Comparison of Performance Limits by Mutual Information and Practical Realizations for Optical Long-Haul Coded Modulation Communication Systems

Tobias Fehenberger[(1)], Norbert Hanik [(1)]

[(1)] *Institute for Communications Engineering, Technische Universität München*

tobias.fehenberger@tum.de

Abstract

Mutual information (MI) is derived as an information-theoretical figure of merit to assess the performance bounds of an optical long-haul communication system when forward error correction is used. This limit, which is valid for any optical channel model, is compared with a trellis coded modulation (TCM) system and with a regular low-density parity-check (LDPC) code in order to evaluate the gap in performance due to imperfection in the code design.

Keywords: Forward error correction, trellis coded modulation, LDPC, mutual information.

1. Introduction

The required data rates in optical networks continue to grow very fast [1]. At the same time, propagation distances of an optical signal without electrical regeneration remain unchanged. High-order modulation formats, which are a potential way to provide higher bit rates, are in general more susceptible to distortions. Hence, strong forward error correction (FEC) is becoming more and more important to ensure sufficiently small bit error rates (BER) as optical long-haul systems are coming closer to their capacity limits. In this paper, we outline mutual information (MI) as an information-theoretical measure that represents a performance limit when coding is considered. Further, a trellis-coded modulation (TCM) scheme and an LDPC coding scheme are evaluated via simulations for an additive white Gaussian noise (AWGN) channel as well as a nonlinear optical channel.

2. Mutual Information

Mutual information (MI) describes the statistical relations between the channel input and output in a single number that represents the rate of reliable communication for long codes. For a formal definition, let X and Y be the input and output with respective realizations x and y of a discrete-input continuous-output memoryless channel. The symbolwise MI $I(X;Y)$ of X and Y is defined as [2]

$$I(X;Y) = \frac{1}{M}\sum_{m=1}^{M}\int_{-\infty}^{\infty} p(y|x_m)\log_2\left[\frac{p(y|x_m)}{p(y)}\right]dy, \tag{1}$$

where M is the modulation order. The values of $I(X;Y)$ lie in the interval $[0, \log_2 M]$. We further define the bitwise MI $I^*(X;Y)$ as

$$I^*(X;Y) = \frac{1}{\log_2 M} I(X;Y). \tag{2}$$

Naturally, $I^*(X;Y) \in [0,1]$. Note that (1) and (2) are appropriate for memoryless channels. Although optical channels are not memoryless, we neglect their memory and rather treat interference as stochastic impairment.

Informational theory, in particular Shannon's noisy-channel coding theorem [3] tells us that $I(X;Y)$ is the capacity C of a memoryless channel with discrete input that is distributed according to X, and that C must be larger than the rate of any code in order for the code to achieve an arbitrarily small BER after decoding, denoted $\mathrm{BER}_{\mathrm{out}}$. Therefore, successful decoding is possible for any code of rate R if $I^*(X;Y)$ at the decoder input is larger than R. Conversely, if $I^*(X;Y) \leq R$, decoding is guaranteed to fail for any code. Note that this theorem guarantees successful decoding only for ideal, i.e. capacity-achieving codes. Practical codes, however, exhibit a gap from the asymptotic Shannon coding limit due to their non-random structure and finite block length. Considering (1), determining the MI at the decoder input is based on calculating the probability distribution $p(y)$ of all received symbols y, as well as the channel transition probability distribution $p(y|x_m)$ for each symbol x_m of the modulation alphabet of size M. If we assume that the sent symbols are known at the receiver, the distributions $p(y|x_m)$ can be estimated using histograms or calculated from the soft input of the decoder.

3. Forward Error Correction

A four-dimensional (4D) TCM scheme and an LDPC coding scheme are explained in this section and compared for AWGN. Apart from their respective coding realizations, a main difference between them is the way redundancy is added when the net bit rate is to be kept constant. Coded modulation schemes in general expand the modulation to generate space for redundancy; however, the symbol rate is not changed. In contrast, the symbol rate of system that uses an LDPC code of rate R must be increased by $1/R$ in order to keep the net bit rate unchanged if the modulation is not to be expanded.

3.1. Multidimensional Trellis Coded Modulation

A TCM scheme employing a 4D constellation space as described in [4] is evaluated. Conventional TCM schemes are based on the concept of mapping by set partitioning developed by Ungerboeck [5]. Redundancy for FEC is added by increasing the constellation size and using the additionally created bit(s) per symbol as coding overhead, which improves the error performance without sacrificing data rate or requiring more bandwidth [6]. The 4D TCM scheme used in this paper introduces and exploits the dependence between symbols in two successive signal intervals. For polarization-multiplexed optical channels, 4D TCM is an interesting way of utilizing coded modulation. Instead of using two successive time slots to transmit the 4D symbol, each 4D symbol spans over two polarizations of a signaling interval.

3.1.1. Encoding

In order to transmit Q information bits per signaling interval using a code of rate $m/(m+1)$, a 4D constellation of 2^{2Q+1} points is needed. Out of the Q information bits of each signaling interval, m bits are encoded with a $m/(m+1)$ trellis encoder. The resulting $m+1$ coded bits specify the subset, and the remaining information bits are used to specify the 4D point to be transmitted. This point is obtained by concatenating a pair of two-dimensional (2D) points taken from constituent 2D constellations such that the intra-subset minimum squared Euclidean distance (MSED) is increased by a factor of four. The pair of 2D points then corresponds to the 4D point transmitted in two time slots or over two polarizations.
Instead of a 192-point constellation as outlined in [4], quadrature amplitude modulation (QAM) with 24 points (24-QAM) as presented in [6] is used. The constellation points on the complex plane are depicted in Fig. 1. The blue dots represent the underlying 16-QAM, the red crosses the additional 8 points. Underneath each point is the three-bit mapping that is appended to a two-bit subset identifier. The four subsets are labeled 'a', 'b', 'c', and 'd' in Fig. 1. The MSED between neighboring points is chosen such that the transmit signal has unit power when taking into account the probability of occurrence of each symbol. For 4D 24-QAM TCM, 8 information bits are encoded to 10 bits contained in every 4D symbol. The resulting coding rate is thus R=0.8.

3.1.2. Decoding

The Viterbi algorithm is used for maximum-likelihood decoding [7]. The decoder determines the point in each of the multi-dimensional subsets which is closest to the received 4D point. The respective squared Euclidean distances are then used as metrics for the trellis paths.

3.2. Low-Density Parity-Check Code

Soft-decision LDPC codes are widely investigated in the field of optics [8]. Throughout this paper, we use the LDPC code of the DVB-S2 standard [9] with rate R=0.8 and at most 50 decoding iterations. Perfect knowledge of the channel transition probabilities is assumed for the soft-decision input to the LDPC decoder.

3.3. Coding Gain

The asymptotic coding gain of the 4D 24-QAM TCM scheme in comparison to uncoded 16-QAM is calculated as

$$10\log_{10}\left(\frac{4\mathrm{d}_0^2}{3.5\mathrm{d}_0^2}/\frac{\mathrm{d}_0^2}{2.5\mathrm{d}_0^2}\right) = 4.56 \text{ dB}, \tag{3}$$

where $4\mathrm{d}_0^2$ is the intra-subset MSED after set partitioning, and $3.5{d_0}^2$ is the average power of the 24 point 4D constellation taking into account the probability of occurrence of the symbols. The average power of the 16-QAM constellation with equiprobable symbols is $2.5\mathrm{d}_0^2$.
In contrast to this analysis, the coding gain of an LDPC code with iterative decoding cannot be determined theoretically. A typical approach to assess its error correcting capability is to perform Monte-Carlo simulations and plot the calculated bit error rate versus the ratio $\mathrm{E_s/N_0}$ of energy per symbol to noise power spectral density.
Figure 2 shows the performance of the 4D TCM with 24-QAM (dashed green line) and the LDPC scheme with

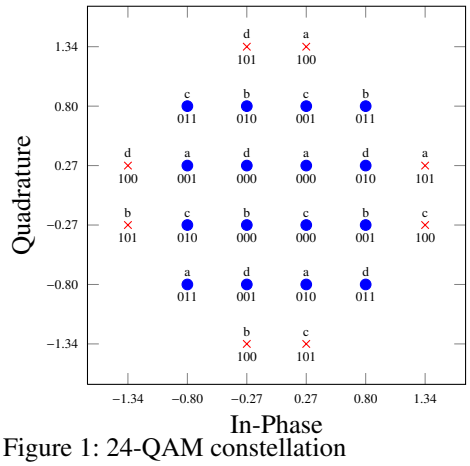

Figure 1: 24-QAM constellation

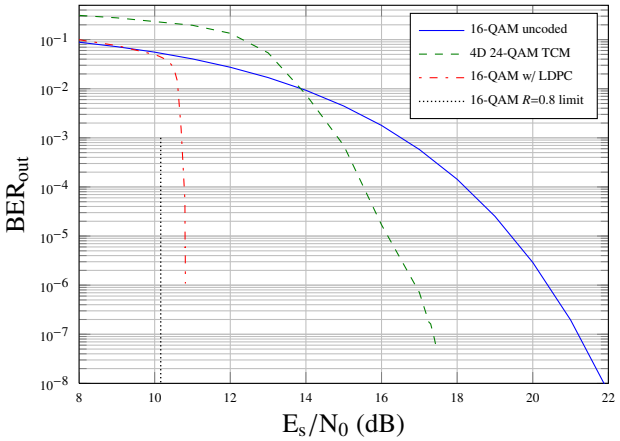

Figure 2: BER$_{out}$ vs. E$_s$/N$_0$ in dB.

16-QAM (dashed-dotted red line) for an AWGN channel in terms of their BER$_{out}$ vs. E$_s$/N$_0$ in dB. As comparison, uncoded 16-QAM is shown as solid blue line. We see that 4D 24-QAM TCM approximately achieves the asymptotic coding gain specified in (3), and 16-QAM with LDPC coding performs better. However, symbol rates are not taken into account in this AWGN simulation, which is clearly in favor of the LDPC scheme. Thus, the two coding schemes are compared in a more realistic scenario in the following. Also, note that LDPC codes usually have an error floor (not shown in Fig. 2) at very low BER$_{out}$. Applying an outer "'clean-up"' code to take care of the error floor is out of the scope of this work. Further shown in Fig. 2 is the asymptotic Shannon limit for R=0.8 as dotted black line. Having a gap of about 0.5dB from this limit, the LDPC code shows excellent performance.

4. Simulations

A single-carrier wavelength- and polarization-division multiplexing (WDM, PDM) system is simulated. Details on the transmitter and receiver subsystem are given in [10] unless stated otherwise in this paper. The center of the three simulated WDM channels are 50Ghz apart when co-propagating over the fiber. All but the center channel are dropped at a coherent receiver that employs digital signal processing (DSP). At the transmitter, 2^{18} bits per polarization are generated pseudo-randomly, encoded using either the 4D 24QAM-TCM scheme or the LDPC coding scheme, and mapped onto 4D 24-QAM (TCM) or regular 16-QAM (LDPC) symbols. The target net baud rate is 28Gbaud. As outlined earlier, the baud rate of the LDPC scheme is increased to 28Gbaud/R=35GBaud. Due to the modulation expansion of TCM, the baud rate of the TCM scheme remains at 28Gbaud. Pulses are shaped as non-return-to-zero in either case. After ideal digital-to-analog conversion (DAC), an IQ-Mach Zehnder modulator drives a laser at 1550nm with 100kHz linewidth to create the optical signal that propagates over a fiber link of spans of 80km standard single mode fiber (SMF) each followed by an erbium-doped fiber amplifier (EDFA). The SMF parameters are attenuation $0.2\frac{\text{dB}}{\text{km}}$, nonlinear coefficient $1.2\frac{1}{\text{W}\cdot\text{km}}$, dispersion $16\frac{\text{ps}}{\text{nm}\cdot\text{km}}$, and PMD of $0.1\frac{\text{ps}}{\sqrt{\text{km}}}$. The EDFA noise figure is 4.5dB.

The optical receiver front-end is a 90° optical hybrid with balanced photodiodes. After ideal analog-to-digital conversion (ADC), dispersion compensation using a FIR filter, equalization using a radially directed scheme [11] with 3 rings for 16-QAM and 4 rings for 24-QAM, and ideal carrier phase recovery, the symbols are demapped and decoded. Finally, the BER is calculated from the resulting bit sequence. As simulating a very low BER$_{out}$ of typically 10^{-15} would take several days even with real-time hardware, the desired BER$_{out}$ is set to 10^{-4} to compare the two coding schemes.

In Fig. 3, three maximum reach curves are depicted for launch powers per channel P_{tx} from -6dBm to 4dBm in steps of 1dBm. The dashed-dotted black and dashed green curve represent LDPC and TCM simulations, respectively. The simulations are carried out until BER$_{out}$ is at most 10^{-4} with transmission over an additional span resulting in exceeding this limit. In the low-power regime, the reach is mostly limited by amplifier noise. After peaking at 0dBm (LDPC) and -2dBm (TCM), the system performance is limited by fiber nonlinearities. We see that the LDPC scheme clearly outperforms the TCM scheme for the simulated parameters. The main reason is the better coding performance as shown for an AWGN channel in Fig. 3. Additionally, the increased modulation size of the TCM scheme is disadvantageous for a power-dependent channel, such as the optical channel. This becomes evident from Fig. 3 as the TCM curve peaks at a lower P_{tx} compared to the LDPC curve, which means that the system is more susceptible to nonlinear distortions due to the additional modulation points. Also, an equalization scheme that is more advanced than the blind method with 4 rings will most likely also lead to better performance for 24-QAM. The baud rate increase of $1/R$ of the LDPC scheme is only disadvantageous concerning the nonlinear fiber effects as the spectral gap between neighboring WDM channels is decreased. There are no bandwidth

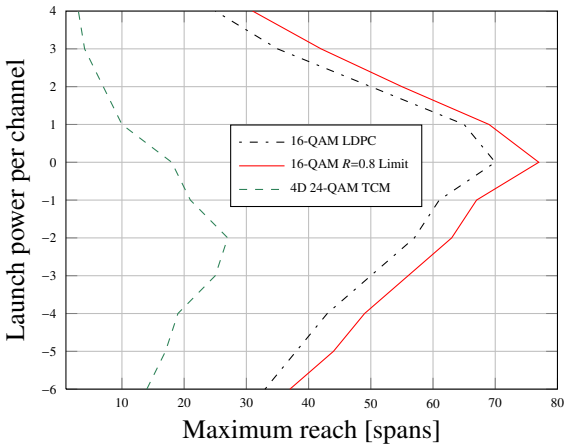

Figure 3: Maximum reach for various P_{tx}.

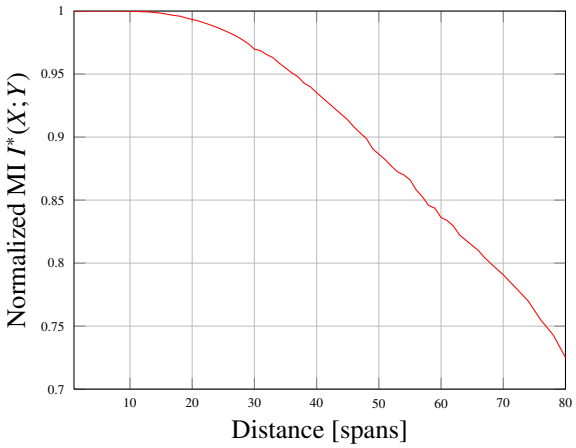

Figure 4: $I^*(X;Y)$ after every span for P_{tx}=-1dBm.

limitations of components that would lead to more severe inter-symbol interference than for TCM. Also 16-QAM and 24-QAM have the same implementation penalty in our setup.

The solid red curve in Fig. 3 represent the maximum distance of the 16-QAM system when $I^*(X;Y)$ falling below R=0.8 for the first time is used as exit criterion. This corresponds to Shannon's noisy-channel coding theorem outlined in Section 2. For the simulated setup, the red solid curve thus represents the maximum distance the signal can propagate over if an arbitrarily small BER is to be achieved. There exists no code that would allow further transmission and still decode to a small BER_{out}.

Figure 4 depicts $I^*(X;Y)$ for the above described LDPC setup with P_{tx}=-1dBm. With increasing distance and thus, more distortions, $I^*(X;Y)$ decreases. At 68 spans, $I^*(X;Y)$ is smaller 0.8 and thus, no code of rate 0.8 could decode the received bits to a small BER_{out}. It is important to understand that this limit is independent of the code under inspection. For example, for distances of 48 spans and more, $I^*(X;Y)$ is smaller than 0.9. Hence, no code of R=0.9 could perform successful decoding for these distances.

5. Conclusion

In this paper, we introduced mutual information (MI) as information-theoretic performance measure. A 4D TCM scheme was presented, and its coding gain was assessed in theory and via simulations over an AWGN channel. It was compared to an LDPC code whose error correcting capability was shown to be superior for AWGN. Simulating an optical communications system showed that the LDPC scheme performs better than TCM. Also, the theoretical transmission limit of the simulated system when FEC is given by MI.

References

[1] R. Essiambre and R. Tkach, "Capacity trends and limits of optical communication networks," *Proceedings of the IEEE*, vol. 100, no. 5, pp. 1035–1055, Jan. 2012.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience New York, 2006.

[3] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, July 1948.

[4] L. Wei, "Trellis-coded modulation with multidimensional constellations," *IEEE Transactions on Information Theory*, vol. I, no. 4, pp. 483–501, July 1987.

[5] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, vol. IT-28, no. 1, pp. 55–67, Jan. 1982.

[6] G. D. Forney Jr. *et al.*, "Efficient modulation for band-limited channels," *IEEE Journal on Selected Areas in Communications*, vol. SAC-2, no. 5, pp. 632–647, Sept. 1984.

[7] G. D. Forney Jr., "The viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.

[8] K. Onohara *et al.*, "Soft-decision-based forward error correction for 100 Gb/s transport systems," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 5, pp. 1258–1267, Sept. 2010.

[9] A. Morello and V. Mignone, "DVB-S2: The second generation standard for satellite broad-band services," *Proceedings of the IEEE*, vol. 94, no. 1, pp. 210–227, Jan. 2006.

[10] S. Kilmurray *et al.*, "Comparison of the nonlinear transmission performance of quasi-Nyquist WDM and reduced guard interval OFDM," *Optics Express*, vol. 20, no. 4, pp. 4198–4205, Feb. 2012.

[11] S. J. Savory, "Digital coherent optical receivers: algorithms and subsystems," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 5, pp. 1164–1179, Sept. 2010.