

Four-Dimensional Estimates of Mutual Information in Coherent Optical Communication Experiments

Tobias A. Eriksson⁽¹⁾, Tobias Fehenberger⁽²⁾, Norbert Hanik⁽²⁾, Peter A. Andrekson⁽¹⁾, Magnus Karlsson⁽¹⁾, and Erik Agrell⁽³⁾

⁽¹⁾ Dept. of Microtechnology and Nanoscience, ⁽³⁾ Dept. of Signals and Systems, Chalmers University of Technology, SE-41296 Gothenburg, Sweden, tobias.eriksson@chalmers.se

⁽²⁾ Institute for Communications Engineering, Technische Universität München, 80333 Munich, Germany.

Abstract *Mutual information is experimentally investigated for long-haul coherent transmission. Receivers that consider memoryless four-dimensional noise distributions can achieve significantly higher rates than receivers assuming two-dimensional symmetric distributions.*

Introduction

Forward error correcting (FEC) codes are an essential part of today's coherent fiber-optical communication systems. Traditionally, hard-decision (HD) codes such as the Reed-Solomon (RS) code have been the dominating codes in long-haul optical communication systems¹. Recently, more advanced coding schemes applying soft-decision (SD) FEC such as low-density parity check (LDPC) codes or turbo-product codes (TPC) have effectively extended the achievable transmission reach of long-haul systems, often in combination with higher-order modulation formats such as polarization-multiplexed 16-ary quadrature amplitude modulation (PM-16QAM).

In coherent fiber-optical experiments, data is often captured and processed off-line using digital signal processing (DSP) structures that should resemble what could be done practically in real-time using application-specific integrated circuits (ASICs) or field-programmable gate arrays (FPGAs). Due to the extremely high sampling rates it is both expensive and time consuming to develop real-time circuits. This, together with the fact that programming languages such as Matlab or C++ offer simple reconfigurability and debugging, has made off-line processing the norm in fiber-optical transmission research experiments. The one part that is most often omitted in the DSP is the FEC, because it is simply not feasible to capture and process enough data with off-line processing to reach post-FEC BER levels that are considered "error free" (often $< 10^{-15}$).

The common practice in the community today is to measure the uncoded BER and depending on which FEC code that is assumed, set a pre-FEC BER target that is considered "error free". For HD codes, there is fixed relationship between pre-FEC BER and the BER after decoding², which makes this practice valid. However, for SD-FEC

codes, there is no such relationship, and in fact the decoders work on soft information. Hence, for experiments assuming an SD-FEC code, the mutual information (MI) is a more logical measure and has been shown to be more reliable than using the pre-FEC BER^{3,4}.

In this paper, we experimentally compare four different estimates of MI assuming either independent, identically distributed (iid) Gaussian noise distribution in 2D or 4D with the same variance for all constellation points or assuming 2D or 4D correlated Gaussian distributions with separate covariance matrices for each constellation point.

Estimate of Mutual Information with Mismatched Decoding

The MI for a memoryless channel $f_{Y|X}$ is given by $I(X; Y) \triangleq \mathbb{E}_{X,Y} \left[\log_2 \frac{f_{Y|X}(Y|X)}{f_Y(Y)} \right]$, where f_Y is the probability density function at the output. However, for the fiber optical channel, $f_{Y|X}$ is not known, which forces us to some assumptions. Following⁵, it can be shown that using a mismatched decoder, substituting $f_{Y|X}$ with an auxiliary function $q_{Y|X}$ gives a lower bound on MI². For a system transmitting uniformly distributed symbols x from a constellation \mathcal{X} with m bits per symbol, an achievable rate is computed as

$$R \approx m + \frac{1}{N} \sum_{n=1}^N \sum_{x \in \mathcal{X}} q_{X|Y}(x|y_n) \log_2 q_{X|Y}(x|y_n),$$

where $q_{X|Y}$ is obtained using Bayes' rule (see², eq. (4)) and Y_1, \dots, Y_N are N observed output samples². Although this is a lower bound on the MI, it can be argued that this is the practically achievable rate since the decoder needs to assume a channel. In many cases, this is assumed to be a memoryless channel with Gaussian distributions of the noise. The better $q_{Y|X}$ resembles the true channel $f_{Y|X}$, the tighter the bound on

the true (although memoryless) MI becomes.

In this work, we compare four different assumptions for $q_{Y|X}$. The distribution parameters were estimated from experimental data as follows:

2D-iidG: 2D iid Gaussian noise in each polarization. The output distribution corresponding to each input point is circularly Gaussian with the same variance for each constellation point.

2D-CG: 2D correlated Gaussian (CG) noise, independent in each polarization. A 2D covariance matrix is estimated for each constellation point.

4D-iidG: 4D iid Gaussian noise, with the same variance for all 4D constellation points.

4D-CG: 4D CG noise. A 4D covariance matrix is estimated per 4D constellation point.

For the 2D estimates, the total MI is the sum of the MI in the x- and the y-polarization.

Note that in an experiment, the channel is not only the optical fiber itself, but the whole communication system from the transmitted constellation to recovered constellation after the DSP. In the transmitter, components such as digital-to-analogue converters, electrical amplifiers and optical modulators can all affect the transmitted constellation with effects such as quantization noise and nonlinear transfer functions. In the receiver, effects such as limited resolution in the analogue-to-digital converters, non-orthogonality of the optical hybrids, and non ideal algorithms in the DSP will influence the constellation.

Experimental Setup and Results

The recirculating loop that was used in all experiments is shown in Fig. 1. The spans consisted of 80 km standard single-mode fiber (SMF) amplified by erbium-doped fiber amplifiers (EDFAs). Inline dispersion compensation could be enabled by the use of dispersion-compensating modules (DCMs) based on fiber Bragg-gratings which compensated the dispersion in each span. For transmission without dispersion compensation, these modules were bypassed. A polarization scrambler that was synchronized to the loop-roundtrip time was used to avoid any nonrealistic accumulation of polarization effects, and a third EDFA was used to compensate for the loss of the scrambler and the loop-switching components.

The first experiment transmitted 7 channels with a spacing of 50 GHz and a symbol rate of 28 Gbaud to emulate wavelength-division multiplexed (WDM) transmission of either PM

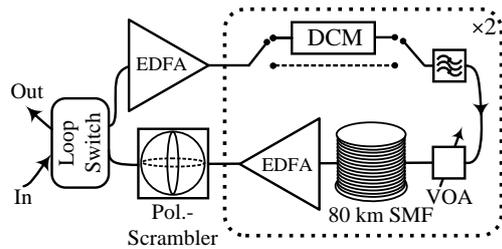


Fig. 1: The recirculating loop used in all experiments. To enable inline dispersion, compensating modules (DCM) based on fiber Bragg-gratings were used. These could be bypassed for uncompensated transmission.

quadrature phase shift keying (PM-QPSK) or polarization-switched QPSK (PS-QPSK). For this experiment, no inline dispersion compensation was used. The DSP is based on the constant modulus algorithm for equalization and Viterbi-Viterbi phase estimation. The achievable rate as a function of transmission distance for the center channel using all four estimates for PM-QPSK and the two 4D estimates for PS-QPSK are shown for different launch powers in Fig. 2. There is no distinctive difference between the different estimates, even at the highest launch power. This shows that for a dispersion unmanaged link with 28 Gbaud WDM channels, assuming an iid Gaussian distribution of the noise with the same variance for all constellation points yields as high rates as with a much more complex 4D receiver. For the optimal launch power, we expect this to be true for higher-order modulation formats and we note that this model has been used to estimate MI for QAM formats up to 256-QAM⁶. However, for high launch powers, this needs verification when the modulation order is increased.

In the second experiment, we transmitted a single channel of 14 Gbaud PM-16QAM, this time using the inline DCMs. The DSP is based on decision-directed least mean square for equalization and blind phase estimation based on 64 test-phases. The launch power was either 0 dBm or -7 dBm. The achievable rate using the four estimates is shown in Fig. 3. For -7 dBm, there is no significant difference in the achievable rate between the four methods. At 1920 km transmission, the difference between the highest estimate, 4D-CG, and the lowest, 2D-iidG, is approximately 0.01 bit/4D-symbol. However, when we increase the launch power to 0 dBm, a clear difference can be noticed between the four estimates. The 4D-CG estimate achieves the highest rate at all transmission distances. Assuming iid Gaussian noise in 2D with the same variance for all constellation points achieves the lowest rate. This is easily understood by the con-

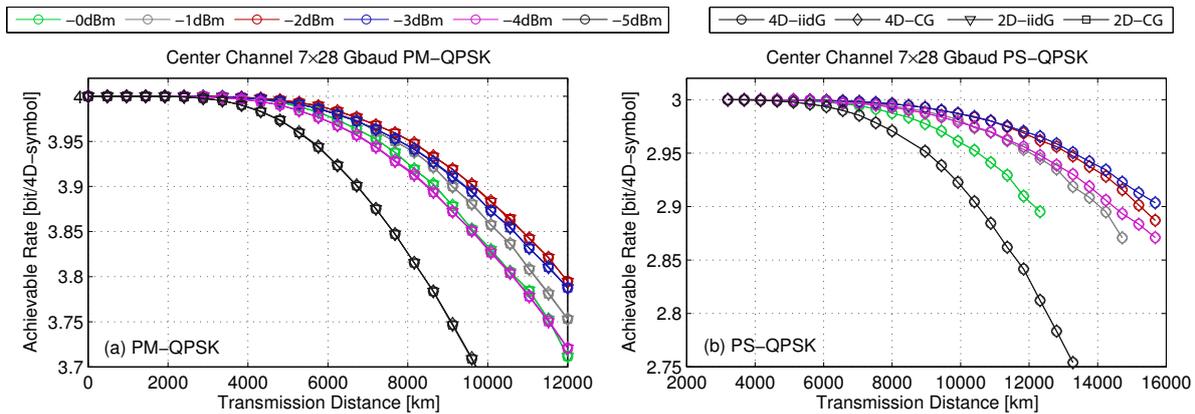


Fig. 2: Achievable rate in dispersion unmanaged WDM transmission using the four different estimates for (a) PM-QPSK and the two 4D estimates for (b) PS-QPSK. Note that the lines for the different estimates are overlapping at the same launch-power.

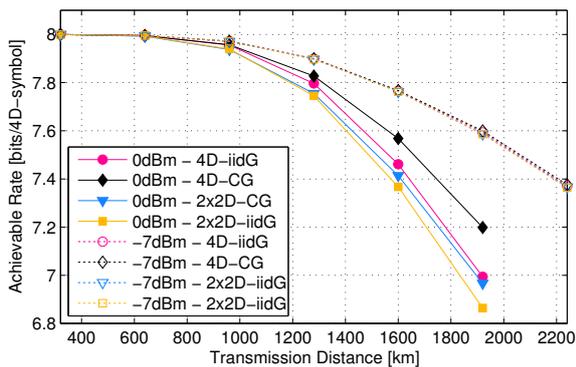


Fig. 3: Achievable rate for single-channel 14 Gbaud PM-16QAM with inline compensation and -7 dBm launch power (open markers, dashed lines) and 0 dBm (solid markers, solid lines).

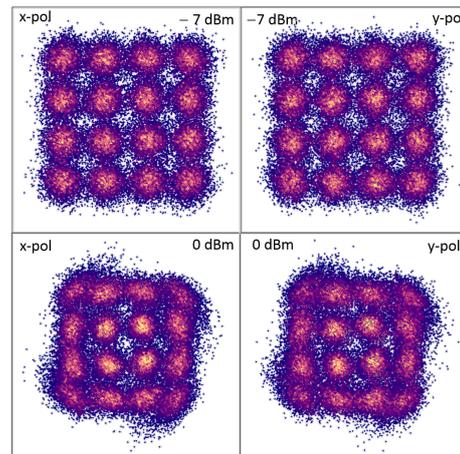


Fig. 4: PM-16QAM constellations after 1920 km for the inline compensated link with -7 or 0 dBm launch-power.

stellations plotted in Fig. 4, showing the signal in each polarization after 1920 km transmission with -7 dBm or 0 dBm launch power. It is common for many FEC decoders to assume 2D iid Gaussian statistics of the noise. However, our results here show that for specific channels such as the one we consider, significant gains can be seen by considering 4D-CG distributions, showing 0.2 bit/4D-symbol higher achievable rate at a transmission distance of 1600 km.

Surprisingly, the 4D-iidG estimate with the same covariance for all constellation points achieves a higher rate than the sum of the 2D estimates with correlated Gaussian distributions with individual estimates of the covariance for each constellation point. This shows that for a highly nonlinear channel, the polarizations are no longer independent and there is loss in information by separating them.

Conclusions

We have experimentally compared four different estimates of MI with mismatched decoders. In long-haul WDM transmission without inline dispersion compensation, assuming iid Gaussian noise in 2D gives a good estimate for PM-QPSK and iid Gaussian noise in 4D for PS-QPSK. For a

highly nonlinear channel, such as single-channel PM-16QAM in an inline dispersion-compensated link with high launch power, there is significant gains in the achievable rate by considering noise distributions over the full 4D optical field. For this channel, a practical conclusion is that a receiver that assumes the noise to be correlated 4D Gaussian distributed with individual covariance matrices per 4D constellation point achieves a higher rate than other memoryless receivers.

Acknowledgements: This work was partially funded by the Swedish Research Council.

References

- [1] ITU, Recommendation G.975: Forward error correction for submarine systems, Oct. 2000.
- [2] T. Fehenberger *et al.*, "On Achievable Rates for Long-Haul Fiber-Optic Communications," *Op. Exp.*, vol. **23**, p. 9183 (2015).
- [3] A. Leven *et al.*, "Estimation of soft FEC performance in optical transmission experiments," *Photon. Technol. Lett.*, vol. **23**, p. 1547 (2011).
- [4] A. Alvarado *et al.*, "LDPC Codes for Optical Channels: Is the "FEC Limit" a Good Predictor of Post-FEC BER?," *OFC'15*, p. Th3E.5 (2015).
- [5] D. Arnold *et al.*, "Simulation-based computation of information rates for channels with memory," *Trans. Inf. Theory*, vol. **52**, p. 3498 (2006).
- [6] D. S. Millar *et al.*, "Transceiver-Limited High Spectral Efficiency Nyquist-WDM Systems," *OFC'15*, p. Th2A.13 (2015).