

Channel Capacity and Non-Uniform Signalling for Free-Space Optical Intensity Channels

Ahmed A. Farid, *Student Member, IEEE*, and Steve Hranilovic, *Senior Member, IEEE*

Abstract—This work considers the design of capacity-approaching, non-uniform optical intensity signalling in the presence of average and peak amplitude constraints. Although it is known that the capacity achieving input distribution is discrete with a finite number of mass points, finding it requires complex non-linear optimization at every SNR. In this work, a simple expression for a capacity-approaching distribution is derived via source entropy maximization. The resulting mutual information using the derived discrete non-uniform input distribution is negligibly far away from the channel capacity. The computation of this distribution is substantially less complex than previous optimization approaches and can be easily computed at different SNRs. A practical algorithm for non-uniform optical intensity signalling is presented using multi-level coding followed by a mapper and multi-stage decoding at the receiver. The proposed signalling is simulated on free-space optical channels and outage capacity is analyzed. A significant gain in both rate and probability of outage is achieved compared to uniform signalling, especially in the case of channels corrupted by fog.

I. INTRODUCTION

POINT-to-point free-space optical (FSO) communication links through the atmosphere provide an economical high-speed link for wireless access [1]–[3]. However, this high data rate can be degraded severely due to the atmospheric turbulence induced fading, scattering and misalignment errors [3]–[6]. Conventional signalling for FSO channels consists of uncoded, equiprobable binary on-off keying. This work considers improving the rates and reliability of FSO links using capacity-approaching, non-uniform signalling. A family of capacity-approaching input distributions are developed along with practical non-uniform signalling which is shown to be essential in maximizing the rate of a wide range of practical FSO channels.

The performance of FSO channels has been extensively studied using bit-error rate (BER) as a metric [7]–[12]. Recently, information and coding theory have been applied to FSO channels for both Poisson [13], [14] and Gaussian channels [15] in the absence of misalignment errors. The performance of detection techniques [16], bounds on the pairwise error probability for a variety of coding schemes [17], [18] and low-complexity codes [13] have been considered for FSO channels. The outage capacity for FSO channels considering beam width optimization and on-off keying is

analyzed in [19]. In all previous work, a uniform source distribution is considered.

In order to design input distributions for the optical channel, amplitude constraints must be considered. In wireless optical systems, data are transmitted by modulating the instantaneous intensity of a laser or a LED. Hence, all signals must be non-negative. Due to eye safety standards, an average optical power, i.e., amplitude, constraint is imposed on the transmitted signal. A peak amplitude constraint is also applied due to safety and physical limitations. It is known that under peak and average power constraints, the capacity-achieving distribution is discrete with a finite number of probability mass points [20], [21]. However, analytical expressions for the capacity-achieving input distributions are not available, and are instead found via complex optimization. Under non-negativity and average optical power constraints, a lower bound on channel capacity was computed using the max-entropic continuous exponential distribution [22]. Additional bounds on the capacity of such links were developed in [23]–[25], however, no explicit source distributions were given. As a result, no clear insights for communication system design can be drawn. In recent work [26], [27], a tight lower bound on channel capacity is given with a closed form expression for a capacity-approaching input distribution, however, peak amplitude constraints were not considered in the analysis.

In this paper, under non-negativity, average and peak power constraints, a family of discrete non-uniform input distributions is developed via entropy maximization. For comparison, the capacity-achieving distributions are computed via complex non-linear optimization [20]. The tightness of the corresponding information rates to the channel capacity is verified numerically over a practical range of signal-to-noise ratios (SNRs) and peak-to-average ratios. Based on these results, the designed distributions are termed *capacity-approaching*. The developed distributions have rates close to the channel capacity over a wide range of SNRs and require fewer amplitude mass points, making implementation simpler. Their simple structure is exploited to develop a practical signalling algorithm. A deterministic mapper, to induce the correct non-uniform distribution, is coupled with multilevel coding (MLC) and multistage decoding (MSD) [28]–[31]. Finally, the proposed non-uniform input distributions are applied to practical FSO channels and the outage capacity is simulated and compared against uniform signalling. The results show that non-uniform signalling is essential for FSO channels, especially those operating at low SNRs, i.e., fog.

Section II presents an overview of the channel model and Sec. III presents the channel capacity definitions. Section IV

Manuscript received 16 January 2009; revised 1 May 2009. This paper was presented in part at the IEEE Int. Symp. Inform. Theory, July 6–11, 2008, Toronto, Canada and the IEEE Int. Conf. Commun., June 14–18, 2009, Dresden Germany.

The authors are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada L8S 4K1 (email: hranilovic@mcmaster.ca).

Digital Object Identifier 10.1109/JSAC.2009.0912xx.

outlines the design of capacity-approaching input distributions via entropy maximization. A practical algorithm to approach the higher available data rates is presented in Sec. V and is applied to slow-fading FSO channels in Sec. VI. The paper concludes in Sec. VII.

II. SYSTEM MODEL

The following notation is used in this paper. Let X be a random variable defined on \mathbb{R} with x as a particular realization. The probability density function (pdf) of X is identified by the subscript X , e.g. $f_X(x)$, and the expectation denoted as $\mathbb{E}\{X\}$. Boldface, e.g., \mathbf{W} , denotes a vector defined on \mathbb{R}^N while sets are defined using blackboard bold font, e.g. \mathbb{R} .

A majority of free-space optical (FSO) communication links employ intensity modulation with direct detection (IM/DD). A consequence of intensity modulation is that all signals input to the channel must be non-negative. Also due to eye safety standards and device limits, both the average power, i.e., average amplitude, and peak amplitude must be limited [2]. A mathematical representation for a pulse amplitude modulated (PAM) IM/DD channel is given by [2]

$$Y = HX + Z, \quad (1)$$

where X represents the transmitted intensity with constraints

$$0 \leq X \leq A, \quad \mathbb{E}\{X\} \leq P,$$

where A is the peak-amplitude limit and P is the average power limit. Channel loss and fading are represented by H with probability density function $f_H(h)$ that is accurately modeled for FSO channels by a Gamma-Gamma density function [32]. The noise, Z , models both thermal noise and ambient light induced shot noise and is well modeled as zero-mean, signal-independent, Gaussian noise with variance σ^2 . The optical SNR for a given channel realization $H = h$ is defined as [22]

$$\text{SNR}(h) = \frac{P h}{\sigma}$$

and the optical peak-to-average power ratio (PAR) is defined as

$$\rho = \frac{A}{P}.$$

For FSO channels the time scale of the fading process, i.e., the coherence time, is on the order of 1–100 msec [16]. Notice that this is significantly slower than typical bit rates which are often shorter than 1 nsec. As a result, the FSO channel is often modeled as a slow-fading channel and the use of extensive interleaving to average fading states is impractical. Based on this fact we assume a block-fading model where H is fixed over a large number of transmitted symbols and varies independently over blocks. Following this assumption the receiver can efficiently estimate the channel and hence we consider a perfect channel state information (CSI) at the receiver. In this scenario, there is a non-zero probability that a transmitted rate exceeds the instantaneous channel capacity. In this case, outage capacity is a suitable metric to characterize the channel capacity where each transmitted rate is associated with an outage probability, i.e., the probability

that the channel cannot support a given rate. In Sec. VI the outage capacity of FSO communication links is considered for both conventional uniform signalling as well as the developed capacity-approaching distributions.

III. CHANNEL CAPACITY

Consider the Gaussian channel model given in (1). The mutual information, in bits/channel use, between channel input and output for a given a channel realization is

$$\mathcal{I}(X; Y|H = h) = \int \int f_{Y|H,X}(y|h, x) f_X(x) \log_2 \frac{f_{Y|H,X}(y|h, x)}{f_{Y|H}(y|h)} dx dy, \quad (2)$$

where

$$f_{Y|H,X}(y|h, x) = \mathcal{N}(h x, \sigma^2),$$

$$f_{Y|H}(y|h) = \int_x f_X(x) f_{Y|H,X}(y|h, x) dx,$$

and $\mathcal{N}(\nu, \gamma^2)$ denotes a Gaussian distribution with mean ν and variance γ^2 . Let \mathbb{F} denote the family of all input distributions $f_X(x)$ satisfying the non-negativity, peak and average optical power constraints. The instantaneous channel capacity, $\mathcal{C}(h)$, for a given channel state h , is the maximum mutual information over the set \mathbb{F} and is given by

$$\mathcal{C}(h) = \max_{f_X(x) \in \mathbb{F}} \mathcal{I}(X; Y|H = h), \quad (3)$$

where the *capacity-achieving* input distribution is

$$f_X^*(x) = \arg \max_{f_X(x) \in \mathbb{F}} \mathcal{I}(X; Y|H = h). \quad (4)$$

In general, $f_X^*(x)$ is different for each PAR and SNR(h), i.e., each channel state h .

For channels with constrained input amplitude and power, it was first shown in [20] that the capacity-achieving input distributions are discrete with a finite number of mass points. Define the family of discrete input distributions

$$\mathbb{P} = \left\{ p_X(x) : p_X(x) = \sum_{k=0}^K a_k \delta(x - x_k), x_k \in [0, A], \right. \\ \left. a_k \geq 0, \sum_{k=0}^K a_k = 1, K \in \mathbb{Z}^+, P \geq \sum_{k=0}^K x_k a_k \right\},$$

where $\delta(\cdot)$ is the delta functional and \mathbb{Z}^+ is the set of positive integers. The number of mass points is $K + 1$, and a_k and x_k are the amplitudes and positions of the k^{th} mass point respectively. Therefore, $f_X^*(x) \in \mathbb{P}$ and the channel capacity for a given h , can be found by solving the following complex non-linear optimization problem,

$$\mathcal{C}(h) = \max_{f_X(x) \in \mathbb{P}} \mathcal{I}(X; Y|H = h), \quad (5)$$

where the free parameters are a_k , x_k and K .

Finding an analytical closed form expression for the optimum distribution is difficult. However, for a given A , P , h , and σ^2 , numerical optimization methods can be used to efficiently solve this non-linear optimization problem to find $f_X^*(x)$. It was shown in [20] that for a given ρ the number

of mass points in $f_X^*(x)$ is monotonically non-decreasing with SNR and a mass point at $x_k = 0$ always exists. The capacity-achieving distribution is found by repeatedly solving the optimization problem in (5) for fixed $K = 1, 2, 3, 4, \dots$ until a stopping criterion is met. Assume that the capacity-achieving distribution has $K^* + 1$ mass points. Solving the optimization problem with $K = K^* + m$, $m > 1$, results in the capacity-achieving distribution where the extra $m - 1$ mass points are assigned zero amplitudes. Consequently, this fact is utilized as a stopping criterion for the optimization problem while incrementing K .

IV. CAPACITY-APPROACHING DISTRIBUTIONS

For every average optical power, peak amplitude and noise variance, the optimization problem (5) must be solved to extract $f_X^*(x)$. In addition to this drawback, the complexity of each run of the optimization problem increases as the number of mass points is increased. In this section, a simple family of capacity-approaching input distributions is developed based on source entropy maximization and termed *capacity-approaching* source distributions. These distributions have mutual information which closely approaches the channel capacity over a practical range of SNRs and are substantially simpler to generate than the capacity-achieving distribution. In addition, these capacity-approaching distributions are fixed over intervals of SNR making practical implementation easier.

A. Definition of Distributions

Consider the set of discrete input distributions $\mathbb{Q}_K \subset \mathbb{P}$ with $K + 1$ equally spaced mass points,

$$\mathbb{Q}_K = \left\{ q(x) \in \mathbb{P} : \ell = \frac{A}{K}, q(x) = \sum_{k=0}^K a_k \delta(x - k\ell) \right\} \quad (6)$$

where ℓ is defined as the mass point spacing. Define the maxentropic input distribution with $K + 1$ mass points in \mathbb{Q}_K as

$$\bar{q}_X(x; K) = \arg \max_{\mathbb{Q}_K} \mathcal{H}(X) \quad (7)$$

where

$$\mathcal{H}(X) = \sum_{k=0}^K a_k \log_2 \frac{1}{a_k}.$$

The input distribution that maximizes $\mathcal{H}(X)$ is considered based on the intuition that this distribution is capacity-approaching at high SNR. Define the collection of entropy maximizing discrete input distributions with different number of mass points as

$$\bar{\mathbb{Q}} = \{ \bar{q}(x) : \forall K \in \mathbb{Z}^+ \bar{q}(x) = \bar{q}_X(x; K) \}$$

For a given A , P and σ^2 , the capacity-approaching input distribution is the distribution in $\bar{\mathbb{Q}}$ which maximizes the mutual information, i.e.,

$$\bar{q}_X^*(x; \bar{K}(\rho, \text{SNR})) = \arg \max_{f_X(x) \in \bar{\mathbb{Q}}} \mathcal{I}(X; Y | H = h), \quad (8)$$

where $\bar{K}(\rho, \text{SNR}) + 1$ is defined as the number of mass points in the capacity-approaching distribution. Notice that while \bar{q}_X is independent of the channel parameters, \bar{q}_X^* is not.

For each number of mass points $K + 1$, an expression for the maxentropic input distribution $\bar{q}_X(x; K)$ in (7) can be found by solving

$$\begin{aligned} & \max_{a_k} \mathcal{H}(X) \\ \text{s.t. } & \sum_{k=0}^K a_k = 1, \quad \sum_{k=0}^K k\ell a_k \leq P, \quad A = K\ell. \end{aligned} \quad (9)$$

Define \mathcal{J} as the Lagrangian associated with the optimization problem as

$$\mathcal{J} = \sum_{k=0}^K a_k \log_2 \frac{1}{a_k} - \lambda_1 \left(\sum_{k=0}^K a_k - 1 \right) - \lambda_2 \left(\sum_{k=0}^K k\ell a_k - P \right).$$

Proposition 4.1: The discrete input distribution with equally spaced mass points that maximizes the entropy is given as,

$$\bar{q}(x; K) = \sum_{k=0}^K \bar{a}_k \delta(x - k\ell) \quad (10)$$

where

$$\begin{aligned} \bar{a}_k &= \frac{1}{K + 1}, & A &\leq 2P \\ \bar{a}_k &= \frac{t_0^k}{1 + t_0 + t_0^2 + \dots + t_0^K}, & A &\geq 2P \end{aligned} \quad (11)$$

t_0 is the unique positive real root of

$$S(t) = \sum_{k=0}^K \left(1 - \frac{k}{K} \rho \right) t^k. \quad (12)$$

Proof: Notice that when $A < 2P$, a uniform distribution over the a_k satisfies both average and peak amplitude constraints. Since the uniform distribution is entropy maximizing among all discrete distributions, it must also be the result of the optimization problem. In this case there is slack in the average constraint, i.e., $\sum_{k=0}^K k\ell a_k < P$ and hence $\lambda_2 = 0$.

When $A \geq 2P$ the above optimization problem is solved analytically considering all constraints. Solving the equations $\partial \mathcal{J} / \partial a_k = 0$ and substituting the constraints given in (9) it is straightforward to obtain (11). Notice that $\bar{a}_k > 0$ in (11) and for a given t_0 the polynomial $1 + t_0 + t_0^2 + \dots + t_0^K$ has a fixed sign. Thus, $t_0 > 0$ since otherwise the sign of \bar{a}_k would alternate. Lemma A.1 in the Appendix demonstrates that $S(t)$ has a unique positive root in $[0, 1]$. For a given K , equation (12) can be solved efficiently to obtain the real root t_0 . As shown in Appendix A, $t_0 = 1$ when $\rho = 2$ and both expressions for the mass point amplitude a_k coincide. ■

An analytic expression for \bar{K} , that maximizes the mutual information is difficult to obtain since it varies with SNR and ρ and since the mutual information expression depends on a nested relation between t_0 and K . Here we provide a simple approximation for \bar{K} based on numerical analysis.

For a given ρ and SNR, $\bar{q}^*(x; \bar{K}(\rho, \text{SNR}))$ has the highest mutual information over all source distributions in $\bar{\mathbb{Q}}$. Consider increasing SNR for a fixed ρ . There exists an $\text{SNR}_o > \text{SNR}$ at which the mutual information using $\bar{q}^*(x; \bar{K}(\rho, \text{SNR}_o))$, is greater than or equal to that obtained using $\bar{q}^*(x; \bar{K}(\rho, \text{SNR}))$. In this case, $\bar{K}(\rho, \text{SNR}_o) = \bar{K}(\rho, \text{SNR}) + 1$. We term SNR_o

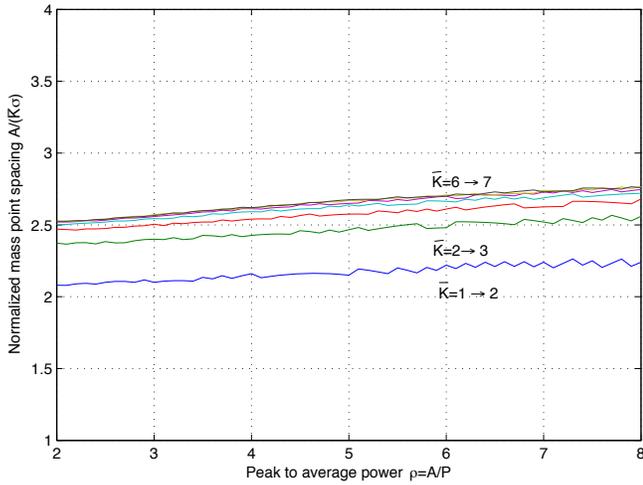


Fig. 1. Normalized spacing ℓ/σ versus ρ at transition SNRs between $\bar{K} + 1 \rightarrow \bar{K} + 2$ mass points in $\bar{q}_X^*(x; \bar{K})$.

as a *transition SNR* where the number of mass points in the capacity-approaching distribution is incremented.

In order to remove the impact of scaling, define the normalized mass point spacing

$$\frac{\ell}{\sigma} = \frac{A}{K\sigma} = \rho \cdot \text{SNR} \cdot \frac{1}{\bar{K}(\rho, \text{SNR})}.$$

The normalized mass point spacing is shown in Fig. 1 versus ρ at the transition SNRs as \bar{K} increases from 1 to 7. From the figure, it is clear that ℓ/σ for a given transition between numbers of mass points changes slowly with ρ . A simple approximation adopted here is to set $\ell/\sigma = c$, for some constant c . This constant can be chosen from Fig. 1 depending on the SNR range of interest. For example, for high SNR cases where \bar{K} is large, a reasonable value for $c \approx 2.7$, whereas, for low SNR cases $c \approx 2.2$. For the purposes of our numerical work we set $c = 2.5$ to yield the simple approximation for \bar{K}

$$\hat{K} = \left\lfloor \frac{A}{2.5\sigma} \right\rfloor. \quad (13)$$

Although many selections of \hat{K} yield acceptable performance, the resulting mutual information using this approximation remains close to the channel capacity as shown in Sec. IV-B.

B. Channel Capacity and Information Rates

The mutual information of the proposed maxentropic input distributions (7) versus SNR for $\rho = 2$ and $K = 1, 2$ and 3 are shown in Fig. 2. Recall that when $\rho = 2$ the maxentropic input distribution is uniform with $K + 1$ probability mass points (*Proposition 4.1*). For comparison, the channel capacity computed via non-linear optimization is also presented. Clearly, a negligible gap can be noticed between the mutual information and the channel capacity for different SNRs. In addition, based on the numerical results obtained from solving the optimization problem, the input distribution $f_X^*(x) = 0.5\delta(x) + 0.5\delta(x - A)$ is a capacity-achieving input distribution at low SNRs which is also the maxentropic distribution obtained in (11). This distribution can be implemented directly using a binary linear codes.

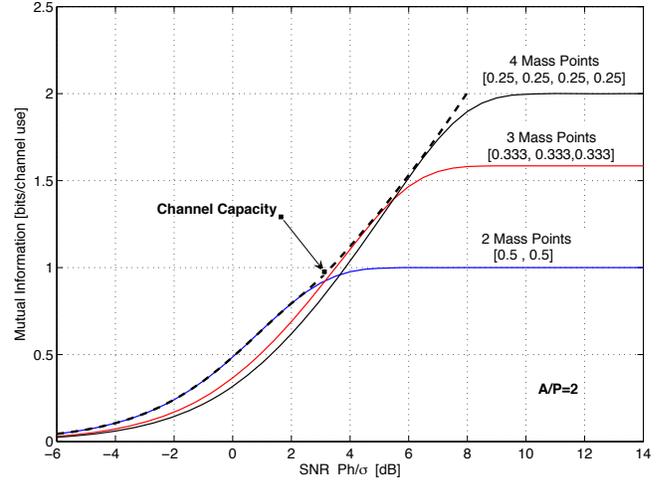


Fig. 2. Channel capacity and mutual information for the maxentropic input distributions (7) with different number of mass points versus SNR for $\rho = 2$.

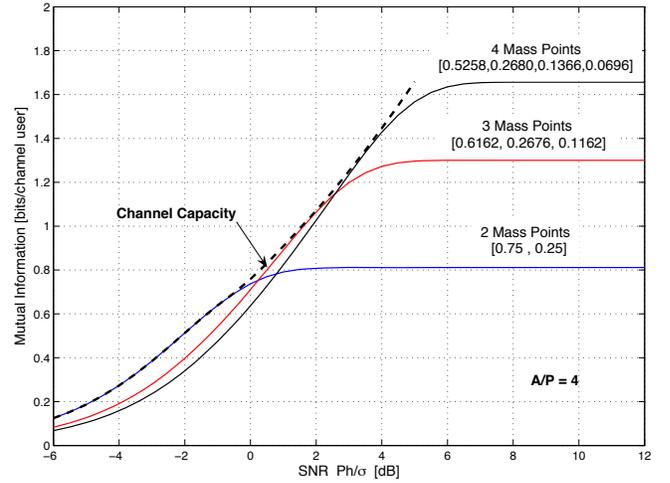


Fig. 3. Channel capacity and mutual information for the maxentropic input distributions (7) with different number of mass points versus SNR for $\rho = 4$.

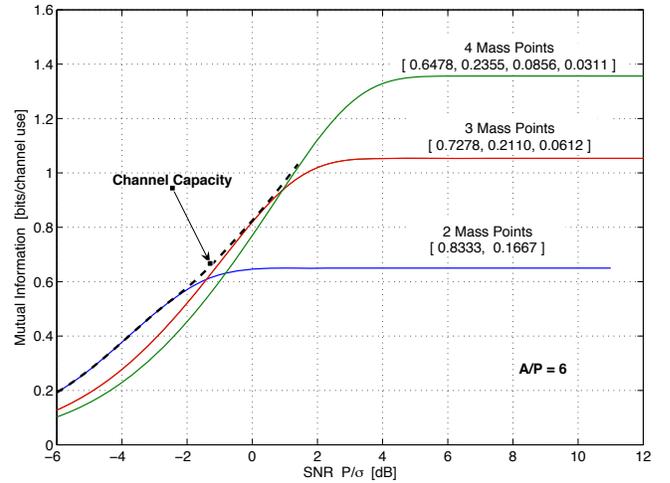


Fig. 4. Channel capacity and mutual information for the maxentropic input distributions (7) with different number of mass points versus SNR for $\rho = 6$.

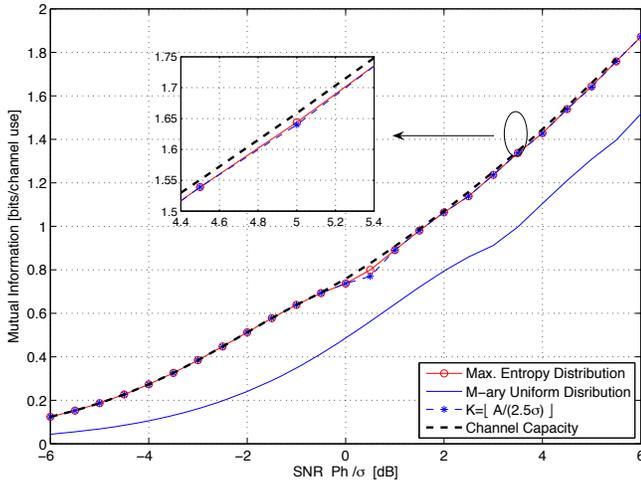


Fig. 5. Channel capacity and mutual information for the proposed input distribution with K and \hat{K} and the uniform input distributions for $\rho = 4$.

Although the uniform distribution is a capacity-approaching distribution when $\rho = 2$, the situation is quite different when $\rho > 2$. As shown in Fig. 3 when $\rho = 4$ the capacity-approaching distributions are non-uniform. The maxentropic input distributions approach the channel capacity over a wide range of SNRs. At low SNRs the maxentropic distribution is a capacity-approaching distribution with $\bar{q}^*(x, 1) = 0.75 \delta(x) + 0.25 \delta(x-A)$ where more weight is assigned to the mass point located at zero. Fig. 4 presents the mutual information and the channel capacity when $\rho = 6$. As shown, as ρ increases the maxentropic input distributions become more non-uniform with increasing mass point amplitude at zero amplitude.

From the previous discussion, at low SNRs, an input distribution with two mass points (i.e., $K = 1$) is sufficient to achieve/approach the channel capacity. The probability mass for these two-level outputs can be found from (10) and are given by

$$[p_0, p_1] = \left[\frac{\rho - 1}{\rho}, \frac{1}{\rho} \right].$$

Notice that as ρ increases the resulting capacity-approaching distributions become increasingly non-uniform with most weight on the zero-amplitude.

The capacity of the wireless optical intensity channel versus SNR (5) is shown in Fig. 5 for $\rho = 4$. In addition, the mutual information with input distribution $\bar{q}_X^*(x; \bar{K})$ in (8) is presented. Clearly, the proposed input distributions achieve nearly all of the data rate offered by the optical channel with a substantial reduction in complexity. Figure 5 also plots the mutual information using the maxentropic source distribution (7) where the number of mass points are approximated to have $\hat{K} + 1$ mass points (13). Notice that the mutual information is also close to the capacity over the SNR range considered, and only differs negligibly from the case where a search is performed to find a good \bar{K} value. Thus, the approximation (13) does not incur a significant penalty in terms of rate.

For comparison, the mutual information using uniform input distributions satisfying both the average and the peak power constraints is also presented. The uniform distributions utilized in the comparison are selected from the set \mathbb{U} defined for $\rho \geq 2$

TABLE I
NUMBER OF MASS POINTS FOR CAPACITY-ACHEIVING/APPROACHING INPUT DISTRIBUTIONS ($\rho = 4$)

SNR [dB]		-1	0	1	2	3	4	5
Capacity-Achieving (5)	$K^* + 1$	2	3	3	4	4	5	7
Max. Entropy (8)	$\bar{K} + 1$	2	2	3	3	4	4	5
Max. Entropy approx K (13)	$\hat{K} + 1$	2	2	3	3	4	5	6
Uniform	$K_u + 1$	2	2	2	2	2	3	3

as

$$\mathbb{U} = \left\{ q_X(x) : \forall K > 0, d = \frac{2P}{K}, \right. \\ \left. q_X(x) = \sum_{k=0}^K \frac{1}{K+1} \delta(x - kd) \right\}.$$

Notice that these distributions have mass points with equal probability that are equally spaced. At each SNR, the distribution in \mathbb{U} which maximizes the mutual information is selected and the information rate plotted. As shown in Fig. 5, a remarkable gap between the mutual information of the uniform distribution and the non-uniform distribution is noticed. Therefore, the use of non-uniform signalling is essential for optical intensity channels, especially as ρ increases.

C. Input Distributions and Numbers of Mass Points

The number of mass points in the capacity-achieving source distribution (5), $K^* + 1$, the capacity-approaching maxentropic distribution (8), $\bar{K} + 1$, the maxentropic distribution with approximated K (13), $\hat{K} + 1$, and the uniform distribution, $K_u + 1$, are presented in Table I for different SNR values.

Note that, although K^* is fixed over a range of SNR, the input distribution, i.e., mass points amplitudes and locations, varies for each SNR value. Unlike the capacity-achieving input distribution, for a given \bar{K} the maxentropic distribution is fixed. Therefore, and as shown in Table I, the maxentropic distribution that maximizes the mutual information is fixed over a range of SNRs. This advantage, in addition to the negligible gap between the mutual information and the channel capacity and the substantial complexity reduction in generating this distribution, renders the maxentropic input distribution more practical for realization over the capacity-achieving distribution obtained via optimization. Also notice that the uniform distribution has the minimum number of mass points as SNR increases simplifying its implementation, however, there is a severe rate degradation as shown in Fig. 5.

The capacity-achieving and the maxentropic input distributions are shown in Fig. 6 for SNRs=[-3, 0, 3, 5] dB at $\rho = 4$. Notice that the mass point spacing is a free parameter in the capacity-achieving distributions while it is fixed in the maxentropic distributions. When SNR=-3 dB, an input distribution with two mass points can achieve the channel capacity and both the capacity-achieving and the maxentropic distributions coincide. As the SNR increases to 0 dB and although the number of mass points is different, the maxentropic distribution is still capable to approach the

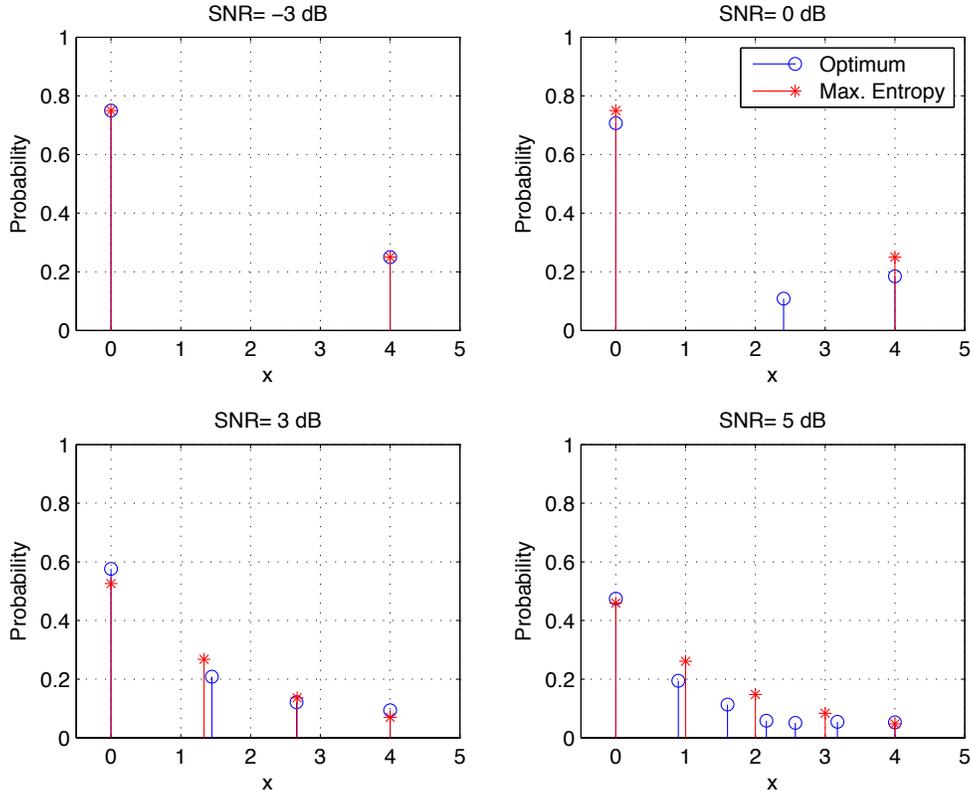


Fig. 6. The optimum and the proposed input distribution for different SNRs with $\rho = 4$.

channel capacity with a negligible gap. When SNR increases further to 5 dB, five and seven mass points are shown for the maxentropic and the capacity-achieving input distributions respectively. Surprisingly, the resulting information rates for both distributions are still very close as shown in Fig. 5.

V. NON-UNIFORM SIGNALLING ALGORITHM

Based on the results obtained in the previous section, non-uniform input distributions achieve higher rates compared to uniform distributions for different SNRs. Although coding for uniform input distributions is simple where binary linear codes can be applied directly, coding for channels with non-uniform input distribution is more complex and different strategies are proposed. A deterministic mapper at the output of a binary encoder was presented in [28]. LDPC codes design over $GF(q)$ is discussed in [33] where higher complexity in both code design and decoding process limits the employment of such techniques. A method to realize a capacity approaching system with a non-uniform source distribution is to employ a multilevel coding (MLC) followed by a deterministic mapper [29]–[31]. In this section we apply the results of [28] and [31] to develop an algorithm for an FSO system with non-uniform input distributions.

A. Generation of a Non-uniform Input Distribution

Here we consider the combination of MLC with a deterministic mapper which is used to ensure that an appropriate

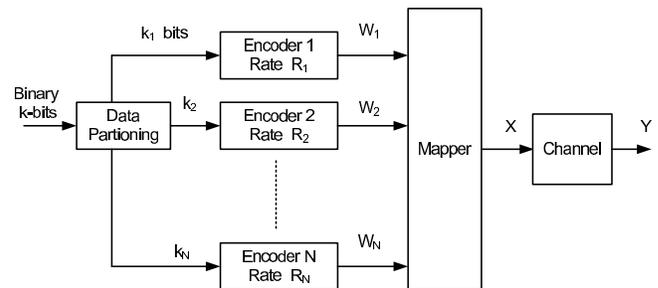


Fig. 7. System model for multilevel coding (MLC) and mapping scheme for optical intensity channels with non-uniform channel input distribution.

non-uniform distribution is induced on the output symbols. Figure 7 shows a block diagram for the MLC system including a deterministic mapper. The inputs are k independent, equally likely bits which are divided into N sub-streams each with k_i bits such that $\sum_{i=1}^N k_i = k$. In each of the sub-streams the i^{th} encoder applies a linear binary code of rate \mathcal{R}_i .

The output of the encoders are the bits $\mathbf{W} = [W_1, \dots, W_N]$, where W_i denotes the i^{th} encoder output bit. The W_i 's are mapped to a constellation point $X \in [0, \ell, 2\ell, \dots, A]$ using a deterministic mapper \mathcal{M} . Note that the mapping $\mathbf{W} = [W_1, \dots, W_N] \xrightarrow{\mathcal{M}} X$ is not necessarily a bijection, as is conventionally assumed in MLC/MSD [29]. In spite of this fact, the mutual information, $\mathcal{I}(X; Y)$, between channel input and output is unaffected as presented in the following Lemma.

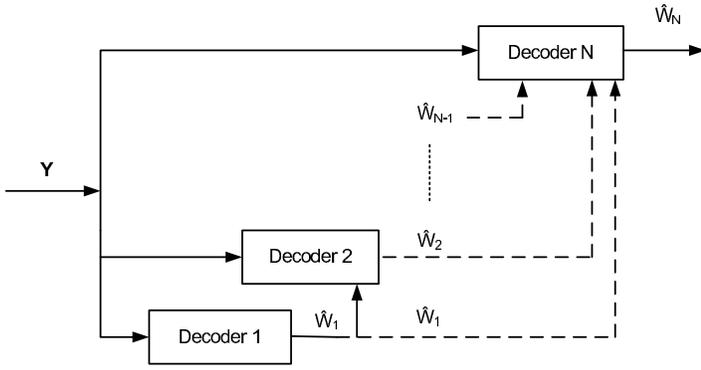


Fig. 8. System model for multi-stage decoders (MSD).

Lemma 5.1: Given a Markov chain $\mathbf{W} \rightarrow X \rightarrow Y$ and a deterministic mapper $\mathcal{M}(\mathbf{W}) = X$, then

$$\mathcal{I}(X; Y) = \mathcal{I}(\mathbf{W}; Y).$$

Proof: See [31]. ■

Applying the chain rule, the mutual information can be expressed in terms of the sub-channels rates as follows,

$$\mathcal{I}(\mathbf{W}; Y) = \sum_{i=1}^N \mathcal{I}(W_i; Y | W_1, \dots, W_{i-1}),$$

where the sub-channel rates are given by

$$\mathcal{R}_i = \mathcal{I}(W_i; Y | W_1, \dots, W_{i-1}).$$

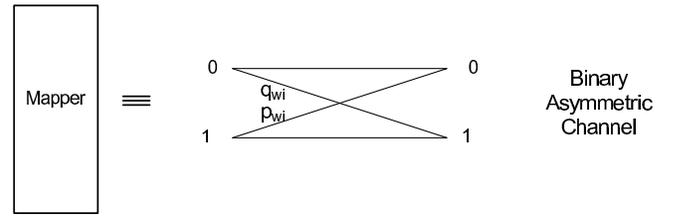
The received codeword is decoded sequentially to extract the transmitted data bits. Figure 8 shows a block diagram for the multi-stage decoding (MSD) technique that describes the rule of the MSD process. The first decoder utilizes the received signal, Y , to obtain the estimate \hat{W}_1 . Given that the first codeword is decoded correctly, i.e., $\hat{W}_1 = W_1$, the second decoder utilizes both Y and \hat{W}_1 to decode the second codeword and obtain the estimate \hat{W}_2 . This process is repeated till \hat{W}_N is estimated. Notice that if an estimate was incorrect, i.e., $\hat{W}_i \neq W_i$, error propagation will occur and lead to decoding error.

The mapper can be modeled with the equivalent channel model shown in Fig. 9 which is in general different for each W_i . As an example, consider two mass points with input alphabet $\{0, A\}$ and $p_0 = 3/4$. This system can be constructed using $N = 2$ encoders where the mapping function \mathcal{M} is defined as,

$$\mathbf{W} = [W_1, W_2] \xrightarrow{\mathcal{M}} X: X = \begin{cases} 1 & \text{if } W_1 = W_2 = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

The equivalent channel seen by bit W_1 is a Z-channel with $p_{W_1} = 1/2$ and $q_{W_1} = 0$. However, the situation is different for W_2 where the equivalent channel is determined based on W_1 . Given $W_1 = 0$, then $p_{W_2} = 1$ and $q_{W_2} = 0$ otherwise $p_{W_2} = q_{W_2} = 0$.

In general the equivalent channel for the proposed system is composed of two cascaded channels, a binary asymmetric channel representing the mapper followed by an optical fading channel with Gaussian noise. LDPC codes optimized for


 Fig. 9. Equivalent channel for the mapper as seen by the individual bit W_i .

Gaussian channels have been shown to perform well over many asymmetric channels [34] and are hence used in our design and simulations in Sec. VI.

B. Quantized-level Distributions

Although the input distribution given in (10) has a simple form and is capacity-approaching, its implementation is impractical due to the real valued mass point amplitudes. In this subsection, a modified version of this distribution is considered with a quantized probability mass amplitudes with a finite number of quantization levels. From (10) and since $t_0 \in [0, 1]$, the elements of $\{\bar{a}_k\}$ in (11) satisfy

$$\forall i < j, \quad \bar{a}_i > \bar{a}_j.$$

Define

$$\check{\mathcal{Q}}_K = \{\check{q}(x; K) \in \mathcal{Q}_K : a_k \in \mathbb{A}, a_k \geq a_{k+1},\}$$

where \mathbb{A} is the set of rational numbers of the form,

$$\mathbb{A} = \left\{ a : a = \frac{i}{2^N}, i \in \{1, \dots, 2^N - 1\} \right\}$$

and 2^N denotes the number of permissible quantization levels. The set $\check{\mathcal{Q}}_K$ consists of all discrete input distributions with $K + 1$ equally spaced mass points with descending quantized mass amplitudes. Note that for a given K , there exist a number of input distributions in $\check{\mathcal{Q}}_K$ which satisfy the peak and the average power constraints. As a result, for a given SNR it is computationally expensive to search for the distribution that maximizes the mutual information over $\check{\mathcal{Q}}_K$. In order to reduce the computation within the set $\check{\mathcal{Q}}_K$, a new set $\tilde{\mathcal{Q}}$ is proposed based on the relative entropy, \mathcal{D} , between the maxentropic input distribution, $\bar{q}(x; K)$, and the quantized distribution, $\check{q}(x; K)$, and is given by,

$$\tilde{\mathcal{Q}} = \left\{ \check{q}(x) : \forall K \in \mathbb{Z}^+ \quad \check{q}(x) = \arg \min_{\check{q} \in \check{\mathcal{Q}}} \mathcal{D}(\bar{q}(x; K) || \check{q}(x; K)) \right\}$$

where

$$\mathcal{D}(\bar{q}(x; K) || \check{q}(x; K)) = \sum_{k=0}^K \bar{a}_k \log \left(\frac{\bar{a}_k}{\check{a}_k} \right).$$

The savings in computation arise since computing the relative entropy is simpler than numerically simulating of the mutual information for all input distributions in $\check{\mathcal{Q}}_K$ over K . Once $\tilde{\mathcal{Q}}$ is formed, the input distribution which maximizes the mutual information is selected, i.e.,

$$\tilde{q}_X^*(x; \tilde{K}) = \arg \max_{f_X(x) \in \tilde{\mathcal{Q}}} \mathcal{I}(X; Y | H = h), \quad (15)$$

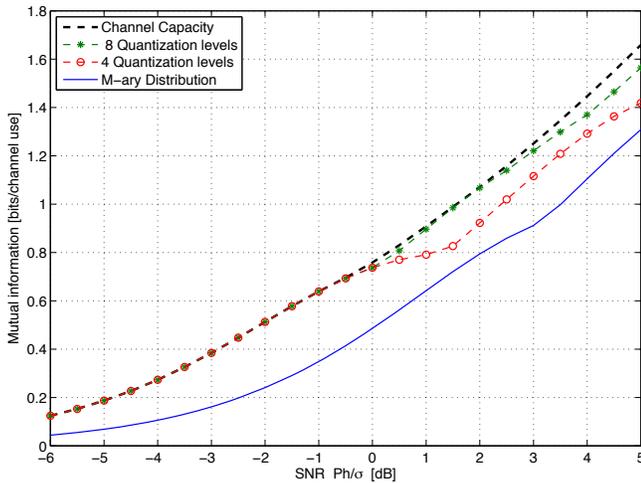


Fig. 10. Channel capacity and mutual information for the quantized input distribution (15) with $\tilde{K} + 1$ mass points for 4- and 8-quantization levels for $\rho = 4$. The M -ary uniform input distributions is presented for comparison.

where $\tilde{K} + 1$ is the number of mass points associated with the input distribution $\tilde{q}_X^*(x; \tilde{K})$. Notice that $\tilde{q}_X^*(x; \tilde{K})$ can be implemented directly by the deterministic mapper discussed in Sec. V-A.

Figure 10 presents the mutual information versus SNR when the quantized-level input distributions are employed for $\rho = 4$. When $N = 2$, i.e., $2^N = 4$ quantization levels. In this example, the corresponding permissible mass points amplitudes are

$$\bigcup_K \tilde{\mathcal{Q}}_K = \left\{ \left[\frac{3}{4}, \frac{1}{4} \right], \left[\frac{1}{2}, \frac{1}{2} \right], \left[\frac{1}{2}, \frac{1}{4}, \frac{1}{4} \right], \left[\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \right] \right\}. \quad (16)$$

The mass point spacing is set such that average and peak constraints are met. The spacing ℓ is increased until one of the constraints is met with equality. In Fig. 10 and for a given SNR, the 4-Quantization levels distribution that maximizes the mutual information is selected among the permissible set (16). Also the information rate using 8-Quantization levels distribution is shown in Fig. 10.

At low SNRs the information rates using the input distribution $q(x) = 0.75 \delta(x) + 0.25 \delta(x - A)$ are negligibly far from the channel capacity. Since both 4- and 8-Quantization levels techniques can both generate this distribution, they have the same mutual information at low SNRs. For comparison the channel capacity obtained by numerically solving (5) and the mutual information using M -ary uniform input distributions from \mathbb{U} are shown. Clearly, the information rates using the 8-Quantization levels distribution approach the channel capacity closely with a small gap while the rates achieved using the 4-Quantization level distributions incur a degradation at moderate SNRs.

VI. CODING SCHEME FOR FSO CHANNELS

Under the slow-fading channel model considered, there is a probability bounded above zero that the realized channel capacity is below the transmitted data rate, i.e., $\mathcal{C}(h) < \mathcal{R}_0$.

An *outage* is that event that the instantaneous capacity of the channel is below the transmission rate. For a given transmission rate, \mathcal{R}_0 , the probability of outage is denoted as

$$P_{\text{out}}(\mathcal{R}_0) = \text{Prob}(\mathcal{C}(h) < \mathcal{R}_0).$$

Considering the monotonic increase of $\mathcal{C}(\cdot)$ with h , the above equation can be rewritten as follows

$$P_{\text{out}}(\mathcal{R}_0) = \text{Prob}(h < h_0),$$

where the threshold h_0 is given by

$$\mathcal{I}(X; Y | H = h_0) = \mathcal{R}_0. \quad (17)$$

The pairs $(\mathcal{R}_0, P_{\text{out}})$ that can be simultaneously realized depend on the statistical properties of the channel which are presented next.

A. Optical Channel Fading

The channel gain H models the random fluctuation of the optical intensity at the receiver due to atmospheric effects and pointing errors. The gain can be factored as $H = g_a H_a H_p$, where g_a is a deterministic loss in the channel due to scattering, H_a models the random turbulence induced fading and H_p models random pointing errors.

Atmospheric turbulence induces fluctuations in the received optical intensity which is well modeled by a Gamma-Gamma distribution [32] providing close agreement with the measurements under different turbulence strengths. The Gamma-Gamma pdf is expressed as [35],

$$f_{H_a}(h_a) = \frac{2(\alpha\beta)^{(\alpha+\beta)/2}}{\Gamma(\alpha)\Gamma(\beta)} h_a^{\frac{(\alpha+\beta)}{2}-1} K_{\alpha-\beta}(2\sqrt{\alpha\beta h_a}), \quad (18)$$

where $K_{\alpha-\beta}(\cdot)$ is the modified Bessel function of the second kind, $\Gamma(\cdot)$ is the gamma function, and $1/\beta$ and $1/\alpha$ are the variances of small and large scale eddies respectively defined in terms of Rytov variance σ_R^2 given by [32],

$$\sigma_R^2 = 1.23 k^{7/6} C_n^2 L^{11/6}$$

where C_n^2 is the weather-dependent index of refraction structure parameter, $k = 2\pi/\lambda$ is the optical wave number, λ is the wavelength and L is the propagation distance. The fading statistical model given in (18) is governed by the parameters α and β . We consider a Gaussian beam with beam waist, w_o , and radius of curvature, F_o . Expressions for α and β , taking into account the aperture average effects at the receiver, are given explicitly in [35, Sec. 10.3.5].

The attenuation factor, g_a , is also weather dependent. It is described by the exponential Beers-Lambert Law [36, Eq. 1],

$$g_a = \exp(-\mu L)$$

where μ is the attenuation coefficient and is expressed in terms of the visibility through empirical formulas [36].

The misalignment error can be modeled by considering a laser beam with a spatial Gaussian profile of beam waist w_L at distance L and a receiver with a circular aperture of radius a . The random displacement at the receiver in both elevation and horizontal directions, due to misalignment, has an identical zero-mean Gaussian distribution with variance σ_s^2 [7]. As a

result, the fraction of collected power by the receiver aperture, h_p , has the following distribution [19],

$$f_{H_p}(h_p) = \frac{\gamma^2}{A_0 \gamma^2} h_p^{\gamma^2-1}, \quad 0 \leq h_p \leq A_0 \quad (19)$$

where

$$v = \frac{\sqrt{\pi}a}{\sqrt{2}w_L}, \quad A_0 = [\text{erf}(v)]^2, \quad \gamma = \frac{\sqrt{\pi} w_L^2 \text{erf}(v)}{4\sigma_s v \exp(-v^2)}.$$

An expression for the beam spot size at the receiver, w_L , is given in [35, Eq. 45, p. 238]. Note that the jitter displacement standard deviation at the receiver σ_s can be expressed in terms of the jitter angle standard deviation at the transmitter σ_θ by the approximation $\sigma_s \cong z \sigma_\theta$.

In general, the statistics of H_a vary with radial distance from the beam center. For the case of mild to moderate jitter, as is the case in Sec. VI-B, it can be demonstrated that the normalized intensity variance changes within 1% at the extreme misalignment deviation point (radial displacement $r \cong 3\sigma_s$) compared to the beam centre point ($r = 0$) [35]. As a result, for small displacements, the statistics of the Gamma-Gamma distribution are set to be equal to those at the beam centre. Thus, the probability density function for atmospheric turbulence and pointing errors fading is given by (see [19] for more details),

$$f_H(h) = \frac{\gamma^2 h^{\gamma^2-1}}{(A_0 g_a)^{\gamma^2}} \int_{h/A_0 g_a}^{\infty} h_a^{-\gamma^2} f_{H_a}(h_a) dh_a. \quad (20)$$

B. Simulation results

The probability of outage is simulated on a typical FSO channel considering both the atmospheric turbulence and the misalignment effects. Atmospheric parameters of a light fog weather condition are used and are listed in Table II. Notice that the statistical channel models depend also on the beam waist, w_o , radius of curvature, F_o , distance, L , receiver radius, a , and the jitter angle variance, σ_θ^2 . The physical parameters for the FSO system are also presented in Table II.

The impact of non-uniform signalling on the pairs $(\mathcal{R}_0, P_{\text{out}})$ can be quantified for modelled FSO channel. Consider the case when $\rho = 4$ and the input distribution $\bar{q}(x; 1) = 0.75 \delta(x) + 0.25 \delta(x - A)$ which is the capacity-approaching distribution at low SNR (see Fig. 3). For comparison the uniform input distribution $q(x) = 0.5 \delta(x) + 0.5 \delta(x - A/2) \in \mathbb{U}$ which has the same average power as the non-uniform case.

The probability of outage $P_{\text{out}}(\mathcal{R}_0)$ versus the rate \mathcal{R}_0 is shown in Fig. 11. The grey area is the unachievable region, i.e., no pair $(\mathcal{R}_0, P_{\text{out}})$ in this region can be realized with any input distribution satisfying the peak and the average power constraints. Notice that the performance with $\bar{q}(x; 1)$ nearly coincides with the capacity boundary. Clearly, for a given rate, an order of magnitude reduction in probability of outage is noticed when employing the non-uniform input distribution compared to the uniform input distribution. Similar gain can be achieved in terms of rate. As an example, at $P_{\text{out}} = 10^{-3}$ the non-uniform distribution achieves $\mathcal{R}_0 = 0.187$ bits/channel use, however, the uniform distribution achieves $\mathcal{R}_0 = 0.07$ bits/channel use.

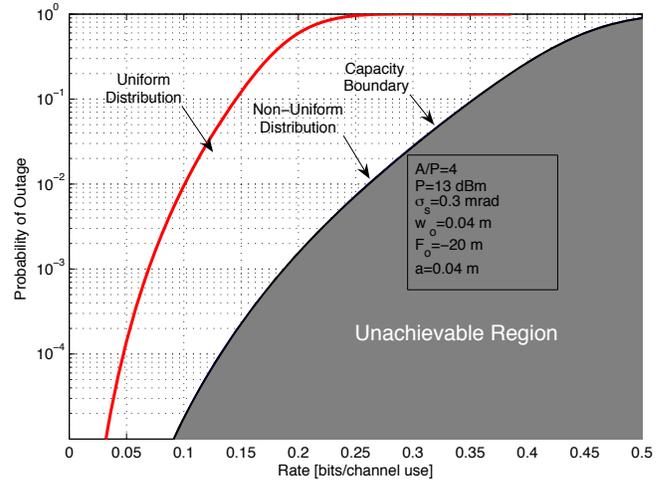


Fig. 11. Probability of outage versus rate using non-uniform $\bar{q}(x; 1) = \{0.75, 0.25\}$ for $\rho = 4$ and uniform input distributions.

C. Code Design Example

The section presents a coding design for the non-uniform signalling algorithm in Sec. V-A and applies it to the light fog FSO channel model.

Consider the non-uniform input distribution in Sec. VI-B. This distribution can be implemented using $N = 2$ encoders and the mapper in (14). In this case the sub-channels rates $\mathcal{R}_1 = \mathcal{I}(Y; W_1)$ and $\mathcal{R}_2 = \mathcal{I}(Y; W_2|W_1)$ are shown in Fig. 12 versus SNR with a total rate $\mathcal{R}_0 = \mathcal{R}_1 + \mathcal{R}_2$.

Consider the FSO channel model and assume that it is required to design the system to operate in a region such that $P_{\text{out}} \leq 10^{-3}$. From Fig. 11 the equivalent rate using the non-uniform input distribution $q(x) = \{0.75, 0.25\}$ is $\mathcal{R}_0 \leq 0.187$ bits/channel use. From Fig. 12 and at $\mathcal{R}_0 = 0.187$, the sub-channel mutual information rates are $\mathcal{R}_1 = \mathcal{I}(W_1; Y) = 0.066$ and $\mathcal{R}_2 = \mathcal{I}(W_2; Y|W_1) = 0.121$.

For convenience, block codes are considered where the output codeword length of each encoder is fixed. Since, in practice, it is not straightforward to design LDPC codes with arbitrary rates, we consider practical LDPC codes with rates $\mathcal{R}_1^{\text{LDPC}} = 0.05$ and $\mathcal{R}_2^{\text{LDPC}} = 0.12$, that is, the system is designed to operate at $\mathcal{R}_0 = 0.17$. The LDPC codes are generated based on degree distributions that are designed for Gaussian channels [38]. As mentioned, these codes have been shown to efficiently operate over many asymmetric channels as well. The codeword length is fixed to $n = 10,000$ for each code. The corresponding probability of outage at rate $\mathcal{R}_0 = 0.17$, bits/channel use, is $P_{\text{out}} = 5 \times 10^{-4}$ for the non-uniform distribution and $P_{\text{out}} = 0.25$ for the uniform distribution.

The performance of this system is shown in Fig. 13 where BER versus SNR is plotted. The Shannon limit at rate $\mathcal{R}_0 = 0.17$ bits/channel use, for both the uniform and the non-uniform input distributions is shown and given as -2.85 dB and -5.25 dB respectively. Clearly the performance of the non-uniform input distribution $q(x) = 0.75\delta(x) + 0.25\delta(x - A)$ when generated using the mapper and coupled with MLC using LDPC codes of length $n = 10,000$ and MSD at the receiver is close to the Shannon limit where this structure

TABLE II
SYSTEM PARAMETERS

Parameter	Symbol	Value
Transmitted optical power	P	20 mW
Combined Tx/Rx optics efficiency	η	0.64
Responsivity	r	0.5 A/W
Wavelength	λ	1550 nm
Noise standard deviation (at 1Gbps)	σ	5×10^{-7} A
Receiver radius	a	4 cm
Distance	z	1000 m
Beam waist radius	w_o	4 cm
Phase front radius	F_o	-20 m
Jitter angle standard deviation	σ_θ	0.3 mrad
Structure Parameter	C_n^2	$0.5 \times 10^{-14} \text{ m}^{-2/3}$ [37]
Attenuation coefficient	μ	2.8 km^{-1} [37]

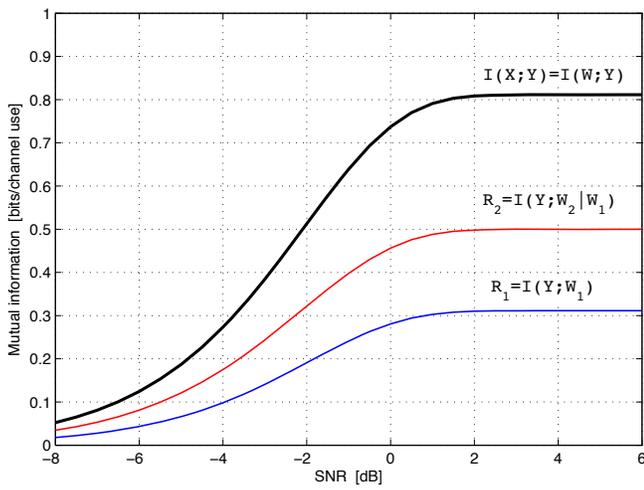


Fig. 12. Sub-channels rate for the MLC system using two mass points with probabilities $[3/4, 1/4]$ at $\rho = 4$.

achieves $\text{BER}=10^{-5}$ at $\text{SNR} = -3.8$ dB, i.e., 1.45 dB greater than the Shannon limit. Furthermore, this structure is approximately 1 dB less than the Shannon limit when using a uniform input distribution.

VII. CONCLUSION

The capacity of optical intensity channels with peak and average optical power constraints is considered. The capacity-achieving distribution is found by numerically solving a non-linear optimization problem. A capacity-approaching distribution based on entropy maximization is developed. In addition to the substantial complexity reduction in generating this distribution compared to the optimum distribution, a negligible gap between the resulting mutual information and the channel capacity is noticed. Unlike the capacity-achieving distributions where for each SNR value a different input distribution is obtained, the proposed maxentropic input distribution is fixed over a range of SNRs. The derived capacity-approaching distributions serve as a useful tool not only to bound the channel capacity but to guide the development of channel coding for optical wireless channels. Simulations of a mapper to realize a non-uniform input distribution at the channel input is presented along with MLC/MSD. Analysis of the

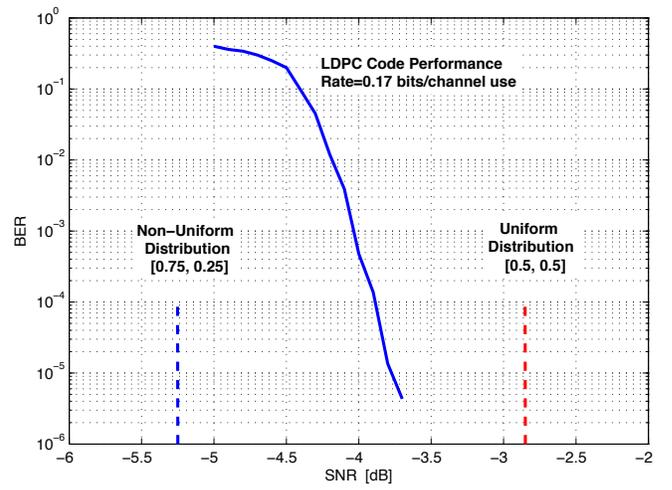


Fig. 13. BER versus SNR for the non-uniform signalling using finite length LDPC codes. Also Shannon limits for both non-uniform and uniform signalling is presented.

outage capacity of FSO channels with non-uniform signalling and atmospheric turbulence and pointing errors is presented. A significant reduction in the outage probability using the proposed non-uniform input distributions is noticed compared to a uniform input distribution. As a result, non-uniform signalling is motivated as an important area of further study to improve the rates and reliability of FSO channels.

APPENDIX A

UNIQUENESS OF THE ROOT t_0

Lemma A.1: The polynomial $S(t)$ in (12) has a unique positive real root $t_0 \in [0, 1]$.

Proof: Consider the polynomial

$$S(t) = \sum_{k=0}^K \left(1 - k \frac{\rho}{K}\right) t^k$$

where $\rho \geq 2$. It is straightforward to show

$$S(0) = 1, \quad \text{and} \quad S(1) = (K + 1) \left(1 - \frac{\rho}{2}\right) \leq 0 \quad (\text{A-1})$$

By the intermediate value theorem of continuous functions, there exists *at least* one real root for $S(t)$ in the interval $[0, 1]$. Note that a root at $t = 1$ exist when $\rho = 2$.

Consider ordering the coefficients of this polynomial in terms of ascending exponents of the variable t . Note that the first coefficient is positive and equal to one independent of ρ and K while the last coefficient is negative since $\rho \geq 2$. Notice also that the coefficients decay linearly with increasing exponent and thus there must exist a $k = k^*$ such that

$$\left(1 - k \frac{\rho}{K}\right) \implies \begin{cases} \text{positive,} & k \leq k^*, \\ \text{negative,} & k > k^*. \end{cases}$$

More precisely, the number of variations in sign in this ordering of the coefficients of $S(t)$ is one. For completeness, consider the following well-known theorem.

Descartes' Rule of Signs [39]: *Let $S(t)$ be a polynomial with real coefficients ordered in terms of ascending powers of the variable. The number of positive roots of $S(t)$ is either equal to the number of variations in sign of consecutive non-zero coefficients of $S(t)$ or less than this by a multiple of 2.*

Applying Descartes' Rule of Signs, $S(t)$ has a single root for $t \geq 0$. Since it has already been demonstrated that a real root exists in $[0, 1]$, thus $S(t)$ has a unique real root $t_0 \in [0, 1]$. ■

REFERENCES

- [1] D. J. Heatley, D. R. Wisely, I. Neild, and P. Cochrane, "Optical wireless: The story so far," *IEEE Commun. Mag.*, pp. 72–82, Dec. 1998.
- [2] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proc. IEEE*, vol. 85, pp. 265–298, Feb. 1997.
- [3] S. Karp, R. Gagliardi, S. Moran, and L. Stotts, *Optical Channels*. New York: Plenum, 1988.
- [4] A. Ishimaru, *Wave Propagation and Scattering in Random Media*. New York: Academic, 1978.
- [5] R. M. Gagliardi and S. Karp, *Optical Communication*. New York: John Wiley, 1976.
- [6] J. H. Churnside and R. J. Hill, "Probability density of irradiance scintillations for strong path-integrated refractive turbulence," *J. Opt. Soc. Am. A*, vol. 4, pp. 727–733, Apr. 1987.
- [7] S. Arnon, "Effects of atmospheric turbulence and building sway on optical wireless communication systems," *Optics Lett.*, vol. 28, pp. 129–131, Jan. 2003.
- [8] D. Kedar and S. Arnon, "Optical wireless communication through fog in the presence of pointing errors," *J. Appl. Opt.*, vol. 42, pp. 4946–4954, Aug. 2003.
- [9] K. Kiasaleh, "On the probability density function of signal intensity in free-space optical communications systems impaired by pointing jitter and turbulence," *Opt. Eng.*, vol. 33, pp. 3748–3757, Nov. 1994.
- [10] O. Steinvall, "Performance of laser tracking of small targets during turbulence and beam jitter," *Opt. Eng.*, vol. 43, pp. 1609–1621, July 2004.
- [11] S. Arnon, "Optimization of urban optical wireless communication systems," *IEEE Trans. Commun.*, vol. 2, pp. 626–629, Nov. 2003.
- [12] S. Arnon, S. Rotman, and N. S. Kopeika, "Beam width and transmitter power adaptive to tracking system performance for free-space optical communication," *J. Appl. Opt.*, vol. 36, pp. 6095–6101, Aug. 1997.
- [13] J. A. Anguita, I. B. Djordjevic, M. Neifeld, and B. V. Vasic, "Shannon capacities and error-correction codes for optical atmospheric turbulent channels," *OSA J. Opt. Net.*, vol. 4, pp. 586–601, Sept. 2005.
- [14] S. Haas and J. H. Shapiro, "Capacity of wireless optical communications," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 1346–1357, Oct. 2003.
- [15] L. Jing and M. Uysal, "Optical wireless communications: system model, capacity and coding," in *IEEE 58th Veh. Tech. Conf.*, Florida, USA, Oct. 2003, pp. 168–172.
- [16] X. Zhu and J. Kahn, "Free space optical communication through atmospheric turbulence channels," *IEEE Trans. Commun.*, vol. 50, pp. 1293–1300, Aug. 2002.
- [17] X. Zhu and J. M. Kahn, "Performance bounds for coded free-space optical communications through atmospheric turbulence," *IEEE Trans. Commun.*, vol. 51, pp. 1233–1239, Aug. 2003.
- [18] M. Uysal, L. Jing, and Y. Meng, "Error rate performance analysis of coded free-space optical links over gamma-gamma atmospheric turbulence channels," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1229–1233, June 2006.
- [19] A. A. Farid and S. Hranilovic, "Outage capacity optimization for free-space optical links with pointing errors," *IEEE J. Lightwave Tech.*, vol. 25, pp. 1702–1710, July 2007.
- [20] J. G. Smith, "The information capacity of amplitude and variance-constrained scalar Gaussian channels," *Inf. Contr.*, vol. 18, pp. 203–219, 1971.
- [21] T. H. Chan, S. Hranilovic, and F. R. Kschischang, "Capacity-achieving probability measure for conditionally Gaussian channels with bounded inputs," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2073–2088, June 2005.
- [22] S. Hranilovic and F. R. Kschischang, "Capacity bounds for power- and band-limited optical intensity channels corrupted by Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 50, pp. 784–795, May 2004.
- [23] S. M. Moser, "Duality-based bounds on channel capacity," Ph.D. dissertation, Swiss Federal Institute of Technology Zürich, 2005.
- [24] A. Lapidoth, S. M. Moser, and M. A. Wigger, "On the capacity of free-space optical intensity channels," in *Proc. IEEE Int. Symp. Inform. Theory*, Toronto, Canada, 2008, pp. 2419–2423.
- [25] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat fading channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2426–2467, Oct. 2003.
- [26] A. A. Farid and S. Hranilovic, "Upper and lower bounds on the capacity of wireless optical intensity channels," in *Proc. IEEE Int. Symp. Inform. Theory*, Nice, France, 2007, pp. 2416–2420.
- [27] —, "Capacity bounds for wireless optical intensity channels with Gaussian noise," *Submitted to IEEE Trans. Inform. Theory*, Nov. 2007.
- [28] R. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley and Sons, 1968.
- [29] H. Imai and S. Hirakawa, "A new multilevel coding method using error-correcting codes," *IEEE Trans. Inform. Theory*, vol. 23, pp. 371–377, May 1977.
- [30] U. Wachsmann, R. Fischer, and J. Huber, "Multilevel codes: theoretical concepts and practical design rules," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1361–1391, 1999.
- [31] J. Jiang and K. R. Narayanan, "Multilevel coding for channels with non-uniform inputs and rateless transmission over the BSC," in *Proc. IEEE Int. Symp. Inform. Theory*, Seattle, Washington, U.S.A., 2006, pp. 518–522.
- [32] L. C. Andrews, R. L. Phillips, C. Y. Hopen, and M. A. Al-Habash, "Theory of optical scintillation," *J. Opt. Soc. Am. A*, vol. 16, pp. 1417–1429, June 1999.
- [33] E. A. Ratzner and D. MacKay, "Sparse low-density parity check codes for channels with cross-talk," in *Proc. IEEE Info. Theory Workshop*, Paris, France, Mar. 2003, pp. 127–130.
- [34] M. Franceschini, G. Ferrari, and R. Raheli, "Does the performance of LDPC codes depend on the channel?" *IEEE Trans. Commun.*, vol. 54, pp. 2129–2132, Dec. 2006.
- [35] L. Andrews and R. Phillips, *Laser Beam Propagation Through Random Media*, 2nd ed. Bellingham, Washington: SPIE, 2005.
- [36] I. Kim, B. McArthur, and E. Korevaar, "Comparison of laser beam propagation at 785 nm and 1550 nm in fog and haze for optical wireless communications," *Proc. SPIE*, vol. 4214, pp. 26–37, Feb. 2001.
- [37] E. J. McCartney, *Optics of the Atmosphere*. New York: John Wiley & Sons, 1976.
- [38] R. Urbanke, "LDPC degree distribution optimizer for Gaussian channel," <http://lthcwwww.epfl.ch/research/ldpcopt/>.
- [39] B. Anderson, J. Jackson, and M. Sitharam, "Descartes' rule of signs revisited," *The American Mathematical Monthly*, vol. 105, no. 5, pp. 447–451, May 1998.



Ahmed A. Farid (S'06) received the B.Sc. and M.Sc. degrees from Alexandria University, Egypt in 1996 and 2000 respectively and the M.A.Sc degree from McMaster University, Canada in 2005 where he is currently pursuing the Ph.D. degree. He worked as a Teaching Assistant with the Department of Electrical Engineering, Faculty of Engineering, Alexandria University, and with the Department of Electrical and Computer Engineering, McMaster University.

His field of interest includes wireless optical communications, digital communications, signal processing, information theory and coding. Mr. Farid was awarded the Canadian Satellite Users Association scholarship in 2007, the Ontario Graduate Scholarship (OGS) in 2008 and the Research in Motion (RIM) Ontario Graduate Scholarship in 2009.



Steve Hranilovic (S'94-M'03-SM'07) received the B.A.Sc. degree with honours in electrical engineering from the University of Waterloo, Canada in 1997 and M.A.Sc. and Ph.D. degrees in electrical engineering from the University of Toronto, Canada in 1999 and 2003 respectively.

He is currently an Associate Professor in the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada.

His research interests are in the areas of free-space and wired optical communications, digital communication algorithms, and electronic and photonic implementation of coding and communication algorithms. He is the author of the book *Wireless Optical Communications Systems* (New York:Springer, 2004).

Dr. Hranilovic is a licensed Professional Engineer in the Province of Ontario. In 2006, he was awarded the Government of Ontario *Early Researcher Award*. He is currently the Chair of the joint IEEE Communications/Signal Processing/Information Theory Societies Chapter in the Hamilton Section.