CHALMERS


# On Signal Constellations and Coding for Long-Haul Fiber-Optical Systems

Christian Häger

*Department of Signals and Systems*
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden, 2014

THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING

# On Signal Constellations and Coding for Long-Haul Fiber-Optical Systems

Christian Häger


CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Signals and Systems
Chalmers University of Technology
Göteborg, Sweden, 2014

**On Signal Constellations and Coding for Long-Haul Fiber-Optical Systems**

Christian Häger



Technical Report No. R008/2014
ISSN 1403-266X

This thesis has been prepared using LaTeX and PSTricks.

Department of Signals and Systems
Chalmers University of Technology
SE-412 96 Göteborg, Sweden
Phone: +46 (0)31 772 1000
www.chalmers.se

Printed by Chalmers Reproservice
Göteborg, Sweden, April 2014

# Abstract

Motivated by the realization that even the enormous bandwidth available in an optical fiber is finite and valuable, the design of *spectrally efficient* long-haul fiber-optical communication systems has become an important research topic. Compared to other wireline technologies, e.g., transmission over coaxial cables, the main challenge comes from the inherent nonlinearity of the underlying communication channel caused by the relatively high signal intensities. In this thesis, we study the design of spectrally efficient fiber-optical systems for both uncoded and coded transmission scenarios.

We consider the problem of designing higher-order signal constellations for a system that is severly impaired by nonlinear phase noise. By optimizing amplitude phase-shift keying constellations, which can be seen as the union of phase-shift keying constellation with different amplitude levels, gains of up to 3.2 dB at a symbol error probability of $10^{-2}$ are shown to be achievable compared to conventional constellations. We also illustrate a somewhat counterintuitive behavior of optimized constellations for very high input powers and nonlinear distortions. In particular, sacrificing a constellation point or ring may be beneficial in terms of the overall performance of the constellation.

Furthermore, we study polarization-multiplexed transmission, where spectral efficiency is increased by encoding data onto both polarizations of the light. For a memoryless fiber-optical channel, we introduce a low-complexity detector which is based on an amplitude-dependent phase rotation and subsequent threshold detection. The complexity compared to the four-dimensional maximum likelihood detector is considerably reduced, albeit at the expense of some performance loss.

Lastly, we consider the design of a coded fiber-optical system operating at high spectral efficiency. In particular, we study the optimization of the mapping of the coded bits to the modulation bits for a polarization-multiplexed fiber-optical system that is based on the bit-interleaved coded modulation paradigm. This technique, which we refer to as bit mapper optimization, is extended to the class of spatially coupled low-density parity-check codes, which have shown outstanding performance over memoryless binary-input channels. For a transmission scenario without optical inline dispersion compensation, the results show that the transmission reach can be extended by roughly up to 8%, without significantly increasing the system complexity.



**Keywords:** Bit-interleaved coded modulation, bit mapper, constellation optimization, detector, fiber-optical communication, low-density parity-check codes, spatial coupling.

# List of Publications

This thesis is based on the following publications:

**Paper A**
Christian Häger, Alexandre Graell i Amat, Alex Alvarado, Erik Agrell, "Design of APSK Constellations for Coherent Optical Channels with Nonlinear Phase Noise", *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3362–3373, Aug. 2013

**Paper B**
Christian Häger, Lotfollah Beygi, Erik Agrell, Pontus Johannisson, Magnus Karlsson, Alexandre Graell i Amat, "A Low-Complexity Detector for Memoryless Polarization-Multiplexed Fiber-Optical Channels", *IEEE Commun. Lett.*, vol. 18, no. 2, pp. 368–371, Jan. 2014

**Paper C**
Christian Häger, Alexandre Graell i Amat, Fredrik Brännström, Alex Alvarado, Erik Agrell "Improving Soft FEC Performance for Higher-Order Modulations by Bit Mapper Optimization", submitted to *Optics Express*, Apr. 2014

Other publications by the author not included in this thesis:

- Christian Häger, Alexandre Graell i Amat, Alex Alvarado, Erik Agrell, "Constellation Optimization for Coherent Optical Channels Distorted by Nonlinear Phase Noise", in Proc. *IEEE Global Communications Conf. (GLOBECOM)*, Anaheim, CA, Dec. 2012
- Christian Häger, Alexandre Graell i Amat, Alex Alvarado, Fredrik Brännström, Erik Agrell, "Optimized Bit Mappings for Spatially Coupled LDPC Codes over Parallel Binary Erasure Channels", in Proc. *IEEE Int. Conf. Communications (ICC)*, Sydney, Australia, June 2014

# Acknowledgements

First of all, I would like to thank my supervisor Alexandre Graell i Amat and my co-supervisors Alex Alvarado, Fredrik Brännström, and Erik Agrell for all the fruitful discussions and the invaluable feedback that I have received throughout the past two and a half years. In particular, I would like to thank Alexandre for his constant support and guidance. I would also like to thank Alex again for being an outstanding host during my research visit in Cambridge.

I would like extend my thanks to Iryna Andriyanova for sharing her insight on spatially coupled LDPC codes during my stay in Paris. Also, one of the persons who taught me a lot during these "early" PhD years is Lotfollah Beygi. I am immensely grateful for all the time that he spent with me discussing various aspects of fiber-optical communication. Paper B in this thesis is a collaborative effort which was initiated by Lotfollah. I would also like to thank Mikhail Ivanov for proofreading parts of the thesis.

Many thanks also to the administrative staff of the department, in particular Lars Börjesson, Agneta Kinnander, Madeleine Persson, and Natasha Adler. Last but not least, I would like to thank everyone in our group, and also everyone involved in FORCE, the fiber-optic research center at Chalmers, for creating a fantastic and stimulating work environment.

Regarding the financial support, I would like to thank and acknowledge the Ericsson Research Foundation for their generous financial support of my research visit to Cambridge.

Christian Häger, Göteborg, March 2014

# Acronyms

| Abbreviation | Meaning |
|---|---|
| **APSK** | amplitude phase-shift keying |
| **ASE** | amplified spontaneous emission |
| **AWGN** | additive white Gaussian noise |
| **BEC** | binary erasure channel |
| **BEP** | bit error probability |
| **BICM** | bit-interleaved coded modulation |
| **BP** | belief propagation |
| **BRGC** | binary reflected Gray code |
| **CM** | coded modulation |
| **CN** | check node |
| **DBP** | digital backpropagation |
| **DCF** | dispersion compensating fiber |
| **DE** | density evolution |
| **DM** | dispersion-managed |
| **EDFA** | erbium-doped fiber amplifier |
| **EXIT** | extrinsic information transfer |
| **FEC** | forward error correction |
| **FFT** | fast Fourier transform |
| **GMI** | generalized mutual information |
| **IID** | independent and identically distributed |
| **LB** | lower bound |
| **LDPC** | low-density parity-check |
| **LLR** | log-likelihood ratio |
| **MAP** | maximum a posteriori |
| **MI** | mutual information |
| **ML** | maximum likelihood |
| **MLCM** | multilevel coded modulation |
| **MLSE** | maximum likelihood sequence estimation |
| **NLPN** | nonlinear phase noise |
| **NLSE** | nonlinear Schrödinger equation |
| **OOK** | on-off keying |
| **PDF** | probability density function |
| **PM** | polarization-multiplexed |
| **PMD** | polarization mode dispersion |
| **PSD** | power spectral density |
| **PSK** | phase-shift keying |

| Abbreviation | Meaning |
|---|---|
| **QAM** | quadrature amplitude modulation |
| **QPSK** | quadrature phase-shift keying |
| **ROADM** | reconfigurable optical add-drop multiplexer |
| **RPS** | Raman pump station |
| **SC-LDPC** | spatially coupled low-density parity-check |
| **SEP** | symbol error probability |
| **SMF** | single-mode fiber |
| **sNLSE** | stochastic nonlinear Schrödinger equation |
| **SNR** | signal-to-noise ratio |
| **SP** | single-polarization |
| **SSFM** | split-step Fourier method |
| **TCM** | trellis coded modulation |
| **TS** | two-stage |
| **VN** | variable node |
| **WD** | windowed decoder |
| **WDM** | wavelength-division multiplexing |

# Contents

# Part I

# Introduction

# CHAPTER 1

# Background

When requesting a website, most internet users are probably unaware that the digital data is modulated onto a light source and transmitted over thousands of kilometers in an optical waveguide, a so-called optical fiber, at some point on the way from the remote server to their home computer or mobile device. In fact, more than 99% of the global intercontinental traffic is carried over optical fiber and such "long-haul" fiber-optical communication systems are the key enabler of high-speed internet data transfer connecting cities, countries, and continents [1].

Optical fiber as a transmission medium is very well suited for sending large amounts of data over long distances. Engineers and physicists have spent a great effort to refine the fiber material to be very transparent for optical light over a large frequency range and to improve the components in an optical transmission system, e.g., lasers and amplifiers. Meanwhile, however, fiber-optical systems traditionally employ digital modulation techniques that are rather wasteful with the available frequency spectrum. As an example, switching the light source on and off according to the digital data stream, referred to as on-off keying (OOK), is highly inefficient from a spectral viewpoint.

To keep up with the increasing data rate demands of current applications, and to enable innovative broadband technologies in the future, it becomes more and more apparent that next generation fiber-optical systems need to use the available spectrum more efficiently. The bandwidth of optical fibers is now considered a limited resource and due to this realization, there is currently a great deal of interest in determining the ultimate limits of optical systems in terms of spectral efficiency [2–4] and developing practical schemes that can achieve these limits [5–7]. The distinctive feature, and at the same time the main challenge in fiber-optical communication, is that the underlying communication

channel is fundamentally nonlinear. The fiber nonlinearity is considered as one of the limiting factors for increasing the data rates in long-haul systems [8–10]. Some important problems that arise in the design of spectrally efficient fiber-optical systems are addressed in this thesis.

To improve the spectral efficiency over OOK, the data can be encoded into multiple amplitude and/or phase levels of the optical carrier. This leads to the problem of designing good signal constellations that are robust to the severe nonlinear distortions. Signal constellation design can be thought of as optimizing the placement of a given number of points in two or more dimensions under some constraints. This problem can be considered as a classical problem in communication theory [11, Ch. 1] and it is revisited in Paper A, paying special attention to fiber nonlinearities.

A further increase in spectral efficiency can be made by properly utilizing both polarizations of the optical light, which we refer to as polarization-multiplexed (PM) transmission. Compared to single-polarization (SP) transmission, signals are now represented as points in a four- rather than a two-dimensional space. Optimal maximum likelihood (ML) detection in four dimensions can be computationally very complex, particularly for constellations with many points. Therefore, in Paper B, we study a detector design for a recently developed four-dimensional, memoryless fiber-optical channel model with the intention to significantly reduce the detection complexity.

The problems described above are related to uncoded transmission schemes and do not consider forward error correction (FEC). However, FEC needs to be considered in order to operate close to the ultimate transmission limits of optical fibers. We therefore also consider the problem of designing coded transmission systems at high spectral efficiencies in Paper C. Here, we restrict ourselves to systems that are based on the bit-interleaved coded modulation (BICM) paradigm. BICM can be seen as a pragmatic way to combine signal constellations consisting of many points with powerful binary error correction codes. In particular, we study how the coded bits should be allocated to the modulation bits. We refer to this problem as bit mapper optimization. As one particular example, we consider spatially coupled low-density parity-check (SC-LDPC) codes which have recently been shown to achieve outstanding performance over a variety of communication channels [12–14].

## 1.1 Thesis Organization

The licentiate degree is an intermediate step for a doctoral student towards the final PhD degree and the licentiate thesis documents the progress that has been made over a period of roughly two to two and a half years. The format of this thesis is a so-called collection of papers. It is divided into two parts, where the first part serves as an introduction to the appended papers in the second part.

The remainder of the first part of this thesis is structured as follows. In Chapter 2, we provide an introduction to fiber-optical channel modeling and describe the origin of

the channel models that are used in the appended papers. Our main goal is to show that all utilized models have the same mathematical foundation, namely the nonlinear Schrödinger equation (NLSE), albeit assuming somewhat different material parameters and system configurations. In Chapter 3, we explain the necessary theoretical background for the optimization problem that we consider in Paper C. For Paper C, the reader is assumed to be somewhat familiar with coding theory and in particular low-density parity-check (LDPC) codes, which we discuss in Chapter 3. Finally, some conclusions are summarized in Chapter 4, where we also briefly discuss future work.

## 1.2 Notation

Throughout this thesis, vectors are denoted by boldface letters $\mathbf{A}$, matrices by blackboard letters $\mathbb{A}$, sets by calligraphic letters $\mathcal{A}$ (except the sets containing the real numbers $\mathbb{R}$, complex numbers $\mathbb{C}$, integers $\mathbb{Z}$, and natural numbers $\mathbb{N}$), and random variables by capital letters $A$. The probability density function (PDF) of a continuous random variable $Y$ is denoted by $f_Y(\,\cdot\,)$ and $\mathbb{E}[\,\cdot\,]$ denotes expectation. A matrix transpose is denoted by $(\,\cdot\,)^\intercal$. $\delta(t)$ denotes Dirac's delta function while $\delta[k]$ denotes the Kronecker delta. Convolution is denoted by $\circledast$. For a continuous-time signal $x(t)$, we write $x(t) \circ\!\!-\!\!\bullet X(f)$ to indicate its Fourier transform. The imaginary unit is denoted by $\jmath \triangleq \sqrt{-1}$.

We also acknowledge the following notational inconsistencies. In Paper A, $x$ and $y$ are used for the channel input and output, respectively, whereas in Paper B and C, $x$ and $y$ are used to differentiate between polarizations. In Paper A, conditioning on particular realization of a random variable is denoted by $f_{Y|X=x}(y)$, whereas in the introductory part of the thesis we use $f_{Y|X}(y|x)$ for readability purposes. In Papers A and B, the spontaneous emission factor $n_{\mathrm{sp}}$ appeared in the context of distributed amplification, which should be replaced with the photon occupancy factor $K_T$ for Raman amplification [2]. Further, the additive noise power spectral density $N_0$ in Paper A is defined per unit length and this definition is inconsistent with the definition used in Part I of the thesis where it has units of [W/Hz]. We use a different font $\mathsf{N}_0$ in Part I to indicate the difference.

# CHAPTER 2

## Fiber-Optical Channel Modeling

A channel model is a mathematical description of the propagation medium and possibly also includes certain elements of the transmitter and receiver (e.g., filters). In the appended papers, the starting point for the analysis is a discrete memoryless channel, which can be specified in the form of a conditional PDF. This chapter is intended to describe the origin of the assumed PDFs and also to give the reader a broader picture about optical channel modeling in general.

We are concerned with coherent, long-haul (i.e., distances exceeding 2000 km) data transmission over single-mode fibers (SMFs) and the main challenge is a nonlinear effect caused by the relatively high signal power in relation to the small cross-section area of the fiber. Without going further into the physical details, a useful way to think about this effect is to imagine that the presence of an optical signal can compress the fiber material (in most cases silica) to such a degree that its propagation properties, in particular the refractive index, are changed in a nonlinear way [9, p. 18].

If nonlinear effects are ignored, an optical fiber can be regarded as a linear dispersive channel. Motivated by this, we start by reviewing some important results for this channel in Section 2.1. In Section 2.2, we discuss the NLSE, which is a deterministic channel model for an SMF. Multi-span links consisting of several SMFs and different amplification types are described in Section 2.3. In Section 2.4, the channel models assumed in the papers are introduced and compared.

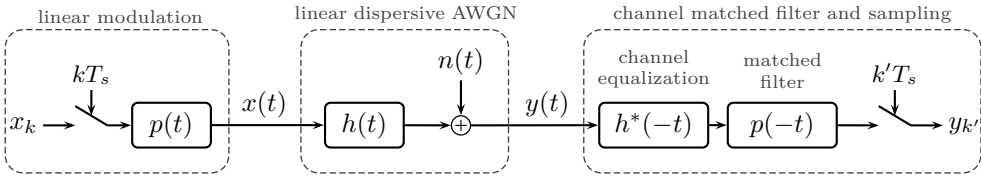



**Figure 2.1:** Block diagram of a system with linear modulation, a linear dispersive AWGN channel, and a receiver that obtains a sufficient statistic for optimal detection.

## 2.1 Linear Dispersive Channels

Consider the complex-valued linear dispersive additive white Gaussian noise (AWGN) channel

$$y(t) = h(t) \circledast x(t) + n(t), \tag{2.1}$$

where $x(t)$ is the baseband representation of the input signal, $y(t)$ is the output signal, $h(t)$ is the channel impulse response, and $n(t)$ is a circularly symmetric complex Gaussian stochastic process with zero mean with power spectral density (PSD) $\mathsf{N}_0$, i.e., $\mathbb{E}[N(t)N^*(t')] = \mathsf{N}_0\delta(t-t')$. Given an appropriate impulse response $h(t)$, equation (2.1) is widely used as a model for, e.g., wireline transmission over coaxial cables. A vast amount of literature exists on the linear dispersive AWGN channel and many important problems such as the optimal receiver/detector structure, the ultimate achievable transmission rates, or practical schemes that achieve these rates can be considered well studied and understood by now, see [15] and references therein.

In the following, we review some important concepts and results under the assumption that $h(t)$ is a unit gain all-pass filter, i.e., $|H(f)| = 1$, where $h(t) \circ\!\!-\!\!\bullet H(f)$. This assumption will turn out to be accurate for the fiber-optical channel later if nonlinear effects are ignored. Henceforth, the receiver is always assumed to have perfect knowledge about $h(t)$, obtained through an appropriate channel estimation technique. Furthermore, it is assumed that perfect carrier and timing synchronization between transmitter and receiver has been achieved.

In Fig. 2.1, a block diagram of the considered system is shown. We start with a linearly pulse-modulated[1] input signal

$$x(t) = \sum_k x_k p(t - kT_s), \tag{2.2}$$

where $x_k \in \mathbb{C}$ are the information symbols for $k \in \mathbb{Z}$, $p(t)$ is the real-valued pulse shape, and $T_s$ is the symbol period. The symbol rate is defined as $R_s = 1/T_s$. The input signal

[1] A different approach to communicate over linear dispersive channels is through multicarrier transmission, which we do not consider in this thesis.

$x(t)$ is assumed to be bandlimited to a bandwidth $W \geq 1/T_s$ and the power of $x(t)$ is

$$P \triangleq \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |x(t)|^2 \, \mathrm{d}t. \tag{2.3}$$

The receiver observes the filtered and noisy version of $x(t)$ according to (2.1) as

$$y(t) = h(t) \circledast x(t) + n(t) = \sum_k x_k \tilde{p}(t - kT_s) + n(t), \tag{2.4}$$

where $\tilde{p}(t) = h(t) \circledast p(t)$ is the convolution of the pulse shape with the channel impulse response. A sufficient statistic[2] for detecting the symbols $x_k$ based on $y(t)$ can be obtained by filtering $y(t)$ with the channel matched filter $\tilde{p}^*(-t)$ and sampling at time instances $t = k'T_s$, $k' \in \mathbb{Z}$ [16, Prop. 28.5.2]. Since $\tilde{p}^*(-t) = h^*(-t) \circledast p(-t)$, one may interpret the application of the channel matched filter as a two-step process. The first step (filtering with $h^*(-t)$) is referred to as "channel equalization" or simply "equalization". Due to the assumption that $h(t)$ is an all-pass filter, one may also think of applying $h^*(-t)$ as a compensation technique[3], in the sense that $h(t) \circledast h^*(-t) = \delta(t)$, which is readily seen by applying the Fourier transform and invoking the all-pass filter assumption, i.e., $H(f)H^*(f) = |H(f)|^2 = 1$. After applying the (pulse) matched filter $p(-t)$ and sampling, one finally obtains the discrete-time channel model [16, Prop. 28.5.2 (ii)]

$$y_{k'} = \sum_k x_k \mathsf{R}_p((k - k')T_s) + n_{k'}, \tag{2.5}$$

where $\mathsf{R}_p(t) = p(t) \circledast p(-t)$ is the self-similarity (or autocorrelation) function of the pulse shape [16, Def. 11.2.1] and $n_{k'}$ is a zero mean Gaussian random variable with $\mathbb{E}[N_k N_{k'}^*] = \mathsf{N}_0 \mathsf{R}_p((k - k')T_s)$. A discrete memoryless channel $y_k = x_k + n_k$ is obtained by choosing the pulse shape such that its self-similarity function $\mathsf{R}_p(t)$ satisfies Nyquist's criterion $\mathsf{R}_p((k-k')T_s) = \delta[k-k']$ [16, Def. 11.3.1], e.g., choosing $p(t)$ to be a root-raised cosine pulse with an arbitrary roll-off factor. In this case, the channel from $x_k$ to $y_k$ is completely characterized by the conditional PDF

$$f_{Y_k|X_k}(y_k|x_k) = \frac{1}{\pi \mathsf{N}_0} \exp\left(-\frac{|y_k - x_k|^2}{\mathsf{N}_0}\right). \tag{2.6}$$

On the other hand, if the Nyquist criterion is not fulfilled (or $h(t)$ is not an all-pass filter), the discrete-time channel model is not memoryless (one may write the first term on the right-hand side of (2.5) as a discrete convolution) and the noise samples are correlated. The optimal detection approach in that case is maximum likelihood sequence estimation (MLSE) either with [17] or without [18] the insertion of a noise whitening filter.

[2] See [16, Ch. 26] for a formal definition.

[3] The term "dispersion compensation" is often used instead of "equalization" in the context of fiber-optical communication systems.

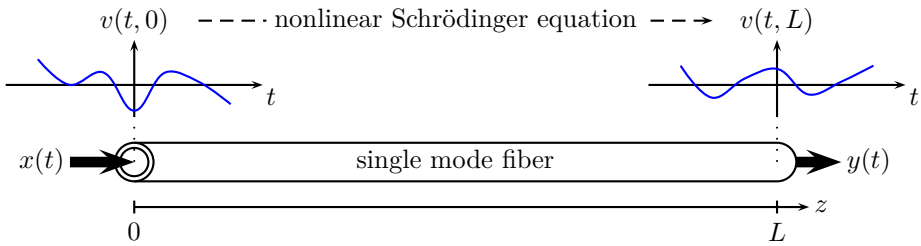



**Figure 2.2:** Conceptual representation of the signal evolution through a SMF. The NLSE describes the relationship between the input signal $x(t) = u(t, 0)$ and the output signal $y(t) = u(t, L)$.

The previous description is relevant to fiber-optical systems for several important reasons. The first one is that, while the actual fiber-optical channel model is nonlinear, the linear (pulse) modulation and receiver structure in Fig. 2.1 are nonetheless ubiquitously used in practical fiber-optical systems. Obviously, they are not necessarily optimal anymore[4], but their performance can still be analyzed and seen as a baseline. Secondly, the previous discussion illustrates how an originally continuous-time channel model (eq. (2.1)) can be simplified to a discrete-time model (eq. (2.5)) which in turn can then be used to study, e.g., detection algorithms or channel coding schemes. This is the approach taken in all the appended papers, where the analysis is based on discrete-time channel models.

## 2.2 The Nonlinear Schrödinger Equation

The starting point for coherent, long-haul fiber-optical channel modeling is the NLSE, which can be derived from the Maxwell equations under some assumptions that are appropriate for SMFs [20]. The NLSE is a partial differential equation that defines the input–output relationship for optical baseband signals[5] propagating through SMFs.

Let us first, in addition to the time parameter $t$, introduce a distance parameter $0 \leq z \leq L$ that denotes the propagation distance of the signal from the beginning of the fiber, where $L$ is the total length of the fiber. The baseband signal of interest is then a function of two parameters, denoted by $v(t, z)$. To be consistent with the previous notation, we define the input and output signals as $x(t) = v(t, 0)$ and $y(t) = v(t, L)$, i.e., $x(t)$ is the signal launched into the fiber at $z = 0$, and $y(t)$ is the signal received after propagating through an SMF of length $L$. This is conceptually illustrated in Fig. 2.2. Before we continue, we also define the instantaneous signal power $P(t, z) \triangleq |v(t, z)|^2$ and the power profile $P(z) \triangleq \lim_{T\to\infty}(\int_{-T}^{T} P(t, z)\,\mathrm{d}t)/(2T)$, where $P(0) = P$ is the power of the input

[4] The author in [19, p. 42] goes so far to say that these methods "borrowed from linear system theories, are inappropriate for communication over optical fiber networks".

[5] Often called "slowly varying envelope" in the literature. The carrier frequency is assumed to be the equivalent of a 1550 nm light wave, corresponding to roughly 193.4 THz.

signal.

The NLSE accounts for signal attenuation, chromatic dispersion, and nonlinear effects in an SMF and can be written as

$$\frac{\partial v(t,z)}{\partial z} = -\frac{\alpha}{2}v(t,z) - \jmath\frac{\beta_2}{2}\frac{\partial^2 v(t,z)}{\partial t^2} + \jmath\gamma v(t,z)|v(t,z)|^2, \tag{2.7}$$

where $\alpha$ is the attenuation coefficient, $\beta_2$ is the chromatic dispersion coefficient, and $\gamma$ is the nonlinear Kerr parameter. If we take into account only the first term on the right-hand side of (2.7), one obtains $v(t,z) = \exp(-\alpha z/2)v(t,0)$ as a solution[6], i.e., we immediately see that the signal amplitude in an SMF decays exponentially with the propagation distance. By defining a renormalized version of $v(t,z)$ as $u(t,z) \triangleq \exp(\alpha z/2)v(t,z)$ and substituting it into (2.7), one obtains an alternative and somewhat simpler version of the NLSE as [20, eq. (4)]

$$\frac{\partial u(t,z)}{\partial z} = -\jmath\frac{\beta_2}{2}\frac{\partial^2 u(t,z)}{\partial t^2} + \jmath\gamma e^{-\alpha z}u(t,z)|u(t,z)|^2. \tag{2.8}$$

Unfortunately, there are no closed-form solutions to the NLSE and one has to use numerical methods in order to solve (2.7) or (2.8).

We proceed by discussing two special cases of (2.8) in Sections 2.2.1 and 2.2.2, for $\gamma = 0$ and $\beta_2 = 0$, respectively. In both cases, a closed-form solution can be obtained. These solutions are also the key ingredients for one of the most widely used numerical methods to solve (2.8), namely, the split-step Fourier method (SSFM), which is described in Section 2.2.3.

### 2.2.1 Absence of Nonlinear Effects

As we will see, when nonlinear effects are ignored, the solution of (2.8) can be more conveniently written as the convolution of the input signal with a dispersive filter. For $\gamma = 0$, (2.8) becomes

$$\frac{\partial u(t,z)}{\partial z} = -\jmath\frac{\beta_2}{2}\frac{\partial^2 u(t,z)}{\partial t^2}, \tag{2.9}$$

which can be solved analytically by first transforming (2.9) into the Fourier domain[7] using the correspondence $\frac{\partial^n}{\partial t^n}x(t) \circ\!\!-\!\!\bullet (\jmath 2\pi f)^n X(f)$ to obtain [20]

$$\frac{\partial U(f,z)}{\partial z} = -\jmath\frac{\beta_2}{2}(\jmath 2\pi f)^2 U(f,z), \tag{2.10}$$

[6] Recall that the solution of $\partial f(z)/\partial z = cf(z)$ is given by $f(z) = \exp(cz)f(0)$.

[7] The solution of (2.9) is sometimes immediately written as $u(t,z) = \exp(z\hat{D})u(t,0)$ with $\hat{D} = -\jmath\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}$. The operator $\exp(z\hat{D})$ may be interpreted with the help of the Taylor expansion of the exponential function around 0, i.e., using $e^x = 1 + x + x^2/2 + \cdots$ with $x = z\hat{D}$.

where $U(f,z) \bullet\!\!-\!\!\circ\, u(t,z)$. The solution of (2.10) can be verified to be

$$U(f,z) = \exp\left(\jmath 2\beta_2\pi^2 f^2 z\right) U(f,0) \tag{2.11}$$

$$= H(f,z)U(f,0), \tag{2.12}$$

where $H(f,z) = \exp(\jmath 2\beta_2\pi^2 f^2 z)$ can be seen as the frequency response of a ($z$-dependent) dispersive filter. Applying the inverse Fourier transform to (2.12) finally leads to the expression

$$u(t,z) = h(t,z) \circledast u(t,0), \tag{2.13}$$

where $h(t,z) = \exp\left(\jmath t^2/(2\beta_2 z)\right)/\sqrt{\jmath 2\pi\beta_2 z} \circ\!\!-\!\!\bullet\, H(f,z)$ is the impulse response of the filter. We further have $|H(f,z)| = 1$. In summary, in the absence of nonlinear effects, chromatic dispersion manifests itself as a unit-gain all-pass filter.

### 2.2.2 Absence of Dispersion

Another special case that allows for an exact and explicit solution of the NLSE is when chromatic dispersion is completely ignored. In this case, i.e., for $\beta_2 = 0$, (2.8) becomes

$$\frac{\partial u(t,z)}{\partial z} = \jmath\gamma e^{-\alpha z}|u(t,z)|^2 u(t,z). \tag{2.14}$$

One may then verify that the solution to (2.14) is given by [20, eq. (17)]

$$u(t,z) = u(t,0)e^{\jmath\gamma L_{\text{eff}}(z)|u(t,0)|^2}, \tag{2.15}$$

where

$$L_{\text{eff}}(z) \triangleq \int_0^z e^{-\alpha z'}\,\mathrm{d}z' = \frac{1-\exp(-\alpha z)}{\alpha} \tag{2.16}$$

is the effective nonlinear length, where $L_{\text{eff}}(z) \le z$ with $L_{\text{eff}}(z) \to z$ as $\alpha \to 0$.

From (2.15), we see that, for a given $z$, the nonlinear effect by itself causes a phase-shift of the signal that is proportional to the instantaneous power $|u(t,0)|^2$, whereas the amplitude $|u(t,z)| = |u(t,0)|$ remains unchanged. This effect is called self-phase modulation and an important consequence of its nonlinear nature is that the bandwidth of $u(t,z)$ may grow during propagation through the fiber.

### 2.2.3 Split-Step Fourier Method

As was mentioned earlier, the NLSE with $\gamma \neq 0$ and $\beta_2 \neq 0$ cannot be solved analytically and one has to resort to numerical methods in order to obtain the relationship between the input and output signals. A popular and computationally efficient numerical method is the SSFM which we describe in the following.[8]

[8] There exist several different versions of the SSFM and the one presented here is referred to as asymmetric and non-iterative. For more details we refer the reader to [9, Sec. 2.4.1].

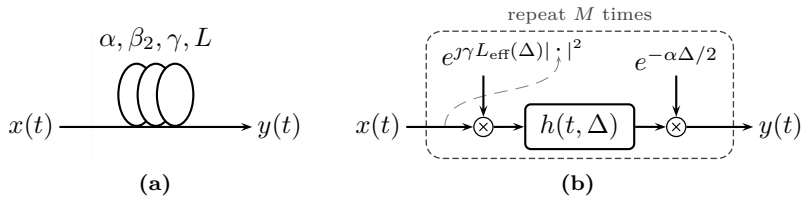



**Figure 2.3:** Symbolic representation of an SMF in (a) and an (approximate) mathematical model via the SSFM in (b). The notation $|\cdot|^2$ stands for the instantaneous power of the signal arriving at the corresponding multiplication block as indicated by the dashed, gray line.

Conceptually, we start by discretizing the spatial dimension and subdividing the entire fiber of length $L$ into small segments of length $\Delta$, where $M = L/\Delta \in \mathbb{N}$ is the total number of segments. For the $i$th segment, $1 \leq i \leq M$, the input signal is denoted by $u(t,(i-1)\Delta)$ and the corresponding output signal by $u(t,i\Delta)$. It is assumed that an approximate solution to obtain $u(t,i\Delta)$ based on $u(t,(i-1)\Delta)$ is given by first applying (2.15) and then (2.13), i.e., for small $\Delta$, we assert that

$$u(t,i\Delta) \approx h(t,\Delta) \circledast \left( u(t,(i-1)\Delta) e^{\jmath\gamma L_{\text{eff}}(\Delta)|u(t,(i-1)\Delta)|^2} \right). \tag{2.17}$$

Then, an approximate solution for an entire SMF of length $L$ is given by repeatedly applying (2.17), starting with the first segment $i = 1$, i.e., with the input signal $u(t,0) = x(t)$. The SSFM step in (2.17) is given in terms of the normalized signal $u(t,z)$ and to incorporate the signal attenuation, the output signal $u(t,i\Delta)$ is multiplied by $\exp(-\alpha\Delta/2)$ to obtain $v(t,i\Delta)$ after each step. The resulting numerical method is shown in terms of a block diagram in Fig. 2.3. In the figure, the notation $|\cdot|^2$ stands for the instantaneous power of the signal that arrives at the corresponding multiplication block (e.g., $|x(t)|^2$ in the first segment, $|u(t,\Delta)\exp(-\alpha\Delta/2)|^2$ in the second, and so on). It has been shown that the above method converges to the true solution for $\Delta \to 0$ [9, p. 42]. Practical guidelines on the choice of the segment size are developed in [21].

The name of the method originates from the fact that the nonlinear phase-shift operation and the linear filtering in Fig. 2.3(b) are commonly carried out in the time and frequency domain, respectively. Therefore, one forward and one inverse Fourier transform have to be performed per segment. In computer implementations, a sampled version of $x(t)$ is considered which facilitates the application of the computationally efficient fast Fourier transform (FFT). Such an implementation is for example provided in [9, App. B].

## 2.3 Optical Amplification and Noise

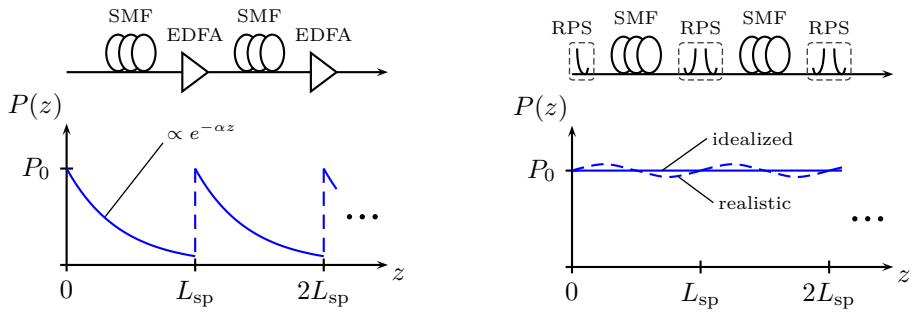
The numerical value of the attenuation coefficient $\alpha$ is typically between 0.2 and 0.4 dB/km. Assuming $\alpha = 0.2$ dB/km and a transmission distance of $L = 2000$ km, the input signal would be attenuated by 400 dB implying that $y(t)$ is practically zero [2, Sec. IX-B]. It is therefore necessary to amplify the signal along the transmission path, which invariably introduces noise into the system.

In this section, we briefly discuss two types of amplification, lumped and distributed, in terms of their effect on the power profile of the signal and the type of noise that they introduce. Modeling the power profile is important due to the dependency of the nonlinear effect on the instantaneous signal power. Thus, one cannot simply ignore attenuation effects and make a link budget analysis as is common for linear channels. Details about the underlying physical aspects of optical amplification can be found in standard textbooks on optical data transmission, e.g., [22, Ch. 6]. It should, however, be pointed out that the optical amplifier noise is in fact the dominant source of noise in long-haul systems meaning that noise from other sources, for example thermal noise from electrical components, is negligible in comparison and can therefore be ignored [2, Sec. IX-A].

To account for amplification and noise, the NLSE (2.7) can be extended by inserting a gain profile $g(z)$ and a complex-valued stochastic process $w(t, z)$, resulting in

$$\frac{\partial v(t,z)}{\partial z} = -\frac{\alpha - g(z)}{2} v(t,z) - \jmath \frac{\beta_2}{2} \frac{\partial^2 v(t,z)}{\partial t^2} + \jmath \gamma v(t,z)|v(t,z)|^2 + w(t,z). \tag{2.18}$$

Equation (2.18) is referred to as the stochastic nonlinear Schrödinger equation (sNLSE) [23]. We first discuss the gain profile $g(z)$ and its effect on the power profile of the signal $v(t, z)$, ignoring all other effects (including $w(t, z)$). Both amplification types are applied periodically, in the sense that the entire transmission distance $0 \leq z \leq L$ is split up into spans of length $L_{\text{sp}}$, varying between 60 and 120 km, where $N_{\text{sp}} = L/L_{\text{sp}} \in \mathbb{N}$ is the total number of spans. In the case of lumped amplification, an optical amplifier, most often an erbium-doped fiber amplifier (EDFA) [2, Sec. IX-B], is inserted after each span, where the amplifier gain $G$ matches the power loss of the signal in that span, i.e., $G = e^{\alpha L_{\text{sp}}}$. In (2.18), this is accounted for by setting $g(z) = \alpha L_{\text{sp}} \sum_{i=1}^{N_{\text{sp}}} \delta(z - iL_{\text{sp}})$. The corresponding power profile is illustrated in Fig. 2.4(a). The signal power decreases exponentially according to the loss coefficient $\alpha$ and is periodically restored to the input power $P$ after each span. In the case of distributed amplification, it is assumed that the signal power can be held at an approximately constant level as shown in Fig. 2.4(b). In order to achieve this, pump waves are launched into the fiber at Raman pump stations (RPSs) which are located at the beginning and after each span [2, Sec. IX-B]. The pump waves co-propagate together with the signal $v(t, z)$ and the nonlinear nature of the fiber is exploited to continuously transfer energy from the pump wave to the signal. In Fig. 2.4(b), the "realistic" power profile (dashed line) is schematically reproduced from [24, Fig. 3] and assumes two pump waves per span, one propagating co-directionally



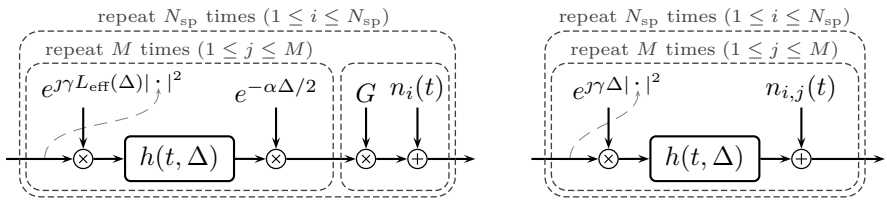

**(a)** lumped **(b)** distributed

**Figure 2.4:** Schematic comparison of the power profile as a function of the transmission distance $z$ for the two considered amplification types.

and one contra-directionally to the signal. From a modeling perspective, the "idealized" constant profile (solid line) is assumed for simplicity, where $g(z) = \alpha$ and hence the first term on the right-hand side of (2.18) simply disappears.

Next, we discuss the noise that is generated by the optical amplification schemes through a process called amplified spontaneous emission (ASE). For lumped amplification, noise can be thought of as being added to the signal at discrete locations $z_i \triangleq iL_{\mathrm{sp}}$, $1 \leq i \leq N_{\mathrm{sp}}$. Thus, if $v(t, z_i^-)$ is the output signal after the $i$th fiber span, the input signal to the next span is given by $v(t, z_i^+) = Gv(t, z_i^-) + n_i(t)$, where $n_i(t)$ is the additive noise originating from the $i$th amplifier [19, p. 36]. It has been experimentally verified that ASE noise can be accurately modeled as circularly symmetric complex Gaussian [2, p. 667] and therefore it remains to specify the autocorrelation function of $n_i(t)$, where processes from different amplifiers are uncorrelated. The most common assumption is white Gaussian noise, i.e., $\mathbb{E}[N_i(t)N_j^*(t')] = \mathsf{N}_\ell \delta(t-t')\delta[i-j]$, where the noise PSD per amplifier for EDFAs is computed as $\mathsf{N}_\ell = (G-1)h\nu_s n_{\mathrm{sp}}$ [2, eq. (54)]. The meaning and values of the quantities appearing in this expression are summarized in Table 2.1 at the end of this section. We further set $\mathsf{N}_0 = N_{\mathrm{sp}}\mathsf{N}_\ell$ in the case of lumped amplification, which one might think of as the cumulative PSD at the end of the transmission link for $N_{\mathrm{sp}}$ amplifiers. Since temporally white noise has infinite instantaneous power, this assumption would, however, lead to infinite phase rotations due to the nonlinear effect. In reality, the noise power is of course finite, and the PSD of ASE noise is comparable to the gain spectrum of the amplifier. For an idealized EDFA that provides flat gain over a certain frequency range $W_N$, one would then replace $\delta(t-t')$ with $\delta_{W_N}(t-t')$ where $\delta_{W_N}(x) = W_N \mathrm{sinc}(W_N x)$ [19]. Further limitations of the optical bandwidth can occur due to the insertion of optical bandpass filters and/or reconfigurable optical add-drop multiplexers (ROADMs) along the transmission line [2].

Based on the previous description, a block diagram of a continuous-time model describing a multi-span transmission link with lumped amplification is depicted in Fig. 2.5(a).




**(a)** lumped **(b)** distributed

**Figure 2.5:** Block diagram for a multi-span link including amplification and noise for (a) lumped and (b) distributed amplification.

The model consists of the concatenation of the deterministic model for an SMF based on the SSFM (cf. Fig. 2.3(b)) with a multiplicative gain and additive noise for the optical amplifier. For completeness, we also indicate how the additive noise terms $n_i(t)$ can be related to $w(t,z)$ in (2.18) for lumped amplification. Note that if we neglect all terms on the right-hand side of (2.18) except $w(t,z)$, we have $\partial v(t,z)/\partial z = w(t,z)$ and integrating this equation leads to

$$v(t,z) = v(t,0) + \int_0^z w(t,\xi)\,\mathrm{d}\xi = v(t,0) + n(t,z). \tag{2.19}$$

Here, $n(t,z)$ represents the noise that is added to the signal up to a certain distance $z$. For lumped amplification, one may set $w(t,z) = \sum_{i=1}^{N_{\mathrm{sp}}} n_i(t)\delta(z - iL_{\mathrm{sp}})$ [20, p. 84], so that $n(t,z) = \sum_{i=1}^{\lfloor z/L_{\mathrm{sp}} \rfloor} n_i(t)$ corresponds the addition of all $n_i(t)$ up to distance $z$ (the upper integral limit in (2.19) is interpreted as $z^+$).

Next, we discuss distributed amplification, where ASE noise is continuously added throughout the entire transmission link. A common assumption is that $w(t,z)$ is a white Gaussian stochastic process in both time and space, and hence [2, eq. (53)] [19, p. 37]

$$\mathbb{E}\left[W(t,z)W^*(t',z')\right] = \bar{\mathsf{N}}_{\mathrm{d}}\delta(t-t')\delta(z-z'), \tag{2.20}$$

where $\bar{\mathsf{N}}_{\mathrm{d}}$ is the distributed PSD per unit length (in [W/km/Hz]) computed as $\bar{\mathsf{N}}_{\mathrm{d}} = \alpha h\nu_s K_T$, where $K_T$ is the photon occupancy factor. Similarly as for lumped amplification, we set $\mathsf{N}_0 = L\bar{\mathsf{N}}_{\mathrm{d}}$ as the PSD at the end of the entire transmission line of length $L$. The expressions for $\mathsf{N}_0$ for the two amplification types are related via $N_{\mathrm{sp}} = L/L_{\mathrm{sp}}$ and letting $L_{\mathrm{sp}} \to 0$, and replacing $n_{\mathrm{sp}}$ with $K_T$. Regarding the temporal correlation of $w(t,z)$, one can make similar arguments as for the lumped case and replace $\delta(t-t')$ with $\delta_{W_N}(t-t')$ for some $W_N$ to account for the bandwidth limitation of physically realistic noise. Under the assumption that $w(t,z)$ is uncorrelated in space, $n(t,z)$ in (2.19) is a Wiener process, i.e., the integral of a white Gaussian process. As pointed out in [25, Sec. III], the sNLSE then has to be interpreted via an equivalent integral representation (similar to (2.19) but including all terms of (2.18)), since a Wiener process is

not differentiable for any $z$ to satisfy $\partial n(t,z)/\partial z = w(t,z)$. Here, we will ignore such issues and rely on some intuition to describe how the Wiener process can be included in the SSFM to numerically solve the sNLSE with distributed noise. In particular, we assert that the distributed nature of the noise can be accurately captured by adding a Gaussian stochastic process *in each segment* of the SSFM. We denote the process that is added in the $j$th segment of the $i$th fiber span by $n_{i,j}(t)$, where processes from different segments and spans are uncorrelated. In order for all $MN_{\mathrm{sp}}$ processes $n_{i,j}(t)$ to produce the same PSD as the Wiener process at the end of the link, i.e., $\mathsf{N}_0 = L\bar{\mathsf{N}}_{\mathrm{d}}$, we require $\mathbb{E}[N_{j,i}(t)N^*_{j,i}(t')] = L\bar{\mathsf{N}}_{\mathrm{d}}/(MN_{\mathrm{sp}})\delta(t-t')$. The corresponding block diagram is shown in Fig. 2.5(b). Compared to the lumped amplification case in Fig. 2.5(a), it can be seen that the attenuation step is removed (also $L_{\mathrm{eff}}(\Delta) = \Delta$, since $\alpha = 0$) and an additive noise term is included in each segment.

All relevant quantities that have been described in this section for the two amplification types are summarized in Table 2.1.

### 2.3.1 Linear Regime

We briefly discuss the case where $\gamma = 0$ but now including ASE noise from the optical amplification schemes. To that end, consider again the two block diagrams in Fig. 2.5(a) and (b). If the nonlinear phase rotations are removed, the model indeed reverts to the linear dispersive AWGN channel (2.1) (see also Fig. 2.1) for both amplification types. To see this, first note that the attenuation and gain factors in Fig. 2.5(a) cancel out due to the linearity of the model. Further, due to the all-pass nature of the dispersive filters, one may freely rearrange the additive noise terms because filtered noise remains Gaussian with the same PSD. Thus, we may assume that all noise processes are added together at the end of the transmission link and the model in fact corresponds to the linear dispersive AWGN channel in both cases [20, Sec. 5.1]. We have $h(t) = h(t,L)$ (convolving $h(t,\Delta)$ $MN_{\mathrm{sp}}$ times with itself) and $n(t) = \sum_{i=1}^{N_{\mathrm{sp}}} n_i(t)$ and $n(t) = \sum_{i=1}^{N_{\mathrm{sp}}}\sum_{j=1}^{M} n_{i,j}(t)$ for lumped and distributed amplification, respectively.

## 2.4 Channel Models in the Appended Papers

### 2.4.1 Paper A: Zero-Dispersion Fiber, Single Polarization

In Paper A, we consider the special case where dispersive effects are absent, i.e., $\beta_2 = 0$, in combination with a distributed amplification scheme. In Section 2.2.2, it has already been shown that it is possible to find an analytical solution for the (deterministic) NLSE if $\beta_2 = 0$. When ASE noise from optical amplifiers is also considered, the signal and noise interact through the fiber nonlinearity and give rise to the phenomenon of nonlinear phase noise (NLPN). Fortunately, it turns out that an exact analytical characterization of a discrete-time channel model can still be found. We proceed by first describing how the discrete-time channel is obtained from the continuous-time channel. Based on this

**Table 2.1:** Comparison of amplification types

| quantity | meaning | lumped | distributed |
|---|---|---|---|
| $g(z)$ | gain profile | $\alpha L_{\mathrm{sp}} \sum_{i=1}^{N_{\mathrm{sp}}} \delta(z - iL_{\mathrm{sp}})$ | $\alpha$ |
| $w(t,z)$ | added to $\partial v(t,z)/\partial z$ | $\sum_{i=1}^{N_{\mathrm{sp}}} n_i(t)\delta(z - iL_{\mathrm{sp}})$ | white Gaussian |
| $n(t,z)$ | $\int_0^z w(t,\xi)\,\mathrm{d}\xi$ | $\sum_{i=1}^{\lfloor z/L_{\mathrm{sp}} \rfloor} n_i(t)$ | Wiener process |
| $\mathbb{E}\left[W(t,z)W^*(t',z')\right]$ | - | $\mathsf{N}_\ell \delta(t-t') \sum_{i=1}^{N_{\mathrm{sp}}} \delta(z - iL_{\mathrm{sp}})\delta(z' - iL_{\mathrm{sp}})$ | $\bar{\mathsf{N}}_{\mathrm{d}} \delta(t-t')\delta(z-z')$ |
| $\mathbb{E}[N(t,z)N^*(t',z')]$ | - | $\mathsf{N}_\ell \delta(t-t') \min(\lfloor z/L_{\mathrm{sp}} \rfloor, \lfloor z'/L_{\mathrm{sp}} \rfloor)$ | $\bar{\mathsf{N}}_{\mathrm{d}} \delta(t-t') \min(z,z')$ |
| $\mathsf{N}_0$ | PSD after distance $L$ | $N_{\mathrm{sp}}\mathsf{N}_\ell = N_{\mathrm{sp}}(e^{\alpha L_{\mathrm{sp}}} - 1)h\nu n_{\mathrm{sp}}$ | $L\bar{\mathsf{N}}_{\mathrm{d}} = N_{\mathrm{sp}} L_{\mathrm{sp}} \alpha h\nu K_T$ |
| $F_n$ | amplifier noise figure | typically 4–7 dB | - |
| $n_{\mathrm{sp}}$ | spontaneous emission factor | $F_n(1 - G^{-1})^{-1}/2$ | - |
| $K_T$ | photon occupancy factor | - | $\approx 1.13$ (at room temp.) |
| $h$ | Planck's constant | $6.626 \cdot 10^{-34}$ [Js] | |
| $\nu_s$ | optical carrier frequency | $1.936 \cdot 10^{14}$ [Hz] | |

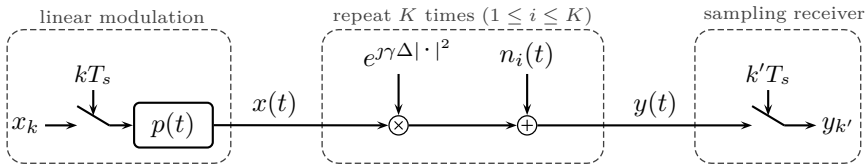



**Figure 2.6:** Block diagram of the transmission system considered in Paper A. Note that the absence of a filter front-end at the receiver means that the noise bandwidth is implicitly assumed to be limited due to inline optical filters.

discrete-time channel, an analytical expression for the conditional PDF is presented, which is essentially the starting point for the discussions in Paper A.

For the case of distributed amplification and in the absence of dispersion, the sNLSE (2.18) reduces to

$$\frac{\partial v(t,z)}{\partial z} = \jmath\gamma v(t,z)|v(t,z)|^2 + w(t,z). \tag{2.21}$$

A block diagram of the corresponding continuous-time channel model (via the SSFM) is then derived from Fig. 2.5(b) by simply removing the dispersive filters. A discrete-time channel model can be obtained by assuming a conventional linear pulse modulation (cf. Fig. 2.1) in combination with a sampling receiver. The resulting block diagram depicting this scenario is shown in Fig. 2.6. For simplicity, and to make the notation consistent with Paper A, we have replaced the double enumeration over segments and spans with a single enumeration over $1 \leq i \leq K = MN_{\mathrm{sp}}$.

Due to the absence of a filter prior to sampling the received signal $y(t)$ (see Fig. 2.6), it is implicitly assumed that the noise bandwidth $W_N$ is somehow limited during propagation, e.g., through inline optical filters (otherwise the samples $y_{k'}$ would have infinite variance) [26]. Furthermore, due to the absence of a matched filter (matched to the pulse shape), the pulse $p(t)$ itself rather than the self-similarity function of the pulse $\mathsf{R}_p(t)$ needs to fulfill the Nyquist criterion $p(kT_s) = \delta[k]$, e.g., one may assume sinc pulses.

It is important to point out that the step from continuous-time to discrete-time is not necessarily optimal, i.e., the samples $y_{k'}$ do not necessarily form a sufficient statistic for detecting $x_k$ based on $y(t)$. Therefore, statements about optimality (e.g., "ML detection") are implicitly understood with respect to the discrete-time channel only, not with respect to the actual waveform channel.

The channel from the transmitted symbols to the received samples is memoryless and the indices can hence be dropped, i.e., $x$ and $y$ denote the (complex-valued) channel input and output, respectively. The system model presented thus far has been extensively studied in the literature and there exist several different derivations of the conditional PDF $f_{Y|X}(y|x)$, some of which we mention in the following in chronological order.

- Gordon and Mollenauer were the first to recognize that the interaction between

the signal and inline amplifier noise leads to a phase noise effect [27] and NLPN is sometimes referred to as the Gordon–Mollenauer effect.

- Their model was made rigorous by Mecozzi in [23] and later in [28].
- In [29], Turitsyn *et al.* proposed a technique to derive the PDF based on the Martin–Siggia–Rose formalism from statistical mechanics. It was also recognized that the capacity of the discrete channel grows unbounded with input power.
- Ho provided an in-depth treatment on the subject of NLPN including a derivation of the PDF based on characteristic functions [10, Ch. 5], see also [30].
- Most recently, Yousefi and Kschischang have considered the zero-dispersion case in a series of papers [25, 26, 31, 32]. They provide several additional derivations of the PDF based on a sum-product and a Fokker–Planck differential equation approach.

The PDF is given as follows. Adopting a polar notation for the channel input and output including a magnitude normalization with $\sigma_{\mathrm{d}}^2 = L\bar{\mathsf{N}}_{\mathrm{d}}W_N = \mathsf{N}_0 W_N$ according to $x/\sigma_{\mathrm{d}} = r_0 e^{\jmath\theta_0}$ and $y/\sigma_{\mathrm{d}} = re^{\jmath\theta}$, one may write the PDF in the form of a Fourier series as[9]

$$f_{\Theta,R|\Theta_0,R_0}(\theta, r|\theta_0, r_0) = \frac{1}{2\pi}\sum_{k\in\mathbb{Z}} C_k(r, r_0)e^{\jmath k(\theta-\theta_0)}, \tag{2.22}$$

where the Fourier coefficients are given by

$$C_k(r, r_0) = \frac{2z_k}{\sin z_k} r \exp\left(-\frac{r^2 + r_0^2}{(\tan z_k)/z_k}\right) \mathrm{I}_k\left(\frac{2z_k}{\sin z_k} r r_0\right), \tag{2.23}$$

where $\mathrm{I}_k(\,\cdot\,)$ is the modified Bessel function of the first kind and $z_k = \sqrt{\jmath\gamma k\sigma_{\mathrm{d}}^2}$. The singularities for $z_0 = 0$ are understood to be resolved as $(\sin 0)/0 = 1$ and $(\tan 0)/0 = 1$. Since the PDF is a real function, the symmetry condition $C_k(r, r_0) = C_{-k}(r, r_0)$ holds.

The zero-dispersion assumption can be motivated by the fact that the dispersion coefficient $\beta_2$ can be physically engineered to take on values over a certain range including (approximately) zero. However, this scenario is generally assumed to be unrealistic due to the severe spectral broadening that may occur during propagation. For example, in [2] the authors state that "the zero dispersion region is generally to be avoided as the effects of fiber nonlinearity are enhanced dramatically". Similarly, in [20] it is noted that "the zero-dispersion regime is not practical for communications". A discussion about this topic can also be found in [26, Sec. VIII].

The model can nonetheless be useful since the assumption of zero dispersion is sometimes fulfilled, at least approximately, in dispersion-managed (DM) transmission links.

[9] The joint density of the magnitude and phase of a complex random variable with density $f_Y(y)$ can be obtained via $f_{\Theta,R}(\theta, r) = r f_Y(re^{\jmath\theta})$ [16, Lem. 17.3.5].

These links consist of the concatenation of a standard SMF with a dispersion compensating fiber (DCF) whose dispersion coefficient $\beta_2'$ has been engineered to be of opposite sign, i.e., $\operatorname{sgn}(\beta_2') = -\operatorname{sgn}(\beta_2)$. If the length of the DCF $L_{\mathrm{sp}}'$ is matched to the length of the standard fiber $L_{\mathrm{sp}}$ according to the simple linear relation $|\beta_2|L_{\mathrm{sp}} = |\beta_2'|L_{\mathrm{sp}}'$, then the net effect of the dispersion after distance $L_{\mathrm{sp}}+L_{\mathrm{sp}}'$ is zero (neglecting nonlinear effects). To see this, consider the concatenation of two dispersive filters $H(f, L_{\mathrm{sp}}) = \exp(\jmath 2\beta_2\pi^2 f^2 L_{\mathrm{sp}})$ (cf. (2.10)) and $H'(f, L_{\mathrm{sp}}') = \exp(\jmath 2\beta_2'\pi^2 f^2 L_{\mathrm{sp}}')$ which exactly cancel in this case. If nonlinear effects are taken into consideration for such DM links, the accuracy of the zero-dispersion model depends essentially on the symbol rate $R_s$ and lengths $L_{\mathrm{sp}}$ and $L_{\mathrm{sp}}'$. In general, the lower the symbol rate and the shorter the fibers, the more accurate is the model.

### 2.4.2 Paper B: Zero-Dispersion Fiber, Polarization Multiplexing

In Paper B, we study an extension of the model used in Section 2.4.1 which was recently derived in [33]. The model takes into account PM transmission, where both polarizations of the light are used to transmit data. For PM transmission, the sNLSE equation can be further extended by considering the vector signal $\boldsymbol{v}(t,z) = (v_{\mathsf{a}}(t,z), v_{\mathsf{b}}(t,z))^{\mathsf{T}}$, where the indices indicate the signals in the two polarizations $\mathsf{a}$ and $\mathsf{b}$.[10] The resulting equation is referred to as the Manakov equation including loss and gain terms and amplifier noise and is given by [34, p. 8]

$$\frac{\partial \boldsymbol{v}(t,z)}{\partial z} = -\frac{\alpha - g(z)}{2}\boldsymbol{v}(t,z) - \jmath\frac{\beta_2}{2}\frac{\partial^2 \boldsymbol{v}(t,z)}{\partial t^2} + \jmath\gamma \boldsymbol{v}(t,z)\|\boldsymbol{v}(t,z)\|^2 + \boldsymbol{w}(t,z), \tag{2.24}$$

where $\boldsymbol{w}(t,z) = (w_{\mathsf{a}}(t,z), w_{\mathsf{b}}(t,z))^{\mathsf{T}}$ are two (independent) stochastic processes describing the ASE noise generated in both polarizations. The major difference between (2.24) and (2.18) is that (2.24) models the nonlinearity that is due to the sum of the instantaneous power in both polarizations $\|\boldsymbol{v}(t,z)\|^2 = P_{\mathsf{a}}(t,z)^2 + P_{\mathsf{b}}(t,z)^2$. We should mention that (2.18) ignores the fact that amplifier noise is always generated "in two polarizations", i.e., even if we assume one of the two signals in $\boldsymbol{v}(t,z)$ to be zero (as was done in Section 2.3), technically the amplifier noise in that polarization still contributes via (2.24) through the fiber nonlinearity.

The derivation presented in [33] makes similar assumptions as in, e.g., [10, 23, 26, 29] for the continuous-to-discrete time conversion and the subsequent analysis. In particular, dispersive effects are ignored. This also includes polarization mode dispersion (PMD), which would cause different group velocities of the signals in polarization $\mathsf{a}$ and $\mathsf{b}$ because of natural imperfections and asymmetries of the fiber cross-section area. The simplified Manakov equation for distributed amplification and neglecting all dispersive effects is

[10] This nonstandard notation for the polarizations is an attempt to avoid confusion with the transmit and received signals. However, we acknowledge inconsistent notation with respect to Paper B, where the polarizations are denoted by $x$ and $y$.

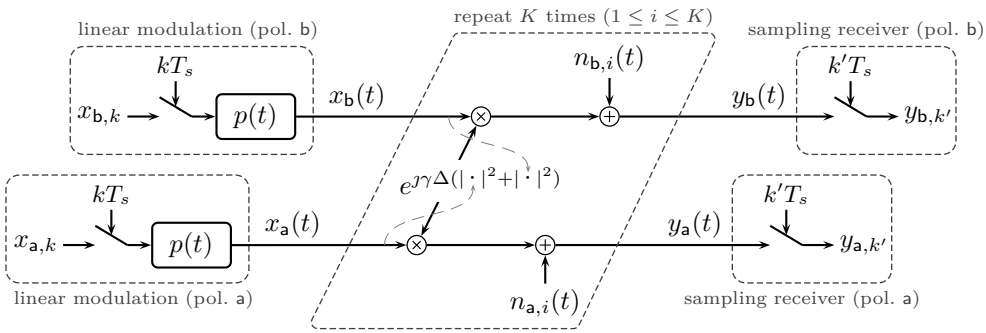



**Figure 2.7:** Block diagram of the transmission system considered in Paper B.

obtained from (2.24) as

$$\frac{\partial \boldsymbol{v}(t,z)}{\partial z} = \jmath\gamma \boldsymbol{v}(t,z)\|\boldsymbol{v}(t,z)\|^2 + \boldsymbol{w}(t,z). \tag{2.25}$$

Again, noise is assumed to be bandlimited and a discrete channel is obtained based on samples that are taken at the receiver at a rate $1/T_s$ in both polarizations. A block diagram of the assumed transmission system including a graphical representation of (2.25) via the SSFM is shown in Fig. 2.7. It can be seen that there are essentially two SP transmission systems which are now coupled due to the fiber nonlinearity (and hence uncoupled when $\gamma = 0$).

As for the SP case, the discrete-time model is memoryless. Dropping the discrete-time indices, the joint input and output in both polarizations are denoted by $\mathbf{y} = (y_{\mathsf{a}}, y_{\mathsf{b}})^{\mathsf{T}}$ and $\mathbf{x} = (x_{\mathsf{a}}, x_{\mathsf{b}})^{\mathsf{T}}$, respectively. A notation based on polar coordinates with the same normalization as before is preferable and we therefore set $x_{\mathsf{a}}/\sigma_{\mathrm{d}} = r_{0\mathsf{a}}e^{\jmath\theta_{0\mathsf{a}}}$, $x_{\mathsf{b}}/\sigma_{\mathrm{d}} = r_{0\mathsf{b}}e^{\jmath\theta_{0\mathsf{b}}}$, $y_{\mathsf{a}}/\sigma_{\mathrm{d}} = r_{\mathsf{a}}e^{\jmath\theta_{\mathsf{a}}}$, $y_{\mathsf{b}}/\sigma_{\mathrm{d}} = r_{\mathsf{b}}e^{\jmath\theta_{\mathsf{b}}}$, and collect the corresponding transmitted and received magnitudes and phases in the vectors $\boldsymbol{r}_0 = (r_{0\mathsf{a}}, r_{0\mathsf{b}})^{\mathsf{T}}$, $\boldsymbol{r} = (r_{\mathsf{a}}, r_{\mathsf{b}})^{\mathsf{T}}$, $\boldsymbol{\theta}_0 = (\theta_{0\mathsf{a}}, \theta_{0\mathsf{b}})^{\mathsf{T}}$, $\boldsymbol{\theta} = (\theta_{\mathsf{a}}, \theta_{\mathsf{b}})^{\mathsf{T}}$. Using this notation, the conditional PDF is given in the form of a two-dimensional Fourier series as

$$f_{\boldsymbol{\Theta},\boldsymbol{R}|\boldsymbol{\Theta}_0,\boldsymbol{R}_0}(\boldsymbol{\theta},\boldsymbol{r}|\boldsymbol{\theta}_0,\boldsymbol{r}_0) = \frac{1}{4\pi^2}\sum_{k_{\mathsf{a}}\in\mathbb{Z}}\sum_{k_{\mathsf{b}}\in\mathbb{Z}} C_{\boldsymbol{k}}(\boldsymbol{r},\boldsymbol{r}_0)e^{\jmath\boldsymbol{k}(\boldsymbol{\theta}-\boldsymbol{\theta}_0)}, \tag{2.26}$$

where $\boldsymbol{k} = (k_{\mathsf{a}}, k_{\mathsf{b}})^{\mathsf{T}}$ and the Fourier coefficients are

$$C_{\boldsymbol{k}}(\boldsymbol{r},\boldsymbol{r}_0) = \left(\frac{2z_{\boldsymbol{k}}}{\sin z_{\boldsymbol{k}}}\right)^2 r_{\mathsf{a}}r_{\mathsf{b}} \exp\left(-\frac{\|\boldsymbol{r}\|^2+\|\boldsymbol{r}_0\|^2}{(\tan z_{\boldsymbol{k}})/z_{\boldsymbol{k}}}\right) \mathrm{I}_{k_{\mathsf{a}}}\left(\frac{2z_{\boldsymbol{k}}}{\sin z_{\boldsymbol{k}}} r_{\mathsf{a}}r_{0\mathsf{a}}\right) \mathrm{I}_{k_{\mathsf{b}}}\left(\frac{2z_{\boldsymbol{k}}}{\sin z_{\boldsymbol{k}}} r_{\mathsf{b}}r_{0\mathsf{b}}\right), \tag{2.27}$$

with $z_{\boldsymbol{k}} = \sqrt{\jmath\gamma(k_{\mathsf{a}}+k_{\mathsf{b}})\sigma_{\mathrm{d}}^2}$. Since the PDF is a real function, the Fourier coefficients satisfy the symmetry condition $C_{\boldsymbol{k}}(\boldsymbol{r},\boldsymbol{r}_0) = C_{-\boldsymbol{k}}(\boldsymbol{r},\boldsymbol{r}_0)$.

Visualizing such a four-dimensional PDF can be difficult and we make an attempt by showing various scatter plots further below and in Paper B. When showing a scatter plot of the received points in only polarization $\mathsf{a}$, one is effectively showing a particle representation of the marginal distribution

$$f_{\Theta_\mathsf{a},R_\mathsf{a}|\boldsymbol{\Theta}_0,\boldsymbol{R}_0}(\theta_\mathsf{a},r_\mathsf{a}|\boldsymbol{\theta}_0,\boldsymbol{r}_0) = \int_0^{2\pi}\int_0^{\infty} f_{\boldsymbol{\Theta},\boldsymbol{R}|\boldsymbol{\Theta}_0,\boldsymbol{R}_0}(\boldsymbol{\theta},\boldsymbol{r}|\boldsymbol{\theta}_0,\boldsymbol{r}_0)\,\mathrm{d}r_\mathsf{b}\,\mathrm{d}\theta_\mathsf{b} \tag{2.28}$$

$$= \frac{1}{4\pi^2}\sum_{k_\mathsf{a}\in\mathbb{Z}}\sum_{k_\mathsf{b}\in\mathbb{Z}} e^{\jmath k_\mathsf{a}\theta_\mathsf{a}}\int_0^{\infty} C_{\boldsymbol{k}}(\boldsymbol{r},\boldsymbol{r}_0)\,\mathrm{d}r_\mathsf{b}\int_0^{2\pi} e^{\jmath k_\mathsf{b}\theta_\mathsf{b}}\,\mathrm{d}\theta_\mathsf{b} \tag{2.29}$$

$$= \frac{1}{2\pi}\sum_{k_\mathsf{a}}\tilde{C}_{k_\mathsf{a}}(r_\mathsf{a},\boldsymbol{r}_0)e^{\jmath k_\mathsf{a}\theta_\mathsf{a}}, \tag{2.30}$$

where the modified Fourier coefficient for the above marginal PDF is [33]

$$\tilde{C}_{k_\mathsf{a}}(r_\mathsf{a},\boldsymbol{r}_0) = \int_0^{\infty} C_{(k_\mathsf{a},0)}(\boldsymbol{r},\boldsymbol{r}_0)\,\mathrm{d}r_\mathsf{b} \tag{2.31}$$

$$= \frac{2z_{\boldsymbol{k}}}{\sin z_{\boldsymbol{k}}} r_\mathsf{a}\exp\left(-\frac{r_\mathsf{a}^2+r_{0\mathsf{a}}^2}{(\tan z_{\boldsymbol{k}})/z_{\boldsymbol{k}}}\right) I_{k_\mathsf{a}}\left(\frac{2z_{\boldsymbol{k}}}{\sin z_{\boldsymbol{k}}} r_\mathsf{a} r_{0\mathsf{a}}\right)\cdot\frac{\exp\left(r_{0\mathsf{b}}^2\cdot z_{\boldsymbol{k}}\tan z_{\boldsymbol{k}}\right)}{\cos z_{\boldsymbol{k}}}. \tag{2.32}$$

Due to symmetry, the polarizations $\mathsf{a}$ and $\mathsf{b}$ have the same marginal distributions. Also, the above distributions are conditioned on a given symbol in *both* polarizations.

In Paper B, we restrict ourselves to $M$-PSK with the same power in both polarizations in order to simplify the detection problem.

### 2.4.3 Paper C: Non-Dispersion-Managed Links with a Linear Receiver

For Paper C, we consider PM transmission without neglecting dispersive effects that are due to $\beta_2$ (we do, however, neglect PMD). The optical transmission link consists of the periodic concatenation of a standard SMF and an EDFA (i.e., a lumped amplification scheme) and there is no optical inline dispersion compensation through DCFs. A block diagram of this setup is shown in Fig. 2.8. The transmitters (TX) employ a linear pulse modulation according to $x_\mathsf{a}(t) = \sum_k x_{\mathsf{a},k}p(t-kT_s)$ for polarization $\mathsf{a}$ and similarly for polarization $\mathsf{b}$. The evolution of the PM signal is described by the Manakov equation (2.24). The received signal in each polarization is processed according to the linear matched filter receiver shown in Fig. 2.1. For polarization $\mathsf{a}$, this amounts to passing $y_\mathsf{a}(t)$ through an equalizer, a pulse-matched filter, and a sampler, to obtain $y_{\mathsf{a},k'} = y_\mathsf{a}(t)\circledast h(t,-L)\circledast p(-t)|_{t=k'T_s}$ and similarly for polarization $\mathsf{b}$.

Characterizing the statistical relationship between the transmitted symbols and received samples is a challenging task due to the complicated interaction of the signal with itself, the noise, and the signal in the orthogonal polarization. The crucial difference with respect to the setup in Paper A and Paper B is the presence of dispersive filtering

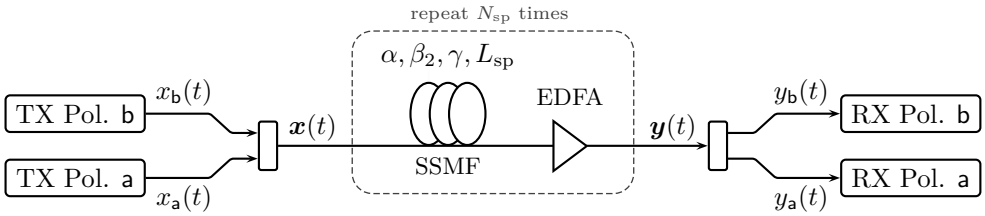



**Figure 2.8:** Block diagram of the PM transmission scheme considered in Paper C.

effects throughout the signal propagation. Optical transmission links without any inline dispersion compensation are referred to as non-DM or uncompensated transmission links. Recently, there has been a substantial amount of work on these types of transmission links with the goal to find such a statistical relationship [35–38].[11]

In [35], it is shown that the discrete-time channel for non-DM links is well modeled by a circularly symmetric complex additive Gaussian channel including a complex scaling factor. In the derivation of the model, the assumption is that dispersive effects are dominant (i.e., the symbol rate is high enough) and that the nonlinear effects are not too strong. The complex scaling accounts for a constant phase offset as well as the fact that part of the signal is converted into noise-like interference through the interaction between the dispersive and nonlinear effects. For simplicity, it is then assumed that the nonlinear noise is additive, Gaussian, and uncorrelated (both in time and across polarizations). Thus, the discrete-time channel model in polarization $\mathsf{a}$ is given according to

$$y_{\mathsf{a}} = \zeta x_{\mathsf{a}} + n_{\mathsf{a}} + \tilde{n}_{\mathsf{a}}, \tag{2.33}$$

where $\zeta \in \mathbb{C}$ is a complex scaling factor, $n_{\mathsf{a}}$ corresponds to the linear ASE noise with $\mathbb{E}[N_{\mathsf{a}} N_{\mathsf{a}}^*] = \mathsf{N}_0/T_s$, and $\tilde{n}_{\mathsf{a}}$ accounts for nonlinear noise with $\mathbb{E}[\tilde{N}_{\mathsf{a}} \tilde{N}_{\mathsf{a}}^*] = \eta P^3$, where the same transmit power $P$ is assumed for both signals in the two polarizations. $\eta$ (and hence the nonlinear noise variance) is a function of the link parameters and the symbol time, i.e., $\eta = f(\alpha, \beta_2, \gamma, L_{\text{sp}}, N_{\text{sp}}, T_s)$ [35, eq. (15)], and $|\zeta|^2 = 1 - |\eta| P^2$.

The main difference with respect to a "conventional" discrete-time additive Gaussian channel is that the signal-to-noise ratio (SNR) (defined as the ratio of the input power to the *additive* noise power) is not sufficient to characterize the operating point of the channel but rather one needs to consider the pair $(P, P_{\text{ASE}})$ or, more practically relevant, the pair $(P, L)$. This parameter pair in turn leads to both a linear and a nonlinear noise variance based on which an effective SNR can be computed.

[11] These links are also of high practical relevance and according to [36], "the current consensus is that green-field installations, as well as major overhauling and refurbishing of existing links, should adopt uncompensated transmission."

### 2.4.4 Comparison

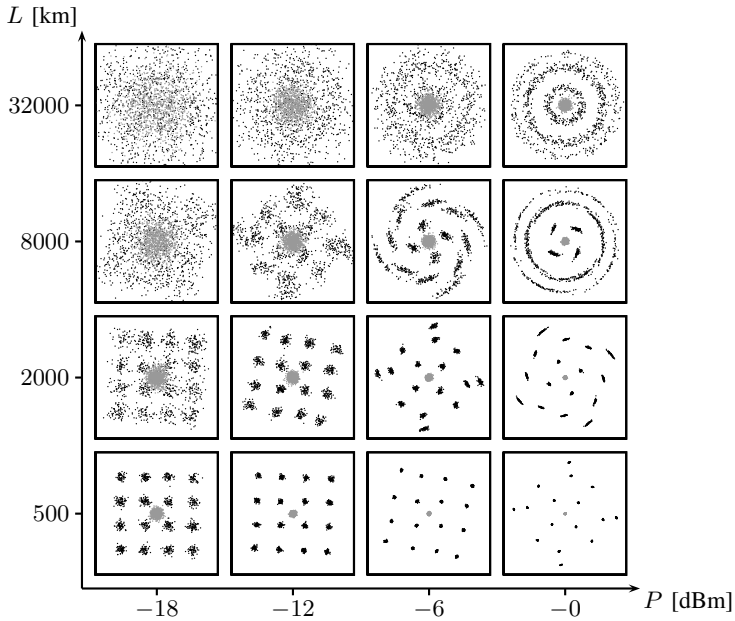
In order to illustrate the previously described models, we show scatter plots of the received symbols for different combinations of input power $P$ and transmission distance $L$ in Fig. 2.9. The transmitted symbols are taken uniformly at random from the 16-QAM signal constellation. For Fig. 2.9(a) and (b), we use the same link parameters as in Paper A. However, when generating the scatter plots shown in Fig. 2.9(a), we take into account the noise in *both* polarizations. In fact, the gray points in Fig. 2.9(a) correspond to the received symbols in the unused polarization and hence they represent only the noise. The scatter plots along diagonals in Fig. 2.9 correspond to the same signal-to-additive-noise ratio. For Fig. 2.9(c), we use the same parameters as in Paper C. We caution the reader that a quantitative comparison between the scatter plots shown for the models in Paper A/B and C is not fair due to the different link parameters and amplification schemes. Furthermore, spectral considerations are not reflected in the scatter plots. For example, operating at high input power and transmission length seems feasible for the memoryless models, since the phase is predominantly distorted. However, in this transmission regime severe spectral broadening of the signal is to be expected. Also, for the PM transmission shown in Fig. 2.9(b), the nonlinear phase rotation in one polarization depends on the selected point in the other polarization. Therefore, the noise clouds at high input power begin to separate into three smaller clouds, where each cloud corresponds to a different magnitude of the transmitted symbol in the orthogonal polarization.

From Fig. 2.9(c), it can be observed that the Gaussian noise assumption for the nonlinear interference appears to be valid for a wide range of transmission parameters. For the scatter plot in the lower right corner, is seems that the outer symbols are more affected by phase noise rather than circularly symmetric Gaussian noise, which can be explained by the fact that the transmission distance (two fiber spans) is relatively short.

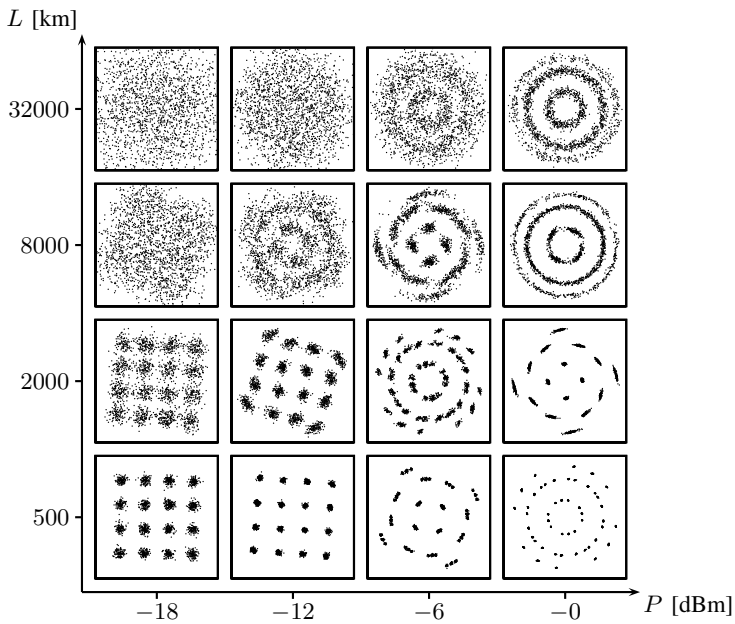
## 2.5 Further Reading

In this chapter, we have presented an overview of the basic concepts regarding fiber-optical channel modeling for long-haul data transmission. However, there are many important issues that have not been addressed, some of which we briefly mention in this section.

One important topic is the study of wavelength-division multiplexing (WDM) systems, where many input signals are multiplexed in the frequency domain at different carrier frequencies and simultaneously transmitted over the fiber. In this case, the underlying channel model, i.e., the sNLSE together with the corresponding block diagrams via the SSFM in Fig. 2.5, are still valid. However, nonlinear effects may be more pronounced due to the additional input power. Furthermore, it was mentioned in the beginning of this chapter that the matched filter receiver can be seen as a baseline. The development of improved receiver structures is an active area of research. For example, digital

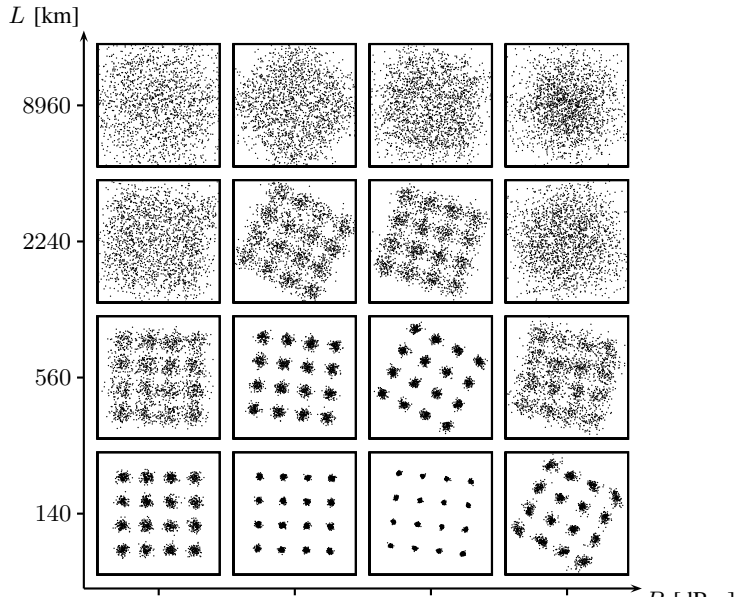



**(a)** Paper A




**(b)** Paper B

**Figure 2.9:** Comparison of scatter plots for different combinations of input power $P$ and transmission length $L$.




**(c)** Paper C

**Figure 2.9:** (cont.)

backpropagation (DBP) can be employed to improve performance, which is based on the invertibility of the Schrödinger equation in the absence of noise. In [39, 40], the authors extend DBP via a factor-graph approach showing significant improvements over DBP for DM links in the nonlinear operating regime. In [41, 42], different detection approaches are compared, with a focus on MLSE. In [19], a new transmission scheme based on the nonlinear Fourier transform is proposed.

We have also ignored hardware imperfections, e.g., laser phase noise, which might potentially have a significant impact on the discrete-time channel model that should be used to design practical modulation and coding schemes.

# CHAPTER 3

# Bit-Interleaved Coded Modulation with Low-Density Parity-Check Codes

In this chapter, we provide a brief introduction on how to reliably transmit data at high spectral efficiencies. Spectrally efficient communication can be achieved in practice by combining forward error correction with higher-order signal constellations, which is commonly referred to as coded modulation (CM). We focus on bit-interleaved coded modulation (BICM), which is a pragmatic approach to CM and often implemented in practice, due to its inherent simplicity and flexibility. This chapter should be seen mainly as supplementary material for the problem statement that is addressed in Paper C, where it is assumed that the reader is somewhat familiar with coding theory and iterative decoding techniques.

We start by outlining the main principles behind coded modulation in Section 3.1. In Section 3.2, we explain the building blocks of a BICM system. In Section 3.3, we review some basic concepts behind LDPC codes and iterative decoding, focusing on protograph-based codes. We also briefly cover SC-LDPC codes, which are one of the code classes considered for the problem statement addressed in Paper C.

## 3.1 Introduction

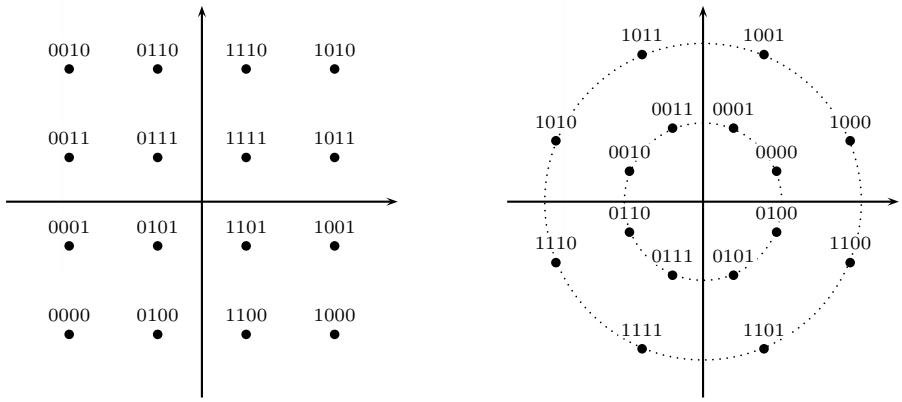
Consider again the discrete memoryless AWGN channel $y_k = x_k + n_k$ (see Section 2.1), where $x_k \in \mathbb{C}$ is the channel input at time instant $k$, $y_k$ is the corresponding channel output, and $n_k$ is the realization of a zero-mean, circularly symmetric complex Gaussian random variable $N_k$ with $\mathbb{E}[|N_k|^2] = P_\mathrm{N}$. The channel input is assumed to satisfy the average power constraint $\mathbb{E}[|X_k|^2] = P$ and the SNR is given by $\mathrm{SNR} = P/P_\mathrm{N}$. The goal is to reliably transmit data at high spectral efficiencies over this channel. To do so, one can formally define an encoder $\mathcal{E} : \{0,1\}^d \to \mathcal{C}_\mathrm{c}$, which maps a vector of $d$ information bits to a codeword in the code $\mathcal{C}_\mathrm{c} \subset \mathbb{C}^N$. Each codeword is a complex vector of length $N$ and its components serve as the input for $N$ consecutive uses of the AWGN channel. Similarly, one can define a decoder $\mathcal{D} : \mathbb{C}^N \to \{0,1\}^d$, which maps a vector of $N$ channel outputs back to a sequence of $d$ estimated bits. Assuming equally likely information bits, the communication rate (measured in [bits/complex symbol]) of such a system is given by $\kappa = \log_2(|\mathcal{C}_\mathrm{c}|)/N = d/N$. Notice that the communication rate of the discrete channel (in [bits/complex symbol]) is intimately related to the spectral efficiency of the continuous-time channel (in [bits/s/Hz]) via the bandwidth of the pulse shape $p(t)$ and the symbol rate. Shannon proved that all rates up to the channel capacity

$$C = \log_2(1 + \mathrm{SNR}) \tag{3.1}$$

are achievable, in the sense that there exists an encoder/decoder pair that can provide an arbitrarily small error probablitity as long as $N \to \infty$ [43].

While Shannon's proof provides communication engineers with an invaluable benchmark, the problem of designing practical encoders and decoders that operate close to the capacity and are implementable with reasonable complexity was not directly addressed by Shannon. In practical systems, the channel input $x_k$ commonly does not take on arbitrary complex values, but is constrained to a discrete signal constellation $\mathcal{X} \subset \mathbb{C}$. Given this premise, it is useful to introduce a soft dividing line between two different operating regimes for this channel. This dividing line is at $\kappa = 2$, where $\kappa < 2$ is referred to as the power-limited regime and $\kappa > 2$ as the bandwidth-limited regime [15]. Roughly speaking, in the power-limited regime, it is sufficient to consider a binary modulation, independently in the real and imaginary part (e.g., Gray-labeled quadrature phase-shift keying (QPSK) according to $\mathcal{X} = \{1 + \jmath, 1 - \jmath, -1 + \jmath, -1 - \jmath\}$ and scaled by $\sqrt{P/2}$), in combination with binary error correction codes in order to operate close to the capacity. On the other hand, spectrally efficient communication, i.e., $\kappa > 2$, requires the use of signal constellations with cardinality larger than 4, which are referred to as higher-order[1] constellations. By invoking the capacity formula, it follows directly that operating at high spectral efficiencies $\kappa > 2$ requires the signal power to be at least three times the noise power. In other words, spectrally efficient communication requires a reasonably high SNR.

[1] One may also classify complex constellations with 4 points as "higher-order", as long as they cannot be viewed as two independent binary modulations per real and complex dimension.




**(a)** 16-QAM **(b)** (8,8)-APSK

**Figure 3.1:** Two examples of higher-order signal constellations with 16 points.

Devising practical encoder/decoder pairs where $x_k$ is constrained to a higher-order signal constellation is commonly referred to as CM design. There exist several different approaches, for example trellis coded modulation (TCM) [44], CM with nonbinary codes [45], multilevel coded modulation (MLCM) [46], or BICM [47]. Our focus is on BICM in combination with (binary) LDPC codes, which is one of the most popular capacity-approaching coding schemes for achieving high spectral efficiency, due to its simplicity and flexibility [48]. BICM is employed as the de-facto standard in many wireless communication standards and has also been studied by many authors for fiber-optical communication systems, see, e.g., [49] or [50] and references therein.

## 3.2 BICM System Model

The transmitted symbols $x_k$ in each time instant $k$ are assumed to take on values from a discrete signal constellation $\mathcal{X} \subset \mathbb{C}$ with $|\mathcal{X}|$ points, where $|\mathcal{X}|$ is a power of two. Furthermore, each point in the constellation is assumed to be labeled with a unique binary string of length $m = \log_2 |\mathcal{X}|$, where $b_i(a)$, $1 \leq i \leq m$, denotes the $i$th bit in the binary string assigned to $a \in \mathcal{X}$ (counting from left to right). Two examples of signal constellations with $|\mathcal{X}| = 16$ points are shown in Fig. 3.1 and referred to as 16-quadrature amplitude modulation (QAM) and $(8, 8)$-amplitude phase-shift keying (APSK). For a detailed definition of APSK constellations, we refer the reader to Paper A.

In the following, we describe the main components of a BICM system. First, consider the block diagram shown in Fig. 3.2(a), where the modulo 2 addition of $d_{i,k}$ and multiplication by $\bar{d}_{i,k} = (-1)^{d_{i_k}}$ are explained further below and can be ignored for now. At each time instant, the modulator $\Phi$ takes $m$ bits $b_{i,k}$, $1 \leq i \leq m$, and maps them to one

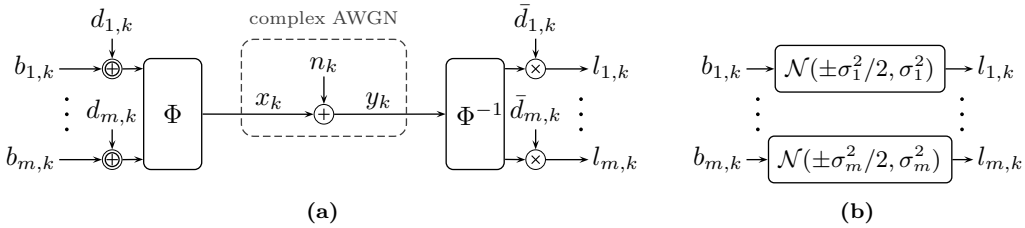



**Figure 3.2:** The modulator $\Phi$, demodulator $\Phi^{-1}$, and channel symmetrization technique in (a). A helpful approximate channel model via parallel symmetric Gaussian LLR channels in (b).

of the constellation points according to the binary labeling of the signal constellation. At the receiver, the demodulator $\Phi^{-1}$ computes soft reliability information about the transmitted bits in the form of the log-likelihood ratios (LLRs)

$$l_{i,k} \triangleq \log \frac{f_{Y_k|B_{i,k}}(y_k|0)}{f_{Y_k|B_{i,k}}(y_k|1)} = \log \frac{\sum_{x \in \mathcal{X}_{i,0}} f_{Y_k|X_k}(y_k|x)}{\sum_{x \in \mathcal{X}_{i,1}} f_{Y_k|X_k}(y_k|x)}, \tag{3.2}$$

where $\mathcal{X}_{i,u} \triangleq \{a \in \mathcal{X} : b_i(a) = u\}$ is the subconstellation where all points have the bit $u$ at the $i$th position of their binary label. The LLR is a function of the observation, and, since the observation is a random variable, the LLR is also a random variable.

One way to interpret the setup depicted in Fig. 3.2(a) is as follows. The concatenation of the modulator $\Phi$, the AWGN channel, and demodulator $\Phi^{-1}$ establishes a binary interface for the complex-valued AWGN channel. It is useful to imagine transmitting data over a set of $m$ parallel binary-input continuous-output channels, or simply bit channels, where one may view the conditional distribution of the LLR $f_{L_{i,k}|B_{i,k}}(\cdot|\cdot)$, $1 \leq i \leq m$, as a bit channel. In the following, a bit channel $f_{L|B}(l|b)$ is called symmetric if $f_{L|B}(l|0) = f_{L|B}(-l|1)$ and referred to as an LLR channel if $f_{L|B}(l|0)e^l = f_{L|B}(l|1)$. The terminology here is used to emphasize that, if the second condition is fulfilled, the output of the channel corresponds to a "true" LLR. This is important because, in practice, low-complexity approximations of (3.2) are often considered, and the resulting bit channel in that case is not necessarily an LLR channel. One can show that $f_{L_{i,k}|B_{i,k}}(\cdot|\cdot)$ is an LLR channel. However, the channel is not necessarily symmetric in general.[2] Symmetry can be enforced by adding modulo 2 independent and identically distributed bits $d_{i,k}$ to the bits $b_{i,k}$ [51]. After the demodulator, the corresponding LLR is multiplied by $\bar{d}_{i,k} = (-1)^{d_{i,k}}$, which implies that the bits $d_{i,k}$ are known to both the transmitter and receiver. The resulting bit channel $f_{L_{i,k}|B_{i,k}}(\cdot|\cdot)$ can be shown to be symmetric.

We proceed by quantifying the quality of the $m$ bit channels, where we rely on the mutual information (MI) as a measure of quality. The MI between the output of a

[2] The symmetry condition will become important when discussing density evolution and LDPC codes, where one relies on the all-zero codeword assumption.

symmetric LLR channel $f_{L|B}(l|b)$ and uniform input bits is given by

$$I(L;B) = \mathbb{E}\left[\log_2\left(\frac{f_{L|B}(L|B)}{f_L(L)}\right)\right] \tag{3.3}$$

$$= 1 - \mathbb{E}\left[\log_2\left(\frac{f_{L|B}(L|B) + f_{L|B}(L|1-B)}{f_{L|B}(L|B)}\right)\right] \tag{3.4}$$

$$= 1 - \mathbb{E}\left[\log_2\left(1 + \frac{f_{L|B}(L|1-B)}{f_{L|B}(L|B)}\right)\right] \tag{3.5}$$

$$= 1 - \mathbb{E}\left[\log_2\left(1 + \exp((-1)^{1-B}L)\right)\right] \tag{3.6}$$

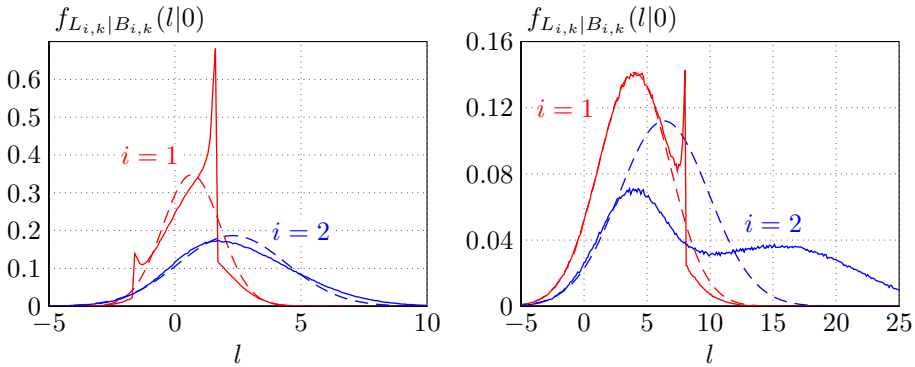
$$= 1 - \int_{-\infty}^{+\infty} f_{L|B}(l|0)\log_2(1+\exp(-l))\,\mathrm{d}l. \tag{3.7}$$

Writing the MI in the form (3.7) can be useful in order to compute the MI with the help of Monte Carlo integration.

It turns out that, while the channel quality of the bit channels can be determined quite efficiently, it is very difficult to find exact analytical expressions for the actual densities $f_{L_{i,k}|B_{i,k}}(\cdot\,|\,\cdot)$. A common approach in the analysis of BICM is to make the simplifying assumption that the densities $f_{L_{i,k}|B_{i,k}}(\cdot\,|\,\cdot)$ are approximately Gaussian. An LLR channel with a Gaussian density is particularly simple, because it can be parametrized by a single parameter. More precisely, we refer to a bit channel $f_{L|B}(l|b)$ as a symmetric Gaussian LLR channel with parameter $\sigma^2$ if $L \sim \mathcal{N}(\sigma^2/2, \sigma^2)$ conditioned on $B = 0$ and $L \sim \mathcal{N}(-\sigma^2/2, \sigma^2)$ conditioned on $B = 1$. The MI between the output of a symmetric Gaussian LLR channel and uniform input bits is denoted by $J(\sigma)$. Under the Gaussian assumption, a helpful approximation of the setup in Fig. 3.2(a) is shown in Fig. 3.2(b), where transmission takes place over $m$ parallel symmetric Gaussian LLR channels with different parameters $\sigma_i^2$. In order to find a correspondence between the LLR channels $f_{L_{i,k}|B_{i,k}}(\cdot\,|\,\cdot)$ and the parameters $\sigma_i^2$, one may match the MI according to $J(\sigma_i) = I_i(\mathrm{SNR}) \Leftrightarrow \sigma_i^2 = J^{-1}(I_i(\mathrm{SNR}))^2$, where $I_i(\mathrm{SNR}) = I(B_{i,k}; L_{i,k})$ is independent of $k$.

While the parallel Gaussian model can be quite useful, one should, however, be aware of the inaccuracies of this simplified model. In particular, the bit channels are not independent as suggested in Fig. 3.2(b) and the true distribution of the LLRs is not Gaussian. To illustrate the latter inaccuracy, in Fig. 3.3, we compare the actual densities with the approximated Gaussian densities for two different SNRs for the first two bit positions of the 16-QAM constellation shown in Fig. 3.1(a).[3] The densities $f_{L_{i,k}|B_{i,k}}(\cdot\,|0)$ are estimated via histograms and shown by the solid lines, whereas the Gaussian densities are shown by the dashed lines. It can be seen that the actual densities are clearly non-Gaussian and the accuracy of the Gaussian approximation therefore depends on the application scenario. For the application in Paper C (predicting the iterative performance

[3]The third and fourth bit positions lead to identical distributions, due to the fact that 16-QAM with the shown labeling can be seen as a product constellation of two one-dimensional constellation.




**(a)** SNR = 3 dB **(b)** SNR = 10 dB

**Figure 3.3:** Comparison of the true LLR channels (including channel symmetrization) with the symmetric Gaussian LLR channels that have the same MI.

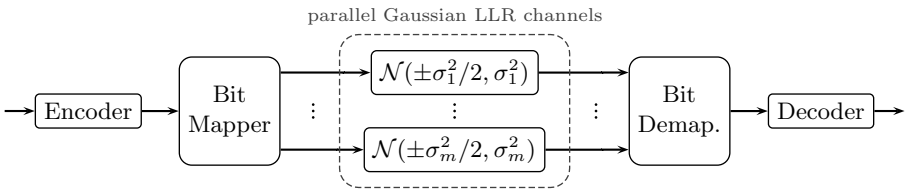



**Figure 3.4:** A useful approximate system model for BICM systems.

behavior of LDPC codes), the approximation is quite accurate and at the same time allows for a major simplification of the analysis, thereby justifying its use.

Consider now the case where we employ a single binary code $\mathcal{C} \subset \{0,1\}^n$ of length $n$, and each codeword is transmitted using $N = n/m$ symbols $x_k$. The allocation of the coded bits to the modulator (i.e., the different bit channels in Fig. 3.2(b)) is determined by a bit mapper as shown in Fig. 3.4. In Paper C, our goal is to find good bit mappers for a given code and signal constellation, where we focus on protograph-based LDPC codes.

As a side note, we remark that the term "bit interleaver" is also commonly used instead of "bit mapper". In fact, the modulator $\Phi$ is sometimes referred to as the (symbol) mapper (and the demodulator $\Phi^{-1}$ as the demapper), which the reader should be aware of in order to avoid confusion. However, the terms "bit mapper", "bit mapping", or "mapping" seem to be preferred in the literature when the allocation of the coded bits to the constellation symbols $\Phi$ is explicitly studied or optimized, see, e.g., [52,53]. Moreover, outside the context of BICM, the terms "mapping device" or "channel mapper" are used when studying parallel channels in combination with binary codes, e.g., in [54, 55].

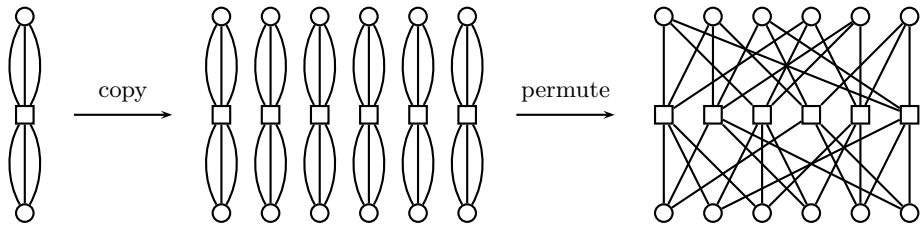



**Figure 3.5:** Illustration of the protograph lifting procedure for $\mathbf{P} = (3, 3)$ and $M = 6$.

## 3.3 Low-Density Parity-Check Codes

LDPC codes were proposed by Gallager in his PhD thesis [56]. They were conceived as practically decodable codes, able to "utilize the long block lengths necessary for low error probability without requiring excessive equipment or computation" [57]. Formally, a binary LDPC code $\mathcal{C}$ of length $n$ is defined as the null space of a sparse parity-check matrix $\mathbf{H} = [h_{i,j}] \in \{0,1\}^{c \times n}$, i.e., $\mathcal{C} = \{\boldsymbol{c} \in \{0,1\}^n : \mathbf{H}\boldsymbol{c}^\intercal = \mathbf{0}\}$, where $n > c$ and operations are over the binary field. Assuming that $\mathbf{H}$ has full rank $c$, one can invoke the fundamental theorem of linear algebra to infer that the code has $|\mathcal{C}| = 2^d$ codewords, where $d = n - c$ is the dimension of the code. The code rate is defined as $R = d/n = 1 - c/n$.

### 3.3.1 Construction via Protographs

There exist different methods to construct "good" LDPC codes, i.e., good matrices $\mathbf{H}$, and one popular method is by using protographs [58]. The parity-check matrix of an LDPC code can be represented by using a bipartite Tanner graph consisting of $n$ variable nodes (VNs) and $c$ check nodes (CNs), where the $i$th CN is connected to the $j$th VN if $h_{i,j} = 1$. A protograph is also a "small" bipartite graph defined by an adjacency matrix $\mathbf{P} = [p_{i,j}] \in \mathbb{N}_0^{c' \times n'}$, called the base matrix. Given $\mathbf{P}$, a parity-check matrix $\mathbf{H}$ is obtained by replacing each entry $p_{i,j}$ in $\mathbf{P}$ with a random binary $M$-by-$M$ matrix which contains $p_{i,j}$ ones in each row and column. This procedure is called lifting and $M \geq \max_{i,j} p_{i,j}$ is the lifting factor. Graphically, it amounts to copying the protograph $M$ times and subsequently permuting edges, in order to obtain the Tanner graph. Parallel edges, i.e., for $p_{i,j} > 1$, are permitted in the protograph and are resolved in the lifting procedure. The design rate of the code is given by $R = 1 - c/n = 1 - c'/n'$, where $c = c'M$ and $n = n'M$. As an example, the lifting procedure for $\mathbf{P} = (3, 3)$ and $M = 6$ is illustrated in Fig. 3.5.

Designing codes via protographs has several practical advantages, e.g., a quasi-cyclic code construction is easily applied by constraining the $M$-by-$M$ matrices to have a circulant structure which in turn allows for hardware-efficient implementation [59, p. 263], suitable for high-speed optical communications [6]. Moreover, codes of different lengths can be obtained simply by adjusting the lifting factor.

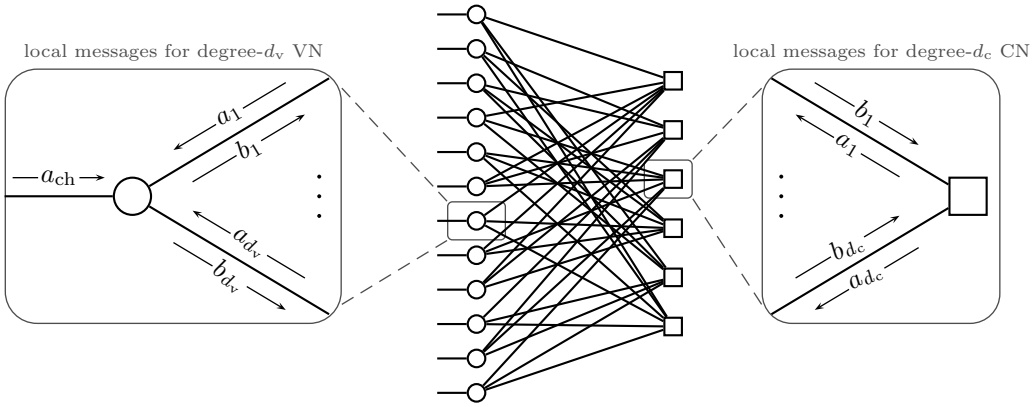



**Figure 3.6:** Illustration of the messages involved in the iterative BP decoding algorithm.

### 3.3.2 Iterative Belief Propagation Decoding

Consider the scenario where each bit in the codeword is transmitted over an LLR channel $f_{L|B}(\cdot\,|\,\cdot)$. The goal of the decoder is to recover the transmitted codeword based on the observation from the channel, which consists of $n$ LLRs. These LLRs can be interpreted as the initial belief about each coded bit. During the decoding process, the decoder tries to iteratively improve the accuracy of the belief, by exchanging messages in the form of extrinsic LLRs between VNs and CNs along the edges of the Tanner graph.

For an excellent and comprehensive description of belief propagation (BP) decoding, we refer the reader to [59, Ch. 5.3]. Here, we will only briefly review the basic steps of the decoding algorithm. We use the following convention. Messages arriving at VNs are denoted by $a$, while messages emanating from VNs are denoted by $b$. For CNs, it is the other way around, i.e., arriving messages are denoted by $b$, while emanating messages by $a$. In an attempt to avoid cluttered notation, only one index is appended to $a$ or $b$ in order to *locally* distinguish between messages along different edges for the same node. The corresponding picture we have in mind is illustrated in Fig. 3.6. By locally we mean that, for example, the message $b_1$ emanating from the magnified VN does not correspond to the message $b_1$ arriving at the magnified CN. (In fact, from the way the figure is drawn, the message $b_1$ arriving at the magnified CN would emanate from the fourth VN, counting from the top.)

Consider now an arbitrary VN of degree $d_v$, where the degree of a VN corresponds to the number of CNs that are connected to it. There are $d_v + 1$ messages arriving at this VN, where $a_1, \ldots, a_{d_v}$ are messages from CNs and $a_{ch}$ corresponds to a channel LLR. The $d_v$ outgoing messages $b_1, \ldots, b_{d_v}$ are computed according to

$$b_i = \sum_{\sim i} a_j + a_{ch}, \tag{3.8}$$

where the summation is over the index set $j \in \{1, \ldots, d_\mathrm{v}\}$ *excluding* the index $i$. Similarly, if we consider an arbitrary CN of degree $d_\mathrm{c}$, there are $d_\mathrm{c}$ messages $b_1, \ldots, b_{d_\mathrm{c}}$ arriving and the outgoing messages are computed according to

$$a_i = 2 \tanh^{-1} \left( \prod_{\sim i} \tanh(b_j/2) \right), \tag{3.9}$$

where the product is over the index set $j \in \{1, \ldots, d_\mathrm{c}\}$ excluding the index $i$. Since the CN operation (3.9) is central in the analysis of LDPC codes under iterative decoding, it is very common to rewrite it in terms of the binary boxplus operator defined as

$$b_1 \boxplus b_2 = 2 \tanh^{-1} \left( \tanh(b_1/2) \tanh(b_2/2) \right), \tag{3.10}$$

where the box-addition of an arbitrary number of terms is evaluated by recursively applying (3.10), e.g., $b_1 \boxplus b_2 \boxplus b_3 = (b_1 \boxplus b_2) \boxplus b_3$. With this convention, one can write the CN operation more concisely as

$$a_i = \boxplus_{\sim i} b_j. \tag{3.11}$$

The decoding process can now be described as follows. Initialize $a_\mathrm{ch}$ for all VNs to the corresponding channel LLR, and set all other messages to 0. Then, repeat the following two steps. First, compute outgoing messages for all VNs according to (3.8). After that, compute the outgoing messages for all CNs according to (3.9). Stop if either a maximum number of iterations has been reached, or the proper combination of the hard decisions on the messages

$$\sum_{j=1}^{d_\mathrm{v}} a_j + a_\mathrm{ch} \tag{3.12}$$

for all VNs forms a valid codeword.

### 3.3.3 Density Evolution

Density evolution (DE) is a powerful tool to analyze the iterative decoding behavior and performance of LDPC codes [60]. DE mimics the decoding process under a cycle-free graph assumption by tracking how the densities of the messages evolve with iterations. DE is commonly used to find so-called decoding thresholds, which can be interpreted as the capacity for LDPC codes under BP decoding. Similar to the channel capacity, the threshold divides the channel quality parameter range (for example the parameter $\sigma^2$ of a symmetric Gaussian LLR channel) into a region where reliable decoding is possible and where it is not.

The main steps in the DE algorithm can be understood by considering the update equations for the VNs (3.8) and CNs (3.9). If we assume that the involved incoming messages

are random variables, then they have a certain probability distribution or density (for example, $a_{\text{ch}}$ is distributed according to LLR channel). The main question is, how can we obtain the densities of the outgoing messages? For the VN update, the answer turns out to be a simple convolution. In particular, for two independent random variables $A$ and $B$ with distributions $f_A(a)$ and $f_B(b)$, their sum $C = A + B$ is distributed according to $f_C(c) = f_A(a) \circledast f_B(b)$. It is convenient to introduce the short notation $\mathsf{a} \circledast \mathsf{b}$, where $\mathsf{a}$ and $\mathsf{b}$ are placeholders for the densities of the random variables $A$ and $B$ [61]. With this notation, the densities of the outgoing messages, given the densities of the incoming messages, can be computed according to

$$\mathsf{b}_i = \bigoplus_{\sim i} \mathsf{a}_j \circledast \mathsf{a}_{\text{ch}}. \tag{3.13}$$

For the CN update, it is somewhat more challenging to obtain the densities of the outgoing messages. The most straightforward approach is by using Monte Carlo techniques and histograms. Consider the case where two messages $b_1$ and $b_2$ with densities $\mathsf{b}_1$ and $\mathsf{b}_2$ are processed according to the boxplus operation $a = b_1 \boxplus b_2$. In order to obtain the density $\mathsf{a}$, one can simply generate many independent realizations of the random variables $B_1$ and $B_2$, perform the boxplus operation, and collect the resulting samples. These samples can be seen as a particle representation of the density $\mathsf{a}$. This method is illustrated in Fig. 3.7, where it is shown how two consistent Gaussian densities "evolve" under the boxplus operation. A density $\mathsf{a}$ is called a consistent Gaussian density[4] with parameter $\sigma^2$, if $A \sim \mathcal{N}(\sigma^2/2, \sigma^2)$. As a short notation, one may introduce the operator $\mathsf{a} = \mathsf{b}_1 \boxtimes \mathsf{b}_2$, referred to as box-convolution [61]. In practice, the box-convolution of two densities can be implemented by using a look-up table approach [62]. Similar to (3.14), the densities of the outgoing CN messages can then be computed according to

$$\mathsf{a}_i = \boxed{\circledast}_{\sim i} \mathsf{b}_j. \tag{3.14}$$

For protograph-based codes, DE can be used to analyze the iterative decoding behavior by tracking one density for each edge in the protograph. This asserts that the messages exchanged during the decoding process over edges belonging to the same edge-type (defined by one protograph edge) have the same density. Assume that the transmission takes place over a symmetric LLR channel with a fixed channel quality and we wish to predict the iterative decoding behavior. Due to the channel symmetry, one may assume the transmission of the all-zero codeword [59, p. 389]. Start by initializing $\mathsf{a}_{\text{ch}}$ for all VNs in the protograph to $f_{L|B}(l|0)$ and set all other densities to $\delta(l)$. Then, repeat the following two steps. First, calculate the outgoing message densities for all VNs in the protograph according to (3.13). After that, calculate the outgoing message densities for all CNs in the protograph according to (3.14). Stop if the error probability associated

[4] Note that the conditional distribution $f_{L|B}(l|0)$ of a symmetric Gaussian LLR channel corresponds to a consistent Gaussian density.

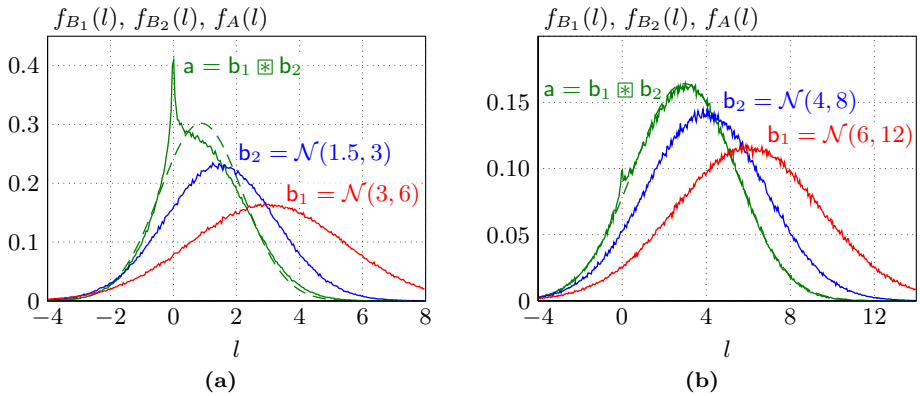



**Figure 3.7:** Illustration of the box-convolution of two consistent Gaussian densities. The green dashed line corresponds to the consistent Gaussian approximation obtained via EXIT functions.

with the density

$$\overset{d_v}{\underset{j=1}{\circledast}} \mathsf{a}_j \circledast \mathsf{a}_{\text{ch}} \tag{3.15}$$

for each VN is below a certain target bit error probability (successful decoding), where the error probability associated with a density $\mathsf{a}$ is given by

$$p_e(\mathsf{a}) = \int_{-\infty}^{0} f_A(a)\,\mathrm{d}a, \tag{3.16}$$

or a maximum number of iterations is reached (decoding failure). In order to find the decoding threshold, the above procedure is repeated many times for decreasing channel quality until the decoding fails, starting from a channel quality where the decoding is successful.

### Approximate Density Evolution via EXIT Functions

Tracking the full densities (or quantized densities in practice) is computationally demanding and extrinsic information transfer (EXIT) functions are usually considered to be a good compromise between computational efficiency and accuracy [63]. Let us assume that the density $\mathsf{a}$ fulfills the condition $f_A(a)e^a = f_A(-a)$. Then, the density can be associated with the MI measure

$$I(\mathsf{a}) = 1 - \int_{-\infty}^{\infty} f_A(a)\log_2(1 + e^{-a})\,\mathrm{d}a. \tag{3.17}$$

Now, instead of tracking the evolution of densities, one may track the evolution of the MI measure associated with the densities (which is just a scalar value for each density). Let us assert that, under the VN operation, this measure evolves approximately according to

$$I(\mathsf{b}_i) \approx \tilde{J}\left(\sum_{\sim i} \tilde{J}^{-1}(I(\mathsf{a}_j)) + \tilde{J}^{-1}(I(\mathsf{a}_{\mathrm{ch}}))\right), \tag{3.18}$$

whereas, under the CN operation it evolves approximately according to

$$I(\mathsf{a}_i) \approx 1 - \tilde{J}\left(\sum_{\sim i} \tilde{J}^{-1}(1 - I(\mathsf{b}_j))\right), \tag{3.19}$$

where $\tilde{J}(x) = J(\sqrt{x})$. These two equations can be motivated as follows. Eq. (3.18) is exact under the assumption that all incoming densities $\mathsf{a}_1, \ldots, \mathsf{a}_{d_\mathrm{v}}$, and $\mathsf{a}_{\mathrm{ch}}$ are consistent Gaussian densities. To see this, note that the convolution of two consistent Gaussian densities with parameters $\sigma_1^2$ and $\sigma_2^2$ is another consistent Gaussian density with parameter $(\sigma_1^2 + \sigma_2^2)/2$. Furthermore, if $\mathsf{a}$ is a consistent Gaussian density with parameter $\sigma^2$, the operation $\tilde{J}^{-1}(I(\mathsf{a}))$ simply returns $\sigma^2$. Without going into details, (3.19) can be heuristically motivated by a duality property that holds for the binary erasure channel (BEC) [59, p. 415]. It is important to point out that (3.19) it is not exact, even if all incoming densities are consistent Gaussians, but it turns out to be surprisingly accurate nonetheless. For example, the green dashed lines in Fig. 3.7 have been obtained using (3.19), where the resulting MI measure is plotted in the form of the associated consistent Gaussian density.

### 3.3.4 Spatially Coupled LDPC Codes

Spatial coupling of regular[5] LDPC codes has emerged as a powerful technique to construct capacity-achieving codes for a large class of channels using iterative BP decoding [13,64]. The main idea is to make several copies of the Tanner graph that defines the regular base code, arrange the copies next to each other, and then interconnect neighboring graphs in a particular way. The key to the outstanding performance of codes constructed in such a way is a boundary effect due to slight irregularities at the two ends of the resulting Tanner graph.

In general, SC-LDPC codes have parity-check matrices with a band-diagonal structure, see, e.g., [64] for a formal definition. Here, we briefly introduce their construction via protographs [65], [66, Sec. II-B]. The base matrix $\mathbf{P}_{[T]}$ of a $(J, K)$ regular, protograph-based SC-LDPC code with termination length $T$ can be constructed by specifying matrices $\mathbf{P}_i$, $0 \leq i \leq m_\mathrm{s} + 1$ of dimension $J'$ by $K'$, where $m_\mathrm{s}$ is referred to as the memory. The matrices are such that $\mathbf{P} = \sum_i \mathbf{P}_i$ has column weight $J$ and row weight $K$ for all columns and rows, respectively. Given $T$ and the matrices $\mathbf{P}_i$, the base matrix $\mathbf{P}_{[T]}$

[5] An LDPC codes is called regular if all VNs have degree $d_\mathrm{v}$ and all CNs have degree $d_\mathrm{c}$.

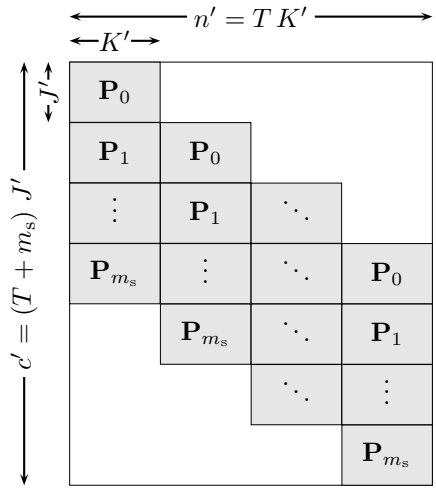



**Figure 3.8:** Illustration of the base matrix $\mathbf{P}_{[T]}$ of a $(J, K)$ regular, protograph-based SC-LDPC code.

is constructed as shown in Fig. 3.8. From the dimensions of $\mathbf{P}_{[T]}$ one can infer a design rate of $R(T) = 1 - (T + m_s)J'/(TK')$. As $T$ grows large, the rate approaches $R(\infty) = 1 - J'/K'$.

Before continuing, it is insightful to recall the following statement from [67], where the design of irregular LDPC codes is studied. (VNs are referred to as message nodes.)

> "[...] we offer some intuition as to why irregular graphs prove useful. [...] Message nodes with high degree tend to their correct value quickly. These nodes then provide good information to the check nodes, which subsequently provide better information to lower degree message nodes. Irregular graph constructions thus *lead to a wave effect*, where high degree message nodes tend to get corrected first, and then message nodes with slightly smaller degree, and so on down the line." [emphasis added]

For SC-LDPC codes, one can give a similar heuristic explanation for their outstanding performance as follows (see [64] for a detailed explanation). By inspecting the structure of the base matrix in Fig. 3.8, one may verify that the CN degrees corresponding to the first and last couple of rows is lower than the CN degrees corresponding to the rows in between. The lower degree CNs lead to a locally better decoding capability which helps decoding neighboring VNs. This local boundary effect turns out to initiate a wave-like behavior and can have a global effect on the decoding capability of the entire code with increasing number of decoding iterations. To illustrate this behavior, in Fig. 3.9, we show the predicted bit error rates $p_e$ via (approximate) DE for the coded bits corresponding to the $j$th column of the SC-LDPC protograph $\mathbf{P}_{[T]}$ with component matrices $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = (1, 1)$ and $T = 100$. We assume transmission over a symmetric

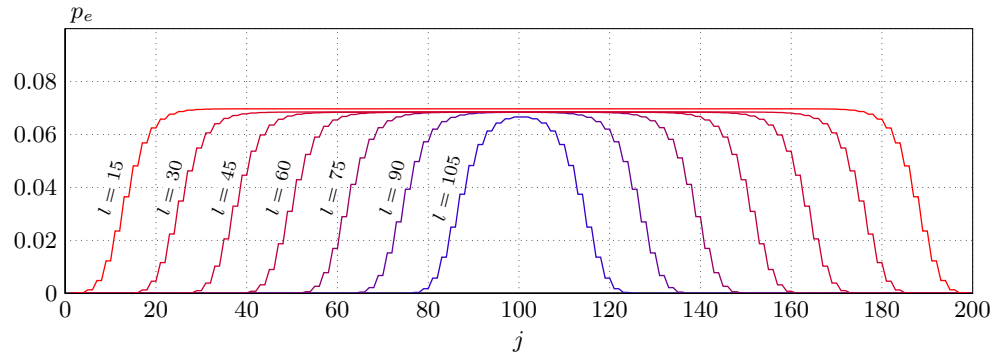



**Figure 3.9:** Illustration of the wave-like decoding behavior of SC-LDPC codes.

Gaussian LLR channel with parameter $\sigma^2 = 4$. In the figure, $l$ denotes the iteration number. It can be observed that the error probability of the VNs at the two ends of the graph converges to zero after 15 iterations. Due to the spatial coupling, this boundary effect propagates inwards all the way to the center of the protograph in a wave-like fashion.

An important reason for the tremendous interest in spatially coupled codes is their universality. While irregular LDPC codes have been optimized for various communication channels, the degree distribution pairs that achieve the best performance usually vary from channel to channel [68]. In contrast, SC-LDPC codes derived from simple regular codes have been shown to universally achieve capacity for a variety of channels. However, there are also many open research problems concerning the practical implementation of SC-LDPC, see [69] for a recent overview. For example, the price to pay for the wave-like decoding behavior is a rate loss with respect to regular, uncoupled codes that are defined by the protograph $\boldsymbol{B} = \sum_i \boldsymbol{B}_i$.

# CHAPTER 4

# Conclusions and Future Work

In this chapter, we summarize the main conclusions from the appended papers and outline some potentially interesting ideas for future work.

### Paper A

In Paper A, we study the design of APSK signal constellations under the assumption of a memoryless fiber-optical channel model with NLPN as the main system impairment. Optimized APSK constellations can offer significant performance advantages over conventional QAM constellations for the assumed channel model and detection scheme. It is also shown that the optimization of the signal constellation in the presence of severe nonlinear distortions can lead to somewhat counterintuitive results in the form of sacrificial points or sacrificial rings. As outlined in [70], these effects become particularly important when studying the channel capacity for such channels. Furthermore, when the bit error probability (BEP) is taken as a performance measure, it is important to consider the joint design of both the constellation and the labeling. In particular, an optimized constellation for symbol error probability (SEP) with an optimal labeling does not necessarily provide the best BEP. In fact, more structured constellations such as the considered rectangular APSK constellations may give better performance due to the possibility of Gray-like labeling methods.

An interesting direction for future work would be to replace the relevant measures for uncoded transmission (SEP and BEP) with the relevant measures for coded transmission (MI and generalized mutual information (GMI)). It may also be rewarding to find an accurate approximate characterizations of the channel PDF for the considered channel

model which may simplify the optimization procedure in this case.

### Paper B

In Paper B, we study the design of a low-complexity detector for a memoryless fiber-optical channel assuming PM transmission. The detector uses a phase compensation scheme based on the received signal amplitudes in both polarizations, followed by a subsequent threshold detection. The complexity can be significantly reduced compared to a four-dimensional ML detector, albeit at some performance loss.

An important conclusion for the corresponding detector for SP transmission is that a nonlinear phase compensation scheme and subsequent threshold detection is equivalent to the ML detector. Unfortunately, a similar conclusion does not hold for the proposed detector for PM signals and it would be interesting to study the design of an improved detector that has this property (or show that this is fundamentally not possible). Furthermore, the detector in Paper B is limited to PM-$M$-PSK constellations and similar detection schemes for constellations with multiple amplitude levels could be investigated.

### Paper C

In Paper C, we study a coded transmission system that operates over a fiber-optical link without inline dispersion compensation. Assuming a linear coherent receiver, the classical AWGN channel with a modified SNR expression is used as a design channel. We propose a method to optimize the bit mapper that determines the allocation of the coded bits from the FEC encoder to the labeling bits of the signal constellation that is applicable to any protograph-based LDPC code. Protograph-based codes are particularly interesting for fiber-optical systems because they allow for an efficient hardware implementation. We also extend the technique to SC-LDPC codes using a windowed decoder. The results show that by using an optimized bit mapper, the transmission reach can be extended by up to 8%, with almost no added system complexity.

An interesting direction for future work could be to study the bit mapper optimization assuming hard-decision decoding. For example, Smith and Kschischang have proposed such a setup in [49] using staircase codes for BICM with an additional shaping unit. An appropriate model in this case would be to study parallel binary symmetric channels with different crossover probabilities. To the best of our knowledge, bit mapper optimization for such a scenario has not yet been considered in the literature.

# Bibliography


[1] K. F. Rauscher, "ROGUCCI study final report," IEEE Communications Society, Tech. Rep., 2010.

[2] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[3] E. Agrell and M. Karlsson, "Satellite constellations: Towards the nonlinear channel capacity," in *Proc. IEEE Photon. Conf.*, Burlingame, CA, Sep. 2012.

[4] A. D. Ellis, J. Zhao, and D. Cotter, "Approaching the non-linear Shannon limit," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 423–433, Feb. 2010.

[5] B. P. Smith and F. R. Kschischang, "Future prospects for FEC in fiber-optic communications," *IEEE J. Sel. Topics. Quantum Electron.*, vol. 16, no. 5, pp. 1245–1257, Oct. 2010.

[6] L. Schmalen, A. J. de Lind van Wijngaarden, and S. ten Brink, "Forward error correction in optical core and optical access networks," *Bell Labs Tech. J*, vol. 18, no. 3, pp. 39–66, Mar. 2013.

[7] L. Beygi, E. Agrell, J. M. Kahn, and M. Karlsson, "Coded modulation for fiber-optic networks," *IEEE Signal Processing Mag.*, vol. 31, no. 2, pp. 93–103, Mar. 2014.

[8] I. Djordjevic, W. Ryan, and B. Vasic, *Coding for Optical Channels*. Springer, 2010.

[9] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. Academic Press, 2006.

[10] K.-P. Ho, *Phase-modulated Optical Communication Systems*. Springer, 2005.

[11] L. Hanzo, W. Webb, and T. Keller, *Single- and Multi-carrier Quadrature Amplitude Modulation: Principles and Applications for Personal Communications, WLANs and Broadcasting*. Wiley, 2000.

[12] A. J. Felström and K. S. Zigangirov, "Time-varying periodic convolutional codes with low-density parity-check matrix," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2181–2191, Sep. 1999.

[13] M. Lentmaier, A. Sridharan, K. S. Zigangirov, and D. J. Costello, Jr., "Terminated LDPC convolutional codes with thresholds close to capacity," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Adelaide, Australia, Sep. 2005.

[14] S. Kudekar, T. Richardson, and R. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Cambridge, MA, Jul. 2012.

[15] G. D. Forney, Jr. and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2384–2415, Oct. 1998.

[16] A. Lapidoth, *A Foundation in Digital Communication*. Cambridge University Press, 2009.

[17] G. D. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inf. Theory*, vol. 18, no. 3, pp. 363–378, May 1972.

[18] G. Ungerboeck, "Adaptive maximum-likelihood receiver for carrier-modulated data-transmission systems," *IEEE Trans. Commun.*, vol. 22, no. 5, pp. 624–636, May 1974.

[19] M. I. Yousefi, "Information transmission using the nonlinear Fourier transform," Ph.D. dissertation, University of Toronto, 2013.

[20] M. Secondini and E. Forestieri, "The nonlinear Schrödinger equation in fiber-optic systems," *Riv. Mat. Univ. Parma*, vol. 8, pp. 69–98, Sep. 2008.

[21] O. V. Sinkin, R. Holzlohner, J. Zweck, and C. R. Menyuk, "Optimization of the split-step Fourier method in modeling optical-fiber communications systems," *J. Lightw. Technol.*, vol. 21, no. 1, pp. 61–68, Jan. 2003.

[22] G. P. Agrawal, *Fiber-optic Communication Systems*, 4th ed. Wiley-Interscience, 2010.

[23] A. Mecozzi, "Limits to long-haul coherent transmission set by the Kerr nonlinearity and noise of the in-line amplifiers," *J. Lightw. Technol.*, vol. 12, no. 11, pp. 1993–2000, Nov. 1994.

[24] T. J. Ellingham, J. D. Ania-Castanón, R. Ibbotson, X. Chen, L. Zhang, and S. K. Turitsyn, "Quasi-lossless optical links for broad-band transmission and data processing," *IEEE Photon. Technol. Lett.*, vol. 18, pp. 268–270, Jan. 2006.

[25] M. I. Yousefi and F. R. Kschischang, "A Fokker-Planck differential equation approach for the zero-dispersion optical fiber channel," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Austin, TX, Jun. 2010.

[26] ——, "On the per-sample capacity of nondispersive optical fibers," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7522–7541, Nov. 2011.

[27] J. P. Gordon and L. F. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Opt. Lett.*, vol. 15, no. 23, pp. 1351–1353, Dec. 1990.

[28] A. Mecozzi, "Probability density functions of the nonlinear phase noise," *Opt. Lett.*, vol. 29, no. 7, pp. 673–675, Apr. 2004.

[29] K. S. Turitsyn, S. A. Derevyanko, I. V. Yurkevich, and S. K. Turitsyn, "Information capacity of optical fiber channels with zero average dispersion," *Phys. Rev. Lett.*, vol. 91, no. 20, p. 203901, Nov. 2003.

[30] K.-P. Ho, "Probability density of nonlinear phase noise," *J. Opt. Soc. Am. B*, vol. 20, no. 9, pp. 1875–1879, Sep. 2003.

[31] M. I. Yousefi and F. R. Kschischang, "A probabilistic model for optical fiber channels with zero dispersion," in *Proc. 25th Biennial Symp. on Communications*, Kingston, ON, May 2010.

[32] ——, "The per-sample capacity of zero-dispersion optical fibers," in *Proc. 12th Canadian Workshop on Information Theory*, Kelowna, BC, May 2011.

[33] L. Beygi, E. Agrell, M. Karlsson, and P. Johannisson, "Signal statistics in fiber-optical channels with polarization multiplexing and self-phase modulation," *J. Lightw. Technol.*, vol. 29, no. 16, pp. 2379–2386, Aug. 2011.

[34] L. Beygi, "Channel-aware multilevel coded modulation for coherent fiber-optic communications," Ph.D. dissertation, Chalmers University of Technology, 2013.

[35] L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, and H. Wymeersch, "A discrete-time model for uncompensated single-channel fiber-optical links," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3440–3450, Nov. 2012.

[36] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links," *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 2012.

[37] P. Poggiolini, A. Carena, V. Curri, G. Bosco, and F. Forghieri, "Analytical modeling of nonlinear propagation in uncompensated optical transmission links," *IEEE Photon. Technol. Lett.*, vol. 23, no. 11, pp. 742–744, Jun. 2011.

[38] P. Johannisson, "Analytical modeling of nonlinear propagation in a strongly dispersive optical communication system," *arXiv:1205.2193v2 [physics.optics]*, May 2012. [Online]. Available: http://arxiv.org/abs/1205.2193

[39] N. V. Irukulapati, H. Wymeersch, P. Johannisson, and E. Agrell, "Extending backpropagation to account for noise," in *Proc. European Conf. Optical Communication (ECOC)*, London, UK, 2013, p. We.3.C.4.

[40] N. V. Irukulapati, "On nonlinear compensation techniques for coherent fiber-optical channel," Licentiate Thesis, Chalmers University of Technology, 2014.

[41] D. Marsella, M. Secondini, E. Forestieri, R. Magri, G. Moruzzi, and I. Pisa, "Detection strategies in the presence of fiber nonlinear effects," in *Proc. European Conf. Optical Communication (ECOC)*, Amsterdam, NL, 2012, p. P4.06.

[42] D. Marsella, M. Secondini, and E. Forestieri, "Maximum likelihood sequence detection for mitigating nonlinear effects," *J. Lightw. Technol.*, vol. 32, no. 5, pp. 908–916, Mar. 2014.

[43] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[44] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.

[45] A. Bennatan and D. Burshtein, "Design and analysis of nonbinary ldpc codes for arbitrary discrete-memoryless channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 549–583, Feb. 2006.

[46] U. Wachsmann, R. Fischer, and J. Huber, "Multilevel codes: theoretical concepts and practical design rules," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1361–1391, Jul. 1999.

[47] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[48] D. J. Costello and G. D. Forney, Jr., "Channel coding: The road to channel capacity," *Proc. IEEE*, vol. 95, no. 6, pp. 1150–1177, Jun. 2007.

[49] B. Smith and F. R. Kschischang, "A pragmatic coded modulation scheme for high-spectral-efficiency fiber-optic communications," *J. Lightw. Technol.*, vol. 30, no. 13, pp. 2047–2053, Jul. 2012.

[50] I. B. Djordjevic, M. Arabaci, and L. L. Minkov, "Next generation FEC for high-capacity communication in optical transport networks," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3518–3530, Aug. 2009.

[51] J. Hou, P. H. Siegel, L. B. Milstein, and H. D. Pfister, "Capacity-approaching bandwidth-efficient coded modulation schemes based on low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2141–2155, Sep. 2003.

[52] G. Richter, A. Hof, and M. Bossert, "On the mapping of low-density parity-check codes for bit-interleaved coded modulation," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Nice, Italy, Jun. 2007.

[53] T. Cheng, K. Peng, J. Song, and K. Yan, "EXIT-aided bit mapping design for LDPC coded modulation with APSK constellations," *IEEE Commun. Lett.*, vol. 16, no. 6, pp. 777–780, Jun. 2012.

[54] R. Liu, P. Spasojevic, and E. Soljanin, "Reliable channel regions for good binary codes transmitted over parallel channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1405–1424, Apr. 2006.

[55] I. Sason and I. Goldenberg, "Coding for parallel channels: Gallager bounds and applications to turbo-like codes," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2394–2428, Jul. 2007.

[56] R. Gallager, "Low-density parity-check codes," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, 1963.

[57] R. G. Gallager, "Low-density parity-check codes," *IRE Trans. Inf. Theory*, vol. 8, no. 1, pp. 21–28, Jan. 1962.

[58] J. Thorpe, "Low-density parity-check (LDPC) codes constructed from protographs," *IPN Progress Report 42-154, JPL*, 2005.

[59] W. Ryan and S. Lin, *Channel Codes Classical and Modern.* Cambridge University Press, 2009.

[60] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.

[61] T. T. Richardson and R. Urbanke, *Modern Coding Theory.* Cambridge University Press, 2008.

[62] Sae-Young Chung, G. D. Forney, T. J. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 db of the Shannon limit," *IEEE Commun. Lett.*, vol. 5, no. 2, pp. 58–60, Feb. 2001.

[63] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: model and erasure channel properties," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2657–2673, Nov. 2004.

[64] S. Kudekar, T. Richardson, and R. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 803–834, Feb. 2011.

[65] D. G. M. Mitchell, M. Lentmaier, and D. J. Costello, Jr., "AWGN channel analysis of terminated LDPC convolutional codes," *Proc. Information Theory and Applications Workshop (ITA)*, 2011.

[66] A. R. Iyengar, M. Papaleo, P. H. Siegel, J. K. Wolf, A. Vanelli-coralli, and G. E. Corazza, "Windowed decoding of protograph-based LDPC convolutional codes over erasure channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2303–2320, Apr. 2012.

[67] M. Luby, M. Mitzenmacher, A. Shokrollah, and D. Spielman, "Analysis of low density codes and improved designs using irregular graphs," in *Proc. 30th Annual ACM Symp. on Theory of Computing (STOC)*, New York, USA, 1998.

[68] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[69] D. J. Costello, L. Dolecek, T. E. Fuja, J. Kliewer, D. G. M. Mitchell, and R. Smarandache, "Spatially coupled sparse codes on graphs - theory and practice," *arXiv:1310.3724v1 [cs.IT]*, Oct. 2013. [Online]. Available: http://arxiv.org/abs/1310.3724

[70] E. Agrell, "On monotonic capacity–cost functions," *arXiv:1209.2820v1 [cs.IT]*, Sep. 2012. [Online]. Available: http://arxiv.org/abs/1209.2820

# Part II

# Papers

# PAPER A

**Design of APSK Constellations for Coherent Optical Channels with Nonlinear Phase Noise**

Christian Häger, Alexandre Graell i Amat, Alex Alvarado, and Erik Agrell

*IEEE Transactions on Communications*,
vol. 61, no. 8, pp. 3362–3373, Aug. 2013

*The layout has been revised.*

**Abstract**

We study the design of amplitude phase-shift keying (APSK) constellations for a coherent fiber-optical communication system where nonlinear phase noise (NLPN) is the main system impairment. APSK constellations can be regarded as a union of phase-shift keying (PSK) signal sets with different amplitude levels. A practical two-stage (TS) detection scheme is analyzed, which performs close to optimal detection for high enough input power. We optimize APSK constellations with 4, 8, and 16 points in terms of symbol error probability (SEP) under TS detection for several combinations of input power and fiber length. For 16 points, performance gains of 3.2 dB can be achieved at a SEP of $10^{-2}$ compared to 16-QAM by choosing an optimized APSK constellation. We also demonstrate that in the presence of severe nonlinear distortions, it may become beneficial to sacrifice a constellation point or an entire constellation ring to reduce the average SEP. Finally, we discuss the problem of selecting a good binary labeling for the found constellations.

# 1 Introduction

Fiber nonlinearities are considered to be one of the limiting factors for achieving high data rates in coherent long-haul optical transmission systems [1–3]. Therefore, a good understanding of the influence of nonlinearities on the system behavior is necessary in order to increase data rates of future optical transmission systems.

The optimal design of a signal constellation, i.e., placing $M$ points in the complex plane such that the symbol error probability (SEP) is minimized under an average or peak power constraint, can be considered a classical problem in communication theory [4, Ch. 1]. The problem was addressed for example by Foschini *et al.* in the early 70s for the additive white Gaussian noise (AWGN) channel with [5] and without [6] considering a random phase jitter. However, only little is known about the influence of fiber nonlinearities on the optimal signal set. In this paper, we consider signal constellation design for a nonlinear fiber-optical channel model assuming single-channel transmission, hence neglecting interchannel impairments. We focus on a specific class of constellations called amplitude phase-shift keying (APSK), which can be defined as the union of phase-shift keying (PSK) signal sets with different amplitude levels. This choice is motivated by the fact that these constellations have long been recognized to be robust formats to cope with nonlinear amplifier distortions prevalent in satellite communication systems, see, e.g., [7–10], [11, pp. 27–28] and references therein.

The input–output relationship of the fiber-optical baseband channel is described implicitly by the stochastic nonlinear Schrödinger equation (sNLSE) [12, Ch. 2]. It is well

recognized that this type of channel model does not lend itself to an easy solution for various communication theoretic problems [13, 14]. We therefore consider a simplified, dispersionless channel model which follows from the sNLSE by neglecting the dispersive term and captures the interaction of Kerr-nonlinearities with the signal itself and the inline amplified spontaneous emission (ASE) noise, giving rise to nonlinear phase noise (NLPN) [15, 16]. A discrete channel is obtained from the waveform channel on a per-sample basis (assuming ideal carrier and timing recovery) [17, Sec. III]. This model has been previously considered by several authors in the literature and different methods have been applied to derive the joint probability density function (PDF) of the received amplitude and phase [3, 16–18]. Since all these derivations neglect dispersion, the resulting PDF should serve as a useful approximation for dispersion-managed (DM) optical links, provided that the local accumulated dispersion is sufficiently low [3, p. 160], [19]. However, if the interaction between dispersion and nonlinearities becomes too strong, the channel model is likely to diverge from the one assumed here.[1] We point out that several studies have addressed the influence of dispersion on the variance of NLPN in the context of DM links using linearization techniques [22–26]. In [27] a comprehensive study on quantifying the parameter space where nonlinear signal-noise interactions including NLPN are dominant impairments for different modulation formats was presented. A brief discussion on the applicability of the assumed channel model in the context of DM links is also provided in [27]. An extensive literature review on the topic of NLPN is included in [25].

Signal constellation design and detection assuming the same channel as here has been studied previously in [28–30]. In [28], the authors applied several predistortion and post-compensation techniques in combination with minimum-distance detection for quadrature amplitude modulation (QAM) to mitigate the effect of NLPN. They also proposed a two-stage (TS) detector consisting of a radius detector, an amplitude-dependent phase rotation, and a phase detector. Moreover, parameter optimization was performed with respect to four 4-point, custom constellations under maximum likelihood (ML) detection. In [29], the TS detector was used to optimize the radii of four 16-point constellations for two power regimes. It was shown that the optimal radii highly depend on the transmit power. In [30], the SEP of $M$-PSK was studied assuming a minimum-distance detector. In [14], a capacity analysis is presented for fiber-optical channels. The authors use bivariate Gaussian PDFs to represent the discrete-time channel where the covariance matrices are obtained through extensive numerical simulation. Continuous-input ring constellations are used to exploit the assumed rotational invariance of the channel and subsequently find lower bounds on the maximum achievable information rates. For the same channel model, in [31] the occupancy frequency and spacing of the ring constellation were optimized. Related work was presented in [32], where the channel output PDF is approximated through numerically obtained histograms and optimized ring con-

[1] As an extreme case, for dispersion-uncompensated links, it was found that the channel is well-modeled by a Gaussian PDF [20, 21].

stellations are found. Discrete constellations are then obtained through quantization. A similar quantization technique was applied in [33].

In this paper, we first analyze the (suboptimal) TS detector developed in [28]. We regard radius detection and phase rotation as a separate processing block that yields a postcompensated observation and we explain how to derive the corresponding PDF for constellations with multiple amplitude levels. To the best of our knowledge, this PDF has not been previously derived, possibly due to the fact that the SEP under TS detection can be calculated with a simplified PDF [34]. The new PDF is used to gain insights into the performance behavior of the TS detector compared to optimal detection. We also show that this PDF is necessary to accurately calculate the average bit error probability (BEP) of the constellation.

We then find optimal APSK constellations in terms of SEP under TS detection for a given input power and fiber length. In contrast to [29], we optimize the number of rings, the number of points per ring, as well as the radii. For the case $M = 4$, we choose identical system parameters as in [28] and a comparison reveals that our approach results in similar, sometimes better, constellations, with the advantage of much less computational design complexity. This allows us to extend the optimization to $M = 8$ and $M = 16$. For the latter case, our results show that the widely employed 16-QAM constellation has poor performance compared to the best found constellations over a wide range of input powers for this channel model and detector. We also provide numerical support for the phenomenon of *sacrificial points* or *satellite constellations*, which arise in the context of constellation optimization in the presence of very strong nonlinearities [35,36]. Our findings show, somewhat counterintuitively, that it is sometimes optimal to place a constellations point (or even an entire constellation ring) far away from all other points in order to improve the average performance of the constellation.

Due to the separation of a hard-decision symbol detector and subsequent error correction in state-of-the-art fiber-optical communication systems, the uncoded BEP is an important figure of merit. Therefore, we also address the problem of choosing a good binary labeling for APSK constellations in the presence of NLPN. We pay special attention to a class of APSK constellation which allows the use of a Gray labeling, which we call rectangular APSK. For this class, we propose a method to choose the phase offsets of the constellation rings resulting in near-optimal performance. The proposed method might also be useful when soft information is passed to a decoder in the form of bit-wise log-likelihood ratios in a bit-interleaved coded modulation (BICM) scheme. For BICM, it is known that the labeling can have a significant impact on the achievable information rate and the system performance [37].

The remainder of the paper is organized as follows. In Sec. 2, we present the channel model and define the generic APSK signal set. In Sec. 3, we briefly review ML detection and then describe and analyze the suboptimal TS detector together with the corresponding PDF. The results of the constellation optimization with respect to SEP are presented and discussed in Sec. 4. Binary labelings are discussed in Sec. 5. Concluding remarks

**Table 1:** Constants and Parameters taken from [28]

| symbol | value | meaning |
|---|---|---|
| $\gamma$ | $1.2\ \mathrm{W}^{-1}\mathrm{km}^{-1}$ | nonlinearity parameter |
| $n_{sp}$ | 1.41 | spontaneous emission factor |
| $h$ | $6.626 \cdot 10^{-34}$ J s | Planck's constant |
| $\nu$ | $1.936 \cdot 10^{14}$ Hz | optical carrier frequency |
| $\alpha$ | $0.0578\ \mathrm{km}^{-1}$ | fiber loss (0.25 dB/km) |
| $\Delta\nu$ | 42.7 GHz | optical bandwidth |

can be found in Sec. 6.

# 2 System Model

## 2.1 Channel

We consider the discrete memoryless channel [3, Ch. 5]

$$Y = (X + Z)e^{-\jmath\Phi_{\mathrm{NL}}}, \tag{A.1}$$

where $\jmath \triangleq \sqrt{-1}$ denotes the imaginary unit, $X \in \mathcal{X}$ the complex channel input, $\mathcal{X}$ the signal constellation, $Z$ the total additive noise, $Y$ the channel observation, and $\Phi_{\mathrm{NL}}$ the NLPN, which is given by [3, Ch. 5]

$$\Phi_{\mathrm{NL}} = \frac{\gamma L}{K} \sum_{i=1}^{K} |X + Z_i|^2. \tag{A.2}$$

In (A.2), $\gamma$ is the nonlinear Kerr-parameter, $L$ is the total length of the fiber, $K$ denotes the number of fiber segments, and $Z_i$ is the noise contribution of all fiber segments up to segment $i$. More precisely, $Z_i \triangleq N_1 + \ldots + N_i$ is defined as the sum of $i$ independent and identically distributed complex Gaussian random variables with zero mean and variance $\sigma_0^2$ per dimension (real and imaginary parts). The total additive noise of all $K$ fiber segments is denoted by $Z \triangleq Z_K$, which has variance $\sigma^2 \triangleq \mathbb{E}\left[|Z|^2\right] = 2K\sigma_0^2$, where $\mathbb{E}[\,\cdot\,]$ is the expected value. For ideal distributed amplification, we consider the case $K \to \infty$. The total noise variance can be calculated as $\sigma^2 = 2n_{sp}h\nu\alpha\Delta\nu L$ [14, Sec. IX-B], where all parameters are taken from [28] and are summarized in Table 1. The additive noise power spectral density as defined in [14] is then given by $N_0 \triangleq n_{sp}h\nu\alpha = 1.04 \cdot 10^{-20}$ W/km/Hz. Note that the total additive noise variance scales linearly with the fiber length.

An important aspect of this channel model is the fact that the variance of the phase noise is dependent on the channel input (cf. (A.2)), or equivalently on the average transmit power $P$, defined as $P \triangleq \mathbb{E}\left[|X|^2\right]$. In Fig. 1 we show received scatter plots for $Y$ (cf. (A.1)) assuming $X \in \mathcal{X}_{\text{16-QAM}}$, where $\mathcal{X}_{\text{16-QAM}} \triangleq \{\sqrt{P/10}(a + \jmath b) \ : \ a, b \in \{\pm 1, \pm 3\}\}$ is

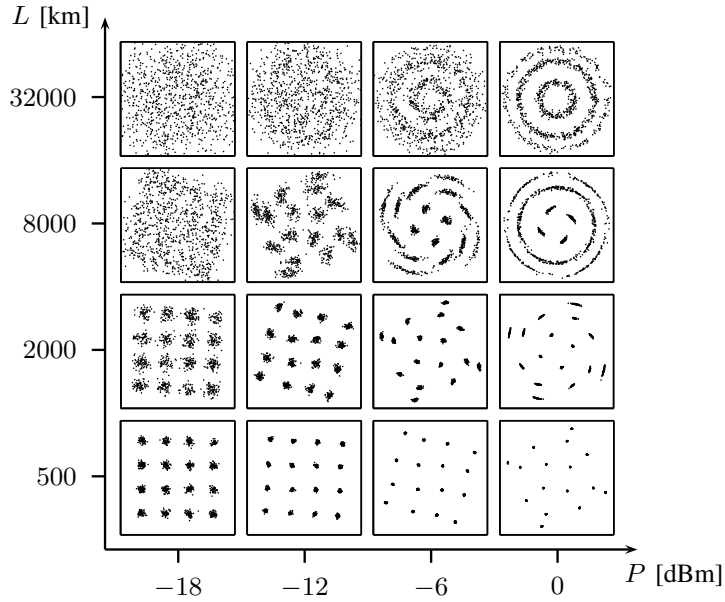



**Figure 1:** Scatter plots of $Y$ assuming $X \in \mathcal{X}_{\text{16-QAM}}$ for several combinations of transmit power $P$ and fiber length $L$.

the 16-QAM constellation, and $K = 100$ for several combinations of $P$ and $L$. The purpose of Fig. 1 is to gain insight into the qualitative behavior of the channel. It can be observed that for very low input power and fiber length, nonlinearities are negligible and the channel behaves as a standard AWGN channel. The scatter plots along a diagonal in Fig. 1 correspond to a constant signal-to-noise ratio (SNR), defined as $\text{SNR} \triangleq P/\sigma^2$. In contrast to an AWGN channel for which the scatter plots along any diagonal would look similar, the received constellation points in Fig. 1 start to rotate in a deterministic fashion and the effect of the NLPN becomes pronounced for large $L$ and $P$. Therefore, in order to specify the operating point of the channel, the SNR alone is not sufficient, because the parameter space of the channel is two-dimensional, cf. [17, Sec. VII]. In this paper, we present performance results assuming a fixed fiber length and variable transmit power.

## 2.2 Amplitude-Phase Shift Keying

We use the term APSK for the discrete-input constellations considered in this paper and focus on constellations with $M = 4$, 8, and 16 points. The APSK signal set is defined as [9]

$$\mathcal{X} \triangleq \left\{ r_k e^{\jmath\left(\frac{2\pi j}{l_k}+\varphi_k\right)} \;:\; 1 \le k \le N, \quad 0 \le j \le l_k - 1 \right\}, \tag{A.3}$$

where $N$ denotes the number of amplitude levels or rings, $r_k$ the radius of the $k$th ring, $l_k \geq 1$ the number of points in the $k$th ring, where $\sum_{k=1}^{N} l_k = M$, and $\varphi_k$ the phase offset in the $k$th ring. Throughout the paper, we assume a uniform distribution on the channel input $X$ over all symbols, and thus $P = (1/M)\sum_{k=1}^{N} l_k r_k^2$. The radii are assumed to be ordered such that $r_1 < \ldots < r_N$ and we use $\boldsymbol{r} \triangleq (r_1, \ldots, r_N)$ to denote the *radius vector*. In this paper, for $l_1 = 1$, the point in the first ring is always placed at the origin, implying $r_1 = 0$. The radius vector is said to be uniform if $r_{k+1} - r_k = \Delta$ for $1 \leq k \leq N-1$, where $\Delta = r_2$ if $l_1 = 1$ and $\Delta = r_1$ if $l_1 \geq 2$. The symbols are assumed to be indexed, i.e., $x_i \in \mathcal{X}$, $i = 1, \ldots, M$. The indexing is done starting in the innermost ring ($k = 1$) by increasing $j$ from 0 to $l_1 - 1$ and then moving to the next ring increasing $j$ from 0 to $l_2 - 1$ and so on. Thus, finally we have $x_1 = r_1 e^{j\varphi_1}, \ldots, x_M = r_N e^{j2\pi(l_{N-1})/l_N + \varphi_N}$.

We also define the vectors $\boldsymbol{l} \triangleq (l_1, \ldots, l_N)$ and $\boldsymbol{\varphi} \triangleq (\varphi_1, \ldots, \varphi_N)$, and use the notation $\boldsymbol{l}$-APSK for an APSK constellation with $N$ rings and $l_k$ points in the $k$th ring, e.g., (4,4,4,4)-APSK. Note that this notation does not specify a particular constellation without ambiguity, due to the missing information about the radii and phase offsets.

# 3 Detection Methods

## 3.1 Symbol Error Probability

Let $\mathcal{R}_i$, $1 \leq i \leq M$, be the decision region implemented by a detector for the symbol $x_i$, i.e., $\hat{X} = x_i$ if $Y \in \mathcal{R}_i$, where $\hat{X}$ denotes the detected symbol. The average SEP is then

$$\mathrm{SEP} = 1 - \frac{1}{M}\sum_{i=1}^{M} P_{i\to i}, \tag{A.4}$$

where $P_{i\to j} \triangleq \Pr[\hat{X} = x_j | X = x_i]$, $1 \leq i, j \leq M$, are the symbol transition probabilities[2] calculated as

$$P_{i\to j} = \int_{\mathcal{R}_j} f_{Y|X=x_i}(y)\,\mathrm{d}y. \tag{A.5}$$

That is, $P_{i\to j}$ is obtained through integration of the conditional PDF of the observation given the channel input $X = x_i$ over the detector region for $x_j$.

## 3.2 Maximum Likelihood Detection

Let the polar representation of the channel input and the observation be given by $X = R_0 e^{j\Theta_0}$ and $Y = Re^{j\Theta}$, respectively. The PDF of $Y$ can be written in the form [17, Sec. III], [3, p. 225], [28]

$$f_{Y|X=x}(y) = \frac{f_{R|R_0=r_0}(r)}{2\pi r} + \frac{1}{\pi r}\sum_{m=1}^{\infty} \Re\left\{C_m(r, r_0)e^{jm(\theta-\theta_0)}\right\}, \tag{A.6}$$

[2]For the SEP, only the cases $j = i$, $1 \leq i \leq M$, need to be considered. In Sec. 5, all transition probabilities are used.

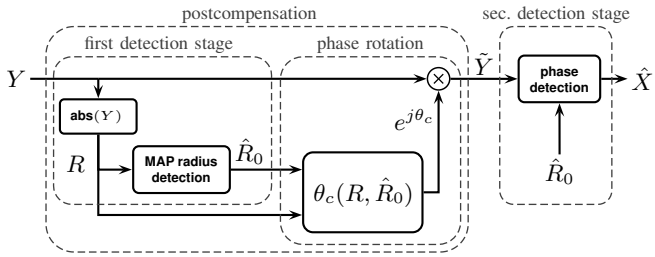



**Figure 2:** Block diagram of the TS detector. Note that the depicted postcompensation of $Y$ to $\tilde{Y}$ is reversible.

where $x = r_0 e^{\jmath\theta_0}$, $y = re^{\jmath\theta}$, $\Re\{z\}$ is the real part of $z \in \mathbb{C}$, and the PDF of the received amplitude $R$ given the transmitted amplitude $R_0 = r_0$ is

$$f_{R|R_0=r_0}(r) = \frac{2r}{\sigma^2} \exp\left(-\frac{r^2 + r_0^2}{\sigma^2}\right) \mathrm{I}_0\left(\frac{2rr_0}{\sigma^2}\right), \tag{A.7}$$

where $\mathrm{I}_0(\,\cdot\,)$ is the modified Bessel function of the first kind. Analytical expressions for the coefficients $C_m(r, r_0)$ can be found in [17, Sec. III]. The ML detector can now be described in the form of decision regions $\mathcal{R}_i^{\mathrm{ML}} \subset \mathbb{C}$ for each symbol $x_i \in \mathcal{X}$:

$$\mathcal{R}_i^{\mathrm{ML}} \triangleq \bigcap_{\substack{j=1 \\ j\neq i}}^{M} \{y \in \mathbb{C} \;:\; f_{Y|X=x_i}(y) \geq f_{Y|X=x_j}(y)\}. \tag{A.8}$$

If we take the ML decision regions defined in (A.8), then (A.4) becomes a lower bound on the achievable SEP with suboptimal detectors.

Evaluating the SEP by numerically integrating over the PDF (A.6) is computationally expensive. Moreover, unlike for an AWGN channel, where the ML regions simply scale proportionally to $\sqrt{P}$, the ML decision regions defined in (A.8) change their shape based on the transmit power $P$ [28]. This renders ML detection rather impractical for the purpose of constellation optimization.

## 3.3 Two-Stage Detection

In this paper, we study a slightly modified version of the suboptimal TS detector proposed in [28]. This detector is a practical alternative to the ML detector because it has much lower complexity. In particular, the TS detector employs one-dimensional decisions: First in the amplitude direction (first detection stage), followed by a phase rotation, and then in the phase direction (second detection stage).

In Fig. 2, we show a block diagram of the TS detector. We refer to the first detection stage together with the phase rotation as *postcompensation*. Based on the absolute value

of the observation $R$, radius detection is performed. In contrast to [28] and [29], we use maximum a posteriori (MAP) instead of ML radius detection and make use of the a priori probability for selecting a certain ring at the transmitter, thereby achieving a small performance advantage. The radius detector implements the rule: Choose $\hat{R}_0 = r_k$, when $\mu_{k-1} \leq R < \mu_k$, where $\mu_k$, $0 \leq k \leq N$, denote the decision radii or thresholds. The MAP decision threshold $\mu_k$, $1 \leq k \leq N-1$, between $r_k$ and $r_{k+1}$ is obtained by solving

$$\Pr[R_0 = r_k] f_{R|R_0=r_k}(\mu_k) = \Pr[R_0 = r_{k+1}] f_{R|R_0=r_{k+1}}(\mu_k), \tag{A.9}$$

where the a priori probabilities are given by $\Pr[R_0 = r_k] = l_k/M$.[3] We always define $\mu_0 \triangleq 0$ and $\mu_N \triangleq \infty$. Based on the radius $\hat{R}_0$ of the detected ring and the received amplitude $R$, a correction angle $\theta_c$ is calculated, by which the observation $Y$ is rotated to obtain the postcompensated observation $\tilde{Y}$ as shown in Fig. 2. The correction angle is given by

$$\theta_c(R, \hat{R}_0) = -\angle C_1(R, \hat{R}_0), \tag{A.10}$$

which is approximately a quadratic function in $R$ [28].

The second detection stage is then performed with respect to $\tilde{Y}$: A phase detector chooses the constellation point with radius $\hat{R}_0$ that is closest to $\tilde{Y}$. Graphically, the TS detector employs so called annular sector regions (or annular wedges) as decision regions for $\tilde{Y}$.

## 3.4 PDF of the Postcompensated Observation

It is shown in [28] that for PSK signal sets (which, in this paper, are denoted by $(M)$-APSK) where $\hat{R}_0 = \sqrt{P} = \text{const.}$, a minimum-distance detector for $\tilde{Y}$ is equivalent to ML detection. In contrast, for constellations with multiple amplitude levels, the receiver structure in Fig. 2 does not perform ML detection. However, in principle, optimal detection of $X$ is still possible based on $\tilde{Y}$ due to the fact that the postcompensation in Fig. 2 is invertible and every invertible function forms a sufficient statistic for detecting $X$ based on $Y$ [38, p. 443]. Thus, the performance loss associated with the TS detection scheme originates solely from suboptimal detection regions, not from the postcompensation itself, which is a lossless operation.[4]

In the following, we show how the PDF of the postcompensated observation $\tilde{Y}$ can be computed. This PDF can then be used to find optimal detection regions for $\tilde{Y}$. It is clear from the block diagram of Fig. 2 that the PDF can be written as

$$f_{\tilde{Y}|X=x}(\tilde{y}) = f_{Y|X=x}\left(\tilde{y} \cdot e^{-\jmath\theta_c(R,\hat{R}_0)}\right). \tag{A.11}$$

[3] Solving (A.9) for $\mu_k$ can be done numerically and for an approximate analytical solution assuming that $r_k \neq 0$, one may apply the high-SNR approximation $\mathrm{I}_0(x) \approx \frac{e^x}{\sqrt{2\pi x}}$ as was done in [29] for the ML radius detector.

[4] An important question that we do not address is whether the phase rotation (A.10) is still the best choice for multilevel constellations, assuming that one is constrained to straight-line phase decision boundaries for $\tilde{Y}$.

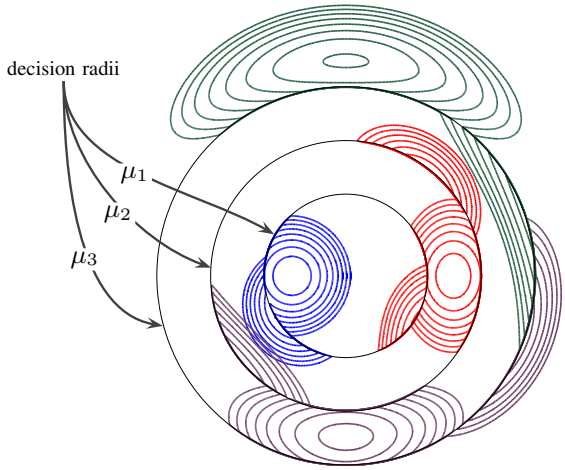



**Figure 3:** PDF of $\tilde{Y}$ for $P = -5$ dBm and $L = 5500$ km conditioned on one particular point in each ring of the uniform (4,4,4,4)-APSK constellation. Color is helpful.

Most importantly, the correction angle $\theta_c$ is a discontinuous function of the amplitude $R$ because it depends on the detected ring $\hat{R}_0$. In general, the correction angle can be written as

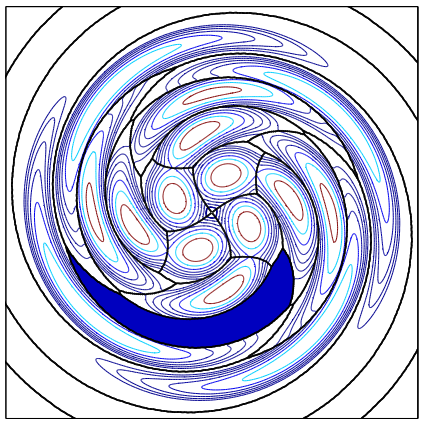
$$\theta_c(R, \hat{R}_0) = \begin{cases} \theta_c(R, r_1) & \text{if } \mu_0 \leq R < \mu_1 \\ \vdots & \vdots \\ \theta_c(R, r_N) & \text{if } \mu_{N-1} \leq R < \mu_N \end{cases} . \tag{A.12}$$

For illustration purposes, we plot in Fig. 3 the PDF resulting from (A.11) and (A.12) conditioned on one particular point in each ring of the uniform $(4, 4, 4, 4)$-APSK constellation[5]. If we consider the PDF conditioned on $X = r_2$, i.e., $R_0 = r_2$ and $\Theta_0 = 0$ (shown in red), it can be observed that the contour lines look as though they have been sliced up along the decision radii of the radius detector. For $R < \mu_1$, the correction angle is calculated with respect to $r_1$, and thus, the phase is undercompensated. On the other hand, for $R \geq \mu_2$, the wrongly detected radius results in an overcompensation.

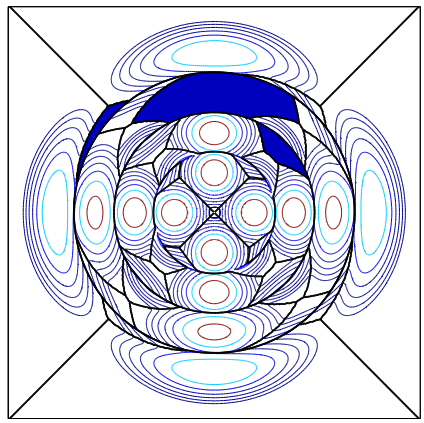
## 3.5 Performance Comparison

For a qualitative performance comparison between the different detectors, in Fig. 4(a), the PDF in (A.6) is plotted for the uniform $(4, 4, 4, 4)$-APSK constellation together with the ML decision regions. In Fig. 4(b), the PDF in (A.11) is used instead. Finally, Fig. 4(c) shows the same PDF as Fig. 4(b) together with the suboptimal decision regions

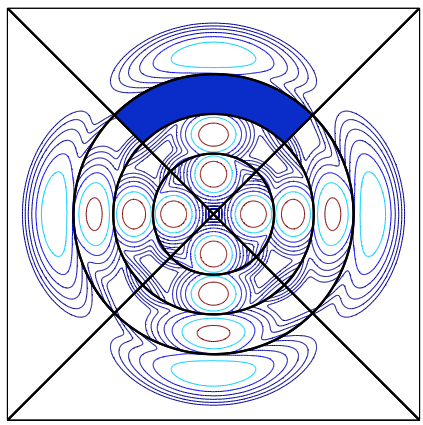
[5]The PDFs of the points which are not shown look identical to the PDF of the corresponding point in the same ring up to a phase rotation.



**(a)** ML decision regions w.r.t. $Y$



**(b)** ML decision regions w.r.t. $\tilde{Y}$



**(c)** TS decision regions

**Figure 4:** Decision regions for the uniform $(4, 4, 4, 4)$-APSK constellation at $P = -4$ dBm. Shaded regions correspond to the point $X = r_3 e^{\jmath\pi/2}$.

implemented by the TS detector. As an example, the decision regions corresponding to the point $X = r_3 e^{\jmath\pi/2}$ are shaded in Fig. 4.

Comparing the optimal decision regions in Fig. 4(b) with the decision regions in Fig. 4(c), it can be seen that TS detection is clearly suboptimal for this constellation and input power. However, one would expect the two small shaded regions in Fig. 4(b) to become smaller for higher power. Intuitively, this is explained by the increasing accuracy of the radius detector for increasing transmit power, due to the Rice distribution (A.7) of the amplitude. More precisely, let $P_k^{(e)} \triangleq \Pr[\hat{R}_0 \neq R_0 | R_0 = r_k]$ be the probability of

an error in the first detection stage, given that a symbol in the $k$th ring is transmitted. Then

$$P_k^{(e)} = 1 - \left(Q_1\left(\tilde{r}_k, \tilde{\mu}_{k-1}\right) - Q_1\left(\tilde{r}_k, \tilde{\mu}_k\right)\right), \tag{A.13}$$

where $Q_1(\cdot, \cdot)$ is the Marcum Q-function, $\tilde{r}_k \triangleq \sqrt{2} r_k/\sigma$, and $\tilde{\mu}_k \triangleq \sqrt{2}\mu_k/\sigma$. It follows that the SEP under TS detection converges to the SEP under ML detection for increasing input power and any APSK constellation with uniform radius vector, since then $P_k^{(e)}$, $1 \leq k \leq N$, tends to zero as $P$ increases.

# 4 Constellation Optimization

## 4.1 Problem Statement

In this section, we optimize the parameters of APSK constellations by minimizing the SEP under TS detection. Formally, the optimization problem can be stated as: Given $M$, $P$, and $L$,

$$\begin{aligned}
\underset{\boldsymbol{l},\boldsymbol{r}}{\text{minimize}} \quad & \text{SEP under TS detection} && \text{(A.14)}\\
\text{subject to} \quad & 1 \leq N \leq M && \text{(A.15)}\\
& l_1 + \ldots + l_N = M && \text{(A.16)}\\
& l_1 r_1^2 + \ldots + l_N r_N^2 = PM && \text{(A.17)}\\
& l_k \geq 1, \text{for } 1 \leq k \leq N. && \text{(A.18)}
\end{aligned}$$

The objective function can be computed analytically using (A.4) with the PDF (A.11) integrated over the TS detector regions [34, Eq. (4)]. Note that the phase offset vector $\boldsymbol{\varphi}$ does not appear in the minimization problem. This is due to the fact that the SEP under TS detection does not change assuming a phase offset in any of the constellation rings: The PDF of $\tilde{Y}$ is simply rotated by the phase offset, but so is the detector region itself, and hence the integrals in (A.5) are not affected.

It is instructive to begin by discussing two special cases of the general optimization problem above. The first case is obtained when $\boldsymbol{r}$ is assumed to be uniform and an optimization is performed only over the number of points in each ring $\boldsymbol{l}$, cf. (A.14)−(A.18). The optimization problem then becomes an integer program which can be solved in an exhaustive fashion for the constellation sizes considered in this paper. The number of ways to distribute $M$ constellation points to $i$ rings is given by $\binom{M-1}{i-1}$. At most, there are $M$ rings, which gives a total of $\sum_{i=1}^{M} \binom{M-1}{i-1} = 2^{M-1}$ possibilities to choose $\boldsymbol{l}$. There are 8, 128, and 32768 possibilities for 4, 8, and 16 points, respectively. It is clear that such a brute-force approach becomes unfeasible for larger constellations. However, based on the obtained results, it might be possible to devise more sophisticated search methods for $M > 16$, e.g., by neglecting unrealistic combinations a priori.

The second special case is given by the optimization of the radius vector $\boldsymbol{r}$ for a certain constellation with fixed parameter $\boldsymbol{l}$. For this case, the optimization problem becomes a

nonlinear program. Due to the power constraint, the dimensionality of the search space is $N-1$ if $l_1 > 1$ and $N-2$ if $l_1 = 1$, respectively. By inspecting the target function, one can verify that this problem is nonconvex, i.e., a local optimum does not necessarily imply a global solution. We tested different nonlinear solvers and obtained very good solutions with the Nelder–Mead simplex method [39]. Even though the solution is not guaranteed to be the globally optimal, we verified the global optimality for several constellations and several combinations of input power and fiber length with a brute-force grid search.

By combining the solutions to the two special cases above, a solution to the general optimization problem can then be found as follows: For each APSK constellation with a certain fixed parameter $\boldsymbol{l}$ one determines the optimal radius vector via the simplex method for a given input power and fiber length, and then optimizes over all possible $\boldsymbol{l}$. We call this approach *joint optimization*.

## 4.2 Results and Discussion

### 4 Points

We start by finding optimal APSK constellations with $M = 4$ points. The fiber length is fixed at $L = 7000$ km and the input power $P$ is varied from $-15$ dBm to 5 dBm in steps of 0.5 dBm. In Fig. 5 we plot the performance of all possible eight 4-point APSK constellations with optimal radius vector (dashed lines) and the curves are labeled with the corresponding $\boldsymbol{l}$. (For (4)-APSK and $(1, 3)$-APSK the radius vector is fixed, i.e., no optimization is performed.) The results of the joint optimization are shown with markers. We also show the SEP of (4)-APSK (or 4-QAM) in an AWGN channel under ML detection as a well-known reference curve (dash-dotted line). Note that for each $P$ the SEP is minimized by a constellation with certain parameters $\boldsymbol{l}$ and $\boldsymbol{r}$. For example, it can be seen that for an input power range between $-15$ dBm and $-7.5$ dBm, (4)-APSK is optimal.

Based on the behavior of the SEP for the individual APSK constellations with optimized radius vector (dashed lines in Fig. 5), it is possible to classify the constellations into three classes. The first class exhibits a well known performance minimum, i.e., an optimal operating power. The second class does not exhibit a minimum, but eventually flattens for very high input power (see, e.g., the performance of $(3, 1)$-APSK). The third class exhibits a performance behavior which is strictly and steadily decreasing with increasing input power.

The flattening of the SEP for the second class is explained by the availability of a *sacrificial point* in the outer constellation ring. The meaning of the term *sacrifical* is best explained with the help of an example. In Fig. 6 we show the results of the radius optimization for $(3, 1)$-APSK (top), together with the optimal values of the parameter $\boldsymbol{r}$ (bottom). It can be observed that, for $P > -3$ dBm, the optimal ring spacing shows a very peculiar behavior. For this power regime, $r_1$ appears to be fixed and any increase in average power is absorbed by putting the outermost point further away from the inner

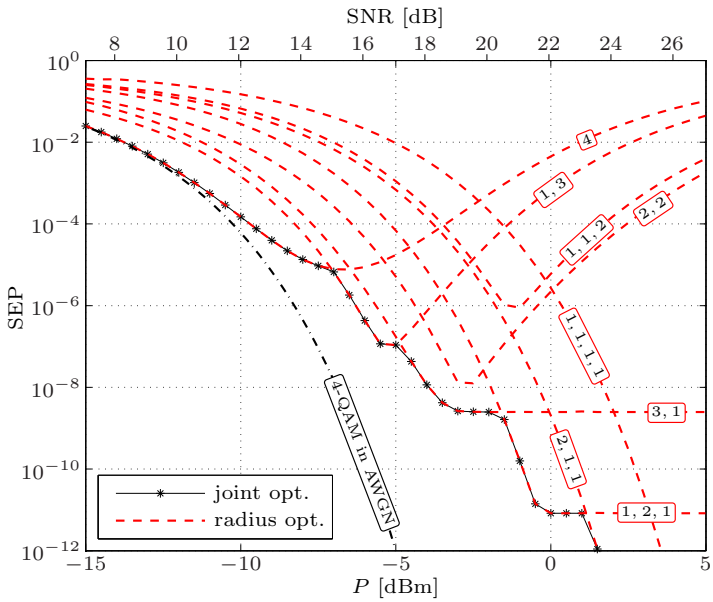



**Figure 5:** Results for the constellation optimization with $M = 4$. The fiber length is $L = 7000$ km. The dash-dotted line is a reference curve, showing the SEP of 4-QAM in an AWGN channel under ML detection.

ring. In some sense, the outer point (experiencing very high NLPN) is sacrificed with the result of saving the average SEP of the constellation. Observe that $(1, 1, 1, 1)$-APSK is the only APSK constellation with 4 points that belongs to the third class[6] and it was already argued in [28], that this constellation is optimal for very high input power.[7]

The system parameters are chosen in such a way that the obtained results are directly comparable to the performance of the optimized constellations presented in [28, Sec. IV]. To facilitate a comparison, in Fig. 7 we provide a digitalized version of [28, Fig. 15] and plot the outcome of the joint optimization in the same figure. All four constellations used in [28] can be seen as APSK constellations and they are depicted in Fig. 7 for convenience. With the notation introduced in this paper, the constellations are (4)-APSK, $(1, 3)$-APSK, $(2, 2)$-APSK with $\boldsymbol{\varphi} = (0, \varphi_2)$, and $(1, 2, 1)$-APSK with $\boldsymbol{\varphi} = (0, 0, \varphi_3)$. The parameters $\varphi_2$ and $\varphi_3$ are determined by a precompensation technique developed in [28], while the radius vector of the two latter constellations was optimized. It is important to point out that the optimization in [28] was performed with respect to ML detection, while for the joint optimization in this paper the suboptimal TS detector is assumed.

[6] The SEP for $(2, 1, 1)$-APSK in Fig. 5 flattens for very high input power.

[7] For APSK constellations with only one point in each ring the SEP under TS detection can be calculated using (A.13) as $\text{SEP} = \frac{1}{M} \sum_{k=1}^{M} P_k^{(e)}$.

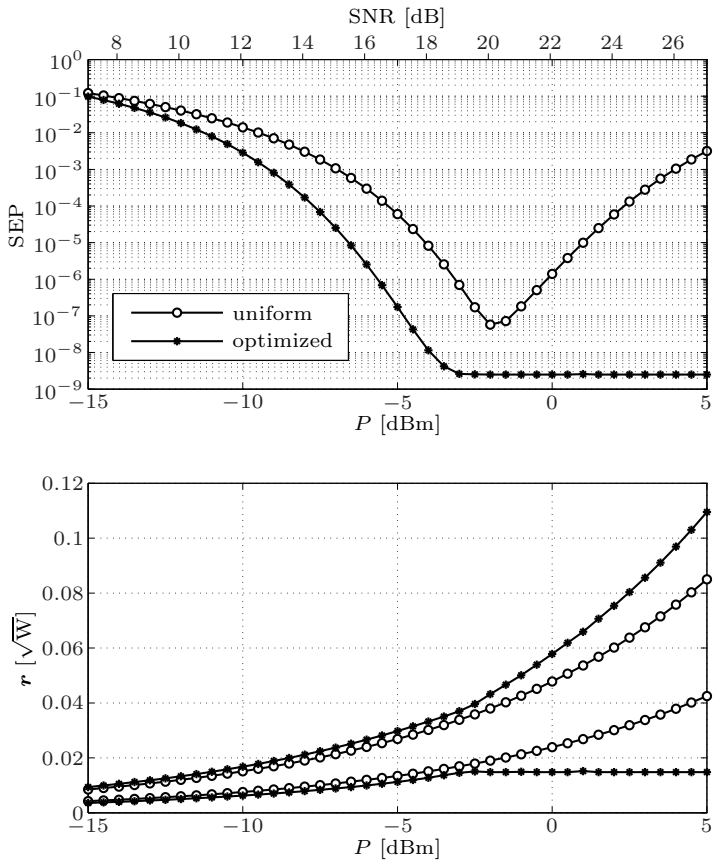



**Figure 6:** Performance of the $(3, 1)$-APSK constellation with a uniform and optimized radius vector (top) and the corresponding radius vector (bottom).

Notice that, for (4)-APSK these two detection schemes are equivalent and hence, the performance results taken from [28] for (4)-APSK (constellation $(a)$) in Fig. 7 overlap with the results of the joint optimization for an input power up to $-7.5$ dBm. Comparing the results in Fig. 7, it can be seen that the jointly optimized APSK constellations perform very close to the optimized constellations in [28]. For certain input powers, e.g., $-7$ dBm or $-5$ dBm, a performance loss is visible, which is explained by the weaker detection technique. On the other hand, for some power regimes, e.g., $-3.5$ dBm or $-2$ dBm, the jointly optimized constellations perform as well as, or even outperform, the best constellations presented in [28]. We attribute this performance gain to the more systematic search which is offered by the APSK framework. Also note that there is no need to find phase precompensation angles as was done in [28], because those are relevant only for ML detection, but irrelevant for the performance under TS detection. This

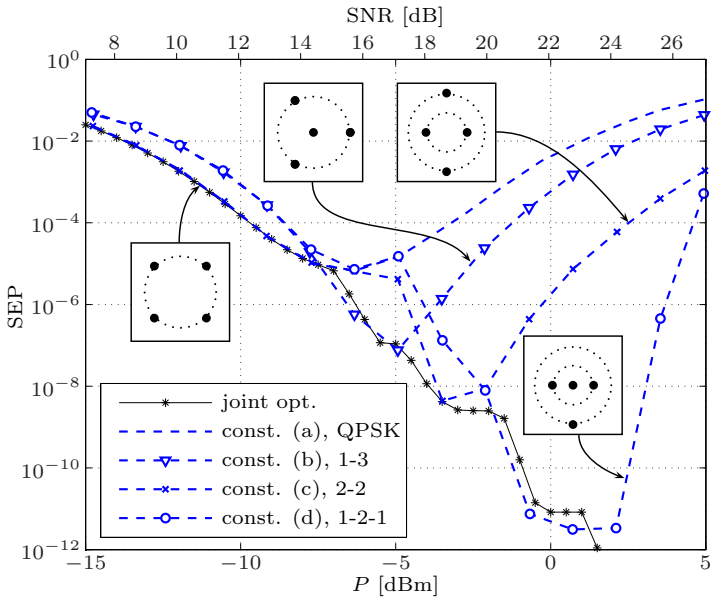



**Figure 7:** Comparison between the joint optimization of the 4-point APSK constellation and the results for optimized constellations (a)–(d) in [28] under ML detection. The legend shows the constellation names originally used in [28].

removes one degree of freedom from choosing a constellation and makes the optimization simpler. As a last point, it is unclear why constellation (d) in [28] does not exhibit a flattening SEP for very high input power, even though the radius vector was optimized and a sacrificial point is available. We conjecture that the SEP results for constellation (d) for 3 dBm and 5 dBm are only locally optimal.

## 8 Points

For $M = 8$ points, we present optimization results for the same system parameters as before. The results for the joint optimization are shown in Fig. 8 with markers. Since it is not very instructive to show the performance of all 8-point APSK constellations in the same figure, we only plot the SEP of those constellations that are optimal somewhere in the considered power range (dashed lines). Thus, the parameter $\boldsymbol{l}$ is indicated by the labeling of the corresponding dashed line. To avoid cumbersome notation, we also define $\mathbf{1}_N \triangleq 1, \ldots, 1$ ($N$ times). To get a more intuitive feeling for the optimal constellations in different power regimes, the inset figures show the actual constellations that are optimal at $-12$ dBm, $-7.5$ dBm, $-4.5$ dBm, and $-1$ dBm. The constellations are shown with their optimized radius vector for the corresponding input power.

We also perform an optimization only over $\boldsymbol{l}$ assuming that the radius vector is uniform.

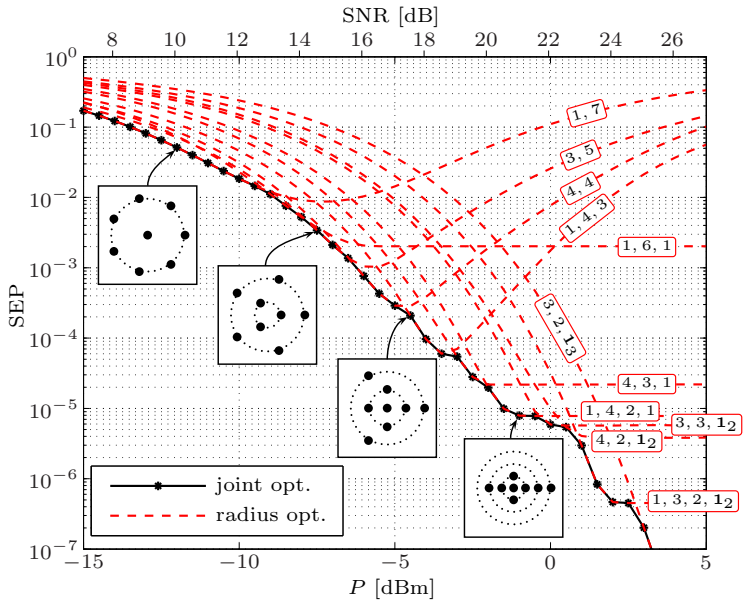



**Figure 8:** Results for the constellation optimization with $M = 8$. The fiber length is $L = 7000$ km.

The results are depicted in Fig. 9. To facilitate a better comparison with the jointly optimized constellations, the solid black line in Fig. 8 is again plotted in Fig. 9. The dashed lines in Fig. 9 show the SEP of the APSK constellation with a fixed parameter $\boldsymbol{l}$ and assuming a uniform radius vector. An important observation here is that up to an input power of $-3.5$ dBm, there is almost no difference between the performance of the jointly optimized constellations and the optimal constellations obtained assuming a uniform radius vector, suggesting that most of the performance improvement is due to optimizing the parameter $\boldsymbol{l}$.

### 16 Points

Motivated by the results obtained for $M = 8$, for $M = 16$ points, we limit ourselves to the case where the radius vector is assumed to be uniform for all constellations. The fiber length is $L = 5500$ km and the input power $P$ is varied from $-14$ dBm to 10 dBm in steps of 2 dBm. The results are shown in Fig. 10, where we indicate the optimal parameter $\boldsymbol{l}$ next to the corresponding marker of the curve. The individual SEP under TS detection for 16-QAM, uniform $(4, 4, 4, 4)$-APSK, and $(\mathbf{1}_{16})$-APSK are shown for comparison, while the SEP under ML detection for 16-QAM in an AWGN channel is shown as a reference. The results in Fig. 10 show that up to an input power of $-8$ dBm, the performance of the optimized constellations follows closely the performance of 16-QAM in AWGN. In other

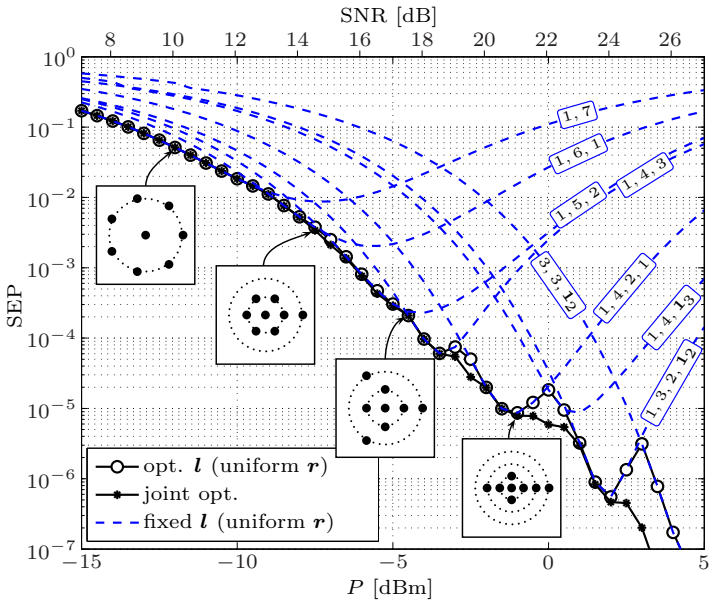



**Figure 9:** Results for the constellation optimization with $M = 8$ (white circles) assuming that a uniform radius vector for all constellations. The fiber length is $L = 7000$ km. The results of the joint optimization from Fig. 8 (black markers) are shown for comparison. The dashed lines correspond to the SEP of the constellations that are indicated by the labels.

words, for this power regime it is possible to find APSK constellations with TS detection that perform as well as 16-QAM for a channel without nonlinear impairments under ML detection. For higher input power, the optimized constellations gradually utilize more amplitude levels, due to the increase in NLPN. If we take as a baseline the minimum SEP achieved by the 16-QAM constellation ($P \approx -2.8$ dBm and SEP $\approx 10^{-2}$), and interpolate the optimal APSK performance for the same SEP, we observe that a performance gain of 3.2 dBm is achieved.

In order to verify the assumption that a joint optimization approach does not lead to significant performance gains, we also perform a reduced complexity approach to the joint optimization, where the optimization is restricted to constellations with at most $N = 6$ rings. This makes the results meaningful only for low and moderate input power ($P \leq 0$ dBm), because, as we have seen previously, for higher input power, the dominance of the NLPN dictates the use of more amplitude levels to achieve good performance. The results are also shown in Fig. 10 for $P \leq 0$ dBm (diamond markers). It can be seen that the jointly optimized constellations with the additional constraint $N \leq 6$ follow closely the performance of the constellations obtained for a uniform radius vector, confirming

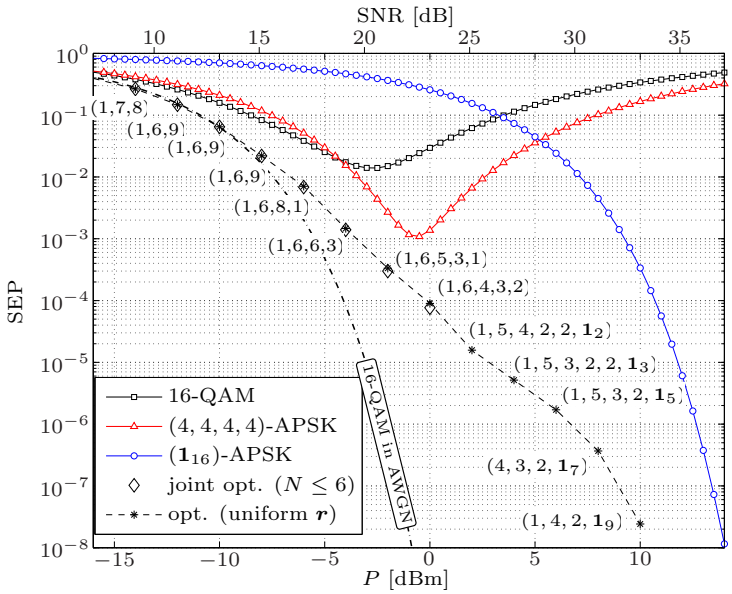



**Figure 10:** Results for the constellation optimization with $M = 16$. The SEP under TS detection for 16-QAM, uniform $(4, 4, 4, 4)$-APSK, and $(\mathbf{1}_{16})$-APSK are shown for comparison. The SEP under ML detection for 16-QAM in an AWGN channel (dash-dotted line) is shown as a reference.

that the joint optimization only yields negligible performance improvements for this power regime. For higher input power, the obtained jointly optimized constellations perform worse than the optimal constellations obtained with a uniform radius vector, which is simply due to the restriction to six rings.

Finally, the phenomenon of sacrificial points discussed previously also generalizes to entire rings, i.e., when optimizing the radius vector of constellations with more than one point in the outer ring. In this case, however, the SEP still exhibits a minimum. As an example, in Fig. 11 we show the result of the radius optimization for $(4, 4, 4, 4)$-APSK (top) together with the optimal radius vector (bottom). The radius vector $\boldsymbol{r}$ is plotted in a normalized fashion $\tilde{\boldsymbol{r}} = \boldsymbol{r}/\sqrt{P}$ (i.e., $\sum_{i=1}^{N} l_i \tilde{r}_i^2 = 1$), so that the effect is more clear. It can be observed that up to an input power of 4 dBm the distance between any two adjacent rings for the optimal radius vector is approximately the same. Moreover the distance decreases with higher input power (like "squeezing accordion pleats" [31]). However, for $P > 4$ dBm the optimal radius vector changes significantly. Somewhat counterintuitively, in this power regime, is is better to place the four points in the outer constellation ring far away from the other rings.

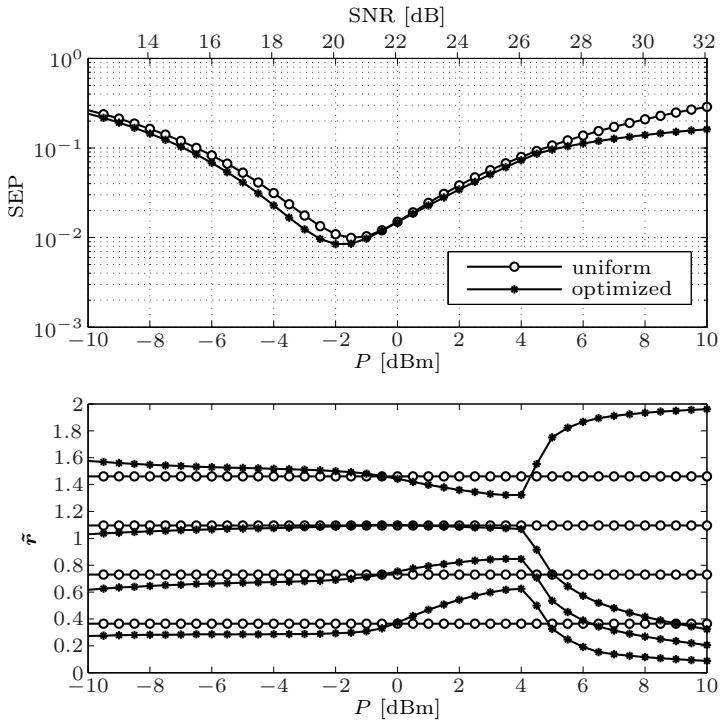



**Figure 11:** Illustration of a sacrificial ring which occurs for the radius optimization of $(4, 4, 4, 4)$-APSK for $P > 4$ dBm. The system length is $L = 7000$ km.

# 5 Binary Labelings

In order to allow for the transmission of binary data, we assume that each symbol $x_i \in \mathcal{X}$ is labeled with a binary vector $\boldsymbol{c}_i = (c_{i,1}, \ldots, c_{i,m}) \in \{0,1\}^m$, where $m = \log_2 M$. The different binary vectors are the binary representations of the integers $\{0, 1, \ldots, M-1\}$.[8] A specific mapping between vectors and constellation symbols is called a binary labeling, which will be denoted by an $M \times m$ matrix $\mathbb{L}_m = (\boldsymbol{c}_1^T, \ldots, \boldsymbol{c}_M^T)^T$.

A *Gray labeling* is obtained if the binary vectors of neighboring symbols, i.e., symbols that are closest in terms of Euclidean distance, differ by only one bit position. As an example, the Gray labeling $\mathbb{G}_m$ for $(M)$-APSK constellations[9] may be constructed by $m-1$ recursive reflections of the trivial labeling $\mathbb{G}_1 = (0,1)^T$. To obtain $\mathbb{G}_{m+1}$ from $\mathbb{G}_m$ by reflection, generate the matrix $(\boldsymbol{c}_1^T, \ldots, \boldsymbol{c}_M^T, \boldsymbol{c}_M^T, \ldots, \boldsymbol{c}_1^T)^T$ and add an extra column

[8] One may arbitrarily choose $c_{i,1}$ as the most significant bit.

[9] Different Gray labelings exist for a given constellation and for simplicity we restrict ourselves to $\mathbb{G}_m$, which is referred to as the binary reflected Gray code (BRGC) in the literature: It is the provably optimal Gray labeling for PSK constellations in an AWGN channel at high SNR [40].

from the left, consisting of $M$ zeros followed by $M$ ones [40].

## 5.1 Bit Error Probability

The average BEP of the signal constellation is given by

$$\mathrm{BEP} = \frac{1}{mM} \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j \neq i}}^{M} d_{\mathrm{H}}(\boldsymbol{c}_i, \boldsymbol{c}_j) \cdot P_{i \to j}, \tag{A.19}$$

where $d_{\mathrm{H}}(\cdot\,,\cdot)$ denotes the Hamming distance between two binary vectors. A lower bound for the BEP is $\mathrm{SEP}/m \leq \mathrm{BEP}$, which directly follows from $1 \leq d_{\mathrm{H}}(\boldsymbol{c}_i, \boldsymbol{c}_j)$, $i \neq j$.

The probabilities $P_{i \to j}$ are fixed for a given constellation and $P$ and $L$ (cf. (A.5)), hence (A.19) depends only on the labeling. However, it is important to realize that the phase offset vector $\boldsymbol{\varphi}$ has an effect on these probabilities for $j \neq i$. Therefore, even though two APSK constellations with the same $\boldsymbol{l}$ and $\boldsymbol{r}$ but different $\boldsymbol{\varphi}$ have the same SEP, they may have a different BEP. In the following, we show how to exploit this new degree of freedom for a class of APSK constellations that permit the use of a Gray labeling.

## 5.2 Rectangular APSK

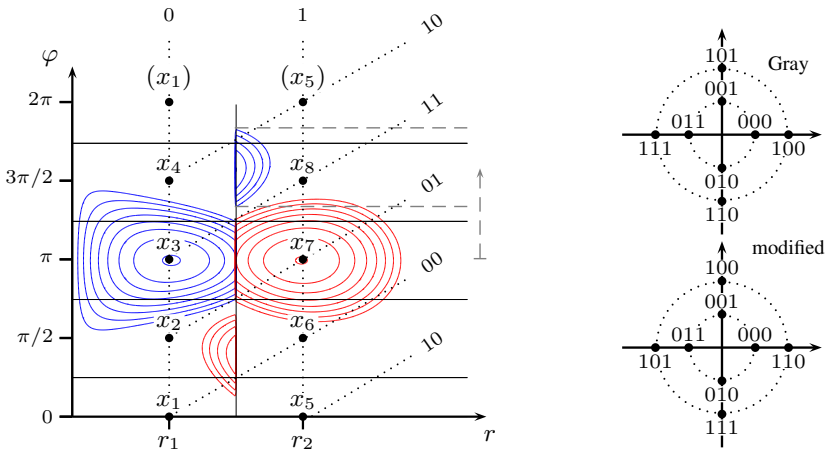
APSK constellations with $1 < N < M$, $\boldsymbol{\varphi} = \boldsymbol{0}$ and $l \triangleq l_k = M/N$, $1 \leq k \leq N$, have a "rectangular" structure when plotted in polar coordinates. For these constellations, a Gray labeling is given by $\mathbb{L}_m = \mathbb{G}_{\log_2 N} \otimes \mathbb{G}_{\log_2 l}$, where $\otimes$ is the ordered direct product, defined as

$$(\boldsymbol{a}_1^T, \ldots, \boldsymbol{a}_p^T)^T \otimes (\boldsymbol{b}_1^T, \ldots, \boldsymbol{b}_q^T)^T = (\boldsymbol{c}_1^T, \ldots, \boldsymbol{c}_{pq}^T)^T, \tag{A.20}$$

where $\boldsymbol{c}_{qi+j} = (\boldsymbol{a}_i, \boldsymbol{b}_j)$, for $1 \leq i \leq p$ and $1 \leq j \leq q$. This amounts to independently choosing a Gray labeling for the radius and phase coordinates of the constellation and then concatenating the binary vectors. In Fig. 12(b) (top) an example for such a construction is shown.

Gray labelings ensure that for a standard AWGN channel the BEP closely approaches the lower BEP bound for high SNR in Sec. 5.1. Using the same labeling in a nonlinear channel, however, does not necessarily ensure good performance since it is constructed without considering the underlying PDF of the observation. From Fig. 3 it is evident that this PDF has a rather unusual shape, due to the slicing effect caused by the post-compensation. To further illustrate this point, consider $(4,4)$-APSK with $r_1/r_2 = 0.424$ and $\boldsymbol{\varphi} = (0,0)$, operating at $L = 7000$ km and $P = -5$ dBm, which is the optimal APSK constellation for these parameters (cf. Fig. 8). The PDF of $\tilde{Y}$ conditioned on one point in each ring[10] is evaluated and plotted *in polar coordinates*, as shown in Fig. 12(a) for $x_3$ and $x_7$. The solid lines correspond to the decision boundaries of the TS

[10] The PDFs conditioned on all other points can be obtained by a phase translation. Note that the PDF is periodic in phase.




**(a)** Illustration in polar coordinates

**(b)** Labeled Const.

**Figure 12:** In (a), the PDF of $\tilde{Y}$ is shown for $(4,4)$-APSK conditioned on $X = x_3$ (blue) and $X = x_7$ (red) for $L = 7000$ km and $P = -5$ dBm. Solid lines correspond to the decision boundaries of the TS detector and dotted lines show connections between neighboring symbols according to the PDF.

detector. Recalling that the symbol transition probabilities are obtained through integration of the PDF over the detector regions (cf. (A.5)), Fig. 12(a) can then be used to identify "neighboring symbols" of $x_3$ and $x_7$ (and consequently of all points) in the sense that the corresponding transition probabilities will dominate in (A.19). The main observation here is that, even though $x_3$ and $x_7$ are adjacent symbols in radial direction, the corresponding transition probabilities, i.e., $P_{3\to 7}$ and $P_{7\to 3}$, are negligible compared to $P_{3\to 8}$ and $P_{7\to 2}$, respectively. The dotted lines in Fig. 12(a) show appropriate connections between neighboring symbols taking into account the nonlinear PDF. In Fig. 12(b) the $(4,4)$-constellation is shown with the Gray labeling $\mathbb{G}_1 \otimes \mathbb{G}_2$ (top) and the modified labeling (bottom) that results from "following" the dotted lines in Fig. 12(a) and concatenating the binary vectors.

Observe that the labeled constellation in the bottom of Fig. 12(b) may be obtained from the Gray labeled constellation by using a phase offset vector of $\boldsymbol{\varphi} = (0, \pi/2)$: In this case, the non-zero phase offset in the second ring does not change the constellation (i.e., the set of symbols), but rather leads to a different indexing of symbols (cf. Sec. 2.2), and consequently to a different mapping between symbols and binary vectors. Going one step further, we now allow for arbitrary phase offsets in all constellation rings[11], with the intention to "steer" the phase decision boundaries such that they are roughly symmetric around the PDF. Starting from a Gray labeled rectangular APSK constellation, a simple

[11] Note that an APSK constellation with $\boldsymbol{\varphi} \neq \mathbf{0}$ is not necessarily rectangular anymore.

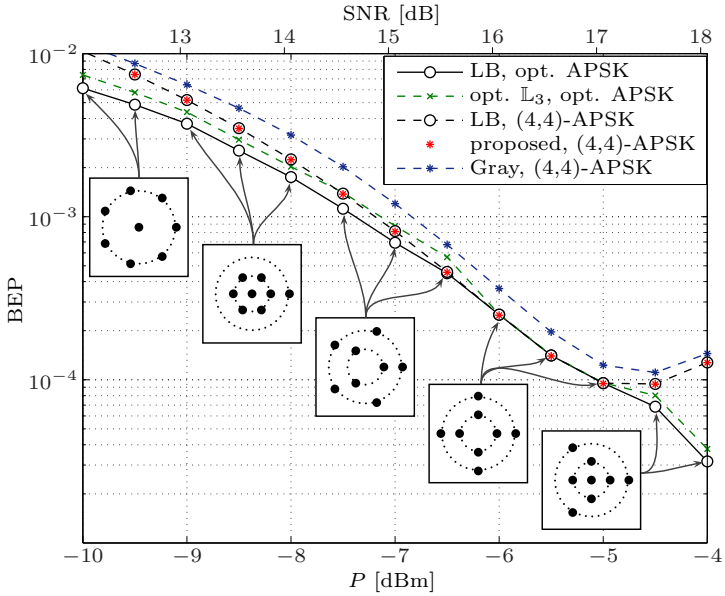



**Figure 13:** Average BEP and lower bounds (LB) for $M = 8$ and $L = 7000$ km for different APSK constellations. The subfigures show the optimized APSK constellations for the corresponding input power.

way to achieve this is by initializing $\varphi_1 = 0$ and then calculating

$$\varphi_i = \theta_c(\mu_{i-1}, r_{i-1}) - \theta_c(\mu_{i-1}, r_i) + \varphi_{i-1}, \tag{A.21}$$

for $i = 2, \ldots, N$. For the previous example, evaluating (A.21) for $i = 2$ results in $\varphi_2 \approx 1.878$, corresponding to the length of the dashed, grey arrow in 12(a). The two dashed, grey lines are the new phase decision boundaries for $x_7$ and it can be seen that they appear roughly symmetric around the blue PDF. We highlight that this proposed method to determine the phase offset vector may be applied to any rectangular APSK constellation of arbitrary constellation size, provided that $M$ is a power of 2.

## 5.3 Results and Discussion

In Fig. 13, the lower bound (LB) for the BEP is plotted for (4,4)-APSK with optimized $\boldsymbol{r}$ and $L = 7000$ km (black, dashed line). The average BEP of the labeled constellation is shown with the proposed phase offset vector (red markers) and $\boldsymbol{\varphi} = \mathbf{0}$, resulting in the conventional Gray labeling (blue, dashed line). The performance with the proposed method is almost indistinguishable from the lower bound and a gain of approximately 0.4 dB is visible at SEP $= 10^{-3}$ over the Gray labeling approach.

Moreover, in Fig. 13 we plot the LB based on the jointly optimized APSK constellations with $M = 8$ (cf. Fig. 8), where the subfigures are provided to show the optimal parameter $\boldsymbol{l}$ for the different input powers. For each input power, the optimal labeling is determined exhaustively[12] and the corresponding BEP is shown by the green line. Note that the LB is tight only for the rectangular $(4, 4)$-APSK. The results demonstrate that first optimizing the constellation with respect to SEP and then choosing an optimal labeling does not guarantee to give the best BEP performance. In particular, for $-7$ dBm and $-6.5$ dBm, (4,4)-APSK with the proposed phase offset vector achieves a lower BEP than the jointly optimized constellations (with respect to SEP) with an optimal labeling.

As a last point, one might argue that the class of rectangular APSK constellations is not particularly interesting, because they appear rarely as optimal APSK constellations with respect to SEP (e.g., for $M = 8$ they only appear in a small power range and for $M = 16$ they do not appear at all). However, the previous results show that if we take the BEP as the main figure of merit, rectangular APSK constellations may be advantageous in certain power regimes, because they can closely approach the lower BEP bound. Moreover, as we described earlier, the proposed labeling method is easily applicable to any constellation size. It would therefore be relatively simple to find optimal *rectangular* APSK constellations for $M > 16$ because in this case not many choices exist. Obviously, these constellations might then be far away from the performance of the true optimal constellation, but they might still offer a significant performance gain over rectangular QAM constellations in the presence of NLPN, with the advantage that a constructive labeling method is readily available.

# 6 Conclusion

In this paper, we optimized APSK constellations for a fiber-optical channel model without dispersion. It was shown how to derive the PDF of the postcompensated observation assuming a TS detection scheme. The PDF was used to gain insight into the performance behavior with respect to optimal detection and to calculate the BEP. Optimal APSK constellations under TS detection have been presented. For $M = 16$ constellation points, significant performance improvements in terms of SEP can be achieved by choosing an optimized APSK constellation compared to a baseline 16-QAM constellation. For very high input power, we showed that sacrificing points or constellation rings may become beneficial. Finally, the binary labeling problem was discussed and a constructive labeling method was presented, which is applicable to rectangular APSK constellations. An important topic for future work would be the investigation of the influence of fiber chromatic dispersion and nonlinearities on the optimal signal set.

[12] It was pointed out in [41] that, in general, the labeling problem falls under the category of quadratic assignment problems, and as such, is NP-hard.

## Acknowledgment

The authors would like to thank L. Beygi and F. Brännström for many helpful discussions.

## References

[1] I. Djordjevic, W. Ryan, and B. Vasic, *Coding for Optical Channels*. Springer, 2010.

[2] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. Academic Press, 2006.

[3] K.-P. Ho, *Phase-modulated Optical Communication Systems*. Springer, 2005.

[4] L. Hanzo, W. Webb, and T. Keller, *Single- and Multi-carrier Quadrature Amplitude Modulation: Principles and Applications for Personal Communications, WLANs and Broadcasting*. Wiley, 2000.

[5] G. J. Foschini, R. D. Gitlin, and S. B. Weinstein, "On the selection of a two-dimensional signal constellation in the presence of phase jitter and gaussian noise," *Bell Syst. Tech. J*, vol. 52, no. 6, pp. 927–965, Jul. 1973.

[6] ——, "Optimization of two-dimensional signal constellations in the presence of gaussian noise," *IEEE Trans. Commun.*, vol. 22, no. 1, pp. 28–38, Jan. 1974.

[7] C. M. Thomas, M. Y. Weidner, and S. H. Durrani, "Digital amplitude-phase keying with M-ary alphabets," *IEEE Trans. Commun.*, vol. 22, no. 2, pp. 168–180, Feb. 1974.

[8] E. Biglieri, "High-level modulation and coding for nonlinear satellite channels," *IEEE Trans. Commun.*, vol. 32, no. 5, pp. 616–626, May 1984.

[9] R. De Gaudenzi, A. Guillén i Fàbregas, and A. Martinez, "Performance analysis of turbo-coded apsk modulations over nonlinear satellite channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2396–2407, Sep. 2006.

[10] W. Sung, S. Kang, P. Kim, D. Chang, and D. Shin, "Performance analysis of APSK modulation for DVB-S2 transmission over nonlinear channels," *Int. J. Commun. Syst. Network*, vol. 27, no. 6, pp. 295–311, Nov. 2009.

[11] European Telecommunications Standards Institute, "Digital video broadcasting (DVB) - second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications (DVB-S2)," Sep. 2009. [Online]. Available: http://www.dvb.org/technology/standards/index.xml

[12] G. P. Agrawal, *Fiber-optic Communication Systems*, 4th ed. Wiley-Interscience, 2010.

[13] B. W. Göbel, "Information-theoretic aspects of fiber-optic communication channels," Ph.D. dissertation, TU Munich, Munich, 2010.

[14] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Göbel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[15] J. P. Gordon and L. F. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Opt. Lett.*, vol. 15, no. 23, pp. 1351–1353, Dec. 1990.

[16] A. Mecozzi, "Limits to long-haul coherent transmission set by the kerr nonlinearity and noise of the in-line amplifiers," *J. Lightw. Technol.*, vol. 12, no. 11, pp. 1993–2000, Nov. 1994.

[17] M. I. Yousefi and F. R. Kschischang, "On the per-sample capacity of nondispersive optical fibers," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7522–7541, Nov. 2011.

[18] K. S. Turitsyn, S. A. Derevyanko, I. V. Yurkevich, and S. K. Turitsyn, "Information capacity of optical fiber channels with zero average dispersion," *Phys. Rev. Lett.*, vol. 91, no. 20, p. 203901, Nov. 2003.

[19] E. Ip, "Nonlinear compensation using backpropagation for polarization-multiplexed transmission," *J. Lightw. Technol.*, vol. 28, no. 6, pp. 939–951, Mar. 2010.

[20] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links," *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 2012.

[21] L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, and H. Wymeersch, "A discrete-time model for uncompensated single-channel fiber-optical links," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3440–3450, Nov. 2012.

[22] A. G. Green, P. P. Mitra, and L. G. L. Wegener, "Effect of chromatic dispersion on nonlinear phase noise," *Opt. Lett.*, vol. 28, no. 24, pp. 2455–2457, Dec. 2003.

[23] S. Kumar, "Effect of dispersion on nonlinear phase noise in optical transmission systems," *Opt. Lett.*, vol. 30, no. 24, pp. 3278–3280, Dec. 2005.

[24] K.-P. Ho and H.-C. Wang, "Effect of dispersion on nonlinear phase noise." *Opt. Lett.*, vol. 31, no. 14, pp. 2109–11, Jul. 2006.

[25] A. Demir, "Nonlinear phase noise in optical-fiber-communication systems," *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2002–2032, Aug. 2007.

[26] S. Kumar, "Analysis of nonlinear phase noise in coherent fiber-optic systems based on phase shift keying," *J. Lightw. Technol.*, vol. 27, no. 21, pp. 4722–4733, Nov. 2009.

[27] A. Bononi, P. Serena, and N. Rossi, "Nonlinear signal-noise interactions in dispersion-managed links with various modulation formats," *Opt. Fiber. Techn.*, vol. 16, no. 2, pp. 73–85, Mar. 2010.

[28] A. P. T. Lau and J. M. Kahn, "Signal design and detection in presence of nonlinear phase noise," *J. Lightw. Technol.*, vol. 25, no. 10, pp. 3008–3016, Oct. 2007.

[29] L. Beygi, E. Agrell, and M. Karlsson, "Optimization of 16-point ring constellations in the presence of nonlinear phase noise," in *Proc. Optical Fiber Communication Conf. (OFC)*, Los Angeles, CA, Mar. 2011.

[30] N. Ekanayake and H. M. V. R. Herath, "Effect of nonlinear phase noise on the performance of m-ary psk signals in optical fiber links," *J. Lightw. Technol.*, vol. 31, no. 3, pp. 447–454, Feb. 2013.

[31] T. Freckmann, R.-J. 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. 2009.

[32] J. Zhang and I. B. Djordjevic, "Optimum signal constellation design for rotationally symmetric optical channel with coherent detection," in *Proc. Optical Fiber Communication Conf. (OFC)*, Beijing, China, Mar. 2011.

[33] I. B. Djordjevic, T. Liu, L. Xu, and T. Wang, "Optimum signal constellation design for high-speed optical transmission," in *Proc. Optical Fiber Communication Conf. (OFC)*, Los Angeles, CA, Mar. 2012.

[34] L. Beygi, E. Agrell, P. Johannisson, and M. Karlsson, "A novel multilevel coded modulation scheme for fiber optical channel with nonlinear phase noise," in *Proc. IEEE Glob. Communication Conf. (GLOBECOM)*, Miami, FL, Dec. 2010.

[35] E. Agrell and M. Karlsson, "Satellite constellations: Towards the nonlinear channel capacity," in *Proc. IEEE Photon. Conf.*, Burlingame, CA, Sep. 2012.

[36] E. Agrell, "On monotonic capacity–cost functions," Sep. 2012. [Online]. Available: http://arxiv.org/abs/1209.2820

[37] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[38] A. Lapidoth, *A Foundation in Digital Communication.* Cambridge University Press, 2009.

[39] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Computer J.*, vol. 7, no. 4, pp. 308–313, Jul. 1965.

[40] E. Agrell, J. Lassing, E. Ström, and T. Ottosson, "On the optimality of the binary reflected gray code," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3170–3182, Dec. 2010.

[41] Y. Huang and J. Ritcey, "Optimal constellation labeling for iteratively decoded bit-interleaved space-time coded modulation," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1865–1871, May 2005.

# PAPER B

**A Low-Complexity Detector for Memoryless Polarization-Multiplexed Fiber-Optical Channels**

Christian Häger, Lotfollah Beygi, Erik Agrell, Pontus Johannisson,
Magnus Karlsson, and Alexandre Graell i Amat

*IEEE Communications Letters*,
vol. 60, no. 11, pp. 3440–3450, Nov. 2012

*The layout has been revised.*

**Abstract**

A low-complexity detector is introduced for polarization-multiplexed $M$-ary phase shift keying modulation in a fiber-optical channel impaired by nonlinear phase noise, generalizing a previous result by Lau and Kahn for single-polarization signals. The proposed detector uses phase compensation based on both received signal amplitudes in conjunction with simple straight-line rather than four-dimensional maximum-likelihood decision boundaries.

# 1 Introduction

The Manakov equation describes the propagation of a polarization-multiplexed signal in a fiber-optical channel. Two major impairments, linear chromatic dispersion and the Kerr nonlinear effect, are modeled by this equation. The nonlinear effect causes a phase rotation proportional to the field instantaneous power. The interaction of the signal and the amplified spontaneous emission (ASE) noise generated by optical amplifiers due to the nonlinear Kerr effect gives rise to nonlinear phase noise (NLPN). NLPN imposes a major degradation in the performance of coherent optical data transmission systems.

Bononi *et al.* [1] investigated the effect of NLPN on popular modulation formats for single-channel and wavelength-division multiplexing systems in a dispersion-managed fiber link. The performance of orthogonal frequency-division multiplexing systems in the presence of NLPN has been evaluated in [2] by theoretical, numerical, and experimental approaches. In [3], [4, ch. 4], comprehensive surveys of known techniques for the analysis and characterization of NLPN and its impact on the system performance are provided.

The statistics of NLPN and the detector design for a channel with NLPN have been studied in [5–7] by analytical approaches and in [8] by numerical methods. The joint probability density function (pdf) of the received amplitude and phase given the initial amplitude and phase of the transmitted signal and the optical signal-to-noise ratio (OSNR) is derived in [5,9,10] [4, pp. 157, 224–225] for a fiber-optical channel with NLPN caused by distributed or lumped amplification. Moreover, compensation of NLPN has been studied in [11] based on the aforementioned pdf.

In this paper, we extend the detector structure introduced for a single polarization (SP) $M$-PSK system in [11] to polarization-multiplexed (PM) $M$-PSK, using the signal statistics derived in [12]. To this end, we first introduce a simplified approach to reproduce the result in [11] for the SP case. This method can be easily used to extend the result to the PM case and can also be applied to both lumped and distributed amplification. For simplicity, we assume single-channel transmission and inter-channel effects are not taken into consideration. The symbol error rate (SER) of the proposed detector is compared to the performance of the maximum-likelihood (ML) detector for PM-4-PSK and to the performance of the ML detector for SP-4-PSK for the same bandwidth as well as for the same data rate.

## 2 System Model and Preliminaries

We assume zero dispersion to make the analysis applicable to memoryless (nondispersive) fiber-optical channels, similarly as in, e.g., [5, 8, 10, 11, 13]. Due to this assumption, the subsequent analysis ignores the interaction of chromatic dispersion and nonlinearity. The resulting model can serve as an approximation for dispersion-managed transmission links provided that the local accumulated dispersion is sufficiently low [1, 12]. For a zero polarization-mode and chromatic dispersion fiber-optical channel, the Manakov equation with loss included reduces to [14, ch. 6]

$$j\frac{\partial \mathbf{E}}{\partial z} + \gamma(\mathbf{E}\mathbf{E}^{\dagger})\mathbf{E} + j\frac{\alpha}{2}\mathbf{E} = \mathbf{0}, \tag{B.1}$$

where $\mathbf{E} = (E_{\mathrm{x}}, E_{\mathrm{y}})$ is the polarization-multiplexed launched envelope signal into the fiber, $\gamma$ is the nonlinear coefficient, $\alpha$ is the attenuation coefficient, $\dagger$ denotes Hermitian conjugation, and $z$ is the distance from the beginning of the fiber. The solution to (B.1) at time $t$ can be written as [14, ch. 4]

$$\mathbf{E}(z,t) = \mathbf{E}(0,t)q(z)\exp\left(j\gamma P_0(t)\int_0^z q^2(\tau)\,\mathrm{d}\tau\right), \tag{B.2}$$

where $P_0(t) = |E_{\mathrm{x}}(0,t)|^2 + |E_{\mathrm{y}}(0,t)|^2$ is the instantaneous launched power into the fiber and $q(z) = \exp(-\alpha z/2)$ is a function that describes the power evolution.

Here, we assume a fiber link with total length $L$ and either distributed or lumped amplification to compensate for the fiber loss perfectly. We consider ASE noise within the optical signal bandwidth, i.e., ignoring the Kerr effect induced from out-of-band signal and noise in the same way as in [13]. If a four-dimensional (4D) signal $\mathbf{S} = (S_{\mathrm{x}}, S_{\mathrm{y}})$, consisting of two two-dimensional (2D) complex signals, is transmitted, it experiences an overall NLPN $\varphi_{\mathrm{NL}} = \varphi_{\mathrm{x}} + \varphi_{\mathrm{y}}$. The terms $\varphi_{\mathrm{x}}$ and $\varphi_{\mathrm{y}}$ are generated by the interaction of the signal and noise due to the Kerr effect in polarizations x and y, respectively. For lumped amplification and a link consisting of $N$ spans, ASE noise $n_{\mathrm{x}}^i$, $i = 1, \ldots, N$, with variance $\sigma_0^2$ is added after each span.[1] One may use (B.2) to obtain $\varphi_{\mathrm{x}} = \gamma L_{\mathrm{eff}} \sum_{i=1}^{N} |S_{\mathrm{x}} + \sum_{l=1}^{i} n_{\mathrm{x}}^l|^2$, where $L_{\mathrm{eff}} = (1 - \exp(-\alpha L/N))/\alpha$ is the effective nonlinear length. It is clearly seen that signals in both polarizations contribute to the generated NLPN $\varphi_{\mathrm{NL}}$. The received electric field $\mathbf{E}$ can be written as $\mathbf{E} = \hat{\mathbf{E}}e^{-j\varphi_{\mathrm{NL}}}$, where $\hat{\mathbf{E}} = \mathbf{S} + \sum_{i=1}^{N} \mathbf{n}^i$ is the linear part of the electric field and $\mathbf{n}^i = (n_{\mathrm{x}}^i, n_{\mathrm{y}}^i)$. One may regard distributed amplification as lumped amplification with an infinite number of spans. This gives $\lim_{N\to\infty} NL_{\mathrm{eff}} = L$. In this case, a continuous amplifier noise vector $\mathbf{n}(z) = (n_{\mathrm{x}}(z), n_{\mathrm{y}}(z))$ is considered. The elements of this vector are zero-mean complex-valued Wiener processes [4, p. 154] with autocorrelation function $\mathrm{E}[n_{\mathrm{x}}(z_1)n_{\mathrm{x}}^*(z_2)] = \sigma_{\mathrm{d}}^2 \min(z_1, z_2)$, where $\sigma_{\mathrm{d}}^2 = 2h\nu_{\mathrm{opt}} W \alpha n_{\mathrm{sp}}$ [11], $h\nu_{\mathrm{opt}}$ is the energy of a photon, $n_{\mathrm{sp}}$ is the spontaneous emission factor, and $W$ is the

[1] Throughout the paper, we give expressions for polarization x only, if polarization y has an equivalent expression.

bandwidth of the optical signal. The SNR vector is defined as $\boldsymbol{\rho} = (\rho_\mathrm{x}, \rho_\mathrm{y})$ where $\rho_\mathrm{x}$ is $|S_\mathrm{x}|^2/(L\sigma_\mathrm{d}^2)$ or $|S_\mathrm{x}|^2/(N\sigma_0^2)$ for distributed or lumped amplification, respectively. The normalized received amplitude $r_\mathrm{x}$ is denoted by $|E_\mathrm{x}|/(\sigma_\mathrm{d}\sqrt{L})$ or by $|E_\mathrm{x}|/(\sigma_0\sqrt{N})$ for distributed and lumped amplifications.

The joint pdf of the received phase vector $\boldsymbol{\theta} = (\theta_\mathrm{x}, \theta_\mathrm{y})$ and the normalized amplitudes $\mathbf{r} = (r_\mathrm{x}, r_\mathrm{y})$ of a zero-dispersion fiber-optical channel is [12]

$$\begin{aligned} f_{\boldsymbol{\Theta},\mathbf{R}}(\boldsymbol{\theta},\mathbf{r}) = \frac{f_\mathbf{R}(\mathbf{r})}{4\pi^2} &+ \frac{1}{2\pi^2}\sum_{k_\mathrm{x}=1}^{\infty} \mathrm{Re}\left\{C_{\mathbf{k}_\mathrm{x}}(\mathbf{r})e^{jk_\mathrm{x}\theta_\mathrm{x}}\right\} \\ &+ \frac{1}{2\pi^2}\sum_{k_\mathrm{x}=1}^{\infty}\sum_{k_\mathrm{y}=1}^{\infty} \mathrm{Re}\left\{C_\mathbf{k}(\mathbf{r})e^{j\mathbf{k}\cdot\boldsymbol{\theta}} + C_{\mathbf{k}^*}(\mathbf{r})e^{j\mathbf{k}^*\cdot\boldsymbol{\theta}}\right\} \\ &+ \frac{1}{2\pi^2}\sum_{k_\mathrm{y}=1}^{\infty} \mathrm{Re}\left\{C_{\mathbf{k}_\mathrm{y}}(\mathbf{r})e^{jk_\mathrm{y}\theta_\mathrm{y}}\right\}, \end{aligned} \tag{B.3}$$

where $f_\mathbf{R}(\mathbf{r})$ is the joint pdf of the two normalized independent Ricean random variables $r_\mathrm{x}$ and $r_\mathrm{y}$, and the Fourier coefficients $C_\mathbf{k}(\mathbf{r})$ are given in [12]. In (B.3), we assume the transmitted phase vector to be $(0, 0)$. Due to the rotational invariance of the channel, the pdf for an arbitrary transmitted phase vector $(\theta_{0,\mathrm{x}}, \theta_{0,\mathrm{y}})$ is obtained by replacing $\theta_\mathrm{x}$ and $\theta_\mathrm{y}$ in (B.3) with $\theta_\mathrm{x} - \theta_{0,\mathrm{x}}$ and $\theta_\mathrm{y} - \theta_{0,\mathrm{y}}$, respectively.

For an SP scheme, the joint pdf of the phase and the normalized amplitude of the received signal in the corresponding polarization is simplified to [4, ch. 5]

$$f_{\Theta,R}(\theta,r) = \frac{f_R(r)}{2\pi} + \frac{1}{\pi}\sum_{k=1}^{\infty} \mathrm{Re}\left\{C_k(r)e^{jk\theta}\right\}, \tag{B.4}$$

where $f_R(r)$ is the Ricean pdf of the amplitude $r$, and the Fourier coefficients $C_k(r)$ are given in [12] for both types of amplifications. Again, the transmitted phase in (B.4) is assumed to be 0, and the pdf for an arbitrary transmitted phase $\theta_0$ is obtained by replacing $\theta$ with $\theta - \theta_0$.

In the following, we consider $M$-PSK constellations with $s_k = \sqrt{E_\mathrm{s}}\exp\left(j\frac{\pi}{M}(2k+1)\right)$, $k = 0, \ldots, M-1$, where $E_\mathrm{s}$ is the average energy of the constellation.

# 3 The ML Receiver for SP-$M$-PSK

For SP-$M$-PSK, the optimal decision (Voronoi) regions for the received constellation have spiral shape (cf. [11, Fig. 1]), and hence ML detection is computationally complex. To decrease the complexity of the detector, Lau and Kahn showed that straight-line decision boundaries can be used, provided that an amplitude-dependent phase rotation $\theta^c$ is applied before detection [11]. The corresponding receiver structure is illustrated in the top half of Fig. 1(a). It can be seen that the phase rotation is solely a function of the

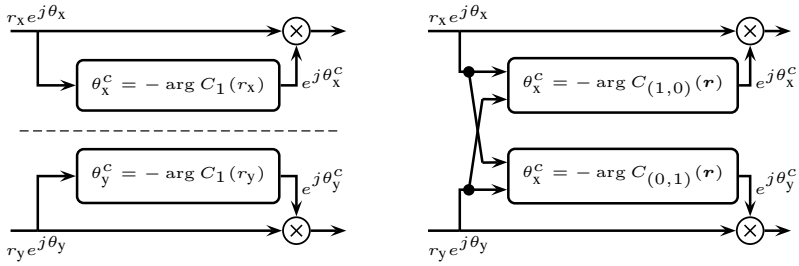



**Figure 1:** Receiver for PM-$M$-PSK with (a) separate detection (PM-Det1) and (b) joint calculation of the amplitude-dependent phase rotations (PM-Det2).

received amplitude in one polarization and a simple ML detection of $M$-PSK for additive white Gaussian noise (AWGN) with straight-line decision boundaries is subsequently performed.

In this section, we introduce a new approach to derive the optimal phase rotation as a function of the received amplitude. In contrast to [11], this approach can be easily extended to PM-$M$-PSK. For a transmitted phase of $\theta_0 = 0$, we assume that the conditional pdf $f_{\Theta|R}(\theta|r)$ of the received phase $\theta$ given the received amplitude $r$ is approximately symmetric around $\theta_{\max}(r)$, where $\theta_{\max}(r)$ denotes the phase value where $f_{\Theta|R}(\theta|r)$ is maximum. This assumption is motivated by inspection of the pdf and its validity is justified later by the obtained results. In fact, an equivalent approximation was also done in [11, App. A]. This assumption is used for both distributed and lumped amplifications.

**Lemma 1:** *Let $f_X(x)$ be the (periodic) pdf of a random angle $X$. Furthermore, let the pdf be symmetric around $x_{max} \in (-\pi, \pi]$, the value where $f_X(x)$ has its maximum. If the pdf decreases monotonically from $x_{max}$ to $x_{max} \pm \pi$, then $x_{max} = -\arg \Psi_X(1)$, where $\Psi_X(\nu)$ is the (discrete) characteristic function (CF) of $X$.*

*Proof.* Define $\tilde{X} = X - x_{\max}$. Since the pdf of $\tilde{X}$ is an even function, its CF is real. Furthermore, the CFs of $X$ and $\tilde{X}$ are related via $\Psi_X(\nu) = \Psi_{\tilde{X}}(\nu)e^{-j\nu x_{\max}}$. Letting $\nu = 1$ and solving for $x_{\max}$ gives $x_{\max} = \arg \Psi_{\tilde{X}}(1) - \arg \Psi_X(1)$. Thus, it needs to be shown that $\arg \Psi_{\tilde{X}}(1) = 0$. Having already established that $\Psi_{\tilde{X}}(1)$ is real, we only need to show that it is also positive. This follows from the definition $\Psi_{\tilde{X}}(1) = \int_{-\pi}^{\pi} f_{\tilde{X}}(x) \cos(x)\, \mathrm{d}x$ and the fact that $f_{\tilde{X}}(x)$ is nonnegative and decreases monotonically from 0 to $\pm\pi$. □

Using Lemma 1, one can compute the rotation of the $M$-PSK ML decision boundaries due to NLPN as described in the following theorem.

**Theorem 1:** *Consider a memoryless fiber-optical channel with NLPN. The decision boundary of the ML detector for SP-$M$-PSK between symbols $s_k$ and $s_{k+1}$ has the polar coordinates $(r, \theta_b(k, r))$, where $\theta_b(k, r) = -C_1(r) + 2k/M$ for $r \geq 0$ and $C_1$ is the first Fourier coefficient in* (B.4).

*Proof.* The ML decision boundary between the two symbols of the $M$-PSK constellation

with $k = 0$ and $k = M - 1$ is determined in such a way as to satisfy

$$f_{\Theta|R,\Theta_0}(\theta^{\mathrm{c}}(r) \mid r, -\tfrac{\pi}{M}) = f_{\Theta|R,\Theta_0}(\theta^{\mathrm{c}}(r) \mid r, \tfrac{\pi}{M}).$$

Using the symmetry of $f_{\Theta|R,\Theta_0}(\theta|r,\theta_0)$ around $\theta_{\max}(r) + \theta_0$, we obtain $\theta^{\mathrm{c}}(r) = \theta_{\max}(r)$. Using Lemma 1, we get $\theta_{\max}(r) = -\arg\ \Psi_{\Theta|R,\Theta_0}(1 \mid r, 0) = -\arg C_1(r)$. □

# 4 Receivers for PM-$M$-PSK

For a fixed state of polarization, we receive two dependent 2D symbols, which have been rotated by the NLPN equally. Using (B.3), the ML detector in this case can be written as

$$\hat{\boldsymbol{\theta}}_0 = \arg\max_{\boldsymbol{\theta}_0} f_{\boldsymbol{\Theta},\boldsymbol{R}|\boldsymbol{\Theta}_0}(\boldsymbol{\theta}, \boldsymbol{r}|\boldsymbol{\theta}_0). \tag{B.5}$$

The optimization is performed over all possible $M^2$ transmitted phase combinations for a PM-$M$-PSK signal. We refer to this detector as "PM-ML".

A simple, but clearly suboptimal, way to reduce the complexity of solving (B.5) is to treat the received signals in both polarizations independently. In other words, the marginal pdfs $f_{\Theta_{\mathrm{x}},R_{\mathrm{x}}}(\theta_{\mathrm{x}}, r_{\mathrm{x}})$ and $f_{\Theta_{\mathrm{y}},R_{\mathrm{y}}}(\theta_{\mathrm{y}}, r_{\mathrm{y}})$ are used to perform detection separately in each polarization, which leads again to spiral-shaped decision boundaries as in Sec. 3. Equivalently, one may extend the receiver structure for SP in a straightforward manner as shown in Fig. 1(a), where a different rotation angle is applied to each received symbol, based on the received amplitude in the corresponding polarization. Using Theorem 1, the computation of the rotation angles is then based on the first Fourier coefficient of the two marginal pdfs. We refer to this detector as "PM-Det1".

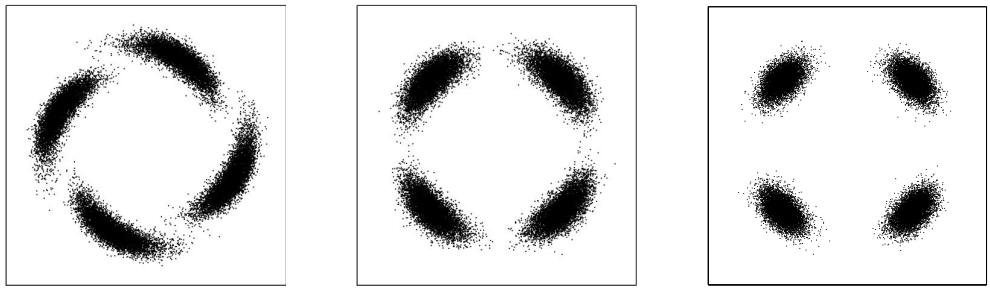
As seen in (B.2), the phase rotation due to the nonlinear Kerr effect is a function of the signal amplitudes in both polarizations. Hence, one may improve the performance of PM-Det1 by taking into account the amplitudes of both polarizations in computing the phase rotation. To this end, we use the same symmetry assumption as in the previous section for $f_{\Theta_{\mathrm{x}}|\mathbf{R}}(\theta|\mathbf{r}) = f_{\Theta_{\mathrm{x}},\mathbf{R}}(\theta, \mathbf{r})/f_{\mathbf{R}}(\mathbf{r})$, where $f_{\Theta_{\mathrm{x}},\mathbf{R}}(\theta, \mathbf{r})$ is the marginal of (B.3) with respect to $\Theta_{\mathrm{y}}$, i.e., we assume that $f_{\Theta_{\mathrm{x}}|\mathbf{R}}(\theta|\mathbf{r})$ is symmetric around the phase for which this pdf is maximum. This assumption allows us to describe the decision boundaries of the PM-$M$-PSK signal distorted by NLPN in each polarization as the rotated version of the straight-line decision boundaries for an AWGN channel.

**Theorem 2:** *The decision boundaries of the detector given by*

$$\hat{\theta}_{0,x} = \arg\max_{\theta_{0,x}} f_{\Theta_x,\boldsymbol{R}|\theta_{0,x}}(\theta_x, \boldsymbol{r}|\theta_{0,x}) \tag{B.6}$$

*for polarization $x$ can be transformed to straight lines using the phase rotation given by*

$$\theta_x^c(\mathbf{r}) = -\arg C_{(1,0)}(\mathbf{r}), \tag{B.7}$$



**(a)** no comp. **(b)** $\theta_{\mathrm{x}}^c = f(r_{\mathrm{x}})$ **(c)** $\theta_{\mathrm{x}}^c = f(\boldsymbol{r})$

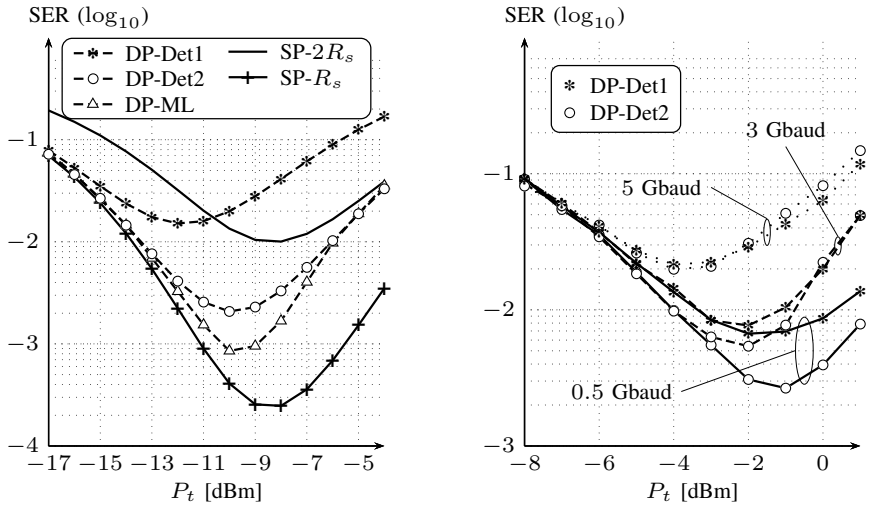
**Figure 2:** Scatter plots in the x polarization for (a) no compensation, (b) compensation according to Fig. 1(a), and (c) according to Fig. 1(b). Decision boundaries in (b) and (c) are straight lines. In (a), the boundaries are spiral shaped and depend on the received amplitude in the y polarization.

*where* $C_{(1,0)}(\mathbf{r})$ *is the Fourier coefficient appearing in* (B.3) *with* $\boldsymbol{k}_x = (1,0)$*. Similarly, the rotation for polarization y is obtained as* $\theta_y^c(\mathbf{r}) = -\arg C_{(0,1)}(\mathbf{r})$*.*

*Proof.* One may follow an analogous approach as in the proof of Theorem 1, by replacing $f_{\Theta_{\mathrm{x}}|\mathbf{R}}(\theta|\mathbf{r})$ with $f_{\Theta|R}(\theta|r)$ to show that the decision boundary between symbols $S_x = s_k$ and $S_x = s_{k+1}$ in polarization x has the parametric description $r_{\mathrm{x}} \exp(j\theta_{\mathrm{x}}^c(\mathbf{r}) + 2jk\pi/M)$. □

The proposed detector implementing (B.6) via this phase rotation method is referred to as "PM-Det2" and shown in Fig. 1(b). It can be seen that the rotation angle in each polarization is computed using the received amplitudes $\mathbf{r}$. It is worth mentioning that since the rotation is an invertible operation, joint 4D demodulation is still possible after the rotation. For complexity reasons, however, we perform hard decision on each 2D soft symbol using simple straight-line decision boundaries as shown in the figure.

In Fig. 2, a qualitative comparison of the two different rotation schemes corresponding to PM-Det1 and PM-Det2 is shown. Fig. 2(a) shows a scatter plot of the received symbols in polarization x directly after the channel, i.e., no phase compensation is assumed. In Fig. 2(b), a phase rotation of each received symbol is applied, which is solely based on the corresponding received amplitude of this symbol (PM-Det1). Lastly, Fig. 2(c) shows the result of applying a phase rotation that is based on the received amplitude of the received symbols in *both* polarizations (PM-Det2). Observe that the second rotation method leads to a notably smaller phase variance compared to the first method. However, it should be mentioned that the receiver structure shown in Fig. 1(b) does *not* correspond to the ML receiver for PM-$M$-PSK, since the residual phases in both polarizations after the rotation are not statistically independent. The performance loss compared to ML detection is quantified in the next section.




**(a)** analytical SER with zero dispersion

**(b)** simulated SER for a dispersion-managed link

**Figure 3:** The SER of SP and PM systems with 4-PSK versus the transmitted power per polarization $P_\mathrm{t}$.

# 5 Performance Analysis

The SER of PM-$M$-PSK for PM-Det1 and PM-Det2 can be computed analytically. After the introduced phase rotations, the marginal pdf of the phase in polarization x, given that the phase of the transmitted signal is zero, is obtained by replacing $\theta_\mathrm{x}$ with $\theta'_\mathrm{x} - \theta^\mathrm{c}_\mathrm{x}$ in $f_{\Theta_\mathrm{x}|\mathbf{R}}(\theta|\mathbf{r})$ and then integrating out the radii $r_\mathrm{x}$ and $r_\mathrm{y}$ over $[0, +\infty)$ to get

$$f_{\Theta'_\mathrm{x}}(\theta) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{k=1}^{\infty}\cos(k\theta)\int_0^\infty\int_0^\infty |C_{(k,0)}(\mathbf{r})|\mathrm{d}r_\mathrm{x}\mathrm{d}r_\mathrm{y}. \tag{B.8}$$

Here, we only show how to compute the SER of PM-Det2 for a PM-$M$-PSK system. An analogous derivation can be applied for PM-Det1. One can write

$$\mathrm{SER}_\mathrm{x} = 1 - \int_{-\frac{\pi}{M}}^{\frac{\pi}{M}} f_{\Theta'_\mathrm{x}}(\theta)\mathrm{d}\theta = \frac{M-1}{M} - \sum_{k=1}^{\infty}\frac{2\mathrm{sinc}\left(\frac{k}{M}\right)}{M}\int_0^\infty |C^\mathrm{x}_{(k,0)}(r_\mathrm{x})|\mathrm{d}r_\mathrm{x}\int_0^\infty |C^\mathrm{y}_{(k,0)}(r_\mathrm{y})|\mathrm{d}r_\mathrm{y},$$

where $C^\mathrm{x}_{(k,0)}$ and $C^\mathrm{y}_{(k,0)}$ are computed using [12, eq. (26)].

In Fig. 3(a), the performance of PM-4-PSK is evaluated using the above analytical approach for PM-Det1 and PM-Det2. The SER of the ML detector defined by (B.6) is given by a four-dimensional integral of the pdf over the ML decision regions. This

SER, estimated by Monte-Carlo integration, is also shown in Fig. 3(a). Moreover, we compute the SER of SP-4-PSK to compare with SP data transmission in two different scenarios: (i) For the same data rate per polarization (i.e., the same bandwidth) and (ii) for the same total data rate as the PM case. This evaluation is done for distributed amplification with channel parameters $L = 9000$ km, $\gamma = 1.4$ W$^{-1}$km$^{-1}$, $R_\mathrm{s} = 28$ Gbaud, $\nu_\mathrm{opt} = 193.55$ THz, $\alpha = 0.25$ dB/km, and $F_n = 6$ dB. As seen in Fig. 3(a), the PM schemes show negligible performance degradation in the linear regime for a fixed bandwidth (case (i)), i.e., for $P_\mathrm{t} < -15$ dBm, compared to the SP scheme. For a fixed data rate (case (ii)), one may observe a 2 dB performance improvement using PM-Det2, at a SER of $1.5 \times 10^{-2}$. In the strongly nonlinear regime, the SP scheme is superior to PM at the expense of losing half of the data rate. Furthermore, in the linear regime, the SP scheme in case (i) and the PM scheme have the same performance and their SER curves overlap, while in the strongly nonlinear regime, the SER of the PM scheme converges to the SER of the SP scheme in case (ii). This is because the system performance is intimately related to the product of the noise variance and the transmit power in the nonlinear regime, which is the same for these two scenarios. Fig. 3(a) also indicates that in the linear regime, the detectors PM-Det2 and PM-Det1 perform similarly. However, the reduction in circular variance observed in Fig. 2 translates into a noticeably better SER in the nonlinear regime for PM-Det2 when compared to PM-Det1. In the region of interest, i.e., SNRs around $-10$ dBm, the performance degradation of PM-Det2 compared to the ML detector is 0.7 dB. This is due to independent detection of phase information in the two polarizations. In Fig. 3(b), we also show the performance of DP-Det1 and DP-Det2 for a dispersion-managed link using the split step Fourier method. The system parameters are the same as before, but now we assume 45 fiber spans of length 90 km and a lumped amplification scheme. Dispersion is compensated after each span using an ideal dispersion-compensating fiber. The symbol rate is varied between 0.5 and 5 Gbaud to determine the robustness of the detector with respect to residual dispersion. The memoryless pdf loses its accuracy for high symbol rates due to the strong interaction between nonlinearities and dispersion and therefore the superiority of the proposed detector disappears for these parameters and symbol rates higher than 3 Gbaud. Similar observations regarding the accuracy of the memoryless model have been made in [12].

# 6 Conclusion

A low-complexity detector is proposed for memoryless polarization-multiplexed fiber-optical channels by compensating the amplitude-dependent NLPN. The compensation is performed by a phase rotation of the received symbols depending on the amplitudes in both polarizations. The performance results confirm the superiority of PM schemes to SP for the same data rate.

## References


[1] A. Bononi, P. Serena, and N. Rossi, "Modeling of signal-noise interactions in nonlinear fiber transmission with different modulation formats," *in Proc. European Conf. Optical Communication (ECOC), Vienna, Austria*, 2009.

[2] X. Yi, W. Shieh, and Y. Ma, "Phase noise effects on high spectral efficiency coherent optical OFDM transmission," *J. Lightw. Technol.*, vol. 26, no. 10, pp. 1309–1316, May 2008.

[3] A. Demir, "Nonlinear phase noise in optical-fiber-communication systems," *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2002–2032, Aug. 2007.

[4] K.-P. Ho, *Phase-Modulated Optical Communication Systems*. Springer, 2005.

[5] A. Mecozzi, "Probability density functions of the nonlinear phase noise," *Opt. Lett.*, vol. 29, no. 7, pp. 673–675, 2004.

[6] S. Kumar, "Analysis of nonlinear phase noise in coherent fiber-optic systems based on phase shift keying," *J. Lightw. Technol.*, vol. 27, no. 21, pp. 4722–4733, Nov. 2009.

[7] Y. Yadin, M. Shtaif, and M. Orenstein, "Nonlinear phase noise in phase-modulated WDM fiber-optic communications," *IEEE Photon. Technol. Lett.*, vol. 16, no. 5, pp. 1307–1309, May 2004.

[8] C. Häger, A. Graell i Amat, A. Alvarado, and E. Agrell, "Design of APSK constellations for coherent optical channels with nonlinear phase noise," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3362–3373, Aug. 2013.

[9] K. S. Turitsyn, S. A. Derevyanko, I. V. Yurkevich, and S. K. Turitsyn, "Information capacity of optical fiber channels with zero average dispersion," *Phys. Rev. Lett.*, vol. 91, no. 20, p. 203901, Nov. 2003.

[10] M. I. Yousefi and F. R. Kschischang, "On the per-sample capacity of nondispersive optical fibers," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7522–7541, Nov. 2011.

[11] A. P. T. Lau and J. M. Kahn, "Signal design and detection in presence of nonlinear phase noise," *J. Lightw. Technol.*, vol. 25, no. 10, pp. 3008–3016, Oct. 2007.

[12] L. Beygi, E. Agrell, M. Karlsson, and P. Johannisson, "Signal statistics in fiber-optical channels with polarization multiplexing and self-phase modulation," *J. Lightw. Technol.*, vol. 29, no. 16, pp. 2379–2386, Aug. 2011.

[13] J. P. Gordon and L. F. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Opt. Lett.*, vol. 15, no. 23, pp. 1351–1353, 1990.

[14] G. P. Agrawal, *Nonlinear fiber optics*, 4th ed. Academic Press, 2007.

# PAPER C

**Improving Soft FEC Performance for Higher-Order Modulations by Bit Mapper Optimization**

Christian Häger, Alexandre Graell i Amat, Fredrik Brännström, Alex Alvarado, and Erik Agrell

submitted to *Optics Express*,
Apr. 2014

*The layout has been revised.*

**Abstract**

Soft forward error correction with higher-order modulations is often implemented in practice via the pragmatic bit-interleaved coded modulation paradigm, where a single binary code is mapped to a nonbinary modulation. In this paper, we study the optimization of the mapping of the coded bits to the modulation bits for a polarization-multiplexed fiber-optical system without optical inline dispersion compensation. Our focus is on protograph-based low-density parity-check (LDPC) codes which allow for an efficient hardware implementation, suitable for high-speed optical communications. The optimization is applied to the ARJ4A protograph family, and further extended to protograph-based spatially coupled LDPC codes assuming a windowed decoder. Full field simulations via the split-step Fourier method are used to verify the analysis. The results show performance gains of up to 0.25 dB, which translate into a possible extension of the transmission reach by roughly up to 8%, without significantly increasing the system complexity.

# 1 Introduction

There is currently a large interest in developing practical coded modulation (CM) schemes that can achieve high spectral efficiency close to the ultimate capacity limits of optical fibers [1]. Pragmatic bit-interleaved coded modulation (BICM) in combination with low-density parity-check (LDPC) codes is one of the most popular capacity-approaching CM techniques for achieving high spectral efficiency, due to its simplicity and flexibility [2]. For a BICM system, a helpful abstraction is to think about transmitting data using a single forward error correction (FEC) encoder over a set of parallel binary-input channels, or simply bit channels, with different qualities. This is due to the fact that bits are not protected equally throughout the signal constellation. With this useful picture, an immediate problem is how to best allocate the coded bits from the encoder to these channels. As a baseline, a random or consecutive/sequential mapper[1] is commonly employed in practice. However, by optimizing the bit mapper, one can improve the system performance, at almost no increased complexity cost. While BICM has been studied for fiber-optical communications by many authors, see e.g., [3] or [4] and references therein, to the best of our knowledge, optimized bit mappers have not yet been considered for such systems.

In this paper, we address the bit mapper optimization for a BICM system based on LDPC codes in the context of long-haul fiber-optical communications. Our target

[1] In the literature, the term "bit interleaver" is also frequently used.

system operates over a communication link with a lumped amplification scheme and without optical inline dispersion compensation. In general, the signal undergoes a complicated evolution and interacts with amplified spontaneous emission (ASE) noise and co-propagating signals through dispersive and nonlinear effects. For dispersion uncompensated transmission, it has been shown that an additive Gaussian noise (GN) model can be assumed, provided that dispersive effects are dominant and nonlinear effects are weak [5, 6]. We use the GN model for our analysis, which accounts for both the ASE noise from inline erbium-doped fiber amplifiers (EDFAs) and nonlinear noise due to the optical Kerr effect.

The starting point for the optimization problem is a fixed modulation format and a given error correction code, i.e., we do not consider the *joint* design of the modulation, bit mapper, and code. This scenario is often encountered in practice when the modulation and code have been designed separately and/or are predetermined according to some communication standard. Our focus is on protograph-based LDPC codes [7], which are very attractive from a design perspective and allow for a high-speed hardware implementation, suitable for fiber-optical communications [8]. A protograph is a (small) bipartite graph, from which the Tanner graph defining the code is obtained by a copy-and-permute procedure. As one illustrative example for protograph-based codes, we consider the ARJ4A protographs developed by researchers from JPL/NASA in [9]. We also consider bit mapper optimization for protograph-based spatially coupled low-density parity-check (SC-LDPC) codes using the windowed decoder (WD) proposed in [10]. SC-LDPC codes, originally introduced as LDPC convolutional codes in [11], have attracted a lot of attention due to their capacity-achieving performance under belief propagation (BP) decoding for a variety of communication channels [12]. SC-LDPC codes can be constructed using protographs and they are considered as viable candidates for future spectrally efficient fiber-optical systems [8].

Most of the literature about bit mapper optimization deals with irregular LDPC codes that are not based on protographs, see e.g., [13, 14]. Attempts to improve the performance of BICM systems with protograph-based codes through bit mapper optimization have been previously made in [15–17]. In [15], a mapping strategy inspired by the water-filling algorithm for parallel channels called variable degree matched mapping (VDMM) is presented. This idea is extended in [16], where the authors exhaustively search over all possible nonequivalent connections between protograph nodes and modulation bits showing performance improvements over VDMM. As pointed out in [17], the above approaches are somewhat restrictive in the sense that only certain protographs can be used with certain modulation formats. A more flexible approach is proposed in [17], which is in principle suitable for any protograph structure and modulation but relies on a larger intermediate protograph.

Our optimization of the bit mapper is based on the decoding threshold over the additive white Gaussian noise (AWGN) channel, assuming a fixed number of decoding iterations. The decoding threshold divides the channel quality parameter range (in our

case the equivalent signal-to-noise ratio (SNR) of the GN model) into a region where reliable decoding is possible and where it is not. In the asymptotic case, i.e., assuming infinite codeword length, density evolution (DE) or one-dimensional simplifications via extrinsic information transfer (EXIT) functions can be used to find the decoding threshold for LDPC codes under BP decoding [18]. Approximate decoding thresholds of protograph-based codes assuming binary modulation can be obtained by using the protograph extrinsic information transfer (P-EXIT) analysis [19]. The approach proposed here is different from the ones in [15–17], in the sense that the P-EXIT analysis is modified to allow for fractional allocations between protograph nodes and modulation bits. This also allows for an unrestricted matching of protographs and modulation formats and additionally does not suffer from an increased design complexity due to a larger intermediate protograph. We also discuss several ways to reduce the optimization complexity, including periodic bit mappers for SC-LDPC codes with a WD. This is based on the results we previously presented in [20], where optimized bit mappers are found for (nonprotograph-based) SC-LDPC codes assuming parallel binary erasure channels (BECs) without considering the WD. The use of a WD in this paper is motivated by the reduced complexity and decoding delay with respect to full decoding. Finally, we provide a simulative verification assuming both linear and nonlinear transmission scenarios. For the latter case, we use the split-step Fourier method (SSFM) to show that the performance improvements predicted from the AWGN analysis can be achieved for a realistic transmission scenario including nonlinear effects.

### 1.1 Notation

Vectors and matrices are typeset in bold font by lowercase letters $\mathbf{a}$ and capital letters $\mathbf{A}$, respectively. Matrix transpose is denoted by $(\cdot)^\intercal$, Hermitian transpose by $(\cdot)^\dagger$, and the squared norm of a complex vector by $\|\mathbf{a}\|^2$. $\mathbf{I}_n$ denotes the identity matrix of size $n$. Complex conjugation is denoted by $(\cdot)^*$. $\delta(t)$ is Dirac's delta function, whereas $\delta[k]$ is the Kronecker delta. Convolution is denoted by $\circledast$. $\mathbb{N}_0$, $\mathbb{R}$, and $\mathbb{C}$ denote the set of nonnegative integers, real numbers, and complex numbers, respectively. Random variables and vectors are denoted by capital letters and their realizations by lowercase letters. The probability density function (PDF) of a random variable $Y$ conditioned on the realization of another random variable $X$ is denoted by $f_{Y|X}(y|x)$ and the expected value by $\mathbb{E}[\,\cdot\,]$.

## 2 System Model

### 2.1 Continuous-Time Channel

We consider transmission of a polarization-multiplexed (PM) signal over a standard single-mode fiber (SSMF) with a lumped amplification scheme as shown in Fig. 1.

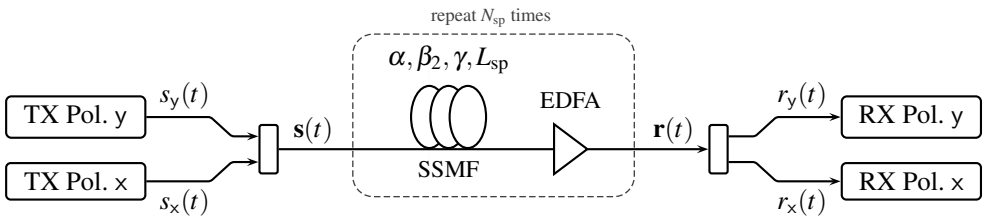



**Figure 1:** Block diagram of the consider fiber-optical transmission system.

The optical link consists of $N_{\text{sp}}$ spans of SSMF with length $L_{\text{sp}}$. The baseband signal in each polarization[2] is generated via a linear pulse modulation according to $s_{\text{x}}(t) = \sum_k s_{\text{x},k} p(t - k/R_s)$, where $s_{\text{x},k} \in \mathbb{C}$ are the information symbols, $p(t)$ the real-valued pulse shape, and $R_s$ the symbol rate. The PM signal $\mathbf{s}(t) = (s_{\text{x}}(t), s_{\text{y}}(t))^\intercal$ is launched into the fiber and propagates according to [21, Ch. 3]

$$\frac{\partial \mathbf{v}(t,z)}{\partial z} = -\frac{\alpha - g(z)}{2}\mathbf{v}(t,z) - \jmath\frac{\beta_2}{2}\frac{\partial^2 \mathbf{v}(t,z)}{\partial t^2} + \jmath\gamma\mathbf{v}(t,z)\|\mathbf{v}(t,z)\|^2 + \mathbf{w}(t,z), \qquad \text{(C.1)}$$

where $\mathbf{v}(t,z)$ is the complex baseband representation of the electric field and the input to the first fiber span and the output signal are $\mathbf{s}(t) = \mathbf{v}(t,0)$ and $\mathbf{r}(t) = \mathbf{v}(t, N_{\text{sp}}L_{\text{sp}})$, respectively. In (C.1), $\alpha$ is the attenuation coefficient, $\beta_2$ the chromatic dispersion coefficient, and $\gamma$ the nonlinear Kerr parameter. The terms $g(z)$ and $\mathbf{w}(t,z) = (w_{\text{x}}(t,z), w_{\text{y}}(t,z))^\intercal$ model the amplifier gain and the generated ASE noise [22, p. 84]. Each EDFA introduces circularly symmetric complex Gaussian noise with two-sided power spectral density (PSD) $\mathsf{N}_\ell = (G-1)h\nu_s n_{\text{sp}}$ [1, eq. (54)] per polarization, where $G = e^{\alpha L_{\text{sp}}}$ is the amplifier gain, $h$ is Planck's constant, $\nu_s$ the carrier frequency, and $n_{\text{sp}}$ the spontaneous emission factor. A standard coherent linear receiver is used, consisting of an equalizer, a pulse-matched filter and a symbol-time sampler. This amounts to $r_{\text{x},k} = r_{\text{x}}(t) \circledast h(t) \circledast p(-t)|_{t=k/R_s}$, where the frequency response of the equalizer $h(t)$ is $H(f) = \exp(\jmath 2\beta_2\pi^2 f^2 N_{\text{sp}}L_{\text{sp}})$.

## 2.2 Discrete-Time Channel

An approximate discrete-time model for the received samples $\mathbf{r}_k = (r_{\text{x},k}, r_{\text{y},k})^\intercal$ based on the transmitted symbols $\mathbf{s}_k = (s_{\text{x},k}, s_{\text{y},k})^\intercal$ is given by $\mathbf{r}_k \approx \zeta\mathbf{s}_k + \mathbf{n}_k + \tilde{\mathbf{n}}_k$, where $\zeta \in \mathbb{C}$ [5]. The term $\mathbf{n}_k = (n_{\text{x},k}, n_{\text{y},k})^\intercal$ accounts for the linear ASE noise with $\mathbb{E}[\mathbf{N}_k\mathbf{N}_{k'}^\dagger] = P_{\text{ASE}}\mathbf{I}_2\delta[k-k']$, where $P_{\text{ASE}} = N_{\text{sp}}\mathsf{N}_\ell R_s$. The term $\tilde{\mathbf{n}}_k = (\tilde{n}_{\text{x},k}, \tilde{n}_{\text{y},k})^\intercal$ accounts for nonlinear noise with $\mathbb{E}[\tilde{\mathbf{N}}_k\tilde{\mathbf{N}}_{k'}^\dagger] = \eta P^3\mathbf{I}_2\delta[k-k']$, where $P = \lim_{T\to\infty}(\int_{-T}^{T} s_{\text{x}}(t)^2\,\mathrm{d}t)/(2T)$ is the transmit power per polarization (assumed to be equal for both polarizations). $\eta$ is a function of the link parameters $\alpha, \beta_2, \gamma, L_{\text{sp}}, N_{\text{sp}}$ and the symbol rate $R_s$ [5, eq. (15)], and $|\zeta|^2 = 1 - |\eta|P^2$. The conditional PDF in this model is assumed to be Gaussian

[2] We give expressions for polarization x only, if polarization y has an equivalent expression.

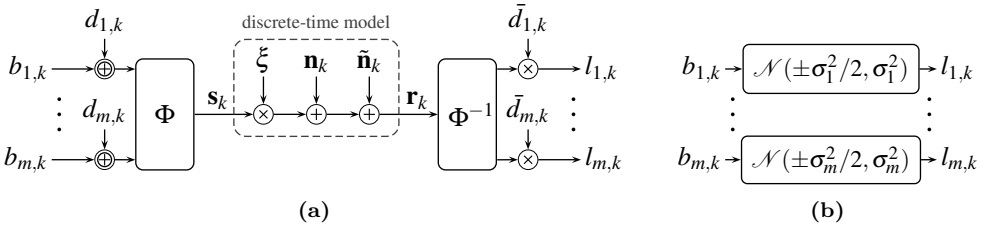



**Figure 2:** (a) BICM block diagram including the channel symmetrization technique. (b) Approximate model with parallel Gaussian LLR channels.

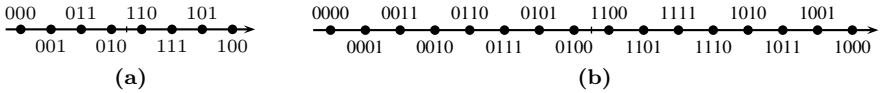



**Figure 3:** The considered signal constellations in each dimension.

according to

$$f_{\mathbf{R}_k|\mathbf{S}_k}(\mathbf{r}_k|\mathbf{s}_k) = \frac{1}{(\pi P_{\mathrm{N}})^2} \exp\left(-\frac{\|\mathbf{r}_k - \zeta \mathbf{s}_k\|^2}{P_{\mathrm{N}}}\right), \tag{C.2}$$

where $P_{\mathrm{N}} = P_{\mathrm{ASE}} + \eta P^3$. The equivalent SNR is defined as $\rho \triangleq |\zeta|^2 P/(P_{\mathrm{ASE}} + \eta P^3)$.

## 2.3 Bit-Interleaved Coded Modulation

The transmitted symbols $\mathbf{s}_k$ in each time instant $k$ take on values from a discrete signal constellation $\mathcal{X} \subset \mathbb{C}^2$. Each point in the constellation is labeled with a unique binary string of length $m = \log_2 |\mathcal{X}|$, where $b_i(\mathbf{a})$, $1 \leq i \leq m$, denotes the $i$th bit in the binary string assigned to $\mathbf{a} \in \mathcal{X}$. Consider now the block diagram shown in Fig. 2(a), where the modulo 2 addition of the independent and identically distributed bits $d_{i,k}$ and multiplication by $\bar{d}_{i,k} = (-1)^{d_{i,k}}$ serves as a symmetrization technique [23]. At each time instant $k$, the modulator $\Phi$ takes $m$ bits $b_{i,k}$, $1 \leq i \leq m$, and maps them to one of the constellation points according to the binary labeling. We consider two product constellations of one-dimensional constellations labeled with the binary reflected Gray code (BRGC) as shown in Fig. 3, which we refer to as PM-64-QAM and PM-256-QAM. At the receiver, the demodulator $\Phi^{-1}$ computes soft reliability information about the transmitted bits in the form of the log-likelihood ratios (LLRs)

$$l_{i,k} \triangleq \log \frac{f_{\mathbf{R}_k|B_{i,k}}(\mathbf{r}_k|0)}{f_{\mathbf{R}_k|B_{i,k}}(\mathbf{r}_k|1)} = \log \frac{\sum_{\mathbf{s}\in\mathcal{X}_{i,0}} f_{\mathbf{R}_k|\mathbf{S}_k}(\mathbf{r}_k|\mathbf{s})}{\sum_{\mathbf{s}\in\mathcal{X}_{i,1}} f_{\mathbf{R}_k|\mathbf{S}_k}(\mathbf{r}_k|\mathbf{s})}, \tag{C.3}$$

where $\mathcal{X}_{i,u} \triangleq \{\mathbf{a} \in \mathcal{X} : b_i(\mathbf{a}) = u\}$ is the subconstellation where all points have the bit $u$ at the $i$th position of their binary label.

A useful way to think about the setup depicted in Fig. 2(a) is to imagine transmitting over a set of parallel bit channels, where one may interpret the conditional distribution

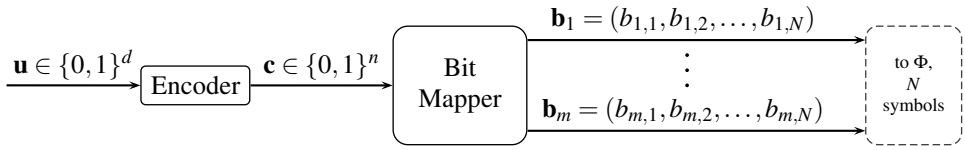



**Figure 4:** Block diagram illustrating the purpose of the bit mapper.

of the LLR $f_{L_{i,k}|B_{i,k}}(\cdot\,|\,\cdot)$ as a bit channel. In the following, we say that a bit channel $f_{L|B}(l|b)$ is symmetric if $f_{L|B}(l|0) = f_{L|B}(-l|1)$ and the channel is referred to as an LLR channel if $f_{L|B}(l|0)e^l = f_{L|B}(l|1)$. To simplify the analysis, the original bit channels are replaced with parallel symmetric Gaussian LLR channels, as shown in Fig. 2(b), where an LLR channel $f_{L|B}(l|b)$ is called a symmetric Gaussian LLR channel with parameter $\sigma^2$ if $L \sim \mathcal{N}(\sigma^2/2, \sigma^2)$ conditioned on $B = 0$ and $L \sim \mathcal{N}(-\sigma^2/2, \sigma^2)$ conditioned on $B = 1$. In order to find a correspondence between the LLR channels $f_{L_{i,k}|B_{i,k}}(\cdot\,|\,\cdot)$ and the parameters $\sigma_i^2$, one may match the mutual information (MI) according to $\sigma_i^2 = J^{-1}(I_i(\rho))^2$, where $I_i(\rho) = I(B_{i,k}; L_{i,k})$ is independent of $k$ and $J(\sigma)$ denotes the MI between the output of a symmetric Gaussian LLR channel and uniform input bits.

Consider now the case where a binary code $\mathcal{C} \subset \{0,1\}^n$ of length $n$ and dimension $d$ is employed and each codeword $\mathbf{c} = (c_1, \ldots, c_n)$ is transmitted using $N = n/m$ symbols $\mathbf{s}_k$. The allocation of the coded bits to the modulation bits (i.e., the different bit channels in Fig. 2(b)) is determined by a bit mapper as shown in Fig. 4, where the vectors $\mathbf{b}_1, \ldots, \mathbf{b}_m$ are of length $N$. Our goal is to find good bit mappers for a fixed code and modulation. As a baseline, we consider a consecutive mapper according to $b_{i,k} = c_{(k-1)m+i}$ for $1 \le i \le m$, $1 \le k \le N$.

# 3 Protograph-Based LDPC Codes

An LDPC code of length $n$ and dimension $d$ is defined via a sparse parity-check matrix $\mathbf{H} = [h_{i,j}] \in \{0,1\}^{c \times n}$, where $c = n - d$. There exist different methods to construct "good" LDPC codes, i.e., good matrices $\mathbf{H}$. One popular method is by using protographs [7]. An LDPC code can be represented by using a bipartite Tanner graph consisting of $n$ variable nodes (VNs) and $c$ check nodes (CNs), where the $i$th CN is connected to the $j$th VN if $h_{i,j} = 1$. A protograph is also a bipartite graph defined by an adjacency matrix $\mathbf{P} = [p_{i,j}] \in \mathbb{N}_0^{c' \times n'}$, called the base matrix. Given $\mathbf{P}$, a parity-check matrix $\mathbf{H}$ is obtained by replacing each entry $p_{i,j}$ in $\mathbf{P}$ with a random binary $M$-by-$M$ matrix which contains $p_{i,j}$ ones in each row and column. This procedure is called lifting and $M \ge \max_{i,j} p_{i,j}$ is the so-called lifting factor. Graphically, this construction amounts to copying the protograph $M$ times and subsequently permuting edges. Parallel edges, i.e., for $p_{i,j} > 1$, are permitted in the protograph and are resolved in the lifting procedure. The design rate of the code is given by $R = 1 - c/n = 1 - c'/n'$, where $c = c'M$ and $n = n'M$.

## 3.1 ARJ4A Codes

As one example to illustrate the bit mapper optimization technique, we consider the ARJ4A code family defined by the protographs in [9, Fig. 8]. The base matrix $\mathbf{P}^{(\ell)}$ of the ARJ4A code ensemble with parameter $\ell \in \mathbb{N}_0$ can be recursively defined via [17]

$$\mathbf{P}^{(\ell)} = \left(\begin{array}{c|cc} \mathbf{P}^{(\ell-1)} & \begin{matrix}0 & 0\\ 3 & 1\\ 1 & 3\end{matrix} \end{array}\right), \qquad \mathbf{P}^{(\ell=0)} = \begin{pmatrix}1 & 2 & 0 & 0 & 0\\ 0 & 3 & 1 & 1 & 1\\ 0 & 1 & 2 & 2 & 1\end{pmatrix} \tag{C.4}$$

with $c' = 3$ and $n' = 5 + 2\ell$. VNs corresponding to the second column of the base matrix are punctured, leading to a design rate of $R = (1 - c'/n') \cdot n'/(n' - 1) = (\ell + 1)/(\ell + 2)$.

## 3.2 Spatially Coupled LDPC Codes

SC-LDPC codes have parity-check matrices with a band-diagonal structure (for a general definition see, e.g., [12]). For completeness, we briefly review the construction via protographs in [24], [10, Sec. II-B]. The base matrix $\mathbf{P}_{[T]}$ of a $(J, K)$ regular, protograph-based SC-LDPC code with termination length $T$ can be constructed by specifying matrices $\mathbf{P}_i$, $0 \leq i \leq m_\mathrm{s}$ of dimension $J'$ by $K'$, where $m_\mathrm{s}$ is referred to as the memory. The matrices are such that $\mathbf{P} = \sum_{i=0}^{m_\mathrm{s}} \mathbf{P}_i$ has column weight $J$ and row weight $K$ for all columns and rows, respectively. Given $T$ and the matrices $\mathbf{P}_i$, the base matrix $\mathbf{P}_{[T]}$ is constructed as

$$\mathbf{P}_{[T]} = \begin{pmatrix} \mathbf{P}_0 & & \\ \mathbf{P}_1 & \ddots & \\ \vdots & \ddots & \mathbf{P}_0 \\ \mathbf{P}_{m_s} & \ddots & \mathbf{P}_1 \\ & \ddots & \vdots \\ & & \mathbf{P}_{m_s} \end{pmatrix}. \tag{C.5}$$

From the dimensions of $\mathbf{P}_{[T]}$ one can infer a design rate of $R(T) = 1 - (T + m_\mathrm{s})J'/(TK')$. As $T$ grows large, the rate approaches $R(\infty) = 1 - J'/K'$.

Since our goal is not to optimize the code, we rely on base matrices that have been proposed elsewhere in the literature, in particular in combination with a WD which we discuss below. We consider $\mathbf{P}_0 = (2, 2, 2)$ and $\mathbf{P}_1 = (1, 1, 1)$ according to [10, Design rule 1], where $J' = 1$, $K' = 3$, $m_\mathrm{s} = 1$, and $R(\infty) = 2/3$.

## 3.3 Decoding and Asymptotic EXIT Analysis

We use a modified version of the P-EXIT analysis as a tool to predict the iterative BP performance behavior of the protograph-based codes [19]. A detailed description

**Algorithm 1:** P-EXIT analysis of the WD for a $(J, K)$ regular SC-LDPC protograph.

**Input**: $l_{\max}$ (max. iterations per window), $p_{\text{tar}}$ (target error probability), $W$, $(J', K')$, $\rho$

**Output**: $S$ (decoding success, either true or false), $l_{\text{s}}$ (iterations until successful decoding)

**for** $i = 1$ **to** $n'$ **do** `/* initialization of channel variances for VNs */`
  **if** `VN` $i$ is punctured **set** $\sigma_i^2 = 0$ `/* treat as an erasure */`
  **else set** $\sigma_i^2 = f(\rho)$ `/* E.g.,` $f(\rho) = 8R\rho$ `if` $\rho = E_b/N_o$ `[25] */`
$l_{\text{s}} = 0$ `/* total iteration counter */`
**for** $j = -W + 2$ **to** $T$ **do**
  $c_{\text{start}} \leftarrow \max((j-1)J' + 1, 1)$ `/* first index of active CNs */`
  $c_{\text{end}} \leftarrow \min((W + j - 1)J', m)$ `/* last index of active CNs */`
  $v_{\text{start}} \leftarrow \max((j-1)K' + 1, 1)$ `/* first index of active VNs */`
  $v_{\text{end}} \leftarrow \min((W + j - 1)K', n')$ `/* last index of active VNs */`
  $t_{\text{start}} \leftarrow \max((j-1)K' + 1, 1)$ `/* first index of target VNs */`
  $t_{\text{end}} \leftarrow \max((j-1)K' + K', K')$ `/* last index of target VNs */`
  $l = 0$
  **while** $l \leq l_{max}$ **do**
    **if** `mean` (error probability of `VN` $t_{\text{start}}$ to $t_{\text{end}}$) $< p_{\text{tar}}$ **break while**
    **for** $i = v_{\text{start}}$ **to** $v_{\text{end}}$ **compute** `Messages` $(\sigma_i^2)$ of VN $i$ `/* Eq. (9.46) [25] */`
    **for** $i = c_{\text{start}}$ **to** $c_{\text{end}}$ **compute** `Messages` of CN $i$ `/* Eq. (9.47) [25] */`
    $l \leftarrow l + 1$ and $l_{\text{s}} \leftarrow l_{\text{s}} + 1$
**if** `mean` (error probability of `VN` 1 to $n'$) $< p_{\text{tar}}$ **set** $S = 1$ **else set** $S = 0$

of this tool for binary modulation is available in [19] and [25, Algorithm 9.2]. Here, we only describe the necessary modifications to account for the WD and the nonbinary modulations. We start with the former and explain the latter in the next section.

We employ the WD scheme developed in [10]. WD helps to alleviate the long decoding delays and high decoding complexity of SC-LDPC codes under full BP decoding by exploiting the fact that two VNs are not involved in the same parity-check equation if they are at least $(m_{\text{s}} + 1)K'$ columns apart [10]. The WD restricts message updates to a subset of VNs and CNs in the entire graph. After a predetermined number of decoding iterations, this subset changes and the decoding window slides to the next position. Pseudocode for the modified P-EXIT analysis of SC-LDPC codes accounting for the WD is presented in Algorithm 1. The main difference with respect to BP decoding is the window size parameter $W$, which specifies the number of active CNs in the protograph considered in each window as a multiple of $J'$. The P-EXIT analysis for the standard BP decoder can be recovered from Algorithm 1 by setting $T = 1$, $W = 1$, $J' = c'$, and $K' = n'$.

# 4 Bit Mapper Optimization

## 4.1 Asymptotic Bit Mapper Model

Each VN in the protograph represents $M$ VNs in the lifted Tanner graph. Since a VN corresponds to one bit in a codeword, the $n'$ VNs in the protograph give rise to $n'$ different classes of coded bits that are treated as statistically equivalent in the P-EXIT analysis. In particular, for binary modulation, each protograph VN is assigned with one input variance, corresponding to either a punctured bit or the Gaussian LLR channel (see lines 2 and 3 in Algorithm 1). For nonbinary modulations, VNs in the same class can in principle have different input densities. Assume for example that a given protograph is lifted with an even lifting factor $M$ and coded bits are mapped consecutively to a 4-ary modulation. Then, $M/2$ VNs in each class are allocated to the first modulation bit and $M/2$ to the second.

We model the bit mapper by specifying the assignment of VN classes to the bit channels via a matrix $\mathbf{A} = [a_{i,j}] \in \mathbb{R}^{m \times n'}$, where $a_{i,j}$, $0 \leq a_{i,j} \leq 1$ $\forall i, j$ denotes the proportional allocation of VNs from the $j$th class (corresponding to the $j$th column in the base matrix) allocated to the $i$th bit in the signal constellation. The approaches in [15–17] can be recovered by considering only nonfractional assignments, i.e., $a_{i,j} \in \{0, 1\}$. In that case, VNs of the original protograph [15, 16] or an intermediate protograph [17] are directly assigned to the modulation bits.

We point out that, instead of interpreting $a_{i,j}$ as a deterministic fraction of VNs in a particular class allocated to a particular channel, one should interpret $a_{i,j}$ as a probability, and study the bit mapper as a probabilistic mapping device that assigns coded bits randomly to channels, similar to [26]. Under this assumption, one may argue that the VNs belonging to a certain class "see" an equivalent bit channel which is the average of the original bit channels $f_{L_{i,k}|B_{i,k}}(l|b)$, weighted according to the probabilities $a_{i,j}$. The MI of each equivalent bit channel is a weighted average of the original channels' MI as shown in the following lemma.

**Lemma 1:** *Let $\{f_{L_i|B_i}(l|b) : 1 \leq i \leq m\}$ be a collection of symmetric LLR channels. Consider a new channel $f_{L|B}(l|b)$, where transmission takes place over the $i$th channel in the collection with probability $\alpha_i$ and $\sum_i \alpha_i = 1$. Then $I(L;B) = \sum_i \alpha_i I(L_i;B_i)$ for uniform input bits.*

*Proof.* The channel $f_{L|B}(l|b)$ is a symmetric LLR channel. The claim then follows from $f_{L|B}(l|b) = \sum_i \alpha_i f_{L_i|B_i}(l|b)$ and the fact that the MI between the output of a symmetric LLR channel $f_{L|B}(l|b)$ and uniform input bits is $I(L;B) = 1 - \int_{-\infty}^{\infty} f_{L|B}(l|0) \log_2(1 + e^{-l})\, \mathrm{d}l$. □

If we collect the MI corresponding to the original $m$ symmetric LLR channels in a vector $\mathbf{I}(\rho) = (I_1(\rho), \ldots, I_m(\rho))$, then, multiplying $\mathbf{I}(\rho)$ by $\mathbf{A}$ leads to a vector $(\tilde{I}_1, \tilde{I}_2, \ldots, \tilde{I}_{n'})$ with the MIs corresponding to the averaged bit channels as seen by the $n'$ VN classes. These averaged bit channels are modeled as symmetric Gaussian LLR channels with parameters $(\sigma_1^2, \ldots, \sigma_{n'}^2)$. In particular, the P-EXIT analysis for nonbinary modulation

is obtained by changing the initialization step in line 3 of Algorithm 1 and assigning $\sigma_i^2 = J^{-1}(\tilde{I}_i)^2$, where the algorithm takes $\mathbf{A}$ as an additional input to compute $\tilde{I}_i$ as described.

In order to have a valid probabilistic assignment, all columns in $\mathbf{A}$ have to sum to one and all rows in $\mathbf{A}$ have to sum to $n'/m$, i.e., we have $mn'$ equality constraints in total. The first condition ensures that, asymptotically, all VNs are assigned to a channel, while the second condition ensures that all parallel channels are used equally often. The set of valid assignment matrices is denoted by $\mathcal{A}^{m\times n'} \subset \mathbb{R}^{m\times n'}$. In the case of punctured VNs, the corresponding columns in $\mathbf{A}$ are removed and $n'$ is interpreted as the number of *unpunctured* VNs.

## 4.2 Optimization

For a given bit mapper, i.e., for a given assignment matrix $\mathbf{A}$, an approximate decoding threshold $\rho^*(\mathbf{A})$ can be found using Algorithm 1 as follows. Fix a certain precision $\delta$, target bit error probability $p_{\text{tar}}$, and maximum number of iterations $l_{\max}$. Starting from some SNR $\rho$ where Algorithm 1 converges to a successful decoding, $S = 1$, iteratively decrease $\rho$ by $\delta$ until the decoding fails. The smallest $\rho$ for which $S = 1$ is declared as the decoding threshold $\rho^*(\mathbf{A})$. For any $\rho \geq \rho^*(\mathbf{A})$, we denote the number of iterations until successful decoding by $l_{\text{s}}(\mathbf{A}, \rho)$.

We are interested in optimizing $\mathbf{A}$ in terms of the decoding threshold for a given protograph and modulation format. The optimization problem is thus

$$\mathbf{A}_{\text{opt}} = \underset{\mathbf{A}\in\mathcal{A}^{m\times n'}}{\operatorname{argmin}} \quad \rho^*(\mathbf{A}), \tag{C.6}$$

where the baseline system realizes a mapping of coded bits to modulation bits such that $a_{i,j} = 1/m$, $\forall i, j$. The corresponding assignment matrix is denoted by $\mathbf{A}_{\text{uni}}$. The search space $\mathcal{A}^{m\times n'}$ can be regarded as a convex polytope $\mathcal{P}$ in $p = (m-1)(n'-1)$ dimensions by removing the last row and column in $\mathbf{A}$, replacing the equality constraints with inequality constraints, and writing the matrix elements in a vector $\mathbf{x} \in \mathbb{R}^p$ according to the prescription $x_{(i-1)n'+j} = a_{i,j}$ for $1 \leq i \leq m-1$ and $1 \leq j \leq n'-1$. While the search space is convex, one can show by simple examples that the objective function is nonconvex in $\mathcal{P}$. In the following, we discuss ways to obtain good bit mappers with reasonable effort. We also remark that some of the optimization approaches proposed previously in the context of bit mapper optimization for irregular LDPC codes are not necessarily appropriate in our case due to the higher number of VN classes, i.e., they can be too complex (for example the iterative grid search in [13]) or do not explore the search space efficiently (simple hill climbing approaches as in [14]).

First, as an alternative to directly optimizing the decoding threshold, we iteratively optimize the convergence behavior in terms of the number of iterations until successful decoding as follows. Initialize $\rho$ to the decoding threshold for the baseline bit mapper, i.e., $\rho = \rho^*(\mathbf{A}_{\text{uni}})$. Find $\mathbf{A}^*$ such that it minimizes the number of decoding iterations

until convergence for the given $\rho$, i.e.,

$$\mathbf{A}^* = \underset{\mathbf{A}\in\mathcal{A}^{m\times n'}}{\operatorname{argmin}} \quad l_{\mathrm{s}}(\mathbf{A}, \rho). \tag{C.7}$$

We we employ differential evolution [27] to solve the optimization problem in (C.7).[3] For the found optimized $\mathbf{A}^*$, calculate the new decoding threshold $\rho^*(\mathbf{A}^*)$. If the threshold did not improve, stop. Otherwise, set $\rho = \rho^*(\mathbf{A}^*)$ and repeat the optimization. A similar iterative approach was proposed in [28, Sec. IV] to find optimized degree distributions for irregular LDPC codes. The above approach was used by the authors to find good bit mappers for SC-LDPC codes in [20] for parallel BECs. The computational complexity can be significantly reduced compared to the threshold minimization (C.6). However, it is not guaranteed to be equivalent to a true threshold optimization, i.e., in general $\mathbf{A}_{\mathrm{opt}} \neq \mathbf{A}^*$.

The optimization complexity is further reduced by constraining the maximum number of iterations $l_{\max}$. Practical systems commonly operate with a relatively small number of BP iterations. For example, in Sec. 5, we assume 50 BP iterations, and hence the decoding thresholds are optimized for the same number of iterations. In the simulative verification, we have observed that the performance of the finite-length codes assuming 50 BP iterations is generally better using a bit mapper that is also optimized for $l_{\max} = 50$ compared to, say, $l_{\max} = 1000$, although the differences were small.

Additionally, for SC-LDPC codes, we take advantage of the structure of the optimized bit mappers for parallel BECs [20], which show a certain form of periodicity. The optimization complexity can then be reduced by assuming that the optimal solution lies in a lower-dimensional subspace of $\mathcal{P}$, defined by assignment matrices that take on a periodic form as $\mathbf{A} = (\mathbf{A}', \mathbf{A}'', \mathbf{A}'', \cdots, \mathbf{A}'', \mathbf{A}''')$, with $m \times V$ matrices $\mathbf{A}'$, $\mathbf{A}''$, and $\mathbf{A}'''$, where $V$ is the periodicity factor. If $V$ is chosen small enough, the dimensionality of the search space (i.e., $(m-1)(3V-1)$) can be substantially reduced, which generally improves the convergence speed of the differential evolution algorithm.

# 5 Results and Discussion

In this section, we present and discuss numerical results, and illustrate the performance gains that can be achieved by employing optimized bit mappers. For the baseline systems, we use a consecutive mapping of coded bits to modulation bits. Alternatively, one may use a uniformly random mapping, which has the same expected performance.

In order to show the flexibility of the technique, we consider four different scenarios, combining both modulation formats with one code based on the ARJ4A protographs and one SC-LDPC code, where the lifting factor is $M = 3000$ in all cases. For simplicity, the

[3] The fact that the search space is a convex polytope makes it relatively straightforward to take uniformly distributed starting points in order to initialize the starting population in the differential evolution algorithm.

codes are randomly generated without further consideration of the graph structure. The protograph lifting procedure can in principle be combined with standard techniques to avoid short graph cycles that may potentially lead to high error floors[4] [25, Ch. 6.3]. A rate $R = 2/3$ code based on the ARJ4A protograph for $\ell = 1$ is used, which is denoted by $\mathcal{C}_{\mathrm{ARJ4A}}$. For the spatially coupled case with $T = 30$, a code based on the protograph described in Sec. 3.2 is used, which is denoted by $\mathcal{C}_{\mathrm{SC}}$. For the given value of $T$, the design rate is $R(30) = 0.656$. For the ARJ4A code, standard BP decoding is assumed with $l_{\max} = 50$, while for the SC-LDPC codes, we employ a WD with $W = 5$ and $l_{\max} = 10$, which again amounts to a total of 50 iterations per decoded bit. We also tried other combinations of $W$ and $l_{\max}$ with a similar total number of iterations and this combination gave the best performance. For the bit mapper optimization and in particular the P-EXIT analysis, we use the same values for $l_{\max}$ and $W$, and additionally $p_{\mathrm{tar}} = 10^{-5}$. The finite-length bit mappers are obtained via the rounded matrix $M\mathbf{A}^*$ from which the index assignment of coded bits to modulation bits is determined.

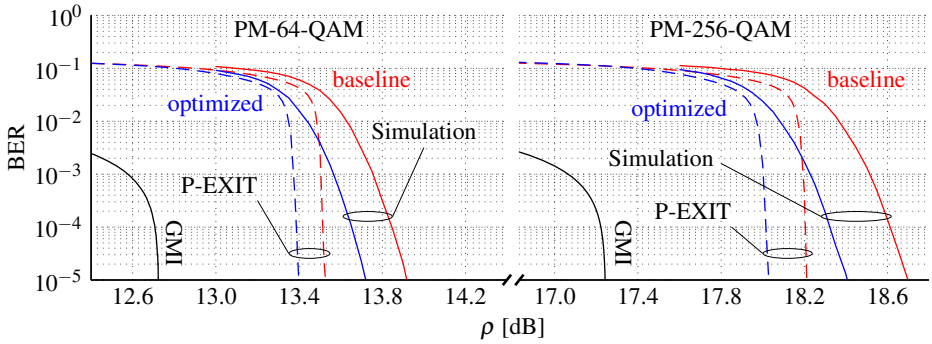
## 5.1 Linear Transmission

We start by providing a verification of the proposed optimization technique assuming an AWGN channel. This case is obtained when nonlinear effects are ignored, i.e., $\gamma = 0$. In this case, the channel PDF (C.2) is valid without approximations.

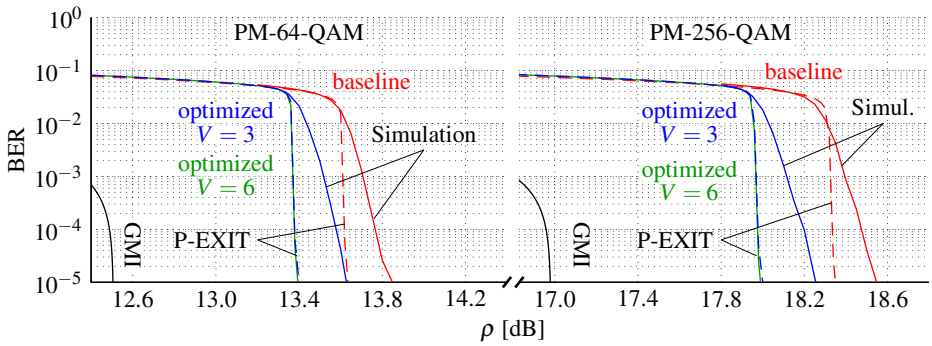
In Fig. 5(a), the predicted bit error rate (BER) of the ARJ4A code via the P-EXIT analysis is shown together with Monte Carlo simulations by the dashed and solid lines, respectively. Performance curves for the baseline bit mappers are shown in red and for the optimized ones in blue. As a reference, we also plot the BER-constrained [25, p. 17] generalized mutual information (GMI) for the corresponding spectral efficiency in each figure (the GMI is also referred to as the BICM capacity [29]). For both scenarios, it can be observed that the optimized bit mappers lead to a significant performance improvement. The gains that can be achieved at a BER of $10^{-5}$ are approximately 0.19 and 0.25 dB for PM-64-QAM and PM-256-QAM, respectively. The predicted gains from the P-EXIT analysis for the same BER is slightly less, i.e., 0.12 and 0.19 dB, respectively. The deviation of the asymptotic analysis from the actual simulation results is to be expected due to the Gaussian approximation of the LLR densities and the finite lifting factor and, hence, finite block lengths of the codes. However, it is important to observe that, even though the optimization was carried out assuming a cycle-free graph structure, the predicted performance gains for the finite-length codes is well preserved.

Similarly, the performance of the SC-LDPC code is shown in Fig. 5(b). The periodicity factor for the bit mapper optimization was set to $V = 3$. The observed gains at a BER of $10^{-5}$ are approximately 0.20 dB for PM-64-QAM and 0.25 dB for PM-256-QAM. We also show the predicted P-EXIT performance obtained for bit mappers that are optimized

[4] Alternatively, an additional outer algebraic code may be assumed, which removes remaining errors to achieve a required target BER of $10^{-15}$.




**(a)** $\mathcal{C}_{\text{ARJ4A}}$ with BP decoding and $l_{\max} = 50$




**(b)** $\mathcal{C}_{\text{SC}}$ with a WD, $W = 5$, and $l_{\max} = 10$

**Figure 5:** Comparison of the optimized bit mappers (blue) with the baseline bit mappers (red) for the linear transmission scenario. Dashed lines correspond to P-EXIT analysis and solid lines to simulation results. In (b), solid green lines correspond to the P-EXIT analysis for $V = 6$.

assuming a larger periodicity factor of $V = 6$ by the solid green curves. It can be seen that for both modulation formats, the additional gains are incremental, i.e., for PM-64-QAM the predicted performance curves virtually overlap, while for PM-256-QAM, the difference is roughly 0.01 dB. This suggests that a full optimization of $\mathbf{A}$ will be only marginally better than with $V = 3$.

We would like to stress that a direct comparison between the two codes is difficult, because of the slightly different code rates (and hence spectral efficiencies) and different decoding complexities and delays. Fair comparisons between SC-LDPC codes and LDPC block codes is an active area of research and beyond the scope of this paper.

## 5.2 Nonlinear Transmission

In this section, we consider a transmission scenario including nonlinear effects, i.e., $\gamma \neq 0$, where the assumed channel PDF (C.2) is only approximately valid. In particular, we

| parameter | meaning | value |
|---|---|---|
| $R_s$ | symbol rate | 40 Gbaud |
| $L_{\text{sp}}$ | span length | 70 km |
| $\alpha$ | attenuation coefficient (0.25 dB/km) | 0.0576 $\text{km}^{-1}$ |
| $\beta_2$ | chromatic dispersion coefficient | -21.668 $\text{ps}^2/\text{km}$ |
| $\gamma$ | nonlinear Kerr parameter | 1.4 $\text{W}^{-1}$ $\text{km}^{-1}$ |
| $\nu_s$ | carrier frequency (1550 nm) | $1.934 \times 10^{14}$ Hz |
| $n_{\text{sp}}$ | spontaneous emission factor | 1.622 |

**Table 1:** System parameters

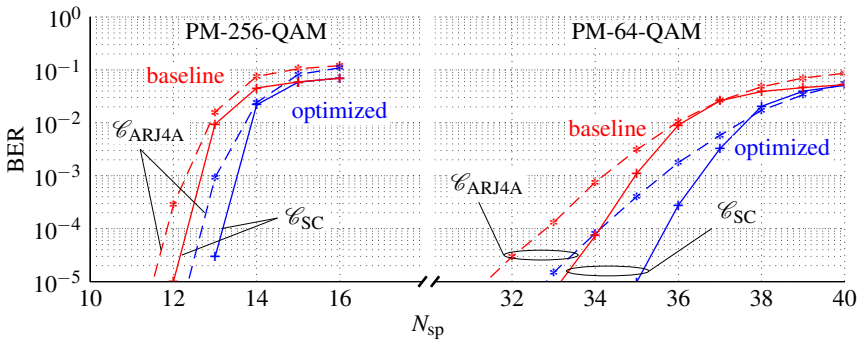



**Figure 6:** Comparison of the optimized bit mappers (blue) with the baseline bit mappers (red) for the nonlinear transmission scenario.

study the potential increase in transmission reach that can be obtained by employing the optimized bit mappers.

We consider a single channel transmission scenario to keep the simulations within an acceptable time. In the simulation model, we assume perfect knowledge about the polarization state, and perfect timing and carrier synchronization. All chosen system parameters are summarized in Table 1. Additionally, we use a root-raised cosine pulse $p(t)$ with a roll-off factor of 0.25. In order to solve (C.1), we employ the symmetric SSFM with two samples per symbol and a fixed step size of $\Delta = (10^{-4} L_{\text{D}}^2 L_{\text{NL}})^{1/3}$, where $L_{\text{D}} = 1/(|\beta_2| R_s^2)$ and $L_{\text{NL}} = 1/(\gamma P)$ is the dispersive and nonlinear length, respectively. The input power that maximizes $\rho$ according to the GN model varies between $-2.2$ dBm for $N_{\text{sp}} = 10$ and $-2.6$ dBm for $N_{\text{sp}} = 40$. For simplicity, the input power per polarization is fixed to $P = -2.5$ dBm for all values of $N_{\text{sp}}$.

In Fig. 6, the simulated BER of the PM systems using $\mathcal{C}_{\text{ARJ4A}}$ and $\mathcal{C}_{\text{SC}}$ is shown as a function of the number of fiber spans $N_{\text{sp}}$ by the dashed and solid lines, respectively. Again, curves corresponding to the baseline bit mappers are shown in red, while curves corresponding to the optimized bit mappers are shown in blue. Notice that the SNR decrease (in dB) is not linear with increasing number of spans, hence the different slopes

compared to the curves shown in Fig. 5. For PM-256-QAM, the transmission reach can be increased by roughly one additional span for both codes, at the expense of a slightly increased BER. For example, for $\mathcal{C}_{\mathrm{SC}}$, the transmission reach can be increased from 12 to 13 spans, while the BER slightly increases from $10^{-5}$ to $3 \cdot 10^{-5}$. For PM-64-QAM, the increase is roughly 1 span for $\mathcal{C}_{\mathrm{ARJ4A}}$ and roughly 2 spans for $\mathcal{C}_{\mathrm{SC}}$. In fact, these gains can be approximately predicted also from the GN model. For example, for the chosen input power and system parameters, the GN model predicts an SNR decrease of roughly 0.3 dB from $N_{\mathrm{sp}} = 12$ to $N_{\mathrm{sp}} = 13$ and 0.15 dB from $N_{\mathrm{sp}} = 34$ to $N_{\mathrm{sp}} = 35$, i.e., one would expect the performance improvements in the linear transmission scenario to translate into roughly one additional span for PM-256-QAM and one to two additional spans for PM-64-QAM. This estimate corresponds to an increase of the transmision reach by 3–8%, which is well in line with the simulation results presented in Fig. 6.

# 6 Conclusion

In this paper, we studied the bit mapper optimization for a PM fiber-optical system. Focusing on protograph-based codes, an optimization approach was proposed based on a fractional allocation of protograph bits to modulation bits via a modified P-EXIT analysis. Extensive numerical simulations were used to verify the analysis for a dispersion uncompensated link assuming both linear and nonlinear transmission regimes. The results show performance improvements of up to 0.25 dB, translating into a possible extension of the transmission reach by up to 8%.

# Acknowledgements

This work was partially funded by the Swedish Research Council under grant #2011-5961 and by the European Community's Seventh's Framework Programme (FP7/2007-2013) under grant agreement No. 271986. The simulations were performed in part on resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE.

# References

[1] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," J. Lightw. Technol. **28**, 662–701 (2010).

[2] D. J. Costello and G. D. Forney, Jr., "Channel coding: The road to channel capacity," Proc. IEEE **95**, 1150–1177 (2007).

[3] B. Smith and F. R. Kschischang, "A pragmatic coded modulation scheme for high-spectral-efficiency fiber-optic communications," J. Lightw. Technol. **30**, 2047–2053 (2012).

[4] I. B. Djordjevic, M. Arabaci, and L. L. Minkov, "Next generation FEC for high-capacity communication in optical transport networks," J. Lightw. Technol. **27**, 3518–3530 (2009).

[5] L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, and H. Wymeersch, "A discrete-time model for uncompensated single-channel fiber-optical links," IEEE Trans. Commun. **60**, 3440–3450 (2012).

[6] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links," J. Lightw. Technol. **30**, 1524–1539 (2012).

[7] J. Thorpe, "Low-density parity-check (LDPC) codes constructed from protographs," IPN Progress Report 42-154, JPL (2005).

[8] L. Schmalen, A. J. de Lind van Wijngaarden, and S. ten Brink, "Forward error correction in optical core and optical access networks," Bell Labs Tech. J **18**, 39–66 (2013).

[9] D. Divsalar, C. Jones, S. Dolinar, and J. Thorpe, "Protograph based LDPC codes with minimum distance linearly growing with block size," in "Proc. IEEE Glob. Communication Conf. (GLOBECOM)," (St. Louis, Missouri, 2005).

[10] A. R. Iyengar, M. Papaleo, P. H. Siegel, J. K. Wolf, A. Vanelli-coralli, and G. E. Corazza, "Windowed decoding of protograph-based LDPC convolutional codes over erasure channels," IEEE Trans. Inf. Theory **58**, 2303–2320 (2012).

[11] A. J. Felström and K. S. Zigangirov, "Time-varying periodic convolutional codes with low-density parity-check matrix," IEEE Trans. Inf. Theory **45**, 2181–2191 (1999).

[12] S. Kudekar, T. Richardson, and R. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC," IEEE Trans. Inf. Theory **57**, 803–834 (2011).

[13] T. Cheng, K. Peng, J. Song, and K. Yan, "EXIT-aided bit mapping design for LDPC coded modulation with APSK constellations," IEEE Commun. Lett. **16**, 777–780 (2012).

[14] G. Richter, A. Hof, and M. Bossert, "On the mapping of low-density parity-check codes for bit-interleaved coded modulation," in "Proc. IEEE Int. Symp. Information Theory (ISIT)," (Nice, Italy, 2007).

[15] D. Divsalar and C. Jones, "Protograph based low error floor LDPC coded modulation," in "Proc. IEEE Military Communications Conf. (MILCOM)," (Atlantic City, NJ, 2005).

[16] Y. Jin, M. Jiang, and C. Zhao, "Optimized variable degree matched mapping for protograph LDPC coded modulation with 16QAM," in "Proc. Int. Symp. Turbo Codes and Iterative Information Processing (ISTC)," (Brest, France, 2010).

[17] T. Van Nguyen, A. Nosratinia, and D. Divsalar, "Threshold of protograph-based LDPC coded BICM for Rayleigh fading," in "Proc. IEEE Glob. Communication Conf. (GLOBECOM)," (Houston, TX, 2011).

[18] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," IEEE Trans. Inf. Theory **47**, 599–618 (2001).

[19] G. Liva and M. Chiani, "Protograph LDPC codes design based on EXIT analysis," in "Proc. IEEE Glob. Communication Conf. (GLOBECOM)," (Washington, DC, 2007).

[20] C. Häger, A. Graell i Amat, A. Alvarado, F. Brännström, and E. Agrell, "Optimized bit mappings for spatially coupled LDPC codes over parallel binary erasure channels," in "Proc. IEEE Int. Conf. Communications (ICC)," (Sydney, Australia, 2014).

[21] G. P. Agrawal, *Lightwave Technology: Telecommunication Systems* (Wiley-Interscience, 2005).

[22] M. Secondini and E. Forestieri, "The nonlinear Schrödinger equation in fiber-optic systems," Riv. Mat. Univ. Parma **8**, 69–98 (2008).

[23] J. Hou, P. H. Siegel, L. B. Milstein, and H. D. Pfister, "Capacity-approaching bandwidth-efficient coded modulation schemes based on low-density parity-check codes," IEEE Trans. Inf. Theory **49**, 2141–2155 (2003).

[24] D. G. M. Mitchell, M. Lentmaier, and D. J. Costello, Jr., "AWGN channel analysis of terminated LDPC convolutional codes," Proc. Information Theory and Applications Workshop (ITA) (2011).

[25] W. Ryan and S. Lin, *Channel Codes Classical and Modern* (Cambridge University Press, 2009).

[26] R. Liu, P. Spasojevic, and E. Soljanin, "Reliable channel regions for good binary codes transmitted over parallel channels," IEEE Trans. Inf. Theory **52**, 1405–1424 (2006).

[27] R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," J. Global Opt. **11**, 341–359 (1997).

[28] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," IEEE Trans. Inf. Theory **47**, 619–637 (2001).

[29] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," IEEE Trans. Inf. Theory **44**, 927–946 (1998).