Capacity Estimation and Near-Capacity Achieving Techniques for Digitally Modulated Communication Systems

Yankov, Metodi Plamenov

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Capacity Estimation and Near-Capacity Achieving Techniques for Digitally Modulated Communication Systems

Metodi Plamenov Yankov

December 2015

DTU Fotonik
Department of Photonics Engineering

Coding & Visual Communication
DTU Fotonik
Technical University of Denmark
2800 Kgs. Lyngby
Denmark

Contents

Abstract

This thesis studies potential improvements that can be made to the current data rates of digital communication systems. The physical layer of the system will be investigated in band-limited scenarios, where high spectral efficiency is necessary in order to meet the ever-growing data rate demand. Several issues are tackled, both with theoretical and more practical aspects.

The theoretical part is mainly concerned with estimating the constellation constrained capacity (CCC) of channels with discrete input, which is an inherent property of digital communication systems. The channels under investigation will include linear interference channels of high dimensionality (such as multiple-input multiple-output), and the non-linear optical fiber channel, which has been gathering more and more attention from the information theory community in recent years. In both cases novel CCC estimates and lower bounds are provided in this thesis. Intuition about the optimal signaling distribution is also provided, which is generally not the standard uniform for high spectral and energy efficiency communications.

The practical part deals with tools to approach the CCC with real-life transceivers. The constellation shaping concept is one such tool. More specifically, the probabilistic shaping concept is of interest to this thesis. A rate-adaptive solution is proposed for designing the mapping function of a probabilistic shaped coded modulation system, which allows for approaching the above mentioned optimal distribution in practice. This results in increased energy and/or spectral efficiency for both linear and non-linear channels, but also increased maximum reach of the optical link at fixed spectral efficiency.

The specific problem of phase noise in digital systems is also stud-

ied in this thesis. Phase noise, and particularly non-linear phase noise is especially detrimental to high-speed, high spectral efficiency optical communications. As part of this work, a low-complexity solution is proposed for tracking, which is able to combat the combined effect of linear and non-linear phase noise in optical fibers, achieving close to the CCC estimate.

The main contribution of the PhD project is providing engineers with limits on the data rates that current digital communication systems can achieve, and also with methods and insights for approaching those rates, thus interconnecting theory and practice.

Resumé

Denne afhandling undersøger potentielle forbedringer af datahastigheden i digitale kommunikationssystemer. Systemets fysiske lag bliver undersøgt i båndbredde-begrænsede scenarier, hvor høj spektral effektivitet er nødvendig for at opfylde kravet om stadigt stigende datahastighed. Afhandlingen løser en række problemstillinger, både af praktisk og mere teoretisk karakter.

Den teoretiske del beskæftiger sig hovedsageligt med estimering af kapaciteten under begrænsning af konstellationen (constellation constrained capacity (CCC)) i kanaler med diskret input, hvilket er en iboende egenskab ved digitale kommunikationssystemer. Kanalerne der er omfattet af undersøgelsen inkluderer lineære interferens kanaler af høj dimensionalitet (såsom multipelt-input multipelt-output), og den ikke-lineære optiske fiberkanal, som har fået stadig større opmærksomhed fra informationsteori fællesskabet i de senere år. I begge tilfælde er nye CCC estimater og nedre grænser fastsat i denne afhandling. Intuition om den optimale signalfordeling gives også, hvilket generelt ikke er standard ligefordelt for højspektral og højenergieffektivitetskommunikation.

Den praktiske del omhandler værktøjer, der muliggør at datahastigheden i virkelige transceivere nærme sig CCC. Formning af konstellationen udgår et sådan værktøj. Mere specifikt er probabilistisk formning genstand for undersøgelse i denne afhandling. Der foreslås en hastighedstilpassende løsning til at designe afbildningfunktionen for et probabilistisk formeret kodet modulationssystem, som gør det muligt at nærme sig den optimale fordeling i praksis. Dette resulterer i forøget energi- og/eller spektral-effektivitet for både lineære og ikke-lineære kanaler, men også forøget maksimal rækkevidde af det optiske link ved fast spektral effektivitet.

Denne afhandling undersøger desuden det specifikke problem med fasestøj i digitale systemer. Fasestøj og især ikke-lineær fasestøj er en særlig udfordring ved højhastigheds og højspektraleffektiv optisk kommunikation. Som en del af dette arbejde, foreslås en lavkompleksitetsløsning til sporing af fasestøj, som er i stand til at bekæmpe den kombinerede effekt af lineær og ikke-lineær fasestøj i optiske fiber systemer hvilke kommer tæt på CCC estimatet.

Ph.d.-projektets hovedbidrag er at give ingeniører de grænseværdier for datahastigheder man kan opnå i nuværende digitale kommunikationssystemer, samt metoder til og indsigt i hvordan man opnår disse hastigheder, og dermed sammenkoble teori og praksis.

Acknowledgments

I would like to start by thanking my supervisors at DTU, prof. Søren Forchhammer and prof. Knud J. Larsen, for giving me the opportunity to work as a PhD student in the Coding and Visual Communications group. I also thank them for the guidance and many fruitful discussions over the last three years.

Thanks to my wonderful fellow PhD students and post-docs at the Coding group - Claire Mantel, Jacob Søgaard, Matteo Salmistraro, Nino Burini and Marco Zamarin - for sharing many good moments during my studies. I also thank my colleagues from the High Speed Optical Communications group - Francesco Da Ros and Edson Porto Da Silva - for the close collaboration during the last few months of the PhD project. I am particularly grateful to assoc. prof. Darko Zibar from the same group for encouraging me to study the optical fiber channel.

Thanks to the team led by Niels Mørch at Renesas Mobile Europe, and later Fingerprint cards, for hosting me during the first several months of my PhD. Special thanks go to Lars P. B. Christensen from the team for co-supervising my project and many, many good insights he has provided me with.

I would like to thank prof. Gerhard Kramer from the Institute for Communications Engineering at the Technical University of Munich for hosting me during my six months external stay. Thanks to the wonderful people from his group for making me feel like at home. Special thanks to Tobias Fehenberger and Luca Barletta for the fruitful discussions on the optical channel and also for co-authoring one of my papers.

I would like to acknowledge my friends back in Bulgaria for never losing touch with me even after a long time apart. Thanks to Ivaylo Kostov for proof-reading this text.

I acknowledge the Oticon, Augustinus and Otto Mønsted foundations for the financial support of my conference travels and the external stay in Munich.

Finally, I would like to thank my family for believing in me, and always supporting me when I decided to study in Denmark. The biggest thanks go to my wife, Ginka, for always standing by me, especially in the most intense of times towards to end.

Thank you!

Metodi P. Yankov

List of Ph.D. Publications

This thesis is based on the following original publications, which can be found in Appendix A:

Peer-reviewed

PAPER 1 M. P. Yankov, S. Forchhammer, K. J. Larsen, and L. P. B. Christensen, "Factorization properties of the optimal signaling distribution of multi-dimensional QAM constellations", *Proc. of 2014 IEEE 6-th International Symposium on Communications, Control and Signal Processing (ISCCSP 2014)*, May 2014, pp. 384 - 387.

PAPER 2 M. P. Yankov, S. Forchhammer, K. J. Larsen, and L. P. B. Christensen, "Rate-adaptive constellation shaping for turbo-coded BICM", *Proc. of 2014 IEEE International Conference on Communications (ICC 2014)*, June 2014, pp. 2112 - 2117.

PAPER 3 M. P. Yankov, D. Zibar, S. Forchhammer, K. J. Larsen, and L. P. B. Christensen, "Constellation shaping for fiber-optic channels with QAM and high spectral efficiency", *IEEE Photonics Technology Letters*, vol. 26, no. 23, pp. 2407-2410, Dec. 2014

PAPER 4 M. P. Yankov, S. Forchhammer, K. J. Larsen, and L. P. B. Christensen, "Approximating the constellation constrained capacity of the MIMO channel with discrete input", *Proc. of 2015 IEEE International Conference on Communications (ICC 2015)*, June 2015, pp. 5634-5639.

PAPER 5 T. Fehenberger, **M. P. Yankov**, L. Barletta, and N. Hanik, "Compensation of XPM interference by blind tracking of the nonlinear phase in WDM systems with QAM input," *Proc. of 2015 European Conference on Optical Communications (ECOC 2015)*, Oct. 2015, p. P.5.8.

PAPER 6 M. P. Yankov, T. Fehenberger, L. Barletta, and N. Hanik, "Low-Complexity tracking of laser and nonlinear phase noise in WDM optical fiber systems", *IEEE Journal of Lightwave Technology*, vol. 33, no. 23, pp. 4975-4984, Dec. 2015

Non-peer-reviewed

PAPER 7 M. P. Yankov and S. Forchhammer, "Achievable information rates on linear interference channels with discrete input", *Proc. of 2015 8th Workshop on Information Theoretic Methods in Science and Engineering (WITMSE 2015)*, June 2015, pp. 46-49.

Contributions of the Ph.D. project, but not included in this thesis:

[P8] A. Rasmussen, **Metodi P. Yankov**, Michael S. Berger, Knud J. Larsen, and Sarah Ruepp, "Improved energy efficiency for optical transport networks by elastic forward error correction", *IEEE Journal on Optical Communications and Networking*, vol. 6, no. 4, pp. 397-407, Apr. 2014.

Chapter 1

Introduction

1.1 Motivation

As we all know, any technology sector nowadays, including the communication and information technology (IT), aims at sustainable operation and low energy consumption. On the other hand, data rate demands have increased dramatically in recent years, with the trends being for steadily increasing throughput every year, which directly increases energy consumption. Recent studies have shown the IT sector to be responsible for around 2% of the global greenhouse gas emission, with around 31% of that contributed to communications [1]. Energy efficiency is therefore of key importance to the future communication systems. Further, in band-limited cases, such as wireless and recently the optical fiber communications, high spectral efficiency is a requirement for high data rate systems. The optimal trade-off between the energy, used for transmission over the channel and spectral efficiency was defined by Shannon for the additive white Gaussian noise (AWGN) channel in his pioneering work [2] from 1948. This trade-off is called the *channel capacity*.

Since then, the communication and information theory communities have been developing new ways of approaching this capacity in practice. These efforts have resulted in very powerful error-correcting codes such as the low-density parity check (LDPC) codes, discovered first by Gallager in the 60's [3] and later re-invented in the mid-90's [4], when technology allowed practical implementation of decoding algorithms. Another type of capacity approaching channel code is the turbo code, invented in the

early 90's [5]. While these codes perform very close to capacity in low signal-to-noise ratio (SNR) scenarios, mid-to-high SNR (high spectral efficiency, respectively) cases are still an open question. The alphabet expansion principle was proposed by Ungerboeck [6] as a way to increase the spectral efficiency, which led to the concept of *coded modulation*, and later to bit-interleaved coded modulation (BICM) [7]. Increasing the spectral efficiency further was made possible by the multiple-input multiple-output (MIMO) principle, which exploits the spatial dimension of the signal in addition to frequency and time, by employing multiple transmit and receive antennas. MIMO is now an integral part of most mobile and fixed communication standards, with the *massive MIMO* concept emerging for the next generation of wireless services - the 5G [8]. While massive MIMO in principle provides the required data rates, there are still a lot of open issues in its practical implementation. Of these we mention the energy efficiency, interference management and computational complexity. This thesis is particularly interested in the problem of capacity estimation for MIMO when digital modulation is employed. While the channel capacity has been found for some boundary cases, such as very low and very high SNR [9], or for continuous input [10], the problem with estimating it for discrete input constellations and large number of antennas still stands.

Looking at the backbone communications networks, which are mostly comprised of optical fibers, the historical development has been a little bit different. Due to the "infinite" bandwidth assumption, very little effort had been put on spectral efficiency for a very long time in optical communications. With the data rate demands increasing exponentially, this assumption becomes more and more obsolete, and fiber optic networks become stressed for bandwidth. Entering the *Zetta byte* era [11] in the next few years, several solutions are possible for meeting the demand:

1. Increase the number of fiber connections, thereby directly increasing the number of parallel channels and the data throughput. This solution is very inefficient from an economic point of view due to the high cost of laying down the fibers.

2. Exploit the spatial dimension for communications. This can be realized by replacing the single-mode fibers with multi-mode fibers and/or multi-code fibers and the adoption of the MIMO principle

for detection. The same economical problem is seen here.

3. Increase the spectral efficiency inside existing fibers. This requires adopting already existing technologies, such as the above-mentioned alphabet expansion and coded modulation. There are a number of immediate problems with increasing the spectral efficiency:

   - First and foremost - the so-called *nonlinear Shannon limit*. Due to the nonlinearities of the fiber, increasing the transmit power does not necessarily lead to increased spectral efficiency. In fact, it decreases with launch power if the above techniques are directly applied without considerations of the nonlinear effects [12]. Finding the true capacity of the optical fiber in the presence of nonlinearities is still an open problem, with only approximations and bounds currently available, e.g. [12–15].
   - Computational complexity. Increasing the throughput per channel inherently requires faster electronics, which could be an issue for the very high speed future communication systems.
   - Energy efficiency. Optical fiber communications will still face the same issue as with wireless and wired communication systems, which tend to operate further away from capacity at high spectral efficiency. Furthermore, operating close to capacity requires sophisticated signal processing, which inevitably results in increased energy consumption.

We conclude this section by mentioning the general energy efficiency problem of communications and IT. Even when solutions are found for meeting the high data rate demand and increasing the spectral efficiency of communication systems, keeping the same energy efficiency per bit will clearly not be sustainable with the projected exponential increase in data rate.

1.2 Structure of the thesis

This thesis is structured as follows. In Chapter 2, the theoretical background of our investigation is given. The main concepts from information

theory that we use are briefly described, and mathematical formulations are provided. We briefly touch upon the difference between channel capacity, mutual information (MI) and achievable information rate (AIR), and we present the general method for obtaining lower bounds on capacity via auxiliary channels. The main channel models of interest to this thesis are presented, including the MIMO channel, the linear impulse response channel, the Wiener phase noise channel and the nonlinear optical fiber channel. The latter is mathematically described via the non-linear Schrödinger equation (NLSE), and the split-step Fourier method (SSFM) for solving it is also presented. We introduce the BICM system and highlight the differences between symbol-wise and bit-wise MI. The concept of constellation shaping is then introduced, and a particular method for achieving shaping gain is presented, which is later built upon in Chapter 5.

Chapter 3, Chapter 4 and Chapter 5 contain the original contributions of this thesis. In Chapter 3, theoretical investigations are carried out into the discrete constellation constrained capacity (CCC) of general linear interference channels, and a simple method for lower bounding the CCC is obtained. AIRs are studied on the nonlinear optical fiber, and some properties are derived for the optimal input distribution in both cases.

In Chapter 4, linear and non-linear phase noise (NLPN) in optical fibers are studied. A tracking method is then designed, which is able to combat the combined effect of NLPN and laser phase noise. Close to a benchmark method's AIRs are reached with this algorithm at much lower complexity.

In Chapter 5, mapping functions are designed for turbo coded BICM, which achieve close to the theoretical shaping gain, and bring the system performance close to the CCC on an AWGN channel. The designed method produces mappings that perform also very well in the nonlinear region of the multi-channel optical fiber.

Chapter 2

Theoretical background and state of the art

This chapter familiarizes the reader with basic concepts from information theory, such as mutual information and channel capacity, and also describes some state of the art methods for estimating the capacity in cases, where it cannot be directly expressed. We also briefly introduce bit-interleaved coded modulation (BICM) and in the last section we present a method for constellation shaping, which is further discussed in Chapter 5.

The quantities mutual information (MI), entropy and capacity will usually be displayed in *bits per symbol*, or *bits per channel use*. Depending on the context, one symbol or channel use may be single- or multi-dimensional, but all cases share the bandwidth and time dimensions. Therefore the most popular metric in this thesis will be *bits/s/Hz*. Under the assumption of ideal Nyquist pulse shaping, as here, the metrics are identical. When we discuss e.g. the multiple-input multiple-output (MIMO) channel, one symbol (or channel use) will contain all spatial dimensions of the signal, unless otherwise stated, and thus the same unit will be used. However, in the optical fiber channel, where depending on the scenario single or double polarization of the light may be used for transmission, a more appropriate unit can be *bits/s/Hz/polarization*.

2.1 Notation

The following general notation is adopted in this thesis:

- Capital letter, e.g. X, represent random variable (RV)s.
- Lower case letters, e.g. x_k, represent realization of the respective variable, and the subscript usually represents time.
- Bold lower case letters, e.g. $\mathbf{x}_k$, represent realization of the respective multidimensional (or vector) RV.
- Bold capital letters, e.g. $\mathbf{H}$, represent realization of the respective matrix RV.
- The indexing $(\cdot)_k^K$ represents the sequence from k to K.
- $p_X(X)$ represents the probability distribution of X, while $p(x) \triangleq p_X(X = x)$ will denote the probability of X taking value x.

Other notation terms will be defined when they occur.

2.2 Capacity, mutual information and achievable information rates

We begin by introducing the very basic channel model, given in Fig. 2.1. The transmitter generates binary data, which is encoded by a channel code, also known as forward error correction (FEC) code, into the binary sequence B. The encoded sequence is modulated into constellation symbols X, taken from a finite size alphabet $\mathcal{X}$, which are then sent on the channel. The elements in $\mathcal{X}$ are assumed to be complex valued, unless otherwise stated. The channel is generally modeled by a transition probability distribution $p_{Y|X}(Y|X)$. The receiver demodulates the corrupted observations $Y \in \mathbb{C}$ into the bit sequence $\hat{B}$, which may either be hard or soft decision (to be defined later). The channel decoder then decodes and recovers the original bit sequence, possibly with errors. When the distributions $p_X(X)$ and $p_{Y|X}(Y|X)$ are fixed, the maximum amount of data, which can be transferred through the channel is given by the MI

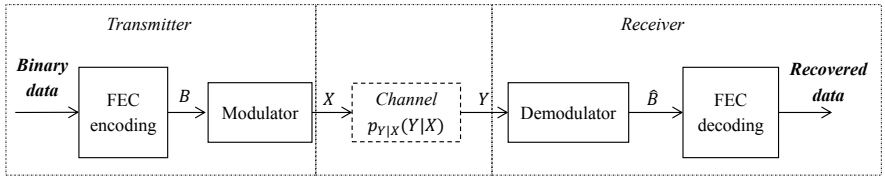


Figure 2.1: Basic communication channel model.

between X and Y

$$\mathcal{I}(X;Y) = \mathcal{H}(X) - \mathcal{H}(X|Y) = \mathcal{H}(Y) - \mathcal{H}(Y|X), \tag{2.1}$$

where $\mathcal{H}(\cdot)$ is the entropy function. When at least one of the arguments is continuous, the entropy is replaced by *differential entropy*. Calculation of differential entropy requires integration, which is not always possible in closed form. The usual method is then to perform numerical integration on a very long sequence, e.g. x_1^K, which was generated by the distribution $p_X(X)$. In that case the entropy of X can be found as [16]

$$\mathcal{H}(X) = \lim_{K\to\infty} \frac{1}{K} \log_2 \frac{1}{p_X(x_1^K)}. \tag{2.2}$$

The entropy $\mathcal{H}(X|Y)$ is calculated in a similar manner: a sequence pair (x_1^K, y_1^K) is generated by the distribution $p_{X,Y}(X,Y) = p_X(X)p_{Y|X}(Y|X)$, and then the distribution $p_{X|Y}(X|Y) = \frac{p_{X,Y}(X,Y)}{\sum_{X_i\in\mathcal{X}^K} p_{X,Y}(X_i,Y)}$ is used as in Eq. (2.2).

The number K needed for the convergence in (2.2) is usually very large, and therefore prevents the calculation of the normalization, needed for calculating $p_{X|Y}(X|Y)$, except for some particular cases. Of these, the most popular is the *memoryless channel*, which when given memoryless input factorizes as

$$p(y_1^K|x_1^K) = \prod_{k=1}^{K} p(y_k|x_1^K, y_1^{k-1}) = \prod_{k=1}^{K} p(y_k|x_k). \tag{2.3}$$

In that case the MI is simply calculated as

$$\mathcal{I}_{memoryless}(X;Y) = \mathbb{E}_k\left[-\log_2 p(x_k) + \log_2 p(x_k|y_k)\right], \tag{2.4}$$

where $\mathbb{E}_k[\cdot]$ is the expectation operator over k, and the probabilities $p(x_k|y_k)$ are found at each k as $p(x_k|y_k) = \frac{p(x_k)p(y_k|x_k)}{\sum_{x\in\mathcal{X}} p(x)p(y_k|x)}$. The most important memoryless channel is probably the additive white Gaussian noise (AWGN) channel, for which $Y = X + W$, where W is zero-mean Gaussian variable with variance σ^2 : $W \sim \mathcal{N}(\sigma^2, 0; W)$. In that case $p(y_k|x_k) = \mathcal{N}(\sigma^2, x_k; y_k)$.

2.2.1 Channel capacity for some channel models

The capacity of the channel is defined as the MI, which is maximized over the input distribution

$$C = \max_{p_X(X)} \mathcal{I}(X;Y). \tag{2.5}$$

If the signal-to-noise ratio (SNR) is defined as SNR $= \mathbb{E}[|X|^2]/\sigma^2$, the capacity of the AWGN channel was found by Shannon to be $C = \log_2(1+$ SNR), and is achieved by $p_X(X) \sim \mathcal{N}(\mathbb{E}[|X|^2], 0; X)$. This result follows from the fact that differential entropy is maximized by the Gaussian distribution [17]. When the input, noise and output of the channel are real-valued, the capacity of the AWGN channel is $C = 0.5\cdot\log_2(1+\text{SNR})$, in which case the SNR is defined as SNR $= \mathbb{E}[|X|^2]/(2\sigma^2)$.

This result is extended to a variety of channels, of which we also focus on the MIMO channel. It is defined as

$$Y = \mathbf{H}X + W, \tag{2.6}$$

where X and Y are $M-$ and $N-$dimensional channel input and output column vectors, respectively, and $\mathbf{H}$ is the $[N\mathbf{x}M]$ channel matrix. The noise W is also $N-$dimensional. When the channel realization is not known at the transmitter but is known at the receiver, and we assume that the noise covariance matrix is the scaled identity $\sigma^2\mathbb{I}^N$ of size $[N\text{x}N]$, the channel capacity is found as [10]

$$C = \log_2 \det\left(\mathbb{I}^N + \frac{\mathbb{E}[X^H X]}{M\cdot\sigma^2}\mathbf{H}\mathbf{H}^H\right), \tag{2.7}$$

where $(\cdot)^H$ is the complex conjugate and transpose operator, and the optimal input distribution is $p_X(X) \sim \mathcal{N}(\frac{1}{M}\mathbb{E}[X^H X]\cdot\mathbb{I}^M, \mathbf{0}; X)$. When the

channel changes over time, the more interesting quantity is the *ergodic* capacity, which is given by

$$C_e = \mathbb{E}_{\mathbf{H}} \left[\log_2 \det \left(\mathbb{I}^N + \frac{\mathbb{E}[X^H X]}{M \cdot \sigma^2} \mathbf{H}\mathbf{H}^H \right) \right], \tag{2.8}$$

and is achieved with the same input. Many variations exist for the expression (2.8). For example, knowing the channel at the transmitter will increase the capacity by finding an optimal *pre-coder*, which introduces dependencies between the dimensions of X.

2.2.2 Capacity with discrete input

When the input is constrained to a finite size alphabet, the maximization in (2.5) cannot be done in closed form. The input in this case is described by its probability mass function (PMF), which is defined as

$$p_X(X, \alpha) = \sum_{x_i \in \alpha \cdot \mathcal{X}} w_i \delta(X - x_i), \tag{2.9}$$

where δ is the Dirac delta function and w_i represent the weights with $\sum_{i=1:|\mathcal{X}|} w_i = 1$. A scaling factor α is also included in order to fulfill the average power constraint $\mathbb{E}\left[X^H X\right] = P_{av}$. Optimizing the values in the set $\mathcal{X}$ leads to geometric shaping, while optimizing the weights leads to probabilistic shaping. In this work, we focus on probabilistic shaping. For fixed set $\alpha \cdot \mathcal{X}$ the MI is strictly concave in the PMF [18], which means that efficient numerical methods can be performed for the optimization. The algorithm for finding the optimal weights of the input PMF and thereby the constellation constrained capacity (CCC) was first found by Blahut [19] and Arimoto [20] for the discrete memoryless channels, and later generalized to AWGN [21] and MIMO channels [22]. It uses expectation-maximization (EM) type rules to update the PMF for a fxied α, increasing the MI on each step, eventually converging to the maximum. It is summarized in Algorithm 1, where we have slightly abused the notation to define the MI as a function of the input distribution and α. In [21], it was proven that as the step-size for brute-force sweeping of the scaling factor α becomes arbitrarily small, capacity is

achieved[1]. When discussing optimized PMFs in the rest of the thesis, the optimization on α is always assumed and is omitted for simplicity.

Algorithm 1 Algorithm for finding the optimal PMF on an AWGN channel with fixed $\mathcal{X}$.

1: Sweep α

Initialize: $p_X(X,\alpha)$, such that the power constraint $\mathbb{E}\left[X^H X\right] \leq P_{av}$ is satisfied and $\sum_{i=1:|\mathcal{X}|} w_i = 1$

2: **while** converged $p_X(X,\alpha)$ **do**
3: Generate $x_1^K \sim p_X(X,\alpha)$
4: Generate $y_1^K \sim p_{Y|X}(Y|X)$
5: $p_{X|Y}(X|Y) = \frac{p_{X,Y}(X,Y)}{\sum_{X_i \in \mathcal{X}} p_{X,Y}(X_i,Y)}$
6: $p_X(X) = \arg\max_{p_X(X)} \mathcal{I}(p_X(X,\alpha))$, s.t. $\mathbb{E}\left[X^H X\right] \leq P_{av}$ and $\sum_{i=1:|\mathcal{X}|} w_i = 1$
7: **end while**
8: $C(\alpha) = \mathcal{I}(p_X(X,\alpha))$ ▷ Maximum for each alpha
9: Go To 1:)
10: $C = \max_\alpha C(\alpha)$ ▷ Global maximum

2.2.3 Auxiliary channel lower bound

The basic Blahut-Arimoto algorithm (BAA) is not always practical, especially in cases of large dimensionality of the input or long memory in the channel. In such cases we must resort to bounds. One of the most popular methods for lower bounding the MI and thereby the capacity of such channels is the auxiliary channel method (also known as the *mismatched decoding* principle). It relies on the fact that the entropy of some variable Z can be upper-bounded by using some auxiliary probability distribution, $\hat{p}_Z$, as such

$$\mathcal{H}(Z) \leq \bar{\mathcal{H}}(Z) = \lim_{K\to\infty} \frac{1}{K} \log_2 \frac{1}{\hat{p}_Z(z_1^K)}, \tag{2.10}$$

[1] It is also easily proven that for fixed power constraint, α is bounded in a certain interval, defined by the peak-amplitude values of $\mathcal{X}$

where the sequence z_1^K was generated by the original $p_Z(Z)$. The proof relies on the non-negativity of the Kullback-Leibler divergence (KLD), and can be found in [16]. The tightness of the bound depends on the quality of the auxiliary PMF, and equality in (2.10) is achieved when $\hat{p}_Z \equiv p_Z$. Calculation of $H(X)$ in communication systems is usually trivial, and by upper-bounding the entropy $\mathcal{H}(X|Y)$ in this manner we achieve a lower bound on the MI.

Auxiliary PMFs $\hat{p}_{X|Y}$ are usually found by using an auxiliary channel $\hat{p}_{Y|X}$ (which is where the method has its name) and then the Bayes theorem. Several such lower bounds will be discussed in the next sections.

Achievable information rate

In practice, all receivers will perform some form of pre-processing before demodulation. This can be e.g. equalization, frequency and phase offset correction, up/down sampling, or any other linear and nonlinear techniques, which aim at improving the quality of the received signal. That is, they attempt to make y_1^K resemble x_1^K as closely as possible. We can model the receiver processing by some function $\hat{Y} = f(Y, \Theta)$ of the channel output and some set of parameters Θ. The channel model from Fig. 2.1 can therefore be replaced by the model from Fig. 2.2. By doing so, we are effectively using an auxiliary channel $\hat{p}_{Y|X}(Y|X) = p_{\hat{Y}|X}(\hat{Y}|X, \Theta)$ to calculate $\mathcal{I}(X;Y)$, and thus the calculated value will actually be a lower bound on the MI. This can also be seen from the data processing inequality [17].

This type of lower bound is called achievable information rate (AIR). It is strictly a lower bound on the MI, but on the other hand it provides an upper bound on the error-free bit rate, which can be achieved after demodulation and decoding. The AIR is thus also a metric of the quality of the specific receiver that we employ, i.e. the function $f(Y, \Theta)$, and of the chosen parameters Θ. In contrast, the MI is a metric of the channel and the input distribution. Both values are generally of interest, however, an AIR is always easier to calculate.

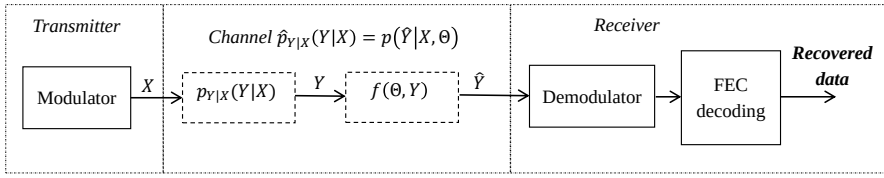


Figure 2.2: Auxiliary channel model.

2.2.4 Constellation constrained capacity estimation

Large input constellation and the MIMO channel

As mentioned, Algorithm 1 is impractical for large dimensionality of the input. For example, when the channel is $[N\mathbf{x}M]$ MIMO, and at least K samples are needed for the convergence in (2.2), the number of calculations per step is $K \cdot |\mathcal{X}|^M$. For example at 64 quadrature amplitude modulation (QAM) and 2x2 channel convergence is seen for $K \geq 10^5$, with this number increasing both with the size of the constellation and the number of transmit antennas. Lots of research has gone into estimating the capacity or even the MI in this case, of which we only mention a few recent works. E.g. in [23], the authors propose an approximation for the MI when the input is uniformly distributed. The approximation is seen to be inaccurate at low and high SNR. Most of the other methods for estimating the capacity rely on the minimum mean squared error (MMSE)-MI relationship [24] [25]. Particularly in the limit of low and high SNR [9] good approximations can be found for uniform input PMF to the MIMO channel. In [26], the authors also use this principle to find the capacity of an AWGN channel in the limit of high SNR for any input distribution. In [24], the authors prove that the CCC converges to the Gaussian capacity in the limit $|\mathcal{X}| \to \infty$. The proof relies on the fact that the MI is continuous in the input distribution when taken in the quadratic Wasserstein space. The quadratic Wasserstein distance between two probability measures μ and v from this space is defined as:

$$W_2(\mu, v) = \inf \left\{ \left(\mathrm{E}_{X,Y} \left[||X - Y||^2 \right] \right)^{1/2} \right\}, \tag{2.11}$$

where X and Y are governed by laws μ and v respectively, and the minimum is taken over all joint distributions of (X, Y) [24]. Since the distance between the Gaussian and a sampled Gaussian vanishes in the limit of large constellation [27], capacity is achieved with sampled Gaussian of infinite size. This is proven for the AWGN channel only.

The linear impulse response channel

Here we define another important class of linear interference channels - the linear impulse response channel. The input-output relation for these channels is

$$Y_k = \sum_{l=0:L} h_l X_{k-l} + W_k, \tag{2.12}$$

where $\mathbf{h} = (h_0, h_1, \dots h_L)^T$ is the impulse response and L is the memory of the channel. Equivalently, the channel may be expressed in its matrix form:

$$Y_1^k = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ h_L & h_{L-1} & \ddots & \vdots \\ 0 & h_L & \ddots & \vdots \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & h_1 & h_0 \end{bmatrix} \times X_1^k + W_1^k. \tag{2.13}$$

Due to the dependencies between the outputs at different times, this channel is difficult to analyze in the time domain. A standard technique for parallelization of the channel is to perform orthogonal frequency division multiplexing (OFDM), which effectively modulates each symbol on a different carrier frequency, which is orthogonal to the others. As long as the channel is known at the receiver, each of the carriers can then be demodulated separately. The optimal strategy at the transmitter is then to use Gaussian distribution on each carrier and allocate the input power according to the water-filling principle [28]. OFDM becomes suboptimal when the channel is not known at the transmitter, and therefore

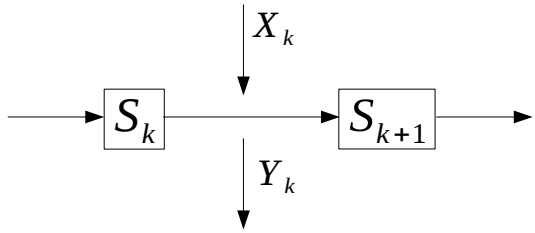


Figure 2.3: Trellis representation of the impulse response channel (2.12)

uniform power allocation should be performed, and when the input is constrained to some finite alphabet.

The MI for such channels is found numerically in [16], where a trellis is used to calculate $p(y_1^K) = \prod_{1:K} p(y_k|y_{1:K-1})$, and thereby the entropy $\mathcal{H}(Y)$. The entropy $\mathcal{H}(Y|X)$ is found as the entropy of the Gaussian noise. Algorithm 1 is later extended in [29] to cover the linear impulse response case. One section of the trellis which is used is given in Fig. 2.3. The interfering symbols are cast into the state: $S_k = \{X_{k-l}, ...X_{k-1}\}$, and the current symbol governs the transition. Marginalizing the state, the desired probability at time k is $p(y_1^k) = \sum_{s_k} p(s_k, y_1^k)$, where each term is calculated recursively [16]

$$p(s_k, y_1^k) = \sum_{x_k} \sum_{s_{k-1}} p(s_{k-1}, y_1^{k-1}) p(y_k|x_k, s_k) p(x_k|s_k). \tag{2.14}$$

The likelihood $p(y_k|x_k, s_k) = \mathcal{N}(\sigma^2, \sum_{l=0:L} h_l X_{k-l}; y_k)$. Since the number of states is given by $|\mathcal{S}| = |\mathcal{X}|^L$, the dimensionality problem is the same as for the MIMO channel, and this method is therefore impractical for large L. Instead, the mismatched receiver lower bound can be applied.

The basic idea is to reduce the cardinality of the state. This can be done by *shortening* the memory to a length $\hat{L}$, and modeling the interference from the remaining channel elements $\{\hat{L}+1, \dots L\}$ as noise. The auxiliary channel in that case is

$$\hat{p}(y_k|x_k, s_k) = \mathcal{N}(\sigma^2 + \mathbb{E}[|X|^2] \cdot \sum_{l=\hat{L}+1}^{L} |h_l|^2, \sum_{l=0:\hat{L}} h_l X_{k-l}; y_k). \tag{2.15}$$

This method is not particularly effective, especially in cases of severe inter symbol interference (ISI). Other methods for shortening find an auxiliary channel of shorter length, which is not necessarily a linear combination of elements of $\mathbf{h}$. For example in [30], a channel of shorter length is found either by numerically minimizing the difference between the achieved lower bound and a previously calculated upper bound, or numerically minimizing the KLD between the true and the auxiliary channel. This method provides AIRs of practical receivers, since the resulting shorter channel is directly used for demodulation. Channel shortening is the impulse response equivalent of the sphere detection for MIMO [28], which can also be described by a trellis. Both methods are equivalently limited in the cardinality of the state which can be used, and the performance improvements will therefore saturate when the cardinality of the input (the memory of the impulse response channel, respectively) is increased beyond some value. A different approach was recently applied in [31], where the authors calculate a lower bound based on the compound capacity model, and show quite tight performance to the MMSE based lower bound.

The Wiener phase noise channel

Another class of channels with memory which is of interest to this thesis is the Wiener phase noise channel. The input-output relation is defined as

$$Y_k = X_k e^{j\theta_k} + W_k, \tag{2.16}$$

where the phase noise θ is modeled by a first order Wiener process with variance Δ^2 by

$$\theta_k = \theta_{k-1} + \Delta \cdot v_k. \tag{2.17}$$

The v_k's are standard i.i.d. Gaussian variables.

The source of the phase noise in a practical system is the non-ideal local oscillator (LO), which in this case is included in the channel model. Local oscillators, especially cost-efficient solutions, have frequency spectrum, which has a non-zero spectral width around its central frequency. When the LO is a laser, which is the focus in this thesis, the more common term for the spectral width is laser linewidth (LLW). When the

LLW at $-3dB$ (also known as full-width half maximum) is f_W and the symbol rate is T_s, the parameter Δ can be found as

$$\Delta^2 = 2\pi f_W T_s. \tag{2.18}$$

The quantity $f_W \cdot T_s$ is how systems are usually parametrized.

The capacity of this channel was found in [32] by discretizing the phase into M values $\theta_k \in \mathcal{S} = \{-\pi\colon \frac{2\pi}{M}\colon \pi\}$, and then using a trellis, similar to the approach in the previous section to upper- and lower-bound the MI. The two bounds coincide, thus giving the CCC. The state in this case represents the phase noise value. This method, while providing an AIR, is rather impractical since it requires knowledge of the past symbols x_1^{k-1} when estimating the probabilities $p(x_k|y_1^K, x_1^{k-1})$. A simplification for this method was made in [33], where the probabilities $p(x_k|y_1^K)$ are estimated instead. First, the posterior distribution of the phase noise is calculated as

$$p(\theta_k|y_1^K) = p(\theta_k, y_1^k)p(y_{k+1}^K|\theta_k), \tag{2.19}$$

where the factorization is a direct consequence of (2.17), and the two factors are calculated by the forward and backward recursions of the BCJR algorithm [34], respectively. The posterior distributions of the symbols are then calculated as

$$p(x_k|y_1^K) = \sum_{\theta_k \in \mathcal{S}} p(\theta_k|y_1^K)p(x_k|\theta_k, y_k), \tag{2.20}$$

where the last term is obtained by the Bayes formula from the Gaussian likelihood $p(y_k|x_k, \theta_k) = \mathcal{N}(\sigma^2, x_k e^{j\theta_k}; y_k)$. This method operates very close to the true CCC at reasonable complexity for small constellations, where the quality of the quantization of the $[-\pi; \pi)$ interval does not need to be very fine. This, however, is not the case for e.g. 256QAM constellation and beyond, where more than 512 states are needed in severe phase noise scenarios (that is, large values of Δ). Recently, this issue has been treated in [35], where reduced complexity trellis was proposed. Even further complexity reduction is possible when the state is represented by a continuous variable, e.g. Gaussian, giving the rise of the Kalman filter solution [36]. The Kalman filter is shown to be near-optimal in some scenarios [37], however, it suffers performance degradation for high information rates and large modulation format. Another method was presented

in [38] for phase noise tracking for constant amplitude modulation formats, such as phase shift keying. All the above methods basically employ the sum-product algorithm for finding marginal and posterior distributions on a chain. In the case of the first order Wiener process the chain is a hidden Markov model.

Another class of algorithms for phase noise cancellation are the decision directed (DD) methods. The basic idea there is to make a decision $\hat{x}_k$ on a single or multiple received noisy observations, and estimate the phase noise value as $\hat{\theta}_k = \angle \hat{x}_k^* y_k$. This can be improved by estimating the average value

$$\hat{\theta}_k = \angle \sum_{i=k-l}^{k+l} \hat{x}_i^* y_i \tag{2.21}$$

assuming the phase noise was constant within the period $[k-l; k+l]$. Single-symbol tracking gives rise to the phase-locked loop technique and modifications, e.g. [39, 40]. Averaging, while more complex, provides a better approximation of the phase noise. It was first proposed in [41], and later modified in [42] to increase the resilience to phase slips. DD methods generally perform poorly in low SNR (low AIR, respectively) scenarios, due to the very high symbol error rate (SER). Further, in severe phase noise cases, the phase noise can no longer be considered constant for a very long time, and DD methods become unable to track these fast changes due to the long sequence averaging.

In order to improve the quality of the decisions, and thus the phase noise estimate, information from the channel code can be included in the decision process [43, 44]. This means that the phase noise tracking is integrated into an iterative receiver, which will be introduced in Section 2.3.

Last, we mention the Viterbi & Viterbi carrier phase estimation for quadrature phase shift keying (QPSK), which relies on estimating the phase noise as $\angle y_k^4$. Modifying this method to higher order QAM constellations requires making a decision on the amplitude of the signal first, and then removing the information in the phase by taking it to the power of four [45]. This method generally also performs poorly in low SNR/high SER due to the DD factor.

The nonlinear optical fiber channel

Finally in this section we describe the modeling and some MI and AIR estimation methods for the *single mode* optical fiber channel. Several phenomena take place during light propagation in the optical fiber which deteriorate the signal quality, of which we mention the most difficult to combat:

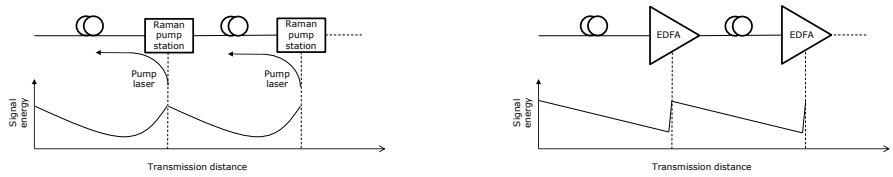
1. The chromatic dispersion (CD). The different spectral components inside the fiber propagate at different velocities. This phenomenon results in pulse broadening, which can be described by the group velocity dispersion parameter $\beta_2 = d^2\beta/d\omega^2$, where β is the propagation constant, and ω is the angular frequency. Alternatively, the CD can be described by the dispersion parameter D, measured in [ps/(nm $\cdot$ km)], and represents how much a pulse with certain bandwidth broadens after a certain transmission distance. It is related to β_2 via the equation

$$D = \frac{2\pi c}{\lambda^2}\beta_2, \tag{2.22}$$

   where c is the velocity of light in vacuum and λ is the center wavelength of the light source. The pulse broadening can be calculated as $\Delta T = DL\Delta\lambda$, where L is the transmission distance, and $\Delta\lambda$ is the bandwidth of the pulse [46].

2. The amplified spontaneous emission (ASE) noise. During propagation, the signal experiences exponential power decay, given by the attenuation constant α, usually in [dB/km]. Transmission at longer distances therefore requires amplification. This is realized by having a high power laser at different central wavelength co-propagate with the signal, which results in signal amplification and spontaneous emission. The latter is simply treated as AWGN [12].

   Two amplification schemes are currently in use - distributed and lumped. In case of distributed amplification the medium, used for co-propagation is the transmission fiber itself. The laser sources, known as Raman pumps, are coupled into the fiber with regular spacing, and provide the required amplification continuously among the fiber length through the stimulated Raman scattering



(a) Raman (b) EDFA

Figure 2.4: Power decay profile for distributed Raman (left) and lumped EDFA (right) power amplification schemes.

effect. Lumped amplification is realized by using a rare-earth element doped fiber, which is much shorter than the transmission length, as amplification medium. The most popular doping element with current systems is Erbium, as its emission band covers the C-band of the optical spectrum. The Erbium doped fiber amplifier (EDFA)s are also inserted regularly, but the power boost they provide is quasi-discrete in space. The power decay profiles are given in Fig. 2.4(a) for backward pumping Raman amplification [2] and Fig. 2.4(b) for EDFA.

The two amplification schemes have different optical SNR (OSNR) performance, with the distributed scheme generally providing higher OSNR at the receiver [12]. The drawback of standard Raman amplifiers is the narrower amplification bandwidth.

3. Non-linear phase shift. The refractive index of the fiber is dependent on the intensity of the electric field, propagating through it. This results in nonlinear phase shift, depending on the instantaneous power of the signal. This effect is described by the nonlinear coefficient γ, defined as

$$\gamma = \frac{2\pi\bar{n}_2}{A_{eff}\lambda}, \tag{2.23}$$

where $\bar{n}_2$ is the nonlinear refractive index and A_{eff} is the effective mode area [46]. The nonlinear phase shift is particularly important when multiple channels are co-propagating in the same fiber, which

[2] Forward and bi-directional pumping are also possible, and beneficial in some cases

is the case in wavelength division multiplexing (WDM) systems. If the power in the $i-$th channel is P_i, the nonlinear phase shift (or non-linear phase noise (NLPN)) in each channel is given by

$$\theta_i^{NL} = \gamma L_{eff}(P_i + 2\sum_{j\neq i} P_j), \tag{2.24}$$

where L_{eff} is the effective length of the interaction, which also accounts for the linear losses. This means that the shift from neighboring channels (called cross-phase modulation (XPM)) is twice as much as the self-induced shift (called self-phase modulation (SPM)) for the same power.

There are several other nonlinear phenomena taking place during propagation, such as stimulated scattering and four wave mixing (FWM). However, there are rather effective ways of suppressing them [46]. We must also mention the effect of polarization mode dispersion (PMD), which occurs due to fiber birefringence. In reality the fiber will not be ideally circular, leading to cross-talk between the otherwise orthogonal in space polarization modes. Neglecting this effect leads to an idealized scenario, which is however easier to analyze. Unless otherwise stated, PMD will NOT be taken into account in this work. We also note that there are higher order dispersive terms, the most important being the third order *dispersion slope* $S = dD/d\lambda$, governed by the corresponding third order dispersion parameter $\beta_3 = d^3\beta/d\omega^3$. This results in different values of D for different WDM channels, but for practical values of the symbol rate it can be considered constant within each channel. Unless otherwise stated, we will assume that $S = 0$ ps/(nm$^2 \cdot$ km).

Making the above simplifications, the wave equation of propagation can be expressed as a differential equation, known as the non-linear Schrödinger equation (NLSE)[3]

$$\frac{dE(z,t)}{dz} = -\frac{1}{2}\alpha E(z,t) - i\frac{\beta_2}{2}\frac{d^2E(z,t)}{dt^2} + i\gamma E(z,t)|E(z,t)|^2, \tag{2.25}$$

where z is the distance, t is time and E is the amplitude of the field. The first term describes the attenuation, the second is the quadratic phase shift in frequency, and the third term describes the second-order

[3]The Rayleigh and Brillouin scattering effects are also neglected.

Algorithm 2 Split-step Fourier method for solving the NLSE

Initialize:

dz ▷ step size
$CD(w, dz)$ ▷ frequency response of the CD after distance dz
z ▷ span distance between EDFAs
N_z ▷ number of spans
$E(z,t) = E(0,t)$ ▷ input sequence to the fiber
$N_0(z, \alpha, F)$ ▷ ASE power spectral density, depending on the distance, required amplification, and the noise figure of the amplifier F
for $n = 1$ **to** N_z **step** 1 **do**
    for $m = dz$ **to** z **step** dz **do**
        $E(z,t) = E(z,t) \cdot \exp(-\alpha \cdot dz/2)$ ▷ attenuation
        $E(z,t) = IFFT\left(FFT\left(E(z,t)\right) \cdot CD(w, dz)\right)$ ▷ CD
        $E(z,t) = E(z,t) \cdot \exp(i \cdot \gamma \cdot dz/2|E(z,t)|^2)$ ▷ NLPN
    end for
    $E(z,t) = E(z,t) \cdot \exp(\alpha \cdot z/2) + \mathcal{N}(N_0(z, \alpha, F)/2, 0)$ ▷ ASE
end for

nonlinearities, such as the SPM, XPM and FWM [46]. Due to the interaction between the terms, no closed form solution exists for the signal at distance z. Instead, numerical solutions are used, the standard approach being the split-step Fourier method (SSFM). It assumes that for very small distances the terms are independent, and can be applied one by one. The SSFM is described for a lumped amplification scheme via pseudo code in Algorithm 2. When distributed amplification scheme is used, the noise loading in Algorithm 2 is performed for each step instead of at the end of each span.

Chromatic dispersion introduces memory in the channel in the form of ISI, which is nonlinear due to the interaction between CD and XPM and SPM effects. A very popular method for canceling the nonlinearities and CD is the digital back-propagation (DBP) [47]. Since all the terms on each step in the SSFM are deterministic, they can be ideally inverted at the receiver, following exactly the same steps, only changing the sign in front of the dispersion, attenuation and NLPN values. Several major issues are seen with this method: 1) complexity is currently

too high for real-time implementation; 2) in a WDM optical network the origins of all received channels are not the same, which means that the transmission distance and other propagation properties of each line are not available; 3) the noise introduces randomness, making the process strictly non-deterministic, and thus DBP is sub-optimal. In [48,49], the authors introduce the stochastic DBP method, which performs a stochastic cancellation on each step by back-propagating not the particular instant of the received signal values, but rather the distribution of the values. Inference tools, such as the belief propagation method can then be used to better estimate the distribution of the input, and thus achieve a better performance than classic DBP. Since the distribution of the output is represented by particles, all of which are back-propagated, the complexity of this method is very high, even compared to the already complex classic DBP.

Another classic approach is to only cancel the effect of CD. As seen from (2.25), CD introduces a quadratic phase shift in frequency, and acts as an all-pass filter. The response of this filter for standard single mode fiber (SSMF)s is known to be $CD(\omega, z) = \exp(i\beta_2\omega^2 z/2)$ [12], and can be matched at the receiver side, either in time or in frequency [50]. Since the required filter is too long in time, and due to the simplicity of performing FFT/IFFT in hardware, the frequency approach is usually preferred.

The CD induced memory and its interaction with the nonlinear effects and the ASE noise makes capacity estimation of the fiber difficult. In probably the most famous paper on the subject [12], the authors attempt to use general methods as described previously in this Chapter, in order to obtain a capacity estimate. Consequently, lower bounds are provided based on a memoryless assumption, and where the input-output relation (or the auxiliary channel) is modeled by a circularly symmetric Gaussian density, with variance depending on the amplitude of each constellation symbol. The constellation they use is a geometrically optimized ring constellation [51]. The AIRs with this method are shown in Fig. 2.5 for several distances and two symbol rates for a SSMF with parameters, as defined in Table 2.1. The amplification scheme is an ideal distributed Raman amplification (IDRA), which means that the attenuation is compensated perfectly on each step, making the profile in Fig. 2.4(a) flat. Single polarization signal is used. Each span is 100km,

Table 2.1: Standard single-mode fiber parameters

Fiber loss	$\alpha = 0.2$ dB/km
Non-linear coefficient	$\gamma = 1.3$ (W·km)$^{-1}$
Dispersion	$D = 17$ ps/(nm·km)
Central wavelenght	$\lambda_0 = 1.55$ μm
SSFM step	0.1 km

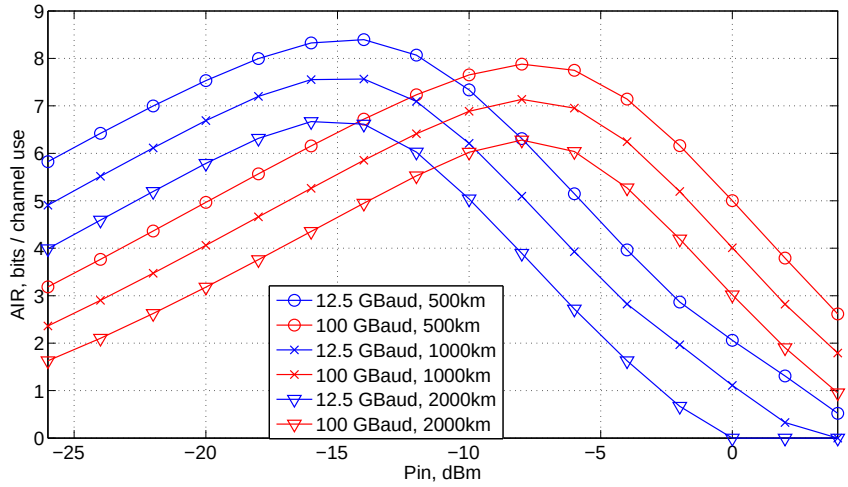


Figure 2.5: AIRs with the optimized ring constellations [51].

and 5 channels are simulated with a guard band of 2GHz. A *sinc* pulse shape is used with oversampling factor of 16.

The results show an AIR, which is gradually increasing with the OSNR for low values of the input power. In such cases the nonlinear phase shift is small, and the channel is mainly corrupted by the ASE. The AWGN model is a good approximation in this regime. Increasing the input power (OSNR, respectively) beyond a certain point results in significant nonlinear effects, thereby increasing the variance of the Gaussian densities of the likelihoods, leading to reduced AIR.

A very popular model of the fiber also assumes Gaussian density as auxiliary functions, hence the Gaussian noise (GN) model, with variance proportional to P_{in}^3 [52], where P_{in} is the launch power. Approximations

of the capacity are provided there, with the assumption of continuous Gaussian input, and also several approximations for QAM constellation input. This method provides a good heuristic for estimating the capacity of the channel, but more importantly circumvents having to solve the Schrödinger equation via the SSFM. The GN model is extended to encompass finite memory of the channel in [53].

The authors in [54] provide a model for the NLPN contribution to the received sequence in WDM systems, and later use this model to improve the above mentioned lower bounds in [55]. The *constellation shaping* (the concept will be introduced in the next section) gains are analyzed with this model in [56]. This set of papers provides a good intuition about the behavior of the NLPN in time and frequency and allows for semi-analytical treatment of the channel memory.

Making use of the above model and the long time and frequency correlation properties of the XPM induced noise, the authors in [57, 58] design a tracking algorithm in frequency domain, which is able to increase the AIRs for Gaussian inputs.

Finally in this section we mention the recent work [14], where an upper bound is found to the capacity of the optical fiber. The bound is found to be the linear AWGN channel capacity of $d_f \cdot \log_2(1 + SNR)$, where d_f represents all the degrees of freedom in the channel - frequency, time, and space in case of dual polarization and multi-mode and multi-core fibers.

2.3 Practical coded modulation schemes

In this section some methods for approaching the AIR in practice are discussed. We focus on the concept of *coded modulation*, first introduced in [59] as trellis coded modulation. The basic idea at the time was to increase the spectral efficiency by using constellation expansion in combination with a channel code, which has to correct all errors, introduced by the resulting increased SER. The concept of BICM [7] was later introduced and was widely developed and deployed in the 90's and 00's in most communication standards, where high spectral efficiency is required.

The general block diagram of BICM with the optional iterative demapping is given Fig. 2.6. At the transmitter, an interleaver is inserted be-

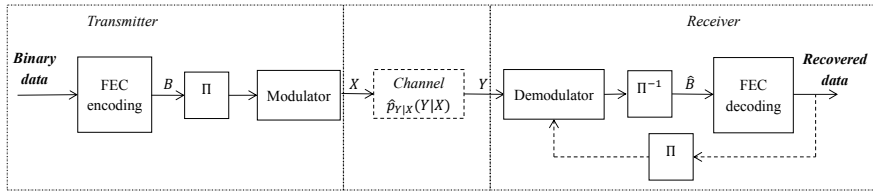


Figure 2.6: General BICM with iterative demapping channel model.

tween the channel code and the modulator, which has the task of *decoupling* the consecutive bits. At the receiver, deinterleaving is performed after demodulation, and the channel code is decoded. This scheme is particularly effective in fast fading channels, where errors occur in bursts, which is something that channel codes generally do not handle very well. The interleaving effectively distributes the errors across the entire codeword. BICM also allows for iterative demapping and decoding. During this process soft decision bits after decoding are sent back to the demapper, which in turn calculates new soft values, to be used by the decoder on the next iteration. Error performance is theoretically improved on each iteration, until some point of convergence.

The soft value L of a transmitted bit B_l at bit time l is defined as

$$L(B_l) = \log \frac{p(B_l = 1|y_1^K)}{p(B_l = 0|y_1^K)}, \tag{2.26}$$

and describes the probabilities of the bit being '1' or '0'. Another useful representation of the soft value is through the log-likelihood ratio (LLR)

$$\begin{aligned} LLR(B_l) &= \log \frac{p(y_1^K|B_l = 1)}{p(y_1^K|B_l = 0)} \\ &= \log \frac{\frac{p(B_l=1)p(B_l=1|y_1^K)}{p(B_l=1)p(B_l=1|y_1^K)+p(B_l=0)p(B_l=0|y_1^K)}}{\frac{p(B_l=0)p(B_l=0|y_1^K)}{p(B_l=1)p(B_l=1|y_1^K)+p(B_l=0)p(B_l=0|y_1^K)}} \\ &= L_a(B_l) + L(B_l), \end{aligned} \tag{2.27}$$

where $L_a(B_l) = \log \frac{p(B_l=1)}{p(B_l=0)}$ is the a-priori knowledge about the bit. If the bit b_l is mapped to the $j-$th position of the labeling of symbol x_k,

the $L-$ value of the bit may be found as

$$L(B_l) = \log \frac{\sum_{x_k : x_k(j)='1'} p(x_k|y_1^K)}{\sum_{x_k : x_k(j)='0'} p(x_k|y_1^K)}, \tag{2.28}$$

where the sums run for the symbols, which have a '1' ('0', respectively) mapped to them at position j in the mapping function. Calculation of soft values can be done very efficiently in hardware and software by using the max-star and max-LOG approximations [60]. Many simplifications exist for these equations, particularly when the input bits and symbols are uniformly distributed. However, we do not go into details, which can be found in e.g. [7, 61].

Soft values basically describe the probability density function (PDF) of each bit, thereby enabling inference methods for channel decoding. For example, the above mentioned belief propagation has been shown to be near-capacity achieving for some of the most powerful channel codes - LDPC [4] and turbo codes [5]. Employing this method for inferring the final PDF allows for calculating the maximum a-posteriori probability (MAP) of each bit, thus making the optimal decision on the transmitted value of the bit. The belief propagation calculates the MAP solution strictly under the assumption of independent bits and errors, which is clearly not the case, since dependencies are already introduced at the channel code, then at the mapping function and at the channel itself. However, if good interleavers are designed, these dependencies vanish for long codewords, and near-optimal inference is possible with belief propagation [4].

2.3.1 Symbol-wise vs. bit-wise AIR and the EXIT chart

Due to the decoupling between modulator and channel encoder BICM theoretically enables the separate design of the channel code and mapping function. In order to characterize the performance of the channel code, we are therefore interested in the average MI per bit provided to it, instead of the AIR after demodulation. Assuming memoryless channel,

[4] We do not go into details of channel code designs, where one of the primary objectives is decoupling between bits, especially for LDPC code design.

for the block diagram from Fig. 2.1 this MI is

$$\mathcal{I}(X;Y) = \sum_{i=0:m-1} \mathcal{I}(B;\hat{B}_i|\hat{B}_{\{0:i-1\}}), \tag{2.29}$$

where B_i is the $i-$th bit in the mapping function, m is the maximum number of bits one symbol can take and we have used the chain rule for MI. This quantity we will refer to as the symbol-wise MI. In (2.29) we have slightly abused the definition of B to express the sequence of bits, related to one constellation symbol, instead of the entire sequence of encoded bits. On a memoryless channel, the MI from (2.4) is identical to the newly defined symbol-wise MI (2.29). Usually, $m = \log_2|\mathcal{X}|$, however, this is not a requirement, as we shall see in the next section. In case of ideal interleaving, the bits are independent, and the condition in Eq. (2.29) disappears. However, ideal interleaving implies infinite codewords, something which cannot be realized in practice. The average MI provided to the decoder is therefore

$$\mathcal{I}(B;\hat{B}) = \frac{1}{m}\sum_{i=0:m-1} \mathcal{I}(B_i;\hat{B}_i) \leq \frac{1}{m}\mathcal{I}(X;Y), \tag{2.30}$$

where the inequality follows from the fact that conditioning does not increase the entropy. This is the BICM bit-wise MI. The MIs from Equations (2.29) and (2.30) are given for an AWGN channel with 16 pulse amplitude modulation (PAM) real-valued input in Fig. 2.7. The mapping function in this case is a binary-reflected Gray labeling function, which was conjectured to maximize the bit-wise MI at the SNR regions of interest [62]. We see that especially at low SNR there is a notable difference between bit-wise and symbol-wise MI[5]. This difference is related to the constellation size. As we get to an AIR of $|\mathcal{X}|$ this difference vanishes. The bit-wise MI is also clearly dependent on the labeling of the symbols, as discussed in e.g. [62,64]. As mentioned, the bit-wise MI is what an ideal channel code can decode. That is, error-free performance can be expected with an ideal code (or capacity achieving code) of rate $R = \mathcal{I}(B;\hat{B})$.

[5] It was proven in [63] that binary-reflected Gray mappings are sub-optimal at *asymptotically low SNR*. However, this is not yet the effect we see in Fig. 2.7, as the bit-wise MI is above 1 in the region of interest.

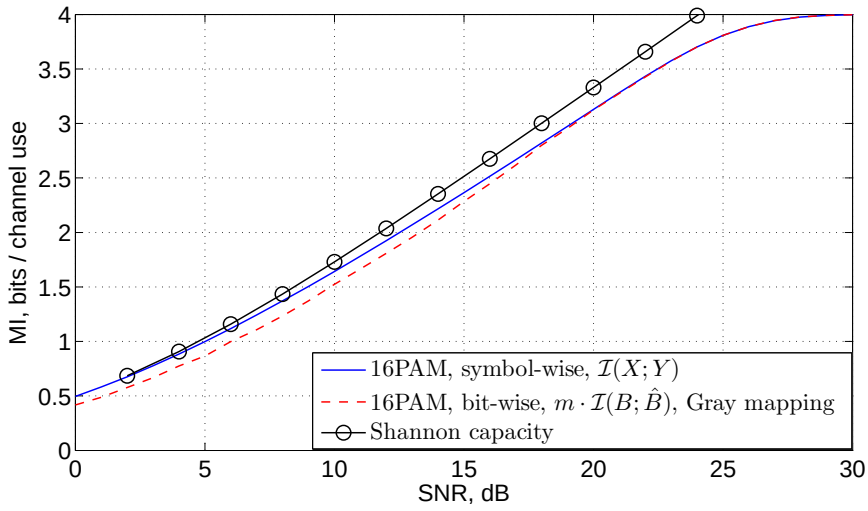


Figure 2.7: Bit-wise vs. symbol wise MI.

Reducing the gap between bit-wise and symbol-wise MI is possible with iterative demapping, where new prior L_a values are provided to the demapper from the decoder. This improvement depends on the extrinsic information transfer (EXIT) function of the demapper. The EXIT function describes a device in terms of the increase of the *extrinsic* information it provides as a function of the prior information it receives. They were first introduced in [65] as a way to characterize the performance of turbo codes with constituent convolutional codes, however, they can be applied to any soft-decision device. In case of the demapper, the extrinsic $L-$values, and thereby the extrinsic information, are found from Equations (2.27) and (2.28) as $L_e(B_l) = L(B_l) - L_a(B_l)$. In general, the extrinsic information can be found as the difference between the complete posterior information and the prior information.

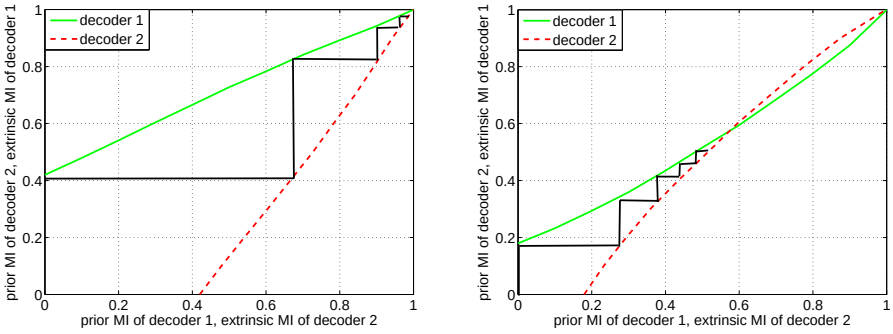
An EXIT chart of a pair of devices is obtained by plotting their EXIT functions against each other, where the extrinsic information of one serves as the a-priori information of the other and vice-versa. Example EXIT chart is given in Fig. 2.8(a) and Fig. 2.8(b) for a turbo code operating on an AWGN channel, where the two devices, iterating between each other are the two constituent convolutional codes. The code rate is $R = 1/3$, and the EXIT functions are plotted for two different

SNRs. The solid black lines represent the evolution of the extrinsic information through iterations. When there is a "tunnel" between the EXIT functions, successful decoding will eventually be achieved, since the MI between the output of one of the decoders and the original uncoded binary signal will be '1'. For low SNR we see the two EXIT functions intersecting before the point $(1, 1)$, and the decoder will therefore "get stuck". EXIT charts are also particularly useful when characterizing the iterative process between the demapper and decoder. We will use this method later in Chapter 5. For more details on how the EXIT charts are generated, the extrinsic information calculation and more, the reader is referred to [65–67].

As we see on Fig. 2.7, increasing the SNR will allow for approaching the symbol-wise AIR without iterations. However, at that point we are moving further away from the channel capacity, due to the limited constellation size. Increasing the constellation size on the other hand will result in increased gap beween bit-wise and symbol-wise MI. Thus we see the inherent trade-off between energy efficiency and complexity of the receiver processing for fixed spectral efficiency. This trade-off can also be seen from the EXIT charts. It can be proven that the area between the two EXIT functions corresponds to the gap to channel capacity for the AWGN channel and the binary symmetric channel [67]. At high SNR, the tunnel will be wide open, and very few iterations will be needed for convergence, however, the area between the EXIT functions will be larger, thus the gap to capacity is also larger.

2.3.2 Turbo coded BICM with constellation shaping

Shannon capacity can be approached with practical codes, such as turbo and LDPC codes, at low SNR (low rates, respectively), and using BPSK / QPSK modulation. However, at high SNR, increased modulation sizes are needed. When the input alphabet is large, uniform input distribution results in a loss to the Gaussian capacity, which is called the *shaping gap*, also known as *shaping gain*. Alternatively, the shaping gap can be described as the energy loss of using uniform input instead of Gaussian at the same information rate. This energy loss can be calculated for constellations of infinite size to be 1.53 dB [68], called the *ultimate* shaping gain.



(a) SNR = 1.5dB (b) SNR = 0dB

Figure 2.8: EXIT chart of a convolutional turbo code for AWGN channel with two different SNRs. Successful decoding is possible when there is an open tunnel (left), and fails otherwise (right).

In order to find the optimal weights of a discrete PMF with a fixed $\mathcal{X}$, Algorithm 1 is used. Doing so on a real-valued AWGN channel results in the CCCs, given in Fig. 2.9. We see that optimizing the discrete PMF gains around 0.8 dB at 8PAM up to 1.4 dB at 64PAM. These gains are directly translated to the complex-valued channel and the constellation equivalents 64QAM and 4096QAM, respectively.

Achieving these gains with practical coded modulation schemes is far from trivial. First attempts were made in the 90's [69, 70]. In [69], a trellis decoding on a few bits of the symbol label was used to select the low-energy points of the constellation more often. Trellis shaping was used in e.g. [71] for mitigation of nonlinearities in optical communications and in [22] for MIMO, where significant shaping gains and near-capacity performance are achieved.

Huffman decoding was proposed in [72] in order to approach a quantized Gaussian distribution. The idea is to use inverse source coding for transformation of the uniform input distribution to a non-uniform output, which is matched to the channel. The problem with such a scheme is the error propagation of the Huffman codes. In this case the shaping must be employed after channel coding, or before channel decoding at the receiver side, which means that the Huffman code will be unpro-

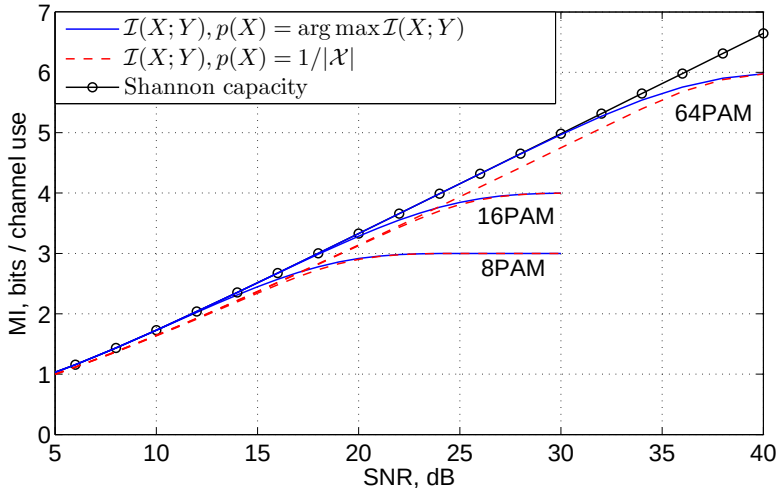


Figure 2.9: Constellation constrained capacity for real-valued AWGN channel.

tected from errors. A single error will therefore make the entire following sequence erroneous.

We mention a few other practical schemes for shaping. In [73], the authors explicitly select low-energy points more often by a nonlinear block code, which can be optimally decoded at the receiver, and achieve significant shaping gains. Similar technique is used in [74] for optical fibers. In [75], the positioning of the points is optimized instead of their probabilities, leading to the so-called *geometric* shaping, as opposed to the probabilistic shaping. Combination of the two approaches is also possible, such as the superposition modulation [76], where the points are found via multiple superpositions of an initial BPSK constellation. This idea was later applied to the optical channel [77]. The target of the above schemes is to approach a quantized Gaussian distribution, or even a simple staircase distribution with only a few steps, for which the low-energy points appear more often. This leads to sub-optimal solution in variety of scenarios. One such case is the high SNR, where the AIR begins to flatten, i.e. close to the point, where the AIR is limited by the entropy of the input. For discrete constellations this entropy is maximized with uniform PMF, making the quantized Gaussian suboptimal.

Recently, the authors in [78] designed a probabilistic shaping scheme,

which for long block lengths can approach any symbol PMF of interest. The authors make use of a modified arithmetic code, circumventing the error-propagation problem by only modifying the probabilities of the systematic bits of a linear block code, thus introducing error protection before the source decoding. The scheme achieves very close to Gaussian capacity for a variety of rates, making it very practical.

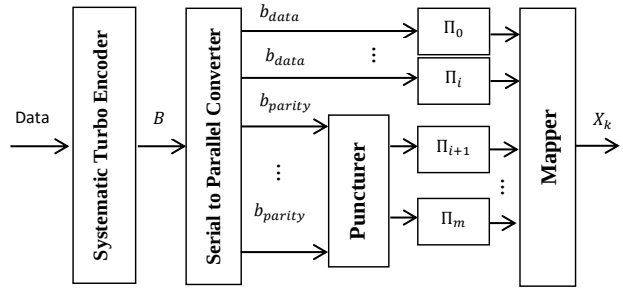
A non-bijective mapping function is used in [79] in order to make some symbols of the constellation more probable. The mapping is combined with a convolutional code and iterative demapping to resolve the ambiguities, introduced during modulation. This is also known as *many-to-one* mapping.

Of particular interest to this thesis is the work from [80], where a turbo coded BICM is used. A fixed many-to-one mapping function for a 16PAM constellation is proposed there. Block diagrams of the transmitter and receiver are given in Fig. 2.10(a) and Fig. 2.10(b), respectively. They represent a general pragmatic turbo coded scheme, where only the mapping function is changed to a non-bijective one. 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. Each stream is then interleaved and the signal is modulated. At the receiver, iterative processing between the decoder and demapper is required, in order to resolve the ambiguities, induced by the non-bijective mapping. Mappings design will be discussed in Chapter 5.

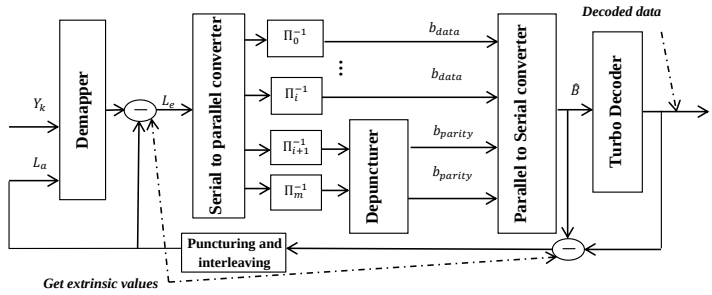
The system was only made for 16PAM, making it unsuitable for higher SNR due to the limited size. Further, the symbol PMF was similar to previously mentioned schemes designed to be quantized Gaussian, and is therefore suboptimal in wide variety of scenarios. However, it still achieves close to the theoretical shaping gain of 1 dB for the 16PAM at the moderate SNR.

2.4 Summary

In this section the major concepts of interest to the thesis were introduced. Information theoretic background and definitions were given for terms like channel capacity, AIR and CCC. The channel models under investigations were introduced, including the MIMO, linear impulse re-



(a) Transmitter



(b) Receiver

Figure 2.10: Pragmatic turbo coded BICM block diagram. Transmitter (a) and receiver (b).

sponse, Wiener phase noise and the nonlinear fiber optic channels. The main issues with estimating the CCC were outlined for each case and the auxiliary channel method for deriving a lower bound and an AIRs was presented. Some state of the art and emerging methods and techniques were described for estimating CCC and AIRs, and also for mitigation of the major impairments.

Section 2.3 outlined the main differences between AIR and what coded modulation schemes can achieve in terms of symbol-wise and bit-wise mutual information. The EXIT chart method was briefly presented, which makes use of the bit-wise MI to predict the performance of iterative receivers in terms of complexity and number of iterations, and also gap to capacity. Finally, the concept of constellation shaping was presented, with one particular method for shaping of uniformly distributed

binary data, which will be of further interest in this thesis.

Chapter 3

Capacity estimation and bounding techniques

In this chapter the novel contributions are described regarding constellation constrained capacity (CCC) estimation. We begin our investigation with the linear interference channels - multiple-input multiple-output (MIMO) and impulse response channel. Some new properties are derived for the optimal input probability mass function (PMF) of MIMO channels and good approximations are proposed for finding it, and thereby the CCC. Novel lower bounds are derived for both MIMO and impulse response channel via the QR decomposition (QRD) of the channel matrix form. Finally, we discuss the nonlinear optical fiber channel and derive new lower bounds based on a modified Blahut-Arimoto algorithm (BAA). Some intuition about the corresponding PMF is also provided in this case.

Unless otherwise stated, quadrature amplitude modulation (QAM) constellation input will be assumed, which is obtained by a product of two identical pulse amplitude modulation (PAM) constellations. We will also assume that the matrix $\mathbf{H}$ has full rank.

We focus on methodology rather than detailed step-by-step mathematical derivation. Theorems are therefore proven only in sketches. The complete proofs can be found in the corresponding papers.

3.1 Linear channels

The novel contributions of papers **PAPER 1**, **PAPER 4** and **PAPER 7** are described in this section. These papers deal with linear interference channels with discrete input, and derive methods for estimating the CCC and the optimal input PMF for fixed constellation $\mathcal{X}$. Of particular interest are the cases, where the large dimensionality of the input prevents Algorithm 1 from being used.

3.1.1 Factorization properties of the optimal PMF

The main contribution of **PAPER 1** is to prove that the optimal input PMF to the MIMO channel factorizes into its marginals. This has been posted as a conjecture in [22], where Algorithm 1 was modified to cover the MIMO channel. Using this conjecture, the authors were able to reduce the complexity by minimizing the degrees of freedom in the optimization process.

We use a two-step proof for this conjecture, given by the two theorems below. The basic idea is to show that if the BAA is initialized with a factorized distribution, the distribution on the next step is also factorized. We recall the notation of the MIMO channel from Chapter 2, where $\mathbf{H}$ is the $[N\mathbf{x}M]$ channel matrix, X is the $M-$dimensional input with discrete PMF $p_X(X)$, and Y is the $N-$dimensional output. We also denote $p_{X_i}(X_i)$ to be the marginal PMF of the $i-$th dimension, which we also refer to as the $i-$th MIMO *layer*. The PMF with optimal weights is denoted $p_X^*(X)$ for the respective channel. We will express the mutual information (MI) both in the standard form $\mathcal{I}(X;Y|\mathbf{H})$ and as a function of the input PMF - $\mathcal{I}(p_X(X))$ or $\mathcal{I}(p_X(X)|\mathbf{H})$, to be differentiated from the context.

Theorem 1 *If* $p_X(X) = \prod_{i=1:M} p_{X_i}(X_i)$, *then for fixed* $p_{X|Y}(X|Y) = \frac{p_X(X)p_{Y|X}(Y|X)}{p_Y(Y)}$ *the conditional entropy on each layer becomes* $\mathcal{H}(X_i|Y,\mathbf{H},X_{\{1:M\}\setminus i}) = \mathcal{H}(X_i|Y,\mathbf{H})$.

Proof sketch. The proof relies on the fact that the conditional entropy may be expressed as a Kullback-Leibler divergence (KLD) between two distributions. The dependency on $X_{\{1:M\}\setminus i}$ is expressed by an equal linear shift of the mass of both distributions, which does not change the

KLD between them. Details are provided in the Appendix of **PAPER 1**. ■

Theorem 2 *If* $p_X^n(X) = \prod_{i=1:M} p_{X_i}^n(X_i)$ *at step* n *of the BAA, then* $p_X^{n+1}(X) = \prod_{i=1:M} p_{X_i}^{n+1}(X_i)$ *at step* $n+1$*, and by induction* $p_X^*(X) = \arg\max \mathcal{I}(X;Y|\mathbf{H}) = \prod_{i=1:M} p_{X_i}^*(X_i)$

Proof sketch. We assume the optimal PMF on step $n+1$ is $p_X^{n+1}(X)$, and the corresponding MI is $\mathcal{I}(p_X^{n+1}(X))$. After Theorem 1, the conditional entropy $\mathcal{H}(X|Y,\mathbf{H})$ on step n for fixed $p_{X|Y}(X|Y)$ can be expressed as a sum of functions of the marginal PMFs $p_{X_i}^n(X_i)$. Then if we take the PMF, which is found as the product of the marginal PMFs $\hat{p}_X(X) = \prod_{i=1:M} p_{X_i}^{n+1}(X_i)$, we can increase the MI

$$\begin{aligned}\mathcal{I}(\hat{p}_X(X)) &= \sum_{i=1:M} \mathcal{H}(X_i) - \mathcal{H}(X|Y,\mathbf{H}) \\ &\geq \sum_{i=1:M} \mathcal{H}(X_i|X_{1:i-1}) - \mathcal{H}(X|Y,\mathbf{H}) \\ &= \mathcal{I}(p_X^{n+1}(X)). \end{aligned} \tag{3.1}$$

Since we assumed that $\mathcal{I}(p_X^{n+1}(X))$ is the optimum, and due to the strict concavity in the input PMF of the MI [18], the distributions $p_X^{n+1}(X)$ and $\hat{p}_X(X)$ must coincide, which proves the theorem. ■

The BAA achieves unique optimum, which must also be factorized. This property allows for optimization only on the basic 1D distribution.

3.1.2 Estimating the CCC in the limit of large constellations

Since the output is multi-dimensional, calculation of MI and achievable information rate (AIR)s still requires much effort due to the increasing number of required samples K for the convergence in Eq. (2.2). Reducing this number is possible when the interference channel can be parallelized, thus factorizing the problem into several layers of lower dimensionality. This can be achieved for MIMO channels via the singular value decomposition (SVD). It is defined as

$$\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^H, \tag{3.2}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, and $\mathbf{S}$ is diagonal. If the input is pre-coded as $\hat{X} = \mathbf{V}X$ and the received signal is pre-processed as $\hat{Y} = \mathbf{U}^H Y$, the resulting channel is a set of parallel Gaussian channels $\hat{Y}_i = \mathbf{S}_{ii}X_i + W_i$, where $\mathbf{S}_{ii}$ is the element on the $i-th$ row and $i-$th column of $\mathbf{S}$. The capacity is then found by water-filling power allocation and Gaussian input. When the channel is not known at the transmitter power allocation cannot be performed. Furthermore, as discussed in Chapter 2, water-filling is sub-optimal for discrete input constellations.

We define the Gaussian distribution of X as $p_G(X)$. Discrete PMFs are defined as (2.9), and the notation $p_X(\mathbf{A}X)$, where A is full rank, will denote a PMF with mass points linearly shifted by $\mathbf{A}$, with the original weights intact. We prove the following theorems for the CCC, derived in **PAPER 4**.

Theorem 3 $\lim_{|\mathcal{X}|\to\infty} \mathcal{I}(p_X^*(X)|H=\mathbf{H}) = \mathcal{I}(p_G(X)|H=\mathbf{H})$

Proof sketch. The following steps are taken:

1. Show that for orthogonal channels $\lim_{|\mathcal{X}|\to\infty} \mathcal{I}(p_X^*(X)|H=\mathbf{S}) = \mathcal{I}(p_G(X)|H=\mathbf{S}) \Leftrightarrow \lim_{|\mathcal{X}|\to\infty} W_2(p_X^*(X), p_G(X)) = 0$, following the AWGN results from [24] ($W_2(\mu, v)$ was defined in Eq. (2.11)).

2. Show that

$$\lim_{|\mathcal{X}|\to\infty} W_2(p_X^*(\mathbf{V}^H X), p_G(\mathbf{V}^H X)) = \lim_{|\mathcal{X}|\to\infty} W_2(p_X^*(X), p_G(X)) = 0, \tag{3.3}$$

   from the fact that the Gaussian distribution is rotationally invariant.

3. Show that

$$\lim_{|\mathcal{X}|\to\infty} W_2(p_X^*(\mathbf{V}^H X), p_G(\mathbf{V}^H X)) = 0 \tag{3.4}$$

$$\Leftrightarrow \lim_{|\mathcal{X}|\to\infty} W_2(p_X^*(\mathbf{S}\mathbf{V}^H X), p_G(\mathbf{S}\mathbf{V}^H X)) = 0 \tag{3.5}$$

$$\Leftrightarrow \lim_{|\mathcal{X}|\to\infty} \mathcal{I}(p_X^*(X)|H=\mathbf{H}) = \mathcal{I}(p_G(X)|H=\mathbf{H}). \tag{3.6}$$

■

This result is extended to the ergodic case in the following.

Theorem 4 $\lim_{|\mathcal{X}|\to\infty} \mathbb{E}_{\mathbf{H}}\left[\mathcal{I}(p_X^*(X)|H=\mathbf{H})\right] = \mathbb{E}_{\mathbf{H}}\left[\mathcal{I}(p_G(X)|H=\mathbf{H})\right] = C_e$, *where* C_e *is defined as (2.8)*

Proof sketch. The proof relies on the fact that the W_2 distance is a distance measure. For the optimal distributions p_i^* and p_j^* of channel realizations $\mathbf{H}_i$ and $\mathbf{H}_j$, respectively, we then have

$$\lim_{|\mathcal{X}|\to\infty} W_2(p_i^*, p_j^*) \leq \lim_{|\mathcal{X}|\to\infty} W_2(p_i^*, p_G) + \lim_{|\mathcal{X}|\to\infty} W_2(p_j^*, p_G) = 0 + 0. \tag{3.7}$$

The distribution on any channel realization is therefore optimal on all other channel realizations, and provides the rate of the Gaussian distribution, which proves the theorem. ■

The main implication of Theorems 3 and 4 is that if some PMF achieves close to the rate of the Gaussian distribution on the equivalent orthogonal channel, obtained by the SVD, that PMF will also achieve close to the Gaussian rate on the interference channel. The PMF and the CCC are straightforward to find on an orthogonal channel due to the complete factorization - the PMF on each layer is found by Algorithm 1, and the CCC is the sum of the CCC on each layer. We can then easily identify the signal-to-noise ratio (SNR) regions, where the above claim is true.

3.1.3 QR decomposition based lower bound

The approximations derived in the previous section are usually accurate for low-to-mid SNR. It is of interest to find good approximations, or at least good lower bounds, in the mid-to-high SNR as well. We derive such bounds based on the QRD of the channel, defined as

$$\mathbf{H} = \mathbf{QR}, \tag{3.8}$$

where $\mathbf{Q}$ is unitary and $\mathbf{R}$ is upper triangular. The *successive interference cancellation* MIMO receiver utilizes the form of $\mathbf{R}$ to cancel the interference from previously detected layers in the following manner: the received samples are pre-processed as $\hat{Y} = \mathbf{Q}^H Y$, and the channel model becomes $\hat{Y}_i = \sum_{j=i}^{M} \mathbf{R}_{i,j} X_j$. Assuming the layers from $i+1$ to M are correctly decoded by the following channel code, the symbols can be re-modulated and subtracted from the current layer i. Here we use a similar technique to derive a lower bound on the channel capacity.

Since $\mathbf{Q}$ is unitary and does not change the entropy of Y, and thus the MI, we can write

$$\begin{aligned}\mathcal{I}(X;Y|\mathbf{H}) &= \mathcal{H}(X) - \mathcal{H}(X|\hat{Y},\mathbf{H}) \\ &= \mathcal{H}(X) - \sum_{i=1:M} \mathcal{H}(X_i|\hat{Y}, X_{i+1}^M, \mathbf{H}) \\ &\geq \mathcal{H}(X) - \sum_{i=1:M} \mathcal{H}(X_i|\hat{Y}_i, X_{i+1}^M, \mathbf{H}),\end{aligned} \tag{3.9}$$

where we have used the fact, that conditioning does not increase the entropy. In order to calculate the terms in the sum, we express the posterior probabilities as

$$p(X_i|\hat{Y}_i, X_{i+1}^M, \mathbf{H}) = \frac{p(X_i)p(\hat{Y}_i|X_i, X_{i+1}^M, \mathbf{H})}{\sum_{X_i} p(X_i)p(\hat{Y}_i|X_i, X_{i+1}^M, \mathbf{H})} \tag{3.10}$$

Since we condition on the following layers, the likelihood above can be expressed as

$$\begin{aligned}p(\hat{Y}_i|X_i, X_{i+1}^M, \mathbf{H}) &= \mathcal{N}(\sigma^2, \sum_{j=i:M} R_{i,j}X_j; \hat{Y}_i) \\ &= \mathcal{N}(\sigma^2, R_{i,i}X_i; \hat{Y}_i - \sum_{j=i+1:M} R_{i,j}X_j),\end{aligned} \tag{3.11}$$

where $R_{i,j}$ is the element on the $i-$th row and $j-$th column of $\mathbf{R}$. Using (3.11), lower bound on the MI on each layer can be calculated independently. When $M <= N$, the achievable rate on the $M-$th layer coincides with the actual capacity for that layer. However, when $M > N$, there is residual interference on the $(N+1)$-st to the $M-$th layers from layers, which are not yet decoded, and the resulting lower bound becomes poorer. In order to improve it, we model the residual interference as noise, which is a standard practice in communications engineering. The likelihood we use on layers $i > N$ is then:

$$\mathcal{N}(\hat{\sigma}_i^2, R_{N,i}X_i; \hat{Y}_N - \sum_{j=i+1:M} R_{N,j}X_j), \tag{3.12}$$

where $\hat{\sigma}_i^2 = \sigma^2 + \sum_{j=N:i-1} |R_{N,j}|^2 \mathrm{E}\left[X_i^2\right]$.

Alternatively, the same lower bound can be derived via the auxiliary channel method. If we use the auxiliary probability distribution $\bar{p}(X|Y,\mathbf{H})$

$$\bar{p}(X|\hat{Y},\mathbf{H}) = \prod_{i=1}^{M} \bar{p}(X_i|\hat{Y},X_{i+1}^{M},\mathbf{H})$$
$$= \prod_{i=1}^{M} \frac{\bar{p}(\hat{Y}|X_i,X_{i+1}^{M},\mathbf{H})p(X_i)}{\sum_{X_i}\bar{p}(\hat{Y}|X_i,X_{i+1}^{M},\mathbf{H})p(X_i)},$$

where

$$\bar{p}(\hat{Y}|X_i,X_{i+1}^{M},\mathbf{H}) =$$
$$\mathcal{N}(\sigma^2,\mathbf{R}_{i,i}X_i;\hat{Y}_i - \sum_{j=i+1:M}\mathbf{R}_{i,j}X_j), \tag{3.13}$$

the same rate is achieved. Details about the above derivations are found in **PAPER 4** and **PAPER 7**.

On Fig. 3.1, CCC approximations are given for several constellation sizes for 8x8 MIMO matrix, which is too complex for the general optimization, and even for MI calculation on a standard computer. Instead, we have used the factorization properties of the optimal input to reduce the degrees of freedom in the optimization, which is then performed on per-layer basis. We report the CCC estimates, obtained via the SVD approximations, and the lower bounds, obtained via the QR decomposition based approach [1]. We can directly read the regions, where the SVD based approximations are accurate: up to SNR = 10dB, 16dB and 24dB for 64QAM, 256QAM and 1024QAM, respectively. This means that the true CCC also approaches Gaussian capacity up to those points. At high SNR, the SVD based approximation becomes worse than the QRD based lower bound. This comes to show that orthogonalizing the channel with discrete inputs can be sub-optimal. The envelope of the two curves can be generally used as a good approximation to the CCC.

The lower bounds are directly extended to the linear impulse response channel. The QRD in this case is performed on the matrix form

[1] Since the QRD method provides lower bounds on each BAA iteration, the final result is a lower bound on the CCC

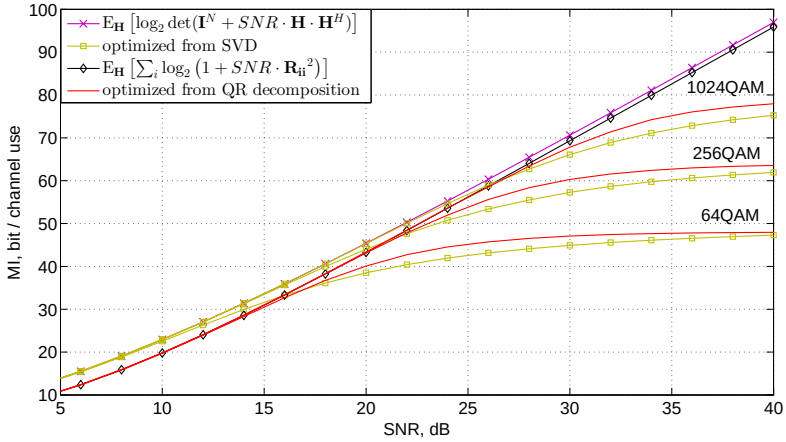


Figure 3.1: CCC approximations via the SVD method and the QRD method. The SNR regions, where the SVD based approximation is true can be directly read from the curves as the regions, where Gaussian capacity is approached.

of the channel (2.13), and the MI is calculated for each of the K channels separately. Since the number K is usually large, an SNR-MI look-up table (LUT) can be used, where the SNR on the $i-$th layer is defined as SNR$=|R_{i,i}|^2\mathbb{E}\left[X_i^2\right]/\sigma^2$. Some results for this case are given in **PAPER 7** for Rayleigh fading channels of different length and 16QAM input. Here we mention that the same behavior as with MIMO channels is observed - at high SNR, the QRD based lower bounds outperform the AIR of orthogonal frequency division multiplexing (OFDM) with uniform power allocation, which once again shows the sub-optimality of orthogonalization of the channel with discrete input.

3.2 Non-linear optical fiber channel

As mentioned in Chapter 2, the nonlinear interaction between chromatic dispersion (CD), amplified spontaneous emission (ASE) noise and self-phase modulation (SPM) and cross-phase modulation (XPM) effects makes the capacity estimation of the nonlinear optical fiber channel difficult. This is mainly due to the memory, introduced in the channel. When operating on memoryless assumption, the lower bounds in Fig. 2.5 are

calculated using ring constellation input [51]. Here we extend the idea of probabilistic shaping to cover this scenario.

3.2.1 Blahut-Arimoto algorithm for the nonlinear channel

We modify the BAA from Section 2.2.2 to Algorithm 3. In Step 4, the mean and covariance matrices are found from the sample mean and covariance matrix of the sequence $Y_{\mathcal{K}} : \mathcal{K} \equiv \{k : X_k = x_i\}$, i.e. the samples Y_k, where $X_k = x_i$. The optimization over the scaling factor α is as in Algorithm 3, and is omitted for simplicity.

Algorithm 3 Algorithm for finding the optimal PMF on a nonlinear optical fiber channel

Initialize: $p_X(X)$, such that the power constraint $\mathbb{E}\left[X^H X\right] \leq P_{av}$ is satisfied and $\sum_{i=1:|\mathcal{X}|} w_i = 1$

1: **while** converged $p_X(X)$ **do**
2: Generate $x_1^K \sim p_X(X)$
3: Generate y_1^K by solving the NLSE via the SSFM
4: Estimate $p(Y|X = x_i) = \mathcal{N}(\Sigma_i, \mu_i; [\mathrm{Re}[Y], \mathrm{Im}[Y]]^T)$
5: $p_{X|Y}(X|Y) = \frac{p_{X,Y}(X,Y)}{\sum_{X_i \in \mathcal{X}} p_{X,Y}(X_i,Y)}$
6: $p_X(X) = \arg\max_{p_X(X)} \mathcal{I}(Y;X)$, s.t. $\mathbb{E}\left[X^H X\right] \leq P_{av}$ and $\sum_{i=1:|\mathcal{X}|} w_i = 1$
7: **end while**
8: $C_{LB} = \mathcal{I}(p_X(X))$

The Gaussian fit represents an auxiliary channel. If we assume that the split-step Fourier method (SSFM) is an accurate representation of the fiber optic channel, the MI on each iteration is an AIR, and the final MI C_{LB} is a lower bound on the capacity. These AIRs are given in Fig. 3.2 for an Erbium doped fiber amplifier (EDFA) link with parameters, given in Table 3.1. The fiber parameters are the same as in Table 2.1. Both polarizations are employed, and we report the AIRs as a function of the input power *per channel per polarization*. We compare to the AIRs with conventional uniformly distributed QAM constellations and the AIRs with ring constellations, optimized as in [51] with 32 rings

Table 3.1: System parameters, EDFA transmission

EDFA noise figure	$F = 3$ dB
Symbol rate	28 GBaud
Pulse shape	Raised cosine, roll-off factor - 0.01
Span length	80km
Total distance	800km

and 1024 points. We see that the maximum AIR can be increased by around 0.5 and 0.2 bits/s/Hz/polarization w.r.t. uniform QAM and ring constellations, respectively. Probabilistic shaping of QAM constellations is more attractive than the geometric optimization of ring constellations for several reasons:

1. The optimization is convex and can be performed via numerical methods (as in Algorithm 3).
2. The original quadrature shape is retained, which allows for standard I/Q modulation/demodulation.
3. Straight-forward Gray mapping can be applied.

Some optimal 1024QAM PMFs are given in Fig. 3.3(a) for the optimal input power of -6 dBm and in Fig. 3.3(b) at the highly nonlinear region of 0 dBm, together with the quantized Gaussian PMF. Only the marginal 1D PMFs are shown, i.e. 32PAM. The points are rescaled so that the same average power is retained with different PMFs. At -6 dBm, the optimal PMF is already slightly non-Gaussian. In the highly nonlinear region this is even more pronounced, as the PMF converges to only a few levels and the points with largest amplitude are given a lot of mass. This is in contrast to shaping for the linear region, where quantized Gaussian is near-optimal. Operating QAM constellations with an optimized PMF will be the subject of Chapter 5.

3.3 Summary

In this section the contributions of this thesis regarding capacity estimation and lower bounding for several channels of interest were presented.

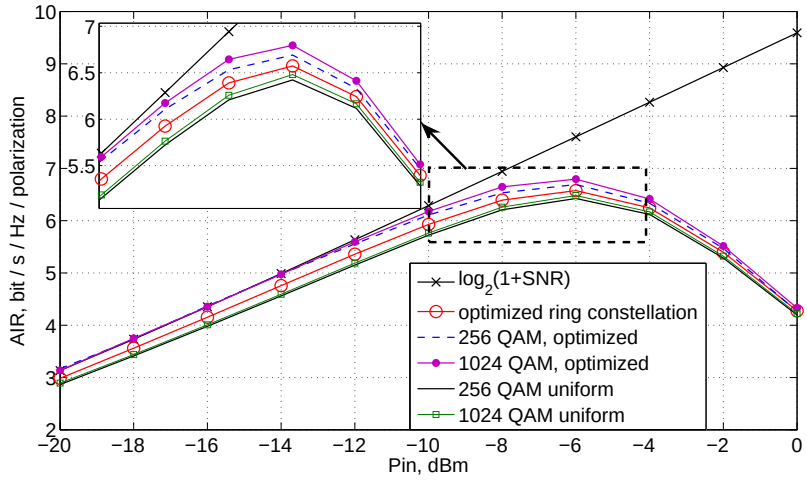
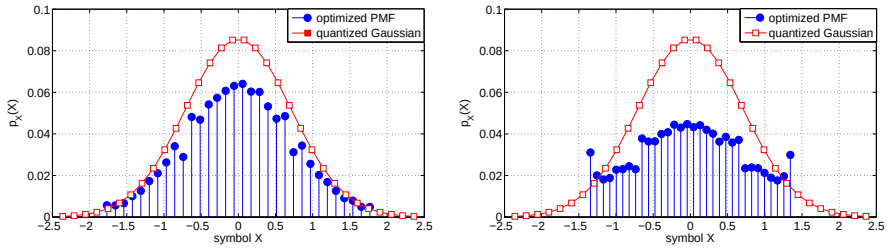


Figure 3.2: AIRs with probabilistic optimized QAM constellations vs. optimized ring constellations [51] and uniform QAM constellations.



(a) Pin = -6 dBm (b) Pin = 0 dBm

Figure 3.3: Optimal 1024QAM PMFs for the optimal and highly nonlinear input power regime.

These include the linear interference channels, such as MIMO and linear impulse response channels, but also the nonlinear optical fiber channel. Factorization properties were derived for the optimal signaling distribution of QAM constellations for MIMO, and SVD based approximations were derived, which are particularly useful in the low-to-mid SNR re-

gions. For the higher SNR, QRD based lower bounds were derived. In the case of optical fiber, the BAA was extended to cover the nonlinear channel, and improvements were shown with probabilistic shaped QAM constellations over conventional uniformly distributed QAM constellations and also ring constellations.

Chapter 4

Phase noise mitigation in digitally modulated systems

In this chapter we study the Wiener phase noise channel model. We investigate some properties of the nonlinearities in optical fibers and show that the first order Wiener process is a suitable model for the cross-phase modulation (XPM) induced non-linear phase noise (NLPN). Then, we propose a practical method for tracking the combined effect of NLPN and regular laser linewidth (LLW) induced phase noise. The main results are available in **PAPER 5** and **PAPER 6**.

4.1 The nonlinear XPM noise

The self-phase modulation (SPM) and XPM were extensively studied in [54], where a model for the auto-correlation function (ACF) of the XPM induced phase noise was proposed. The authors show that the NLPN is highly correlated in time and frequency. Exploiting this correlation, it was shown in [55,57] that the achievable information rate (AIR)s with Gaussian input to the optical fiber can be significantly improved by subtracting the previously estimated NLPN values with a genie approach, similar to Eq. (2.21), where instead of decisions the actual samples x_k are used. Using this approach also allows for computing the empirical ACF, which was shown to agree with the theoretical model in point-to-point links [55].

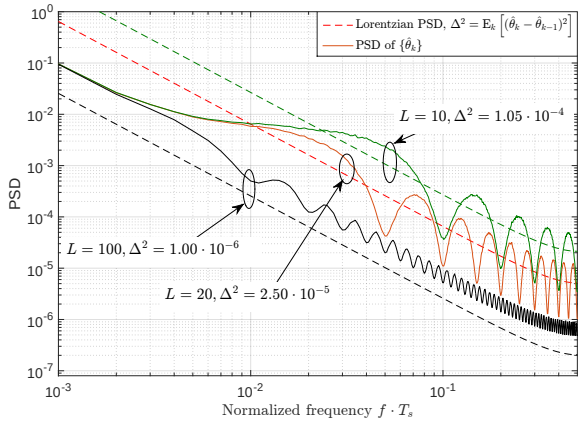


Figure 4.1: Lorentzian PSD of a Wiener process (dashed lines) with process noise variance as in (4.1), together with the PSD of the phase noise estimates $\{\hat{\theta}_k\}$ (solid lines) for a WDM link as described in the text. Depending on the choice of the window L a good match can be found to the theoretical model.

In order to validate the Wiener process for modeling the NLPN, we study the power spectral density (PSD) of the sequence $\{\hat{\theta}_k\}$. The sequence is estimated from the above mentioned genie approach for a wavelength division multiplexing (WDM) single polarization system with 5 channels, 100 GBaud each, after 40 Raman amplification spans of 100km. The input constellation is 256QAM, and the input power is optimized to -4 dBm per channel. In order to estimate the XPM induced noise only, we perform digital back-propagation (DBP) for the channel of interest, thus removing the SPM contribution. Empirical PSDs are given in Fig. 4.1 for several lengths L of the sliding window averaging. We compare them to the Lorentzian PSD of a true Wiener process with parameter

$$\Delta^2 = \mathbb{E}_k\left[(\hat{\theta}_k - \hat{\theta}_{k-1})^2\right]. \tag{4.1}$$

We see that the slope of the empirical PSD closely follows the Lorentzian function. For discussion on the window length and the periodic behavior of the empirical PSD, please refer to **PAPER 6**.

We conclude this section with a note that the correlation properties of the NLPN become much less significant in a WDM network, where

channels are added and dropped on the way. These properties were studied in **PAPER 5**.

4.2 Wiener phase noise tracking based on Tikhonov distributions

The general framework for mitigating the phase noise in digital systems is to first estimate the phase noise samples $\hat{\theta}_k$ with some algorithm, preprocess the received samples as $\hat{y}_k = y_k \cdot \exp(-j\hat{\theta}_k)$, and then operate the receiver on a memoryless assumption. Most classes of such receivers were briefly presented in Section 2.2.4. Here we propose a method that directly estimates the posterior probabilities $p(x_k|y_1^K)$, thus taking into account the memory of the channel, and implicitly calculates the posterior probability density function (PDF) of the phase noise values $p(\theta_k|y_1^K)$. The basic idea is to model the forward and backward recursions, $p(\theta_k|y_1^{k-1})$ and $p(y_{k+1}^K|\theta_k)$, respectively, by mixtures of Tikhonov distributions. The general belief propagation inference model can then be applied for finding the posteriors. The method is thus closely related to the Kalman filter [36], the trellis model [33] and the phase noise tracking method for phase shift keying constellations from [38].

Details of the method can be found in **PAPER 6**. Here we briefly describe the differences to the other mentioned models and the benefits of using the mixture of Tikhonov distributions. The factor graph that we use for inference is given in Fig. 4.2, where circles represent variables, and rectangles represent factors. The directions of the messages is indicated by the arrows, and the messages themselves are calculated as

- $\xi_1 = p(\theta_{k-1}|y_1^{k-1}) = \sum_{m=1}^{M} \alpha_{m,k-1} t(w_{m,k-1}; \theta_{k-1})$
- $\xi_2 = p(\theta_k|y_1^{k-1}) = \int_{-\pi}^{\pi} \xi_1 \cdot p(\theta_k|\theta_{k-1}) d\theta_{k-1}$
- $\xi_3 = p(y_{k+1}^K|\theta_k) = \int_{-\pi}^{\pi} \xi_4 \cdot p(\theta_{k+1}|\theta_k) d\theta_{k+1}$
- $\xi_4 = p(y_{k+1}^K|\theta_{k+1}) = \sum_{n=1}^{N} \beta_{n,k+1} t(u_{n,k+1}; \theta_{k+1})$
- $\xi_5 = \xi_2 \cdot \xi_3$
- $\xi_6 = \int_{-\pi}^{\pi} \xi_5 \cdot p(y_k|\theta_k, x_k) d\theta_k$

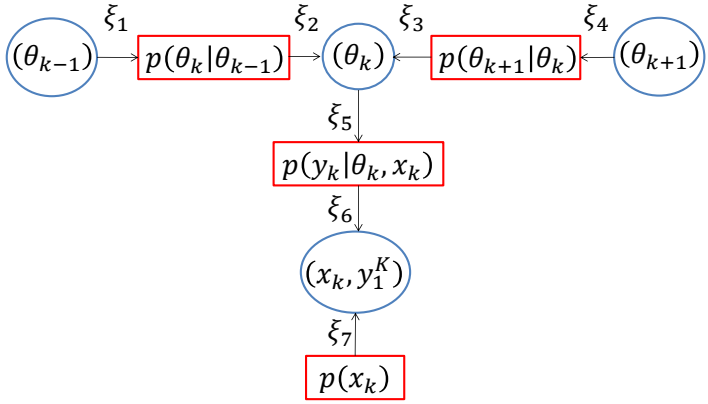


Figure 4.2: Graph of a first-order hidden Markov model, used for inference on the Wiener phase noise channel. The direction of the messages is indicated by the arrows.

- $\xi_7 = p(x_k)$

The desired posteriors are finally calculated as

$$p(x_k|y_1^K) \propto p(x_k, y_1^K) = \xi_6 \cdot \xi_7. \tag{4.2}$$

The distribution $p(y_k|\theta_k, x_k) \propto \mathcal{N}(\sigma^2, x_k \exp(j\theta_k); y_k)$. The coefficients $\alpha_{m,k}$ and $\beta_{n,k}$ above satisfy $\sum_{m=1}^{M} \alpha_{m,k} = 1$ and $\sum_{n=1}^{N} \beta_{n,k} = 1$. The factors $p(\theta_k|\theta_{k-1})$ are simply the Gaussian PDFs $\mathcal{N}(\Delta^2, \theta_{k-1}; \theta_k)$. The Tikhonov distribution at θ with complex parameter w in messages ξ_1 and ξ_4 is defined as

$$t(w;\theta) = \frac{\exp(\mathrm{Re}\,[w \cdot \exp(-j\theta)])}{2\pi I_0(|w|)}, \qquad \theta \in [-\pi;\pi), \tag{4.3}$$

and 0 elsewhere, and I_0 is the zero-th order modified Bessel function of the first kind. When the phase is discrete as in [33], the integration in the messages ξ_2, ξ_3 and ξ_6 must be done numerically, which is very complex for fine quantization. If the messages are represented by a Gaussian, the function under the integral in the message ξ_6 is a product of Gaussians, where $p(y_k|\theta_k, x_k)$ is not a linear function of the phase. The integral is therefore not solved in closed form.

Algorithm 4 Algorithm for phase noise tracking, based on Tikhonov parametrization

Initialize:

$\alpha_{m,0} = 1/M$
$\beta_{n,K+1} = 1/N$
$w_{m,0} = \frac{1}{\Delta^2} e^{j(m \cdot \frac{2\pi}{M} - \pi)}$
$u_{n,K+1} = \frac{1}{\Delta^2} e^{j(n \cdot \frac{2\pi}{N} - \pi)}$
for $k = 1$ **to** K **step** 1 **do**
    messages ξ_1 and ξ_2 ▷ Forward recursion
end for
for $k = K$ **to** 1 **step** -1 **do**
    messages ξ_3 and ξ_4 ▷ Backward recursion
    messages ξ_5 and ξ_6
    $p(x_k|y_1^K) \propto p(x_k, y_1^K) = \xi_6 \cdot \xi_7$ ▷ Posterior calculation
end for

Tikhonov parametrization allows the likelihood to be represented directly as a distribution in the phase as

$$p(y_k|x_k, \theta_k) = \frac{2 \cdot \mathrm{SNR} \cdot I_0(2 \cdot \mathrm{SNR}|y_k x_k^*|) t(2 \cdot \mathrm{SNR} \cdot y_k x_k^*; \theta_k)}{\exp(\mathrm{SNR}(|y_k|^2 + |x_k|^2))}. \quad (4.4)$$

Further, the convolution of a Tikhonov and a Gaussian is again a scaled Tikhonov of the same variable, making the analytical integration and coefficient updates in messages ξ_3 and ξ_4 straightforward. The complete algorithm is summarized in Algorithm 4. The initialization values are chosen such that each Tikhonov component corresponds to a Gaussian distribution of the phase with variance Δ^2, and the means of the components are uniformly spaced within $[-\pi; \pi)$.

The AIRs obtained with this method are shown in Fig. 4.3 for a selected dual polarization Erbium doped fiber amplifier (EDFA) fiber link with parameters, given in Table 4.1. In this case DBP is not performed, only chromatic dispersion compensation in frequency domain. Extensive simulations results are found in **PAPER 6**. We compare with the trellis method [33] with 256 states (256 quantization levels of the phase), and a pseudo-ideal phase noise removal, which is obtained by a genie approach. We see that the trellis based AIRs are closely approached with

Table 4.1: System parameters, EDFA transmission for phase noise tracking performance estimation

Symbol rate	28 GBaud
Number of channels	17
Guardband	0.56 GHz (2% of symbol rate)
EDFA noise figure	4 dB
LLW	100 kHz
Span length	100km

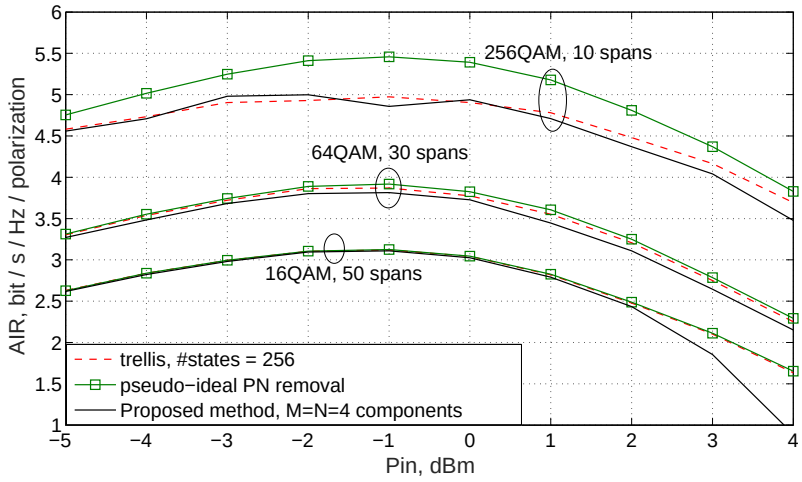


Figure 4.3: AIRs with the proposed phase noise tracking method, compared with the trellis method [33] and a genie approach phase noise removal.

the proposed algorithm at much lower complexity. General decision directed (DD) methods cannot be used at such LLWs and received signal-to-noise ratio (SNR)s due to the very high symbol error rate (SER), as discussed in Section 2.2.4.

Finally in this section we note the $\pi/2$ ambiguity of the quadrature amplitude modulation (QAM) constellations. In mild phase noise scenarios, if an initial phase value is reliably estimated and no phase-slips are expected, the algorithm performs really well blindly. However, in

more severe phase noise cases, minor pilot rate is needed for ensuring the phase does not slip to a neighboring mixture component. This pilot rate in the above EDFA scenario was kept at 0.5%.

4.3 Summary

In this chapter the novel contributions of the thesis were presented regarding Wiener phase noise channels. We showed that the Wiener process is suitable for modeling the NLPN in optical fiber systems via PSD fitting to a Lorentzian spectrum. We then proposed a novel method for tracking of the phase noise based on this model. The method relies on Tikhonov parametrization of the phase, and is able to effectively combat the combined effect of laser phase noise and the NLPN in optical fibers.

Chapter 5

Probabilistic shaping of QAM Constellations

In this chapter we describe the contributions of this thesis to the field of constellation shaping. As mentioned in Chapter 2, when high spectral efficiency (SE) is targeted, constellation shaping must be performed in order to approach capacity. We will mainly be interested in probabilistic shaping for quadrature amplitude modulation (QAM) constellations, and particularly - the many-to-one scheme, which was first proposed in [80]. We propose a rate adaptive method for designing the mapping function, which is key to achieving shaping gain. Especially in dynamic scenarios, where the channel conditions change rapidly, rate adaptation is a necessary feature of the communication system. In static channel conditions the rate flexibility is also necessary in order to cover a wide range of high SE services.

We will analyze the potential shaping gains in a linear additive white Gaussian noise (AWGN) channel, but also in an optical fiber system. In the latter case energy efficiency gains and increased maximum reach of the link for fixed SE will be of interest.

Most of the contributions of this Chapter can be found in **PAPER 2** and **PAPER 3**. Section 5.3 contains also currently unpublished material, mainly in terms of results.

5.1 The mapping function

The many-to-one method basically means that multiple bit sequences are mapped to the same modulation symbol. One way to achieve this is to map all sequences of certain length with the same prefix to the same symbol, leading to the Huffman decoding procedure [81], which was originally used for source coding. The suffix of the label is filled up with the remaining bits of the sequence, which are treated as ambiguous at the receiver, that is both '1' and '0'. Their L values are thus zero on the first demapping iteration. The bit sequence length for each symbol in this case is generally $m \geq \log_2 |\mathcal{X}|$. The block diagrams of the transmitter and receiver are given in Fig. 2.10(a) and Fig. 2.10(b), respectively.

Since all labels have the same length, the synchronization and error propagation problems of source codes are circumvented. The resulting probability mass function (PMF)s have *dyadic* form: $p_X(X = x_i) = 2^{-l_i}$, where l_i is the length of the unique prefix of the $i-$th symbol. The optimal dyadic approximation to a PMF was found in [82] to be achieved by the geometric Huffman codes (GHC).

A given dyadic PMF generally can be assigned different bit labellings. In **PAPER 2**, a labeling method was proposed that achieves Gray-like property of the unique parts of the labels. The design of a mapping function for a certain channel (including the nonlinear optical fiber) is as follows:

1. Obtain the optimal PMF via Algorithm 1 (Algorithm 3 in case of nonlinear channels).

2. Run the GHC to obtain a dyadic approximation to the optimal PMF.

3. Run the algorithm from **PAPER 2** to obtain the labeling.

All mapping functions that we use in this thesis are given in Appendix B. We design only the pulse amplitude modulation (PAM) mapping functions for the marginal 1D PAM PMFs, and the QAM labeling is obtained by taking an outer product of two PAM labelings.

The combination of the Gray-like property with the fact that ambiguous bits are inserted along the stream makes this labeling scheme

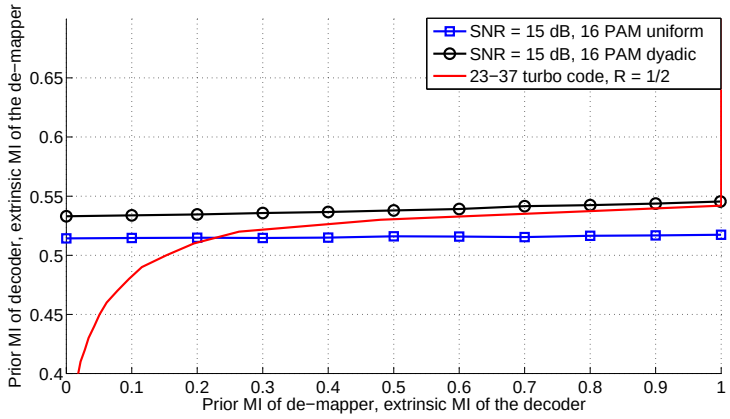


Figure 5.1: EXIT chart of demapper and turbo decoder. The EXIT function of the many-to-one mapping achieves higher extrinsic information with no prior information than the conventional Gray labeled PAM, and furthermore is slightly inclined, leading to a better match to the EXIT function of the turbo decoder.

particularly useful when combined with turbo codes. An extrinsic information transfer (EXIT) chart of such combination is given in Fig. 5.1 for signal-to-noise ratio (SNR), as given in the legend. The mapping function is given in Fig. B.4, and the channel code is a turbo code with constituent convolutional codes with generator polynomials $(23, 37)$ [83] in standard octal notation, punctured to rate $R = 1/2$. The Gray-like property ensures maximum extrinsic information transferred from the demapper on the first iteration, while the ambiguities results in an EXIT function of the demapper, which is slightly inclined. This makes the fit between the EXIT functions of the demapper and the turbo decoder much tighter than with uniformly distributed PMF with Gray labeling. As mentioned in Chapter 2, depending on the SNR, iterative processing can be beneficial for Gray labelings, even though their EXIT functions appear flat, due to the difference between bit-wise and symbol-wise MI. This effect is even more pronounced for the proposed many-to-one mapping. While shaping gains can generally be expected with one-shot processing, they can be further increased by performing iterative demapping and decoding.

Table 5.1: Summary of used mapping functions

Constellation	m	Type
256QAM	12	Raphaeli's original mapping [80]
64QAM	8	Designed in **PAPER 2**
256QAM	10	
1024QAM	14	
1024QAM	16	
64QAM	6	Binary-reflected Gray mappings with uniform PMF
256QAM	8	
1024QAM	10	

5.2 Gains in AWGN channel

We analyze the shaping gains on an AWGN channel in terms of the achieved block error rate (BLER) based throughput, which is calculated as $T = (1 - BLER) \cdot \eta$, where η is the input SE. A specific SE is achieved by puncturing the channel code to rate $R = \eta/m$. The mapping functions we examine are given in Appendix B, and summary of their size is given in Table 5.1. We use 20 turbo iterations, 5 demapping iterations, and the block length is 1000 data bits. Simulated SEs are between 3 and 9, with a step of 0.5 bits/channel use, and the throughput envelope is reported, that is, the throughput of the best performing mapping function from Table 5.1 at the respective SNR. The results are shown in Fig. 5.2. Shaping gains between 0.6 dB and 1.2 dB are seen depending on the SNR. We clearly see the limitations of the original scheme [80], which can only operate above and below a certain SNR, due to the constellation size and the value of m. We note, however, that the performance in that region is near-optimal. For more extensive results, including shaping gains with non-iterative processing please refer to **PAPER 2**.

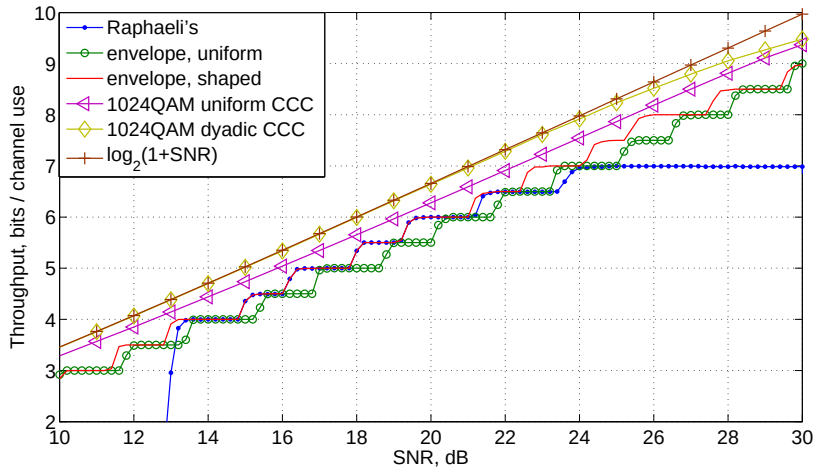


Figure 5.2: Throughput envelope of the designed mapping functions, showing the shaping gains w.r.t. conventional uniformly distributed Gray labeled QAM. The performance of Raphaeli's original design [80] is also shown. The AIRs with the dyadic approximations to the optimal 1024 QAM PMF at each SNR are also shown for comparison.

5.3 WDM optical transmission

The 1024QAM with $m = 16$ from Table 5.1 was used in **PAPER 3** for a single channel optical link, and gains of up to 1dB were reported w.r.t. uniform QAM. Here we apply the methodology from the previous section to a wavelength division multiplexing (WDM) link with the same parameters as in Section 3.2, and design the 1024QAM and the 256QAM mappings, given in Appendix B. The specific PMFs are obtained by optimizing the PMF for the optimal input power (see Fig. 3.2), and then finding the GHC approximation.

In our simulations, the main impairments, present in digital communication systems are included, summarized in Table 5.2. In order to combat these impairments, quadrature phase shift keying (QPSK) pilot symbols are interleaved with the data at rate $P = 1\%$.

The block diagram of the receiver is given in Fig. 5.3. Frequency offset and phase noise are induced on the up-sampled signal. After analog-

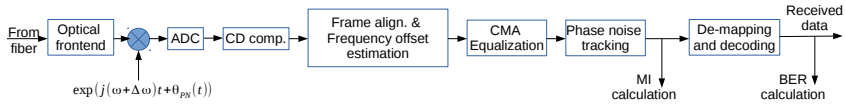


Figure 5.3: Receiver block diagram.

Table 5.2: Transmission impairments

Frequency offset $\Delta\omega$	50 MHz
Laser linewidth	10 kHz
ADC accuracy	6 bits per real dimension
Receiver sampling frequency	56 GSa/s

to-digital conversion (ADC), chromatic dispersion (CD) compensation is performed in frequency domain. Frame alignment is performed based on a short Zaddof-Chu sequence [84], inserted before the first transmitted block. Frequency offset is compensated for via maximum likelihood brute force search, and the symbol sequence is sent for constant modulus equalization, based on the QPSK pilots only. Finally, phase noise tracking is performed with the algorithm from Chapter 4, and the received sequence is sent for demapping and decoding, which is described in Fig. 2.10(b).

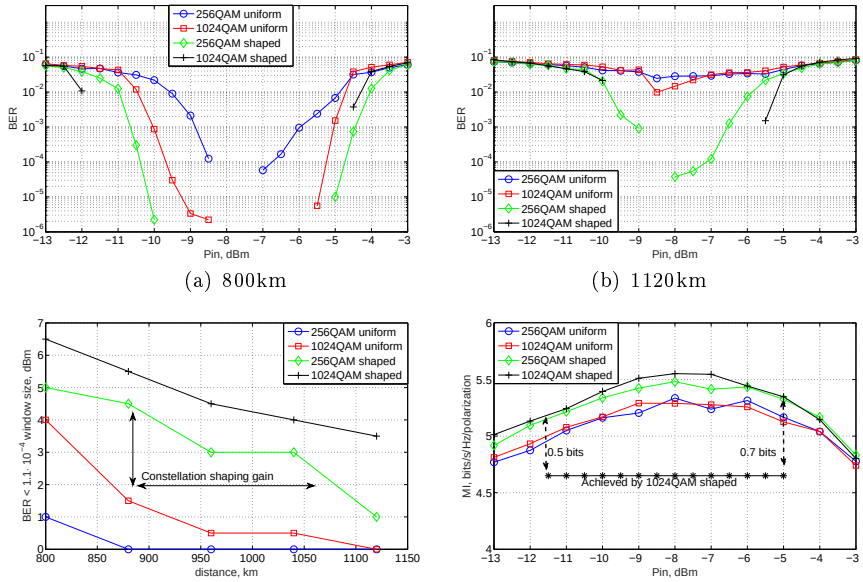
We evaluate the achievable information rate (AIR) before decoding, which is calculated by the phase noise tracking algorithm, and also the achieved bit error rate (BER). The input SE is $\eta = 5$ bits / s / Hz / polarization.

Figures 5.4(a) and 5.4(b) show BER performance after the turbo decoding as a function of the input power per channel per polarization. We sweep the input power in steps of 0.5 dBm, and the absence of a point on a curve means that no errors were found for the duration of the simulation. The block length is 6000 symbols, and 30 blocks were simulated in each case, which at $\eta = 5$ means a total of $9{\cdot}10^6$ information bits are transmitted. The "error-free" region thus corresponds to actual BER $< 1.1 \cdot 10^{-6}$. We assume that the BER results are reliable above $1.1 \cdot 10^{-4}$, since at least 100 errors are seen. 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. For example, the Reed-Solomon code of rate $R_{RS} = 239/255$, defined for the optical transfer networks protocol can operate at such low output BER, when the inner code is subject to the above mentioned error-floor, and when independence of the error events may be assumed (see **PAPER 8**). This assumption is reasonable if interleaveing is performed between the inner and outer codes.

As seen from Fig. 5.4(a) shaping gain is achieved both in the linear regime (around 1.5dB), and slightly less in the nonlinear regime of transmission at a distance of 800 km, i.e. 10 spans of fiber. Due to the very low code rate of $R = 5/14$ for the 1024QAM shaped system, the BER slope is steeper than for the rest of the schemes, and values of the BER between around 10^{-3} and the minimum detectable $1.1 \cdot 10^{-6}$ are not present at the given resolution of the sweep of 0.5 dBm. We also study distances between 960 and 1120 km. A summary of these results is given in Fig. 5.4(c), where the size of the region, for which BER $< 1.1 \cdot 10^{-4}$ is given as a function of the link distance for the examined mappings. The achieved shaping gain may be seen as the increase of the size of this region, but also as the increased maximum distance, where an "error-free" region is found. At 960 km the 1024QAM conventional system may still be used, however, a very small window of "error free" performance is found, which would make it unreliable. When increasing the distance further, no "error-free" region is found for both 1024QAM and 256QAM with uniform PMF. The shaped 256QAM can be used up to 1040 km, and the shaped 1024QAM is able to achieve an "error-free" performance also at 1120 km. If the pilot symbols and the rate R_{RS} of the above mentioned outer code are taken into account, the SE is reduced to $\hat{\eta} = R_{RS} \cdot (1-P) \cdot \eta = 4.65$ bits/s/Hz/polarization, making the the achieved data rate around 260 Gbps per channel.

In Fig. 5.4(d) we plot the received AIR after 800km, calculated between the transmit symbols and the received symbols after the phase noise tracker. We note that due to the pilot overhead, the entropy $\mathcal{H}(X)$ in Eq. (2.1) must be scaled to $(1-P)\mathcal{H}(X)$ in order to get an accurate estimate of the achievable rate. The region, where the final SE of $\hat{\eta} = 4.65$ is achieved by the 1024QAM shaped system (the region, where



(a) 800km (b) 1120km

(c) BER $< 1.1 \cdot 10^{-4}$ (d) Received MI

Figure 5.4: Performance in presence of impairments. (a) and (b) : BER at 800 and 1120km, resp., (c) : size of the window, where BER $< 1.1 \cdot 10^{-4}$, (d) : received MI after 800km by the analyzed transmission formats, together with the window, where the input SE is achieved

BER $< 1.1 \cdot 10^{-4}$ after turbo decoding) is also given in Fig. 5.4(d). We see that performance is close to the theoretical limit - 0.5 and 0.7 bits/s/Hz/polarization at the two edges of the "error free" region, respectively.

5.4 Summary

In this section the novel contributions of this thesis regarding probabilistic shaping were presented. We have introduced a new algorithm for designing mapping functions for many-to-one shaping methodology. The mapping functions achieve Gray-like properties, and are thus particularly suitable for integration with turbo codes. This was demonstrated via EXIT chart analysis of the many-to-one demapper. Shaping gains of up to 1.2 dB were achieved with this method for a wide variety of input

spectral efficiencies on an AWGN channel. Furthermore, the maximum achievable distance of WDM optical fiber links at fixed spectral efficiency can be increased by more than 300km in some cases, at a reference distance of 800km.

Chapter 6

Description of Ph.D. Publications

In this Chapter a description of the publications included in the thesis is provided, highlighting novelties and main results. The papers are divided into three groups, following the order used in the previous chapters. This order is not necessarily the same as the chronological order, in which the papers were published. This section is designed to be as self-contained as possible, therefore overlaps with previous chapters exist.

6.1 Constellation constrained capacity estimation and bounding

PAPER 1: Factorization properties of the optimal signaling distribution of multi-dimensional QAM constellations

The motivation behind this work was the conjecture, posted in [22], that the optimal probability mass function (PMF) of quadrature amplitude modulation (QAM) constellations input to a multiple-input multiple-output (MIMO) channel factorizes into the product of its marginals. This allowed the authors to reduce the complexity of the modified Blahut-Arimoto algorithm (BAA) for finding the optimal PMF on a MIMO channel.

PAPER 1 proves this conjecture by an induction method and exploiting the concavity of mutual information (MI) in the input

PMF. It also uses the expectation-maximization (EM) property of the BAA and the fact that the MI is increased on each EM step, eventually converging to the maximum. Recall the basic steps of the BAA, which are iterated until convergence:

1. **Expectation:** Fix $p_X(X)$, and calculate $p_{X|Y}(X|Y)$.
2. **Maximization:** Fix $p_{X|Y}(X|Y)$, and find $p_X^*(X) = \arg\max \mathcal{I}(X;Y|\mathbf{H})$.

The proof is divided in two major parts, represented by Theorem 1 and Theorem 2. In the expectation step, we prove that for fixed factorized $p_X(X) = \prod_i p_{X_i}(X_i)$ the conditional entropy $\mathcal{H}(X|Y,\mathbf{H})$ separates into a sum of the marginal entropies $\mathcal{H}(X_i|Y,\mathbf{H})$. In the maximization step, we show that by taking the product of the marginals of the optimal PMF $p_X^*(X)$, the MI is increased, which is in contradiction with the strict concavity of the MI. The optimal PMF and the product of its marginals must therefore coincide. Unique optimum is found on the last step, which is also factorized.

The paper provides numerical evidence of the theoretical results in terms of achieved constellation constrained capacity (CCC). We run the full-complexity BAA and the reduced complexity BAA for a 64QAM 2x2 MIMO system, and show that the achieved rates coincide within the numerical precision of the simulations. We also show that the Kullback-Leibler divergence (KLD) between the optimal PMFs found by the two algorithms is very small, and can be attributed to the limited input-output sequence length that was simulated and the convergence thresholds, set in the algorithm implementation.

PAPER 4: Approximating the constellation constrained capacity of the MIMO channel with discrete input

The motivation for this work was the set of papers [24, 25], where the minimum mean squared error (MMSE)-MI relation was found and it was shown that the quantized Gaussian PMF is capacity achieving on an additive white Gaussian noise (AWGN) channel as the constellation size grows to infinity. In **PAPER 4**, the MIMO channel is studied, where similar property was derived. The goal of

the paper was to estimate the CCC in scenarios, where numerical methods are too complex, such as large MIMO matrix. This was done by calculating the CCC on the equivalent orthogonal channel, obtained from the singular value decomposition (SVD), and proving that for large constellation the CCC on the interference channel is the same. The proof relies on the fact that the MI is continuous on the quadratic Wasserstein space in the input PMF, and showing that the quadratic Wasserstein distance from the optimal discrete PMF to the Gaussian disappears.

Then, using the fact that the quadratic Wasserstein distance is a distance measure, we are able to extend this result to the ergodic case. This is done by showing that in the limit of infinitely large constellations, the optimal PMF to any channel realization from the channel distribution is also optimal on any other channel from the same distribution.

The QR decomposition (QRD) is also studied in the paper as means to derive a lower bound on the CCC. The orthogonal channel, obtained from the diagonal elements of the **R** matrix of the QRD is used in this case, instead of the **S** channel of the SVD. The **R** channel provides a better lower bound than the above SVD-based estimate for high signal-to-noise ratio (SNR). Furthermore, it holds for any input PMF, which factorizes into the product of its marginals, and not only the optimal one. This includes the standard uniform PMF.

The result section of the paper provides CCC estimates on a variety of channels, where the MI on the true interference channel can not be calculated due to high complexity, e.g. 1024QAM on 8x8 MIMO. It was shown that the envelope of the SVD-based estimate and the QRD-based lower bound serves as a good approximation to the CCC for the entire SNR region.

PAPER 7: Achievable information rates on linear interference channels with discrete input

The main goal of the paper was to provide meaningful lower bounds for the linear impulse response channels. This is done by extending the QRD based bound from **PAPER 4** to cover the general linear

interference case. To that end, the bound is expressed via the auxiliary channel method, which was presented in detail in Chapter 2. It was shown that the same bound can be achieved if auxiliary channel functions are used, for which the interference from previously decoded layers are subtracted in the current layer. This is similar to the approach, taken in successive interference cancellation receivers. The auxiliary channel method allows to bound the CCC also in cases, where the number of transmit antennas is larger than the number of receive antennas, although the bound quality is poor in that setting.

If the linear impulse response channel is then expressed in its matrix form, we can directly apply the QRD based bound. In this case the $\mathbf{H}$ matrix is of size $[K\text{x}K]$, where K is the length of the input-output sequence. The QRD can still be performed, and the lower bounds can be estimated by a SNR-MI look-up table for each element on the diagonal of the $\mathbf{R}$ matrix.

Results for uniform input PMF are provided for several MIMO channels and two different Rayleigh fading impulse response channels of length $l = 3$ and $l = 6$. In the first case, the QRD bound approaches the true CCC, which is obtained by the trellis method from [16]. In the second case, the trellis method cannot be applied due to the high complexity associated with the exponentially increasing number of states. The QRD-based bound, however, can be used, and achieves close to the Gaussian capacity. It was shown that OFDM with uniform power allocation is outperformed by this bound. This result is in agreement with the result on the MIMO channel from **PAPER 4**, where the SVD based approximation was outperformed at high SNR by the QRD-based bound. Both these results show that orthogonalizing the channel with constellation constrained input can be sub-optimal.

6.2 Phase noise in optical fiber systems

PAPER 5: Compensation of XPM interference by blind tracking of the nonlinear phase in WDM systems with QAM input

The motivation of this paper was previous work by Dar [55], where the auto-correlation function (ACF) of the non-linear phase noise (NLPN) is modeled, and then used to find improved bounds on the optical channel capacity assuming Gaussian input. The goal of **PAPER 5** is to study the ACF of the NLPN with QAM input and estimate potential gains from tracking the NLPN in a practical receiver. To that end, standard models for phase noise are studied, including the Wiener process. In order to validate the Wiener process for modeling the NLPN, the latter is estimated by a genie approach, and the power spectral density (PSD) of the process is compared to the PSD of a true Wiener process, which is a Lorentzian function. A good match was observed, which allowed us to apply existing methods for phase noise tracking under this model.

The chosen method was the trellis method [33]. We optimized the method in terms of number of states in the trellis, and showed that with reasonable complexity the genie based achievable information rate (AIR)s are approached. While the genie approach assumes knowledge of the entire past when estimating the current phase noise sample, the trellis method applies a forward-backward recursion for estimation, which can be implemented in a practical receiver.

We study point-to-point links, but also network scenarios, where the interfering channels are changed along the transmission path. It was shown that for 1000km link and 1024QAM input, the maximum AIR can be increased by around 0.5 bits/channel use in the point-to-point case, similar to the result in [55] with Gaussian input.

These gains disappear in a network scenario due to the very quickly diminishing ACF. The NLPN samples in that case cannot be considered constant even for short duration, and both the genie approach and the trellis method fail to provide significant gains.

PAPER 6: Low-Complexity tracking of laser and nonlinear phase noise in WDM optical fiber systems

In this paper the Wiener process is further studied as a potential tool for modeling the NLPN in optical fibers. Similar to **PA-**

PER 5, a genie approach is used to estimate the NLPN samples via window averaging $\hat{\theta}_k = \angle \sum_{i=k-l}^{k+l} x_i^* y_i$, where x_k are the input symbols, assumed known, and y_k are the channel output samples. We experimented with the window length $2 \cdot l$, and showed that a very good match of the empirical PSD is possible to the theoretical Lorentzian PSD of a Wiener process with variance, given by the sample variance $\Delta^2 = \mathbb{E}_k \left[(\hat{\theta}_k - \hat{\theta}_{k-1})^2 \right]$.

In the main part of the paper we propose a general low-complexity phase noise tracking algorithm. The method models the phase noise distribution at each time by a mixture of Tikhonov distributions, and applies forward and backward recursions for estimating the posterior probabilities of the input symbols $p(x_k|y_1^K)$. Incidentally, the phase noise distribution $p(\theta_k|y_1^K)$ is also estimated. It was shown that the forward and backward recursions can actually be cast directly into the sum-product algorithm for estimating densities on a graph, in this case - a hidden Markov chain. The benefit of using Tikhonov distributions instead of Gaussians is the possibility for analytical integration at the variable nodes of the graph during the recursions. This analytical integration is also beneficial in terms of complexity compared to the trellis method [33], where the phase is discretized and the integration is numerical.

The performance of the algorithm is evaluated on a standard Wiener phase noise channel, on wavelength division multiplexing (WDM) ideally distributed Raman amplified optical fiber link with single polarization and on an Erbium doped fiber amplifier (EDFA) link with dual polarization. The rates approach the AIRs of the significantly more complex trellis method, and were superior to the decision directed method from [41] in most cases.

6.3 Constellation shaping for near-capacity digital communications

PAPER 2: Rate-adaptive constellation shaping for turbo-coded BICM

This paper studies many-to-one probabilistic shaping strategies for

bit-interleaved coded modulation (BICM). The focus is on the pragmatic turbo-coded approach from [80]. Intuition about the method's performance is provided via extrinsic information transfer (EXIT) charts. By analyzing the bit-wise and symbol-wise MI after demapping, it is possible to predict where iterative demapping and decoding will be necessary in order to achieve shaping gain, and where single-shot demapping is sufficient.

Instead of using the quantized Gaussian as in [80], the input PMF is optimized via the BAA and then a dyadic approximation is found, using geometric Huffman coding (GHC) [82]. This allows for performance improvements in many SNR regions, particularly high SNR, where the optimal PMF starts resembling the uniform. A dyadic PMF is achieved with a many-to-one mapping by having each constellation symbol carry many labeling sequences, which share the same prefix. The length of the unique prefix determines the probability of each symbol.

We then propose a method for the construction of the mapping function for symmetric dyadic PMFs, which achieves a Gray-like property of the unique parts of the labels. This property is crucial for the near-optimal performance of the mapping functions for both iterative and non-iterative demapping reception, especially when combined with a turbo code. This is shown via an EXIT chart analysis of the demapper.

Several mapping functions were designed with the proposed method, and evaluated extensively on an AWGN channel and a Rayleigh fading channel in terms of the block error rate based throughput. We show up to 1.2 dB shaping gain with the proposed mappings at high SNR. The iterative performance prediction is confirmed via throughput measurements, where it is shown that iterations are only needed at low-to-mid SNR in order to achieve the complete shaping gain.

PAPER 3: Constellation shaping for fiber-optic channels with QAM and high spectral efficiency

The main idea of the paper is to apply the concept of constellation shaping for NLPN mitigation in optical fibers. The fo-

cus was on QAM constellations, due to their implementation simplicity. Further, when the receiver is subject to practical constrains, such as limited analog-to-digital and digital-to-analog conversion, the regularity of QAM constellations is preferred in I/Q modulation/demodulation. The main contribution of this paper is two-fold:

- A modified BAA is used which allows for optimizing the input PMF on an optical fiber channel. The optimized QAM PMFs achieve a rate, higher than the previously considered most suitable for this channel ring constellations;
- The turbo-coded constellation shaping method from **PAPER 2** is applied on the optical fiber channel, and gains of around 1dB of energy efficiency are achieved.

The modified BAA relies on a Gaussian auxiliary channel assumption. The likelihood functions are modeled as bi-variate Gaussian distributions with mean and covariance matrix obtained from the sample mean and covariance matrices of each constellation symbol. The achieved rates are thus lower bounds on the channel capacity.

The paper implements the optical fiber channel in software via the split-step Fourier method for solving the nonlinear Schrödinger equation (2.25). The system is evaluated on a 800km EDFA link, with single channel double polarization signal. Gains of around 1 dB are reported in bit error rate (BER) in both the linear and nonlinear regions of transmission w.r.t. standard, non-shaped QAM system.

Chapter 7

Conclusion

This thesis was concerned with capacity analysis of, and near-capacity achieving techniques for digital communication systems. The PhD project was focused on both these aspects for a variety of channels with discrete input of finite size. The channels of interest included:

- General additive white Gaussian noise (AWGN) channel;
- Linear interference channels, such as the multiple-input multiple-output (MIMO) channel and impulse response channel;
- The Wiener phase noise channel;
- The nonlinear optical fiber channel.

Particularly for the linear channels, novel techniques were developed for estimating the constellation constrained capacity (CCC) in the low-to-mid signal-to-noise ratio (SNR) and lower-bounding it at high SNR as well. The CCC estimates do not require multi-dimensional numerical integration and can therefore be easily computed for the desired channel, thereby allowing for simple performance evaluation of different receiver processing methods. Instead of the CCC, achievable information rate (AIR)s were analyzed for the phase noise and the optical fiber channels. Applying a modified Blahut-Arimoto algorithm, probabilistic shaping can be performed for quadrature amplitude modulation (QAM) constellations and the AIR of wavelength division multiplexing (WDM)

systems can be increased w.r.t. conventional uniformly distributed QAM inputs.

The non-linear phase noise (NLPN) in optical fibers was studied in Chapter 4, and the first order Wiener process model was validated for the cross-phase modulation (XPM) induced phase noise. A low-complexity method was then proposed based on Tikhonov parametrization of the phase, which effectively combats the combined effect of NLPN and laser phase noise in WDM optical systems. The method exploits the Wiener property and achieves significant performance improvements over standard decision directed methods in variety of scenarios at lower computational complexity.

Constellation shaping for high spectral efficiency digital communications was discussed in Chapter 5. A new method was introduced for designing many-to-one mapping functions for rate adaptive systems, which achieve close to the optimal shaping gain in terms of energy efficiency for the AWGN channel. The proposed method is particularly suitable for systems, which operate turbo-codes (such as the 4G-LTE mobile communications), for which the gains can be achieved with minimal effort by simply modifying the mapping function of the transmitter. The method performs very well on the nonlinear WDM channel, where up to 1.5 dB energy gains can be seen, or equivalently the maximum error-free transmission distance can be increased w.r.t. conventional QAM input constellations. Alternatively, the maximum AIR can be increased by 0.5 bits/channel use on a 800 km link w.r.t. conventional uniformly distributed QAM constellations.

Future Work and Discussion

Several topics of interest have arisen during the course of the PhD study, which are outlined below:

- Constellation shaping for the optical fiber channel. While the scheme, which was designed for the AWGN channel, was successfully transferred to the optical fiber, many improvements may be possible in that setting:

    1. Integration with low-density parity check (LDPC) - convolutional codes (CC). Turbo codes are not particularly suitable

for very high speed communications, due to the inherent latency of the successive BCJR decodings. LDPC-CC have previously been shown to be superior in that regard. However, combining LDPC-CC with high order modulation and many-to-one mappings effectively introduces cycles into the Tanner graph of the code. Therefore, the code needs to be carefully designed with these considerations in mind. Latency is also becoming a problem in wireless communications, e.g. wireless control. Integrating the shaping scheme from Chapter 5 with a LDPC code is therefore of general interest.

2. Joint probabilistic and geometric shaping. Probabilistic shaping is attractive due to its simple and hardware-transparent implementation. However, as we saw in Chapter 3, the gains especially in the nonlinear fiber channel are limited. Jointly optimizing the probabilities and the positions of the points on the complex plane is an interesting area of future research. Further, in Chapters 3 and 5 a memoryless channel is considered, which is far from a realistic model of the nonlinear optical fiber. Temporal constellation shaping is therefore also of interest.

- Iterative phase noise tracking, de-mapping and decoding. Phase noise is still a dominant limitation to the AIRs of cost-efficient optical communication systems. Integrating the tracking scheme from Chapter 4 into the iterative loop might further improve the performance. While this integration seems straightforward, potential improvements, especially in the complexity of the tracking, are of interest for a practical receiver.

- Memory of the optical fiber. As mentioned above, temporal constellation shaping is an interesting area for research. Exploiting the memory in the fiber in general is also of interest. While it was shown in Chapter 4 that the NLPN is highly correlated in phase, such correlation in the amplitude noise are not studied in this thesis. Extending the noise model, and subsequently the tracking algorithm, to cover memory in the amplitude might provide even higher gains in AIRs.

Appendices

Appendix A

Ph.D. Publications

FACTORIZATION PROPERTIES OF THE OPTIMAL SIGNALING DISTRIBUTION OF MULTI-DIMENSIONAL QAM CONSTELLATIONS

Metodi P. Yankov, Søren Forchhammer, Knud J. Larsen

Technical University of Denmark
Department of Photonics Engineering
2800 Kgs Lyngby, Denmark, meya@fotonik.dtu.dk

Lars P. B. Christensen

Fingerprint Cards
Lyskær 3CD, 2730 Herlev, Denmark

ABSTRACT

In this work we study the properties of the optimal Probability Mass Function (PMF) of a discrete input to a general Multiple Input Multiple Output (MIMO) channel. We prove that when the input constellation is constructed as a Cartesian product of 1-dimensional constellations, the optimal PMF factorizes into the product of the marginal 1D PMFs. This confirms the conjecture made in [1], which allows for optimizing the input PMF efficiently when the rank of the MIMO channel grows. The proof is built upon the iterative Blahut-Arimoto algorithm. We show that if the initial PMF is factorized, the PMF on each successive step is also factorized. Since the algorithm converges to the optimal PMF, it must therefore also be factorized.

Index Terms— MIMO, QAM, Constellation shaping

I. INTRODUCTION

As the data rate demand increases, the physical links in band limited scenarios are pushed to operate at high SNR and with high order of modulation in order to achieve high spectral efficiency. Furthermore, it is well known that at high SNR, uniformly distributed signaling achieves the Shannon capacity rate for 1.53dB more energy, called the *shaping gap* or *shaping gain*. In order to close that gap, continuously distributed Gaussian signaling is required [2]. While it is clear that such signals are completely described by their mean and variance, nothing explicit can be said about the shape of the optimal *Probability Mass Function* (PMF) of a given discrete signaling set.

We consider a standard *Multiple Input Multiple Output* (MIMO) channel model:

$$Y = HX + W \tag{1}$$

where X is $2M$ dimensional vector random variable, taking values from the discrete, real-valued set $\mathcal{X}^{2M}$, H is the $2Nx2M$ real-valued equivalent of the NxM complex channel matrix, W is the usual $2N$ dimensional AWGN and Y is the $2N$ dimensional channel output. The set $\mathcal{X}$ is the basic PAM constellation set. The channel is assumed to be perfectly known at the receiver unless otherwise stated.

The algorithm for finding the optimal PMF input to an AWGN channel was derived independently by Blahut [3] and Arimoto [4]. It uses *Expectation-Maximization* (EM) type update rules, sequentially increasing the concave cost function (in this case the *Mutual Information* (MI) as a function of the input distribution) subject to average power constraint $\sum_{i=1}^{|\mathcal{X}^{2M}|} p(\mathbf{x_i})|\mathbf{x_i}|^\mathbf{2} \leq P_{av}$, and normalization constraint $\sum_{i=1}^{|\mathcal{X}^{2M}|} p(\mathbf{x_i}) = 1$, where $\mathbf{x_i}$ is the $i'th$ element of the signaling set and $P(X = \mathbf{x_i}) = p(\mathbf{x_i})$ is its probability.

In [5], the authors replace the power constraint with equality, and then sweep all possible scaled versions of the signaling set, i.e. $\hat{\mathcal{X}^{2M}} = \alpha\mathcal{X}^{2M}$. The EM algorithm is then run for each of these sets, and the one achieving maximum MI is chosen as optimal. The optimal PMF of $\mathcal{X}^{2M}$ is the PMF of the optimal set, and its respective MI is the channel capacity. This algorithm was later extended to cover MIMO scenarios in [1]. Our work is largely based on the derivations made in [1], and so we briefly introduce the mathematical notation of the algorithm used there.

I-A. EM for finding the optimal distribution on MIMO channel with average power constraint

The channel capacity when signaling with $\mathcal{X}^{2M}$ and averaging among the possible channel realizations can be expressed as [1]:

$$C = \max_{p(X)} \mathcal{I}(X;Y|H) =$$

$$= \max_{p(X)} \sum_{i=1:|\mathcal{X}^{2M}|} p(\mathbf{x_i}) \left(\log_2 \frac{1}{p(\mathbf{x_i})} + T_i \right) \tag{2}$$

where,:

$$T_i = \int_{\mathbf{H}} p(\mathbf{H}) \int_{\mathbf{y}} p(\mathbf{y}|\mathbf{x_i},\mathbf{H}) \log_2 p(\mathbf{x_i}|\mathbf{y},\mathbf{H}) d\mathbf{y} d\mathbf{H} \tag{3}$$

We have replaced $p(H = \mathbf{H})$ with $p(\mathbf{H})$ for simplicity and $\mathcal{I}$ is the MI. The EM algorithm is as follows:

Sweep $\alpha \in [\alpha_{min}; \alpha_{max}]$

- **Initialization:**
  Initialize $p(\mathbf{x})$, so that it satisfies $\sum_{i=1}^{|\mathcal{X}^{2M}|} p(\mathbf{x_i}) = 1$, and $\sum_{i=1}^{|\mathcal{X}^{2M}|} p(\mathbf{x_i})|\alpha\mathbf{x_i}|^\mathbf{2} = P_{av}$

- **E:** For each point i and fixed $p(\mathbf{x_i})$, compute T_i.
- **M:** Given the set of fixed T_i values, maximize $\mathcal{I}(X;Y|H)$ w.r.t. the probability of each point. This is a strictly concave optimization problem, solved using Lagrange multipliers.
- **Go to 'E' until convergence**

As the order of modulation and number of dimensions of the signal grow, this algorithm is impractical even when performed offline due to the exponential increase in the number of parameters to be optimized. In [1] the authors conjecture, that the optimal PMF factorizes into the PMFs of each dimension, however, no theoretical proof has been provided. In this work we prove the conjecture. We also show how the Blahut-Arimoto (BA) algorithm can be used to solve the power-allocation problem for discrete signaling sets.

I-B. Notation

In the rest of the paper we use the following notation: x_i denotes the $i'th$ value from the set $\mathcal{X}$. The one-dimensional variable X_i represents dimension i, and $p(X_i = x_j) = p(\mathbf{x_{ij}})$ is the probability of that variable taking value x_j. Capacity achieving PMF will be denoted as $p(X)^*$, and $p(X)^n$ will be the PMF at the beginning of the **E** step on the $n'th$ iteration. The entropy function is denoted as $\mathcal{H}(\cdot)$. Dependency on the noise variance is omitted in the expressions for probability and entropy for brevity.

II. FACTORIZING THE OPTIMAL QAM DISTRIBUTION

In the following, without loss of optimality of the EM algorithm, we assume it has been initialized with a distribution, which is symmetric around 0, and which factorizes as $p(X)^1 = \prod_{k=1:2M} p(X_k)^1$.

Theorem 1. *If $p(X)^n = \prod_{k=1:2M} p(X_k)^n$, then at step n, for fixed $p(X|Y,H)$, the conditional entropy $\mathcal{H}(X_k|X_{\{1:2M\}\setminus k}, Y, H = \mathbf{H})$ is independent of $p\left(X_{\{1:2M\}\setminus k}\right)^{n+1}$*

Proof. See Appendix A □

Theorem 2. *If $p(X)^n = \prod_{k=1:2M} p(X_k)^n$, then $p(X)^{n+1} = \prod_{k=1:2M} p(X_k)^{n+1}$, and therefore $p(X)^* = \prod_{i=1:2M} p(X_k)^*$*

Proof. The MI after step n for real-valued 2x2 MIMO channel can be written as:

$$\mathcal{I}(X;Y|H) = \mathcal{H}(X|H) - \mathcal{H}(X|Y,H) = \mathcal{H}(X_1) + \mathcal{H}(X_2|X_1) - \mathcal{H}(X_1|Y,H) - \mathcal{H}(X_2|X_1,Y,H) \tag{4}$$

where we have used the chain rule for entropy and conditional entropy. It is clear that $\mathcal{H}(X_1|Y,H)$ is independent of $p(X_2)^{n+1}$. After Theorem I, $\mathcal{H}(X_2|X_1,Y,H)$ is also independent of $p(X_1)^{n+1}$, or the conditional entropy $\mathcal{H}(X|Y,H)$ is separated into functions of the marginal PMFs. Let's assume that the optimal PMF at this step was found as $p(X)^*$, and its respective marginals are $p(X_k)^* = \sum_{X_{\{1:2M\}\setminus k}} p(X)^*$. We then consider the PMF, obtained as product of those marginals $p(X)^{\tilde{}} = \prod_{k=1:2M} p(X_k)^*$. The entropy of X as a function of this PMF is:

$$\mathcal{H}(p(X)^{\tilde{}}) = \mathcal{H}(X_1) + \mathcal{H}(X_2) \geq \mathcal{H}(p(X)^*) \tag{5}$$

Then for the MI as a function of the PMF we have:

$$\mathcal{I}(p(X)^{\tilde{}}) = \mathcal{H}(p(X)^{\tilde{}}) - \mathcal{H}(X|Y,H) \geq \mathcal{H}(p(X)^*) - \mathcal{H}(X|Y,H) = \mathcal{I}(p(X)^*) \tag{6}$$

However, $\mathcal{I}(p(X)^*) \geq \mathcal{I}(p(X))$ for any $p(X)$, and therefore $\mathcal{I}(p(X)^*) = \mathcal{I}(p(X)^{\tilde{}})$. In [6] the authors prove, that the MI is strictly concave in $p(X)$, and therefore the optimal distribution is unique. The optimal distribution therefore must be the same as the product of its marginals. The extension to $M > 2$ is straight-forward. The MI after step n can be expressed as:

$$\mathcal{I}(p(X)^{n+1}) = \mathcal{H}(X) - \sum_{k=1:2M} \mathcal{H}(X_k|X_{\{1:k-1\}}, Y, H) \tag{7}$$

where each element in the sum only depends on its respective marginal PMF. Then if $\mathcal{I}(p(X)^{n+1}) = \max \mathcal{I}(p(X)^{n+1}) \Rightarrow \mathcal{H}(X_m|\{X_{1:2M}\}\setminus X_m) = \max \mathcal{H}(X_m|\{X_{1:2M}\}\setminus X_m)$, and therefore $p(X_m|\{X_{1:2M}\}\setminus X_m)^{n+1} = p(X_m)^{n+1}$. Since $p(X)^{n+1} = \prod_{i=1:M} p(X_i)^{n+1}$, the theorem is proven. □

II-A. Modified BA algorithm

The channel capacity of the simple 2x2 real-valued channel can now be expressed as:

$$C = \max_{p(X)} \mathcal{I}(X;Y|H) = \max_{p(X_1)} (\mathcal{I}(X_1;Y|H)) + \max_{p(X_2)} (\mathcal{H}(X_2) - \mathcal{H}(X_2|X_1,Y,H)) \tag{8}$$

The maximization problem is separated into 2 maximization problems of much lower dimensionality, and power constrain $P_{av}^k = \sum_{l=1:|\mathcal{X}|} p(x_l)\alpha_k |x_l|^2$, where α_k is the scaling coefficient for the $k'th$ dimension. The degrees of freedom in the maximization problem are reduced from $|\mathcal{X}|^{2M}$ to only $|\mathcal{X}|$. If we assume symmetric distribution on H, it is clear that $p(X_1)^* = p(X_2)^* = \cdots = p(X_{2M})^*$, and $P_{av}^1 = P_{av}^2 = \cdots = P_{av}^{2M} = P_{av}/2M$. In that case the modified BA algorithm from Section I takes the form:

Sweep $\alpha \in [\alpha_{min}; \alpha_{max}]$

- **Initialization:**
  Initialize $p(X_1)$, so that it satisfies $\sum_{i=1}^{|\mathcal{X}|} p(\mathbf{x_{1i}}) = 1$, and $\sum_{i=1}^{|\mathcal{X}|} p(\mathbf{x_{1i}})|\alpha x_i|^2 = \frac{P_{av}}{2M}$

- **E:** For each point $x_i \in \mathcal{X}$ and fixed $p(x_i)$, compute $\hat{T}_i$:

$$\hat{T}_i = \int_{\mathbf{H}} p(\mathbf{H}) \int_{\mathbf{y}} p(\mathbf{y}|\mathbf{x_{1i}}, \mathbf{H}) \log_2 p(\mathbf{x_{1i}}|\mathbf{y}, \mathbf{H}) \quad (9)$$

- **M:** Given the set of fixed $\hat{T}_i$ values, find $p(X_1)$ which maximizes $I(X_1; Y|H)$ as:

$$p(X_1) = \arg\max \mathcal{I}(X_1; Y|H) =$$
$$\arg\max \sum_{i=1:|\mathcal{X}|} p(\mathbf{x_{1i}}) \left(\log_2 \frac{1}{p(\mathbf{x_{1i}})} + \hat{T}_i \right) \quad (10)$$

- **Go to 'E' until convergence**

III. FURTHER DISCUSSION

III-A. Orthogonal channel

Let us consider the case, when $\mathbf{H}$ is diagonal. When the noise is i.i.d. Gaussian, the likelihood on this channel can be expressed as a product of the likelihoods on each dimension:

$$p(Y|X, \mathbf{H}) = \prod_{k=1:2M} p(Y_k|X_k, \mathbf{H_{kk}}) \quad (11)$$

where $\mathbf{H_{kk}}$ is the element on the $k'th$ row and $k'th$ column (we assume the channel has full rank). In this case it is straightforward to prove that the entropy $\mathcal{H}(X|Y, H)$ can be expressed as a sum of functions of the marginal PMFs, namely $\mathcal{H}(X|Y, H) = \sum_{k=1:2M} \mathcal{H}(X_m|Y_m, H)$. It suffices to show that at step n, the conditional distribution $p(X|Y, H)$ factorizes as:

$$p(X|Y, H) = \frac{p(Y|X, \mathbf{H}) p(X)^n}{\sum_{k=1:|\mathcal{X}^{2M}|} p(Y|\mathbf{x}_k, \mathbf{H}) p(\mathbf{x}_k)^n} =$$
$$\frac{\prod_{k=1:2M} p(Y_k|X_k, \mathbf{H_{kk}}) p(X_k)}{\prod_{k=1:2M} \sum_{i=1:|\mathcal{X}|} p(Y_k|\mathbf{x_{ki}}, \mathbf{H}) p(\mathbf{x_{ki}})^n} =$$
$$\prod_{k=1:2M} p(X_k|Y_k, \mathbf{H}) \quad (12)$$

III-B. Non-symmetric channel matrix distribution

When the channel distribution is not symmetric, the power constraint on each maximization problem is not necessarily the same. In order to find these constraints, a power allocation solution is needed. However, this can be circumvented if the scaling coefficient α can be chosen differently for each dimension, and the maximization problem is run on the full-rank system for each allowed $\alpha = [\alpha_1, ...\alpha_{2M}]^T$ and the set $\mathcal{X}^{2M} = \prod_{k=1:2M} \alpha_k \mathcal{X}$. As usual, the set, achieving the maximum MI will be chosen as optimal. Since for fixed $p(X)$, α_k directly gives the power, allocated to dimension k, it is clear that the power allocation is obtained by the modified Blahut-Arimoto algorithm. When channel state information at the transmitter is not available or is imperfect, and symmetry in the channel distribution cannot be assumed, the optimization problem cannot be separated into problems of reduced dimensionality, because the optimal power allocation is not known a-priori, and only the total power constraint:

$$P_{av} = \sum_{k=1:|\mathcal{X}^{2M}|} p(\mathbf{x_k})[\alpha_1, ...\alpha_{2M}][|x_k^1|^2, ...|x_k^{2M}|^2]^T \quad (13)$$

can be considered (here $\mathbf{x_k} = [x_k^1, ...x_k^{2M}]^T$). In this case the degrees of freedom in the optimization process for each α are reduced from $|\mathcal{X}|^{2M}$ to $2M|\mathcal{X}|$, but the number of values of the vector α that need to be swept grows exponentially with M.

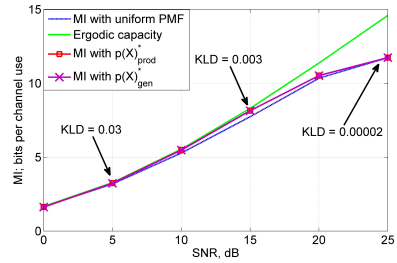


Fig. 1. MI with uniform PMF, the optimal PMFs obtained from the algorithms in Sections I and II, and the ergodic capacity. The KLD between the two optimal PMFs is indicated for selected SNR values

IV. SOME RESULTS ON THE OPTIMAL PMFS

We compare the PMF, obtained by the general algorithm from Section I, with the PMF, obtained by the modified algorithm from Section II by means of the *Kullback-Leibler Distance* (KLD). If $p(X)^*_{gen}$ is the former, and $p(X)^*_{prod}$ is the latter PMF, the KLD is defined as:

$$D(p(X)^*_{gen} || p(X)^*_{prod}) = \sum_{i=1:|\mathcal{X}^{2M}|} p(\mathbf{x_i})^*_{gen} \log_2 \frac{p(\mathbf{x_i})^*_{gen}}{p(\mathbf{x_i})^*_{prod}}$$

In Fig. 1 the MI is given for both PMFs, together with the MI with uniform PMF and the ergodic capacity [7]. The MI is practically the same for both. The KLD is also indicated in the figure for selected SNR values. We see that it is practically zero (we attribute any error to numerical inaccuracy) at the low SNR (where the shaping gain is very small), the medium SNR (where the largest shaping gain can be expected) and at high SNR (where the uniform PMF is near-optimal).

V. CONCLUSION

In this paper the factorization properties of the optimal PMF input to a MIMO channel were considered. It was proven, that the optimal PMF factorizes into the product of its marginal PMFs, confirming the conjecture, made in [1]. The proof relies on the iterative BA algorithm, showing

that if the initial PMF factorizes, it stays factorized on each subsequent iteration, evidently reaching the unique optimum, which must also be factorized. Using the factorization property, it was also shown how the power allocation problem can be solved by the BA algorithm.

VI. APPENDIX A

We prove Theorem 1 for $M = 2$ and real-valued channel for simpler notation, and then show that it can be extended to higher order MIMO and complex-valued channel. Let us denote $f(p(X_1)^{n+1}, p(X_2)^{n+1}) = -\mathcal{H}(X_2|X_1, Y, H = \mathbf{H})$ at iteration n for fixed $p(X|Y, \mathbf{H})$:

$$f = \sum_{k=1:|\mathcal{X}|} p(\mathbf{x_{1k}})^{n+1} \int_{\mathbf{y}} p(\mathbf{y}|\mathbf{H}, \mathbf{x_{1k}}) \cdot$$

$$\sum_{j=1:|\mathcal{X}|} p(\mathbf{x_{2j}}|\mathbf{x_{1k}}, \mathbf{y}, \mathbf{H}) \log_2 p(\mathbf{x_{2j}}|\mathbf{x_{1k}}, \mathbf{y}, \mathbf{H}) d\mathbf{y} =$$

$$= \sum_{k=1:|\mathcal{X}|} p(\mathbf{x_{1k}})^{n+1} \sum_{j=1:|\mathcal{X}|} p(\mathbf{x_{2j}}|\mathbf{x_{1k}})^{n+1} \cdot g_j(\mathbf{x_{1k}}) \quad (14)$$

where the function $g_j(X_1)$ is defined as:

$$g_j(X_1) = \int_{\mathbf{y}} p(\mathbf{y}|\mathbf{H}, \mathbf{x_{2j}}, X_1).$$

$$\log_2 \frac{p(\mathbf{y}|\mathbf{H}, \mathbf{x_{2j}}, X_1) p(\mathbf{x_{2j}})^n}{\sum_{i=1:|\mathcal{X}|} p(\mathbf{y}|\mathbf{H}, \mathbf{x_{2i}}, X_1) p(\mathbf{x_{2i}})^n} d\mathbf{y} =$$

$$= \log_2 p(\mathbf{x_{2j}})^n + D(p(Y|\mathbf{H}, \mathbf{x_{2j}}, X_1)||p(Y|\mathbf{H}, X_1)) \quad (15)$$

In (14) the Bayes theorem is used to express the conditional distribution $p(X_2|X_1, \mathbf{y}, \mathbf{H})$, we remove the dependency on H where it is not relevant, and we have used the fact that $p(\mathbf{x_{2j}}|\mathbf{x_{1k}})^n = p(\mathbf{x_{2j}})^n$.

Let us examine the PDFs in the KLD. The first one is a Gaussian with mean $\mathbf{H}\,[X_1, \mathbf{x_{2j}}]^T$ and covariance matrix given by the noise. The second is a *Mixture of Gaussians* (MoG) with the same covariance matrices, and mixing coefficients given by the marginal PMF $p(X_2)^n$. For different values of X_1, the shape of both PDFs is unchanged, only the mass is linearly shifted. We note that this shift is the same for both PDFs. This means that the relative offset between them is the same regardless of X_1, or in other words - the KLD is unchanged for different X_1. This property is illustrated in Fig. 2. We plot the likelihood and the marginal PDF for a random $\mathbf{H}$, 2 randomly chosen values of $X_1 : X_1 = x_k$ and $X_1 = x_m$, fixed $X_2 = x_j$ and a random valid PMF $p(X_2)$. The likelihood (the white Gaussian curve) sits on top of one of the components of the MoG (plotted in gray). The shift is along the line, defined by the points $\mathbf{H}[X_1, x_j]^T$.

The logarithm in (15) is clearly independent of X_1, therefore g_j is independent of X_1. Equation (14) can now be rewritten as:

$$f = \sum_{j=1:|\mathcal{X}|} p(\mathbf{x_{2j}}) \sum_{k=1:|\mathcal{X}|} p(\mathbf{x_{1k}}|\mathbf{x_{2j}}).g_j = \sum_{j=1:|\mathcal{X}|} p(\mathbf{x_{2j}}).g_j$$

which proves the theorem.

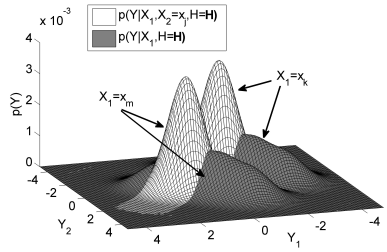


Fig. 2. Illustration of the KLD independence of X_1. The masses of the likelihood and the marginal PDF are shifted by the same factor, retaining their relative offset, and therefore the KLD

It is straight-forward to extend the proof to $M > 2$ and complex-valued input and channel. The upper-mentioned linear shift will be across multiple dimensions, instead of just 1, still keeping the KLD $D(p(Y|\mathbf{H}, \mathbf{x_{2j}}, X_{\{1:2M\}\setminus 2})||p(Y|\mathbf{H}, X_{\{1:2M\}\setminus 2}))$ independent of $X_{\{1:2M\}\setminus 2}$. Then $g_j(X_1, X_3, ...X_{2M})$ only depends on $p(\mathbf{x_{2j}})$, and $f = \sum_{j=1:|\mathcal{X}|} p(\mathbf{x_{2j}}).g_j$.

VII. REFERENCES

[1] J. Bellorado, S. Ghassemzadeh and A. Kavcic, "Approaching the capacity of the MIMO Rayleigh flat-fading channel with QAM constellations, independent across antennas and dimensions", *IEEE Trans. Commun.*, vol. 5, no. 6, pp. 1322-1332, 2006

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

[3] R. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 460-473, 1972.

[4] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 14-20, 1972.

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

[6] I. C. Abou-Faycal, M. D. Trott and S. Shamai, "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1290-1301, 2001.

[7] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj and H. V. Poor, *MIMO Wireless Communications*, Cambridge University Press 2007

Rate-adaptive Constellation Shaping for Near-capacity Achieving Turbo Coded BICM

Metodi Yankov, Søren Forchhammer and Knud J. Larsen
Department of Photonics Engineering
Technical University of Denmark
2800 Kgs Lyngby, Denmark
Email: {meya, sofo, knjl}@fotonik.dtu.dk

Lars P. B. Christensen
Wireless Algorithm Design
Renesas Mobile Europe
2450 Copenhagen, Denmark
Email: lars.christensen@renesasmobile.com

***Abstract*—In this paper the problem of constellation shaping is considered. Mapping functions are designed for a many-to-one signal shaping strategy, combined with a turbo coded *Bit-interleaved Coded Modulation* (BICM), based on symmetric Huffman codes with binary reflected Gray-like properties. An algorithm is derived for finding the Huffman code with such properties for a variety of alphabet sizes, and near-capacity performance is achieved for a wide SNR region by dynamically choosing the optimal code rate, constellation size and mapping function based on the operating SNR point and assuming perfect channel quality estimation. Gains of more than 1dB are observed for high SNR compared to conventional turbo coded BICM, and it is shown that the mapping functions designed here significantly outperform current state of the art Turbo-Trellis Coded Modulation and other existing constellation shaping methods.**

I. Introduction

It is well known that at high SNR, uniformly distributed constellation input to a Gaussian channel achieves a rate that is $1.53dB$ away from channel capacity, as given by Shannon. In order to close that gap, a continuous Gaussian distributed input is required. For discrete modulation used in any practical digital communication system, that gap can be closed for very high order constellations (high number of coded bits per constellation symbol), which have a sampled Gaussian distribution [1]. *Signal shaping* is the process of achieving non-uniform input *Probability Mass Function* (PMF), from the uniformly distributed data. Since this usually comes at the cost of additional redundancy, the channel code rate has to be increased in order to compensate and keep the same data rate, or the constellation size has to be increased to accommodate the extra redundancy bits and retain the high spectral efficiency. These in turn degrade the performance and/or lead to increased complexity at the receiver. However, the energy savings resulting from signal shaping can compensate for the performance degradation. These energy savings at a given data rate are referred to as *Shaping gain*.

In [2] an algorithm is derived for finding the optimal PMF on an AWGN channel for a certain SNR, which was later extended to cover fading scenarios in [3]. The exact optimal PMF is however very hard to achieve in a practical communication system, transmitting coded binary data. In [4], the authors use a shaping code to explicitly select points with low energy more often. This method results in a PMF with only a few levels, and thus suffers a significant loss in *Mutual Information* (MI) between the input and the output of the channel w.r.t. a continuous Gaussian distribution. In [5] superposition modulation is used together with an optimized irregular convolutional code to combat this problem and achieve a near-capacity performance. However, the decoding complexity grows exponentially with the number of levels. Furthermore, different transmission rates would require different codes, which is impractical in a rate adaptive system, such as e.g. LTE. In [6] the authors approximate the optimal PMF with a *dyadic* one (to be defined in Section II). Dyadic PMFs are attractive because they can be achieved with binary data in a number of ways, e.g. with *Variable Length Codes* (VLC) [6] or a *many-to-one* strategy [7]. In [7] Raphaeli proposed a simple method for achieving a dyadic PMF, using the notion of pragmatic turbo codes [8]. However, the PMF is fixed and is designed for a certain SNR region, making it sub-optimal if a change in the spectral efficiency is required.

In this work a rate-adaptive system is designed, by tuning the transmit PMF depending on the SNR value and the desired spectral efficiency. The desired PMF is achieved by a many-to-one mapping function, and near-capacity performance is reported for a wide SNR region when the mapping function is combined with a turbo coded *Bit Interleaved Coded Modulation* (BICM) [9].

II. System overview and many-to-one shaping

In [10] it is shown how the optimal discrete PMF can be approximated with a dyadic one. A dyadic PMF has the form $p(x_i) = 2^{-l_i}$, where l_i is a positive integer, and x_i is the $i'th$ symbol of the alphabet. The authors in [10] prove that the dyadic PMF which minimizes the $KLD = D(p_{approx}||p_{optimum})$ is found by *Geometric Huffman Codes* (GHC), where KLD is the Kullback-Leibler Divergence, $p_{optimum}$ is the capacity achieving PMF, and p_{approx} is any dyadic PMF. The MI between the input and the output of a real-valued system (i.e. PAM), when signaling with the optimal input PMF is given in Fig. 1, together with the Shannon capacity, the MI with an uniform input PMF, and the MI with GHC approximation. We see that even though the dyadic PMF cannot achieve capacity, significant gains can still be expected

978-1-4799-2003-7/14/$31.00 ©2014 IEEE

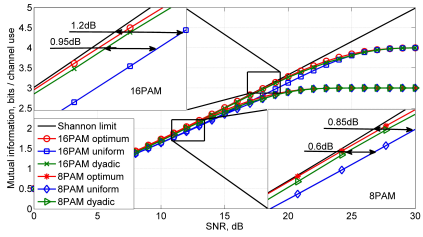


Fig. 1. MI for different input PMFs

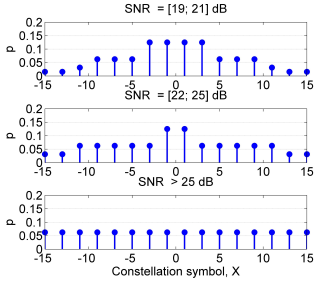


Fig. 2. Dyadic PMFs, approximating the optimal PMF for selected SNR regions

at the examined SNR regions and modulation orders. We also see that after a certain SNR, increase in the modulation size is needed (around 15dB for e.g. 8PAM, or equivalently 64QAM) if Gaussian capacity is targeted. This is due to the limited entropy of the small constellations. Examples of dyadic approximations to the optimal PMF at different SNRs are given in Fig. 2. The PMF for $SNR \in [19; 21] dB$ is the one the authors used in [7]. Higher or lower SNR will have different optimal PMFs and dyadic GHC approximations.

The transmitter from [7], which we also use here, 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, where the size of the modulation format is 2^m. Each stream is then interleaved, and the signal is modulated. The receiver is given in Fig. 4. The signal is demodulated and the data are deinterleaved. The parity streams are de-punctured and sent for decoding together with the data bits. In case of iterative demapping, the decoder sends extrinsic information back to the demapper, after proper puncturing and interleaving.

Dyadic PMF of the output symbols can be achieved by decoding the binary stream into the constellation symbols using an entropy rate source code, e.g. VLC [11]. The VLC needs to be an entropy code, because otherwise there will exist forbidden sequences in the code, and thus a part of

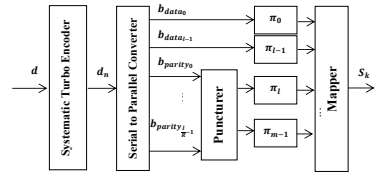


Fig. 3. Turbo coded BICM transmitter

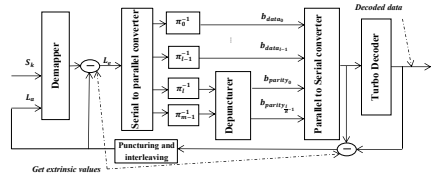


Fig. 4. Turbo coded BICM system receiver

the data will be impossible to decode to any symbol. The probabilities of the output symbols depend on the length of the bit labels assigned to them. For dyadic PMFs, the prefix-free Huffman code is an entropy rate code [11][12]. However, the lack of any structure in the Huffman code results in a severe synchronization problem, especially when there is no error protection before it (as is the case if the VLC is used as an inner code, or in this case - a mapping function).

In [7], the authors tackle this problem by appending ambiguous bits (taking a value of both '1' and a '0') to each bit sequence, after a constellation symbol is chosen. In this way the dyadic PMF is retained, but the label length is constant across symbols. This process can also be viewed as a many-to-one mapping, where the ambiguous bits have the effect of assigning multiple bit sequences with the same prefix (defined by the Huffman code) to the same output symbol. An example of such a mapping function is given in Table I, where the ambiguous bits are labeled as 'X'. For example, constellation symbol '-5' will carry all sequences of length 5, that start with '0010', namely '00100' and '00101'. If the data are i.i.d., the probability of encountering this prefix sequence, and thus this constellation symbol, is clearly 2^{-4}. Compared to

TABLE I
EXAMPLE OF 8PAM MAPPING FUNCTION WITH OPTIMAL DYADIC PMF

Constellation symbol X	-7	-5	-3	-1	1	3	5	7
Bit Label	0	0	0	1	1	0	0	0
	0	0	1	0	1	1	0	0
	1	1	0	X	X	1	0	1
	1	0	X	X	X	X	X	1
	1	X	X	X	X	X	X	0
$p(X)$	2^{-5}	2^{-4}	2^{-3}	2^{-2}	2^{-2}	2^{-3}	2^{-3}	2^{-5}

conventional turbo coded BICM, the only difference in the receiver is the demapper, which needs to take the many-to-one ambiguities into account. Since the complexity of the turbo-decoder dominates over that of the de-mapper, the complexity of the receiver overall is not significantly increased.

A. Mapping design criteria for iterative and non-iterative receivers

In a general BICM, when no iterations are allowed between the de-mapper and decoder, binary-reflected Gray coding is chosen since it provides maximum MI at the output of the demapper with no a-priori knowledge at its input [13]. A popular assumption is that the MI after the de-mappper in that case does not improve with iterative de-mapping (or that estimation of the bit at a certain position in the labeling scheme does not benefit from knowing the rest of the bits), i.e. [14]:

$$I(Y;X) = \sum_m I(Y;b_m|b_1,...b_{m-1}) \approx \sum_m I(Y;b_m) \quad (1)$$

where X is the channel input, Y is the channel output, and b_m is the $m'th$ bit in the labeling scheme. However, (1) does not hold in the general case. Having a channel code will introduce dependencies between the bits, and (1) therefore has the form:

$$I(Y;X) = \sum_m I(Y;b_m|b_1,...b_{m-1}) \geq \sum_m I(Y;b_m) \quad (2)$$

Where the last inequality turns into equality for binary-reflected Gray mapping *only in perfect conditions*, such as ideal interleaving, perfectly white noise, perfectly white data, optimal de-mapper etc. The usual methods for generating an EXIT chart, e.g. [15], do not take this into account, and the EXIT function of a Gray mapper therefore appears completely flat. In a practical system it will be slightly inclined. In Fig. 5, the bit-wise $(\sum_m I(Y;b_m))$ and symbol-wise $(I(Y;X))$ MI are given for a 16PAM binary-reflected Gray mapping, with the optimal maximum a-posteriori probability de-mapper. We see that at low-to-medium SNR, gains can be expected from iterative de-mapping even with Gray mappings. As the SNR increases and the MI curve starts to level out, this gain disappears. The many-to-one ambiguities will have the effect of further inclining the EXIT function of the de-mapper, resulting in a very good match with the EXIT function of the turbo code, as can be seen from Fig. 6. The shaped system will thus benefit from iterative de-mapping and decoding. Moreover, it is well known that convolutional and turbo codes perform best, when the systematic bits are not punctured. Adding ambiguities on the bit positions can be considered as puncturing, because at the receiver those bits will have equal probability for a '1' and a '0'. We are therefore designing the mapper in a way, so as the unique bit positions resemble a binary-reflected Gray code, and the systematic bits are mapped to them. These criteria also ensure very good performance, when iterations between the decoder and the de-mapper are not allowed.

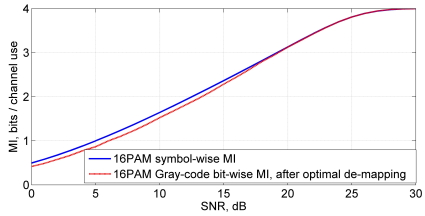


Fig. 5. Bit-wise vs. symbol-wise MI for binary-reflecetd Gray coded 16 PAM

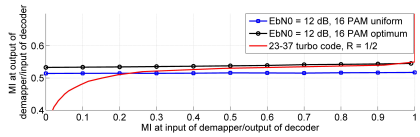


Fig. 6. EXIT chart of 16PAM de-mappers

III. Obtaining binary reflected Gray-like many-to-one mapping function

In [7], the authors use a single 16PAM mapping function with label length of 6 bits, including the ambiguous bits. As shown in Section I, that PMF of the symbols is optimal only for a limited SNR range (Fig. 2, optimum PMF for $SNR \in [19;21]dB$). A straight-forward way of obtaining a mapping function for any dyadic PMF is to use a Huffman tree [12]. However, this method will generally produce a mapping function without any structure, and very far from the target binary reflected Gray code. Here we propose a method for designing the mapping function for a 1D constellation, with the additional constraint that the dyadic PMF is symmetric (we note that the GHC approximation from [10] does not necessarily yield a symmetric dyadic PMF, even though the optimal PMF is always symmetric).

A. Obtaining the optimal symmetric dyadic PMF

Let $\mathbf{p}$ be the optimum PMF of the constellation points and the constellation have N points, $\mathbf{p}_{1:N}$ is the PMF of the points from 1 to N, and p_i is the probability of point i. Let $\mathbf{p}^{half}$ be the normalized version of $\mathbf{p}_{1:N/2}$ (we assume even number of constellation points). For symmetric PMFs, it is clear that $\mathbf{p}^{half} = 2\mathbf{p}_{1:N/2}$.

Theorem 1: The symmetric dyadic PMF $\mathbf{q}$ which minimizes $D(\mathbf{q}||\mathbf{p})$ is the one, for which $\mathbf{q}^{half}$ is the GHC approximation of $\mathbf{p}^{half}$

Proof: $D(\mathbf{q}||\mathbf{p}) = \sum_{i=1}^{N} q_i \log_2 \frac{q_i}{p_i} = \sum_{i=1}^{N/2} 2q_i \log_2 \frac{2q_i}{2p_i} = D(\mathbf{q}^{half}||\mathbf{p}^{half})$, where the last two equalities follow from the fact that $\mathbf{q}$ and $\mathbf{p}$ are symmetric. As proven in [6], $D(\mathbf{q}^{half}||\mathbf{p}^{half})$ is minimized by GHC, which proves the theorem. ■

The loss in MI from constraining the dyadic PMF to a symmetric one is given in Fig. 7 for a 16PAM constellation. We only see loss in the low SNR region, up to 14 dB. Surprisingly, Raphaeli's PMF [7], even though optimal for around $20dB$, performs very well also for lower SNRs. We note that the GHC dyadic approximation from [6], and the symmetric constraint we apply are considered optimal in the sense that they minimize $D(p_{approx}||p_{optimum})$. This optimality does not guarantee that these approximations result in the minimum loss in MI between input and output (unless $D(p_{approx}||p_{optimum}) \to 0$), which is also what we can see in Fig. 7. Finding such an approximation is currently an NP hard problem [6]. This figure also shows that the PMF from [7] suffers significant loss at high SNR.

The proposed algorithm for finding the binary reflected Gray-like Huffman code is as follows:

- **Consider the PMF of half the constellation points** (e.g. negative points in a symmetric PAM constellation)

- Assign an all-ones string to the point with shortest label (symbol number $N/2$, where N is the number of constellation points)
- **for** $i = N/2 - 1$ down to 1 **do**
  1) Copy label $i+1$ to label i
  2) For label i, flip the right-most possible bit, making sure the resulting table is prefix free (bit is highlighted in the example in Table II)
  3) For label i, set bits $(l_{i+1}; l_i]$ to 1, where l_i is the length of label i (bits are underlined in the example in Table II)

- Once the final label is reached, mirror the table to the other half of the PMF, while flipping the left-most bit, ensuring the labels are prefix free

An example of the obtained Huffman code is given in Table II. The PMF we used there is an example 16PAM dyadic, namely, optimized for $SNR \in [22; 25]dB$ (see also Fig. 2). As discussed above, the mapping function is obtained by appending ambiguous bits to the short labels. Shaped QAM PMFs are obtained as product of PAM PMFs, preserving the Gray-code property.

IV. Results

In this section we present results for the mapping functions designed in Section III, and compare them with the system, originally designed by Raphaeli in [7], and also other near-capacity achieving schemes. The systems are compared in terms of *Bit Error Rate* (BER) and *Block Error Rate* (BLER). The BER analysis is chosen for easier comparison with other systems and capacity. However, the BLER gives a better idea of the achievable throughput in a real communication system over a wide SNR range. Consequently, we report the achieved throughput T in bits per channel use. It is calculated as $T = (1 - BLER)\eta$, where η is the spectral efficiency in bits per channel use $[b/cu]$. Table III summarizes the analyzed modulation schemes. QAM constellations are designed as a product of two identical PAM constellations. Therefore, $l_{max}^{QAM} = 2l_{max}^{PAM}$, where l_{max} is the length of the longest label in the QAM/PAM modulation table.

The system parameters are 5000 information bits block length, 20 turbo iterations, 5 demapping iterations, two convolutional component codes with polynomials (23, 37) in standard octal notation and uniformly distributed puncturing

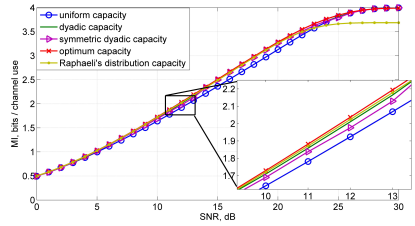


Fig. 7. MI loss from symmetric dyadic approximation

TABLE II
Algorithm for finding the binary reflected Gray-like Huffman code for a 16PAM dyadic PMF

X	i	Bit label					l_i
-15	1	1	0	1	0	0	5
-13	2	1	0	1	0	1	5
-11	3	1	0	1	1		4
-9	4	1	0	0	1		4
-7	5	1	0	0	0		4
-5	6	1	1	0	0		4
-3	7	1	1	0	1		4
-1	8	1	1	1			3
1	9	0	1	1			3
3	10	0	1	0	1		4
5	11	0	1	0	0		4
7	12	0	0	0	0		4
9	13	0	0	0	1		4
11	14	0	0	1	1		4
13	15	0	0	1	0	1	5
15	16	0	0	1	0	0	5

TABLE III
Summary of used mapping functions

Constellation	l_{max}	Type
256QAM	12	Raphaeli's [7]
64QAM	8	Designed in Section III
256QAM	10	
1024QAM	14	
1024QAM	16	
64QAM	6	Binary-reflected Gray mappings with uniform PMF
256QAM	8	
1024QAM	10	

pattern. Simulated data rates (modulation spectral efficiency) are between 3 and 9, with a step of 0.5 b/cu. We note, that not all formats from Table III support all possible rates. This is because the lowest possible code rate of the turbo code in question is $1/3$, and therefore constellation with bit label length of e.g. 16 bits cannot support spectral efficiency of less than 5.5 b/cu. We simulated 10000 blocks for each code rate at each relevant SNR.

A. AWGN channel

As we saw in Fig. 7, Raphaeli's PMF suffers from a severe loss above a certain SNR. In Fig. 8, we see the performance of the mapping functions at $\eta = 7b/cu$. The best performing 256QAM PMF is the one designed based on the Huffman code from Table II. However, at this point the ambiguities have a significant effect on the waterfall properties of the turbo code. We see that Raphaeli's mapping function (which employs more ambiguities) performs worse than the conventional system with uniform 256QAM, if BER below 10^{-3} is targeted. When we increase the modulation order (equivalently decrease the code rate), the gains are significantly improved to about 0.9 dB. We see that even for these rather short block lengths and very high spectral efficiency, we are very close to uniform capacity (within $0.6dB$ at $BER \approx 10^{-5}$).

In Fig. 9 the throughput envelope curves are given. The envelope is obtained by taking the throughput of the best-performing combination code-rate/mapping function/modulation size at that SNR operating point. We see that the performance is within $1dB$ from uniform capacity ($0.3dB$ at $8b/cu$), and that steady shaping gains of $1dB$ and above are achieved for high spectral efficiency ($\eta > 5b/cu$). Raphaeli's system, even though applicable for a wide SNR range, cannot be used for $\eta > 6.5b/cu$, and at $\eta > 6b/cu$ is already outperformed by the mapping functions designed here. The near-capacity *Turbo Trellis Coded Modulation* (TTCM) from [16] at $\eta = 5b/cu$ is also outperformed by about $1dB$ (for that system we only report the error-free transmission SNR point, or the SNR, for which the BER $< 10^{-5}$). We note that TTCM is generally limited by the constellation size (namely increasing the constellation size increases the complexity exponentially). For comparison we also plot the performance of the sub-constellation shaping system from [4] at $\eta = 3b/cu$ with block length of 64800 bits, which employs an LDPC code and 32APSK modulation. This system is also significantly outperformed. We also note, that the explicit shaping code in such a system significantly increases both the receiver and transmitter complexity, whereas the mapping functions designed here (as discussed in Section II) only add complexity at the de-mapper, which is not so crucial overall.

B. Rayleigh fading channel

Our preliminary results suggested that fading channels will require stronger codes, and therefore increase in modulation size, if the spectral efficiency is to be kept high. This can also be related to the design criteria of trellis codes. Namely increasing the Euclidian distance is not the main design

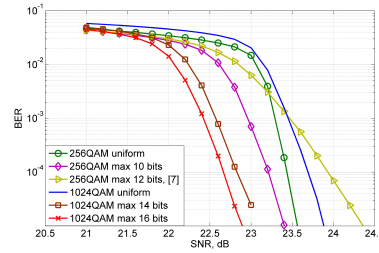


Fig. 8. BER at 7 bits/symbol. Shannon capacity at $21.04dB$, 1024QAM uniform capacity at $22.28dB$, 256QAM uniform capacity at $22.48dB$

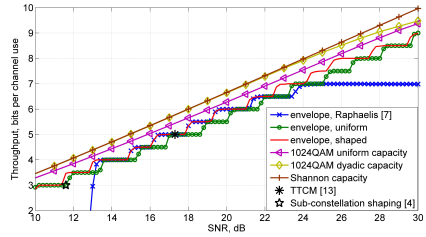


Fig. 9. Throughput envelope, AWGN channel

criterion, but maximizing the product of the metrics on the shortest error-event path and maximizing the length of that path [17]. Decreasing the code rate (reducing the puncturing) directly increases this product for trellis (convolutional) codes. In Fig. 10 the BER performance is given at $\eta = 5b/cu$. We see that already here Raphaeli's system is outperformed. This is due to the required increase in modulation size. The same follows for the TTCM from [18], designed specifically for fading channels, where a 64QAM is used at this spectral efficiency, and the 32 APSK from [4] at $\eta = 3b/cu$. In Fig. 11 a throughput envelope curve is given. We see that our mappings perform close to capacity also on the fading channel (within $1 - 1.5dB$, depending on the modulation size and SNR). The gain over the uniform system vanishes at high SNR due to the limited size of the analyzed constellations (in this case - 1024QAM). As seen from Fig. 10, large constellations with low code rate experience an error-floor at around 10^{-4}. This may be partly contributed to the rather short block lengths considered here. We note that the reported performance of both TTCM systems from [13] (for AWGN channel) and [14] (for fading channels) is achieved for much longer codes - 15000 information bits in [16] and 10000η information bits in [18], whereas our system employs 5000 information bits only.

C. Iterative vs. non-iterative de-mapping

As discussed in Section II, even though usually employed in non-iterative receivers, binary-reflected Gray mapping can

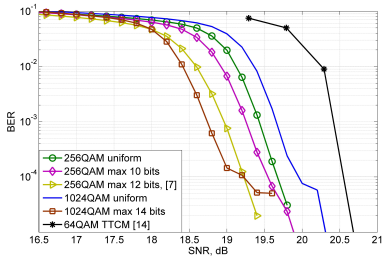


Fig. 10. BER at 5 bits/symbol. Ergodic capacity at $17.15dB$, 64QAM uniform capacity at $19.70dB$, 256QAM uniform capacity at $18.18dB$, 1024QAM uniform capacity at $18.12dB$

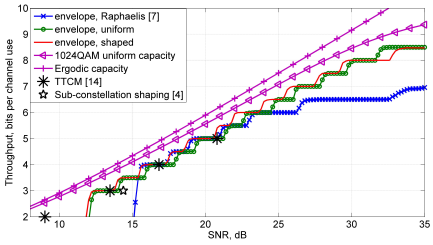


Fig. 11. Throughput envelope, i.i.d. Rayleigh fading channel

still be expected to benefit from iterations when used together with many-to-one shaping and a turbo code. On Fig. 12 we see this effect over a wide range of SNR (i.e. the throughput curves) for AWGN and i.i.d. fading channels. The iterative gain is kept for a wide region of SNR, due to the fact that as the SNR increases, the envelope follows the performance of the higher order modulation formats. i.e. the operating point always falls in the region, where the MI curve appears linear in the log-SNR domain, and the bit-wise MI is smaller than the symbol-wise (see Fig. 5). Towards the very high SNR this gain is not so pronounced because we are approaching the point, where the MI for the largest constellation considered (here 1024QAM), starts to level out.

V. Conclusion

In this paper a family of mapping functions were designed for a turbo coded BICM system employing many-to-one constellation shaping. The mappings were designed based on a Huffman code with binary-reflected Gray code properties. Performance of the mappings was analyzed on AWGN and i.i.d. Rayleigh fading channels over a wide range of SNR and spectral efficiencies. It was shown that they outperform current state of the art TTCM, as well as the original many-to-one shaping from [7], further reducing the gap to capacity. It was also shown that iterative de-mapping improves the

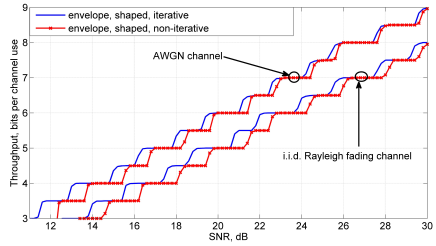


Fig. 12. Illustration of the iterative processing gain

performance of binary-reflected Gray coded mappings, when combined with a turbo code.

References

[1] R. F. H. Fischer, *"Precoding and Signal Shaping for Digital Transmission"*, New York, John Wiley & Sons, 2002

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

[3] J. Bellorado, S. Ghassemzadeh and A. Kavcic, "Approaching the Capacity of the MIMO Rayleigh Flat-Fading Channel with QAM Constellations, Independent across Antennas and Dimensions", *IEEE Trans. Commun.*, vol. 5, no. 6, pp. 1322-1332, 2006

[4] M. C. Valenti, X. Xiang, "Improving DVB-S2 performance through constellation shaping and iterative demapping," *IEEE Military Communications Conference*, pp. 549-554, 2011.

[5] M. Noemm, A. Mourad and P. A. Hoeher, "Superposition Modulation with Irregular Convolutional Coding", GLOBECOM, pp. 2346–2350, 2012.

[6] G. Bocherer, F. Altenbach, and R. Mathar, "Capacity Achieving Modulation for Fixed Constellations with Average Power Constraint", *IEEE International Conference on Communications*, pp. 1-5, 2011.

[7] 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.

[8] S. Le Goff, A. Glavieux, and C. Berrou, "Turbo-Codes and High Spectral Efficiency Modulation" *IEEE International Conference on Communications*, vol. 2, pp. 645-649, 1994.

[9] G. Caire, G. Taricco, and E. Biglieri, "Bit-Interleaved Coded Modulation" *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927-946, 1998.

[10] G. Bocherer and R. Mathar, "Matching Dyadic Distributions to Channels", *Data Compression Conference (DCC)*, pp. 23-32, 2011.

[11] L. Hanzo, *"Near-capacity variable-length coding : regular and exit-chart-aided irregular designs"*. Chichester, West Sussex, U.K, John Wiley & Sons, 2011

[12] D. A. Huffman, "A Method for the Construction of Minimum-Redundancy Codes," Proc. IRE, vol. 40, no. 9, pp. 1098-1101, 1952.

[13] C. Stierstorfer and R. F. H. Fischer, "(Gray) Mappings for bit-interleaved coded modulation," *IEEE Vehicular Technology Conference*, pp. 1703-1707, Dublin, Ireland, Apr. 2007.

[14] F. Schreckenbach, *Iterative Decoding of Bit-Interleaved Coded Modulation*, Ph.D. dissertation, Technical University of Munich, 2007.

[15] J. Hagenauer, "The EXIT Chart", Signal Processing - European Conference, vol. 2, no. 12, pp. 1541-1548, 2004.

[16] P. Robertson and T. Wörz, "Bandwidth-Efficient Turbo Trellis-Coded Modulation Using Punctured Component Codes", *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 206-218, 1998

[17] L. Hanzo, T. H. Liew, B. L. Yeap, R. Y. S. Tee and S. X. Ng *"Turbo Coding, Turbo Equalisation and SpaceTime Coding"*. Chichester, West Sussex, U.K, John Wiley & Sons, 2011

[18] S. X. Ng, O. R. Alamri, Y. Li, J. Kliewer and L. Hanzo, "Near-Capacity Turbo Trellis Coded Modulation Design Based on EXIT Charts and Union Bounds," *IEEE Trans. Commun.*, vol. 56, no. 12, pp. 2030-2039, 2008.

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 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]:

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

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

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

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 $\mathrm{E}[\cdot]$ 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 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.

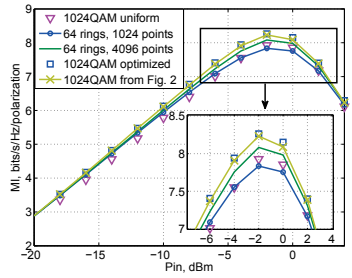


Fig. 1. MI achieved by QAM and ring constellations.

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

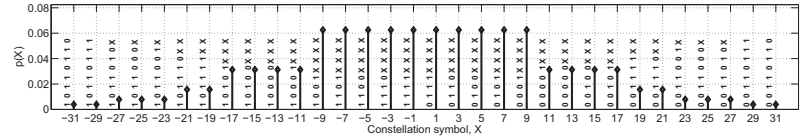


Fig. 2. 32 PAM mapping function for many-to-one constellation shaping

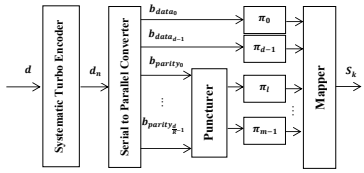


Fig. 3. Turbo coded BICM transmitter for many-to-one constellation shaping

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.

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

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

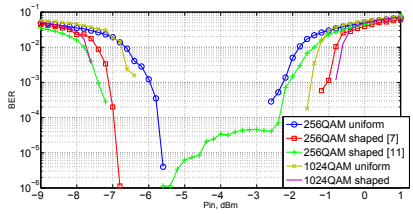


Fig. 4. BER performance with 6 bits/s/Hz/polarization input data rate.

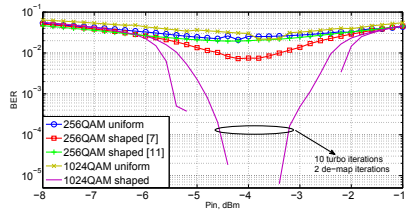


Fig. 5. BER performance with 6.5 bits/s/Hz/polarization input data rate.

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.

Approximating the Constellation Constrained Capacity of the MIMO Channel with Discrete Input

Metodi P. Yankov*, Søren Forchhammer*, Knud J. Larsen* and Lars P. B. Christensen†
* Technical University of Denmark, Department of Photonics Engineering, 2800 Kgs Lyngby, Denmark
† Fingerprint Cards, Lyskær 3CD, 2730 Herlev, Denmark
Email: meya@fotonik.dtu.dk

***Abstract*—In this paper the capacity of a Multiple Input Multiple Output (MIMO) channel is considered, subject to average power constraint, for multi-dimensional discrete input, in the case when no channel state information is available at the transmitter. We prove that when the constellation size grows, the QAM constrained capacity converges to Gaussian capacity, directly extending the AWGN result from [1]. Simulations show that for a given constellation size, a rate close to the Gaussian capacity can be achieved up to a certain SNR point, which can be found efficiently by optimizing the constellation for the equivalent orthogonal channel, obtained by the singular value decomposition. Furthermore, lower bounds on the constrained capacity are derived for the cases of square and tall MIMO matrix, by optimizing the constellation for the equivalent channel, obtained by QR decomposition.**

I. Introduction

In next generation communication systems, high spectral efficiency will be needed in order to satisfy the exponential increase in data rate. The Multiple Input Multiple Output (MIMO) principle with large number of transmit and receive antennas (massive MIMO) will be a key technology to achieving this high spectral efficiency [2]. The ergodic capacity of the MIMO channel was found in [3], and is achieved when the input is a continuous Gaussian with variance given by the power constraint. However, practical transceivers demand signaling with constellations from a finite alphabet, making the analytical calculation of the Constellation Constrained Capacity (CCC) difficult. In [4][5] Blahut and Arimoto derive an iterative algorithm for finding the capacity and the capacity achieving input distribution on an AWGN channel. This algorithm was later modified in [6] to cover MIMO fading channels. However, the complexity of the algorithm grows exponentially both with the constellation size and the number of transmit antennas, making it impractical to calculate the CCC beyond e.g. 2x2 64QAM transmission. In order to cope with this problem, the authors in [6] conjecture, that when the input is taken as a Cartesian product of 1D PAM constellations, the Probability Mass Function (PMF) of the discrete points factorizes into the PMFs of each dimension, thus reducing the complexity of the optimization to 1D. As part of [7] we proved this conjecture to be true. On the other hand, Mutual Information (MI) is usually calculated using Monte-Carlo estimation of the normalized likelihood functions. When the number of receive antennas (the dimensionality of the observations) grows, the complexity still increases dramatically. In [8] the authors derive an analytical approximation of the MI when the input is uniformly distributed. However, it is seen to be inaccurate at low and high SNR. In [9] a better approximation is derived for high SNR via expansion of the MI. Lower and upper bounds are derived in [9] [10] based on the relation between the MI and the Minimum Mean Squared Error. Those bounds are only valid for uniform PMF, and are also quite inaccurate for low-to-moderate SNR. High SNR asymptotic behavior of the MI is studied for arbitrary input distribution in [11], where only AWGN channel is considered.

To our knowledge, the CCC of MIMO channels in the moderate SNR region, which is where practical communication systems tend to operate, is yet to be characterized. Furthermore, in that region the largest shaping gain can be expected for constellations of practical size [12]. The main contributions of this paper are as follows:

- It is proven, that as the constellation size grows, the CCC of MIMO channels approaches the Gaussian capacity, directly extending the AWGN result from [1]. The convergence rate is also the same as in [1]. For a given constellation size, information rates close to Gaussian capacity can be achieved up to a certain SNR point, which can be efficiently found by optimizing the constellation for the equivalent orthogonal channel, obtained by the Singular Value Decomposition (SVD).
- Lower bounds on the CCC of MIMO channels are derived for any SNR, based on the QR decomposition of the channel, using the diagonal elements of the upper-triangular $\mathbf{R}$ matrix. The bounds hold for the cases of square and tall MIMO matrix.
- It is shown empirically that in the low-to-mid SNR region the CCC is the same as the capacity of the equivalent orthogonal channel, obtained by the SVD, whereas in the mid-to-high SNR the above-mentioned lower bound can be used to characterize the MI.

II. Channel model and CCC of orthogonal channels

We consider a standard MIMO channel model:

$$Y = \mathbf{H}X + W, \tag{1}$$

where X is M-dimensional complex random variable vector $X = [X_1, X_2, \dots X_M]^T$, which can be continuous, or discrete taking values from the complex-valued set $\mathcal{X}^M$, obtained as

the Cartesian product of the basic 1D set $\mathcal{X}$. This can be a QAM, APSK, PAM etc. complex-valued set. The matrix $\mathbf{H}$ represents the $[N\mathbf{x}M]$ complex-valued channel, assumed here to have full rank $R = \min(M, N)$. The N dimensional complex AWGN W is assumed here to have unit variance and Y is the N dimensional channel observation. The received SNR with these assumptions is defined as SNR$= P_{av}/M$. We assume the channel realization is known at the receiver, but not at the transmitter. When the input is continuous, the ergodic capacity is achieved by Gaussian input distribution and can be found as [13]:

$$C_G = \mathrm{E}_{\mathbf{H}}\left[\log_2 \det\left(\mathbf{I}^N + \frac{P_{av}}{M}\mathbf{H}\mathbf{H}^H\right)\right], \tag{2}$$

where $\mathrm{E}_{\mathbf{H}}[\cdot]$ denotes the expectation operator w.r.t. $\mathbf{H}$, $\mathbf{I}^N$ is a $[N\mathbf{x}N]$ identity matrix, $(\cdot)^H$ means conjugate transpose and $P_{av} = \sum_{i=1}^{|\mathcal{X}^M|} p(\mathbf{x_i})\alpha|\mathbf{x_i}|^2$ is the average power constraint at the transmitter. Here α is some scaling coefficient, and $\mathbf{x_i}$ is the $i-th$ point in the constellation. In this paper the focus is on finding the capacity and capacity achieving PMF when a fixed input constellation is used and without channel state information at the trasmitter, i.e. uniform power allocation and no pre-coding are employed. The channel capacity when signaling with $\mathcal{X}^M$ and averaging among the possible channel realizations can be expressed as [6]:

$$\begin{aligned} C &= \max_{p(X),\alpha} \mathrm{E}_{\mathbf{H}}\left[\mathcal{I}(X;Y|\mathbf{H})\right] = \\ &= \max_{p(X),\alpha} \sum_{i=1}^{|\mathcal{X}^M|} p(\mathbf{x_i})\left(\log_2 \frac{1}{p(\mathbf{x_i})} + T_i\right), \\ \text{s.t. } &\sum_{i=1}^{|\mathcal{X}^M|} p(\mathbf{x_i})\alpha|\mathbf{x_i}|^2 = P_{av} \text{ and } \sum_{i=1}^{|\mathcal{X}^M|} p(\mathbf{x_i}) = 1 \end{aligned} \tag{3}$$

where $\mathcal{I}(\cdot;\cdot)$ is the MI and:

$$T_i = \mathrm{E}_{\mathbf{H}}\left[\int_{\mathbf{y}} p(\mathbf{y}|\mathbf{x_i},\mathbf{H}) \log_2 p(\mathbf{x_i}|\mathbf{y},\mathbf{H}) d\mathbf{y}\right]. \tag{4}$$

As mentioned before, when the order of modulation and number of dimensions of the signal grow, maximizing (3) is impractical even when performed offline due to the exponential increase in the number of parameters to be optimized, and because it involves numerically calculating the expectation and integration in (4). We found that sufficiently accurate calculation of the MI on e.g. 64QAM 2x2 requires more than 10^5 samples of the observation space, resulting in a likelihood matrix of size [10^5x$2^{2\cdot 6}$]. Assuming larger constellations and larger antenna arrays, e.g. 256QAM and 8x8 set-up, which is already of interest in practical scenarios for single user MIMO in the e.g. IEEE 802.11ac WLAN standard, the number of samples in the output, needed for accurate estimation of the MI grows exponentially, thus making the calculations challenging for a standard computer.

A. Capacity of orthogonal channels

In this section we consider the CCC of orthogonal channels (or set of parallel channels). This means that the channel matrix can be expressed as diagonal, for which $M = N = R$. For each channel realization, the likelihood is Gaussian and factorizes as:

$$p(Y|X,\mathbf{H}) = \prod_{k=1}^{M} p(Y_k|X_k, \mathbf{H_{kk}}), \tag{5}$$

where $\mathbf{H_{kk}}$ is the element on the $k-th$ row and $k-th$ column of $\mathbf{H}$, and X_k is a random variable, representing the $k-th$ dimension of X, taking values from $\mathcal{X}$. As we prove in [7], when the input constellation is constructed as a Cartesian product of M 1D constellations $\mathcal{X}$, the capacity achieving PMF factorizes into its marginal PMFs on each dimension. The conditional distribution $p(X|Y,\mathbf{H})$ then also factorizes as $p(X|Y,\mathbf{H}) = \prod_{k=1:M} p(X_k|Y_k, \mathbf{H_{kk}})$ [7]. The capacity can then be expressed as:

$$\begin{aligned} \hat{C} &= \max_{p(X)} \mathrm{E}_{\mathbf{H}}\left[\mathcal{I}(X;Y|\mathbf{H})\right] = \\ &\max_{p(X)} \mathrm{E}_{\mathbf{H}}\left[\mathcal{H}(X|\mathbf{H}) - \mathcal{H}(X|Y,\mathbf{H})\right] = \\ &\max_{p(X)} \mathrm{E}_{\mathbf{H}}\left[\sum_{k=1}^{M}\left[\mathcal{H}(X_k|\mathbf{H_{kk}}) - \mathcal{H}(X_k|Y_k,\mathbf{H_{kk}})\right]\right] = \\ &\sum_{k=1}^{M} \max_{p(X_k)} \mathrm{E}_{\mathbf{H_{kk}}}\left[\mathcal{I}(X_k;Y_k|\mathbf{H_{kk}})\right], \end{aligned} \tag{6}$$

where $\mathcal{H}$ is the entropy function. When the channel matrix elements are identically distributed, (6) simplifies to:

$$\hat{C} = M \max_{p(X_i)} \mathrm{E}_{\mathbf{H_{ii}}}\left[\mathcal{I}(X_i;Y_i|\mathbf{H_{ii}})\right] \tag{7}$$

for any $i \in [1;M]$, subject to power constraint on the $i-th$ input $P_{av}^i = P_{av}/M$.

III. Capacity of interference channels

In this section the core results of the paper are derived. Let $\mathbf{U}, \mathbf{S}$ and $\mathbf{V}$ be the SVD components of $\mathbf{H}$: $\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^H$. We assume that $\mathbf{S}$ is ordered, so that its first R diagonal elements are non-zero. Let us then consider 3 different channel models:

1) $Y = \mathbf{H}X + W$: one realization of the channel from (1)
2) $\hat{Y} = \mathbf{S}\hat{X} + W$: the channel, obtained by the SVD, where $\hat{Y} = \mathbf{U}^H Y$ and $\hat{X} = \mathbf{V}^H X$
3) $\tilde{Y} = \mathbf{S}X + W$: orthogonal channel, where $\mathbf{S}$ is the diagonal channel matrix.

We denote the MI on each channel as a function of the input distribution with $\mathcal{I}_1(\cdot), \mathcal{I}_2(\cdot)$ and $\mathcal{I}_3(\cdot)$, respectively. Let $\delta_{(\cdot)}$ denote the Dirac-delta function. We will also need the following PDFs:

1) $p_1(X) = \sum_{i=1:|\mathcal{X}^M|} w_i \delta_{\mathbf{x_i}}$
2) $q_1(X) = \sum_{i=1:|\mathcal{X}^M|} w_i \delta_{\mathbf{S}\mathbf{x_i}}$
3) $p_2(X) = \sum_{i=1:|\mathcal{X}^M|} w_i \delta_{\mathbf{V}^H\mathbf{x_i}}$
4) $q_2(X) = \sum_{i=1:|\mathcal{X}^M|} w_i \delta_{\mathbf{S}\mathbf{V}^H\mathbf{x_i}}$
5) $p_G(X) = \mathcal{N}(0, diag(P_{av}/M))$
6) $q_G(X) = \mathcal{N}(0, \mathbf{S}\mathbf{S}^H diag(P_{av}/M))$

In the first 4 PDFs, $w_i \geq 0$ for all i and $\sum_{i=1:|\mathcal{X}^M|} w_i = 1$. In the last 2 PDFs, $diag(P_{av}/M)$ is the covariance matrix

of the Gaussian, which is a diagonal matrix with elements P_{av}/M or $\mathbf{S_{ii}}^2 P_{av}/M$, respectively. Let p_1^* denote the optimal PMF (or the PMF with optimal weights w_i) for Channel 3), i.e. $p_1^* = \arg\max \mathcal{I}_3(p_1(X))$. Likewise, p_2^* is the PMF with the same weights on the rotated version of the original QAM constellation. We will need the following auxiliary theorems:

Theorem 1: For any input PDF, $p_1(X)$, the mutual information on the non-orthogonal channel is the same as on the equivalent orthogonal channel with rotated input: $\mathcal{I}_1(p_1(X)) = \mathcal{I}_2(p_2(X))$

Proof: Given in the Appendix. ■

Theorem 2: When $P_{av}^1 = P_{av}^2 = \cdots = P_{av}^M = P_{av}/M$, the MI on all three channels with a continuous Gaussian input is the same: $\mathcal{I}_1(p_G(X)) = \mathcal{I}_2(p_G(X)) = \mathcal{I}_3(p_G(X))$

Proof: Given in the Appendix. ■

In [1] the authors prove that as the size of the constellation grows, Shannon capacity can be approached for AWGN channels. The proof relies on the fact that the MI is continuous in the *quadratic Wasserstein space*. The loss, incurring from the discrete nature of the input is then continuous in the quadratic Wasserstein distance W_2 between the discrete PMF and the continuous Gaussian distribution. The quadratic Wasserstein distance between two probability measures μ and v from the same space is defined as:

$$W_2(\mu, v) = \inf \left\{ \left(\mathrm{E}_{X,Y} \left[||X - Y||^2 \right] \right)^{1/2} \right\}, \qquad (8)$$

where X and Y are governed by laws μ and v respectively, and the minimum is taken over all joint distributions of (X, Y) [1][14]. In Section II-A we showed that the capacity of an orthogonal channel is the sum of the capacities on each parallel channel, and therefore:

$$\lim_{|\mathcal{X}|\to\infty} \mathcal{I}_3(p_1^*(X)) = \mathcal{I}_3(p_G(X)). \qquad (9)$$

Using the above results, we state the following theorem:

Theorem 3: Consider the interference channel model 1). As the size of the constellation grows, the CCC is $\lim_{|\mathcal{X}|\to\infty} \mathcal{I}_1(p_1^*(X)) = \mathcal{I}_1(p_G(X))$

Proof: The continuity of the MI in the W_2 space result from [1], together with the strict concavity of the MI in the input distribution [15] mean that when SNR> 0:

$$\lim_{|\mathcal{X}|\to\infty} \mathcal{I}_3(p_1^*(X)) = \mathcal{I}_3(p_G(X)) \Leftrightarrow \lim_{|\mathcal{X}|\to\infty} W_2(q_1(X), q_G(X)) = 0. \qquad (10)$$

When $\mathbf{S}$ is full rank, if for some joint PDF $p(X, Y)$ we have that $\mathrm{E}_{X,Y}\left[||X - Y||^2\right] = 0 \Rightarrow \mathrm{E}_{\mathbf{S}X,\mathbf{S}Y}\left[||\mathbf{S}X - \mathbf{S}Y||^2\right] = \mathrm{E}_{\mathbf{S}X,\mathbf{S}Y}\left[||\mathbf{S}(X - Y)||^2\right] = 0$. Then the following is true:

$$\lim_{|\mathcal{X}|\to\infty} W_2(q_1(X), q_G(X)) = 0 \Leftrightarrow \lim_{|\mathcal{X}|\to\infty} W_2(p_1(X), p_G(X)) = 0. \qquad (11)$$

The distribution $p_G(X)$ is rotationally invariant, and therefore $W_2(p_1^*(X), p_G(X)) = W_2(p_2^*(X), p_G(X))$. Then applying (10) we have:

$$\lim_{|\mathcal{X}|\to\infty} W_2(q_2(X), q_G(X)) = 0 \Rightarrow \lim_{|\mathcal{X}|\to\infty} \mathcal{I}_2(p_2^*(X)) = \mathcal{I}_3(p_G(X)). \qquad (12)$$

By Theorems 1 and 2 we have $\mathcal{I}_1(p_1^*(X)) = \mathcal{I}_2(p_2^*(X))$ and $\mathcal{I}_3(p_G(X)) = \mathcal{I}_1(p_G(X))$ which proves the theorem. ■

Consequences of Theorem 3: The main consequence is that as the constellation size grows, Gaussian capacity can be approached for the interference channel for any SNR> 0. This is a direct extension of the result in [1] for AWGN channel. However, there is also a very practical application of Theorem 3: if for a fixed constellation size the CCC of the orthogonal channel approaches the Gaussian capacity, similar CCC can be expected on the equivalent interference channel *without precoding*. Using similar arguments as in [1], the rate at which the gap to Gaussian capacity vanishes can be estimated as $\mathcal{O}(1/|\mathcal{X}|)$.

Theorems 2 and 3 are proven for one realization of the channel. However, they can be extended to cover the ergodic case:

Theorem 4: $\lim_{|\mathcal{X}|\to\infty} \max_{p_1(X)} \mathrm{E}_{\mathbf{H}}\left[\mathcal{I}_1(p_1(X))\right] = C_G$

Proof: Let us re-define $p_1^*(X|\mathbf{H}_k) = \arg\max \mathcal{I}_1(p_1(X)|H = \mathbf{H}_k)$ as the optimal PMF for the $k-th$ channel realization. By Theorem 3 we have that:

$$\lim_{|\mathcal{X}|\to\infty} W_2(p_1^*(X|\mathbf{H}_k), p_G(X)) = 0 \qquad (13)$$

for all k. The Wasserstein distance is a distance measure, and therefore [14]:

$$W_2(p_1^*(X|\mathbf{H}_k), p_1^*(X|\mathbf{H}_j)) \leq W_2(p_1^*(X|\mathbf{H}_k), p_G(X)) + W_2(p_1^*(X|\mathbf{H}_j), p_G(X)). \qquad (14)$$

Taking the limit of large constellations, we get:

$$\lim_{|\mathcal{X}|\to\infty} W_2(p_1^*(X|\mathbf{H}_k), p_1^*(X|\mathbf{H}_j)) \leq \lim_{|\mathcal{X}|\to\infty} W_2(p_1^*(X|\mathbf{H}_k), p_G(X)) + \lim_{|\mathcal{X}|\to\infty} W_2(p_1^*(X|\mathbf{H}_j), p_G(X)) = 0. \qquad (15)$$

The continuity of the MI means that due to the vanishing Wasserstein distance, in the limit of infinitely large constellations, if the optimal PMF on channel j achieves Gaussian capacity, it must also achieve similar capacity on channel k:

$$\lim_{|\mathcal{X}|\to\infty} \mathcal{I}_1(p_1^*(X|\mathbf{H}_k)|H = \mathbf{H}_k) = \mathcal{I}_1(p_1^*(X|\mathbf{H}_j)|H = \mathbf{H}_k) = \mathcal{I}_1(p_G(X)|H = \mathbf{H}_k) \qquad (16)$$

for any j and k. Then for the average MI we have:

$$\lim_{|\mathcal{X}|\to\infty} \mathrm{E}_{\mathbf{H}}\left[\mathcal{I}_1(p_1^*(X, \mathbf{H}_k)) - \mathcal{I}_1(p_G)\right] = 0 \qquad (17)$$

for any k, which proves the theorem. ■

A. Lower bounds via QR decomposition

Let $\mathbf{H} = \mathbf{QR}$ be the QR decomposition of $\mathbf{H}$, where $\mathbf{Q}$ is unitary and $\mathbf{R}$ is upper-triangular. A well known method for detection of MIMO signals uses the form of $\mathbf{R}$ to successively detect each layer by removing the interference from the previously detected layers - Successive Interference Cancellation (SIC). In this section we analyze the maximum rate which can be achieved by SIC under uniform power allocation and i.i.d. on the elements of the channel matrix, for the case of $M \leq N$, and therefore $R = M$.

We introduce two more channel models - Channels 4) and 5), with MI $\mathcal{I}_4(\cdot)$ and $\mathcal{I}_5(\cdot)$, respectively:

4) $\dot{Y} = \mathbf{R}X + W$, where $\dot{Y} = \mathbf{Q}^H Y$
5) $\ddot{Y} = diag(\mathbf{R})X + W$,

where $diag(\mathbf{R})$ means the matrix with the diagonal elements of $\mathbf{R}$ on its diagonal, and zeros elsewhere. Rotation does not change the multivariate Gaussian with i.i.d. on each dimension, and the noise distribution is therefore unchanged. Similarly to $p_1^*(X)$, we define the optimal discrete PMF input to Channel 5) as $p_5^*(X) = \arg\max \mathcal{I}_5(p(X))$.

Theorem 5: $\mathcal{I}_1(p_5^*(X)) \geq \mathcal{I}_5(p_5^*(X))$

Proof: We express the MI on Channel 4) with input $p_5^*(X)$ as:

$$\begin{aligned}\mathcal{I}_4(p_5^*(X)) &= \mathcal{H}(X) - \mathcal{H}(X|\dot{Y}) = \\ &= \mathcal{H}(X) - \mathcal{H}(X_M|\dot{Y}) - \sum_{i\in\{1:M-1\}} \mathcal{H}(X_i|\dot{Y}, X_{\{i+1:M\}}) \geq \\ &\geq \mathcal{H}(X) - \mathcal{H}(X_M|\dot{Y}) - \sum_{i\in\{1:M-1\}} \mathcal{H}(X_i|\dot{Y}), \qquad (18)\end{aligned}$$

where the last inequality follows from the fact, that conditioning does not increase the entropy. Using this argument again, we can write:

$$\mathcal{H}(X_M|\dot{Y}) \leq \mathcal{H}(X_M|\dot{Y}_M) = \mathcal{H}(X_M|\ddot{Y}_M), \qquad (19)$$

where we have also used the fact, that on the $M-th$ layer of Channel 4) there is no interference, and the conditional distributions $p(\dot{Y}_M|X_M)$ and $p(\ddot{Y}_M|X_M)$ for Channels 4) and 5) are the same. Consequently, for the same input $p_5^*(X)$, $p(X_M|\dot{Y}_M)$ and $p(X_M|\ddot{Y}_M)$ are also the same.

Due to the i.i.d. of the channel elements, and applying the chain rule for entropy multiple times, for any i we have:

$$\begin{aligned}\mathcal{H}(Y|X_M, H) &= \mathcal{H}(Y|X_i, H) = \\ \mathcal{H}(\mathbf{Q}^H Y|X_M, H) &= \mathcal{H}(\mathbf{Q}^H Y|X_i, H) \Rightarrow \\ \mathcal{H}(\dot{Y}|X_M, H) + \mathcal{H}(X_M) &= \mathcal{H}(\dot{Y}|X_i, H) + \mathcal{H}(X_i) \Rightarrow \quad (20) \\ \mathcal{H}(\dot{Y}, X_M|H) &= \mathcal{H}(\dot{Y}, X_i|H) \Rightarrow \\ \mathcal{H}(X_M|\dot{Y}, H) + \mathcal{H}(\dot{Y}) &= \mathcal{H}(X_i|\dot{Y}, H) + \mathcal{H}(\dot{Y}) \Rightarrow \\ \mathcal{H}(X_M|\dot{Y}, H) &= \mathcal{H}(X_i|\dot{Y}, H) \leq \\ \leq \mathcal{H}(X_M|\ddot{Y}, H) &= \mathcal{H}(X_i|\ddot{Y}_i, H). \qquad (21)\end{aligned}$$

Equation 20 follows from the fact, that due to the i.i.d. of the channel elements the marginal distributions on each dimension of X are identical. In (21) we have used that by definition, Channel 5) is orthogonal. Inserting (21) in (18) we have:

$$\begin{aligned}\mathcal{I}_4(p_5^*(X)) &\geq \mathcal{H}(X) - \sum_{i\in\{1:M\}} \mathcal{H}(X_i|\dot{Y}) \geq \\ &= \mathcal{H}(X) - \sum_{i\in\{1:M\}} \mathcal{H}(X_i|\ddot{Y}_i) = \mathcal{I}_5(p_5^*(X)). \qquad (22)\end{aligned}$$

Similarly to Theorem 1, we have that $\mathcal{I}_4(p_5^*(X)) = \mathcal{I}_1(p_5^*(X))$, which proves the theorem. ■

We have arrived at a lower bound for the channel capacity. The MI $\mathcal{I}_5(p(X))$ can be easily optimized and calculated, in a manner, similar to $\mathcal{I}_3(p(X))$, since the channel is orthogonal. We only need the elements on the diagonal of the $\mathbf{R}$ matrix. Even though Theorem 5 was proven for $p_5(X)^*$, it actually follows for any $p(X)$, for which the dimensions of X are i.i.d., e.g. the uniform PMF.

In the case of continuous Gaussian input with uniform power allocation, the proof of Theorem 5 can be simplified. We can notice that the outputs of Channels 4) and 5) in that case are Gaussians, with respective covariance matrices:

$$\dot{\Sigma} = \mathrm{Cov}\left[\dot{Y}\right] = \frac{P_{av}}{M}\mathbf{R}\mathbf{R}^H + \Sigma_W \qquad (23)$$

$$\ddot{\Sigma} = \mathrm{Cov}\left[\ddot{Y}\right] = \frac{P_{av}}{M}diag(\mathbf{R})diag(\mathbf{R})^H + \Sigma_W, \qquad (24)$$

where Σ_W is the diagonal covariance matrix of the noise. It is then trivial to show that:

$$\det\dot{\Sigma} \geq \det\ddot{\Sigma} \Rightarrow \mathcal{I}_4(p_G(X)) \geq \mathcal{I}_5(p_G(X)), \qquad (25)$$

with equality if SNR $= \infty$.

B. Discussion of the theorems in Section III

The main implication of the theorems in this section is that while the ergodic Gaussian capacity of the orthogonal channels, obtained from the SVD of each channel realization can be approached with a finite size constellation, it can be expected that the ergodic Gaussian capacity of the interference channel is also approached with the same constellation, having the same PMF. As discussed in Section II-A, the CCC and the capacity achieving PMF of the orthogonal channel are easily calculated by the Blahut-Arimoto algorithm, taking $\mathbf{H}_{\mathrm{ii}}$ in (7) to have the distribution of the singular values of the MIMO matrix. For large MIMO it is shown in [16] that the singular values distribution of Gaussian distributed channel matrix coefficients follows a quarter-quadratic law, which can be used to generate singular values for the 1D optimization. For small matrices the SVD is simple to calculate, and the distribution can be accurately approximated by Monte Carlo methods. When calculating the QR decomposition based lower bounds, the distribution of the elements on the diagonal of the $\mathbf{R}$ matrix is needed. Even though this distribution is not known, similar approach can be taken - draw matrices $\mathbf{H}$ from their known distribution, perform the QR decomposition on each $\mathbf{H}$, and use the diagonal elements of $\mathbf{R}$ instead of $\mathbf{H}_{\mathrm{ii}}$ when maximizing (7).

As we mentioned in the introduction, complex-valued input sets are the focus here. When $\mathcal{X}$ is the popular QAM set, which is a product of two real-valued PAM sets, the reduction in complexity is relevant further down to the PAM set. Theorems 4 and 5 can then be used without loss of generality. Equation (7) becomes $\hat{C} = 2M \max_{p(X_i)} \mathrm{E}_{\hat{\mathbf{H}}_{\mathbf{ii}}} \left[\mathcal{I}(X_i; Y_i | \hat{\mathbf{H}}_{\mathbf{ii}}) \right]$, where $\hat{\mathbf{H}}$ is the real-valued equivalent of $\mathbf{H}$, and each dimension of the input is taken from the corresponding PAM set. This is the model we consider in the following sections.

C. Some near-optimal input PMFs

Since we will exclusively use orthogonal channels to approximate the capacity in (3), it is worth examining the implications the distribution of the singular values has on the optimal PMF of X. Figure 1 depicts the optimal 8PAM component PMFs, i.e. $p(X) = \arg\max \mathrm{E}_{\mathbf{S_{11}}} \left[\mathcal{I}(X_1; \hat{Y}_1 | \mathbf{S_{11}}) \right]$, for transmission of different rank $R = M = N$ at the same average SNR, for which $\mathrm{E}_{\mathbf{H}} \left[\mathcal{I}_3(p_1^*(X)) \right] \approx C_G$. It is interesting to see how the shape of the optimal PMF changes when we increase the rank. This can be contributed to the fact, that the distribution of the singular values broadens. The AWGN channel can be considered as MIMO with zero variance of the singular values. The optimal input PMF on the AWGN channel therefore approaches the continuous Gaussian for this SNR. On the scalar fading channel, the singular values are Rayleigh distributed. As the rank of transmission grows, the distribution of singular values approaches the broader quarter-circular law. The optimal PMF in that case must account for broader range of SNR. It is well known, and can also be seen in [6], that uniform PMF approaches capacity at low and high SNR. The optimal PMF is therefore pushed to uniform when the rank of transmission increases. In Fig. 2, the histogram of the elements on the diagonal of the $\mathbf{R}$ matrix is shown. We see that it is narrower than the distribution of the singular values. We will use this fact in the next section to analyze the rates, achieved on the orthogonal Channels 3) and 5).

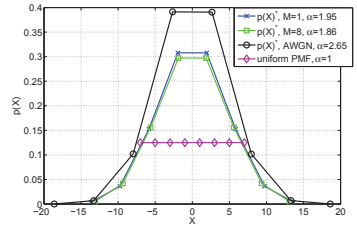


Fig. 1. 8PAM PMFs for different channels at average SNR $= 8dB$. The scaling coeficient α, resulting in the optimal PMF is indicated in the legend. As α increases, more and more weight needs to be put on points with lower energy, in order to keep the average power the same.

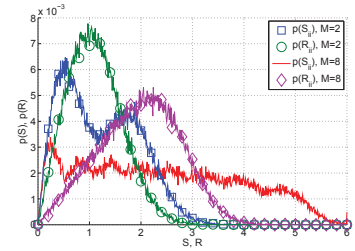


Fig. 2. Histogram of the singular values and the elements on the diagonal of the $\mathbf{R}$ matrix for MIMO of different rank

IV. Numerical calculation of capacity

In this section we provide Monte Carlo based calculation of the capacity for various channels. In Fig. 3, the ergodic capacity for the 2x2 i.i.d. MIMO Rayleigh fading channel, i.e. $\mathrm{E}_{\mathbf{H}} \left[\mathcal{I}_3(p_G(X)) \right]$ is shown, together with the 64QAM CCC, i.e. $\max \mathrm{E}_{\mathbf{H}} \left[\mathcal{I}_1(p(X)) \right]$, the capacity of the SVD based orthogonal channel, i.e. $\max \mathrm{E}_{\mathbf{H}} \left[\mathcal{I}_3(p(X)) \right]$, and the QR decomposition based lower bound - $\max \mathrm{E}_{\mathbf{H}} \left[\mathcal{I}_5(p(X)) \right]$. We directly see the region, where the limits in Theorem 4 are approached: up to around SNR$= 10dB$. As the SNR increases, the gap to capacity also increases due to the limited size of the constellation. As shown in [10], when the input to an orthogonal channel is discrete, orthogonal inputs can be suboptimal. In Fig. 3 this effect can be seen, as $\max \mathrm{E}_{\mathbf{H}} \left[\mathcal{I}_1(p_1(X)) \right] = \max \mathrm{E}_{\mathbf{H}} \left[\mathcal{I}_2(p_2(X)) \right] > \max \mathrm{E}_{\mathbf{H}} \left[\mathcal{I}_3(p_1(X)) \right]$, or a rotated version of the original QAM performs better on the orthogonal channel. The QR based lower bound in the low SNR is seen as a worse approximation than the SVD based capacity. In this regime the noise is the limiting factor, and the inequality in (25) becomes more and more strict. In the moderate to high SNR we see that the lower bound becomes tighter, and exceeds the SVD based channel capacity. This is due to the distribution of the diagonal elements of the $\mathbf{R}$ matrix (see Fig. 2). Tighter distribution means that the optimal PMF does not need to account for high and low instantaneous SNR, where uniform PMF is optimal, i.e. the channel is more stable. The SVD based channel on the other hand has optimal PMF, which must be robust to deep fades and vanishing fades. It is therefore pushed to uniform PMF, resulting in lower average MI.

In Fig. 4, the SVD based channel capacity for a 8x8 setup is given, together with the QR decomposition based lower bounds. For SNR $< 10, 16$, and $24dB$ for 64, 256 and 1024 QAM, respectively, the SVD based capacity is approaching Gaussian capacity, and we can therefore expect that in those SNR regions, $\max \mathrm{E}_{\mathbf{H}} \left[\mathcal{I}_1(p_1(X)) \right]$ is also close to the Gaussian capacity. As before, the QR based lower bounds are worse at low SNR, but become better at moderate to high SNR. The envelope of the two curves - the SVD and QR based capacities - can therefore serve as an approximation to the CCC of the MIMO channel for the entire SNR region.

V. Future work

As mentioned before, in this paper lower bounds are derived only for the case of $M \leq N$. When $M > R$, the last layer

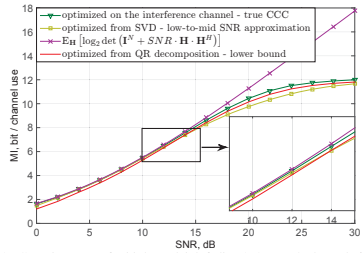


Fig. 3. Capacity curves for i.i.d. Rayleigh fading 2x2 MIMO channel, 64QAM

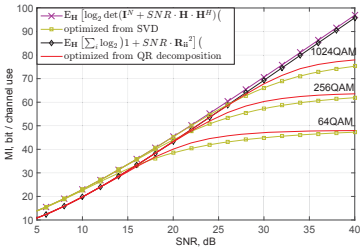


Fig. 4. Capacity curves for i.i.d. Rayleigh fading 8x8 MIMO channel

of Channel 4) will no longer be interference free, which was a necessary condition for stating (19). An interesting area for future research is to provide non-trivial lower bounds for the case of $M > R$.

VI. Conclusion

In this paper the capacity of a MIMO channel with discrete input was considered in the case when no channel knowledge at the transmitter is available. It was proven that as the constellation size grows, the capacity of the interference channel converges to the capacity of the equivalent orthogonal channel, obtained by the SVD, and consequently approaches Gaussian capacity. Simulations showed that values close to that rate can be achieved up to a certain SNR point for QAM constellations of a given size. Using the Blahut-Arimoto algorithm, the capacity of the orthogonal channel can be easily calculated, and the SNR threshold can be obtained, together with the capacity achieving PMFs for each SNR point. These PMFs can then be used on the ergodic interference channel, and similar capacity can be expected up to the SNR threshold. Lower bounds on the constellation constrained capacity were also derived for MIMO channels with square or tall matrix, using the capacity of the QR decomposition based channel. The envelope of the SVD based approximation and the QR based lower bounds can then be used to estimate the constellation constrained capacity for the entire SNR region.

VII. Appendix

A. Proof of Theorem I

Since $\mathbf{U}$ is a rotation matrix, we have:

$$\mathcal{H}(\hat{Y}|\cdot) = \mathcal{H}(\mathbf{U}\hat{Y}|\cdot) - \log_2|\det \mathbf{U}| = \mathcal{H}(\mathbf{U}\hat{Y}|\cdot) = \mathcal{H}(Y|\cdot).$$

Similarly $\mathcal{H}(X|\cdot) = \mathcal{H}(\hat{X}|\cdot)$. Then it is clear that:

$$\mathcal{I}(X;Y|\mathbf{H}) = \mathcal{I}(\hat{X};Y|\mathbf{H}) = \mathcal{I}(Y;\hat{X}|\mathbf{H}) = \mathcal{I}(\hat{Y};\hat{X}|\mathbf{H}).$$

B. Proof of Theorem II

The distribution $p_G(X)$ is rotationally invariant, i.e. $X \equiv \mathbf{V}^H X \Rightarrow \mathcal{I}_2(p_G(X)) = \mathcal{I}_3(p_G(X))$ and by Theorem 1 $\mathcal{I}_2(p_G(X)) = \mathcal{I}_1(p_G(X))$.

Acknowledgment

The authors would like to thank Prof. Dr. sc. techn. Gerhard Kramer at the Technical University of Munich for the fruitful discussion towards deriving the QR decomposition based lower bounds on the constellation constrained MIMO capacity.

References

[1] Y. Wu and S. Verdú, "Functional Properties of Minimum Mean-Square Error and Mutual Information", *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1289-1301, 2012

[2] T. Marzetta, O. Edfors, F. Tufvesson, E. G. Larsson, "Massive MIMO for Next Generation Wireless Systems", *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186-195, 2014

[3] A. Goldsmith and P. Varaiya, "Capacity of Fading Channels With Channel Side Information," *IEEE Trans. Inf. Theory*, vol. 43, pp. 1986-1992, 1997.

[4] R. Blahut, "Computation of Channel Capacity and Rate-Distortion Functions," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 460-473, 1972.

[5] S. Arimoto, "An Algorithm for Computing the Capacity of Arbitrary Discrete Memoryless Channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 14-20, 1972.

[6] J. Bellorado, S. Ghassemzadeh and A. Kavcic, "Approaching the Capacity of the MIMO Rayleigh Flat-Fading Channel with QAM Constellations, Independent across Antennas and Dimensions", *IEEE Trans. Commun.*, vol. 5, no. 6, pp. 1322-1332, 2006

[7] M. Yankov, S. Forchhammer, K. J. Larsen and L. P. B. Christensen, "Factorization Properties of the Optimal Signaling Distribution of Multi-dimensional QAM Constellations", *IEEE 6-th International Symposium on Communications, Control and Signal Processing* pp. 384 - 387, 2014

[8] J. Akhtman and L. Hanzo, "Closed-form Approximation of MIMO Capacity," *IET Electronics Letters*, vol. 45, no. 1, pp. 68-69, 2009.

[9] M. Rodrigues, "Multiple-Antenna Fading Channels With Arbitrary Inputs: Characterization and Optimization of the Information Rate", *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 569-585, 2014

[10] F. Prez-Cruz, M. R. D. Rodrigues and S. Verdú, "MIMO Gaussian Channels With Arbitrary Inputs: Optimal Precoding and Power Allocation", *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1070-1084, 2010

[11] A. Alvarado, F. Brännström, E. Agrell and T. Koch, "High-SNR Asymptotics of Mutual Information for Discrete Constellations With Applications to BICM", *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1061-1076, 2014

[12] M. Yankov, S. Forchhammer, K. J. Larsen and L. P. B. Christensen, "Rate-adaptive Constellation Shaping for Turbo-coded BICM", *IEEE International Conference on Communications*, pp. 2112 - 2117, 2014

[13] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj and H. V. Poor, *MIMO Wireless Communications*, Cambridge University Press 2007

[14] C. Villani, *Optimal Transport: Old and New*, Berlin, Germany, Springer-Verlag, 2008.

[15] I. C. Abou-Faycal, M. D. Trott and S. Shamai, "The Capacity of Discrete-Time Memoryless Rayleigh-Fading Channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1290-1301, 2001.

[16] A. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*, Now Publishers, 2004

Compensation of XPM Interference by Blind Tracking of the Nonlinear Phase in WDM Systems with QAM Input

Tobias Fehenberger[(1)], Metodi P. Yankov[(2)], Luca Barletta[(1)], and Norbert Hanik[(1)]

[(1)] Institute for Communications Engineering, Technische Universität München, 80333 Munich, Germany
[(2)] Department of Photonics Engineering, Technical University of Denmark, Kgs. Lyngby 2800, Denmark
tobias.fehenberger@tum.de

Abstract *Exploiting temporal correlations in the phase, achievable rates are studied and a blind trellis-based receiver is presented. Gains of 0.5 bit per symbol are found in point-to-point links irrespective of the symbol rate. These gains disappear in network configurations.*

Introduction

Lower bounds on the capacity of optical fiber systems have been studied [1] by treating the fiber nonlinearities as additive and white, i.e., uncorrelated, Gaussian noise (AWGN). Recent studies consider the dependence on modulation format [2], and temporal [3] and additionally spectral [4] correlations due to cross-phase modulation (XPM). By exploiting the correlations of the phase noise, revised rates have been reported [3,4] that improve considerably on the previous bounds [1]. These increased achievable rates have been calculated for Gaussian input and point-to-point links. To our knowledge, only one implementation to obtain these rates has been presented [4], yet it uses knowledge of the sent symbols and is therefore not immediately applicable in practice.

In this paper, in place of Gaussian input, we consider uniformly distributed quadrature amplitude modulation (QAM) symbols to show achievable rates for a correlation-aware receiver that exploits the strong temporal correlations within a block of symbols. A blind trellis-based phase tracking algorithm is also presented. Point-to-point links and optical fiber networks are studied.

Phase Noise Receiver

We consider the transmission of a block of N symbols x^N over an optical fiber system and model the input-output relation at time k as

$$y_k = x_k \exp(j\phi_k) + n_k^{\mathrm{NLI}} + n_k^{\mathrm{ASE}}, \qquad (1)$$

where the complex Gaussian variates n_k^{NLI} and n_k^{ASE} denote the additive nonlinear interference noise (NLI) and the amplified spontaneous emission (ASE) noise from amplifiers, respectively. The phase noise term ϕ_k has temporal correlation between neighboring symbols mainly due to XPM, and is further assumed to be block-wise constant [3]. This assumption is justified by the auto-correlation function (ACF) shown in Fig. 2, which is discussed later in detail.

We exploit the intra-block correlation to calculate achievable rates as follows, without making explicit use of frequency correlation [4]. At the receiver, the angle between x_k and y_k is estimated for every k. This angle includes both the correlated phase noise ϕ_k as well as the additive noise terms n_k^{NLI} and n_k^{ASE}. Note that there might be correlations in n_k^{NLI}, but they are not considered in this work. The mean angle $\bar{\phi}_k$ of w past symbols is calculated as a moving-window average,

$$\bar{\phi}_k = \angle \sum_{l=1}^{w} x_{k-l}^* y_{k-l}, \qquad (2)$$

where x^* is the complex conjugate of x. We are allowed to use all y_k but only past x_k to obtain an achievable rate, which becomes apparent by the chain rule of mutual information between input and output sequences [1]. If the phase noise samples $\angle(x_k^* y_k)$ are not or are only weakly correlated, then $\bar{\phi}_k \approx 0$. The XPM-induced phase noise of the received symbol y_k is compensated by a phase rotation,

$$y_k' = y_k \cdot \exp(-j\bar{\phi}_k). \qquad (3)$$

Finally, an achievable rate is calculated from the symbols y_k' using circular Gaussian statistics on a symbol-by-symbol basis [5]. We call this correlation-aware processing a phase noise (PN) receiver, while we speak of an AWGN receiver when correlations are neglected and the rates are calculated directly from y_k with circular Gaussian statistics.

Blind Trellis-Based Phase Tracking

While the PN receiver produces an achievable rate, it requires knowledge of the past symbols, which makes it impractical. In order to build a practical receiver we model the phase noise as

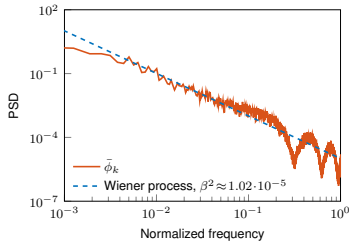


Fig. 1: Power spectral density (PSD) of the estimated phase noise $\bar{\phi}_k$, and a true Wiener process. The estimate is obtained from symbols after fiber transmission of 9 WDM channels.

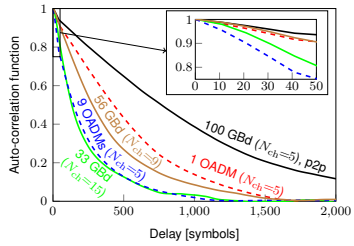


Fig. 2: ACF of point-to-point (p2p) WDM systems at different symbol rates (solid), and different network configurations at 100 GBd (dashed). Inset: Zoomed ACF for up to 50 symbols.

a Wiener process,

$$\phi_k = \phi_{k-1} + \beta v_k, \tag{4}$$

where the samples v_k's are i.i.d. and drawn from a standard Gaussian distribution. The scalar β^2 is obtained offline from the PN receiver:

$$\beta^2 = \mathrm{E}_k[(\bar{\phi}_k - \bar{\phi}_{k-1})^2]. \tag{5}$$

Figure 1 shows a good match between the power spectral density (PSD) of the phase noise process $\bar{\phi}_k$ obtained from simulation data, and the theoretical PSD of a Wiener process with $\beta^2 \approx 1.02 \cdot 10^{-5}$, thus justifying the Wiener phase noise model. Based on it, we implement a trellis-based receiver[5] where each state represents the distribution of the phase noise given the channel output samples, $p(\phi_k|y_1^N)$. The phase is discretized into a finite number of bins within a limited range that is obtained offline from the PN receiver. A large number of bins improves the phase estimate ϕ_k but increases the complexity. The model (4) allows for the factorization of $p(\phi_k|y_1^N)$ into $p(\phi_k, y_1^k)p(y_{k+1}^N|\phi_k)$, which can be efficiently calculated by the BCJR algorithm. The posterior distribution of the input is then

$$p(x_k|y_1^N) = \sum_{\phi_k} p(\phi_k|y_1^N)p(x_k|\phi_k, y_k). \tag{6}$$

The second term in Eq. (6) is calculated using Bayes' theorem, where the likelihood $p(y_k|\phi_k, x_k)$ is circular Gaussian with zero mean and variance estimated offline. The posterior shown in Eq. (6) is used in the achievable rate calculation and in the demodulation process[5]. The PN receiver and this trellis algorithm work for any input distribution.

Numerical Analysis

We investigate a single-polarization wavelength division multiplexing (WDM) system where each WDM channel uses 1024-QAM and a sinc pulse shape. The guard band between neighboring channels is 2% of their bandwidth and the overall spectral width is constant at 510 GHz. We study N_{ch}=5, 9, 15 WDM channels, with a symbol rate of 100, 56, and 33 GBd, respectively.

The signal propagates over 1000 km of standard single-mode fiber with γ=1.3 (W·km)$^{-1}$ and D=17 ps/nm/km. Ideal distributed amplification is employed to compensate for the fiber loss of α=0.2 dB/km. Fiber propagation is simulated using the split-step Fourier method with 32 samples per symbol and 0.1 km step size. Optical add-drop multiplexers (OADMs) are inserted into the link when a network setup is studied. An OADM is modeled by ideal band-pass filtering of the center WDM channel, creating new WDM neighbors and combining the old center channel and the new neighbors. We consider point-to-point (p2p) links, links with one OADM at 500 km, and with an OADM every 100 km, i.e., 9 OADMs in total.

At the receiver, the center WDM channel is ideally band-pass filtered, digitally back-propagated (DBP) to remove self-phase modulation, and down-sampled. The received symbols are either not processed further (AWGN receiver), processed with the PN receiver, or with the blind trellis phase tracking, and achievable rates are calculated on a symbol-by-symbol basis using circular Gaussian noise statistics. We also tested conditional bivariate Gaussian statistics and found no significant difference in achievable rates. The parameter w of Eq. (2) is set to 40 for the considered optical system parameters. Simulations show that w in the range between 30 and 80 is not critical for calculating $\bar{\phi}_k$.

The temporal ACF of the phase noise is shown in Fig. 2. We use simulation data and the blockwise phase noise model[3] for computing the ACF. We observe that temporal correlations are re-

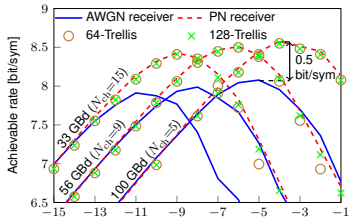


Fig. 3: Achievable rates for three WDM systems. The gain from the PN receiver (dashed) and the trellis processing (markers) over the AWGN receiver (solid) is 0.5 bit/sym at the optimum power for every WDM configuration.

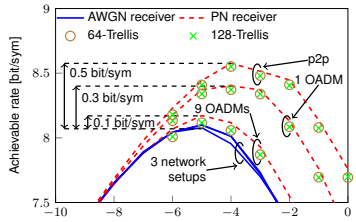


Fig. 4: Achievable rates for N_{ch}=5 WDM channels and three link configurations: point-to-point (p2p), one OADM after 500 km, and one OADM every 100 km (9 OADMs in total). The fiber length is always 1000 km.

duced when the symbol rate per channel is decreased (recall that the total WDM bandwidth is kept constant), and also when network elements are inserted in the fiber. We will next analyze whether the correlations that are apparent from the ACF translate into rate gains.

Achievable rates for the three different WDM setups are compared in Fig. 3. For the AWGN receiver, the maximum rate for 15 channels is 7.9 bits per symbol (bit/sym) and increases to 8.1 bit/sym for 5 channels. This is because for fewer WDM channels, single-channel (SC) DBP is able to cancel a larger amount of nonlinearities. Exploiting the block-wise correlations in the receiver improves the achievable rate by 0.5 bit/sym, which is comparable to simulations with Gaussian input[3]. The gain from the PN receiver is found to be constant for all three WDM setups, despite the dependence of the ACF on the per-channel symbol rate shown in Fig. 2.

The rates obtained with blind phase tracking (markers in Fig. 3) closely approach the rates of the PN receiver for all considered configurations. We also observe that 64 trellis states are sufficient to get full gains at relevant powers.

Achievable rates for N_{ch}=5 channels and three different network setups are shown in Fig. 4. For the AWGN receiver, the maximum rates are about 8.1 bit/sym, independent of the network configuration. When one OADM is inserted in the center of the link, the rate of the PN receiver is reduced from 8.6 bit/sym to 8.4 bit/sym, which is a decrease in gain by 0.2 bit/sym in comparison to the 0.5 bit/sym for the p2p case without OADMs. This is because the coherent build-up of correlations is effectively terminated half-way during propagation. Further simulations show that the center of the link is the worst among all potential locations of one OADM with respect to achievable rates. In a network with 9 OADMs, the gain is reduced to less than 0.1 bit/sym due to repeated filtering of the co-propagating WDM channels and the addition of new, uncorrelated channels. For all network setups, the trellis processing gives rates very close to the PN receiver.

For a long-haul p2p fiber system[6] with dual-polarization 16-QAM at 28 GBd, 60 spans of 100 km each, lumped amplification and SC DBP, we found that the gain by applying the PN receiver was 0.1 bit/sym. This limited gain is mainly attributed to lumped amplification leading to less temporal correlation of the XPM phase noise[4].

Conclusions

We show gains in achievable rate of up to 0.5 bit/sym by exploiting correlation in the nonlinear phase noise. These gains are also obtained by trellis-based processing without knowledge of input symbols. Larger gains are expected by using models that resemble the XPM-induced correlations better than a block-wise constant phase. Especially correlations in the additive NLI noise term could be investigated.

References

[1] R.-J. Essiambre *et al.*, "Capacity limits of optical fiber networks," JLT **28**(4), 2010.

[2] A. Mecozzi and R.-J. Essiambre, "Nonlinear shannon limit in pseudolinear coherent systems," JLT **30**(12), 2012.

[3] R. Dar *et al.*, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," Opt. Exp. **21**(22), 2013.

[4] M. Secondini and E. Forestieri, "On XPM mitigation in WDM fiber-optic systems," PTL **26**(22), 2014.

[5] L. Barletta *et al.*, "Pilot-aided trellis-based demodulation," PTL **25**(13), 2013.

[6] T. Fehenberger *et al.*, "On achievable rates for long-haul fiber-optic communications," Opt. Exp. **23**(7), 2015.

Low-Complexity Tracking of Laser and Nonlinear Phase Noise in WDM Optical Fiber Systems

Metodi P. Yankov, *Student Member, IEEE,* Tobias Fehenberger, *Student Member, IEEE,* Luca Barletta, *Member, IEEE,* and Norbert Hanik, *Senior Member, IEEE*

***Abstract*—In this paper the wavelength division multiplexed (WDM) fiber optic channel is considered. It is shown that for ideal distributed Raman amplification (IDRA), the Wiener process model is suitable for the non-linear phase noise due to cross phase modulation from neighboring channels. Based on this model, a phase noise tracking algorithm is presented. We approximate the distribution of the phase noise at each time instant by a mixture of Tikhonov distributions, and derive a closed form expression for the posterior probabilities of the input symbols. This reduces the complexity dramatically compared to previous trellis based approaches, which require numerical integration. Further, the proposed method performs very well in low-to-moderate signal-to-noise ratio (SNR), where standard decision directed (DD) methods, especially for high order modulation, fail. The proposed algorithm does not rely on averaging, and therefore does not experience high error floors at high SNR in severe phase noise scenarios. The laser linewidth (LLW) tolerance is thereby increased for the entire SNR region compared to previous DD methods. In IDRA WDM links the algorithm is shown to effectively combat the combined effect of both laser phase noise and non-linear phase noise, which cannot be neglected in such scenarios. In a more practical lumped amplification scheme we show near-optimal performance for 16QAM, 64QAM and 256QAM with LLW up to 100kHz, and reasonable performance for LLW of 1MHz for 16QAM and 64QAM, at the moderate received SNR region. The performance in these cases is close to the information rate achieved by the above mentioned trellis processing.**

***Index Terms*—Phase noise, WDM, optical fiber communication, Wiener process, Cross phase modulation**

I. Introduction

THE non-linear phase noise (NLPN) is the main factor for the currently limited achievable rates on the wavelength division multiplexed (WDM) optical fiber channels [1]. Due to the interaction between self and cross phase modulation (SPM and XPM, resp.) effects, the amplified spontaneous emission (ASE) noise and chromatic dispersion in the fiber, the non-linear phase rotation due to SPM and XPM generally cannot be canceled completely at the receiver. Furthermore, XPM and SPM introduce memory in the channel, which makes optimal detection even more challenging. Recent works have gone into modeling this memory [2]–[4], and shown that the phase noise is highly correlated in time and frequency. A model is also proposed that allows to calculate the auto correlation function (ACF) of the phase noise. Using the time domain ACF properties, the authors in [2] separate the contribution of the noise into linear and non-linear part, and thereby are able to show potential increase in the maximum achievable rate in a point to point WDM link. A strategy for canceling the phase noise is proposed in [5], based on a frequency domain equalizer. This work is extended in [6] to a multi-carrier modulation, which is shown to be beneficial in combating XPM interference. Gaussian input is assumed in [2] [5] [6], which is not currently realizable in practice.

On top of the NLPN, fiber-optic systems generally suffer from phase noise due to imperfect lasers. The non-zero laser line width (LLW) results in time-varying carrier phase offset. The Wiener process has previously been shown to be accurate in modeling the laser phase noise, and also the NLPN for a few particular cases [7]. For QPSK constellations simple methods such as Viterbi and Viterbi are shown to be effective for carrier phase recovery. Extensions of this method to larger order constellations are possible, e.g. [8]. However, they generally require hard decision on the signal's amplitude before the phase is estimated. Other types of decision directed (DD) methods are available for higher order constellations, e.g. [10] [11] [12]. In [12], a second order DD phase-locked loop (PLL) method is employed, which is shown to be effective also in the presence of frequency fluctuations. We highlight the method from [10] as very popular among engineers, due to its simple implementation. DD methods typically require an SNR relatively high to the order of modulation (equivalently, low symbol error rate (SER)). Alternatively, the phase offset may be estimated from very long window averages, which makes the system unreliable in severe phase noise scenarios, since the phase varies significantly within the window. Another problem with the method from [10] is its vulnerability to phase slips. A modulation size independent method was derived in [14] for quadrature amplitude modulation (QAM) constellations, which aims at forcing the received symbols in each quadrant to the original square shape, and uses a PLL-like circuit to track the phase offset. This method also suffers greatly at low SNR/high SER, however it is very simple to implement. The above methods do not generally exploit the Wiener process model for the phase noise in order to improve the estimate. The authors in [13] propose an extension to the algorithm from [10], where a pilot-based coarse estimation is combined with the sub-sequent DD estimation, in order to combat the phase-slip problem.

M. Yankov is with the Department of Photonics Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark: e-mail: meya@fotonik.dtu.dk

T. Fehenberger, L. Barletta and N. Hanik are with the Institute for Communications Engineering, Technische Universität München, 80333 Munich, Germany.

Iterative decoding and laser phase-noise cancellation is proposed in [15], later extended to cover the above mentioned NLPN in WDM systems [16]. Up to 16QAM is considered there. The clear problems with iterative receivers are complexity and latency.

In [17], a Kalman filter is used that needs to be linearized since the output is not a linear function of the phase noise. The Kalman filter approach has low complexity and was shown to be near-optimal in some cases [18]. However, it suffers performance degradation for high information rates and SNRs. In a more general approach [19] the phase is discretized into bins, and trellis processing is used for phase tracking, Mutual Information (MI) calculation and demodulation. This method allows for processing of non-linear functions of the phase. Both [17] and [19] are basically a special case of the sum-product algorithm for finding marginal and posterior distributions, which can be represented with a graph. In [20] this algorithm is presented in the context of phase noise, where the graph may be constructed as a Markov chain. Several approximations are also proposed there for constant amplitude modulation, such as phase-shift keying (PSK). Of these approximations we mention the modeling of the distribution of the phase noise as a Tikhonov (also known as von Mises) distribution.

In this paper we show that the Wiener process is suitable for modeling the phase noise due to XPM in WDM systems with ideal distributed raman amplification (IDRA). Then we propose a simple approximate detector based on phase noise tracking, which is able to combat the combined effect of laser phase noise and NLPN in the fiber-optic channel. Focus is given on discrete input constellations, particularly QAM.

II. Channel Model

The fiber channel model under investigation is given in Fig. 1. Data is modulated into constellation symbols x, which are drawn from a finite size alphabet $\mathcal{X}$. The modulated symbols are then passed through a Nyquist pulse shaping filter and up-converted to the desired carrier frequency on the frequency grid. During upconversion laser phase noise is introduced, which is modeled by a Wiener process. The signal is then combined with the other channels, and sent on N_{span} spans of optical fiber. In order to compensate for attenuation, IDRA is employed [1]. At the receiver, the WDM channel is down-converted to baseband, while introducing laser phase noise identically distributed to the transmitter's. The interfering channels are filtered out, and the desired channel is sent for baseband processing. This includes digital back-propagation (DBP) of the channel of interest only, in order to remove SPM, and the subsequent phase noise tracking algorithm. The signal after DBP is denoted as y. We are interested in the MI between x and y

$$\mathcal{I}(X;Y) = \mathcal{H}(X) - \mathcal{H}(X|Y) = \mathcal{H}(X) + \lim_{K\to\infty} \frac{1}{K} \log_2(p(x_1^K|y_1^K)), \quad (1)$$

where x_1^K and y_1^K denote the input and output sequences from time 1 to K, respectively, and we have used the convergence properties of the entropy function for long sequences [21]. Evaluating the joint and/or conditional probability of the entire sequences is an exponentially complex problem, and so a typical receiver will usually approximate $p(x_1^K|y_1^K) \approx \prod_k p(x_k|y_1^K)$, or even $p(x_1^K|y_1^K) \approx \prod_k p(x_k|y_k)$. When such a mismatched receiver is employed, the result is an upper bound on the entropy $\mathcal{H}(X|Y)$ [21], that gives an achievable information rate (AIR), which is a lower bound on the MI rate in (1).

A. Simplified Channel Model

In order to design a phase noise tracking algorithm, we employ the following simplified model of a standard phase noise channel

$$y_k = x_k \exp(j(\theta_k^{tx} + \theta_k^{rx} + \theta_k^{NL})) + w_k, \quad (2)$$

where $j = \sqrt{-1}$ is the imaginary unit. The phase noise contribution from the transmitter, receiver and the non-linearity (θ^{tx}, θ^{rx} and θ^{NL}, respectively) are all modeled by a Wiener process, e.g., for the transmitter term

$$\theta_k^{tx} = \theta_{k-1}^{tx} + \Delta_{tx} v_k, \quad (3)$$

where the v_k's are standard i.i.d. Gaussian variables, and Δ_{tx}^2 is the process noise variance. If we assume independent sources, they can be combined into a single phase noise process θ with parameter $\Delta^2 = \Delta_{tx}^2 + \Delta_{rx}^2 + \Delta_{NL}^2$, and the channel model becomes

$$y_k = x_k \exp(j\theta_k) + w_k. \quad (4)$$

The term w_k is a sample of additive white Gaussian noise (AWGN) with zero mean and variance assumed to be known at the receiver. When the phase noise is generated by a local oscillator (LO) with a certain spectral width around the central frequency (in case of lasers the more popular term is linewidth), the parameter Δ is found as

$$\Delta^2 = 2\pi f_W T_s, \quad (5)$$

where f_W is the width of the spectrum in Hz at -3dB of the maximum value (also known as full-width half-maximum bandwidth), and T_s is the sampling time in seconds.

B. The Non-Linear Term

In order to validate the Wiener model for the NLPN, we simulate an IDRA link, and examine the power spectral density (PSD) of the phase noise process. For demonstration, we choose a link with 5 channels at 100 GBaud each, of total length 4000km (the other fiber parameters are standard, and are listed in Table II), *without* transmitter and receiver phase noise. The input constellation in this case is 256QAM and the input power is -4dBm, which we found to be optimal at this distance. As shown in [2] [22], the NLPN is highly correlated within a window of several tens of symbols, which can be exploited in order to estimate the actual phase noise samples as

$$\hat{\theta}_k = \angle \sum_{l=k-L/2}^{k+L/2} y_l x_l^*, \quad (6)$$

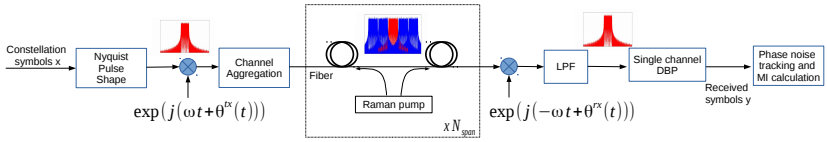


Fig. 1. Optical fiber channel model. Fiber parameters are defined in Section IV.

where $L+1$ is the window size, $(\cdot)^*$ denotes complex conjugate, and $\angle(\cdot)$ denotes the angle of a complex number. We then compare the estimated PSD of $\{\hat{\theta}_k\}$ to the theoretical PSD of a Wiener process given by the Lorentzian function with process noise variance

$$\Delta^2 = \mathrm{E}_k\left[(\hat{\theta}_k - \hat{\theta}_{k-1})^2\right]. \tag{7}$$

Example PSDs are given in Fig. 2 for different values of L. The PSD of $\{\hat{\theta}_k\}$ is obtained by Welch's method [23] with 10^6 samples, which were divided into 1999 blocks, 1000 samples each, with 50% overlap. Depending on the window size reasonable match can be obtained to the theoretical model. Long window results in relatively smooth PSD due to the better estimation in presence of noise. This would translate to better modeling of $\{\theta_k\}$ by the Wiener process. However, it leads to underestimation of Δ^2_{NL}. Decreasing L results in stronger oscillations at high frequency, and a large bias in the low frequency. In the rest of the paper the window size for estimating the samples θ_k, and thereby Δ^2, is chosen to be $L=20$, which as we see in Fig. 2 is a reasonable compromise between modeling accuracy and bias.

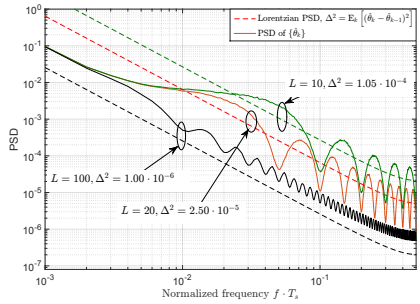


Fig. 2. Lorentzian PSD of a Wiener process (dashed lines) with process noise variance as in (7), together with the PSD of the phase noise $\{\hat{\theta}_k\}$ (solid lines), for the link given in Section II-B. Depending on the choice of the window L a good match can be found to the theoretical model.

III. Phase Noise Tracking

As mentioned in Section II, in order to compute AIRs, we need to compute the posterior probability of the input sequence $p(x_1^K|y_1^K)$, which we approximate as $\prod_k p(x_k|y_1^K)$, resulting in a lower bound on the MI rate. In this section we propose an efficient algorithm for calculating the posterior distributions at each time recursively.

Marginalizing the phase noise at time k, the posterior can be re-written as

$$\begin{aligned} p(x_k|y_1^K) &= \int_{-\pi}^{\pi} p(x_k,\theta_k|y_1^K)d\theta_k \\ &= \int_{-\pi}^{\pi} p(x_k|\theta_k,y_1^K)p(\theta_k|y_1^K)d\theta_k \\ &= \int_{-\pi}^{\pi} \frac{p(x_k)p(y_k|x_k,\theta_k)}{p(y_k|\theta_k)}p(\theta_k|y_1^K)d\theta_k \qquad (8) \\ &\propto \int_{-\pi}^{\pi} \frac{p(x_k)p(y_k|x_k,\theta_k)}{p(y_k|\theta_k)}p(y_k|\theta_k)p(y_{k+1}^K|\theta_k)p(\theta_k|y_1^{k-1})d\theta_k \qquad (9) \\ &= \int_{-\pi}^{\pi} p(x_k)p(y_k|x_k,\theta_k)p(y_{k+1}^K|\theta_k)p(\theta_k|y_1^{k-1})d\theta_k. \qquad (10) \end{aligned}$$

To get to (8) we used the fact that given the state θ_k, the input samples are independent of the past and future: $p(x_k|\theta_k,y_1^K) = p(x_k|\theta_k,y_k)$. To get to (10) we have used the fact that output samples are independent of the past given the phase: $p(y_{k+1}^K|\theta_k,y_1^{k-1}) = p(y_{k+1}^K|\theta_k)$, and we have removed the factor $p(y_{k+1}^K|y_1^{k-1})$ that does not depend on x_k and θ_k. The first and second term under the integral in Eq. (10) are the prior distribution of the constellation symbols and the likelihood of the output at time k, respectively. In order to derive expressions for the last two terms, we first define forward and backward recursions for the posterior distributions $p(\theta_k|y_1^k)$ and $p(y_k^K|\theta_k)$, which we model by mixtures of M and N Tikhonov distributions, respectively

$$p(\theta_k|y_1^k) = \sum_{m=1}^{M} \alpha_{m,k}\, t(w_{m,k};\theta_k), \tag{11}$$

$$p(y_k^K|\theta_k) = \sum_{n=1}^{N} \beta_{n,k}\, t(u_{n,k};\theta_k). \tag{12}$$

The terms $\alpha_{m,k}$ and $\beta_{n,k}$ are non-negative mixing coefficients, and are such that $\sum_m \alpha_{m,k} = 1$ and $\sum_n \beta_{n,k} = 1$. The Tikhonov distribution at θ with a complex parameter w is defined as

$$t(w;\theta) = \frac{\exp(\mathrm{Re}\,[w\cdot\exp(-j\theta)])}{2\pi I_0(|w|)}, \qquad \theta\in[-\pi;\pi), \tag{13}$$

and 0 elsewhere. In (13), I_0 is the zero-th order modified Bessel function of the first kind and $\mathrm{Re}\,[z]$ is the real part of z. We now derive mixture parameters for the predictive backward recursion term that appears in Eq. (10)

$$\begin{aligned} p(y_{k+1}^K|\theta_k) &= \sum_{n=1}^{N} \bar{\beta}_{n,k} t(\bar{u}_{n,k};\theta_k) \\ &= \int_{-\pi}^{\pi} p(y_{k+1}^K|\theta_{k+1},\theta_k) p(\theta_{k+1}|\theta_k) d\theta_{k+1} \\ &= \int_{-\pi}^{\pi} p(y_{k+1}^K|\theta_{k+1}) g(\theta_k,\Delta;\theta_{k+1}) d\theta_{k+1} \\ &= \sum_{n=1}^{N} \beta_{n,k+1} \int_{-\pi}^{\pi} t(u_{n,k+1};\theta_{k+1}) g(\theta_{k+1},\Delta;\theta_k) d\theta_{k+1}, \end{aligned} \tag{14}$$

where $g(\mu,\sigma;z)$ is the Gaussian probability density function (PDF) at z with mean μ and standard deviation σ. The update parameters can then be found as

$$\bar{\beta}_{n,k} = \beta_{n,k+1}, \qquad \bar{u}_{n,k} = \frac{u_{n,k+1}}{1+\Delta^2|u_{n,k+1}|}, \tag{15}$$

where we have used the fact that the convolution of a Gaussian and Tikhonov distributions is a Tikhonov with a modified complex parameter [20]. In order to complete the recursion, the updates for $\beta_{n,k}$ and $u_{n,k}$ are found from the following:

$$\begin{aligned} p(y_k^K|\theta_k) &= p(y_k|\theta_k, y_{k+1}^K) p(y_{k+1}^K|\theta_k) \\ &= \sum_{n=1}^{N} \bar{\beta}_{n,k} \sum_{x_k\in\mathcal{X}} p(x_k) p(y_k|x_k,\theta_k) t(\bar{u}_{n,k};\theta_k) \\ &\propto \sum_{n=1}^{N} \sum_{x_k\in\mathcal{X}} \mu_{n,k}(x_k) t(\bar{u}_{n,k} + 2\cdot\mathrm{SNR}\cdot y_k x_k^*;\theta_k), \end{aligned} \tag{16}$$

where the likelihood $p(y_k|x_k,\theta_k)$ is expressed as a Tikhonov approximation to the Gaussian

$$p(y_k|x_k,\theta_k) \approx \frac{2\cdot\mathrm{SNR}\cdot I_0(2\cdot\mathrm{SNR}|y_k x_k^*|) t(2\cdot\mathrm{SNR}\cdot y_k x_k^*;\theta_k)}{\exp(\mathrm{SNR}(|y_k|^2+|x_k|^2))}. \tag{17}$$

In (16) we have used the fact that the product of two Tikhonov distributions may also be expressed as a Tikhonov distribution in order to calculate the sub-component mixture coefficient

$$\mu_{n,k}(x_k) = \frac{\bar{\beta}_{n,k}\cdot p(x_k) I_0(|\bar{u}_{n,k} + 2\cdot\mathrm{SNR}\cdot y_k x_k^*|)}{I_0(|\bar{u}_{n,k}|)\exp(\mathrm{SNR}\cdot|x_k|^2)}. \tag{18}$$

Due to the discrete nature of the input signal, the number of components needed for tracking the phase noise grows exponentially with time. In order to avoid this problem, we propose an approximation to the inner sum in (16), where at each step we only take the sub-component with the largest mixing coefficient

$$\hat{x}_{n,k} = \arg\max_{x_k\in\mathcal{X}} \mu_{n,k}(x_k), \tag{19}$$

$$u_{n,k} = \bar{u}_{n,k} + 2\cdot\mathrm{SNR}\cdot y_k \hat{x}_{n,k}^*, \tag{20}$$

$$\beta_{n,k} = B\cdot\mu_{n,k}(\hat{x}_{n,k}), \tag{21}$$

where B is such that $\sum_{n=1}^{N}\beta_{n,k} = 1$.

Similarly we express the predictive forward distribution appearing in (10), as

$$p(\theta_k|y_1^{k-1}) = \sum_{m=1}^{M} \bar{\alpha}_{m,k} t(\bar{w}_{m,k};\theta_k), \tag{22}$$

where

$$\bar{\alpha}_{m,k} = \alpha_{m,k-1}, \qquad \bar{w}_{m,k} = \frac{w_{m,k-1}}{1+\Delta^2|w_{m,k-1}|}. \tag{23}$$

The update parameters are found as

$$\rho_{m,k}(x_k) = \frac{\bar{\alpha}_{m,k}\cdot p(x_k) I_0(|\bar{w}_{m,k} + 2\cdot\mathrm{SNR}\cdot y_k x_k^*|)}{I_0(|\bar{w}_{m,k}|)\exp(\mathrm{SNR}\cdot|x_k|^2)}, \tag{24}$$

$$\hat{x}_{m,k} = \arg\max_{x_k\in\mathcal{X}} \rho_{m,k}(x_k), \tag{25}$$

$$w_{m,k} = \bar{w}_{m,k} + 2\cdot\mathrm{SNR}\cdot y_k \hat{x}_{m,k}^*, \tag{26}$$

$$\alpha_{m,k} = A\cdot\rho_{m,k}(\hat{x}_{m,k}), \tag{27}$$

where A is such that $\sum_{m=1}^{M}\alpha_{m,k} = 1$. We are now ready to express the posterior distribution (10) as

$$\begin{aligned} p(x_k|y_1^K) = p(x_k) \sum_{m=1}^{M} \bar{\alpha}_{m,k} \sum_{n=1}^{N} \bar{\beta}_{n,k} \times \\ \int_{-\pi}^{\pi} p(y_k|x_k,\theta_k) t(\bar{u}_{n,k};\theta_k) t(\bar{w}_{m,k};\theta_k) d\theta_k. \end{aligned} \tag{28}$$

Using the expression for the likelihood (17), the integrand in (28) becomes a product of three Tikhonov distributions in θ_k, which is again a scaled Tikhonov of θ_k

$$\begin{aligned} p(y_k|x_k,\theta_k) t(\bar{u}_{n,k};\theta_k) t(\bar{w}_{m,k};\theta_k) \propto \\ \frac{I_0(|\bar{w}_{m,k} + \bar{u}_{n,k} + 2\cdot\mathrm{SNR}\cdot y_k x_k^*|)}{I_0(|\bar{w}_{m,k}|) I_0(|\bar{u}_{n,k}|)\exp(\mathrm{SNR}(|y_k|^2+|x_k|^2))} \times \\ t(\bar{w}_{m,k} + \bar{u}_{n,k} + 2\cdot\mathrm{SNR}\cdot y_k x_k^*;\theta_k), \end{aligned} \tag{29}$$

where we have removed the proportionality factors, independent of θ_k and x_k. The scaling factor goes out of the integral in (28), and the remaining Tikhonov distribution integrates to one. The expression for the posterior is then

$$\begin{aligned} p(x_k|y_1^K) = p(x_k) \sum_{m=1}^{M} \bar{\alpha}_{m,k} \sum_{n=1}^{N} \bar{\beta}_{n,k} \times \\ \frac{I_0(|\bar{w}_{m,k} + \bar{u}_{n,k} + 2\cdot\mathrm{SNR}\cdot y_k x_k^*|)}{I_0(|\bar{w}_{m,k}|) I_0(|\bar{u}_{n,k}|)\exp(\mathrm{SNR}(|y_k|^2+|x_k|^2))}. \end{aligned} \tag{30}$$

The expression (30) is a sum of $N\cdot M$ elements, which is very efficiently calculated in the log domain, using the large value approximation of the modified Bessel function

$$I_0(w) \approx \frac{\exp(w)}{\sqrt{2\pi w}}. \tag{31}$$

Typical values of $|\bar{w}|$ and $|\bar{u}|$ are above 200, for which the approximation in (31) is accurate.

A. Summary

The complete algorithm is summarized in the following steps:

1) **Initialize**:

$$\alpha_{m,0} = 1/M, \qquad \beta_{n,K+1} = 1/N,$$
$$w_{m,0} = \frac{1}{\Delta^2} e^{j(m\cdot\frac{2\pi}{M}-\pi)}, \qquad u_{n,K+1} = \frac{1}{\Delta^2} e^{j(n\cdot\frac{2\pi}{N}-\pi)} \tag{32}$$

2) **Forward recursion** - equations (23),(25)-(27)
3) **Backward recursion** - equations (15),(19)-(21)
4) **Posteriors calculation** - Eq. (30)

The initialization values are chosen such that each component corresponds to a Gaussian distribution of the phase with variance Δ^2, and the means of the components are uniformly spaced within $[-\pi;\pi)$.

B. Phase Noise Distribution

The phase noise PDF at each time is not explicitly calculated by the algorithm, but can be found, if needed for further processing, as

$$p(\theta_k|y_1^k) \propto p(y_{k+1}^K|\theta_k)p(\theta_k|y_1^k) = \sum_{n=1}^{N} \bar{\beta}_{n,k}\, t(\bar{u}_{n,k};\theta_k) \sum_{m=1}^{M} \alpha_{m,k}\, t(w_{m,k};\theta_k). \tag{33}$$

An example distribution update is given in the surface plot in Fig. 3 for a standard Wiener phase noise channel, simulated via (4), with parameter $\Delta = 0.05$ and 256QAM input. The distribution is given as a surface plot, where the height of the surface is described by the color bar. We evaluate the phase noise distribution for all phases within $[-\pi;\pi)$, with a resolution of $2\pi/512$ radians. We also plot the actual phase noise realizations. In this case the number of mixture components in the forward and backward recursions is $M = N = 4$. We clearly see the different components, one of which tracks the true phase noise realization. Due to the unknown initial phase and the 4-fold symmetry of the constellation, there is phase ambiguity in multiples of $\pi/2$ radians. In order to combat this problem, we insert pilot symbols at rate P, which give an absolute reference for the phase. Pilot symbols are accounted for in the model by changing the PMF $p(x_k)$ at the pilot instants to an indicator function, which is '1' for the true sent symbol, and '0' otherwise. Even though this is a sub-optimal pilot design, it suffices for our further analysis. The resulting phase noise distribution after pilot insertion is given in Fig. 4.

C. Relation to Other Algorithms

When the PDFs $p(\theta_k|y_1^K)$, $p(\theta_k|y_1^k)$ and $p(y_k^K|\theta_k)$ are modeled by a Gaussian, the solution of the recursions is the Kalman filter [9]. The difference between our approach and the Kalman filter is visible in Eq. (28). If the densities are modeled as Gaussians, the likelihood $p(y_k|\theta_k, x_k)$, which is a non-linear function of the phase θ_k, needs to be linearized in order to make the integration simple, as done here. The linearization leads to sub-optimal performance of the Kalman filter. We note that the Tikhonov distribution may be seen as a Gaussian "wrapped" around $\pm\pi$. Example of the distribution with mean

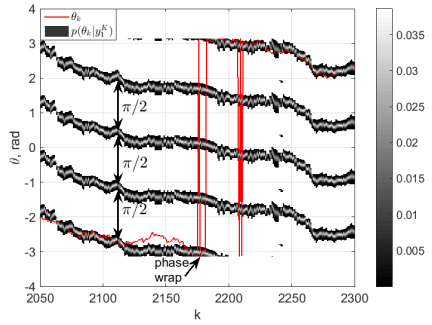


Fig. 3. Probability distribution of the phase noise at time k (given by the surface plot), together with actual phase noise samples, given by the red line (pale line on black and white printer). Brighter color represents higher probability, given by the side color bar.

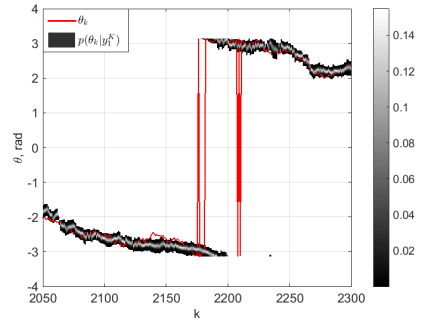


Fig. 4. Probability distribution of the phase noise at time k after insertion of $P = 0.2\%$ pilot symbols. Brighter color represents higher probability, given by the side color bar. Note the different scale w.r.t Fig. 3 due to the absence of the $\pi/2$ ambiguity.

$-\pi/4$ and different variances is given in Fig. 5, together with the Gaussian and a wrapped Gaussian, as defined in the legend. The latter distribution and the Tikhonov have support set $[-\pi;\pi)$. For small variance the three distributions coincide due to the exponentially vanishing tail of the Gaussian. For large variance we see the wrapping becoming significant and the distributions diverging from each other.

The trellis approach from [19] replaces the above distributions with histograms. This means that the integrations for the updates (Eq. (14) and its forward recursion analogue) and the posterior calculation (Eq. (28)) become sums. For large constellations and fine resolution of the phase, the complexity increases significantly.

The approach in [20] also employs Tikhonov distributions, however, only for PSK constellations. There, the information is

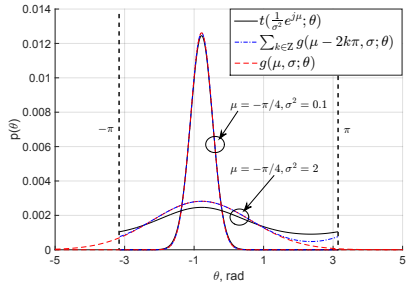


Fig. 5. Example Tikhonov vs. Gaussian and wrapped Gaussian distributions with the same parameters. For small variance the three distributions are almost identical.

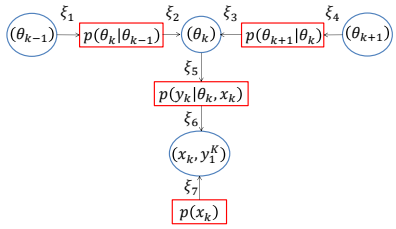


Fig. 6. One section of the graph, used for estimating the posterior distributions. Factors are given as rectangles, and variables with circles. The factors are represented by the corresponding distributions.

already included in the phase and therefore phase tracking and posterior calculation can be performed simultaneously. The focus here is on QAM constellations.

We note that all of the above approaches, including the one proposed in this paper, in principle employ the sum-product rules for estimating densities in graphs. In this particular case, the graph represents a first-order hidden Markov chain. One section of the graph for this algorithm is given in Fig. 6. The variable nodes are given as circles, and the factor nodes as rectangles. The direction of the messages ξ is given by the arrows, and the messages themselves are calculated as:

- $\xi_1 = p(\theta_{k-1}|y_1^{k-1})$
- ξ_2 - Eq. (22)
- ξ_3 - Eq. (14)
- $\xi_4 = p(y_{k+1}^K|\theta_{k+1})$
- $\xi_5 = \xi_2 \cdot \xi_3$
- $\xi_6 = \int_{-\pi}^{\pi} \xi_5 \cdot p(y_k|\theta_k, x_k) d\theta_k$
- $\xi_7 = p(x_k)$

According to the sum-product rule, the unnormalized distribution at each variable is the product of all incoming messages. Then we find $p(x_k|y_1^K) \propto p(x_k, y_1^K) = \xi_6 \cdot \xi_7$, which is exactly the expression in Eq. (10).

D. Complexity

In the following we refer to the proposed algorithm as Tikhonov mixture model (TMM) algorithm.

The complete algorithm can be cast into the log-domain with standard max-log approximations. Thus the complexity is dominated by the computation of Euclidean distances (ED), needed for calculating the largest mixing coefficients in (19) and (25). There, the complexity is linear in the constellation size and the number of mixing coefficients. The forward and backward recursions require one max-log operation at each time across M and N elements (for the normalization in (21) and (27), respectively), in order to calculate the updates in (15) and (23). The calculation of the posterior requires one max-log across $N \cdot M$ elements, bringing the total number of max-log operations to $\approx \mathcal{O}(K \cdot (M \cdot N + M + N))$. Typical values of M and N are as small as 2 or 4, so we conclude that the complexity is still dominated by the EDs calculation, which requires complex multiplication. We compare these numbers to the DD algorithm [10], where a cost function of the EDs to the closest constellation symbol is calculated for N_p candidate phases, and the minimum is taken as the phase rotation. The number of EDs (complex multiplications) needed is therefore $\mathcal{O}(K \cdot N_p \cdot |\mathcal{X}|)$. Typical number of test phases (taken directly from [10]) is $N_p \approx 16$ for $|\mathcal{X}| \leq 16$ and $N_p \approx 64$ for $|\mathcal{X}| \leq 256$. Our method requires $\mathcal{O}(K \cdot (M+N) \cdot |\mathcal{X}|)$, which for 256QAM and $M = N = 2$ Tikhonov components is more than an order smaller.

One issue with the sequential processing of our algorithm is the latency. We note that even though very long sequences are needed for the convergence $\lim_{K\to\infty} -\frac{1}{K}\log_2 p(x_1^K|y_1^K) = \mathcal{H}(X|Y)$, a real receiver does not aim at computing the entropy. The value of K may therefore be kept at a reasonable value, while keeping the posteriors in (30) accurate.

IV. Results

We examine the performance of the TMM algorithm mainly in terms of AIR in *bits/channel use*. In case of a standard Wiener phase noise channel 1 bit/channel use means 1 bit/s/Hz/complex dimension. In case of the WDM link, 1 bit/channel use means 1 bit/s/Hz/polarization/complex dimension. The input power in the latter case is defined as input power per channel. In all cases, the AIR is estimated from a block of length 10^5 symbols.

A. Standard Wiener Phase Noise Channel

We start by analyzing a standard Wiener phase noise channel, which is simulated via (4). In Fig. 7, the AIRs are given for a channel with $f_W \cdot T_s = 8 \cdot 10^{-5}$ and 256QAM input. This is beyond the capabilities of the DD algorithm [10], and our simulations also confirm that the DD algorithm fails with such parameters, giving lower bounds on the MI well below zero. We examine the performance of the TMM algorithm with $M = N = 2$ and 4 mixture components, and pilot rates of $P = 0.005, 0.01, 0.02$, and 0.05. For reference, we also plot the AIR on an AWGN channel without phase noise, and what is achieved by the trellis algorithm from [19] with pilot rate $P = 0.002$ and 128 states. The rate loss due to the pilot

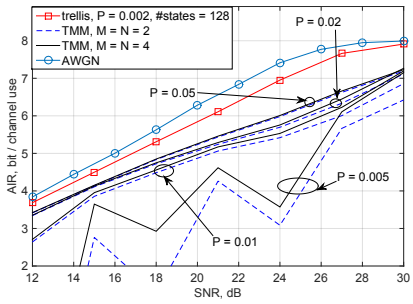


Fig. 7. Comparison of pilot rates for the TMM algorithm. From bottom curve to top: $P = 0.005, 0.01, 0.02, 0.05$. Standard Wiener phase noise channel, $f_W \cdot T_s = 8 \cdot 10^{-5}$, 256QAM input. For small pilot rates the algorithm is unstable.

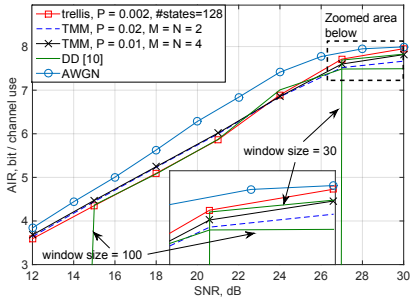


Fig. 8. Comparison with the DD algorithm. The trellis and the TMM algorithm have optimized pilot rates, given in the legend. The DD algorithm is tested with sliding window length of 30 and 100 symbols. Inset: very high rate (low BER, respectively). Standard Wiener phase noise channel, $f_W \cdot T_s = 8 \cdot 10^{-6}$, 256QAM input.

symbols is taken into account in the AIR calculation by scaling the entropy $\mathcal{H}(X)$ in (1) by $1 - P$. For such large values of Δ^2 and insufficient pilot rate the algorithm is unstable, and often switches to adjacent $\pi/2$ components. This leads to very poor lower bounds on the MI, sometimes even negative. At pilot rate above 1% the algorithm is stable, and we see only marginal improvement going from 2 to 4 mixture components. The performance is sub-optimal at high rates, however a stable behavior is observed over the entire SNR range.

In Fig. 8 we show the performance at a smaller value $f_W \cdot T_s = 8 \cdot 10^{-6}$. In this case the algorithm from [10] is able to operate with minor SNR penalty at BER $= 10^{-3}$ with an averaging window of size around 30. This window represents the length over which test phases are evaluated, and should not be confused with the window L we used in Section II-B for estimating Δ^2. As seen from Fig. 8, the achievable rate with

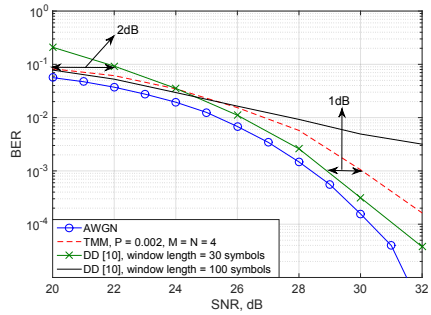


Fig. 9. BER at high SNR for the TMM algorithm and the DD algorithm [10] with 256QAM input, standard Wiener phase noise channel, $f_W \cdot T_s = 8 \cdot 10^{-6}$. For the DD algorithm, at low SNR short windows result in high BER, whereas at high SNR long windows result in high error floor. The TMM algorithm performs well for the entire region.

this window size is close to the maximum $\mathcal{H}(X) = \log_2 |\mathcal{X}|$, which corresponds to the above mentioned low BER and SER. However, when the SNR is reduced, the algorithm fails due to the higher SER. Increasing the window length helps at the lower SNRs, and at 100 symbols a stable behavior is observed down to SNR$=$ 15dB. Long averaging windows, however, result in a sub-optimal performance at high SNR, where the AIR appears to achieve a maximum value, smaller than $\mathcal{H}(X)$. This is because small-scale, fast variations of the phase noise are not captured by the long window averaging. This would correspond to an error floor after de-mapping. The TMM algorithm on the other hand achieves stable performance, close to the reference trellis based AIR at significantly reduced complexity, for the entire SNR region.

To further illustrate the problem with the error floor, we analyze the BER at high SNR. On Fig. 9, the BER is given at $f_W \cdot T_s = 8 \cdot 10^{-6}$. We assume Gray labeling of the symbols. The algorithm from [10] requires differential coded modulation, which technically will increase the BER slightly. Therefore we can argue that the comparison is fair, with a slight advantage given to the DD algorithm. At high SNR, the TMM algorithm experiences up to 1dB loss compared to the DD algorithm with averaging window of length 30 symbols. However, at low SNR (BER $\approx 10^{-1}$), such window length results in around 2dB loss compared to the TMM algorithm. Increasing the window length helps, but results in very high BER at high SNR. For completeness, in Table I the SNR penalty to the AWGN channel performance at BER$=10^{-3}$ is given for several combinations of modulation format and LLW. The window length of the DD algorithm is 30, and we simulated SNRs up to 40dB. The TMM algorithm employs 4 components and $P = 0.002$ pilot rate. For high values of $f_W \cdot T_s$ the DD algorithm completely fails with this window length, while reception with the TMM algorithm is still possible. As mentioned above, decreasing the window size of the DD algorithm might improve the penalty and allow for

TABLE I
SNR PENALTY @ BER=10^{-3}. INFINITE PENALTY MEANS THAT THE REQUIRED BER IS NOT ACHIEVED.

$f_W \cdot T_s$	$5 \cdot 10^{-6}$	$1 \cdot 10^{-5}$	$5 \cdot 10^{-5}$	$1 \cdot 10^{-4}$	$5 \cdot 10^{-4}$
256QAM, TMM	1.1dB	1.3dB	4.9dB	7.8dB	∞
256QAM, DD [10]	0.6dB	0.9dB	∞	∞	∞
64QAM, TMM	1.0dB	1.2dB	2.2dB	4.4dB	9.8dB
64QAM, DD [10]	0.5dB	0.7dB	1.7dB	∞	∞
16QAM, TMM	0.9dB	1.1dB	1.2dB	1.3dB	3.1dB
16QAM, DD [10]	0.1dB	0.2dB	0.3dB	0.9dB	∞

transmission, but will result in increased penalty at lower SNR (higher BER, respectively).

B. IDRA WDM Optical Fiber Link

Next we evaluate the TMM algorithm in a WDM optical link. The link is simulated using the split-step Fourier method (SSFM). The fiber and transceiver parameters are given in Table II. Single polarization transmission is used and single-channel digital back-propagation is performed on the central channel, which is the channel of interest.

In Fig. 10, the AIRs are given as a function of the launch power per channel. For reference, we calculate the AIR with a pseudo-ideal phase noise removal (PIPNR). This is achieved by pre-processing the output samples as $\dot{y}_k = y_k e^{-j\dot{\theta}_k}$. The estimates of the phase noise $\dot{\theta}_k$ are obtained similar to (6), but from a window of past samples only, i.e. $\dot{\theta}_k = \angle \sum_{l=k-L-1}^{k-1} y_l x_l^*$. This is done in order for the calculated values to be AIRs. To better understand this, observe that the probability $p(x_1^K|y_1^K)$ may be expressed from the product rule as $p(x_1^K|y_1^K) \propto \prod_k p(x_k|y_1^K, x_1^{k-1})$, which means that if lower bounds on the MI are targeted, only past samples may be used to obtain mismatched probability distribution.

In Fig. 10 we also plot the AIR in the idealized case of no laser phase noise without any processing. In this case we assume memoryless channel and model the likelihoods $p(y_k|x_k)$ as Gaussian distributions with known mean and variance. To put the values of the LLW in the perspective of Section IV-A, we estimated the value of Δ_{NL}^2 at the optimal input power in the idealized case of $f_W = 0$ to be $\Delta_{NL}^2 \approx 2.5 \cdot 10^{-5}$. The term $f_W \cdot T_s$ then becomes $\approx 3.9 \cdot 10^{-6}, 4.2 \cdot 10^{-6}$ and $6 \cdot 10^{-6}$ in the 3 cases of LLW, respectively. The window size for the DD algorithm is optimized to 500 (we note that the AIR at the optimal input power increases very slowly for window sizes from 200 to 500, and then starts to decrease). The pilot rate for the trellis and the TMM algorithm is fixed to a minimal value of $P = 0.002$. The TMM algorithm has $M = N = 4$ mixture components. We see that the performance of the TMM algorithm is close to the pseudo-ideal one. The AIR in the idealized case without processing is below what is achieved in the 10kHz case when phase noise tracking is performed. This is due to the non-zero value of Δ_{NL}. The consequence of this observation is two-fold:

1) Even if ideal lasers are used, correlations in the NLPN can still be exploited by tracking it in order to improve the performance.
2) The algorithm for tracking should be able to account for the combined effect of laser phase noise and NLPN.

The latter observation suggests that even when perfect knowledge is available for the lasers, the combined phase noise process variance should always be estimated as $\Delta^2 = \Delta_{tx}^2 + \Delta_{rx}^2 + \Delta_{NL}^2$ for the desired link set-up. In case the lasers are not ideally characterized, the process noise variance may be estimated from training data via (7). The latter approach is also more robust to instabilities.

TABLE II
SYSTEM PARAMETERS, IDRA TRANSMISSION

Span length	100 km
Symbol rate	100 GBaud
Number of channels	5
Guardband	2 GHz (2% of symbol rate)
Total bandwidth	510 GHz
Oversampling factor	32
Pulse shape	$sinc$
LLW	0 kHz, 10 kHz and 100 kHz
Fiber loss	$\alpha = 0.2$ dB/km
Non-linear coefficient	$\gamma = 1.3$ (W·km)$^{-1}$
Dispersion	$D = 17$ ps/(nm·km)
Central wavelenght	$\lambda_0 = 1.55$ μm
SSFM step	0.1 km

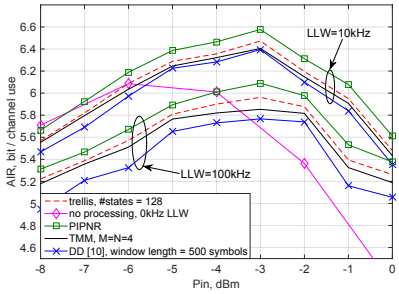


Fig. 10. AIRs after 40 spans, 256QAM, at different LLWs at transmitter and receiver, 100GBaud.

C. EDFA WDM Optical Fiber Link

In order to assess the performance of the TMM algorithm in a more practical scenario, we consider a lumped amplification scheme, where Erbium doped fiber amplifiers (EDFA) are inserted at the end of each fiber span, instead of the Raman pump. The new system parameters are given in Table III. The fiber parameters are the same as in Table II.

In this case we employ both polarizations. DBP is not performed here, but only electronic chromatic dispersion compensation in the frequency domain. Polarization mode dispersion is neglected in the SSFM.

TABLE III
SYSTEM PARAMETERS, EDFA TRANSMISSION

Symbol rate	28 GBaud
Number of channels	17
Guardband	0.56 GHz (2% of symbol rate)
EDFA noise figure	4 dB
LLW	10 kHz, 100 kHz and 1 MHz

We also study the performance of the algorithm for smaller constellations, where the LLW tolerance is generally higher. AIRs are given for 256QAM, 64QAM and 16QAM, at 10 spans, 30 spans and 50 spans, respectively, in Fig. 11 for 10kHz LLW, Fig. 12 for 100kHz LLW and Fig. 13 for 1MHz LLW. The distances and constellation sizes are chosen such that the maximum AIR at the optimal input power is smaller than $\approx \frac{3}{4}\log_2|\mathcal{X}|$, which is the desired operating point for energy and spectral efficient communications. This point can be seen as the maximum SNR, at which the slope of the MI with discrete input on a Gaussian channel is the same as for Gaussian input, i.e., where the AIR is not yet limited by the size of the constellation. For such information rates and sufficiently large values of Δ, DD methods generally perform poorly due to the high SER, regardless of the modulation format.

At 10kHz LLW the TMM algorithm achieves near-optimal performance in all cases up to the optimal input power for the respective distances and constellations. We only see an instability for very high input power and low SNR in the case of 16 QAM after 50 spans. Similar to the IDRA case, the DD algorithm from [10] requires around 500 samples for averaging out the noise. At 100kHz LLW, the phase noise cannot be considered constant for such a long period, and the DD algorithm fails. The TMM method requires increase in the pilot rate to $P = 0.005$, and achieves stable performance, close to that of the trellis method. We note that the number of states in the trellis was increased to 256, and we see that in the case of 256 QAM there is still around 0.5 bits/channel use gap to the PIPNR rate. For smaller constellations 256 states are enough to see convergence in the performance. At 1MHz LLW the TMM method becomes sub-optimal, however, still achieves reasonable performance for 16 QAM and 64 QAM. The pilot rate in this case is increased to $P = 0.05$. The gap to the trellis method is around 0.4 and 0.2 bits/channel use for 64 QAM and 16 QAM, respectively.

V. DISCUSSION AND FUTURE WORK

As shown earlier, the trellis method from [19] achieves near-PIPNR performance in nearly all cases of interest, however, it is very complex. This is particularly the case for large values of Δ, where the entire range $[-\pi;\pi)$ must be covered with very fine precision. Recently, the authors in [24] studied this problem, and proposed a low-complexity solution, which basically reduces the state space and thus the complexity of the algorithm. A comparison in terms of complexity and performance between the TMM and the solution from [24] would be of interest, but out of the scope of this paper.

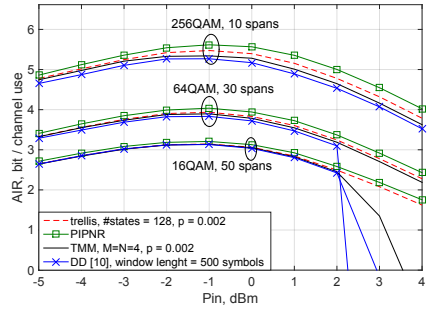


Fig. 11. AIRs with 10kHz LLW for different modulation formats at different distances. The DD method requires long averaging window. The TMM closely approaches the trellis method in all cases.

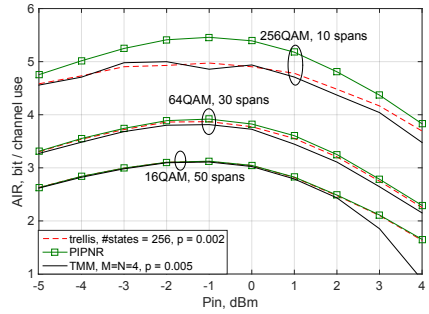


Fig. 12. AIRs with 100kHz LLW for different modulation formats at different distances. The DD method cannot be used. The trellis method requires more states than before. The TMM requires a slight increase in the pilot rate, but still achieves a performance close to that of the trellis method in all cases.

As mentioned in Section II-B, the choice of the window length L for estimating Δ^2 may improve the quality of the Wiener process model for the NLPN. We found that the AIRs in Section IV vary only slightly for $L \in [10;100]$, which means that the proposed method is generally robust to variations and instabilities in the estimation of Δ. We note that Eq. (6) may be seen as a convolution of the signal $y_k \cdot x_k^*$ with a rectangular pulse. Optimizing the pulse shape may further improve the quality of the model.

A remark on the pilot symbols assumptions in our design follows. In cases of very narrow linewidth, we have confirmed that if the initial $\pi/2$ ambiguity is avoided, the algorithm is stable in tracking the phase noise and does not require subsequent pilots. Even though the $P = 0.002$ is negligible in terms of reduced spectral efficiency, if the initial phase noise value is known, the pilots may be entirely removed.

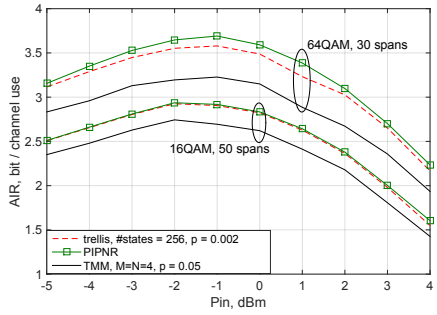


Fig. 13. AIRs with 1MHz LLW for different modulation formats at different distances. The DD method cannot be used. The TMM method requires an increased pilot rate, and achieves stable performance, with a loss compared to the trellis method of 0.4 and 0.2 bits/channel use for 64 QAM and 16 QAM, respectively.

This is a reasonable assumption when e.g. previous blocks are decoded correctly, and such information can be extracted. We note that due to the increasing order of modulation in optical fiber systems, inserting pilot symbols is becoming a more and more popular approach in the research community due to the improved equalization they provide. The same pilots that are used for equalization may generally be used for phase noise tracking.

VI. Conclusion

An algorithm was proposed for tracking the phase noise in wavelength division multiplexed optical fiber channels. It was shown that in ideal distributed raman amplification (IDRA) links, the proposed method can effectively combat the combined effect of laser phase noise and non-linear phase noise, outperforming previous decision directed methods, at significantly reduced complexity compared to previous trellis methods. Near optimal performance can be achieved for IDRA links, but also in more practical lumped amplification links with dual polarization input. Depending on the severity of the phase noise, pilot symbols may be introduced, which allow for stable performance on a wide variety of SNRs, achievable information rates and laser linewidths.

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] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Optics Express*, vol. 21, no. 22, pp. 25685-25699, Nov. 4, 2013.

[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] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Inter-channel nonlinear interference noise in WDM systems: modeling and mitigation," *Journal of Lightwave Technology*, vol. 33, no. 5, pp. 1044-1053, Mar. 1, 2015.

[5] M. Secondini and E. Forestieri, "On XPM mitigation in WDM fiber-optic systems," *Photonics Technology Letters*, vol. 26, no. 22, pp. 2252-2255, Nov. 15, 2014.

[6] D. Marsella, M. Secondini, E. Agrell, and E. Forestieri, "A simple strategy for mitigating XPM in nonlinear WDM optical systems," in Proceedings of Optical Fiber Communication Conference, Paper TH4D.3, 2015.

[7] M. Magarini, A. Spalvieri, F. Vacondio, M. Bertolini, M. Pepe, and G. Gavioli, "Empirical modeling and simulation of phase noise in long-haul coherent optical systems," *Optics Express*, vol. 19, no. 23, pp. 22455-22461, Nov. 7, 2011.

[8] I. Fatadin, D. Ives, and S. J. Savory, "Laser linewidth tolerance for 16-QAM coherent optical systems using QPSK partitioning," *Photonics Technology Letters*, vol. 22, no. 9, pp. 631-633, May. 1, 2010.

[9] L. Barletta, M. Magarini, and A. Spalvieri, "The information rate transferred through the discrete-time Wiener's phase noise channel," *Journal of Lightwave Technology*, vol. 30, no. 10, pp. 1480-1486, May 15, 2012.

[10] T. Pfau, S. Hoffman, and R. Noé, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *Journal of Lightwave Technology*, vol. 27, no. 8, pp. 989-999, Apr. 15, 2009.

[11] S. Zhang, P. Kam, C. Yu, and J. Chen, "Decision-aided carrier phase estimation for coherent optical communications," *Journal of Lightwave Technology*, vol. 28, no. 11, pp. 1597-1607, June 1, 2010.

[12] T. N. Huynh, A. T. Nguyen, W. Ng, L. Nguyen, L. A. Rusch, and L. P. Barry, "BER performance of coherent optical communications systems employing monolithic tunable lasers with excess phase noise," *Journal of Lightwave Technology*, vol. 32, no. 10, pp. 1973-1980, May 15, 2014.

[13] M. Magarini, L. Barletta, A. Spalvieri, F. Vacondio, T. Pfau, M. Pepe, M. Bertolini, and G. Gavioli, "Pilot-symbols-aided carrier-phase recovery for 100-G PM-QPSK digital coherent receivers," *Photonics Technology Letters*, vol. 24, no. 9, pp. 739-741, May 1, 2012.

[14] Y. Gao, A. P. T. Lau, and C. Lu, "Modulation-format-independent carrier phase estimation for square M-QAM systems," *Photonics Technology Letters*, vol. 25, no. 11, pp. 1073-1076, Jun. 1, 2013.

[15] C. Pan and F. R. Kschischang, "Coded-aided phase tracking for coherent fiber channels," *Journal of Lightwave Technology*, vol. 32, no. 6, pp. 1041-1047, Mar. 15, 2014.

[16] C. Pan, H. Bülow, W. Idler, L. Schmalen, and F. R. Kschischang, "Optical nonlinear-phase-noise compensation for 9x32 Gbaud polDM-16 QAM transmission using a code-aided expectation-maximization algorithm," *Journal of Lightwave Technology*, vol. 33, no. 17, pp. 3679-3686, Sep. 1, 2014.

[17] L. Barletta, M. Magarini, and A. Spalvieri, "A new lower bound below the information rate of Wiener phase noise channel based on Kalman carrier recovery," *Optics Express*, vol. 20, no. 23, pp. 25471-25477, Nov. 5, 2012.

[18] L. Barletta, M. Magarini, and A. Spalvieri, "Tight upper and lower bounds to the information rate of the phase noise channel," IEEE International Symposium on Information Theory (ISIT), pp. 2284-2288, July 7-12, 2013.

[19] L. Barletta, F. Bergamelli, M. Magarini, N. Carapellese, and A. Spalvieri, "Pilot-aided trellis-based demodulation," *Photonics Technology Letters*, vol. 25, no. 13, pp. 1234-1237, July 1, 2013.

[20] A. Barbieri and G. Colavolpe. "Soft-output decoding of rotationally invariant codes over channels with phase noise," *IEEE Transactions on Communications*, vol. 55, no. 11, pp. 2125-2133, Nov. 2007.

[21] D. M. Arnold, H. Loeliger, P. O. Vontobel, A. Kavčić, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Transactions on Information Theory*, vol. 52, no. 8, Aug. 2006.

[22] T. Fehenberger, M. P .Yankov, L. Barletta, and N. Hanik, "Compensation of XPM interference by blind tracking of the nonlinear phase in WDM systems with QAM input," Proceedings of European Conference on Optical Communications (ECOC), Paper P.5.8, Oct. 2015.

[23] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. AU-15, no. 2, June 1967.

[24] S. Pecorino, S. Mandelli, L. Barletta, M. Magarini, and A. Spalvieri, "Bootstrapping iterative demodulation and decoding without pilot symbols," *Journal of Lightwave Technology*, vol. 33, no. 17, pp. 3616-3622, Sep. 1, 2015.

ACHIEVABLE INFORMATION RATES ON LINEAR INTERFERENCE CHANNELS WITH DISCRETE INPUT

Metodi Yankov and Søren Forchhammer [1]

[1]Department of Photonics Engineering, Technical University of Denmark,
Bldng. 343, Ørsteds Plads, 2800 Kgs. Lyngby, Denmark
meya@fotonik.dtu.dk

ABSTRACT

In this paper lower bound on the capacity of multi-dimensional linear interference channels is derived, when the input is taken from a finite size alphabet. The bounds are based on the QR decomposition of the channel matrix, and hold for any input distribution that is independent across dimensions. Calculation of the bounds can be performed on a per-dimensions basis via look-up tables of the information rates of 1D channels.

1. INTRODUCTION

The capacity of a set of linearly interfering channels when the input is taken from a finite size alphabet has been a long standing problem in information theory. In the case of the Multiple Input Multiple Output (MIMO) channel with Gaussian input the capacity has been found [1]. When the transmitter has perfect knowledge of the channel, it can align the input to the channel eigen modes and allocate the power based on the water-filling strategy. When the channel is known at the receiver only, i.i.d. Gaussian input is optimal. It has been shown in [2] that when the input is discrete, both orthogonalization and water-filling power allocation are sub-optimal. Low and high SNR asymptotic expressions for the capacity in the discrete case are derived based on the Mutual Information (MI) - Minimum Mean Squared Error (MMSE) relation [3][4][5]. Due to the requirement for high spectral efficiency on current communication systems, the mid-SNR is usually where they operate. The MIMO Constellation Constrained Capacity (CCC) in this region remains unknown. The capacity of a standard impulse response channel with discrete input is another open problem in the area of linear interference channels. The general method for computing it relies on trellis processing [6], which quickly becomes intractable when the channel memory increases. Some extensions and simplifications exist, e.g. [7], which usually attempt to shorten the memory length, however, they still suffer from the inherent complexity of the trellis description.

In [8] we derived a lower bound on the CCC of the ergodic MIMO channel with i.i.d. matrix elements using the QR Decomposition (QRD) of the channel. Here we generalize this result to the single channel realization case, and we use it to also bound the Achievable Information rate (AIR) on a general impulse response channel.

2. CHANNEL MODEL AND COMPLEXITY PROBLEM

Consider a standard MIMO channel model:

$$Y = \mathbf{H}X + W, \tag{1}$$

where X is M-dimensional complex random variable vector $X = [X_1, X_2, \dots X_M]^T$, which is discrete and takes values from the complex-valued set $\mathcal{X}^M$, obtained as the Cartesian product of the basic 1D set $\mathcal{X}$. This can be a QAM, APSK, etc. complex-valued set. The matrix $\mathbf{H}$ represents the $[N \mathbf{x} M]$ complex-valued channel, W is N dimensional complex AWGN, assumed here to have unit variance and Y is the N dimensional channel observation. We assume the channel realization is known at the receiver, but not at the transmitter. The realization of a random variable, e.g. X, at time k will be denoted as $\mathbf{x}_k$ (x_k in the case of 1D variable), and the sequence from time t to k as $\mathbf{x}_t^k = [\mathbf{x}_t, \mathbf{x}_2, \dots \mathbf{x}_k]^T$.

The AIR on the channel when signaling with $\mathcal{X}^M$, having Probability Mass Function (PMF) $p(X)$, and averaging among the possible channel realizations is given by the MI:

$$\begin{aligned}\mathcal{I}(X;Y) = \mathrm{E}_{\mathbf{H}}\left[\mathcal{I}(X;Y|\mathbf{H})\right] = \\ \mathcal{H}(X) - \mathrm{E}_{\mathbf{H}}\left[\mathcal{H}(X|Y,\mathbf{H})\right].\end{aligned} \tag{2}$$

The standard method for calculating the MI is to generate a long enough pair of input-output sequences, and use the fact, that the entropy converges [6]:

$$\mathcal{H}(X|Y,\mathbf{H}) = -\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\log_2 p(\mathbf{x}_k|\mathbf{y}_k,\mathbf{H}). \tag{3}$$

The probability above is calculated from Bayes theorem:

$$p(\mathbf{x}_k|\mathbf{y}_k,\mathbf{H}) = \frac{p(\mathbf{y}_k|\mathbf{x}_k,\mathbf{H})p(\mathbf{x}_k)}{\sum_{\mathbf{x}_k\in\mathcal{X}^M} p(\mathbf{y}_k|\mathbf{x}_k,\mathbf{H})p(\mathbf{x}_k)} \tag{4}$$

Since the normalization term in (4) must be calculated, the complexity grows exponentially with M. Furthermore, in order to see the convergence in (3), K must also be increased with M. Going beyond e.g. 64QAM on a 3x2 channel on a standard computer becomes challenging.

3. LOWER BOUNDS

Let $\mathbf{H} = \mathbf{QR}$ be the QR decomposition of $\mathbf{H}$, where $\mathbf{Q}$ is unitary and $\mathbf{R}$ is upper-triangular. A well known MIMO receiver utilizes the form of $\mathbf{R}$ to successively cancel the interference from previously detected layers, hence Successive Interference Cancellation (SIC), in the following manner: the received samples are pre-processed as $\hat{Y} = \mathbf{Q}^H Y$, and the channel model becomes $\hat{Y}_i = \sum_{j=i}^{M} \mathbf{R}_{i,j} X_j$. Assuming the layers $i+1$ to M are correctly decoded by the following channel code, the symbols can be re-modulated and subtracted from the current layer i. Here we use a similar technique to derive a lower bound on the channel capacity.

Since $\mathbf{Q}$ is unitary and doesn't change the entropy of Y, and thus the MI, we can write:

$$\mathcal{I}(X;Y|\mathbf{H}) = \mathcal{H}(X) - \mathcal{H}(X|\hat{Y}|\mathbf{H}) = \mathcal{H}(X) - \sum_{i=1:M} \mathcal{H}(X_i|\hat{Y}, X_{i+1}^M, \mathbf{H}) \geq \mathcal{H}(X) - \sum_{i=1:M} \mathcal{H}(X_i|\hat{Y}_i, X_{i+1}^M, \mathbf{H}) = \underline{\mathcal{I}}(X;Y|\mathbf{H}), \quad (5)$$

where we have used the fact, that conditioning does not increase the entropy. In order to calculate the terms in the sum, we express the posterior probabilities similar to (4):

$$p(X_i|\hat{Y}_i, X_{i+1}^M, \mathbf{H}) = \frac{p(X_i)p(\hat{Y}_i|X_i, X_{i+1}^M, \mathbf{H})}{\sum_{X_i} p(X_i)p(\hat{Y}_i|X_i, X_{i+1}^M, \mathbf{H})} \quad (6)$$

Since we condition on the following layers, the likelihood above can be expressed as:

$$p(\hat{Y}_i|X_i, X_{i+1}^M, \mathbf{H}) = \mathcal{N}(\hat{Y}_i| \sum_{j=i:M} R_{i,j}X_j, 1) = \mathcal{N}(\hat{Y}_i - \sum_{j=i+1:M} R_{i,j}X_j|X_i, 1), \quad (7)$$

where $R_{i,j}$ is the element on the $i-$th row and $j-$th column of $\mathbf{R}$, and $\mathcal{N}(x|\mu, \sigma^2)$ is a 1D Gaussian function at x, with mean and variance μ and σ^2, respectively. Using (7), lower bound on the MI on each layer can be calculated independently from an SNR-MI Look-Up Table (LUT), where the SNR is given by $|R_{i,i}|^2 \mathrm{E}\left[X_i^2\right]$. When $M <= N$, the achievable rate on the $M-$th layer coincides with the actual capacity for that layer. However, when $M > N$, there is residual interference on the $N+1$-st to the $M-$th layers from layers, which are not yet decoded, and the resulting lower bound becomes poorer. In order to improve it, we model the residual interference as noise, which is a standard practice in communications engineering. The likelihood we use on layers $i > N$ is then:

$$\mathcal{N}(\hat{Y}_N - \sum_{j=i+1:M} R_{N,j}X_j|R_{N,i}X_i, \hat{\sigma}_i), \quad (8)$$

where $\hat{\sigma}_i = 1 + \sum_{j=N:i-1} |R_{N,j}|^2 \mathrm{E}\left[X_i^2\right]$. In this case it is clear, that in the asymptotically high SNR we have:

$$\lim_{\mathrm{E}[X_i^2]\to\infty} \mathcal{I}(X;Y|\mathbf{H}) = \mathcal{H}(X), \quad (9)$$

whereas:

$$\lim_{\mathrm{E}[X_i^2]\to\infty} \underline{\mathcal{I}}(X;Y|\mathbf{H}) = \mathcal{H}(X) - \sum_{i=N+1:M} \mathcal{H}(X_i|Y, H, SNR_i), \quad (10)$$

where the conditional entropy is larger than zero, because $\lim_{\mathrm{E}[X_i^2]\to\infty} SNR_i = \frac{R_{N,i}}{\sum_{j=N:i-1} R_{N,j}}$, which is a finite number.

3.1. Relation to auxiliary channel lower bounds

A simple upper bound on the entropy of a variable X with PDF $p(X)$ can be obtained by using an *auxiliary* probability function $\bar{p}(X) \neq p(X)$. If X is generated by its original PDF, then the upper bound is found by calculating the entropy function from X, but using $\bar{p}(X)$ [6]:

$$\bar{\mathcal{H}}(X) = -\frac{1}{K}\sum_k \log_2 \bar{p}(x_k) \geq \mathcal{H}(X)$$

A lower bound on the MI is derived in a similar manner. Say there is a channel with input-output sequence pair $x \to y$, governed by the laws $p_{Y|X}(Y|X)$, and $p_{X|Y}(X|Y) = \frac{p_{Y|X}(Y|X)p_X(X)}{\sum_X p_{Y|X}(Y|X)p_X(X)}$. Then if y is generated by the law $p_{Y|X}(Y|X)$, the lower bound is calculated as:

$$\underline{\mathcal{I}}(X;Y) = \mathcal{H}(X) - \bar{\mathcal{H}}(X|Y) \leq \mathcal{H}(X) - \mathcal{H}(X|Y) = \mathcal{I}(X;Y), \quad (11)$$

where $\bar{\mathcal{H}}(X|Y)$ is calculated using some valid PMF $\bar{p}_{X|Y}(X|Y) \neq p_{X|Y}(X|Y)$.

Turning back to the $\mathbf{R}$ channel, we use the auxiliary probability distribution $\bar{p}(X|Y, \mathbf{H})$:

$$\bar{p}(X|\hat{Y}, \mathbf{H}) = \prod_{i=1}^{M} \bar{p}(X_i|\hat{Y}, X_{i+1}^M, \mathbf{H}) = \prod_{i=1}^{M} \frac{\bar{p}(\hat{Y}|X_i, X_{i+1}^M, \mathbf{H})p(X_i)}{\sum_{X_i} \bar{p}(\hat{Y}|X_i, X_{i+1}^M, \mathbf{H})p(X_i)},$$

where:

$$\bar{p}(\hat{Y}|X_i, X_{i+1}^M, \mathbf{H}) = \mathcal{N}(\hat{Y}_i - \sum_{j=i+1:M} \mathbf{R}_{i,j}X_j|\mathbf{R}_{i,i}X_i, 1), \quad (12)$$

which leads to the same lower bound.

3.2. Impulse response channels

Consider a standard impulse response channel:

$$y_k = \sum_{i=0:l} h_i x_{k-i} + w_k, \quad (13)$$

where $\mathbf{h} = [h_0, h_1, \dots h_l]^T$ is the impulse response. Equivalently, the channel may be expressed in its matrix form:

$$y_1^k = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ h_l & h_{l-1} & \ddots & \vdots \\ 0 & h_l & \ddots & \vdots \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & h_1 & h_0 \end{bmatrix} \times x_1^k + w_1^k. \quad (14)$$

The MI with channel knowledge at the receiver is now calculated as:

$$\mathcal{I}(X;Y|\mathbf{h}) = \mathcal{H}(Y|\mathbf{h}) - \mathcal{H}(Y|X,\mathbf{h}) = \\ = -\frac{1}{K}\log_2 p(y_1^K|\mathbf{h}) - H(W) \quad (15)$$

The standard approach to calculating (15) is to use a trellis to calculate $p(y_1^K) = \prod_{1:K} p(y_k|y_{1:K-1})$. One section of such trellis is given in Fig. 1. The interfering symbols are cast into the state: $S_k = \{X_{k-l}, ...X_{k-1}\}$, and the current symbol governs the transition. Marginalizing the state, the desired probability at time k is $p(y_1^k) = \sum_{s_k} p(s_k, y_1^k)$, where each term is calculated recursively [6]:

$$p(s_k, y_1^k) = \sum_{x_k} \sum_{s_{k-1}} p(s_{k-1}, y_1^{k-1}) p(y_k|x_k, s_k) p(x_k|s_k)$$

Since the number of states is given by $|\mathcal{S}| = |\mathcal{X}|^l$, the dimensionality problem is the same as for the MIMO channel. The equivalent of the above mentioned 3x2 64QAM here is 64QAM with maximum 2 taps, or similarly - 16QAM with maximum 3 taps, for a standard PC. Trellis pruning techniques may be utilized both in case of MIMO and impulse response channels, leading to the so-called sphere detection [9]. Sphere detection is popular, but is still limited in the number of nodes which can be pruned before the performance degrades significantly. Another approach for the impulse response channel is to use an auxiliary channel of shorter length [7]. The same problem exist here - the more the channel is shortened, the worse auxiliary channel we can find, and thus worse lower bounds.

Instead we can use the QRD based lower bounds. If the channel is expressed as in (14), the QR decomposition may be performed, and a bound may be obtained by the above mentioned LUT. In this case $M = N$, and so Eq. (7) is used. This method is independent of the memory length. The only bottleneck is the QRD computation, which for very long sequences may become problematic. In this paper we used $K = 10^4$, which we found was enough to see convergence for 16QAM constellations. The QRD on the $\left[10^4 \text{x} 10^4\right]$ matrix was computed in a few seconds on the PC we used. We note that the channel matrix in this case is highly structured and periodic, and the $\mathbf{R}$ matrix therefore may be expected to also hold some structure. For example, in all our simulations the diagonal elements of the $\mathbf{R}$ matrix either converged to some value, or to some periodic pattern. However, exploiting this periodicity is left for future research.

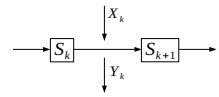


Figure 1. Trellis representation of the impulse response channel (13)

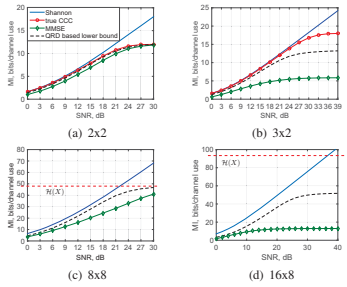
4. RESULTS

4.1. MIMO channel

In Fig. 2(a) the lower bounds from Eq. (5) are shown for a 2x2 MIMO with 64QAM input, together with the true MI, as calculated from Eq. (2). The input PMF is uniform. For comparison, we also plot the AIRs with the popular linear MMSE receiver processing [9]. We see that the true information rate is closely approached by the proposed method. The MMSE processing also calculates a lower bound, however, poorer than the QRD based one. As mentioned in Section 3, in the case of $M > N$, the bounds will not be as tight. In Fig. 2(b) the AIRs are shown for a 3x2 MIMO with 64QAM input. We see a significant underestimation, especially in the high SNR region. However, we note that the transmit diversity system is generally not used for maximizing throughput, and therefore a practical system would not operate at this high SNR region with an input of rank, which is larger than $rank(H) \leq \min(M, N)$. In the low-to-mid SNR, the QRD based bound may still be used. In Fig. 2(c) the AIRs on a 8x8 system are shown, where the full-complexity algorithm can no longer be used. The QRD based lower bound follows the slope of the Gaussian capacity, and converges to $\mathcal{H}(X)$. When we further increase M, more terms are added in the conditional entropy in Eq. (10), and the lower bound becomes worse. However, the slope at low-to-mid SNR is still the same as the Gaussian capacity. Finally in this section we note, that the uniform PMF is not a requirement. The bounds hold for any PMF, which is independent across dimensions. The consequence is that optimization can also be performed using the auxiliary function (7). The PMF, which is optimized for the auxiliary channel can then be used on the true channel, and the AIR in that case is still bounded by what is achieved in (5). Some results obtained by the well known Blahut-Arimoto algorithm for optimization of the input PMF on an ergodic MIMO channel may be found in [8].

4.2. Impulse response channel

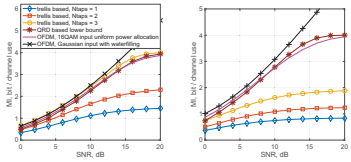
We also analyze the QRD based lower bound on a fixed impulse response channel, where $\mathbf{h}$ is obtained from standard Gaussian distribution. In Figures 3(a) and 3(b) we see the AIRs on an impulse response channel with $l = 3$ and $l = 6$, respectively (channel as given in the caption). Without loss of generality, we sort the channel elements



(a) 2x2 (b) 3x2

(c) 8x8 (d) 16x8

Figure 2. AIRs on MIMO channels of different size



(a) $l = 3$ (b) $l = 6$

Figure 3. Achievable information rates for the channels: a) $\mathbf{h} = [0.37 - 0.18i, -0.35 + 0.05i, -0.20 - 0.26i, -0.23 - 0.17i]^T$; b) $\mathbf{h} = [-0.18 - 0.47i, -0.35 - 0.26i, 0.37 - 0.06i, -0.26 + 0.25i, -0.17 + 0.27i, -0.23 + 0.18i, -0.20 + 0.05i]^T$

in descending order of their amplitude. The input symbols are i.i.d., and so this does not change the AIRs, but makes the implementation of the trellis simpler, since the state actually represents previous symbols. In the general case, we would like our state to represent the symbols, responsible for largest interference. We compare the QRD based bounds with the trellis based method, which casts $Ntaps$ previous symbols into the state, and the rest $l - Ntaps$ symbols are modeled as noise, similar to Eq. (8). When $Ntaps = l$ the AIR is the true CCC. For comparison we also include the AIRs using OFDM. We note that OFDM with Gaussian input and water-filling power allocation is the power constrained channel capacity. On the short channels, the QRD based lower bound closely approaches the constrained capacity, achieved with the trellis algorithm. It is slightly outperformed in the low SNR by OFDM, and slightly outperforms OFDM in the mid-to-high SNR. When we increase the channel length, the trellis based algorithm can be used with up to 3 taps on the PC we used for simulations. The QRD bound in this case is able to provide larger improvement over the OFDM. Both figures show that orthogonalization of the channel when the input is discrete can be sub-optimal, confirming the results from [2][8].

5. CONCLUSION

In this paper some of the more popular linear interference channels are studied. Lower bounds on the AIRs are derived using the QR decomposition of the channel. In case of linear 1D channels with memory, the QRD is performed on the matrix form of the channel. Based on the diagonal elements of the $\mathbf{R}$ matrix, an SNR-AIR look-up table can be efficiently used to find lower bounds on capacity. These bounds were shown to closely approach the true constellation constrained capacity, where the latter can be computed by standard PC, and were also shown to have good performance in terms of slope and distance to Gaussian capacity in most cases of interest.

6. REFERENCES

[1] G. J. Foschini and Gans M. J, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[2] F. Pereź-Cruz, M. R. D. Rodrigues, and S. Verdú, "MIMO gaussian channels with arbitrary inputs: Optimal precoding and power allocation," *IEEE Trans. Inf. Theory*, vol. 56, pp. 1070–1084, 2010.

[3] Y. Wu and S. Verdú, "Functional properties of minimum mean-square error and mutual information," *IEEE Trans. Inf. Theory*, vol. 58, pp. 1289–1301, 2012.

[4] A. Alvarado, F. Brännström, E. Agrell, and T. Koch, "High-SNR asymptotics of mutual information for discrete constellations with applications to BICM," *IEEE Trans. Inf. Theory*, vol. 60, pp. 1061–1076, 2014.

[5] M. Rodrigues, "Multiple-antenna fading channels with arbitrary inputs: Characterization and optimization of the information rate," *IEEE Trans. Inf. Theory*, vol. 60, pp. 569–585, 2014.

[6] D. M. Arnold, H. Loeliger, P. O. Vontobel, A. Kavčić, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inf. Theory*, vol. 52, pp. 3498–3508, 2006.

[7] F. Rusek and D. Fertonani, "Bounds on the information rate of intersymbol interference channels based on mismatched receivers," *IEEE Trans. Inf. Theory*, vol. 58, pp. 1470–1482, 2012.

[8] M. P. Yankov, S. Forchhammer, K. J. Larsen, and L. P. B. Christensen, "Approximating the constellation constrained capacity of the MIMO channel with discrete input," in *Proc. IEEE International Conference on Communications.*, Jun. 2015, pp. 5634–5639.

[9] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*, Cambridge University Press, 2007.

Appendix B

Mapping functions for many-to-one turbo-coded BICM

In this Chapter we provide the mapping functions that were used in this thesis. The symbol 'X' in the mapping represents ambiguity - that is, both '0' and '1' are mapped to that bit position. We only plot the 1D marginal mapping functions, as obtained by the algorithm from **PAPER 2**. Note that the 256QAM for AWGN channel and the optimized 256QAM for the WDM channel are identical.

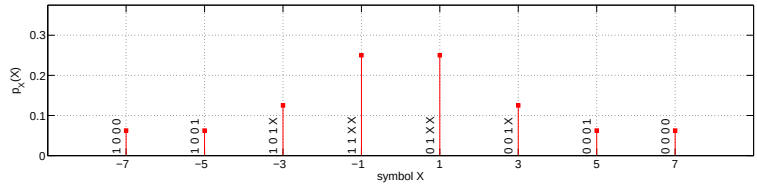


Figure B.1: 8PAM marginal mapping function for the 64QAM constellation used in Section 5.2. $m = 8$.

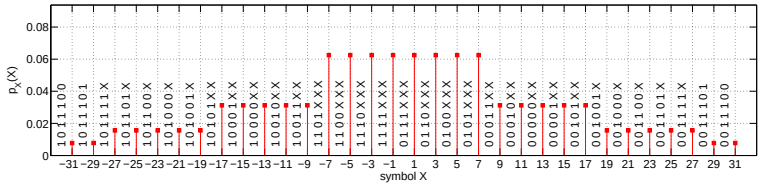


Figure B.2: 32PAM marginal mapping function for the 1024QAM constellation used in Section 5.2. $m = 14$.

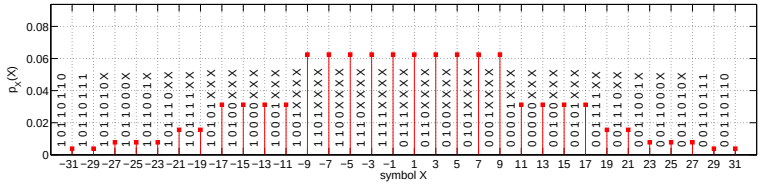


Figure B.3: 32PAM marginal mapping function for the 1024QAM constellation used in Section 5.2. $m = 16$.

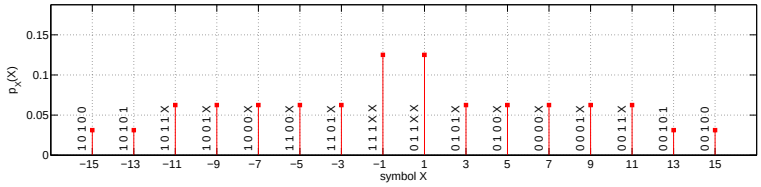


Figure B.4: 16PAM marginal mapping function for the 256QAM constellation used in Section 5.3 and Section 5.2. $m = 10$.

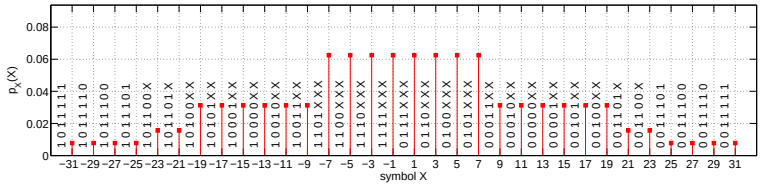


Figure B.5: 32PAM marginal mapping function for the 1024QAM constellation used in Section 5.3. $m = 14$.

List of Acronyms

ACF	auto-correlation function
ADC	analog-to-digital conversion
AIR	achievable information rate
ASE	amplified spontaneous emission
AWGN	additive white Gaussian noise
BER	bit error rate
BICM	bit-interleaved coded modulation
BLER	block error rate
BAA	Blahut-Arimoto algorithm
CCC	constellation constrained capacity
CD	chromatic dispersion
DBP	digital back-propagation
DD	decision directed
EDFA	Erbium doped fiber amplifier
EM	expectation-maximization
EXIT	extrinsic information transfer
FEC	forward error correction

FWM	four wave mixing
GHC	geometric Huffman codes
IDRA	ideal distributed Raman amplification
ISI	inter symbol interference
IT	information technology
KLD	Kullback-Leibler divergence
LDPC	low-density parity check
LLW	laser linewidth
LO	local oscillator
LUT	look-up table
MAP	maximum a-posteriori probability
MI	mutual information
MIMO	multiple-input multiple-output
MMSE	minimum mean squared error
NLPN	non-linear phase noise
NLSE	non-linear Schrödinger equation
OFDM	orthogonal frequency division multiplexing
OSNR	optical SNR
PAM	pulse amplitude modulation
PDF	probability density function
PMD	polarization mode dispersion
PMF	probability mass function
PSD	power spectral density

QAM	quadrature amplitude modulation
QPSK	quadrature phase shift keying
QRD	QR decomposition
RV	random variable
SE	spectral efficiency
SER	symbol error rate
SNR	signal-to-noise ratio
SPM	self-phase modulation
SSFM	split-step Fourier method
SSMF	standard single mode fiber
SVD	singular value decomposition
WDM	wavelength division multiplexing
XPM	cross-phase modulation

Bibliography

[1] "ITU background report," ITU Symposium on ICTs and Climate Change, Quito, Ecuador, Tech. Rep., July 2009.

[2] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal,*, vol. 27, no. 3, pp. 379–423, July 1948.

[3] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.

[4] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 399–431, Mar. 1999.

[5] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error - correcting coding and decoding : Turbo-codes," in *Proc. of IEEE International Conference on Communications*, May 1993, pp. 1064–1070.

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

[7] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[8] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[9] M. Rodrigues, "Multiple-antenna fading channels with arbitrary inputs : characterization and optimization of the information rate," *IEEE Transactions on Information Theory*, vol. 60, no. 1, pp. 569–585, Jan. 2014.

[10] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[11] "The zettabyte era: Trends and analysis," Cisco Systems, Inc., Tech. Rep., May 2015.

[12] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *IEEE Journal of Lightwave Technology*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[13] R. Dar, M. Shtaif, and M. Feder, "New bounds on the capacity of the nonlinear fiber-optic channel," *Optics letters*, vol. 39, no. 2, pp. 398–401, Jan. 2014.

[14] G. Kramer, M. I. Yousefi, and F. R. Kschischang, "Upper bound on the capacity of a cascade of nonlinear and noisy channels," Mar. 2015. [Online]. Available: http://arxiv.org/abs/1503.07652

[15] E. Agrell and M. Karlsson, "Influence of behavioral models on multiuser channel capacity," *IEEE Journal of Lightwave Technology*, vol. 33, no. 17, pp. 3507–3515, June 2015.

[16] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavčić, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3498–3508, Aug. 2006.

[17] T. M. Cover and J. A. Thomas, *elements of information theory, 2nd edition.* Hoboken, NJ: John Wiley & Sons, Inc., 2006.

[18] A.-F. C. Ibrahim, M. D. Trott, and S. Shamai, "The capacity of discrete-time memoryless rayleigh-fading channels," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1290–1301, May. 2001.

[19] R. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Transactions on Information Theory*, vol. 18, no. 4, pp. 460–473, July 1972.

[20] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 14–20, Jan. 1972.

[21] N. Varnica, X. Ma, and A. Kavčić, "Capacity of power constrained memoryless AWGN channels with fixed input constellations," in *Proc. of IEEE Globecom*, Nov. 2002, pp. 1339–1343.

[22] J. Bellorado, S. Ghassamzadeh, and A. Kavčić, "Approaching the capacity of the MIMO rayleigh flat-fading channel with QAM constellations, independent across antennas and dimensions," *IEEE Transactions on Information Theory*, vol. 5, no. 6, pp. 1322–1332, June 2006.

[23] J. Akhtman and L. Hanzo, "Closed-form approximation of MIMO capacity," *IET Electronics Letters*, vol. 45, no. 1, pp. 1–2, Jan. 2009.

[24] Y. Wu and S. Verdú, "Functional properties of minimum mean-square error and mutual information," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1289–1301, Mar. 2012.

[25] ——, "Functional properties of MMSE," in *Proc. of IEEE International Symposium on Information Theory*, June 2010, pp. 1453–1457.

[26] A. Alvarado, F. Brännström, E. Agrell, and T. Koch, "High-SNR asymptotics of mutual information for discrete constellations with applications to BICM," *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 1061–1076, Feb. 2014.

[27] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[28] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*. New York: Cambridge University Press, 2007.

[29] P. O. Vontobel, A. Kavčić, D. M. Arnold, and H.-A. Loeliger, "A generalization of the Blahut Arimoto algorithm to finite-state channels," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 1887–1918, May 2008.

[30] F. Rusek and D. Fertonani, "Bounds on the information rate of intersymbol interference channels based on mismatched receivers," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1470–1482, Mar. 2012.

[31] Y. Carmon and S. Shamai, "Lower bounds and approximations for the information rate of the ISI channel," *IEEE Transactions on Information Theory*, vol. 61, no. 10, pp. 5417–5431, Oct. 2015.

[32] L. Barletta, M. Magarini, and A. Spalvieri, "The information rate transferred through the discrete-time Wiener's phase noise channel," *IEEE Journal of Lightwave Technology*, vol. 30, no. 10, pp. 1480–1486, May 2012.

[33] L. Barletta, F. Bergamelli, M. Magarini, N. Carapellese, and A. Spalvieri, "Pilot-aided trellis-based demodulation," *IEEE Photonics Technology Letters*, vol. 25, no. 13, pp. 1234–1237, July 2013.

[34] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, no. 1, pp. 284–287, Mar. 1974.

[35] S. Pecorino, S. Mandelli, L. Barletta, M. Magarini, and A. Spalvieri, "Bootstrapping iterative demodulation and decoding without pilot symbols," *IEEE Journal of Lightwave Technology*, vol. 33, no. 17, pp. 3613–3622, Sep. 2015.

[36] L. Barletta, M. Magarini, and A. Spalvieri, "A new lower bound below the information rate of Wiener phase noise channel based on Kalman carrier recovery," *Optics Express*, vol. 20, no. 23, pp. 25 471–25 477, Nov. 2012.

[37] ——, "Tight upper and lower bounds to the information rate of the phase noise channel," in *Proc. of IEEE International Symposium on Information Theory*, July 2013, pp. 2284–2288.

[38] A. Barbieri and G. Colavolpe, "Soft-output decoding of rotationally invariant codes over channels with phase noise," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 2125–2133, Mar. 2007.

[39] S. Zhang, P.-Y. Kam, C. Yu, and J. Chen, "Decision-aided carrier phase estimation for coherent optical communications," *IEEE Journal of Lightwave Technology*, vol. 28, no. 11, pp. 1597–1607, June 2010.

[40] T. N. Huynh, A. T. Nguyen, W.-C. Ng, L. Nguyen, L. A. Rusch, and L. P. Barry, "BER performance of coherent optical communications systems employing monolithic tunable lasers with excess phase noise," *IEEE Journal of Lightwave Technology*, vol. 32, no. 10, pp. 1973–1980, May 2014.

[41] T. Pfau, S. Hoffman, and R. Noé, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for QAM constellations," *IEEE Journal of Lightwave Technology*, vol. 27, no. 8, pp. 989–999, Apr. 2009.

[42] M. Magarini, L. Barletta, A. Spalvieri, F. Vacondio, M. Pfau, T amd Pepe, and G. Bertolini, M amd Gavioli, "Pilot-symbols-aided carrier-phase recovery for 100-G PM-QPSK digital coherent receivers," *IEEE Photonics Technology Letters*, vol. 24, no. 9, pp. 739–741, May 2012.

[43] C. Pan and F. R. Kschischang, "Coded-aided phase tracking for coherent fiber channels," *IEEE Journal of Lightwave Technology*, vol. 32, no. 6, pp. 1041–1047, Mar. 2014.

[44] C. Pan, H. Büllow, W. Idler, L. Schmalen, and F. R. Kschischang, "Optical nonlinear-phase-noise compensation for 32 Gbaud PolDM-16 QAM transmission using a code-aided expectation-maximization algorithm," *IEEE Journal of Lightwave Technology*, vol. 33, no. 17, pp. 3679–3686, Sep. 2015.

[45] I. Fatadin, D. Ives, and S. J. Savory, "Laser linewidth tolerance for 16-QAM coherent optical systems using QPSK partitioning," *IEEE Photonics Technology Letters*, vol. 22, no. 9, pp. 631–633, May 2010.

[46] G. P. Agrawal, *Fiber-optic communication systems.* Hoboken, NJ: John Wiley & Sons, Inc., 2010.

[47] E. Ip, "Nonlinear compensation using backpropagation for polarization-multiplexed transmission," *IEEE Journal of Lightwave Technology*, vol. 28, no. 6, pp. 939–951, Mar. 2010.

[48] N. Jiang, Y. Gong, J. Karout, H. Wymeersch, P. Johannisson, M. Karlsson, E. Agrell, and P. A. Andrekson, "Stochastic backpropagation for coherent optical communications," in *Proc. of European Conference on Optical Communications*, Sep. 2011, p. We.10.P1.81.

[49] N. V. Irukulapati, H. Wymeersch, P. Johannisson, and E. Agrell, "Stochastic digital backpropagation," *IEEE Journal of Lightwave Technology*, vol. 62, no. 11, pp. 3956–3968, Nov. 2014.

[50] S. J. Savory, "Digital filters for coherent optical receivers," *Optics Express*, vol. 16, no. 2, pp. 804–817, Jan. 2008.

[51] T. Freckmann, R.-J. Essiambre, P. J. Winzer, G. J. Foschini, and G. Kramer, "Fiber capacity limits with optimized ring constellations," *IEEE Photonics Technology Letters*, vol. 21, no. 20, pp. 1496–1498, Oct. 2009.

[52] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *IEEE Journal of Lightwave Technology*, vol. 32, no. 4, pp. 694–721, Feb. 2014.

[53] E. Agrell, A. Alvarado, G. Durisi, and M. Karlsson, "Capacity of a nonlinear optical channel with finite memory," *IEEE Journal of Lightwave Technology*, vol. 32, no. 16, pp. 2862–2876, Aug. 2014.

[54] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Optics Express*, vol. 21, no. 22, pp. 25 685–25 699, Nov. 2013.

[55] ——, "Inter-channel nonlinear interference noise in WDM systems : Modeling and mitigation," *IEEE Journal of Lightwave Technology*, vol. 33, no. 5, pp. 1044–1053, Mar. 2015.

[56] ——, "On shaping gain in the nonlinear fiber-optic channel," in *Proc. of IEEE International Symposium on Information Theory*, July 2014, pp. 2794–2798.

[57] M. Secondini and E. Forestieri, "On XPM mitigation in WDM fiber-optic systems," *IEEE Photonics Technology Letters*, vol. 26, no. 22, pp. 2252–2255, Nov. 2014.

[58] D. Marsella, M. Secondini, E. Agrell, and E. Forestieri, "A simple strategy for mitigating XPM in nonlinear WDM optical systems," in *Proc. of Optical Fiber Communication Conference*, Mar. 2015, p. Th4D.3.

[59] G. Ungerboeck and I. Csajka, "On improving data link performance by increasing the channel alphabet and introducing sequence coding," in *Proc. of IEEE International Symposium on Information Theory*, June 1977.

[60] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 429–445, Mar. 1996.

[61] S. Lin and D. J. Costello, *Error Control Coding, 2nd edition.* Upper Saddle River, NJ: Pearson Education Inc., 2004.

[62] C. Stierstorfer and R. F. H. Fischer, "Gray mappings for bit-interleaved coded modulation," in *Proc. of IEEE 65th Vehicular Technology Conference*, Spring 2007, pp. 1703–1707.

[63] E. Agrell and A. Alvarado, "Optimal alphabets and binary labelings for BICM at low SNR," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6650–6672, Oct. 2011.

[64] S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping and decoding for multilevel modulation," in *Proc. of IEEE Globecom*, Nov. 1998, pp. 579–584.

[65] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[66] L. Hanzo, R. G. Maunder, J. Wang, and L.-L. Yang, *Near-Capacity Variable-Length Coding: Regular and EXIT-Chart-Aided Irregular Designs.* The Atrium, Southern Gate, Chichester, West Sussex: John Wiley and sons, ltd., 2011.

[67] F. Schreckenbach, "Iterative decoding of bit-interleaved coded modulation," Ph.D. dissertation, Technischen Universität München, 2007.

[68] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission.* New York: John Wiley and sons, inc., 2002.

[69] D. G. Forney, "Trellis shaping," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 281–300, Mar. 1992.

[70] A. R. Calderbank and L. H. Ozarow, "Nonequiprobable signaling on the Gaussian channel," *IEEE Transactions on Information Theory*, vol. 36, no. 4, pp. 726–740, July 1990.

[71] B. P. Smith and F. R. Kschischang, "A pragmatic coded modulation scheme for high-spectral-efficiency fiber-optic communications," *IEEE Journal of Lightwave Technology*, vol. 30, no. 13, pp. 2047–2053, July 2012.

[72] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 913–929, May 1993.

[73] M. C. Valenti and X. Xiang, "Constellation shaping for bit-interleaved coded APSK," in *Proc. of IEEE International Conference on Communications*, June 2011, pp. 1–5.

[74] L. Beygi, E. Agrell, J. M. Kahn, and M. Karlsson, "Rate-adaptive coded modulation for fiber-optic communications," *IEEE Journal of Lightwave Technology*, vol. 32, no. 2, pp. 333–343, Jan. 2014.

[75] C. Fragouli, R. D. Wesel, D. Sommer, and G. P. Fettweis, "Turbo codes with non-uniform constellations," in *Proc. of IEEE International Conference on Communications*, June 2001, pp. 70–73.

[76] M. Noemm, A. Mourad, and A. Hoeher, "Superposition modulation with irregular convolutional coding," in *Proc. of IEEE Globecom*, Dec. 2012, pp. 2346–2350.

[77] J. Estarán, D. Zibar, and I. T. Monroy, "Capacity-approaching superposition coding for optical fiber links," *IEEE Journal of Lightwave Technology*, vol. 32, no. 17, pp. 2960–2972, Sep. 2014.

[78] G. Böcherer, P. Schulte, and F. Steiner, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," Apr. 2015. [Online]. Available: http://arxiv.org/abs/1502.02733

[79] F. Schreckenbach, "Approaching AWGN channel capacity using non-unique symbol mappings," in *Proc. of 7th Australian Communications Theory Workshop*, Mar. 2006.

[80] D. Raphaeli and A. Gurevitz, "Constellation shaping for pragmatic turbo-coded modulation with high spectral efficiency," *IEEE Transactions on Communications*, vol. 52, no. 3, pp. 341–345, Mar. 2004.

[81] D. A. Huffman, "A method for the construction of minimum-redundancy codes," in *Proc. of IRE*, 1951, pp. 1098–1101.

[82] G. Böcherer, F. Altenbach, and R. Mathar, "Capacity achieving modulation for fixed constellations with average power constraint," in *Proc. of IEEE International Conference on Communications*, June. 2011, pp. 1–5.

[83] S. Benedetto, R. Garello, and G. Montorisi, "A search for good convolutional codes to be used in the construction of turbo codes," *IEEE Transactions on Communications*, vol. 46, no. 9, pp. 1101–1105, Sep. 1998.

[84] H.-J. Zepernick and A. Finger, *Pseudo random signal processing : theory and application*. The Atrium, Southern Gate, Chichester, West Sussex: John Wiley and sons, inc., 2005.