

Improved Markov Models for Terrestrial Free-Space Optical Links

Zifeng Wu and Berthold Lankl

Institute for Communications Engineering, University of the Federal Armed Forces Munich, 85577 Neubiberg, Germany

zifeng.wu@unibw.de, berthold.lankl@unibw.de

Dirk Giggenbach

Institute of Communications and Navigation, German Aerospace Center, 82234 Oberpfaffenhofen-Wessling, Germany

dirk.giggenbach@dlr.de

Abstract

Finite-state Markov chains are a useful tool for modelling communication channels with correlated fading and have recently also been applied with success to terrestrial free-space optical (FSO) communication channels. However, the issue of how such Markov models should be optimized in order to accurately approximate the original continuous fading channel has not been addressed in a systematic manner. In this paper, we investigate the properties and approximation accuracy of Markov models which are optimized according to information-theoretic considerations. We validate and evaluate our approach using a set of experimental measurements over a 12km link distance. The obtained results confirm that optimized Markov models can provide better accuracy at lower state complexity, yet there remain shortcomings in capturing the autocovariance of the fading process.

1. Introduction

Free-space optical (FSO) communication systems transmit information through the air by means of modulated laser beams. A single-wavelength point-to-point FSO link can be considered as a non-frequency-selective ("flat") fading channel where the total optical power incident on the detector at the receiver side fluctuates over time. The mechanism behind this fading is atmospheric turbulence which leads to the formation of inhomogeneous air pockets ("eddies") along the propagation path [1]. In their entirety, these eddies act like a randomly curved lens which produces far-field intensity speckles at the receiver. The fading process in FSO links has comparatively long coherence times of up to several milliseconds which are correlated with the typical wind speeds in transverse direction of the link [2]. It is commonly assumed in this context that the wind sweeps a static speckle pattern over the receiver aperture without substantially changing its structure (Taylor's "frozen" turbulence hypothesis.) Therefore, any image on the receiver plane requires on the order of D/v_{\perp} seconds to traverse an aperture of diameter D if v_{\perp} is the transversal wind speed.

The aforementioned coherence times can seriously degrade transmission at Gbit/s rates typical for FSO communications as the channel may remain in a deep fade for the duration of several million bits. To maintain high quality-of-service throughout varying channel conditions, measures such as link adaptation and cross-layer performance optimization are widely employed in radio-frequency wireless communications, and they are increasingly being investigated also for FSO communications. In order to implement effective link adaptation strategies and analyse their performance, a concise statistical model of the underlying physical channel is highly useful.

While a great deal of work on FSO channel models has focused on the probability density function (PDF) of the fading process – an introductory survey can be found e.g. in [3] –

comparatively fewer works have addressed its temporal correlation in a satisfactory manner. However, accurate modelling of the latter could be quite useful e.g. for predicting how the channel state will evolve given knowledge of the current state. A common approach for generating random samples according to a given PDF and autocovariance function (ACF) operates in the spectral domain by use of the power spectral density. This approach is presented and applied to a number of FSO-relevant scenarios e.g. in [4] and suggests a systems-theoretic model which describes the fading process as being obtained through linear filtering and subsequent nonlinear transformation of a Gaussian process. While accurate and generally applicable, such a model can be too unwieldy for real-time applications. On the other hand, finite-state Markov chains have long been recognized as a useful tool for modelling communication channels with correlated fading [5] and can offer a greater degree of conciseness. For example, in [6] a 2-state continuous-time Markov model was applied to experimental power measurements obtained from a 250m FSO link, with the results displaying a good fit to the observed behaviour. However, since a 2-state model can only model the extreme states of a communication system, i.e. working or failing ("outage"), it is not suitable for evaluating more sophisticated adaptation strategies such as adaptive modulation and coding. In [7], discrete-time Markov models with up to 64 states were derived from power measurements in a 1.87km terrestrial FSO link. Therein, it was concluded that the channel state realizations generated by the respective models can accurately reproduce both the PDF and ACF of the measured fading process, given that a sufficient number of states is employed. However, the supplementary material provided for download by the authors suggests that increasing the number of states does not always result in a more accurate ACF [8]. Furthermore, the authors did not suggest a systematic procedure for choosing the number of states or attempt an optimization of the state partitioning for a given number of states.

The main contribution of this paper compared to the prior work is to systematically investigate the number of states required in a Markov model to achieve a satisfactory approximation to the original continuous fading channel. We demonstrate that by suitable partitioning of the state space, the approximation accuracy can be increased for the same number of states. To quantify the goodness-of-fit of our models and optimize the state partitioning, we evaluate in addition to the PDF and ACF the information-theoretic capacity of the obtained Markov channels. Markov models have previously been applied to the study of capacity in FSO channels [9]; however, our work differs in that it is concerned with model estimation and optimization on the basis of measurements rather than with obtaining information-theoretic limits. In particular, we derive Markov models from power measurements in a 12km terrestrial FSO link which represents a longer distance than was considered in prior work. Another contribution of this paper is to improve the ACF of the previously considered models by means of a simple modification. In particular, we demonstrate that the ACF of the fading process can be reproduced more accurately by making use of multi-step transition matrices instead of the commonly employed one-step transition matrix.

2. Theoretical Framework

2.1 Capacity of continuous fading channels

In information theory, the capacity of a channel is the highest data rate for which reliable communication (i.e. with an arbitrarily low probability of error) is possible. A well-known result due to Shannon is the capacity C of an additive white Gaussian noise (AWGN) channel for constrained average transmit power [10]:

$$C = B \cdot \log_2(1 + \gamma), \quad [\text{bits/s}] \quad (1)$$

where $\gamma = P_r/\sigma^2$ denotes a constant signal-to-noise ratio (SNR) for average received power P_r and AWGN variance σ^2 , and B is the transmission bandwidth. Note that this result assumes that the distribution of the input signal to the channel is Gaussian.

A fading channel can be considered as a time-varying AWGN channel where γ is an ergodic stochastic process with PDF $f(\gamma)$. If only the receiver has knowledge about the instantaneous γ (referred to as channel side information, CSI), capacity is reduced in the presence of fading when compared to an AWGN channel with constant SNR equal to $E[\gamma]$ [11], where

$$E[\gamma] = \int_{-\infty}^{+\infty} \gamma \cdot f(\gamma) d\gamma. \quad (2)$$

However, if both the transmitter and receiver have perfect CSI, the capacity of the fading channel can potentially exceed that of the reference AWGN channel. The optimal strategy in the latter case is to transmit with a time-varying data rate according to the instantaneous SNR [11]. In particular, if γ_0 is a "cutoff" value satisfying

$$\int_{\gamma_0}^{\infty} \left(\frac{1}{\gamma_0} - \frac{1}{\gamma} \right) f(\gamma) d\gamma = 1, \quad (3)$$

then one can achieve a capacity of

$$C_{CSI} = B \cdot \int_{\gamma_0}^{\infty} \log_2 \left(\frac{\gamma}{\gamma_0} \right) f(\gamma) d\gamma \quad (4)$$

by transmitting with rate $B \cdot \log_2(\gamma/\gamma_0)$ when $\gamma \geq \gamma_0$, and rate zero when $\gamma < \gamma_0$. Due to the assumed ergodicity of the process, the integrals in (3) and (4) can also be replaced with a time average over all γ for which $\gamma \geq \gamma_0$, i.e.

$$\int_{\gamma_0}^{\infty} g(\gamma) f(\gamma) d\gamma = E_{\gamma \geq \gamma_0} [g(\gamma)], \quad (5)$$

where $g(\gamma)$ denotes a general function. This property is useful for evaluating capacity when the analytical PDF $f(\gamma)$ is not exactly known.

We note that the assumption of perfect transmitter and receiver CSI is not entirely impractical as under certain circumstances, point-to-point FSO links can exhibit a high degree of reciprocity [12]. Therefore, even in the absence of a feedback channel, a transceiver in a bi-directional link can infer the channel quality at the other end of the link from its own received power.

2.2 Finite-state Markov channels

The time-variation of the channel in a finite-state Markov model is described using a discrete-time Markov chain [13]. Different channel qualities, e.g. in terms of SNR, are represented using a finite set of states $\mathcal{S} = \{s_0, s_1, \dots, s_{K-1}\}$. Transitions between the states are governed by a stationary Markov process S_n , where n is the time index. The Markov property implies that the transition probabilities p_{jk} only depend on the current state and not on any preceding states, while stationarity implies that they are independent of n , i.e.

$$p_{jk} = \Pr(S_{n+1} = s_k | S_n = s_j) \quad (6)$$

for all n and $j, k \in \{0, 1, 2, \dots, K-1\}$. In practice, given a sufficiently large number of channel state measurements, the transition probabilities can be estimated as [7]

$$\hat{p}_{jk} = \frac{N_{jk}}{N_j}, \quad (7)$$

where N_{jk} is the number of observed transitions from state s_j to state s_k , and N_j is the total number of times that state s_j was observed. The set of transition probabilities can be collected

into a $K \times K$ transition matrix \mathbf{P} , where a valid transition matrix has the property that each row sums to 1.

\mathbf{P} is also referred to as the one-step transition matrix as it describes how the current state evolves after a single time step. Given that S_n is stationary, the m -step transition matrix describing how the state evolves after m time steps is simply obtained as $\mathbf{P} \cdot \mathbf{P} \cdot \dots = \mathbf{P}^m$.

Now assume there exists a vector $\boldsymbol{\pi}$ for which

$$\boldsymbol{\pi}^T \mathbf{P} = \boldsymbol{\pi}^T, \quad (8)$$

i.e. $\boldsymbol{\pi}$ is an eigenvector of \mathbf{P} with eigenvalue 1. Then each component π_k of $\boldsymbol{\pi}$ represents the probability of being in state s_k as $n \rightarrow \infty$ (steady state), i.e.

$$\pi_k = \Pr(S_n = s_k), \quad k \in \{0, 1, 2, \dots, K-1\} \quad (9)$$

where the π_k must also sum to 1.

If perfect transmitter and receiver CSI are assumed, the capacity C_M of a Markov channel is simply the average capacity over all states [13], i.e.

$$C_M = \sum_{k=0}^{K-1} c_k \cdot \pi_k = \boldsymbol{\pi}^T \mathbf{c}, \quad (10)$$

where c_k denotes the capacity of state s_k .

3. Approaches for state partitioning

Assume that a Markov model is to be parameterized based on measured samples of the receiver SNR γ . Then it is first necessary to decide how to suitably choose the set of states \mathcal{S} . The most straightforward choice is to partition the range of observed SNRs $\gamma_{\min} \leq \gamma \leq \gamma_{\max}$ into K uniform intervals such that each state s_k corresponds to an observed SNR of

$$\gamma_{\min} + k\Delta \leq \gamma \leq \gamma_{\min} + (k + 1)\Delta, \quad (11)$$

where $\Delta = (\gamma_{\max} - \gamma_{\min})/K$ is the step size [7]. However, this choice may not be optimal depending on the desired properties of the model.

Arguably, the most basic criterion that should be fulfilled is that the model accurately reproduces the PDF and ACF of the measured γ . In addition, it could be desirable e.g. to optimize the sparseness of \mathbf{P} in order to achieve the most compact representation of the fading process. While this is an interesting topic in itself, this shall not be our focus here. Due to the proven usefulness of channel capacity as a quality measure, we want to consider how the states of a Markov model should be optimized if the capacity of the original continuous channel is to be preserved as far as possible. To this end, it is useful to imagine the scenario that the transmitter adapts its rate as outlined in Section 2.1, but that only a finite set of rates can be employed, which is a justified assumption in practice.

Evidently, the transmitter should adapt its rate according to the instantaneous state S_n of the Markov channel. We further assume that the receiver noise is AWGN and that a capacity-achieving modulation and coding scheme is employed for each state. Now if s_k corresponds to some SNR range $\gamma_k \leq \gamma < \gamma_{k+1}$, then the capacity c_k of s_k computed using (1) must necessarily be given by

$$c_k = B \cdot \log_2(1 + \gamma_k), \quad (12)$$

i.e. the highest rate which can be reliably transmitted is limited by the lowest SNR value of that state. To approach the capacity of the original channel as calculated using (3) and (4), the set of thresholds $\Theta = \{\gamma_0, \gamma_1, \dots, \gamma_{K-1}, \gamma_K\}$ with $\gamma_0 = \gamma_{\min}$ and $\gamma_K = \gamma_{\max}$ must be optimized as to maximize the capacity of the Markov channel given by (10).

We briefly describe an iterative manner in which this optimization can be performed: start with a single state s_0 corresponding to $\gamma_{\min} \leq \gamma \leq \gamma_{\max}$. Now split this state into two new states by adding an additional threshold γ' in the interval $[\gamma_{\min}; \gamma_{\max}]$ such that the overall capacity is maximized for all possible choices of γ' . Continue this procedure by further splitting the newly obtained states such that in each step, the number of states is doubled. For example, the second iteration would consist of adding two additional thresholds in the intervals $[\gamma_{\min}; \gamma')$ and $(\gamma'; \gamma_{\max}]$, thus increasing the number of states from two to four.

4. Validation and evaluation

In order to validate and evaluate the proposed state partitioning approach, we parameterized a number of finite-state Markov channels based on a set of terrestrial FSO measurements performed by the German Aerospace Center (DLR) in cooperation with the Institute of Telecommunications Research at the University of South Australia in 2011 [12]. In each case, the received signal power detected by a photodiode was recorded with a sampling rate of 1 ksamples/sec, with every trace having a useable duration of 80 seconds. To emulate different receiver sensitivities, we scaled the power samples by different values of σ^2 leading to different average SNRs $E[\gamma]$. Transition matrices for the Markov models were calculated directly from the scaled traces and optimization of the state partition was performed as described in Section 3. The maximum number of states hereby utilized was $K = 64$. For comparison purposes, we also performed the same procedure using a uniform state partition as given by (11). As very similar results were obtained for all considered measurements, we will only present the results for one specific scenario in the following. The selected measurement corresponds to scenario C from [12]. In particular, the link distance was 12 km with the receiver employing a 127 mm diameter aperture.

4.1 Channel capacity

Figures 1 and 2 show the obtained Markov channel capacities when a uniform and an optimized state partition are applied, respectively. Here, the normalized capacity C/B is plotted over $E[\gamma]$ for different values of K . For reference, the capacity of an AWGN channel with constant SNR $E[\gamma]$ and the capacity of the continuous fading channel computed using (3)-(5) are also shown. The "optimized" curve for $K = 64$ is not plotted as it offers only negligible improvement compared to $K = 32$. It can be seen that the optimized state partition can require substantially fewer states to approximate the capacity of the continuous channel than the uniform partition, the advantage being especially pronounced for $K \leq 16$. In particular, the optimized partition for $K = 2$ performs nearly as well as the uniform partition for $K = 8$, while the optimized partition for $K = 8$ performs better than the uniform partition for $K = 16$. For $K \geq 16$, the difference between both partitioning approaches begins to vanish. However, we note that in this region, the optimized state partition requires only half the number of states to achieve similar capacity as the uniform partition, which is a significant reduction in terms of transition matrix size.

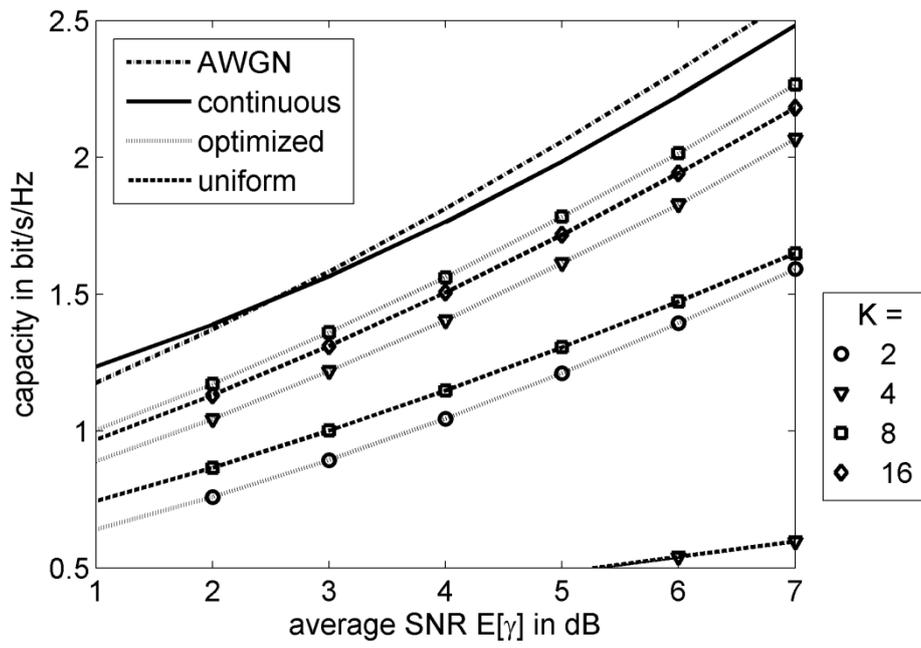


Fig. 1. Markov channel capacities for uniform and optimized state partitions ($K = 2$ to 16)

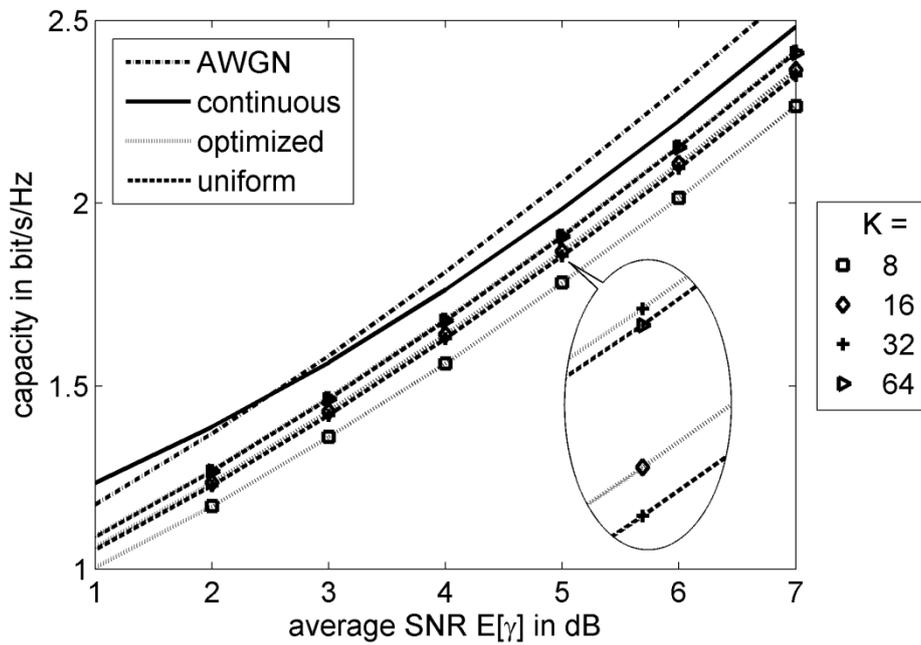


Fig. 2. Markov channel capacities for uniform and optimized state partitions ($K = 8$ to 64)

4.2 Transition matrix

Figure 3(a) illustrates an example result for a transition matrix obtained using an optimized $K = 32$ state partition. In this representation, darker spots correspond to higher probability values. It can be seen that the entries of the transition matrix are concentrated around the diagonal which signifies that the probability of transitioning to adjacent states dominates. This implies that the channel quality changes rather gradually and smoothly over time. For comparison, Fig. 3(b) shows the transition matrix when applying a uniform state partition to the same measurement. It can be seen here that the entries tend to disperse towards the bottom-right of the matrix, which suggests that the step size in this region is chosen too finely and thus introduces noise. However, contrary to what the visual impression may imply, the "uniform" matrix is actually somewhat sparser than the "optimized" matrix, with the latter having about 32% more non-zero entries.

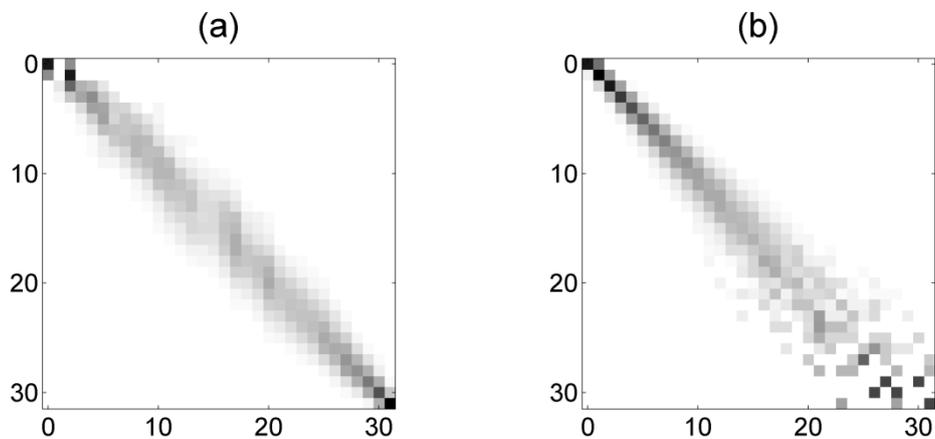


Fig. 3. Transition matrices for optimized (a) and uniform (b) state partition

4.3 Probability density function

In order to simulate the original continuous fading channel, the obtained Markov models can be used to produce random samples that are approximately distributed according to the PDF

of the original process. In the simplest case, this involves generating a sequence of states according to the estimated one-step transition matrix. For each state $S_n = s_k$, it is necessary to output a value γ from the continuous interval $\gamma_k \leq \gamma < \gamma_{k+1}$. One possibility is to output the value corresponding to the lower limit γ_k . This simulates a worst-case scenario and the PDF of the generated samples will be discrete-valued.

A simple approach for generating continuous-valued samples is to randomly output a value γ according to a uniform distribution over the interval $\gamma_k \leq \gamma < \gamma_{k+1}$ when $S_n = s_k$. Figures 4 and 5 demonstrate for a $K = 32$ model that the hereby obtained PDF can approximate the empirical PDF quite well. When comparing the PDF obtained from an "optimized" model (Fig. 5) to that obtained from a "uniform" model (Fig. 4), it can be seen that the optimized partition is more faithful to the empirical PDF for low γ values, while for higher γ values, the "steps" introduced by the approximation can be clearly seen. The uniform partition on the other hand offers a visually equal fit over the entire observed range. It can be argued however that the goodness-of-fit in the high SNR region is less critical than in the low SNR region since the probability of error in a communication system drops off sharply towards high SNR.

To evaluate the goodness-of-fit in a more quantitative manner, we computed the continuous capacity as given by (3)-(5) using both the "optimized" and "uniform" PDFs. The obtained capacity curves were practically indiscernible from the one calculated using the original measurement which implies that the goodness-of-fit is adequate. We note that a satisfactory fit was also obtained for $K = 16$ but that our simulations suggest that K should not be chosen smaller than 16 to avoid any significant deviation from the actual capacity. Compared to the results shown in Section 4.1, we can conclude that the benefit of using the optimized state partition is reduced when the "reconstructed" continuous PDF is considered. We note however that this is not the case for the discrete PDF which yields a capacity identical to the respective Markov channel.

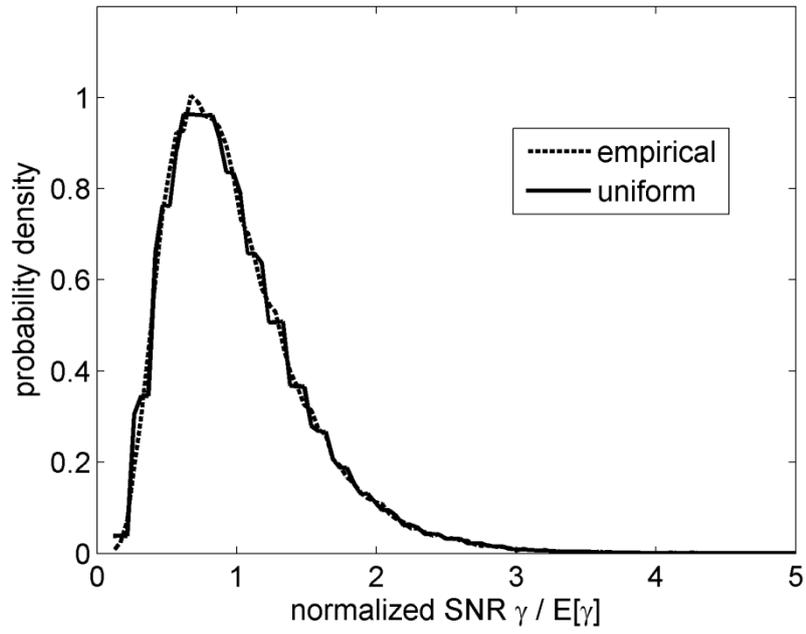


Fig. 4. Probability density function for $K = 32$ uniform state partition

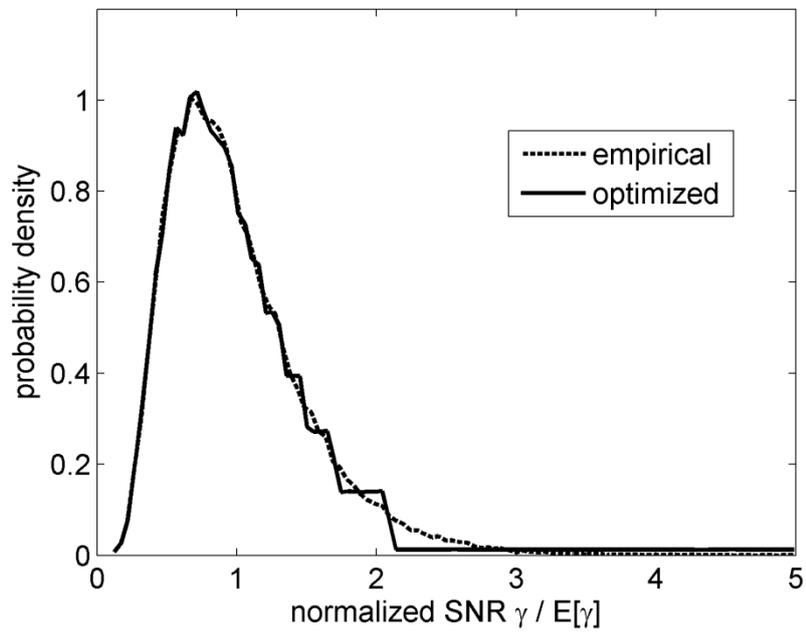


Fig. 5. Probability density function for $K = 32$ optimized state partition

4.4 Autocovariance function

Finally, we want to address the issue of accurately reproducing the autocovariance function of the original process. We observed in our work that the Markov models are not particularly accurate in capturing the autocovariance of our measurements, as is shown exemplarily in Fig. 6. Here we have plotted the normalized ACFs for samples generated using an “optimized” model with $K = \{16, 32, 64\}$. In each case, we used the approach described in Section 4.3 to generate 10^7 continuous-valued samples. The common observed behaviour is that the ACF of the generated samples is much wider than that of the measurement, the width increasing as K increases. This behaviour is also seen with the “uniform” models and implies that the correlation between neighbouring samples is too large. We note that the Markov models provided on [8] display much the same behaviour, therefore this seems to be a general issue rather than an effect specific to the selected measurements.

A straightforward method to reduce the excess correlation is to pick only every m -th generated sample and discard the remaining ones. This can be achieved more efficiently by generating a sequence of states using the m -step transition matrix \mathbf{P}^m of the model rather than computing samples that would be discarded anyway. As can be seen exemplarily in Fig. 7 for the “optimized” case, this modification can significantly improve the accuracy of the ACF though the fit is still not ideal. We note in particular that the ACF of the measurement displays a more rounded “dome” shape near zero lag which the Markov models are not capable of reproducing. The behaviour for the “uniform” case is very similar and is thus not plotted.

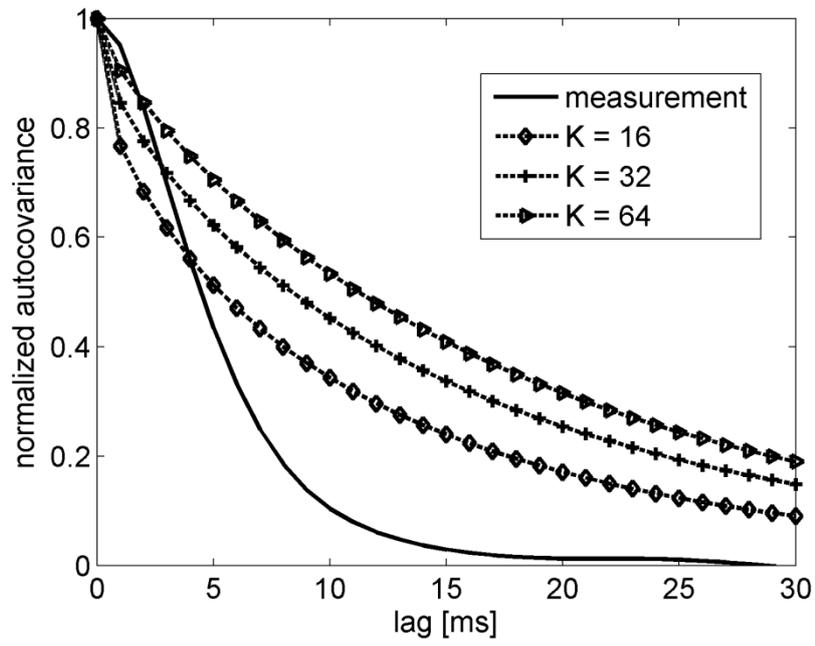


Fig. 6. Normalized autocovariance for one-step transition matrix

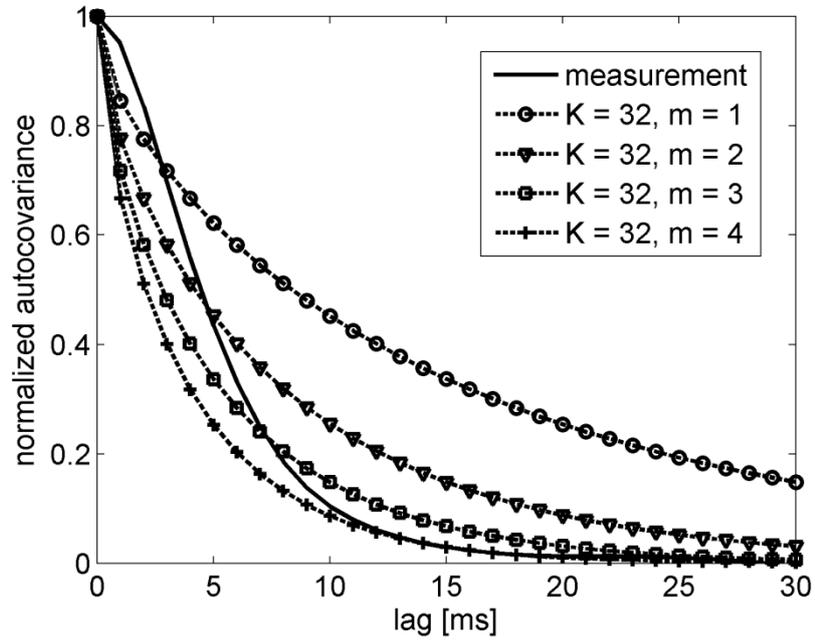


Fig. 7. Normalized autocovariance for multi-step transition matrix

5. Conclusion

We have demonstrated in this contribution that Markov models employing an optimized state partition as compared to a uniform state partition can perform better in terms of approximating the capacity of the original continuous fading channel in a terrestrial FSO link. This benefit is especially pronounced if the employed number of states is less than 16. For an increasing number of states, the accuracy of the two partitioning approaches becomes more similar.

On the other hand, we have observed that if we generate continuous-valued samples using the Markov models according to an interval-wise uniform approximation, both state partitioning approaches approximate the original channel capacity equally well. Furthermore, the number of states in this case should not be chosen smaller than 16 to avoid significant deviation from the actual capacity.

Finally, we have also addressed the issue of accurately reproducing the autocovariance function of the fading process and found that by using the multi-step transition matrix of the model, the performance can be improved as compared to the one-step transition case. However, it must be stated that the Markov models considered here are not ideal in terms of their autocovariance behaviour as they are incapable of reproducing the true shape of the measured autocovariance but can only approximate the rate of decay. Whether a different modification may improve upon this problem remains the subject of further study.

Acknowledgment

The authors would like to thank the anonymous reviewers for their helpful suggestions. This work was supported by the Initiative and Networking Fund of the Helmholtz Association via Munich Aerospace e.V. within the "Helmholtz-Alliance DLR@Uni" program.

References

- [1] Andrews, L.C., and Phillips R.L.: "Laser beam propagation through random media" (SPIE Press, Bellingham, 2nd edn. 2005), pp. 57-76
- [2] Dravins, D., Lindegren, L., Mezey, E., Young, A.T.: „Atmospheric intensity scintillation of stars. I. statistical distributions and temporal properties", Pub. of the Astron. Soc. of the Pacific, 1997, 109, pp. 173-207
- [3] Jurado-Navas, A., Garrido-Balsells, J.M., Paris, J.F., Puerto-Notario, A.: "A unifying statistical model for atmospheric optical scintillation", in Awrejcewicz, J. (Ed.): "Numerical simulations of physical and engineering processes" (InTech, 2011), pp. 181-206
- [4] Yura, H.T., Hanson, S.G.: "Digital simulation of an arbitrary stationary stochastic process by spectral representation", JOSA A, 2011, 28, (4), pp. 675-685
- [5] Sadeghi, P., Kennedy, R.A., Rapajic, P.B., Shams, R.: "Finite-state Markov modeling of fading channels – a survey of principles and applications", IEEE Sig. Proc. Mag., 2008, 25, (5), pp. 57-80
- [6] Puryear, A., Jin, R., Lee E., Chan, V.W.S.: "Experimental analysis of the time dynamics of coherent communication through turbulence: Markovianity and channel prediction". Proc. Int. Conf. Space Optical Systems and Applications (ICSOS), IEEE, Santa Monica, USA, May 2011, pp. 28-37
- [7] Mostafa, A., Hranilovic, S.: "Channel measurement and Markov modeling of an urban free-space optical link", IEEE/OSA J. Opt. Commun. Netw., 2012, 4, (10), pp. 836-846
- [8] "Free-Space Optical Communication Algorithms Laboratory (FOCAL), McMaster University", <http://focal.mcmaster.ca/>, accessed Feb. 2015
- [9] Denic, S.Z., Djordjevic, I., Anguita, J., Vasic, B., Neifeld, M.A.: "Information theoretic limits for free-space optical channels with and without memory", IEEE J. of Lightwave Tech., 2008, 26, (19), pp. 3376-3384
- [10] Shannon, C.E.: "A mathematical theory of communication", The Bell System Tech. J., 1948, 27, pp. 379-423, 623-656
- [11] Goldsmith, A.: "Wireless communications", (Cambridge University Press, New York, 1st edn. 2005), pp. 99-111
- [12] Giggenbach, D., Cowley, W., Grant, K., Perlot, N.: "Experimental verification of the limits of optical channel intensity reciprocity", Applied Optics, 2012, 51, (16), pp.3154-3152
- [13] Wang, H.S., Moayeri, N.: "Finite-state Markov channel – a useful model for radio communication channels," IEEE Trans. on Vehicular Tech., 1995, 44, (1), pp. 163-174