Mitigation of nonlinear transmission effects for OFDM 16-QAM optical signal using adaptive modulation

ANTON S. SKIDIN,[1,2] OLEG S. SIDELNIKOV,[1,2] MIKHAIL P. FEDORUK,[1,2] AND SERGEI K. TURITSYN[1,3,*]

[1]*Novosibirsk State University, Novosibirsk, 2 Pirogova str., 630090 Russia*
[2]*Institute of Computational Technologies SB RAS, Novosibirsk, 6 acad. Lavrentjev av., 630090 Russia*
[3]*Aston Institute of Photonic Technologies, Aston University, B4 7ET Birmingham, UK*
[*]*s.k.turitsyn@aston.ac.uk*

Abstract: The impact of the fiber Kerr effect on error statistics in the nonlinear (high power) transmission of the OFDM 16-QAM signal over a 2000 km EDFA-based link is examined. We observed and quantified the difference in the error statistics for constellation points located at three power-defined rings. Theoretical analysis of a trade-off between redundancy and error rate reduction using probabilistic coding of three constellation power rings decreasing the symbol-error rate of OFDM 16-QAM signal is presented. Based on this analysis, we propose to mitigate the nonlinear impairments using the adaptive modulation technique applied to the OFDM 16-QAM signal. We demonstrate through numerical modelling the system performance improvement by the adaptive modulation for the large number of OFDM subcarriers (more than 100). We also show that a similar technique can be applied to single carrier transmission.

References and links

1. C. E. Shannon, "A mathematical theory of communication," Bell Syst. Tech. J. **27**, 623–656 (1948).
2. A. Splett, C. Kurtzke, and K. Petermann, "Ultimate transmission capacity of amplified fiber communication systems taking into account fiber nonlinearities," Proc. of 19th European Conference on Optical Communication (ECOC), MoC2.4 (1993).
3. A. D. Ellis, Z. Jian, and D. Cotter, "Approaching the non-linear Shannon limit," J. Lightwave Technol. **28**(4), 423–433 (2010).
4. D. J. Richardson, "Filing the Light Pipe," Science **330**(6002), 327–328 (2010).
5. E. Temprana, E. Myslivets, B. P.-P. Kuo, V. Ataie, N. Alic, and S. Radic, "Overcoming Kerr-induced capacity limit in optical fiber transmission," Science **348**(6242), 1445–1448 (2015).
6. P. J. Winzer, "Scaling Optical Fiber Networks: Challenges and Solutions," Opt. Photon. News **26**, 28–35 (2015).
7. R. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," J. Lightwave Technol. **28**(4), 662–701 (2010).
8. E. Agrell, G. Durisi, and P. Johannisson, "Information-theory-friendly models for fiberoptic channels: A primer," IEEE Information Theory Workshop (2015).
9. E. Agrell, A. Alvarado, G. Durisi, and M. Karlsson "Capacity of a nonlinear optical channel with finite memory," J. Lightwave Technol. **16**, 2862–2876 (2014).
10. B. P. Smith and F. R. Kschischang, "A pragmatic coded modulation scheme for high-spectral-efficiency fiber-optic communications," J. Lightwave Technol. **30**(13), 2047–2053 (2012).
11. L. Beygi, E. Agrell, J. M. Kahn, and M. Karlsson, "Rate-adaptive coded modulation for fiber-optic communications," J. Lightwave Technol. **32**(2), pp. 333–343 (2014).
12. M. P. Yankov, D. Zibar, K. J. Larsen, L. P. B. Christensen, "Constellation shaping for fiber-optic channels with QAM and high spectral efficiency," IEEE Photon. Technol. Lett. **26**(23), 2407–2410 (2014).
13. T. Fehenberger, G. Böcherer, A. Alvarado, and N. Hanik, "LDPC coded modulation with probabilistic shaping for optical fiber systems," Proc. of Optical Fiber Communication Conference (OFC), Th.2.A.23 (2015).
14. T. Fehenberger, D. Lavery, R. Maher, A. Alvarado, P. Bayvel, N. Hanik, "Sensitivity gains by mismatched probabilistic shaping for optical communication systems," IEEE Photon. Technol. Lett. **28**(7), 786–789 (2016).
15. F. Buchali, F. Steiner, G. Bocherer, L. Schmalen, P. Schulte, and W. Idler, "Rate adaptation and reach increase by probabilistically shaped 64- QAM: An experimental demonstration," J. Lightwave Technol. **34**(7), 1599–1609 (2016).
16. C. Diniz, J. H. Junior, A. Souza, T. Lima, R. Lopes, S. Rossi, M. Garrich, J. D. Reis, D. Arantes, J. Oliveira, and D. A. Mello, "Network cost savings enabled by probabilistic shaping in DP-16QAM 200-Gb/s systems," Proc. Optical

Fiber Communication Conference (OFC), Tu3F.7, (2016).

17. A. Shafarenko, A. Skidin, and S. K. Turitsyn, "Weakly-constrained codes for suppression of patterning effects in digital communications," IEEE Trans. Commun. **58**(10), 2845–2854 (2010).

18. A. Shafarenko, K. S. Turitsyn, S. K. Turitsyn, "Information-theory analysis of skewed coding for suppression of pattern-dependent errors in digital communications," IEEE Trans. Commun. **55**(2), 237–241 (2007).

19. A. Alvarado, E. Agrell, D. Lavery, R. Maher, and P. Bayvel, "Replacing the Soft-Decision FEC Limit Paradigm in the Design of Optical Communication Systems," J. Lightwave Technol. **33**(20), 4338–4352 (2015).

20. B. Djordjevic, and B. Vasic, "Nonlinear BCJR equalizer for suppression of intrachannel nonlinearities in 40 Gb/s optical communications systems," Opt. Express **14**, 4625–4635 (2006).

21. N. Kashyap, P. H. Siegel, and A. Vardy, "Coding for the optical channel: the ghost-pulse constraint," IEEE Trans. Inf. Theory **52**(1), 64–77 (2006).

22. S. K. Turitsyn, M. P. Fedoruk, O. V. Shtyrina, A. V. Yakasov, A. Shafarenko, S. R. Desbruslais, K. Reynolds, and R. Webb, "Patterning effects in a WDM RZ-DBPSK SMF/DCF optical transmission at 40Gbit/s channel rate," Opt. Commun. **277**(2), 264–268 (2007).

23. B. Slater, S. Boscolo, A. Shafarenko, and S. K. Turitsyn, "Mitigation of patterning effects at 40 Gbits/s by skewed channel pre-encoding," J. Opt. Netw. **6**(8), 984–990 (2007).

24. S. T. Le, M. E. McCarthy, S. K. Turitsyn, "Optimized hybrid QPSK/8QAM for CO-OFDM transmissions," Proc. of 9th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP), 763–766 (2014).

25. X. Zhou, L. E. Nelson, P. Magill, R. Isaac, B. Zhu, D. W. Peckham, P. I. Borel, and K. Carlson, "High Spectral Efficiency 400 Gb/s Transmission Using PDM Time-Domain Hybrid 32-64 QAM and Training-Assisted Carrier Recovery," J. Lightwave Technol. **31**(7), 999–1005 (2013).

26. S. O. Zafra, X. Pang, G. Jacobsen, S. Popov, S. Sergeyev, "Phase noise tolerance study in coherent optical circular QAM transmissions with Viterbi-Viterbi carrier phase estimation," Opt. Express **22**(25), 30579–30585 (2014).

27. P. J. Winzer, "High-spectral-efficiency optical modulation formats," J. Lightwave Technol. **30**(24), 3824–3835 (2012).

28. C. Pan and F. R. Kschischang, "Probabilistic 16-QAM Shaping in WDM Systems," J. Lightwave Technol. **34**(18), 4285–4292 (2016).

1. Introduction

In modern optical fiber links the nonlinear transmission effects are one of the major factors limiting system performance. As opposed to linear channels, where performance degradation due to noise can be mitigated by a signal power increase [1], in optical fiber communications the increased signal power leads to new (nonlinear) sources of distortions and loss of information (see e.g. [2–7] and numerous references therein). Exploitation of modern communication systems using a large number of WDM-channels (or super-channels) assumes an increase of the total signal power in the fiber leading to a growing impact of nonlinear transmission effects. The operation of optical communication systems in such nonlinear regimes is rather different from the conventional lower power mode. This calls for the development of new approaches and techniques to better understand the peculiarities of high signal power transmission regimes and the root cause of errors.

From the view point of classical information theory, the important challenge is to evaluate the Shannon capacity of optical communication channels and to find the capacity-achieving input signal distribution [1, 7–9]. Most optical systems currently in operation use the uniform input signal distribution with relatively simple alphabets. The performance of such systems can be improved by using more advanced input signal distributions, though this might require additional complexity of the transceivers and receivers. For instance, improvement can be achieved by modifying the shape of the transmitted signal by changing the constellations or by using non-uniform distribution for the occurrence probability of the symbols in a pre-selected constellation. These two kinds of signal shaping are often distinguished as geometric and probabilistic shaping [10–16]. Geometric shaping corresponds to non-uniformly distributed constellation points with equiprobable symbols, while probabilistic shaping corresponds to standard uniform constellations with varying probabilities of constellation points. In optical communication probabilistic shaping can be used to reduce the frequency of occurrence of high-power symbols in order to suppress nonlinear transmission effects. Probabilistic shaping can be performed using modified forward-error correction (FEC) codes [12–14]. Forward error

correction and modulation methods can be jointly employed to improve practically achievable rates.

The powerful FEC techniques are critically important in the modern optical communications systems for the provision of an ultra-low bit-error rate (BER) requirement. FEC allows optical engineers to operate systems at relatively high BER that after FEC decoding is improved to $BER = 10^{-12}$ or even $BER = 10^{-15}$. Most of the FEC techniques are developed for the memoryless and linear additive white Gaussian noise channel. A nonlinear fiber channel is not Gaussian and it has memory that results in inter-symbol interference and patterning effects. At low BER levels the FEC methods can cope with errors due to both the channel noise and patterning effects. However, as the inter-symbol interference (and corresponding patterning effects) grows stronger, at some point (BER threshold) the FEC scheme starts to deteriorate fast and it is at that vital point that error prevention becomes an essential issue in maintaining the low-BER operational regimes. Operation close to the BER threshold is not desirable and any additional margins on top of the FEC are vitally important. The application of weakly-constrained codes (skewed coding) [17,18] was proposed to provide extra margin in addition to FEC. In this approach a weakly constrained code is employed to decrease the frequency of occurrence of undesirable patterns in order to reduce the part of the BER that occurs due to inter-symbol interference and, consequently, bring the combined BER back under the FEC break-down threshold. Since the FEC threshold is typically very sharply defined (see an important discussion concerning the BER threshold in [19]), the most economical pre-encoding scheme has to be tuneable: any extra redundancy below the threshold is more effective when it is utilized by the FEC itself. Weakly constrained codes decrease the frequency of occurrence of patterns of a certain types. The amount of reduction is defined by a trade-off with the code redundancy and is controllable by a parameter that can be varied almost continuously. This makes this technique [17,18] ideally suited for the control of patterning effects reduction for the purposes of FEC operation close to the BER threshold.

In this work we examine the power-affected error statistics for both the single-carrier 16-QAM and for the multiple-carrier OFDM 16-QAM signal and explore the possibilities to mitigate the nonlinear effects at high signal powers through specific modulation and coding approaches. We propose and apply here the adaptive modulation technique that aims to reduce the impact of the nonlinear transmission effects at high signal powers, when statistics of errors are affected by nonlinear interactions. Our approach has some similarities with the use of weakly-constrained stream and block codes with tunable pattern-dependent statistics [17, 18, 20–23], rate-adaptive coded modulation [10, 11], the hybrid QAM [24, 25], and probabilistic signal shaping using low-parity-density codes [12–14]. Our focus here is on the relatively high power signal regimes that substantially affect error statistics. We demonstrate the feasibility of the technique and quantify the improvement for the OFDM 16-QAM modulation format. We also analyze the trade-off between the reduction of symbol error rates and redundancy due to adaptive signal modulation.

2. Simulation set-up and system parameters

To examine the impact of nonlinear effects on error statistics we consider two types of systems: the single carrier 16-QAM and the OFDM 16-QAM with a varying number of subcarriers. We study the transmission link shown in Fig. 1. Each span of the transmission system includes the standard-monomode fiber (SMF) of 100 km and the Erbium-doped fiber amplifier (EDFA) that exactly compensates the signal power attenuation in the fiber preceding the amplifier. The transmitter generates the 16-QAM signal with the baud rate R_s; the number of samples per one OFDM symbol is 16 (i.e., the sampling rate is $16R_s$). For pulse shaping the raised-cosine filter is employed. The noise loading is performed after each amplifier such that it corresponds to the amplified spontaneous emission added by EDFA. The chromatic dispersion of the link is

compensated at the receiver before signal processing is used to recover the signal phase. The fiber links consisting of 10, 15, and 20 spans are examined.
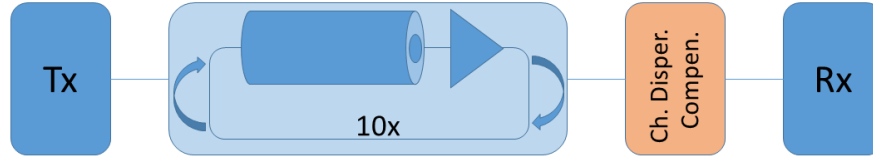

Fig. 1. The scheme of the 1000 km fiber link under consideration.

The propagation of a signal along the fiber span is modelled by the standard nonlinear Schrödinger equation (NLSE):

$$\frac{\partial A}{\partial z} = -\frac{\alpha}{2}A - i\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} + i\gamma|A|^2 A,$$

where $A(z,t)$ is the slowly-varying envelope of a signal. The NLSE is solved numerically using the well-known split-step Fourier method (SSFM). The following parameters are used in the numerical simulations: the over-sampling factor $q = 16$, the number of symbols $N_S = 2^{18}$, fiber losses $\alpha = 0.2$ dB/km, the fiber nonlinearity coefficient $\gamma = 1.4$ W^{-1}km^{-1}, the chromatic dispersion $\beta_2 = -25$ ps^2/km, the signal wavelength $\lambda = 1.55$ nm, the amplifier noise figure is $NF = 4.5$.

Using numerical modelling the dependence of BER on input signal power (without using any coding) has been computed for a varying number of spans, as is depicted in Fig. 2. It can be seen, that the optimal power is about 3 dBm.
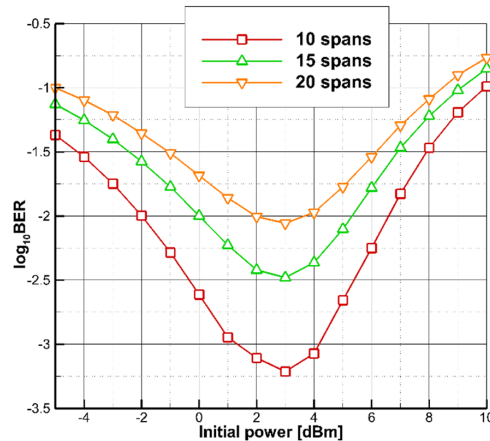

Fig. 2. The bit-error rate (BER) as a function of the initial signal power.

3. Nonlinear distortion of the 16-QAM signal

Next we numerically examine the error statistics in highly nonlinear regimes. To yield statistically significant results, we have performed 100 runs (with different noise realisations) with 2^{18} 16-QAM OFDM symbols in each run. Thus, the total number of transmitted 16-QAM symbols is $100 \cdot 2^{18} \cdot K = 2.62 \cdot 10^7 \cdot K$, where K - number of modulated subcarriers. The propagation

distance was 1000 km (i.e. 10-span link). Symbol time interval was varying with the number of subcarriers following the relation $T_s = K/BW$ with $BW = 100$ GHz, i.e. $T_s = 1$ ns for $K = 100$. We would like to stress that here we did not use the standard signal optimization over power to choose the best operational point with the lowest bit-error rate, as shown in Fig. 2. Instead, the initial power of a signal is chosen from Fig. 2 to have bit-error rate level $BER = 10^{-2}$ at high powers. This makes it possible to study error statistics in a somewhat artificially created "nonlinear" regime, where the signal power is higher than the optimal one that is around 3 dBm. Also the threshold of $BER = 10^{-2}$ is chosen as the level from which we aim to reduce using the proposed approach BER down to the hard FEC threshold [19], where errors can be corrected using the forward-error correction methods.

Figure 3 illustrates the dependence between the symbol-error rate (SER) and the baud rate for the nonlinear regime in case of a single carrier ($K = 1$) 16-QAM transmission. All the errors are divided into three categories by the powers of constellation points and corresponding three power rings. These rings are shown as dashed circles in Fig. 4. In these simulations all the constellation points have equal probabilities, this corresponds to uniform input signal distribution.
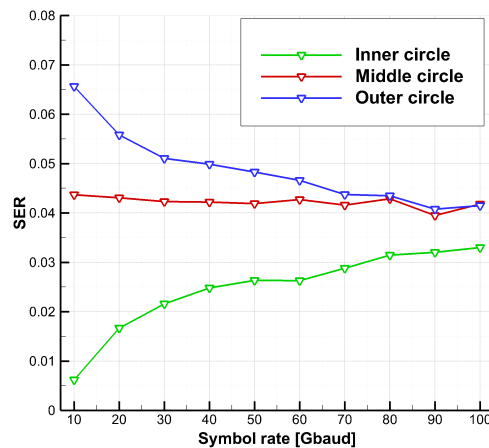

Fig. 3. SER as a function of the baud rate for the single-carrier transmission.

As can be seen in Fig. 3, as expected, the outer ring is the most error-prone, followed by the middle ring. This can be qualitatively understood through the observation that in the outer ring the high-power symbols are more affected by the nonlinear effects. When the baud rate grows, the difference in errors between rings becomes smaller due to spreading pulses and an effective averaging over the data stream.

Next, we simulate the propagation of an OFDM 16-QAM signal in the same fiber link depicted in Fig. 1. Again, here all the constellation points have equal probabilities. The only difference is that the OFDM modulator and demodulator are used as the transmitter and the receiver, respectively. For the OFDM modulator we use the following parameters: the maximum number of subcarriers is 1024, the bandwidth is of $BW = 100$ GHz. The actual number of the modulated subcarriers K is varied in this case.

Figure 5 shows how SER for different rings depends on the number of modulated OFDM-subcarriers. It looks similar to Fig. 3: for the relatively small number of subcarriers the distinction is significant between the different rings. When the number of subcarriers increases, the difference disappears.

We would like to stress again, that the results of massive numerical modelling presented in Figs. 3 and 5 should be understood not as optimization modelling. In all points presented in these figures we choose power from the target condition of having BER=0.01 in the nonlinear regime.
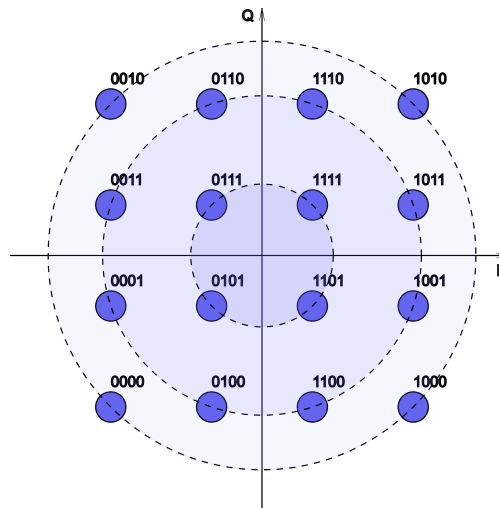
Fig. 4. The constellation points of "squared" 16-QAM format and three corresponding power rings.
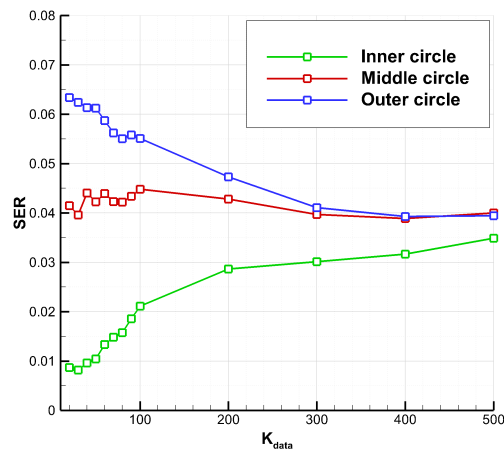

Fig. 5. SER as a function of the number of subcarriers of the OFDM-signal.

This gives us a possibility to analyze the error statistics in such highly nonlinear transmission. One can see that for a single carrier the main difference between error probabilities of rings is observed at lower baud rates. In the case of the OFDM 16-QAM the greater variations of the error statistics are for a lower number of subcarriers. The observed asymmetry in the error probabilities for different rings calls for applications of constrained coding to improve system performance operating near the FEC BER threshold.

4. Theoretical analysis

The 16-QAM constellation points can be divided into three sets with different powers, i.e. three power rings. For the sake of clarity, we enumerate the constellation sets in the ascending order of amplitude of the points they consist of (i.e. the first set consists of the points that belong to the "inner" ring on the constellation diagram, the second set contains the points from the "middle" ring, and finally the third set includes the points from the "outer" ring). This can easily be seen in

Fig. 4, $s_1 = 4$, $s_2 = 8$, $s_3 = 4$, where s_i is the number of constellation points in the ith set.

It should be noted that the constellation points of a 16-QAM modulation format can be put on the phase plain in different ways (see e.g. discussions in [26, 27]). For instance, the constellation points of 16-QAM formats that are widely employed in practice, form either a square or a circle. Below we consider only the "squared" 16-QAM modulation format. However, the theoretical approach proposed here can be applied to any modulation format irrespective of the way the constellation points are arranged on a phase plain.

To estimate the impact of the nonlinear effects on a QAM-modulated optical signal we assume that the error rate of a symbol depends only on its power. Below the error rate of a symbol from the i-th set is denoted by q_i, and the probability of a symbol from the i-th set to appear in a data stream is denoted by P_i.

The symbol error rate (SER) in a data stream can be found as follows:

$$SER(P_1, P_2, P_3) = P_1 \cdot q_1 + P_2 \cdot q_2 + P_3 \cdot q_3. \tag{1}$$

Since $P_3 = 1 - P_1 - P_2$, the SER value effectively depends only on two unknown probabilities.

Particularly if $q_1 = q_2 = q_3 = q$, formula (1) becomes trivial, and $SER = q$. This is the case, when the error rate does not depend on the signal power as it is observed, for example, in linear or effectively linear channels. As we have shown above, in a nonlinear optical communication channel the error probabilities for various symbol sets differ from each other as a result of the impact of the nonlinear Kerr effect.

Our goal is to reduce the number of errors by varying the probabilities that the symbols from the i-th set appear in a data stream. This process can generally be referred to as the adaptive modulation. It is also sometimes named the hybrid QAM modulation format [24, 25]. In this consideration we prefer the term "adaptive modulation", because the proposed approach can be used with any modulation format, not only with QAM. Our approach, changing the input signal distribution, is a version of the probabilistic shaping technique (see recent publication [28] and references therein). The advantage of the proposed adaptive modulation is the use of the error statistics to modify in a flexible and adaptive way probabilities of occurrences of symbols. The use of the detailed error statistics enables to take into account subtle difference in the transmission of various constellation symbols that, in turn, allows to mitigate the channel impairments with a relative small redundancy.

The adaptive modulation is that the symbols at various positions in a data stream are modulated by different "virtual" modulation rules derived from the original modulation format simply by excluding the constellation points that are more prone to the nonlinear induced errors. Here the symbol position means its position in time (in the data stream). For example, odd symbols can be modulated using only the four symbols from the "inner" 16-QAM ring, and the even symbols can be modulated using the 16-QAM format itself, without any restriction. Of course, this example is mentioned for illustrative purposes only, and the adaptive modulation scheme that enables the system performance to be improved is in general more complex.

It should be noted that when we vary the probabilities that the symbols from the i-th class appear in a data stream, we, in general, reduce the information entropy of a data stream and, in turn, increase the redundancy of a transmitting message. This results in the reduction of an actual channel rate. The entropy of a data stream per one symbol can be found as follows:

$$H(\overrightarrow{P}) = -p_1 \log p_1 - p_2 \log p_2 - ... - p_{16} \log p_{16}, \tag{2}$$

where $\log x = \log_{16} x$, p_j is the probability of 16-QAM symbol j ($j = 1, 2, ..., 16$). Since it is assumed in our consideration that the error rate of a 16-QAM symbol depends only on its power, then $p_j = p_k$ if symbols j and k belong to the same set i, thus $P_i = s_i p_j$. Consequently, for the 16-QAM format the information entropy $H(P_1, P_2)$ can be expressed using the following

equation:

$$H(P_1, P_2) = -P_1 \log \frac{P_1}{4} - P_2 \log \frac{P_2}{8} - \\ (1 - P_1 - P_2) \log \frac{1 - P_1 - P_2}{4} \tag{3}$$

5. Reduction of SER through adaptive modulation of the 16-QAM channel

As was explained above, the nonlinear effects might result in the dependence of the symbol error rate on the symbol power. This gives a possibility to reduce the symbol error rate by means of the constrained encoding reducing the number of error prone symbols. Encoding is understood here in a broad sense, as a method to process and alter the data to be transmitted, regardless of the implementation details.

From Eq. (1) it can be derived that the initial symbol error rate (i.e. the symbol error rate before any encoding is applied) is as follows:

$$SER_0 = \frac{1}{4} \cdot q_1 + \frac{1}{2} \cdot q_2 + \frac{1}{4} \cdot q_3. \tag{4}$$

In general the encoded signal has a different symbol error rate SER_C.

Our goal is to find the input signal distribution probability vector $\vec{P} = (P_1, P_2, P_3)$ that minimizes the symbol-error rate, for a given set $\vec{q} = (q_1, q_2, q_3)$ (i.e. for a given error distribution across 16-QAM constellation rings), and for a given entropy $0 \le H_0 \le 1$, i.e. for a given code rate $C_0 = 1 - H_0$. This allows to evaluate a trade-off between system performance improvement and data redundancy. The problem can be solved by using the Lagrange multipliers method.

Let us consider the Lagrange function

$$L(P_1, P_2, \lambda) = SER(P_1, P_2) + \lambda \cdot (H(P_1, P_2) - H_0). \tag{5}$$

We assume that q_is in function $SER(P_1, P_2)$ are not equal to each other. This assumption does not imply the loss of generality, however it makes analysis easier.

The stationary points of function (5) can be found by solving the following system of equations:

$$\frac{\partial L}{\partial P_1} = q_1 - q_3 + \lambda \frac{\partial H(P_1, P_2)}{\partial P_1} = 0 \\ \frac{\partial L}{\partial P_2} = q_2 - q_3 + \lambda \frac{\partial H(P_1, P_2)}{\partial P_2} = 0 \\ H(P_1, P_2) = H_0. \tag{6}$$

One can establish that $\frac{\partial H(P_1, P_2)}{\partial P_1} = \log \frac{1 - P_1 - P_2}{P_1}$, and $\frac{\partial H(P_1, P_2)}{\partial P_2} = \log \frac{2(1 - P_1 - P_2)}{P_2} = 1/4 + \log \frac{1 - P_1 - P_2}{P_2}$. From the first two equations of system (6) one can find the relationship between P_1 and P_2:

$$\frac{1 - P_1 - P_2}{P_1} = 2^\alpha + \left(\frac{1 - P_1 - P_2}{P_2}\right)^\alpha, \tag{7}$$

where $\alpha = \frac{q_3 - q_1}{q_3 - q_2}$. Since $q_1 \ne q_2 \ne q_3$ (as it was assumed above), we do not have divergence of α. If $\alpha \le 0$, it can be shown that equation (7) has a single root $P_2(P_1)$ for any fixed P_1; on the other hand, if $\alpha > 0$, there exists only one root $P_1(P_2)$ for any fixed P_2. Given this, the dependence between P_1 and P_2 can be quickly estimated numerically without the need for an exhaustive search.

Any solution of equation (7) yields the stationary point of function (5), if $H(P_1, P_2) = H_0$ is met. Thus, the SER minimum value can be found by substituting the stationary points into equation (1). However, it can be derived that the stationary points of function (5) are always of the same type, i.e. they are all either the minimum or the maximum points. This is because the

sign of a second differential strongly depends on the sign of λ that can generally take both the negative and positive values depending on the values q_1, q_2, and q_3.

It is noteworthy that the same approach can be used for other modulation formats with a constellation diagram consisting of many distinct power levels. In such a case, the Lagrange function would look like this: $L(P_1, P_2, ..., P_N, \lambda) = SER(P_1, P_2, ..., P_N) + \lambda \cdot (H(P_1, P_2, ..., P_N) - H_0)$, and the entropy $H(P_1, P_2, ..., P_N) = -\sum P_i \log(c_i \cdot P_i)$, i.e. in mathematical sense these formulae are almost identical to equations (5) and (3), respectively.

6. Adaptive modulation scheme

To illustrate how the theoretical optimization can be practically utilised, we use the simple adaptive modulation scheme in which different time slots are modulated using different modulation formats that include both the 16-QAM format itself, and the "restricted" modulation formats that are obtained from the 16-QAM format as shown in Fig. 6. In this scheme, the number of symbols that use a specific modulation pattern may vary according to the desired distribution of the 16-QAM symbols in a resulting data stream.

Unlike the probabilistic shaping method [11, 12], the proposed technique improves the data transmission by varying not only the average power of a signal, but also the average distance between different constellation points. The latter also affects the symbol error rate, because the smaller average distance between adjacent points on a constellation diagram (i.e. the smaller average power of a signal) results in an increased number of errors due to the "linear" noise. On the contrary, the large distance between constellation points means that the main cause of the errors is the prevalence of nonlinear effects.
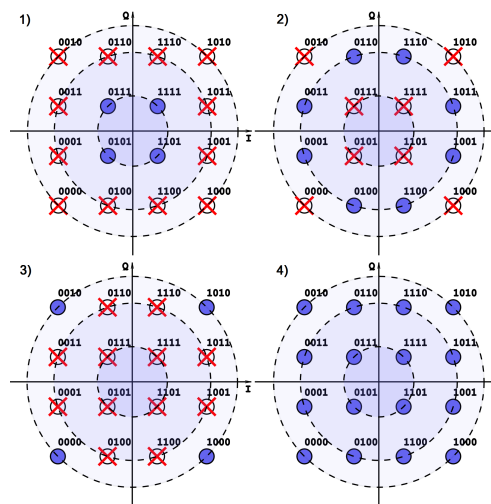

Fig. 6. The constellation diagrams of the 16-QAM-based modulation pattern set: diagrams 1) and 2) transmit two bits per one time slot, 3) transmits three bits, 4) is the original 16-QAM format.

The adjustable SER reduction can be achieved by using the block-based approach to produce adaptively modulated data. To accomplish this, the output data stream is divided into separate data blocks of length N symbols, where i-th symbol in a data block, $i = 1, .., N$, is modulated by a deliberately selected modulation pattern with the number m_i. These patterns are selected out of the patterns shown in Fig. 6. Obviously, $m_i \in \{1, 2, 3, 4\}$.

Let us denote by C the desirable code rate of our adaptive modulation scheme. It can be

selected in such a way as to obtain the desired SER. That is, the code rate is treated as one of the input parameters for the adaptive modulation scheme [10, 11]. Another input parameter is the optimal probability vector $\overrightarrow{P}$.

Denote by n_i the number of symbols in a data block that use the i-th modulation pattern from Fig. 6. Obviously, $n_1 + n_2 + n_3 + n_4 = N$, and $0 \leq n_i \leq N$. It can easily be seen that

$$C = C(N) = \sum n_i c_i / 4N,$$

where c_i is the number of bits that the i-th modulation pattern conveys. Since the number of constellation points used in one data block is proportional to N, the capacity of such scheme $C_\infty = \lim_{N \to \infty} C(N) = C(N_0)$ for any N_0 that is a multiple of N ($N_0 = k \cdot N$, $k \geq 1$).

The values n_i for a given $\overrightarrow{P}$ can be obtained by solving the following linear system:

$$
\begin{aligned}
2n_1 + 3n_2 + 2n_3 + 4n_4 &= 4NC \\
\frac{n_1 + n_4/4}{N} &= P_1 \\
\frac{n_2 + n_4/2}{N} &= P_2 \\
\frac{n_3 + n_4/4}{N} &= P_3 = 1 - P_1 - P_2.
\end{aligned}
\tag{8}
$$

It can be found that

$$
\begin{aligned}
n_4 &= \frac{2N}{3}(4C - P_2 - 2) \\
n_3 &= \frac{N}{3}\left(4 - 2C - 3P_1 - \frac{5}{2}P_2\right) \\
n_2 &= \frac{N}{3}(2 + 4P_2 - 4C) \\
n_1 &= \frac{N}{6}(2 + 6P_1 + P_2 - 4C).
\end{aligned}
\tag{9}
$$

The system of equations (9) can be solved if the right-hand sides of equations (9) are positive. However, if we deal with a particular set of probabilities $\overrightarrow{P}$, it cannot be expected that this requirement is met for any $0 \leq C \leq 1$. In fact, this means there are the probability vectors $\overrightarrow{P}$ for which it is impossible to build a code of the desirable code rate. In this case it is necessary either to obtain the values n_i that give the probability distribution close to $\overrightarrow{P}$, or to vary the code rate in order to make system (9) consistent.

Note that though the presented theoretical results are strict and give exact trade-off between redundancy and improvement of performance in systems with symbol dependent errors, this simple theory is not applied directly to optical fiber systems because statistics of errors is also affected by the change of probabilities of different input signal power (rings). This can be taken into account, but consideration of this effect is beyond the scope of our current paper. Here, instead, we use analytical results only as a qualitative guidance in the direct numerical optimization of system performance.

7. Numerical modeling results and discussion

In this section we apply the adaptive modulation in order to reduce the nonlinear transmission impairments in the OFDM-system employing the 16-QAM modulation format. Denote by κ the target SER reduction rate ($0 \leq \kappa \leq 1$) that is defined as follows:

$$\kappa = \frac{SER_C}{SER_0}. \tag{10}$$

This coefficient can be treated as a measure of the encoding performance. Evidently, there is a trade-off between reduction in SER and redundancy in the data stream required to implement such coding.

The performance of the nonlinear transmission systems is defined by the interplay between the effects of noise and nonlinear effects on the signal. Optical signal power always corresponds to the minimal BER.

To estimate the efficiency of the adaptive modulation in a practical implementation, we consider the signal propagation after 2000 km. We have selected the transmission distance in such a way as to reach the bit-error rate close to the forward-error correction limit. Currently this value lies in the range between $5 \cdot 10^{-3}$ to 10^{-2} [19], depending on the error correction code. From Fig. 2 we have found the transmission distance where the minimum BER is about 10^{-2}. The signal power is set to an optimal one, i.e. $P_{in} = 3$ dBm. After transmitting the signal, we have obtained that $q_1 = 0.030$, $q_2 = 0.037$, $q_3 = 0.035$. At first glance, there is no significant difference between error rates from the various QAM modulation "rings". However, when applying the adaptive modulation, it turns out that even a small skew in the error rates yields a significant symbol-error rate improvement. Figure 7(a) shows that the number of errors can be reduced by half at the cost of 12% redundancy. These results are averaged over 100 numerical runs with different noise realisations. In Fig. 7(b) the mutual information dependence on the adaptive modulation redundancy is shown. As it is expected, the mutual information gradually decreases as the redundancy grows. However, for small values of redundancy (less than 5%) the mutual information falls slowly compared to the reduction of the actual code rate. Consequently, the low-redundant adaptive modulation can be the optimal choice for the systems where the bit-error rate of the QAM-modulated signal is near to the FEC code limit.
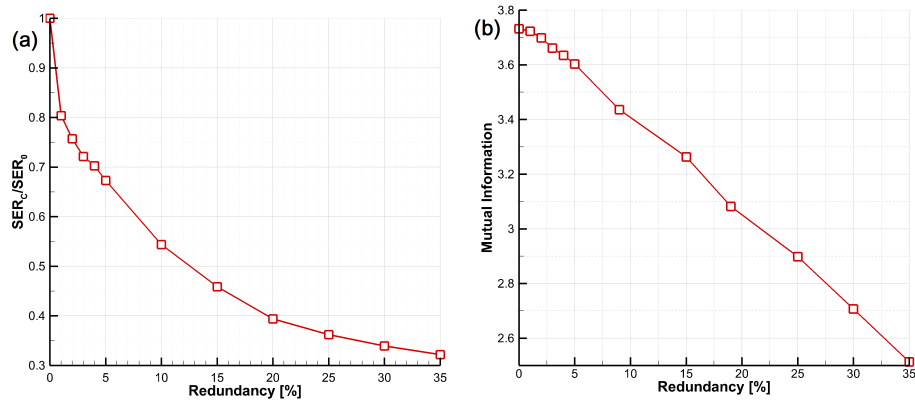

Fig. 7. (a) The SER improvement rate $\kappa = SER_0/SER$ as a function of the code redundancy, (b) the mutual information as a function of the code redundancy.

Figure 8 shows the possible BER improvement as a function of the signal power. It can be seen that even if the redundancy is relatively small, the bit-error rate can be reduced significantly. It should also be noted that the optimal power gradually increases as the redundancy grows. Thus, the adaptive modulation makes it possible to effectively use large signal powers. As it can be seen from Fig. 8 (red curve), the BER improvement from $BER = 10^{-2}$ to $BER = 10^{-2.5}$ can be achieved using the 12%-redundant adaptive modulation. This enables to apply the FEC encoding with an overhead from 5 to 12% to the adaptively modulated data. The main difference between $BER = 10^{-2}$ and $BER = 10^{-2.5} = 3 \cdot 10^{-3}$ is that for small BERs (below $5 \cdot 10^{-3}$), any modern FEC code is able to reduce BER to 10^{-9} and less. For larger BERs (especially for $BER > 10^{-2}$),
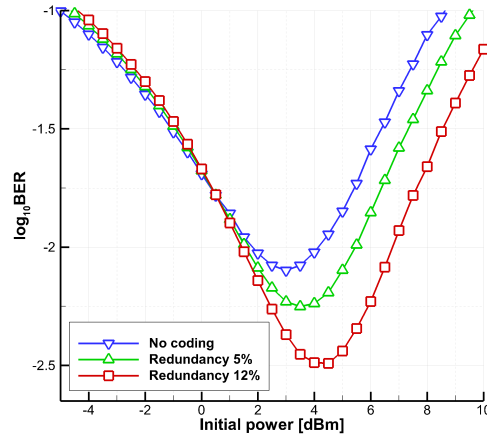
Fig. 8. The dependence between BER and the signal power for various redundancy of the adaptive modulation.

the correction code ability falls drastically, and more sophisticated coding should be applied. Adaptive modulation allows to use more practical and well established codes.

The system improvement is also shown in Fig. 9 as a Q-factor improvement. The Q-factor is calculated from BER using the standard formula:

$$Q = 20 \cdot \log_{10}(\sqrt{2} \ \text{erfcinv}(2 \cdot BER)). \tag{11}$$

From Fig. 9 we see that a Q-factor improvement of 1 dB can be achieved for any transmission distance between 1000 and 2000 km. It is also to be noted that the using of the adaptive modulator allows to increase the propagation distance up to 500 km compared to the signal without coding for Q-factor close to the forward-error correction limit.
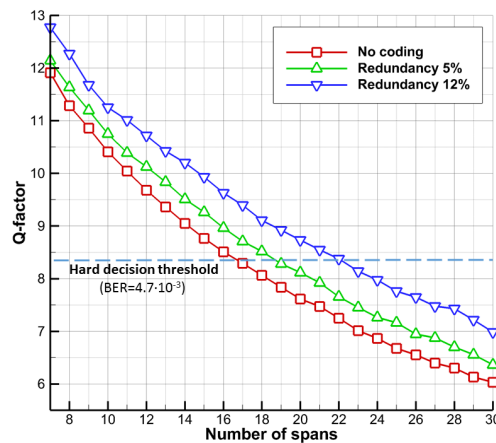

Fig. 9. The Q-factor for various transmission distances.

8. Conclusion

Rate-adaptive coded modulation [10, 11], probabilistic signal shaping [10–14] and skewed signal coding [17, 18] are techniques used to mitigate nonlinear effects and improve system performance

either by modifying the size of the alphabet (and probabilities) of the transmitted constellation points or by applying non-uniform distribution for the occurrence probability of the symbols of a given constellation. In optical communication these approaches are used to remove most error prone patterns or symbols, that typically occur due to power dependence of the error probabilities. Probabilistic shaping of input signal can be implemented using modified FEC codes [12–14] or using reshaping of constellations [11]. We first examined here the impact of the fiber Kerr effect on error statistics in a highly nonlinear transmission of the OFDM 16-QAM signal over a 1000 km EDFA-based link. Based on these observations, we presented the theoretical framework for the probabilistic coding of three constellation power rings to minimize the symbol-error rate of such a signal. We proposed the adaptive modulation technique to produce the OFDM 16-QAM signal that is more tolerant to the nonlinear impairments compared to the initial signal. We demonstrated that the significant performance improvement can be achieved for a large number of OFDM subcarriers (more than 100) using the proposed adaptive modulation scheme. Similar techniques can be applied to single carrier transmission and various modulation formats. The proposed theoretical optimization approach can be applied to the polarization-multiplexed data formats and various correlated data streams.