UNIVERSITÄT DER BUNDESWEHR MÜNCHEN
Fakultät für Elektrotechnik und Informationstechnik

Enhancing the Performance of Few-Mode Fiber Systems Impaired by Mode-Dependent Loss

M.Sc. Adriana Patricia Lobato Polo

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Universität der Bundeswehr München zur Erlangung des akademischen Grades eines

Doktor-Ingenieur
(Dr.-Ing.)

genehmigten Dissertation

Vorsitzender:	Prof. Dr. Walter Hansch
1. Gutachter:	Prof. Dr. Berthold Lankl
2. Gutachter:	Prof. Dr. Norbert Hanik

Die Dissertation wurde am 29.03.2019 bei der Universität der Bundeswehr München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 05.11.2019 angenommen. Die mündliche Prüfung fand am 18.11.2019 statt.

Abstract

The demand for the transport of vast amounts of information in long-haul optical communication has driven the research into spatial-division multiplexing. A promising subcategory of spatial-division multiplexing, mode-division multiplexing employing few-mode fibers, uses the modes for transmitting parallel streams of information. Despite the complexity escalation due to mode-multiplexers and additional digital signal processing (DSP) to unravel the inter-modal effects, mode-division multiplexed systems show in early studies energy saving potential compared to parallel single-mode-based systems. However, dealing with multiple modes has shown so far the occurrence of the undesirable effect of mode-dependent loss (MDL). In the extreme case, losses may affect selectively certain modes causing system outage. Current approaches to reduce the MDL focus either on improving the design of single inline components or balancing the modal gains and losses throughout the optical link. Although both approaches aid to improve the system performance, significant amount of MDL remains in the system. By means of simulations and experiments, this thesis describes an innovative approach to further improve the system tolerance towards MDL. The approach combines an improvement in the optical domain and the use of alternative DSP schemes. For the former enhancement, the impact of mode coupling on MDL is first studied with simulations. With three- and six-spatial-mode transmission, it is shown that inducing strong mode coupling along the fiber or using mode scramblers helps to increase the tolerance towards MDL. Next, the use of alternative receiver DSP schemes such as near maximum-likelihood (ML) detection and their combination with different levels of mode coupling is evaluated. It is demonstrated that strongly coupled modes and ML techniques greatly improve the performance of MDL-impaired systems. At this point, the complexity of such algorithms is also computed and analyzed. Finally, experiments with fiber supporting six spatial modes in single and multiple wavelength-division multiplexing channels are described. The experiments clearly show performance degradation caused by MDL, which is improved through a near-ML detector.

Kurzfassung

Die Nachfrage nach dem Transport großer Informationsmengen in der optischen Langstreckeübertragung hat die Forschung nach Raummultiplexverfahren vorangetrieben. Eine vielversprechende Unterkategorie davon, Modenmultiplexen mit Wenigmodenfasern, verwendet die Modi zur Übertragung paralleler Informationsströme. Trotz der Komplexitätssteigerung durch Modenmultiplexer und zusätzliche digitale Signalverarbeitung (DSP) zum Entzerren der intermodalen Effekte zeigen Modenmultiplexen-Systeme in frühen Studien Energieeinsparpotenziale im Vergleich zu parallelen Singlemode-Systemen. Der Umgang mit mehreren Modi hat jedoch bisher das Auftreten des unerwünschten Effekts eines modenabhängigen Verlusts (MDL) gezeigt. Im Extremfall können Verluste selektiv bestimmte Modi beeinflussen und damit einen Systemausfall verursachen. Aktuelle Ansätze zur Reduzierung des MDL konzentrieren sich entweder auf die Verbesserung des Designs einzelner Inline-Komponenten oder auf das Ausbalancieren der modalen Verstärkungsgrade und Verluste über die gesamte optische Verbindung. Obwohl beide Ansätze zur Verbesserung der Systemleistung beitragen, verbleibt eine erhebliche Menge an MDL im System. Diese Arbeit beschreibt anhand von Simulationen und Experimenten einen innovativen Ansatz zur weiteren Verbesserung der Systemtoleranz gegenüber MDL. Der Ansatz kombiniert eine Verbesserung im optischen Bereich und die Verwendung alternativer DSP-Schemen. Zuerst wird der Einfluss der Modenkopplung auf MDL mit Simulationen untersucht. Bei der Übertragung mit drei und sechs räumlichen Moden wird gezeigt, dass das Induzieren einer starken Modenkopplung entlang der Faser oder die Verwendung von Moden-Scramblern dazu beiträgt, die Toleranz gegenüber MDL zu erhöhen. Außerdem wird die Verwendung alternativer Empfänger-DSP-Schemen wie *near maximum-likelihood* (ML) Detektion und ihre Kombination mit verschieden starker Modenkopplung bewertet. Es wird gezeigt, dass stark gekoppelte Modi und ML-Techniken die Leistung von MDL-beeinträchtigten Systemen erheblich verbessern. Hier wird auch die Komplexität solcher Algorithmen berechnet und analysiert. Schließlich werden Experimente mit Fasern beschrieben, die sechs räumliche Moden beim einfachen und mehrfachen Wellenlängenmultiplexen untersützen. Die Experimente zeigen deutlich eine Leistungsverschlechterung durch MDL, die durch einen *near*-ML Detektor verbessert wird.

Contents

1. Introduction

Communication systems based on fiber optics plays a fundamental role in today's transport network. Already in 1950s optical waveguides started to be considered an alternative medium to replace coaxial cables and microwave systems used mainly for worldwide telephone networks [1]. However, in order to use optical waveguides, transmission impairments from the optical medium needed to be overcome. Over the years, research has gradually overcome obstacles to reach today's vast and complex optical network. In 1960 the laser was invented, which signified the origin of a suitable optical source [1]. At that time, parallel investigations in optical waveguide manufacturing delivered a fiber loss of 1000 dB/km. It was only until 1979 that the scientific community proofed a fiber attenuation close to today's values of 0.2 dB/km [1].

Another breakthrough in the field was constituted by the erbium-doped fiber amplifiers (EDFA), which were developed to overcome the power-budget limitations and became a mature technology in 1994 [2]. Thanks to this invention, from this point on long-haul optical communication was feasible. EDFAs enabled the implementation of the first transatlantic commercial optical link in 1995, named as TAT-12/13 cable network, where TAT stands for Trans-Atlantic Telephone [2,3]. In the following year the first transpacific optical link was completed [2, 4].

In 1996 the commercial breakthrough of the wavelength-division multiplexing (WDM) occurs [2]. This technology consists of simultaneously transmitting information through several channels at different frequencies, analogous to the well known frequency-division multiplexing (FDM) in wireless communications. Before, for enabling WDM systems there was a need of placing a repeater for every channel, which is very costly. The invention of the flat gain EDFA made WDM a cost-effective deployable technology [2, 5].

Transmission over fiber optics has become a mature technology, which facilitated communication featuring low-loss, low-latency and high bandwidth [6,7]. Nowadays, fiber optics networks are organized as a hierarchy in three categories depending on the transmission reach: long-haul, metro and access networks [8–10]. Fig. 1.1 illustrates the optical network structure with their corresponding transmission reaches. Long-haul networks, also called backbone networks, are in charge of the submarine or terrestrial communication between countries and continents reaching distances longer than 1000 km [10]. Networks carrying the data traffic over 3000 km are called ultra long-haul. For distances between 100 km and 1000 km the metro network is employed interconnecting the long-haul and access networks [10]. It aggregates data to the long-haul network or demultiplexes the data for distribution [8]. In the metro network layer, there are more exchange points of information and, therefore, the main challenge in such networks is to design a cost-effective switching architecture [10]. Access or 'last mile' networks provide connectivity to an extensive number of costumers comprising distances smaller than 100 km [9, 10]. Recently, fiber-to-the-x connections are becoming popular enabling higher data rates, this includes all possible 'x', e. g., fiber-to-the-curb, fiber-to-the-node, or fiber-to-the-home [8].

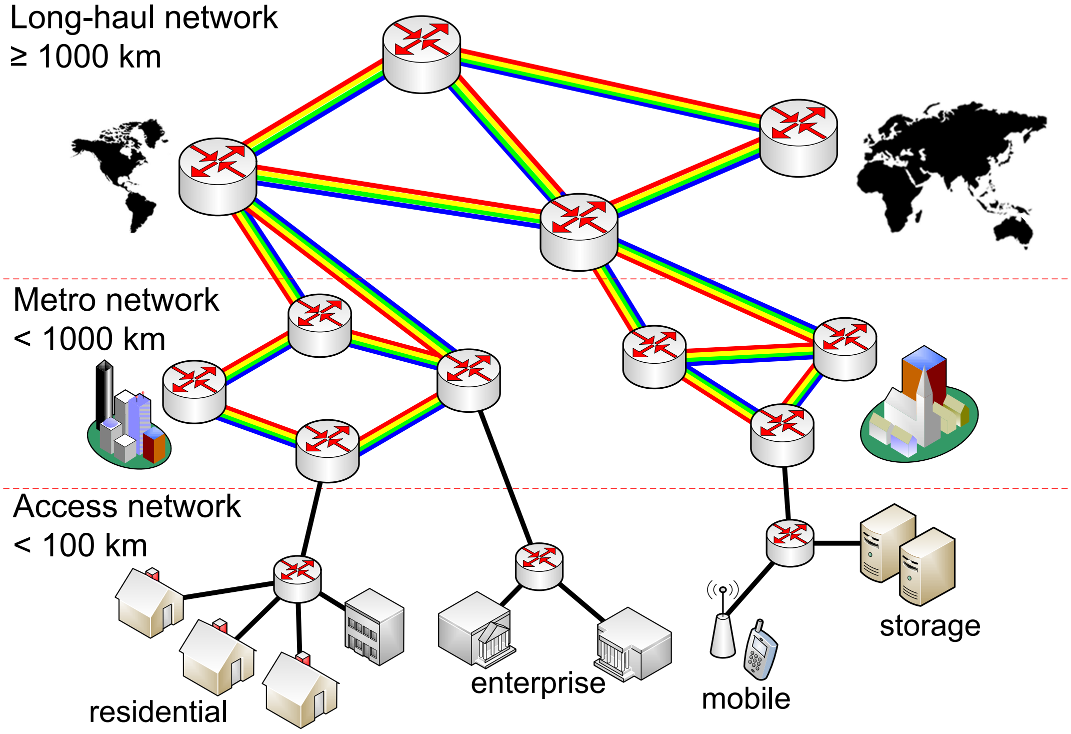


Figure 1.1.: Optical network hierarchy [5, 10].

Nowadays, fiber optic communication carry over 137 Tb/s of data rate [11]. Figure 1.2 suggests that this value is about to increase in the next years. The reasons for this phenomenon are diverse: bandwidth-intensive applications, such as Netflix [12], Youtube [13], Facebook [14], the introduction of new markets from emerging nations, the growing number of Internet-enabled devices, increasing volumes of web-based storage and cloud computing, amongst others [11]. Figure 1.2 also shows that the annual growth rate has been decreasing over the years. According to the Internet traffic report in Fig. 1.2 the growth has slowed down due to the tendency of different applications to store the content close to the end-users causing little impact on the international data traffic.

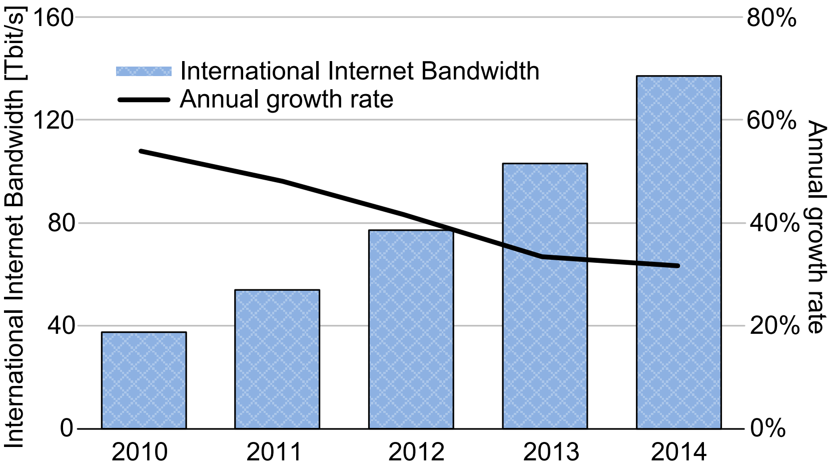


Figure 1.2.: International Internet bandwidth data rate and annual growth rate [11].

This dampening effect on long-haul data rates is also present in research. Since the deployment of WDM there has been the need of occupying the approximately 5 THz of bandwidth covered by the flat gain of the optical amplifiers with the highest possible spectral efficiency (SE), defined as the ratio between the information rate and the used spectral bandwidth [6]. To increase the SE in single-mode fibers (SMFs), which is the current medium of transmission for long-distance optical communications, information is transmitted in two orthogonal polarizations, high-level modulation formats and advanced digital signal processing (DSP) techniques are being investigated. However, the higher the SE, the lower is the achievable transmission reach. On the one hand, optical fiber systems are limited by the Gaussian noise introduced by optical amplifiers. On the other hand, the channel is limited by the nonlinearities generated from high optical launch powers, as shown in [6]. In [6, 15] the authors predict that in the near future the capacity of long-haul communications reaches its fundamental limit leading to the so-called *capacity crunch*. Even recent experiments are already close to this limit by a factor of two [16].

Spatial-division multiplexing (SDM) has emerged as a possible solution to overcome the *capacity crunch*. In the past years, all other physical dimensions available in the optical channel have been studied [17]. By transmitting on two orthogonal polarization, this dimension allows for increasing the transmission capacity by a factor of two compared to single-polarization transmission. The time dimension is being exploited not only by transmitting symbols one after another at a certain symbol rate, but also by advanced techniques of pulse shaping that allow for an efficient bandwidth utilization [18]. The carrier of the transmitted symbols can be modulated in phase and amplitude, which give rise to high-level quadrature-amplitude modulation formats and therefore, exploiting the amplitude and phase dimension. The frequency dimension is being employed for transmission since the 1990s with the WDM technology. SDM aims to exploit the dimension of space and consists of achieving a linear increase in the capacity by transmitting parallel paths of information through fiber bundles, multiple cores or multiple modes [6, 16, 17, 19].

Transmitting information in multiple modes, referred as mode-division multiplexing (MDM), is particularly interesting since it differentiates from other solutions by a possible integration of inline optical components. Developing optical components that operate simultaneously on all modes shows an advantage with respect to savings of cost per bit [20, 21]. Additionally, the system tolerance to nonlinear impairments in the fiber increases since the fiber utilized for MDM systems has a higher core diameter in order to support several modes [22, 23].

However, for developing deployable MDM systems there are several challenges that need to be addressed. Although until now MDM-based experiments have achieved transmission over only few modes, namely over few-mode fiber (FMF), the different modes propagate at quite different speeds and interact exchanging power with each other. These phenomena, namely the differential mode delay (DMD) and mode coupling, cause longer channel impulse responses than conventional single-mode systems. This has a negative impact on the receiver (RX) DSP complexity for recovering the transmitted data streams.

Additionally, there is mode-dependent loss or gain, commonly referred as loss, (MDL). MDL originates from all kind of inline optical components as the signal experiences different losses on every mode. It is a performance-limiting factor, that in extreme cases might lead to systems outage or the partial loss of the information carried over one or several modes [19]. Amplifiers, mode (de)multiplexers and the fiber itself are some of the elements that introduce MDL to the system.

This thesis describes the study of different conditions present in the optical link and the RX DSP that contributes to enhance the reach of FMF transmission systems in the presence of MDL. Both simulations and experiments were carried out in order to observe the different factors that increase the transmission reach.

1.1. Structure of this dissertation

This thesis is structured in seven chapters. Chapters 2 and 3 compile the necessary theoretical concepts to develop this thesis. Chapter 2 describes the properties of the transmission channel, which is the optical fiber. The third chapter aims to explain the different parts of optical systems: transmitter (TX), RX and inline optical components excluding the fiber.

Chapter 4 presents the investigation on the impact of mode coupling on the system tolerance towards MDL. Here, mode coupling is studied and classified mainly in two categories: distributed and discrete mode coupling. The former referring to the one, which is inherently produced constantly in the fiber itself; the latter, to the one introduced in specific parts of the link.

Chapter 5 studies different RX DSP configurations that contribute to an enhanced system performance, i.e. an increase in the transmission reach, in the presence of MDL. Apart from their performance, its computational effort is computed and compared. Here, trade-offs are analyzed.

Chapter 6 shows experimental results on a system supporting six spatial modes. The system performance is compared with different equalization and detection schemes. Single and multiple WDM channels are transmitted. In Chapter 7 the conclusions are drawn and possible future research is discussed.

2. Few-mode fiber optical channel

The evolution of the transmission losses in fiber optics from 1000 dB/km in the 1960s to current loss coefficients lower than 0.2 dB/km and their posterior worldwide deployment replacing copper coaxial cable for terrestrial communications has demonstrated that optical fibers are a suitable medium for long-distance transmission.

This chapter is devoted to the description of the optical medium properties of few-mode fiber systems. A review of different phenomena existing in the fiber is presented at this point in order to understand the dynamic of signal propagation and mode multiplexing. The fundamentals described here are the basis for the next chapters explaining the link components in Chapter 3 and the results in Chapters 4 to 6. First, the principles of light propagation inside a fiber are explained. Afterward, starting off from the Maxwell's equations, the origin of the modes in the fiber is explained and exemplary field distributions with their corresponding phase and intensity profile are shown. Next, the linear and nonlinear mechanisms of distortions are described. The fiber impairments description is considered fundamental for the fiber modeling used in the simulations (Chapters 4 and 5) and for understanding the measurement setup (Chapter 6) presented in this thesis.

2.1. Propagation principle in optical fibers

Fibers are cylindrical dielectric waveguides composed of the low loss material silica (SiO_2) glass, which are structured as illustrated in Fig. 2.1. They are mainly composed of a core, which is surrounded by a cladding and protective coatings. The primary coating depicted in Fig. 2.1 is already applied at the manufacturing process of the fiber (commonly referred as drawing) and its functionality is to protect the fiber from damage during manufacturing and installation process. Furthermore, the primary coating provides stability over a temperature range and in moist environments. Secondary coatings can also be employed to protect the fiber from corrosion and damage, when fibers are exposed to natural and man-made hostilities (e.g. wind, animals, torsion, earthquakes) [24].

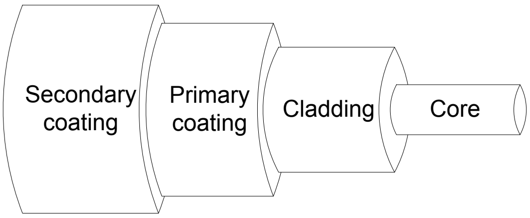


Figure 2.1.: Optical fiber structure.

In order to guide the light rays inside the fiber, the core has slightly higher refractive index than the cladding. The desired refractive index profile is realized by using dopants in different regions of the fiber, as Germanium-Oxide (GeO_2) for increasing the refractive index in the core or, fluoride (F) or boron (B) for decreasing the refractive index in the cladding [25]. Because of this refractive index difference the light remains confined in the core.

Generally, the fiber refractive index profile can be classified in two categories: step-index and graded-index profile. In the former category, the refractive index takes a constant value in the core region. Whereas the graded-index corresponds to a refractive index, which monotonically decreases from the core center towards the cladding. The refractive index profile $n(r)$ in both cases can be characterized with Eq. (2.1), where r is the radial coordinate, n_{co}, n_{clad}, α, and Δn, are the core refractive index at the center of the fiber, the cladding refractive index, the core refractive index decay factor and the relative refractive index profile (see Eq. (2.2)), respectively [1]. The radial coordinate is specified as in Fig. 2.4, where $r = 0$ corresponds to the center of the fiber and $r = a$ to the radius of the fiber core.

$$n(r) = \begin{cases} n_{\text{co}} \left[1 - \Delta n \left(\frac{r}{a}\right)^{\alpha}\right], & \text{for } r < a \\ n_{\text{clad}}, & \text{for } r \geq a \end{cases} \tag{2.1}$$

$$\Delta n = \frac{n_{\text{co}} - n_{\text{clad}}}{n_{\text{co}}} \tag{2.2}$$

If $\alpha = \infty$ the refractive index profile has an abrupt change at the core-cladding interface, thus, the refractive index corresponds to a step-index profile. Otherwise the refractive index is graded. When $\alpha = 2$ the refractive index is said to be parabolic. This is illustrated in Fig. 2.2 for different values of decay factor α.

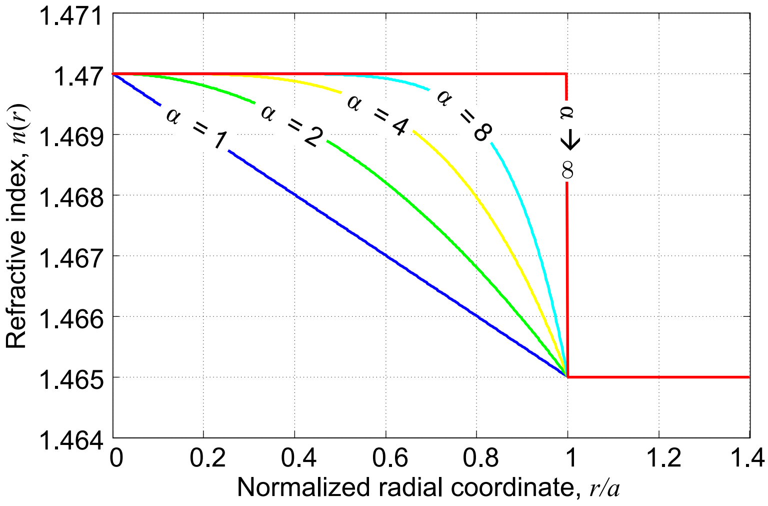


Figure 2.2.: Refractive index profile for different values of α with $n_{\mathrm{co}} = 1.47$ and $n_{\mathrm{clad}} = 1.465$.

2.1.1. Comparison between step-index and graded-index profile

According to Snell's law if a ray of light passes through two isotropic media of refractive indexes n_1 and n_2, the light is refracted following the relation

$$n_1 \sin \theta_1 = n_2 \sin \theta_2, \tag{2.3}$$

where θ_1 and θ_2 are the angles of incidence in medium one and two, respectively. Given that $n_1 > n_2$, if the angle of incidence θ_1 increases, θ_2 also increases, but there is a maximum angle that θ_2 can take, which is $\pi/2$. If $\theta_2 = \pi/2$, θ_1 in Eq. (2.3) is called critical angle θ_{c}. For $\theta_1 > \theta_{\mathrm{c}}$, Eq. (2.3) does not hold anymore, thus, the incident ray is reflected into the originating medium with high efficiency (approximately 99.9 % [26]) and *total internal reflection* takes place. These phenomena is illustrated in Fig. 2.3.

The principle of propagation in step-index fibers relies on the *total internal reflection* of the rays at the interface between the core and the cladding. As shown in Fig. 2.4, the light is confined in the cladding if the angle θ_1 is greater than the critical angle θ_{c}. For angles lower than θ_{c} the light is refracted to the cladding, therefore, there is loss. In order to fulfill the condition for *total internal reflection*, there is a maximum angle $\theta_{0_{\max}}$ that the fiber can accept for guiding light into the core. Associated to the angle $\theta_{0_{\max}}$ is the *numerical aperture* (NA), which is defined in Eq. (2.4). The

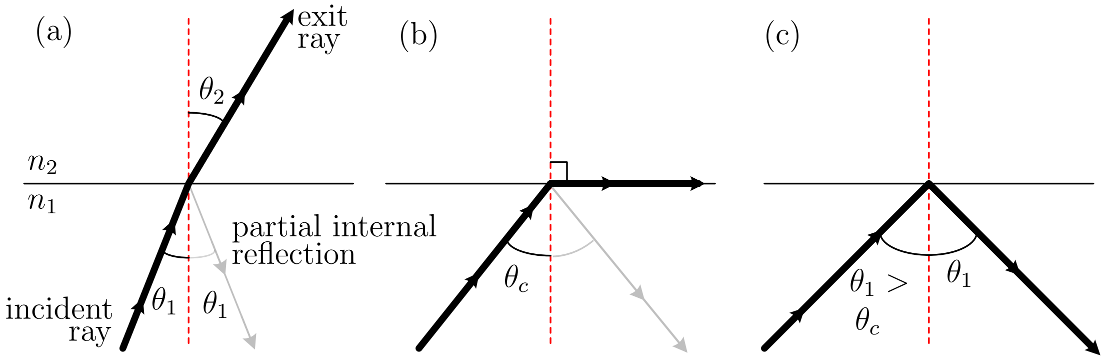


Figure 2.3.: Principle of reflection and refraction [26]. Three cases are illustrated: (a) incident light is reflected and refracted ($\theta_1 < \theta_\mathrm{c}$), (b) refraction at critical angle ($\theta_1 = \theta_\mathrm{c}$) and (c) total internal reflection ($\theta_1 > \theta_\mathrm{c}$).

derivation of the right part of Eq. (2.4) can be found in Appendix A.1. *NA* is a dimensionless number that represents the light acceptance or gathering capability of the fiber.

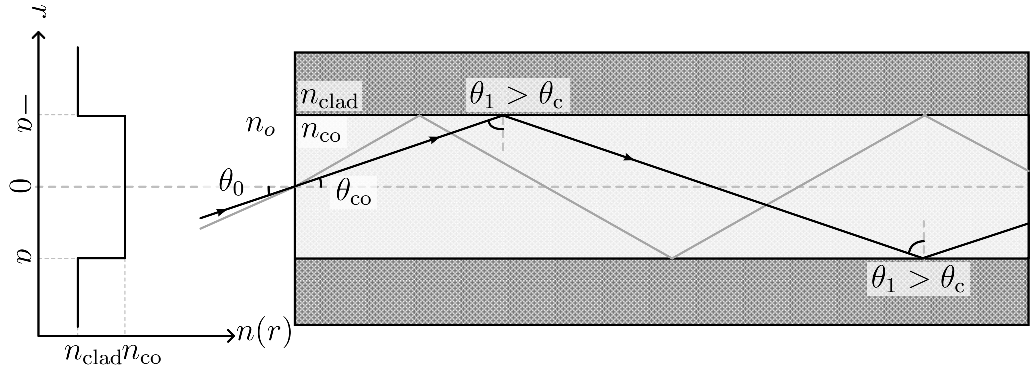


Figure 2.4.: Total internal reflection in a step-index fiber.

$$NA = n_0 \sin \theta_{0_\mathrm{max}} = \sqrt{n_\mathrm{co} - n_\mathrm{clad}} \tag{2.4}$$

In graded-index fibers, *total internal reflection* also takes place, but not necessarily at the core-cladding interface. The ray propagation in graded-index fibers can be schematically depicted as in Fig. 2.5. The refractive index profile of the fiber in Fig. 2.5 has been sliced in several steps, in order to show how the ray is refracted in such fibers. Note that as the ray propagates, the refracted angle constantly increases, as indicated in Fig. 2.5 with the exemplary angles θ_4 and θ_5, until the incident ray in the next refractive index region is totally reflected.

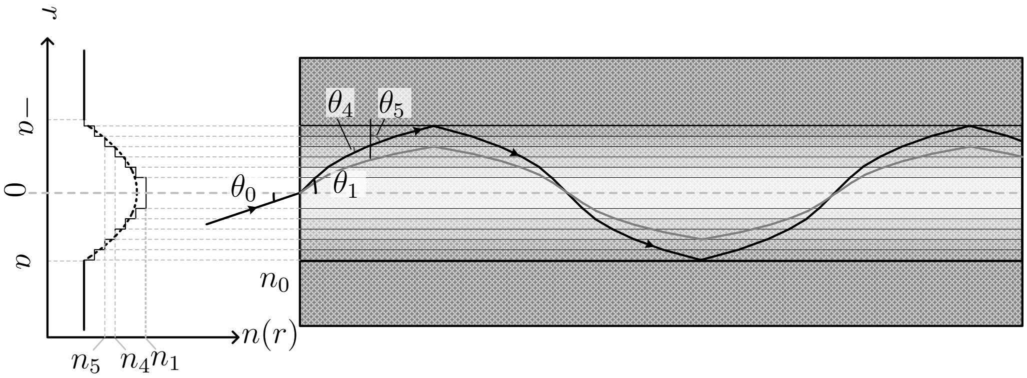


Figure 2.5.: Light propagation in graded-index fibers.

When there is more than one ray propagating along the fiber, multipath dispersion takes place, as a consequence of rays arriving at the end of the fiber at different times. In the step-index fiber case the rays have the same velocity of propagation or group velocity $v_{\mathrm{gr}} = c/n_{\mathrm{co}}$, where c is the speed of light in the vacuum ($c \cong 2.998 \times 10^8$ m/s), since the refractive index is constant. In the case of graded-index fibers the velocity of propagation changes depending on the region where the ray is propagating. Rays traveling along the fiber axis experience lower propagation velocity, but propagate through a shorter path. The opposite happens with those oblique rays that travel a longer path close to the cladding, however, they propagate faster. Therefore, by designing a suitable refractive index profile and adjusting its curvature, the delay between the rays can be reduced.

Consequently, step-index fibers are generally used for single-mode long-haul transmission, since their flat refractive index profile in the core is relatively simple to fabricate, and graded-index fibers are preferred when the fiber allows more than one ray to be guided, namely multimode. Currently, multimode graded-index fiber are used for high speed data transferring between computers and Ethernet applications [1]. Furthermore, graded-index multimode fibers are also preferred for MDM, where the data from several beams are multiplexed and a RX is able to demultiplex the transmitted information. For MDM a low modal delay is beneficial, since it is directly related with the complexity of the RX digital signal processing. Accordingly, the equations that described the modal propagation assume that there is a radial dependency of the core refractive index.

2.2. Modal propagation

As any electromagnetic phenomenon, the wave propagation inside optical fibers is governed by the well known Maxwell's equations. Modes in a fiber are solutions of the Maxwell's equations constrained to a set of assumptions and boundary conditions given by a certain refractive index profile. The set of Eqs. (2.5a) to (2.5d) are the Maxwell's equations, where **E**, **H**, **D** and **B** are the electric field vector in [V/m], magnetic field vector in [A/m], electric displacement vector in [C/m^2] and magnetic induction vector in [Wb/m^2], respectively. Additionally, in Eqs. (2.5b) and (2.5c), **J** and ρ are the electric current density vector in [A/m^2] and electric charge density [C/m^3], respectively.

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \tag{2.5a}$$

$$\nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J} \tag{2.5b}$$

$$\nabla \cdot \mathbf{D} = \rho \tag{2.5c}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.5d}$$

As optical fibers are made of a dielectric material, no electric current can flow (**J** is zero) and there are no free charges (ρ is zero). Thus, making use of the relations in Eqs. (2.6a) and (2.6b), Eqs. (2.5a) and (2.5b) can be transformed into Eqs. (2.7a) and (2.7b), where ϵ, $\epsilon_0 = 8.854 \times 10^{-12}$ C^2/Nm^2 and $\mu_0 = 4\pi \times 10^{-7}$ Ns^2/C^2 are the dielectric permittivity of the medium, permittivity and permeability of a vacuum, respectively.

$$\mathbf{B} = \mu_0 \mathbf{H} \tag{2.6a}$$

$$\mathbf{D} = \epsilon \mathbf{E} = \epsilon_0 n^2(r) \mathbf{E} \tag{2.6b}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \tag{2.7a}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} = \epsilon_0 n^2(r) \frac{\partial \mathbf{E}}{\partial t} \tag{2.7b}$$

As shown in the previous equations, the refractive index profile $n(r)$ is assumed to be longitudinal invariant, therefore independent from the z coordinate. The refractive index profile $n(r)$ is assumed to be symmetric with respect to the core axis, therefore having no azimuthal dependency. Thus, by using the cylindrical coordinate system, where $\mathbf{r}$ corresponds to the radial, azimuthal and longitudinal components (r, φ, z), the refractive index is reduced to $n(\mathbf{r}) = n(r)$.

In order to have separated expressions for the electric and magnetic field, Eqs. (2.7a) and (2.7b) are further treated. Applying the curl $\nabla\times$ at the right and left sides of Eq. (2.7a), Eq. (2.8a) is obtained and further extended in Eqs. (2.8b) and (2.8c) by using Eq. (2.7b) and the identity $\nabla \times (\nabla \times \mathbf{E}) = \nabla (\nabla \cdot \mathbf{E}) - \nabla^2\mathbf{E}$.

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla \times \left(-\mu_0 \frac{\partial \mathbf{H}}{\partial t}\right) \tag{2.8a}$$

$$\nabla (\nabla \cdot \mathbf{E}) - \nabla^2\mathbf{E} = -\mu_0 \frac{\partial}{\partial t} \nabla \times \mathbf{H} \tag{2.8b}$$

$$= -\mu_0 \epsilon_0 n^2(r) \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{2.8c}$$

As for the left side of Eqs. (2.8b) and (2.8c), an equivalent expression for $\nabla \cdot \mathbf{E}$ is derived in Eqs. (2.9a) and (2.9b) by employing Eqs. (2.5c) and (2.6b).

$$0 = \nabla \cdot \mathbf{D} = \epsilon_0 \nabla \cdot \left(n^2(r)\mathbf{E}\right) = \epsilon_0 \left(\nabla n^2(r) \cdot \mathbf{E} + n^2(r) \nabla \cdot \mathbf{E}\right) \tag{2.9a}$$

$$\nabla \cdot \mathbf{E} = \frac{-\nabla n^2(r) \cdot \mathbf{E}}{n^2(r)} \tag{2.9b}$$

The resulting differential equation for the electric field is obtained by replacing the derivation of Eq. (2.9b) in Eq. (2.8c), yielding to

$$\nabla^2\mathbf{E} - \nabla \left(\frac{-\nabla n^2(r) \cdot \mathbf{E}}{n^2(r)}\right) - \mu_0 \epsilon_0 n^2(r) \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0 \tag{2.10}$$

The treatment of the magnetic field is similar to the one developed for the electric field. As shown with Eq. (2.11a), the starting point is Eq. (2.7b) to which the curl of both sides is applied. Then, in Eqs. (2.11b) to (2.11d), the curl properties, Eqs. (2.7a) and (2.7b) are applied.

$$\begin{aligned}
\nabla \times (\nabla \times \mathbf{H}) &= \nabla \times \left(\epsilon_0 n^2(r) \frac{\partial \mathbf{E}}{\partial t}\right) && \text{(2.11a)}\\
\nabla(\nabla \cdot \mathbf{H}) - \nabla^2 \mathbf{H} &= \epsilon_0 \frac{\partial}{\partial t}\left[\nabla \times n^2(r)\mathbf{E}\right] && \text{(2.11b)}\\
&= \epsilon_0 \frac{\partial}{\partial t}\left[\nabla n^2(r) \times \mathbf{E} + n^2(r)\left(\nabla \times \mathbf{E}\right)\right] && \text{(2.11c)}\\
&= \epsilon_0 \nabla n^2(r) \times \frac{\partial \mathbf{E}}{\partial t} + \epsilon_0 n^2(r) \frac{\partial}{\partial t}\left(\nabla \times \mathbf{E}\right) && \text{(2.11d)}\\
&= \epsilon_0 \nabla n^2(r) \times \frac{\partial \mathbf{E}}{\partial t} - \mu_0 \epsilon_0 n^2(r) \frac{\partial^2 \mathbf{H}}{\partial t^2} && \text{(2.11e)}\\
&= \frac{\nabla n^2(r)}{n^2(r)} \times (\nabla \times \mathbf{H}) - \mu_0 \epsilon_0 n^2(r) \frac{\partial^2 \mathbf{H}}{\partial t^2} && \text{(2.11f)}
\end{aligned}$$

In the same way as for the electric field, the term $\nabla \cdot \mathbf{H}$ at the left side of Eqs. (2.11b) to (2.11f) is expanded as presented in Eq. (2.12) by employing Eqs. (2.5d) and (2.6a). Eq. (2.12) evidences that the term $\nabla \cdot \mathbf{H}$ is null.

$$0 = \nabla \cdot \mathbf{B} = \nabla \cdot (\mu_0 \mathbf{H}) = \mu_0 \nabla \cdot \mathbf{H} \tag{2.12}$$

Therefore, reorganizing Eq. (2.11f) and taking into account that the divergence of $\mathbf{H}$ is null, the following differential equation is obtained for the magnetic field

$$\nabla^2 \mathbf{H} + \frac{\nabla n^2(r)}{n^2(r)} \times (\nabla \times \mathbf{H}) - \mu_0 \epsilon_0 n^2(r) \frac{\partial^2 \mathbf{H}}{\partial t^2} = 0 \tag{2.13}$$

For both electric and magnetic field in Eqs. (2.10) and (2.13), the second term involving $\frac{\nabla n^2(r)}{n^2(r)}$ can be neglected because of the weakly guiding approximation, which states that for practical fibers the core and cladding refractive index are very similar, i.e. Δn is much smaller than one [27], yielding

$$\begin{aligned}
\nabla^2 \mathbf{H} - \mu_0 \epsilon_0 n^2(r) \frac{\partial^2 \mathbf{H}}{\partial t^2} &= 0 && \text{(2.14a)}\\
\nabla^2 \mathbf{E} - \mu_0 \epsilon_0 n^2(r) \frac{\partial^2 \mathbf{E}}{\partial t^2} &= 0 \ , && \text{(2.14b)}
\end{aligned}$$

for the magnetic and electric field, respectively. With this approximation, Eqs. (2.10) and (2.13) have the same structure and so their solution. It can be identified that the equations are composed of one part dependent on the transverse components and the other part on the time. Additionally, as mentioned before, taking into account that the $n(r)$ is longitudinal invariant the electric or magnetic field can be expressed as the product of a transverse component $\psi(r,\varphi)$ and a propagation phasor $e^{j(\beta z-\omega t)}$ as

$$\Psi(r,\varphi,z,t) = \psi(r,\varphi)e^{j(\beta z-\omega t)} \quad , \tag{2.15}$$

where $\omega = 2\pi f = 2\pi c/\lambda$, β, f and λ are the angular frequency in [rad/s], propagation constant in [rad/m], frequency in [Hz] and wavelength in [m], respectively.

By replacing Eq. (2.15) into Eqs. (2.14a) and (2.14b), taking into account that the speed of light c can be expressed as $1/\sqrt{\mu_0\epsilon_0}$ and the Laplacian operator ∇^2 is decomposed into its transversal and longitudinal part as in Eq. (2.16), Eqs. (2.17a) and (2.17b) are obtained for the electromagnetic field where $\kappa = \omega/c = 2\pi/\lambda$ is the wavenumber in [rad/m].

$$\nabla^2 = \nabla_t^2 + \frac{\partial^2}{\partial z^2} \quad , \quad \nabla_t^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2} \tag{2.16}$$

$$\left[\nabla_t^2 + \kappa^2 n^2(r) - \beta^2\right]\psi(r,\varphi) = 0 \tag{2.17a}$$

$$\frac{\partial^2\psi(r,\varphi)}{\partial r^2} + \frac{1}{r}\frac{\partial\psi(r,\varphi)}{\partial r} + \frac{1}{r^2}\frac{\partial^2\psi(r,\varphi)}{\partial\varphi^2} + \left[\kappa^2 n^2(r) - \beta^2\right]\psi(r,\varphi) = 0 \tag{2.17b}$$

Equation 2.17b can be solved by the method of variable separation taking advantage of the fiber azimuthal symmetry [28,29]. Accordingly, the solution corresponds to the multiplication of a radial-dependent function with an azimuthal-dependent function: $\psi(r,\varphi) = R(r)\Phi(\varphi)$. Using this product and rearranging Eq. (2.17b), Eq. (2.18) is obtained, where ℓ is a constant.

$$\frac{r^2}{R}\left(\frac{\partial^2 R}{\partial r^2} + \frac{1}{r}\frac{\partial R}{\partial r}\right) + r^2\left(kn^2(r) - \beta^2\right) = -\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\varphi^2} = \ell^2 \tag{2.18}$$

In order to guarantee that the azimuthal part Φ has angular consistency, i.e. $\Phi(\varphi + 2\pi) = \Phi(\varphi)$, the constant l has to be an integer or zero. Hence, the harmonic functions $\cos l\varphi$, $\sin l\varphi$ or the complex representation $e^{jl\varphi}$ represent possible solutions for the azimuthal part, Φ. In the case $l = 0$, there is no azimuthal dependency, thus, the electromagnetic field is two-fold degenerate accounting for two states of polarization. If $l \geq 1$, the electromagnetic field has azimuthal dependency becoming four-fold degenerate: two polarizations and two degenerate spatial modes (one solution involving $\cos l\varphi$ and another involving $\sin l\varphi$). The spatial degenerate modes have the same field distribution with the only difference that they are rotated from each other $\pi/(2l)$, as the relation between the $\cos l\varphi$ and $\sin l\varphi$ solutions suggest.

The solution for the radial part of ψ, R, are Bessel functions in the case of step-index fibers and, in the general case, a linear combination of first-kind and second-kind Bessel functions, which are characterized by a radial mode number m. However, the radial part of the solution has to satisfy also the azimuthal part. Therefore, there is a set of $\{l, m\}$ mode pairs that make up the solution as $\psi_{l,m}(r,\varphi) = R_{l,m}(r)\Phi_l(\varphi)$. The set of modes can be summarized in Eq. (2.19), where the first term corresponds to the so-called discrete, guided or bounded modes, and the second term, to the continuum, radiation or unbounded modes. The latter category corresponds to modes that are not confined in the fiber core, i.e. they can propagate in the cladding. Nevertheless, radiation modes are strongly attenuated by the fiber coating. In Eq. (2.19), the constants c_{lm} and $c(\beta_\mathrm{r})$ are proportional the the power carried by the corresponding mode and, ψ_{β_r} and β_r are the transverse field component and the propagation constant of the radiation modes, respectively.

$$\Psi_{lm}(r,\varphi,z) = \sum_{lm} c_{lm}\psi_{lm}(r,\varphi)e^{-j\beta_{lm}z} + \int_{-\infty}^{\kappa n_\mathrm{clad}} c(\beta_\mathrm{r})\psi_{\beta_\mathrm{r}}(r,\varphi)e^{-j\beta_\mathrm{r}z}d\beta_\mathrm{r} \tag{2.19}$$

Note that each one of the guided modes is characterized by field distribution and a related propagation constant β_{lm}, which satisfy the relation Eq. (2.20a) or, in terms of the effective refractive index $\bar{n}_{lm} = \beta_{lm}/\kappa$, Eq. (2.20b).

$$\kappa n_\mathrm{co} > \beta_{lm} > \kappa n_\mathrm{clad} \tag{2.20a}$$

$$n_\mathrm{co} > \bar{n}_{lm} > n_\mathrm{clad} \tag{2.20b}$$

Guided modes also satisfy the orthonormality condition Eq. (2.21) [28], where * is the complex conjugate operator, $\delta_{ll'}$ and $\delta_{mm'}$ are Kronecker delta functions, i.e. its value is one for $l = l'$ or $m = m'$, otherwise is zero. The orthogonality of the modes is an important property and is the principle of signal multiplexing and demultiplexing for MDM as explained in Chapter 3.

$$\int_0^{\infty}\int_0^{2\pi}\psi^*_{lm}(r,\varphi)\psi_{l'm'}(r,\varphi)rdrd\varphi=\delta_{ll'}\delta_{mm'} \tag{2.21}$$

As mentioned before, the previous derivation is based on the the weakly guiding approximation first introduced by Gloge [27] with the assumption that $\Delta n << 1$. If this was not the case, the solution of the scalar wave equations Eqs. (2.10) and (2.13) would yield the true or exact mode solutions TE_{0m} (transverse electric), TM_{0m} (transverse magnetic), HE_{lm} (hybrid electric) and EH_{lm} (hybrid magnetic). The TE_{0m} and TM_{0m} modes have a null component along the longitudinal axis for the electric and magnetic field, respectively. In contrast, the hybrid modes have a non-zero longitudinal component and are classified in EH_{lm} or HE_{lm} depending weather the electric or magnetic field is dominant, respectively. The weakly guided approximation enables to group the exact modes with virtually equal propagation constants in approximate linear combinations to form the linearly polarized modes, LP_{lm}. Table 2.1 shows which true modes are superimposed to form each LP_{lm} mode and the degeneracy accounting for polarization and spatial modes.

Table 2.1.: Exact-mode content in LP modes and their degeneracy.

LP modes	Exact modes	Degeneracy (Number of spatial and polarization modes)
LP_{0m}	HE_{1m}	2
LP_{1m}	HE_{2m}, TM_{0m}, TE_{0m}	4
LP_{lm} $(l > 1)$	$\mathrm{HE}_{(l+1)m}$, $\mathrm{EH}_{(l+1)m}$	4

As an example, a graded-index fiber with $n_{\mathrm{co}} = 1.46$, $\frac{n_{\mathrm{co}}^2-n_{\mathrm{clad}}^2}{2n_{\mathrm{co}}^2} = 0.01$, at $\lambda = 1550$ nm is considered. In [28] an approximation of the propagation constants β_{lm} and the radial field distributions $R_{lm}(r)$ are given. These relations are used to depict Fig. 2.6, where V is a dimensionless number called normalized frequency or V number. V relates the wavelength, core radius and the core and cladding refractive indices as in Eq. (2.22). The normalized frequency is an indicator of the number of guided modes supported by the fiber. Figure 2.6 indicates the so-called cutoff normalized frequency V_{c} for single mode operation. In general, V_{c} is the normalized frequency up to which the propagation of certain number of modes can be guaranteed. Furthermore, it can be seen from Eq. (2.22) and Fig. 2.6 that the guidance of multiple modes is directly related to the core radius: the greater the core radius, the more modes that can be guided. Note that Fig. 2.6 graphically shows that the effective refractive indices $\bar{n}_{lm}$ of the guided modes are bounded by the core and cladding refractive indices, as also shown in Eq. (2.20b).

$$V = a\kappa\sqrt{n_{\text{co}}^2 - n_{\text{clad}}^2} = \frac{2\pi a}{\lambda}\sqrt{n_{\text{co}}^2 - n_{\text{clad}}^2} \tag{2.22}$$

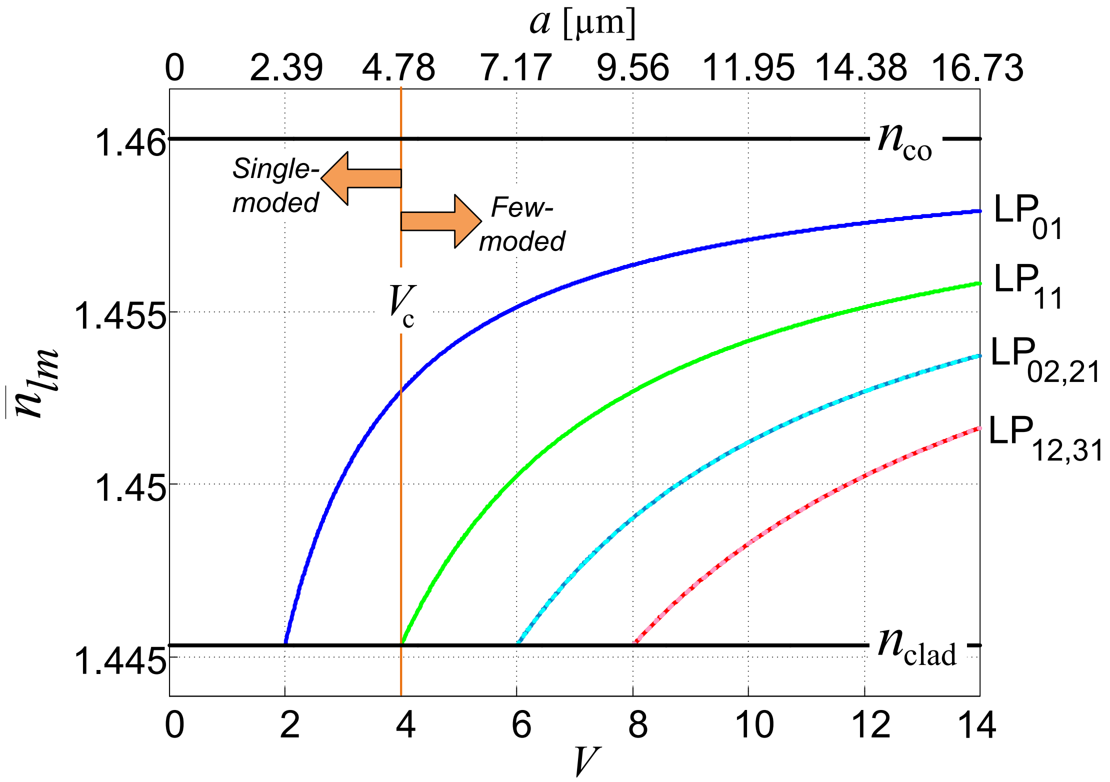


Figure 2.6.: $\bar{n}_{lm}$ vs. V for a graded-index fiber.

The field distributions of the graded-index fiber are also illustrated in Fig. 2.7 for the first four mode groups, also following [28]. Note that l is related to two properties of the field distributions. First, there are $2l$ maxima around the circumference of the fiber core; and second, there is a $\pi/(2l)$ rotation between degenerate modes. The constant m corresponds to the maxima along the radial axis. Although Figs. 2.6 and 2.7 represent the field characteristics for a graded-index fiber, the distributions are very similar to other profile variations. Therefore, these exemplary field properties can be considered as typical.

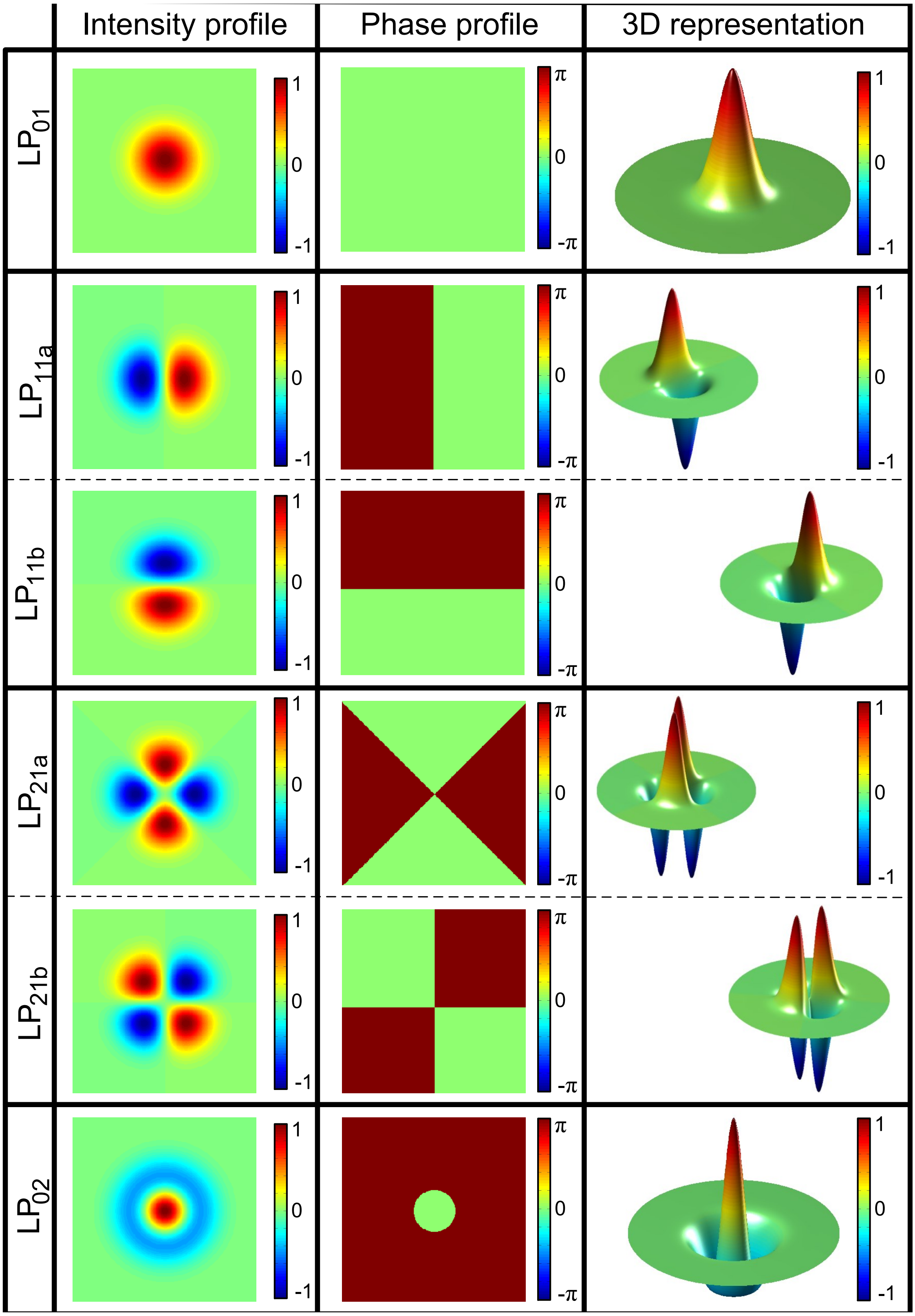


Figure 2.7.: Intensity, phase profile and 3D representation of the first six spatial modes for a graded-index fiber.

2.3. Fiber impairments

This section reviews the different effects that causes distortion during pulse propagation including attenuation, different kinds of dispersion, propagation delay between modes, mode coupling and fiber nonlinearities.

2.3.1. Fiber losses

Any physical medium, except for vacuum, has power losses depending on the material they are constituted of. In the case of optical fibers, the power after certain distance can be described as an exponential decrease along the distance, as defined in Eq. (2.23), where the P_{in}, α_{p}, L and P_{out}, are the fiber input power in [W], attenuation coefficient in Neper per kilometer [Np/km], the transmission distance in [km] and the fiber output power after the distance L [W], respectively [30].

$$P_{\mathrm{out}} = P_{\mathrm{in}} e^{-\alpha_{\mathrm{p}} L} \tag{2.23}$$

It is also possible, and more usual, to describe the fiber attenuation coefficient in [dB/km] as in Eq. (2.24), where the relation between the coefficient in [Np/km] and [dB/km] is given. These loss relations are for both single-mode and multimode fibers, for the total power or for each one of the modes when there is no power interchange between the modes, namely mode coupling (see Section 2.3.4).

$$\alpha_{\mathrm{dB}} = -\frac{10}{L}\log_{10}\left(\frac{P_{\mathrm{out}}}{P_{\mathrm{in}}}\right) = -\frac{10}{L}\log_{10}(e^{-\alpha_{\mathrm{p}} L}) = 10\alpha_{\mathrm{p}}\log_{10}(e) \approx 4.343\alpha_{\mathrm{p}} \tag{2.24}$$

The attenuation in optical fibers is frequency dependent and its origin can be classified in three categories: intrinsic, extrinsic and bending losses. The first category, intrinsic losses, originates from the silica glass itself, whose molecules interacts with the light yielding absorption in the ultraviolet and infrared region. In the ultravilolet region, close to 0.1 μm, there is an electronic resonance, and in the infrared region molecule oscillation takes place, centered between 7 and 9 μm, both with tails extending to the visible region of the spectrum [30].

The second category, extrinsic losses, originates from imperfections in the fabrication process. As during the manufacturing process the silica glass is exposed to thermal diffusion, it can be contaminated with water vapor, i.e. OH-ions. The OH-ions have a vibrational resonance centered between 2.7 μm and 3 μm, although is far from the wavelength range of interest, they have harmonics at 1.383, 1.24, and 0.95 μm. With the current technology the tones produced by these impurities can be lowered, which has given origin to the so-called *dry fiber* [24]. On the other hand, also during the melting process of the glass, silica molecules move randomly and take some place during solidification. Consequently, fluctuations in the density of the silica molecules and of the refractive index can take place. Due to this random variations of the material density, scattering of some of the propagation light occurs, giving rise to the so-called *Rayleigh scattering*.

The last category, bending or radiation losses, are determined after installation and can be divided in micro- and macrobending losses. Microbending results from irregularities in the core-cladding interface, such as diameter fluctuations and axis misalignment. These are also called waveguide imperfections. Furthermore, it can also be originated from external mechanisms such as pressure, twist and tension [31]. Macrobending is caused by bending the fiber on a larger scale. Normally, for *total internal reflection* the rays have an angle of incidence, which is greater than the critical angle. However, bending the fiber might cause that the incident angle becomes smaller than the critical angle, thus, causing some power to be refracted into the cladding.

The different contributions to the fiber loss and its frequency dependance can be observed in Fig. 2.8. Note that the optical frequency is related to wavelength according to $f = c/\lambda$. Figure 2.8 also shows the wavelength bands for transmission over fiber optics according to the International Telecommunication Union (ITU) [24]. There are six transmission bands ranging from 1260 nm to 1675 nm: original- (O), extended- (E), short wavelength- (S), conventional- (C), long wavelength- (L) and ultralong wavelength-band (U). For long-haul communications the C-band, sometimes extended to L-band, are generally used since those frequency bands have the lowest loss.

2.3.2. Dispersion

This section reviews the effect of fiber dispersion, which plays a very important role in the study of optical transmission systems. In general, dispersion is defined as the frequency dependance of a quantity [33]. In this case, the dispersive medium is the fiber and the frequency dependent quantity is the propagation constant β_{lm}.

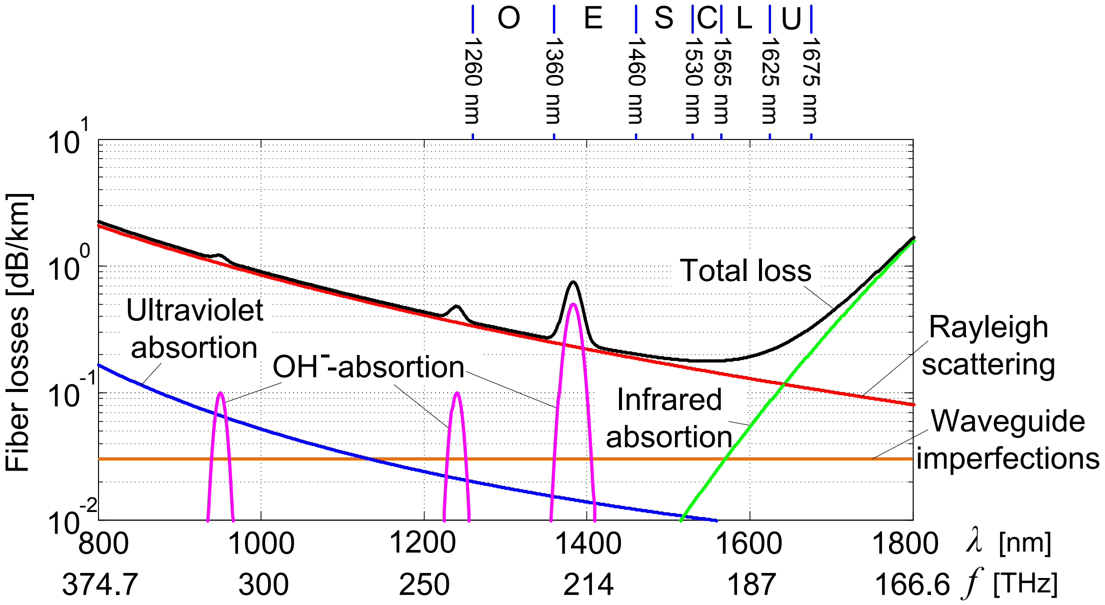


Figure 2.8.: Fiber losses as a function of the frequency [32].

Dispersion can be classified in intra-modal and inter-modal dispersion, when it arises within a mode or relates more than one mode, respectively. Only for single-moded fibers no inter-modal dispersion takes place. In the following, these two varieties of dispersion are described.

Intra-modal dispersion

Consider the transmit electrical field $\mathbf{E}_{lm}$ for a specific mode LP$_{lm}$, which is modulated by the data signal $s_{lm}(z,t)$. In the frequency domain, the modulated data signal corresponds to $S_{lm}(z,\omega-\omega_{\mathrm{c}})$, where ω_{c} is the carrier angular frequency. The angular frequency ω_{c} can be also written as $\omega_{\mathrm{c}} = 2\pi f_{\mathrm{c}}$, where f_{c} is the carrier frequency. The propagating field can be written as Eq. (2.25a), where $\beta_{0_{lm}}$ is the propagation constant at the carrier frequency and $\hat{E}_{lm}$ is the transverse component of the electrical field [30, 34]. A single spectral component of $S_{lm}(z,\omega)$ propagating through the fiber can be expressed in terms of the input signal $S_{lm}(z=0,\omega)$, the corresponding propagation constant β_{lm} and the term representing the attenuation in the fiber $e^{-\frac{\alpha_{\mathrm{p}}}{2}z}$ as in Eq. (2.26) [34]. Based on Eq. (2.26), Eq. (2.25a) can be expressed as Eq. (2.25b).

$$\mathbf{E}_{lm}(\mathbf{r}, z, \omega) = \hat{E}_{lm}(\mathbf{r}) S_{lm}(z, \omega - \omega_\mathrm{c}) e^{j\beta_{0_{lm}} z} \tag{2.25a}$$

$$\mathbf{E}_{lm}(\mathbf{r}, z, \omega) = \hat{E}_{lm}(\mathbf{r}) S_{lm}(z = 0, \omega - \omega_\mathrm{c}) e^{j\beta_{lm} z} e^{-\frac{\alpha_\mathrm{p}}{2} z} \tag{2.25b}$$

$$S_{lm}(z, \omega) = S_{lm}(z = 0, \omega) e^{-j\left(\beta_{0_{lm}} - \beta_{lm}\right) z} e^{-\frac{\alpha_\mathrm{p}}{2} z} \tag{2.26}$$

As the spectral occupancy of $S_{lm}(z, \omega)$ is much smaller than the carrier frequency, β_{lm} can be expanded in a Taylor series as

$$\beta_{lm}(\omega) = \beta_{0,lm}(\omega_\mathrm{c}) + (\omega - \omega_\mathrm{c}) \beta_{1,lm}(\omega_\mathrm{c}) + \frac{1}{2} (\omega - \omega_\mathrm{c})^2 \beta_{2,lm}(\omega_\mathrm{c}) + \frac{1}{6} (\omega - \omega_\mathrm{c})^3 \beta_{3,lm}(\omega_\mathrm{c}) + \ldots\ , \tag{2.27}$$

where

$$\beta_{g,lm}(\omega_\mathrm{c}) = \left(\frac{d^g \beta_{lm}(\omega)}{d\omega^g} \right)_{\omega = \omega_\mathrm{c}} \quad \text{for } g = 1, 2, 3, \ldots \tag{2.28}$$

The source of intra-modal dispersion is this frequency dependance of β_{lm}, as observed in the previous equations. Usually, it is sufficient to represent $\beta_{lm}(\omega)$ up to the third-order term in the Taylor series, higher order terms can be neglected.

Each one of the terms in Eq. (2.27) is associated to a quantity having a physical meaning. The first term, $\beta_{0,lm}(\omega_\mathrm{c})$, corresponds to a constant phase shift, which is therefore irrelevant for the study of pulse distortion. The second term, proportional to $\beta_{1,lm}(\omega_\mathrm{c})$, is related to the time delay of the signal envelope, i.e. to the group velocity $v_{\mathrm{gr},lm}$. This implies that $v_{\mathrm{gr},lm}$ is frequency-dependent. The group velocity is defined as the inverse of $\beta_{1,lm}(\omega_\mathrm{c})$, as shown in Eq. (2.29).

$$v_{\mathrm{gr},lm} = \left(d\beta_{lm} / d\omega \right)^{-1} \tag{2.29}$$

As a consequence of $v_{\mathrm{gr},lm}$ being frequency-dependent, after propagation different frequency components of the signal arrive at the end of the fiber with different delays. This phenomenon is known as chromatic dispersion, group velocity dispersion (GVD) or simply dispersion. The latter term is adopted in this thesis. The dispersion is responsible for the pulse broadening and, consequently, inter-symbol interference (ISI) between neighboring symbols.

Furthermore, the change of $\beta_{lm}(\omega)$ with the frequency evidences that the effective refractive index $\bar{n}_{lm} = \beta_{lm}/\kappa$ is also frequency-dependent. By using this relation and expressing κ in terms of ω and c, Eq. (2.29) can be further expanded as shown in Eqs. (2.30a) and (2.30b). The quantity $\bar{n}_{\mathrm{gr},lm}$ is the group effective refractive index for the mode LP_{lm}, which is the sum of $\bar{n}_{lm}$ and a term related to the slope of $\bar{n}_{lm}$ with respect to the frequency, as indicated in Eq. (2.30b).

$$v_{\mathrm{gr},lm} = \left(\frac{d\bar{n}_{lm}\kappa}{d\omega}\right)^{-1} = c\left(\frac{d\bar{n}_{lm}\omega}{d\omega}\right)^{-1} = c\left(\bar{n}_{lm} + \omega\frac{d\bar{n}_{lm}}{d\omega}\right)^{-1} \tag{2.30a}$$

$$v_{\mathrm{gr},lm} = \frac{c}{\bar{n}_{\mathrm{gr},lm}} \tag{2.30b}$$

The third term in Eq. (2.27), the second order term proportional to $\beta_{2,lm}(\omega)$, is related to the rate of change of the group velocity $v_{\mathrm{gr},lm}$ with the frequency. Here, it is useful to introduce the dispersion parameter $D_{lm}(\lambda)$ defined for a particular mode LP_{lm}, usually in $[\mathrm{ps}\cdot\mathrm{nm}^{-1}\cdot\mathrm{km}^{-1}]$. The dispersion parameter is an important quantity to calculate the pulse broadening and precises the rate of change of the group velocity with the frequency. Consider a signal that propagates over a fiber of length L in [km] with a spectral width of $\Delta\lambda$ in [nm], the pulse broadening ΔT produced by the dispersion is defined as in Eq. (2.31).

$$\Delta T = LD_{lm}(\lambda)\Delta\lambda = LD_{lm}(\lambda)\frac{-\lambda^2}{2\pi c}\Delta\omega \tag{2.31}$$

The dispersion parameter is defined as

$$D_{lm}(\lambda) = \frac{d}{d\lambda}\left(\frac{1}{v_{\mathrm{gr},lm}}\right) = \frac{d\omega}{d\lambda}\frac{d}{d\omega}\left(\frac{1}{v_{\mathrm{gr},lm}}\right) = \frac{-2\pi c}{\lambda^2}\frac{d}{d\omega}\left(\frac{1}{v_{\mathrm{gr},lm}}\right) = \frac{-2\pi c}{\lambda^2}\beta_{2,lm}\ . \tag{2.32}$$

Equation 2.32 can be further expanded by replacing $v_{\mathrm{gr},lm}$ from Eq. (2.30a) in terms of the effective refracted index as

$$D_{lm}(\lambda) = \frac{-2\pi c}{\lambda^2}\frac{d}{d\omega}\left(\bar{n}_{lm} + \omega\frac{d\bar{n}_{lm}}{d\omega}\right) = \frac{-2\pi}{\lambda^2}\left(2\frac{d\bar{n}_{lm}}{d\omega} + \omega\frac{d^2\bar{n}_{lm}}{d\omega^2}\right)\ . \tag{2.33}$$

Zero dispersion is possible at a certain wavelength. However, even if $D_{lm}(\lambda)$ is equal to zero, the wavelength dependance of $D_{lm}(\lambda)$ leads to dispersive effects in the transmission over the optical fiber. These dispersive effects are contributed by the higher-order terms in Eq. (2.27). Usually the terms up to the third order in the Taylor expansion are considered. Third-order dispersive effects are characterized by the dispersion slope $S_{\mathrm{lo},lm}(\lambda)$, which describes the change of $D_{lm}(\lambda)$ with the wavelength as evident in its definition in Eq. (2.34). The units of $S_{\mathrm{lo},lm}(\lambda)$ are $[\mathrm{ps}\cdot\mathrm{nm}^{-2}\cdot\mathrm{km}^{-1}]$.

$$S_{\mathrm{lo},lm}(\lambda) = \frac{dD_{lm}(\lambda)}{d\lambda} = \frac{d\omega}{d\lambda}\frac{dD_{lm}(\lambda)}{d\omega} = \frac{-2\pi c}{\lambda^2}\frac{dD_{lm}(\lambda)}{d\omega} = \frac{-\omega^2}{2\pi c}\frac{dD_{lm}(\lambda)}{d\omega} \tag{2.34}$$

As $\beta_{2,lm}$ is the second derivative of β_{lm} with respect to the angular frequency ω, it is convenient to expand Eq. (2.34) in terms of ω as in Eqs. (2.35a) to (2.35d). By using the definition of $D_{lm}(\lambda)$ in Eq. (2.32), Eq. (2.34) can be further developed in Eqs. (2.35a) to (2.35d).

$$S_{\mathrm{lo},lm}(\lambda) = \frac{-\omega^2}{2\pi c}\frac{d}{d\omega}D_{lm}(\lambda) = \frac{-\omega^2}{2\pi c}\frac{d}{d\omega}\left(\frac{-2\pi c}{\lambda^2}\beta_{2,lm}\right) \tag{2.35a}$$

$$S_{\mathrm{lo},lm}(\lambda) = \frac{-\omega^2}{2\pi c}\frac{d}{d\omega}\left(\frac{-\omega^2}{2\pi c}\beta_{2,lm}\right) = \frac{\omega^2}{(2\pi c)^2}\frac{d}{d\omega}\left(\omega^2\beta_{2,lm}\right) \tag{2.35b}$$

$$S_{\mathrm{lo},lm}(\lambda) = \frac{\omega^2}{(2\pi c)^2}\left(2\omega\beta_{2,lm} + \omega^2\beta_{3,lm}\right) = \frac{2\omega^3}{(2\pi c)^2}\beta_{2,lm} + \frac{\omega^4}{(2\pi c)^2}\beta_{3,lm} \tag{2.35c}$$

$$S_{\mathrm{lo},lm}(\lambda) = \frac{4\pi c}{\lambda^3}\beta_{2,lm} + \left(\frac{2\pi c}{\lambda^2}\right)^2\beta_{3,lm} \tag{2.35d}$$

Dispersion is the result of the contribution of material and waveguide dispersion depending on its origin. Material dispersion is the responsible for the spectral dependance of the refractive index. The refractive index in the frequency domain is defined with the following relation [1]

$$n(\omega) = \left(1 + Re\left\{\chi^{(1)}(\mathbf{r},\omega)\right\}\right)^{1/2} , \tag{2.36}$$

where the $\chi^{(1)}(\mathbf{r},\omega)$ is the first order susceptibility of the medium. The susceptibility characterizes the interaction of the dielectric medium, silica in the case of fiber optics, with the incident electromagnetic field, which transfers some of its energy to the bound electrons in the medium. As indicated in Eq. (2.36) by the frequency dependance of $\chi(\mathbf{r}, \omega)$, this interaction is different for different frequency components, which makes the value of the refractive index frequency-dependent. At certain frequencies the material absorbs the electromagnetic radiation, leading to the appearance of K discrete resonances at $\omega_1, \omega_2, ...\omega_K$. For fused silica glass, the Sellmeier equation in Eq. (2.37) gives a good approximation of this behavior, where B_i are weighting factors obtained from fitting this relation to measurements on the medium [35].

$$n(\omega) = \left(1 + \sum_{i=1}^{K} \frac{B_i}{1 - (\omega/\omega_i)^2}\right)^{1/2} \tag{2.37}$$

Whereas the material dispersion is attributed to the properties of the material itself, the waveguide dispersion is attributed to the shape of the waveguide. By modifying the fiber characteristics, e.g. its core radius, difference between the core and cladding refractive index or the shape of the refractive index profile, the waveguide dispersion can be modified and therefore the total dispersion profile can be tailored for a specific purpose. An example of these specially tailored fibers are the dispersion-shifted fibers (G.653) and dispersion-compensating fibers [1].

Inter-modal dispersion

As explained in the previous section, each mode possesses a different group velocity $v_{\mathrm{gr},lm}$. The inter-modal dispersion, commonly also known as differential mode delay (DMD), results from the fact that pulses from different modes propagate with different group velocities, therefore, they arrive at the end of the fiber with different propagation delays. Thus, DMD exclusively takes place in fiber supporting multiple modes.

The DMD between mode LP_{lm} and $\mathrm{LP}_{l'm'}$ is given by the difference of each mode's delay per unit length, as shown in Eq. (2.38) [36].

$$DMD_{lm,l'm'} = \frac{1}{v_{\mathrm{gr},lm}} - \frac{1}{v_{\mathrm{gr},l'm'}} \tag{2.38}$$

As mentioned in the former section, in the presence of dispersion the group velocity becomes frequency-dependent. Each mode may have a slightly different dispersion, therefore the group velocities of each mode would change by different amounts with the frequency. This implies also that the difference between the group velocities, i.e. the DMD, changes with the frequency.

In [37], it is shown that by means of the proper fiber design the DMD can be virtually flat over the bands of transmission C and L. The authors in [37] used a graded-index fiber with a cladding trench and optimized its refractive index profile in order to minimize the DMD and its variation over frequency. For this reason, the variation of the DMD with the frequency is neglected in the simulations.

The frequency dependency of the modal delay is referred to as higher-order modal delay. For further information on this topic see [38].

2.3.3. Birefringence and polarization-mode dispersion

For an ideal fiber with perfectly circular core and isotropic characteristics, orthogonally polarized modes have the same propagation properties, i.e. same propagation constant. However, this is unlikely to happen. Real fibers have a slightly elliptical core, suffer from all kinds of stresses, asymmetries, anisotropy, and harmful effects from the outside such as twisting and temperature changes. All these effects have an impact on the propagation constant of each polarization, causing the degeneracy of the polarization modes to break, and on the orientation of the polarization axis, causing polarization scrambling and power leakage from one polarization to the other and vice versa. Modal birefringence is the result of these effects and is characterized by the difference between the effective refractive index of each polarization $\Delta\bar{n}_{lm,\mathrm{xy}}$ as shown with different equivalences in Eq. (2.39) [25, 39], where x and y represent two arbitrary transversal axis in the fiber. The axis x is said to be the fast axis if $\bar{n}_{lm,\mathrm{x}} < \bar{n}_{lm,\mathrm{y}}$, since it will have a shorter delay.

$$\Delta\bar{n}_{lm,\mathrm{xy}} = |\bar{n}_{lm,\mathrm{x}} - \bar{n}_{lm,\mathrm{y}}| = \frac{|\beta_{lm,\mathrm{x}} - \beta_{lm,\mathrm{y}}|}{\kappa} = \frac{c\,|\beta_{lm,\mathrm{x}} - \beta_{lm,\mathrm{y}}|}{\omega} = \frac{\lambda\,|\beta_{lm,\mathrm{x}} - \beta_{lm,\mathrm{x}}|}{2\pi} \tag{2.39}$$

Constant birefringence leads to the exchange of energy between the polarizations with a period commonly called beat length L_B equals to $2\pi/\,|\beta_{lm,\mathrm{x}} - \beta_{lm,\mathrm{y}}| = \lambda/\Delta\bar{n}_{lm,\mathrm{xy}}$ [25]. This also implies that after a distance of L_B the input and output state of polarization coincide. The state of polarization is the pattern traced by the relation of the x and y electric fields. Propagate x and y without any phase difference, then there is a linear polarization; if the phase difference is $\pi/2$ then it is called circular polarization and elliptical otherwise [39].

Furthermore, at a constant birefringence the difference in propagation constants generates a difference in the group velocity. This implies that the pulses propagating along each polarizations arrive with different delays at the end facet of the fiber, which is commonly referred as differential group delay (DGD) denoted in Eq. (2.40) as $\Delta\tau$. By using Eq. (2.29), in terms of the first derivative of the propagation constant $\beta_{1,lm,\mathrm{x/y}}$, or Eq. (2.30b), in terms of the group effective refractive index $\bar{n}_{\mathrm{gr},lm,\mathrm{x/y}}$, the delay $\Delta\tau$ can be further developed in Eqs. (2.41a) to (2.41c).

$$\Delta\tau = \left| \frac{L}{v_{\mathrm{gr},lm,\mathrm{x}}} - \frac{L}{v_{\mathrm{gr},lm,\mathrm{y}}} \right| \tag{2.40}$$

$$\begin{aligned}
\Delta\tau &= \frac{L}{c}\left|\bar{n}_{\mathrm{gr},lm,\mathrm{x}} - \bar{n}_{\mathrm{gr},lm,\mathrm{y}}\right| = \frac{L}{c}\left|\Delta\bar{n}_{\mathrm{gr},lm,\mathrm{xy}}\right| && \text{(2.41a)}\\
= L\left|\beta_{1,lm,\mathrm{x}} - \beta_{1,lm,\mathrm{y}}\right| &= L\left|\frac{d\beta_{lm,\mathrm{x}}}{d\omega} - \frac{d\beta_{lm,\mathrm{y}}}{d\omega}\right| = L\left|\frac{d\kappa\bar{n}_{lm,\mathrm{x}}}{d\omega} - \frac{d\kappa\bar{n}_{lm,\mathrm{y}}}{d\omega}\right| && \text{(2.41b)}\\
= L\frac{d\kappa\Delta\bar{n}_{lm,\mathrm{xy}}}{d\omega} &= \frac{L}{c}\frac{d\omega\Delta\bar{n}_{lm,\mathrm{xy}}}{d\omega} = \frac{L}{c}\left(\omega\frac{d\Delta\bar{n}_{lm,\mathrm{xy}}}{d\omega} + \Delta\bar{n}_{lm,\mathrm{xy}}\right) && \text{(2.41c)}
\end{aligned}$$

DGD is graphically represented in Fig. 2.9, where is evident that this effect is one of the sources of distortion that introduces delay spread in the impulse response of the transmission channel.

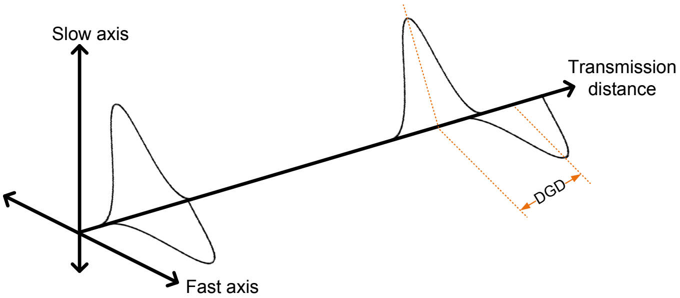


Figure 2.9.: Differential group delay between orthogonal polarizations.

Polarization-mode dispersion (PMD) can be defined as the DGD per unit distance. The PMD is characterized by the PMD coefficient PMD_c, which corresponds to $\Delta\tau/L$. Usually, in fibers of more than several kilometers long, as in the case of this thesis, the PMD_c is not given by a linear relation between time and distance, but it is given in [ps/$\sqrt{\text{km}}$]. The reason for this nonlinear behavior is that the PMD is the product of the concatenation of multiple random events [39, 40]. Each of these random event can be represented as splitting the fiber into several small sections with different but constant DGD values. The DGDs may seem longer or shorter by comparing the different sections, i.e. the DGDs do not sum up linearly.

DGD is the origin of the so-called first-order PMD. The frequency dependance of the states of polarization (depolarization) and the DGD (polarization chromatic dispersion) arises from higher-order PMD [10], which originates from including higher-order terms of the propagation constant derivative and not only $\beta_{1,lm,\mathrm{x/y}}$ in Eq. (2.40). Depolarization and polarization chromatic dispersion can also be described by a Taylor expansion as shown in [10]. Analogous to the previous section on inter-modal dispersion, in the presence of dispersion it may be that the polarizations have slightly different dispersion values leading to a change by different amounts of the group velocities of each polarization.

As there are other effects that are more dominant in few-mode fibers such as DMD and MDL, higher-order PMD is out of the scope of this thesis.

2.3.4. Mode coupling

Just as PMD, mode coupling takes place when there is a geometrical or environmental perturbation in the fiber. These perturbations can be intended or unintented. An example of intentional introduction of mode coupling is the utilization of a mode scrambler as proposed in the late 70's with [41, 42]. Mode scramblers are employed in standard multimode fibers to launch power evenly over all modes, which is useful for making DMD tolerance measurements and ensure the repeatability of the measurements.

The consequence of the perturbations mentioned before is the power transfer between modes [43]. Mode coupling takes place especially, and strongly, between degenerate modes or modes that have similar propagation constants. Marcuse [44] and Olshansky [45] developed the coupled mode theory for optical waveguides with imperfections and perturbations. Both found that the strength of mode coupling is directly related to the difference in the propagation constants $\Delta\beta = \beta_a - \beta_b$, in this case between mode a and b.

In addition, and also similar to PMD, random mode coupling can effectively reduce the impulse response length in a transmission system, namely the DMD [45]. Therefore, the relation impulse response length due to DMD is no longer proportional to the time delay between the slowest and fastest pulse multiplied by the fiber length, but multiplied by the square root of the fiber length. In [45], Olshansky explains this phenomenon in a very intuitive way by regarding the transmission as the propagation of different energy packages. At the end of the transmission, the different energy packages will have traveled through different modes due to mode coupling, and consequently, they arrive with a delay, which corresponds to the average delay over all modes, and whose distribution depends of the strength and nature of the mode coupling [45].

Mainly two kinds of fiber model are used in the literature for describing the mode coupling behavior: power and field coupling models [46]. The former only considers the real-valued exchange of energy between modes and was used in the past to simplify the coupling equations so that various effect of the fiber could be easily studied, such as power distribution and pulse responses [44, 46]. The field coupling model is in contrast a complex-valued electrical field model containing amplitude and phase information. This is therefore the appropriate model for coherent detection and the one used in this thesis.

The modeling of mode coupling and linear impairments was provided by the authors in [22, 47]. It is based on the coupled mode theory presented by Marcuse in [44]. Waveguide imperfections such as bends and surface corrugation can be modeled as a small core displacement from the transversal coordinates. Therefore the fiber is divided into small sections simulating this behavior as illustrated in Fig. 2.10. The core displacement is represented by a small perturbation of the refractive index or permittivity, taking into account that the refractive index relates to the permittivity by $n^2 = \frac{\epsilon\mu}{\epsilon_0\mu_0}$, where μ is the permeability of the medium in [H/m].

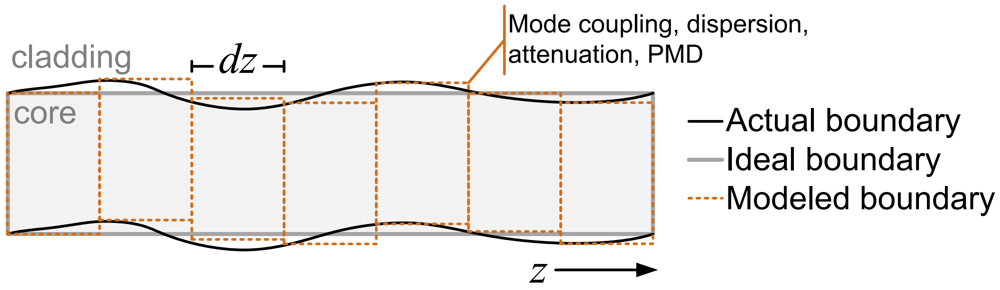


Figure 2.10.: Fiber model illustration.

The permittivity of the perturbed medium is then given by $\epsilon(x,y) = \epsilon_{\mathrm{a}}(x,y) + \Delta\epsilon(x,y)$, where $\epsilon_{\mathrm{a}}(x,y)$ is the permittivity of the unperturbed medium and $\Delta\epsilon(x,y)$ is the perturbation. As mentioned above, the permittivity $\epsilon(x,y)$ can be also represented as a deviation of the permittivity in the core center as $\epsilon\left(x,y\right) = \epsilon_{\mathrm{a}}\left(x+\delta x, y+\delta y\right)$, where δx and δy is the deviation of the abscissa and ordinate coordinates, respectively.

According to the coupled mode theory, the perturbed permittivity is incorporated in the coupling coefficients C_{u,u_1} and the propagation equation Eq. (2.42), where $S_u\left(z,\omega-\omega_{\mathrm{c}}\right)$, C_{u,u_1} and $\Delta\beta_{0_{u,u_1}}$ are the Fourier transform of the complex-valued field envelope, the mode coupling coefficients between mode u and u_1, the difference between the propagation constants $\Delta\beta_{0_{u,u_1}} = \beta_{0_u} - \beta_{0_{u_1}}$, respectively. The first part of the right hand side of Eq. (2.42) corresponds to the dispersion effects, and the second part, to the mode coupling. Only linear impairments were considered in this thesis to concentrate the attention on the effect of MDL in the system assuming the launch power is low enough to neglect the nonlinear impairments. The derivation of Eq. (2.42) has been included in Appendix A.2.

$$\frac{\partial}{\partial z}S_u\left(z,\omega-\omega_{\mathrm{c}}\right) = -j\left(\beta_{0,u}-\beta_u\right)S_u\left(z,\omega-\omega_{\mathrm{c}}\right)+j\sum_{u_1\neq u}C_{u,u_1}S_{u_1}\left(z,\omega-\omega_{\mathrm{c}}\right)e^{-j\Delta\beta_{0_{u,u_1}}z} \tag{2.42}$$

The coupling coefficients between the mode u and u_1 are given by the overlap integral modified by the perturbation $\Delta\epsilon\left(x,y\right)$ as in Eq. (2.43).

$$C_{u,u_1} = \frac{\omega_{\mathrm{c}}}{4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{E}_u^*\left(x,y\right)\Delta\epsilon\left(x,y\right)\hat{E}_{u_1}\left(x,y\right)dxdy \tag{2.43}$$

Solving Eq. (2.42) analytically can be very complicated and numerical methods require powerful computational resources. For this reason, the authors in [47] proposed a semi-analytical solution for Eq. (2.42). The idea behind it is to divide the solution of Eq. (2.42) into two parts where only a dispersive effect or mode coupling is effective for each fiber section, similar to the split-step Fourier method [25]. The sections are small enough so that the continuous wave approximation is valid, i.e. the delay spread caused by modal dispersion should be much smaller than the pulse duration [47]. The solution of the first dispersive part of Eq. (2.42) leads to Eq. (2.26) at the carrier frequency ω_c, which additionally includes the fiber attenuation. Therefore, Eq. (2.26) can be rewritten to include until the term related to the dispersion slope as Eq. (2.44), where $\beta_{1_{lm}}$, $\beta_{2_{lm}}$ and $\beta_{3_{lm}}$ are derived from Eqs. (2.29), (2.32) and (2.35d) in Eq. (2.45). In Eq. (2.45a), apart from the modal group delay represented by $v_{\mathrm{gr},lm}$, the group velocity difference between polarization x and y is included in the term related to $\Delta v_{\mathrm{gr},lm,\mathrm{xy}}$. The value of $\Delta v_{\mathrm{gr},lm,\mathrm{xy}}$ can be derived using Eqs. (2.41a) and (2.41b) resulting in Eq. (2.46).

$$\begin{aligned} S_{lm}(z,\omega-\omega_\mathrm{c}) &= S_{lm}(z=0,\omega-\omega_\mathrm{c})e^{-j\left(\beta_{0_{lm}}-\beta_{lm}\right)z}e^{-\frac{\alpha_\mathrm{p}}{2}z} \\ &= S_{lm}(z=0,\omega-\omega_\mathrm{c})e^{j\left((\omega-\omega_c)\beta_{1,lm}+\frac{1}{2}(\omega-\omega_c)^2\beta_{2,lm}+\frac{1}{6}(\omega-\omega_c)^3\beta_{3,lm}\right)z}e^{-\frac{\alpha_\mathrm{p}}{2}z} \end{aligned} \tag{2.44}$$

$$\beta_{1_{lm}} = \frac{1}{v_{\mathrm{gr},lm} \pm \frac{\Delta v_{\mathrm{gr},lm,\mathrm{xy}}}{2}} \tag{2.45a}$$

$$\beta_{2_{lm}} = -\frac{\lambda^2 D_{lm}}{2\pi c} \tag{2.45b}$$

$$\beta_{3_{lm}} = \left(S_{\mathrm{lo},lm} + \frac{2D_{lm}}{\lambda}\right)\left(\frac{\lambda^2}{2\pi c}\right)^2 \tag{2.45c}$$

$$\Delta v_{\mathrm{gr},lm,\mathrm{xy}} = \frac{\Delta\bar{n}_{\mathrm{gr},lm,\mathrm{xy}}c}{\bar{n}^2_{\mathrm{gr},lm}} \tag{2.46}$$

For computing $\beta_{1_{lm}}$, $\beta_{2_{lm}}$ and $\beta_{3_{lm}}$ and the attenuation factor in Eq. (2.44), the dispersion D_{lm}, dispersion slope $S_{\mathrm{lo},lm}$, attenuation α_p and birefringence $\Delta v_{\mathrm{gr},lm,\mathrm{xy}}$ parameters were set to typical values for doped silica [36]. The group velocities $v_{\mathrm{gr},lm}$ and group effective indexes $\bar{n}_{\mathrm{gr},lm}$ are obtained from optimizing the fiber refractive index profile in order to minimize the DMD as described in [37]. In [37], Maxwell's equations were solved numerically to find the most suitable fiber profile.

The model also considers random orientation of the polarizations from section to section. This can be represented as a concatenation of Jones matrices, i.e. their successive multiplication (see e.g. [48] chapter 7).

Given the coupling coefficients and the propagation constant differences for a DMD-optimized fiber design, the mode coupling part of Eq. (2.42) can be solved. The process of solving this part of the equation can be found in [47].

2.3.5. Nonlinear impairments

Operating the fiber at high power levels gives rise to a nonlinear behavior in this medium. Nonlinearities can be classified in Kerr nonlinearities, or the ones that causes a nonlinear dependance of the refractive index in the fiber, and nonlinear scattering. In the following, these two kinds of nonlinearities are explained.

Kerr effect nonlinearities

The Kerr nonlinearities were named after the Scottish physicist John Kerr (1824-1907), who discovered this effect in 1875 [49]. In Section 2.2, the medium was considered isotropic, linear, nonconducting and nonmagnetic. The fact that the medium is linear is reconsidered here. Dielectrics, including fiber optic, respond in a nonlinear way to intense electromagnetic fields.

For considering the nonlinear terms, Eq. (2.6b) can be rewritten in the frequency domain so that the permittivity of the medium is defined in terms of the product $\epsilon_0\epsilon_\mathrm{r}$, where ϵ_r is the dielectric constant, also called relative permittivity, as in Eq. (2.47a). ϵ_r in turn can be expressed as $1+\chi_\mathrm{e}$, where χ_e is the susceptibility of the medium. In a linear medium, the χ_e is independent of the applied field strength. However, for nonlinear media χ_e is a field-intensity-dependent tensor as shown in Eqs. (2.47a) and (2.47b) [30].

$$\begin{aligned}\mathbf{D}(\mathbf{r},\omega)=\epsilon(\mathbf{r},\omega)\mathbf{E}(\mathbf{r},\omega)=\epsilon_0\epsilon_\mathrm{r}(\mathbf{r},\omega)\mathbf{E}(\mathbf{r},\omega)&=\epsilon_0(1+\chi_\mathrm{e}(\mathbf{E},\omega))\mathbf{E}(\mathbf{r},\omega) && (2.47a)\\ &=\epsilon_0\mathbf{E}(\mathbf{r},\omega)+\epsilon_0\chi_\mathrm{e}(\mathbf{E},\omega)\mathbf{E}(\mathbf{r},\omega) && (2.47b)\\ &=\epsilon_0\mathbf{E}(\mathbf{r},\omega)+\mathbf{P}(\mathbf{r},\omega) && (2.47c)\end{aligned}$$

Equation 2.47c introduces the polarization density vector $\mathbf{P}=\epsilon_0\chi_\mathrm{e}(\mathbf{E},\omega)\mathbf{E}(\mathbf{r},\omega)$, which is the vector representing the induced dipole moments per unit volume. $\mathbf{P}$ is not linear and is often written as a power expansion as shown in Eq. (2.48a) [30], where $\chi^{(n)}$ is the n^{th} order susceptibility, as well as a tensor of rank $n+1$.

$$\begin{aligned}\mathbf{P}(\mathbf{r},\omega)=\epsilon_0\chi_e(\mathbf{E},\omega)\mathbf{E}(\mathbf{r},\omega)&=\epsilon_0\left[\chi^{(1)}\mathbf{E}+\chi^{(2)}\mathbf{E}\mathbf{E}+\chi^{(3)}\mathbf{E}\mathbf{E}\mathbf{E}+\cdots\right] && (2.48a)\\ &=\mathbf{P}_\mathrm{L}(\mathbf{r},\omega)+\mathbf{P}_\mathrm{NL}(\mathbf{r},\omega) && (2.48b)\\ \mathbf{P}_\mathrm{L}(\mathbf{r},\omega)&=\epsilon_0\chi^{(1)}\mathbf{E} && (2.48c)\\ \mathbf{P}_\mathrm{NL}(\mathbf{r},\omega)&=\epsilon_0\left[\chi^{(2)}\mathbf{E}\mathbf{E}+\chi^{(3)}\mathbf{E}\mathbf{E}\mathbf{E}+\cdots\right] && (2.48d)\end{aligned}$$

In Eq. (2.48b)), $\mathbf{P}_\mathrm{L}(\mathbf{r},\omega)$ and $\mathbf{P}_\mathrm{NL}(\mathbf{r},\omega)$ are described as in Eqs. (2.48c) and (2.48d). The term associated with $\chi^{(1)}$, $\mathbf{P}_\mathrm{L}(\mathbf{r},\omega)$, corresponds to the dominant contribution and is the responsible for the linear effects, such as the material dispersion (as shown with Eq. (2.36)) and the attenuation. The higher order terms, described by

$\mathbf{P}_{\mathrm{NL}}(\mathbf{r},\omega)$, contribute to the nonlinear response. However, the second-order term is only relevant for unsymmetrical molecule structures, which is not the case for silica glasses. The first and most relevant nonlinear contribution is given by $\chi^{(3)}$, which is responsible for the Kerr effect. Higher order terms in Eq. (2.48d) can be neglected. Therefore, Eq. (2.48d) can be approximated as in Eq. (2.49).

$$\mathbf{P}_{\mathrm{NL}}(\mathbf{r},\omega) \approx \epsilon_0\chi^{(3)}\mathbf{EEE} \tag{2.49}$$

Note from Eqs. (2.48a) and (2.48b) that $\mathbf{P}_{\mathrm{L}}(\mathbf{r},\omega)$ and $\mathbf{P}_{\mathrm{NL}}(\mathbf{r},\omega)$ have their corresponding counterparts in time domain, $\mathbf{P}_{\mathrm{L}}(\mathbf{r},t)$ and $\mathbf{P}_{\mathrm{NL}}(\mathbf{r},t)$, whose origin is not the multiplication with the electrical field but their convolution.

For describing the effect of these nonlinear terms, the wave equation is now used. By replacing the time domain version of Eqs. (2.47c) and (2.48b) in Eq. (2.14b) and taking into account the weakly guiding approximation ($\nabla\cdot\mathbf{E}=0$), the wave relations in Eq. (2.50) can be obtained.

$$\nabla^2\mathbf{E} - \mu_0\epsilon_0\frac{\partial^2\mathbf{E}}{\partial t^2} - \mu_0\frac{\partial^2\mathbf{P}}{\partial t^2} = 0 \tag{2.50a}$$

$$\nabla^2\mathbf{E} - \mu_0\epsilon_0\frac{\partial^2\mathbf{E}(\mathbf{r},\omega)}{\partial t^2} - \mu_0\frac{\partial^2\mathbf{P}_{\mathrm{L}}}{\partial t^2} - \mu_0\frac{\partial^2\mathbf{P}_{\mathrm{NL}}}{\partial t^2} = 0 \tag{2.50b}$$

As $\mathbf{P}_{\mathrm{L}}$ corresponds to $\epsilon_0\chi^{(1)}\mathbf{E}$, the linear term of the equation can be joined in one single term by introducing the refractive index to Eq. (2.50b), as expressed before in Eqs. (2.10) and (2.36). Thus, the wave equation can be written as

$$\nabla^2\mathbf{E} - \mu_0\epsilon_0 n^2(r)\frac{\partial^2\mathbf{E}}{\partial t^2} - \mu_0\frac{\partial^2\mathbf{P}_{\mathrm{NL}}}{\partial t^2} = 0 \tag{2.51}$$

Solving Eq. (2.51) can be quite complicated, therefore the term corresponding to $\mathbf{P}_{\mathrm{NL}}$ is treated in the literature as a small perturbation, which is appropriate for a weak nonlinear medium such as silica [2, 25]. Thus, this term generates a nonlinear perturbation of the phase of the signal proportional to the nonlinear change of the refractive index $\frac{3}{8n}Re\left\{\chi^{(3)}_{xxxx}\right\}|\mathbf{E}|^2$ [25], where $\chi^{(3)}_{xxxx}$ is one component of the tensor of rank four.

Note that the phase distortion produced by the Kerr effect is proportional to the power of the electrical field. Thus, the modulated signal in the optical field is interfering with itself. Form the point of view of a specific transmitted WDM channel, when the interference is provoked by the power coming from the channel under test, the distortion is called self-phase modulation (SPM); if it comes from the power of other WDM channels, this is called cross-phase modulation (XPM). Four-wave mixing is also another kind of distortion due to the Kerr effect, which describes the interaction of three waves generating a fourth wave.

Nonlinear light scattering

Nonlinear scattering processes arise from the interaction between the incident optical field and the silica molecules such that energy transfer from the incoming light to the dielectric medium takes place. Brillouin and Raman scattering are two kinds of such nonlinear effect.

Both nonlinear processes are generated spontaneously or stimulated by a pump light source upon the use of high power levels. In contrast to linear elastic scattering processes, i.e. Rayleigh scattering, where the scattered wave has the same frequency as the incident wave, these nonlinear processes are inelastic. Namely, the scattered photon from both processes is shifted to a higher (anti-stokes) or a lower (Stokes) frequency. Although anti-stokes processes can occur, it requires previous conditioning of the medium and therefore, is irrelevant for optical communications [50].

In the case of Raman scattering, the incident light interplays with the vibrations of the silica molecules in the fiber [31]. Thus, part of the energy from the incident photon is transferred to a mechanical vibrational state of the silica molecules, namely the optical phonons [50]. This energy dissipation in form of optical phonons corresponds to the energy difference between the incoming and scattered photons (stokes photons). The consequence of this molecule vibration is the energy transfer to a lower frequency: around 100 nm (approximately 13 THz) away from the incident light.

Brillouin scattering originates from the reflections at periodical variations in the density of the silica [31]. The product of this interaction is the energy dissipation in form of acoustical phonons (phonons at sound velocity) and the generation of the stokes photon in the opposite direction as the incident light at lower frequency approximately 0.1 nm$\approx$10 GHz away from the incident photon.

Both Brillouin and Raman scattering can produce gain with an appropriate pump light source. Compared to Raman scattering, Brillouin achieves a higher gain. However, the bandwidth provided by the stimulated Raman scattering is higher than the Brillouin scattering. The gain bandwidth of the Brillouin scattering is of the order of 20 MHz, whereas for Raman is in the order of 20 THz to 30 THz [1, 31]. Another difference between Raman and Brillouin scattering is that stokes photon of the former effect can co- and counter-propagate, whereas the stokes photon due to the latter can only counter-propagate [1].

3. Mode-division multiplexing

The *capacity crunch* requires exploring new transmission schemes and concepts. SDM exploits the spatial dimension by means of simultaneously transmitting several streams of information through the fiber.

The realization of SDM can be done in different ways. The most intuitive way is to send the information through several single-mode fibers or, a more compact variant, through fiber bundles. Other alternatives are depicted in Fig. 3.1, which illustrates the cross-section of different types of fibers with respect to their core position and size. These alternatives can basically be classified in three categories: FMF, multicore fiber (MCF) and the combination of both.

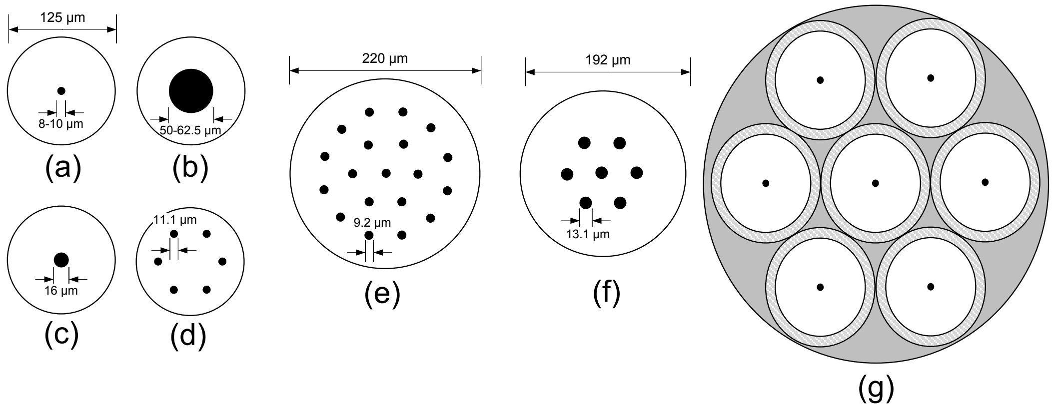


Figure 3.1.: Cross section examples of candidates for SDM. The black dots represent the core and the outer circle the cladding. (a) Standard single-mode and (b) multimode fiber [24], (c) few-mode fiber [51], (d) 6-core fiber [52], (e) 19-core fiber [32], (f) 7-core fiber [53] and (g) fiber bundles.

For comparison purposes, commercial SMFs and multimode fibers (MMFs) are also illustrated in Fig. 3.1(a) and (b). SMFs have substantially lower core diameters than MMFs. Whereas the former type is the standard for today's long-haul and metro networks, the latter type is used for short-reach applications, e.g. in data centers, because of the low cost of the laser technology used for illuminating such fibers and the low loss splices between fiber segments [45, 54]. The main reach limitation of commercial MMFs is the modal delay between hundreds of modes that propagate down this fiber. For such systems the information is launched in form of pulses, whose modal content is a mix of several modes. However, the propagating pulse is affected by the DMD, i.e. the delay among modes, which causes interference between the neighboring pulses and limits the transmission to a few hundreds of meters.

An example of one of the candidates for SDM is shown in Fig. 3.1(c). The figure depicts a fiber that has slightly greater core diameter than SMFs and supports only few modes (FMF). By supporting a low number of modes, the refractive index profile design can be better controlled to achieve a lower DMD. Another type of SDM trend is shown in Fig. 3.1(d), (e) and (f): multicore fibers (MCF). Figure 3.1(e) shows a 19-core fiber, where each one of the cores supports a single mode. In this case, the cores are placed as far as possible to minimize crosstalk between them. In turn, the cores of the fiber depicted in Fig. 3.1(d) are placed closer in order to introduce strong coupling between them. The strongly coupled cores are also called supermodes, since their energy profile resembles the modes in a multimode fiber. Similar to strongly coupled modes, as mentioned in Section 2.3.4, having strong core coupling or supermodes helps to reduce the channel impulse response length and MDL (see Chapter 4). Furthermore, unlike the uncoupled core configuration, coupled cores have the advantage of being unaffected by possible external factors that may cause core coupling, e.g. splicing misalignment. The fiber shown in Fig. 3.1(f) has uncoupled cores with a slightly larger diameter than the other MCFs in Fig. 3.1(d) and (e), which makes it a combination of multicore and multimode fiber supporting three modes in each core.

Fiber bundles are considered also a SDM candidate and are depicted in Fig. 3.1(g). This is the simplest and most straightforward approach to increase the transmission data rates. The fiber bundles option does not have any impact on the system structure, since the same transceivers and optical components can be used. Deploying such systems with N single-mode fibers would increase the capacity N times, but would also increase the operation cost by N times. Therefore, it is necessary to find cost-efficient solutions to the *capacity crunch*.

Multiplexing the signal into different modes or cores constitutes a more attractive alternative. As illustrated in Fig. 3.1, both MCF and FMF achieve more information density per unit area than fiber bundles. Especially FMFs have the potential to offer the advantage of being more energy efficient than MCF systems with respect to line equipment, which is dominated by the energy consumption of optical amplifiers [20,55,56]. There is also a potential advantage in the nonlinear tolerance of few-mode transmission with respect to single-mode. Since nonlinearities in the fiber originates from the confinement of the light in the silica, the slightly higher diameter of FMFs enables better transmission performance in this respect than single moded fibers [22].

However, MDM faces challenges that need to be addressed for launching out commercial systems. Crosstalk between the modes distorts the signal. This fact combined with modal delay affects negatively the RX digital signal processing complexity [57]. Unequal attenuation or amplification of the modes, MDL, is also an issue that needs to be minimized. Quantification of MDL is defined in Section 3.2 and methods to improve the system tolerance against MDL are described in Chapters 4 and 5.

The structure of an MDM system is depicted in Fig. 3.2. The block diagram shows several data streams that are multiplexed into several modes (MDM MUX) and frequency by WDM channels (WDM MUX). In the optical link the signal is transmitted through FMF and is amplified periodically. Reconfigurable optical add/drop multiplexers also take part of the optical link. They are in charge of incorporating or diverting WDM channels from the optical link in order to reach other destinations in the optical networks. After the signal is propagated through the optical link, it is demultiplexed in frequency and space with the WDM and MDM demultiplexer (DMUX), respectively. The depicted WDM MUX and DMUX need to support multiple modes. Theses devices are being developed and investigated in e.g. [21, 58]. At the RX the effects of the channel are unraveled by means of a multiple-input multiple-output (MIMO) RX.

This chapter describes the system configuration taken into account for the analysis of the MDM systems. The first part of the chapter is devoted to the TX and the conversion of the baseband signal into the transmission band in the optical domain. Before going into detail on optical components in the link, the impairment MDL is explained, since MDL corresponds to one of the performance indicators for optical components. Then, the RX is described, which includes a review of the optical to electrical conversion and the DSP of different modulation schemes, and a brief description of forward-error correction (FEC) techniques.

3.1. Optical TX

Nowadays, signals are modulated not only into one but two orthogonal dimensions: quadrature (data Q) and in-phase (data I). The modulated signals are generated digitally and, subsequently, converted into analog signals with a digital-to-analog

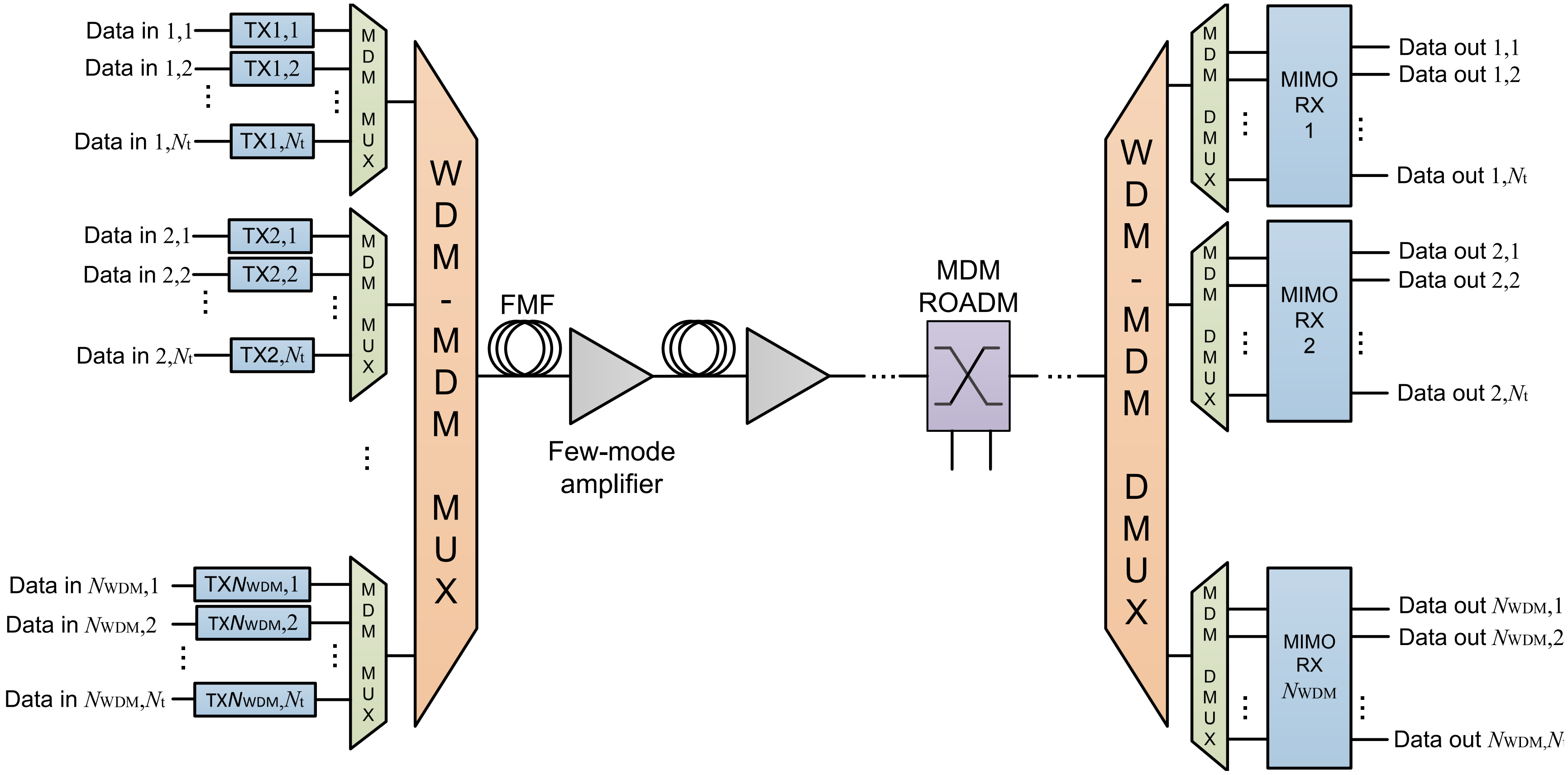


Figure 3.2.: Structure and main components of a system based on MDM, where N_t and N_WDM are the number of tributaries (spatial and polarization modes) and the number of WDM channels, respectively.

converter (DAC). These electrical analog signals are converted into an optical signal by means of an IQ modulator. This component uses a light source (laser) to combine the pair of orthogonal IQ signals into one complex optical signal. Figure 3.3 shows this procedure. In the following, the principle of the optical source and IQ modulator are explained.

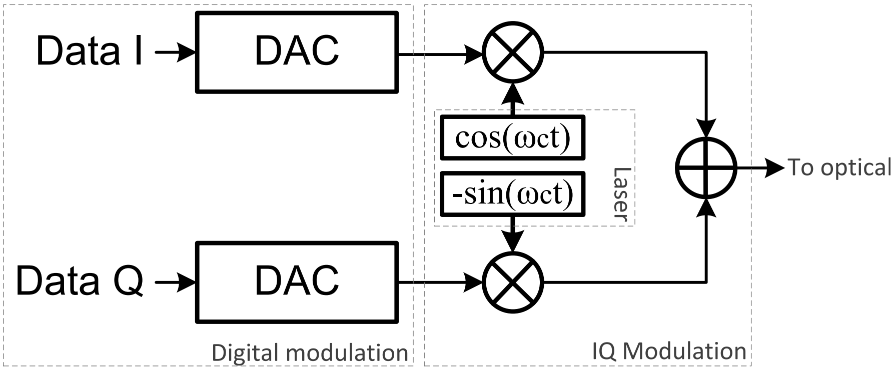


Figure 3.3.: TX basic diagram.

3.1.1. Optical source

Absorption, spontaneous and stimulated emission are the three processes that can occur in an atomic system. These are depicted in Fig. 3.4, which describes a system having two energy levels W_1 (ground state, valence band or equilibrium state) and W_2 (excited state or conduction band) [31]. In Fig. 3.4, the system is radiated by an electrical field of frequency f_{21}.

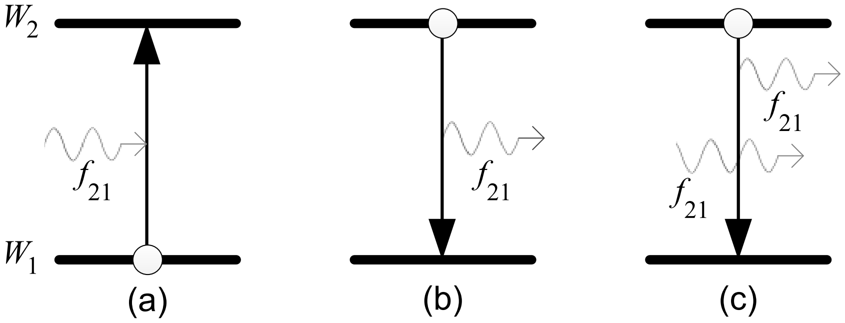


Figure 3.4.: Processes in an atomic system: (a) absorption, (b) spontaneous and (c) stimulated emission.

Absorption takes place when the atom in the ground level goes to the excited level causing the incident light to attenuate. In turn, stimulated emission takes the atom from the excited energy level to the ground level generating a new photon of the same characteristics as the stimulating photon, i.e. same frequency, phase, polarization and direction of propagation. Similar to absorption, stimulated emission is induced by the radiation field. In the case of spontaneous emission, the atom that has been taken to the excited level by absorption only remains for a certain time in this state and returns spontaneously without the need of a radiation field. In contrast with stimulated emission, the generated photons from spontaneous emission have random and uncorrelated phase and polarization, i.e. are incoherent.

As the name suggests, the principle of lasers (Light Amplification by Stimulated Emission of Radiation) is based on the process of stimulated emission. There are other optical sources, such as light emitting diodes (LEDs), which are based on spontaneous emission of light. However, LEDs are not relevant for coherent optical systems since their spectral width is significantly above lasers (at least 50 times broader than lasers corresponding to tens of nanometers of spectral width [31, 59]).

There are several types of lasers such as gas, solid state and semiconductor lasers. Semiconductor lasers are mostly used in long-haul optical fiber communications, since they provide high efficiency and a compact size in the wavelength window of interest for optical long-haul communications. Semiconductor lasers are composed of a p- and n-junction, where electrons and holes recombine as a response to an external stimulus (bias current). The light emission is caused during recombination of holes and electrons by an active material (e.g. $GaAlAs$, $InGaAsP$), which is usually sandwiched between the p- and n-junction [31,60]. Most common used lasers include distributed feedback laser (DFB) and external cavity laser (ECL) with a spectral width in the sub-MHz region [59].

Although the main mechanism of light emission in lasers is the stimulated emission, spontaneous emission is unavoidable in such devices. This is the source of the so-called laser phase and amplitude noise. Laser phase and amplitude noise refers to time-varying phase and amplitude. This phenomenon can be represented in Eq. (3.1) with $\phi_\mathrm{c}(t)$ and $E_\mathrm{c}(t)$, respectively. The carrier frequency, initial phase, time-varying phase, electrical field, and its magnitude are represented in Eq. (3.1) as f_c, ϕ_0, $\phi_\mathrm{c}(t)$, $\mathbf{E}_\mathrm{c}(t)$ and $E_\mathrm{c}(t)$, respectively.

$$\mathbf{E}_\mathrm{c}(t) = E_\mathrm{c}(t)e^{j(\phi_0+\phi_\mathrm{c}(t))}e^{-j2\pi f_c t} \tag{3.1}$$

Phase and amplitude noise is a concatenation of events or a random walk. Figure 3.5 represents the evolution of the laser electrical field, whose phase and amplitude have been distorted by three spontaneous emission events. In Fig. 3.5 the initial phase and magnitude are illustrated as ϕ_0 and E_0, respectively. The subsequent phases and magnitudes correspond to ϕ_i and E_i, where $i \in \mathbb{N}$. The spontaneous emission events are imposed by ϕ_{sp_i} and E_{sp_i}. Subsequently, the resulting electrical field is the sum of the initial state $E_0 e^{j\phi_0} e^{-j2\pi f_\mathrm{c} t}$ and the spontaneous emission events $E_{\mathrm{sp}_i} e^{j\phi_{\mathrm{sp}_i}} e^{-j2\pi f_\mathrm{c} t}$.

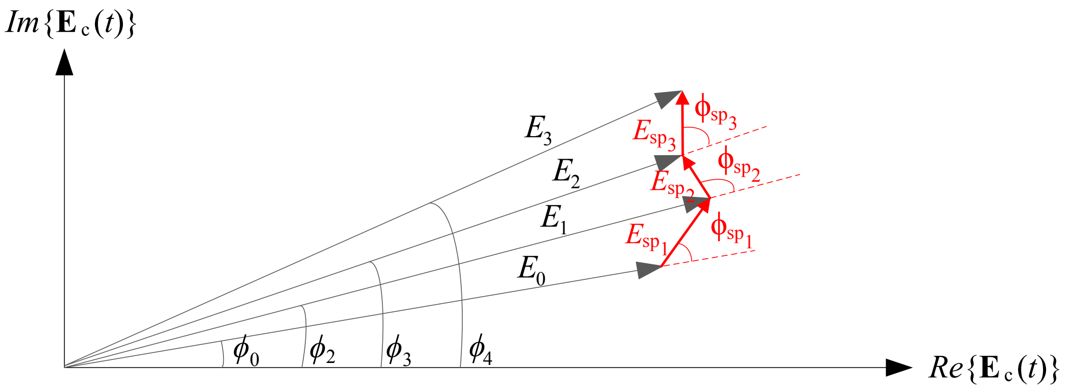


Figure 3.5.: Evolution of phase and amplitude of the laser electrical field disturbed by three spontaneous emission events.

Laser amplitude noise can be found in the laser data sheet as relative intensity noise (RIN). The RIN indicates the intensity fluctuations per unit of bandwidth relative to the average intensity. It is usually given in [Hz^{-1}] or [dBc/Hz]. The intensity noise is a less relevant effect since the RIN typical values are very low, less than -150 dBc/Hz [31].

For modeling laser phase noise it is appropriate to define its statistical properties. It is convenient to characterize the temporal change of phase during a time interval ΔT denoted by $\Delta\phi_\mathrm{c}(t, \Delta T) = \phi_\mathrm{c}(t + \Delta T) - \phi_\mathrm{c}(t)$. In the literature, these phase changes are also referred as random walk [59]. Since the origin of the random walk is numerous independent noise events caused by spontaneous emissions, these phase changes can be characterized by a Gaussian probability density function (PDF).

The variance of the phase changes $\Delta\phi_\mathrm{c}(t, \Delta T)$ can be demonstrated to be directly proportional to time [31]. The longer the observation period the greater the variance. This makes the stochastic process of phase noise non-stationary. The variance corresponds to $2\pi\Delta f\,|\Delta T|$ [31,59], where Δf is the spectral width, i.e. the so-called laser 3-dB linewidth (LW), spectral half-width or LW at full-width-half-maximum [31].

Modeling laser phase noise can be realized by generating independent Gaussian distributed phase noise values with variance $2\pi\Delta f\,|\Delta T|$, and adding up these events successively as $\phi_{\mathrm{c}}([n+1]\Delta T) = \phi_{\mathrm{c}}(n\Delta T) + \Delta\phi_{\mathrm{c}}(n\Delta T, \Delta T)$, where $n \in \mathbb{N}$ represents a discrete instant of time. This results in a power spectral density of the optical source with the Lorentzian line shape [59]

$$S_{\mathrm{laser}}(f) = \frac{2}{\pi\Delta f}\frac{1}{1+\left(\frac{f-f_c}{\Delta f/2}\right)^2}\ . \tag{3.2}$$

3.1.2. IQ Modulator

Signal modulation in optical systems is the conversion of the electrical signal into the optical domain. It can be realized by directly modulated lasers (DMLs), electro-absorption modulators (EAMs) and Mach-Zehnder modulators (MZMs).

Signal modulation with DML is a simple and cost-effective solution, since it is compact and it has no additional components other than the optical source. This modulation technique is as simple as turning on and off the laser according to the electrical drive signal. Nevertheless, modulating the laser directly produces a residual phase modulation or frequency chirp, and consequently, introduces unwanted frequency components to the signal [1, 10]. This is unacceptable for WDM systems. Furthermore, the dynamic response of DMLs shows that they are only suitable for low-bit rate applications up to 2.5 Gb/s [10, 61].

EAMs and MZMs are external modulators created to overcome the drawbacks of the DMLs. The former modulator uses its material to absorb or let pass the incoming light depending on the electrical drive signal. Compared to DMLs, EAMs have less frequency chirp, however, their main disadvantage lies on their typically wavelength-dependent behavior [10].

MZMs are nearly wavelength-independent, have high extinction ratio (>20 dB) and low insertion loss (approximately 4 dB) [10]. Note that extinction ratio refers to the ratio between the maximum and minimum signal power at the output of the modulator. The higher the extinction ratio the better the signal amplitude levels are distinguished from each other.

IQ modulators in long-haul optical systems are composed of an arrangement of MZMs. To understand its concept, a simple phase modulator is depicted in Fig. 3.6(a). Certain materials have the property of changing their refractive index proportional to the strength of the applied electric field. This effect corresponds to the so-called electro-optic effect. Lithium Niobate ($LiNbO_3$) is commonly used for this purpose [62], which is illustrated in Fig. 3.6(a) as the electro-optic substrate. Thus, the drive voltage $u_1(t)$ is applied to the electrodes in Fig. 3.6(a), changing the refractive index by Δn_{PM} and causing a delay τ_{PM} in the propagating wave of $\Delta n_{\mathrm{PM}} L_{\mathrm{e}}/c$, where L_{e} is the length of the electrodes. This delay is translated into phase changes $\Delta\phi_{\mathrm{PM}}$ by multiplying it by the angular frequency ω_{c} of the incoming laser source. This is summed up in Eq. (3.3), where λ_{c} is the wavelength of the optical carrier [62].

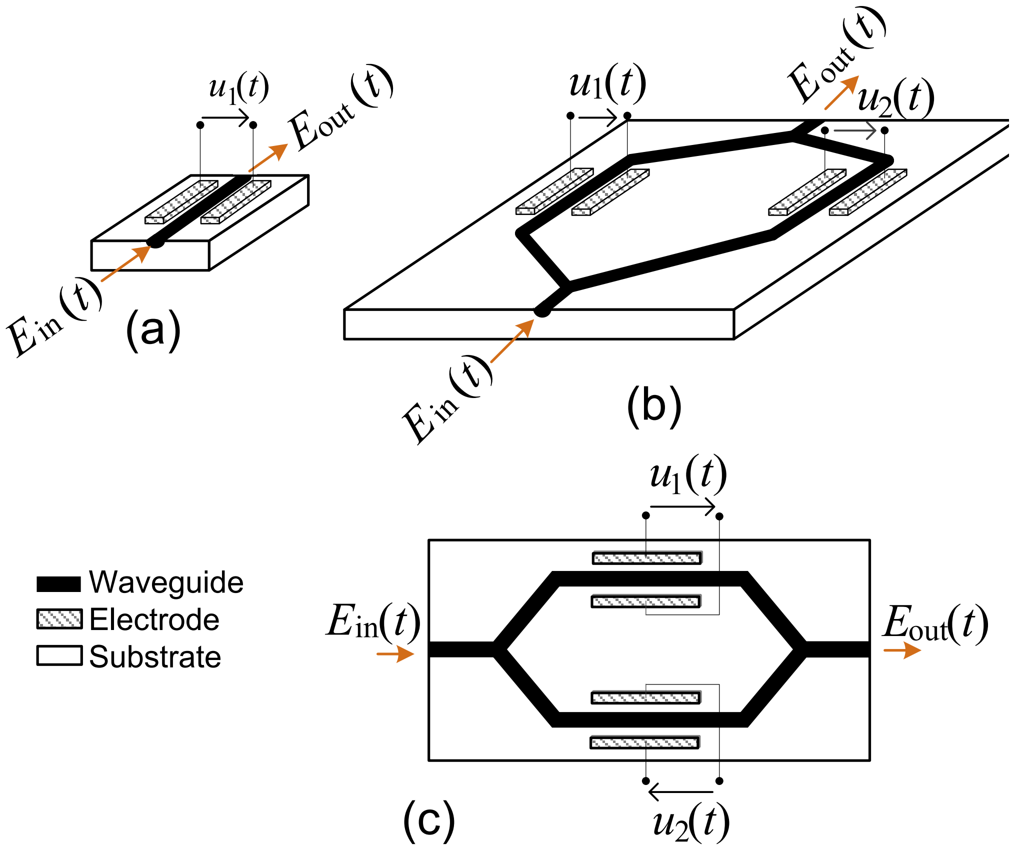


Figure 3.6.: (a) Phase modulator and (b) MZM in perspective view, and (c) MZM in top view.

$$\Delta\phi_{\mathrm{PM}} = \omega_{\mathrm{c}}\tau_{\mathrm{PM}} = \frac{2\pi}{\lambda_{\mathrm{c}}}\Delta n_{\mathrm{PM}} L_{\mathrm{e}} \tag{3.3}$$

As mentioned, the refractive index changes are proportional to the applied voltage on the electrodes. The relation describing this behavior is given by Eq. (3.4) [62,63], where Γ and n_{PM} are both dimensionless and represent the overlap integral between the modulating electric field and the confined optical field in the waveguide and the unperturbed refractive index of the waveguide, respectively; r_{33} and d_e are the electro-optic coefficient in [m/V] (typically 30.9 pm/V for $LiNbO_3$) and the distance between the electrodes in [m], respectively.

$$\Delta n_\mathrm{PM} = \frac{1}{2}\Gamma n_\mathrm{PM}^3 r_{33}\frac{u_1(t)}{d_\mathrm{e}} \tag{3.4}$$

By replacing Eq. (3.4) in Eq. (3.3), the resulting phase changes can be written as Eq. (3.5). It can be observed that the phase changes $\Delta\phi_\mathrm{PM}$ depends on the applied voltage and a constant factor depending on the material and construction characteristics of the phase modulator. Using Eq. (3.5), the relation between the input and output electrical field of a phase modulator can be written as Eq. (3.6) [59], where the constant factor in Eq. (3.5) is represented by π/V_π. V_π is the voltage required to achieve a phase shift of π.

$$\Delta\phi_\mathrm{PM} = -\frac{\pi L_\mathrm{e}}{\lambda_\mathrm{c}}\Gamma n^3 r_{33}\frac{u_1(t)}{d_\mathrm{e}} \tag{3.5}$$

$$E_\mathrm{out} = E_\mathrm{in}e^{j\Delta\phi_\mathrm{PM}} = E_\mathrm{in}e^{-j\frac{\pi L_\mathrm{e}}{\lambda_\mathrm{c}}\Gamma n^3 r_{33}\frac{u_1(t)}{d_\mathrm{e}}} = E_\mathrm{in}e^{j\frac{\pi u_1(t)}{V_\pi}} \tag{3.6}$$

With Eq. (3.6), the relation between the input and output electrical field of a MZM can be derived. As shown in Fig. 3.6(b) and (c), the MZM splits the incoming electrical field into two paths and changes the phase of the incoming field with the drive voltages $u_1(t)$ and $u_2(t)$. Then, the electrical fields are combined. According to Eq. (3.6), the output electrical field of a MZM is expressed as in Eq. (3.7a) or with further mathematical treatment as in Eq. (3.7b).

$$E_\mathrm{out} = \frac{E_\mathrm{in}}{2}\left(e^{j\frac{\pi u_1(t)}{V_\pi}} + e^{j\frac{\pi u_2(t)}{V_\pi}}\right) \tag{3.7a}$$

$$E_\mathrm{out} = E_\mathrm{in}\cos\left(\frac{\pi}{2V_\pi}\left[u_1(t) - u_2(t)\right]\right)e^{j\frac{\pi}{2V_\pi}(u_1(t)+u_2(t))} \tag{3.7b}$$

According to Eq. (3.7b), if $u_1(t) = u_2(t)$, the MZM works as the phase modulator illustrated in Fig. 3.6(a). This configuration of the driving voltages is called *push-push* [59]. If $u_1(t) = -u_2(t)$, the MZM works in *push-pull* configuration, which leads to a pure amplitude modulation.

Equation 3.7b demonstrates an amplitude and phase modulation of the incoming optical power. However, the amplitude and the phase are dependent on each other by using a combination of the driving voltages. In order to solve this issue, the IQ modulator is composed of two MZMs in *push-pull* configuration and a phase shifter as shown in Fig. 3.7. In this way, the data from the quadrature and inphase path are independent from each other. Furthermore, an additional phase shifter is introduced to achieve a 90° shift with a driving voltage of $V_\pi/2$ as required for orthogonal operation of the I and Q branches.

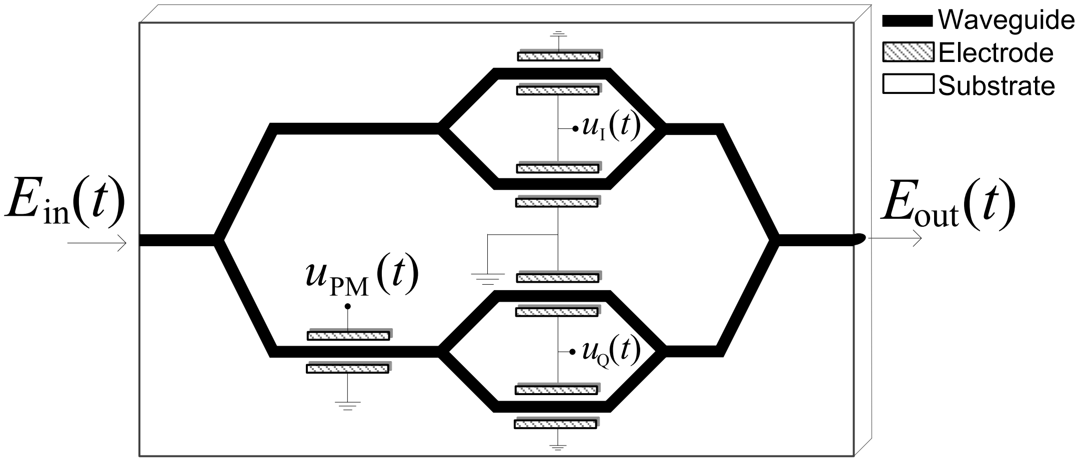


Figure 3.7.: Top view of an optical IQ modulator based on MZMs.

From Fig. 3.7, the relation between the input and output light can be derived as in Eqs. (3.8a) to (3.8c), where the drive voltages of the I and Q branches are $u_{\text{I}}(t)$ and $u_{\text{Q}}(t)$, respectively. In Eq. (3.8c), it can be seen that the output field is related to the driving voltages by cosine functions. If the outputs of the cosine functions are close enough to zero the relation between the input drive voltages $u_{\text{I}}(t)$ and $u_{\text{Q}}(t)$ and output electrical field is nearly linear.

$$E_{\text{out}} = \frac{E_{\text{in}}}{2}\left[\cos\left(\frac{\pi}{V_\pi}u_{\text{I}}(t)\right) + e^{ju_{\text{PM}}\frac{\pi}{V_\pi}}\cos\left(\frac{\pi}{V_\pi}u_{\text{Q}}(t)\right)\right] \tag{3.8a}$$

$$E_{\text{out}} = \frac{E_{\text{in}}}{2}\left[\cos\left(\frac{\pi}{V_\pi}u_{\text{I}}(t)\right) + e^{j\frac{V_\pi}{2}\frac{\pi}{V_\pi}}\cos\left(\frac{\pi}{V_\pi}u_{\text{Q}}(t)\right)\right] \tag{3.8b}$$

$$E_{\text{out}} = \frac{E_{\text{in}}}{2}\left[\cos\left(\frac{\pi}{V_\pi}u_{\text{I}}(t)\right) + j\cos\left(\frac{\pi}{V_\pi}u_{\text{Q}}(t)\right)\right] \tag{3.8c}$$

3.2. Mode-dependent loss

As studied in Chapter 2, multimode fiber poses new challenges compared to systems using SMF. Mode coupling and differential mode delay have to be addressed and its effect on the transmitted signal can be unraveled with proper signal equalization at the RX, as mentioned in Section 3.5. Therefore, mode coupling and DMD adds complexity to the RX compared to single-mode RXs, but it is not a fundamental performance limitation.

A source of fundamental performance limitation and channel capacity decrease is mode-dependent loss (MDL) [64]. MDL, also referred as mode-dependent gain, is the responsible for unequal attenuation or amplification of the modes. In a pessimistic scenario, MDL can cause the completely extinction of one or several modes after transmission. It can be introduced for all kinds of optical inline components such as amplifiers [65–67], (de)multiplexers [68] and even from the FMF itself if the modes have different attenuation coefficients [69]. Additionally, unideal splices, which are unavoidable in optical links, can also generate MDL [70].

MDL definition is directly related with the channel transfer function $\mathbf{H} \in \mathbb{C}^{N_\mathrm{t} \times N_\mathrm{t}}$, where N_t is the number of tributaries of the transmission system (total number of polarization and spatial modes). Each one of the entries of the channel matrix $\mathbf{H}$, h_{ij}, represent the gain from the transmit tributary j to the received tributary i. Transmit and received tributaries are inputs and outputs of the (D)MUX, and not necessarily correspond to spatial and polarization modes as shown in Section 3.3.1.

The channel $\mathbf{H}$ is assumed to be slowly time-variant (for a certain period of time) disturbed by additive white Gaussian noise (AWGN), considering low enough launch power to avoid the presence of nonlinear distortions. The channel transfer function includes the effect of the optical channel linear impairments, (D)MUX, TX and RX filters. Hence, the channel output $\mathbf{r}_\mathrm{o} \in \mathbb{C}^{N_\mathrm{t}}$ and channel input $\mathbf{s} \in \mathbb{M}^{N_\mathrm{t}}$ can be related by Eq. (3.9), where $\mathbb{M}$ is the set of transmitted symbols and $\boldsymbol{\eta}_\mathrm{ASE} \in \mathbb{C}^{N_\mathrm{t}}$ is the random AWGN component, assuming it is normally distributed with zero mean and with diagonal covariance matrix, i.e. uncorrelated among tributaries.

$$\mathbf{r}_\mathrm{o} = \mathbf{H}\mathbf{s} + \boldsymbol{\eta}_\mathrm{ASE} \tag{3.9}$$

MDL is computed using the singular value decomposition (SVD) on the channel matrix [71]. With the SVD, $\mathbf{H}$ can be decomposed as indicated in Eq. (3.10) [71,72], where $\mathbf{U}$ and $\mathbf{V}$ are unitary square matrices in $\mathbb{C}^{N_\mathrm{t} \times N_\mathrm{t}}$, and $\boldsymbol{\Lambda} \in \mathbb{R}^{N_\mathrm{t} \times N_\mathrm{t}}$ is a diagonal matrix, whose diagonal is composed by the non-negative and decreasing singular values of $\mathbf{H}$, λ_i. As decreasing, the relation $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_{N_\mathrm{t}}$ is meant. The diagonal component of Λ, $\lambda_1, \lambda_2, ..., \lambda_{N_\mathrm{t}}$, correspond to the singular values of $\mathbf{H}$ or, equivalently, the square root of the eigenvalues of $\mathbf{HH}^*$.

$$\mathbf{H} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^\mathrm{H} \tag{3.10}$$

A difference in the singular values of $\mathbf{H}$, λ_i, describes MDL. MDL is generally defined as the ratio between the maximum and minimum eigenvalues of $\mathbf{HH}^*$ [19, 64]. It is usually expressed in dB, as shown mathematically in Eq. (3.11).

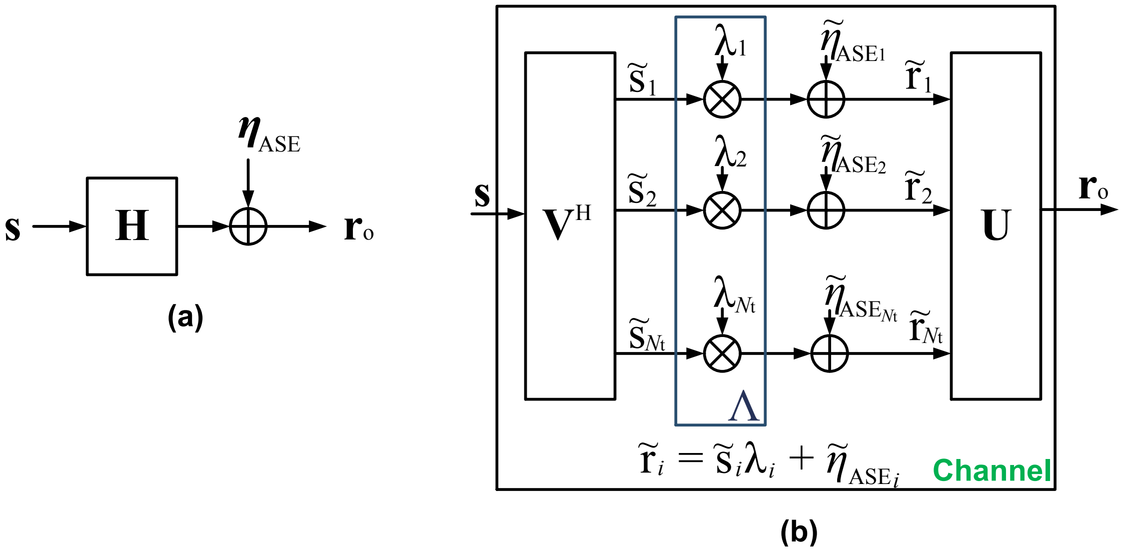
$$\Upsilon[\mathrm{dB}] = 10\,\log_{10}\left(\frac{\max\left\{\lambda_1^2, \lambda_2^2, ..., \lambda_{N_\mathrm{t}}^2\right\}}{\min\left\{\lambda_1^2, \lambda_2^2, ..., \lambda_{N_\mathrm{t}}^2\right\}}\right) \tag{3.11}$$

MDL is better visualized by multiplying both sides of Eq. (3.9) by $\mathbf{U}^\mathrm{H}$ and using Eq. (3.10). Further derivation of this operation is shown in Eqs. (3.12a) to (3.12d), where the matrix products $\mathbf{U}^\mathrm{H}\mathbf{r}_\mathrm{o}$, $\mathbf{V}^\mathrm{H}\mathbf{s}$ and $\mathbf{U}^\mathrm{H}\boldsymbol{\eta}_\mathrm{ASE}$ become $\widetilde{\mathbf{r}}$, $\widetilde{\mathbf{s}}$ and $\widetilde{\boldsymbol{\eta}}_\mathrm{ASE}$, respectively.

$$\begin{aligned}
\mathbf{U}^\mathrm{H}\mathbf{r}_\mathrm{o} &= \mathbf{U}^\mathrm{H}\mathbf{H}\mathbf{s} + \mathbf{U}^\mathrm{H}\boldsymbol{\eta}_\mathrm{ASE} && \text{(3.12a)}\\
\mathbf{U}^\mathrm{H}\mathbf{r}_\mathrm{o} &= \mathbf{U}^\mathrm{H}\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^\mathrm{H}\mathbf{s} + \mathbf{U}^\mathrm{H}\boldsymbol{\eta}_\mathrm{ASE} && \text{(3.12b)}\\
\mathbf{U}^\mathrm{H}\mathbf{r}_\mathrm{o} &= \boldsymbol{\Lambda}\mathbf{V}^\mathrm{H}\mathbf{s} + \mathbf{U}^\mathrm{H}\boldsymbol{\eta}_\mathrm{ASE} && \text{(3.12c)}\\
\widetilde{\mathbf{r}} &= \boldsymbol{\Lambda}\widetilde{\mathbf{s}} + \widetilde{\boldsymbol{\eta}}_\mathrm{ASE} && \text{(3.12d)}
\end{aligned}$$

An illustration of the channel model, SVD and the relation in Eq. (3.12d) is shown in Fig. 3.8, where $\widetilde{\mathrm{r}}_i$, $\widetilde{\mathrm{s}}_i$ and $\widetilde{\eta}_{\mathrm{ASE}_i}$ are the i-th component of the corresponding vectors. In Fig. 3.8(b), parallel paths or sub-systems, where the input and output are $\widetilde{\mathrm{s}}_i$ and $\widetilde{\mathrm{r}}_i$, respectively, can be identified. MDL can be visualized as when the ratio among the gains of those sub-systems, λ_i, is different than one. That is, in the presence of MDL, those parallel paths have different gains.

Note that in the definitions of the transmitted and received vectors, the number of tributaries used for transmission will always be in this work equal to the number of detected tributaries, and at the same time is equal to the maximum number of spatial and polarization modes supported by the fiber. This is mainly because of two reasons. Firstly, this is done in order to fully exploit the spatial diversity offered by



(a)

(b)

Figure 3.8.: (a) Channel model and (b) its conversion into parallel, independent scalar Gaussian sub-channels.

the fiber; and secondly, if less modes are launched and detected than the number of modes that the fiber is able to support, energy might be transferred to those modes that are not being detected due to mode coupling. Not detecting all the modes leads to a sub-optimum performance of the RX, i.e. higher bit error rate (BER) or diminished transmission reach.

3.3. Optical link components

As illustrated in Fig. 3.2, the optical link is composed of mainly four types elements: (D)MUX, FMFs, FM-amplifier and MDM ROADMs. After conversion from electrical to the optical signal by the IQ modulator, different single-mode polarization-multiplexed (POLMUX) signals need to be multiplexed into different modes by the MUX. The inverse process is carried out by the DMUX. The signal is transported through the FMF, whose principle is described in Chapter 2. Because of losses in the optical fiber, the signal has to be amplified in order to transmit it further. This is the task of the few-mode amplifiers (FMAs). MDM ROADMs make the network more flexible and re-direct WDM channels to a specific destination.

This section describes the principle of the optical components in the link, such as mode (D)MUX, FMF amplifier and ROADMs.

3.3.1. Mode multiplexer and demultiplexer

Mode multiplexer and demultiplexers (MUXs and DMUXs) are responsible for launching the modulated information from the SMFs to the FMFs and vice versa, respectively. There are mainly two types of devices that accomplish this job: mode-selective and spatial (D)MUXs.

For uniquely multiplexing and demultiplexing the signals in and from the fiber, the signals have to be orthogonal in at least one dimension, e.g. in frequency, polarization, quadrature, space, time. For example, WDM channels are orthogonal in the frequency domain, since WDM channels do not overlap in the frequency domain. Mode (D)MUXs take advantage of the property of LP modes of being an orthogonal basis, as shown in Eq. (2.21). In the following, the two types of mode (D)MUXs are addressed.

Mode-selective (D)MUXs

Mode-selective MUXs target to excite each one of the modes separately. Within this category there are several types of mode (D)MUXs, which are based on: phase masks, spatial light modulators (SLMs) and long-period gratings.

The principle of the phase-mask-based (D)MUXs is initially proposed in [73,74] and was adapted for MDM systems first in [51], and then followed by the experiments reported in [69,75]. This type of (D)MUX takes advantage of the unique properties of the LP modes with respect to their phase profile. As shown in Fig. 2.7, each mode has a phase profile which takes the value of 0 or π. The idea behind this type of MUX is to shape the phase of the incoming beam from the SMF with a binary phase mask or plate which causes a phase shift of 0 or π according to the desired mode to be multiplexed into the FMF. As an example, Figure 3.9(a) illustrates the selective excitation of one of the LP_{11} modes. In order to excite the other LP_{11} mode, it is sufficient to rotate the phase plate 90°, whereas no phase plate is required for multiplexing LP_{01}. Note that every mode can be uniquely addressed, due to the orthogonality property of the linearly polarized modes.

The DMUX operates similar to the MUX, but the beam flows in the opposite direction. This is shown in Fig. 3.9(b). The principle of the DMUX can be compared by an optical matched filter for pattern recognition [76]. Usually, a linear combination of the modes is to be multiplexed. This beam containing the linear combination of all modes is collimated and then passes through a phase plate which resembles the

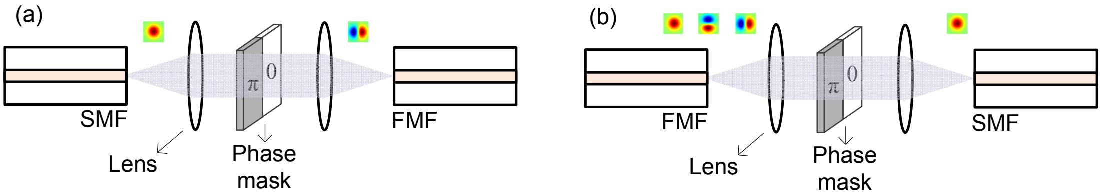


Figure 3.9.: Operating principle of phase-plate-based (a) MUX and (b) DMUX.

complex conjugate of the mode phase profile. The beam passing through the phase plate is represented as the multiplication of input beam with the complex conjugate of the mode phase profile. As the linearly polarized modes as orthogonal, by means of this optical multiplication each mode can be uniquely demultiplexed.

For selectively multiplexing several modes at the same time with the phase plate scheme it is necessary to use one SMF per spatial mode to be multiplexed and combine the phase modulated beams passively with beam combiners, as shown in Fig. 3.10. The mode DMUX has the same structure as the MUX, as mentioned before, but letting the signal flow from right to left in Fig. 3.10.

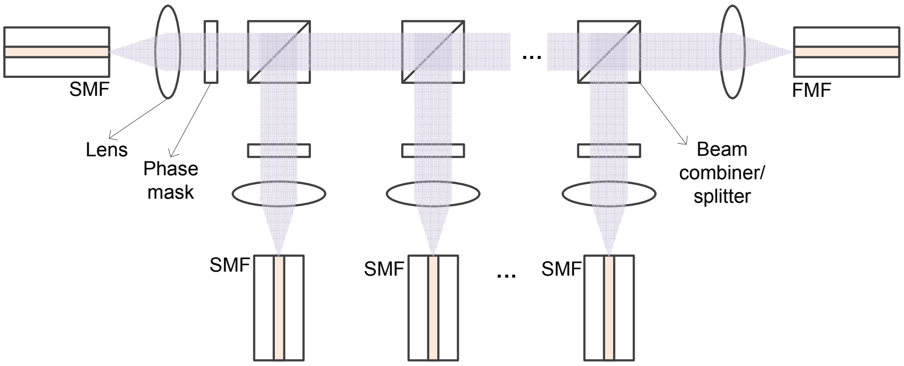


Figure 3.10.: Mode MUX for simultaneously exciting several spatial modes.

One of the disadvantages of the phase plate based (D)MUX is the high insertion loss related to this device [46, 51]. The loss is related to the number of beam splitters/combiners used in the setup and thus, is also proportional to the number of modes to be (de)multiplexed. This is due to the fact that one port of the beam splitter/combiner remains unused. Other issue that needs to be addressed is the phase and amplitude similarities of the supported modes when there are more than three supported modes [46]. For example, the modes LP_{01} and LP_{02} have a similar phase and amplitude profile, therefore in order to uniquely address them not only the phase has to be shaped but also the amplitude.

SLMs work in a similar way as (de)multiplexers based on phase masks. Usually, liquid crystal on silicon (LCoS) devices are used as the element which changes the beam phase front according to the mode to be (de)multiplexed [77, 78]. LCoS are programmable devices, which are composed of a 2D pixel array containing e.g. 1920x1080 pixels as in [77]. The phase of each pixel can be addressed and modified accordingly. However, they are frequency-dependent, which is problematic when dealing with WDM systems [46]. These devices are also polarization dependent, therefore every polarization is converted separately to a higher order mode, which makes the setup more complex.

Mode converters based on long-period gratings in turn work differently. They take advantage of resonant coupling or the longitudinal phase matching of gratings, which states that mode a will couple into mode b if the fiber is pressed against a grating with a perturbation period of Λ equivalent to $2\pi/(\beta_\mathrm{a}-\beta_\mathrm{b})$ or in terms of the effective refractive index $\bar{n}$, $\lambda/(\bar{n}_\mathrm{a}-\bar{n}_\mathrm{b})$ [48]. The effectiveness of the mode conversion has already been proven in [79–81]. However, as shown before with the perturbation period equation, the mode conversion is directly influenced by the wavelength, which makes it not convenient to implement for the whole C and L-band.

Spatial (D)MUXs

In contrast with mode-selective (D)MUXs, spatial multiplexers use each modulated signal from the SMF to illuminate a portion of the end facet of the fiber. Hence, each signal is spatially separated and at the same time excites a linear combination of the modes. The implementation of spatial multiplexers does not require beam splitters/combiners, which reduces the insertion loss of the (D)MUX. Spatial (D)MUXs can be implemented in two different ways: as spot couplers or photonic lanterns.

Spot couplers were proposed in [82–84] constructed with bulky optical pieces. The mode MUX illuminates the end facet of the fiber forming an arrangement of non-overlapping spots. The mode DMUX does the inverse operation and takes the phase and amplitude of a portion of the optical field coming out of the fiber.

The coupling matrix describing the spot-coupler (D)MUX can be determined by evaluating the overlap integral between the spots and the supported linearly polarized modes [83]. The authors in [83] propose the use of another base for the degenerate modes ($l \neq 0$ for the mode LP_{lm}) to ease the representation of the coupling matrix. This corresponds to the complex notation mentioned in Section 2.2, where $\Phi(\varphi)$ in the electrical field $\psi(r,\varphi) = R(r)\Phi(\varphi)$ takes the value of $e^{jl\varphi}$, instead of $\cos l\varphi$ or $\sin l\varphi$. The complex notation allows to have a field distribution, which is independent of the angular coordinate suggesting the use of rings to arrange the spots. Thus, each real LP_{lm} degenerate mode (notated with a or b) can be expressed in terms of the complex modes $\hat{\psi_{lm}}(r,\varphi)$ and its conjugate $\hat{\psi_{lm}}^*(r,\varphi)$ and vice versa as shown in Eqs. (3.13a) to (3.13d). The square root term in Eqs. (3.13a) to (3.13d) indicates that half of the power is given by each one of the addends. The conversion to complex notation is also graphically shown in Fig. 3.11 with $\mathrm{LP}_{11\mathrm{c}}$ and $\mathrm{LP}_{21\mathrm{c}}$, where c denotes the complex notation.

$$\psi_{lm\mathrm{b}}(r,\varphi) = \left(\hat{\psi}_{lm}(r,\varphi) + \hat{\psi}^*_{lm}(r,\varphi)\right)/\sqrt{2} \tag{3.13a}$$

$$\psi_{lm\mathrm{a}}(r,\varphi) = \left(-j\hat{\psi}_{lm}(r,\varphi) + j\hat{\psi}^*_{lm}(r,\varphi)\right)/\sqrt{2} \tag{3.13b}$$

$$\hat{\psi}_{lm}(r,\varphi) = \left(\psi_{lm\mathrm{b}}(r,\varphi) + j\psi_{lm\mathrm{a}}(r,\varphi)\right)/\sqrt{2} \tag{3.13c}$$

$$\hat{\psi}^*_{lm}(r,\varphi) = \left(\psi_{lm\mathrm{b}}(r,\varphi) - j\psi_{lm\mathrm{a}}(r,\varphi)\right)/\sqrt{2} \tag{3.13d}$$

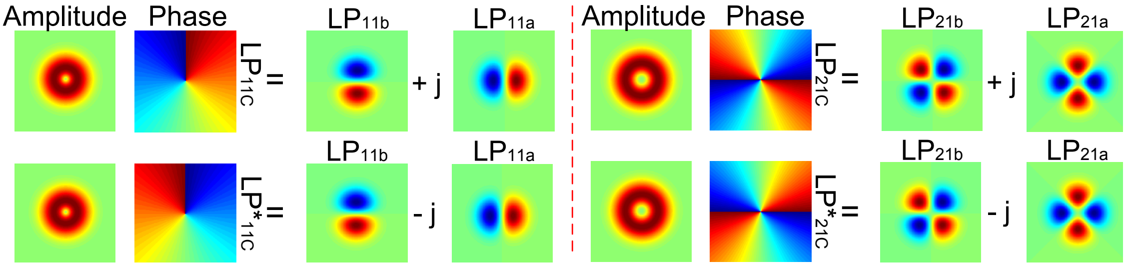


Figure 3.11.: Conversion from real to complex notation of LP_{11} and LP_{21} modes.

As mentioned before, the complex notation suggests the arrangement of the spots ring-wise. In [82, 83], two rules are presented, for the spots arrangement as the number of modes increases. First, the number of spots corresponds to the number of spatial modes, thus, one spot is added per non-degenerate mode and two spots are added per degenerate mode group. Secondly, the spots are first arranged on a circle; as the number of modes increases new spots are added to the circle, but if the modes to be supported by the FMF increases in radial number (m increases for the mode LP_{lm}), then a new spot in the center is added. Different spot arrangements proposed in [82, 83] are depicted in Fig. 3.12.

Using the complex notation, the coupling matrix of the spot coupler for three modes is given by Eq. (3.14) [83], where a_0 and a_1 represent the overlap integral between a spot and LP_{01} and $\mathrm{LP}_{11\mathrm{c}}$, respectively, and ψ_{spot_n} is the n^{th} launch field from the

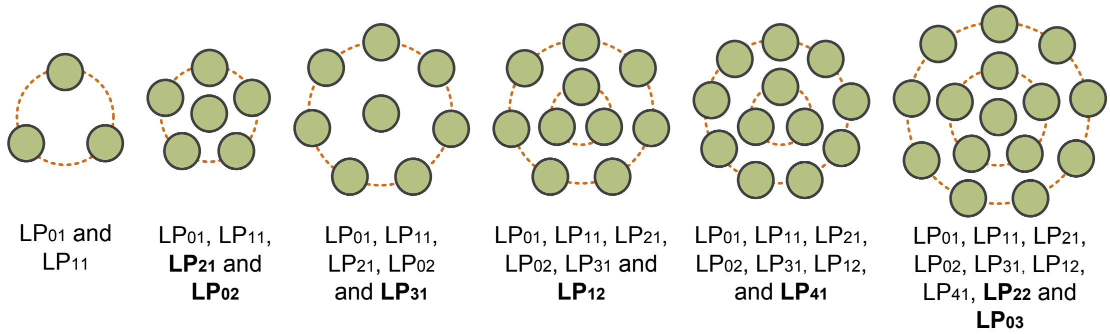


Figure 3.12.: Proposed spot arrangements in [82, 83] for up to 15 spatial modes (9 mode groups) supported by the FMF.

SMF. Each element of the first row of the coupling matrix in Eq. (3.14) corresponds to a_0 since there is neither azimuthal variation of the amplitude nor of the phase. This applies in general for LP_{0m} modes, whose overlap integral is common for spots of the same ring. The second row is described by a reference overlap integral a_1 and a rotated version of it for the rest of the matrix positions.

$$\begin{pmatrix} a_0 & a_0 & a_0 \\ a_1 & a_1 e^{j2\pi/3} & a_1 e^{-j2\pi/3} \\ a_1 & a_1 e^{-j2\pi/3} & a_1 e^{j2\pi/3} \end{pmatrix} \begin{pmatrix} \psi_{\mathrm{spot}_1} \\ \psi_{\mathrm{spot}_2} \\ \psi_{\mathrm{spot}_3} \end{pmatrix} = \begin{pmatrix} \hat{\psi}_{01} \\ \hat{\psi}_{11} \\ \hat{\psi}^*_{11} \end{pmatrix} \tag{3.14}$$

The coupling matrix in Eq. (3.15) represents the case when the fiber supports six spatial modes [83]. The elements of this matrix are computed similar to the three-mode coupling matrix. Here, a_0, a_1, a_2 and a_3 correspond to the overlap integral of a spot on the outer ring with the field of the modes LP_{01}, $\mathrm{LP}_{11\mathrm{c}}$, $\mathrm{LP}_{21\mathrm{c}}$ and $\mathrm{LP}_{02\mathrm{c}}$, respectively. b_0 and b_3 are the overlap integral with the center spot and the LP_{01} and LP_{02}, respectively.

$$\begin{pmatrix} a_0 & a_0 & a_0 & a_0 & a_0 & b_0 \\ a_1 & a_1 e^{j2\pi/5} & a_1 e^{j4\pi/5} & a_1 e^{j6\pi/5} & a_1 e^{j8\pi/5} & 0 \\ a_1 & a_1 e^{-j2\pi/5} & a_1 e^{-j4\pi/5} & a_1 e^{-j6\pi/5} & a_1 e^{-j8\pi/5} & 0 \\ a_2 & a_2 e^{j4\pi/5} & a_2 e^{j8\pi/5} & a_2 e^{j12\pi/5} & a_2 e^{j16\pi/5} & 0 \\ a_2 & a_2 e^{-j16\pi/5} & a_2 e^{-j8\pi/5} & a_2 e^{-j12\pi/5} & a_2 e^{-j16\pi/5} & 0 \\ a_3 & a_3 & a_3 & a_3 & a_3 & b_3 \end{pmatrix} \begin{pmatrix} \psi_{\mathrm{spot}_1} \\ \psi_{\mathrm{spot}_2} \\ \psi_{\mathrm{spot}_3} \\ \psi_{\mathrm{spot}_4} \\ \psi_{\mathrm{spot}_5} \\ \psi_{\mathrm{spot}_6} \end{pmatrix} = \begin{pmatrix} \hat{\psi}_{01} \\ \hat{\psi}_{11} \\ \hat{\psi}^*_{11} \\ \hat{\psi}_{21} \\ \hat{\psi}^*_{21} \\ \hat{\psi}_{02} \end{pmatrix} \tag{3.15}$$

Photonic lanterns are the second kind of spatial (D)MUX used for the study of FMF in optical long-haul communications. They are similar to the spot coupler described above, however, here the idea is instead of using free space optics use compact optical devices. Photonic lanterns are employed in the field of astrophysics and, more specific, in telescopes for coupling the light coming from large core fibers (50 to 300 μm of core diameter) into SMFs [85]. For optical communications they achieve virtually the same purpose of coupling all the light of a FMF into single SMFs and vice versa.

There are mainly two ways of implementing photonic lanterns: all-fiber based and 3-dimensional waveguide lanterns (3DW) [86, 87]. All fiberized lanterns consist of SMFs, which are wrapped by a low refractive index capillary. The fibers are tapered down so that the originally SMF cores vanish, the SMFs cladding become the new FMF core and the low index capillary become the new FMF cladding. This implementation of photonic lantern is illustrated in Fig. 3.13. In the case of the MUX, the number of SMFs to be tapered down correspond to the number of spatial modes supported by the FMF to be used. Due to this gradual transition down to the FMF, the light propagate first uncoupled as in a conventional SMF bundle and then as the taper reduces, the light from the SMF cores strongly couples as the cores become closer forming the so-called supermodes; subsequently, the light couples into the cladding of these SMFs, which is now the core of the FMF. This adiabatic tapering allows that no light is lost, i.e. the light couples completely into the FMF. The demultiplexing process is explained by the inverse of the process describe before.

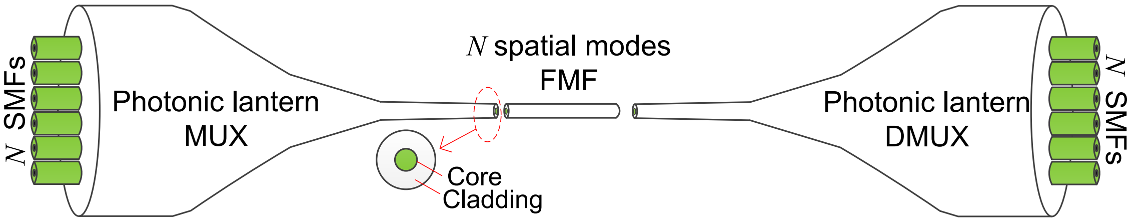


Figure 3.13.: Fiber-based photonic lantern.

Alternatively, photonic lanterns can be implemented as a 3DW. It consists of glass substrate (silica), which has been laser-inscribed to form N waveguides. In the case of a MUX, the input of the waveguides are N SMFs, whose light is transported through the waveguides and brought together to illuminate the core of the FMF. Figure 3.14 depicts schematically a 3DW that multiplexes the light into a FMF that supports six modes. In practice, the last stage of the 3DW MUX to couple the light into the FMF, the 3DW requires an additional down tapering stage [68] or free space optics [88] to match light coming from the waveguide into the FMF core and avoid loss.

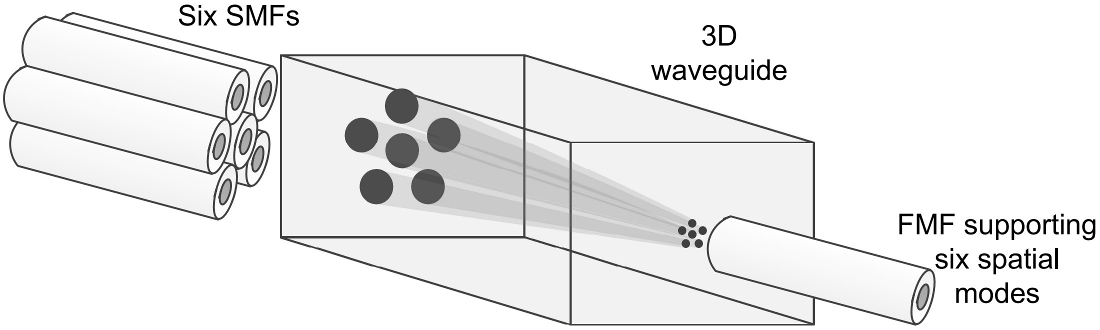


Figure 3.14.: Example of 3DW with a fiber that support six spatial modes.

3.3.2. Optical amplification

As mentioned in Section 2.3.1, fiber loss is one of the fiber impairments that needs to be addressed. Long-haul communication along hundreds of kilometers is impossible without proper compensation of the signal attenuation due to the fiber itself and other inline passive optical components such as multiplexers and couplers. In the past, to accomplish this task optical signal regenerators were used. They are devices that converted the optical signal into electrical, reconstructed the signal electrically and converted it back to the optical domain. There are two main problems with regenerators: they are costly and their design and structure heavily depend on the modulation format and symbol rate. After 1995 optical amplifiers were developed to solve this situation and to enable the simultaneous amplification of WDM channels all over C and L-band. In 1996, they were already part of transoceanic submarine links across the Pacific and the Atlantic [1].

This section describes the working principle of optical amplifiers, their impact on the signal-to-noise ratio (SNR) and the particular case of few-mode amplification.

Principle of amplification

There are principally two types of optical amplifies: lumped and distributed amplifiers. Lumped amplifiers are represented by EDFAs, which amplify the signals after propagation through a fiber of certain length (typically between 60 and 80 km). Raman amplifiers perform the distributed amplification, which takes place inside the fiber used for data transmission and uses the nonlinear effect of stimulated Raman scattering (SRS). The principle of operation and their upgrade to MDM systems is described in this section.

Amplification in EDFAs occurs inside a fiber, which is doped with ions of a rare-earth element called erbium (Er^{3+}). There are other rare-earth elements eligible for amplification of signals at wavelengths between 0.5 and 3.5 μm such as holmium, neodymium, samarium, thulium and ytterbium [1]. However, the use of erbium became popular because of its amplification properties near the 1.55 μm window, coinciding with the wavelength region where the fiber exhibits low attenuation.

Similar to the working principle of lasers, EDFAs amplify the signal via stimulated emission. This can be explained with help of the three-level energy diagram of the erbium-doped fiber depicted in Fig. 3.15. The process of amplification starts with a laser pump, which excites the electrons inside the erbium ions taking them from the ground to an excited energy level. For this purpose, the laser pump uses its photons whose energy matches the energy difference hf_{pump} between the ground and the excited state, being h Planck's constant (6.626068×10^{-34} J s) and f_{pump} the frequency of the pump laser. The ground level is denoted in Fig. 3.15 as $^4\mathrm{I}_{15/2}$ following the Russel-Saunders coupling terminology [2]. Two superior energy levels are also shown in Fig. 3.15: $^4\mathrm{I}_{11/2}$ and $^4\mathrm{I}_{13/2}$. The level $^4\mathrm{I}_{11/2}$ is achieved with a pump laser at 980 nm, whereas the $^4\mathrm{I}_{13/2}$ with 1480 nm. Pumps lasers at 980 nm and 1480 nm are commonly used because of their ease of fabrication and because at these frequencies the erbium ions can absorb easily energy from the pump; however, the former frequency is commonly used due to their high gain efficiency (10 dB per mW of pump power) and low noise figure [10, 89].

Using the 980-nm pump, the electrons at the energy level $^4\mathrm{I}_{11/2}$ have a relatively short lifetime (around 1 μs [90]), which makes them to quickly decay nonradiatively to the energy level $^4\mathrm{I}_{13/2}$. The energy level $^4\mathrm{I}_{13/2}$ is metastable, which implies that the electrons stay in this state for a relatively long time (around 10 ms [90]). At this point the electron population in the higher energy level is higher than at ground, this is called population inversion and it is a requirement for amplification. The electron in $^4\mathrm{I}_{13/2}$ is stimulated by a signal photon to release its energy in form of another photon with the same characteristics of the stimulating photon, i.e. same phase, wavelength and propagation direction. These two photons can each stimulate the transition of another electron to the ground state causing the emission of two further photons. This process is repeated many times and leads to an increase in the signal power. Note that the emission is possible since the separation between

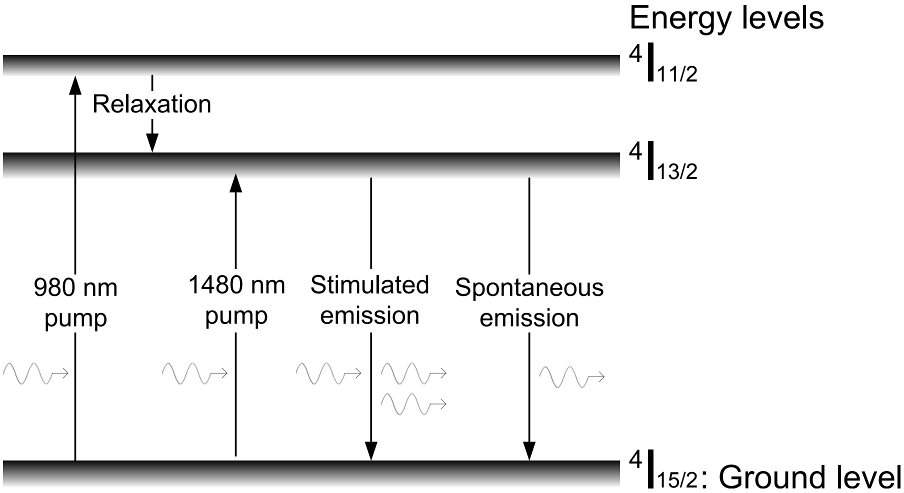


Figure 3.15.: Three-level energy diagram of the erbium-doped silica fiber [2, 90].

the ground state and $^4I_{13/2}$ corresponds to the wavelength range between 1.46 μm to 1.62 μm, which comprises the transmission bands of interest, C and L- band [34].

Electrons which spontaneously go to the ground state emit a photon with random phase, wavelength and propagation direction. These electrons are not used for amplification, and subsequently this process decreases the gain. Spontaneous emissions are also amplified along the erbium-doped fiber and also by subsequent amplifiers. Such spurious emission is called amplified spontaneous emission (ASE).

Figure 3.16 shows a basic diagram of an EDFA. As depicted in Fig. 3.16 the pump is combined with the signal by means of a coupler and it can counter (backward pumping) or co-propagate (forward pumpimg) with the signal, or both (bidirectional pumping). The erbium-doped fiber can be from few meters long to 100 m depending on the doping level [31]. Insufficient pump power and too long erbium-doped fiber can lead to loss rather than gain, that is why these parameters should be properly dimensioned [31]. Isolators are part of the amplifier structure to let the signal pass in only one direction. The isolator at the input prevents ASE noise to disturb the incoming signal and at the output prevents the output light to propagate back to the amplifier, i.e. to avoid back-reflections. A gain equalizing filter is also placed after amplification since, as depicted schematically in Fig. 3.16, the typical gain profile of an EDFA is not flat in the C-band. An EDFA can have a gain of up to 40 dB at an output power of 23 dBm [10].

Raman amplification is based on SRS. As mentioned in Section 2.3.5, SRS is a nonlinear effect that is initiated by a strong pump laser or spontaneously, similar to the case of EDFAs. Raman amplification can be explained with the energy level

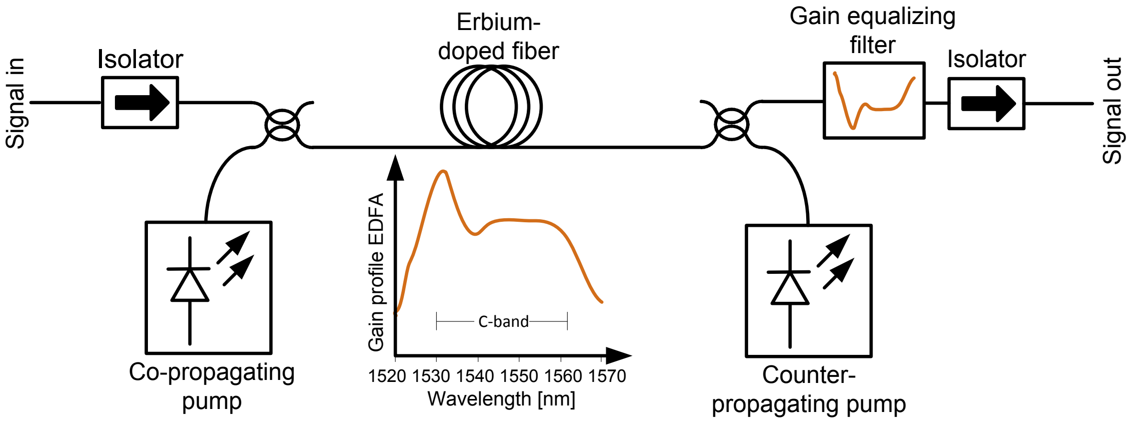


Figure 3.16.: Basic EDFA structure.

diagram in Fig. 3.17. Photons from the pump signal are absorbed by the silica molecules, which are taken to a higher excited energy level denoted in Fig. 3.17 as W_3. The process of stimulated light emission takes place when the incoming signal photons stimulate the molecules to emit photons at a lower frequency than the pump and at the same frequency, phase and polarization as the stimulated signal photons. The Raman gain achieves its maximum when the pump is approximately 13.2 THz above (or 100 nm below) the frequency of the propagating signal. This frequency difference, $f_{\text{pump}} - f_{\text{signal}}$, is called the Stokes shift [50]. The remaining energy difference between the pump and signal photons is absorbed by the silica molecules via phonons (molecular vibrations).

The main difference between the stimulated emission in the EDFAs and the SRS is that in the former process an incident photon stimulates the emission of another photon without loosing its energy; in the case of the SRS, the pump photon gives up its energy (scatters light) to produce another photon [1].

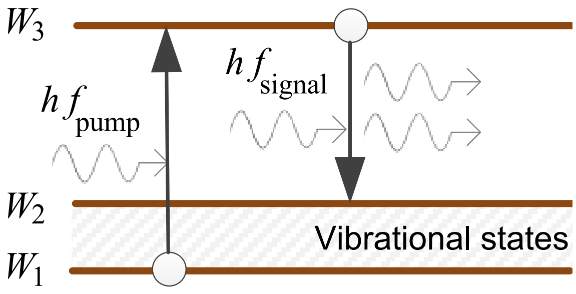


Figure 3.17.: Diagram of energy levels of Raman amplification via SRS [50, 61].

The structure of the Raman amplifiers is similar than that shown in Fig. 3.16 for

EDFAs, with the difference that the amplification medium is the transmission fiber instead of an erbium-doped fiber. The Raman amplification is done in a distributed manner along the length of the transmission fiber. This has the advantage of having a better noise figure than EDFAs, since the lowest power along the link is not as low as in the case of EDFAs, which leads to less noise amplification [50]. Nevertheless, Raman amplifiers might not compensate for all the link loss and are therefore combined with EDFAs, which gives a noise figure improvement close to three decibels [10].

Another advantage of Raman amplification is that it can take place in any frequency band. Additionally, its amplification band extends up to 30 THz [1,31]. However, the Raman gain is not uniform. This problem is solved by using several pumps at slightly different wavelength and power levels such that the total gain is virtually flat [50]. A significant disadvantage of the Raman amplifiers is that it requires higher pump power than its counterpart, the EDFA, which makes them less cost-efficient [50, 61].

Optical signal-to-noise ratio

Optical amplification introduces noise in the transmitted signal due to photons that have random phase, wavelength and propagation direction due to ASE. Each one of the amplifiers along the optical link contributes to noise accumulation and limits the system performance. The optical signal-to-noise ratio (OSNR) is frequently used as a measure of system performance in both single-mode and few-mode systems. The OSNR is commonly measured per mode and corresponds to the ratio between the average power of the transmitted signal after amplification $P_{\text{sig,out}}$ and the noise power at certain fixed reference bandwidth as defined in Eq. (3.16), where N_0 is the single-polarization one-sided ASE noise power spectral density, B_{ref} is the bandwidth reference and N_{spans} is the number of spans. B_{ref} has been standardized in the optical community as 0.1 nm or 12.5 GHz at 1550 nm [6, 9]. In Eq. (3.16), it is assumed that all spans have the same characteristics: same loss, amplification, added noise and length. The factor of two in the denominator of Eq. (3.16) is a constant value indicating that the noise is present in both polarizations.

$$OSNR = \frac{P_{\text{sig,out}}}{2N_0 B_{\text{ref}} N_{\text{spans}}} \tag{3.16}$$

The noise power spectral density in Eq. (3.16) is given by [34, 91]

$$N_0 = n_{\text{sp}} h f_0 (G-1) \ , \tag{3.17}$$

where n_{sp}, f_0 and G are the spontaneous emission coefficient, a reference frequency and the gain of a single amplifier, respectively. Replacing Eq. (3.17) in Eq. (3.16), OSNR can be expressed as Eq. (3.18).

$$OSNR = \frac{P_{\mathrm{sig,out}}}{2n_{\mathrm{sp}} h f_{\mathrm{c}} (G-1) B_{\mathrm{ref}} N_{\mathrm{spans}}} \tag{3.18}$$

It is helpful for the OSNR analysis, that the OSNR is expressed in terms of the noise figure NF of an amplifier, which is given by [1, 91]

$$NF = 2n_{\mathrm{sp}} \frac{G-1}{G} \ . \tag{3.19}$$

By replacing the term $(G-1)$ from Eq. (3.19) in Eq. (3.18) and considering that $P_{\mathrm{sig,out}}$ is equal to the power of the signal at the input of the amplifier $P_{\mathrm{sig,in}}$ multiplied by the amplifier gain G, Eq. (3.18) can be re-writen in linear scale and in dB as

$$OSNR = \frac{P_{\mathrm{sig,in}}}{h f_{\mathrm{c}} NF B_{\mathrm{ref}} N_{\mathrm{spans}}} \tag{3.20a}$$

$$OSNR_{\mathrm{dB}} = P_{\mathrm{sig,in}_{\mathrm{dBm}}} - 10\log_{10}\left(h f_{\mathrm{c}} B_{\mathrm{ref}}\right) - NF_{\mathrm{dB}} - 10\log_{10}\left(N_{\mathrm{spans}}\right) \tag{3.20b}$$

$$OSNR_{\mathrm{dB}} = P_{\mathrm{sig,out}_{\mathrm{dBm}}} - \alpha_{\mathrm{dB}} L_{\mathrm{span}} + 58\ \mathrm{dBm} - NF_{\mathrm{dB}} - 10\log_{10}\left(N_{\mathrm{spans}}\right) \tag{3.20c}$$

where the subscript dB or dBm indicates the units of the variable, L_{span} is the length of a single span in [km], $10\log_{10}\left(h f_{\mathrm{c}} B_{\mathrm{ref}}\right)$ is approximately -58 dBm, taking the reference frequency and bandwidth as as 193.4 THz (1550 nm) and 12.5 GHz at 1550 nm, respectively.

Equation 3.20c suggests that a shorter span length will contribute to a higher OSNR. However, the shorter the span length, the more amplifiers that have to be placed, which turns up to be not economically viable. Consequently, the number of amplifiers have to be minimized for cost reasons and at the same time the performance have to be reliable enough to operate correctly [34].

Few-mode amplifiers

Few-mode fiber amplifiers are key elements in MDM long-haul WDM systems. The challenge here is to provide a more efficient implementation of these devices, i.e. amplify N POLMUX signals using less energy than N times single mode amplifiers.

This is exactly what MDM promises: a power consumption, which increases in less proportion than a linear increase with the number of modes [20, 55, 56].

EDFA and Raman amplification has been studied, being EDFA the most popular choice due to its higher power efficiency. The principle of the EDFA is the same as in single-mode transmission. However, in the case of few-mode amplification, the dimensionless parameter of overlap factor plays an important role. The overlap factor characterizes the portion of the incoming optical mode intensity that overlaps with the erbium ion distribution of the erbium-doped fiber in the amplifier. Only that overlapping portion will stimulate emission and lead to amplification [91]. Therefore, this parameter becomes relevant when the fiber in the transmission link supports several modes, which suggests that the optimum Er^{3+} distribution is no longer a flat top distribution. The overlap factor is given by Eq. (3.21) as the overlap integral $\eta_{\mathrm{p}j,\mathrm{s}i}$ between the normalized optical intensity distribution of the i-th signal mode $\Gamma_{\mathrm{s},i}(r,\varphi)$ and the normalized Er^{3+} distribution of the j-th pump mode $\Gamma_{\mathrm{p},j}(r,\varphi)$ [91]. The better the match between signal and pump (the higher the $\eta_{\mathrm{p}j,\mathrm{s}i}$) the more the amplification that is obtained for the i-th signal mode. The erbium-doped fiber at the frequency of 980 nm supports several modes besides its fundamental mode, and even some more than the FMF used for transmission [92].

$$\eta_{\mathrm{p}j,\mathrm{s}i} = \int_0^{\infty} \int_0^{2\pi} \Gamma_{\mathrm{s},i}(r,\varphi)\Gamma_{\mathrm{p},j}(r,\varphi) r dr d\varphi \tag{3.21}$$

Based on the value of $\eta_{\mathrm{p}j,\mathrm{s}i}$, there have been studies based on simulations which demonstrate that achieving equal gain for all modes depends on different factors such as the modal content of the pump, power given to each one of the supported mode in the pump and the erbium-doping concentration profile. The pump modal content and its power per mode are studied in [92, 93], where the authors show that pumping with only the fundamental mode will favor the gain in the signal LP_{01} mode leading to the so-called mode-dependent gain (MDG), equivalent to MDL. In [93], it is shown that three spatial modes can be virtually equally amplified by using the LP_{21} mode with a small amount of LP_{01} mode as pumps. Furthermore, combined with different pump modes power, specially tailored profile of the erbium ions concentration is found in [92] in order to achieve gain equalization. Other research groups have proposed alternatives to replace the specially tailored erbium distribution, which is difficult to fabricate, by much simpler distribution profiles: ring-doped [94, 95], concentric ring-doped (ring-doped and the center is also doped) [96] and raised profile (slightly higher doping of erbium at the core edge) [94]. These concepts for gain equalization in EDFAs are illustrated in Fig. 3.18, where the pump laser is split into N paths that are converted to higher order modes with phase plates and their power is regulated with optical attenuators.

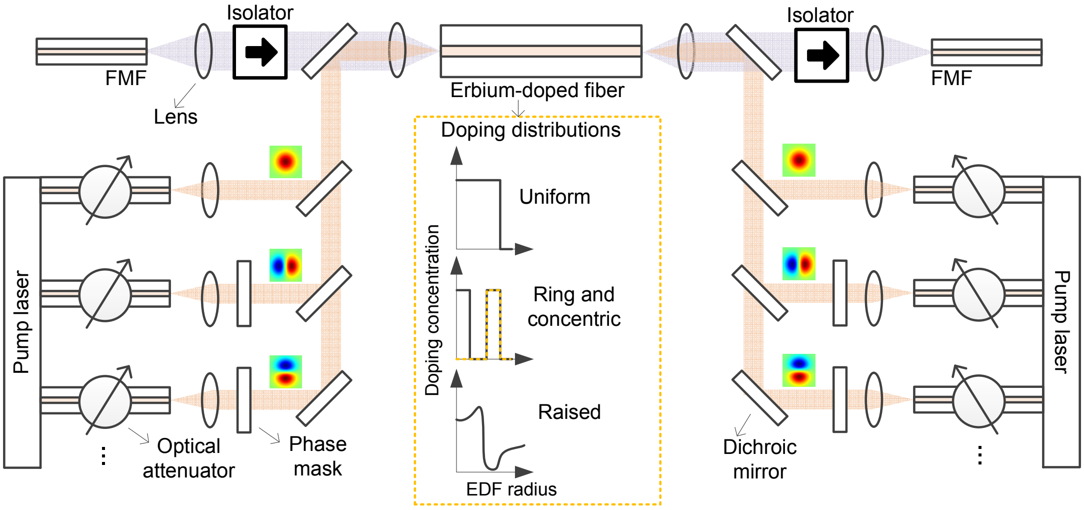


Figure 3.18.: Structure of the few-mode EDFA equalizing the gain per mode with co- and counter-propagating pumps (adapted from the proposal in [93]).

Experimental studies also apply concepts of regulation of the modal content and dopant distribution for gain equalization. The authors in [67, 97] investigated the raised and ring doped profile combined with an offset launch of the pump laser, which allows the excitation of higher order modes. By using the offset launch technique of the pump laser the modal content of the pump cannot precisely be controlled as in the phase plates approach of Fig. 3.18, but avoids a complicated set-up and the high amount of power being lost in splitting the pump laser and using phase plates. In [67,97] the authors experimentally demonstrated an MDG of approximately 2 dB with optimum pump and signal powers. Similar studies were also performed by the authors in [98], where they employed in their three-mode amplifier an erbium-doped fiber with a ring profile achieving similar values of MDG.

Amplifiers supporting six spatial modes have also been investigated in [65, 66, 99]. In [65] the authors studied the MDG of the amplifier by using a combination of 3.5 m of ring-doped fiber for enhancing the higher order modes gain and 1 m of uniformly doped fiber for enhancing the gain of LP_{01}. The authors achieved an MDG of approximately 4 dB. In [66,99] the reported MDG is slightly lower by using a LP_{02} pump and in a bi-directional manner to allow for a gain of more than 20 dB. In [100] it is predicted that the more the propagating modes are, the more difficult it will be to equalize the gains in the amplification stage. This is due to the fact that higher order modes have similar amplitude profile in the radial direction, therefore pumping in a certain mode will simultaneously amplify more modes making MDG difficult to control.

Few-mode Raman amplification has also been studied in [101, 102], where modal content of the pump plays also an important role. In [102] the Raman amplification provides about 1 dB of MDG, although at relatively low gain for long-haul applications, at about 8 dB. It is also found in [101, 102] that distributed amplification might be beneficial for averaging the MDG aided by the mode coupling.

3.3.3. MDM ROADM

Reconfigurable optical add/drop multiplexers (ROADMs) are key elements of today's long-haul and metro WDM networks. They allow to add or redirect channels at a network node without making use of costly electronic regeneration. Figure 3.19 shows a typical network node architecture interconnected by ROADMs for adding and dropping WDM channels. The wavelength selective switches (WSSs) shown in Fig. 3.19 are key optical devices in a ROADM. They can select WDM channels from a common input into different output ports, combine input WDM channels to one common output port or block certain WDM channels according to the network rules.

ROADMs can be remotely reconfigured, which is an attractive feature in a continuously growing network. Apart from WSSs, ROADMs are typically conformed of amplifiers and channel monitors. The amplifiers guarantee that output signals have the proper signal levels. The channel monitor plays an important role for channel equalization, since it is in charge of measuring the power of the WDM channels in a ROADM.

WSSs are the key component of a ROADM. Basically, WSSs consist of beam shaping and polarization diversity optics, a dispersion element and a switching engine [103]. Firstly, the beam passes through a beam shaping stage and a polarization splitter, since there are polarization-dependent components in the WSS. Then, a dispersive element separates the incoming beam into its different frequency components. This step is generally performed by a diffraction grating. The diffracted light is projected spatially onto a switching engine, which does the beam steering, i.e. the redirection of light to the required output port. There are two main technologies for beam steering: micro-electromechanichal system (MEMS) mirror arrays and LCoS. MEMS mirror arrays redirect the optical beam by rotating the mirrors, whereas the LCoS

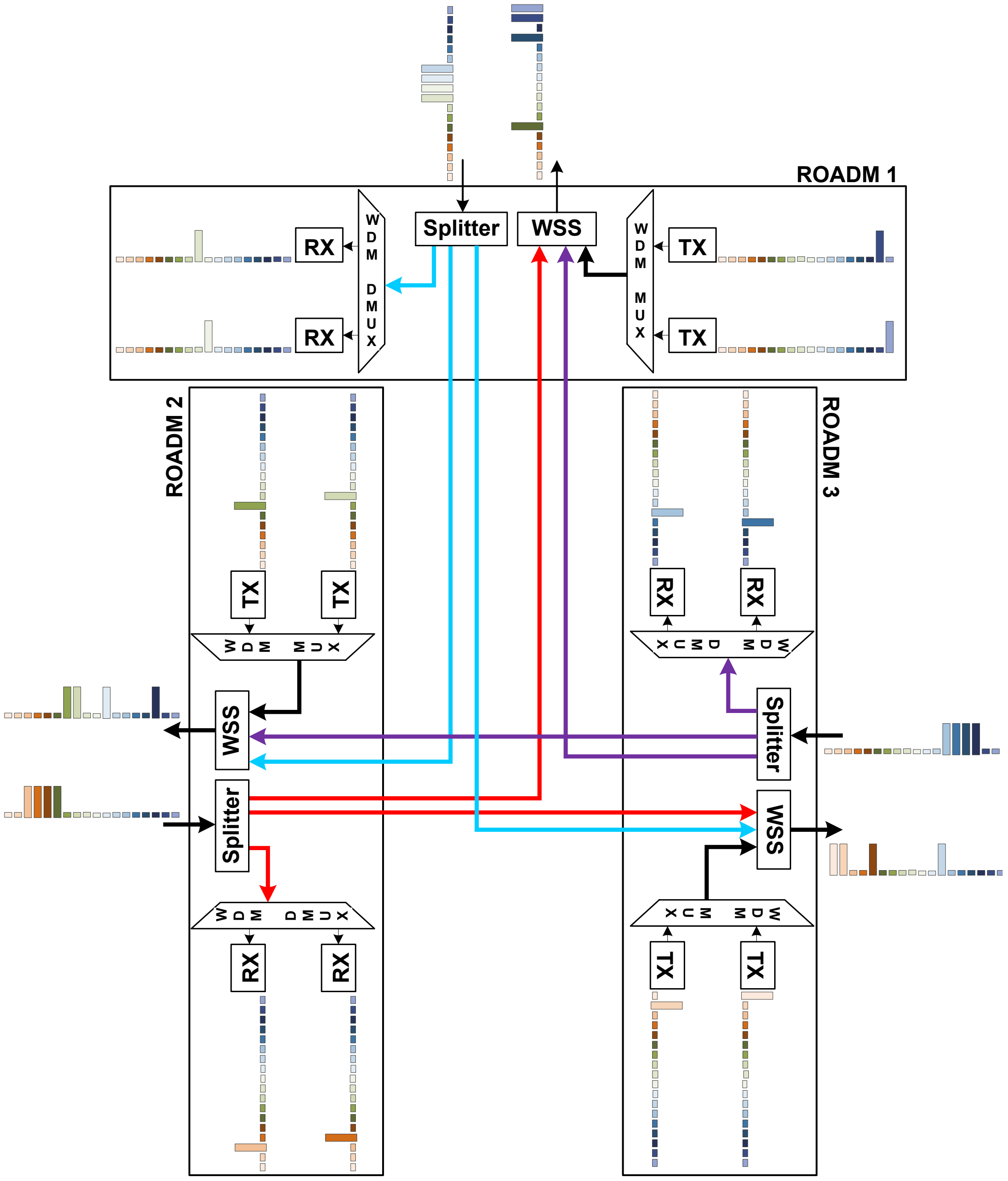


Figure 3.19.: Three interconnected ROADMs.

controls the phase of the light pixelwise. While MEMS technology counts with faster switching than LCoS technology, LCoS offers better granularity to address dense WDM channel spacing and flexible WDM channel grid [103, 104]. The beams are then recombined again by the dispersive element, passed through the beam shaping and polarization diversity stage and finally, reach the corresponding output port.

Studies have been performed to adapt single-mode to few-mode WSSs. The task of this device would be to switch spatial superchannels, i.e. redirect and combine information on wavelengths and modes simultaneously. As it would be costly to demultiplex at the same time modes and wavelengths, since this requires optical to electrical conversion and MIMO digital signal processing, the work that have been proposed on WSSs addresses the switching of information fiber wise, i.e. switching the information on all the modes simultaneously.

The simplest way to accomplish a few-mode switching device would be to replace the input SMFs by FMFs as in [105]. The idea here is to use for N spatial modes the same single-mode optics to steer, which leads to cost savings compared to separately switching N independent data streams. Similar to the case of FMAs, the implementation of few-mode ROADMs should result in more cost-efficient devices than the equivalent single-mode ROADM network. However, the disadvantage of this approach is that the modes have different passband characteristics. This is due to the fact that each mode has different spatial power distributions occupying a different portion of space on the switching engine [106–108]. A consequence of this negative aspect is the poor performance of the switch at the edges of the passband spectrum, which also lead to MDL and poor spectral efficiency of the device, i.e. WDM channels require larger guard bands than when using single-mode WSSs.

A promising approach for a few-mode switch that still has the advantage of using the same single-mode optics and, additionally, does not compromise its spectral filtering capabilities is proposed by [106–108]. In order to overcome the mode-dependent frequency responses, it is proposed in [106, 107] the use of photonic lanterns at the input and output ports to demultiplex and multiplex the beams. With this solution the incoming beams would have all Gaussian shapes as the fundamental mode, which enables all the beams to be treated in the switching engine evenly. MDL values for this device ranging from 3 dB to 4.5 dB are reported in [106]. In [108], a similar proposal is made by employing instead of photonic lanterns, 3DWs.

3.4. Optical RX

After transmission over a certain distance and demultiplexing N optical beams, where N coincides with the number of spatial modes, it is necessary to transform the optical signal of each beam to an electrical signal in order to decode its information bits. The photodiode is the transducer in charge of this task. Photodiodes convert the optical intensity into electrical current following the relation in Eq. (3.22), where

$I_{\mathrm{pho}_{\mathrm{out}}}$, R_{pho}, η_{pho}, e, $P_{\mathrm{pho}_{\mathrm{in}}}$ and $f_{\mathrm{pho}_{\mathrm{in}}}$ correspond to the output electrical current in [A], the responsivity in [A/W], the quantum efficiency (average number of electrons generated per photon by the photodiode), charge of an electron (1.6×10^{-19} C), input power to the photodiode in [W], frequency of the incident photons in [Hz], respectively [26, 60]. The shot and thermal noise of the photodetector has been ignored in Eq. (3.22) since the dominant source of noise in long-haul systems is mainly composed of ASE noise from the optical amplifiers.

$$I_{\mathrm{pho}_{\mathrm{out}}} = R_{\mathrm{pho}} P_{\mathrm{pho}_{\mathrm{in}}} = \frac{\eta_{\mathrm{pho}} e P_{\mathrm{pho}_{\mathrm{in}}}}{h f_{\mathrm{pho}_{\mathrm{in}}}} \tag{3.22}$$

As evident in Eq. (3.22), the photodetector senses the strength of the incident optical power. Therefore, to recover the imaginary and real part on both polarizations it is necessary to do a down-conversion similar to the inverse function of the diagram in Fig. 3.3, nevertheless, using an analog-to-digital converter (ADC). The so-called coherent detection is the current technology to detect not only the power but the full optical field including amplitude, phase and polarization information of the signals. In Fig. 3.20 the structure of the coherent RX is depicted, where E_{LO}, $\mathbf{E}_{\mathrm{r}}$ and $E_{\mathrm{r}_{\mathrm{x/y}}}$ correspond to the electrical field of the LO, received electrical field after the optical link and demultiplexing and its two orthogonal polarizations components, respectively; LO and TIA stand for local oscillator and transimpedance amplifier, respectively. The TIA does the conversion from electrical current to voltage to properly feed the ADC.

The data and the LO signal are first mixed with a 2 × 4 90° optical hybrid, whose structure is illustrated in Fig. 3.22. Taking into account that the input and output relationship of a single 3-dB coupler is given by

$$\begin{pmatrix} out_1 \\ out_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} in_1 \\ in_2 \end{pmatrix} \text{, as illustrated in Fig. 3.21.}$$

From the schematic in Fig. 3.22, it can be deduced that the input and output electrical fields follow the relation

$$\begin{pmatrix} E_1 \\ E_2 \\ E_3 \\ E_4 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & j \\ 1 & -j \end{pmatrix} \begin{pmatrix} E_{\mathrm{r}_{\mathrm{x}}} \\ E_{\mathrm{LO}} \end{pmatrix} . \tag{3.23}$$

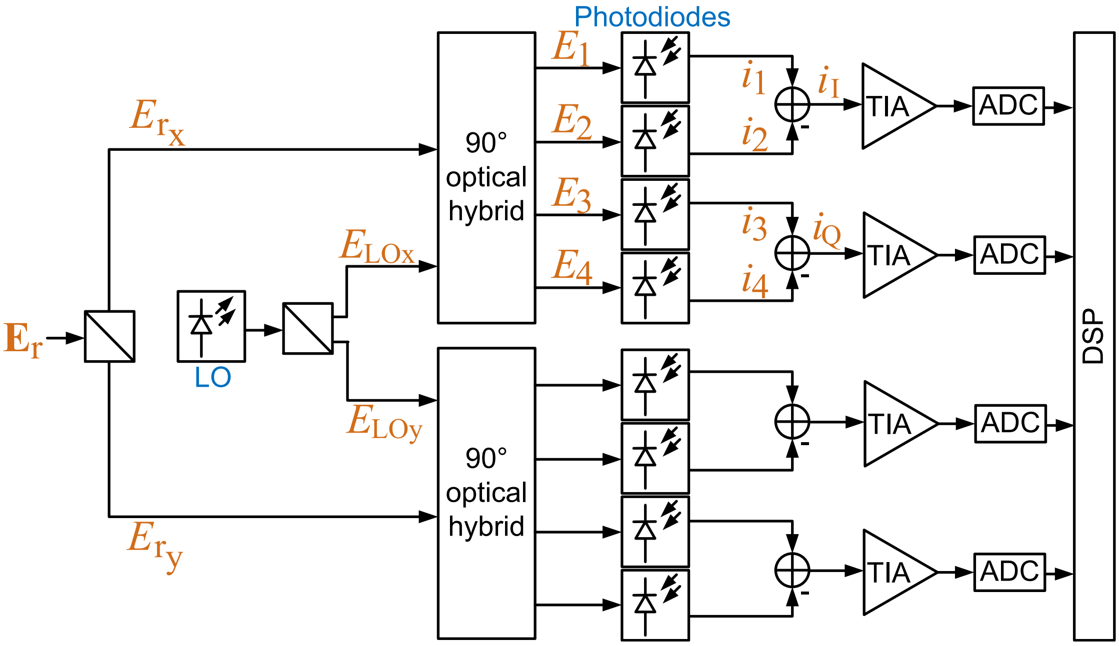


Figure 3.20.: Structure of polarization diversity coherent RX.

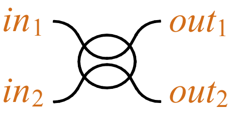


Figure 3.21.: 3-dB coupler.

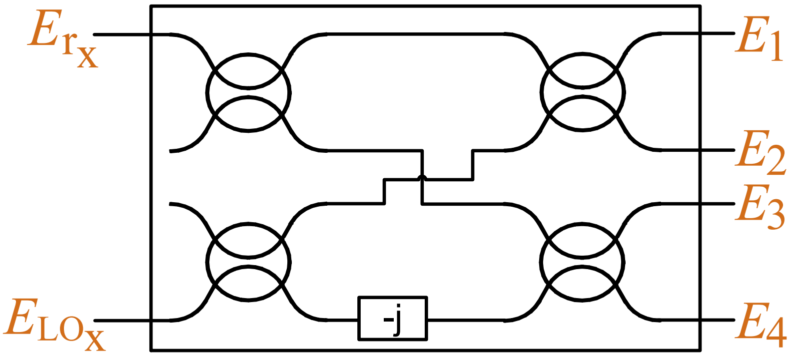


Figure 3.22.: 2 × 4 90° optical hybrid.

With these considerations and being $E_{r_x} = s(t)e^{j(\omega_c t+\phi_s(t))}$ and $E_{LO} = a_{LO}(t)e^{j(\omega_{LO} t)}$, then i_1 and i_2 from Fig. 3.20 correspond to Eqs. (3.24a) to (3.24d). $\phi_s(t)$ is the phase of the data signal relative to the LO phase. The following photodiode current equations are written in terms of the incoming electrical fields, since their square absolute value is proportional to the incoming optical power [1, 26, 61].

$$i_1 \propto |E_1|^2 = \left|\frac{E_{\mathrm{r_x}}}{2} + \frac{E_{\mathrm{LO}}}{2}\right|^2 \tag{3.24a}$$

$$= \frac{1}{4}\left[s(t)^2 + a_{\mathrm{LO}}(t)^2 + 2s(t)a_{\mathrm{LO}}(t)\cos\left(\left(\omega_{\mathrm{c}} - \omega_{\mathrm{LO}}\right)t + \phi_{\mathrm{s}}(t)\right)\right] \tag{3.24b}$$

$$i_2 \propto |E_2|^2 = \left|\frac{E_{\mathrm{r_x}}}{2} - \frac{E_{\mathrm{LO}}}{2}\right|^2 \tag{3.24c}$$

$$= \frac{1}{4}\left[s(t)^2 + a_{\mathrm{LO}}(t)^2 - 2s(t)a_{\mathrm{LO}}(t)\cos\left(\left(\omega_{\mathrm{c}} - \omega_{\mathrm{LO}}\right)t + \phi_{\mathrm{s}}(t)\right)\right] \tag{3.24d}$$

In the same way, i_3 and i_4 can be derived leading to Eqs. (3.25a) and (3.25b).

$$i_3 \propto \frac{1}{4}\left[s(t)^2 + a_{\mathrm{LO}}(t)^2 + 2s(t)a_{\mathrm{LO}}(t)\cos\left(\left(\omega_{\mathrm{c}} - \omega_{\mathrm{LO}}\right)t + \phi_{\mathrm{s}}(t) - \pi/2\right)\right] \tag{3.25a}$$

$$i_4 \propto \frac{1}{4}\left[s(t)^2 + a_{\mathrm{LO}}(t)^2 - 2s(t)a_{\mathrm{LO}}(t)\cos\left(\left(\omega_{\mathrm{c}} - \omega_{\mathrm{LO}}\right)t + \phi_{\mathrm{s}}(t) - \pi/2\right)\right] \tag{3.25b}$$

Consequently, the in-phase i_{I} and quadrature i_{Q} components of one of the polarizations of the signal are represented by Eqs. (3.26a) and (3.26b), where $\omega_{\mathrm{IF}} = \omega_{\mathrm{c}} - \omega_{\mathrm{LO}}$ is called the intermediate angular frequency. The components for the other polarization can be derived similarly. These currents are translated into voltage by the TIAs, then the signals are sampled by the ADC and the DSP is finally in charge of compensating the optical link impairments and recovering the transmitted bits.

$$i_{\mathrm{I}}(t) = i_1 - i_2 \propto s(t)a_{\mathrm{LO}}(t)\cos\left(\omega_{\mathrm{IF}}t + \phi_{\mathrm{s}}(t)\right) \tag{3.26a}$$

$$i_{\mathrm{Q}}(t) = i_3 - i_4 \propto s(t)a_{\mathrm{LO}}(t)\sin\left(\omega_{\mathrm{IF}}t + \phi_{\mathrm{s}}(t)\right) \tag{3.26b}$$

If the LO and the transmitter laser operate at the same frequency and phase, the RX is configured as homodyne. Whereas if ω_{IF} is greater than the bandwidth of the signal at the RX, the RX is called heterodyne. Intradyne RXs are those whose ω_{IF} is almost zero.

In the case of homodyne RXs, the RX is obligated to actively control any phase and frequency differences between TX and RX lasers [10]. With the heterodyne and intradyne RX such frequency and phase control is not mandatory.

The heterodyne RX has the following advantage: it is enough to detect the in-phase or quadrature components on each of its polarization to recover the entire field.

Although it is necessary to further down-convert the signal from the intermediate frequency to baseband after the analog-to-digital conversion. It should be also considered that for WDM transmission the channel under test (CUT) should be optically filtered out in order to avoid crosstalk from other neighboring WDM channels. Alternatively, the CUT can be down-converted to ω_{IF} so that its spectrum lies in-between the image signal spectra from neighboring WDM channels, subsequently, the CUT can be filtered out in the electrical domain by the DSP.

Currently, the type of RX that is mostly used for experiments and commercial products is the intradyne RX, since neither phase and frequency of the LO has to be controlled nor the optical and electrical components bandwidth requirements are as stringent as in the heterodyne RX case.

3.5. DSP and modulation formats

As shown in Fig. 3.3, it is customary to modulate signals in quadrature (data Q) and in-phase (data I) and combine them as a pair of orthogonal signals. IQ modulation exploits the quadrature physical dimension allowing for higher data rates. After their optical detection and the sampling process carried out by the ADC, the DSP takes care of recovering the distorted data by the compensation of the transmission link impairments such as dispersion, modal delay, mode coupling and PMD. The DSP structure highly depends on the modulation format chosen for transmission. Single-carrier and orthogonal-frequency division multiplexing (OFDM) are two schemes that are being investigated for MDM systems and are described in the following subsections.

3.5.1. Single-carrier modulation

Single-carrier modulation schemes for single mode and FMF channel are similar since they have the same structure. Single-carrier DSP for MDM can be upgraded as depicted in Fig. 3.23, where N_{m} represents the number of spatial modes. The TX has the same arrangement of functionality blocks as single-mode fiber with POLMUX signals. Also as in single-mode systems, after optical-to-electrical conversion and sampling at at least the Nyquist rate, the discrete received electrical signals $r_i[k]$ of each tributary, with i ranging from one to N_{t} and k being the discrete time variable, passes through the first equalization stage, clock recovery, second equalization stage, carrier recovery, detection and demapping. The number of tributaries N_{t} refers to the number of sent data streams and corresponds to two times (two polarizations) the supported spatial modes N_m.

The task of the first equalization stage is to compensate for intra-modal dispersion. Since this distortion does not induce crosstalk between the modes, it can be compensated per tributary. In contrast, the combined effect of PMD, mode coupling and DMD induces crosstalk between the data streams and therefore requires an equalization where are all received tributaries participate. As the intra-modal dispersion can be represented as a linear filter, the inverse response of this filter can be applied to undo its effect. As mentioned in Section 2.3.2, the intra-modal dispersion and its slope are related to $\beta_{2,lm}$ and $\beta_{3,lm}$, which can be derived from Eqs. (2.32) and (2.35d), leading to Eqs. (3.27a) and (3.27b) for the reference wavelength λ_c. The compensation of different dispersion values individually for each mode is not possible due to the mode mixing caused by mode coupling. Nevertheless, as intra-modal dispersion is very similar among modes, most of the dispersion can be compensated with the average dispersion and slope, and the posterior MIMO equalization will compensate for the remaining dispersion. Hence, the transfer function of the average intra-modal dispersion is the all-pass transfer function given by Eq. (3.28) for a fiber of length L. In Eq. (3.28), $\bar{\beta}_2$, $\bar{\beta}_3$, $\bar{D}$ and $\bar{S_\mathrm{lo}}$ are the mode-averaged second and third derivative of the propagation constant with respect to the angular frequency and, mode-averaged dispersion and slope parameter, respectively.

$$\beta_{2,lm} = \frac{-\lambda_\mathrm{c} D_{lm}}{2\pi c} \tag{3.27a}$$

$$\beta_{3,lm} = \frac{\lambda_\mathrm{c}^2}{(2\pi c)^2} \left(\lambda_\mathrm{c}^2 S_{\mathrm{lo},lm} + 2\lambda_\mathrm{c} D_{lm}\right) \tag{3.27b}$$

$$H_{\bar{\mathrm{D}}}(f) = e^{\frac{1}{2}j(2\pi f)^2 \bar{\beta}_2 L + \frac{1}{6}j(2\pi f)^3 \bar{\beta}_3 L} = e^{-j\frac{\pi\lambda_\mathrm{c}^2 \bar{D} f^2 L}{c} + j\frac{\pi\lambda_\mathrm{c}^2 f^3 L}{3c^2}\left(\lambda_\mathrm{c}^2 \bar{S_\mathrm{lo}} + 2\lambda_\mathrm{c}\bar{D}\right)} \tag{3.28}$$

The average dispersion compensation is thus done by multiplying the incoming signal in frequency domain by the conjugate of $H_{\bar{\mathrm{D}}}(f)$. The techniques overlap-add or overlap-save aid to do this operation blockwise in order to reduce the processing time [109]. Dispersion compensation is also possible in time domain with finite impulse response (FIR) filters or infinite impulse response (IIR) filters, being the latter less common [110]. However, time domain compensation adds a tremendous amount of complexity to the signal processing and therefore the compensation is mostly done in frequency domain [57, 110].

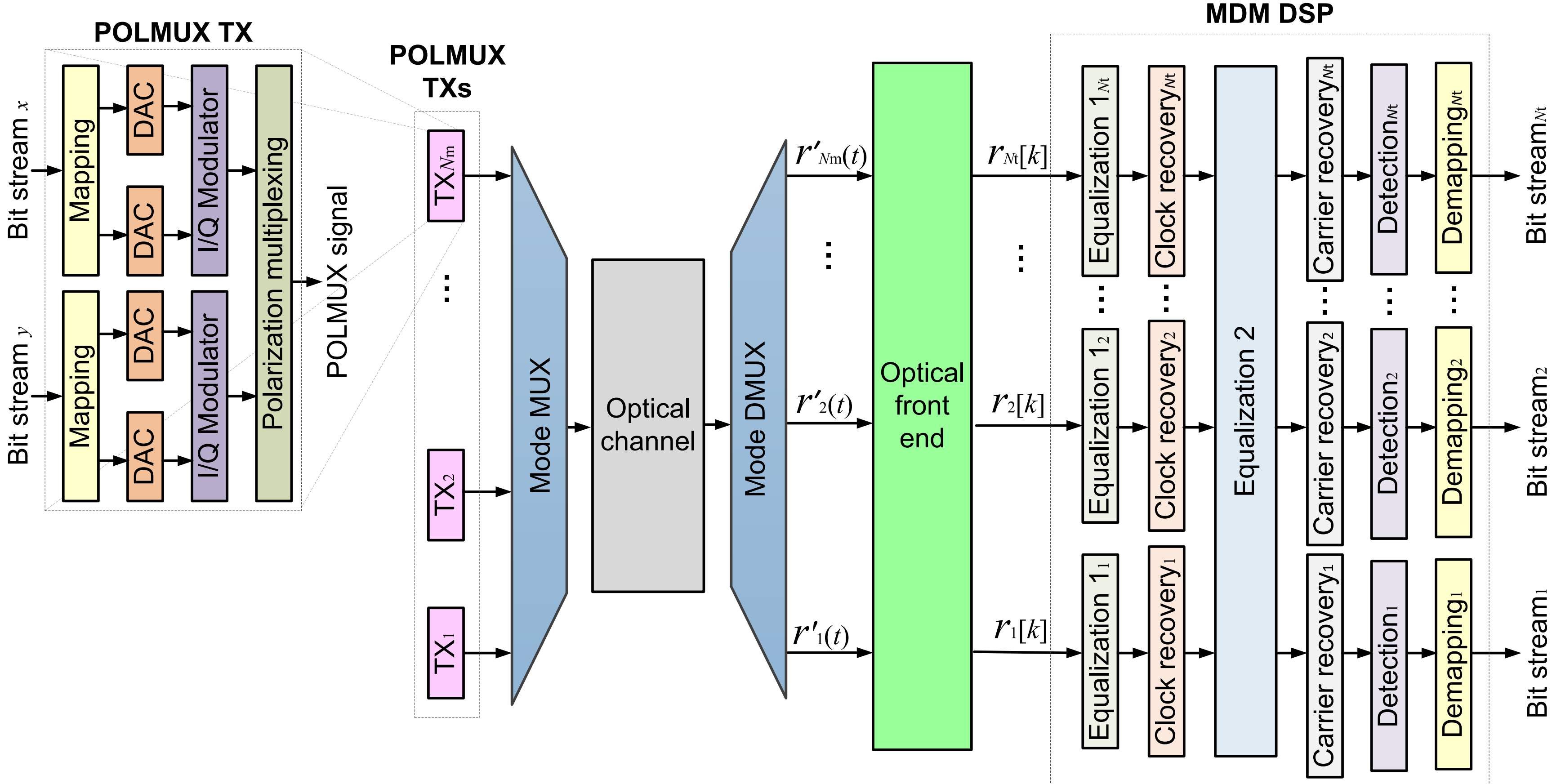


Figure 3.23.: Single carrier DSP structure.

As illustrated in Fig. 3.23, after compensating most of the intra-modal dispersion it follows a clock recovery stage, also commonly called timing recovery. This stage is necessary since the TX and RX clocks are not synchronized, leading to sampling frequency offset and phase mismatch. Different techniques are compiled in [111]. Note that at this stage the signals are still distorted by the channel effects and for this reason the timing recovery has to be designed so that it is tolerant towards them. The approach of doing the timing recovery before MIMO equalization is usually done to enable convergence stability in a posterior blind equalization. Experiments reporting first timing recovery and then MIMO equalization can be found in [69, 75, 112]. Alternatively, the timing recovery can be done after a data-aided MIMO equalization stage as in the DSP scheme presented in [113].

Equalization number two illustrated in Fig. 3.23 is in charge of the MIMO equalization to unravel the rest of linear channel impairments such as modal crosstalk, PMD, DMD and, if there is any, residual intra-modal dispersion. In order to find the filter parameters channel estimation is required; it can be data-aided via training symbols, blindly via algorithms such as the least-mean square (LMS) algorithm and constant-modulus algorithm or a combination of both: first data-aided channel estimation to guarantee convergence and then adaptively changing the filter coefficients according to the channel changes. The latter option is the most common variant among the reported experiments, whose filter adaptation mechanism is based on the decision-directed LMS algorithm [69, 75, 114–116]. Thus, on the basis of the channel estimation, the compensating filter can be implemented as a butterfly equalizer with N_{t}^2 FIR filters in the time or frequency domain. The structure of this filter bank is depicted in Fig. 3.24 for six tributaries, i.e. a fiber that supports three spatial modes.

The MIMO equalization depicted in Fig. 3.24 is an extension of the 2×2 MIMO equalization for single-mode fibers. The difference is that in the case of FMFs the channel memory is significantly greater than in the single-mode case due to the combined effect of the DMD and mode coupling. The number of taps that have to be used in the time domain equalization is directly proportional to the delay spread caused by the channel. Current transmission experiments show that DMD of less than 100 ps/km is possible [69, 115, 117]; additionally, the delay spread due to DMD can be reduced by using FMF with positive and negative DMD parameters [69,117,118] and inducing mode coupling [119]. However, the DMD value is very sensible to small changes or imperfection in the FMF profile fabrication process [120]. Reducing the delay spread will also reduce the complexity of time domain equalization. Nevertheless, complexity studies report that even with a DMD of 27 ps/km,

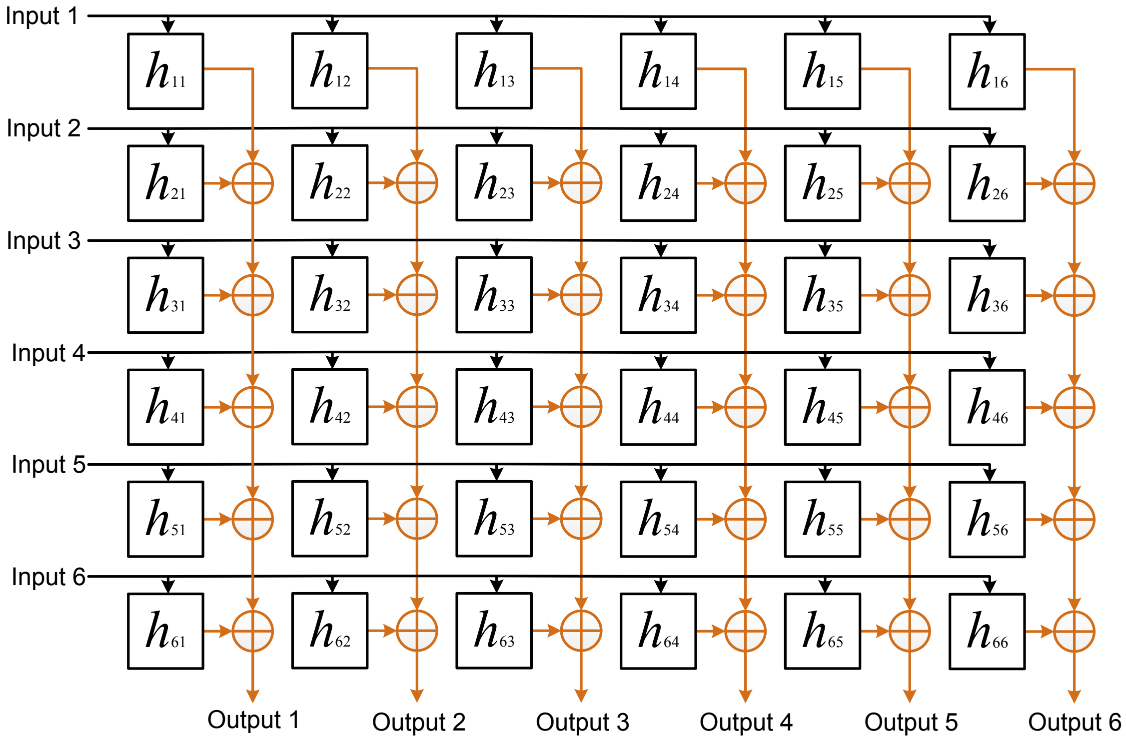


Figure 3.24.: Structure of the second equalizer for $N_\mathrm{t} = 6$.

time domain equalization have a complexity that surpasses the frequency domain equalization by several orders of magnitude [57]. The same result is also reported for single-mode systems [110]. The lower complexity of frequency domain equalization comes at the cost of higher overhead due to a cyclic prefix (CP) to enable circular convolution of the incoming signal and the channel response, similar to OFDM (see Section 3.5.2).

The last stages of the single-carrier DSP are the carrier recovery, detection and demapping. The carrier recovery tracks the frequency and phase differences between the TX laser and the LO via algorithms used also in single-mode systems as summed up in [111]. The detection stage assigns the equalized symbols to a symbol in the constellation diagram and the demapping stage translates symbols into bits.

3.5.2. OFDM

OFDM is a multicarrier modulation format, which was first introduced in 1966 by Chang in [121] and appeared with its current name for wireless transmission in the 70's [62]. Nowadays, OFDM is the basis for the wireless standards for digital audio/video broadcasting, wireless local area network (IEEE 802.11a/g), wireless metropolitan area networks (IEEE 802.16e), asymmetric digital subscriber line (ITU G.992.1) and the forth generation of mobile communications, long-term evolution [62, 122]. OFDM is an attractive modulation format since it is:

- Robust against the delay spread, in wireless caused by fading and in the optical communication caused by the dispersion, PMD and DMD.
- Spectrally efficient since OFDM is characterized by a virtually rectangular spectrum without the need of filtering. In optical systems, this allows to closely pack the WDM channels to transmit more information per unit bandwidth.
- Computationally less complex than time domain equalization in single-carrier due to the use of the efficient implementation of fast Fourier transformation. Its complexity is similar to single-carrier frequency domain equalization [57].
- Flexible considering that OFDM allows the manipulation of the information on each subcarrier according to the systems needs.

On the other hand, the relatively high peak-to-average power ratio (PAPR) and its sensitivity to phase noise, frequency and phase offset belong to its well-know drawbacks [62, 123, 124].

The principle of OFDM is to transmit information at a certain data rate via simultaneous modulation of lower rate data streams. Therefore, each data stream occupies a small portion of the entire spectrum. Figure 3.25 shows this principle, where the data streams are upconverted to different frequencies, multiplexed and, at the RX, demultiplexed by downconverting the signal with mixers of opposite sign as in the TX and integrated within an OFDM symbol duration. Mathematically, the OFDM signal represented in Fig. 3.25 corresponds to Eq. (3.29), where $s(t)$ is the transmitted OFDM signal, $s_{i+N_{\mathrm{FFT}}/2}$ are the symbols modulated in each subcarrier, T is the OFDM symbol duration, t_{s} is the starting time of the OFDM symbol and N_{FFT} represents for the moment the number of subcarriers. The received OFDM signal is represented by Eq. (3.30) for recovering the symbol $s_{w+N_{\mathrm{FFT}}/2}$ at the frequency w/T.

$$s(t) = \begin{cases} \sum\limits_{i=-\frac{N_{\mathrm{FFT}}}{2}}^{\frac{N_{\mathrm{FFT}}}{2}-1} s_{i+N_{\mathrm{FFT}}/2} e^{j2\pi \frac{i}{T}(t-t_{\mathrm{s}})} & t_{\mathrm{s}} \le t \le T + t_{\mathrm{s}} \\ 0 & t < t_{\mathrm{s}} \land t > t_{\mathrm{s}} + T \end{cases} \tag{3.29}$$

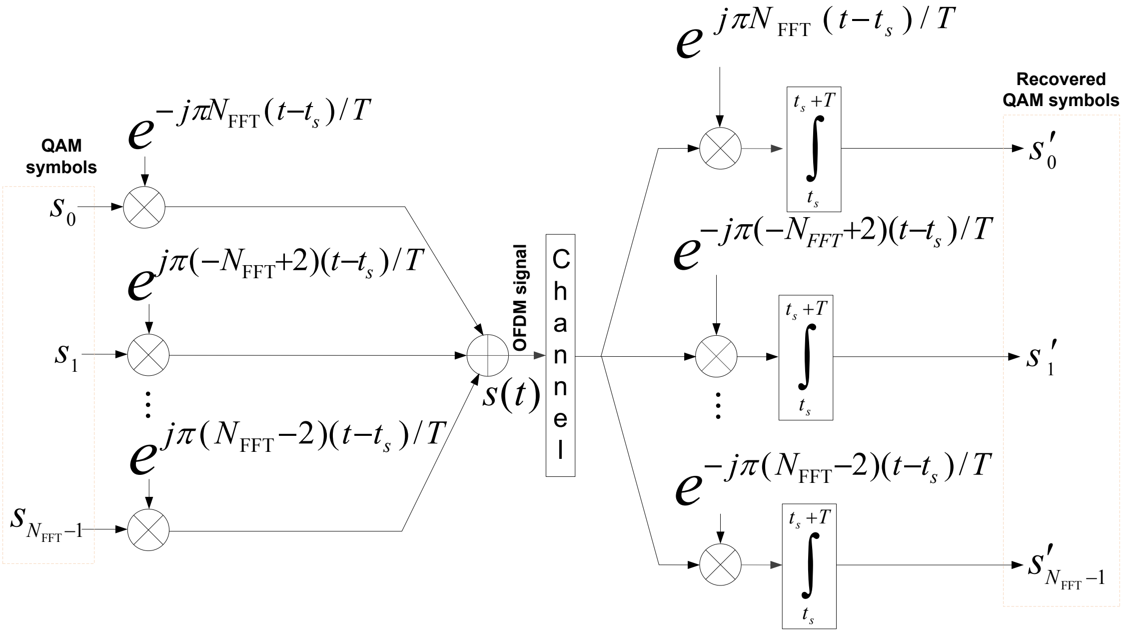


Figure 3.25.: Principle of OFDM.

$$\begin{aligned}
&\int_{t_\mathrm{s}}^{t_\mathrm{s}+T} e^{-j2\pi\frac{j}{T}(t-t_\mathrm{s})} \sum_{i=-\frac{N_\mathrm{FFT}}{2}}^{\frac{N_\mathrm{FFT}}{2}-1} s_{i+N_\mathrm{FFT}/2} e^{j2\pi\frac{i}{T}(t-t_\mathrm{s})}\mathrm{d}t \\
&= \sum_{i=-\frac{N_\mathrm{FFT}}{2}}^{\frac{N_\mathrm{FFT}}{2}-1} s_{i+N_\mathrm{FFT}/2} \int_{t_\mathrm{s}}^{t_\mathrm{s}+T} e^{j2\pi\frac{i-w}{T}(t-t_\mathrm{s})}\mathrm{d}t \\
&= s_{j+N_\mathrm{FFT}/2}T
\end{aligned} \tag{3.30}$$

Equation 3.30 shows clearly the origin of the orthogonality in OFDM. The integral results in a constant factor T only if $i = w$; whereas for $i \neq w$ the integral is null, since the sinusoidal functions that are being integrated consist of integer number of cycles. The orthogonality can be also be seen in the frequency domain. The multiplication or modulation at the TX illustrated in Fig. 3.25 between the symbol $s_{i+N_\mathrm{FFT}/2}$ for a period of time T and the exponentials corresponds in the frequency domain to the convolution between sinc functions of the form $\mathrm{sinc}(\pi f T)$ as shown graphically in Fig. 3.26. It can be observed in Fig. 3.26, for example, the maximum value of subcarrier located at frequency f_1 lies where the amplitude of the spectra of other subcarriers centered at f_2 and f_3 are zero. This is due to the fact that the subcarriers are separated in the frequency domain from each other by a multiple of the inverse of the symbol duration, just when the sinc functions crosses the frequency axis by zero.

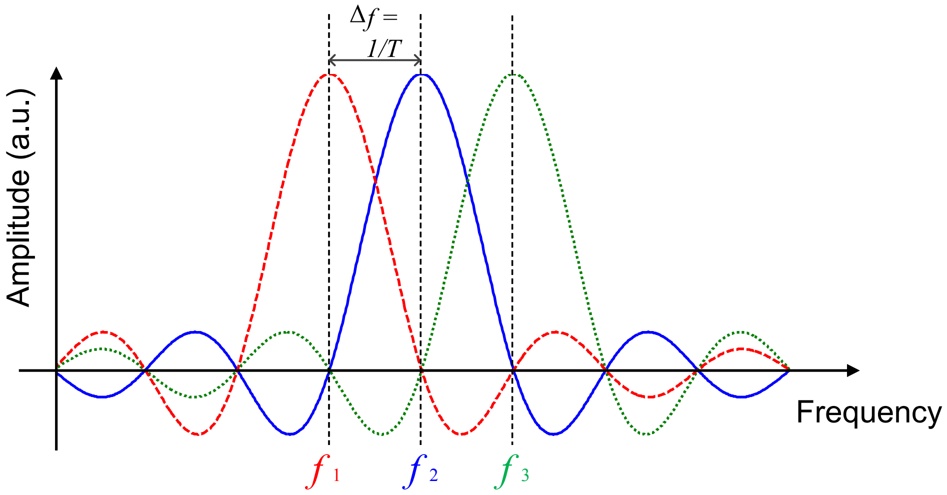


Figure 3.26.: Transmitted OFDM signal for three subcarriers.

An efficient way of realizing the scheme in Fig. 3.25 was found by Weinstein and Ebert [125] without the use of numerous oscillators and filters. Weinstein and Ebert proposed the used of the inverse discrete Fourier transform (IDFT) at the TX side and the discrete Fourier transform at the RX side (DFT). By incorporating the IDFT in Eq. (3.29), it results in the discrete form of the transmitted signals

$$s\left(t_\mathrm{s} + \frac{kT}{N_\mathrm{FFT}}\right) = s[k] = \sum_{i=-\frac{N_\mathrm{FFT}}{2}}^{\frac{N_\mathrm{FFT}}{2}-1} s_{i+N_\mathrm{FFT}/2}\, e^{j2\pi\frac{ik}{N_\mathrm{FFT}}} \ , \tag{3.31}$$

where k is the discrete time variable.

The application of the (I)DFT has nowadays evolved to the use of the more efficient implementation: the fast Fourier transform (FFT) and its inverse (IFFT). While the (I)DFT requires N^2 complex multiplications, the (I)FFT requires only $N/2\ \log_2(N)$ [62, 123].

A mechanism used in OFDM to combat the ISI is the addition of a CP. The CP consists of N_CP samples having a duration of T_CP and corresponds to a copy of the last part of the OFDM symbol as indicated in Fig. 3.27 with three unmodulated subcarriers. In the presence of delay spread, the CP enables the OFDM symbols to be easily separable in time windows of N_FFT samples, so that each recovered OFDM symbol preserves orthogonality, i.e. each subcarrier still has an integer number cycles. The delay spread originates from the dispersion, making the different frequency components of the signal to have different delays, and DMD making the signals on different modes to have different propagation delays. Note that CP should be at least the delay spread duration. Thus, by removing the ISI, the symbols are separated from each other therefore easing the posterior equalization.

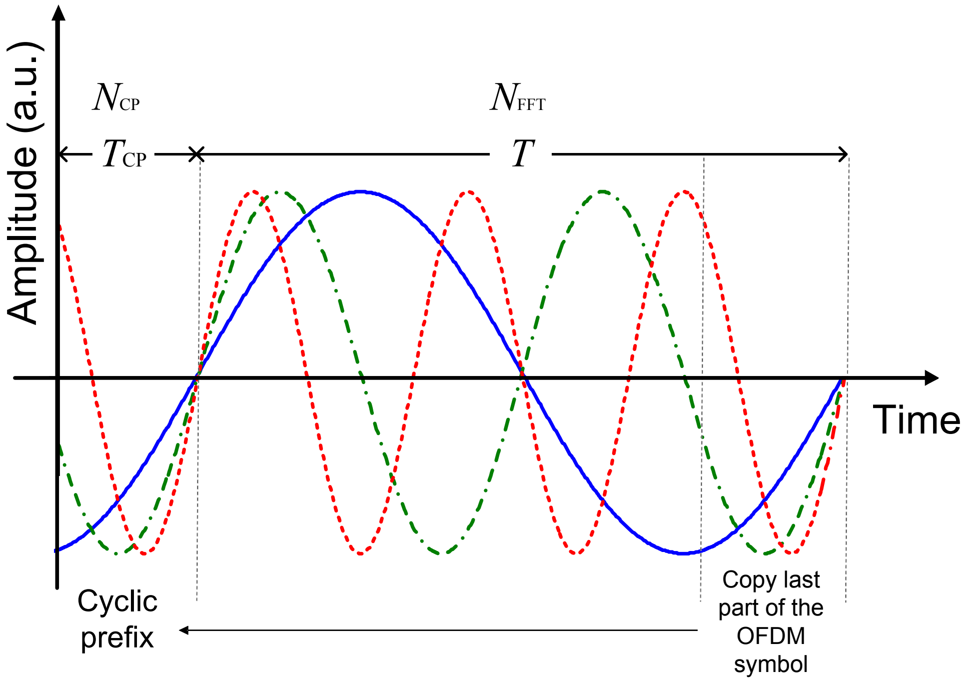


Figure 3.27.: Exemplary illustration of CP incorporation to the OFDM symbol.

The typical OFDM block diagram is represented in Fig. 3.28, which has been modified to support multiple modes. The transmitter is composed of POLMUX TXs, which perform the IFFT and insert the CP and training symbols (TSs). TSs are sent periodically for channel estimation and frame synchronization. Before the IFFT is executed, zeros are inserted at the edges of the OFDM spectrum in order to isolate aliasing products originated by the DAC from the signal spectrum, so these products can be easily filtered out. Appending these zeros, N_{ZP} zeros, to the spectrum is called zero padding (ZP).

At the RX after the signal is downconverted and sampled, there is an optional equalizer 1. Equalizer 1 is optional since it compensates for dispersion electrically in order to reduce the CP overhead. Subsequently, frame synchronization is performed. There are several ways to accomplish synchronization such as with the technique introduced by Schmidl and Cox [126] or by correlating known sequences with received ones [127]. Then, the CP is removed, the FFT is applied and the ZP is removed. Equalization 2 is mostly done in the optical communication with the one-tap equalization: zero-forcing (ZF) [128]. More equalizers are investigated in Chapter 5. Afterwards, symbols are detected and demapped as indicated in Fig. 3.28.

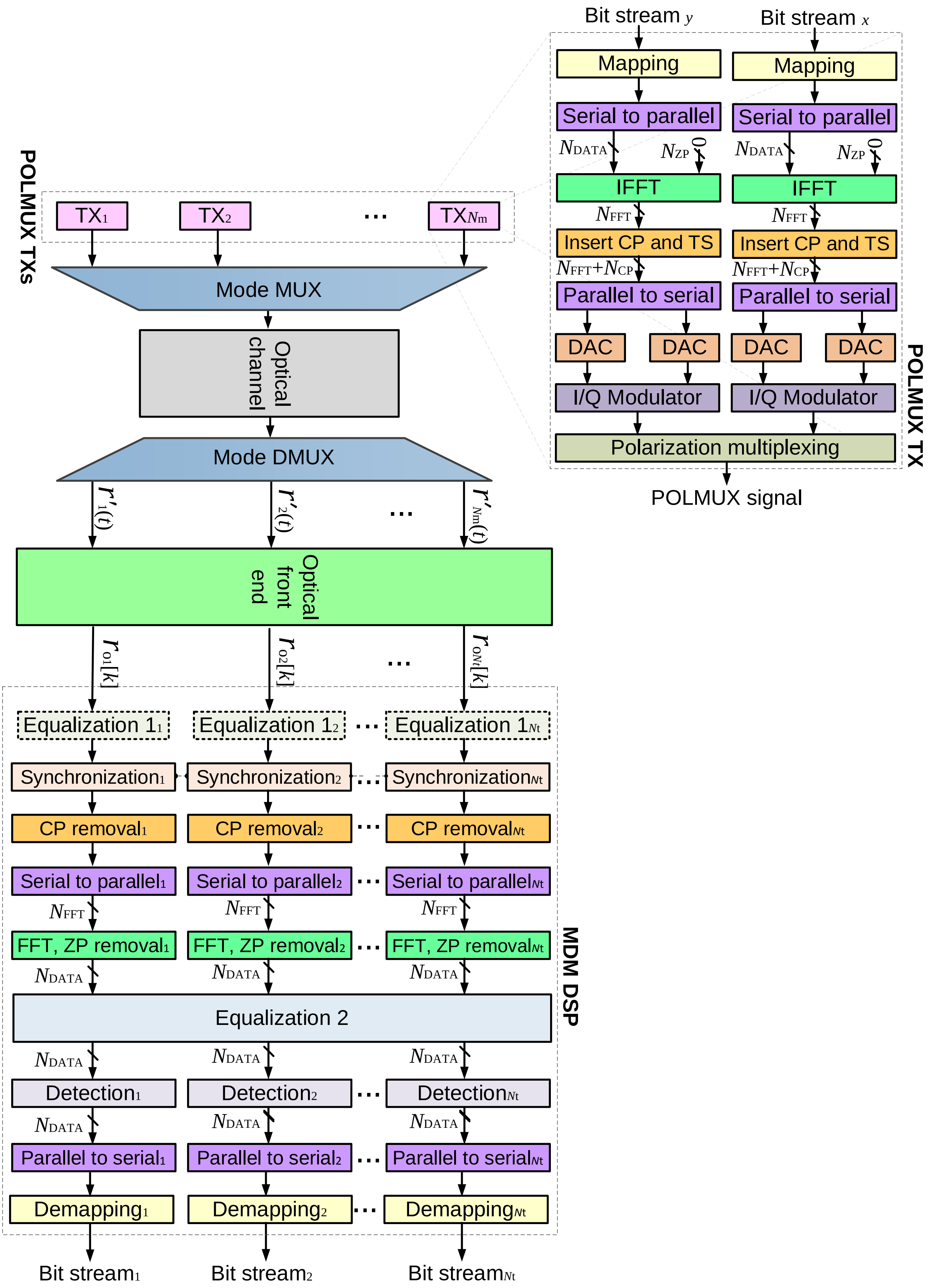


Figure 3.28.: OFDM block diagram adapted for MDM.

OFDM appears in the optical communication at the beginning of the 2000s and now is being considered as a possible candidate for the modulation of MDM systems [81, 117, 129, 130]. It has already been shown that equalization in the frequency domain is computationally less complex than time-domain equalization for single-mode fibers [110]. A similar study was presented in [57], in which it is shown that the complexity difference between these type of equalization schemes is even larger by several orders of magnitude if DMD is considered. However, the use of OFDM and frequency domain equalization have to be traded off with the overhead inserted by the CP and TSs.

As mentioned before OFDM suffers from relatively high PAPR ratios, which can be a disadvantage when transmitting through dispersion-managed links [131]. However, inline dispersion compensation is not the better option for modern coherent optical system since it enhances nonlinear effects [132]. High PAPR is rather harmful at the DAC and ADC stages, where most of the signal (close to the average) is quantized only within the middle of the quantization range and therefore the OFDM signal would exhibit a reduced SNR due to quantization noise [62]. Nevertheless, there are methods for reducing the PAPR such us clipping and coding [123]. Furthermore, OFDM needs a proper frequency and phase noise compensation since these effects generate distortion through intercarrier interference (ICI). Methods for correcting frequency offset and phase noise can be implemented using pilot tones [124,129], the method of Schmidl and Cox [126] or common phase error (CPE) correction [62,133].

Despite the fact the OFDM imposes challenges for its implementation, OFDM is the modulation format with the attractive characteristic of being able to be implemented in combination with per-subcarrier advanced detection schemes such as maximum-likelihood (ML) detection (see Chapter 5). For this reason, OFDM was preferred as the modulation scheme used in the simulations and experiments in this dissertation rather than single-carrier.

3.6. Forward error correction

To achieve virtually error-free transmission (BER smaller than 1e-13), FEC techniques have become popular and are nowadays essential in such systems. Basically, by adding redundant information, bits are encoded at the TX and at the RX the decoder exploits the redundant information to identify and correct errors. This leads to an FEC overhead, which is translated into an increase of the nominal data rate.

Apart from the FEC overhead, the FEC codes are characterized by the error correction capability measured by the FEC limit and coding gain, if the code is iterative, or not and if they are soft or hard-decision-capable. The FEC limit is the decoder input signal BER at which a certain target BER after decoding is achieved. Above the FEC limit a decoder output BER below the target BER cannot be guaranteed. The coding gain is the difference in OSNR between the uncoded and coded BER vs. OSNR curves at a certain target BER, which is typically lower than 1e-12. The net coding gain is a measure used to compare different codes with different overheads, since it takes into account the extra OSNR requirement due to the insertion of redundant information. A code uses iteration in the decoding to improve the error correction capability. It can be based on hard-decision or soft-decision. The input of hard decision decoders are symbols after detection. In contrast, soft decision decoders are capable to receive the demodulated symbols before detection allowing to collect more information from the symbols, i.e. how reliable is the input to the decoder.

Reed-Solomon (RS) codes are non-binary cyclic codes and were the first codes implemented for long-haul optical communication. Especially the code RS(255,239) has been included in the ITU recommendation ITU-T G.975 [134]. The nomenclature RS(N_{FEC},K_{FEC}) refers to the a code that consist of N_{FEC} symbols from which K_{FEC} are information symbols and $N_{\mathrm{FEC}} - K_{\mathrm{FEC}}$ are parity symbols. The RS(255,239) code provides around 5 dB of coding gain using 7 % of overhead and an FEC limit of 1e-4 for decoder output BER lower than 1e-15. Since the RS code is the code that was first implemented in commercial systems, it represents the first generation of FEC codes in optical communications.

The second generation of FEC codes are concatenated codes. The concatenated codes use different FEC code stages from the same code or different codes in order to reduce the implementation complexity and latency of the codes. They are often constituted of two RS or Bose-Chaudhuri-Hocquenhem (BCH) codes, or a combination of both [10, 135, 136]. The ITU recommendation ITU-T G.975.1 [136] specifies the concatenated codes, some of them also supporting soft decision schemes, for submarine and terrestrial optical systems, whose typical overhead is around 7 %, FEC limit is slightly above 1e-3 and net coding gain ranges between 7 and 9 dB. The net coding gain of concatenated codes can be improved by iterative decoding, which has the advantage of not requiring additional overhead to achieve better performance.

The third generation is characterized by soft-decision capable decoders. Mainly two classes of codes are introduced: turbo and Low-density parity check (LDPC) codes. Turbo codes were first introduced by Berrou et al. in 1993 [137] and applied in the optical communication by Mizuochi et al. in 2004 [138]. They have the advantage of being high parallelizable and have low error floors, however, targeting a smaller overhead requires large block lengths, which increases the latency [139]. Low-density parity check (LDPC) codes also became popular in the early 90's, although Gallanger proposed this kind of codes in the early 60's in [140]. LDPC codes are attractive, since they allow parallelization, shorter code words than the above mentioned codes and are flexible with respect to their implementation. Recently, a promising LDPC-based FEC code capable of nearly error free transmission (decoder output BER of 1e-15) at an input BER of 2.4e-2 using 20 % overhead has been proposed in [141]. This overhead value and FEC threshold is assumed for the experiments of this thesis.

4. Impact of mode coupling in MDL impaired systems

SDM has emerged as an attractive solution for overcoming the so-called *capacity crunch* by exploring the physical dimension of space of the fiber as mentioned in Chapter 1. As mentioned in Chapter 3, from the two SDM tendencies (multiple modes or multiples cores), FMF has gained considerable interest because of its advantage of carrying more data per unit area, its power efficiency in terms of optical amplification [20,55,56] and its higher tolerance against nonlinearities over MCF [22].

The implementation of MDM systems brings several challenges that need to be addressed for the realization of MDM commercial systems. One of the challenges is the computational complexity of the RX DSP because of the combined effect of DMD and mode coupling. The channel impulse response in this case is longer, which increases the computational complexity, i.e. more filter taps are needed in a time domain equalizer than for SMF systems. The RX DSP complexity can be reduced by frequency domain equalization or OFDM schemes by using a CP, however, a longer channel impulse response translates directly into a longer CP, i.e. larger CP overhead [57].

Modal dispersion is not a cause of fundamental performance degradation, but makes the RX DSP more complex or increases the transmission overhead. In turn, MDL (described in Section 3.2) fundamentally degrades the performance and therefore is another challenge that needs to be addressed. This negative effect originates from all kinds of inline optical components and the FMF itself, but mainly from the multimode amplification. Too much MDL can cause, in the extreme case, information from one or more modes to vanish, which is directly reflected in a reduction of the channel capacity. Recent studies show efforts on reducing the MDL from optical amplifiers (in the context of optical amplification referred as MDG) by using active fibers with specially tailored Erbium distribution and by modifying the modal content of the pump [65, 66, 99]. However, a constant and equal amplifier gain for all the modes over a wide range of input power and/or over frequency has not been demonstrated yet.

The detrimental effect of MDL is studied in [19, 64]. The authors in [19, 64] showed that the channel capacity is negatively affected by the effects of MDL. In [19] the authors also demonstrated through simulations that the system should address all possible supported modes by the fiber to avoid system outage due to power leakage to unaddressed modes. MDL in frequency-selective channels is investigated in [142]. In [142], it is found that delay spread alters the modal gains and also makes it frequency-dependent, which has a favorable effect on system performance.

This chapter is based on the work published in [143–149] and describes a study of the impact of mode coupling strength on the MDL in the system. The system performance is evaluated and analyzed for three mode coupling scenarios and for three and six spatial modes.

4.1. Simulation setup

The simulated system follows the structure of an MDM OFDM system as illustrated in Fig. 3.28. With OFDM, the ISI caused by DMD and dispersion is suppressed in the RX DSP by removing the cyclic part of the OFDM symbol added with the CP. Using a CP for ISI elimination allows for much less complex multiplications than time-domain equalization [57]. Additionally, OFDM permits a subcarrier-based post-processing at the RX, which is of advantage when using advanced equalization schemes such as ML detection (see Chapter 5).

Each TX generates a dual polarized signal with a net data rate of 100 Gb/s, modulated with quadrature phase shift keying (QPSK) OFDM. Since the simulated systems support three and six spatial modes, and each one of them supports two polarizations, there is a total of $N_\mathrm{t} = 6$ and $N_\mathrm{t} = 12$ tributaries, respectively. The former system supports the modes: LP_{01}, $\mathrm{LP}_{11\mathrm{a}}$ and $\mathrm{LP}_{11\mathrm{b}}$. The latter system supports: LP_{01}, $\mathrm{LP}_{11\mathrm{a}}$, $\mathrm{LP}_{11\mathrm{b}}$, $\mathrm{LP}_{21\mathrm{a}}$, $\mathrm{LP}_{21\mathrm{b}}$ and LP_{02}. The net data rate can be written as 3×100 Gb/s and 6×100 Gb/s for three and six spatial modes, respectively.

From a total of 4096 subcarriers (N_FFT), 81.3% of the subcarriers are dedicated to the actual payload (N_DATA) and the remaining 18.7% are used for zero-padding (N_ZP). An overhead of 24% for FEC was added, according ITU recommendation ITU-T G.975.1 for a concatenated, low latency and complexity RS and convolutional self-orthogonal code (CSOC) [136].

The TS preamble is constituted in each tributary by two consecutive OFDM symbols and the payload by three OFDM symbols as depicted in Fig. 4.1. The frame structure in Fig. 4.1 results in a TS overhead of 66%, which can be reduced by adding more OFDM symbols to the payload. The payload was kept relatively short in order to reduce simulation time, but it can be extended to lower the TS overhead. According to a study of the FMF channel dynamics in [150], the channel is expected to change in response of a mechanical perturbation as fast as every 250 μs. If the minimum TS repetition period is dimensioned so that is one order of magnitude faster than the period of the channel variations, the TSs would have to be sent every 25 μs. With an OFDM symbol duration of 97.6 ns, this minimum TS repetition period would allow the TS overhead to be reduced down to less than 1% with a frame of two OFDM TS and 254 OFDM symbols for the payload. The OFDM symbol duration is the result of multiplying the OFDM symbol size, including the CP, with the time spacing between two samples. According to the previously mentioned channel dynamics and following the maximum TS overhead constrain in [57], it is assumed that the number of OFDM symbols for the payload can be incremented and the TS overhead can be set to 10%.

To evaluate a target BER of 1e-3, up to three frames were simulated with different noise realizations. This results in the evaluation of more than 350 and 718 thousand bits for three and six spatial modes, respectively. The number of evaluated bits can be calculated from multiplying the number of modulated subcarriers in one OFDM symbol, the total number of tributaries, the number of transmitted OFDM symbols and the number of bits modulated in one subcarrier.

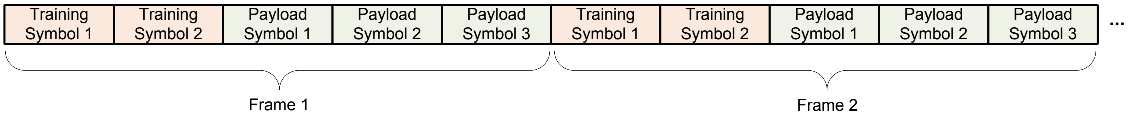


Figure 4.1.: Frame structure.

As mentioned in Chapter 3, the CP is dimensioned according to the delay spread caused by intra- and inter-modal dispersion. As described by τ_{DMD} in [ps] in Eq. (4.1), the maximum delay spread caused by inter-modal dispersion results from multiplying the DMD_{p} in [ps/km] (see Table 4.1) by the maximum simulated distances L_{max} in [km], i.e. 1200 km (three spatial modes) and 800 km (six spatial modes). This results in 10.8 ns and 6.08 ns of delay spread.

$$\tau_{\mathrm{DMD}} = \left| DMD_{\mathrm{p}} \right| L_{\mathrm{max}} \tag{4.1}$$

As for the delay spread caused by intra-modal dispersion, it is suggested in [57,110, 151] to use Eq. (4.2) to compute it. In Eq. (4.2), $\tau_{\max_\mathrm{D}}$, $D_{\max_{lm}}$, Δf represent the maximum delay spread due to the mode with the largest dispersion parameter in [ps], maximum dispersion parameter among the modes [ps $\cdot$ nm^{-1} $\cdot$ km^{-1}], spectral width in [Hz], respectively. Equation 4.2 is similar to the pulse broadening relation in Eq. (2.31) shown in Chapter 2 with the difference that here the worst case spread due to the maximum dispersion is considered. An approximate calculation of the spectral width Δf in Eq. (4.2) can be done as in [124] by calculating the spectral occupancy of the modulated subcarriers, i.e.

$$\frac{N_\mathrm{DATA} R_\mathrm{DAC}}{N_\mathrm{FFT}} \,.$$

The DAC rate R_DAC is equal to 49 GSamples/s. This results in a spectral width of 40 GHz. Considering a central frequency of 193.4 THz, and the maximum intra-modal dispersion parameters in Table 4.1, for three spatial modes, the $\tau_{\max_\mathrm{D}}$ results in 8.5 ns; for six spatial modes, $\tau_{\max_\mathrm{D}}$ corresponds to 5.6 ns. These delay spread values would add up to the duration of the CP. To avoid a larger overhead, the average dispersion is compensated in Equalization 1 in Fig. 3.28. As the dispersion parameters are very similar among the modes, most of the intra-modal dispersion can be compensated by this first equalizer. The residual dispersion is computed similar as in Eq. (4.2), but instead of $D_{\max_{lm}}$, the maximum deviation from the dispersion average is used. For three spatial modes, the delay spread due to residual dispersion is 76.2 ps; for six modes, 69.8 ps. For these reasons, the used CP length corresponds to 660 samples, i.e. a length of 13.5 ns and an overhead of 16.1%, which is larger than the delay spread caused by DMD and residual intra-modal dispersion.

$$\tau_{\max_\mathrm{D}} = \frac{c \left| D_{\max_{lm}} L_{\max} \right| \Delta f}{f_\mathrm{c}^2} \tag{4.2}$$

The total gross bit rate results from multiplying the net data rate by all aforementioned overhead from the CP, FEC and TSs. For three spatial modes, this value is 3×158.4 Gb/s, i.e. 475.1 Gb/s; and for six spatial modes is 6×158.4 Gb/s, i.e. 950.3 Gb/s.

The six and 12 tributaries are ideally multiplexed into the FMF channel without introducing crosstalk. The FMF channel, as depicted in Fig. 4.2 inset (a), is composed of 80-km spans. A total of N_{spans} form the link. The length of one span is a typical value for long-haul terrestrial links since it trades off the OSNR degradation with the distance and the number of amplifiers required for the link [24,152]. The channel model corresponds to the one described in Section 2.3.4 provided by the authors of [22,37,153]. Low launch power (negligible nonlinear effects) and no laser PN are considered here. The parameter of the FMFs are listed in Table 4.1, where $A_{\text{eff}_{lm}}$ and DMD_{p} are the effective area of mode LP_{lm} and the DMD parameter, respectively. The parameter $\Delta\bar{n}_{lm,\text{xy}}$ is the birefringence, which was set to a typical value for current fiber manufacturing technologies [1]. The DMD parameter is the difference between the inverse of the group delay of the fastest and the slowest modes, which gives a hint of the maximum impulse response duration attributed to the DMD. The two fiber profiles correspond to a graded-index with depressed cladding profile in order to minimize the DMD [120]. The mode coupling is modeled as in described in Section 2.3.4 by dividing the fiber into sections of 200 m, each having constant coupling coefficients due to a random displacement of the core center position. The displacement of the core center position was assumed to follow a uniform distribution within a certain displacement range, which will be specified in the next paragraph for the different simulation scenarios. It was shown in [36] that a uniform distribution for the core displacement provided a good fit with the experimental data in [154] regarding the evolution of the modal crosstalk with the transmission reach.

The continuous-wave approximation on which the model described in [47] is supported, is valid for fiber sections with a delay spread that is much smaller than the pulse duration. If a net symbol rate of 25 GBaud is used (equivalent single-carrier system), with a pulse duration of 40 ps, the maximum accumulated delay spread in a section may be 4 ps. For a DMD of 9 ps/km, this gives sections that have to be at most 444 m long.

Two fiber scenarios where modeled: fiber with strongly and weakly coupled modes. To generate strong mode coupling the maximum allowed core displacement was 4.7 % of the core radius. To emulate weakly coupled modes the maximum allowed core displacement was 0.6 % of the core radius. To visualize the mode coupling strength, a 40-ps width Gaussian pulse was transmitted through one of the polarization of the LP_{01} mode and the mean power evolution through all modes is depicted in Fig. 4.3(a) and (b). In Fig. 4.3, the power per mode is averaged over 200 channel realizations taking into account that the mode coupling is a random event. The fiber attenuation

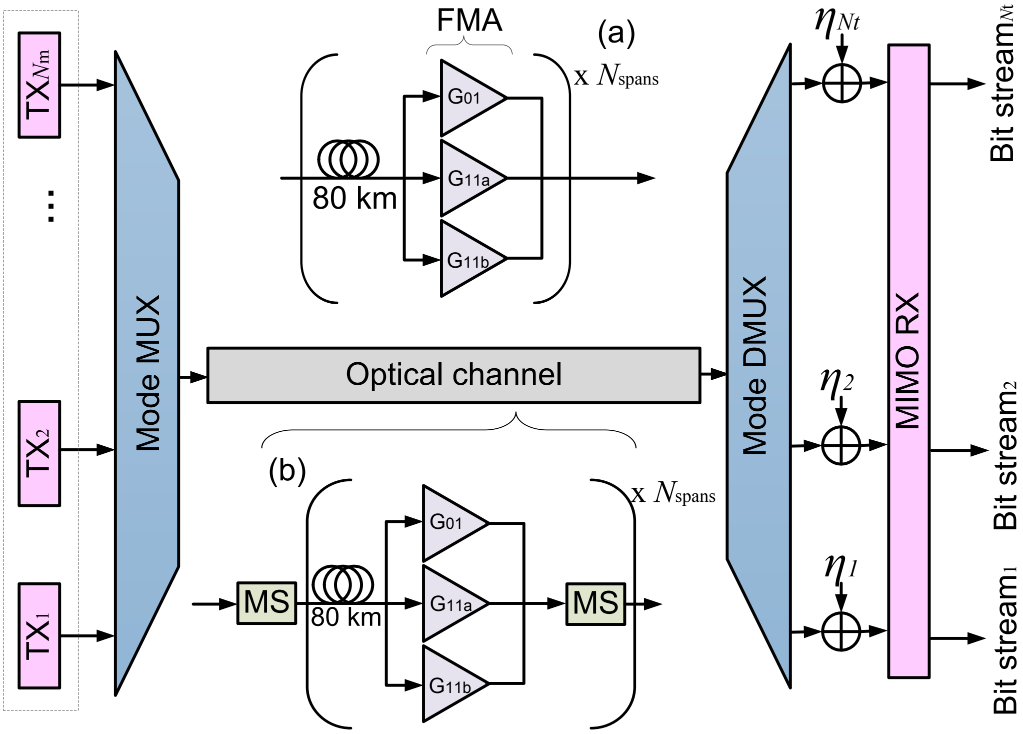


Figure 4.2.: FMF channel structure for transmission (a) without mode scramblers (MSs) and (b) with MSs.

Table 4.1.: Three and six spatial mode fiber parameters.

	LP$_{01}$	LP$_{11\text{a,b}}$	LP$_{01}$	LP$_{11\text{a,b}}$	LP$_{02}$	LP$_{21\text{a,b}}$
$A_{\text{eff}_{lm}}$ [μm^2]	101.6	137.1	142.9	191.8	297.7	256.3
D_{lm} [ps/km/nm]	22.2	21.8	22.2	22.1	21.7	21.9
$S_{\text{lo},lm}$ [ps/km/nm^2]	0.0663	0.0674	0.0665	0.0663	0.0626	0.0644
V	5.1		7.25			
DMD_{p} [ps/km]	9		7.6			
α_{dB} [dB/km]	0.22					
$\Delta\bar{n}_{lm,\text{xy}}$	1e-7					

was not taken into account in order to observe only the effect of the two simulated types of mode coupling. It can be seen from Fig. 4.3 that for both FMFs in the case of strongly coupled modes the power from LP$_{01}$ leaks faster to the rest of the modes with the distance than for the weakly coupled modes. In which degree the power leaks from one mode to another is directly related to the propagation constants of the modes, the closer they are the stronger the mode coupling will be [43].

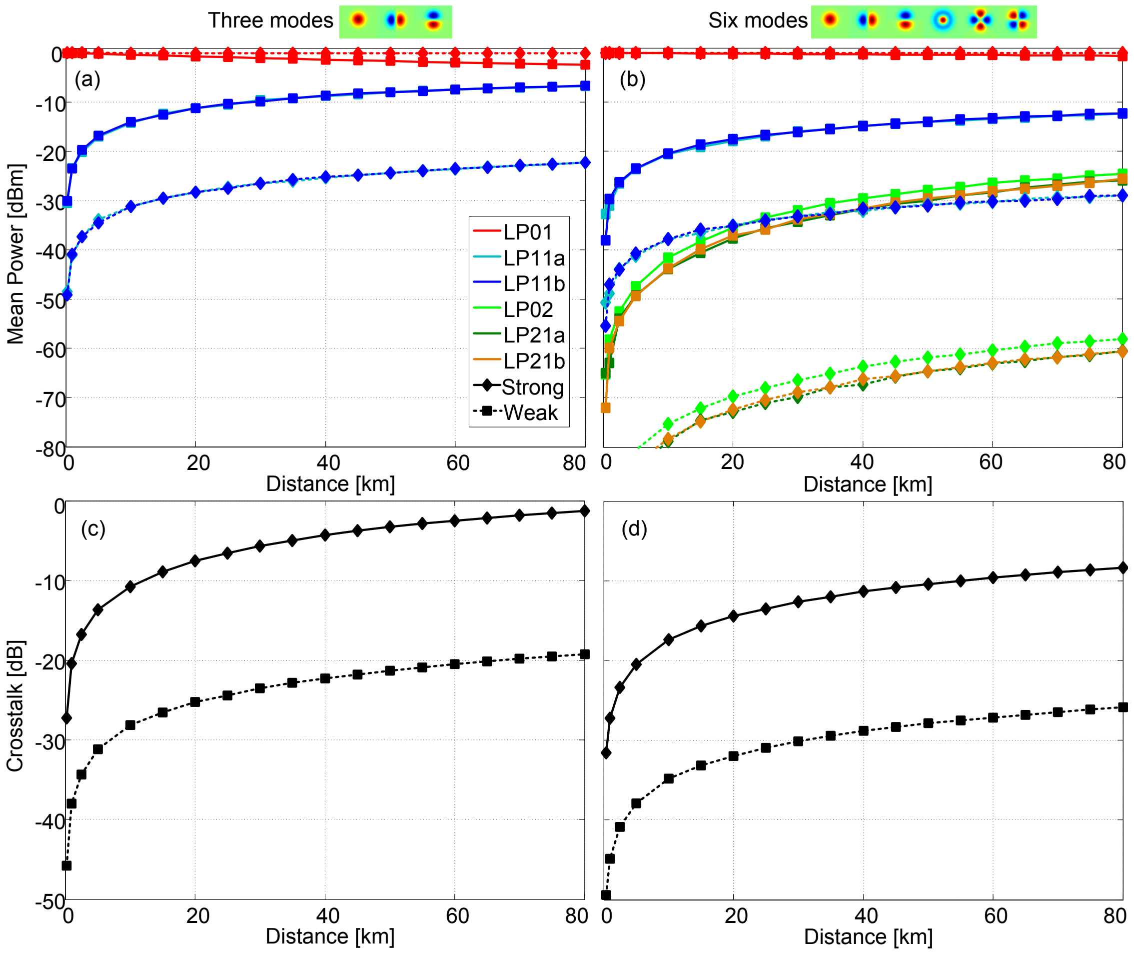


Figure 4.3.: Average power and crosstalk as a function of the transmission distance for strongly and weakly coupled modes for three spatial modes in (a) and (c), and for six spatial modes in (b) and (d).

Figure 4.3(c) and (d) show the crosstalk in [dB] as a function of the distance. The crosstalk is defined as Eq. (4.3) in linear units, where $\bar{P}_u$ is the average power of the mode u, $\bar{P}_{u1}$ is the average power of the mode $u1$ from which the pulse is launched, $\mathbb{U}$ is the set of all modes except for $u1$ and $\mathbb{U}_{\text{total}}$ is the set of all supported modes by the FMF. Crosstalk is a measure of strength of mode coupling and is given after a certain distance. In the three spatial modes case, after 80 km the induced average mode crosstalk between LP_{01} and the degenerate modes $\text{LP}_{11\text{a}}$ and $\text{LP}_{11\text{b}}$ is -2.3 dB for the modes which are strongly coupled and -19.2 dB for the weakly coupled modes. For the six mode fiber, the crosstalk values are similar corresponding to -

8.3 dB and -26.9 dB, respectively. The FMF supporting three weakly coupled spatial modes resembles the mode coupling behavior of the fibers presented in [51,155]. At 30 km, the FMF used in [155] has -25 dB of crosstalk, whereas in [51] for 33 km the crosstalk is -18 dB. For six spatial modes there were no crosstalk values found in the literature, presumably since to measure it a mode-selective (D)MUX setup is needed to measure the power of each one of the modes separately. The use of mode-selective (DMUX) is not a common practice for more than three spatial modes since the experimental setup is much more complex than a setup with spatial (D)MUX.

$$\begin{aligned} XT &= 10\ \log_{10}\frac{\sum\limits_{u\in\mathbb{U}}\bar{P}_u}{\bar{P}_{u1}} \\ \mathbb{U} &= \{u \mid u \in \mathbb{U}_{\text{total}} \wedge u \neq u1\} \\ \mathbb{U}_{\text{total}} &= \{\text{all supported spatial modes}\} \end{aligned} \tag{4.3}$$

Weakly and strongly coupled modes can also be distinguished from the estimated impulse responses as shown in Fig. 4.4. The figure illustrates exemplary impulse responses out of a total of 36 after 80 km of transmission for the three-mode fiber. For the sake of clarity, only the impulse response related to LP_{01} and $\text{LP}_{11\text{a}}$ are depicted in Fig. 4.4; the impulse responses related to LP_{01} and $\text{LP}_{11\text{b}}$ are very similar to the plots shown in Fig. 4.4. In the case of weak mode coupling, it can be seen in Fig. 4.4(e)-(h) that most of the transmitted energy belonging to one mode is received in the same mode; only a small part couples into other modes as shown in Fig. 4.4(g) and (f). In the case of strongly coupled modes the cross-mode estimated impulse responses in Fig. 4.4(b) and (c) have larger energy than in the weakly coupled case. The estimated impulse response representing the direct path from one mode to itself, Fig. 4.4(a) and (d), show two peaks each having additional components between their positions as a product of a stronger mode coupling. The fiber supporting six modes shows similar information.

The channel transfer function of the two mode coupling strengths also presents differences. Exemplary histograms of the normalized magnitude of the complex-valued channel matrix are depicted in Fig. 4.5 for all simulated frequencies, 1200 km and three spatial modes. For strongly coupled modes all 36 histograms are very similar, therefore solely one histogram is shown. The histogram representing the strong mode coupling presents a good fit to a Rayleigh distribution, which is an indication that the real an imaginary values of the channel matrix are statistically independent, Gaussian distributed with zero-mean and equal variance. This can be proven

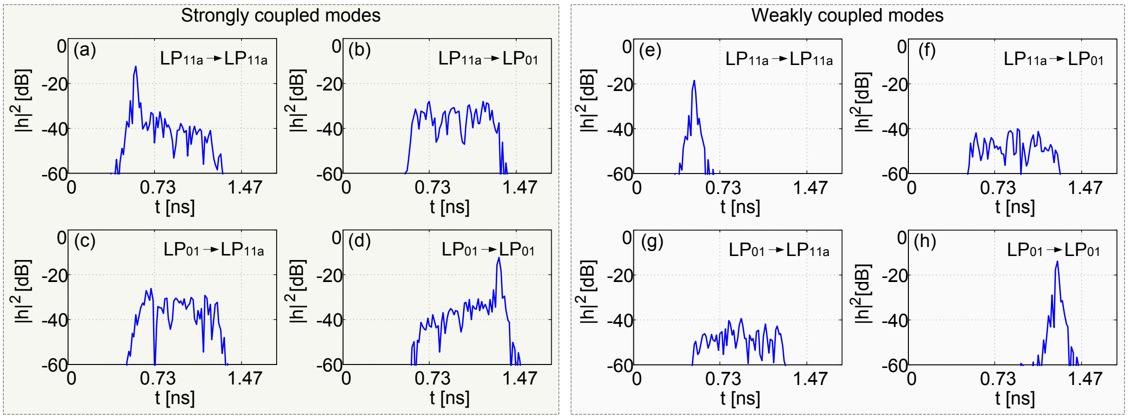


Figure 4.4.: Estimated impulse responses after 80 km of transmission for a fiber presenting strong mode coupling from: a) LP_{11a} to LP_{11a}, b) LP_{11a} to LP_{01}, c) LP_{01} to LP_{11a} d) LP_{01} to LP_{01}, and weak mode coupling from: e) LP_{11a} to LP_{11a}, f) LP_{11a} to LP_{01}, g) LP_{01} to LP_{11a}, h) LP_{01} to LP_{01}.

by determining the joint PDF of the real and imaginary parts of a random variable assuming them to be statistically independent Gaussian distributed with zero mean and equal variance. By changing the coordinates from Cartesian to polar, i.e. magnitude and phase, and then determining the marginal probability distribution of the magnitude, it can be proven that the magnitude is Rayleigh distributed. The interested reader is referred to the derivation shown e.g. in [156].

The histograms for weakly coupled modes show also shows a Rayleigh distribution shape in the cross-mode transfer functions with small magnitude values. In turn, the histograms corresponding to the transfer functions from LP_{11} to the LP_{11} modes do not have a specific shape. As the mode coupling in this case is weak, the channel is almost flat and the magnitude of the 16 transfer functions related to the LP_{11} fluctuates from high to low levels nearly uniformly. Similarly the histograms related to the LP_{01} mode do not have a structure.

The FMA is the element in the span generating MDL. Let G_{lm} be the gain for mode LP_{lm}. As reported in [97] for 1550 nm, which is the carrier frequency for these simulations, for three spatial modes G_{11} is set to $G_{01} - 2$ dB. Whereas, for six spatial modes the gains were assigned as in [65] also for 1550 nm as follows: $G_{11} = G_{01} - 1.5$ dB, $G_{02} = G_{01} - 4$ dB, and $G_{21} = G_{01} - 3.5$ dB. In both cases, G_{01} compensates the attenuation in one span.

As illustrated in Fig. 4.2, after demultiplexing the signals from the FMF, white Gaussian noise is added and the signals are processed in the receiver DSP. The DSP has the structure shown in Fig. 3.28. The signal is compensated for dispersion in the

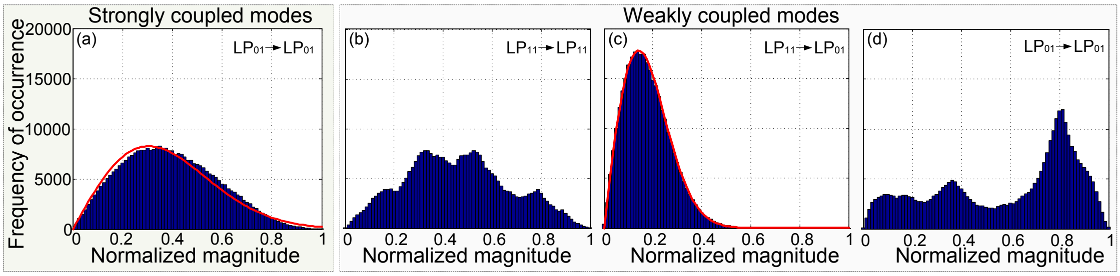


Figure 4.5.: Exemplary histograms of the channel matrix magnitudes after 1200 km of transmission for (a) strongly couple modes and (b)-(d) weakly coupled modes. The red continuous line represents the fitted histogram to a Rayleigh distribution.

first stage of the equalization, synchronized, the CP is removed in order to eliminate the interference of the OFDM symbol with other neighboring symbols, equalized, symbols are detected and demapped into bits. The dispersion compensation is done as described in Section 3.5.2 in the frequency domain.

For synchronization and channel estimation constant amplitude zero autocorrelation (CAZAC) sequences are employed [127, 157]. These sequences are also called Frank-Zadoff or Chu codes. This type of preamble offers relatively short preambles and ideal autocorrelation properties, which is beneficial for the channel estimation through the correlation of the received preamble and the stored preamble at the RX. An additional advantage is the low PAPR, which is good for avoiding nonlinearities in the system. In order to create orthogonality between the training sequences, the sequences are time shifted with respect to one another. The shift is related to the maximum delay spread in the system. A typical time shift is the length of the sequence divided into the number of tributaries, this value should be at least the delay spread of the channel impulse response. The preamble is composed of two identical CAZAC sequences, which creates periodicity for the synchronization algorithm. Delaying the received sequence by the length of one CAZAC sequence and crosscorrelating it with the original received sequence will result in a peak indicating the beginning of the preamble. The synchronization procedure is based on the algorithm proposed by Schmidl and Cox in [126] and is the algorithm adopted for these simulations. Another option for synchronization is to crosscorrelate the received sequence with a previously stored transmitted preamble. However, the mode coupling represents a major distortion that hinders a clear identification of the preamble.

The channel equalization is performed by a ZF equalizer. Other kinds of equalizers are studied in Chapter 5. ZF is a linear equalizer that aims to remove completely the multi-stream interference. As illustrated in Fig. 4.6, the data streams are first decoupled and separately detected. This model is valid for each OFDM subcarrier, since each one of them can be seen as an independent frequency-flat channel.

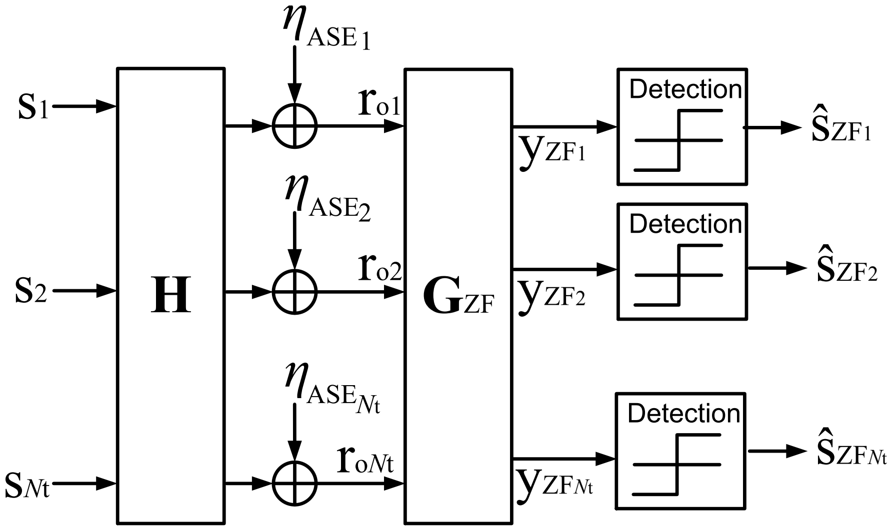


Figure 4.6.: Linear transmission and detection model with AWGN

As described in Eq. (4.4), the ZF filter matrix $\mathbf{G}_{\text{ZF}} \in \mathbb{C}^{N_t \times N_t}$ accomplishes the elimination of the interference by making the product $\mathbf{G}_{\text{ZF}}\mathbf{H}$ an identity matrix, where $\mathbf{H} \in \mathbb{C}^{N_t \times N_t}$ is the channel matrix. In Eq. (4.4), $\mathbf{y}_{\text{ZF}} \in \mathbb{C}^{N_t}$ represents the symbols after the zero-forcing equalization, $\mathbf{r}_o \in \mathbb{C}^{N_t}$ the received symbols after the channel distortion and ASE noise addition, and $\boldsymbol{\eta}_{\text{ASE}} \in \mathbb{C}^{N_t}$ the ASE noise terms. For this purpose the matrix $\mathbf{G}_{\text{ZF}}$ is defined as the inverse of the channel matrix $\mathbf{H}$, however, ignoring the noise term. Therefore, if the channel matrix presents small values, the inversion and multiplication with the noise vector will result in noise amplification. If the channel was noiseless, the ZF would represent an ideal equalizer. Other alternatives that overcome this drawback are presented in Chapter 5. After ZF equalization the symbols in $\mathbf{y}_{\text{ZF}}$ are detected and demapped into bit streams.

$$\begin{aligned} \mathbf{y}_{\text{ZF}} &= \mathbf{G}_{\text{ZF}}\mathbf{r}_o = \mathbf{s} + \mathbf{G}_{\text{ZF}}\boldsymbol{\eta}_{\text{ASE}} \\ \mathbf{G}_{\text{ZF}} &= \mathbf{H}^{-1} \end{aligned} \tag{4.4}$$

4.2. Simulation results

This sections describes the simulations results related to the impact of the mode coupling strength in the system performance and the amount of MDL in the system. The simulation results are divided in three subsections assessing the mode coupling from three different sources: when its generated in a distributed manner along the fiber, from mode (D)MUXs and from mode scramblers at specific locations along the optical link.

4.2.1. Distributed mode coupling

Distributed mode coupling originates from the fiber itself as the mode coupling described in the previous section with Figs. 4.3 and 4.4. The performance of the two simulated distributed mode coupling strengths as a function of the distance is depicted in Fig. 4.7 for three spatial modes. The OSNR penalty is the difference of the actual OSNR and a reference OSNR which corresponds to the back-to-back (B2B) OSNR for a target BER (TBER) of 1e-3. Figure 4.7 also illustrates the OSNR penalty for different levels of modal gain. The vertical bars in Fig. 4.7 represent +/− the standard deviation of the OSNR penalty from the simulated 200 channel realizations.

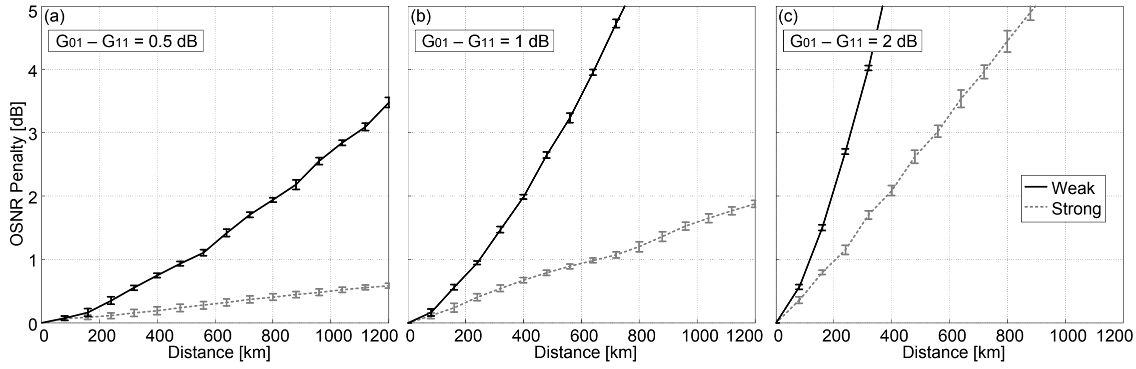


Figure 4.7.: OSNR penalty at a BER of 1e-3 as a function of the distance for the fiber supporting three spatial modes and various modal gain differences: (a) 0.5 dB, (b) 1 dB and (c) 2 dB.

Figure 4.7 shows, the OSNR penalty increases with both distance and modal gain difference. In addition, for weakly coupled modes the OSNR penalty is higher than for strongly coupled modes. Allowing an OSNR penalty of 1 dB the fiber generating strong coupling provides an improvement of 164 % and 76 % in the maximum transmission distance for 1 dB and 2 dB modal gain differences, respectively. For the 0.5-dB gain difference, in the strongly coupled case the average OSNR penalty does not reach the 1 dB limit after 1200-km transmission, which means that at least a one-fold increase in the transmission distance is expected. The OSNR penalty for the simulated FMF supporting six spatial modes shows also that the performance of strongly coupled modes outperforms the weakly coupled modes as shown in Fig. 4.8 following the modal gain differences in [65]. As the modal gain differences are greater than in the three mode cases, the performance improvement is reduced.

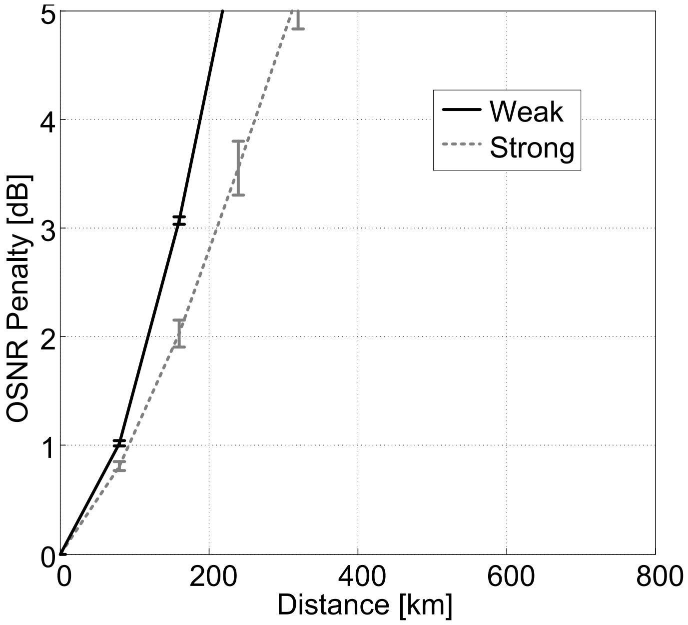


Figure 4.8.: OSNR penalty at a BER of 1e-3 as a function of the distance for the fiber supporting six spatial modes with the modal gain differences according to the experiments in [65]: $G_{11} = G_{01} - 1.5$ dB, $G_{02} = G_{01} - 4$ dB, and $G_{21} = G_{01} - 3.5$ dB. G_{01} compensates the attenuation of one span.

As observed in Fig. 4.4, during propagation a power leakage from one mode to another takes place due to mode coupling. The amount of power leakage depends on the strength of the mode coupling and the propagation distance. As shown in Fig. 4.3, when strongly coupled modes propagate, the power interchange is faster with the distance compared to when weakly coupled modes traverse the fiber. Mode coupling does not allow to concentrate the loss in one or certain modes, thus, the loss generated by the modal gain difference distributes into other modes. This is the reason why in the presence of MDL strongly coupled modes outperforms the weakly coupled modes.

MDL limits the system performance and mode coupling increases the tolerance against it. This can also be seen by means of the channel capacity and the mutual information. Here, it is considered that there is no previous information about the channel at the transmitter, i.e. there is no channel state information, e.g. no water filling is used. The transmission of channel state information back to the transmitter is difficult to implement, since it requires a round-trip delay that may be greater than the channel dynamics [19]. For this reason, it is assumed that each tributary is launched with equal power. The channel is assumed to be linear time-invariant and disturbed by AWGN as described in Sections 3.2 and 4.1. For a certain number of modes and without MDL, the channel capacity is defined as Eq. (4.5a), where SNR is the signal-to-noise ratio (total signal power from all tributaries over noise per tributary) [71]. When the number of tributaries tend to infinity Eq. (4.5a) can be rewritten as Eq. (4.5b), where e is Euler's number [64]. The capacity in Eqs. (4.5a) and (4.5b) is illustrated as a function of SNR and for $N_\mathrm{t} = 6$ in Fig. 4.9. These two upper curves in Fig. 4.9 serve as a reference for the two lower curves corresponding to the computed average mutual information (MI) for strongly and weakly coupled modes. The MI was averaged from 200 channel realizations after the propagation through 1200 km of FMF supporting three spatial modes. The link had a modal gain difference per span of 2 dB. The computed MI for a single channel realization follows Eq. (4.5c), where λ_{ij} is the i-th singular value of the channel matrix $\mathbf{H}$ for the j-th OFDM data subcarrier [71, 158].

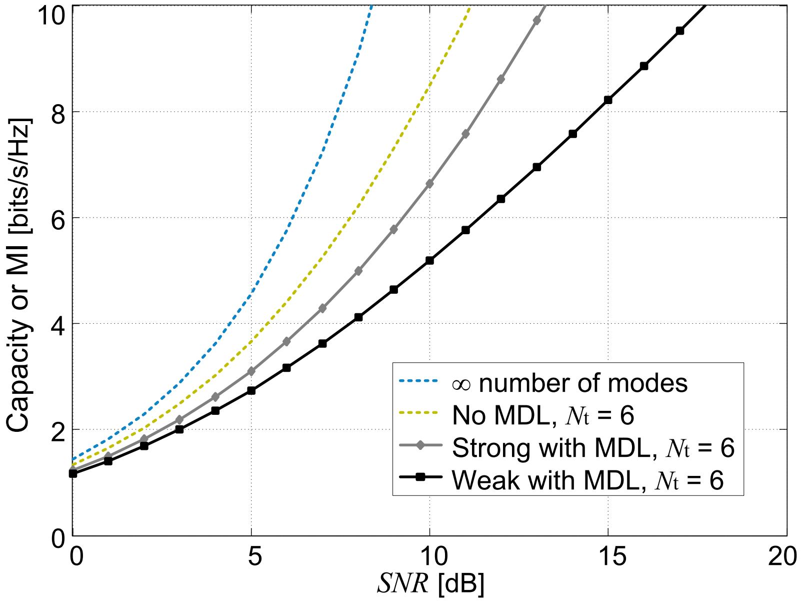


Figure 4.9.: Channel capacity for infinite number of modes and six tributaries with and without MDL for weakly and strongly coupled modes after 1200 km of propagation.

$$C = N_\mathrm{t} \log_2 \left(1 + \frac{SNR}{N_\mathrm{t}}\right) \tag{4.5a}$$

$$C_\infty = SNR \log_2 (e) \tag{4.5b}$$

$$\hat{I} = \frac{1}{N_\mathrm{DATA}} \sum_{j=1}^{N_\mathrm{DATA}} \sum_{i=1}^{N_\mathrm{t}} \log_2 \left(1 + \frac{SNR}{N_\mathrm{t}} \lambda_{ij}^2\right) \tag{4.5c}$$

Figure 4.9 shows that the average MI for strongly coupled modes is the closest to the curve which represents the capacity. Similar to the system performance plots in Fig. 4.7, strongly coupled modes have less impact on the system MI of an MDL-impaired system than weakly coupled modes.

The effect of mode coupling on the performance of an MDL impaired system is similar to the effect of PMD on single-mode systems impaired by polarization-dependent loss (PDL) [159,160]. Similarly to the MDL case for FMF systems, in a signal from a single-mode system that transverses a PDL-impaired link without the effect of PMD, all OFDM subcarriers or channel frequency components will experience the same loss; with PMD these frequency components will evolve differently (different alignments with the axis of the PDL element), therefore the probability that all of them experience the same loss is greatly reduced. The consequence of this is that power levels of these frequency components and the associated OSNR diversifies [159]. In the best case, in presence of PDL and two polarizations, PMD rotates both polarizations so they are equally attenuated. In the worst case, PDL attenuates only one of the polarizations. For MDM, the best case would be when the modes have equal losses and all the inline optical components have the same gain or attenuation; the worst case would take place when there is no mode coupling and some of the modes are more attenuated than the others, until the point where the information on these modes cannot longer be recovered.

Although the strong mode coupling reduces the OSNR penalty in the system, it is worth to note that the absence of mode coupling is also desirable for two main reasons. First, the spatial modes propagate independently, which would avoid a complex MIMO equalization scheme higher than 2×2. However, even the propagation of weakly coupled modes through short distances, the complexity reduction based on an equalizer of smaller dimension, i.e. a MIMO equalizer of less than $N_\mathrm{t} \times N_\mathrm{t}$, would come at a cost of system performance [161, 162]. Secondly, mode coupling is not desirable in DMD compensated transmission (as in the experiments in Chapter 6).

DMD compensation is achieved with sections of FMFs with opposite DMD characteristics. This would avoid a large number of taps for time-domain equalizers or high overhead due to a long CP for frequency-domain equalization. DMD compensation is performed by employing fibers of opposite DMD signs. In this kind of link architecture, mode coupling would introduce artifacts in the impulse response as in Fig. 4.4, which would expand the impulse response with the propagation through the negative and positive DMD fibers. For DMD compensation, mode coupling is avoided not only from the fiber but also from inline optical components and splices [163,164].

The degradation of the system performance is closely related to an increase of the MDL with the propagation distance. This is shown in Fig. 4.10 for the weakly and strongly coupled modes, three and six spatial modes. The modal gain differences in Fig. 4.10 are set according to the experiments in [65, 97], i.e. for three spatial modes G_{11} is set to $G_{01} - 2$ dB and for six spatial modes $G_{11} = G_{01} - 1.5$ dB, $G_{02} = G_{01} - 4$ dB, and $G_{21} = G_{01} - 3.5$ dB; in both cases G_{01} compensates the fiber attenuation in each span. The MDL Υ is computed per subcarrier and is averaged in Fig. 4.10 over the all subcarriers and the 200 channel realizations. Figure 4.10 shows for the case of weak mode coupling that the average MDL accumulates linearly with the distance. On the other hand, in the case of strong mode coupling the average MDL increases proportional to the square root of the transmission distance.

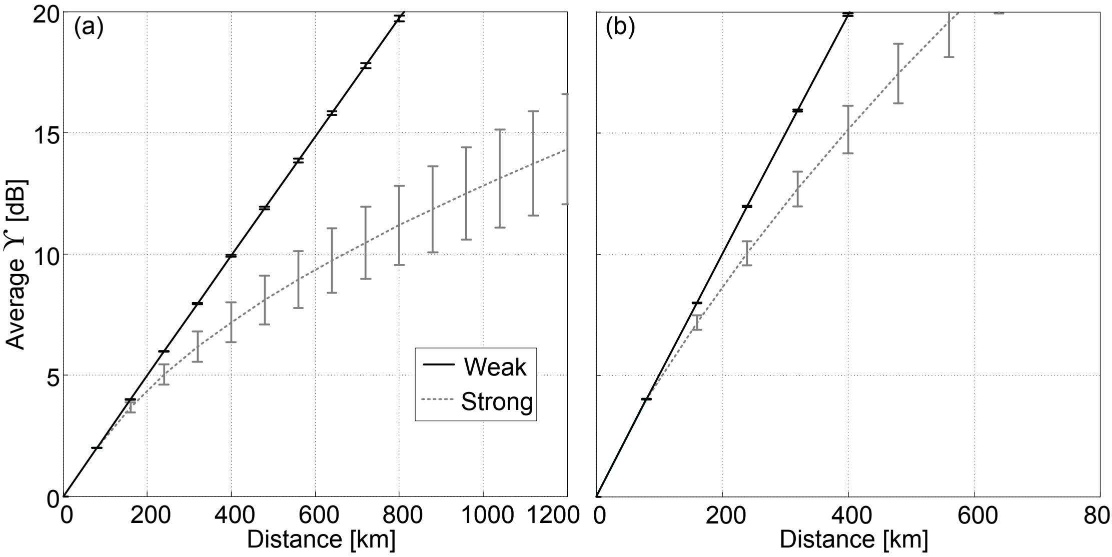


Figure 4.10.: Average MDL as a function of the transmission distance for the two mode coupling strengths, (a) three and (b) six spatial modes, and the maximum modal gain differences of 2 dB [97] and 4 dB [65], respectively.

It can be seen in Fig. 4.10, Υ has larger variations with the distance for strongly coupled modes than weakly coupled modes. The increase in the variance is a product of the effect of mode coupling and the accumulated DMD, which at the same time makes the channel transfer function frequency-selective. As a consequence, the gains of the Gaussian subchannels (see Section 3.2) fluctuate as shown in [142] and MDL also fluctuates randomly. The frequency dependance of Υ can be seen in two exemplary signals in Fig. 4.11.

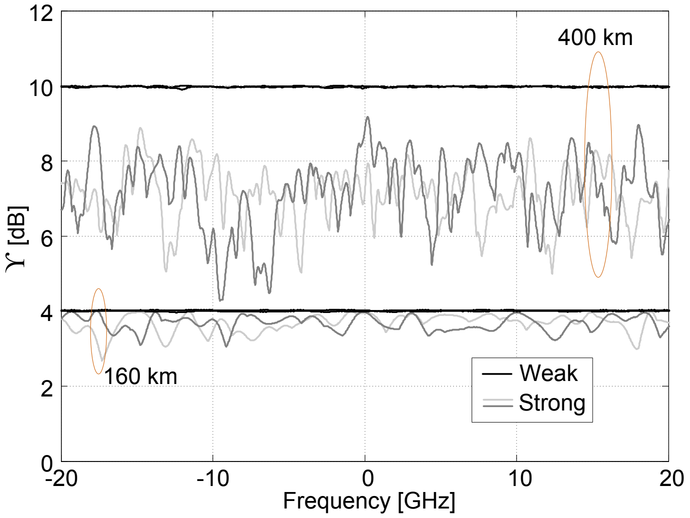


Figure 4.11.: Exemplary MDL values Υ in [dB] as a function of the frequency after 400 km and 160 km for weak and strong mode coupling for the three spatial modes. The modal gain difference between LP_{01} and LP_{11} is 2 dB. Also here, G_{01} compensates for the attenuation in one span.

Designing a fiber with inherently strongly coupled modes may be difficult to realize. In Chapter 8 of [43] the authors describe with the help of the illustration in Fig. 4.12, some of the design criteria for creating a fiber with low attenuation constant difference among the modes. The V parameter is first chosen slightly higher than needed to support certain number of modes. This ensures that the lowest propagation constant among the supported modes (β_{min}) is higher than some threshold, which is a requirement for low macrobending loss. It also implies to guide a higher order mode, but this mode is highly lossy. The propagation constant of this bounded but lossy mode is called β_{lossy}. Another requirement is to maximize the difference between β_{min} and β_{lossy}, $\Delta\beta_{\text{lossy}}$, in order to guarantee low microbending loss. For strong mode coupling, it is required that the difference between propagation constants $\Delta\beta_{ij}$ is small. Fulfilling the criteria for macro and microbending results in $\Delta\beta_{\text{lossy}}$ and $\Delta\beta_{ij}$ being large and very similar. Note that $\Delta\beta_{\text{lossy}}$ and $\Delta\beta_{ij}$ are difficult to manipulate independently [43]. Since well bounded modes with small modal loss differences is a fundamental feature of fibers, most of the fabricated fibers have large $\Delta\beta_{ij}$, i.e. weakly coupled modes.

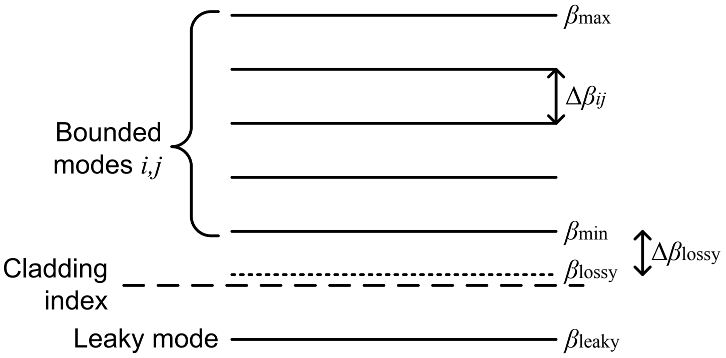


Figure 4.12.: Schematic for propagation constant design.

Other alternatives to induce distributed mode coupling is to introduce a perturbation during the manufacturing process of the fiber. In [165] the authors proposed to spin the fiber during the drawing process to increase the mode coupling strength. For SMFs, this procedure reduces the delay spread caused by PMD; in the case of FMF, it should decrease the MDL and the DMD. However, [165] only studies a change in delay spread between the degenerate LP_{11} modes and shows that spinning the fiber might not be as effective as for SMF. The effect of a spun fiber regarding the mode coupling of different mode groups has not been studied yet.

The authors in [166] proposed to introduce distributed mode coupling by means of long-period gratings. The gratings are incorporated as a periodical perturbation in form of ultraviolet (UV) illumination in the fiber fabrication. The UV beam introduces a refractive index perturbation that produces mode coupling. Although the work in [166] is based on simulations, it shows promising results for a fiber supporting three spatial modes.

4.2.2. Mode coupling introduced by mode (D)MUXs

Spatial mode MUXs excite the modes by illuminating a portion of the FMF core. As shown in Section 3.3.1, at the moment the fiber is illuminated at the beginning of the transmission link a linear combination of the signal and the spatial modes is created. As mode coupling is also a linear combination of the modes, in this section the impact of the use of spatial (D)MUXs is studied. By using the (D)MUX model presented in Section 3.3.1, simulations to asses the system performance are carried out and depicted in Fig. 4.13.

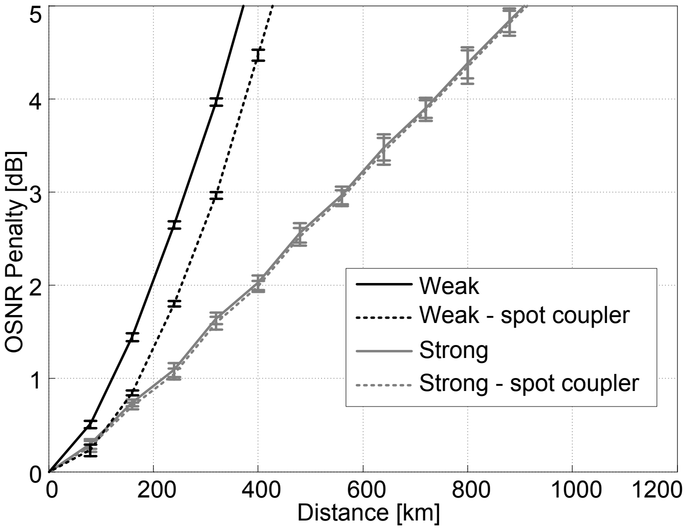


Figure 4.13.: OSNR penalty at a BER of 1e-3 as a function of the distance for a system using spot couplers as (D)MUX, three spatial modes, 2 dB of modal gain difference per span, weakly and strongly coupled modes.

The plot shows for the weakly coupled modes an increase in maximum transmission distance of more than 40 % allowing an OSNR penalty of 1 dB. For the fiber with strongly coupled modes, the transmission distance increment is not greater than 10 %. This shows that the spatial MUXs aid to improve the system performance, but not significantly. Since there is more modal gain difference and the weakly and strongly coupled mode curves are closer to each other in OSNR penalty plots in Fig. 4.8, for the six spatial mode case it is shown that the performance will not be much affected by the use of spatial MUXs.

4.2.3. Mode-scramblers-induced mode coupling

As seen in the previous sections mode coupling decreases the accumulated MDL in the link and improves the system performance. In this section it is proposed using MSs to increase the mode coupling strength.

As illustrated in Fig. 4.2(b), MSs were included after the mode MUX and each FMA, since the best results were obtained by mixing the signal before and after the MDL sources. In order to introduce mode scrambling, it is necessary to introduce a mechanical perturbation in the fiber. In the literature, several options of mode scrambling can be found. These are depicted schematically in Fig. 4.14.

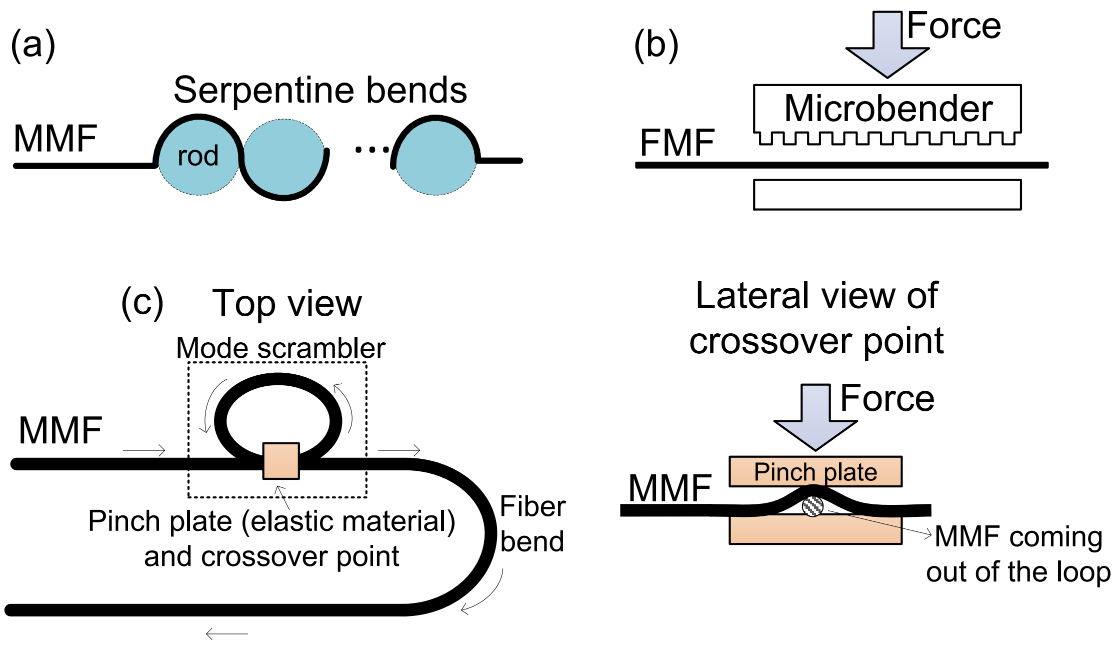


Figure 4.14.: Different options for the realization of MSs represented schematically by: (a) sepentine bends [42], (b) microbenders [81] and (c) pressing the fiber against an elastic material [167].

MSs are mostly used for measuring the bandwidth-distance product of commercial multimode fibers and the insertion loss when multiple multimode fibers are concatenated [168, 169]. Bandwidth-distance of a multimode fiber refers to the product of the transmission reach and 3-dB bandwidth of the baseband frequency response relating the input and output signal that directly modulates the light source feeding the fiber. The DMD is a key factor that influences the bandwidth-distance product, since the larger the delay spread caused by the DMD, the lower the bandwidth. The pulse distortion due to DMD depends highly on the modes that are being excited in the fiber, and therefore, on the light launch conditions. Standard methods for measuring the bandwidth-distance product use MSs to guarantee equal launch conditions in every measurement [169]. In the case of connector insertion loss measurements, the launch conditions are also fundamental, since the power loss due to misalignment in the connectors variates depending on the power distribution in the core. A higher insertion loss is measured, if the entire core and, partially, the cladding is illuminated (over-filled launch); in consequence, there is light propagating close to the core-cladding boundary, which makes the connector insertion loss higher due to misalignment. On the contrary, a lower insertion loss is measured if the launch is restricted to only a small portion of the core.

Both the bandwidth-distance product and the connector insertion loss measurements require a controlled light launch into the fiber. MSs have been studied to generate an stable and reproducible power distribution in the core [42, 167, 170]. This enables reproducibility and comparability of the measurement results.

The method of mode scrambling via sepertine bends is proposed in [42]. As shown in Fig. 4.14(a), rods are used to create a curvature in the fiber; in [42], by varying the rods in number and diameter, and their separation the far field intensity pattern is studied to create a stable mode distribution. By doing a sweep of the frequency of the oscillator modulating a laser source, the authors show that by using the proposed MS, the baseband attenuation of a fiber can be predicted by measuring and adding up the baseband attenuation of shorter fibers.

Mode scrambling by means of microbending can also be realized [81,170]. As depicted in Fig. 4.14(b), it consist of creating a perturbation in the fiber via a deformation. In [170], the authors showed that the greater the depth of the deformation the stronger the mode coupling. In [81], the authors demonstrate a two-spatial mode transmission over 4.5-km of FMF. Microbending was used in [81] as a mode MUX at the transmitter to achieve mode conversion from LP_{01} to both LP_{01} and LP_{11}; and as a mode DMUX at the receiver to convert LP_{11} into LP_{01}. The force applied to the fiber and the position of the grating is optimized in [81] to obtained the desired conversion ratio in the transmitter, equal power for both spatial modes, and the highest possible attenuation of LP_{01} at the receiver.

The authors in [167] proposed a MS in combination with a mode filter to achieve reproducibility of connector insertion loss measurements. As illustrated in Fig. 4.14(c), the MS consisted of a loop, whose crossover point is pressed and sandwiched between an elastic material. In [167], it is shown that increasing the applied force to the elastic material increases the mode coupling and achieves a uniform near-filled pattern at the output of the MS. Subsequently, in order to increase the loss in the higher-order modes, the light goes through mode filtering consisting of a half turn bend. The mode filter enables an intensity distribution in the core of the fiber that complies with the International Electrotechnical Commission (IEC) standard for attenuation measurement in multimode fibers IEC61280-4-1 [167]. The proposed device in [167] is shown to be capable of reproducing the measured connector insertion loss independently of the intensity profile of the input light.

In this work, MSs were modeled as a 12-cm fiber (similar to [42]), in which mode coupling is introduced every 500 μm (as the nominal grating pitch in [81]). In order to generate enough mode-mixing at the end of the 12-cm fiber the maximum core displacement was adapted, so that when a Gaussian pulse is launched in one of the modes, its power distributes nearly equally over all the modes. In the case of three and six spatial modes, the maximum core displacement was 18.9 % of the core radius; for six spatial modes, 37.8 %. As also implemented in the simulations in 4.2.1, the displacement of the core center position was assumed to follow a uniform distribution within the maximum core displacement. The power evolution of a pulse through the length of the MS was simulated for 200 channel realizations and depicted in Fig. 4.15.

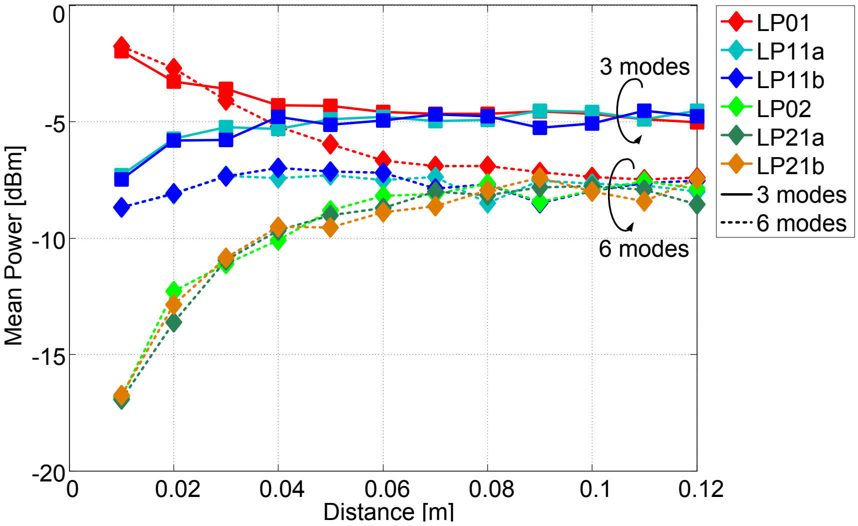


Figure 4.15.: Mean power evolution of a pulse as a function of distance for a three- and six-mode MS.

The results regarding the average system performance and accumulated MDL are depicted in Fig. 4.16 for the simulated 200 channel realizations. The OSNR penalty curves show that for the three spatial modes system using MSs increases the maximum transmission reach by virtually two-fold with respect to the weakly coupled modes and allowing 1 dB OSNR penalty; for six spatial modes there is an increase of more than 70 % in the transmission distance. The MDL plots show likewise positive results in form of a reduction on the accumulated MDL, even for strongly coupled modes.

Note that in Fig. 4.16(a)-(b), the average performance of the system with MSs is very similar for weakly and strongly coupled modes. Therefore, it would not be necessary to design and fabricate fibers with inherently strongly coupled modes for MDL-tolerant systems, but only place inline MSs. It can be seen also in Fig. 4.16 that when MSs are used, the standard deviation of the OSNR penalty is larger for weakly than for strongly coupled modes. The distributed mode coupling in the fiber

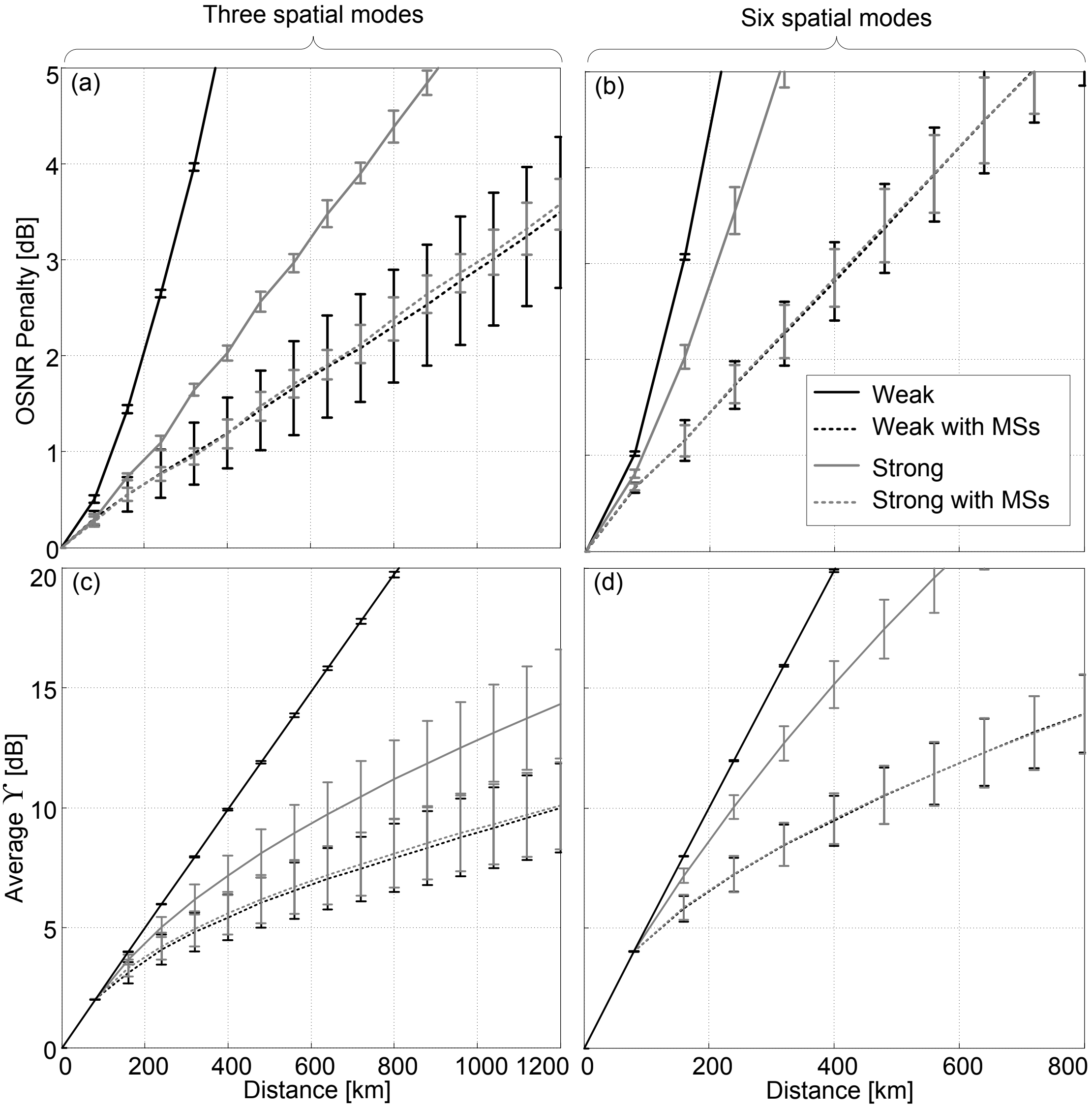


Figure 4.16.: (a)-(b) OSNR penalty at a BER of 1e-3 and (c)-(d) average MDL as a function of the distance for three and six spatial modes, respectively, and a system containing MSs in the position illustrated in Fig. 4.2(b). The modal gain differences are set here also according to the experiments in [65, 97], i.e. for three spatial modes $G_{11} = G_{01} - 2$ dB, and for six spatial modes to $G_{11} = G_{01} - 1.5$ dB, $G_{02} = G_{01} - 4$ dB, and $G_{21} = G_{01} - 3.5$ dB. In both cases, G_{01} compensates for the fiber attenuation of one span.

reinforces the effect of the mode coupling introduced by the MSs and the system performance fluctuates less in this case.

4.3. Discussion

By means of simulations of three and six spatial modes, it is shown that strongly coupled modes increase the system tolerance against MDL. Mode coupling enables the exchange of energy between modes causing the loss from MDL to distribute among all modes. This situation is significantly more favorable than the loss harming a selected group of spatial and polarization modes, as in the weakly coupled mode case.

Three sources of mode coupling have been investigated: distributed, from mode couplers and MSs. In principle, every source of strong mode coupling is beneficial for the system performance and the accumulation of MDL. Inherently coupled modes by distributed mode coupling might be difficult to realize. However, a new promising approach using gratings during the fabrication process has been introduced in the literature [166]. This needs to be studied for more mode groups. Spatial mode couplers also introduce a small portion of mode coupling. MSs represent an option to replace fiber with inherently strongly coupled modes. By employing MSs in particular positions along the link makes a great improvement in the system performance and MDL accumulation.

5. Receiver schemes for reducing the impact of MDL

As seen in the previous chapter, strong mode coupling is advantageous to improve the system performance of MDL-impaired systems. In Chapter 4, the different mode coupling scenarios were simulated with zero-forcing equalization. This equalizer is simple, but the matrix inversion that is involved in the equalization (see Eq. (4.4)) amplifies the noise if the magnitude of the estimated channel matrix contains small values.

This chapter is based on the work published in [143,146,171] and studies the effect on system performance when different equalization and detection schemes are employed. Three RX scheme alternatives are compared with linear equalization: ML detection, improved reduced-search ML (IRSML) detection [172,173] and sphere decoding (SD) [174, 175]. The two latter schemes were initially proposed for wireless systems, but they represent good candidates for improving the performance of MDL-impaired systems. Despite the fact that the channel matrix in wireless systems is different from the one in FMF systems, they both share the particularity of being nonunitary matrices, in wireless because of reflections, refraction and scattering, and in FMF systems because of MDL.

Based on simulations, three- and six-spatial-moded-OFDM RXs are assessed in the presence of MDL in terms of performance and computational complexity, which is a fundamental aspect for practical implementation. Two different modulation formats are presented resulting in 3×158.4 Gb/s and 6×158.4 Gb/s MDM-QPSK-OFDM and 3×316.8 Gb/s and 6×316.8 Gb/s MDM 16-level quadrature amplitude modulation (16QAM) OFDM (MDM-16QAM-OFDM).

5.1. Simulation setup

The simulated system transmits an OFDM signal using three or six spatial modes for MDM. Through its CP, OFDM allows for the removal of the ISI introduced by mainly the DMD. Therefore, equalization and detection can be done separately on each subcarrier, and thus, it is less complex than when using single-carrier schemes. The system setup is similar to the one presented in Chapter 4, i.e. similar to the diagram in Figs. 3.28 and 4.2.

Each TX generates a POLMUX signal with a net data rate of 3×100 Gb/s in case of QPSK and three spatial modes. Accordingly, the six-mode system with twelve tributaries allows for a total of 6×100 Gb/s by using QPSK. For 16QAM, bit rates of 3×200 Gb/s and 6×200 Gb/s, respectively, are achieved. These calculations follow the convention *Number of spatial modes* × *Net data rate per POLMUX signal*. With the same overhead described in Section 4.1 the gross bit rate results in 3×158.4 Gb/s and 6×158.4 Gb/s for three and six-moded QPSK transmission. This results from multiplying the net data rate by the overheads originated from TS (10 %), CP (16.1 %) and FEC (24 %). In the same manner and using the same overhead mentioned before, for 16QAM the gross symbol rate correspond to 3×316.8 Gb/s and 6×316.8 Gb/s, respectively.

The block diagram of the simulated system is illustrated in Fig. 5.1. The TX is identical as in Fig. 3.28. The number of modulated subcarrier is maintained as in Chapter 4, i.e. from a total of 4096 subcarriers (N_{FFT}), 81.3% corresponding to 3328 subcarriers are modulated (N_{DATA}). The remaining non-modulated subcarriers are used for zero-padding. The RX structure is similar but not the same. As illustrated in the figure the MMSE equalization requires a separate detection stage, while the output of the other ML and near-ML schemes studied in this chapter delivers already detected symbols.

The optical link structure follows the schematic in Fig. 4.2. As seen in Chapter 4 strongly coupled modes are favorable for system performance and inherently coupled modes is challenging to achieve [8]. Here, the simulations are carried out with a FMF supporting weakly coupled modes reinforced by MSs as Fig. 4.2(b). The FMF is modeled identically as in Chapter 4.

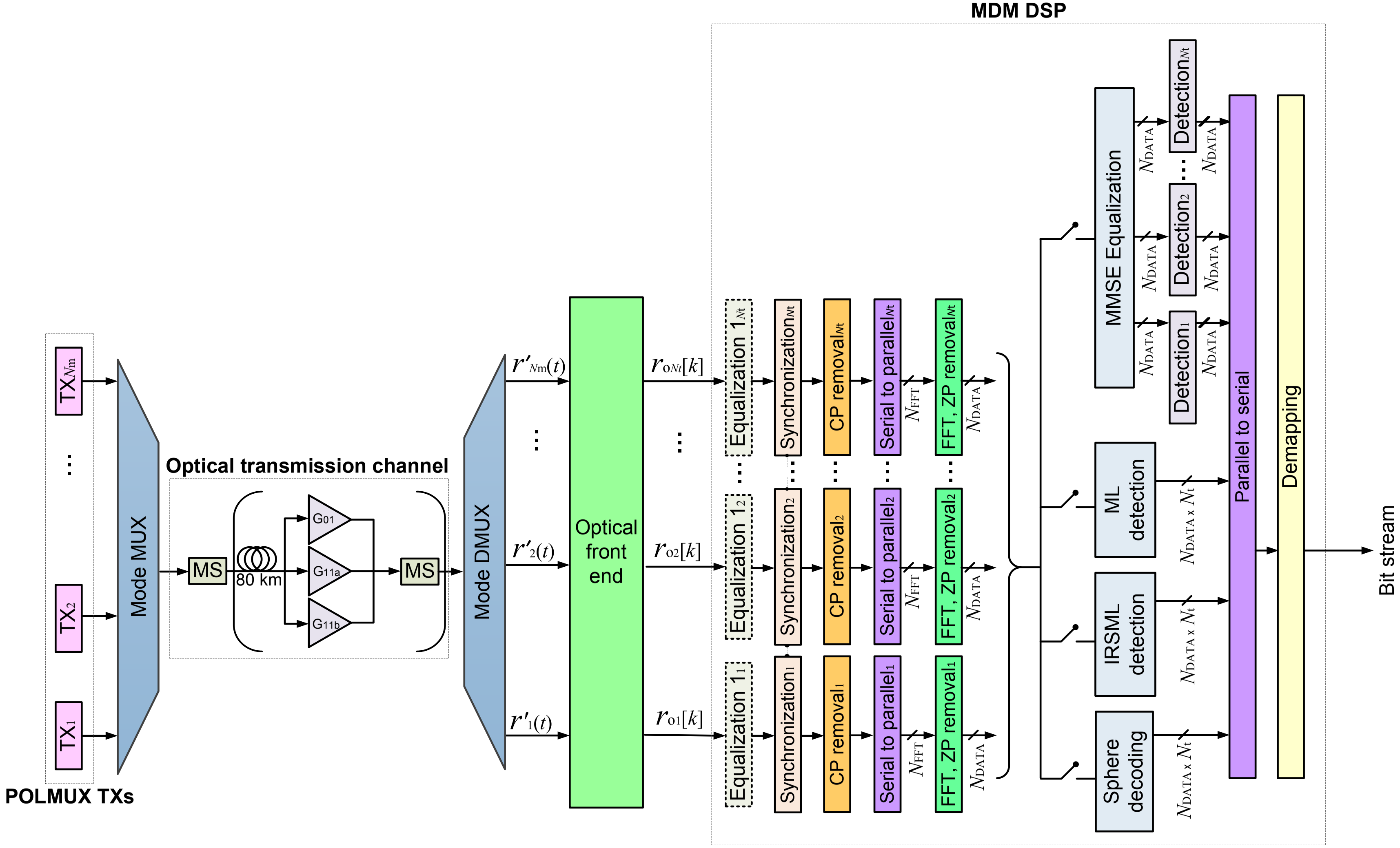


Figure 5.1.: Block diagram of the simulated system.

5.1.1. Detection schemes

This section describes the different RX schemes in charge of equalization and detection. First, the linear minimum-mean square error (MMSE) equalizer and ML detector are presented. Afterwards, two options that achieve near-ML detection, namely IRSML and SD, are investigated. The channel model is assumed to be linear time-invariant and disturbed by AWGN as described in Section 3.2. For the purpose of exploiting the advantage of one-tap channel equalization provided by OFDM, channel and the subsequent equalization and detection definitions apply subcarrier-wise.

MMSE Equalizer

For the equalization of the spatial and polarization modes, the linear MMSE equalizer is used in this chapter because it suffers less from noise amplification than the ZF equalizer. It allows for the best possible compromise between canceling the distortion and noise amplification. Its equalizer matrix $\mathbf{G}_{\mathrm{MMSE}} \in \mathbb{C}^{N_{\mathrm{t}} \times N_{\mathrm{t}}}$ is given as (e.g. [176])

$$\mathbf{G}_{\mathrm{MMSE}} = \left(\mathbf{H}^{\mathrm{H}}\mathbf{H} + \frac{\sigma_{\boldsymbol{\eta}}^2}{\sigma_{\mathbf{s}}^2}\mathbf{I}_{N_{\mathrm{t}}} \right)^{-1} \mathbf{H}^{\mathrm{H}} \ . \tag{5.1}$$

In Eq. (5.1), $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{t}} \times N_{\mathrm{t}}}$, $\sigma_{\mathbf{s}}^2$ and $\sigma_{\boldsymbol{\eta}}^2$ correspond to the channel matrix, variances of the transmitted signal and noise, respectively. The equalization is performed by multiplying the equalizer matrix $\mathbf{G}_{\mathrm{MMSE}}$ with the received signal $\mathbf{r}_{\mathrm{o}} \in \mathbb{C}^{N_{\mathrm{t}}}$ as

$$\mathbf{y}_{\mathrm{MMSE}} = \mathbf{G}_{\mathrm{MMSE}}\mathbf{r}_{\mathrm{o}} \ . \tag{5.2}$$

In Eq. (5.2), $\mathbf{y}_{\mathrm{MMSE}} \in \mathbb{C}^{N_{\mathrm{t}}}$ are the equalized symbols from all tributaries.

Maximum-likelihood detection

The optimum receiver is provided by the maximum *a posteriori* (MAP) criterion [176], which is equivalent to the ML decisions if the input symbols have equal *a priori* probabilities, as in this case. The ML detector is based on the maximization of the probability $P(\mathbf{s} \mid \mathbf{r}_{\mathrm{o}})$ for all input hypersymbols $\mathbf{s}$. In case of an AWGN channel, this leads to the well-known form of the ML detection scheme

$$\hat{\mathbf{s}}_{\mathrm{ML}} = \arg\min_{\mathbf{s}\in\mathbb{M}^{N_{\mathrm{t}}}} \|\mathbf{r}_{\mathrm{o}} - \mathbf{H}\mathbf{s}\|^2 \tag{5.3}$$

Equation 5.3 implies the extensive search of all possible combinations of input vectors $\mathbf{s} \in \mathbb{M}^{N_{\mathrm{t}}}$, where $\mathbb{M}$ is the set of symbols from the QPSK or 16QAM constellation. For this reason the computational effort rises exponentially with the constellation size and the number of tributaries (see Section 5.2.2). For the number of tributaries and constellation sizes used, this approach is considered to be unfeasible in a practical scenario.

Improved reduced-search ML detection

Algorithms having near-ML performance and at the same time having less computational complexity than ML have been a focus of research in the past years [172, 173, 177, 178]. In this thesis, the approach IRSML is investigated [172, 173], since it achieves near-ML performance with only few preprocessing stages before the ML stage and it has a complexity that can be fixed independently of the channel conditions. The goal of the IRSML algorithm is to find a most reliable search set with a smaller set of hypersymbol hypotheses than the brute-force ML detection. It is referred to as IRSML detector as it is an improved version of the algorithm proposed earlier by the authors in [179]. IRSML offers better performance and higher flexibility regarding different constellation formats compared to the RSML algorithm.

The structure of the IRSML detector is shown in Fig. 5.2. The output of the MMSE equalizer feeds a module that constructs the search set for the reduced search ML stage of the algorithm characterized by

$$\hat{\mathbf{s}}_{\mathrm{IRSML}} = \arg\min_{\mathbf{s}\in\mathbb{S}} \|\mathbf{r}_{\mathrm{o}} - \mathbf{H}\mathbf{s}\|^2 \tag{5.4}$$

where $\mathbb{S}$ is a subset of $\mathbb{M}^{N_{\mathrm{t}}}$ and corresponds to the Cartesian product of the tributary search sets as $\mathbb{S}_1 \times \ldots \times \mathbb{S}_{N_{\mathrm{t}}}$, where $\mathbb{S}_i$ denotes the search set of the i-th tributary [172, 173]. The algorithm aims to gather those constellation points of each tributary into the search set that includes the ML solution with high probability.

The search set $\mathbb{S}$ is fed into the ML detector, where $\mathbb{S}$ is constructed and expanded according to the MAP criterion by finding the next hypothetical symbol s_i with

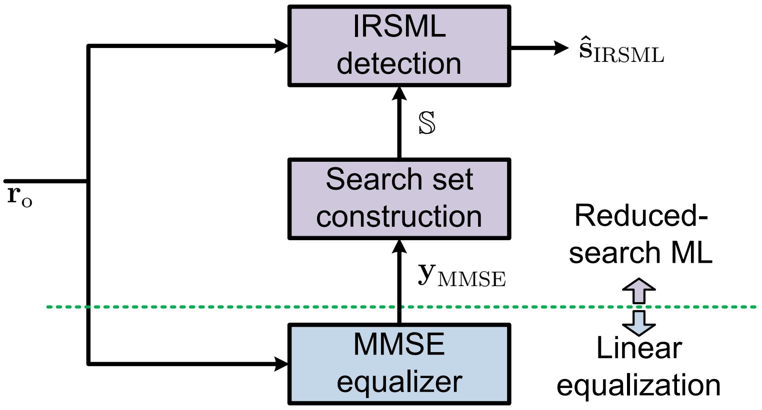


Figure 5.2.: Block diagram of the IRSML detection algorithm.

the highest probability $P(s_i \mid y_{\mathrm{MMSE}_i})$ for $i = 1, ..., N_{\mathrm{t}}$ as derived in [172, 173]. $P(s_i \mid y_{\mathrm{MMSE}_i})$ is the probability that the symbol s_i is transmitted given that y_{MMSE_i} is received for tributary i. Using Bayes' rule, $P(s_i \mid y_{\mathrm{MMSE}_i})$ for $i = 1, ..., N_{\mathrm{t}}$ can be expanded as

$$P(s_i \mid y_{\mathrm{MMSE}_i}) = \frac{p(y_{\mathrm{MMSE}_i} \mid s_i)P(s_i)}{p(y_{\mathrm{MMSE}_i})} \ , \tag{5.5}$$

where $p(\cdot)$ and $P(\cdot)$ denote probability density function and probability, respectively. The probability density $p(y_{\mathrm{MMSE}_i})$ and denominator in Eq. (5.5) can be expanded according to the law of total probability [180] as

$$p(y_{\mathrm{MMSE}_i}) = \sum_{s'_i \in \mathbb{S}} p(y_{\mathrm{MMSE}_i} \mid s'_i)P(s'_i) \ . \tag{5.6}$$

Equation 5.6 is replaced in Eq. (5.5) resulting in

$$P(s_i \mid y_{\mathrm{MMSE}_i}) = \frac{p(y_{\mathrm{MMSE}_i} \mid s_i)}{\sum_{s' \in \mathbb{S}} p(y_{\mathrm{MMSE}_i} \mid s'_i)} \ . \tag{5.7}$$

Note that in Eq. (5.7) the probability $P(s_i)$ is canceled out, since all the hypotheses s_i have the same probability of being sent. The denominator in Eq. (5.7) involves all the hypotheses and cannot be maximized. The numerator in Eq. (5.7) can be written as [172, 173]

$$p(y_{\mathrm{MMSE}_i} \mid s_i) = \frac{1}{\pi} e^{-\left|y_{\mathrm{MMSE}_i} - s_i\right|^2} \ . \tag{5.8}$$

Equation 5.7 can be maximized by maximizing the probability density in Eq. (5.8). This implies the computation of the Euclidean distances $\delta_{\mu,i}$ between the MMSE equalizer output $\mathbf{y}_{\mathrm{MMSE}_i}$ and the hypothesis $\mathbf{s}_{\mathrm{MMSE}}$

$$\delta_{\mu,i} = \left| y_{\mathrm{MMSE}_i} - s_{\mu,i} \right| , \tag{5.9}$$

where $\mu = 1, \ldots, M$ is the index for all possible constellation points; accordingly, $s_{\mu,i}$ refers to one possible constellation point of index μ and tributary i [172, 173]. The search set is constructed by including in the set $\mathbb{S}_i$ the hypothesis $s_{\mu,i}$ that gives the smallest $\delta_{\mu,i}$. Excluding the already selected hypothesis in $\mathbb{S}_i$, the process of adding another hypersymbol candidate to the corresponding set $\mathbb{S}_i$ is repeated numerous times until the size of the set $\mathbb{S}$ reaches a maximum number of hypersymbol candidates $N_{\max}$ or until the size of $\mathbb{S}$ exceeds $N_{\max}$ by adding another candidate. The number $N_{\max}$ is chosen according to required performance and allowed computational complexity.

Note that the algorithm described above requires the computation of all possible $\delta_{\mu,i}$, i.e. per tributary for all constellation points. If there are complexity constraints to perform this operation due to a large number of tributaries and constellation points, it should be considered to reduce the effort of computing all these distances. In [173], it is proposed a method for considering only a limited predefined number of neighboring constellation points to y_{MMSE_i}. Thus, $\delta_{\mu,i}$ is calculated with the closest constellation points to y_{MMSE_i}.

Sphere decoding

The idea behind the SD algorithm is to achieve ML-performance by searching the closest point in a lattice to the received signal with the condition that it should lie within the radius R of a hypersphere [174, 175]. Equation 5.10 and 5.11 represent this constraint [181, 182], where $\mathbb{S}_{R'}(\mathbf{r}_{\mathrm{o}}, \mathbf{H}) \subset \mathbb{M}^{N_{\mathrm{t}}}$, $\mathbf{r}_{\mathrm{o}} \in \mathbb{C}^{N_{\mathrm{t}}}$, $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{t}} \times N_{\mathrm{t}}}$, $\mathbf{s} \in \mathbb{C}^{N_{\mathrm{t}}}$ and $\hat{\mathbf{s}}_{\mathrm{SD}} \in \mathbb{C}^{N_{\mathrm{t}}}$ correspond to the set of possible candidates for performing the SD algorithm, received signal, channel matrix, input hypersymbol candidate and output detection vector of the SD algorithm, respectively.

$$\hat{\mathbf{s}}_{\mathrm{SD}} = \underset{\mathbf{s} \in \mathbb{S}_{R'}(\mathbf{r}_{\mathrm{o}}, \mathbf{H})}{\arg\min} \left\| \mathbf{r}_{\mathrm{o}} - \mathbf{H}\mathbf{s} \right\|^2 . \tag{5.10}$$

$$\mathbb{S}_{R'}(\mathbf{r}_\mathrm{o},\mathbf{H}) = \{\mathbf{s} \in \mathbb{M}^{N_\mathrm{t}}, \|\mathbf{r}_\mathrm{o} - \mathbf{H}\mathbf{s}\|^2 \leq R'^2\} \tag{5.11}$$

By decomposing the channel matrix $\mathbf{H}$ into a unitary matrix $\mathbf{Q} \in \mathbb{C}^{N_\mathrm{t}\times N_\mathrm{t}}$ and an upper triangular matrix $\mathbf{R} \in \mathbb{C}^{N_\mathrm{t}\times N_\mathrm{t}}$ with the so-called QR decomposition [72], another perspective of Eq. (5.10) is obtained. Furthermore, after replacing $\mathbf{H}$ by $\mathbf{QR}$, both addends in Eq. (5.10) can be multiplied by $\mathbf{Q}^\mathrm{H}$, since the multiplication with a unitary matrix does not change the l_2-norm, leading to an equivalent form of Eq. (5.10) and Eq. (5.11):

$$\hat{\mathbf{s}}_\mathrm{SD} = \underset{\mathbf{s}\in\mathbb{S}_R(\mathbf{w},\mathbf{R})}{\arg\min} \|\mathbf{w} - \mathbf{R}\mathbf{s}\|^2 . \tag{5.12}$$

$$\mathbb{S}_{R}(\mathbf{w},\mathbf{R}) = \{\mathbf{s} \in \mathbb{M}^{N_\mathrm{t}}, \|\mathbf{w} - \mathbf{R}\mathbf{s}\|^2 \leq R^2\} \tag{5.13}$$

where $\mathbf{w} = \mathbf{Q}^\mathrm{H}\mathbf{r}_\mathrm{o}$.

Now with $\mathbf{R}$ being triangular, i.e. $R_{kl} = 0\,\forall\, l < k$, $\mathbf{Q}^\mathrm{H}\mathbf{r}_\mathrm{o} - \mathbf{R}\mathbf{s}$ can be rewritten as Eq. (5.14), where $\mathbf{q}_k$ corresponds to the k-th column of the matrix $\mathbf{Q}$.

$$\left\|\mathbf{Q}^\mathrm{H}\mathbf{r}_\mathrm{o} - \mathbf{R}\mathbf{s}\right\|^2 = \sum_{k=1}^{N_\mathrm{t}} \left| \mathbf{q}_k^\mathrm{H}\mathbf{r}_\mathrm{o} - \sum_{l=1}^{N_\mathrm{t}} R_{kl}s_l \right|^2 = \sum_{k=1}^{N_\mathrm{t}} \left| \mathbf{q}_k^\mathrm{H}\mathbf{r}_\mathrm{o} - \sum_{l=k}^{N_\mathrm{t}} R_{kl}s_l \right|^2 \tag{5.14}$$

As evident from Eq. (5.14) the addend for tributary N_t only depends on s_{N_t}, i.e. for $k = N_\mathrm{t} : \left|\mathbf{q}_{N_\mathrm{t}}^\mathrm{H}\mathbf{r}_\mathrm{o} - R_{N_\mathrm{t},N_\mathrm{t}}s_{N_\mathrm{t}}\right|^2$. The rest of the addends for $N_\mathrm{t} - 1$, $N_\mathrm{t} - 2$ and so on, only depend on candidate symbols in their corresponding tributary number and higher tributary number, e.g. as

$$\begin{aligned}
k =& N_\mathrm{t} : && \left|\mathbf{q}_{N_\mathrm{t}}^\mathrm{H}\mathbf{r}_\mathrm{o} - R_{N_\mathrm{t},N_\mathrm{t}}s_{N_\mathrm{t}}\right|^2 \\
k =& N_\mathrm{t} - 1 : && \left|\mathbf{q}_{N_\mathrm{t}-1}^\mathrm{H}\mathbf{r}_\mathrm{o} - R_{N_\mathrm{t}-1,N_\mathrm{t}-1}s_{N_\mathrm{t}-1} - R_{N_\mathrm{t}-1,N_\mathrm{t}}s_{N_\mathrm{t}}\right|^2 \\
k =& N_\mathrm{t} - 2 : && \left|\mathbf{q}_{N_\mathrm{t}-2}^\mathrm{H}\mathbf{r}_\mathrm{o} - R_{N_\mathrm{t}-2,N_\mathrm{t}-2}s_{N_\mathrm{t}-2} - R_{N_\mathrm{t}-2,N_\mathrm{t}-1}s_{N_\mathrm{t}-1} - R_{N_\mathrm{t}-2,N_\mathrm{t}}s_{N_\mathrm{t}}\right|^2 \\
&\vdots && \vdots \\
k =& 1 : && \left|\mathbf{q}_1^\mathrm{H}\mathbf{r}_\mathrm{o} - R_{1,1}s_1 - \ldots - R_{1,N_\mathrm{t}}s_{N_\mathrm{t}}\right|^2 .
\end{aligned}$$

It is important to note that only the last addend for $k = 1$ depends on all the elements of the candidate hypersymbol $\mathbf{s}$ and because of the absolute value in Eq. (5.14) every term adds up monotonically. For this reason the outer summation in Eq. (5.14) can be performed starting from $k = N_\mathrm{t}$ and can be interrupted as soon as it reaches a number which surpasses the radius of the hypersphere R. Thus, a significant number of candidates can be discarded. This algorithm is also described in the literature as a tree search [183]. The decision tree is graphically represented in Fig. 5.3, where an example is illustrated for $N_\mathrm{t} = 3$ and QPSK modulation. The leaves at the bottom of the tree represent all possible symbol combinations from all tributaries and constellation points belonging to $\mathbb{M}^{N_\mathrm{t}}$. The red lines Fig. 5.3 symbolize that the tree was pruned above the node, since the partial Euclidean distance calculated until then surpassed SD constrain in Eq. (5.13).

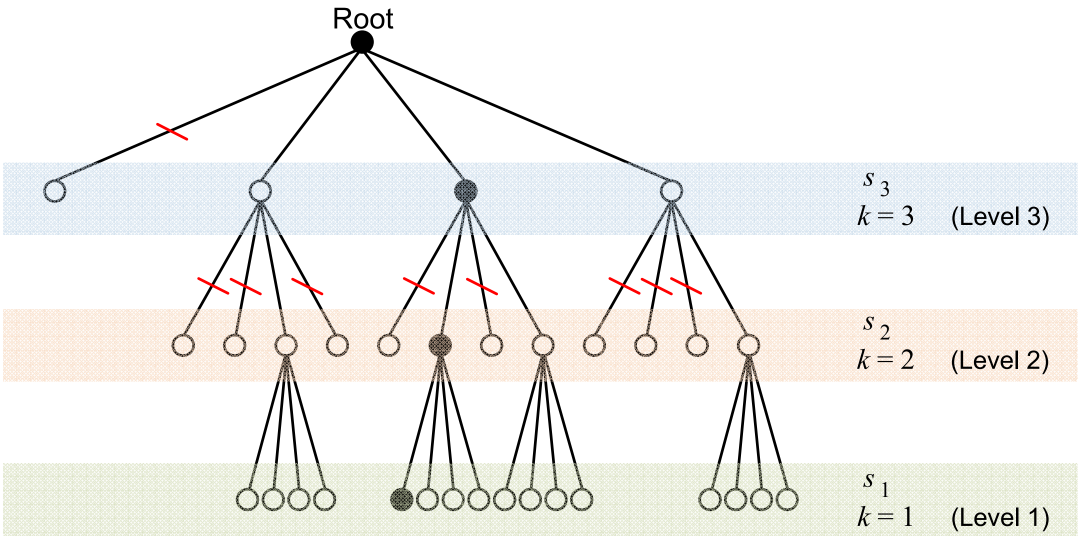


Figure 5.3.: Symbol vector combinations $\mathbf{s}$ organized in the SD decision tree. Here, as an example, $N_\mathrm{t} = 3$ and QPSK modulation is used. Black filled circles symbolize the symbol vector combination $\hat{\mathbf{s}}_\mathrm{SD}$ with less metric that satisfies the SD constrain in Eq. (5.13).

When traversing the tree at certain level, it is convenient to choose the most appropriate symbol to calculate the next metric with. A method proposed by Schnorr and Euchner tackles this topic in [184]. The Schnorr-Euchner enumeration aims to order the constellation points in a way such that always the next symbol to be considered gives the smallest partial Euclidean distance among the remaining symbols with the same parent[1]. In this way, the most probable symbol to belong to the ML solution

---

[1] The term *parent* refers to a node at a higher level. Lower level symbols are referred as *children*.

is examined, and therefore, less symbol candidates are considered making the tree search efficient. This enumeration method is used in this thesis.

The appropriate choice of the initial hypersphere radius R is crucial. If R is too small, no solution will satisfy Eq. (5.13) and $\hat{\mathbf{s}}_{\mathrm{SD}}$ will result in an empty vector. R has to be selected so that at least one solution $\hat{\mathbf{s}}_{\mathrm{SD}}$ can be found. If R is oversized the complexity of the algorithm can be even greater than the ML algorithm due to the additional operations that the SD has to perform, such as the QR decomposition. Following the approach in [185], a starting point for the tree search corresponds to $\mathbf{e} \in \mathbb{C}^{N_{\mathrm{t}}}$ Eq. (5.15), where $\lfloor \cdot \rceil$ denotes rounding to the closest constellation point. The hypersymbol $\mathbf{e}$ might not achieve the ML solution because of the noise amplification through $\mathbf{H}^{-1}$, resembling the ZF approach, but can be close to it. Accordingly, the search starts with an infinite radius R, which is shrunk firstly with the radius resulting from the starting point $\mathbf{e}$ and then is updated if a lower value is found with another symbol combination.

$$\mathbf{e} = \lfloor \mathbf{R}^{-1}\mathbf{w} \rceil = \lfloor \mathbf{R}^{-1}\mathbf{Q}^{\mathrm{H}}\mathbf{r}_{\mathrm{o}} \rceil = \lfloor \mathbf{H}^{-1}\mathbf{r}_{\mathrm{o}} \rceil \tag{5.15}$$

The search stops when no other hypersymbol candidates $\mathbf{s}$ satisfy the SD criterion in Eq. (5.13). At the end, $\hat{\mathbf{s}}_{\mathrm{SD}}$ will contain the same solution as in $\hat{\mathbf{s}}_{\mathrm{ML}}$.

5.2. Simulation results

In this section the receiver schemes described in Section 5.1 are compared from the system performance and the computational complexity point of view.

5.2.1. System performance

The system performance is evaluated via the OSNR penalty at a TBER of 1e-3 and the maximum transmission distance for an OSNR penalty of 1 dB. The modulation formats QPSK-OFDM and 16QAM-OFDM on three and six spatial modes were simulated with the maximum modal differences of 2 dB and 4 dB, respectively, in the FMA stage as in Chapter 4.

Figure 5.4 compares the performance of the different equalization and detection algorithms, using weakly coupled modes with and without MSs in the optical link, QPSK and three spatial modes. As also seen in Chapter 4, the OSNR penalty increases with the distance because of the accumulated MDL. Figure 5.4 depicts two example parameters $N_{\max}$, since the IRSML algorithm offers the possibility to vary the size of the search set. As mentioned in Section 5.1, for SD the initial radius is set to infinity and then is reduced with the radius resulting from the first candidate $\mathbf{e}$.

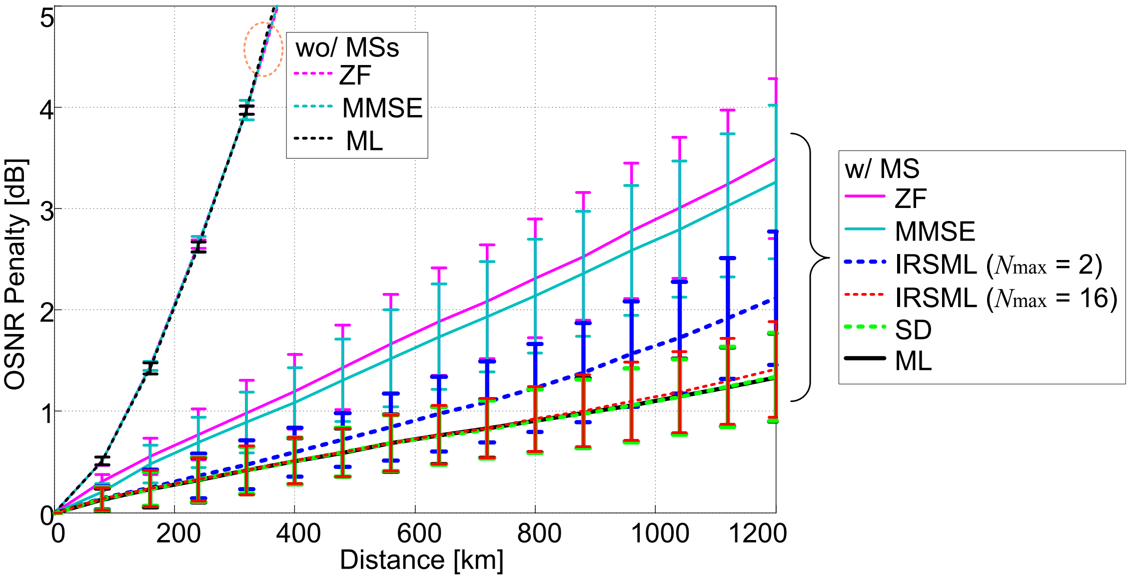


Figure 5.4.: OSNR penalty at a TBER of 1e-3 for the different equalization and detection schemes, three spatial modes, QPSK, weakly coupled modes with and without MSs.

In Fig. 5.4, it can be observed that when no MSs are used the curves from the linear and nonlinear equalizer schemes overlap. In this case, the ML detection cannot outperform the linear equalization, since the modes remain virtually orthogonal, i.e. make up independent data streams. It can also be seen in the plot that when there is strong mode coupling there is a tremendous increase in the distance of almost two-fold the maximum transmission distance at 1 dB OSNR by changing the equalizer from ZF to ML. Comparing ZF and MMSE, MMSE offers slightly better results than ZF. ML and near-ML schemes provide even better results than MMSE, since these detection schemes do not generate noise amplification as MMSE with its inversion operation.

Figure 5.4 shows that both SD and IRSML with sufficient number of hypersymbol candidates $N_{\max}$ achieve near-ML performance. As $N_{\max}$ increases, the IRSML detection curve approaches the performance of the ML detection. In this case of the system simulated for Fig. 5.4, six tributaries and QPSK, the total number of hypersymbol candidates and hence, the largest possible $N_{\max}$, is $4^6 = 4096$. Already allowing for only two candidates yields an increase of the transmission distance of more than 80 % at 1 dB of OSNR penalty with respect to the MMSE curve.

The results for the rest of the cases are summarized in Fig. 5.5, where the different detection schemes and modulation formats are illustrated. The curves represent the average reach allowing an OSNR penalty of 1 dB. Figure 5.5 shows how the reach of the IRSML detection approaches the ML performance with increasing search set size. In Fig. 5.5(a), the MMSE equalization poses the lower bound around 365 km, while ML and SD allow for transmission over more than double this distance.

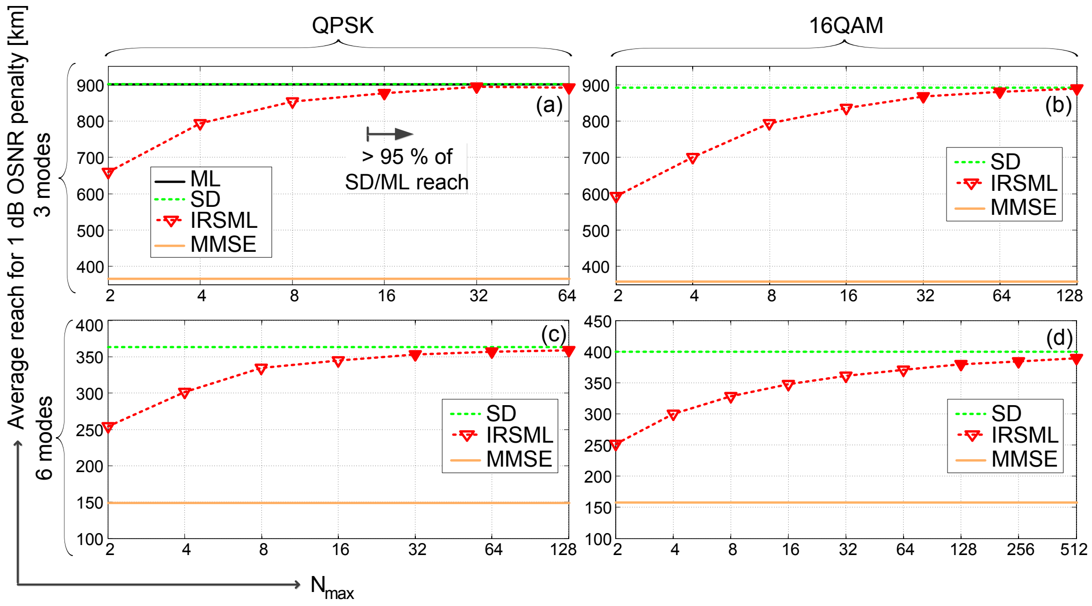


Figure 5.5.: Average reach at an OSNR penalty of 1 dB. Solid markers corresponds to a reach greater than 95 % of the SD/ML reach.

Because of the tremendous computational complexity, it was possible to employ the brute-force ML detection only for the case of three modes and QPSK. Therefore, the SD algorithm will serve as a reference for ML performance. This can be supported by its theoretical derivation (see Section 5.1.1), and it was furthermore confirmed with the near-ML performance achieved in the case of three modes and QPSK, as shown in Fig. 5.4 and 5.5(a).

In addition, Fig. 5.5 highlights with the solid markers the points in the IRSML curve which achieve more than 95 % of the corresponding ML or SD reach. This is used as a criterion for choosing $N_{\max}$ with near-ML performance. For three spatial modes, IRSML achieves this transmission distance with minimum 16 and 32 candidates for

QPSK and 16QAM, respectively; for spatial six modes, these values correspond to 32 and 128 candidates for QPSK and 16QAM, respectively.

5.2.2. Computational complexity

This section shows a comparison of the computational complexity of the different equalization and detection schemes described in the previous sections. Here, the complexity is computed per subcarrier and is defined as the number of complex multiplications needed to detect one bit [57, 110].

MMSE equalizer

According to Eq. (5.1), the MMSE equalizer needs one complex matrix inversion and two matrix multiplications to calculate the filter matrix $\mathbf{G}_{\mathrm{MMSE}} \in \mathbb{C}^{N_\mathrm{t} \times N_\mathrm{t}}$. The matrix inversion is assumed to be performed using the LU matrix factorization method, which is a well-known method with good stability and computational efficiency properties [72]. By means of the LU matrix factorization, a total of

$$\frac{N_\mathrm{t}^3}{3} + N_\mathrm{t}^2 =: c_1 \tag{5.16}$$

complex multiplications are needed for the matrix inversion [72]. As for the two matrix multiplications in Eq. (5.1) ($\mathbf{H}^\mathrm{H}\mathbf{H} \in \mathbb{C}^{N_\mathrm{t} \times N_\mathrm{t}}$ and the multiplication of $\mathbf{H}^\mathrm{H} \in \mathbb{C}^{N_\mathrm{t} \times N_\mathrm{t}}$ on its inverse) $2N_\mathrm{t}^3$ complex multiplications are needed. It is assumed that the computation of $\mathbf{G}_{\mathrm{MMSE}}$ is only done every time the channel is estimated, i.e. every N_f OFDM symbols. Whereas the operation in Eq. (5.2) is necessary for every input vector, which requires N_t^2 complex multiplications. Therefore, The MMSE equalizer computes c_2 complex multiplications per bit as in Eq. (5.17). The denominator in Eq. (5.17) corresponds to the total number of bits that can simultaneously be equalized per subcarrier.

$$\frac{\frac{1}{N_\mathrm{f}}(c_1 + 2N_\mathrm{t}^3) + N_\mathrm{t}^2}{N_\mathrm{t} \log_2 M} =: c_2 \tag{5.17}$$

ML detection

The ML detection described in Eq. (5.3) requires the computation of the Euclidean distances between the M^{N_t} possible input vectors and the received signal vector. The multiplication of the vector $\mathbf{s} \in \mathbb{C}^{N_\mathrm{t}}$ with the channel matrix $\mathbf{H} \in \mathbb{C}^{N_\mathrm{t} \times N_\mathrm{t}}$ requires N_t^2 complex multiplications for each different candidate $\mathbf{s}$. These operations allow also for the simultaneous detection of $N_\mathrm{t} \log_2 M$ bits. Therefore, the ML detector computes

$$\frac{M^{N_\mathrm{t}} N_\mathrm{t}^2}{N_\mathrm{t} \log_2 M} \tag{5.18}$$

complex multiplications per subcarrier and per bit.

IRSML detection

As described in Section 5.1.1, the IRSML scheme can be divided into two stages. First, it equalizes the signal with the MMSE equalizer, which requires c_2 complex multiplications per bit. Then, ML detection is performed, whose complexity corresponds to Eq. (5.18), although the factor M^{N_t} is replaced by N_max due to the reduced number of hypothesis employed in the IRSML algorithm. Therefore the complexity of the IRSML is the sum of the complexity of these two stages as shown in Eq. (5.19).

$$c_2 + \frac{N_\mathrm{max} N_\mathrm{t}^2}{N_\mathrm{t} \log_2 M} \tag{5.19}$$

SD algorithm

The complexity of the SD algorithm has been studied since it was proposed in [181, 186]. In [181], the authors demonstrate that the average complexity grows exponentially, not with M^{N_t} as the ML detection, but proportional to $M^{\gamma N_\mathrm{t}}$ for $\gamma \in (0, 1]$. According to [181], $M^{\gamma N_\mathrm{t}}$ corresponds to the number of nodes visited in the search tree or number of hypothesis considered. If the factor γ is much smaller than one, the SD may be feasible even for large numbers of tributaries and high order constellation formats; if γ is one, the algorithm has visited all nodes in the search tree, as the ML detection, however, with higher computational complexity because of additional operations e.g. QR decomposition.

The complexity of the SD algorithm corresponds to the complexity of the QR decomposition as in Eq. (5.20) [72], the inversion of $\mathbf{R}$ for computing the starting point in Eq. (5.15) with the complexity described in Eq. (5.16), and the multiplications required to compute the metrics in Eq. (5.14) for every hypothesis $\mathbf{s}$. Since the first two terms depend only on the channel matrix and not on the received signal, they are calculated only every N_f OFDM symbols. The latter term, the number of multiplications to compute the metrics in Eq. (5.14) denoted by B, has been monitored during the simulations since it depends on the noise level and the channel conditions, i.e. how much MDL is there in the system. B is the sum over the total number of visited nodes of the complex multiplications per visited node. The total complexity of the SD algorithm corresponds to Eq. (5.21).

$$\frac{2N_\mathrm{t}^3}{3} =: c_3 \tag{5.20}$$

$$\frac{\frac{c_3+c_1}{N_\mathrm{f}} + B}{N_\mathrm{t}\log_2 M} \tag{5.21}$$

Comparison

In this section, the computational complexity of the different detection schemes is compared. The parameter N_f for computing Eq. (5.17) and (5.21) is assumed to be 20, allowing 10 % of TS overhead and sending two TS per frame for synchronization and channel estimation purposes. That is, the channel is assumed to be constant over 20 OFDM symbols or 2 μs, which is in accordance with the experimental results shown in [150], where the minimum rate at which the channel estimation should be updated was determined to be approximately 40 kHz, or every 25 μs. Accordingly, N_f can be incremented even further up to 254 OFDM symbols. With such values of N_f for the IRSML detection and SD, the complex multiplications per bit and subcarrier affected by N_f in Eq. (5.19) and (5.21) are virtually negligible. For the MMSE equalizer, the contribution of the part dependent on N_f in Eq. (5.17) is comparable to the rest of the equation.

The complexity for the IRSML detection is depicted in Fig. 5.6. N_max is assumed to be the minimum value in Fig. 5.5, which accomplish 95 % of the reach of ML or SD. For example, for 3 modes and QPSK the N_max employed is 16. Figure 5.6 shows that IRSML achieves near-ML performance with much less complex multiplications than the ML algorithm. For three spatial modes, the IRSML algorithm reduces the complexity of the ML detection by two and five orders of magnitude for QPSK and 16QAM, respectively; for six spatial modes, the complexity reduction is higher reaching five and 12 orders of magnitude for QPSK and 16QAM, respectively.

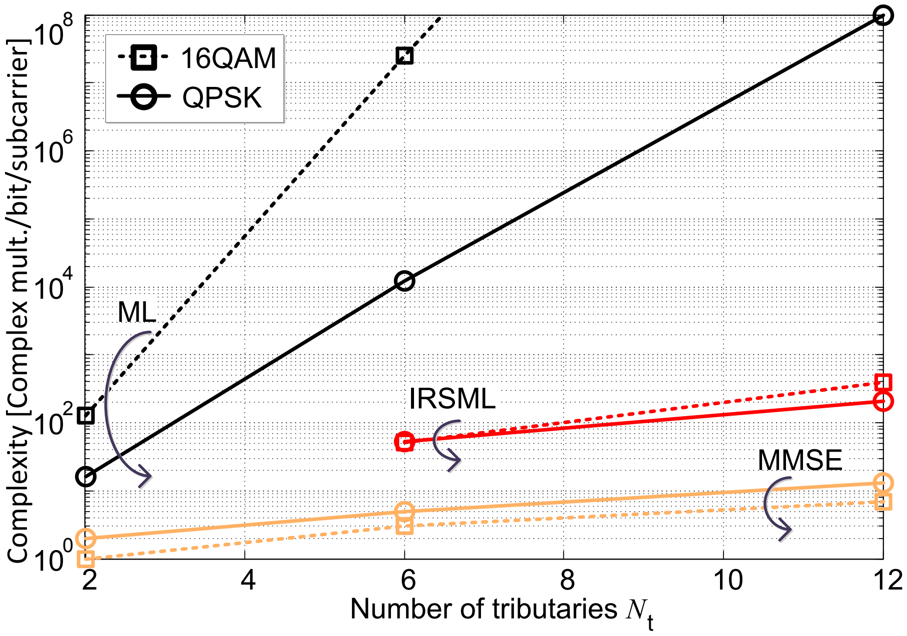


Figure 5.6.: Number of complex multiplications needed for the detection of one bit for one subcarrier.

As the complexity of the SD algorithm is variable it is represented with error bars in Fig. 5.7, where the length of the vertical lines are twice the standard deviation of the data. Figure 5.7(a) shows with the solid lines the average complex multiplications for a TBER of 1e-3; the dotted lines represent the average complexity for values above and below a BER of 1e-3 for six modes and 16QAM. It can be seen that the average complexity of the SD increases proportional to the degradation of the channel conditions. As the inset in Fig. 5.7(a) shows, the MDL accumulates with the square root of the distance, due to the strong mode coupling generated by the MSs as shown in Chapter 4. This demonstrates that both the BER and the MDL influence the complexity of the SD algorithm. The greater the BER and MDL the less unlikely is that the initial hypothesis in Eq. (5.15) is the one that fulfills the minimum argument in Eq. (5.10). Consequently, with badly-conditioned channels the algorithm searches extensively for the ML solution leading to the execution of more complex multiplications.

Whereas Fig. 5.7(a) indicates the average complexity, and therefore, gives an indication of the energy consumption of the algorithm implemented in hardware, the

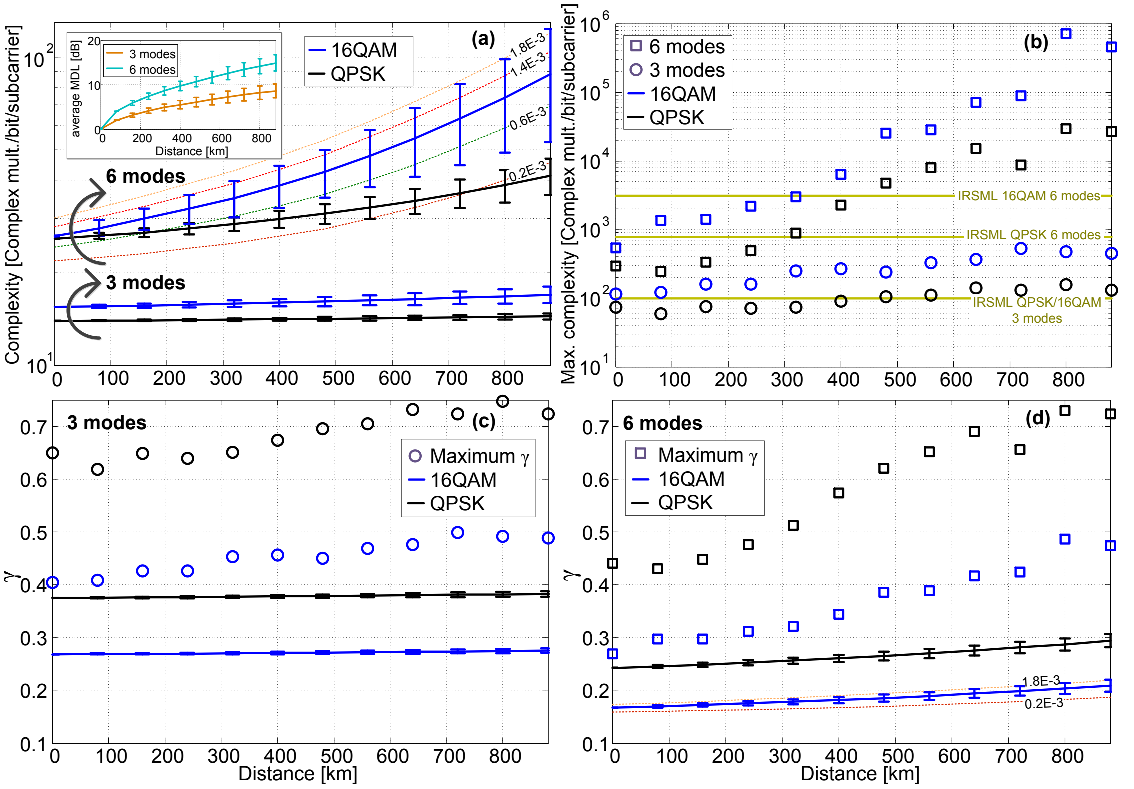


Figure 5.7.: SD complexity as a function of the distance. (a) Average complexity, (b) maximum registered complexity, (c) average and maximum γ for three modes, (d) average and maximum γ for six modes.

maximum complexity indicates the amount of hardware resources that need to be available. As Fig. 5.7(b) shows, the maximum number of complex multiplications are several orders of magnitude above the average complexity, often exceeding the complexity of the IRSML.

Figure 5.7(c) and 5.7(d) illustrates the value of γ, which has also been computed during simulation from the number of nodes visited by the SD algorithm at 1e-3 of BER. The error bars correspond to the average γ and the isolated markers correspond to the maximum γ registered. Additionally, the dotted curves in Fig. 5.7 (d) correspond to the average γ for six modes, 16QAM and a BER of 1.8e-3 and 0.2e-3. Similar to Fig. 5.7(a) and 5.7(b) and as mentioned in [181], the average and the maximum γ increase as the channel conditions worsen and are not longer much smaller than one.

A promising approach to reduce the complexity of SD has been proposed and compared with existing schemes in [182, 187]. The decoder proposed in [182, 187] uses a fixed effort, i.e. without relying on the variable complexity of the SD. However, its complexity reduction has to be traded-off with its performance degradation.

5.3. Discussion

The detection algorithms IRSML and SD have been compared with each other and the linear MMSE equalizer in terms of performance and complexity in 3 × 158.4-Gb/s and 6 × 158.4-Gb/s MDM-QPSK-OFDM and, 3 × 316.8-Gb/s and 6 × 316.8-Gb/s MDM-16QAM-OFDM systems in the presence of MDL.

It has been shown that mode coupling plays an important role for the efficiency of the near-ML algorithms and the reduction of the accumulated MDL. Otherwise, with weakly coupled modes, the data streams are largely orthogonal, which makes the performance of linear and nonlinear equalization very similar.

Both studied algorithms, IRSML and SD, outperform the linear equalizer MMSE. The SD algorithm achieves ML performance with a complexity that in average is lower than the ML detection. However, the complexity of the SD is variable since it depends on the channel conditions, i.e. MDL and SNR, which makes it difficult to implement in hardware. In contrast, the IRSML algorithm has a fixed complexity and can achieve near-ML performance with a fixed and considerable less computational effort than the ML detection. Already by using two hypersymbol candidates for the IRSML detection the 16QAM and six mode system can achieve an improvement of more than 60 % of the average transmission reach over the MMSE equalizer allowing 1 dB OSNR penalty.

6. Experimental demonstration of system tolerance against MDL

So far the beneficial effects of mode coupling and ML-based algorithms on the MDL system tolerance were demonstrated in Chapters 4 and 5 via simulations. As mentioned in Chapter 5, it is important to consider the computational complexity of the equalizer since it is directly related with the RX energy consumption. In Chapter 5, it is demonstrated that the near-ML algorithm IRSML detection not only outperforms the linear equalizer MMSE but also keeps the complexity fixed and at low levels compared to ML detection. In this chapter, a system supporting six spatial modes impaired by MDL is implemented and described to show the better performance of IRSML detection as observed in the previous chapter. Both single-channel (Section 6.1) and WDM channel (Section 6.2) experiments were carried out and described in this chapter based on the publications [130] and [117,188], respectively.

6.1. Single-channel experiments

6.1.1. Experimental setup

Figure 6.1 shows the block diagram of the experimental setup. At the transmitter, the baseband OFDM I and Q signals modulated with QPSK and 8-level quadrature amplitude modulation (8QAM) are loaded into an arbitrary waveform generator (AWG) whose sampling rate is $R_{\mathrm{DAC}} = 10$ GS/s. Throughout this thesis, the abbreviation for *samples* is denoted with the letter S. A picture of the AWG and MZM setup can be observed in Fig. 6.2(b).

From 256 subcarriers (N_{FFT}), 174 are used for data (N_{DATA}), three for quasi-pilot aided (QPA) [133], and three for pilot-aided (PA) phase noise (PN) estimation and compensation [189]. The rest of the subcarriers are employed for ZP, isolating the spectrum from the aliasing products generated by the AWG. These aliasing products are filtered out by a 3.3-GHz 3-dB bandwidth electrical low pass filter (LPF).

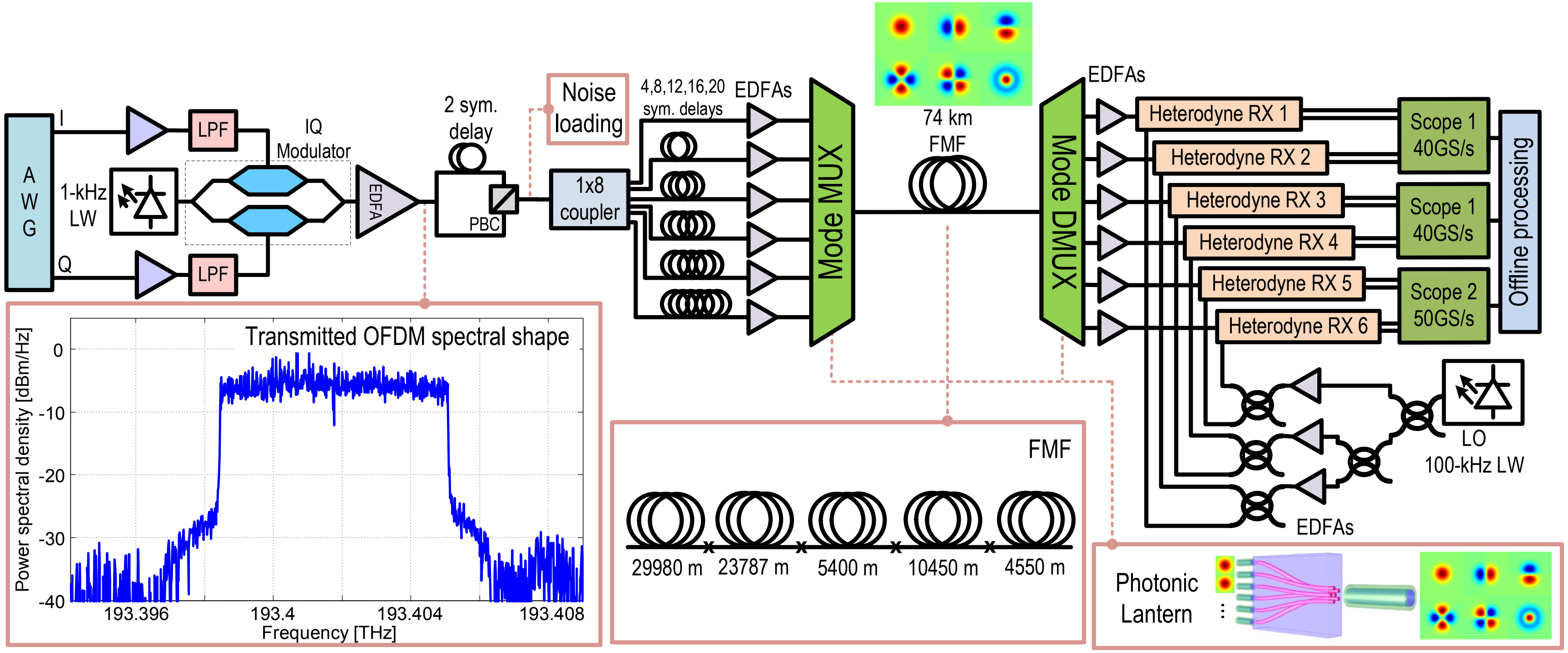


Figure 6.1.: Block diagram of the experimental MDM-OFDM systems (PBC: Polarization beam combiner).

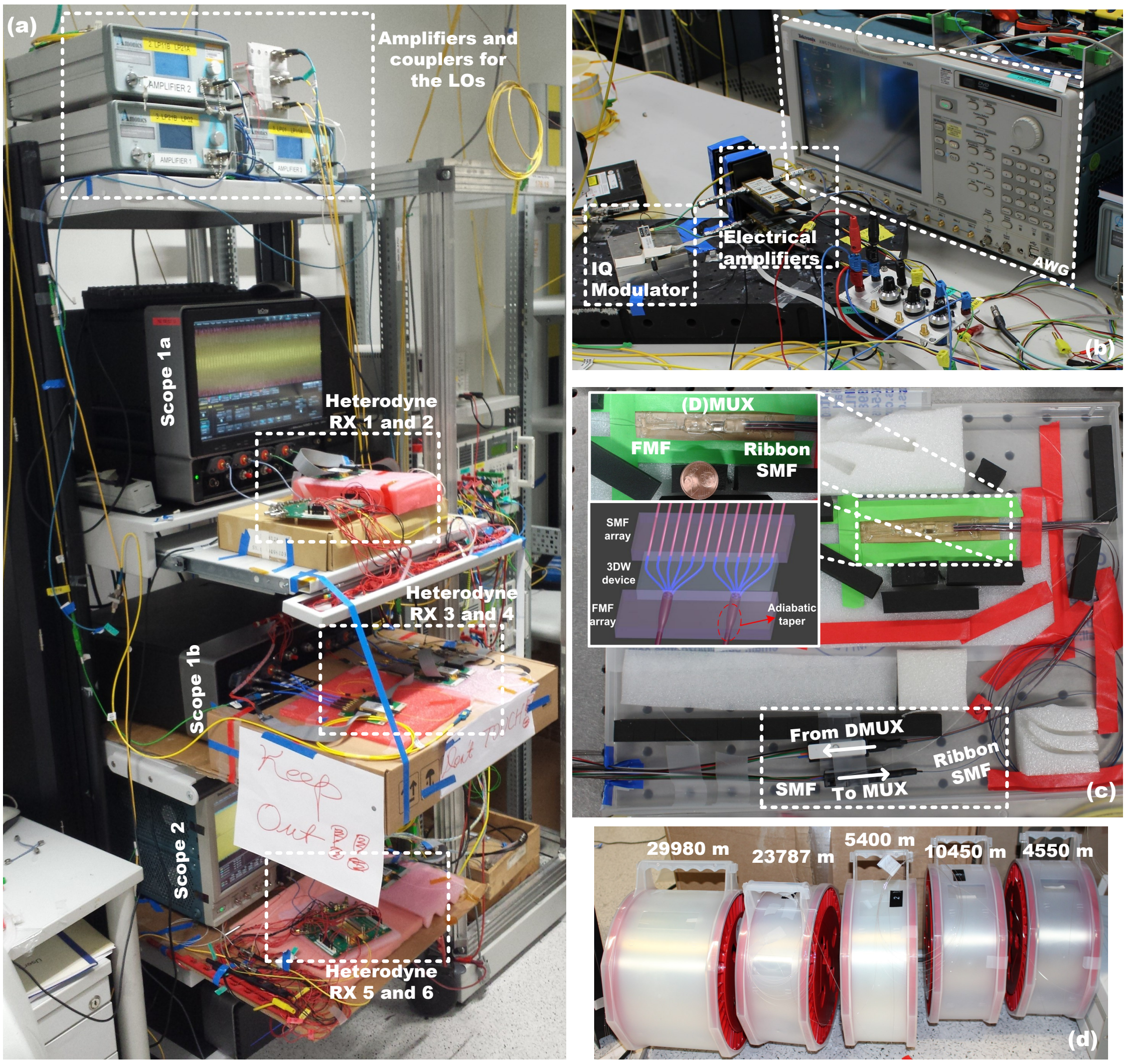


Figure 6.2.: Laboratory setup pictures: (a) Scopes for sampling the received electrical signals from six coherent heterodyne RXs, (b) AWG and IQ modulator setup, (c) photonic lanterns and (d) the five fiber spools used for transmission.

Taking into account the sampling rate $R_{\mathrm{DAC}} = 10$ GS/s, the nominal data rate results in 163 Gb/s and 245 Gb/s for the MDM-QPSK-OFDM and MDM-8QAM-OFDM signals, respectively. The nominal data rate is calculated by multiplying $\frac{N_{\mathrm{DATA}}}{N_{\mathrm{FFT}}} \times R_{\mathrm{DAC}} \times N_{\mathrm{t}} \times N_{\mathrm{bits}}$, where N_{bits} is the number of bits per symbol, i.e. two and three bits for QPSK and 8QAM, respectively. The signal is converted to the optical domain by an IQ MZM fed by a laser at an f_{c} of 193.4 THz, followed by a POLMUX stage having a two OFDM-symbol delay, i.e. a delay of 67 ns. For multiplexing the modes into the FMF, delay lines are included to create time-multiplexed training sequences [127], here, with 4, 8, 12, 16 and 20 OFDM-symbol delay. Every POLMUX TS consists of one pair of identical binary phase-shift keying OFDM (BPSK-OFDM) symbols for synchronization and carrier frequency offset (CFO) compensation purposes described later on.

As mode (D)MUX the fully packaged photonic lanterns described in [68] were used. A picture of the photonic lanterns can be seen in Fig. 6.2(c). The insets of this figure are a close up version of the (D)MUX and a schematic diagram taken from [68]. As shown in Fig. 6.2(c), the photonic lantern consist of three pieces of transparent substrate, which makes the transition from the one-dimensional SMF array to the adiabatically tapered FMF array and vice versa. The MUX and DMUX are integrated in the same transparent substrate. The insertion loss of the (D)MUX ranges from 3 to 7.5 dB per port.

The link consists of five spools of graded-index FMF with depressed cladding [164] supporting six spatial modes namely LP_{01}, LP_{11a} , LP_{11b}, LP_{21a}, LP_{21b}, and LP_{02}. The spools used for transmission are illustrated in Fig. 6.2(d). As shown in Table 6.1, the fibers have positive and negative DMD to achieve DMD compensation. In Table 6.1, $\mathrm{DMD}_{\mathrm{A-B}}$ is the modal delay between the modes LP_{A} and LP_{B}.

Table 6.1.: Relative mode delay with respect to the mode LP_{01}.

Length [m]	DMD_{01-11} [ns]	$\mathrm{DMD}_{01-02/21}$ [ns]
29980	3.45	5.2
23787	−2.75	−5.95
5400	0.95	1.95
10450	~ 0	−0.65
4550	0.3	0.95
$\sum$ length: 74167 m	$\sum$ DMD_{01-11} [ns]: 1.95 ns	$\sum$ DMD_{01-02}: 1.5 ns

Figure 6.3 depicts the time-of-flight measurements used to compute the values in Table 6.1. The measurement was taken by transmitting a 62.5 ps pulse through one input of the MUX, then the pulse was detected by a photodiode and a scope at an output of the DMUX. As Fig. 6.3 shows, the DMD_{01-02} and DMD_{01-21} cannot be individually identified, since the group velocity of both modes is very similar; therefore, they are registered in Table 6.1 in a single column. Reference values for measuring correctly the DMD values were the specifications provided by the fiber manufacturer. The total residual DMD is 1.95 ns. A CP of 8 ns was chosen to completely compensate for the ISI caused by the modal delay and dispersion. The dispersion for 74 km corresponds to approximately 0.08 ns, which can be calculated

as in Eq. (6.1) [151], where the Δf_{OFDM} is the spectral width of the OFDM spectrum and is computed according to Eq. (6.2) and D_{acc} is the accumulated dispersion after propagation, which was calculated with the maximum dispersion parameter of all spools 19 ps/(nm·km) (worst case scenario). In Eq. (6.2), the factor in the numerator corresponds to the number of subcarriers occupied by the OFDM spectrum including the data subcarriers N_{DATA}, the number of pilot subcarrier for the QPA and the PA PN compensation techniques, N_{pQPA} and N_{pPA}, and the number one corresponds to the subcarrier located in DC. A total attenuation of 17.5 dB was measured after splicing all spools together.

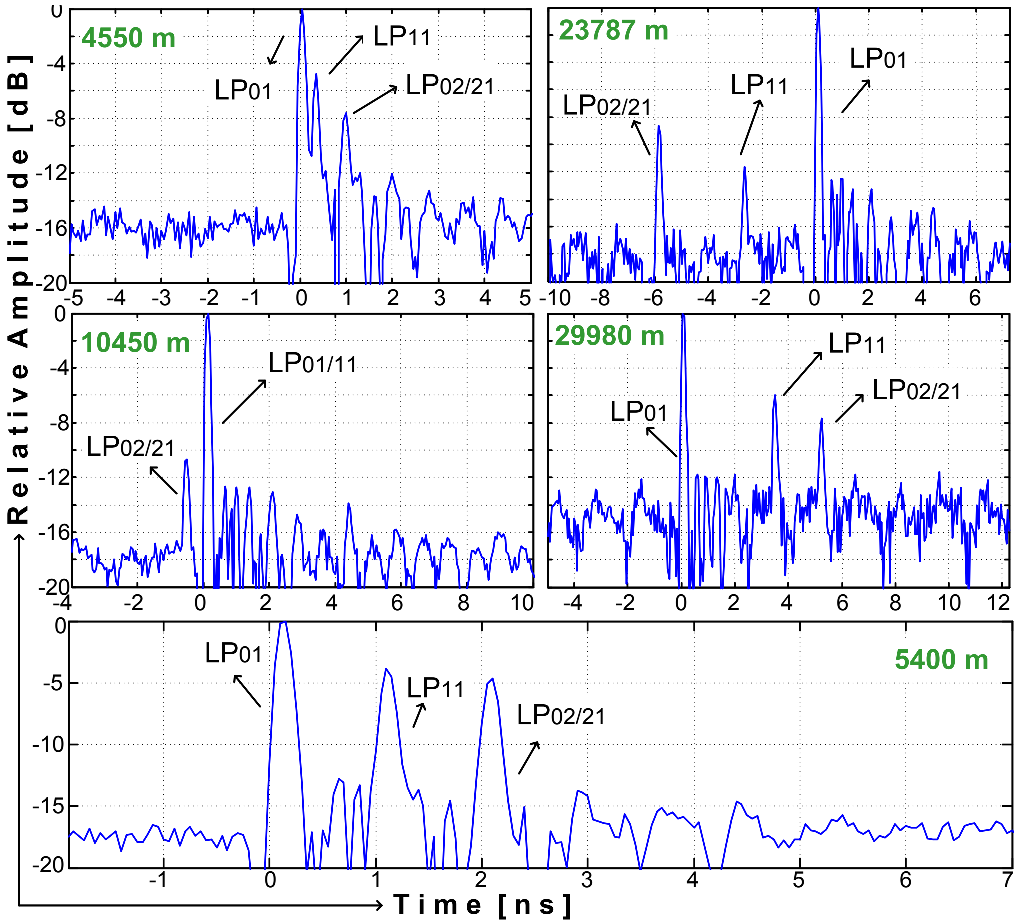


Figure 6.3.: Time-of-flight measurements for the fiber spools in Table 6.1.

$$\tau_{\mathrm{CD}} = \frac{c}{f_c^2} \left| D_{\mathrm{acc}} \right| \Delta f_{\mathrm{OFDM}} \tag{6.1}$$

$$\Delta f_{\mathrm{OFDM}} = \frac{\left(N_{\mathrm{DATA}} + N_{\mathrm{pQPA}} + N_{\mathrm{pPA}} + 1 \right) R_{\mathrm{DAC}}}{N_{\mathrm{FFT}}} \tag{6.2}$$

As mentioned previously, the CP duration is 8 ns (80 samples of an OFDM symbol) corresponding to $\mathrm{CP}_{\mathrm{ov}} = 31.2$ %. The TS contribute to a larger overhead which can be reduced substantially if different DACs for each tributary are used. With 46 OFDM symbols for the TS preamble per frame, i.e. 44 empty OFDM symbols and 2 TSs, and 52 OFDM symbols for the payload, the TS overhead is $\mathrm{TS}_{\mathrm{ov}} = 88.5$ %. If only two TSs are generated with different DACs per tributary, which is the scheme implemented for the simulations, the TS overhead can be reduced to 3.8 %.

After the multimode signal is demultiplexed into six POLMUX signals, it is down-converted by a LO to an intermediate frequency (IF, 6.6 GHz). The POLMUX signals can be received by using only two outputs of a coherent heterodyne RX, thus, the number of required RF inputs in the scopes is 12 for all tributaries. Subsequently, the 12 signals are sampled by two scopes at 40 and 50 GS/s, and processed offline. A picture of the scopes, the coherent heterodyne RXs and the amplifiers and coupler for the LO input of the RXs can be seen in Fig. 6.2(a). Scope 1 is labeled as *Scope 1a* and *Scope 1b*, since it is composed of two modules.

A block diagram illustrating the DSP structure at the RX is shown in Fig. 6.4. The first stage of the RX DSP is a coarse CFO compensation, which is performed by identifying a minimum around the central region of the OFDM spectrum. For this purpose, the central subcarrier was left unmodulated, which can be clearly identified from the inset of Fig. 6.1.

After the course CFO compensation, the sampling frequency offset (SFO) is compensated by resampling the signal by interpolation at a fixed frequency difference between TX and RX clock, since the SFO do not seem to vary much in time. Since the signals from the scopes have higher sampling rate (40 and 50 GS/s) than at the TX and they are also different among them, the signals are downsampled to the TX sampling rate, i.e., 10 GS/s. Afterward, the signals are synchronized by the method proposed by Schmidl and Cox in [126], which consist of the correlation between the received tributary with a time-shifted version of itself. The time shift corresponds to one OFDM symbol, which allows to find two consecutive repeated OFDM symbols marked by peaks in the correlation. However, the information from a single tributary is not enough for synchronization because considerable amount of power of the transmitted training symbols might have leaked to other tributary due to mode coupling; the consequence is that the correlation is conformed by peaks of different magnitudes. For this reason, the correlation of all tributaries is combined leading to clearly identifiable peaks.

The fine CFO compensation is applied via the Schmidl and Cox method [126].

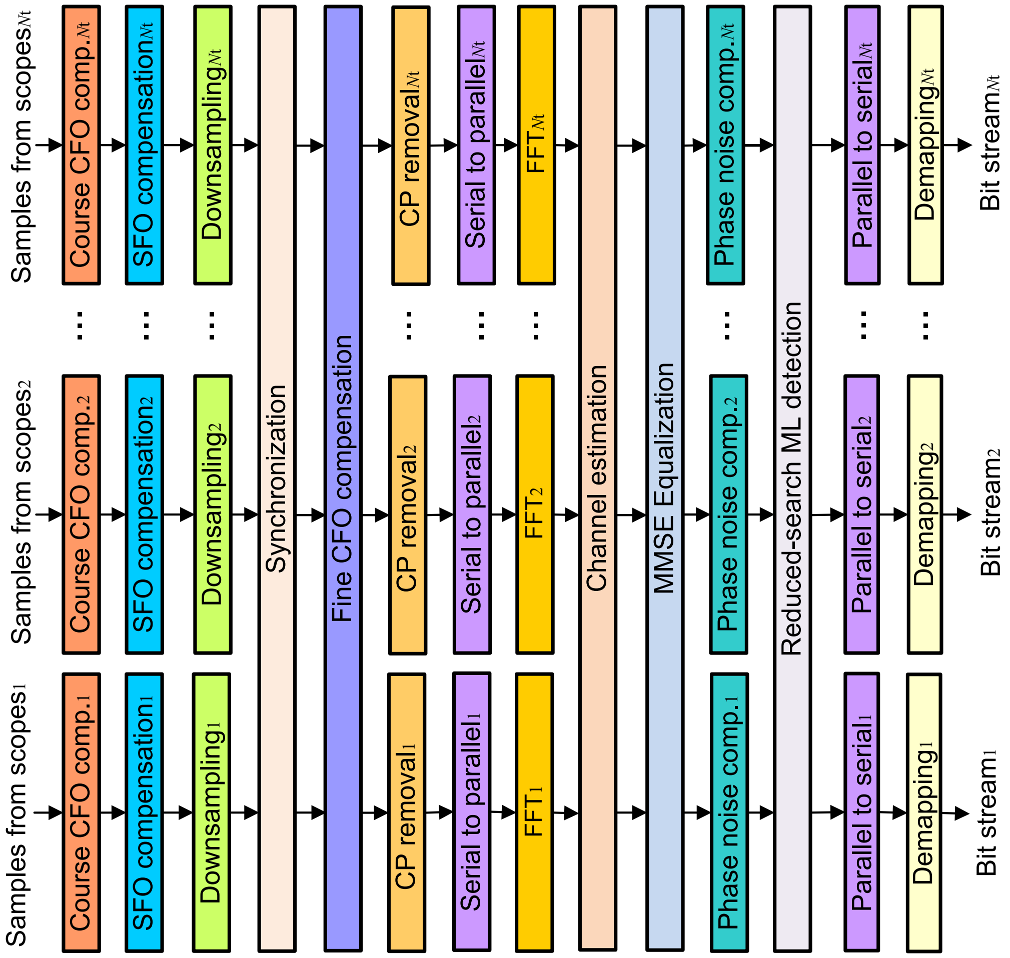


Figure 6.4.: Block diagram of the RX DSP for the experiments.

The coarse frequency offset estimation done by finding the middle of the OFDM spectrum is complemented with this method to improve the overall accuracy of the CFO estimation. After removing the CP, the signals are converted into the frequency domain with the FFT, followed by the channel estimation described for time-multiplexed TSs in [127]. In order to reduce the noise in the channel estimation, a moving averaging filter is applied to TSs of consecutive frames. Then, MMSE equalization is performed.

Afterward, PN estimation and compensation is performed. It is composed of two stages, which are carried out on an OFDM symbol basis. First, a coarse common phase rotation is estimated per OFDM symbol via QPA [133] and PA techniques [189]. The different expressions for QPA and PA PN estimation are shown in Eqs. (6.3a) and (6.3b), where $k \in \mathbb{K}_{\mathrm{QPA/PA}}$ is the OFDM subcarrier index of the pilots, $\mathbb{K}_{\mathrm{QPA/PA}}$ is the set of pilot subcarrier indexes used for QPA and PA, $N_{\mathrm{P_{QPA/PA}}}$

is the cardinality of $\mathbb{K}_{\mathrm{QPA/PA}}$, $\Phi_{\mathrm{QPA/PA}}$ are the QPA and PA phase estimation and, R_k and S_k are the k^{th} received and transmitter pilots symbols, respectively. The subcarrier are indexed so that the subcarrier in the middle of the OFDM spectrum (DC) has the index zero, $k = 0$; left and right of the middle subcarrier correspond to negative- and positive-indexed subcarriers, respectively.

$$\Phi_{\mathrm{QPA}} = \arg\left(\sum_{k\in\mathbb{K}_{\mathrm{QPA}}} R_k R_{-k}\right)/2 \tag{6.3a}$$

$$\Phi_{\mathrm{PA}} = \frac{1}{N_{\mathrm{pPA}}}\sum_{k\in\mathbb{K}_{\mathrm{PA}}} \left[\arg\left(R_k\right) - \arg\left(S_k\right)\right] \tag{6.3b}$$

In the case of QPA, the phase of N_{pQPA} equalized received symbols R_k are compared with their mirror subcarrier, R_{-k}. Instead of using pilot symbols, QPA codes the transmitted symbols as $S_k = S_{-k}^*$, where $*$ denotes the complex conjugate operator. This condition allows using the information of $2N_{\mathrm{pQPA}}$ symbols for the PN estimation and, thus, averages the noise of $2N_{\mathrm{pQPA}}$ subcarriers [133]. The advantage of coding the transmitted symbols with this technique is that similar results can be obtained by using only half of the pilot tones compared to the PA technique computed as Eq. (6.3b) [133].

As Eq. (6.3a) shows, for the QPA PN estimation the phase of two symbols have to be added, but if $|\arg\left(R_k R_{-k}\right)| > \pi$, Eq. (6.3a) results in phase ambiguity. For this reason the QPA scheme is combined with the PA scheme. As Eq. (6.3b) shows, residual phase rotation due to PN and possible cyclic slips generated with the QPA technique are extracted by averaging the phase difference between N_{pPA}-transmitted pilot symbols and the corresponding received symbols [189].

The second stage of the PN compensation is the expectation maximization (EM) [190–193]. EM is an iterative machine learning process, which is used to find the maximum likelihood estimates of the mean and co-variance of clusters of information described by a two-dimensional Gaussian distribution. Two dimensions is referred here to the in-phase and quadrature dimensions of the constellation. In this case, each cluster of information corresponds to an agglomerate of constellation points (e.g., 8QAM has eight clusters). The EM algorithm is especially beneficial when the constellation has an imprinted pattern, as in the case of highly coherent nonlinear and laser PN [190, 193]. The EM algorithm is implemented and provided by the author in [191] and applied on every OFDM symbol. It estimates the center of each cluster, comparing it with the corresponding ideal constellation point, and rotating back the symbols from all subcarriers.

As mentioned before, EM assumes a bi-variate Gaussian model, i.e. the conditional distribution of the received symbol at the input of the EM stage $\mathbf{y}_{\mathrm{EM}} \in \mathbb{R}^2$ given that the symbol $z_k \in \mathbb{C}$ belonging to the k-th cluster was transmitted, can be represented by

$$p(\mathbf{y}_{\mathrm{EM}} \mid z_k) = \mathcal{N}(\mathbf{y}_{\mathrm{EM}} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \ , \tag{6.4}$$

where $\mathbf{y}_{\mathrm{EM}}$ is written as a bidimensional vector since its first and second row corresponds to the real and imaginary part of the received symbol as in Eq. (6.5), respectively.

$$\mathbf{y}_{\mathrm{EM}} = \begin{pmatrix} y_{\mathrm{EM_I}} \\ y_{\mathrm{EM_Q}} \end{pmatrix} \tag{6.5}$$

In Eq. (6.4), $\mathcal{N}(\mathbf{y}_{\mathrm{EM}} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ denotes the Gaussian distribution of $\mathbf{y}_{\mathrm{EM}}$ with mean $\boldsymbol{\mu}_k \in \mathbb{R}^2$ and covariance matrix $\boldsymbol{\Sigma}_k \in \mathbb{R}^{2\times 2}$ and is defined in Eq. (6.6). In Eq. (6.6), the operator $\det\{\cdot\}$ denotes determinant. In the same manner as $\mathbf{y}_{\mathrm{EM}}$, the first row of $\boldsymbol{\mu}_k$ is the real component mean value of the symbols belonging to the k-th cluster $\mu_{\mathrm{I},k}$ and the second row, the mean value of their imaginary component $\mu_{\mathrm{Q},k}$ as represented in Eq. (6.7). Similarly, $\boldsymbol{\Sigma}_k$ is defined as Eq. (6.8), where $\Sigma_{\mathrm{I,I},k}$, $\Sigma_{\mathrm{Q,Q},k}$ and $\Sigma_{\mathrm{I,Q},k}$ are the variance of the real, imaginary part and the cross-covariance terms for a certain cluster k, respectively.

$$\mathcal{N}(\mathbf{y}_{\mathrm{EM}} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{1}{2\pi\sqrt{\det\{\boldsymbol{\Sigma}_k\}}} e^{-\frac{1}{2}(\mathbf{y}_{\mathrm{EM}}-\boldsymbol{\mu}_k)^{\mathrm{T}}\boldsymbol{\Sigma}_k^{-1}(\mathbf{y}_{\mathrm{EM}}-\boldsymbol{\mu}_k)} \tag{6.6}$$

$$\boldsymbol{\mu}_k = \begin{pmatrix} \mu_{\mathrm{I},k} \\ \mu_{\mathrm{Q},k} \end{pmatrix} \tag{6.7}$$

$$\boldsymbol{\Sigma}_k = \begin{pmatrix} \Sigma_{\mathrm{I,I},k} & \Sigma_{\mathrm{I,Q},k} \\ \Sigma_{\mathrm{I,Q},k} & \Sigma_{\mathrm{Q,Q},k} \end{pmatrix} \tag{6.8}$$

The marginal distribution of $\mathbf{y}_{\mathrm{EM}}$ is given by the superposition of K Gaussian densities as in Eq. (6.9) [192], where K is the total number of clusters or the constellation cardinality and π_k represent the mixing coefficients, respectively. As shown in

Eq. (6.9), the mixing coefficients π_k correspond to $P(z_k)$, i.e. the probability that the transmitted symbol belongs to the k-th cluster.

$$p(\mathbf{y}_{\mathrm{EM}}) = \sum_{k=1}^{K} P(z_k) p(\mathbf{y}_{\mathrm{EM}} \mid z_k) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{y}_{\mathrm{EM}} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \tag{6.9}$$

The equations shown before refer to just one received constellation point. The EM algorithm considers an observation period, namely, a certain number of received symbols, which in this case is chosen to be the length of an OFDM symbol allowing the finest possible phase noise estimation. The number of symbols received in an OFDM symbol is denoted as N_{DATA}. The EM algorithm aims to find the parameters $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ that maximizes the log of the likelihood function in Eq. (6.10) for the length of this observation period [192]. In Eq. (6.10), $\mathbf{Y}_{\mathrm{EM}} \in \mathbb{R}^{N_{\mathrm{DATA}} \times 2}$ contains the received symbols at the input of the EM algorithm, whose first and second column correspond the real and imaginary part of the these symbols, respectively.

$$\ln p(\mathbf{Y}_{\mathrm{EM}} \mid \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{n=1}^{N_{\mathrm{DATA}}} \ln \left\{ \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{y}_{\mathrm{EM}_n} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \right\} \tag{6.10}$$

The EM algorithm addresses the problem of maximizing Eq. (6.10) by computing $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ iteratively alternating an E-step and M-step [192]. A block diagram of the process is depicted in Fig. 6.5.

The first step is the initialization, at iteration number $q = 0$. Here, the first cluster association is done for each one of the received symbols at the input of the EM stage. The incoming symbols are compared with ideal constellation points and clusters are formed based on the closest Euclidean distance. Also, initial values for $\boldsymbol{\mu}_k^q$,$\boldsymbol{\Sigma}_k^q$ and π_k^q are computed, which are used to calculate the first log likelihood (LLH) as in Eq. (6.11) with $q = 0$.

$$\ln p(\mathbf{Y}_{\mathrm{EM}} \mid \boldsymbol{\pi}^q, \boldsymbol{\mu}^q, \boldsymbol{\Sigma}^q) = \sum_{n=1}^{N_{\mathrm{DATA}}} \ln \left\{ \sum_{k=1}^{K} \pi_k^q \mathcal{N}(\mathbf{y}_{\mathrm{EM}_n} \mid \boldsymbol{\mu}_k^q, \boldsymbol{\Sigma}_k^q) \right\} \tag{6.11}$$

After incrementing the iteration number q, in the E-step, the *a posteriori* probabilities $\gamma(z_{k_n})^q$ of each received symbol n are computed using the Bayes' theorem as in Eq. (6.12) [192]. These *a posteriori* probabilities can be understood as the probability the symbol z_{k_n} was transmitted given that $\mathbf{y}_{\mathrm{EM}_n}$ was received.

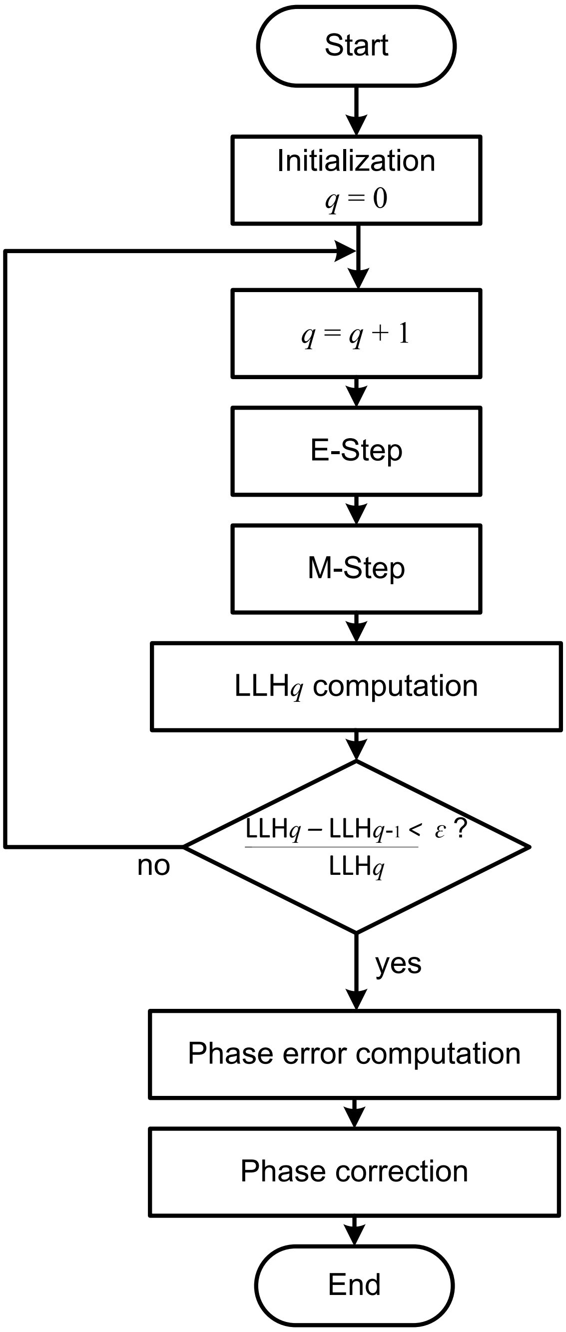


Figure 6.5.: EM algorithm block diagram.

$$\gamma(z_{k_n})^q = P(z_{k_n} \mid \mathbf{y}_{\mathrm{EM}_n})^q = \frac{p(z_k)^{q-1} p(\mathbf{y}_{\mathrm{EM}_n} \mid z_{k_n})^{q-1}}{p(\mathbf{y}_{\mathrm{EM}_n})^{q-1}} = \frac{\pi_k^{q-1} \mathcal{N}(\mathbf{y}_{\mathrm{EM}_n} \mid \boldsymbol{\mu}_k^{q-1}, \boldsymbol{\Sigma}_k^{q-1})}{\sum_{j=1}^{K} \pi_j^{q-1} \mathcal{N}(\mathbf{y}_{\mathrm{EM}_n} \mid \boldsymbol{\mu}_j^{q-1}, \boldsymbol{\Sigma}_j^{q-1})} \tag{6.12}$$

In the M-step, the *a posteriori* probabilities $\gamma(z_{k_n})^q$ are used to compute new $\boldsymbol{\mu}_k^q$,

$\boldsymbol{\Sigma}_k^q$ and π_k^q as [192]

$$\boldsymbol{\mu}_k^q = \frac{1}{N_k^q} \sum_{n=1}^{N_{\text{DATA}}} \gamma(z_{k_n})^q \mathbf{y}_{\text{EM}_n} \tag{6.13a}$$

$$\boldsymbol{\Sigma}_k^q = \frac{1}{N_k^q} \sum_{n=1}^{N_{\text{DATA}}} \gamma(z_{k_n})^q \left(\mathbf{y}_{\text{EM}_n} - \boldsymbol{\mu}_k^q\right) \left(\mathbf{y}_{\text{EM}_n} - \boldsymbol{\mu}_k^q\right)^{\text{T}} \tag{6.13b}$$

$$\pi_k^q = \frac{N_k^q}{N_{\text{DATA}}} \tag{6.13c}$$

where N_k^q is defined as

$$N_k^q = \sum_{n=1}^{N_{\text{DATA}}} \gamma(z_{k_n})^q \ . \tag{6.14}$$

Then, the LLHs from iteration q and $q-1$, denoted in the block diagram in Fig. 6.5 as LLH_q and LLH_{q-1} are compared, i.e. the LLH increase relative to the LLH of the current iteration q is examined. If this relative increase is greater than a predefined threshold ϵ, the algorithm keeps iterating, otherwise, the algorithm ends the iteration and proceeds with the phase error calculation and correction.

The phase error estimated via EM Φ_{EM_k} is referred to the comparison of the computed cluster means and the ideal constellation points as in Eq. (6.15), where g_k are the ideal constellation points closest to the cluster k.

$$\Phi_{\text{EM}_k} = \angle \frac{\mu_{\text{I},k} + j\mu_{\text{Q},k}}{g_k} \tag{6.15}$$

The last step of the EM algorithm is the phase correction. It is done by reverting the estimated phase error for each one of the received symbols as shown in Eq. (6.16), where $x_{\text{EM}_{n_k}}$ and $y_{\text{EM}_{n_k}}$ are the n-th output and input symbol of the EM algorithm associated to cluster k, respectively. The input symbols $y_{\text{EM}_{n_k}}$ are associated to k according to the cluster k that provides the highest *a posteriori* probability $\gamma(z_{k_n})$ at the last iteration.

$$x_{\text{EM}_{n_k}} = (y_{\text{EM}_{I\,n_k}} + j y_{\text{EM}_{Q\,n_k}}) e^{-j\Phi_{\text{EM}_k}} = y_{\text{EM}_{n_k}} e^{-j\Phi_{\text{EM}_k}} \tag{6.16}$$

The beneficial effect of the EM algorithm on the compensation of residual phase noise is depicted in Fig. 6.6 for QPSK, 8QAM and, back-to-back transmission using the (D)MUX.

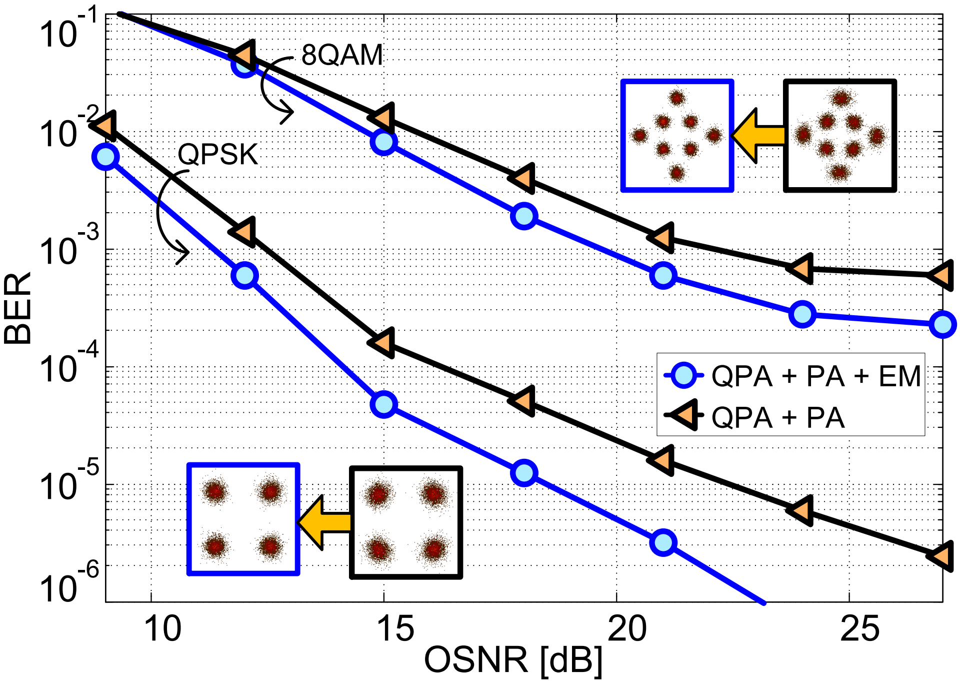


Figure 6.6.: BER versus OSNR with and without applying the EM algorithm.

6.1.2. Experiment results

We compare the performance of the linear equalizer MMSE and the IRSML detection in the experiments. On the one hand, the MMSE equalizer is the better linear equalizer examined in Chapter 5 outperforming the ZF equalization. On the other hand, the IRSML is the near-ML algorithm characterized by a more suitable practical implementation than the SD as mentioned in Chapter 5.

In contrast to Chapter 5, in the experiments there is PN from the TX and RX lasers, which has to be taken into account for the IRSML detection. Therefore, Eq. (5.4) turns into Eq. (6.17), where $\circ$ is the Hadamard product or element-wise operator and the vector $\mathbf{\Phi}_{\text{CPE}} \in \mathbb{C}^{N_\text{t}}$ is the total estimated common phase error from the QPA, PA and EM methods for each one of the tributaries.

$$\hat{\mathbf{s}}_{\text{IRSML}} = \arg\min_{\mathbf{s}\in\mathbb{S}} \|\mathbf{r}_\text{o} - \mathbf{H}\mathbf{s} \circ \mathbf{\Phi}_{\text{CPE}}\|^2 \tag{6.17}$$

The search set $\mathbb{S}$ is determined in the same way as explained in Chapter 5. The maximum size of $\mathbb{S}$ is $N_{\max}$ hypersymbols, which is chosen according to the allowed computational complexity in the system. In this case, $N_{\max}$ is chosen to be eight, since increasing further the size of $\mathbb{S}$ does not change significantly the performance as the inset in Fig. 6.7(a) shows for the transmission over 74 km and QPSK.

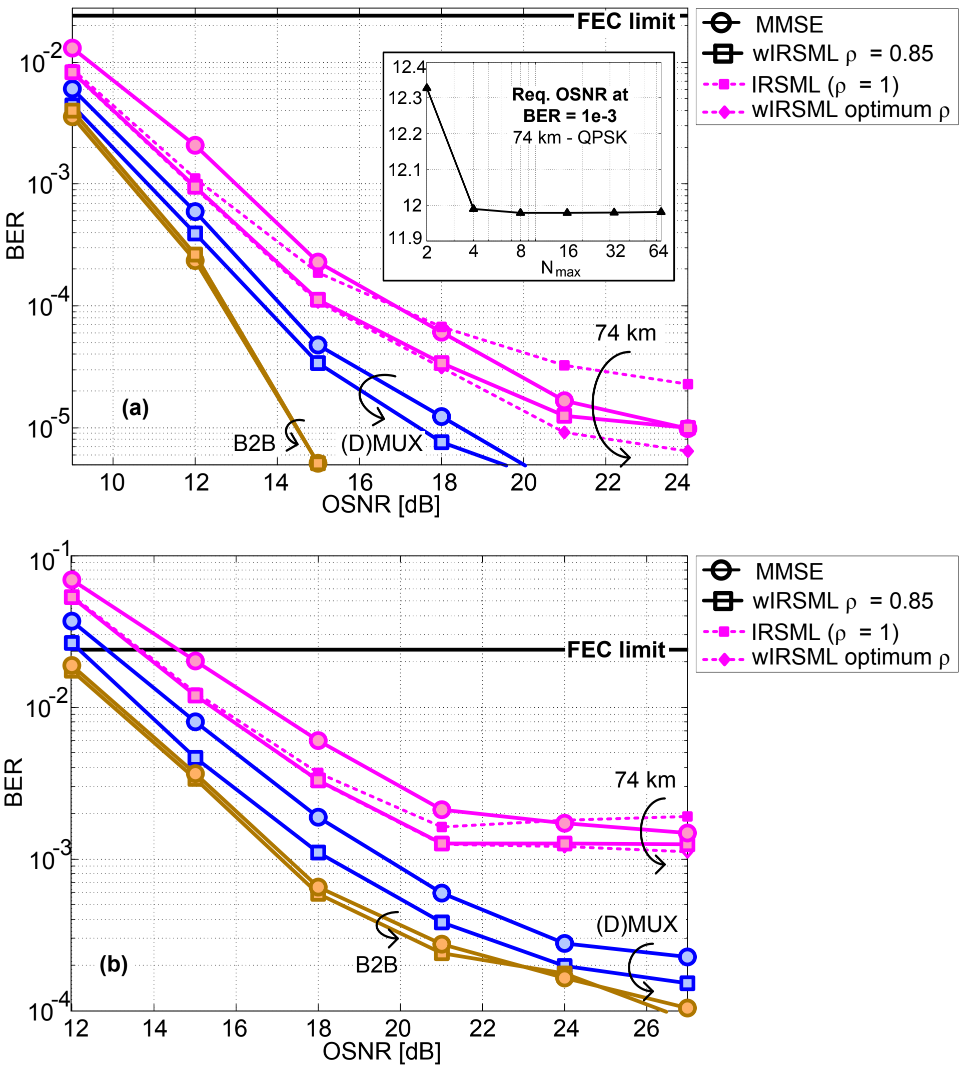


Figure 6.7.: Average BER over all tributaries versus OSNR for (a) MDM-QPSK-OFDM and (b) MDM-8QAM-OFDM.

Ideally, if the channel estimation and the PN estimation are accurate, the detection output corresponds to Eq. (6.17). In this case, perfect channel knowledge is not possible to obtain and there is residual inter-carrier interference from the PN estimation. With the purpose of improving the detection criteria in Eq. (6.17) and make it robust against such inaccuracies, the output of the IRSML algorithm was modified by assigning a weight to the metric corresponding to the MMSE equalizer output after detection $\hat{\mathbf{s}}_{\text{MMSE}}$ as represented in Eq. (6.18) by ρ. In Eq. (6.18), $\hat{\mathbf{s}}_{\text{wIRSML}}$, $\mathbb{A}$ and $\hat{\mathbf{s}}'$ are output of the proposed weighted IRSML (wIRSML) algorithm, the set of hypersymbol candidates excluding $\hat{\mathbf{s}}_{\text{MMSE}}$ and the hypersymbol candidates themselves, respectively. The value of ρ was chosen as a global optimum for all the measurements, corresponding to 0.85. An alternative to further optimize ρ would be to optimize it blindly by observing the corrected errors of the FEC stage at the RX. Simulations to corroborate that the performance of the IRSML is influenced by PN are shown in Section 6.1.2

$$\begin{gathered}\hat{\mathbf{s}}_{\mathrm{wIRSML}} = \underset{\mathbf{s}' \in \mathbb{A} \,\vee\, \hat{\mathbf{s}}_{\mathrm{MMSE}}}{\arg\min} \begin{cases} \rho \left\| \mathbf{r}_{\mathrm{o}} - \mathbf{H}\hat{\mathbf{s}}_{\mathrm{MMSE}} \circ \mathbf{\Phi}_{\mathrm{CPE}} \right\|^2 \\ \left\| \mathbf{r}_{\mathrm{o}} - \mathbf{H}\mathbf{s}' \circ \mathbf{\Phi}_{\mathrm{CPE}} \right\|^2 \end{cases} \\ \mathbb{A} = \{\mathbf{s} \mid \mathbf{s} \in \mathbb{S} \wedge \mathbf{s} \neq \hat{\mathbf{s}}_{\mathrm{MMSE}}\}\end{gathered} \tag{6.18}$$

Figure 6.7 compares the performance of the MMSE equalizer and the IRSML approaches for a set of measurements using MDM-QPSK-OFDM and MDM-8QAM-OFDM. The curves show the performance of B2B with and without (D)MUX and after the 74-km transmission. More than four million bit were evaluated for the BER computation. The FEC limit is also shown for 20 % overhead (2.4e-2 of BER), which is a promising approach and has become a popular choice for researchers in the field [69, 129, 141]. As an example, Fig. 6.7 depicts the wIRSML detection upper (IRSML) and lower (wIRSML with optimum ρ) bound for the 74-km transmission. Fig. 6.7 shows that the wIRSML scheme allows for superior performance in comparison with the MMSE equalizer. Its improvement for QPSK is 0.5 and 1.1 dB for the (D)MUX and 74 km curves, respectively, at 1e-3 BER with respect to the MMSE equalizer. For 8-QAM, the performance improves by 0.9 and 1.3 dB at 1e-2 BER with respect to the MMSE equalization. As shown in Chapter 5, in the presence of MDL, ML and near-ML detection enhances the system performance and its improvement depends on the amount of MDL in the system.

The MDL in this system was computed and is depicted in Fig. 6.8. Similar to [68], an MDL of approximately 8 dB is observed by using the photonic lanterns. An additional MDL of 2 dB is observed after the transmission over 74 km, which is attributed to splicing imperfections and slightly different fiber attenuation coefficients per mode.

Simulations with phase noise and IRSML

Simulations were carried out to find out in which way the IRSML is affected by residual PN. The simulation setup is illustrated in Fig. 6.9. The signal is generated in the same manner as in the experiments. Afterward, TX PN is added as a Gaussian random walk of the phase as described in Section 3.1.1. 12 copies of the signal are

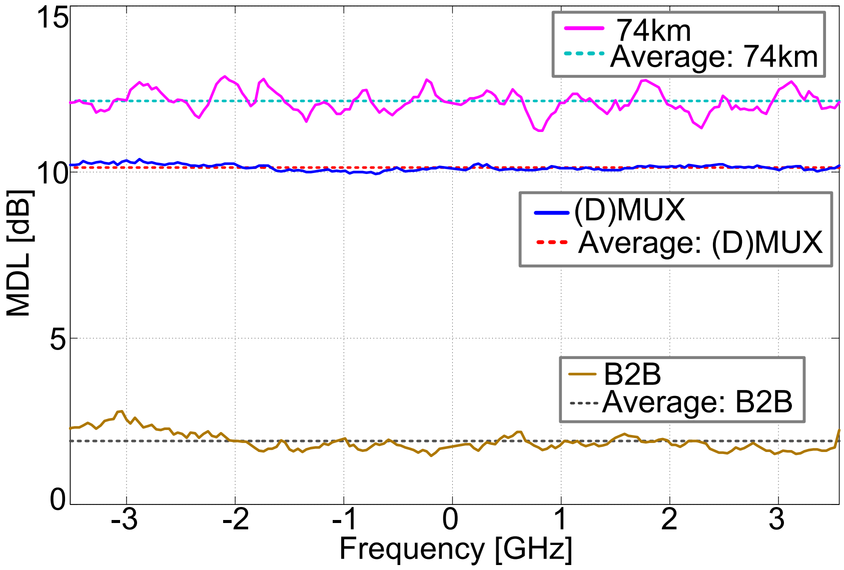


Figure 6.8.: MDL as a function of the frequency in the experiments.

then generated and delayed to form time multiplexed TS as in the experiments. All tributaries are combined by a photonic lantern, which is modeled according to the spatial (D)MUXs described in Section 3.3.1. Then, MDL is added based on the experimental results of the channel matrix SVD. Additional MDL is added as well by setting different attenuation coefficients in the fiber model as presented in [164]. According to [164], LP_{01} has an attenuation coefficient of 0.223 dB/km and the rest of the modes has 0.2 dB/km. The total MDL applied to the transmission resulted in 10.5 dB.

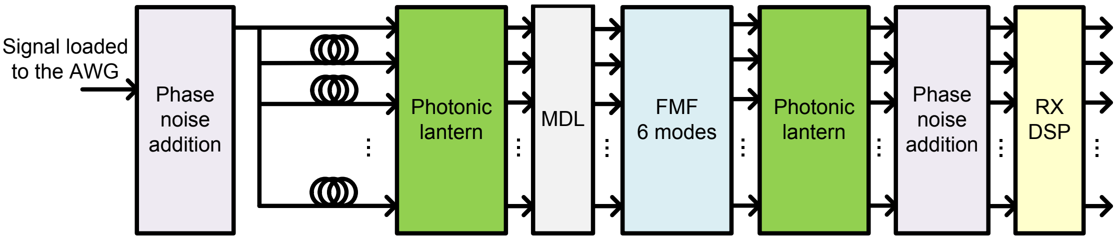


Figure 6.9.: Simulation block diagram with PN addition.

The fiber channel corresponds to the model used for the simulations in Chapters 4 and 5. The demultiplexing and PN addition in the RX are modeled in the same way as at the TX. The DSP for recovering the transmitted bits is similar to the one shown in Fig. 6.4. However, the downsampling, CFO and SFO compensation were bypassed since no upsampling and no frequency offsets are added.

BER curves as a function of OSNR are illustrated in Fig. 6.10 for the MMSE equalizer, IRSML and wIRSML for different values of ρ. PN was generated with a LW of 600 kHz at the TX and RX, respectively. This considerable amount of PN was added to observe a clearer effect of PN on the near-ML detection. The graphic in Fig. 6.10 shows that the IRSML outperforms the MMSE equalizer for OSNR lower than approximately 15 dB. For larger values of OSNR MMSE presents a lower BER. It can be said that in the regions where AWGN is dominant (low OSNR region) the IRSML performs better than MMSE, however, for the region where the residual PN is dominant (high OSNR region) the IRSML generates more errors than the MMSE. This behavior is also observed in the performance of the experimental data in Fig. 6.7.

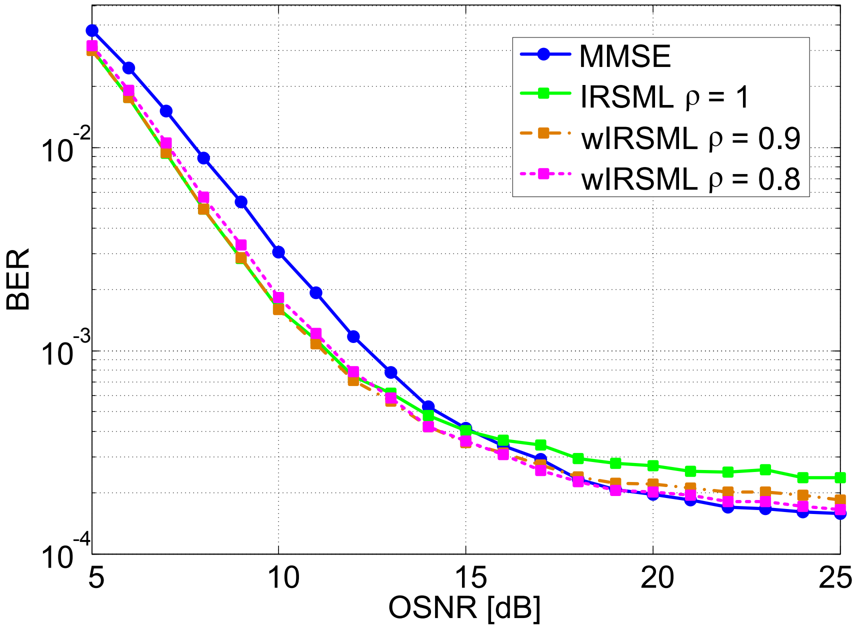


Figure 6.10.: BER as a function of OSNR for transmission impaired by PN and wIRSML.

As shown in Fig. 6.7 the factor ρ aids to overcome this problem and to balance the beneficial effect of near-ML detection and linear equalization in the evaluated OSNR range. A favorable balance is found when the weight ρ is not that low so that the wIRSML curve does not approach too much to the MMSE curve in the low OSNR region and is not that high so that it does not approach the IRSML curve in the high OSNR region.

The effect of the factor ρ was observed also by inspecting the number of errors caused by two different cases: when the output of the detector is $\hat{\mathbf{s}}_{\mathrm{MMSE}}$ but the correct output corresponds to the second line of Eq. (6.18) for one of the elements

$\mathbf{s}' \in \mathbb{A}$ and the other way around, i.e. when the detector output is $\mathbf{s}' \in \mathbb{A}$ but the correct detection corresponds to $\hat{\mathbf{s}}_{\mathrm{MMSE}}$. The metrics were also inspected by means of Δ in Eq. (6.19) in order to measure how close are the detector outputs produced by the first and the second line of Eq. (6.18).

$$\Delta = \frac{\rho \left\| \mathbf{r} - \mathbf{H}\hat{\mathbf{s}}_{\mathrm{MMSE}} \circ \mathbf{\Phi}_{\mathrm{CPE}} \right\|^2 - \arg\min_{\mathbf{s}' \in \mathbb{A}} \left\| \mathbf{r} - \mathbf{H}\mathbf{s}' \circ \mathbf{\Phi}_{\mathrm{CPE}} \right\|^2}{\rho \left\| \mathbf{r}_{\mathrm{o}} - \mathbf{H}\hat{\mathbf{s}}_{\mathrm{MMSE}} \circ \mathbf{\Phi}_{\mathrm{CPE}} \right\|^2} \tag{6.19}$$

The histograms in Fig. 6.11 show the number of errors produced by the two cases under study mentioned before, for different values of ρ and an OSNR of 21 dB. Figure 6.11(a) shows the case of the classical IRSML (Eq. (6.17)), in which it can be observed that there is a considerable amount of errors in the positive part of Δ and close to zero. Note that Δ close to zero in the histogram of errors means that the two metrics compared in Eq. (6.19) were very similar and the detector takes the wrong decision from Eq. (6.18).

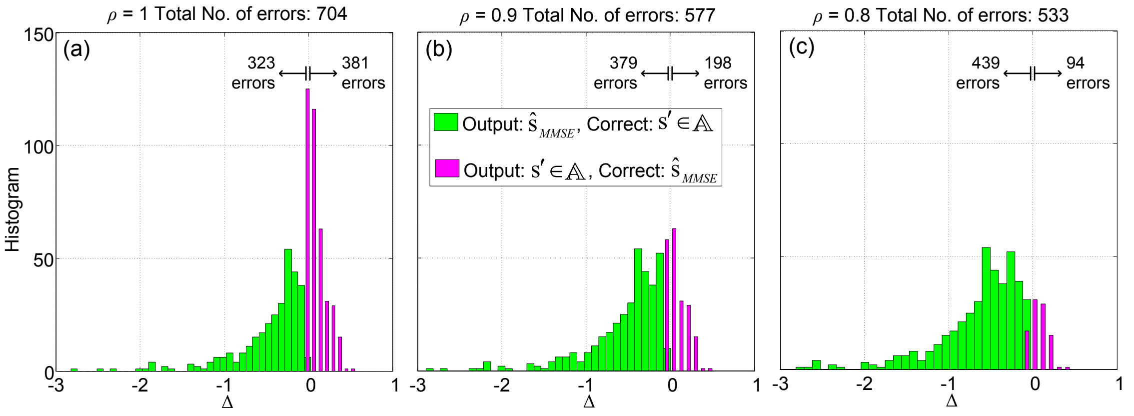


Figure 6.11.: Symbol error histogram for simulations, different values of ρ and for an OSNR of 21 dB.

The prominent peak in Fig. 6.11(a) is decreased by decreasing ρ as shown in (b) and (c). At the same time, by reducing the value of ρ, the errors caused by the detector output $\hat{\mathbf{s}}_{\mathrm{MMSE}}$ (green part of the histogram) increases and the ones caused by $\mathbf{s}' \in \mathbb{A}$ (magenta part of the histogram) decreases. The number of total number of errors written on the top of the histograms decreases as ρ decreases. In accordance with Fig. 6.10, the histograms for the lower OSNR region show an increase of errors for values of ρ different than one.

Also the experimental data was used to compute these histograms and is depicted in Fig. 6.12 for 15 dB of OSNR, which an exemplary OSNR value since the other values in the experiment present similar characteristics. The histograms in Fig. 6.12 illustrate a behavior similar to Fig. 6.11.

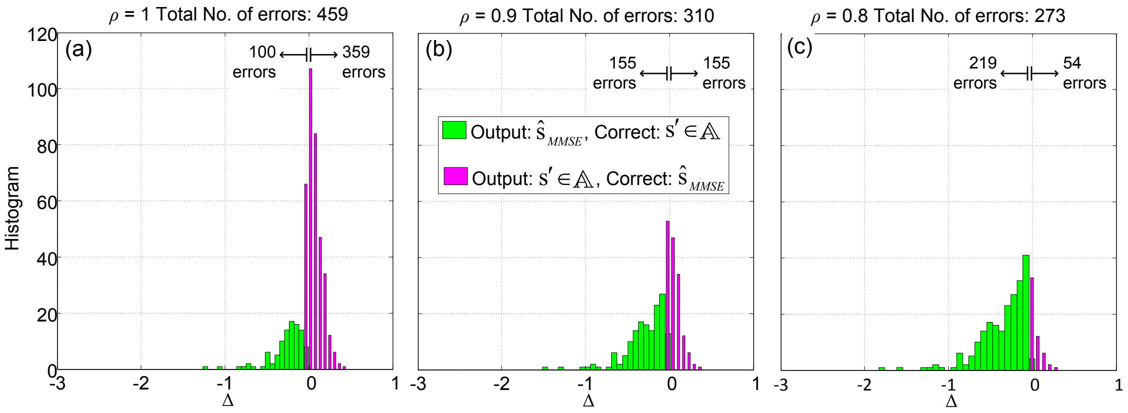


Figure 6.12.: Symbol error histogram for experiments, different values of ρ and for an OSNR of 15 dB.

From these simulations it is inferred that due to the remaining PN in the system the metrics compared in Eq. (6.19) might not be accurate enough and therefore the criterion in Eq. (6.17) might not be optimum in this case. An alternative for PN compensation other than correcting the CPE in the OFDM symbol would be to use a pilot tone in the middle of the OFDM spectrum as in [129]. The pilot tone tracks the PN and at the receiver it is filtered, conjugated and multiplied by the incoming signal. In the setup used in this thesis the pilot tone method is not a feasible option, since the signals from all the tributaries are greatly delayed from one another. However, in a real scenario a PN compensation with pilot tone should be feasible since the signals in all tributaries are different and not delayed.

6.2. Multi-channel experiments

Apart from the single channel experiments also multi-channel experiments with 255 WDM channels were carried out across the C-band of transmission (see Fig. 2.8). As here QPSK is used, the nominal bit rate comes to 41.6 Tb/s. Removing the overhead from the FEC, CP and TS gives a net bit rate of 14 Tb/s. As mentioned before the TS overhead can be reduced to 3.8 %, if two OFDM symbols are used as TSs, which comes to a net bit rate of 25.4 Tb/s. The block diagram of the WDM experiments is shown in Fig. 6.13. Basically, two changes in the structure of Fig. 6.1 can be observed. The first one is the generation of the neighboring 254 WDM channels. The second change is the optical filtering after the demultiplexing stage.

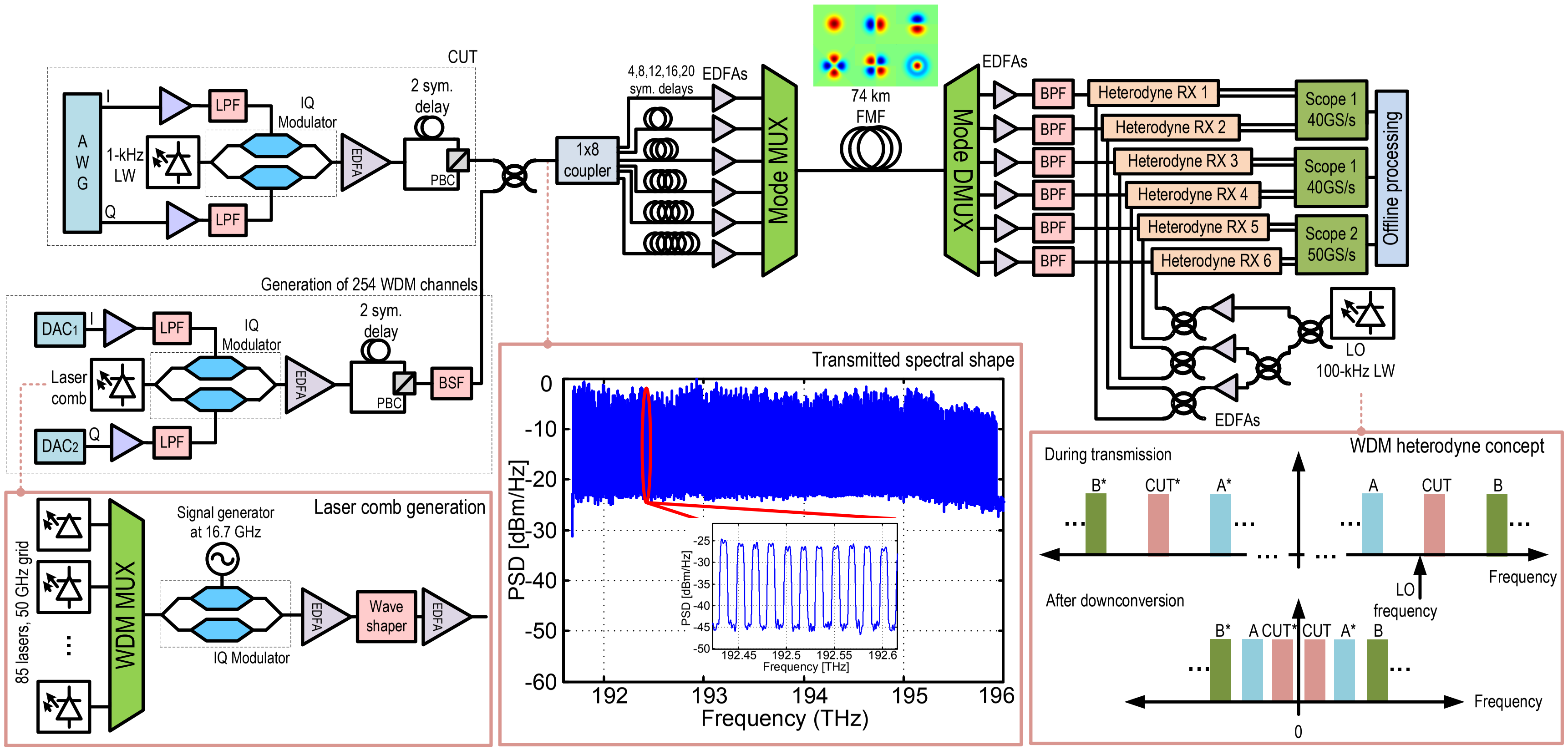


Figure 6.13.: Block diagram of the experimental WDM-MDM-OFDM system.

For generating the neighboring WDM channels essentially the same TX structure as used for the CUT is needed, however, with a laser source containing 254 peaks. For this purpose a laser comb was constructed as described in [117, 188].

The schematic of the laser comb is depicted in an inset of Fig. 6.13. An array of 85 laser sources were available multiplexed with an array-waveguide grating labeled in Fig. 6.13 as WDM multiplexer. The 85-laser-comb was placed at the input of an IQ modulator, which was driven by a single signal generator characterized by

$$u_{\mathrm{RF}}(t) = a_{\mathrm{RF}} \cos\left(\omega_{\mathrm{RF}} t\right) \ , \tag{6.20}$$

where $\omega_{\mathrm{RF}} = 16.7$ GHz. This structure allowed to have a frequency modulation and to control the spectral content at the output of the IQ modulator. The output of the IQ modulator can be described as in Eq. (6.21), where $E_{\mathrm{out/in}_{\mathrm{cmb}}}$ are the electrical field at the output/input of the comb IQ modulator, J_ν are the Bessel functions of first kind of order ν, $\beta_{\mathrm{a}} = \frac{\pi a_{\mathrm{RF}}}{V_\pi}$, $\Theta_{\mathrm{PM}} = \frac{\pi u_{\mathrm{PM}}}{V_\pi}$ and $\Theta_{\mathrm{I/Q}} = \frac{\pi u_{\mathrm{BIAS}_{\mathrm{I/Q}}}}{V_\pi}$, where u_{PM} is the driving DC voltage of the IQ modulator phase shifter part (see Fig. 3.7) and $u_{\mathrm{BIAS}_{\mathrm{I/Q}}}$ is the DC part of the IQ modulator electrical inputs. The derivation of Eq. (6.21) can be found in Appendix A.3.

$$\begin{aligned} E_{\mathrm{out}_{\mathrm{cmb}}} = \frac{E_{\mathrm{in}_{\mathrm{cmb}}}}{2} \Bigg[& \cos\Theta_{\mathrm{I}} J_0(\beta_{\mathrm{m}}) - 2\cos\Theta_{\mathrm{I}} \sum_{n=1}^{\infty} (-1)^{n-1} J_{2n}(\beta_{\mathrm{m}}) \cos\left(2n\omega_{\mathrm{RF}}\right) \\ & - 2\sin\Theta_{\mathrm{I}} \sum_{n=0}^{\infty} (-1)^{n-1} J_{2n+1}(\beta_{\mathrm{m}}) \cos\left(\left(2n+1\right)\omega_{\mathrm{RF}} t\right) + e^{j\Theta_{\mathrm{PM}}} \cos\Theta_{\mathrm{Q}} \Bigg] \end{aligned} \tag{6.21}$$

The target of this laser comb is that the 85 carriers of the laser bank and the spectral components produced by the first order Bessel function ±16.7 GHz survive after the IQ modulator. It is convenient that the rest of the products generated by higher order Bessel functions remain suppressed to avoid that they interfere with other carriers. As observed in Eq. (6.21), the amplitude of the 85 carriers from the laser can be controlled by changing the first and the last term, i.e. tuning β_{m}, Θ_{I}, Θ_{Q} and Θ_{PM}. Hence, to boost this amplitude it is necessary to lower β_{m}, Θ_{I}, Θ_{Q} and Θ_{PM}. In turn, to boost the amplitude of the first order Bessel spectral component is it necessary to increase β_{m} and Θ_{I}. However, β_{m} should be kept low since a large value will increase the amplitude of the spectral components ruled by $J_{\nu>1}$. It should be noted that increasing Θ_I has also the effect of increasing the amplitude of the terms multiplied by J_ν, if ν is odd, and decreasing the amplitude of the even terms. Therefore it is important to carefully tune the bias voltages of the IQ modulator and the amplitude of the RF source and find an amplitude balance between the 85 carriers from the laser bank and the first order Bessel products.

After the IQ modulation in the laser comb generation, the signal is amplified and the amplitude of the laser comb is equalized. For the latter, a wave shaper or a programmable filter is used.

The laser comb fed another IQ modulator, which was driven by the in-phase and quadrature OFDM signal of the same characteristics as the CUT, i.e. same number of modulated subcarriers. The OFDM I and Q signals are generated by two DACs each one of them controlled by a field programmable gate array (FPGA), where the OFDM signal is loaded. The DAC boards are fed with a 16 GHz clock. The setup of these DACs can be seen in the picture in Fig. 6.14.

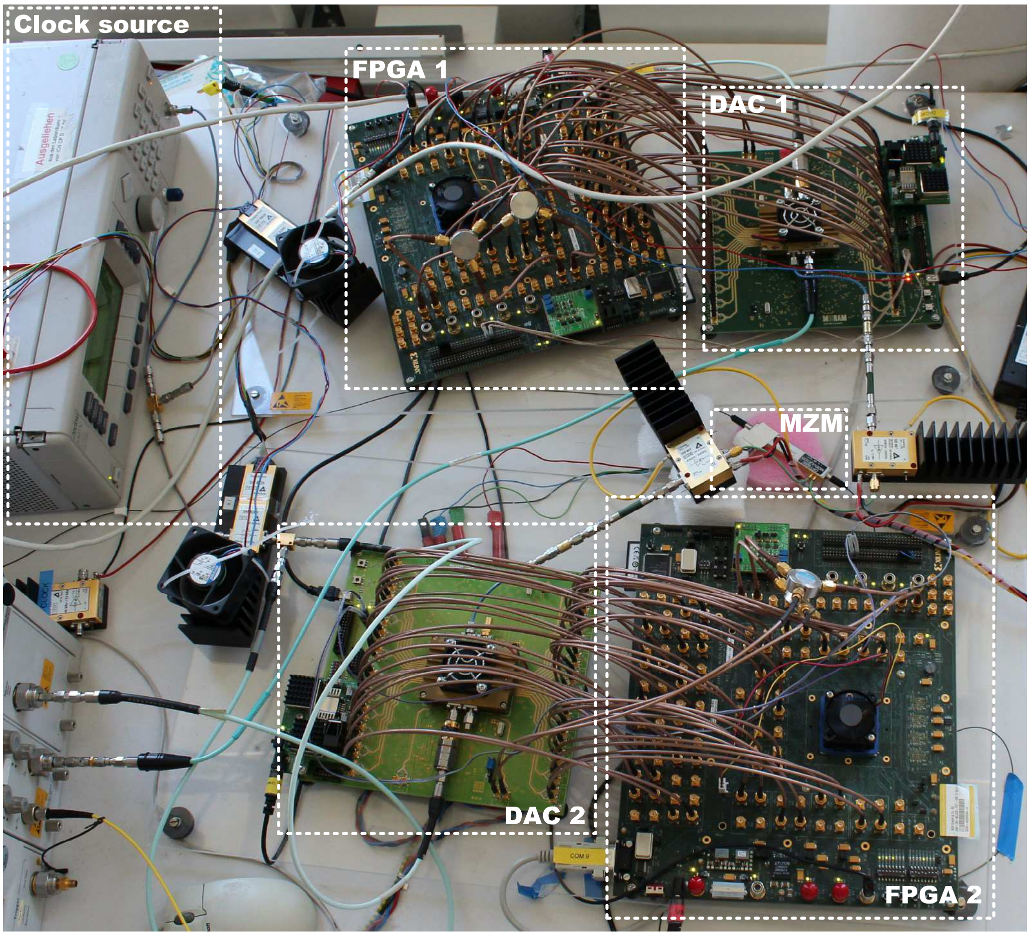


Figure 6.14.: Setup for generating the neighboring WDM channels of the CUT.

After the POLMUX stage of the neighboring WDM channels, one of the generated channels is replaced by the CUT by means of a band-stop filter (BSF) and a coupler. The BSF was realized again by programming a wave shaper. The power spectral density (PSD) of the transmitted spectrum can be observed in the inset in Fig. 6.13.

At the receiver after the mode DMUX, the signals are downconverted by coherent heterodyne RXs similar as in the single channel experiments. The signals from each tributary are first optically filtered by a bandpass filter (BPF) with a 3 dB bandwidth of 0.4 nm, which was the type of filter that was available with the smallest bandwidth and in quantity. With this bandwidth of the optical BPF it is not possible to filter a single WDM channel, i.e. there is power from others channels at the output of the optical BPF. For this reason, in the downconversion process, the channels from the positive and negative part of the spectrum were interleaved and located in the spectral gap of around 9.6 GHz of width between neighboring WDM channels as illustrated in the inset of Fig. 6.13. This interleaved structure was possible with a difference between the LO and TX laser of the CUT of 3.8 GHz, i.e. $f_{\mathrm{c}} - f_{\mathrm{LO}} = 3.8$ GHz, where f_{LO} is the frequency of the LO. Afterward, the sampled signal is processed by the offline DSP similarly as for the single channel experiments.

6.2.1. Experiment results

All 255 channels were measured and processed to compute the BER using the MMSE equalizer and IRSML detection using also eight hypersymbol candidates. The transmission results are shown in Fig. 6.15, where (a) is the received power spectral density after single-mode EDFA amplification and (b) the BER of all channels. The elevation in the power spectral density of Fig. 6.15(a) is attributed to the typical gain profile of EDFAs, where frequencies in this range have higher gain and also slightly lower SNR than lower frequencies.

In Fig. 6.15(b), it can be seen that both schemes, MMSE and IRSML, delivered BERs below the FEC limit. It can be observed that the IRSML detection outperforms the MMSE. Here, the wIRSML was not needed since the dominant noise is not from the residual PN but from the ASE noise from the amplifiers.

6.3. Discussion

In this chapter, an experimental demonstration of a system supporting six spatial modes in single channel configuration and multi-channel configuration is shown.

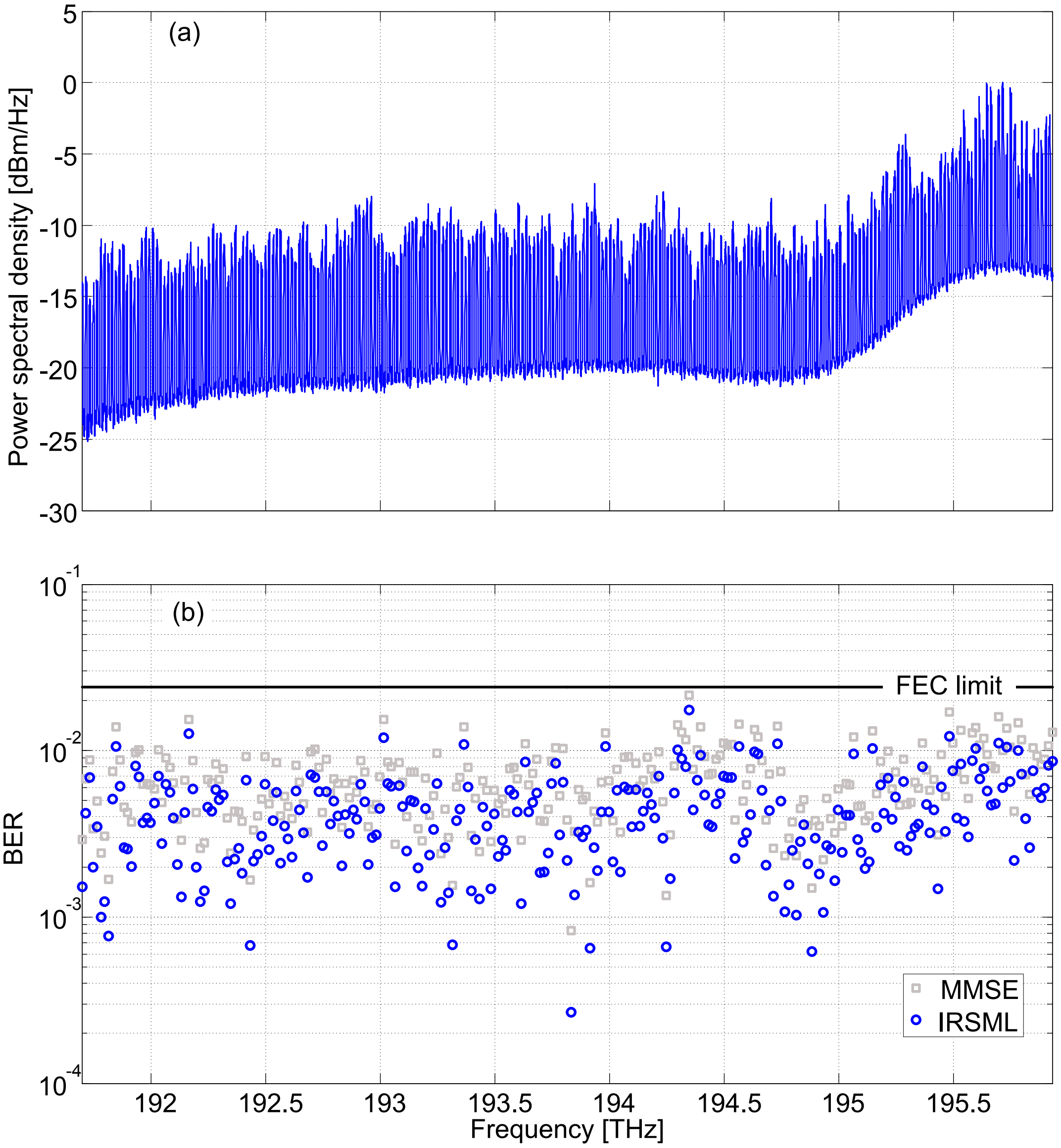


Figure 6.15.: (a) Received power spectral density after amplification and (b) average BER of the 255 measured WDM channels.

A single channel of MDM-QPSK-OFDM and MDM-8QAM-OFDM was transmitted at 163-Gb/s and 245-Gb/s, respectively, over 74 km. 255 WDM channels were also transmitted at 41.6 Tb/s in MDM-QPSK-OFDM configuration. The system is clearly impaired by the MDL introduced by the (D)MUX stage and imperfections in the splicing process and the different attenuation coefficients in the FMF itself. The performance of the IRSML detection algorithm was assessed and compared with the linear equalization provided by MMSE.

A new approach named wIRSML was also implemented to improve the performance of the IRSML in the low BER region or the region where the residual PN is dominant. By improving the PN estimation, it is believed that there is no need of the wIRSML scheme. The PN estimation can be improved by reducing the OFDM symbol duration and using schemes based on CPE correction. However, with a CPE scheme there will always be some ICI from the residual PN. The RF-pilot tone compensation is also another alternative for a better PN compensation, where a pilot tone is inserted in the middle of the spectrum [129]. At the RX the pilot is filtered and the incoming signal is multiplied by its conjugated version. In the case of these experiments, this scheme is not effective because of the large optical delays in the signal generation part at the TX combined with the effect of the mode coupling, which occasionally makes the pilot to leak to other tributaries reducing the pilot power and negatively affecting the PN estimation.

By comparing the MMSE and the IRSML, it was observed that the IRSML scheme outperforms the MMSE equalization by using only eight hypersymbol candidates. Thus, the number of candidates is reduced from more than 16 million (QPSK) and 68 billion (8QAM) hypersymbol candidates to eight compared to ML detection, which is a tremendous computational complexity reduction. Similar to the simulations shown in Chapter 5, the improvement in the performance was observed to be proportional to the amount of MDL in the system.

7. Conclusions

Optical communication systems based on FMF represent an attractive option for upgrading the data rate of long-haul systems. By means of the integration of inline components, FMF systems can potentially reduce substantially the power consumption compared to a system supporting the same number of data streams with independent SMFs. However, MDL posses a fundamental limitation when dealing with more than one mode. In this work, we have studied different aspects that yield an improved performance of MDL-impaired FMF systems for optical long-haul communications.

One of the investigated aspect was the impact of mode coupling on the system, shown in Chapter 4. Although weakly coupled modes could make possible a less complex MIMO equalizer scheme with lower order than $N_\mathrm{t} \times N_\mathrm{t}$, strongly coupled modes makes the MDL to accumulate at a lower rate with the distance than weakly coupled modes. With strongly coupled modes, the MDL accumulates proportionally with the square root of the distance; whereas with weakly coupled modes, MDL increases linearly with the distance. Either with distributed mode coupling or inducing mode coupling in certain positions along the link, e.g. with MSs, the transmission distance can be increased with strong mode coupling. The design of fibers with strong distributed mode coupling might be challenging, however, it might be possible with new techniques such as incorporating a grating in the fiber during its fabrication process [166]. The practical implementation of MSs for FMFs would have to be explored. This can be done with e.g. spatial (D)MUXs joint together or gratings.

The proper RX DSP can also help to improve the performance of MDL-impaired FMF systems. In Chapter 5 linear equalizers and two near-ML algorithms were compared in terms of performance and computational complexity. It was found than near-ML algorithms greatly outperform the linear equalizers. The SD approach achieves ML performance, but its computation requires a variable number of mathematical operations. The IRSML represents a good alternative to achieve near-ML performance at a fixed computational effort, which can be adjusted according to the required performance and the allowed computational complexity. The RX scheme proposed in [182, 187] is also interesting and should be tested in the optical communications for future work, since it administrates the allowed computational effort among all OFDM subcarriers providing more candidates to specific subcarriers when needed.

The experiments presented in Chapter 6 demonstrated the valuable use of near-ML techniques for the RX DSP. Six spatial modes were transmitted in single channel and WDM configuration. The algorithm IRSML was implemented and tested with this systems that after 74 km presented an MDL of 12 dB. It was found that IRSML improves the system performance only by using eight candidates from more than 16 million (QPSK) and 68 billion (8QAM) possible candidates. It was also found that the estimation and compensation of PN plays an important role for the system performance mostly in the high OSNR region, where the residual PN is dominant over the ASE noise. It is believed that a more accurate PN estimation would improve the detector output. The PN estimation can be improved by reducing the OFDM symbol length or by having independent data streams and combining the phase estimation of isolated pilot tones such as the so-called RF pilot method described in [124] for single-moded systems.

Undoubtedly, it is desirable to reduce MDL by eliminating the imperfection in optical components and splices. However, this is even more challenging when dealing with larger amount of supported group modes. Other alternative, to mitigate MDL has been proposed in [194]. In [194], a combination of space-time codes and mode scrambling is used resulting in a reduction of OSNR penalty in MDL impaired systems. A technique called waterfilling can be also used to enhance the performance of systems with MDL if there is information on the gain and losses along the link [64]. Waterfilling assigns different launch power to the modes according to the channel conditions. However, this techniques presents two drawbacks. Firstly, there is usually no information at the TX about the channel, since the delay of the round-trip from TX to RX might be longer than the channel dynamics [19]. Secondly, assigning more launch power to certain modes might arise severe nonlinearity penalty for those modes or all modes due to mode coupling.

In summary, the combination of strong mode coupling with a near-ML algorithm aids to increase the reach of MDL impaired systems. On the one hand, this requires to manufacture new FMFs with distributed strong mode coupling and/or MSs. Near-ML algorithms, on the other hand, require higher computational complexity than linear equalizer, however, significantly lower than ML detection.

A. Appendix

A.1. Numerical aperture

This appendix shows the derivation that leads to the definition of the NA in Eq. (2.4). According to Snell's law and following the notation in Fig. 2.4, the relation between the incident ray outside the fiber with maximum angle for total internal reflection and the angle of the refracted ray in the core is

$$n_0 \sin \theta_{0_{\max}} = n_{\text{co}} \sin \theta_{\text{co}}. \tag{A.1}$$

Note by observing Fig. 2.4 that at $\theta_1 = \theta_{\text{c}}$, the following equation is satisfied

$$\theta_{\text{c}} + \theta_{\text{co}} + \pi/2 = \pi. \tag{A.2}$$

Therefore, θ_{co} can be re-written as $\theta_{\text{co}} = \pi/2 - \theta_{\text{c}}$. Thus, Eq. (A.1) can be also re-written as

$$n_0 \sin \theta_{0_{\max}} = n_{\text{co}} \sin(\pi/2 - \theta_{\text{c}}) = n_{\text{co}} \cos \theta_{\text{c}}. \tag{A.3}$$

Then the relations in Eq. (A.4) are obtained by squaring both sides of Eq. (A.3) and taking into account that for total internal reflection $n_{\text{co}} \sin(\theta_{\text{co}}) = n_{\text{clad}}$.

$$n_0^2 \sin^2 \theta_{0_{\max}} = n_{\text{co}}^2 \cos^2 \theta_{\text{c}} = n_{\text{co}}^2 (1 - \sin^2 \theta_{\text{c}}) = n_{\text{co}}^2 (1 - n_{\text{clad}}^2 / n_{\text{co}}^2) = n_{\text{co}}^2 - n_{\text{clad}}^2 \tag{A.4}$$

Finally, Eq. (A.5) can be obtained and NA is defined by taking the square root of the rightmost and leftmost side of Eq. (A.4).

$$NA = n_0 \sin \theta_{0_{\max}} = \sqrt{n_{\text{co}}^2 - n_{\text{clad}}^2} \tag{A.5}$$

A.2. Propagation equation: dispersive effects and mode coupling

This appendix describes the derivation of equation Eq. (2.42). The derivation is divided in order to obtain one addend of Eq. (2.42) at the time, i.e. first it is assumed that there is no mode coupling (ideal waveguide) and secondly, it is assumed that there is no material wavelength dependency, therefore there are no dispersive effects.

For obtaining the dispersive part of Eq. (2.42), the propagation equation, Eq. (2.14b), is rewritten in the frequency domain as in Eq. (A.6). The last equation of Eq. (A.6) is the result of considering that $\kappa = \omega/c$, $c = 1/\sqrt{\epsilon_0\mu_0}$ and $n^2 = \epsilon_\mathrm{r}$.

$$\begin{aligned}\nabla^2\mathbf{E}_{lm}(x,y,z,\omega) - \mu_0\epsilon_0 n^2\left(j\omega\right)^2\mathbf{E}_{lm}(x,y,z,\omega) &= 0\\ \nabla^2\mathbf{E}_{lm}(x,y,z,\omega) + \mu_0\epsilon_0 n^2\omega^2\mathbf{E}_{lm}(x,y,z,\omega) &= 0\\ \nabla^2\mathbf{E}_{lm}(x,y,z,\omega) + \epsilon_\mathrm{r}\kappa^2\mathbf{E}_{lm}(x,y,z,\omega) &= 0\end{aligned} \tag{A.6}$$

Equation A.6 can be further expanded as in Eqs. (A.7a) to (A.7c), where the Laplace operator ∇^2 is expanded in the Cartesian coordinate system and the electric field is assumed to be of the form $\mathbf{E}_{lm}(x,y,z,\omega) = \hat{E}_{lm}(x,y)S_{lm}(z,\omega-\omega_\mathrm{c})e^{j\beta_{0_{lm}}z}$ as in Eq. (2.25a). In Eq. (A.7c), the common exponential term $e^{j\beta_{0_{lm}}z}$ has been canceled out from Eq. (A.7b). Additionally, the term related to the second derivative with respect to z of the slowly varying field S_{lm} in Eq. (A.7b) is neglected assuming the first derivative is more significant, since the change rate of S_{lm} is low with respect to z.

$$0 = S_{lm}e^{j\beta_{0_{lm}}z}\frac{\partial^2\hat{E}_{lm}}{\partial x^2} + S_{lm}e^{j\beta_{0_{lm}}z}\frac{\partial^2\hat{E}_{lm}}{\partial y^2} + \hat{E}_{lm}\frac{\partial^2}{\partial z^2}\left(S_{lm}e^{j\beta_{0_{lm}}z}\right) + \epsilon_\mathrm{r}\kappa^2\hat{E}_{lm}S_{lm}e^{j\beta_{0_{lm}}z} \tag{A.7a}$$

$$\begin{aligned}0 = {}& S_{lm}e^{j\beta_{0_{lm}}z}\frac{\partial^2\hat{E}_{lm}}{\partial x^2} + S_{lm}e^{j\beta_{0_{lm}}z}\frac{\partial^2\hat{E}_{lm}}{\partial y^2} + \\ & \hat{E}_{lm}\left(e^{j\beta_{0_{lm}}z}\frac{\partial^2 S_{lm}}{\partial z^2} + 2j\beta_{0_{lm}}e^{j\beta_{0_{lm}}z}\frac{\partial S_{lm}}{\partial z} - \beta_{0_{lm}}^2 S_{lm}e^{j\beta_{0_{lm}}z}\right) + \epsilon_\mathrm{r}\kappa^2\hat{E}_{lm}S_{lm}e^{j\beta_{0_{lm}}z}\end{aligned} \tag{A.7b}$$

$$0 = S_{lm}\frac{\partial^2\hat{E}_{lm}}{\partial x^2} + S_{lm}\frac{\partial^2\hat{E}_{lm}}{\partial y^2} + \hat{E}_{lm}\left(2j\beta_{0_{lm}}\frac{\partial S_{lm}}{\partial z} - \beta_{0_{lm}}^2 S_{lm}\right) + \epsilon_\mathrm{r}\kappa^2\hat{E}_{lm}S_{lm} \tag{A.7c}$$

Then, the method of separation of variables is applied, i.e. all terms in Eq. (A.7c) are divided by the product $S_{lm}\hat{E}_{lm}$, which leads to Eq. (A.8). In Eq. (A.8), both equations for S_{lm} and $\hat{E}_{lm}$, respectively, are equated to the eigenvalue β_{lm}^2, which corresponds to the propagation constant.

$$\frac{1}{\hat{E}_{lm}}\frac{\partial^2 \hat{E}_{lm}}{\partial x^2}+\frac{1}{\hat{E}_{lm}}\frac{\partial^2 \hat{E}_{lm}}{\partial y^2}+\epsilon_{\mathrm{r}}\kappa^2=-\frac{1}{S_{lm}}2j\beta_{0_{lm}}\frac{\partial S_{lm}}{\partial z}+\beta_{0_{lm}}^2=\beta_{lm}^2 \tag{A.8}$$

Then, Eq. (A.8) yields Eqs. (A.9a) and (A.9b) after separating the equations. According to [25], $\beta_{lm}^2-\beta_{0_{lm}}^2$ can be approximated to $2\beta_{0_{lm}}\left(\beta_{lm}-\beta_{0_{lm}}\right)$ leading to the right hand side part of Eq. (A.9b). With Eq. (A.9b), the derivation for the dispersive part of Eq. (2.42) is concluded.

$$\frac{\partial^2 \hat{E}_{lm}}{\partial x^2}+\frac{\partial^2 \hat{E}_{lm}}{\partial y^2}+\hat{E}_{lm}\left(\epsilon_{\mathrm{r}}\kappa^2-\beta_{lm}^2\right)=0 \tag{A.9a}$$

$$\frac{\partial S_{lm}}{\partial z}=\frac{-S_{lm}}{2j\beta_{0_{lm}}}\left(\beta_{lm}^2-\beta_{0_{lm}}^2\right)=-jS_{lm}\left(\beta_{0_{lm}}-\beta_{lm}\right) \tag{A.9b}$$

As for the derivation of the mode coupling part of Eq. (2.42), the equation to start is also Eq. (A.6), which is rewritten in Eq. (A.10) as a function of the permittivity and permeability.

$$\nabla^2\mathbf{E}(x,z,\omega)+\epsilon\mu_0\omega^2\mathbf{E}(x,z,\omega)=0 \tag{A.10}$$

Now it is assumed that the total electrical field is given by the sum of all contributions from all supported modes u_1 as in Eq. (A.11).

$$\mathbf{E}(x,z,\omega)=\sum_{u_1}\hat{E}_{u_1}(\mathbf{r})S_{u_1}(z,\omega-\omega_{\mathrm{c}})e^{j\beta_{0_{u_1}}z} \tag{A.11}$$

Equation A.11 is replaced in Eq. (A.10) as in Eq. (A.12a). In Eqs. (A.12b) to (A.12d), Eq. (A.12a) is further expanded by reorganizing the terms, replacing the second derivative with respect of z and replacing ϵ by the sum of the permittivity of the unperturbed medium ϵ_{a} and the perturbation $\Delta\epsilon$.

$$\begin{aligned}0 = {} & \frac{\partial^2}{\partial x^2}\left[\sum_{u_1}\hat{E}_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\right] + \frac{\partial^2}{\partial y^2}\left[\sum_{u_1}\hat{E}_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\right] + \frac{\partial^2}{\partial z^2}\left[\sum_{u_1}\hat{E}_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\right] + \\ & \epsilon\mu_0\omega^2\left[\sum_{u_1}\hat{E}_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\right]\end{aligned} \tag{A.12a}$$

$$\begin{aligned}0 = {} & \sum_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\frac{\partial^2}{\partial x^2}\hat{E}_{u_1} + \sum_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\frac{\partial^2}{\partial y^2}\hat{E}_{u_1} + \sum_{u_1}\hat{E}_{u_1}\frac{\partial^2}{\partial z^2}\left[S_{u_1}e^{j\beta_{0_{u_1}}z}\right] + \\ & \epsilon\mu_0\omega^2\left[\sum_{u_1}\hat{E}_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\right]\end{aligned} \tag{A.12b}$$

$$\begin{aligned}0 = {} & \sum_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\frac{\partial^2}{\partial x^2}\hat{E}_{u_1} + \sum_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\frac{\partial^2}{\partial y^2}\hat{E}_{u_1} + \\ & \sum_{u_1}\hat{E}_{u_1}\left[2j\beta_{0_{u_1}}e^{j\beta_{0_{u_1}}z}\frac{\partial}{\partial z}S_{u_1} - \beta_{0_{u_1}}^2S_{u_1}e^{j\beta_{0_{u_1}}z}\right] + \epsilon\mu_0\omega^2\left[\sum_{u_1}\hat{E}_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\right]\end{aligned} \tag{A.12c}$$

$$\begin{aligned}0 = {} & \sum_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\frac{\partial^2}{\partial x^2}\hat{E}_{u_1} + \sum_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\frac{\partial^2}{\partial y^2}\hat{E}_{u_1} + \\ & \sum_{u_1}\hat{E}_{u_1}\left[2j\beta_{0_{u_1}}e^{j\beta_{0_{u_1}}z}\frac{\partial}{\partial z}S_{u_1} - \beta_{0_{u_1}}^2S_{u_1}e^{j\beta_{0_{u_1}}z}\right] + \epsilon_{\mathrm{a}}\mu_0\omega^2\left[\sum_{u_1}\hat{E}_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\right] + \\ & \Delta\epsilon\mu_0\omega^2\left[\sum_{u_1}\hat{E}_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\right]\end{aligned} \tag{A.12d}$$

Reorganizing the terms in Eq. (A.12d), Eqs. (A.13a) and (A.13b) are obtained. The first term of the right hand side of Eq. (A.13a) is canceled out since it resembles the propagation in an unperturbed medium as in Eq. (A.9a), yielding Eq. (A.13b).

$$\begin{aligned}0 = {} & \sum_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\left[\frac{\partial^2}{\partial x^2}\hat{E}_{u_1} + \frac{\partial^2}{\partial y^2}\hat{E}_{u_1} + \hat{E}_{u_1}\left(\epsilon_{\mathrm{a}}\mu_0\omega^2 - \beta_{0_{u_1}}^2\right)\right] + \\ & \sum_{u_1}\hat{E}_{u_1}\left[2j\beta_{0_{u_1}}e^{j\beta_{0_{u_1}}z}\frac{\partial}{\partial z}S_{u_1}\right] + \Delta\epsilon\mu_0\omega^2\left[\sum_{u_1}\hat{E}_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\right]\end{aligned} \tag{A.13a}$$

$$0 = \sum_{u_1}\hat{E}_{u_1}\left[2j\beta_{0_{u_1}}e^{j\beta_{0_{u_1}}z}\frac{\partial}{\partial z}S_{u_1}\right] + \Delta\epsilon\mu_0\omega^2\left[\sum_{u_1}\hat{E}_{u_1}S_{u_1}e^{j\beta_{0_{u_1}}z}\right] \tag{A.13b}$$

Equation A.13b can be simplified by multiplying by $\hat{E}_u^*$ and integrating over all x and y. As the modes follow the orthonormality relationship Eq. (A.14) [48], there is only one addend that survives in the first term of the right hand side of Eq. (A.13b), i.e. $u = u_1$. In Eq. (A.14), the factor $\frac{2\omega\mu_0}{\beta_{0_u}}$ corresponds to the normalization of the electrical field to 1 W/m^2 and $\delta_{u_1 u}$ is the Kronecker delta function, which is one when $u = u_1$ and zero otherwise. Hence Eq. (A.13b) is rewritten as Eq. (A.15a). The equations are further simplified in Eqs. (A.15b) to (A.15d), where C_{u,u_1} and $\beta_{0_{u,u_1}}$ are defined in Section 2.3.4.

$$\frac{2\omega\mu_0}{\beta_{0_u}} \int\int \left(\hat{E}_{u_1} \cdot \hat{E}_u^*\right) dxdy = \delta_{u_1 u} \tag{A.14}$$

$$2j\beta_{0_u} \left[\frac{2\omega\mu_0}{\beta_{0_u}}\right] \frac{\partial}{\partial z} S_u e^{j\beta_{0_u} z} = -\mu_0\omega^2 \sum_{u_1} S_{u_1} e^{j\beta_{0_{u1}} z} \int\int \hat{E}_u^* \Delta\epsilon \hat{E}_{u_1} dxdy \tag{A.15a}$$

$$\frac{\partial}{\partial z} S_u e^{j\beta_{0_u} z} = -\frac{\omega}{4j} \sum_{u_1} S_{u_1} e^{j\beta_{0_{u1}} z} \int\int \hat{E}_u^* \Delta\epsilon \hat{E}_{u_1} dxdy \tag{A.15b}$$

$$\frac{\partial}{\partial z} S_u = j\frac{\omega}{4} \sum_{u_1} S_{u_1} e^{j\left(\beta_{0_{u1}} - \beta_{0_u}\right) z} \int\int \hat{E}_u^* \Delta\epsilon \hat{E}_{u_1} dxdy \tag{A.15c}$$

$$\frac{\partial}{\partial z} S_u = j \sum_{u_1} C_{u,u_1} S_{u_1} e^{-j\Delta\beta_{0_{u,u_1}} z} \tag{A.15d}$$

The propagation equation Eq. (2.42) is then the result of the sum of Eqs. (A.9b) and (A.15d). The term in the summation in Eq. (A.15d) corresponding to $u = u_1$ is often in the literature neglected since it only contributes to a phase rotation [36,46].

A.3. Derivation of the IQ modulator output for laser comb generation

This appendix describes the derivation of Eq. (6.21). The starting point is the equation for the IQ modulator Eq. (3.8a). In the case of the laser comb u_{I} is the sum of a DC part and a sinusoidal function from the signal generator, i.e. $u_{\mathrm{BIAS_I}} + a_{\mathrm{RF}} \cos(\omega_{\mathrm{RF}} t)$. Thus, Eq. (3.8a) is transformed into Eq. (A.16) for the laser comb taking into account that the driving voltages of the I and Q parts are described in Eq. (A.17).

$$E_{\mathrm{out_{cmb}}} = \frac{E_{\mathrm{in_{cmb}}}}{2} \left[\cos\left(\frac{\pi}{V_\pi} u_{\mathrm{BIAS_I}} + \frac{\pi}{V_\pi} a_{\mathrm{RF}} \cos\left(\omega_{\mathrm{RF}} t\right)\right) + e^{j\Theta_{\mathrm{PM}}} \cos\left(\frac{\pi}{V_\pi} u_{\mathrm{BIAS_Q}}\right)\right] \tag{A.16}$$

$$u_{\mathrm{I}}(t) = u_{\mathrm{BIAS_I}} + u_{\mathrm{RF}}(t) = u_{\mathrm{BIAS_I}} + a_{\mathrm{RF}} \cos\left(\omega_{\mathrm{RF}} t\right) \tag{A.17a}$$

$$u_{\mathrm{Q}}(t) = u_{\mathrm{BIAS_Q}} \tag{A.17b}$$

Applying the trigonometric identity $\cos\left(x+y\right) = \cos x \cos y - \sin x \sin y$, Eq. (A.16) can be rewritten as in Eq. (A.18). In Eq. (A.19), $\frac{\pi u_{\mathrm{BIAS_I}}}{V_\pi}$, $\frac{\pi u_{\mathrm{BIAS_I}}}{V_\pi}$ and $\frac{\pi a_{\mathrm{RF}}}{V_\pi}$ are replaced for convenience by Θ_{I}, Θ_{Q} and β_{a}, respectively.

$$\begin{aligned} E_{\mathrm{out_{cmb}}} = \frac{E_{\mathrm{in_{cmb}}}}{2} \Bigg[& \cos\left(\frac{\pi u_{\mathrm{BIAS_I}}}{V_\pi}\right) \cos\left(\frac{\pi a_{\mathrm{RF}}}{V_\pi} \cos\left(\omega_{\mathrm{RF}} t\right)\right) \\ & - \sin\left(\frac{\pi u_{\mathrm{BIAS_I}}}{V_\pi}\right) \sin\left(\frac{\pi a_{\mathrm{RF}}}{V_\pi} \cos\left(\omega_{\mathrm{RF}} t\right)\right) + e^{j\Theta_{\mathrm{PM}}} \cos\left(\frac{\pi}{V_\pi} u_{\mathrm{BIAS_Q}}\right) \Bigg] \end{aligned} \tag{A.18}$$

$$\begin{aligned} E_{\mathrm{out_{cmb}}} = \frac{E_{\mathrm{in_{cmb}}}}{2} & \left[\cos\left(\Theta_{\mathrm{I}}\right) \cos\left(\beta_{\mathrm{a}} \cos\left(\omega_{\mathrm{RF}} t\right)\right) \right. \\ & \left. - \sin\left(\Theta_{\mathrm{I}}\right) \sin\left(\beta_{\mathrm{a}} \cos\left(\omega_{\mathrm{RF}} t\right)\right) + e^{j\Theta_{\mathrm{PM}}} \cos\left(\Theta_{\mathrm{Q}}\right)\right] \end{aligned} \tag{A.19}$$

Cosine and sine functions of another sinusoidal function can be expressed in terms of Bessel functions of first kind as in Eqs. (A.20a) and (A.20b).

$$\cos\left(x \sin\left(\phi\right)\right) = J_0\left(x\right) + 2\sum_{n=1}^{\infty} J_{2n}\left(x\right) \cos\left(2n\phi\right) \tag{A.20a}$$

$$\sin\left(x \sin\left(\phi\right)\right) = 2\sum_{n=0}^{\infty} J_{2n+1}\left(x\right) \sin\left(\left(2n+1\right)\phi\right) \tag{A.20b}$$

Applying this property, $\cos\left(\beta_{\mathrm{a}} \cos\left(\omega_{\mathrm{RF}} t\right)\right)$ and $\sin\left(\beta_{\mathrm{a}} \cos\left(\omega_{\mathrm{RF}} t\right)\right)$ can be expanded as in Eqs. (A.21a) and (A.21b), and Eqs. (A.21c) and (A.21d), respectively.

$$\cos\left(\beta_{\mathrm{a}} \cos\left(\omega_{\mathrm{RF}} t\right)\right) = \cos\left(\beta_{\mathrm{a}} \sin\left(\omega_{\mathrm{RF}} t + \frac{\pi}{2}\right)\right) \tag{A.21a}$$

$$= J_0(\beta_{\mathrm{a}}) + 2\sum_{n=1}^{\infty} J_{2n}\left(\beta_{\mathrm{a}}\right) \cos\left(2n\left(\omega_{\mathrm{RF}} t + \frac{\pi}{2}\right)\right) \tag{A.21b}$$

$$\sin\left(\beta_{\mathrm{a}} \cos\left(\omega_{\mathrm{RF}} t\right)\right) = \sin\left(\beta_{\mathrm{a}} \sin\left(\omega_{\mathrm{RF}} t + \frac{\pi}{2}\right)\right) \tag{A.21c}$$

$$= 2\sum_{n=0}^{\infty} J_{2n+1}\left(\beta_{\mathrm{a}}\right) \sin\left(\left(2n+1\right)\left(\omega_{\mathrm{RF}} t + \frac{\pi}{2}\right)\right) \tag{A.21d}$$

By observing Eq. (A.21b), it can be noticed that the phase of the cosine is $n\frac{\pi}{2}$. This means that if n is odd $\cos\left(2n\omega_{\mathrm{RF}}t + \frac{n\pi}{2}\right) = -\cos\left(2n\omega_{\mathrm{RF}}t\right)$; if n is even $\cos\left(2n\omega_{\mathrm{RF}}t + \frac{n\pi}{2}\right) = \cos\left(2n\omega_{\mathrm{RF}}t\right)$. Hence, the term in the sumation in Eq. (A.21b) can be multiplied by $-(-1)^{n-1}$ and the phase of the cosine function can be removed. Equation A.21d can be similarly be processed. The phase of the sine is $\left(2n+1\right)\frac{\pi}{2}$, i.e. if n is odd $\sin\left(\left(2n+1\right)\omega_{\mathrm{RF}}t + \frac{n\pi}{2}\right) = -\cos\left(2n\omega_{\mathrm{RF}}t\right)$ and if n is even $\sin\left(\left(2n+1\right)\omega_{\mathrm{RF}}t + \frac{n\pi}{2}\right) = \cos\left(2n\omega_{\mathrm{RF}}t\right)$. According to this considerations, the sine function can be converted to cosine and the term in the summation can is multiplied by $-(-1)^{n-1}$. Thus, Eqs. (A.21b) and (A.21d) are replaced in Eq. (A.19) and the phase considerations just mentioned above result in Eq. (A.22), which agree with Eq. (6.21).

$$\begin{aligned}E_{\mathrm{out_{cmb}}} = \frac{E_{\mathrm{in_{cmb}}}}{2}\Bigg[&\cos\Theta_{\mathrm{I}}J_0(\beta_{\mathrm{m}}) - 2\cos\Theta_{\mathrm{I}}\sum_{n=1}^{\infty}\left(-1\right)^{n-1}J_{2n}(\beta_{\mathrm{m}})\cos\left(2n\omega_{\mathrm{RF}}\right)\\ &- 2\sin\Theta_{\mathrm{I}}\sum_{n=0}^{\infty}\left(-1\right)^{n-1}J_{2n+1}(\beta_{\mathrm{m}})\cos\left(\left(2n+1\right)\omega_{\mathrm{RF}}t\right) + e^{j\Theta_{\mathrm{PM}}}\cos\Theta_{\mathrm{Q}}\Bigg]\end{aligned} \tag{A.22}$$

B. List of Abbreviations, Symbols and Operators

B.1. List of Abbreviations

16QAM	16-level quadrature amplitude modulation
3DW	3-dimensional waveguide
8QAM	8-level quadrature amplitude modulation
ADC	Analog-to-digital converter
ASE	Amplified spontaneous emission
AWG	Arbitrary waveform generator
AWGN	Additive white Gaussian noise
B2B	Back-to-back
BER	Bit error rate
BPF	Bandpass filter
BSF	Band-stop filter
CAZAC	Constant amplitude zero auto-correlation
CFO	Carrier frequency offset

CP	Cyclic prefix
CPE	Common phase error
CSOC	Convolutional self-orthogonal code
CUT	Channel under test
DAC	Digital-to-analog converter
DFB	Distributed feedback laser
DFT	Discrete Fourier transform
DGD	Differential group delay
DMD	Differential mode delay
DML	Directly modulated lasers
DMUX	Demultiplexer
DSP	Digital signal processing
EAM	Electro-absortion modulators
ECL	External cavity laser
EDFA	Erbium-doped fiber amplifier
EM	Expectation maximization
FDM	Frequency-division multiplexing
FEC	Forward-error correction
FFT	Fast Fourier transform
FIR	Finite impulse response
FMA	Few-mode amplifier
FMF	Few-mode fiber
FPGA	Field programmable gate array
GVD	Group velocity dispersion
ICI	Intercarrier interference
IDFT	Inverse discrete Fourier transform
IEC	International Electrotechnical Commission
IF	Intermediate frequency
IFFT	Inverse fast Fourier transform
IIR	Infinite impulse response
IRSML	Improved reduced-search ML
ISI	Inter-symbol interference
ITU	International Telecommunication Union
LCoS	Liquid crystal on silicon
LDPC	Low-density parity check
LED	Light emitting diode
LLH	Log likelihood
LO	Local oscillator
LPF	Low pass filter
LW	Linewidth
MAP	Maximum *a posteriori*
MCF	Multicore fiber

MDG	Mode-dependent gain
MDL	Mode-dependent loss
MDM	Mode-division multiplexing
MEMS	Micro-electromechanichal system
MI	Mutual information
MIMO	Multiple-input multiple output
ML	Maximum-likelihood
MMF	Multimode fiber
MMSE	Minimum-mean square error
MS	Mode scrambler
MUX	Multiplexer
MZM	Mach-Zehnder modulator
NA	Numerical aperture
OFDM	Orthogonal-frequency division multiplexing
OSNR	Optical signal-to-noise ratio
PA	Pilot-aided
PAPR	Peak-to-average power ratio
PBC	Polarization beam combiner
PDF	Probability density function
PDL	Polarization-dependent loss
PMD	Polarization-mode dispersion
PN	Phase noise
POLMUX	Polarization-multiplexed
PSD	Power spectral density
QPA	Quasi-pilot aided
QPSK	Quadrature phase shift keying
RIN	Relative intensity noise
ROADM	Reconfigurable optical add/drop multiplexer
RS	Reed-Solomon
RX	Receiver
SD	Sphere decoding
SDM	Spatial-division multiplexing
SE	Spectral efficiency
SLM	Spatial light modulator
SMF	Single-mode fiber
SNR	Signal-to-noise ratio
SRS	Stimulated Raman scattering
SVD	Singular value decomposition
TBER	Target BER
TIA	Transimpedance amplifier
TS	Training symbol
TX	Transmitter
UV	Ultraviolet

WDM	Wavelength-division multiplexing
wIRSML	Weighted IRSML
WSS	Wavelength selective switch
ZF	Zero-forcing
ZP	Zero padding

B.2. List of Operators

$\det\{\cdot\}$	Determinant
$(\cdot)^{-1}$	Inverse
$(\cdot)^{*}$	Complex conjugate
$(\cdot)^{\mathrm{H}}$	Conjugate transposition
$(\cdot)^{\mathrm{T}}$	Transpose of a vector or matrix
$\circ$	Hadamard product operator
$\lfloor\cdot\rceil$	Round to the closest constellation point
$\log_{a}(\cdot)$	Logarithm in base a
$\arg(\cdot)$	Argument
∇	Vector differential operator
$\arg\min_{\mu}\{\cdot\}$	Minimum argument among μ
$\times$	Cartesian product
$Im\{\cdot\}$	Imaginary part of a complex-valued matrix, scalar or vector
$Re\{\cdot\}$	Real part of a complex-valued matrix, scalar or vector

B.3. List of Symbols

(r, φ, z)	Radial, azimuthal and longitudinal components in the cylindrical coordinate system
α	Refractive index decay factor in the core region
α_{dB}	Fiber attenuation in [dB/km]
α_{p}	Fiber attenuation in [Np/km]
$\bar{\beta}_{g}$	Mode-averaged g-th derivative of the propagation constant with respect to the frequency
$\bar{D}$	Mode-averaged dispersion parameter in $[\mathrm{ps}\cdot\mathrm{nm}^{-1}\cdot\mathrm{km}^{-1}]$
$\bar{n}_{\mathrm{gr},lm,\mathrm{x/y}}$	Group effective refractive index for mode LP_{lm} and polarization x or y
$\bar{n}_{\mathrm{gr},lm}$	Group effective refractive index related to the mode LP_{lm}
$\bar{n}_{lm}$	Effective refractive index related to the mode LP_{lm}
$\bar{P}_{u}$	Average power in mode u

$\bar{S}_{\mathrm{lo}}$	Mode-averaged intra-modal dispersion transfer function in $[\mathrm{ps} \cdot \mathrm{nm}^{-2} \cdot \mathrm{km}^{-1}]$
β_{a}	Abbreviation of $\frac{\pi a_{\mathrm{RF}}}{V_\pi}$
β_{lossy}	Propagation constant of a mode whose propagation constant is close to the cladding index in [rad/m]
β_{max}	Maximum propagation constant among the bounded modes in [rad/m]
β_{min}	Minimum propagation constant among the bounded modes in [rad/m]
β_{r}	Propagation constant of the radiation modes in [rad/m]
$\beta_{g,lm,\mathrm{x/y}}$	g-th derivative of the propagation constant for mode LP_{lm} and polarization x or y
β_{lm}	Propagation constant related to the mode lm in [rad/m]
χ^n	n^{th} order susceptibility
χ_{e}	Susceptibility of the medium
$\Delta\bar{n}_{\mathrm{gr},lm,\mathrm{xy}}$	Difference between the polarizations group effective refractive index
$\Delta\bar{n}_{lm,\mathrm{xy}}$	Birefringence
$\Delta\beta_{\mathrm{lossy}}$	Difference between β_{min} and β_{lossy}
$\Delta\beta_{0_{u,u_1}}$	Propagation constant difference between mode u and u_1 at the carrier frequency
$\Delta\beta_{ij}$	Difference between the propagation constants of mode i and j
$\Delta\epsilon$	Permittivity perturbation
$\Delta\lambda$	Spectral width in [nm]
$\Delta\phi_{\mathrm{c}}(t, \Delta T)$	Temporal change of phase during a time interval ΔT
Δf	Spectral width in [Hz]
Δn	Relative refractive index profile
Δn_{PM}	Refractive index change in a phase modulator
ΔT	Pulse broadening from dispersion in [ps]
$\Delta v_{\mathrm{gr},lm,\mathrm{xy}}$	Group velocity difference between x and y polarizations
Δ	Difference between the two metrics taken into account for the wIRSML computation
$\delta_{ll'}$	Kronecker delta function, $\delta_{ll'} = 1$ for $l = l'$ and $\delta_{ll'} = 0$ otherwise
ϵ	Dielectric permittivity of the medium in $\mathrm{C}^2/\mathrm{Nm}^2$
ϵ_0	Permittivity of a vacuum ($\epsilon_0 = 8.854 \times 10^{-12}\ \mathrm{C}^2/\mathrm{Nm}^2$)

ϵ_{a}	Permittivity of the ideal medium
ϵ_{r}	Dielectric constant or relative permittivity
η_{pho}	Quantum efficiency of a photodiode
$\eta_{\mathrm{p}j,\mathrm{s}i}$	Overlap integral between signal of i-th mode and pump distributions
Γ	Overlap integral between the modulating electric field and the confined optical field in the waveguide of a phase modulator
$\gamma(z_{k_n})$	*A posteriori* of a received symbol n related to cluster k in EM algorithm context
$\Gamma_{\mathrm{p},j}(r,\varphi)$	Normalized optical intensity distribution of the signal
$\Gamma_{\mathrm{s},i}(r,\varphi)$	Normalized Er^{3+} distribution
$\hat{E}_{lm}(\mathbf{r})$,$\hat{E}_{lm}((x,y)$	Transverse component of the electric field for the mode LP_{lm}
$\hat{I}$	Mutual information in the presence of MDL in [bits/s/Hz]
κ	Wavenumber in [rad/m]
λ	Wavelength in [m]
λ_{c}	Wavelength of the optical carrier

$\mathbb{A}$	Set of hypersymbol candidates for the IRSML algorithm excluding the output of the detector after the MMSE equalization $\hat{\mathbf{s}}_{\mathrm{MMSE}}$
$\mathbb{K}_{\mathrm{QPA/PA}}$	Set of pilot subcarrier indexes used for QPA and PA PN compensation
$\mathbb{M}$	Set of symbols candidates for the ML detector
$\mathbb{S}$	Set of hypersymbol candidates for the IRSML algorithm
$\mathbb{S}_i$	Set of hypersymbol candidates for the tributary i
$\mathbb{S}_{R'}(\mathbf{r}_{\mathrm{o}},\mathbf{H})$	Set of possible hypersymbols within the hypersphere radius R' for SD
$\mathbb{S}_{R}(\mathbf{Q}^{\mathrm{H}}\mathbf{r}_{\mathrm{o}},\mathbf{R})$	Set of possible hypersymbols within the hypersphere radius R for SD
$\mathbb{U}$	Set of all possible supported modes by the FMF except for the mode from where a pulse is launched to study modal crosstalk
$\mathbb{U}_{\mathrm{total}}$	Set of all possible supported modes by the FMF
$\mathit{DMD}_{\mathrm{p}}$	DMD parameter in [ps/km]
NF	Noise figure of an optical amplifier

NF_{dB}	Noise figure of an optical amplifier in [dB]
$OSNR$	Optical signal-to-noise ratio
$OSNR_{\mathrm{dB}}$	Optical signal-to-noise ratio in [dB]
PMD_c	PMD coefficient in $[\mathrm{ps}/\sqrt{\mathrm{km}}]$
$\mathrm{CP}_{\mathrm{ov}}$	CP overhead
LP_{lm}	linearly polarized mode lm
$\mathrm{TS}_{\mathrm{ov}}$	TS overhead
μ	Permeability of a medium in [H/m]
μ_0	Permeability of a vacuum ($\mu_0 = 4\pi \times 10^{-7}\ \mathrm{Ns^2/C^2}$)
ω	Angular frequency in [rad/s]
ω_{c}	Carrier angular frequency in [rad/s]
ω_{IF}	Intermediate angular frequency in [rad/s]
ω_{RF}	Angular frequency of $u_{\mathrm{RF}}(t)$ in [rad/s]
$\Phi(\varphi)$	Azimuthal component of the transverse electromagnetic field
ϕ_0	Initial phase of the optical source electric field
Φ_{CPE}	Total estimated common phase error from the QPA, PA and EM methods
$\phi_{\mathrm{c}}(t)$	Laser electric field phase
Φ_{PA}	Estimated common phase error due to PN via the PA method
Φ_{QPA}	Estimated common phase error due to PN via the QPA method
ϕ_{sp_i}	Phase of the electric field of the laser spontaneous emission sp_i
π_k	Mixing coefficients associated to the k-th cluster in EM algorithm context
$\Psi(r,\varphi)$	Electric or magnetic field
$\psi(r,\varphi)$	Transverse component of the electromagnetic field
ψ_β	Transverse field component of radiation modes
ψ_{spot_n}	n^{th} launch field of a spot coupler
ρ	Electric charge density
$\sigma^2_{\boldsymbol{\eta}}$	ASE noise variance
$\sigma^2_{\mathbf{s}}$	Transmit signal variance
τ_{DMD}	Delay spread caused by DMD [ps]
$\tau_{\mathrm{max_D}}$	Maximum delay spread due to the mode with the largest dispersion parameter in [ps]
τ_{PM}	Delay caused by phase modulator

$\Theta_{\mathrm{I/Q}}$	Abbreviation of $\frac{\pi u_{\mathrm{BIAS}_{\mathrm{I/Q}}}}{V_\pi}$
θ_{k}	angle of incidence in medium k
Θ_{PM}	Abbreviation of $\frac{\pi u_{\mathrm{PM}}}{V_\pi}$
$\theta_{0_{\max}}$	Maximum angle of incidence between the ray and the fiber axis for total internal reflection
Υ	Mode-dependent loss
$\boldsymbol{\eta}_{\mathrm{ASE}}$	Additive noise vector
$\hat{\mathbf{s}}_{\mathrm{IRSML}}$	Signal vector after IRSML detection
$\hat{\mathbf{s}}_{\mathrm{ML}}$	Signal vector after ML detection
$\hat{\mathbf{s}}_{\mathrm{SD}}$	Detector output vector after performing the SD algorithm
$\mathbf{\Lambda}$	Diagonal matrix from SVD with elements $\lambda_1, \lambda_2, ..., \lambda_{N_{\mathrm{t}}}$
$\boldsymbol{\mu}_k$	Mean vector of the constellation points associated to the k-th cluster in EM algorithm context
$\mathbf{\Sigma}_k$	Covariance matrix of the constellation points associated to the k-th cluster in EM algorithm context
$\mathbf{B}$	Magnetic induction vector
$\mathbf{D}$	Electric displacement vector
$\mathbf{E}$	Electric field vector
$\mathbf{e}$	First hypersymbol candidate for the SD algorithm
$\mathbf{E}_{\mathrm{c}}(t)$	Laser electric field
$\mathbf{E}_{\mathrm{r}}$	Received electrical field of a single-moded beam after spatial demultiplexing
$\mathbf{G}_{\mathrm{MMSE}}$	MMSE filter matrix
$\mathbf{G}_{\mathrm{ZF}}$	ZF filter matrix
$\mathbf{H}$	Channel transfer function including optical channel, (D)MUX, TX and RX filters with entries h_{ij}
$\mathbf{H}$	Magnetic field vector
$\mathbf{J}$	Electric current density vector
$\mathbf{P}$	Polarization density vector
$\mathbf{Q}$	Q matrix for QR decomposition of H
$\mathbf{R}$	R matrix for QR decomposition of H
$\mathbf{r}$	Vector of the cylindrical coordinate system components
$\mathbf{r}_{\mathrm{o}}$	Channel output vector
$\mathbf{s}$	Channel input vector
$\mathbf{U}$	One of the two unitary matrix from SVD
$\mathbf{V}$	One of the two unitary matrix from SVD

$\mathbf{y}_{\mathrm{EM}}$	Vector containing the imaginary and real part of the received symbol at the input of the EM stage
$\mathbf{y}_{\mathrm{MMSE}}$	Signal vector after MMSE equalization
$\mathbf{y}_{\mathrm{ZF}}$	Signal vector after ZF equalization
$a_{\mathrm{RF}} \cos(\omega_{\mathrm{RF}} t)$	Amplitude of $u_{\mathrm{RF}}(t)$
a_n and b_n	n^{th} overlap integral between the launch fields of a spot coupler and the complex modes in the FMF. a and b correspond to the outer and inner ring of the spot arrange, respectively
$A_{\mathrm{eff}_{lm}}$	Effective area of mode lm in [μm^2]
B_{ref}	Reference bandwidth for OSNR measurements in [Hz]
C	Capacity in absence of MDL in [bits/s/Hz]
c	Speed of light in the vacuum ($c \cong 2.998 \times 10^8$ m/s)
$c(\beta)$	Scaling constant for radiation modes
c_1	Complex multiplications required for a matrix inversion
c_2	Complex multiplications required for performing the MMSE equalization
C_∞	Capacity for infinite number of modes in [bits/s/Hz]
c_{lm}	Scaling constant for guided modes
C_{u,u_1}	Mode coupling coefficients between mode u and u_1
d_{e}	Distance between two electrodes of a phase modulator in [m]
$D_{\mathrm{max}_{lm}}$	Maximum dispersion parameter among the modes [ps$\cdot$nm$^{-1}\cdot$ km^{-1}]
$D_{lm}(\lambda)$	Dispersion parameter for the mode lm in [ps $\cdot$ nm^{-1} $\cdot$ km^{-1}]
e	Charge of an electron ($e = 1.6 \times 10^{-19}$ C)
E_0	Initial magnitude of the optical source electric field
$E_{\mathrm{c}}(t)$	Laser electric field magnitude
E_{LO}	Electrical field of the local oscillator
$E_{\mathrm{out/in}_{\mathrm{cmb}}}$	Electrical field at the output/input of the comb IQ modulator
$E_{\mathrm{r}_{\mathrm{x/y}}}$	Single polarization component of the received electrical field of a single-moded beam after spatial demultiplexing
E_{sp_i}	Electric field of the laser spontaneous emission sp$_i$

f	Frequency in [Hz]
f_0	Reference frequency in [Hz]
f_{c}	Carrier frequency in [Hz]
f_{pump}	Frequency of the pump laser in [Hz]
$f_{\mathrm{pho_{in}}}$	Frequency of the incoming optical signal to a photodiode in [Hz]
G_{lm}	Amplifier gain for mode LP_{lm}
G_{lm}	Optical amplifier gain for mode LP_{lm}
h	Planck's constant ($h = 6.626068 \times 10^{-34}$ J s)
$H_{\bar{\mathrm{D}}}(f)$	Mode-averaged intra-modal dispersion transfer function
$I_{\mathrm{pho_{out}}}$	Output current of a photodiode in [A]
J_{ν}	Bessel functions of first kind of order ν
K	Total number of clusters or the constellation cardinality in EM algorithm context
k	Discreet time variable
L	Fiber length in [km]
l	Azimuthal mode number
L_{e}	Electrodes length of a phase modulator

L_{span}	Length of a single fiber span
L_B	Beat length
M	Constellation cardinality
m	Radial mode number
n	Discrete time variable
$n(r)$,$n\,(x, y)$	Refractive index profile
N_0	Single-polarization one-sided ASE noise power spectral density
n_0	Refractive index of the medium outside the fiber
n_{clad}	Cladding refractive index
n_{co}	Core refractive index
N_{CP}	Number of samples in the CP
N_{DATA}	Data subcarriers
N_{FFT}	Number of subcarrier used for an OFDM symbol corresponding to the FFT size
N_{f}	Number of OFDM symbols between two TS preambles
n_{k}	refractive index of medium k
N_{m}	Number of spatial modes
$N_{\mathrm{PQPA/PA}}$	Number of pilot subcarriers for QPA and PA PN compensation
n_{PM}	Unperturbed refractive index of the waveguide in a phase modulator

N_{spans}	Number of spans in the optical link
n_{sp}	Spontaneous emission coefficient
N_{t}	Number of tributaries (total number of polarization and spatial modes)
N_{WDM}	Number of WDM channels
N_{bits}	Number of bits per symbol
P_{in}	Fiber input power
P_{out}	Fiber output power
$P_{\text{pho}_{\text{in}}}$	Input power to a photodiode in [W]
$P_{\text{sig,out}}$	Optical transmitted signal power at the end of the link
$R(r)$	Radial component of the transverse electromagnetic field
R_{DAC}	DAC sampling rate
R_{pho}	Responsivity of a photodiode in [A/W]
$r_i[k]$	Received signal after quantization and sampling with i ranging from one to N_{t}
r_{33}	Electro-optic coefficient of a phase modulator ($r_{33} = 30.9$ pm/V for *LiNbO*$_3$)
R_k	Received pilot symbol after equalization for pilot-aided PN compensation, where $k \in \mathbb{K}$
$S_{\text{laser}}(f)$	Laser power spectral density
$S_{\text{lo},lm}(\lambda)$	Dispersion slope related to mode lm in $[\text{ps} \cdot \text{nm}^{-2} \cdot \text{km}^{-1}]$
s_i	Modulated symbol on an OFDM subcarrier
S_k	Transmitted pilot symbol for QPA and PA PN compensation, where $k \in \mathbb{K}_{\text{QPA/PA}}$
$S_{lm}(z, \omega)$	Data signal in frequency domain for the mode LP$_{lm}$
T	OFDM symbol duration
t	Time variable in [s]
T_{CP}	Duration of the CP in [s]
t_{s}	Starting time of the OFDM symbol
u	Mode belonging to the set of modes $\mathbb{U}$
$u1$	Mode from where a pulse is launched to study modal crosstalk
$u_{\text{BIAS}_{\text{I/Q}}}$	DC part of the IQ modulator electrical inputs
$u_{\text{I/Q}}(t)$	Electrical drive voltages of an IQ modulator for the I and Q branches
u_{PM}	Electrical drive voltage of an IQ modulator for the phase shifter (PM)

$u_{\mathrm{RF}}(t)$	Comb electrical drive signal for one of the IQ modulator branches
$u_{1/2}(t)$	Driving voltage 1 or 2 of a phase modulator
V	Normalized frequency
V_{c}	Cutoff normalized frequency
v_{gr}	Group velocity
V_{π}	Drive voltage for a phase shift of π in a phase modulator
$v_{\mathrm{gr},lm}$	Group velocity of mode LP_{lm}
W_n	n^{th} energy level
x_{EM}	Output symbol of the EM stage
XT	Modal crosstalk
z_k	Transmitted symbol belonging to the k-th cluster in EM algorithm context
$P_{\mathrm{sig,in}_{\mathrm{dBm}}}$	Signal power at the input of the amplifier in [dBm]
$P_{\mathrm{sig,out}_{\mathrm{dBm}}}$	Optical transmitted signal power at the end of the link in [dBm]
NA	Numerical aperture
SNR	Signal-to-noise ratio

Bibliography

[1] G. P. Agrawal, *Fiber-optic communication systems.* Hoboken, NJ: John Wiley & Sons, Inc., 2010.

[2] N. Grote and H. Venghaus, *Fiber optic communication devices.* Berlin, Germany: Springer-Verlag Berlin Heidelberg, 2001.

[3] P. Trischitta, M. Colas, M. Green, G. Wuzniak, and J. Arena, “The TAT-12/13 cable network,” *IEEE Commun. Mag.*, vol. 34, no. 2, pp. 24–28, 1996.

[4] W. C. Barnett, H. Takahira, J. C. Baroni, and Y. Ogi, “The TPC-5 cable network,” *IEEE Commun. Mag.*, vol. 34, no. 2, pp. 36–40, 1996.

[5] R. C. Alferness, P. A. Bonenfant, C. J. Newton, K. A. Sparks, and E. L. Varma, “A practical vision for optical transport networking,” *Bell Labs Technical Journal*, vol. 4, no. 1, pp. 3–18, 1999.

[6] R. Essiambre and R. W. Tkach, “Capacity trends and limits of optical communication networks,” *Proc. IEEE*, vol. 100, no. 5, pp. 1035–1055, 2012.

[7] H. Kolimbiris, *Fiber optics communications.* Upper Saddle River, NJ: Pearson Prentice Hall, 2004.

[8] I. P. Kaminow and L. Tingye, *Optical fiber telecommunications IV A: Components.* San Diego, CA: Academic Press, 2002.

[9] ——, *Optical fiber telecommunications IV B: Systems and impairments.* San Diego, CA: Academic Press, 2002.

[10] D. v. d. Borne, “Robust optical transmission systems,” Ph.D. dissertation, Technische Universiteit Eindhoven, Eindhoven, 2008.

[11] TeleGeography. (2014) Executive summary. [Online]. Available: http://www.telegeography.com/page_attachments/products/website/research-services/global-internet-geography/0005/1382/GIG_Executive_Summary.pdf

[12] [Online]. Available: https://www.netflix.com

[13] [Online]. Available: www.youtube.com

[14] [Online]. Available: https://www.facebook.com

[15] R. W. Tkach, "Scaling optical communications for the next decade and beyond," *Bell Labs Technical Journal*, vol. 14, no. 4, pp. 3–9, 2010.

[16] P. J. Winzer, "Optical networking beyond WDM," *IEEE Photon. J.*, vol. 4, no. 2, pp. 647–651, 2012.

[17] ——, "Making spatial multiplexing a reality," *Nature Photonics*, vol. 8, no. 5, pp. 345–348, 2014.

[18] G. Bosco, V. Curri, A. Carena, P. Poggiolini, and F. Forghieri, "On the performance of Nyquist-WDM terabit superchannels based on PM-BPSK, PM-QPSK, PM-8QAM or PM-16QAM subcarriers," *J. Lightw. Technol.*, vol. 29, no. 1, pp. 53–61, 2011.

[19] P. J. Winzer and G. J. Foschini, "MIMO capacities and outage probabilities in spatially multiplexed optical transport systems," *Opt. Express*, vol. 19, no. 17, pp. 16 680–16 696, 2011.

[20] P. M. Krummrich, "Optical amplification and optical filter based signal processing for cost and energy efficient spatial multiplexing," *Opt. Express*, vol. 19, no. 17, pp. 16 636–16 652, 2011.

[21] X. Chen, A. Li, Y. Jia, A. Al Amin, and W. Shieh, "Demonstration of few-mode compatible optical add/drop multiplexer for mode-division multiplexed superchannel," *J. Lightw. Technol.*, vol. 31, no. 14, pp. 641–647, 2013.

[22] F. Ferreira, S. Jansen, P. Monteiro, and H. Silva, "Nonlinear semi-analytical model for simulation of few-mode fiber transmission," *IEEE Photon. Technol. Lett.*, vol. 24, no. 4, pp. 240–242, 2012.

[23] F. Yaman, N. Bai, B. Zhu, T. Wang, and G. Li, "Long distance transmission in few-mode fibers," *Opt. Express*, vol. 18, no. 12, pp. 13 250–13 257, 2010.

[24] I. T. S. Sector. (2009) Optical fibres, cables and systems. [Online]. Available: http://www.itu.int/pub/T-HDB-OUT.10-2009-1

[25] G. P. Agrawal, *Nonlinear fiber optics.* San Diego, CA: Academic Press, 2001.

[26] J. M. Senior, *Optical fiber communications: Principles and practice.* Harlow, UK: Pentice hall, 1992.

[27] D. Gloge, "Weakly guiding fibers," *Applied Optics*, vol. 10, no. 10, pp. 2252–2258, 1971.

[28] A. Ghatak and K. Thyagarajan, *An introduction to fiber optics.* New York, NY: Cambridge university press, 1998.

[29] S. Bottacchi, *Multi-gigabit transmission over multimode optical fibre: theory and design methods for 10GbE systems.* The Atrium, UK: John Wiley & Sons Ltd., 2006.

[30] J. A. Buck, *Fundamentals of optical fibers.* Hoboken, NJ: John Wiley & Sons Ldt., 2004.

[31] J. H. Franz and V. K. Jain, *Optical communications: Components and systems.* Oxford, UK: Alpha science, 2008.

[32] J. Sakaguchi, W. Klaus, B. J. Puttnam, J. M. Mendinueta, Y. Awaji, N. Wada, Y. Tsuchida, K. Maeda, M. Tadakuma, K. Imamura *et al.*, "19-core MCF transmission system using EDFA with shared core pumping coupled via free-space optics," *Opt. Express*, vol. 22, no. 1, pp. 90–95, 2014.

[33] E. Voges and K. Petermann, *Optische kommunikationstechnik.* Berlin, Germany: Springer-Verlag Berlin Heidelberg, 2002.

[34] G. P. Agrawal, *Lightwave technology.* Hoboken, NJ: John Wiley & Sons, Inc., 2005.

[35] I. H. Malitson, "Interspecimen comparison of the refractive index of fused silica," *J. Opt. Soc. Am.*, vol. 55, no. 10, pp. 1205–1208, 1965.

[36] F. Ferreira, "High capacity optical transmission systems based on mode division multiplexing," Ph.D. dissertation, Universidade de Coimbra, 2014.

[37] F. Ferreira, D. Fonseca, and H. Silva, "Design of few-mode fibers with arbitrary and flattened differential mode delay," *IEEE Photon. Technol. Lett.*, vol. 25, no. 5, pp. 438–441, 2013.

[38] M. B. Shemirani and J. M. Kahn, "Higher-order modal dispersion in graded-index multimode fiber," *J. Lightw. Technol.*, vol. 27, no. 23, pp. 5461–5468, 2009.

[39] I. P. Kaminow and T. L. Koch, *Optical fiber telecommunications IV B: Systems and impairments.* San Diego, CA: Academic Press, 1997.

[40] A. Galtarossa and C. R. Menyuk, *Polarization mode dispersion.* New York, NY: Springer Science & Business Media, 2005.

[41] M. Ikeda, Y. Murakami, and K. Kitayama, "Mode scrambler for optical fibers," *Applied optics*, vol. 16, no. 4, pp. 1045–1049, 1977.

[42] M. Tokuda, S. Seikai, K. Yoshida, and N. Uchida, "Measurement of baseband frequency response of multimode fibre by using a new type of mode scrambler," *Electronic letters*, vol. 13, no. 5, pp. 146–147, 1977.

[43] I. P. Kaminow, L. Tingye, and A. E. Willner, *Optical fiber telecommunications VI A: Components and subsystems.* San Diego, CA: Academic Press, 2013.

[44] D. Marcuse, *Theory of dielectric optical waveguides.* New York, NY: Academic Press, 1974.

[45] R. Olshansky, "Mode coupling effects in graded-index optical fibers," *Applied optics*, vol. 14, no. 4, pp. 935–945, 1975.

[46] I. P. Kaminow, L. Tingye, and A. E. Willner, *Optical fiber telecommunications VI B: Systems and networks.* Oxford, UK: Academic Press, 2013.

[47] F. Ferreira, P. Monteiro, and H. Silva, "Semi-analytical model for linear modal coupling in few-mode fiber transmission," in *Proc. International Conference on Transparent Optical Networks (ICTON)*, 2012.

[48] A. Yariv and P. Yeh, *Optical electronics in modern communications.* New York, NY: Oxford university press, 2007.

[49] J. K. Petersen, *Fiber optics illustrated dictionary.* Boca Raton, FL: CRC Press, 2003.

[50] C. Headley and G. Agrawal, *Raman amplification in fiber optical communication systems.* Burlington, MA: Elsevier Academic Press, 2005.

[51] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaeelpour, E. C. Burrows, R. Essiambre, P. J. Winzer *et al.*, "Mode-division multiplexing over 96 km of few-mode fiber using coherent 6 x 6 MIMO processing," *J. Lightw. Technol.*, vol. 30, no. 4, pp. 521–531, 2012.

[52] R. Ryf, N. K. Fontaine, B. Guan, R. Essiambre, S. Randel, A. H. Gnauck, S. Chandrasekhar, A. Adamiecki, G. Raybon, B. Ercan *et al.*, "1705-km transmission over coupled-core fibre supporting 6 spatial modes," in *Proc. European Conference on Optical Communication (ECOC)*, 2014.

[53] R. G. H. Van Uden, R. Amezcua Correa, E. Antonio Lopez, F. M. Huijskens, C. Xia, G. Li, A. Schülzgen, H. de Waardt, A. M. J. Koonen, and C. M. Okonkwo, "Ultra-high-density spatial division multiplexing with a few-mode multicore fibre," *Nature Photonics*, vol. 8, no. 11, pp. 865–870, 2014.

[54] C. Cole, “Beyond 100G client optics,” *IEEE Communications magazine*, vol. 50, no. 2, pp. S58–S66, 2012.

[55] P. Krummrich, “Optical amplification for capacity increase in long haul transmission systems by spatial multiplexing,” in *Proc. IEEE Photonic Society 24th Annual Meeting*, 2011.

[56] P. M. Krummrich and K. Petermann, “Evaluation of potential optical amplifier concepts for coherent mode multiplexing,” in *Proc. Optical Fiber Communication Conference (OFC)*, 2011.

[57] B. Inan, B. Spinnler, F. Ferreira, D. van den Borne, A. Lobato, S. Adhikari, V. Sleiffer, M. Kuschnerov, N. Hanik, and S. L. Jansen, “DSP complexity of mode-division multiplexed receivers,” *Opt. Express*, vol. 20, no. 10, pp. 10 859–10 869, 2012.

[58] H. Chen, N. K. Fontaine, R. Ryf, B. Huang, A. Velzquez-Bentez, C. Jin, B. Ercan, and D. T. Neilson, “Wavelength selective switch for commercial multimode fiber supporting 576 spatial modes,” in *Proc. European Conference on Optical Communication (ECOC)*, 2016.

[59] M. Seimetz, *High-order modulation for optical fiber transmission.* Berlin, Germany: Springer-Verlag Berlin Heidelberg, 2009.

[60] K. Ho, *Phase-modulated optical communication systems.* New York, NY: Springer Science & Business Media, 2005.

[61] N. Hanik, “Lecture notes for optical communication systems,” Technische Universität München, 2013.

[62] W. Shieh and I. Djordjevic, *OFDM for optical communications.* Burlington, MA: Academic Press, 2009.

[63] S. Michel, J. Zyss, I. Ledoux-Rak, and C. T. Nguyen, “High-performance electro-optic modulators realized with a commercial side-chain DR1-PMMA electro-optic copolymer,” in *Proc. SPIE Organic Photonic Materials and Devices XII*, vol. 7599, 2010, pp. 759 901–1 – 759 901–14.

[64] K.-P. Ho and J. M. Kahn, “Mode-dependent loss and gain: statistics and effect on mode-division multiplexing,” *Opt. Express*, vol. 19, no. 17, pp. 16 612–16 635, 2011.

[65] M. Salsi, R. Ryf, G. Le Cocq, L. Bigo, D. Peyrot, G. Charlet, N. K. Fontaine, M. A. Mestre, S. Randel, X. Palou, C. Bolle *et al.*, “A six-mode erbium-

doped fiber amplifier," in *Proc. European Conference on Optical Communication (ECOC)*, 2012.

[66] Y. Jung, Q. Kang, J. K. Sahu, B. Corbett, R. Winfield, F. Poletti, S. U. Alam, and D. J. Richardson, "Few-mode EDFA supporting 5 spatial modes with reconfigurable differential modal gain control," in *Proc. European Conference on Optical Communication (ECOC)*, 2013.

[67] Y. Jung, Q. Kang, V. A. J. M. Sleiffer, B. Inan, M. Kuschnerov, V. Veljanovski, B. Corbett, R. Winfield, Z. Li, P. S. Teh *et al.*, "Three mode Er^{3+} ring-doped fiber amplifier for mode-division multiplexed transmission," *Opt. Express*, vol. 21, no. 8, pp. 10 383–10 392, 2013.

[68] H. Chen, N. K. Fontaine, R. Ryf, B. Guan, S. J. B. Yoo, and A. M. J. Koonen, "A fully-packaged 3D-waveguide based dual-fiber spatial-multiplexer with up-tapered 6-mode fiber pigtails," in *Proc. European Conference on Optical Communication (ECOC)*, 2014.

[69] V. A. J. M. Sleiffer, Y. Jung, V. Veljanovski, R. G. H. Van Uden, M. Kuschnerov, H. Chen, B. Inan, L. Grüner Nielsen, Y. Sun, D. J. Richardson *et al.*, "73.7 Tb/s (96 x 3 x 256-Gb/s) mode-division-multiplexed DP-16QAM transmission with inline MM-EDFA," *Opt. Express*, vol. 20, no. 26, pp. B428–B438, 2012.

[70] F. Ferreira, D. Fonseca, A. Lobato, B. Inan, and H. Silva, "Reach improvement of mode division multiplexed systems using fiber splices," *IEEE Photon. Technol. Lett.*, vol. 25, no. 12, pp. 1091–1094, 2013.

[71] D. Tse and V. Pramod, *Fundamentals of wireless communication.* New York, NY: Cambridge university press, 2005.

[72] L. Trefethen and D. Bau, *Numerical linear algebra.* Philadelphia, PA: Society for industrial and applied mathematics, 1997.

[73] W. Mohammed, M. Pitchumani, A. Mehta, and E. G. Johnson, "Selective excitation of the LP_{11} mode in step index fiber using a phase mask," *Optical Engineering*, vol. 45, no. 7, pp. 074 602–1–074 602–7, 2006.

[74] W. Q. Thornburg, B. J. Corrado, and X. D. Zhu, "Selective launching of higher-order modes into an optical fiber with an optical phase shifter," *Optics letters*, vol. 19, no. 7, pp. 454–456, 1994.

[75] E. Ip, M. Li, K. Bennett, S. Bickham, Y. Huang, A. Tanaka, E. Mateo, J. Hu, T. Wang, A. Korolev *et al.*, "Few-mode fiber transmission with inline few-

mode erbium-doped fiber amplifier," in *Proc. SPIE 8647, Next-Generation Optical Communication: Components, Sub-Systems, and Systems II*, 2013, pp. 864 709–1–864 709–10.

[76] J. W. Goodman, *Introduction to Fourier optics.* USA: McGraw-Hill, Inc., 1968.

[77] M. Salsi, C. Koebele, D. Sperti, P. Tran, P. Brindel, H. Mardoyan, S. Bigo, A. Boutin, F. Verluise, P. Sillard *et al.*, "Transmission at 2x100Gb/s, over two modes of 40km-long prototype few-mode fiber, using LCOS-based mode multiplexer and demultiplexer," in *Proc. Optical Fiber Communication Conference (OFC)*, 2011.

[78] J. Carpenter, B. C. Thomsen, and T. D. Wilkinson, "Degenerate mode-group division multiplexing," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3946–3952, 2012.

[79] I. Giles, A. Obeysekara, R. Chen, D. Giles, F. Poletti, and D. Richardson, "Fiber LPG mode converters and mode selection technique for multimode SDM," *IEEE Photon. Technol. Lett.*, vol. 24, no. 21, pp. 1922–1925, 2012.

[80] N. Hanzawa, K. Saitoh, T. Sakamoto, T. Matsui, S. Tomita, and M. Koshiba, "Demonstration of mode-division multiplexing transmission over 10 km two-mode fiber with mode coupler," in *Proc. Optical Fiber Communication Conference (OFC)*, 2011.

[81] A. Li, A. Al Amin, X. Chen, and W. Shieh, "Reception of mode and polarization multiplexed 107-Gb/s CO-OFDM signal over a two-mode fiber," in *Proc. Optical Fiber Communication Conference (OFC)*, 2011.

[82] R. Ryf, M. A. Mestre, A. H. Gnauck, S. Randel, C. Schmidt, R.-J. Essiambre, P. J. Winzer, R. Delbue, P. Pupalaikis, A. Sureka *et al.*, "Low-loss mode coupler for mode-multiplexed transmission in few-mode fiber," in *Proc. Optical Fiber Communication Conference (OFC)*, 2012.

[83] R. Ryf, N. K. Fontaine, and R.-J. Essiambre, "Spot-based mode couplers for mode-multiplexed transmission in few-mode fiber," *IEEE Photon. Technol. Lett.*, vol. 24, no. 21, pp. 1973–1976, 2012.

[84] H. Chen, V. Sleiffer, R. van Uden, C. Okonkwo, M. Kuschnerov, F. Huijskens, L. Grüner-Nielsen, Y. Sun, H. de Waardt, and T. Koonen, "3 MDM×8 WDM×320-Gb/s DP-32QAM transmission over a 120km few-mode fiber span employing 3-spot mode couplers," in *Proc. OptoElectronics and Communications Conference (OECC)*, 2013.

[85] S. Leon-Saval, A. Argyros, and J. Bland-Hawthorn, “Photonic lanterns: a study of light propagation in multimode to single-mode converters,” *Opt. Express*, vol. 18, no. 8, pp. 8430–8439, 2010.

[86] N. K. Fontaine, “Photonic lantern spatial multiplexers in space-division multiplexing,” in *Proc. IEEE Photonics Society Summer Topical Meeting Series*, 2013.

[87] N. K. Fontaine, R. Ryf, J. Bland-Hawthorn, and S. G. Leon-Saval, “Geometric requirements for photonic lanterns in space division multiplexing,” *Opt. Express*, vol. 20, no. 24, pp. 27 123–27 132, 2012.

[88] N. K. Fontaine and R. Ryf, “Characterization of mode-dependent loss of laser inscribed photonic lanterns for space division multiplexing systems,” in *Proc. OptoElectronics and Communications Conference (OECC)*, 2013.

[89] L. Kazovsky, S. Benedetto, and A. Willner, *Optical fiber communication systems.* Norwood, MA: Artech house, 1996.

[90] S. Pachnicke, *Fiber-optic transmission networks:Efficient design and dynamic operation.* Berlin, Germany: Springer-Verlag Berlin Heidelberg, 2012.

[91] P. M. Becker, N. A. Olsson, and J. R. Simpson, *Erbium-doped fiber amplifiers: fundamentals and technology.* San Diego, CA: Academic Press, 1999.

[92] E. Ip, “Gain equalization for few-mode fiber amplifiers beyond two propagating mode groups,” *IEEE Photon. Technol. Lett.*, vol. 24, no. 21, p. 1933, 2012.

[93] N. Bai, E. Ip, T. Wang, and G. Li, “Multimode fiber amplifier with tunable modal gain using a reconfigurable multimode pump,” *Opt. Express*, vol. 19, no. 17, pp. 16 601–16 611, 2011.

[94] Q. Kang, E. Lim, Y. Jung, J. K. Sahu, F. Poletti, S.-U. Alam, and D. J. Richardson, “Modal gain control in a multimode erbium doped fiber amplifier incorporating ring doping,” in *Proc. European Conference on Optical Communication (ECOC)*, 2012.

[95] G. Le Cocq, L. Bigot, A. Le Rouge, M. Bigot-Astruc, P. Sillard, C. Koebele, M. Salsi, and Y. Quiquempois, “Modeling and characterization of a few-mode edfa supporting four mode groups for mode division multiplexing,” *Opt. Express*, vol. 20, no. 24, pp. 27 051–27 061, 2012.

[96] Q. Kang, E. Lim, Y. Jung, F. Poletti, S.-U. Alam, and D. J. Richardson, “Design of four-mode erbium doped fiber amplifier with low differential modal

gain for modal division multiplexed transmissions," in *Proc. Optical Fiber Communication Conference (OFC)*, 2013.

[97] Y. Jung, S. Alam, Z. Li, A. Dhar, D. Giles, G. I. P., J. K. Sahu, F. Poletti, L. Grüner-Nielsen, and D. J. Richardson, "First demonstration and detailed characterization of a multimode amplifier for space division multiplexing," *Opt. Express*, vol. 19, no. 26, pp. B952–B957, 2011.

[98] E. Ip, M.-J. Li, K. Bennett, K. Korolev, W. Wood, C. Montero, and J. Liñares, "Experimental characterization of a ring-profile few-mode erbium-doped fiber amplifier enabling gain equalization," in *Proc. Optical Fiber Communication Conference (OFC)*, 2013.

[99] Y. Jung, Q. Kang, J. Kumar Sahu, B. Corbett, J. O'Callagham, F. Poletti, S.-U. Alam, and D. J. Richardson, "Reconfigurable modal gain control of a few-mode EDFA supporting six spatial modes," *IEEE Photon. Technol. Lett.*, vol. 26, no. 11, pp. 1100–1103, 2014.

[100] E. Ip, M.-J. Li, R. Y. Gu, and J. Kahn, "Multimode transmission technologies," in *Proc. OptoElectronics and Communications Conference (OECC)*, 2014.

[101] R. Ryf, R.-J. Essiambre, J. von Hoyningen-Huene, and P. J. Winzer, "Anaysis of mode-dependent gain in Raman amplified few-mode fiber," in *Proc. Optical Fiber Communication Conference (OFC)*, 2012.

[102] R. Ryf, A. Sierra, R.-J. Essiambre, S. Randel, A. H. Gnauck, C. Bolle, M. Esmaeelpour, P. J. Winzer, R. Delbue, P. Pupalaikise, A. Sureka, D. W. Peckham, A. McCurdy, and R. J. Lingle, "Mode-equalized distributed Raman amplification in 137-km few-mode fiber," in *Proc. European Conference on Optical Communication (ECOC)*, 2011.

[103] JDSU. (2009) A performance comparison of wss switch engine technologies. [Online]. Available: http://www.jdsu.com/productliterature/wsscomp_wp_cms_ae.pdf

[104] Finisar. (2012) Balancing performance, flexibility, and scalability in optical networks. [Online]. Available: https://www.finisar.com/sites/default/files/resources/balancing_performance_flexibility_scalability_in_optical_networks_wss_whitepaper_march2012.pdf

[105] R. Ryf, N. K. Fontaine, J. Dunayevsky, D. Sinefeld, M. Blau, M. Montoliu, S. Randel, C. Liu, B. Ercan, M. Esmaeelpour *et al.*, "Wavelength-selective

switch for few-mode fiber transmission," in *Proc. European Conference on Optical Communication (ECOC)*, 2013.

[106] J. Carpenter, S. G. Leon-Saval, J. R. Salazar-Gil, J. Bland-Hawthorn, G. Baxter, L. Stewart, S. Frisken, M. A. F. Roelens, B. J. Eggleton, and J. Schröder, "1x11 few-mode fiber wavelength selective switch using photonic lanterns," *Opt. Express*, vol. 22, no. 3, pp. 2216–2221, 2014.

[107] J. Carpenter, S. Leon-Saval, B. J. Eggleton, and J. Schroder, "Spatial light modulators for sub-systems and characterization in SDM," in *Proc. OptoElectronics and Communications Conference (OECC)*, 2014.

[108] N. K. Fontaine, R. Ryf, C. Liu, B. Ercan, J. R. Salazar Gil, S. G. Leon-Saval, J. Bland-Hawthorn, D. T. Neilson *et al.*, "Few-mode fiber wavelength selective switch with spatial-diversity and reduced-steering angle," in *Proc. Optical Fiber Communication Conference (OFC)*, 2014.

[109] M. C. Jeruchim, P. Balaban, and K. S. Shanmungan, *Simulation of communication systems.* New York, NY: Plenum press, 1992.

[110] B. Spinnler, "Equalizer design and complexity for digital coherent receivers," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1180–1192, 2010.

[111] M. Kuschnerov, F. N. Hauske, K. Piyawanno, B. Spinnler, M. S. Alfiad, A. Napoli, and B. Lankl, "DSP for coherent single-carrier receivers," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3614–3622, 2009.

[112] N. Bai, C. Xia, and G. Li, "Adaptive frequency-domain equalization for the transmission of the fundamental mode in a few-mode fiber," *Opt. Express*, vol. 20, no. 21, pp. 24 010–24 017, 2012.

[113] T. Rahman, D. Rafique, B. Spinnler, S. Calabro, E. De Man, U. Feiste, A. Napoli, M. Bohn, G. Khanna, N. Hanik *et al.*, "Long-haul transmission of PM-16QAM-, PM-32QAM-, and PM-64QAM-based terabit superchannels over a field deployed legacy fiber," *J. Lightw. Technol.*, vol. 34, no. 13, pp. 3071–3079, 2016.

[114] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, R.-J. Essiambre, P. Winzer, D. W. Peckham, A. McCurdy, R. Lingle, and R. Lingle Jr., "Space-division multiplexing over 10 km of three-mode fiber using coherent 6× 6 MIMO processing," in *Proc. Optical Fiber Communication Conference (OFC)*, 2011.

[115] S. Randel, R. Ryf, A. Sierra, P. J. Winzer, A. H. Gnauck, C. A. Bolle, R.-J. Essiambre, D. W. Peckham, A. McCurdy, and R. Lingle Jr., "6× 56-Gb/s

mode-division multiplexed transmission over 33-km few-mode fiber enabled by 6× 6 MIMO equalization," *Opt. Express*, vol. 19, no. 17, pp. 16 697–16 707, 2011.

[116] E. Ip, M.-J. Li, K. Bennett, S. Bickham, Y.-K. Huang, A. Tanaka, E. Mateo, J. Hu, T. Wang, A. Korolev, K. Koreshkov *et al.*, "6× 28-Gbaud few-mode recirculating loop transmission with gain-equalized inline few-mode fiber amplifier," in *Proc. Optical Fiber Communication Conference (OFC)*, 2013.

[117] Y. Chen, A. Lobato, Y. Jung, H. Chen, V. A. J. M. Sleiffer, M. Kuschnerov, N. K. Fontaine, R. Ryf, D. J. Richardson, B. Lankl, and N. Hanik, "41.6 Tbit/s C-Band SDM OFDM transmission through 12 spatial and polarization modes over 74.17 km few mode fiber," *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1440–1444, 2015.

[118] R. Ryf, S. Randel, N. K. Fontaine, X. Palou, E. Burrows, S. Corteselli, S. Chandrasekhar, A. H. Gnauck, C. Xie, R.-J. Essiambre, P. J. Winzer, R. Delbue, P. Pupalaikis, A. Sureka, Y. Sun, L. Grüner-Nielsen, R. Jensen, and R. Lingle, "708-km combined WDM/SDM transmission over few-mode fiber supporting 12 spatial and polarization modes," in *Proc. European Conference on Optical Communication (ECOC)*, 2013.

[119] S. O. Arik, D. Askarov, and J. M. Kahn, "Effect of mode coupling on signal processing complexity in mode-division multiplexing," *J. Lightw. Technol.*, vol. 31, no. 3, pp. 423–431, 2013.

[120] L. Grüner-Nielsen, Y. Sun, R. V. S. Jensen, J. W. Nicholson, and R. Lingle, "Recent advances in low dgd few-mode fibre design, fabrication, characterization and experiments," in *Proc. Optical Fiber Communication Conference (OFC)*, 2015.

[121] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell Labs Technical Journal*, vol. 45, no. 10, pp. 1775–1796, 1966.

[122] M. Engels, *Wireless OFDM Systems: How to make them work?* Norwell, MA: Kluwer Academic Publishers, 2002.

[123] R. van Nee and R. Prasad, *OFDM for wireless multimedia communications.* USA: Artech House, 2000.

[124] S. L. Jansen, I. Morita, T. C. W. Schenk, N. Takeda, and H. Tanaka, "Coherent optical 25.8-Gb/s OFDM transmission over 4160-km SSMF," *J. Lightw. Technol.*, vol. 26, no. 1, pp. 6–15, 2008.

[125] S. B. Weinstein and P. M. Ebert, “Data transmission by frequency-division multiplexing using the discrete Fourier transform,” *IEEE Trans. Commun. Technol.*, vol. 19, no. 5, pp. 628–634, 1971.

[126] T. M. Schmidl and D. C. Cox, “Robust frequency and timing synchronization for OFDM,” *IEEE Trans. Commun.*, vol. 45, no. 12, pp. 1613–1621, 1997.

[127] T. Schenk, *RF imperfections in high-rate wireless systems: impact and digital compensation.* Dordrecht, The Netherlands: Springer Science & Business Media, 2008.

[128] J. Armstrong, “OFDM for optical communications,” *J. Lightw. Technol.*, vol. 27, no. 3, pp. 189–204, 2009.

[129] B. Inan, Y. Jung, V. Sleiffer, M. Kuschnerov, L. Grüner-Nielsen, S. Adhikari, S. L. Jansen, D. J. Richardson, S.-U. Alam, B. Spinnler, and N. Hanik, “Low computational complexity mode division multiplexed OFDM transmission over 130 km of few mode fiber,” in *Proc. Optical Fiber Communication Conference (OFC)*, 2013.

[130] A. Lobato, Y. Chen, Y. Jung, H. Chen, B. Inan, M. Kuschnerov, N. K. Fontaine, R. Ryf, B. Spinnler, and Lankl, “12-mode OFDM transmission using reduced-complexity maximum likelihood detection,” *Optics letters*, vol. 40, no. 3, pp. 328–331, 2015.

[131] A. Lobato, M. Kuschnerov, A. Diaz, A. Napoli, B. Spinnler, and B. Lankl, “Performance comparison of single carrier and OFDM in coherent optical long-haul communication systems,” in *Proc. Communications and Photonics Conference and Exhibition (ACP)*, 2011.

[132] K. Forozesh, S. L. Jansen, S. Randel, I. Morita, and H. Tanaka, “The influence of the dispersion map in coherent optical ofdm transmission systems,” in *Proc. IEEE/LEOS Summer Topical Meetings*, 2008.

[133] S. T. Le, T. Kanesan, M. E. McCarthy, E. Giacoumidis, I. D. Phillips, M. F. C. Stephens, M. Tan, N. J. Doran, A. D. Ellis, and S. K. Turitsyn, “Experimental demonstration of data-dependent pilot-aided phase noise estimation for CO-OFDM,” in *Proc. Optical Fiber Communication Conference (OFC)*, 2014.

[134] I.-T. G.975. (2000) Forward error correction for submarine systems. [Online]. Available: https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-G.975-200010-I!!PDF-E&type=items

[135] I. Djordjevic, W. Ryan, and B. Vasic, *Coding for optical channels.* New York, NY: Springer Science & Business Media, 2010.

[136] I.-T. G.975.1. (2004) Forward error correction for high bit-rate DWDM submarine systems. [Online]. Available: https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-G.975.1-200402-I!!PDF-E&type=items

[137] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE Int. Conf. Commun.*, 1993.

[138] T. Mizuochi, Y. Miyata, T. Kobayashi, K. Ouchi, K. Kuno, K. Kubo, K. Shimizu, H. Tagami, H. Yoshida, H. Fujita *et al.*, "Forward error correction based on block turbo code with 3-bit soft decision for 10-gb/s optical communication systems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 10, no. 2, pp. 376–386, 2004.

[139] A. Leven and L. Schmalen, "Status and recent advances on forward error correction technologies for lightwave systems," *J. Lightw. Technol.*, vol. 32, no. 16, pp. 2735–2750, 2014.

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

[141] D. Morero, M. A. Castrillon, F. Ramos, T. Goette, O. E. Agazzi, and M. R. Hueda, "Non-concatenated FEC codes for ultra-high speed optical transport networks," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, 2011.

[142] K.-P. Ho and J. M. Kahn, "Frequency diversity in mode-division multiplexing systems," *J. Lightw. Technol.*, vol. 29, no. 24, pp. 3719–3726, 2011.

[143] A. Lobato, F. Ferreira, J. Rabe, M. Kuschnerov, B. Spinnler, and B. Lankl, "Mode scramblers and reduced-search maximum-likelihood detection for mode-dependent-loss-impaired transmission," in *Proc. European Conference on Optical Communication (ECOC)*, 2013.

[144] A. Lobato, F. Ferreira, M. Kuschnerov, D. van den Borne, S. Jansen, A. Napoli, B. Spinnler, and B. Lankl, "Impact of mode coupling on the mode-dependent loss tolerance in few-mode fiber transmission," *Opt. Express*, vol. 20, no. 28, pp. 29 776–29 783, 2012.

[145] A. Lobato, F. Ferreira, M. Kuschnerov, D. van den Borne, S. Jansen, B. Spinnler, and B. Lankl, "Impact of mode coupling on the mode-dependent

loss tolerance in few-mode fiber transmission," in *Proc. European Conference on Optical Communication (ECOC)*, 2012.

[146] A. Lobato, F. Ferreira, B. Inan, S. Adhikari, M. Kuschnerov, A. Napoli, B. Spinnler, and B. Lankl, "Maximum-likelihood detection in few-mode fiber transmission with mode-dependent loss," *IEEE Photon. Technol. Lett.*, vol. 25, no. 12, pp. 1095–1098, 2013.

[147] A. Lobato, F. Ferreira, J. Rabe, B. Inan, S. Adhikari, M. Kuschnerov, A. Napoli, B. Spinnler, and B. Lankl, "Mode-dependent-loss mitigation for mode-division multiplexed systems," in *Proc. Signal Processing in Photonic Communications (SPPCom)*, 2013.

[148] ——, "On the mode-dependent loss compensation for mode-division multiplexed systems," in *Proc. International Conference on Transparent Optical Networks (ICTON)*, 2013.

[149] A. Lobato, F. Ferreira, J. Rabe, M. Kuschnerov, B. Spinnler, and B. Lankl, "Enhanced performance for MDL-impaired few-mode fiber transmission," in *Proc. OptoElectronics and Communications Conference (OECC)*, 2014.

[150] X. Chen, J. He, A. Li, J. Ye, and W. Shieh, "Characterization and analysis of few-mode fiber channel dynamics," *IEEE Photon. Technol. Lett.*, vol. 25, no. 18, pp. 1819–1822, 2013.

[151] W. Shieh and C. Athaudage, "Coherent optical orthogonal frequency division multiplexing," *Electronics Letters*, vol. 42, no. 10, pp. 587–589, 2006.

[152] G. P. Agrawal, *Applications of nonlinear fiber optics.* San Diego, CA: Academic Press, 2008.

[153] F. Ferreira, D. Fonseca, and H. Silva, "On the dependence of differential mode delay in few-mode fibers on the number of modes," in *Proc. International Conference on Transparent Optical Networks (ICTON)*, 2013.

[154] J. Noda, K. Okamoto, and Y. Sasaki, "Polarization-maintaining fibers and their applications," *J. Lightw. Technol.*, vol. 4, no. 8, pp. 1071–1089, 1986.

[155] L. Grüner-Nielsen, Y. Sun, J. W. Nicholson, D. Jakobsen, K. G. Jespersen, R. J. Lingle, and B. Pálsdóttir, "Few mode transmission fiber with low DGD, low mode coupling, and low loss," *J. Lightw. Technol.*, vol. 30, no. 23, pp. 3693–3698, 2012.

[156] K. Harris. (2009) Intro to probability from the Department of Mathematics, University of Michigan. [Online]. Available: http://kaharris.org/teaching/425/Lectures/lec30.pdf

[157] N. Benvenuto and N. Cherubini, *Algorithms for communications systems and their applications.* Atrium, England: John Wiley & Sons Ltd, 2002.

[158] A. J. Paulraj, D. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications-a key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, 2004.

[159] W. Shieh, "PMD-supported coherent optical OFDM systems," *IEEE Photon. Technol. Lett.*, vol. 19, no. 3, pp. 134–136, 2007.

[160] M. Kuschnerov, M. Chouayakh, K. Piyawanno, B. Spinnler, M. S. Alfiad, A. Napoli, and B. Lankl, "On the performance of coherent systems in the presence of polarization-dependent loss for linear and maximum likelihood receivers," *IEEE Photon. Technol. Lett.*, vol. 22, no. 12, pp. 920–922, 2010.

[161] X. Chen, J. Ye, Y. Xiao, A. Li, J. He, Q. Hu, and W. Shieh, "Equalization of two-mode fiber based MIMO signals with larger receiver sets," *Opt. Express*, vol. 20, no. 26, pp. B413–B418, 2012.

[162] C. Koebele, M. Salsi, L. Milord, R. Ryf, C. Bolle, P. Sillard, S. Bigo, Charlet, and G. Charlet, "40km transmission of five mode division multiplexed data streams at 100Gb/s with low MIMO-DSP complexity," in *Proc. European Conference on Optical Communication (ECOC)*, 2011.

[163] T. Mori, t. Sakamoto, M. Wada, T. Yamamoto, and F. Yamamoto, "Low DMD four LP mode transmission fiber for wide-band WDM-MIMO system," in *Proc. Optical Fiber Communication Conference (OFC)*, 2013.

[164] L. Grüner-Nielsen, Y. Sun, R. V. Jensen, J. W. Nicholson, and R. Lingle, "Splicing of few mode fibers," in *Proc. European Conference on Optical Communication (ECOC)*, 2014.

[165] L. Palmieri and A. Galtarossa, "A preliminary analysis of spin in few-mode optical fibers," in *Proc. European Conference on Optical Communication (ECOC)*, 2013.

[166] J. Fang, A. Li, and W. Shieh, "Low-DMD few-mode fiber with distributed long-period grating," *Optics letters*, vol. 40, no. 17, pp. 3937–3940, 2015.

[167] A. G. Hallam, D. A. Robinson, and I. Bennion, “Mode control for emerging link performance standards,” *IET Optoelectronics*, vol. 2, no. 5, pp. 175–181, 2008.

[168] R.-O. G. . C. OHG. (2009) Ef encircled flux multimode anregungsbedingungen. [Online]. Available: https://www.rosenberger-osi.com/fileadmin/pdf_downloads/Whitepaper/Whitepaper-EF_Multimode_Anregungsbedingungen_DE.pdf

[169] I. Cisco Systems. (2008) Limitations of transmission distances over multimode fiber. [Online]. Available: https://www.cisco.com/c/en/us/products/collateral/interfaces-modules/transceiver-modules/white_paper_c11-463677.html

[170] L. Su, K. S. Chiang, and C. Lu, “Microbend-induced mode coupling in a graded-index multimode fiber,” *Applied optics*, vol. 44, no. 34, pp. 7394–7402, 2005.

[171] A. Lobato, J. Rabe, F. Ferreira, M. Kuschnerov, B. Spinnler, and B. Lankl, “Near-ML detection for MDL-impaired few-mode fiber transmission,” *Opt. Express*, vol. 23, no. 8, pp. 9589–9601, 2015.

[172] M. Chouayakh, A. Knopp, and B. Lankl, “Fixed effort MIMO decoders for wireless indoor channels: Theory and practical field trials,” in *Proc. IEEE Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2008.

[173] M. Chouayakh, “Wireless indoor MIMO Systems: Theoretical concept and prectical implementation with focus on low effort data detection,” Ph.D. dissertation, Universitaet der Bundeswehr Muenchen, 2009.

[174] U. Fincke and M. Pohst, “Improved methods for calculating vectors of short length in a lattice, including a complexity analysis,” *Mathematics of computation*, vol. 44, no. 170, pp. 463–471, 1985.

[175] E. Viterbo and J. Boutros, “A universal lattice code decoder for fading channels,” *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1639–1642, 1999.

[176] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO wireless communications.* Cambridge, England: Cambridge University press, 2007.

[177] J.-S. Kim, S.-H. Moon, and I. Lee, “A new reduced complexity ML detection

scheme for MIMO systems," *IEEE Trans. Commun.*, vol. 58, no. 4, pp. 1302–1310, 2010.

[178] K. Honjo and T. Ohtsuki, "Computational complexity reduction of MLD based on SINR in MIMO spatial multiplexing systems," *IEICE Transactions on Communications*, vol. 89, no. 3, pp. 914–921, 2006.

[179] M. Chouayakh, A. Knopp, and B. Lankl, "Low-effort near maximum likelihood MIMO detection with optimum hardware resource exploitation," *Electronics letters*, vol. 43, no. 20, pp. 1104–1106, 2007.

[180] H. Orloff and J. Bloom. (2014) Bayesian updating with continuous priors, Michigan Technological University. [Online]. Available: https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/readings/MIT18_05S14_Reading13a.pdf

[181] J. Jaldén and B. Otterste, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1474–1484, 2005.

[182] N. Tax and B. Lankl, "Fixed effort sphere decoder for MIMO OFDM systems," in *Proc. International Conference on Wireless Communications & Signal Processing (WCSP)*, 2013.

[183] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bolcskei, "VLSI implementation of MIMO detection using the sphere decoding algorithm," *IEEE J. Solid-State Circuits*, vol. 40, pp. 1566–1577, 2005.

[184] C. P. Schnorr and M. Euchner, "Lattice basis reduction: Improved practical algorithms and solving subset sum problems," *Mathematical programming*, vol. 66, pp. 181–199, 1994.

[185] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inf. Theory.*, vol. 48, no. 8, pp. 2201–2214, 2002.

[186] B. Hassibi and B. Vikalo, "On the expected complexity of integer least-squares problems," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2002.

[187] N. Kraska, "Data detection for MIMO OFDM systems with reduced complexity in a frequency selective environment," Ph.D. dissertation, Universitaet der Bundeswehr Muenchen, 2017.

[188] Y. Chen, A. Lobato, Y. Jung, H. Chen, B. Inan, V. A. J. M. Sleiffer, M. Kuschnerov, R. V. Jensen, Y. Sun, L. Grüner-Nielsen, I. P. Giles, R. Chen, V. Garcia-Munoz, N. K. Fontaine, R. Ryf, S. U. Alam, D. J. Richardson, A. M. J. Koonen, B. Lankl, and N. Hanik, "41.6 Tbit/s C-Band SDM OFDM Transmission Through 12 Spatial and Polarization Modes Over 74.17 km Few Mode Fiber," in *Proc. European Conference on Optical Communication (ECOC)*, 2014.

[189] X. Yi, W. Shieh, and W. Tang, "Phase estimation for coherent optical OFDM," *IEEE Photon. Technol. Lett.*, vol. 19, no. 12, pp. 919–921, 2007.

[190] D. Zibar, O. Winther, N. Franceschi, R. Borkowski, A. Caballero, V. Arlunno, M. N. Schmidt, N. Guerrero Gonzales, B. Mao, Y. Ye, K. J. Larsen, and I. Tafur Monroy, "Nonlinear impairment compensation using expectation maximization for dispersion managed and unmanaged PDM 16-QAM transmission," *Opt. express*, vol. 20, no. 26, pp. B181–B196, 2012.

[191] Y. Chen, "Orthogonal frequency-division multiplexing in fiber-optic communication," Ph.D. dissertation, Technische Universitaet Muenchen, 2016.

[192] C. M. Bishop, *Pattern recognition and machine learning.* New York, NY: Springer, 2006.

[193] Y. Chen, C. Ruprecht, W. Rosenkranz, and N. Hanik, "Fiber nonlinearity compensation for dispersion unmanaged PDM 8-QAM CO-OFDM using expectation maximization," in *Proc. OptoElectronics and Communications Conference (OECC)*, 2013.

[194] E. Awwad, G. Rekaya-Ben Othman, Y. Jaouën, and Y. Frignac, "Space-time codes for mode-multiplexed optical fiber transmission systems," in *Proc. Signal Processing in Photonic Communications (SPPCom)*, 2014.

Acknowledgement

First of all, I would like to express my gratitude to Prof. Lankl for giving me the opportunity to do a Ph.D. under his supervision at the Institute for Communications Engineering. His great knowledge and understanding of optical communications make possible an excellent guidance along my Ph.D.

This work was sponsored by the telecommunications company Coriant GmbH (former Nokia Siemens Networks GmbH & Co KG) in Munich, Germany. The long-haul communications research and development team have supported me since 2006 when I did an internship with Erik De Man. Later on, the same department gave me the opportunity to do the master thesis under the supervision of Dr. Sander Jansen. He believed in me and encouraged me to do a Ph.D., for which I am now very thankful. At Coriant I had the opportunity to meet excellent experts in the field and work in one of the greatest optical communication research labs in the world. I would like to thank Dr. Marc Bohn and Dr. Stefan Spälter to giving me the opportunity of doing my Ph.D. at Coriant and for always giving me green light to publish the results of my work. I am greatly indebted to Dr. Maxim Kuschnerov for supervising my project. He guided me with his great supervision, knowledge on signal processing and inspiring spirit. Even though he was not working in the same department in Coriant at the end of my Ph.D., he took the time to discuss work issues. Together with Maxim, I counted with the supervision of Dr. Bernhard Spinnler, who was always available to answer my questions and discuss in which direction the Ph.D. can go. I would like to thank Dr. Stefano Calabrò and Antonio Napoli for the helpful technical discussions. The lab work would not have been possible without the kind support of Erik De Man and Dr. Guido Saathoff on components and lab equipment, and Dr. Ulrich Gaubatz, from whom I learned to operate the amplifier cards and enjoyed his classes on fiber physics together with the rest of Ph.D students at Coriant.

The support of the German Federal Ministry of Education and Research (01BP12300A, EUREKA-Project SASER), and the European Committee 7th Framework Program under grant agreement 258033 (MODE-GAP) is gratefully acknowledged. I would like to especially thank some of the members of the MODE-GAP project, namely Dr. Filipe Ferreira from Coriant Portugal whose great collaboration made possible to integrate the TX and RX DSP with his fiber model and who I would like to thank for partially proof-read my thesis; Dr. Vincent Sleiffer and Dr. Haushuo Chen from the Technical University Eindhoven for their support in the lab; Dr. Yongmin Jung from the Optical Research Centre (ORC) Southampton for his help with the few-mode components including a prototype of six-mode amplifier; and Dr. Lars Grüner Nielsen from OFS for manufacturing the six-mode fiber used in the experiments.

I also would like to express my gratitude to my Ph.D. students colleges from Coriant and the university, who were very friendly and always willing to help. Especially to Beril Inan and Susmita Adhikari, from whom I learned a lot about research methodology, presentations, OFDM and lab, just to name a few things. Beril and Susmita became soon not only excellent research partners but very good friends. I am also very grateful to my college from the Technical University München Yingkan Chen, with whom I worked in the lab through many long work days. I enjoyed the time we worked together both from the technical point of view and the working atmosphere, since it was a very stressful time but we managed to do not give up. Sincere thanks to Johannes Rabe, who learned very fast about the simulation environment and afterward could support me on the simulations. Thanks to my colleges from Coriant Ginni Khana, Tahla Rahman, Carlos Castro, Xiaozhou Wang and Paolo Leoni, and from our University Zifeng Wu, Vito Dantona, Nora Tax, Stefan Ludwig, Daniel Schmidt, Tim Hälsig and Anselm Karl for making up such a good and positive working environment and the Kicker matches. My sincere thanks to Karl-Heinz Besthorn from our University, who was most of the time available, supportive and made a lot easier the administrative work.

I am very grateful to my family: my parents, Pedro and Yadira, my brother and his wife, Javier and Karym, and my sister and his husband, Iveth and Dirk. From the beginning, since I moved "temporarily" to Germany, they gave me the support I needed to carry on, since it takes way more than a financial support. Especially, many thanks to Dirk and Iveth for their wise advise, proof-reading my papers and part of this thesis. Last but not least (on the contrary), I would like to deeply thank my lovely husband for his support, his unfailing love and particularly, for his patience. This work is dedicated to you, Rafael.