Technical University of Denmark

DTU

Constellation Shaping for Fiber-optic Channels with QAM and High Spectral Efficiency

Yankov, Metodi Plamenov; Zibar, Darko; Larsen, Knud J.; Christensen, Lars P. B.; Forchhammer, Søren

DTU Library
Technical Information Center of Denmark

Constellation Shaping for Fiber-optic Channels with QAM and High Spectral Efficiency

Metodi P. Yankov, *Student member, IEEE*, Darko Zibar, Knud J. Larsen, Lars P. B. Christensen and Søren Forchhammer, *Member, IEEE*

Abstract—In this letter the fiber-optic communication channel with Quadrature Amplitude Modulation (QAM) input constellation is treated. Using probabilistic shaping, we show that high order QAM constellations can achieve and slightly exceed the lower bound on the channel capacity, set by ring constellations in [1]. We then propose a mapping function for turbo coded bit interleaved coded modulation based on optimization of the mutual information between the channel input and output. Using this mapping, spectral efficiency as high as 6.5 bits/s/Hz/polarization is achieved on a simulated single channel long-haul fiber-optical link excluding the pilot overhead, used for synchronization, and taking into account frequency and phase mismatch impairments, as well as laser phase noise and analog-to-digital conversion quantization impairments. The simulations suggest that major improvements can be expected in the achievable rates of optical networks with high order QAM.

Index Terms—Constellation shaping, fiber-optic communication, non-linear channel capacity.

I. INTRODUCTION

IN recent years, theoretical investigations into the fiber optic channel capacity are gathering more and more attention, as high Spectral Efficiency (SE) must be achieved to meet the ever-growing demand for high data rates. It is well known, that the non-linear Kerr effect in optical fibers limits the capacity in such links, causing a drop in the maximum achievable rate when increasing the input power [1]. In [1], the authors outline the problem and calculate a lower bound on the channel capacity, by optimizing ring constellations. Details on the optimization can be found in [2]. Ring constellations, however, are challenging for implementation and can only serve as guidelines to the achievable rate. In [3], the lower bound is increased significantly by making use of the long time correlation of the Cross and Self Phase Modulation (XPM and SPM) noise in Wavelength Division Multiplexing (WDM) channels. Continuous Gaussian input is assumed there, which is not realizable in practice. Constellation shaping is considered in [4] [5] as a tool to increase the Mutual Information (MI) between the channel input and output. Specifically in [5] trellis shaping is used to achieve very high SE, close to the theoretical lower bound, which adds some latency both in the transmitter and receiver. In [4] shell-mapping is used to achieve shaping

M. Yankov, D. Zibar, K. J. Larsen and S. Forchhammer are with the Department of Photonics Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Copenhagen: e-mail: meya@fotonik.dtu.dk

L. P. B. Christensen is with Fingerprint Cards, 2730 Herlev, Denmark

gain, which quickly becomes impractical as the constellation size is increased. A successful transmission over a long-haul link was demonstrated in [6] using low density parity check codes in combination with polar modulation and iterative receiver, achieving high spectral efficiency. Polar modulation is able to achieve a significant shaping gain over Quadrature Amplitude Modulation (QAM) constellations, but it is more difficult to realize when the transceiver is subject to limited resolution and other practical constraints. QAM constellations are desirable due to their simplicity and the possibility to directly apply I/Q modulation/demodulation. In [6], digital back-propagation is performed, something which will often be difficult, especially in WDM systems, where the coupling between the channels must be considered.

As part of our previous work [7], near-capacity achieving coded modulation scheme was designed for AWGN channel, based on turbo-coded Bit-Interleaved Coded Modulation (BICM), employing QAM constellations and constellation shaping. In this paper, a feasibility study is performed for that method on the fiber-optic communication channel, and a bit-to-symbol mapping function is proposed that achieves high SE. A comparison is also made between the achievable lower bounds on the MI with QAM constellations and ring constellations [1], which are often considered to be the best suited option for highly non-linear channels. Unless stated otherwise, no digital back-propagation will be performed on the received signal.

II. LOWER BOUNDS ON THE CHANNEL CAPACITY

In information theory, channel capacity is defined as the maximum Mutual Information (MI) between the channel input and output, where the maximization is performed over all input Probability Density Functions (PDF) (Probability Mass Functions (PMF) in case of discrete input). When the input is constrained to some finite size alphabet, e.g. QAM constellation, the constrained capacity and the capacity achieving PMF are found by the Blahut-Arimoto (BA) algorithm [8] [9]. For a channel with memory, the capacity is defined as [1]:

$$C = \frac{1}{N} \max_{p(\mathbf{X}_N)} \mathcal{I}(\mathbf{X}_N; \mathbf{Y}_N) \qquad (1)$$

where $\mathbf{X}_N$ and $\mathbf{Y}_N$ are input and output vectors of length N, respectively, $\mathcal{I}(\cdot; \cdot)$ is the MI between input and output, and $p(\mathbf{X}_N)$ is the N−dimensional input PDF. Each dimension of the input vector takes a value from the signal set. For accurate MI estimation, N needs to be very large, however, even moderate values would make (1) too complex to calculate, since it requires Monte-Carlo estimation of the input-output

relation. What can be calculated instead is the MI of the averaged channel when the input symbols are independent and identically distributed (i.i.d.), leading to a lower bound on capacity. Applying average power constraint P_{av} to the input to the channel, the lower bound is [1]:

$$\hat{C} = \max_{p(X)} \mathcal{I}(X;Y) \leq C, \ \text{s.t. } \mathrm{E}\left[X^2\right] \leq P_{av} \qquad (2)$$

where X is a scalar, and $\mathrm{E}\left[\cdot\right]$ is the expectation operator.

A. MI with QAM Constrained Input

A requirement for the MI optimization algorithm is that the input-output relation of the channel is known, i.e. we have an expression for the probability $p(Y|X)$. Since this is not available, the usual practice is to numerically solve the non-linear Schrödinger equation [10] for a long sequence of input symbols, and then approximate the output density $p(Y|X)$ for each symbol of the input set with a non-circularly symmetric Gaussian distribution, with mean and variance given by the sample mean and variance. For a detailed description of estimating the output density, the reader is referred to [1]. We maximize the MI lower bound (2) w.r.t. the probabilities of the points of a fixed 1024 QAM constellation, which is referred to as *probabilistic shaping*. A single channel link is chosen for the optimization, where both polarizations were employed. The usual lumped amplification is considered, using Erbium Doped Fiber Amplifiers (EDFA). The link parameters are as follows:

- Total length : 800 km
- Fiber span length : 80 km
- EDFA noise figure : 3 dB
- Fiber dispersion : 17 ps / nm · km
- Fiber loss : 0.2 dB/km

The central wavelength is $\lambda = 1.55\ \mu m$, and the signal bandwidth is 28 GHz. The dispersion slope is assumed to be zero. We simulate 10^5 symbols for MI estimation, which makes the variations around the reported values negligible. The achieved lower bounds are given in Fig. 1. For comparison we also plot the lower bounds achieved with ring constellations, optimized as in [2]. This optimization is only performed on the radius of the first ring, and the rest of the rings are equally spaced and equally occupied. Constellations of $K = 64$ rings and maximum of 4096 and 1024 points are considered. Going beyond that makes the brute force geometric optimization of the rings challenging, and furthermore does not improve the MI significantly, when full optimization on the location of each point, the radius of each ring and the ring occupancy is not performed [2]. Probabilistic shaping on the other hand is a convex problem, and therefore tractable for large number of rings. For example 1024 QAM can be seen as a special type of ring constellation with 109 rings and unequal ring occupancy, where the probability of each ring can further be optimized through the PMF. We see in Fig. 1 that the 1024 QAM probabilistic optimization studied here can achieve and slightly exceed the lower bound on the channel capacity, set by the ring constellations from [2]. The exact optimal PMF, however, is very hard to achieve in practice. Furthermore, the optimum is
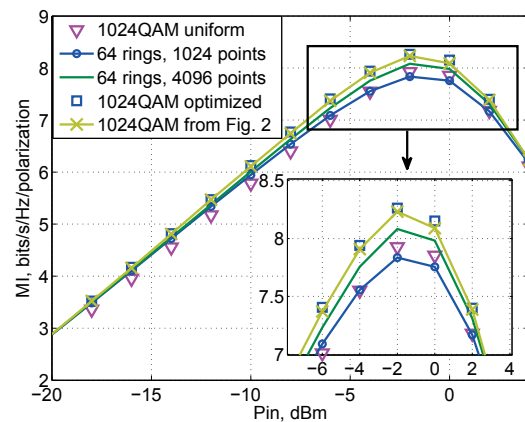

Fig. 1. MI achieved by QAM and ring constellations.

different at different input power values. Approximations are therefore needed in a flexible design. The last curve in Fig. 1 is the performance with the PMF, given in Fig. 2. A detailed description of that PMF and how it is achieved in practice is given in the next section.

III. CONSTELLATION SHAPING FOR PRAGMATIC TURBO CODES

In [11], the authors propose a dyadic PMF approximation. In a dyadic PMF all points have a probability, which can be expressed as 2^{-d}, where d is some integer. The method for achieving a dyadic PMF relies on pragmatic turbo coding. The transmitter for this method is given in Fig. 3. The data are encoded, and then serial to parallel converted. Puncturing is then applied to the parity streams, so that the remaining parity and data bits can be rearranged into m streams, m being the maximum number of coded bits per symbol. Each stream is then interleaved, and the signal is modulated. Dyadic PMF is achieved by a *many-to-one* mapping function. We found that the original design from [11] performs poorly if the required SE increases beyond a certain point [7]. This is partly due to the 256QAM size constraint in [11], and also due to the PMF and mapping not being optimized for high SNR. We therefore propose the mapping function in Fig. 2, where the achieved PMF is plotted, together with the symbol labels (we only plot a 1-D PMF and mapping, i.e. 32PAM, and the 1024QAM structure is achieved by taking a product of PAM with itself). For this mapping function $m = 8$. Each symbol can take multiple bit labels, making some symbols more likely to be produced when the input binary data are i.i.d. For example symbol '-1' takes all bit strings of length 8, which begin with the prefix '1111' (16 different bit strings in total), making its probability 2^{-4}. Symbol '-31' only takes one bit string, and its probability is therefore 2^{-8}. It is crucial that the ambiguous bit positions (marked with 'X') are taken from the parity bits of the turbo code, and the positions, where no ambiguities are found for any symbol, are associated with the systematic bits (see Fig. 3). This requirement is due to the fact, that these ambiguities can be seen as non-stationary puncturing of the bits, and the turbo codes respond better if the puncturing is performed on the parity bits, instead of the data bits. Due to space limitations, the algorithm for obtaining the mapping
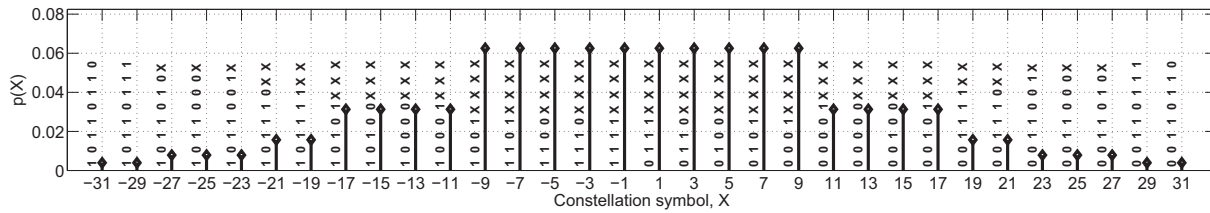
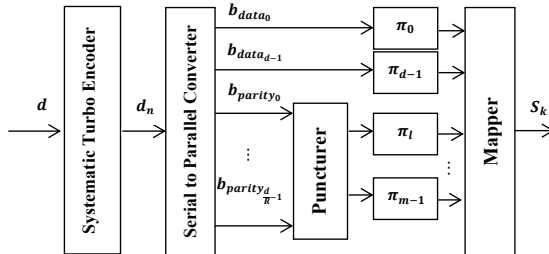Fig. 2. 32 PAM mapping function for many-to-one constellation shaping

Fig. 3. Turbo coded BICM transmitter for many-to-one constellation shaping

TABLE I
TRANSCEIVER PARAMETERS

Local oscillator linewidth	10kHz
Frequency mismatch	50MHz
ADC resolution	7 bits per real dimension
ADC sampling frequency	56GHz
Baudrate	28Gbaud
# turbo iterations	10
# de-mapping iterations	5

function is omitted here, but can be found in [7], together with a more detailed explanation and performance analysis. Here we only state that *binary reflected Gray-like* mapping structure is one of the requirements. Even though the achieved PMF is a heavily quantized version of the optimal one, we saw in Fig.1 that significant shaping gains can still be expected for wide range of input power levels.

At the receiver, iterative processing between the decoder and de-mapper is sometimes required, depending on the operating SNR, achieved PMF, etc., in order to resolve the ambiguities, induced by the many-to-one mapping, and achieve the full shaping gain.

IV. PERFORMANCE IN PRESENCE OF IMPAIRMENTS

While Fig. 1 outlines theoretically achievable rates, the main reason for analyzing QAM constellations is their practicality. In this section we evaluate the Bit Error Rate (BER) performance in the presence of various impairments, present in optical systems, such as frequency offset, local oscillator phase noise and limited Analogue to Digital Converters (ADC) precision. All of them are crucial when constellations as large as 1024 QAM must be detected. In order to evaluate the system from the previous section, simulations are made for the link from Section II, which take all of these impairments into account. The transceiver parameters are given in Table I.

Data-aided carrier synchronization and equalization are performed. To that end, QPSK pilot symbols are interleaved with the coded symbol sequence. The pilots are spread throughout the block in chunks of 20 symbols, in order to mitigate the phase noise contribution to the frequency offset. A brute force search is performed on the frequency offset, and the maximum-likelihood solution is obtained. The *constant modulus algorithm* is used for equalization, based only on the pilot sequences. The total amount of pilot overhead used is around 10%. This number can be decreased if more sophisticated synchronization algorithms are used, however, this is outside the scope of the present paper. Both polarizations are used for transmission.

In Figs. 4 and 5 BER performance is shown as a function of the input power for the link from Section II, at 6 and 6.5 bits/s/Hz/polarization input SE. We sweep the input power with a step of 0.2 dBm in all cases, and the absence of a point on a curve means that no errors were found for the duration of the simulation. A specific SE is achieved by properly selecting the code rate for each constellation. If the desired spectral efficiency is η, the code rate must be $R = \eta/m$. The turbo code we use here has two identical constituent convolutional codes with generator polynomial (23, 37) in standard octal notation, and memory $M = 4$. Code rate adjustments are performed via uniformly distributed puncturing of the turbo code. The block length used is $L = 6000$ symbols, and 50 blocks were simulated in each case, making the total number of transmitted information bits $3.6 \cdot 10^6$ in the first case, and $3.9 \cdot 10^6$ in the second. The error free region in both cases therefore corresponds to actual BER $< 2.7 \cdot 10^{-7}$ and $2.5 \cdot 10^{-7}$, respectively. As with most iterative systems, here we expect an error floor to appear somewhere below these numbers. The usual requirement for optical communication systems is BER $< 10^{-15}$. Therefore an outer code will be necessary. High rate Reed-Solomon codes, e.g. the OTN standard defined (239,255) code can bring the BER from 10^{-4} to 10^{-15} at the cost of small additional redundancy [12]. As this will be required by all modulation formats listed here, their relative performance would not change, only the actual achieved data rates will be reduced correspondingly. Simulation of this outer code is omitted in this paper.

We examine the 1024QAM from Fig. 2, the original mapping function from [11], and also a 256QAM mapping we designed as part of our previous work [7]. At $\eta = 6$ bits/s/Hz/polarization (Fig. 4), the mapping from [11] results in a rather high error floor when the non-linearities start to take effect. The error floor fluctuates due to the limited number of simulated blocks, however, the effect is still clearly visible. The 1024QAM shaped system we propose here is clearly
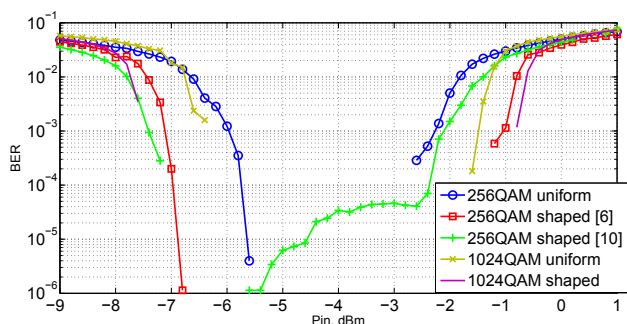
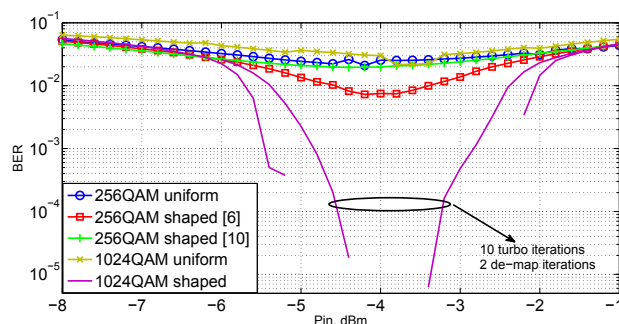Fig. 4. BER performance with 6 bits/s/Hz/polarization input data rate.

Fig. 5. BER performance with 6.5 bits/s/Hz/polarization input data rate.

outperforming the rest. All of them, however, are able to reach an "error-free" performance. When the data rate is increased to 6.5 bits/s/Hz/polarization, only the mapping from Fig. 2 has an "error-free" region, where the other constellations fail (Fig. 5). When operating at 28 Gbaud the achieved throughput is 336 and 364 Gbps, respectively. As mentioned, the pilot overhead used for synchronization in this design is around 10%, and taking into account the potential additional redundancy from the above-mentioned outer code, the actual data throughput achieved here is 283.4 and 307.1 Gbps, respectively.

Multi-channel systems are out of the scope of this paper. It is clear that in a WDM environment the XPM will dominate the SPM effects. The overall performance in that case will clearly be degraded, but whether shaping gain can still be achieved and how the XPM affects the shaped system is left as future research.

V. COMPLEXITY ANALYSIS

The main problem with running turbo coded systems at high rates is the complexity of the decoder. Extreme parallelization would be needed for the proposed design. Parallel architectures do exist for turbo codes, but a problem remains that the throughput of the decoder generally does not scale linearly with the degree of parallelization, see e.g. [13]. The proposed design would therefore require hundreds of parallel BCJR decoders, in order to achieve the above-mentioned speeds at state of the art clock frequencies. Memory is another issue for turbo decoders. The module, dominating the memory requirement is the branch metric calculator. If the accuracy in bits of the stored messages is Q, this requirement is $\propto Q \cdot L \cdot M \cdot \eta$, which is of the order of a few hundred of kilobytes for practical values of Q for each BCJR decoder. We note that all of these parameters can be optimized, in order to reduce the complexity. The number of turbo and decoding iterations can also be reduced. We found that going for example to 2 de-mapping iterations and 10 turbo iterations it is still possible to reach an "error-free" performance (Fig. 5). Even with these optimizations, implementing turbo coded schemes for optical communications remains a challenging task.

VI. CONCLUSION

In this paper constellation shaping for QAM input to a fiber optic channel was considered. We showed that when the optimization of ring constellations is restricted as in [1], QAM with probabilistic shaping is able to achieve a better lower bound on the achievable rate than the ring constellations. A QAM communication system was then simulated, employing a specific shaping method, where the performance limiting channel impairments and transceiver imperfections were taken into account. Many-to-one mapping was combined with a turbo code in order to optimize the PMF of a 1024QAM constellation. This constellation was then used to demonstrate how spectral efficiency up to 6.5 bits/s/Hz/polarization can be achieved on a 800km link, suggesting even higher efficiencies on shorter distances.

REFERENCES

[1] R. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini and B. Goebel, "Capacity Limits of Optical Fiber Networks", *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 662-701, Feb. 15, 2010

[2] T. Freckmann, R. Essiambre, P. J. Winzer, G. J. Foschini and G. Kramer, "Fiber capacity limits with optimized ring constellations", *IEEE Photon. Technol. Lett.*, vol. 21, no. 20, pp. 1496-1498, Oct. 15, 2009

[3] R. Dar, M. Shtaif and M. Feder, "New Bounds on the Capacity of the Non-linear Fiber-optic Channel", *Optics Letters*, Vol. 39, No. 2, pp. 398-401, Jan. 15, 2014

[4] L. Beygi, E. Agrell, J. M. Kahn and M. Karlsson, "Rate-Adaptive Coded Modulation for Fiber-Optic Communications", *Journal of Lightwave Technology*, vol. 32, no. 2, pp. 333-343, Jan. 15, 2014

[5] B. P. Smith and F. R. Kschischang, "A Pragmatic Coded Modulation Scheme for High-Spectral-Efficiency Fiber-Optic Communications", *Journal of Lightwave Technology*, vol. 30, no. 13, pp. 1-7, Jul. 1, 2012

[6] T. H. Lotz, X.Liu, S. Chandrasekhar, P. J. Winzer, H. Haunstein, S. Randel, S. Corteselli, B. Zhu, and D. W. Peckham, "Coded PDM-OFDM Transmission With Shaped 256-Iterative-Polar-Modulation Achieving 11.15-b/s/Hz Intrachannel Spectral Efficiency and 800-km Reach", *Journal of Lightwave Technology*, vol. 31, no. 4, pp. 538-545, Feb. 15, 2013

[7] M. Yankov, S. Forchhammer, K. J. Larsen and L. P. B. Christensen, "Rate-adaptive Constellation Shaping for Near-capacity Achieving Turbo Coded BICM", *ICC proceedings*, pp. 2118-2123, 2014

[8] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, New York, Wiley, 1991.

[9] N. Varnica, X. Ma and A. Kavcic, "Capacity of power constrained memoryless AWGN channels with fixed input constellations," *GLOBECOM proceedings*, pp. 1339-1343 vol.2, 2002.

[10] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. San Diego, CA: Elsevier, 2006.

[11] D. Raphaeli and A. Gurevitz, "Constellation shaping for pragmatic turbo-coded modulation with high spectral efficiency," *IEEE Trans. Commun.*, vol. 52, no. 3, pp. 341-345, 2004.

[12] A. Rasmussen, M. Yankov, M. S. Berger, K. J. Larsen and S. Ruepp, "Improved Energy Efficiency for Optical Transport Networks by Elastic Forward Error Correction," *Journal of Optical Communications and Networking*, vol. 6, no. 4, pp. 397-407, 2014.

[13] C. Roth, S. Belfanti, C. Benkeser and H. Qiuting, "Efficient Parallel Turbo-Decoding for High-Throughput Wireless Systems," *IEEE Transactions on Circuits and Systems*, vol. 61, no. 6, pp. 1824-1835, 2014.