

Challenges in Estimating the Information Capacity of the Fiber-Optic Channel

This article reviews fiber-optic capacity scaling from an information-theoretic perspective, including the difficulties imposed by the nonlinear nature of fiber-optic transmission.

By MARK SHTAIF¹, Fellow IEEE, CRISTIAN ANTONELLI², Senior Member IEEE, ANTONIO MECOZZI², Fellow IEEE, AND XI CHEN³, Senior Member IEEE

ABSTRACT | Since its early commercial deployment in the late 1980s, optical fiber has evolved to become the predominant carrier of the globe's communications. Yet, after accommodating the world's exponentially growing appetite for transmitted data for more than three decades, its ability to continue doing so is being challenged by fundamental factors. In this article, we review these factors and examine their consequences in terms of information capacity. In particular, we review the difficulties that are imposed by the nonlinear nature of fiber-optic transmission on the assessment of the capacity and on the definition of fundamental concepts, such as bandwidth and spectral efficiency. We discuss relevant approximations and regimes of operation in which bounds for the capacity can be effectively assessed while covering a broad range of applications ranging from interdatacenter communications to links spanning transoceanic distances. We relate to a broad variety of transmission schemes and discuss the potential benefits of spatial multiplexing with multimode and multicore fibers.

State-of-the-art transmission experiments are also reviewed and compared with theoretical capacity bounds.

KEYWORDS | Information capacity; optical communications; space-division multiplexing (SDM).

I. INTRODUCTION

For more than three decades, fiber-optic communications have been the unchallenged champion in satisfying the world's exponentially growing appetite for transmitted data [1]. Today, the vast majority of exchanged information passes through an optical fiber somewhere along its journey from source to destination. The bandwidth that is available for communication in optical fibers is larger by many orders of magnitude than in other communications media, such as copper cables and wireless. Today's optical fiber-communications' networks carry well over one exabit (10^{18} bit/s/Hz) per second over multiple billions of kilometers of glass fiber, which is wrapped around the globe [1].

When fiber communications were first commercially adopted in the late 1980s, the available bandwidth appeared to be infinite relative to the needs in those days, and the challenges that needed to be met were related to the physical aspects of signal generation and transmission. Since bandwidth was abundant, no efforts were invested into spectrally efficient signaling, and information was typically encoded by directly modulating the source laser power, while the reception was based on the direct detection (DD) of intensity [2], [3]. Characteristic communication rates over standard single-mode fibers (SMFs) that were commercial in those days were of the order of 1 Gb/s, and the typical system reach was below 100 km [2],

Manuscript received 21 October 2021; revised 19 May 2022; accepted 30 July 2022. Date of publication 24 August 2022; date of current version 11 November 2022. The work of Mark Shtaif was supported by the Israel Science Foundation under Grant 1090/20. The work of Cristian Antonelli and Antonio Mecozzi was supported by the Italian Government in part under Cipe Resolution n. 135 (December 21, 2012) through project INnovating City Planning through Information and Communication Technologies (INCIPICT), and in part through the PRIN 2017 project Fiber Infrastructure for Research on Space-division Multiplexed Transmission (FIRST). (Corresponding author: Mark Shtaif.)

Mark Shtaif is with the Department of Physical Electronics, Tel Aviv University, Tel Aviv 69978, Israel (e-mail: shtaif@tauex.tau.ac.il).

Cristian Antonelli and **Antonio Mecozzi** are with the Department of Physical and Chemical Sciences, University of L'Aquila, 67100 L'Aquila, Italy, and also with the National Laboratory of Advanced Optical Fibers for Photonics (FIBERS), CNIT, 67100 L'Aquila, Italy.

Xi Chen is with Nokia Bell Labs, Murray Hill, NJ 07974 USA.

Digital Object Identifier 10.1109/JPROC.2022.3197188

whereas communications over longer distances required signal reception and retransmission (i.e., electro-optical regeneration).

The first important turning point in the evolution of fiber-communications' systems was the commercialization of erbium-doped fiber amplifiers (EDFAs) [4], [5], [6], which enabled in-fiber optical amplification within a bandwidth of 4 THz around the highest-transparency region of the optical fiber (near 1.5 μm). This invention was accompanied by the development of optical dispersion-compensating modules [7], [8], which lifted the reach limit of fiber-communications' systems making nonregenerated optical transmission over transoceanic distances of 10 000 km possible [9].

The next turning point was the transition to wavelength-division-multiplexed (WDM) transmission [10]. In this paradigm, a single fiber is shared by multiple transmitter and receiver pairs, each operating within a uniquely prescribed frequency band. With the adoption of WDM, the data rates transmitted over deployed optical fiber-systems increased at a rate of 100% per year [11], until toward the end of the 1990s, when the available optical bandwidth was essentially exhausted. At that stage, a single fiber is typically supported on the order of 80 WDM channels, each operating at 10 Gb/s in the conventional EDFA band (C-band). The subsequent deployment of the enhanced bandwidth (C + L band) EDFAs led to the doubling of the transmitted data rates, but it became obvious that further increase in throughput requires the replacement of intensity modulation (IM) and detection with more spectrally efficient signaling and detection techniques.

This notion was the driving force behind the transition to coherent communications schemes, which was the next important turning point in the evolution of fiber-communication systems. The shift to coherent was made possible by advancements in high-speed electronics and digital signal processing (DSP) [12] and allowed communications to take advantage of both quadratures of the transmitted electric fields and its two orthogonal polarization components, thereby increasing the number of exploited degrees of freedom by a factor of four [13], [14], [15]. The transition to coherent transmission and detection opened the door to a plethora of communication methods that have been devised over decades in wireless and wire-line radio frequency communications, including the use of advanced modulation formats, and digital compensation for physical propagation phenomena [14]. In particular, it enabled digital compensation of chromatic dispersion after detection [16], which resulted in abandoning inline optical dispersion compensating modules and considerably simplified the link design.¹ Other propagation phenomena, such as polarization-mode dispersion (PMD), which was a major obstacle to increasing data

rates in intensity-modulated DD systems [18], are now also routinely mitigated in the digital domain. As a result, coherent communication systems were able to comfortably accommodate the continuing growth in demand until the end of the first decade of this millennium when it became clear that a capacity crunch was becoming imminent [19].

The reasons for this capacity crunch lie in the unique properties of light propagation through optical fibers, in particular, in the fact that, at high optical powers, signal propagation becomes nonlinear, resulting in waveform distortions that hinder data recovery. This type of nonlinearity is unique to glass optical fibers [20], and in its presence, the usual reasoning applied in linear communication channels no longer applies. The nonlinearity of the fiber is responsible for the fact that, with known transmission schemes, there is always a system-specific launch power limit beyond which the communication performance of the system reduces. A significant fraction of research conducted in optical communications over the past decades has focused on the understanding and characterization of the fiber nonlinearity and its consequences, with the attempt of mitigating the nonlinear distortions and pushing the launch power limit farther away.

In the search for schemes that allow continued economically viable growth of data rates that are transmitted over fiber-communication systems, space-division multiplexing (SDM) has established itself as a prominent candidate in the past decade [21]. The principle of SDM is that SMFs are replaced with multimode or multicore fibers in which multiple information streams can propagate simultaneously in orthogonal spatial modes [22], [23], [24], [25], [26], [27], [28], [29]. The idea is to encourage the integration of end-equipment and inline optical components that will allow growth in the information throughput without a corresponding increase in cost, as would be the case when simply deploying parallel systems.

In this article, we discuss the ultimate limits to the rate at which information can be reliably communicated over the fiber-optic channel. In the framework of information theory, this quantity is known as the channel capacity, a concept that was introduced and formulated by Shannon [30] in 1948. While Shannon provided a formal expression for the capacity of a generic channel, applying it to a given physical channel is often a very challenging task, particularly so in the presence of nonlinearities, such as those characterizing fiber-optic transmission. Indeed, in spite of the numerous attempts that have been conducted over the years [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], determining the fiber-optic channel information capacity remains an open problem. Nonetheless, methods for assessing capacity lower bounds have been proposed and studied [44], [45], [46], [47], [48], [49], [50].

In what follows, we review the relevant phenomena involved in fiber-optic transmission and discuss their impact on capacity. We include considerations related to ultralong-haul and under-sea cable systems, on the one

¹Interestingly, the abandoning of inline chromatic dispersion compensation turned out to be advantageous also in terms of the system tolerance to nonlinear distortions, as demonstrated in [17].

hand, and short interdata-center systems, on the other hand. In the process, we discuss the main past reported attempts of assessing the fiber channel capacity and review the state of the art of transmission rates that have been demonstrated experimentally.

We stress that, in this article, we do not account for any practical constraints on coding and modulation. The effect of such constraints on the achievable information rate is extensively treated in [51] and references therein. Other fundamental implications of coding are addressed in [52].

This article is organized as follows. In Section II, we review the essential preliminaries that are needed for understanding the subject. These include the necessary concepts from information theory, on the one hand, and from the theory of fiber-optic propagation, on the other hand. In Section III, we discuss the difficulty of assessing the fiber-optic channel capacity and review some of the reported attempts. In Section IV, we introduce the concept of nonlinear interference noise (NLIN) and discuss its modeling in the context of information capacity. Section V discusses the construction of a lower bound for the nonlinear fiber-optic channel capacity, while the capacity upper bound is reviewed in Section VI. Section VII discusses the effects of polarization-related propagation phenomena on the assessments of the information capacity, and in Section VIII, we provide some concluding remarks regarding the capacity of long-haul transmission systems. Section IX is devoted to reviewing fundamental capacity considerations of direct-detection systems, which are often the solution of choice for short-reach and interdata-center communications. In Section X, we review capacity-related aspects in space-division multiplexed transmission in the submarine and terrestrial systems. Finally, Section XI reviews state-of-the-art results in laboratory experiments and commercial systems.

II. ESSENTIAL PRELIMINARIES

Initially, the field of optical communications started as a branch of physics engineering, gradually adopting concepts from communications theory, signal processing, and information sciences, to the extent that it has eventually become a truly interdisciplinary science. Individuals interested in understanding fiber communications today need to be simultaneously familiar with concepts of information and communication theories, as well as with the physical properties of optical fibers that govern signal propagation. In order to provide the reader with the essential background, in this section, we introduce the key concepts of information theory, on the one hand, and fiber optic transmission, on the other hand.

A. Information Capacity

Fundamentally, communications' systems contain three essential components: a transmitter that encodes information onto some transmitted physical object, a channel through which the object propagates over distance, and a receiver whose role is to extract the encoded information

from the physical object once it arrives. What makes the process of communications nontrivial is the fact that, while propagating through the channel, the transmitted object is affected by distortions and noise. In most cases of interest to modern telecommunications, the physical object onto which information is encoded is the electromagnetic field. Denoting by $X(t)$ the signal that is generated by the transmitter and launched into the channel and by $Y(t)$ the signal that emerges from the channel and impinges upon the receiver, the effect of the channel is described by the conditional probability density of receiving $Y(t)$, given that $X(t)$ was transmitted. As information is usually encoded on the complex envelope of electric fields, the elements of $X(t)$ and $Y(t)$ are complex-valued. Moreover, in such cases, $X(t)$ and $Y(t)$ are vectors whose components represent orthogonal polarizations or propagation modes of the electric field. In what follows, we start from the consideration of scalar complex-valued $X(t)$ and $Y(t)$ signals, leaving the generalization to the multidimensional case for a later stage, after the main results are established.

Formally, in order to relate to the conditional probability of receiving $Y(t)$ given that $X(t)$ is transmitted, the continuous scalar time entities $X(t)$ and $Y(t)$ are represented by vectors X and Y having a finite number of elements. Such representation is always possible when dealing with band-limited signals whose spectrum is constrained to be within a bandwidth B and which are considered within a finite time frame T_f . Sampling at a rate that satisfies Nyquist's condition is perhaps the most obvious example of a procedure through which such representation can be achieved. Since the number of degrees of freedom in the continuous-time signals is $2BT_f$ (the factor of two accounts for the two signal quadratures), the dimension of the vectors X and Y is BT_f , as each vector element is complex-valued and, therefore, carries two degrees of freedom. The elements of X and Y are referred to as the transmitted and received symbols, respectively, and the effect of the channel is represented by the conditional probability density function $P_{Y|X}(y|x)$.² Similarly, the operation of the transmitter is fully characterized by the probability density function $P_X(x)$.

One of Shannon's most groundbreaking achievements was to demonstrate that the largest amount of information per symbol that can be reliably communicated between the transmitter and the receiver is given by

$$I(X; Y) = -\frac{\mathbb{E}[\log_2(P_X)] - \mathbb{E}[\log_2(P_{X|Y})]}{BT_f} \quad (1)$$

where \mathbb{E} denotes ensemble averaging, and $P_{X|Y}(x|y) = P_{Y|X}(y|x)P_X(x)/P_Y(y)$. This quantity is known as *mutual information per symbol*, and it is measured in units of bit/s/Hz. The related quantity known as *channel capacity*

²Since X and Y are complex-valued, the symbol $P_{Y|X}(y|x)$ denotes a bivariate conditional probability density function whose entries are the real and imaginary parts of X and Y .

is obtained by optimizing the mutual information with respect to the transmitter in the limit of an infinitely long time frame, namely,

$$C = \lim_{T_f \rightarrow \infty} \sup_{P_X} \{I(X; Y)\}. \quad (2)$$

Notice that, different from some of the existing literature, normalization with respect to BT_f is included in our definition of mutual information. In particular, the division by B implies that, in our definition, C is capacity per unit bandwidth, which is equivalent to what is sometimes referred to as spectral efficiency (SE) and often expressed in bit/s/Hz rather than simply in bits, although, of course, the two coincide with each other. In this article, for brevity, we will often simply refer to C as capacity and reserve the term SE for Section XI, where we use it to relate to the information throughput per unit bandwidth that is successfully transmitted in a given experimental system.

It should also be noted that, while, in linear channels, the total capacity is given by the product of the capacity per unit bandwidth C and the total bandwidth allocated for transmission, this relation does not hold in the case of nonlinear channels, as will be discussed in Section III.

While (1) and (2) provide a clear path for determining the capacity of a known channel, establishing knowledge of the channel law $P_{Y|X}(y|x)$ is often a dauntingly difficult task, and as we shall see, it is particularly difficult in the case of the fiber-optic channel, which is the focus of this article. A famous and important special case in which an explicit formula for the channel capacity exists is that of the memoryless additive Gaussian noise (GN) channel where the relation between the k th input and output symbols is given by $Y_k = X_k + N_k$, with the noise elements N_k being identically distributed and statistically independent complex-circular zero-mean Gaussian variables. In this case, under the constraint of a given average power, the optimal distribution of each of the symbols X_k is Gaussian, and the capacity per unit transmission bandwidth is given by Shannon's famous formula [30]

$$C = \log_2 \left(1 + \frac{\mathbb{E}[|X_k|^2]}{\mathbb{E}[|N_k|^2]} \right). \quad (3)$$

This result can now be generalized to the case where information is encoded simultaneously over multiple modes (such as polarizations or spatial propagation modes in fibers). Under the assumption that the noise is independent and identically distributed in the various signal dimensions, the capacity is achieved by dividing the transmitted energy equally between all dimensions, with the result

$$C = M \log_2 (1 + \text{SNR}) \quad (4)$$

where M is the number of scalar modes and SNR is the signal-to-noise ratio defined as the ratio between

the average signal and noise powers in any of the modes

$$\text{SNR} = \frac{\mathbb{E}[|X_k|^2]}{\mathbb{E}[|N_k|^2]}. \quad (5)$$

1) *Useful Lower Bound for Capacity:* Apart from being attractive from an analytical standpoint, an important practical feature of the Gaussian memoryless channel is that it can be used to obtain a *lower bound* for the capacity of a generic communications channel, as discussed in [53] and [54]. The idea is based on the notion that white GN has the highest entropy for a given noise power, and in that sense, it is the "noisiest" noise possible. In order to obtain the bound, one constructs an auxiliary memoryless additive GN channel

$$Y_k = cX_k + N_k \quad (6)$$

with a scaling factor

$$c = \frac{\mathbb{E}[X_k^* Y_k]}{\mathbb{E}[|X_k|^2]} \quad (7)$$

whose role is to ensure that the auxiliary channel satisfies the same correlation relation $\mathbb{E}[X_k^* Y_k]$ as the original one. Setting the variance of N_k to be $\mathbb{E}[|N_k|^2] = \mathbb{E}[|Y_k - cX_k|^2]$, the desired lower bound is obtained by applying Shannon's formula

$$C \geq M \log_2 \left(1 + \frac{\mathbb{E}[|cX_k|^2]}{\mathbb{E}[|Y_k - cX_k|^2]} \right) \quad (8)$$

which is identical to (4) and (5), except for the scaling factor c .

B. Fiber-Optic Channel

Optical communications' systems that are used today rely on single-mode optical fibers. These fibers are dielectric waveguides supporting propagation of only one spatial mode with two orthogonal polarizations. The two field polarizations can be—and in fact are—addressed as two scalar communication channels operating simultaneously. Systems of this type are referred to as *polarization-multiplexed* (PM).

In the past decade, the use of fibers supporting multiple spatial modes has been proposed and demonstrated in numerous laboratory experiments [22], [23], [24], [25], [26], [27], [28], [29]. In these fibers, the number of simultaneously transmitted communication channels is $2N$, where N is the number of spatial modes and where the factor of two accounts for two orthogonal polarizations. Systems of this type are referred to as *space-division multiplexed*. For simplicity, we focus the fundamental description on the case of PM systems, postponing the generalization to the case of SDM systems to Section X.

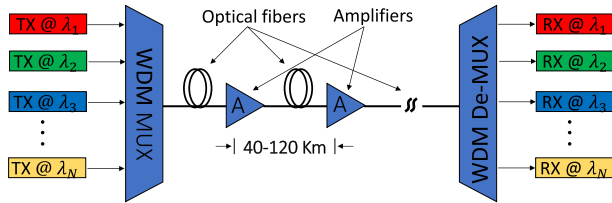


Fig. 1. Schematic of a generic WDM system. Transmitters (TX) generating information-carrying signals at carrier wavelengths $\lambda_1, \dots, \lambda_N$ are multiplexed into a link consisting of multiple fiber spans separated by optical amplifiers. The signals are then demultiplexed at the output of the link and sent to the respective receivers (RX).

The transparency window of available optical fibers is centered around the wavelength of $1.5 \mu\text{m}$, which is where almost all fiber communications take place.³ The transmission bandwidth is determined mostly by the amplification bandwidth of available optical amplifiers, and in the most common cases, using EDFAs [4], [5], [6], it is of the order of 4 THz although extended-bandwidth systems in which the amplification bandwidth reach 10 THz also exist and are attracting significant attention [55]. Typically, fiber-optic systems are wavelength-division multiplexed, which means that an array of transmitters is deployed, where each transmitter operates within a unique spectral window. When combined with polarization multiplexing, each transmitter imposes two complex-valued information-carrying signals onto the two orthogonal polarizations of the generated electric field. A schematic description of a typical WDM system is illustrated in Fig. 1. The signals generated by the various transmitters are combined with the help of an optical multiplexer and launched into the link, which consists of multiple spans of optical fiber separated by optical amplifiers. At the receiver, an optical demultiplexer separates the individual WDM channels and feeds each channel into its corresponding receiver.

Since our focus in this article is on estimates of the fiber-optic channel capacity, we will assume that the transmitters and receivers are ideal. Namely, the transmitter is assumed to be capable of producing any desirable waveform whose spectrum is contained in a specified channel bandwidth B , and the receiver is capable of measuring any waveform that impinges upon it with infinite precision.

Within this framework, the performance of fiber communications is limited solely by the combination of fiber-propagation distortions and amplification noise, to which we refer in more detail in what follows. The length of the fiber spans varies between ~ 40 km in legacy transoceanic systems and ~ 100 km in terrestrial transmission and some modern transoceanic links (e.g., the Curie cable and the Google-owned Subcom system [56]). As the fiber loss is of the order of 0.2 dB/km, the loss

of the individual span is between 8 and 20 dB. The amplifiers placed at the end of each span compensate for this loss. The less desirable, but fundamental property of amplification that is dictated by quantum mechanics, is the generation of noise.

An optical amplifier with power gain G ($G \geq 1$) produces complex-circular GN, which is referred to as amplified spontaneous emission (ASE) noise, whose spectral density in each polarization is given by [57]

$$\hbar\omega_0 n_{\text{sp}}(G - 1) \quad (9)$$

where ω_0 is the central angular frequency of the amplified signal, \hbar is the modified Planck's constant, and n_{sp} is the so-called inversion factor, which is equal to 1 for ideal amplifiers and is greater than 1 in all practical cases (in typical amplifiers, n_{sp} is between 1.6 and 2). Although the presence of ω_0 in (9) implies a direct dependence of the noise power on the optical frequency, this dependence is negligible because the range in which ω_0 varies within the amplification bandwidth is very small relative to the central frequency of the amplification band. Therefore, within regions in which the amplification spectrum is reasonably flat (as required in communications' systems), the noise can be legitimately approximated as white.

The unique features of optical communications are related to the propagation properties of optical fibers. Most importantly, fibers are characterized by loss, chromatic dispersion, and nonlinear propagation distortions resulting from the dependence of the refractive index of glass on optical power. These phenomena interact nontrivially with each other and with the amplification noise that accompanies the signal, thereby challenging the process of extracting the transmitted information. The equation that captures these phenomena is known as the Manakov equation [58], [59], [60]

$$\frac{\partial \vec{E}}{\partial z} = -\frac{\alpha}{2} \vec{E} - i \frac{\beta_2}{2} \frac{\partial^2 \vec{E}}{\partial t^2} + i \frac{8\gamma}{9} |\vec{E}|^2 \vec{E} \quad (10)$$

where $\vec{E}(z, t)$ is a complex-valued 2-D column vector, describing the electric field as a function of position and time. As is customary in the description of signal propagation, t is defined with respect to a moving reference frame.⁴ The orientation of the vector \vec{E} is the signal's state of polarization, and the units of \vec{E} are chosen such that $|\vec{E}|^2$ is the optical power. The term α is the fiber-loss coefficient, β_2 is the chromatic-dispersion coefficient, and the coefficient γ accounts for the fiber nonlinearity. Equation (10) is derived from the more fundamental coupled Schrödinger equations [60], [61], and it accounts for the fact that propagation in long optical fibers is accompanied by rapid random polarization rotations that

³Very short distance communications, such as communications within data centers, are less sensitive to loss and are often performed around $1.3 \mu\text{m}$.

⁴The actual time is $t + z/v_g$, where v_g is the group velocity at the chosen central frequency.

average the nonlinear effect. As a result of this averaging, the nonlinearity enters (10) only through the optical power $|\vec{E}|^2$, and hence, its effect is independent of the signal polarization. Yet, we should stress that the Manakov equation does not include other polarization effects, such as PMD and polarization-dependent loss (PDL), to which we refer separately in Sections VII-A and VII-B, respectively.

At low signal powers, where $\gamma|\vec{E}|^2$ is small to the extent that the nonlinear term can be neglected, the propagating field is given by

$$\vec{E}(z, t) = e^{-\frac{\alpha}{2}z} e^{-\frac{i}{2}\beta_2 z \frac{\partial^2}{\partial t^2}} \vec{E}(0, t) \quad (11)$$

which can be implemented in the frequency domain as

$$\tilde{\vec{E}}(z, \omega) = e^{-\frac{\alpha}{2}z} e^{\frac{i}{2}\beta_2 \omega^2 z} \tilde{\vec{E}}(0, \omega) \quad (12)$$

where $\tilde{\vec{E}}(z, \omega) = \int_{-\infty}^{\infty} \vec{E}(z, t) \exp(i\omega t) dt$ is the Fourier transform of the complex envelope of the field in the time domain so that ω is the angular frequency.⁵ The roles of attenuation and chromatic dispersion become evident when looking at (12), which also shows that chromatic dispersion is an all-pass filter. It is clear that amplifiers can compensate for the effect of attenuation, whereas the effect of dispersion can be undone by applying an all-pass filter with dispersion of the opposite sign. This can be done either in the optical domain, or electronically, by DSP, after coherent detection.

Some insight into the effect of nonlinearity can be gained by solving the Manakov equation in the absence of chromatic dispersion ($\beta_2 = 0$), yielding

$$\vec{E}(z, t) = e^{-\frac{\alpha}{2}z} e^{i\gamma|\vec{E}(0,t)|^2 z_{\text{eff}}(z)} \vec{E}(0, t) \quad (13)$$

where $z_{\text{eff}} = [1 - \exp(-\alpha z)]/\alpha$ is known as the *effective* propagation distance [62]. Equation (13) shows that, in the absence of dispersion, nonlinearity imposes power-dependent phase modulation onto the propagating signal, and thereby, it affects its spectrum, typically spreading it far beyond the original bandwidth.

In all practical situations, when both dispersion and nonlinearity are present, the propagation dynamics become more involved, and in general, the evolution of $E(z, t)$ along the fiber can only be evaluated by solving the Manakov equation numerically. Effective numerical methods for solving the Manakov equation are well-known [63], and they are deployed routinely when simulating fiber communications' systems in the process of system analysis

⁵Note that the sign convention in the definition of the Fourier transform is consistent with the optical communication literature and is opposite to what is common in some other fields of electrical engineering. Similarly, following the same convention, all analytical expressions in this article contain the angular frequency ω as opposed to the frequency $\omega/2\pi$. Nonetheless, bandwidths and frequency separations are specified in units of Hz.

and design. A relevant and interesting property of this equation and the numerical methods used for its solution is that it can also be solved backward. This property has important fundamental and practical consequences in the operation of the system and in the estimation of its capacity. The idea is that the electric field that is recovered by a receiver positioned at the end of the fiber ($z = L$) can be digitized and used as the boundary condition for a real-time numerical backward solution of (10), so as to approximate the field at the beginning of the fiber at $z = 0$. This procedure is known as digital back propagation (DBP) [64], and while it is still absent in commercial systems owing to its high computational complexity, it is occasionally implemented in laboratory systems with off-line processing. The reason why DBP can only approximately reconstruct the input waveform is that it also unavoidably back-propagates amplification noise generated along the system.

In addition to numerical solutions of the Manakov equation, useful approximate analytical solutions that rely on perturbation analysis can be applied in cases of weak or moderate nonlinearity. These will be discussed in Section IV.

III. ASSESSING THE FIBER-OPTIC CHANNEL CAPACITY

Prior to delving into the details of assessing the capacity of the fiber-optic channel, it is useful to set up our expectations based on the signal propagation properties represented in (10). The expected features of the dependence of capacity on the average power of the launched optical signal are illustrated schematically in Fig. 2. At low powers, signal propagation in the fiber is essentially linear. Moreover, since dispersion is an all-pass filter and loss is frequency-independent, the amplification noise remains white Gaussian, and Eq. (4) for the capacity holds. In this regime, by increasing the signal power, the SNR increases accordingly and so does the capacity. However, as the optical power increases to the extent that nonlinearity can no longer be neglected, distortions are introduced both into the signal and into the noise, whose statistical distribution changes eventually. In this regime, the capacity is unknown, and only bounds for it are available. An upper bound is the capacity of the AWGN channel [65], [66], whereas, in the case of the lower bound, multiple system-specific estimates have been reported [44], [45], [46], [47], [48].

In Section IV, we discuss some of these lower bounds. Here, we wish to illuminate the fundamental difficulties of addressing the estimation of the fiber-channel capacity in the regime of nonlinear propagation. One such difficulty is the absence of a closed form relation between the input and output waveforms [except in the special cases represented by (12) and (13)], which implies that an expression for $P_{Y|X}(y|x)$, which is a crucial component in the determination of the capacity in (1) and (2), also does not exist. The second difficulty follows from the fact that, in the

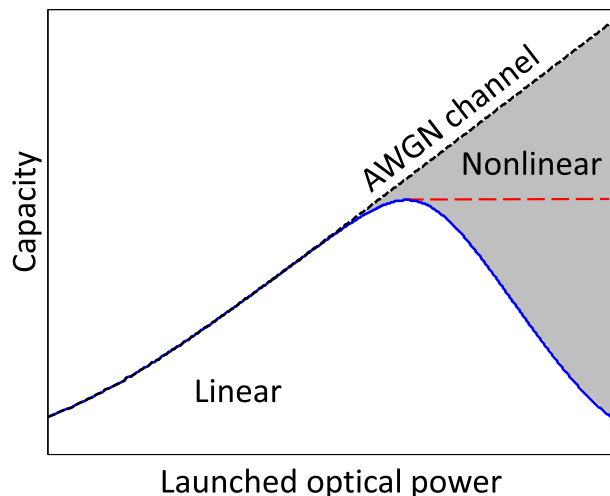


Fig. 2. Illustration of the dependence of the fiber-channel capacity on the launched signal power. At low power levels, propagation is linear, and the fiber channel is equivalent to an AWGN channel, whose capacity is given by (4). When the power increases to an extent that nonlinear effects become significant, the capacity is unknown. It is bounded from above by the AWGN capacity, whereas lower bounds are the subject of ongoing research, as discussed in Section V. The flat dashed line relates to the option of clipping the power coupled into the link to the value at which the system performance peaks [67].

presence of nonlinear propagation, bandwidth—another crucial component in the concept of capacity—cannot be properly defined. The information-carrying signal changes its bandwidth in the process of propagation according to a nontrivial pattern that depends on the shape of the launched waveform, its average power, and the physical properties of the link [62]. Curiously, in principle, spectral broadening even allows one to construct schemes in which the capacity of a nonlinear fiber system within the bandwidth allocated to the transmitter exceeds the capacity of a linear channel of that same bandwidth [68] (although the resulting scheme’s complexity makes its actual implementation impractical). Nonlinearity also implies interference between spectrally separated information-carrying signals, from which it follows that the information rate that can be realized within a given frequency band depends on what is being transmitted outside of it.

For all of these reasons, information capacity estimates of fiber-communication systems have almost always been performed in the so-called pseudolinear regime [69], [70], [71], in which nonlinearity can be treated as a small perturbation to a mostly linear propagation. Even the early works by Mitra and Stark [31] and Turitsyn *et al.* [37], which do not formally limit their analysis to weak nonlinearities, do not hold at very high powers since they too ignore the spectral broadening that characterizes this regime.

The most influential study dealing with the assessment of capacity in the practical context of fiber communication

systems came out of Bell-Labs in 2008 [35] and 2010 [34], and outlined the broad framework that has been used in almost all capacity estimates ever since. This framework focuses on the practical setting of WDM transmission in which one of the channels is declared as the channel of interest (COI), whereas all other channels are referred to as interfering channels (ICs). The assumption is that different WDM channels are transmitted and received independently so that the receiver of the COI has no knowledge of the data transmitted in the other channels. Finally, the COI is assumed to be ideally back-propagated so that nonlinear distortions that are experienced in its propagation are eliminated. We emphasize that, since the entire system is assumed to operate in the pseudolinear regime, the difference between the input and output bandwidths is negligible.

The relevance of this framework of study follows from the fact that the vast majority of fiber-communication systems deployed in the past two decades are indeed operated in the regime of pseudolinear propagation. The system studied in [34] also assumed that chromatic dispersion is compensated for only at the receiver (as opposed to being compensated in a distributed fashion at the amplification sites), but this assumption is not critical to the validity of the analysis, except that it extends the range of signal powers for which the assumption of pseudolinearity holds. Nonetheless, most modern communications’ systems are of this type. In this framework, the interference that the ICs impose on the COI is experienced as noise, to which we refer as NLIN. The NLIN comes on top of the ASE noise that is contributed by amplification. The approach in [34] was to evaluate the channel distribution numerically. To this end, the authors considered discrete constellation points arranged on multiple concentric rings in the complex constellation space.

The number of rings and the number of points in each ring were chosen to be such that, in the absence of nonlinearity, the capacity of the system was well approximated by (4) for the considered range of SNRs. The numerical evaluation of the channel distribution in the presence of nonlinearity, which was performed by means of Monte Carlo simulations, revealed that, for each input-symbol energy, $P_{Y|X}(y|x)$ could be well approximated by a bivariate Gaussian distribution, specified by the best-fitting covariance matrix. It was also evident from the results of [34] that the covariance matrix of the NLIN was signal dependent, and in particular, it appeared as containing a visible contribution of phase noise (e.g., see [34, Fig. 2]). As the study in [34] was concerned with assessing capacity, the system operating conditions were idealized in various ways. In particular, the study assumed ideal distributed amplification.⁶ Furthermore, in order to alleviate the computational effort, this study was performed for a scalar

⁶In this regime, the gain and loss balance each other locally, and the power spectral density of the noise produced within a fiber section of length dz becomes $\hbar\omega_0 n_{sp} \alpha dz$, consistently with (9).

(single polarization) case, as opposed to PM transmission. Essiambre *et al.* [34], [35] provided an important data point in the study of the fiber-optic channel capacity as they reported a lower bound for capacity that was higher by approximately a factor of two than the data rates observed experimentally around that time (see [34, Fig. 38]).

Subsequent studies related to the capacity of fiber-communications channels [40], [43], [45], [48], [49], [50], [72], as well as [54] and references therein, adopted the general framework of [34] but deepened the understanding of the underlying physical phenomena and expanded the consideration to more realistic settings. Further improvements in the assessment of capacity lower bounds were obtained by developing analytical models of NLIN [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86] and through attempts to account for its temporal correlation properties [48], [73], [74], [75], [76], [77], [78].

IV. MODELING THE NONLINEAR INTERFERENCE NOISE

As noted earlier, empirical evidence suggests that the NLIN can be well approximated as a Gaussian process, and hence, all modeling efforts are aimed at characterizing its second-order statistics. These are the power and correlation properties, both with respect to quadratures and polarizations, and with respect to time. Practically, all analytical studies of the NLIN statistics rely on a perturbation analysis, in the sense that they are accurate only to the first order with respect to the nonlinearity coefficient γ . Furthermore, as fiber communications' systems operate at an SNR that is much higher than unity, the nonlinear mixing between signal and noise is assumed to be negligible. Within this framework, the propagating field is first evaluated under the assumption of purely linear transmission (i.e., setting γ to 0), where the solution is given by (11). We call this the zeroth-order solution and denote it by $\vec{E}^{(0)}(z, t)$. Then, in first-order with respect to the nonlinearity coefficient γ , the field of the NLIN is obtained by solving the Manakov equation (10) with $\vec{E}^{(0)}(z, t)$ substituted into the nonlinear term as follows:

$$\frac{\partial \Delta \vec{E}}{\partial z} = -\frac{\alpha}{2} \Delta \vec{E} - i \frac{\beta_2}{2} \frac{\partial^2 \Delta \vec{E}}{\partial t^2} + i \frac{8\gamma}{9} |\vec{E}^{(0)}|^2 \vec{E}^{(0)} \quad (14)$$

with $\Delta E(0, t) = 0$. The solution to (14) can be conveniently expressed as

$$\Delta \vec{E}(z, t) = i \frac{8\gamma}{9} \int_0^z dz' e^{-\frac{1}{2}(\alpha + i\beta_2 \frac{\partial^2}{\partial t^2})(z-z')} \times |\vec{E}^{(0)}(z', t)|^2 \vec{E}^{(0)}(z', t) \quad (15)$$

where we made use of (11) to express $\vec{E}^{(0)}(z, t)$.

At this point, one may consider two distinct approaches for characterizing the NLIN. One is to proceed in the

time domain, as we shall review in what follows, and the other is to derive $\Delta \vec{E}(z, \omega)$ by converting (15) to the frequency domain, where the product $|\vec{E}^{(0)}(z', t)|^2 \vec{E}^{(0)}(z', t)$ is replaced by its Fourier transform (which consists of two convolution integrals), and $\partial/\partial t$ is replaced by $-i\omega$. The frequency-domain approach is equivalent to describing the NLIN as resulting from four-wave mixing processes involving all of the propagating-signal frequency components. This approach, which was first proposed by Splett *et al.* [79] in 1993, underpins the development of the well-known GN model, led by the Optcom group of the Politecnico di Torino [80], [81], which is rigorously valid under the assumption of Gaussian modulation. Removing this assumption resulted in the so-called enhanced GN (EGN) model developed in [73] and [82].

A conspicuous disadvantage of the frequency-domain approach is that it makes the extraction of temporal correlations of the NLIN, as well as some of its other temporal features, less transparent. For this reason, in what follows, we concentrate on the time-domain analysis of NLIN, first introduced in [87]. We focus on the case where all WDM channels use linear single-carrier modulation so that the complex signal $\vec{E}^{(0)}(z, t)$ appearing in (15) can be expressed as

$$\vec{E}^{(0)}(z, t) = \sum_n \vec{a}_n g(z, t - nT) + \sum_{m,n} e^{-i\Omega_m t} e^{i\frac{\beta_2}{2}\Omega_m^2 z} \vec{b}_n^{(m)} g(L, t - nT - \beta_2 \Omega_m z) \quad (16)$$

where the first summation describes the COI and the second represents the ICs. For convenience, the central frequency of the COI is defined as the reference optical frequency, and the central frequencies of the ICs are denoted by Ω_m . By the term \vec{a}_n , we denote the 2-D column vector whose elements are the complex-valued constellation symbols transmitted in the n th symbol duration over the two polarizations. Similarly, $\vec{b}_n^{(m)}$ represents the data transmitted in the n th symbol time slot over the m th IC. The symbol duration is denoted by T , and $g(z, t) = \exp(-i(\beta_2/2)z(\partial^2/\partial t^2))g(0, t)$ describes the propagation of the fundamental pulse waveform. We assume the usual case where the launched pulse $g(0, t)$ is chosen such that it satisfies the orthogonality condition $\int dt g(0, t - jT)g^*(0, t - kT) = \delta_{j,k}$, so that, in the absence of propagation-induced distortions, no intersymbol interference (ISI) is present when the signal is received after matched filtering. The field of the NLIN is obtained by substituting $\vec{E}^{(0)}(z, t)$ of (16) into (15).

The effect of the NLIN on the received n th constellation point after matched filtering is then given by

$$\Delta \vec{a}_j = \int_{-\infty}^{\infty} g^*(L, t - jT) \Delta \vec{E}(L, t) dt \quad (17)$$

where L is the system length and where we assume that the sampling time after the matched filter has been chosen ideally. Notice that the use of the propagated matched filter $g(L, t)$ is equivalent to performing dispersion compensation and then matched filtering with respect to the launched waveform $g(0, t)$. While the procedure of expressing $\Delta \vec{a}_n$ and its statistical properties explicitly in terms of the physical parameters of the problem is somewhat involved, the structure of the solution can be easily anticipated. Given that $\Delta \vec{E}$ is proportional to $|\vec{E}^{(0)}|^2 \vec{E}^{(0)}$ [see (15)] and given $\vec{E}^{(0)}$ of (16), it is clear that the expression for $\Delta \vec{a}_n$ contains all of the following triplets. The triplet $\vec{a}_l^\dagger \vec{a}_m \vec{a}_n$, which represents a single channel nonlinearity involving only the COI. The triplets $\vec{b}_l^{(k)\dagger} \vec{b}_m^{(k)} \vec{a}_n$ and $\vec{b}_l^{(k)\dagger} \vec{a}_m \vec{b}_n^{(k)}$, which represent a two-channel interaction, involving the COI and one of the ICs. The triplet $\vec{a}_l^\dagger \vec{b}_m^{(k)} \vec{b}_n^{(-k)}$ that represents a three-channel interaction involving the COI and two different ICs, and finally, the triplet $\vec{b}_l^{(h)\dagger} \vec{b}_m^{(k)} \vec{b}_n^{(h-k)}$ that accounts for a four-channel interaction, where three different ICs produce interference at a fourth frequency that coincides with the COI. In the special case where $k = h/2$, there are only two different ICs in the latter triplet, and it represents a three-channel interaction.⁷

In the jargon of optical nonlinearities, the three- and four-channel interactions belong to the class of four-wave-mixing (FWM) processes, whose effectiveness strongly depends on so-called phase-matching conditions [88]. These conditions are strictly fulfilled only in the absence of chromatic dispersion, which is never the case in contemporary fibers. For this reason, three- and four-channel interactions can be safely neglected in the vast majority of cases⁸ so that only one- and two-channel interactions need to be taken into account. In this case, the symbol perturbation that is generated by NLIN can be accurately described as follows:

$$\Delta \vec{a}_j = i \frac{8}{9} \gamma \sum_{l,m,n} S_{l-j,m-j,n-j} \vec{a}_l^\dagger \vec{a}_m \vec{a}_n + i \frac{8}{9} \gamma \sum_{k \neq 0l,m,n} \sum X_{l-j,m-j,n-j}^{(k)} \left(\vec{b}_l^{(k)\dagger} \vec{b}_m^{(k)} \mathbf{I} + \vec{b}_m^{(k)} \vec{b}_l^{(k)\dagger} \right) \vec{a}_n. \quad (18)$$

The terms $S_{l,m,n}$ and $X_{l,m,n}^{(k)}$ are proportionality coefficients that are obtained by following through

⁷All other triplets represent interference fields whose central frequencies coincide with that of other channels that are not the COI. Triplets whose central frequency coincides with channels that are immediately adjacent to the COI may contribute to the NLIN owing to the spectral broadening induced by nonlinearity, as pointed out in [82]. This contribution is very small in high-dispersion links carrying high symbol rates, and we neglect it in this article. Triplets whose central frequency coincides with farther away channels do not contribute at all to the NLIN.

⁸The contributions of three- and four-channel interactions to NLIN may not be negligible in dispersion-managed systems, systems with low-dispersion fiber, or low-baud rate systems (typically well below 10 Gbd) [83].

the perturbation analysis procedure presented in (15) and whose expressions can be found in [84]. These coefficients are uniquely determined by the fiber parameters, span lengths, fundamental pulse waveform $g(0, t)$, and the frequencies of the WDM channels. In particular, the two-channel interaction coefficients $X_{l,m,n}^{(k)}$ reduce in magnitude monotonically with the separation between the COI and the IC. This reduction reflects the effect of dispersion, which causes the interacting channels to propagate at different velocities and, thereby, reduces the effectiveness of the nonlinear interaction.

From the standpoint of capacity, the role of single-channel effects is immaterial because the interference can be removed, at least in principle [50], [64], by means of back propagation, as we discussed earlier. For this reason, we will omit this term in what follows.

A. ISI Model for the NLIN

In order to better appreciate the effect of nonlinear signal distortions on the fiber channel capacity, the NLIN can be conveniently interpreted as ISI. To this end, we define the quantity [85]

$$\mathbf{R}_n(j) = \frac{8}{9} \gamma \sum_{k \neq 0} \sum_{l,m} X_{l-j,m-j,-n}^{(k)} \left(\vec{b}_l^{(k)\dagger} \vec{b}_m^{(k)} \mathbf{I} + \vec{b}_m^{(k)} \vec{b}_l^{(k)\dagger} \right) \quad (19)$$

which allows expressing the second term on the right-hand side of (18), which accounts for two-channel interactions, as a time-dependent ISI process

$$\Delta \vec{a}_j = i \sum_n \mathbf{R}_n(j) \vec{a}_{j-n}. \quad (20)$$

Describing NLIN in terms of (20) is meaningful when the dependence of the ISI coefficient matrices $\mathbf{R}_n(j)$ on the time j is slow relative to their dependence on n . The reason for which this is so is that, in the presence of chromatic dispersion, different WDM channels propagate at different velocities, and each symbol of the COI is passed by a large number of symbols in each of the ICs. This implies that adjacent or nearby symbols of the COI typically experience very similar interference.

The contribution of the zeroth-order matrix $\mathbf{R}_0(j)$ to NLIN has a unique nature that differs from the nature of the higher-order ISI contributions. Examination of (19) reveals that the matrix $\mathbf{R}_0(j)$ is Hermitian, and therefore, its contribution manifests itself as phase and polarization-rotation noise (PPRN) [86]. In order to see that, note that, within the framework of first-order analysis

$$\vec{a}_j + \Delta \vec{a}_j = e^{i\mathbf{R}_0(j)} \vec{a}_j + i \sum_{n \neq 0} \mathbf{R}_n(j) \vec{a}_{j-n} \quad (21)$$

where $\exp[i\mathbf{R}_0(j)]$ is a unitary matrix, and thereby, it only produces a rotation of polarization and phase.

An important point here is that the mean value of the PPRN matrix is

$$\mathbb{E}[\mathbf{R}_0(j)] = \mathbf{I} \frac{4\gamma}{3} \sum_{k \neq 0;l} X_{l-j, l-j, 0}^{(k)} = \Delta \bar{\theta} \mathbf{I} \quad (22)$$

which is responsible for the presence of a deterministic phase shift, which is immaterial to the process of detection and needs to be removed in the evaluation of the NLIN variance. Hence, the NLIN variance is obtained in a straightforward, albeit rather tedious procedure, yielding the following result for the case where all channels are characterized by the same average input power P and modulation format [73], [84]:

$$\sigma_{\text{NLIN}}^2 = \mathbb{E} \left[|\Delta \vec{a}_j - i \Delta \bar{\theta} \vec{a}_j|^2 \right] = \xi P^3 \quad (23)$$

where

$$\xi = \chi_1 + \chi_2 \left(\frac{\mathbb{E}[|b|^4]}{\mathbb{E}[|b|^2]^2} - 2 \right) \quad (24)$$

with b representing a single-polarization constellation point of an IC, and where the averaging is performed over the data constellation points. The first term on the right-hand side of (23) coincides with the GN model result [81], whereas the second term, which was first reported in [73], contains an explicit dependence on the modulation format, and it vanishes in the case of Gaussian modulation, where $\mathbb{E}[|b|^4] = 2\mathbb{E}[|b|^2]^2$. Another cumbersome, but straightforward calculation, produces the results

$$\chi_1 = \frac{32}{27} \gamma^2 \sum_k \sum_{l,m,n} |X_{l,m,n}^{(k)}|^2 \quad (25)$$

$$\chi_2 = \frac{80}{81} \gamma^2 \sum_k \sum_{l,m} |X_{l,m,m}^{(k)}|^2 \quad (26)$$

whose computation is discussed in [84].

An additional important feature of the NLIN that will be useful in the section that follows is that its components are uncorrelated with each other and, as should be expected from symmetry

$$\mathbb{E}[\Delta \vec{a}_j \Delta \vec{a}_j^\dagger] = \frac{1}{2} E [|\Delta \vec{a}_j|^2] \mathbf{I} \quad (27)$$

Similarly, no cross-polarization correlations exist between the NLIN $\Delta \vec{a}_j$ and the signal \vec{a}_j , as can be seen by evaluating their correlation matrix using (20) and (22)

$$\mathbb{E}[\Delta \vec{a}_j \vec{a}_j^\dagger] = i \mathbb{E}[|a|^2] \Delta \bar{\theta} \mathbf{I} \quad (28)$$

where $\mathbb{E}[|a|^2] = PT/2$ is the mean symbol energy per polarization.

V. CAPACITY LOWER BOUND

We are now ready to deploy the formalism established in Section II-A1 in order to obtain a lower bound for the capacity of the fiber-optic channel. To this end, consistently with (20), we express the input–output relation of our channel as

$$\hat{\vec{a}}_j = \vec{a}_j + \Delta \vec{a}_j + \vec{n}_{j,\text{ASE}} \quad (29)$$

where $\hat{\vec{a}}_j$ and \vec{a}_j are associated with the quantities denoted by Y_k and X_k in Section II-A1, respectively. We now wish to establish an auxiliary additive GN channel

$$\hat{\vec{a}}_j = c \vec{a}_j + \vec{n}_j \quad (30)$$

whose capacity constitutes a lower bound for that of the actual channel. Formally, one may argue that, since $\hat{\vec{a}}_j$ and \vec{a}_j are vectors, the arguments presented in Section II-A1 should be generalized to the case of vector symbols, possibly allowing the scaling coefficient c to be a matrix. However, such generalization is unnecessary in our case, owing to the fact that there is no correlation between the polarization components of the vectors involved, and they have identical variances, as implied by (27) and (28). This yields

$$c = \frac{\mathbb{E}[\vec{a}_j^\dagger \hat{\vec{a}}_j]}{\mathbb{E}[\vec{a}_j^\dagger \vec{a}_j]} = 1 + i \Delta \bar{\theta} \quad (31)$$

so that

$$C \geq 2 \log_2 \left[1 + \frac{(1 + \Delta \bar{\theta}^2) \mathbb{E}[|\vec{a}_j|^2]}{\mathbb{E}[|\Delta \vec{a}_j - i \Delta \bar{\theta} \vec{a}_j + \vec{n}_{j,\text{ASE}}|^2]} \right] \quad (32)$$

Since the ASE noise $\vec{n}_{j,\text{ASE}}$ is statistically independent of \vec{a}_j and $\Delta \vec{a}_j$, and using (23), it is evident that the denominator in the square parentheses of (32) becomes

$$\mathbb{E}[|\Delta \vec{a}_j - i \Delta \bar{\theta} \vec{a}_j + \vec{n}_{j,\text{ASE}}|^2] = \sigma_{\text{NLIN}}^2 + \sigma_{\text{ASE}}^2 \quad (33)$$

with $\sigma_{\text{ASE}}^2 = \mathbb{E}[|\vec{n}_{j,\text{ASE}}|^2]$. Moreover, within the perturbation analysis considered here, the term $\Delta \bar{\theta}^2$ in the numerator should be neglected so that the capacity lower bound simplifies to the form

$$C \geq 2 \log_2(1 + \text{SNR}_{\text{eff}}) \quad (34)$$

where

$$\text{SNR}_{\text{eff}} = \frac{P}{\sigma_{\text{ASE}}^2 + \sigma_{\text{NLIN}}^2} = \frac{P}{\sigma_{\text{ASE}}^2 + \xi P^3} \quad (35)$$

is referred to as the effective SNR.

An important point that must be stressed at this stage is that, although the concept of effective SNR is meaningful in general, i.e., regardless of the deployed modulation

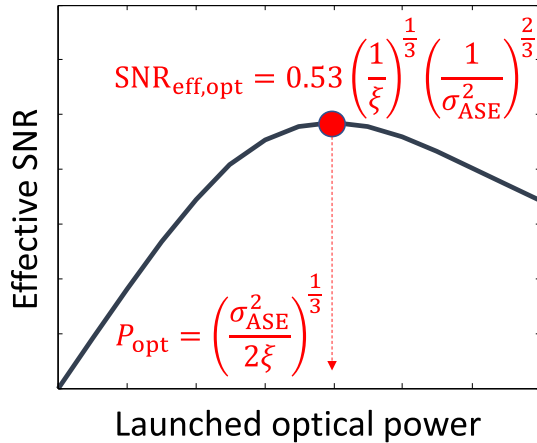


Fig. 3. Effective SNR of (35) as a function of the launched average optical power P . The effective SNR peaks at the launch power denoted by P_{opt} and reaches the value $\text{SNR}_{\text{eff,opt}}$.

format, its use for extracting a capacity lower bound according to (34) is legitimate only when the modulation is Gaussian, in which case $\xi = \chi_1$. The reason for this is that the right-hand side of (34) constitutes the capacity of the auxiliary AWGN channel only when the input symbols \bar{a}_j are Gaussian distributed.

As is evident from (34), the capacity lower bound is determined exclusively by the effective SNR, and therefore, in what follows, we concentrate on characterizing this quantity. In particular, in what follows, we will discuss ways of exploiting correlations that are present within the NLIN, so as to reduce ξ (or χ_1 , as we are using Gaussian modulation) and increase the SNR_{eff} . Before that, it is instructive to examine the direct consequences of (35), which are instrumental for setting one's expectations regarding the capacity.

Fig. 3 shows a plot of the effective SNR of (35), as a function of the launched average optical power P . The peak value of the effective SNR and the power at which it is obtained is indicated in the figure. A critical feature of the peak effective SNR is that it is proportional to $\xi^{-1/3}$, which implies that ξ (and equivalently the NLIN power) needs to be reduced considerably in order to produce a noticeable effect on SNR_{eff} [89], [90]. For example, reducing the NLIN by 3 dB (namely by a factor of 2) results in merely 1 dB of effective SNR improvement. The corresponding increase in the capacity lower bound would be 0.35 bit/s/Hz per single polarization in the high SNR regime (i.e., when $\text{SNR}_{\text{eff}} \gg 1$). As is demonstrated in what follows, a 3-dB reduction of ξ on the basis of correlations present in the NLIN process turns out to be highly nontrivial.

We now move to evaluate the potential effective-SNR improvement that can be achieved by taking advantage of correlations present in the NLIN. As investigated in [77], this improvement can be achieved by means of an adaptive equalizer that takes advantage of the slow

temporal evolution of the matrices $\mathbf{R}_n(j)$, so as to learn the channel and mitigate the effects of ISI. The improvement in performance due to mitigation obviously depends on the temporal correlation properties of the $\mathbf{R}_n(j)$ matrices, which are, in turn, dependent upon the system design and most importantly on the dispersion coefficient, frequency separation between channels, baud rate, and the loss/amplification profile. In addition, as can be seen from [77, Fig. 3], the characteristic correlation time is different for different ISI matrices. In particular, the PPRN matrix $\mathbf{R}_0(j)$ makes the largest contribution to NLIN, and therefore, its mitigation is the most beneficial. Luckily, it is also characterized by the longest correlation time of all ISI matrices, and hence, its equalization is less demanding in terms of the required equalizer adaptation speed than the equalization of the other ISI terms. For example, in the case studied in [77], the correlation length of $\mathbf{R}_0(j)$ is approximately 150 symbols, whereas that of $\mathbf{R}_{\pm 1}(j)$ and $\mathbf{R}_{\pm 2}(j)$ is approximately one order of magnitude smaller.

The potential effect of ISI compensation can be bounded by removing the corresponding terms in the ISI description of NLIN [see (20)]. To illustrate this, we consider an example of a ten-span link over standard SMF (loss of 0.2 dB/km, $\beta_2 = -21$ ps²/km, and $\gamma = 1.3$ W⁻¹ km⁻¹), with a span length of 100 km, assuming a noise figure of 5 dB for the inline EDFAs. The system consisted of 51 WDM Gaussian-modulated channels spaced by 80 GHz with ideal Nyquist pulses, at a symbol rate of 75 Gbaud. In Fig. 4, we plot the effective SNR as a function of the launched average power per channel. The thick curve shows the effective SNR, including all ISI terms, the dashed curve shows the effect of removing the PPRN term proportional to \mathbf{R}_0 , the dotted curve shows the effect of removing also the terms proportional to $\mathbf{R}_{\pm 1}$, and the dashed-dotted curve shows the case in which also the terms proportional to $\mathbf{R}_{\pm 2}$ are removed. The gain observed by simply removing the PPRN term is approximately 1 dB, and it increases to 1.8 dB by removing the other terms listed above. Equation (34) implies that the capacity lower bound may be increased accordingly by ~ 0.6 and ~ 1.1 bit/s/Hz in the relevant regime of high SNR, respectively. Mitigation of higher ISI orders would improve this result, yet addressing them via adaptive equalization is extremely challenging, given the shortening of the correlation times with the ISI order [77].

VI. CAPACITY UPPER BOUND

An important result that was demonstrated in [65] and [66] is that the capacity of a nonlinear system is upper bounded by that of a linear AWGN system operating with the same average input power, i.e.,

$$C \leq 2 \log_2(1 + \text{SNR}_{\text{lin}}) \quad (36)$$

with $\text{SNR}_{\text{lin}} = P/\sigma_{\text{ASE}}^2$. For the optimal power P_{opt} , which is expressed in Fig. 3, the ratio between the two SNRs is

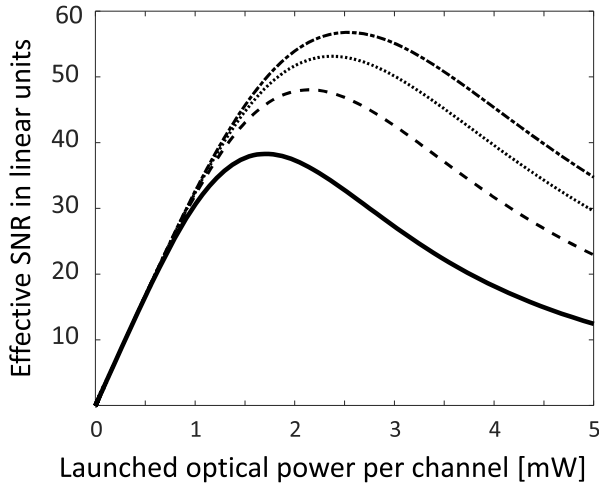


Fig. 4. Effective SNR as a function of the launched average power per channel. The thick curve includes all ISI terms, the dashed curve shows the effect of removing the term proportional to R_0 (PPRN), the dotted curve shows the effect of removing also the terms proportional to $R_{\pm 1}$, and the dashed-dotted curve shows the case in which also the terms proportional to $R_{\pm 2}$ are removed.

given by

$$\frac{\text{SNR}_{\text{lin}}}{\text{SNR}_{\text{eff,opt}}} = \frac{1}{2^{1/3} \times 0.53} \simeq 1.5 \quad (37)$$

so that, in cases where the SNR is much greater than 1, the difference between the upper and lower capacity bounds is

$$\Delta C \simeq 2 \log_2 \left(\frac{\text{SNR}_{\text{lin}}}{\text{SNR}_{\text{eff,opt}}} \right) \simeq 1.17 \text{ bit/s/Hz}. \quad (38)$$

Notice that this difference is a fixed number, independent of any system parameter or mode of operation, as long as the first-order perturbation analysis remains valid, and the SNR is sufficiently large. Intuitively, this result can be attributed to the fact that improving the tolerance of a system to nonlinear effects allows the transmission of higher signal powers, so that both the lower and upper capacity bounds increase by the same amount.

VII. POLARIZATION EFFECTS

Throughout the analysis, we have not accounted for the effect of polarization-related phenomena. In SMFs, these are the phenomena of PMD [18], [91], [92] and PDL [93], [94], [95], [96], which we treat separately in what follows.

A. Polarization-Mode Dispersion

PMD results from structural imperfections in the fiber that perturbs its circular symmetry, thereby causing the polarization of different frequency components of the light in the fiber to evolve differently in the process of propagation. Since PMD is a strictly unitary phenomenon (polarizations rotate as a result of it, but they do not decay), fundamentally, it does not have any effect on the channel capacity in the linear propagation regime. Even in practical

contexts, with the PMD levels that characterize modern SMFs, digital equalizers that are commonly deployed in fiber-optic receivers are very effective in mitigating the effects of PMD in the case of linear propagation, particularly in view of the slow time dynamics of PMD [97], [98], [99], [100], [101], [102], [103], [104], [105], compared to available signal processing speeds.

The effect of PMD becomes much more involved when its interplay with NLIN needs to be taken into account. It is particularly complicated in the highly nonlinear regime [106], [107], where very little can be said about it. Yet, in the regime of pseudolinear transmission, which characterizes the majority of fiber-communications' systems today, the effect of PMD on NLIN appears to be negligible, as has been pointed out through extensive numerical simulations in [108]. Some physical intuition for this can be gained by considering the process through which NLIN is created. Fundamentally, interchannel NLIN is formed when pulses representing symbols in different WDM channels overlap with one another in time. Owing to the different propagation velocities of different WDM channels in the presence of chromatic dispersion, the temporal overlap between given pulses is maintained over a limited-length section in the fiber, to which we refer as collision length [86]. When PMD scrambles the relative states of polarization of the nonlinearly interfering pulses significantly within the collision length, the effectiveness of the nonlinear process and, consequently, the magnitude of NLIN reduce. The amount of PMD that is required in order to produce this result can be assessed based on the polarization autocorrelation function between two tones with a frequency separation Ω [109], [110], [111]

$$\mathbb{E} \left[\hat{E}_1 \cdot \hat{E}_2 \right] = e^{-\frac{1}{8} \Omega^2 \kappa_{\text{PMD}}^2 l} \quad (39)$$

where κ is a PMD coefficient and l is the distance that the signals travel in the fiber. For the relative polarization to be significantly scrambled, the propagation length needs to satisfy $l \gg 8 \Omega^{-2} \kappa_{\text{PMD}}^{-2}$. In modern SMFs, κ_{PMD} is of the order of 0.05 ps/ $\sqrt{\text{km}}$, or less, implying that, for channel separations ranging between 50 and 500 GHz, l ranges between more than 30 000 km and 300 km. These are values that exceed typical collision lengths in relevant fiber communication settings. In the case of farther separated channels, l shortens, and the suppression of NLIN may become more significant, but the combined contribution of all such channels to NLIN is typically negligible. As we show in Section X-B, in the case of fibers used in spatially multiplexed transmission, the effect of modal dispersion (MD) (which generalizes PMD) on the NLIN becomes considerable.

B. Polarization-Dependent Loss

Polarization-dependent loss is another linear polarization phenomenon. It describes a situation in which the attenuation that is experienced by a signal propagating

through the fiber depends on its state of polarization [93]. Since, unlike PMD, PDL is not a unitary effect, it implies a fundamental reduction in the channel capacity, as has been studied in [94] and [112].

In modern fiber-communication systems, PDL turns out to be dominated by lumped inline optical elements, such as optical amplifiers, much more than by the transmission fiber itself.

Each PDL element has two orthogonal polarization axes that apply different levels of attenuation to the propagating signal. Since the PDL elements are separated by sections of fiber that randomly rotate the polarization of the signal propagating between them, the PDL axes of different elements can be viewed as randomly rotated relative to one another [93], [113]. The effect of PDL on capacity must be considered separately for the linear and nonlinear propagation regimes. In the regime of linear propagation, PDL leads to the distortion of the propagating waveform, and it also randomizes the received SNR. It is interesting that, in what concerns the quality of communications, the effect of waveform distortion turns out to be totally negligible relative to the effect of SNR randomization, as has been demonstrated in [95]. The randomization of the SNR occurs because the signal does not experience the same PDL as the ASE noise. The signal passes through all of the PDL elements, whereas the noise that consists of ASE contributions produced at the various amplification sites only sees the PDL of elements that are present after the point at which it is generated.⁹ This difference introduces randomness into the SNR and, thereby, affects the communication capacity of the system. Since the PDL-induced variations of the SNR are very slow, it is appropriate to treat the channel capacity as random and assess the penalty due to PDL in terms of a power margin that needs to be allocated in order to prevent the system capacity from reducing below the desired value with a prescribed outage probability [115].

As may be expected, the assessment of the impact of PDL becomes much more difficult in the regime of nonlinear propagation. Randomly varying signal powers randomize the generation of NLIN and affect its statistics in a rather nontrivial manner. This problem has been addressed in a recent study by Serena *et al.* [116], where the GN model is extended so as to include the effects of PDL.

VIII. CONCLUDING REMARKS ON LONG-HAUL SYSTEMS CAPACITY

Since, fundamentally, the regime of long-haul transmission is limited by nonlinear propagation phenomena, a rigorous capacity assessment is unlikely to be achievable in any practical setting.

It seems that the only known tractable analyses are the ones relying on the first-order perturbation approach. Within this framework, one can treat the effect of nonlin-

ear propagation in terms of NLIN and obtain a useful lower bound for the fiber channel capacity. This lower bound is the capacity of an AWGN channel, where the noise power is the sum of the amplification noise power and the NLIN power obtained from the simplest GN model [81]. As discussed in Section VI, this lower bound is always 1.17 bit/s/Hz below the capacity upper bound in (36), which is the capacity of a linear system operating at the same optical power that is optimal in the nonlinear system [65], [66].

Since the nature of NLIN is similar to that of slowly varying ISI, the NLIN power can be reduced by means of ISI mitigation techniques that leverage the temporal correlations existing in the NLIN process. The potential benefit of ISI compensation depends on the specific system settings, and in a characteristic example considered in Section V, ideal zeroth-order ISI (PPRN) mitigation was shown to increase the capacity lower bound by up to 0.6 bit/s/Hz. An increase of ~ 1.1 bit/s/Hz was shown to be achievable when the five lowest ISI orders are mitigated.

Finally, it should be noted that the modeling of the NLIN reviewed in this article relies on the use of the Manakov equation, which is valid for transmission bandwidths of the order of a few terahertz, covering the case of C-band systems. However, the NLIN analysis becomes more involved in the case where high-throughput transmission is pursued by extending the transmission bandwidth beyond the conventional C-band. In this regime, the Manakov equation needs to be supplemented with additional terms accounting for stimulated Raman scattering [61], which manifests itself in the form of power transfer from high-frequency channels to low-frequency channels. The extent of this power transfer increases with the frequency separation between the ICs, peaking at a separation of approximately 13 THz. As a result, in ultrawideband systems, the Raman scattering modifies the z -dependence of the individual channels' powers, thereby modifying the formation of the NLIN. The extension of the NLIN model accounting for Raman scattering is discussed in [117], [118], and [119].

IX. DIRECT-DETECTION SYSTEMS

In short optical links, system design is usually dominated by cost-related considerations, in which case the assessment of information capacity can only be meaningful when accounting for specific cost and complexity constraints. In what follows, we outline the consideration of the capacity in systems constrained to the use of direct detection, which, in short-reach transmission, is typically preferred over the more costly alternative of coherent detection. This is often the situation in the case of data-center interconnecting links, which constitutes an increasingly significant fraction of the global fiber-optic network. Typically, data-center interconnects (DCIs) extend over a few tens to 100 km. Like in coherent systems, noise is dominated by optical amplification, but the noise in the received power waveform is neither additive nor Gaussian, and hence, the

⁹It is important to stress that, by ignoring the effects of PDL on the ASE noise, as is sometimes done in the literature, one overestimates the impact of PDL on the system performance [114].

capacity expression (4) does not apply. Interestingly, the actual capacity is only 1 bit lower than that of a coherent system [38], [120], provided that the bandwidth of the intensity receiver is twice as large as that of the coherent receiver,¹⁰ and electrical receiver noise is negligible. Otherwise, if the bandwidth is the same as that of the coherent system, the capacity reduces to 1 bit less than the capacity of a single quadrature, as shown in [33]. This latter result can be intuitively understood by noting that intensity constitutes a single degree of freedom (comparable to just one of the two quadratures of the electric field). In addition, since intensity detection does not allow distinguishing between positive and negative signal values, an additional 1 bit of information is lost.

The old generation of intensity-modulated direct-detection systems used one sample per symbol, and therefore, in the high-SNR limit at which they are typically operated, their capacity was 1 bit less than the capacity of a single-quadrature AWGN channel of (3), as demonstrated in [33]. In modern systems, direct detection is used for the purpose of reducing the optical complexity of receivers, typically in short links [121] (albeit not exclusively [122]), and it is complemented with digital processing in the electrical domain. These implementations are commonly referred to as advanced direct-detection schemes. In these schemes, it is possible for the receiver bandwidth to be consistent with the bandwidth of the intensity waveform (i.e., twice the optical bandwidth), in which case the capacity is only 1 bit lower than that of a coherent channel [38].

Most advanced direct-detection schemes involve the transmission of a continuous-wave (CW) carrier A together with the information carrying signal $s(t)$ accompanied by Gaussian amplification noise, which we ignore for simplicity of illustration. Such systems are frequently referred to as *self-coherent*. The photocurrent from which the data are to be extracted is given by

$$I_{PD}(t) = \eta [A^2 + 2A\text{Re}\{s(t)\} + |s(t)|^2] \quad (40)$$

where η is a proportionality coefficient accounting for the quantum efficiency of the detector [123], and where we have assumed with no loss of generality that A is real-valued. It is clear that the information resides exclusively in the second and third terms on the right-hand side of (40), whereas the first term A^2 involves only the carrier and is, therefore, immaterial to the data-recovery process. When $|A| \gg |s(t)|$, the third term $|s(t)|^2$ becomes negligible, and $\text{Re}\{s(t)\}$ is readily extracted from I_{PD} . If $s(t)$ is a single-sideband signal (i.e., its spectrum resides only on one side of the CW carrier), then knowledge of $\text{Re}\{s(t)\}$ implies knowledge of the entire signal $s(t)$ because the real and imaginary parts are related by the Hilbert transform.

¹⁰This result may appear as counterintuitive at first. Notice that, since the bandwidth of the intensity is twice larger than that of the signal, there is a similar number of degrees of freedom in the intensity and the complex field. This point is detailed in [38].

This is contrary to a situation in which the spectrum of $s(t)$ is centered at the carrier frequency, in which case only the real part of $s(t)$ can be exploited for communicating, thereby halving the SE.¹¹ The main issue with the above approach is that the neglect of $|s(t)|^2$ often requires an impractically high carrier power, whereas, for practical values of A , $|s(t)|^2$ is nonnegligible, and it is referred to as signal-to-signal beat noise (SSBN). Many studies conducted on advanced direct-detection systems focused on deploying DSP algorithms that increase the tolerance of the system to SSBN and, thereby, allow operation with lower carrier powers [124], [125], [126], [127], [128], [129], [130], [131], [132]. Many of these studies relied on iterative approaches, where, in each iteration, the SSBN is estimated and subtracted in the iteration that follows.

An alternative scheme that has been reported more recently is that of the Kramers–Kronig (KK) receiver [133]. This scheme allows the exact reconstruction of the received complex-valued optical signal, provided that the information carrying signal is a single sideband with respect to the carrier, and the carrier amplitude A exceeds a well-defined finite value, which ensures that the entire waveform (signal and carrier) is the minimum phase. Numerical and experimental demonstrations using the KK receiver have shown that excellent performance is achieved with the carrier-to-signal power ratios exceeding the order of 6 dB [134], [135] (notably lower than that required by previously studied advanced direct-detection schemes). A number of variants of the KK receiver have been proposed and demonstrated in recent years, as reviewed in [136].

One major limitation of direct-detection schemes is in the fact that they do not easily accommodate polarization multiplexing. Various schemes to overcome this shortcoming have been proposed, the most prominent of which are the ones based on the Stokes-space receiver [137] (sometimes in combination with the KK algorithm [138], [139]) or carrier-assisted differential detection [140]. These schemes have been used to achieve record data transmission rates with polarization multiplexing [141], albeit at the expense of significantly higher optical complexity.

For completeness, another category of direct-detection systems that should be mentioned in this section is the one in which the transmitted waveform is obtained by directly modulating a CW laser. In this case, the receiver consists of a single photodiode without optical chromatic dispersion compensation and without DSP. These systems are attractive for their very low cost, robustness, and low power consumption compared to all other alternatives. Because of the low-cost constraints, such systems resort to the use of temperature-unstabilized free-running lasers. Since the central optical frequency of such lasers wanders by hundreds of GHz, it is not meaningful to refer to the

¹¹Notice that the SE of the two schemes is the same if one refers it to the required electrical bandwidth because photodetection doubles the bandwidth of only the single-sideband signal when the term $|s(t)|^2$ is negligible.

system information capacity or SE. The performance of such systems is limited primarily by the fiber chromatic dispersion, which, in combination with the direct-detection process, corrupts the received intensity waveform to the extent that the information cannot be recovered. Indeed, for a modulated power waveform of the kind $P(t) = P_0[1 + m(t)]$, where $|m(t)| \ll 1$, the transfer function experienced by $m(t)$ after propagation in a fiber link of length L is given by [142]

$$H(\omega) = \cos^2\left(\frac{1}{2}\beta_2\omega^2L\right) \quad (41)$$

which has notches at $\omega^2 = (1 + 2k)\pi L^{-1}|\beta_2|^{-1}$, $k = 0, 1, 2, \dots$. For this reason, it is very challenging to transmit signals $m(t)$ with a bandwidth of 20 GHz over more than 10 km of standard SMF.

X. SPACE-DIVISION MULTIPLEXED TRANSMISSION

As hero experiments reported by the end of the first decade of the present millennium appeared to be gradually closing the gap to early estimates of the SMF channel capacity (see [143] and references therein), SDM came into the spotlight of optical communications, as a candidate for enabling sustainable scaling of fiber-optic communication rates [1], [144]. It is implemented by taking advantage of fiber structures supporting the propagation of multiple spatial modes, where all the modes are used for the transmission of information. Depending on the specific fiber design, spatial modes can be coupled or uncoupled. For example, in a multicore fiber where the cores are sufficiently far apart, the coupling between them is negligible, and each constitutes an independent information channel. On the other hand, in multimode fibers or in multicore fibers containing a larger number of cores (so that the intercore distances are smaller), the spatial modes couple to each other during propagation. In this situation, the extraction of the information at the receiver becomes more complicated, as it requires the use of multiple-input–multiple-output (MIMO) techniques, but, at the same time, the nonlinear distortions discussed in Section IV are reduced relative to the uncoupled case (or equivalently, relative to in the case of separate SMFs), as we discuss in what follows.

In this section, we review some of the main implications of the use of SDM transmission on channel capacity. We use N to denote the number of spatial modes that are supported by the fiber so that the total number of modes M used in (4) is equal to $M = 2N$, where the factor of two accounts for the fact that each spatial mode is twofold degenerate with respect to polarization.

A. Uncoupled-Mode SDM Transmission in Power-Limited Links

In transoceanic transmission, the rules of the game are different from what they are in terrestrial transmission. One of the biggest issues is with the power supply to

the undersea amplifiers. The use of multiple spatial paths introduces a degree of freedom that can be taken advantage of in order to increase the power efficiency, which is defined as

$$\eta_p = \frac{C}{P_{\text{tot}}} \quad (42)$$

where C is the capacity and P_{tot} is the total optical power that needs to be fed into the system

$$P_{\text{tot}} = NN_s P_{\text{out}} \left(1 - \frac{1}{G}\right) \simeq NN_s P_{\text{out}} \quad (43)$$

where P_{out} is the output power of the amplifiers, under the assumption that they are operated in a saturation mode, and the second equality holds for $G \gg 1$, as happens in most cases of interest. As shown by Sinkin *et al.* [145], the power efficiency, under the assumption of linear transmission, is given by the expression

$$\begin{aligned} \eta_p &= \frac{2N \log_2(1 + \text{SNR})}{NN_s P_{\text{out}}} \\ &= \frac{2}{N_s P_{\text{ASE}}} \frac{\log_2(1 + \text{SNR})}{1 + \text{SNR}} \end{aligned} \quad (44)$$

where P_{ASE} is the noise power per mode at the output of the last amplifier, and we have used the fact that $P_{\text{out}} = P_{\text{ASE}} + P_{\text{sig}}$, with P_{sig} denoting the signal launch power in each spatial mode ($P_{\text{sig}}/2$ in each polarization). The power efficiency η_p can be seen to be largest when the SNR is equal to $e - 1 \simeq 1.72$. The relation

$$\text{SNR} = \frac{P_{\text{tot}}/(N_s N) - P_{\text{ASE}}}{P_{\text{ASE}}} \quad (45)$$

allows the extraction of the optimal number of spatial modes for a prescribed available total optical power and the number of spans. Notice that η_p is the largest for a relatively low SNR value, which suggests that the assumption of linear transmission is justified in this case. This observation is in contrast to traditional single-mode transoceanic systems that were designed with the goal of optimizing the system capacity, and hence, they are operated in the nonlinear regime.

Since the approach of [145] focuses only on the optimization of power efficiency, other factors that play an important role in the design, operation, and cost of transoceanic systems are not taken into account. These include the costs of deployment, cable and fiber, amplifiers, and transponders. Some of these aspects are taken into account in [146] and [147]. Additional aspects having to do with nonlinear propagation, as well as with the physics and operation of the EDFAs, are addressed in [72], [148], and [149].

B. Coupled-Mode SDM Transmission

An important aspect of transmission in SDM fibers has to do with the effect of spatial mode coupling on the

nonlinear signal distortion and performance impairments. It has been found in [150] and [151] that, as the random mode coupling increases, the nonlinear penalty to performance increases initially and then reduces to the extent that it becomes lower than that characterizing the case of no intermode coupling, thereby outperforming single-mode transmission. This finding encourages intentionally introducing strong random coupling between modes, as means of mitigating nonlinear transmission impairments. The benefit of strong mode coupling has been first reported in [152] on the basis of numerical studies, and the physical mechanism responsible for it has been described in [153], while experimental evidence was later presented in [154] and [155]. The main idea can be summarized by inspecting the equation describing nonlinear propagation in SDM fibers with strong mode mixing, which is known as the multicomponent Manakov equation [156]

$$\frac{\partial \vec{E}}{\partial z} = -\frac{\alpha}{2} \vec{E} - i \frac{\beta}{2} \frac{\partial^2 \vec{E}}{\partial t^2} + i \kappa \gamma |\vec{E}|^2 \vec{E}. \quad (46)$$

The vector \vec{E} appearing in this equation contains $M = 2N$ complex-valued elements so that each element describes the excitation of a specific space and polarization mode. This equation represents an idealized case where all modes are degenerate; namely, they are characterized by the same propagation constant β and the same loss coefficient α . It is identical in form to the Manakov equation of (10), and it is derived from a set of $2N$ coupled nonlinear Schrodinger equations, taking into account the fast and random mixing occurring between modes during propagation. Here too, $|\vec{E}|^2 = \sum_{n=1}^{2N} |E_n|^2$ represents the total optical power, and $|E_n|^2$ is the optical power in the n th mode. The coefficient κ accounts for the mode multiplicity, and in the case of $N = 1$, it is equal to $8/9$, consistently with (10). In the case of a multicore fiber with N strongly coupled cores, κ is given by [153]

$$\kappa = \frac{8}{3} \frac{1}{2N + 1} \simeq \frac{4}{3N} \quad (47)$$

where the approximate equality holds for large core counts. The above implies that the effective SNR scales as $(2N + 1)^{2/3}$, consistently with the discussion in Fig. 3, so that a lower bound for the capacity is given by

$$C \geq N \log_2 \left(1 + \text{SNR}_{\text{eff,opt}}^{(N)} \right) \quad (48)$$

where $\text{SNR}_{\text{eff,opt}}^{(N)} = (2N + 1)^{2/3} \text{SNR}_{\text{eff,opt}}^{(1)}$ with $\text{SNR}_{\text{eff,opt}}^{(N)}$ denoting the optimal SNR in a fiber with N strongly coupled cores. We note, however, that experiments so far have reported a benefit lower than that predicted by (48) when using coupled-core multicore fibers [155]. The reason for this discrepancy requires further investigation and could be attributed to various factors, such as mode-dependent loss (MDL).

To develop an intuition for the inverse dependence on N , we emphasize that the Manakov equation is obtained by averaging the nonlinear terms of the coupled nonlinear Schrödinger equations with respect to the distributed random mode coupling [156]. This is equivalent to averaging the nonlinear terms on the length scale over which random coupling is effective (which is much shorter than the length scale over which the nonlinear effects become appreciable). As the field in each core becomes a random mixture of the fields transmitted in all of the cores, the effect of this averaging is to suppress the coherent (i.e., phase-dependent) products between the field contributions originating from different cores, with the result that the power transmitted in each fiber core equalizes rapidly between all cores. We stress that this mixing occurs continuously along the fiber. Although (47) was derived for multicore fibers, a similar scaling also holds for multimode fibers, provided that all mode groups are strongly mixed over a sufficiently short length scale [153].

The way in which the $1/N$ dependence of κ translates into the observed reduction of the NLIN can be understood as follows. The nonlinear perturbation that is imposed on the electric field in the j th mode ΔE_j is proportional to $\kappa |\vec{E}|^2$ or equivalently to $\sum_n |E_n|^2 / N$. Since the powers $|E_n|^2$ of the individual modes are statistically independent, the variance of ΔE_j is proportional to the sum of their variances divided by N^2 . This implies that the variance of ΔE_j is smaller by a factor of N than it would be in the absence of mode coupling where the $1/N$ dependence of κ is absent. The effective SNR increases accordingly, and the effect on the capacity can then be assessed by considering the lower bound given in (4). Achievable rates obeying the multicomponent Manakov equation are currently under investigation [157]

Another important aspect of SDM transmission is related to MD and MDL, two concepts that generalize PMD and PDL, respectively, to the case of SDM fibers.

MD, just like PMD, is responsible for introducing frequency dependence into the random mode coupling process, and being a unitary phenomenon, in principle, it does not affect the system capacity in the linear transmission regime. Yet, since the average delay spread produced by MD is greater by two or three orders of magnitude than that produced by PMD in the same length of fiber, the increase in receiver complexity that it entails is considerable [158]. The large MD implies that the frequency correlation of mode-related phenomena becomes very short. For example, after propagation in a 1000-km link with a typical average delay spread coefficient of $5 \text{ ps/km}^{1/2}$, the evolutions of the modal content of two different spectral components of the propagating signal separated by 1 GHz are already uncorrelated from each other. This is in contrast to what happens in a typical SMF, where, for the same propagation distance, the polarization states of two spectral components become similarly decorrelated only when the frequency separation between them exceeds the

order of 1 THz. For this reason, the effect of MD on the propagating signal is averaged over the signal spectrum, with the result that the output intensity waveform can be expressed as a convolution of the input intensity waveform with a deterministic impulse response—the *intensity impulse response* of the fiber [159], [160], [161], [162].

In the regime of nonlinear propagation, on the other hand, MD can be beneficial as it contributes to further reducing the accumulation of NLIN [163], [164]. Indeed, while, as noted in Section VII-A, the relative polarization rotation of nonlinearly interfering frequency channels is too small to affect the build-up of nonlinearity on the length scale of pulse collisions in SMFs, the equivalent effect of MD is much larger, and the build-up of the nonlinear interference is effectively reduced. This reduction comes on top of the above-discussed NLIN reduction due to strong mode mixing.

We conclude this section by briefly discussing the effect of MDL on the capacity of SDM systems. Like PDL in single-mode systems, MDL is a nonunitary phenomenon, and hence, it is responsible for a fundamental reduction in the information capacity. This reduction can be quantified either in terms of the ratio between the actual capacity and some reference capacity [21], [165] or, in terms of the difference between the two, normalized to the total number of modes [114], [166]. The latter form has been shown to be more convenient in the high-SNR limit, where its average value and variance become independent of the SNR itself and the number of modes. In general, we note that the modeling of how MDL affects the system capacity is rather involved, and there are subtleties that require special attention. These include the choice of a reference system with respect to which the capacity loss is assessed, the operation mode of the amplifiers, and the way in which noise loading is described. It should be noted, however, that, because of the fast frequency decorrelation imposed by MD discussed earlier in this section, the fluctuations of capacity induced by the randomness of MDL are effectively suppressed, and the relevant figure of merit becomes either the average capacity ratio or the average capacity loss per mode. As a reference, the average capacity loss per mode ranges between 0.1 and 0.8 bit/s/Hz when the mean link MDL ranges between 5 and 20 dB, respectively [114].

XI. STATE OF THE ART IN THE FIELD AND IN LABORATORY EXPERIMENTS

In this section, we review here fiber-transmission experiments. In the first part, we focus on systems ranging from metro to transoceanic distances [167], [168], [169], [170], [171], [172], [173], [174], [175], [176], [177], [178], [179], [180], [181], [182], [183], [184], [185], [186], [187], [188], [189], [190], [191], [192], [193], [194], [195], [196], [197], [198], [199], [200], [201], [202], [203], [204], [205], [206], whereas, in the second part, we review shorter links where high-speed transmission happens primarily for datacenter

interconnect applications [207], [208], [209], [210], [211], [212], [213].

Since the various experiments were performed with different system settings (fiber type, span length, amplification type, and so on), a meaningful comparison between their results is not always straightforward or even possible. Nonetheless, displaying the results of these record-achieving experiments on a single plot, as we do in Fig. 5, provides valuable insight as to the trends relating between system reach and its SE. In this context, we use the term SE in order to refer to the ratio between the demonstrated information throughput per spatial mode and the optical bandwidth that is allocated for its transmission. The data points in Fig. 5 represent experimental results reported since 2012. The various experiments differ in terms of the transmitted number of WDM channels, starting from five or ten channels occupying a few hundred GHz to fully loaded C-band ($\sim 1530\text{--}\sim 1565$ nm) (e.g., [170], [172], and [176]), C + L band ($\sim 1530\text{--}\sim 1625$ nm) (e.g., [169], [181], and [186]), or even S + C + L band ($\sim 1460\text{--}\sim 1625$ nm) transmission [173], [200]. The fiber types vary between standard single-mode (e.g., [169]), large-effective-area (e.g., [182]), and SDM fibers (e.g., [204]). The amplification schemes vary between EDFA-only (e.g., [180] and [181]), hybrid Raman-EDFA (e.g., [171] and [185]), and ultrawideband semiconductor optical amplifiers [173]. Large differences also characterize the receiver-side DSP, which varies from conventional linear adaptive equalization (e.g., [168]) to complex nonlinearity-compensation processing (e.g., [172], [176], [180], and [181]).

As seen in Fig. 5, the SE drops rather quickly with the transmission distance. This reduction is consistent with the theoretical bounds discussed in Section V. The dashed curves represent the lower bound $2\log_2(1 + \text{SNR}_{\text{eff,opt}})$, where SNR_{eff} is the effective SNR given by (35), which accounts for all NLIN contributions [82] (intrachannel and interchannel) in the absence of nonlinearity compensation, and it is evaluated at the optimal launch power under the assumption of Gaussian modulation, where $\text{SNR}_{\text{eff}} = \text{SNR}_{\text{eff,opt}}$.^{12,13} While the data points shown in the figure were obtained for a variety of system settings, the lower bound curves refer to a system with 50 tightly packed 32-Gbaud WDM channels transmitted over an SMF with a loss coefficient of 0.15 dB/km, $\beta = 21$ ps²/km, and $\gamma = 1.3$ W⁻¹ km⁻¹. The solid curves represent the capacity upper bounds for the two considered span lengths, and as

¹²As noted in the context of (23), the assumption of Gaussian modulation implies that $\text{SNR}_{\text{eff,opt}}$ coincides with the optimal effective SNR obtained with the GN model.

¹³The lower bounds displayed in the figure were obtained under the assumption of SMF transmission. The corresponding lower bound for SDM systems should account for the dependence of the SNR on the number of modes. Depending on the degree of mode coupling, the SDM capacity lower bound may be higher or lower than the corresponding bound in the case of SMF. In the limit of strong coupling, the lower bound would be higher, consistently with (48).

explained in Section VI, they are 1.17 bit/s/Hz higher than their corresponding lower bounds.

The green curves were obtained for span lengths of 100 km and should be compared with experimental results achieved in the range of terrestrial links (roughly up to 2000 km). Transoceanic experiments (above 2000 km), where short spans were used, should be compared with the red curves, which were obtained for span lengths of 50 km. The inset in Fig. 5 shows the effect of ideal PPRN compensation on the capacity lower bounds. Consistently with [86], the effect of PPRN mitigation is more significant in short-span systems (which are closer to the case of distributed amplification [73]). It is largest in the single-span case but rapidly reduces with the number of spans reaching the order of ~ 1 -bit/s/Hz improvement for the 50-km-span system and ~ 0.5 bit/s/Hz in the system using 100-km spans.

It is interesting to note that the largest SE values in all cases were obtained in SMF transmission systems, which is consistent with the higher maturity of SMF technology.

Because some of the SE data points shown in Fig. 5 were obtained with a small number of channels, it is important to provide a more complete picture showing record results for the aggregate system throughput. Fig. 6 summarizes experimentally demonstrated record throughputs per fiber. We group the demonstrations into four types: 1) SMF transmission with WDM channels fully loaded on the C-band; 2) SMF transmission with the fully loaded extended band (C + L or S + C + L); 3) SDM fiber (multicore fiber and/or few-mode fiber) transmission with fully loaded C-band; and 4) SDM fiber with the fully loaded extended band (C + L or S + C + L). It is evident that, when operating only in the C-band, the throughput in the case of SMF saturates at ~ 41 Tb/s for ~ 500 -km transmission [172] and ~ 35 Tb/s for ~ 6000 -km transmission [176]. Extending the wavelengths to the L-band yields at most a factor of two in the throughput, reaching ~ 70 Tb/s for ~ 7000 -km transmission [181]. Using S-, C-, and L-bands, the achieved link throughput is seen to increase to 74 Tb/s for 6300-km transmission [190] and 115 Tb/s for 100-km transmission [173] (with the S-band emulated by ASE noise). Higher rates have been demonstrated for shorter distances [209].

A further increase in throughput is achieved by using multiple spatial modes. In [193], a 1-Pt/s throughput was demonstrated with a 32-core fiber. Many of the record-achieving experiments relied on the use of advanced modulation formats, with probabilistically shaped (PS) QAM [210] playing a prominent role. This technique produces Gaussian-like amplitude distributions, offering shaping gain and the flexibility for continuously tuning the constellation entropy. An example of using PS-QAM to achieve shaping gain and SE flexibility can be seen in [168], where PS-256-QAM is used and the SE is varied from 12.6 bit/s/Hz for 500-km transmission to 10.1 bit/s/Hz for 2000-km transmission. Other coded modulation schemes, such as 64-APSK [176], are

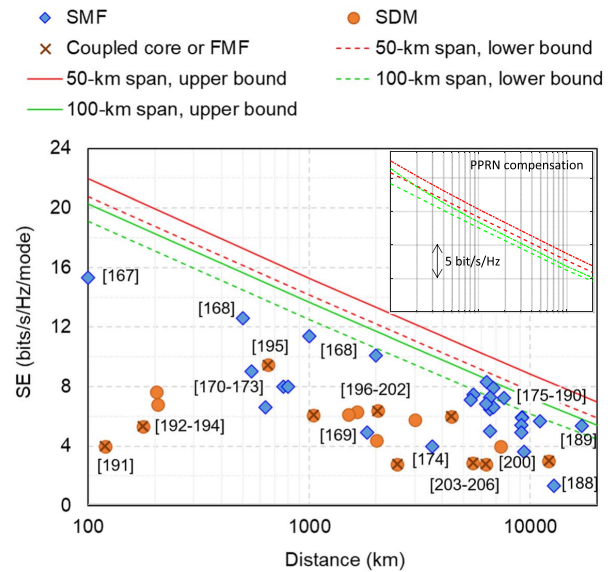


Fig. 5. Summary of record experiment on high SE per spatial mode in WDM transmission experiments. Symbols refer to experiments reported in the displayed references. The dashed curves show the lower bound $2\log_2(1 + \text{SNR}_{\text{eff,opt}})$ in the absence of nonlinearity compensation, and it is evaluated at the optimal launch power under the assumption of Gaussian modulation (details are provided in the main text). The solid curves represent the upper bound $2\log_2(1 + \text{SNR})$ [65], [66], with the SNR evaluated at the same launch power. The green curves were obtained for span lengths of 100 km and should be compared with experimental results achieved in the range of terrestrial links (roughly up to 2000 km). The transoceanic experiments (above 2000 km), which used short spans, should be compared with the red curves, which were obtained for span lengths of 50 km. The dash-dotted curves in the inset show the effect of compensating for PPRN on the lower bounds for the two span lengths. The upper bounds are not shown in the inset.

also used to improve the SE. For capacity-approaching experiments, it is also common to use advanced DSP, such as nonlinear noise compensation (NLC) algorithms. Zhang *et al.* [176] and Cai *et al.* [181] are two examples where DBP is used for nonlinearity compensation. In particular, Zhang *et al.* [176] show that NLC improves the Q factor by ~ 0.8 dB in a fully loaded C-band system with 6375-km transmission, and Cai *et al.* [181] report NLC gain ranging from 1.1 to 1.7 dB for the 295 measured C + L band channels after 7600-km transmission.

Similar to SMF links, advanced modulation formats, NLC, and extended bands are also used in SDM systems. For example, in [200], S-, C-, and L-bands are fully loaded with 552 wavelength channels in each fiber core. The combination of the four-core fiber and the ultrawideband utilization results in 319 Tb/s achieved in a 3001-km link. As mentioned in Section X-B, one of the advantages of SDM fibers is a higher tolerance to nonlinearity. This aspect of SDM transmission is studied systematically through experiments in [201] and [203]. In particular, Ryf *et al.* [201] compare transmission results of QPSK and 16-QAM signals with coupled four-core fiber, coupled seven-core fiber, and

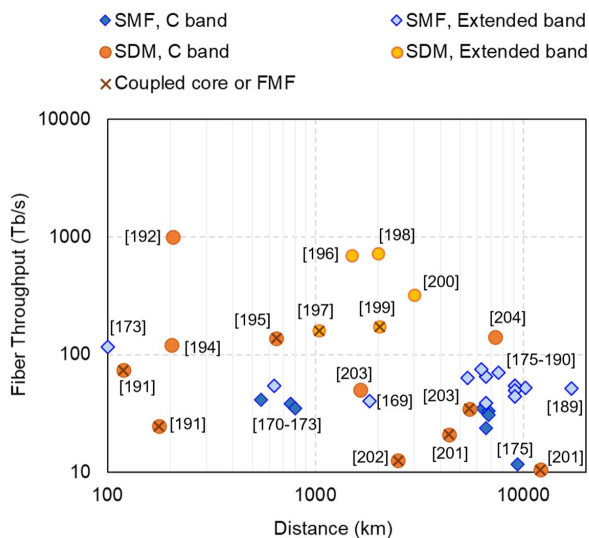


Fig. 6. Summary of record experiment on high-throughput WDM transmission experiments. Most of the data points refer to the same experiments referenced in Fig. 5.

ultralarge effective area SMFs. The coupled four- and seven-core fibers show, as noted earlier, up to 1 dB higher Q factors and higher optimal launch power at all distances ranging from ~ 1000 to $\sim 12\,000$ km. Naturally, SDM fibers exploit the spatial degree of freedom to further increase the link throughput. For example, the transmission of tens of Tb/s has been achieved either by fully loading the C-band of SMF with high order QAM [171], or by transmitting channels with lower order QAM modulation in a few spatial modes fiber [203], or by means of even lower order QAM transmission over an SDM fiber with a large number of spatial modes [206].

We now switch to briefly describing high-speed links extending over distances shorter than 100 km, as would be the case with DCIs. In contrast to long-distance systems, where the link throughput is typically limited by the fiber's capacity, the throughput of short-reach systems is usually limited by cost constraints that are applied to the transceivers. Indeed, commercial short-reach transceivers typically rely on IM and DD (IM-DD) and support rates between 10 and 50 Gb/s per wavelength. The lasers in such modules are usually uncooled, implying that the wavelength drifts significantly with temperature variations, thereby preventing the use of dense WDM (DWDM). Hence, short-link DCIs (e.g., up to several kilometers) typically carry four or eight wavelengths per fiber and provide a link throughput within the range of a few hundreds of Gb/s. IM-DD signals with longer transmission distances suffer from dispersion-induced spectral fading (see Section IX). In practice, IM-DD transmission of more than 50 Gb/s over distances exceeding 10 km is very challenging in the absence of optical dispersion compensation.

Communications between data centers that are separated by more than 10 km require more advanced transceiver configurations. Advanced, self-coherent DD schemes, such as the self-homodyne scheme of [207] and [208], show the real-time demonstration of 600 and 800 Gb/s per wavelength, respectively. KK receiver experiments achieving 104 and 279 Gb/s per wavelength are also reported in [134] and [211], and a Stokes receiver system carrying 186 Gb/s per wavelength is demonstrated in [212]. Combining KK and Stokes receiver techniques, the per-wavelength data rate was shown to achieve 400 Gb/s and 1 Tb/s using a superchannel [213]. Combining DWDM, the potential per-fiber link throughput for self-coherent systems can be tens of Tb/s with transmission distances approaching ~ 100 km. Nonetheless, since self-coherent transmission relies on a new transceiver design that is usually not backward or forward compatible with IM-DD or coherent transmission, they have not yet been adopted in commercial links. Instead, DCIs extending over distances longer than 10 km use simplified coherent transceiver technology. In particular, such transceivers use simpler modulation formats than in long-haul systems (e.g., uniform 16-ary QAM instead of high order PS-QAM) and simpler digital processing algorithms (e.g., simpler adaptive filtering using a much smaller number of taps). Currently, commercially available DCI coherent modules support ~ 60 WDM channels operating at ~ 60 GBaud with 400 Gb/s per channel, corresponding to a total fiber throughput of ~ 24 Tb/s. More details about such links can be found in [214].

XII. CONCLUSION

We reviewed the problem of assessing the information capacity of fiber-communication systems. As the complexity of nonlinear fiber propagation prevents the extraction of information capacity by directly applying Shannon's formulation, most efforts focus on the search after useful bounds that allow its assessment. We reviewed these bounds in a variety of applications, ranging from interdata center links extending over tens of kilometers to submarine systems of thousands of kilometers in length. We also considered a variety of transmission schemes and discussed the potential benefits of spatial multiplexing using multi-mode or multicore fibers. Finally, state-of-the-art transmission experiments addressing the various operation regimes have been reviewed and presented. While the complexity of the fiber-optic channel prevents the extraction of clear numbers for its information capacity, current technology appears to be rapidly approaching capacity bounds, making the continuing growth of information throughputs increasingly more challenging.

Acknowledgment

Mark Shtaif acknowledges important discussions with R. Zamir of Tel Aviv University regarding capacity lower bounds. ■

REFERENCES

- [1] P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, "Fiber-optic transmission and networking: The previous 20 and the next 20 years," *Opt. Exp.*, vol. 26, no. 18, pp. 24190–24239, 2018. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-26-18-24190>
- [2] D. C. Gloge and I. Jacobs, "Terrestrial intercity transmission systems," in *Optical Fiber Telecommunications II*, S. E. Miller and I. P. Kaminow, Eds. New York, NY, USA: Academic, 1988, ch. 23, pp. 855–867. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128165027000129>
- [3] P. K. Runge and N. S. Bergano, "Undersea cable transmission systems," in *Optical Fiber Telecommunications II*, S. E. Miller and I. P. Kaminow, Eds. New York, NY, USA: Academic, 1988, ch. 24, pp. 879–905. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128165027000129>
- [4] R. Mears, L. Reekie, I. Jauncey, and D. Payne, "Low-noise erbium-doped fibre amplifier operating at 1.54 μm ," *Electron. Lett.*, vol. 23, no. 19, pp. 1026–1028, 1987.
- [5] E. Desurvire, J. R. Simpson, and P. C. Becker, "High-gain erbium-doped traveling-wave fiber amplifier," *Opt. Lett.*, vol. 12, no. 11, pp. 888–890, Nov. 1987. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-12-11-888>
- [6] C. R. Giles and E. Desurvire, "Modeling erbium-doped fiber amplifiers," *J. Lightw. Technol.*, vol. 9, no. 2, pp. 271–283, Feb. 1991.
- [7] B. Jopson and A. Gnauck, "Dispersion compensation for optical fiber systems," *IEEE Commun. Mag.*, vol. 33, no. 6, pp. 96–102, Jun. 1995.
- [8] A. H. Gnauck and R. M. Jopson, "Dispersion compensation for optical fiber systems," in *Optical Fiber Telecommunications IIIA*, I. P. Kaminow and T. Koch, Eds. New York, NY, USA: Academic, 1997, ch. 7, pp. 162–190. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128165027000129>
- [9] N. S. Bergano, "Undersea amplified lightwave system design," in *Optical Fiber Telecommunications IIIA*, I. P. Kaminow and T. Koch, Eds. New York, NY, USA: Academic, 1997, ch. 10, pp. 302–332. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128165027000129>
- [10] H. Kogelnik, "High-capacity optical communications: Personal recollections," *IEEE J. Sel. Topics Quantum Electron.*, vol. 6, no. 6, pp. 1279–1286, Nov. 2000.
- [11] R. W. Tkach, "Scaling optical communications for the next decade and beyond," *Bell Labs Tech. J.*, vol. 14, no. 4, pp. 3–9, 2010. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/bltj.20400>
- [12] M. G. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 674–676, Feb. 2004.
- [13] F. Derr, "Coherent optical QPSK intradyne system: Concept and digital receiver realization," *J. Lightw. Technol.*, vol. 10, no. 9, pp. 1290–1296, Sep. 1992.
- [14] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Exp.*, vol. 16, no. 2, pp. 804–817, Jan. 2008. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-16-2-804>
- [15] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, Sep. 2010.
- [16] S. J. Savory, G. Gavioli, R. I. Killey, and P. Bayvel, "Electronic compensation of chromatic dispersion using a digital coherent receiver," *Opt. Exp.*, vol. 15, no. 5, pp. 2120–2126, Mar. 2007. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-15-5-2120>
- [17] V. Curri, P. Poggolini, A. Carena, and F. Forghieri, "Dispersion compensation and mitigation of nonlinear effects in 111-Gb/s WDM coherent PM-QPSK systems," *IEEE Photon. Technol. Lett.*, vol. 20, no. 17, pp. 1473–1475, Sep. 15, 2008.
- [18] H. Kogelnik, R. M. Jopson, and L. E. Nelson, "Polarization-mode dispersion," in *Optical Fiber Telecommunications IV-B* (Optics and Photonics), 4th ed. I. P. Kaminow and T. Li, Eds. Burlington, VT, USA: Academic, 2002, ch. 15, pp. 725–861. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780123951731500153>
- [19] A. Chraplyvy, "Plenary paper: The coming capacity crunch," in *Proc. 35th Eur. Conf. Opt. Commun.*, 2009, p. 1.
- [20] G. P. Agrawal, *Nonlinear Fiber Optics*, 6th ed. New York, NY, USA: Academic, 2019.
- [21] P. J. Winzer and G. J. Foschini, "MIMO capacities and outage probabilities in spatially multiplexed optical transport systems," *Opt. Exp.*, vol. 19, no. 17, pp. 16680–16696, Aug. 2011. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-19-17-16680>
- [22] R. Ryf et al., "Mode-division multiplexing over 96 km of few-mode fiber using coherent 6 \times 6 MIMO processing," *J. Lightw. Technol.*, vol. 30, no. 4, pp. 521–531, Feb. 12, 2012.
- [23] S. Randel et al., "6 \times 56-gb/s mode-division multiplexed transmission over 33-km few-mode fiber enabled by 6 \times 6 MIMO equalization," *Opt. Exp.*, vol. 19, no. 17, pp. 16697–16707, Aug. 2011. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-19-17-16697>
- [24] S. G. Leon-Saval et al., "Mode-selective photonic lanterns for space-division multiplexing," *Opt. Exp.*, vol. 22, no. 1, pp. 1036–1044, Jan. 2014. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-22-1-1036>
- [25] T. Morioka, Y. Awaji, R. Ryf, P. Winzer, D. Richardson, and F. Poletti, "Enhancing optical communications with brand new fibers," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. s31–s42, Feb. 2012.
- [26] N. K. Fontaine et al., "30 \times 30 MIMO transmission over 15 spatial modes," in *Proc. Opt. Fiber Commun. Conf. Post Deadline Papers*, 2015, p. Th5C.1. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2015-Th5C.1>
- [27] C. Koebel et al., "40km transmission of five mode division multiplexed data streams at 100Gb/s with low MIMO-DSP complexity," in *Proc. 37th Eur. Conf. Expo. Opt. Commun.*, 2011, pp. 1–2. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=ECOC-2011-Th.13.C.3>
- [28] G. Rademacher et al., "Long-haul transmission over few-mode fibers with space-division multiplexing," *J. Lightw. Technol.*, vol. 36, no. 6, pp. 1382–1388, Mar. 15, 2018.
- [29] B. J. Puttnam et al., "0.715 Pb/s transmission over 2,009.6 km in 19-core cladding pumped EDFA amplified MCF link," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2019, p. Th4B.1. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2019-Th4B.1>
- [30] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948.
- [31] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fiber communications," *Nature*, vol. 411, pp. 1027–1030, Jun. 2001. [Online]. Available: <https://doi.org/10.1038/35082518>
- [32] J. M. Kahn and K.-P. Ho, "A bottleneck for optical fibres," *Nature*, vol. 411, no. 6841, pp. 1007–1009, Jun. 2001, doi: [10.1038/35082671](https://doi.org/10.1038/35082671).
- [33] A. Mecozzi and M. Shttaif, "On the capacity of intensity modulated systems using optical amplifiers," *IEEE Photon. Technol. Lett.*, vol. 13, no. 9, pp. 1029–1031, Sep. 2001.
- [34] R. J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 15, 2013.
- [35] R.-J. Essiambre, G. Foschini, G. Kramer, and P. J. Winzer, "Capacity limits of information transport in fiber-optic networks," *Phys. Rev. Lett.*, vol. 101, Oct. 2008, Art. no. 163901. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.101.163901>
- [36] S. A. Derevyanko, J. E. Prilepsky, and S. K. Turitsyn, "Capacity estimates for optical transmission based on the nonlinear Fourier transform," *Nature Commun.*, vol. 7, Jun. 2016, Art. no. 12710. [Online]. Available: <https://doi.org/10.1038/ncomms12710>
- [37] K. S. Turitsyn, S. A. Derevyanko, I. V. Yurkevich, and S. K. Turitsyn, "Information capacity of optical fiber channels with zero average dispersion," *Phys. Rev. Lett.*, vol. 91, no. 20, Nov. 2003. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.91.203901>
- [38] A. Mecozzi and M. Shttaif, "Information capacity of direct detection optical transmission systems," *J. Lightw. Technol.*, vol. 36, no. 3, pp. 689–694, Feb. 1, 2018. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-36-3-689>
- [39] L. G. L. Wegener, M. L. Piovelli, A. G. Green, P. P. Mitra, J. B. Stark, and P. B. Littlewood, "The effect of propagation nonlinearities on the information capacity of WDM optical fiber systems: Cross-phase modulation and four-wave mixing," *Phys. D, Nonlinear Phenomena*, vol. 189, nos. 1–2, pp. 81–99, Feb. 2004. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167278903003890>
- [40] G. Bosco, P. Poggolini, A. Carena, V. Curri, and F. Forghieri, "Analytical results on channel capacity in uncompensated optical links with coherent detection," *Opt. Exp.*, vol. 19, no. 26, p. B440, Dec. 2011. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-19-26-B440>
- [41] J. Tang, "The channel capacity of a multispan DWDM system employing dispersive nonlinear optical fibers and an ideal coherent optical receiver," *J. Lightw. Technol.*, vol. 20, no. 7, p. 1095, Jul. 15, 2002. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-20-7-1095>
- [42] J. Tang, "A comparison study of the Shannon channel capacity of various nonlinear optical fibers," *J. Lightw. Technol.*, vol. 24, no. 5, pp. 2070–2075, May 2006.
- [43] A. D. Ellis, J. Zhao, and D. Cotter, "Approaching the non-linear Shannon limit," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 423–433, Feb. 15, 2010.
- [44] A. Alvarado, T. Fehenberger, B. Chen, and F. M. J. Willems, "Achievable information rates for fiber optics: Applications and computations," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 424–439, Jan. 15, 2018.
- [45] E. Agrell, "Capacity bounds in optical communications," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2017, pp. 1–3.
- [46] A. D. Ellis, M. E. McCarthy, M. A. Z. Al-Khateeb, and S. Sygletos, "Capacity limits of systems employing multiple optical phase conjugators," *Opt. Exp.*, vol. 23, no. 16, pp. 20381–20393, Aug. 2015. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-23-16-20381>
- [47] M. A. Sorokina and S. K. Turitsyn, "Regeneration limit of classical Shannon capacity," *Nature Commun.*, vol. 5, p. 3861, May 2014. [Online]. Available: <https://doi.org/10.1038/ncomms4861>
- [48] R. Dar, M. Shttaif, and M. Feder, "New bounds on the capacity of the nonlinear fiber-optic channel," *Opt. Lett.*, vol. 39, no. 2, pp. 398–401, Jan. 2014. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-39-2-398>
- [49] F. J. Garcia-Gomez and G. Kramer, "Mismatched models to lower bound the capacity of optical fiber channels," *J. Lightw. Technol.*, vol. 38, no. 24, pp. 6779–6787, Dec. 15, 2020. [Online].

- Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-38-24-6779>
- [50] F. J. Garcia-Gomez and G. Kramer, "Mismatched models to lower bound the capacity of dual-polarization optical fiber channels," *J. Lightw. Technol.*, vol. 39, no. 11, pp. 3390–3399, Jun. 1, 2021. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-39-11-3390>
- [51] A. Alvarado, E. Agrell, D. Lavery, R. Maher, and P. Bayvel, "Replacing the soft-decision FEC limit paradigm in the design of optical communication systems," *J. Lightw. Technol.*, vol. 33, no. 20, pp. 4338–4352, Oct. 15, 2015.
- [52] E. Agrell, A. Alvarado, G. Durisi, and M. Karlsson, "Capacity of a nonlinear optical channel with finite memory," *J. Lightw. Technol.*, vol. 32, no. 16, pp. 2862–2876, Aug. 15, 2014.
- [53] M. Secondini and E. Forestieri, "Scope and limitations of the nonlinear Shannon limit," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 893–902, Feb. 1, 2017.
- [54] M. Secondini, "Information capacity of optical channels," in *Optical Fiber Telecommunications VII*, A. E. Willner, Ed. New York, NY, USA: Academic, 2020, ch. 20, pp. 867–920. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128165027000233>
- [55] A. Sano et al., "69.1-Tb/s (432 × 171-Gb/s) C- and extended L-band transmission over 240 km using PDM-16-QAM modulation and digital coherent detection," in *Proc. Opt. Fiber Commun. Conf.*, 2010, p. PDPB7. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2010-PDPB7>
- [56] Z. Zhan et al., "Optical polarization-based seismic and water wave sensing on transoceanic cables," *Science*, vol. 371, no. 6532, pp. 931–936, Feb. 2021. [Online]. Available: <https://www.science.org/doi/abs/10.1126/science.abe6648>
- [57] C. M. Caves, "Quantum limits on noise in linear amplifiers," *Phys. Rev. D, Part. Fields*, vol. 26, no. 8, pp. 1817–1839, 1982. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevD.26.1817>
- [58] SV Manakov, "On the theory of two-dimensional stationary self-focusing of electromagnetic waves," *JETP, Sov. Phys.*, vol. 38, no. 2, pp. 248–253, 1974.
- [59] D. Marcuse, C. R. Manyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997.
- [60] S. G. Evangelides, Jr., L. F. Mollenauer, J. P. Gordon, and N. S. Bergano, "Polarization multiplexing with solitons," *J. Lightw. Technol.*, vol. 10, no. 1, pp. 28–35, Jan. 1992.
- [61] G. P. Agrawal, "Polarization effects," in *Nonlinear Fiber Optics*, 6th ed., G. P. Agrawal, Ed. New York, NY, USA: Academic, ch. 6, 2019, pp. 189–244. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128170427000130>
- [62] G. P. Agrawal, "Self-phase modulation," in *Nonlinear Fiber Optics*, 6th ed., G. P. Agrawal, Ed. New York, NY, USA: Academic, ch. 4, 2019, pp. 85–125. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128170427000117>
- [63] O. V. Sinkin, R. Holzlohner, J. Zweck, and C. R. Menyuk, "Optimization of the split-step Fourier method in modeling optical-fiber communications systems," *J. Lightw. Technol.*, vol. 21, no. 1, pp. 61–68, Jan. 2003.
- [64] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, Oct. 15, 2008. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-26-20-3416>
- [65] G. Kramer, M. I. Yousefi, and F. R. Kschischang, "Upper bound on the capacity of a cascade of nonlinear and noisy channels," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Apr. 2015, pp. 1–4.
- [66] M. I. Yousefi, G. Kramer, and F. R. Kschischang, "Upper bound on the capacity of the nonlinear Schrödinger channel," in *Proc. IEEE 14th Can. Workshop Inf. Theory (CWIT)*, Jul. 2015, pp. 22–26.
- [67] E. Agrell, "Conditions for a monotonic channel capacity," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 738–748, Mar. 2015.
- [68] E. Meron, M. Shtaiif, and M. Feder, "Beneficial use of spectral broadening resulting from the nonlinearity of the fiber-optic channel," *Opt. Lett.*, vol. 37, no. 21, pp. 4458–4460, Nov. 2012. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-37-21-4458>
- [69] A. Mecozzi, C. B. Clausen, and M. Shtaiif, "Analysis of intrachannel nonlinear effects in highly dispersed optical pulse transmission," *IEEE Photon. Technol. Lett.*, vol. 12, no. 4, pp. 392–394, Apr. 2000.
- [70] A. Mecozzi, C. B. Clausen, M. Shtaiif, S.-G. Park, and A. H. Gnauck, "Cancellation of timing and amplitude jitter in symmetric links using highly dispersed pulses," *IEEE Photon. Technol. Lett.*, vol. 13, no. 5, pp. 445–447, May 2001.
- [71] R.-J. Essiambre, G. Raybon, and B. Mikkelsen, "Pseudo-linear transmission of high-speed TDM signals: 40 and 160 gb/s," in *Optical Fiber Telecommunications IV-B (Optics and Photonics)*, 4th ed., I. P. Kaminow and T. Li, Eds. Burlington, VT, USA: Academic, 2002, ch. 6, pp. 232–304. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780123951731500062>
- [72] A. Bononi, P. Serena, and J.-C. Antona, "A state-variable approach to submarine links capacity optimization," *J. Lightw. Technol.*, vol. 39, no. 18, pp. 5753–5765, Sep. 22, 2021.
- [73] R. Dar, M. Feder, A. Mecozzi, and M. Shtaiif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Exp.*, vol. 21, no. 22, pp. 25685–25699, 2013. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-21-22-25685>
- [74] R. Dar, M. Feder, A. Mecozzi, and M. Shtaiif, "Accumulation of nonlinear interference noise in fiber-optic systems," *Opt. Exp.*, vol. 22, no. 12, pp. 14199–14211, 2014. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-22-12-14199>
- [75] M. Secondini and E. Forestieri, "Analytical fiber-optic channel model in the presence of cross-phase modulation," *IEEE Photon. Technol. Lett.*, vol. 24, no. 22, pp. 2016–2019, Nov. 15, 2012.
- [76] D. Piliori, F. Forghieri, and G. Bosco, "Residual non-linear phase noise in probabilistically shaped 64-QAM optical links," in *Proc. Opt. Fiber Commun. Conf.*, 2018, pp. 1–3.
- [77] O. Golani et al., "Experimental characterization of nonlinear interference noise as a process of intersymbol interference," *Opt. Lett.*, vol. 43, no. 5, pp. 1123–1126, Mar. 2018. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-43-5-1123>
- [78] O. Golani, D. Piliori, F. P. P. Guiomar, G. Bosco, A. Carena, and M. Shtaiif, "Correlated nonlinear phase-noise in multi-subcarrier systems: Modeling and mitigation," *J. Lightw. Technol.*, vol. 38, no. 6, pp. 1148–1156, Mar. 15, 2020.
- [79] A. Splett, C. Kurtzke, and K. Petermann, "Ultimate transmission capacity of amplified optical fiber communication systems taking into account fiber nonlinearities," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 1993, Art. no. MoC2.4.
- [80] P. Poggiolini, A. Carena, V. Curri, G. Bosco, and F. Forghieri, "Analytical modeling of nonlinear propagation in uncompensated optical transmission links," *IEEE Photon. Technol. Lett.*, vol. 23, no. 11, pp. 742–744, Jun. 15, 2011.
- [81] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3857–3879, Dec. 15, 2012. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-30-24-3857>
- [82] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Exp.*, vol. 22, no. 13, pp. 16335–16362, 2014. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-22-13-16335>
- [83] P. Poggiolini, Y. Jiang, A. Carena, G. Bosco, and F. Forghieri, "Analytical results on system maximum reach increase through symbol rate optimization," in *Proc. Opt. Fiber Commun. Conf.*, 2015, pp. 1–3.
- [84] R. Dar, M. Feder, A. Mecozzi, and M. Shtaiif, "Inter-channel nonlinear interference noise in WDM systems: Modeling and mitigation," *J. Lightw. Technol.*, vol. 33, no. 5, pp. 1044–1053, Mar. 1, 2015.
- [85] O. Golani, R. Dar, M. Feder, A. Mecozzi, and M. Shtaiif, "Modeling the bit-error-rate performance of nonlinear fiber-optic systems," *J. Lightw. Technol.*, vol. 34, no. 15, pp. 3482–3489, Aug. 1, 2016.
- [86] R. Dar, M. Feder, A. Mecozzi, and M. Shtaiif, "Pulse collision picture of inter-channel nonlinear interference in fiber-optic communications," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 593–607, Jan. 15, 2016.
- [87] A. Mecozzi and R.-J. Essiambre, "Nonlinear Shannon limit in pseudolinear coherent systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 2011–2024, Jun. 15, 2012.
- [88] G. P. Agrawal, "Four-wave mixing," in *Nonlinear Fiber Optics*, 6th ed., G. P. Agrawal, Ed. New York, NY, USA: Academic, 2019, ch. 6, pp. 401–462. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128170427000178>
- [89] G. Bosco, A. Carena, R. Cigliutti, V. Curri, P. Poggiolini, and F. Forghieri, "Performance prediction for WDM PM-QSK transmission over uncompensated links," in *Proc. Opt. Fiber Commun. Conf., Nat. Fiber Optic Eng. Conf.*, 2011, pp. 1–3.
- [90] E. Grellier and A. Bononi, "Quality parameter for coherent transmissions with Gaussian-distributed nonlinear noise," *Opt. Exp.*, vol. 19, no. 13, pp. 12781–12788, Jun. 2011. [Online]. Available: <http://opg.optica.org/oe/abstract.cfm?URI=oe-19-13-12781>
- [91] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," in *Proc. Nat. Acad. Sci. USA*, Washington, DC, USA, 2000, vol. 97, no. 9, pp. 4541–4550. [Online]. Available: <https://www.pnas.org/content/97/9/4541>
- [92] M. Brodsky, N. J. Frigo, and M. Tur, "Polarization mode dispersion," in *Optical Fiber Telecommunications (Optics and Photonics)*, 5th ed., I. P. Kaminow, T. Li, and A. E. Willner, Eds. Burlington, VT, USA: Academic, 2008, ch. 17, pp. 605–669. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780123741714000174>
- [93] N. Gisin and B. Huttner, "Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers," *Opt. Commun.*, vol. 142, nos. 1–3, pp. 119–125, Oct. 1997. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0030401897002368>
- [94] A. Mecozzi and M. Shtaiif, "Signal-to-noise-ratio degradation caused by polarization-dependent loss and the effect of dynamic gain equalization," *J. Lightw. Technol.*, vol. 22, no. 8, p. 1856, Aug. 2004. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-22-8-1856>
- [95] M. Shtaiif and O. Rosenberg, "Polarization-dependent loss as a waveform-distorting mechanism and its effect on fiber-optic systems," *J. Lightw. Technol.*, vol. 23, no. 2, p. 923, Feb. 2005. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-23-2-923>
- [96] M. Shtaiif, "Performance degradation in coherent polarization multiplexed systems as a result of polarization dependent loss," *Opt. Exp.*, vol. 16, no. 18, pp. 13918–13932, 2008. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-16-18-13918>
- [97] M. Brodsky, N. J. Frigo, M. Boroditsky, and M. Tur, "Polarization mode dispersion of installed fibers," *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4584–4599, Dec. 2006. [Online]. Available:

- <http://jlt.osa.org/abstract.cfm?URI=jlt-24-12-4584>
- [98] P. Magill and M. Brodsky, "PMD of installed fiber—An overview," in *Proc. LEOS Summer Topical Meeting*, 2003, pp. MB2.2/7–MB2.2/8.
- [99] M. Boroditsky, M. Brodsky, N. J. Frigo, P. Magill, and H. Rosenfeldt, "Polarization dynamics in installed fiberoptic systems," in *Proc. IEEE LEOS Annu. Meeting Conf.*, 2005, pp. 414–415.
- [100] M. Brodsky, M. Boroditsky, P. Magill, N. Frigo, and M. Tur, "Persistence of spectral variations in DGD statistics," *Opt. Exp.*, vol. 13, no. 11, pp. 4090–4095, May 2005. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-13-11-4090>
- [101] M. Brodsky, P. Magill, and N. J. Frigo, "Polarization-mode dispersion of installed recent vintage fiber as a parametric function of temperature," *IEEE Photon. Technol. Lett.*, vol. 16, no. 1, pp. 209–211, Jan. 2004.
- [102] C. Antonelli, C. Colamarino, A. Mecozzi, and M. Brodsky, "A model for temporal evolution of PMD," *IEEE Photon. Technol. Lett.*, vol. 20, no. 12, pp. 1012–1014, Jun. 15, 2008.
- [103] C. Antonelli and A. Mecozzi, "Theoretical characterization and system impact of the Hinge model of PMD," *J. Lightw. Technol.*, vol. 24, no. 11, pp. 4064–4074, Nov. 22, 2006. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-24-11-4064>
- [104] A. Mecozzi, C. Antonelli, and M. Brodsky, "Statistics of the polarization mode dispersion dynamics," *Opt. Lett.*, vol. 32, no. 20, pp. 3032–3034, Oct. 2007. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-32-20-3032>
- [105] M. Boroditsky, M. Brodsky, N. J. Frigo, P. Magill, C. Antonelli, and A. Mecozzi, "Outage probabilities for fiber routes with finite number of degrees of freedom," *IEEE Photon. Technol. Lett.*, vol. 17, no. 2, pp. 345–347, Feb. 2005.
- [106] M. Boroditsky, M. Bourd, and M. Tur, "Effect of nonlinearities on PMD," *J. Lightw. Technol.*, vol. 24, no. 11, pp. 4100–4107, Nov. 2006. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-24-11-4100>
- [107] E. Alperovich, A. Mecozzi, and M. Shtaiif, "PMD penalties in long nonsoliton systems and the effect of inline filtering," *IEEE Photon. Technol. Lett.*, vol. 18, no. 10, pp. 1179–1181, May 15, 2006.
- [108] M. Cantono, D. Piloni, A. Ferrari, A. Carena, and V. Curri, "Observing the interaction of PMD with generation of NLI in uncompensated amplified optical links," in *Proc. Opt. Fiber Commun. Conf.*, 2018, pp. 1–3.
- [109] M. Karlsson and J. Brentel, "Autocorrelation function of the polarization-mode dispersion vector," *Opt. Lett.*, vol. 24, no. 14, pp. 939–941, Jul. 1999. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-24-14-939>
- [110] M. Shtaiif, A. Mecozzi, and J. A. Nagel, "Mean-square magnitude of all orders of polarization mode dispersion and the relation with the bandwidth of the principal states," *IEEE Photon. Technol. Lett.*, vol. 12, no. 1, pp. 53–55, Jan. 2000.
- [111] M. Shtaiif and A. Mecozzi, "Study of the frequency autocorrelation of the differential group delay in fibers with polarization mode dispersion," *Opt. Lett.*, vol. 25, no. 10, pp. 707–709, May 2000. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-25-10-707>
- [112] M. Shtaiif and A. Mecozzi, "Polarization-dependent loss and its effect on the signal-to-noise ratio in fiber-optic systems," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 671–673, Feb. 2004.
- [113] C. Antonelli, A. Mecozzi, L. E. Nelson, and P. Magill, "Autocorrelation of the polarization-dependent loss in fiber routes," *Opt. Lett.*, vol. 36, no. 20, pp. 4005–4007, Oct. 2011. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-36-20-4005>
- [114] C. Antonelli, A. Mecozzi, M. Shtaiif, and P. J. Winzer, "Modeling and performance metrics of MIMO-SDM systems with different amplification schemes in the presence of mode-dependent loss," *Opt. Exp.*, vol. 23, no. 3, pp. 2203–2219, Feb. 2015. [Online]. Available: <http://www.osapublishing.org/oe/abstract.cfm?URI=oe-23-3-2203>
- [115] A. Nafta, E. Meron, and M. Shtaiif, "Capacity limitations in fiber-optic communication systems as a result of polarization-dependent loss," *Opt. Lett.*, vol. 34, no. 23, pp. 3613–3615, Dec. 2009. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-34-23-3613>
- [116] P. Serena, C. Lasagni, and A. Bononi, "The enhanced Gaussian noise model extended to polarization-dependent loss," *J. Lightw. Technol.*, vol. 38, no. 20, pp. 5685–5694, Oct. 15, 2020. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-38-20-5685>
- [117] D. Semrau, R. I. Killely, and P. Bayvel, "The Gaussian noise model in the presence of inter-channel stimulated Raman scattering," *J. Lightw. Technol.*, vol. 36, no. 14, pp. 3046–3055, Jul. 15, 2018. [Online]. Available: <http://opg.optica.org/jlt/abstract.cfm?URI=jlt-36-14-3046>
- [118] D. Semrau, R. I. Killely, and P. Bayvel, "A closed-form approximation of the Gaussian noise model in the presence of inter-channel stimulated Raman scattering," *J. Lightw. Technol.*, vol. 37, no. 9, pp. 1924–1936, Jan. 24, 2019. [Online]. Available: <http://opg.optica.org/jlt/abstract.cfm?URI=jlt-37-9-1924>
- [119] M. R. Zeffeh, F. Forghieri, S. Picciaccia, and P. Poggiolini, "A closed-form nonlinearity model for forward-Raman-amplified WDM optical links," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2021, p. M5C.1. [Online]. Available: <http://opg.optica.org/abstract.cfm?URI=OFC-2021-M5C.1>
- [120] A. Tasbihi and F. R. Kschischang, "On the capacity of waveform channels under square-law detection of time-limited signals," *IEEE Trans. Inf. Theory*, vol. 66, no. 11, pp. 6682–6687, Nov. 2020.
- [121] X. Chen, C. Antonelli, A. Mecozzi, D. Che, and W. Shieh, "High-capacity direct-detection systems," in *Optical Fiber Telecommunications VII*, A. E. Willner, Ed. New York, NY, USA: Academic, 2020, ch. 10, pp. 419–441. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128165027000129>
- [122] R. Rios-Müller, J. Estaran, J. Renaudier, and G. Charlet, "Dual polarization in-phase and quadrature high speed submarine transmission with only two photodiodes, ADCs, MZMs and DACs," in *Proc. IEEE Photon. Conf. (IPC)*, Oct. 2017, pp. 1–2.
- [123] B. E. A. Saleh and M. C. Teich, *Semiconductor Photon Detectors*. Hoboken, NJ, USA: Wiley, ch. 17, pp. 644–695.
- [124] L. Blech, Y. Eldar, C. Antonelli, A. Mecozzi, and M. Shtaiif, "The enhanced Kramers Kronig receiver," in *Proc. Opt. Fiber Commun. Conf.*, 2018, p. Tu2D.7. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2018-Tu2D.7>
- [125] L. Blech, C. Antonelli, A. Mecozzi, Y. C. Eldar, and M. Shtaiif, "Enhancing the Kramers-Kronig receiver via dispersion-based spatial diversity," *Opt. Lett.*, vol. 45, no. 13, pp. 3494–3497, Jul. 2020. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-45-13-3494>
- [126] H. Huang, H. Chen, Y. Huang, N. K. Fontaine, R. Ryf, and Y. Song, "Carrier-less phase retrieval receiver leveraging digital upsampling," *Opt. Lett.*, vol. 45, no. 21, pp. 6070–6073, Nov. 2020. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-45-21-6070>
- [127] H. Chen, N. K. Fontaine, J. M. Gene, R. Ryf, D. T. Neilson, and G. Raybon, "Dual polarization full-field signal waveform reconstruction using intensity only measurements for coherent communications," *J. Lightw. Technol.*, vol. 38, no. 9, pp. 2587–2597, May 1, 2020. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-38-9-2587>
- [128] S. Randel, D. Piloni, S. Chandrasekhar, G. Raybon, and P. Winzer, "100-Gb/s discrete-multitone transmission over 80-km SSMF using single-sideband modulation with novel interference-cancellation scheme," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2015, pp. 1–3.
- [129] K. Zou, Y. Zhu, F. Zhang, and Z. Chen, "Spectrally efficient terabit optical transmission with Nyquist 64-QAM half-cycle subcarrier modulation and direct detection," *Opt. Lett.*, vol. 41, no. 12, pp. 2767–2770, Jun. 2016. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-41-12-2767>
- [130] Z. Li et al., "Comparison of digital signal-signal beat interference compensation techniques in direct-detection subcarrier modulation systems," *Opt. Exp.*, vol. 24, no. 25, pp. 29176–29189, Dec. 2016. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-24-25-29176>
- [131] A. Yekani, S. Amirizadeh, and L. A. Rusch, "Analytical study of optical SSB-DMT with IMDD," *J. Lightw. Technol.*, vol. 36, no. 3, pp. 666–674, Feb. 1, 2018.
- [132] M. Schuster et al., "Spectrally efficient compatible single-sideband modulation for OFDM transmission with direct detection," *IEEE Photon. Technol. Lett.*, vol. 20, no. 9, pp. 670–672, May 1, 2008.
- [133] A. Mecozzi, C. Antonelli, and M. Shtaiif, "Kramers-Kronig coherent receiver," *Optica*, vol. 3, no. 11, pp. 1220–1227, 2016. [Online]. Available: <http://www.osapublishing.org/optica/abstract.cfm?URI=optica-3-11-1220>
- [134] Z. Li et al., "SSBI mitigation and the Kramers-Kronig scheme in single-sideband direct-detection transmission with receiver-based electronic dispersion compensation," *J. Lightw. Technol.*, vol. 35, no. 10, pp. 1887–1893, May 15, 2017.
- [135] X. Chen et al., "218-Gb/s single-wavelength, single-polarization, single-photodiode transmission over 125-km of standard singlemode fiber using Kramers-Kronig detection," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2017, p. Th5B.6. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2017-Th5B.6>
- [136] A. Mecozzi, C. Antonelli, and M. Shtaiif, "Kramers-Kronig receivers," *Adv. Opt. Photon.*, vol. 11, no. 3, pp. 480–517, Sep. 2019. [Online]. Available: <http://aop.osa.org/abstract.cfm?URI=aop-11-3-480>
- [137] D. Che, A. Li, X. Chen, Q. Hu, Y. Wang, and W. Shieh, "Stokes vector direct detection for linear complex optical channels," *J. Lightw. Technol.*, vol. 33, no. 3, pp. 678–684, Feb. 1, 2015.
- [138] T. M. Hoang et al., "Single wavelength 480 Gb/s direct detection over 80km SSMF enabled by Stokes vector Kramers Kronig transceiver," *Opt. Exp.*, vol. 25, no. 26, pp. 33534–33542, Dec. 2017. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-25-26-33534>
- [139] C. Antonelli, A. Mecozzi, M. Shtaiif, X. Chen, S. Chandrasekhar, and P. J. Winzer, "Polarization multiplexing with the Kramers-Kronig receiver," *J. Lightw. Technol.*, vol. 35, no. 24, pp. 5418–5424, Dec. 15, 2017.
- [140] W. Shieh, C. Sun, and H. Ji, "Carrier-assisted differential detection," *Light, Sci. Appl.*, vol. 9, no. 1, p. 18, Feb. 2020, doi: [10.1038/s41377-020-0253-8](https://doi.org/10.1038/s41377-020-0253-8).
- [141] D. Che et al., "Single-channel direct detection reception beyond 1 Tb/s," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2019, pp. 1–3.
- [142] G. Meslener, "Chromatic dispersion induced distortion of modulated monochromatic light employing direct detection," *IEEE J. Quantum Electron.*, vol. JQE-20, no. 10, pp. 1208–1216, Oct. 1984.
- [143] R. Essiambre and R. W. Tkach, "Capacity trends and limits of optical communication networks," *Proc. IEEE*, vol. 100, no. 5, pp. 1035–1055, May 2012.
- [144] B. J. Puttnam, G. Rademacher, and R. S. Luis, "Space-division multiplexing for optical fiber communications," *Optica*, vol. 8, no. 9,

- pp. 1186–1203, Sep. 2021. [Online]. Available: <http://www.osapublishing.org/optica/abstract.cfm?URI=optica-8-9-1186>
- [145] O. V. Sinkin, A. V. Turukhin, W. W. Patterson, M. A. Bolshtyansky, D. G. Foursa, and A. N. Pilipetskii, "Maximum optical power efficiency in SDM-based optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 29, no. 13, pp. 1075–1077, Jul. 1, 2017.
- [146] R. Dar et al., "Cost-optimized submarine cables using massive spatial parallelism," *J. Lightw. Technol.*, vol. 36, no. 18, pp. 3855–3865, Sep. 15, 2018.
- [147] O. V. Sinkin et al., "SDM for power-efficient undersea transmission," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 361–371, Jan. 15, 2018.
- [148] J. K. Perin, J. M. Kahn, J. D. Downie, J. Hurley, and K. Bennett, "Importance of amplifier physics in maximizing the capacity of submarine links," *J. Lightw. Technol.*, vol. 37, no. 9, pp. 2076–2085, May 1, 2019.
- [149] H. Srinivas et al., "Modeling and experimental measurement of power efficiency for power-limited SDM submarine transmission systems," *J. Lightw. Technol.*, vol. 39, no. 8, pp. 2376–2386, Apr. 15, 2021.
- [150] S. Buch, S. Mumtaz, R.-J. Essiambre, A. M. Tulino, and G. P. Agrawal, "Averaged nonlinear equations for multimode fibers valid in all regimes of random linear coupling," *Opt. Fiber Technol.*, vol. 48, pp. 123–127, Mar. 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1068520018305662>
- [151] F. M. Ferreira, C. S. Costa, S. Sygletos, and A. D. Ellis, "Nonlinear performance of few-mode fiber links with intermediate coupling," *J. Lightw. Technol.*, vol. 37, no. 3, pp. 989–999, Feb. 1, 2019. [Online]. Available: <http://opg.optica.org/jlt/abstract.cfm?URI=jlt-37-3-989>
- [152] S. Mumtaz, R.-J. Essiambre, and G. P. Agrawal, "Reduction of nonlinear penalties due to linear coupling in multicore optical fibers," *IEEE Photon. Technol. Lett.*, vol. 24, no. 18, pp. 1574–1576, Sep. 15, 2012.
- [153] C. Antonelli, M. Shtaiif, and A. Mecozzi, "Modeling of nonlinear propagation in space-division multiplexed fiber-optic transmission," *J. Lightw. Technol.*, vol. 34, no. 1, pp. 36–54, Jan. 1, 2016.
- [154] R. Ryf et al., "Long-haul transmission over multi-core fibers with coupled cores," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2017, pp. 1–3.
- [155] R. Ryf et al., "Coupled-core transmission over 7-core fiber," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2019, p. Th4B.3. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2019-Th4B.3>
- [156] A. Mecozzi, C. Antonelli, and M. Shtaiif, "Nonlinear propagation in multi-mode fibers in the strong coupling regime," *Opt. Exp.*, vol. 20, no. 11, pp. 11673–11678, 2012. [Online]. Available: <http://www.osapublishing.org/oe/abstract.cfm?URI=oe-20-11-11673>
- [157] F. J. Garcia-Gomez and G. Kramet, "Regular perturbation and achievable rates of space-division multiplexed optical channels," in *Proc. 17th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Sep. 2021, pp. 1–5.
- [158] R. Ryf and C. Antonelli, *Space-Division Multiplexing*. Cham, Switzerland: Springer, 2020, pp. 353–393.
- [159] R. Ryf et al., "Space-division multiplexed transmission over 4200-km 3-core microstructured fiber," in *Proc. Nat. Fiber Optic Eng. Conf.*, 2012, p. PDP5C.2.
- [160] A. Mecozzi, C. Antonelli, and M. Shtaiif, "Intensity impulse response of SDM links," *Opt. Exp.*, vol. 23, no. 5, pp. 5738–5743, Mar. 2015. [Online]. Available: <http://www.osapublishing.org/oe/abstract.cfm?URI=oe-23-5-5738>
- [161] C. Antonelli, A. Mecozzi, and M. Shtaiif, "The delay spread in fibers for SDM transmission: Dependence on fiber parameters and perturbations," *Opt. Exp.*, vol. 23, no. 3, pp. 2196–2202, Feb. 2015. [Online]. Available: <http://www.osapublishing.org/oe/abstract.cfm?URI=oe-23-3-2196>
- [162] C. Antonelli, A. Mecozzi, M. Shtaiif, N. K. Fontaine, H. Chen, and R. Ryf, "Stokes-space analysis of modal dispersion of SDM fibers with mode-dependent loss: Theory and experiments," *J. Lightw. Technol.*, vol. 38, no. 7, pp. 1668–1677, Apr. 1, 2020.
- [163] C. Antonelli, O. Golani, M. Shtaiif, and A. Mecozzi, "Nonlinear interference noise in space-division multiplexed transmission through optical fibers," *Opt. Exp.*, vol. 25, no. 12, pp. 13055–13078, Jun. 2017. [Online]. Available: <http://www.osapublishing.org/oe/abstract.cfm?URI=oe-25-12-13055>
- [164] P. Serena, C. Lasagni, A. Bononi, C. Antonelli, and A. Mecozzi, "The ergodic GN model for space-division multiplexing with strong mode coupling," *J. Lightw. Technol.*, vol. 40, no. 10, pp. 3263–3276, May 15, 2022.
- [165] K.-P. Ho and J. M. Kahn, "Mode-dependent loss and gain: Statistics and effect on mode-division multiplexing," *Opt. Exp.*, vol. 19, no. 17, pp. 16612–16635, 2011. [Online]. Available: <http://www.osapublishing.org/oe/abstract.cfm?URI=oe-19-17-16612>
- [166] A. Andrusier, M. Shtaiif, C. Antonelli, and A. Mecozzi, "Assessing the effects of mode-dependent loss in space-division multiplexed systems," *J. Lightw. Technol.*, vol. 32, no. 7, pp. 1317–1322, Apr. 15, 2014. [Online]. Available: <http://www.osapublishing.org/jlt/abstract.cfm?URI=jlt-32-7-1317>
- [167] S. L. I. Olsson, J. Cho, S. Chandrasekhar, X. Chen, E. C. Burrows, and P. J. Winzer, "Record-high 17.3-bit/s/Hz spectral efficiency transmission over 50 km using probabilistically shaped PDM 4096-QAM," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2018, pp. 1–3.
- [168] S. Chandrasekhar et al., "High-spectral-efficiency transmission of PDM 256-QAM with parallel probabilistic shaping at record rate-reach trade-offs," in *Proc. 42nd Eur. Conf. Opt. Commun. (ECOC)*, 2016, pp. 1–3.
- [169] T. J. Xia et al., "High capacity field trials of 40.5 Tb/s for LH distance of 1,822 km and 54.2 Tb/s for regional distance of 634 km," in *Proc. Opt. Fiber Commun. Conf., Nat. Fiber Optic Eng. Conf.*, 2013, pp. 1–3.
- [170] T. Kobayashi et al., "35-Tb/s C-band transmission over 800 km employing 1-Tb/s PS-64QAM signals enhanced by complex 8×2 MIMO equalizer," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2019, pp. 1–3.
- [171] T. Rahman et al., "Record field demonstration of C-band multi-terabit 16 QAM, 32 QAM and 64 QAM over 762 km of SSMF," in *Proc. Opto-Electron. Commun. Conf. (OECC)*, Jun. 2015, pp. 1–3.
- [172] T. J. Xia et al., "41.5 Tb/s data transport over 549 km of field deployed fiber using throughput optimized probabilistic-shaped 144 QAM to support metro network capacity demands," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2018, pp. 1–3.
- [173] J. Renaudier et al., "First 100-nm continuous-band WDM transmission system with 115Tb/s transport over 100km using novel ultra-wideband semiconductor optical amplifiers," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2017, pp. 1–3.
- [174] J. Zhang et al., "Transmission of 20×440 -Gb/s super-nyquist-filtered signals over 3600 km based on single-carrier 110-Gbaud PDM QPSK with 100-GHz grid," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2014, pp. 1–3.
- [175] J. Cho et al., "Trans-atlantic field trial using high spectral efficiency probabilistically shaped 64-QAM and single-carrier real-time 250-Gb/s 16-QAM," *J. Lightw. Technol.*, vol. 36, no. 1, pp. 103–113, Jan. 1, 2018.
- [176] S. Zhang et al., "Capacity-approaching transmission over 6375 km at spectral efficiency of 8.3 bit/s/Hz," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2016, pp. 1–3.
- [177] F. Yaman et al., "First quasi-single-mode transmission over transoceanic distance using few-mode fibers," in *Proc. Opt. Fiber Commun. Conf. Post Deadline Papers*, 2015, p. Th5C.7. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2015-Th5C.7>
- [178] F. Yaman et al., "30.6-Tb/s full-duplex bidirectional transoceanic transmission over 75×90.9 -km fiber spans," in *Proc. Opt. Fiber Commun. Conf., Postdeadline Papers*, 2014, pp. 1–3.
- [179] J. X. Cai et al., "64 QAM based coded modulation transmission over transoceanic distance with >60 Tb/s capacity," in *Proc. Opt. Fiber Commun. Conf. Post Deadline Papers*, 2015, p. Th5C.8. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2015-Th5C.8>
- [180] A. Ghazisaeidi et al., "65 Tb/s transoceanic transmission using probabilistically-shaped PDM-64QAM," in *Proc. 42nd Eur. Conf. Opt. Commun.*, 2016, pp. 1–3.
- [181] J.-X. Cai et al., "70.4 Tb/s capacity over 7,600 km in C+L band using coded modulation with hybrid constellation shaping and nonlinearity compensation," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2017, p. Th5B.2. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2017-Th5B.2>
- [182] J.-X. Cai et al., "112 \times 112 Gb/s transmission over 9,360 km with channel spacing set to the baud rate (360% spectral efficiency)," in *Proc. 36th Eur. Conf. Exhib. Opt. Commun.*, Sep. 2010, pp. 1–3.
- [183] J.-X. Cai et al., "Transmission over 9,100 km with a capacity of 49.3 Tb/s using variable spectral efficiency 16 QAM based coded modulation," in *Proc. Opt. Fiber Commun. Conf., Postdeadline Papers*, 2014, p. Th5B.4. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2014-Th5B.4>
- [184] D. G. Foursa et al., "44.1 Tb/s transmission over 9,100 km using coded modulation based on 16 QAM signals at 4.9 bits/s/Hz spectral efficiency," in *Proc. 39th Eur. Conf. Exhib. Opt. Commun. (ECOC)*, 2013, pp. 1–3.
- [185] J.-X. Cai et al., "54 Tb/s transmission over 9,150 km with optimized hybrid Raman-EDFA amplification and coded modulation," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2014, pp. 1–3.
- [186] A. Ghazisaeidi et al., "52.9 Tb/s transmission over transoceanic distances using adaptive multi-rate FEC," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2014, pp. 1–3.
- [187] M. Salsi et al., "38.75 Tb/s transmission experiment over transoceanic distance," in *Proc. 39th Eur. Conf. Exhib. Opt. Commun. (ECOC)*, 2013, pp. 1–3.
- [188] X. Liu, A. R. Chraplyvy, P. J. Winzer, R. W. Tkach, and S. Chandrasekhar, "Phase-conjugated twin waves for communication beyond the Kerr nonlinearity limit," *Nature Photon.*, vol. 7, pp. 560–568, May 2013. [Online]. Available: <https://doi.org/10.1038/nphoton.2013.109>
- [189] J.-X. Cai et al., "51.5 Tb/s capacity over 17,107 km in C+L bandwidth using single-mode fibers and nonlinearity compensation," *J. Lightw. Technol.*, vol. 36, no. 11, pp. 2135–2141, Jun. 1, 2018. [Online]. Available: <http://www.osapublishing.org/jlt/abstract.cfm?URI=jlt-36-11-2135>
- [190] M. Ionescu et al., "74.38 Tb/s transmission over 6300 km single mode fibre enabled by C+L amplification and geometrically shaped PDM-64 QAM," *J. Lightw. Technol.*, vol. 38, no. 2, pp. 531–537, Jan. 15, 2020.
- [191] V. Sleiffer et al., "73.7 Tb/s ($96 \times 3 \times 256$ -Gb/s) mode-division-multiplexed DP-16 QAM transmission with inline MM-EDFA," *Opt. Exp.*, vol. 20, no. 26, pp. B428–B438, Dec. 2012. [Online]. Available: <http://www.osapublishing.org/oe/abstract.cfm?URI=oe-20-26-B428>
- [192] R. Ryf et al., "32-bit/s/Hz spectral efficiency WDM transmission over 177-km few-mode fiber," in *Proc. Opt. Fiber Commun. Conf., Nat. Fiber Optic Eng. Conf.*, 2013, p. PDP5A.1. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2013-PDP5A.1>
- [193] T. Kobayashi et al., "1-Pb/s (32 SDM/46

- WDM/768 Gb/s) C-band dense SDM transmission over 205.6-km of single-mode heterogeneous multi-core fiber using 96-Gbaud PDM-16 QAM channels,” in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2017, p. Th5B.1. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2017-Th5B.1>
- [194] H. Takara et al., “120.7-Tb/s (7 SDM/180 WDM/95.8 Gb/s) MCF-ROPA un epeateded transmission of PDM-32 QAM channels over 204 km,” in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2014, pp. 1–3.
- [195] J. van Weerdenburg et al., “138 Tbit/s transmission over 650 km graded-index 6-mode fiber,” in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2017, pp. 1–3.
- [196] T. Kobayashi et al., “2 × 344 Tb/s propagation-direction interleaved transmission over 1500-km MCF enhanced by multicarrier full electric-field digital back-propagation,” in *Proc. 39th Eur. Conf. Exhib. Opt. Commun. (ECOC)*, 2013, pp. 1–3.
- [197] G. Rademacher et al., “159 Tbit/s C+L band transmission over 1045 km 3-mode graded-index few-mode fiber,” in *Opt. Fiber Commun. Conf. Postdeadline Papers*, 2018, p. Th4C.4. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2018-Th4C.4>
- [198] B. J. Puttnam et al., “0.715 Pb/s Transmission over 2,009.6 km in 19-core cladding pumped EDFA amplified MCF link,” in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2019, p. Th4B.1. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2019-Th4B.1>
- [199] G. Rademacher et al., “172 Tb/s C+L band transmission over 2040 km strongly coupled 3-core fiber,” in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2020, p. Th4C.5. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2020-Th4C.5>
- [200] B. J. Puttnam, R. S. Luis, G. Rademacher, Y. Awaji, and H. Furukawa, “319 Tb/s Transmission over 3001 km with S, C and L band signals over > 120 nm bandwidth in 125 μm wide 4-core fiber,” in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2021, p. F3B.3. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2021-F3B.3>
- [201] R. Ryf et al., “Coupled-core transmission over 7-core fiber,” in *Proc. Opt. Fiber Commun. Conf. Postdeadline*, 2019, p. Th4B.3. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2019-Th4B.3>
- [202] K. Shibahara et al., “DMD-unmanaged long-haul SDM transmission over 2500-km 12-core × 3-mode MC-FMF and 6300-km 3-mode FMF employing intermodal interference cancelling technique,” in *Proc. Opt. Fiber Commun. Conf.*, 2018, p. Th4C.6. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2018-Th4C.6>
- [203] R. Ryf et al., “Long-distance transmission over coupled-core multicore fiber,” in *Proc. 42nd Eur. Conf. Opt. Commun. (ECOC)*, 2016, pp. 1–3.
- [204] K. Igarashi et al., “1.03-exabit/skm super-Nyquist-WDM transmission over 7,326-km seven-core fiber,” in *Proc. 39th Eur. Conf. Exhib. Opt. Commun. (ECOC)*, 2013, pp. 1–3.
- [205] J. Sakaguchi et al., “Realizing a 36-core, 3-mode fiber with 108 spatial channels,” in *Proc. Opt. Fiber Commun. Conf. Post Deadline Papers*, 2015, p. Th5C.2. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2015-Th5C.2>
- [206] T. Mizuno et al., “32-core dense SDM unidirectional transmission of PDM-16 QAM signals over 1600 km using crosstalk-managed single-mode heterogeneous multicore transmission line,” in *Proc. Opt. Fiber Commun. Conf.*, 2016, p. Th5C.3. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2016-Th5C.3>
- [207] T. Gui, X. Wang, M. Tang, Y. Yu, Y. Lu, and L. Li, “Real-time demonstration of homodyne coherent bidirectional transmission for next-generation data center interconnects,” *J. Lightw. Technol.*, vol. 39, no. 4, pp. 1231–1238, Feb. 15, 2021.
- [208] L. Wang et al., “First real-time MIMO-free 800 Gb/s DP-64QAM demonstration using bi-directional self-homodyne coherent transceiver,” in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2021, p. Th3C2.
- [209] L. Galdino et al., “Optical fibre capacity optimisation via continuous bandwidth amplification and geometric shaping,” *IEEE Photon. Technol. Lett.*, vol. 32, no. 17, pp. 1021–1024, Sep. 1, 2020.
- [210] G. Böcherer, P. Schulte, and F. Steiner, “Probabilistic shaping and forward error correction for fiber-optic communication systems,” *J. Lightw. Technol.*, vol. 37, no. 2, pp. 230–244, Jan. 1, 2019.
- [211] X. Chen et al., “Single-wavelength, single-polarization, single- photodiode Kramers–Kronig detection of 440-Gb/s entropy-loaded discrete multitone modulation transmitted over 100-km SSMF,” in *Proc. IEEE Photon. Opt. (IPC)*, Oct. 2017, pp. 1–2.
- [212] D. Che, F. Yuan, and W. Shieh, “200-Gb/s polarization-multiplexed DMFT using Stokes vector receiver with frequency-domain MIMO,” in *Proc. Opt. Fiber Commun. Conf.*, 2017, pp. 1–3.
- [213] D. Che, C. Sun, and W. Shieh, “Single-channel 480-Gb/s direct detection of POL-MUX IQ signal using single-sideband Stokes vector receiver,” in *Proc. Opt. Fiber Commun. Conf.*, 2018, pp. 1–3.
- [214] *Implementation Agreement for a 400ZR Coherent Optical Interface*. Accessed: Aug. 17, 2022. [Online]. Available: <https://www.oiforum.com/technical-work/hot-topics/400zr-2/>

ABOUT THE AUTHORS

Mark Shtaiif (Fellow, IEEE) is currently a Professor of electrical engineering with Tel Aviv University, Tel Aviv, Israel. His fields of research focus primarily on fiber optics and optical communication systems. Within this general area of activity, he integrates the fields of optics, quantum theory, nonlinear systems, communications theory, information theory, and signal processing. Over the years, he has contributed to a variety of topics including optical amplification, analysis of nonlinear propagation, polarization-related phenomena, analyses of noise and signal detection, quantum information in fiber systems, and fundamental limits to optical communications.

Dr. Shtaiif is a Fellow of the Optica (formerly the Optical Society of America).

Cristian Antonelli (Senior Member, IEEE) is currently an Associate Professor of electromagnetic fields with the Department of Physical and Chemical Sciences, University of L'Aquila, L'Aquila, Italy, where he heads the Laboratory of Optics and Photonics of INCIPICT. His research interests are in the area of fiber-optic communications and include primarily the study of propagation effects and novel transmission schemes and also in the context of quantum communications.

Dr. Antonelli is a Fellow of the Optica (formerly the optical society of America). He also serves as a Topical Editor for *Optics Letters*.

Antonio Mecozzi (Fellow, IEEE) is currently a Professor with the Department of Physical and Chemical Sciences, University of L'Aquila, L'Aquila, Italy. His areas of interest include studies on soliton transmission, laser mode-locking, nonlinear propagation in single-mode and multimode fibers, polarization mode dispersion, physics and applications of semiconductor optical amplifiers, and optical amplification and noise.

Dr. Mecozzi is a Fellow of the Optica (formerly the Optical Society of America). He also serves as an Associate Editor for *Optica*.

Xi Chen (Senior Member, IEEE) received the Ph.D. degree in optical fiber communication from The University of Melbourne, Melbourne, VIC, Australia, in 2012.

From 2013 to 2015, she was a Research Fellow with The University of Melbourne, conducting research on optical fiber transmission. Since 2015, she has been with the Nokia Bell Labs, Murray Hill, NJ, USA. Her current research interests include advanced digital signal processing and fiber transmission.

Dr. Chen serves as an Associate Editor for *Optics Express*.