by a circle.

The performance of the first-order IIR low-pass largely follows a U-shaped behavior over the filter cutoff frequency. Over a relatively wide range, error rates are near the minimum: Here, the filter successfully averages over enough values to suppress the noise, but reacts sufficiently fast to track the variation of the fading coefficient. The high-frequency end has poor error rates since the filter output (i.e. the estimated channel coefficient) closely follows the received symbols, with little to no averaging of noise introduced by u and v.

Too low cutoff frequencies essentially lead to a mean-based estimator whose performance strongly depends on the scintillation index. This arises as a consequence of Equation 2.55 – recall that, based on Definition 2.13, $\sigma_{\mathsf{h}}^2 = ln\left(1+\sigma_{\mathrm{SI}}^2\right)$. The exact error probabilities are given by Equations 2.56a and 2.56b, for high SNR, the Q-function can be substituted by 1 and 0 for negative and positive numerators, respectively. This leads to the approximations in Equations 2.57a and 2.57b.

$$E\left\{\mathsf{uh}+\mathsf{v}\right\} = \mu_{\mathsf{v}} + \frac{1}{2}E\left\{\mathsf{h}\right\} \tag{2.55a}$$

$$= \mu_{\mathsf{v}} + \frac{1}{2}\,exp\left(\mu_{\mathsf{h}} + \frac{\sigma_{\mathsf{h}}^2}{2}\right) \tag{2.55b}$$

$$P_{\mathrm{FA}} = Q\left(\frac{exp\left(\mu_{\mathsf{h}} + \frac{\sigma_{\mathsf{h}}^2}{2}\right)}{2\sigma_{\mathsf{v}}}\right) \tag{2.56a}$$

$$P_{\mathrm{M}} = \int_0^\infty f_{\mathsf{h}}\left(h\right) Q\left(\frac{2h - exp\left(\mu_{\mathsf{h}} + \frac{\sigma_{\mathsf{h}}^2}{2}\right)}{2\sigma_{\mathsf{v}}}\right) dh \tag{2.56b}$$

$\sigma_{\mathsf{v}} \ll exp\left(\mu_{\mathsf{h}} + \frac{\sigma_{\mathsf{h}}^2}{2}\right)$ leads to:

$$P_{\mathrm{FA}} \approx 0 \tag{2.57a}$$

$$P_{\mathrm{M}} \approx Q\left(\frac{ln\left(2\right) - \frac{\sigma_{\mathsf{h}}^2}{2}}{\sigma_{\mathsf{h}}}\right) \tag{2.57b}$$

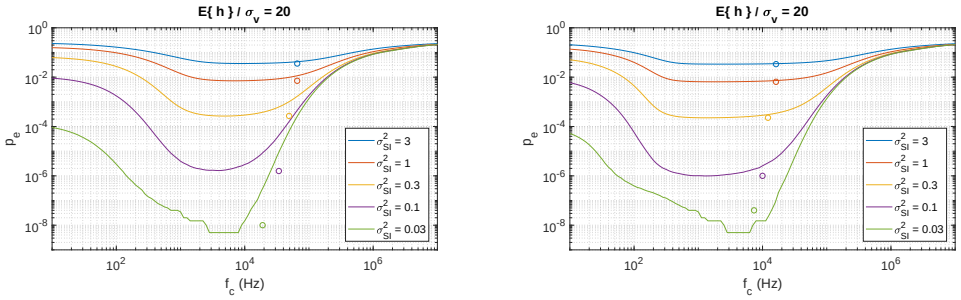


Figure 2.7: Error probability of a temporally correlated log-normal channel coefficient vs. filter cutoff frequency for a SNR of 20, for (a) SMC (first-order) and (b) Bessel filter model. Circles indicate Kalman-based operating points, lines refer to the threshold derived by the first-order lowpass.

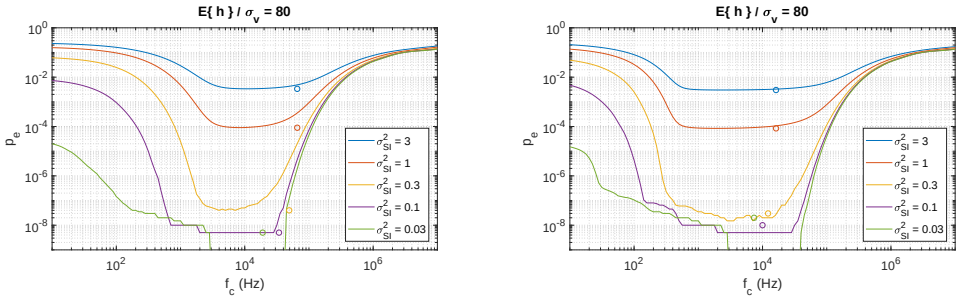


Figure 2.8: Error probability of a temporally correlated log-normal channel coefficient vs. filter cutoff frequency for a SNR of 80, for (a) SMC (first-order) and (b) Bessel filter model. Circles indicate Kalman-based operating points, lines refer to the threshold derived by the first-order lowpass.

3

Channel Measurements

During the course of this thesis, measurements of different optical channels have been performed to study their behavior and generate, verify or disconfirm hypotheses. The first experiments have been conducted solely on the optical table, setting up the data acquisition (DAQ) together with the laser diode and the sensor. Turbulence and temperature variations have been introduced by using a blow dryer directed at the optical path. No discernible intensity fluctuations were observed at the oscilloscope or in the received data.
Further measurements have been carried out on the FSO channel. Preliminary results confirmed the sufficiency of the log-normal and Gaussian models for the fading coefficient h. After several changes in the hardware, the setup used for the final measurement campaign in February 2019 has been settled on the combination laid out in the next section.

3.1 Measurement System

3.1.1 Hardware Setup

The optical wavelength used is 650 nm. The setup consists of a PC connected to a data acquisition instrument (U2531A, Keysight Technologies, Santa Rosa, California, USA) with 14 bit, 2 MSa/s analog input with 10 V fullscale and 1 MSa/s output rate. Its output is connected to a laser diode pointed at a retro-reflector in 300 m distance exposing the beam to urban environment. The reflected beam has been captured by a 14 in (354 mm) telescope, which also serves as mounting for the aforementioned laser diode. The optical path on the receiving side includes a collimating lens, a static tip/tilt mirror, a beam splitter, an unused (flat) deformable mirror, an optional neutral-density (ND) filter as attenuator, a band-pass filter with 10 nm FWHM (full-width at half maximum) and the sensor (APD120A, Thorlabs, Newton, New Jersey, USA), which is also connected to the DAQ device.
The setup was originally planned for use with adaptive optics and to characterize its influence on the communication channel [28]. During the course of this work, the adaptive optics have not been used. The hardware setup is depicted in Figure 3.1.

3.1.2 Software Framework

The Matlab script used to provide the stimulus and measure the channel response is summarized in Figures 3.2 to 3.4. Communication between the PC and the DAQ device has been established via the Virtual Instrument Software Architecture (VISA) standard over USB. The command language of the U2531A DAQ card is Standard Commands for Programmable Instruments (SCPI). The propagation delay of the optical link (2 µs, 300 m distance to retro-reflector) has been corrected for by specifying delayed sampling of the sensor's signal.
The output data is a pseudo-random bit sequence (PRBS) of m-Sequence [29] type generated by a linear shift register realized in code. This type of PRBS offers the advantages of being deterministic, (almost) balanced between the binary states and also containing consecutive "runs" of equal states.

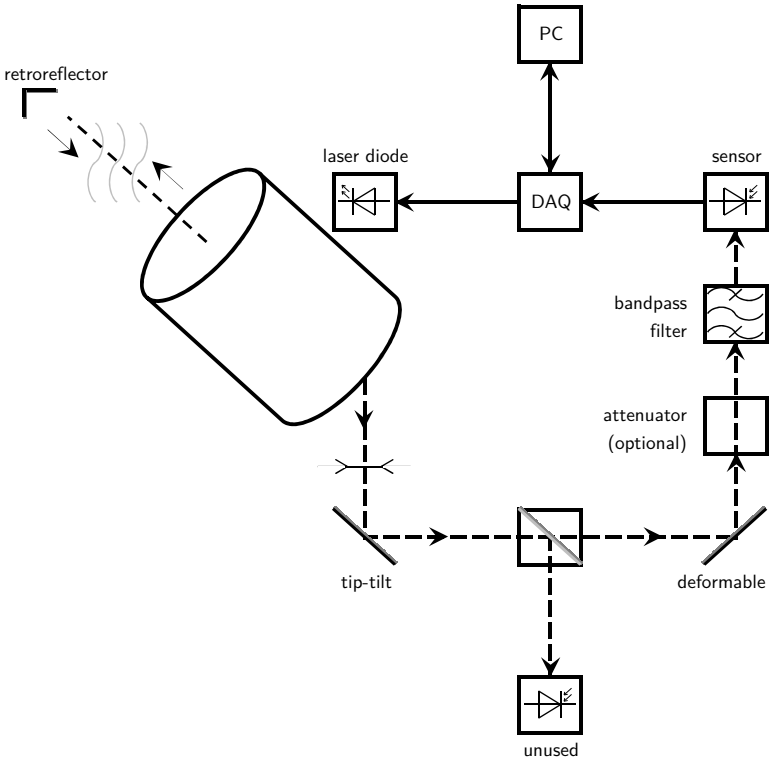


Figure 3.1: Measurement equipment hardware setup, dashed lines indicate the optical path. The laser diode was mounted to the telescope and pointed directly at the retro-reflector (top left).

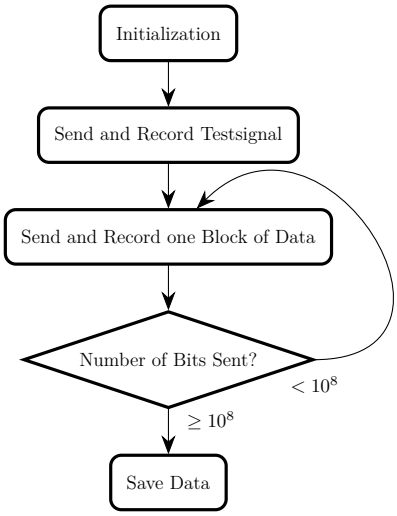


Figure 3.2: Flowchart of the channel measurement script.

Before sending 10^8 bits of data and recording the measured intensity, the script sends and records the testsignal, based on which preliminary statistics (conditional means and variances) are displayed on the PC. It also consists of m-Sequence-generated pseudo-data.
The difference between the input and output sampling rates (2 MSa/s output vs. 1 MSa/s input) is not addressed by the channel measurement script, both signals are saved with the corresponding rate. For the bit error rate (BER) calculation, two methods are possible: viewing the system as 2 MSa/s and assuming each sent bit as duplicated; and averaging the two received signal chips corresponding to the same output bit, effectively increasing the resolution by one bit. Due to the already high SNR observed, the first method was used.
Communication with the DAQ instrument is performed in blocks of 250,000 output bits and twice as many received values, yielding a duration of 0.25 s per block. Overall, 400 blocks have been captured for each scenario, totaling 10^8 output bits and $2 \cdot 10^8$ received values over 100 seconds. The testsignals consist of a single block.

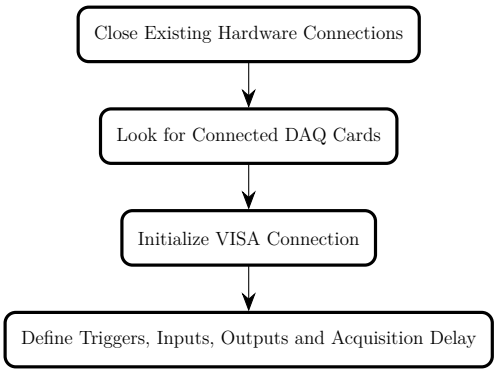


Figure 3.3: Flowchart of the initialization stage.

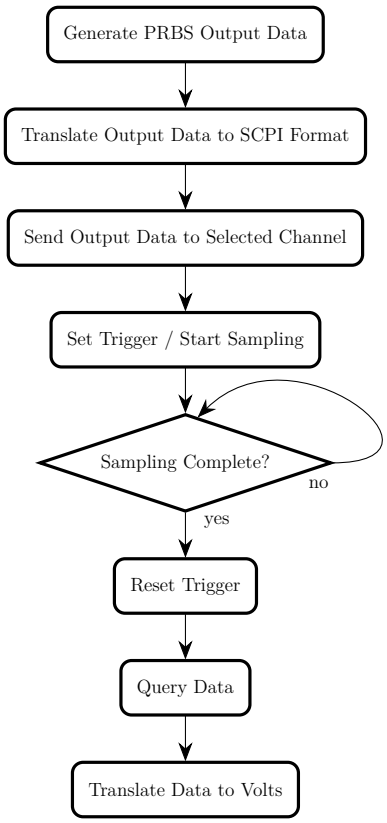


Figure 3.4: Flowchart of the send/record script.

Table 3.1: Summary of the three measurement scenarios.

Scenario	μ_L (mV)	μ_H (mV)	σ_L (mV)	σ_H (mV)	$\mu_H - \mu_L$ (mV)	σ^2_{SI}
500 mV	48.2	543	5.35	100.8	494.5	0.0414
13 mV	8.28	21.2	1.23	2.99	12.9	0.0443
7 mV	4.47	11.7	1.27	1.63	7.24	0.0199

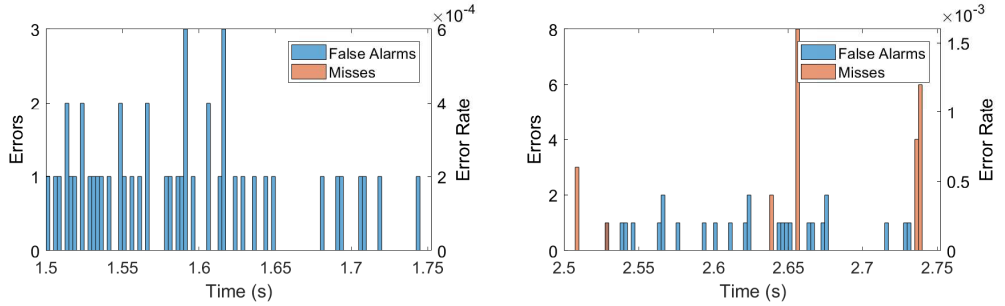
3.2 Measurement Results

The ND filter is used to attenuate the optical signal in order to simulate path loss (FSPL or atmospheric extinction). Three exemplaric scenarios have been created: one without the ND filter, amounting to a estimated mean signal strength $\mu_H - \mu_L$ of approximately 500 mV, and two with varying degrees of attenuation, referred to as 13 mV and 7 mV signal scenario hereafter. Various properties of the three scenarios are summarized in Table 3.1.

The variation of measured μ_L, which should equal the additive noise mean under ideal conditions, can be caused by inter-symbol interference (ISI) and by different ambient light conditions. Since the value varies by more than a factor 10 and no change in background light justifying such a prominent effect was observed by the naked eye, ISI is deemed as the prime cause. This also applies to the increased σ_L of the 500 mV scenario. An example of the influence the sent data had on the measured values via ISI is given in Figure 3.6a. While some ISI after transition from high to low state is observed, another (nonlinear) effect is more prominent: The received value after a rising edge of the data is consistently greater than the following values, appearing like an exponentially decaying high-pass overshoot. After negative edges, no such effect is observed, only the small ISI noted above.

While the observed scintillation index σ^2_{SI} is approximately equal for the 500 mV and the 13 mV scenario, it differs by more than a factor of 2 for the 7 mV case. The cause of this is not known, though the higher σ_L (compared to the 13 mV scenario) seems at least in part responsible – note that as per Definition 2.13, $\sigma^2_{\mathrm{SI}} = \left(\sigma_H^2 - \sigma_L^2\right) / \left(\mu_H - \mu_L\right)^2$.

In terms of field-quantity SNR $\left(\mu_H - \mu_L\right)/\sigma_L$, the scenarios yield SNR values of 92.4, 10.5 and 5.70, while $\mathrm{SNR}_{\mathrm{dB}}$ values are 39.3, 20.4 and 15.1 dB respectively.

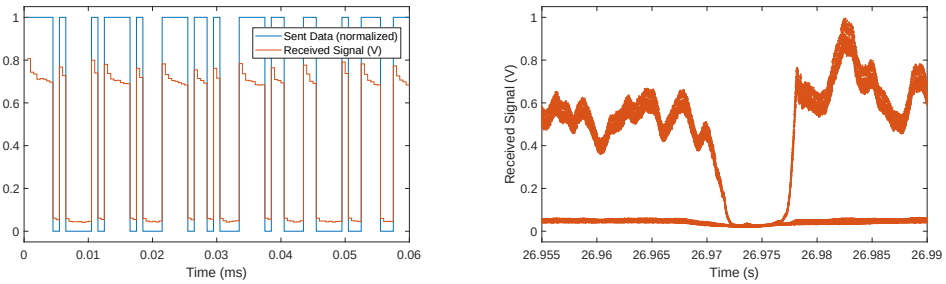


(a) (b)

Figure 3.5: Example error occurrences based on channel measurements during (a) high-SNR interval and (b) including low-SNR intervals. Note the similar false alarm rate.

An anomaly that happened during the measurement of the 500 mV scenario is shown in Figure 3.6b. The received value for present laser power shows a decline for about 2 ms, then remains essentially at μ_L for 5 ms and reverts to normal levels within 1 ms. The reason for this glitch is not known, one possible explanation is line-of-sight blocking by a bird, of which there are plenty in the area around the FSO link studied. Apart from this short digression, the 500 mV scenario is virtually error free: The ML symbol detector threshold for the log-normal model leads to not a single false alarm, but 5677 misses during the event shown in Figure 3.6 while the ML threshold for the Gaussian model accumulates 219 false alarms in total ($P_{\mathrm{FA}} = 1.10 \cdot 10^{-6}$) and 5447 misses during said excursion.

Figure 3.5 shows example error occurrences during high-SNR and low-SNR intervals in the 13 mV measurement, with detection threshold γ based on log-normal model for h and without using the temporal correlation of the fading coefficient. A more or less constant, low false alarm rate is observed, with occasional bursts of several misses in a short period. This is caused by a temporarily low value of the fading coefficient – near enough to the threshold that the additive noise at the receiver corrupts some bits where laser power was present by pushing the received value below the threshold. Due to the temporal correlation of the fading state, these errors occur in bundles.

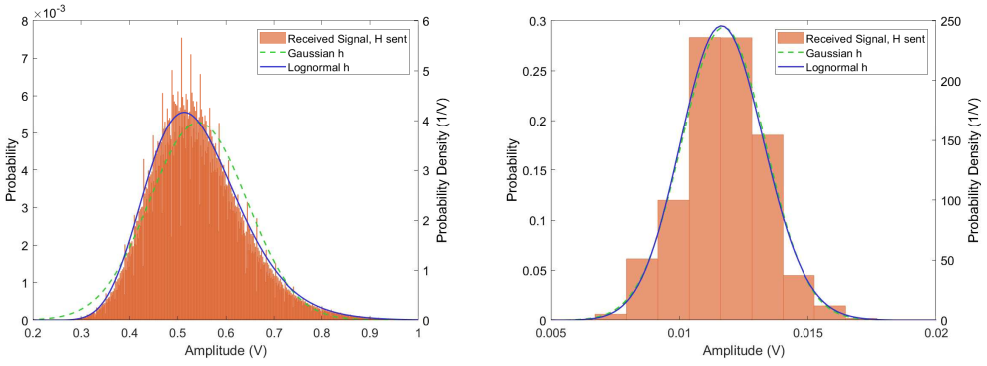


(a) (b)

Figure 3.6: Deviations of measurement data from ideal behavior: (a) ISI and high-pass overshoot and (b) suspected avian line-of-sight blocking.

3.2.1 Relative Model Quality

In order to rank the Gaussian and log-normal models for h according to their relative quality, the Akaike information criterion (AIC) has been used. The AIC is based on an estimator of relative Kullback-Leibler (KL) information [30, 31]. Definition 3.1 states the AIC in this thesis' nomenclature and the Δ_i values used for model selection (lower AIC indicates better models). Evidently, the best model i_{opt} has $\Delta_{i_{\text{opt}}} = 0$ and all others show $\Delta_i \geq 0$.



(a) (b)

Figure 3.7: Histogram of received laser on pulse with estimated pdfs of Gaussian and log-normal model for 100 s measurement, for (a) 500 mV and (b) 7 mV scenario.

Table 3.2: Observed AIC values for Gaussian and log-normal channel coefficient.

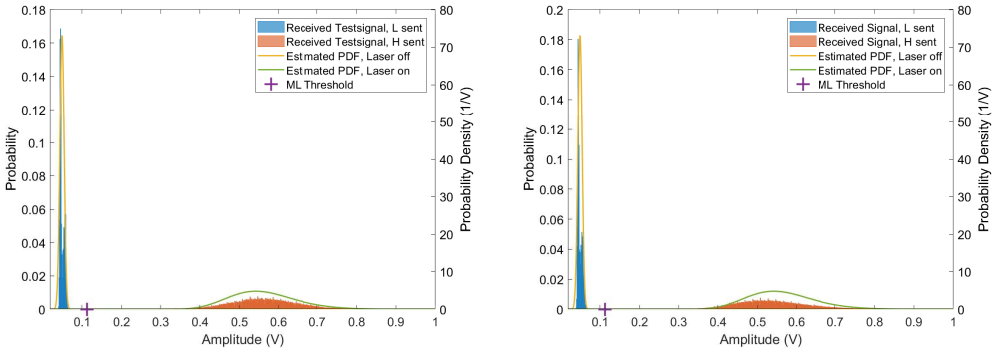
Scenario	$\boldsymbol{AIC}_{\mathbf{Gauss}}$	$\boldsymbol{AIC}_{\mathbf{LN}}$	$\boldsymbol{\Delta}_{\mathbf{Gauss}}$	$\boldsymbol{\Delta}_{\mathbf{LN}}$
500 mV	$-1.75 \cdot 10^8$	$-1.83 \cdot 10^8$	$7.72 \cdot 10^6$	0
13 mV	$-8.79 \cdot 10^8$	$-8.85 \cdot 10^8$	$5.83 \cdot 10^6$	0
7 mV	$-1.00 \cdot 10^9$	$-9.99 \cdot 10^8$	0	$1.48 \cdot 10^5$

Note that the AIC approach solely compares models and identifies the best among them, but needs "good" models to work with. The theoretic considerations regarding the noise sources in this thesis together with the comparison between the histograms and the estimated pdfs (e.g. Figure 3.7) justify the use of the AIC in order to quantify which model produces a better fit.

Definition 3.1. Akaike information criterion (AIC) for model i with k ML-estimated parameters $\hat{\boldsymbol{\theta}}_i$ and Δ_i value for model selection [30, 31]. The data $\boldsymbol{y}$ is considered fixed.

$$\begin{aligned} AIC_i &= -2\,ln\left(f_{\mathbf{y};\boldsymbol{\theta}_i}\left(\boldsymbol{y};\hat{\boldsymbol{\theta}}_i\right)\right) + 2k \\ &= -2\,ln\left(\max_{\boldsymbol{\theta}_i}\left\{f_{\mathbf{y};\boldsymbol{\theta}_i}(\boldsymbol{y};\boldsymbol{\theta}_i)\right\}\right) + 2k \\ \Delta_i &= AIC_i - \min_j\left\{AIC_j\right\} \end{aligned}$$

Since the models only differ in their respective density functions $f_{\mathsf{y}\,|\,\mathsf{u}}(y \,|\, 1)$, the AIC corresponding to the received values given "high" bits are calculated. For the AWGN and the Gaussian model, the AIC is easily computed analytically, while for the log-normal model numeric integration is necessary. In order to circumvent potential distortions arising from this unequal treatment, all AICs were calculated by this numeric approach.
For all possible received values (based on 14 bit resolution of the DAQ instrument), the log-likelihood has been calculated, multiplied with the respective number of occurrences in the data and summed. For the Gaussian model, the likelihood per possible received value obtained analytically matched those

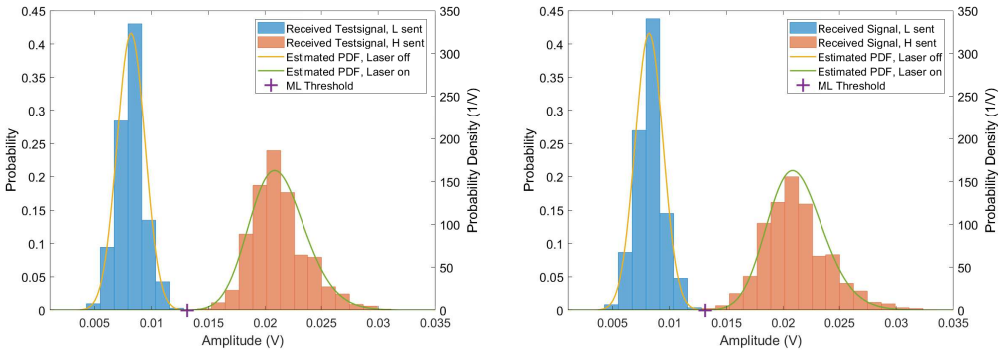


(a) (b)

Figure 3.8: Estimated probability density functions and detection threshold for log-normal channel coefficient. Histograms show (a) testsignal and (b) 100 s measurement of 500 mV scenario.

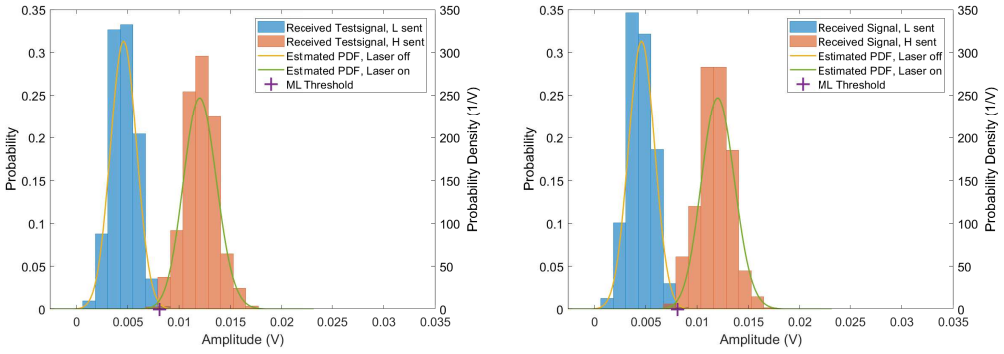
acquired by numeric integration. The results are shown in Table 3.2. For the 500 mV and the 13 mV scenario, the log-normal model yields the better fit, for the 7 mV case, the Gaussian model leads to lower AIC. This is likely due to the receiver noise becoming dominant over the weak signal.

Figures 3.8 to 3.10 show histograms of the received signal for the testsignal and the whole 100 s of measurement. The superimposed pdfs and the ML threshold are based on the respective testsignals. For the 500 mV scenario, Figure 3.8b shows a slight mismatch between the testsignal and the long measurement, indicating that the former was captured while the channel had an above-average fading coefficient. For the other two scenarios, the granularity resulting from the quantization is clearly visible.



(a) (b)

Figure 3.9: Estimated probability density functions and detection threshold for log-normal channel coefficient. Histograms show (a) testsignal and (b) 100 s measurement of 13 mV scenario.



(a) (b)

Figure 3.10: Estimated probability density functions and detection threshold for log-normal channel coefficient. Histograms show (a) testsignal and (b) 100 s measurement of 7 mV scenario.

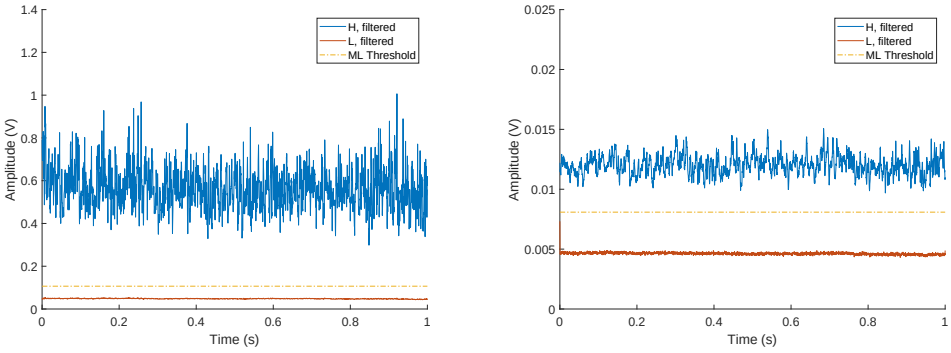
3.2.2 Channel Coefficient Estimation

Figures 3.11 and 3.12 show estimations of the true channel coefficient by acausal, zero-phase low-pass filtering in order to suppress the additive receiver noise. Figure 3.11 displays a whole second and thus does not exhibit clearly visible correlations, while Figure 3.12 reveals them by ranging over 100 ms only. The ML threshold is based on the log-normal model for the channel coefficient h. The latter diagram also shows the noisy received values as translucent dots, helping to visualize how often the corresponding values appear in the data. For Figure 3.12b, they form distinct lines due to the finite quantization (resolution of 0.61 mV).

The properties of the received signal have been calculated block-wise for the 500 mV scenario in Figure 3.13 – showing the temporal evolution of each block's estimated μ_L, μ_H, σ_L and σ_H. The resulting parameters for the log-normal model are given in Figure 3.14a while the observed scintillation index σ_{SI} is depicted in Figure 3.14b.

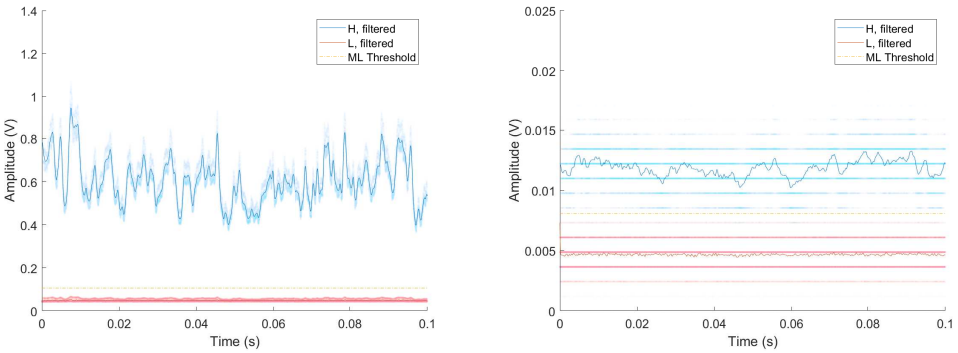
Finally, the frequency content of the fading coefficient has been analyzed by calculating the autocorrelation (Figure 3.15a) and the spectrogram of the received signal points where the laser was on. In order to circumvent the influence of the high-pass overshoot (Figure 3.6a), only each second measured value is included, thereby excluding these peaks. The coherence time τ_c where the correlation decays to e^{-2} is observed to be 2.59 ms, placing the corresponding cutoff frequency in the order of 1 kHz, about twice the value assumed in Section 2.3. Note that the orthogonal wind speed is directly proportional to the cutoff frequency as per Equation 2.38.

The spectrogram shown in Figure 3.15b is based on the AC component of the signal. By this manipulation, low frequency components do not get masked by the strong DC value. As expected, the influence is stronger for low frequencies, but a distinct cutoff frequency is not visible.



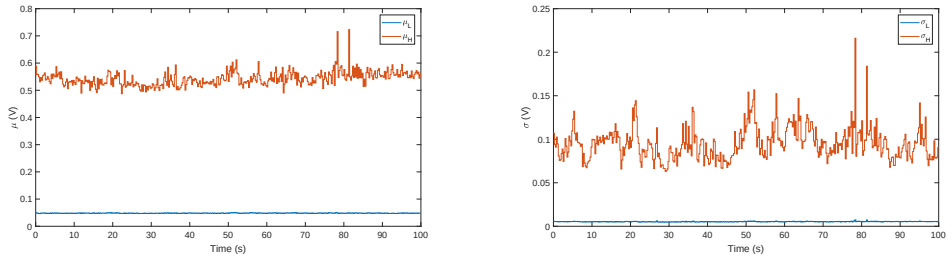
(a) (b)

Figure 3.11: Estimated channel coefficient for 1 s of measurement: (a) 500 mV signal, (b) 7 mV signal.



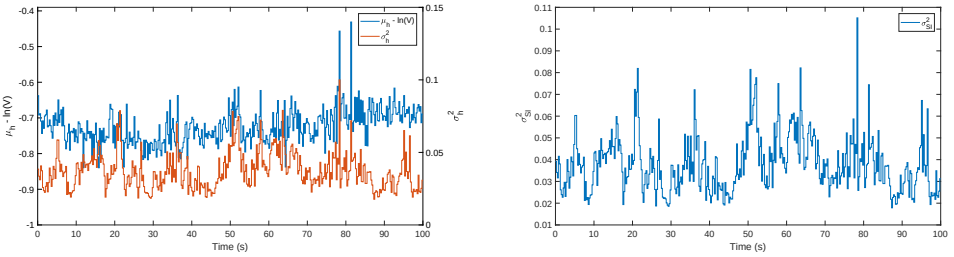
(a) (b)

Figure 3.12: Estimated channel coefficient for 0.1 s of measurement: (a) 500 mV signal, (b) 7 mV signal. Translucent dots represent received values.



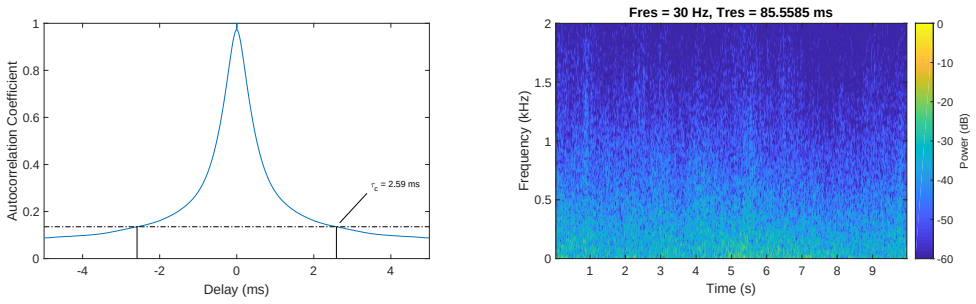
(a) (b)

Figure 3.13: Long-term behavior of (a) mean and (b) standard deviation in the 500 mV scenario.



(a) (b)

Figure 3.14: Long-term behavior of (a) log-normal parameters and (b) scintillation index σ_{SI} in the 500 mV scenario.



(a) (b)

Figure 3.15: 500 mV scenario (laser on pulses only): (a) autocorrelation coefficient and (b) spectrogram.

4

Channel Model and Capacity

The theoretically attainable capacity C of a communication channel is of fundamental importance in assessing whether a realizable/realized modulation and coding scheme is "good" in terms of loss compared to the ideal system. The fundamental work on the topic of information-theoretic limits to the amount of data that can be communicated was done by Shannon [32, 33]. The main insight was the fact that asymptotically error-free communication is possible in the presence of noise. The channel capacity C gives an upper bound for the data rate/code rate of all coding schemes with arbitrarily small error probabilities and the error rate is bounded away from zero for all schemes whose code rate exceeds C.
When binary channels are discussed, the channel capacity and the code rates are typically viewed time-free, i.e. these quantities are unitless fractions as opposed to bits per second. The rate of a code is the ratio of its dimension (the number of data bits) to its word length and the capacity is simply the tight upper bound of the code rate for which asymptotically error-free operation is achievable.

$$\begin{aligned} C_{\mathrm{BSC}} &= 1 - H_2\left(p_e\right) \\ &= 1 + p_e\, log_2\left(p_e\right) + \left(1 - p_e\right) log_2\left(1 - p_e\right) \end{aligned} \tag{4.1}$$

For the binary symmetric channel with error probability p_e, the capacity is given by Equation 4.1 [32]. It is a common benchmark for the performance of various error-correcting codes, by inserting the code's rate instead of C into Equation 4.1 and solving for the maximum p_e that allows communication at this rate. If the code's output error rate becomes sufficiently low at an error rate near the maximum p_e identified for the code rate, it is a good code. Typically, codes need a certain threshold portion of uncorrupted bits in order to converge on the correct codewords. Around this threshold input error rate, the code's output error rate drops with a very steep slope until it reaches the code's error floor, for example caused by indistinguishable error patterns.

The second widely-known channel capacity concerns the AWGN channel: For mean signal power P and additive white Gaussian noise of power N within the bandwidth B, its capacity (in bits per second) is given by Equation 4.2 [33]. Since the noise power also depends on the bandwidth, it is sometimes written in terms of energy per bit E_b absolutely necessary for reliable communication. This lower bound is called the Shannon bound and given by Equation 4.3. Here, $\frac{N_0}{2}$ is the noise power spectral density and ν is the spectral efficiency.

$$C_{\text{AWGN}} = B\,log_2\left(1 + \frac{P}{N}\right) \tag{4.2}$$

$$\frac{E_b}{N_0} \geq \frac{2^\nu}{\nu} \tag{4.3}$$

4.1 Performance Bounds for the Models

Determining the capacity of the FSO link discussed in the previous chapter is out of scope for this thesis though there is literature on optical intensity channels with additive Gaussian noise (and data-dependent noise) e.g. [15]. Nevertheless, several approximations and lower bounds can be given. Evidently, any realizable operation forms a lower bound on the theoretic performance. Known sub-optimal approaches, e.g. hard-decision decoding or OOK instead of PAM, also generate lower bounds on the capacity.
The Gaussian model for h (Definition 2.15) and hard-decision ML decoding lead to the error probabilities in Equation 2.31 to 2.33. If linear codes are used, the unequal error rates are not exploited, yielding the capacity of a BSC with the average error rate of the channel. However, the true capacity of the resulting binary *asymmetric* channel (BAC) is given by Equation 4.4 [34], where without loss of generality the two error probabilities satisfy the condition $0 \leq p_0 \leq p_1 \leq 1$. The corresponding optimal input distribution is given by Equation 4.5 and $P_1^* = 1 - P_0^*$ (where P_0^* refers to the a priori probability of the bit with error rate p_0).

$$C_{\mathrm{BAC}} = \frac{p_0 H_2(p_1) - (1-p_1) H_2(p_0)}{1-p_0-p_1} + log_2\left(1 + 2^{\frac{H_2(p_0)-H_2(p_1)}{1-p_0-p_1}}\right) \tag{4.4}$$

$$P_0^* = \frac{1}{(1-p_0-p_1)\left(1+2^{\frac{H_2(p_0)-H_2(p_1)}{1-p_0-p_1}}\right)} - \frac{p_1}{1-p_0-p_1} \tag{4.5}$$

The resulting capacity for both the average error rate and exploiting the unequal error rates is shown in Figure 4.1. For high scintillation indices, the capacity of the BAC shows a sizable gain over the average-error based capacity – but no relevant effect occurs in the low-scintillation regime. The corresponding plot for the log-normal model is shown in Figure 4.2 and displays nearly no gain of using the asymmetric error rates when designing error correcting codes.
The channel with known perfect fading information forms the idealized version of the estimation schemes discussed in Section 2.2. Though such a channel will still produce burst errors after the threshold decision, the presented analysis is limited to the average error rate. For a marginal distribution of the channel coefficient given by $f_{\mathsf{h}}(\cdot)$, the average error rate follows imme-

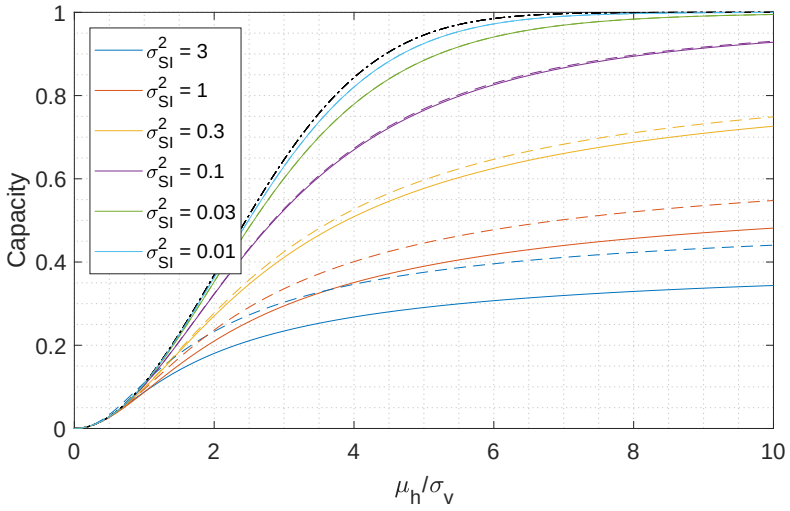


Figure 4.1: Capacity of OOK and hard-decision for Gaussian channel coefficient model vs. SNR in field quantities. Dashed lines indicate BAC performance, the black dot-dashed line is the capacity of the AWGN channel with OOK and hard-decision.

diately from Equations 2.11b and 2.10 – note that ML decision based on the known fading coefficient is always symmetric, thus no consideration of the BAC channel's gain is necessary and the expression for C simplifies to Equation 4.1 with the error probability given by Equation 4.6. The resulting performance for log-normal $f_{\mathsf{h}}(\cdot)$ is depicted in Figure 4.3 and establishes a lower bound on the capacity of a channel with temporally correlated, known log-normal fading coefficients.

$$p_e = E_{\mathsf{h}}\left\{Q\left(\frac{\mathsf{h}}{2\sigma_{\mathsf{v}}}\right)\right\} \tag{4.6a}$$

$$= \int_{R_{\mathsf{h}}} f_{\mathsf{h}}(h)\, Q\left(\frac{h}{2\sigma_{\mathsf{v}}}\right) dh \tag{4.6b}$$

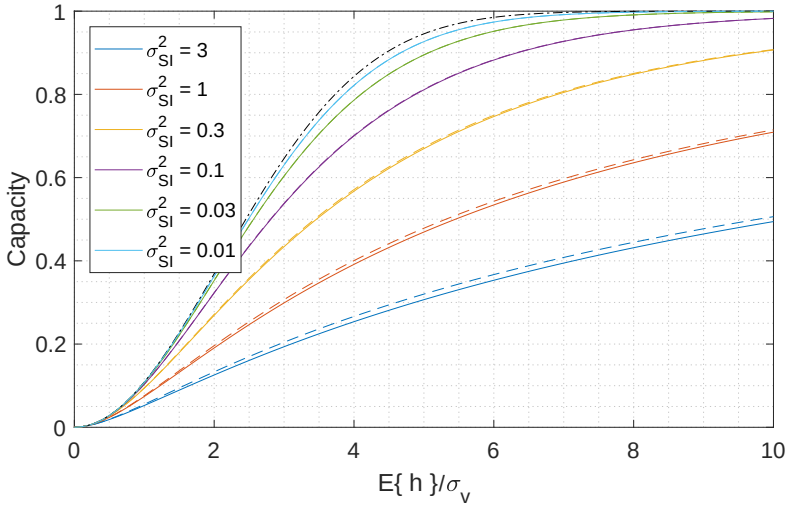


Figure 4.2: Capacity of OOK and hard-decision for log-normal channel coefficient model vs. SNR in field quantities. Dashed lines indicate BAC performance, the black dot-dashed line is the capacity of the AWGN channel with OOK and hard-decision. Only the distribution of h is known to the detector.

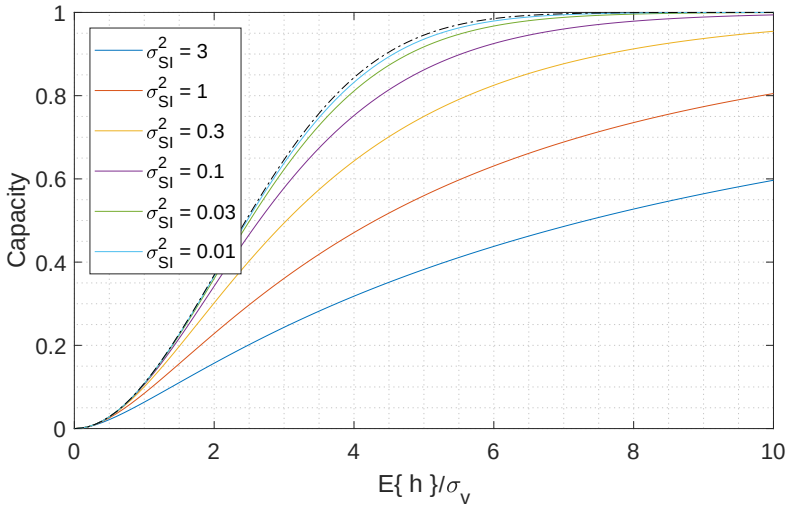


Figure 4.3: Capacity of OOK and hard-decision for known log-normal channel coefficient vs. SNR in field quantities, the black dot-dashed line is the capacity of the AWGN channel with OOK and hard-decision. Each instantaneous value of h is known to the detector.

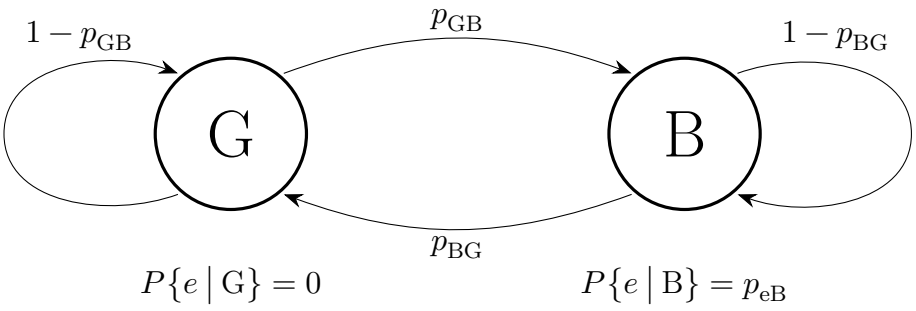


Figure 4.4: Gilbert model state transition graph.

4.2 Gilbert(-Elliott) Burst Error Channel

The Gilbert burst error channel is modeled by a binary Markov chain [7]. Its states correspond to a "good" and a "bad" channel state, with small transition probabilities $p_{\mathrm{GB}} \ll 1 - p_{\mathrm{GB}}$ and $p_{\mathrm{BG}} \ll 1 - p_{\mathrm{BG}}$. In the good state, all transmissions are without errors while the bad state exhibits a nonzero error probability p_{eB}. The Gilbert-Elliott burst error model [8] is the generalization where the good state also exhibits a nonzero error probability $p_{\mathrm{eG}} < p_{\mathrm{eB}}$. Both cases are hidden Markov models (HMMs) in the sense that the state sequence is not directly observable.
The mean lifetime of each state is the inverse of the transition probability away from this state, leading to extended periods of being in the same channel state. The a priori probabilities of being in each state are given by Equation 4.7 [7] and the average error rate for the Gilbert-Elliott model is expressed in Equation 4.8 [8]. For the original Gilbert model, this of course simplifies to $p_e p_{\mathrm{B}}$.

$$p_{\mathrm{G}} = \frac{p_{\mathrm{BG}}}{p_{\mathrm{BG}} + p_{\mathrm{GB}}} \tag{4.7a}$$

$$p_{\mathrm{B}} = \frac{p_{\mathrm{GB}}}{p_{\mathrm{BG}} + p_{\mathrm{GB}}} \tag{4.7b}$$

$$p_e = \frac{p_{\mathrm{eG}} p_{\mathrm{BG}} + p_{\mathrm{eB}} p_{\mathrm{GB}}}{p_{\mathrm{BG}} + p_{\mathrm{GB}}} \tag{4.8}$$

Table 4.1: Gilbert model fit based on trigram statistics (occurrences of error patterns spanning up to three symbols) for all data points. Results for the 500 mV scenario allow no interpretation.

Scenario	$p_{\mathbf{GB}}$	$p_{\mathbf{BG}}$	$p_{\mathbf{eB}}$	p_e
500 mV	$7 \cdot 10^{-6}$	0.2482	$-8.8 \cdot 10^{-11}$	$-3.1 \cdot 10^{-6}$
13 mV	$3.2 \cdot 10^{-4}$	0.624	0.1268	$6.6 \cdot 10^{-5}$
7 mV	0.0154	0.8899	0.7384	0.0126

Table 4.2: Gilbert model fit based on trigram statistics for each second laser-on pulse.

Scenario	$p_{\mathbf{GB}}$	$p_{\mathbf{BG}}$	$p_{\mathbf{eB}}$	p_e
500 mV	$1.35 \cdot 10^{-6}$	0.0246	0.4833	$2.66 \cdot 10^{-5}$
13 mV	$4.7 \cdot 10^{-4}$	0.442	0.5907	$6.28 \cdot 10^{-4}$
7 mV	0.0587	0.5747	0.9540	0.0884

Tables 4.1 and 4.2 show the Gilbert model parameters for the three measurement scenarios obtained from trigram statistics (occurrences of error patterns spanning up to three symbols) as described in the original paper [7]. The detection threshold is determined from the test signal according to the log-normal model. There are two versions of the Gilbert model fit: Table 4.1 includes all data while the error statistics for Table 4.2 are computed on each second measured value of laser-on pulses after calculations that included all signal points have led to nonsensical results (negative probabilities) and to mitigate the influence of the high-pass-like behavior (cf. Figure 3.6a). The Gilbert-Elliott model fits by the Baum-Welch algorithm [35] have been unsuccessful, they either led to indistinguishable states (i.e. had the same error probabilities) or had a transition probability equal to zero, leading to ending in one state almost sure. The two suspected reasons behind this unsuccessful parameter estimation lie in the overgeneralization by the Gilbert-Elliott model: For one, it is linear while the optical channel leads to different error probabilities depending on the sent bit. Second, it prescribes the existence two distinct error probabilities, which would correspond to two possible values for the fading coefficient; while in reality both the fading coefficient and the error probability vary over a continuous space, as opposed to two points.

Table 4.3: Capacities estimated [36] for the studied scenarios.
m: measured errors distributed randomly, f: for Gilbert model in Table 4.1, r: based on Table 4.2, fit: based on CCDF estimate.

Scenario	$C_{\mathbf{m}}$	$C_{\mathbf{f}}$	$C_{\mathbf{r}}$	$C_{\mathbf{fit}}$
500 mV	0.9943	–	1.000	–
13 mV	0.9995	0.992	0.998	0.9989
7 mV	0.9592	0.569	0.808	0.9868

The measured error rates for the three scenarios are $4.53{\cdot}10^{-4}$ for the 500 mV case, $2.83{\cdot}10^{-5}$ for the 13 mV and $4.4{\cdot}10^{-3}$ for the 7 mV scenario respectively. Table 4.3 compares the capacities obtained by treating these error rates as randomly distributed with the capacity obtained by the coin-toss method [36] that also allows estimation of the capacity of a Gilbert-Elliott channel.
These results are inconclusive. For the 500 mV scenario, the capacity is $2{\cdot}10^{-6}$ below 1, but all other Gilbert-based capacity estimates are worse than the measurement-based lower bound. For the 7 mV scenario, the discrepancy is especially dramatic. If, on the other hand, the method of fitting a curve to the run distribution (lengths of error-free periods) from the original paper [7] is used, the obtained capacities are plausible for the 13 mV and 7 mV scenario (for the 500 mV case, the only errors occurred during suspected blocking of the line of sight) – but this approach seems rather arbitrary in terms of how the fits can be chosen. To provide reproducible results grounded on some optimality criterion, an approach based on nonlinear least squares has been chosen (the results do not differ dramatically from the heuristic approach).
The heuristic approach consists of an exponential fit to the complementary cumulative distribution function (CCDF) of error-free periods of length K in the form of AJ^K, followed by a modification for small K as given by Equation 4.9 [7]:

$$U(K) = AJ^K + (1 - A)\, L^K \tag{4.9}$$

The parameter L is obtained by an exponential fit as well, but only to the values of the CCDF for "small" K. The boundary is chosen such that the difference of the CCDF and AJ^K-fit exceeds A for small K. This approach seems consistent and robust. The optimality-based fit minimizes the distance between Equation 4.9 and the CCDF in the log domain directly – convergence

Table 4.4: Gilbert model fit based on run distribution.

Scenario	$p_{\mathbf{GB}}$	$p_{\mathbf{BG}}$	$p_{\mathbf{eB}}$	p_e
13 mV	$8.91 \cdot 10^{-5}$	0.0042	0.9783	0.0204
7 mV	$1.53 \cdot 10^{-3}$	$5.06 \cdot 10^{-4}$	0.9948	0.748

issues are resolved by placing lower and upper bounds on the parameters of the fit ($A, L \in [0, 1]$, $J \in [0.99, 1]$). The lower bound on J is of highest importance and has to be chosen such that J^K_{min} shows a steeper decline than the CCDF but remains above zero for the given numerical accuracy. The optimality-based fit's resulting Gilbert model parameters are summarized in Table 4.4 and the CCDF plots with the corresponding fits are shown together with the heuristic fits in Figures 4.5 and 4.6.
The deviation of the fitted models from the measurement at the bottom right end is a result of the small number of long error-free runs together with the distance metric used. The semi-logarithmic plot conceals the high number of data points in the top-left section, where model and measurement are practically congruent.

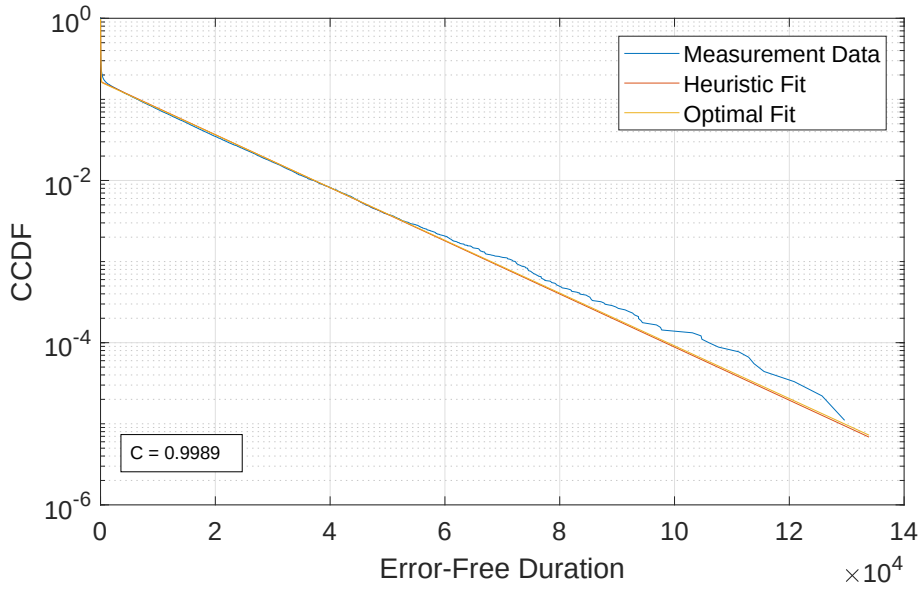


Figure 4.5: Complementary cumulative distribution function of error-free periods and Gilbert model fit for 13 mV scenario.

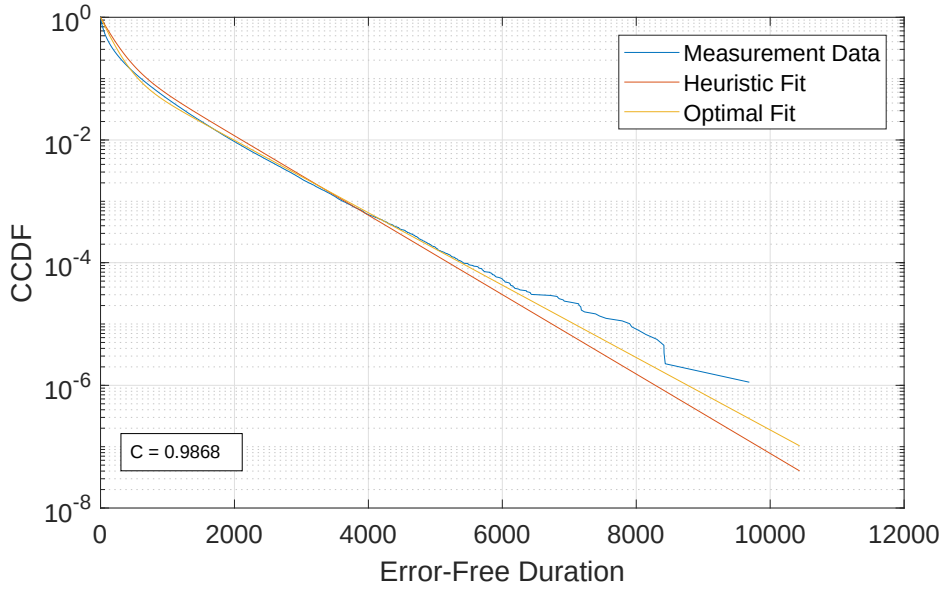


Figure 4.6: Complementary cumulative distribution function of error-free periods and Gilbert model fit for 7 mV scenario.

4.3 Discussion

The capacity of the uncorrelated models (Gaussian as well as log-normal) is obtained in a straightforward manner. For the temporally correlated channels, less strong statements can be made: The performance of any GLRT ML detector or suboptimal approximation thereof is upper-bounded by the detector with perfect knowledge of the fading coefficient. The bit errors incurred by those detectors are temporally correlated as well, hence the true capacity is yet higher than the performance shown in Figure 4.3.
The exploration of Gilbert (-Elliott) burst error channels has not yielded great advance in the quest to determine the physical channel's capacity. While trigram statistics are appealing in their elegance, the results obtained from them do not stand up to scrutiny: The capacity of a burst error channel is lower bounded by the channel's uncorrelated equivalent, hence an estimate that does not surpass the average-error-based capacity can be discarded immediately.
The CCDF-based model fits offer degrees of freedom in how the parameters of Equation 4.9 are matched to the run distribution. The proposed least-squares approaches differ slightly – while the results do not deviate by much, the heuristic approach is straightforward and robust but the optimality-based fit may need intervention depending on the channel in order to ensure convergence of the nonlinear least squares solver. The obtained Gilbert model parameters show two notable peculiarities for both scenarios: The average error rate p_e exceeds the measured values by two orders of magnitude and the error probability in the bad state p_{eB} is near 1, while the expected maximum error probability is 0.5, corresponding to total uncertainty. Furthermore, the fading coefficient only affecting "H" bits' error rate also contradicts the results in Table 4.4.
The reasons suspectedly causing this behavior are congruent with those discussed concerning the Baum-Welch algorithm: The Gilbert-Elliott model assumes symmetrical errors, independent of the data and it operates on the assumption that both the fading coefficient and the instantaneous error probability are members of binary sets as opposed to the continuous space.

5

Conclusion and Outlook

5.1 Conclusion

Atmospheric links are modeled at different levels of abstraction: From general stochastic models without specified pdfs over additive Gaussian noise motivated by the physical processes happening in receivers above a certain photon count to the log-normal and Gaussian models for the fading coefficient. This thesis analyzes and compares the performance of optimal detectors at the symbol level (for OOK). The influence of scintillation on the performance of optimal symbol detectors is quantified and compared to the AWGN case.
Models for the temporal and spatial correlation of atmospheric turbulence induced fading of FSO channels have been surveyed and optimal as well as sub-optimal detectors exploiting this correlation are compared wrt. the properties of FSO communication. Two low-complexity methods for online estimation of the ideal detection threshold are contrasted and their performance is quantified.

Research Question 1.
Is the studied 600 m link sufficiently well-characterized by the Gaussian and log-normal models for the channel coefficient, and which model yields a better fit to the measured values? – Yes, the log-normal model outperforms the Gaussian model in all but the heavily dampened scenario.

A fully functional FSO link was set up and its various parameters are extracted. At up to 39.3 dB signal-to-noise ratio, scintillation indices between 0.02 and 0.1 are observed, averaging at $\sigma_{\mathrm{SI}}^2 = 0.04$. The relative likelihood between the Gaussian and the log-normal model for the fading coefficient is evaluated and rules in favor of the log-normal model for the scenario without ND filter as well as for medium damping. The fading coefficient's autocorrelation yields a coherence time shorter than expected, but within the range of values generally agreed upon by the FSO literature.

Research Question 2.
Are optimal detectors for OOK over channels with iid or correlated Gaussian or log-normal fading coefficient practically realizable? – Optimal detectors for iid Gaussian and log-normal channel coefficients are realizable; a suboptimal detector for correlated log-normally distributed fading coefficients is presented.

Research Question 3.
Can the capacity of the studied link and the discussed models be determined? – Partially. A lower bound on the link's capacity is given by the average-error based capacity estimate. For the uncorrelated models, the capacity is calculated and a lower bound to the temporally correlated model's capacity is presented.

The capacity of the FSO channel with OOK when using the optimal symbol detectors (for the Gaussian and log-normal fading coefficient models) is quantified and compared: In the low-scintillation case $\sigma_{\mathrm{SI}}^2 = 0.03$, the capacity is no more than 0.058 below the AWGN channel's capacity (for equal average signal power, OOK and hard decision). The gain of exploiting different error rates based on the sent data bit is shown to be irrelevant in the low-scintillation case and for the log-normal model in general. The capacity of the channel with perfect knowledge of the instantaneous fading state shows dramatic improvement in the high-scintillation regime: At $\mathrm{SNR}_{\mathrm{dB}} = 20$, the capacity increases from 0.5 to 0.6 and from 0.7 to 0.8 for the $\sigma_{\mathrm{SI}}^2 = 3$ and $\sigma_{\mathrm{SI}}^2 = 1$ case respectively.
Another layer of abstraction is added by fitting Gilbert burst error model parameters against the measurement results. The capacity of the FSO channel has been bounded from below by treating the measured error rates as uniformly distributed, leading to a capacity achievable without exploiting the correlated nature of the burst errors. The capacity obtained from the

Gilbert model fits is of mixed satisfaction – the trigram-based values are worse than the lower bound of random errors and the Baum-Welch algorithm for the Gilbert-Elliott model converges to degenerate models where at least one transition probability is zero. A systematic approach for the parameter estimation based on the run distribution is presented and leads to reproducible results that do not fully agree with the measurements they are based on. This work thus presents evidence and reasoning against using the Gilbert (-Elliott) burst error model on the symbol level for FSO channels with OOK. Packet-/block-level formulations of this class of channel models are likely better suited to help model optical communications on a reduced-complexity basis, given that the packet/block duration is lower than the channel coherence time.

To summarize: The studied link exposed to 600 m of urban environment is best characterized by the log-normal model, but the Gaussian model leads to almost identical detection threshold and hence error statistics. The capacity for OOK with hard decision is near to 1, even when the signal is damped by 19 dB. Dampening the signal by 24 dB still yields a capacity above 0.95. Optimal symbol detectors are easily found and implemented, but optimal sequence detection is not feasible due to excessive computational load. A simple estimator for the optimal decision threshold is proposed, even with the possibility of being implemented in hardware in order to save resources. The design does not need exact knowledge of the fading coefficient's temporal correlation and is robust against variation of the cutoff frequency. Most importantly, it shows good approximation of the performance of the detector with perfect fading state knowledge, which offers a sizable gain in capacity compared to the static threshold of the ML symbol detector.

5.2 Outlook

Future research on the topic of FSO communications should consider using and studying estimators of the optimal decision threshold in order to approximate the channel with perfect knowledge of the fading state and improve channel capacity.
Further work is needed on the nature of temporally correlated errors that arise from threshold detection of channels with real-valued fading coefficients as opposed to the two states of the Gilbert (-Elliott) model. Most importantly, the capacity of such channels needs to be determined.
More generally, the capacity of real-valued, temporally correlated channels needs to be quantified in order to provide a tight upper bound for the achievable data rates over optical links.
Finally, more research concerning forward error correction codes, especially variable-rate codes, is necessary to find not only possible, but also realizable codes that approach the capacity of the channel. Good candidates include product-like codes like staircase codes [37], e.g. in a rate-adaptive configuration as proposed in [38].

Bibliography

[1] R. Lange, F. Heine, M. Motzigemba, M. Lutzer, and R. Meyer, "Roadmap to wide band optical GEO relay networks," in *Proceedings - IEEE Military Communications Conference MILCOM*, 2012.

[2] G. A. Koepf, R. G. Marshalek, and D. L. Begley, "Space laser communications: A review of major programs in the United States," *AEU-Archiv fur Elektronik und Ubertragungstechnik*, 2002.

[3] T. Tolker-Nielsen and G. Oppenhauser, "In-orbit test result of an operational optical intersatellite link between ARTEMIS and SPOT4, SILEX," in *Free-Space Laser Communication Technologies XIV*, 2002.

[4] Z. Sodnik, H. Lutz, B. Furch, and R. Meyer, "Optical satellite communications in Europe," in *Free-Space Laser Communication Technologies XXII*, vol. 7587, p. 758705, SPIE, 2010.

[5] M. Reyes Garcia-Talavera, Z. Sodnik, P. Lopez, A. Alonso, T. Viera, and G. Oppenhauser, "Preliminary results of the in-orbit test of ARTEMIS with the Optical Ground Station," in *Free-Space Laser Communication Technologies XIV*, vol. 4635, pp. 38–49, SPIE, 2002.

[6] A. Alonso, M. Reyes, and Z. Sodnik, "Performance of satellite-to-ground communications link between ARTEMIS and the Optical Ground Station," in *Optics in Atmospheric Propagation and Adaptive Systems VII*, 2004.

[7] E. N. Gilbert, "Capacity of a Burst-Noise Channel," *Bell System Technical Journal*, vol. 39, no. 5, pp. 1253–1265, 1960.

[8] E. O. Elliott, "Estimates of Error Rates for Codes on Burst-Noise Channels," *Bell System Technical Journal*, vol. 42, no. 5, pp. 1977–1997, 1963.

[9] H. Kaushal and G. Kaddoum, "Optical communication in space: Challenges and mitigation techniques," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 57–96, 2017.

[10] M. A. Khalighi and M. Uysal, "Survey on free space optical communication: A communication theory perspective," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 2231–2258, 2014.

[11] J. Ma, Y. Jiang, S. Yu, L. Tan, and W. Du, "Packet error rate analysis of OOK, DPIM and PPM modulation schemes for ground-to-satellite optical communications," *Optics Communications*, vol. 283, no. 2, pp. 237–242, 2010.

[12] V. W. Chan, "Free-space optical communications," *Journal of Lightwave Technology*, vol. 24, no. 12, pp. 4750–4762, 2006.

[13] A. Viswanath, H. Kaushal, V. K. Jain, and S. Kar, "Evaluation of performance of ground to satellite free space optical link under turbulence conditions for different intensity modulation schemes," in *Free-Space Laser Communication and Atmospheric Propagation XXVI*, 2014.

[14] R. H. Hadfield, "Single-photon detectors for optical quantum information applications," *Nature Photonics*, vol. 3, no. 12, pp. 696–705, 2009.

[15] A. Lapidoth, S. M. Moser, and M. Wigger, "Channel capacity," in *Advanced Optical Wireless Communication Systems*, vol. 9780521197878, pp. 146–174, Cambridge University Press, 2012.

[16] H. Henniger and O. Wilfert, "An introduction to free-space optical communications," *Radioengineering*, vol. 19, 2010.

[17] B. Epple, "Simplified channel model for simulation of free-space optical communications," *J. Opt. Commun. Netw.*, vol. 2, no. 5, pp. 293–304, 2010.

[18] Xiaoming Zhu and J. M. Kahn, "Markov chain model in maximum-likelihood sequence detection for free-space optical communication through atmospheric turbulence channels," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 509–516, 2003.

[19] T. Blazek and C. F. Mecklenbrauker, "Measurement-Based Burst-Error Performance Modeling for Cooperative Intelligent Transport Systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 1, pp. 162–171, 2019.

[20] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I.* John Wiley & Sons, Inc., 2001.

[21] M. L. B. Riediger, R. Schober, and L. Lampe, "Blind detection of on-off keying for free-space optical communications," in *Canadian Conference on Electrical and Computer Engineering*, pp. 1361–1364, 2008.

[22] A. Prokeš, "Modeling of atmospheric turbulence effect on terrestrial FSO link," *Radioengineering*, vol. 18, pp. 42–47, apr 2009.

[23] X. Zhu and J. M. Kahn, "Free-space optical communication through atmospheric turbulence channels," *IEEE Transactions on Communications*, vol. 50, no. 8, pp. 1293–1300, 2002.

[24] M. A. Khalighi, N. Aitamer, N. Schwartz, and S. Bourennane, "Turbulence mitigation by aperture averaging in wireless optical systems," in *2009 10th International Conference on Telecommunications, ConTEL 2009*, pp. 59–66, 2009.

[25] Z. Chen, "Bayesian filtering: From kalman filters to particle filters, and beyond," *Statistics*, vol. 182, 2003.

[26] F. Hlawatsch and O. Hlinka, "Sequential Bayesian Estimation and Particle Filtering," in *Parameter Estimation Methods*, pp. 67–75, Vienna: TU Wien, 2016.

[27] A. H. Gray and J. D. Markel, "A Spectral-Flatness Measure for Studying the Autocorrelation Method of Linear Prediction of Speech Analysis," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1974.

[28] A. Sinn, T. Riel, P. Kremsner, and G. Schitter, "Analysis of tip-tilt compensation for reflective free-space optical satellite communication," in *Proceedings Volume 10910, Free-Space Laser Communications XXXI*, p. 52, SPIE-Intl Soc Optical Eng, 2019.

[29] F. J. Macwilliams and N. J. Sloane, "Pseudo Random Sequences and Arrays," *Proceedings of the IEEE*, vol. 64, no. 12, pp. 1715–1729, 1976.

[30] K. P. Burnham and D. R. Anderson, "Multimodel inference: Understanding aic and bic in model selection," *Sociological Methods & Research*, vol. 33, no. 2, pp. 261–304, 2004.

[31] H. Akaike, "Information Theory and an Extension of the Maximum Likelihood Principle," in *Kotz S., Johnson N.L. (eds) Breakthroughs in Statistics*, pp. 199–213, New York: Springer, 1992.

[32] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[33] C. E. Shannon, "Communication in the Presence of Noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[34] S. M. Moser, P.-n. Chen, and H.-y. Lin, "Error Probability Analysis of Binary Asymmetric Channels," tech. rep., final report of project "Finite Blocklength Capacity" (proj.-no.: NSC 97–2221–E–009–003–MY3), 2012.

[35] L. R. Rabiner and B. H. Juang, "An Introduction to Hidden Markov Models," *IEEE ASSP Magazine*, vol. 3, no. 1, pp. 4–16, 1986.

[36] M. Rezaeian, "Computation of capacity for Gilbert-Elliott channels, using a statistical method," in *Proceedings 6th Australian Communications Theory Workshop 2005*, pp. 56–61, 2005.

[37] B. Smith, A. Farhood, A. Hunt, F. Kschischang, and J. Lodge, "Staircase codes: FEC for 100 Gb/s OTN," *Lightwave Technology, Journal of*, vol. 30, pp. 110 – 117, 02 2012.

[38] P. Kukieattikool and N. Goertz, "Variable-rate staircase codes with RS component codes for optical wireless transmission," *Transactions on Emerging Telecommunications Technologies*, 2017.

Erklärung

Hiermit erkläre ich, dass die vorliegende Arbeit gemäß dem Code of Conduct, insbesondere ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel, angefertigt wurde. Die aus anderen Quellen oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.
Die Arbeit wurde bisher weder im In– noch im Ausland in gleicher oder in ähnlicher Form in anderen Prüfungsverfahren vorgelegt.

Wien, im November 2019

Johannes Steinbach