INVITED PAPER

The Role of Parallelism in the Evolution of Optical Fiber Communication Systems

This article serves as an introductory paper, highlighting the state of the art in fiber-optic transmission systems, pointing to capacity scalability issues, and laying out various capacity scalability options.

By Werner Klaus, Peter J. Winzer, *Fellow IEEE*, and Kazuhide Nakajima, *Member IEEE*

ABSTRACT | In order to overcome the capacity limitations of current lightwave systems based on the single-mode optical fiber, massively parallel transmission in the spatial domain [space-division multiplexing (SDM)] supported by extended parallelism in the frequency domain (ultrawideband (UWB) systems) must be used. This article reviews key aspects of parallel transmission systems as the only significant capacity scaling option going forward and discusses the various tradeoffs on an architectural level and a hardware integration level. In doing so, this article also serves as an introduction to the more detailed accounts of fiber-optic systems and their future scaling within this Special Issue of the Proceedings of the IEEE.

KEYWORDS | Fiber capacity limits; hollow-core fiber (HCF); multicore fiber (MCF); multimode fiber (MMF); optical communications; space-division multiplexing (SDM); ultrawideband (UWB) transmission; wavelength-division multiplexing (WDM).

I. INTRODUCTION

Some 25 years ago, it was still widely believed that optical fiber would provide enough capacity for all practical purposes of data exchange that one could ever think of, frequently supported by estimates of the maximum information that a human brain would be able to process. This perception quickly changed during the first decade of this millennium [1], with the introduction of smartphones and broadband Internet services and, most importantly, with the rapid growth of the Internet of Things (IoT), where machine-to-machine traffic has become the dominant source of data exchange across both wireline and wireless applications, taking human physiology out of the picture. Fueled by the increasing popularity of video-streaming and cloud computing, the demand for higher data rates in communication networks has been growing at an unabated annual growth rate of 40%–60% year over year, i.e., network traffic has been doubling every 1.5 to two years [2], [3]. Only a portion of this network traffic growth is accommodated by new fiber deployments, with an estimated six billion kilometers of single-mode fiber (SMF) installed worldwide today, growing by about 15% each year [3], [4]. The rest of the 40%–60% of annual network traffic growth is accommodated by putting into high-capacity service the already installed base of currently unlit ("dark") fiber and/or by upgrading outdated lightwave systems to state-of-the-art capacities. The scaling deficiencies of these systems, however, became clear to researchers around 2010 when a good understanding of the fiber-optic channel's transmission capacity limit had been obtained [5] and concerns about the possibility of a capacity crunch had been clearly voiced [6]. Each year, optical fiber networks are more closely approaching their Shannon limit estimates, i.e., estimates for the maximum rate at which data can be reliably transmitted, leaving only little room for further capacity improvements [7], [8]. There is no other transmission medium in sight that can handle as massive traffic demands as optical fiber, and it is highly unlikely that the transition to order-of-magnitude higher carrier frequencies, which was accomplished when low-loss optical fiber replaced copper cables and directional microwave links starting in the late 1970s, can be successfully repeated [9]. Hence, optical fiber communications

Manuscript received 7 December 2021; revised 6 July 2022; accepted 4 August 2022. Date of current version 11 November 2022. *(Corresponding author: Werner Klaus.)*

Werner Klaus is with the National Institute of Information and Communications Technology (NICT), Tokyo 184-8795, Japan (e-mail: klaus@nict.go.jp).
Peter J. Winzer is with Nubis Communications, New Providence, NJ 07974 USA.
Kazuhide Nakajima is with the Access Network Service Systems Laboratories, Nippon Telegraph and Telephone (NTT) Corporation, Ibaraki 305-0805, Japan.

Digital Object Identifier 10.1109/JPROC.2022.3207920

technologies will have to continue to ensure network traffic scalability for future communications services, and their further substantial scalability has become one of the most crucial topics within the fiber-optics community.

The tremendous progress in the field of fiber-optic communications over the last 50 years is due to an armada of researchers and engineers with wide-ranging expertise encompassing optics, electronics, and optoelectronic device physics, computer science, analog and digital signal processing (DSP), information theory, and network design, who have been relentlessly looking for ways to better understand and overcome the underlying physics and engineering challenges. Record transmission rates over a single strand of fiber achieved in today's research labs range from ~250 Tb/s in SMF [10] to ~10 Pb/s using multicore fibers (MCFs) [11], [12]. Commercial terrestrial systems carry up to ~70 Tb/s per fiber [8], [13], and the currently highest-capacity submarine cable in operation, the Dunant system owned by Google [14], can carry an aggregate of >300 Tb/s across a 6600-km trans-Atlantic distance using 12 fiber pairs [15]. In addition, the Grace Hopper system, also built for Google and to be completed in 2022, will carry an aggregate of >350 Tb/s over 6300 km across the Atlantic using 16 fiber pairs [16]. Furthermore, NEC announced that it will build a 24-fiber pair submarine cable for Meta, with a total capacity of 500 Tb/s [17], and HMN Technologies Company announced a next-generation submarine repeater prototype designed for 32 fiber pairs and a total capacity transmission beyond 700 Tb/s [18], indicating that it is only a matter of a few years until submarine cables will approach the 1-Pb/s mark per cable.

The purpose of this Special Issue of the PROCEEDINGS OF THE IEEE is to give an overview of the state of the art in fiber-optic communications technologies and an outlook on how these technologies will likely evolve in the future. While each paper in this Special Issue addresses a specific facet in more detail, this introductory paper examines general limits to optical fiber capacity scaling and approaches to overcome these limits. Throughout our discussions, it will become evident that *massive parallelism* is the only viable upgrade strategy to deal with a permanently looming capacity crunch. In this context, parallelism refers to the efficient use of spatial and spectral diversities to maximize transmission capacity while minimizing a transmission system's overall cost and energy consumption.

II. HISTORICAL PERSPECTIVE

The history of optical fiber communications to this day can be roughly divided into three major eras [2], [3], [6]. One may argue that the first era was set off in the second half of the 1970s with inventions such as vapor-phase axial deposition (VAD) that enabled mass production of high-quality optical fibers [19], [20] and the announcement of the first 0.2-dB/km SMF at a wavelength of 1.55 μm [21], following the pioneering predictions of Nobel laureate Charles Kao in 1966 [22] and the pioneering work at Corning in 1970 on bringing down the attenuation of optical fibers for the first time to below 20 dB/km at a wavelength of 0.633 μm [23]. Toward the end of the 1970s, fiber losses were already remarkably close to the fundamental limit of pure silica core fiber losses with the best fibers made today achieving losses of 0.142 dB/km at a wavelength of 1.56 μm [24], [25]. During this first era, optical transmission systems employed ON–OFF keying (OOK), where laser light was simply switched ON and OFF to represent logical ones and zeros. The corresponding direct-detection receivers, i.e., optical power detectors, enabled optoelectronic regeneration and had to be periodically placed at distances of some 10 km. Data rates ranged from several 10 Mb/s to ~1 Gb/s. The transmission medium was originally based on multimode fibers (MMFs) but quickly transitioned to SMFs to eliminate modal dispersion, thereby achieving better transmission performance.

The second era was ushered in by the commercial introduction of the erbium-doped fiber amplifier (EDFA) [26], [27] in the mid-1990s [28]. In combination with chromatic dispersion (CD) management techniques [29] to reduce the impact of otherwise detrimental nonlinear propagation effects (in particular four-wave mixing due to the optical Kerr effect [30]), full exploitation of wavelength-division multiplexing (WDM), i.e., *parallelism* in the spectral domain, across the EDFA's gain spectrum was made possible. In an amplified WDM system, individual communications signals are transmitted simultaneously on sufficiently separated optical carrier frequencies. Due to increasingly more frequency-stable laser and optical filter technologies, initial WDM channel spacings of several 100 GHz were constantly reduced to as little as 25 GHz. Today, WDM spacings only slightly exceed the transponder's symbol rate and are typically between 50 and 200 GHz. By periodically compensating for optical losses in the optical domain, the EDFA enabled unregenerated transmission over many 1000 kilometers of optical fiber, i.e., the need for optoelectronic conversion was largely pushed toward the end nodes of a link. The above innovations led to unprecedented growth in long-haul fiber capacities in the 1990s and enabled cost-effective high-speed data transmission over transoceanic distances. This not only largely improved voice communications (many international phone calls were still relayed via satellites at the time) but also became the key enabler of a universally accessible and widely affordable Internet over the following years. To a large part within that second era, transmission capacities per fiber of commercial WDM systems increased by a factor of 2 (i.e., by 100%) every year, from a few 10 Gb/s in the mid-1990s to about 5 Tb/s in the 2000s [31]. These systems employed bit rates of up to 40 Gb/s per wavelength using binary modulation [OOK and differential phase shift keying (DPSK)], which led to severe problems due to: 1) polarization mode dispersion (PMD) [32] causing pulse broadening due to randomly varying optical birefringence in the fiber and 2) the need to carefully adjust residual CD through tunable optical dispersion compensators. The reduced tolerance to noise from in-line optical amplifiers (OAs) was counteracted by distributed (Raman) amplification,

advanced modulation (DPSK), and more powerful forward error correction (FEC). It was clear at the time that the step to 100 Gb/s per wavelength could not be accomplished in a commercially meaningful way through binary modulation at 100 GBaud[1] but rather required higher order modulation, such as differential quadrature phase shift keying (DQPSK), which was initially used in conjunction with direct-detection receivers to demonstrate 100-Gb/s optical networking [33]. Fortunately, the inherently faster paced progress of CMOS integrated circuits at 40% per year as opposed to the 20% progress in per-wavelength interface rates [3] enabled high-speed analog-to-digital conversion (ADC) and digital-to-analog conversion (DAC) together with the associated DSP to digitize and digitally process signals in the multi-GBaud range. In the early 2000s, high-speed ADCs started to be used for advanced 10-Gb/s detection of OOK using maximum-likelihood sequence estimation (MLSE) [34] and high-speed DACs were used for digitally predistorted 10-Gb/s OOK [35].

From then on, DSP rapidly enabled the third era of digital-coherent systems [36], [37], [38], [39], [40], [41], [42], [43]. The concept of digitally coherent optical receivers dates back to the early 1990s [44] and is, today, almost exclusively used by high-speed optical transmission systems exceeding a reach of $\sim$100 km. Digital-coherent systems phase-lock a free-running local oscillator (LO) laser to the received signal through DSP in an application-specific integrated circuit (ASIC) after high-speed sampling/digitizing, thereby eliminating the need for analog LO frequency/phase locking that posed a major practical problem in earlier coherent (heterodyne and homodyne) systems, which did not make it into volume production [45]. In addition to digital phase-locking, modern coherent receivers also digitally undo any polarization rotations occurring along the transmission fiber, which allows the use of polarization-division multiplexing (PDM) and digitally compensates for nearly arbitrary amounts of CD and PMD. High-speed coherent systems also obviate the need for periodic in-line dispersion compensation to manage fiber nonlinearities, which considerably simplifies system design. In addition, the use of coherent detection allows one to modulate simultaneously the real and imaginary parts of the optical field via its two quadratures (i.e., sine and cosine components) and, thereby, enables the use of modulation formats at higher spectral efficiency (SE) than OOK and DPSK, such as quadrature amplitude modulation (QAM) and phase shift keying (PSK). A further doubling of spectral efficiencies is achieved by using the optical field's two polarization states by means of PDM. Together this has enabled transmission at rates well above one bit of information per pulse (or equivalently one bit per symbol, one bit per channel use, or one bit per second per Hertz of channel bandwidth). Sophisticated modulation and FEC schemes can be used in concert with DSP to enable systems to operate close to their theoretical (Shannon) limits. Using all these advanced communications techniques, aggregate per-fiber capacities in commercial products have increased by roughly a factor of 7 over the last decade (i.e., $\sim$20%/year) to a total of $\sim$70 Tb/s for $C+L$-band system operation (see Fig. 9 for detailed information on the fiber-optic transmission band nomenclature).

Record research experiments based on $C+L$-band SMF transmission reported at top conferences for distances above 1000 km have hardly exceeded 100 Tb/s, which let the 100-Tb/s mark often be quoted as the "capacity limit" of SMF. Note, however, that the 100-Tb/s mark certainly does *not* represent a hard capacity limit of SMF-based systems, as we will discuss in Section III-B. Nevertheless, operating long-haul systems at this capacity level indicates the proximity to operating near the Shannon limit and the near exhaustion of the EDFA bandwidth, which makes it increasingly difficult to deal effectively with further growth in data traffic. These limitations suggest that we are on the verge of entering the fourth era of optical communications, which may become known as the era of "massive parallelism" and will manifest itself in space-division multiplexing (SDM), complemented in as far as economically beneficial by ultrawideband (UWB) transmission. SDM describes the multiplexing of signals in the spatial domain (in the form of parallel fibers, parallel fiber cores, or copropagating spatial modes), while UWB (or multiband) transmission exploits the simultaneous use of multiple transmission bands in a fiber's low-loss transmission window, known as the O-, E-, S-, and U-bands to complement the C- and L-bands supported by EDFAs and used in today's products. Such UWB systems come at the cost of more than twice the minimum possible fiber propagation loss and, as we will show in Section IV, are not nearly as scalable as SDM systems. Nevertheless, UWB approaches may be used as a stop-gap solution where the deployment of new cables cannot keep up with traffic demand in a short term, as well as in conjunction with SDM approaches in future systems, subject to applicable technoeconomic tradeoffs (see [46] and [47] for an example of technoeconomic tradeoffs in submarine systems and [48] for technoeconomic tradeoffs in data-center interconnect systems).

In the research community, SDM began to attract attention with the announcement of the first successful data transmission experiments based on direct detection through a 100-m multimode MCF using VCSELs [49], succeeded shortly thereafter by single-mode MCF transmission experiments over more than 10 km [50] with direct detection, and over more than 16 km [51] as well as 76 km [52] with coherent detection. In terms of exploiting multiple fiber modes in an MMF as parallel data transmission paths, early proposals go back to the 1980s [53], with multiple-input multiple-output (MIMO) processing of fiber modes proposed in the early 2000s [54]. The first few-mode fiber (FMF) transmission experiment that combined mode-selective optical signal launch and detection of a fiber's complete mode basis with digital coherent detection and MIMO DSP was reported

[1]Baud is the unit of the symbol rate, which refers to the number of pulses per second, i.e., 1 Baud = 1 symbol per second. One symbol (i.e., pulse) can represent one bit or several bits.

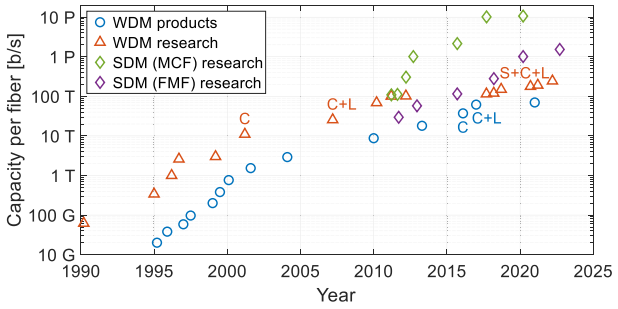

Fig. 1. ***Evolution of transmission capacity per fiber. Letters indicate transmission bands.***

in 2011 [55]. These and all subsequent SDM experiments (based on MFCs and FMFs, as well as variations and combinations thereof) demonstrated a clear advantage over SMF in terms of aggregate per-fiber spectral efficiencies and ultimately per-fiber capacities, as illustrated in Fig. 1. Moreover, coupled-core MCFs (CCFs) were also shown to outperform SMFs in terms of transmission distance, owing to a higher tolerance to nonlinear signal distortions [56].

In terms of field installations, a unique field-deployed SDM testbed is available in the city of L'Aquila, Italy [57], [58], [59], and SDM fibers are expected to make their first appearance in operating networks over the next decade. Laboratory research progress on SDM fibers and related components, as well as network designs based on SDM concepts, have received tremendous research attention over the past decade. Furthermore, considerable efforts toward standardization and refined fabrication processes [60] are under way to bring SDM technologies closer to their adoption in commercial systems.

III. FIBER-OPTIC SYSTEM CAPACITY CONSIDERATIONS

Estimating capacity limits of various fiber-optic transmission systems has been a topic of increasing interest within the fiber-optics community over the past 15 years [1], [7], [61], [62], [63], [64], [65], [66]. An invited paper on this topic written by leading experts in the field is included in this Special Issue [67]. This section summarizes key considerations relevant to the capacity limit of various types of optical fiber by taking into account frequently ignored but practically highly relevant limitations due to a transponder's practically unavoidable noise floor.

A. Shannon Limit Estimate

The total capacity C of a polarization-multiplexed WDM transmission system can be written as

$$C = 2\sum_{m=1}^{M}\sum_{n=1}^{N} R_{mn}. \tag{1}$$

Note that, in line with common industry terminology, we loosely use the term "capacity" throughout this article to denote the net information throughput of a fiber-optic transmission system. In this context, the term "capacity" as used here may denote a commercially specified system capacity, an experimentally measured system capacity, or may more rigorously refer to a system's achievable information rate (AIR), whose upper bound is given by the information-theoretically stricter definition of "capacity" (see [67] and [68] for an overview of various capacity metrics used in the analysis of fiber-optic transmission systems). The factor of 2 in (1) represents polarization multiplexing,[2] M is the number of parallel spatial paths (spatial parallelism), N is the number of WDM channels (spectral parallelism), and R_{mn} is the net information bit rate of a WDM channel. Modern systems can adapt their per-channel bit rate to the conditions of the respective WDM channel using, e.g., probabilistic constellation shaping (PCS) [68], [70], [71], [72]: increasing the probability of the lower power inner constellation points and decreasing the probability of the higher power outer constellation points of a high-order modulation format (e.g., a 64-QAM template), a Gaussian constellation distribution can be approximated. A Gaussian constellation distribution is optimum for an additive white Gaussian noise (AWGN) channel. Using PCS, the width of the Gaussian constellation distribution can be finely adapted to the channel's noise conditions. Note, in this context, that the fiber-optic channel in many modern optical transmission systems of practical interest can be modeled in good approximation by an AWGN channel, as we also assume throughout this article.

In the idealized case of a homogenous WDM system, where each WDM channel carries the same net information bit rate, i.e., $R_{mn} = R$, (1) can be simplified to

$$C = 2MNR. \tag{2}$$

A homogeneous WDM system is defined as a system where each WDM channel: 1) uses the same modulation and coding at a uniform channel spacing; 2) has a constant channel bandwidth of $B_{\text{ch}} = B/N$; 3) has a constant launch power of $P_{\text{ch}} = P/N$, with P being the total launch power; and 4) encounters the same transmission-line properties such as propagation loss, chromatic dispersion, amplifier noise figure etc.

The SE of an optical communication system is defined as the capacity per unit bandwidth, usually per spatial path and frequently (as also used in this article) per

[2]As polarization is a fundamental modal property of waveguides, various authors sometimes consider 2M spatial modes by folding the polarization property of light into the spatial distribution of the electromagnetic mode field. A separation between polarization and spatial mode properties is possible for circular-symmetric fibers with a small index difference between the core and cladding, which leads to quasi-linearly polarized (LP) pseudomodes [69].

polarization, i.e.,

$$\mathrm{SE} = \frac{C}{2MB} \tag{3}$$

where B is the optical system bandwidth encompassing all WDM channels. For a homogeneous WDM system, the SE can be written as

$$\mathrm{SE} = \frac{C}{2MNB_{\mathrm{ch}}}. \tag{4}$$

Using (2), the per-channel net information bit rate in such a uniform WDM system is

$$R = B_{\mathrm{ch}}\mathrm{SE}. \tag{5}$$

Under the assumption of an AWGN channel, the SE is upper bounded[3] by the Shannon limit [74], [75]

$$\mathrm{SE} = \frac{R_S}{B_{\mathrm{ch}}}\log_2\left(1+\mathrm{SNR}\right) \tag{6}$$

where R_S represents the symbol rate and SNR is the *electrical* signal-to-noise-ratio (i.e., the total signal power divided by the total noise power[4] within a bandwidth corresponding to the symbol rate R_S, assuming the use of Nyquist pulses). For a detailed discussion of the information theory of fiber-optic transmission systems, the reader may refer to, e.g., [7] and [67]. In practical systems, the SNR is furthermore affected by an implementation penalty $\eta\,(\eta \leq 1)$ accounting for effects, such as nonideal hardware and nonideal FEC decoding in the receiver, which are difficult to capture in the SNR (see, e.g., [76]), resulting in

$$\mathrm{SE} = \frac{R_S}{B_{\mathrm{ch}}}\log_2\left(1+\eta\mathrm{SNR}\right). \tag{7}$$

The symbol rate is typically chosen smaller than the WDM channel bandwidth, $R_S \leq B_{\mathrm{ch}}$. With ideal Nyquist pulses (i.e., sinc pulses), one can in principle achieve $R_S = B_{\mathrm{ch}}$ without incurring WDM crosstalk between neighboring channels. Practical systems use pulse shaping (sometimes also called spectral shaping) to generate root-raised-cosine (RRC) shaped pulses. These become Nyquist pulses after matched filtering at the receiver. RRC pulses have a finite spectral roll-off that broadens their spectral extent beyond R_S. Pulse shaping involves a balance between digital shaping complexity, WDM channel spacing, WDM crosstalk, and bandwidth requirements of the transponder electronics. Note that pulse shaping applies to the shaping of the *analog signal waveform* and should not be confused with constellation shaping (mentioned earlier in this section), which applies to the shaping of the *digital symbol alphabet* [77]. Together with the use of a guard band between WDM channels, to accommodate laser and optical filter frequency drifts, $R_S < B_{\mathrm{ch}}$ in practical systems.[5] The relationship between R_S and B_{ch} is given by

$$R_S = \frac{1}{(1+\rho_{ro})\,(1+\rho_{gb})}B_{\mathrm{ch}} \tag{8}$$

where ρ_{ro} $(\rho_{ro} \geq 0)$ denotes a roll factor and ρ_{gb} $(\rho_{gb} \geq 0)$ is a guard band ratio. Today's commercial systems typically use R_S/B_{ch} of ∼0.6 for datacenter interconnect (DCI) applications (<100 km), ∼0.6–0.9 for terrestrial mesh networking applications, and even slightly above 0.9 for some submarine applications.

Spatial parallelism is achieved by multiplexing each data signal onto a different spatial path (i.e., parallel fibers, parallel cores in an MCF, or multiple modes in an FMF or CCF). *Spectral parallelism* uses different carrier wavelengths for each WDM channel. The bit rate R of a WDM channel is determined by the underlying modulation and coding within a channel bandwidth B_{ch}. The symbol rate that may be used by a WDM channel is determined by the speed of optical modulators and detectors, including the associated analog, ADC, and DAC electronics. Commercial systems operate at symbol rates in the 25–150-GBaud range, leading to typical WDM channel bandwidths B_{ch} between 50 and 200 GHz. In research experiments, symbol rates of ∼200 GBaud have been demonstrated [78], [79]. Interestingly, symbol rates in research experiments have not shown any notable progress much beyond the 200-GBaud mark since 2016 and seem to have saturated around that level, as the associated electronics with electrical bandwidths beyond 100 GHz are exceedingly difficult to build.[6]

In the following discussion, we assume that $R_S = B_{\mathrm{ch}}$, i.e., a roll-off factor of 0 and no guard bands. The SE then also represents the information rate, i.e., the number of information bits per modulation symbol. While it may seem straightforward to increase the SE [and, thus, the bit rate at a fixed symbol rate; cf. (5)] by using a higher order QAM template for PCS, the logarithmic scaling of the SE with the SNR [cf. (6)] makes a K-fold increase in SE require at least a 2^K-fold increase in SNR, which is increasingly difficult to achieve for higher K.

To appreciate the difficulties in increasing the SNR, we examine the factors that influence the SNR in a coherent

[3]If the channel is not AWGN, higher capacities are in principle achievable [67]. The fact that the exact statistics of some optical noise sources are unknown hints to the difficulties in speaking of a true information-theoretic "Shannon limit" of the fiber-optic channel and let expressions such as (6) only be Shannon limit estimates [73].

[4]In a single-polarization system, the SNR takes into account signal power and noise power in a single polarization. In a polarization-multiplexed (or dual-polarization) system, the SNR takes into account signal and noise powers in both polarizations. In this article, we consider only the practically more relevant dual-polarization case.

[5]In some system experiments, partial-response modulation, faster-than-Nyquist modulation, and single-sideband modulation are being used, all of which can operate at $R_S > B_{\mathrm{ch}}$. However, none of these modulation schemes offers an SE advantage over complex modulation using ideal Nyquist pulses, i.e., $\mathrm{SE} \leq \log_2\left(1+\mathrm{SNR}\right)$ remains the limit for all such systems.

[6]In fact, the most recent announced record experiment with a symbol rate of 260 GBaud [80] is also limited by 75-GHz bandwidth electronics and therefore had to resort to optical equalization. Also note that the highest available symbol rate does not necessarily lead to the highest bit rate [77].

polarization multiplexed system. In such a system, the SNR is given by

$$\mathrm{SNR} = \frac{P_{\mathrm{ch}}}{P_{\mathrm{SH}} + P_{\mathrm{ASE}} + P_{\mathrm{TRX}} + P_{\mathrm{NLIN}} + P_{\mathrm{XT}}} \tag{9a}$$

where P_{ch} denotes the dual-polarization signal power per WDM channel launched into the fiber at the transmitter and restored by every OA following a transmission span (see Fig. 7). Alternatively, the SNR can also be expressed as

$$\begin{aligned}\frac{1}{\mathrm{SNR}} &= \frac{1}{\mathrm{SNR}_{\mathrm{SH}}} + \frac{1}{\mathrm{SNR}_{\mathrm{ASE}}} + \frac{1}{\mathrm{SNR}_{\mathrm{TRX}}} + \frac{1}{\mathrm{SNR}_{\mathrm{NLIN}}} + \frac{1}{\mathrm{SNR}_{\mathrm{XT}}} \\ &= \frac{1}{\mathrm{SNR}_1} + \frac{1}{\mathrm{SNR}_0} + \frac{1}{\mathrm{SNR}_{-2}} \\ &= \frac{1}{\mathrm{SNR}_{\mathrm{L}}} + \frac{1}{\mathrm{SNR}_{\mathrm{NLIN}}}. \end{aligned} \tag{9b}$$

The first line of (9b) separates the total system SNR into its individual physical effects. The second line separates the SNR according to its dependency on P_{ch}: a term SNR_1 that depends linearly on P_{ch} ($\mathrm{SNR}_{\mathrm{SH}}$ and $\mathrm{SNR}_{\mathrm{ASE}}$), a term SNR_0 that is independent of P_{ch} ($\mathrm{SNR}_{\mathrm{TRX}}$ and $\mathrm{SNR}_{\mathrm{XT}}$), and a term SNR-$_2$ that depends inverse quadratically on P_{ch} ($\mathrm{SNR}_{\mathrm{NLIN}}$).[7] The third line then groups SNR_0 and SNR_1 into a linear SNR term $\mathrm{SNR}_{\mathrm{L}}$ and rewrites SNR-$_2$ as the nonlinear SNR term $\mathrm{SNR}_{\mathrm{NLIN}}$. P_{SH} is shot noise (typically dominated by LO shot noise in a coherent receiver), and P_{ASE} is noise due to accumulated amplified spontaneous emission (ASE) generated by in-line optical amplification within the WDM channel's signal bandwidth. Expressions for P_{SH} and P_{ASE} are given in the Appendix. Practical transponders exhibit a noise floor resulting in a maximally achievable SNR, denoted $\mathrm{SNR}_{\mathrm{TRX}}$, which is reflected by $P_{\mathrm{TRX}} = \kappa P_{\mathrm{ch}}$ (i.e., $\mathrm{SNR}_{\mathrm{TRX}} = 1/\kappa$) in (9) with $\kappa \ll 1$ [82], [83], [84]. The nonlinear interference noise (NLIN) $P_{\mathrm{NLIN}} = \chi P_{\mathrm{ch}}^3$ is an equivalent noise term due to the presence of Kerr nonlinearities in the optical transmission fiber. The NLI coefficient χ can be derived using a time-domain model [65], [85], [86], or from analytic equations based on the frequency-domain Gaussian Noise (GN) model [87], [88]. Key equations underlying the results shown in this paper are summarized in the Appendix. Finally, $P_{\mathrm{XT}} = \xi P_{\mathrm{ch}}$ captures uncompensated linear crosstalk effects, such as crosstalk from imperfect optical add/drop multiplexers (MUXs), crosstalk from multipath interference (MPI), crosstalk from neighboring WDM channels, length-dependent crosstalk from guided acoustic-wave Brillouin scattering (GAWBS) [89], [90], [91], and crosstalk between parallel spatial paths in an SDM system. Some transponders actively compensate for linear crosstalk through digital echo cancellation or MIMO DSP. P_{XT} can be relevant in form of GAWBS at ultralong transmission distances or in an SDM system using, e.g., nominally uncoupled MCFs (as discussed in more detail in Section V-A1). Although some noise terms in (9) may not be strictly additive Gaussian, assuming them to be additive Gaussian has so far not contradicted any measured data from practical systems in terms of their capacity limit estimates.

At sufficiently low optical powers, both P_{NLIN} and P_{TRX} are negligible, and the SNR increases linearly with optical launch power until the SNR saturates due to P_{TRX} and/or P_{XT}. At optical signal powers that are high enough to let P_{NLIN} dominate all other noise terms, the SNR becomes inversely proportional to the square of the signal power. These dependencies of the SNR on the optical signal power are illustrated in Fig. 2(a). Typical optical signal launch powers in terrestrial long-haul systems are chosen to maximize the SNR, at a point where P_{NLIN} is not negligible but also not dominant, and the SNR peaks at an optimum launch power [7], [67]. Using this peak SNR in the Shannon formula [cf. (6)] has become widely known as the "nonlinear Shannon limit" or the "fiber-optic Shannon limit," but stated more precisely, due to the many underlying assumptions, this process only yields an *estimate* of the practically achievable capacity of whatever particular fiber transmission system is being investigated.

The transponder noise floor P_{TRX} is mainly due to practical limitations from the transmitter and receiver electronics. This includes quantization noise from ADCs and DACs, an imbalance (skew) between the signals in the modulator's in-phase and quadrature arms used to generate QAM formats [92], [93], as well as any magnitude and phase mismatch in the frequency responses between nominally identical high-speed analog electronic components, which generally becomes more pronounced at high symbol rates. Other factors contributing to a residual noise floor comprise phase noise of the transmitter and LO laser, as characterized by the respective laser linewidths, and the inherent dithering of digital filter parameters in adaptive DSP control loops. This practically relevant noise floor, which is rarely considered in theoretical fiber capacity analyses, caps the system capacity even if the fundamental limitations due to ASE and NLIN would still allow for higher system throughputs. Typical high-speed optical transponders in commercial products operate at an $\mathrm{SNR}_{\mathrm{TRX}}$ of $\sim$20 dB [94], [95], record high-SE research results at much reduced symbol rates have achieved an $\mathrm{SNR}_{\mathrm{TRX}}$ of $\sim$30 dB [96], [97], and laboratory results exploiting special techniques, such as single-sideband modulation (SSB) in conjunction with image-band rejecting heterodyning at the receiver, have shown to reach an $\mathrm{SNR}_{\mathrm{TRX}}$ of $\sim$38 dB, albeit only at low symbol rates [93]. For the system analyses presented in this article, it is reasonable to assume that realistic transponders in the foreseeable future will not exceed an $\mathrm{SNR}_{\mathrm{TRX}}$ of 30 dB (i.e., $\kappa = 10^{-3}$). The limiting effect on the SNR due to $\mathrm{SNR}_{\mathrm{TRX}}$ is clearly visible in Fig. 2(a), where the SNR peak reflecting the tradeoff between ASE and NLIN is substantially

[7] Note that SNR terms with other signal power dependencies exist as well but are neglected here as they are typically not dominant. For example, a term SNR_{-1} can include an insufficiently suppressed interference term due to self-mixing (direct detection term) of multiple channels being detected by an intradyne receiver, which can become important, e.g., in colorless wavelength drop situations with an insufficient common mode rejection ratio [81].

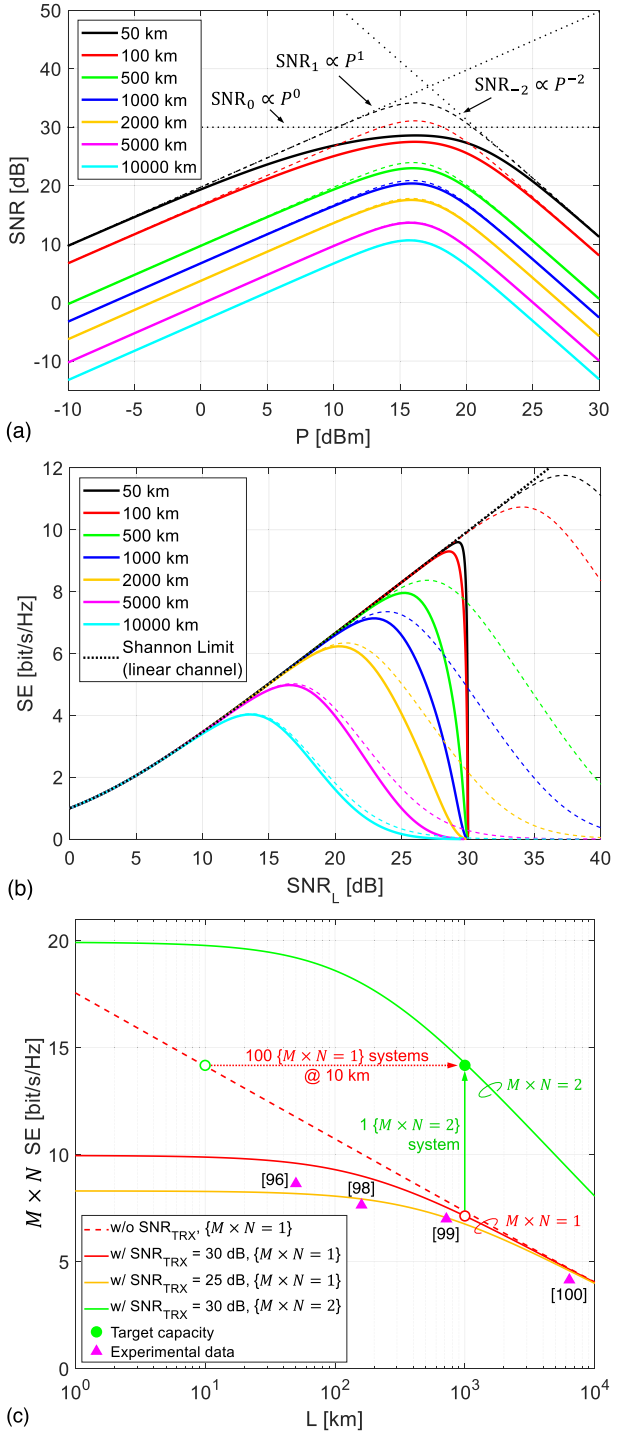

Fig. 2. *(a) SNR as a function of total launch power P (= NP_{ch}) with (solid lines) and without (dashed lines) a transponder SNR_{TRX} of 30 dB. (b) SE per polarization as a function of SNR_L with (solid lines) and without (dashed line) a transponder SNR_{TRX} of 30 dB. (c) Aggregate SE over M spatial and N spectral channels as a function of transmission distance L. The red and orange solid lines represent the cases for SNR_{TRX} of 30 and 25 dB, respectively.*

flattened for transmission distances below a few hundred kilometers. The severe impact of SNR_{TRX} can also be seen at longer transmission distances for ultralow-loss and ultralow-nonlinearity fibers, as discussed in Section IV-A1. The calculations shown in Fig. 2 assume distributed amplification, no implementation penalty (i.e., $\eta = 1$) and a homogenous WDM system across the $C + L$-band (11.5 THz) operating over a standard SMF. Key formulas and parameter assumptions used in these calculations are summarized in the Appendix.

Fig. 2(b) shows the achievable SE as a function of SNR_{L}, the portion of the SNR that comprises SNR_0 and SNR_1, i.e., is only due to linear effects. Neglecting, first, all nonlinearities, transponder noise floor, and crosstalk (i.e., $P_{\text{NLIN}} = 0$, $P_{\text{TRX}} = 0$, and $P_{\text{XT}} = 0$), SNR is linearly proportional to the signal launch power P_{ch}. As reflected by the black dotted curve in Fig. 2(b), this lets the SE grow unbounded with increasing P_{ch} at a rate of $\sim$1 bit/s/Hz per 3-dB increase in P_{ch} (or SNR_{L}) for $\text{SNR}_{\text{L}} \gg 1$ (i.e., SE $\approx \text{SNR}_{\text{L,dB}}/3$ for $\text{SNR}_{\text{L}} \gg 1$). The dashed curves include NLIN but still do not include a transponder noise floor or crosstalk yet. Increasing the signal launch power eventually lets P_{NLIN} become the dominant noise term, and the SE peaks at an SNR_{L} that is referred to as the nonlinear optimum. This peak SNR decreases approximately by 3 dB for each doubling of system reach L. Finally, when including P_{TRX} or a mathematically equivalent crosstalk term P_{XT}, SNR_{L} saturates at its respective floor (SNR_0) with increasing P_{ch}. In the case of Fig. 2(b), this happens at an SNR_{TRX} of 30 dB, as we assume no additional crosstalk here, i.e., $P_{\text{XT}} = 0$. This saturation can occur at an SNR_{L} that is significantly below the NLIN-ASE (nonlinear) optimum. At short distances, the SE drops sharply near $\text{SNR}_{\text{L}} = \text{SNR}_{\text{TRX}}$, which is because SNR_{TRX} can only be approached for very large P_{ch} where NLIN lets the SE drop to 0. This implies that practical systems up to about 100 km are capacity-limited by their transponder noise floors rather than by NLIN. This capacity limitation may, however, extend to many thousands of kilometers for ultralow-loss and ultralow-nonlinearity fibers, as discussed in Section IV-A1.

Fig. 2(c) displays the tradeoff between SE and transmission distance. The y-axis represents the aggregate SE over M spatial and N spectral channels. The red dashed line represents a system with $M \times N = 1$ without transponder noise floor, and the red and orange solid curves apply for an SNR_{TRX} of 30 and 25 dB, respectively. An SE below these curves is achievable by proper modulation and coding, while values above these curves are not achievable without spatial (M) or spectral (N) parallelism. The optical launch power is optimized for each transmission distance so that the system always operates at its peak SE [cf. Fig. 2(b)]. At the respective optimum signal launch power and in the absence of P_{TRX} (dashed line), the interplay of ASE and NLIN deteriorates the SE approximately linearly on a logarithmic scale, with a slope of almost exactly 1 bit/s/Hz for every doubling in transmission distance (equivalent to the above-mentioned $\sim$3-dB drop in SNR) [7]. This slope applies down to an SE of $\sim$2 bit/s/Hz (i.e., SNR of $\sim$5 dB) where the curve starts to level off as the "1+"

term within the logarithm of the Shannon formula [cf. (6)] starts to become nonnegligible [this regime is beyond the axis ranges of Fig. 2(c)].

In practical systems, effects such as electronically enhanced phase noise (EEPN) [101], [102] and GAWBS, both of which contribute to the SNR_0 term [cf. (9b)], also reduce SE (and, thus, capacity), but their effects are more prominent at large transmission distances. At distances below 1000 km, $\mathrm{SNR}_{\mathrm{TRX}}$ starts to dominate capacity and manifests in a capacity saturation. This saturation effect is also evident from the record experimental results shown as markers in Fig. 2(c). From these observations, it is also clear that the often-cited *capacity-distance product* cannot be used for an objective performance comparison of optical communication systems: increasing system reach by a factor m at a fixed SE is, in general, easier to achieve than increasing SE by the same factor m at a fixed reach. Instead of the frequently used capacity-distance product (or equivalently the SE $\times$ L product), the correct metric to compare optical fiber systems (for SNR $\gg$ 1 and at transmission distances where noise is not dominated by P_{TRX}) is $2^{\mathrm{SE}} \times L$ [48], [103]. In this regime, doubling transmission distance at fixed SE is a similar achievement as increasing the SE by 1 bit/s/Hz at a fixed distance.

In terms of system capacity scaling, consider as a target the aggregate SE of 14 bit/s/Hz at 1000 km [solid green marker in Fig. 2(c)], which implies a doubling in SE relative to a state-of-the-art system and is well beyond the 7 bit/s/Hz at 1000-km limit that is achievable through modulation and coding (red open circle marker). The SE of 14 bit/s/Hz could, in principle, be achieved only by modulation and coding if serially concatenating $100\{M \times N = 1\}$ systems at 14 bit/s/Hz with an individual reach of 10 km per system (green open marker on the red dashed line). Apart from being nonviable from an energy and cost perspective [103], this approach goes far beyond the capabilities of practical transponder electronics in terms of $\mathrm{SNR}_{\mathrm{TRX}}$ (red or orange solid curve) and must, therefore, be ruled out. Hence, the only practical option to double capacity is to employ *multiplexing parallelis*m, either by doubling the spectral channels ($N \rightarrow 2N$), or by doubling the number of spatial channels ($M \rightarrow 2M$), or more generally by doubling the product of spatial and spectral channels ($MN \rightarrow 2MN$). All these options are reflected by the green curve in Fig. 2(c). To accommodate even the lowest predicted traffic growth rates of 40% per year, such a doubling would be needed every other year. It is, thus, evident that the exploitation of massive spatial and spectral parallelism is mandatory to deal with exponential traffic growth throughout all network segments.

B. What Is the Actual Capacity Limit of an SMF?

Various estimates exist today for the nonlinear Shannon limit of different optical fiber types as a function of system bandwidth B and transmission distance L [1], [7], [61], [62], [63], [64], [65], [66]. A paper devoted to the information-theoretic aspects of the Shannon capacity is also part of this Special Issue [67]. However, an exact answer to this question with practically relevant parameter sets has not yet been and is unlikely to be derived, even under realistically idealized conditions. In addition, almost all accounts of fiber capacity limits assume ideal transponders without a noise floor represented by $\mathrm{SNR}_{\mathrm{TRX}}$. Furthermore, the fundamental ability and the realistic practicality to digitally compensate for fiber nonlinearities depend on whether point-to-point systems or optical mesh networks are being considered [7] and what kinds of WDM nonlinearity compensation (NLC) strategies are being pursued [104], [105]. Finally, the maximum total power $P_{\mathrm{tot}} \geq P$ (where P is the total signal launch power and represents the case without any Raman pumps) that can be safely propagated in an optical fiber is limited by the fiber fuse effect [106], a destructive fusion phenomenon occurring in the fiber center due to hot plasma formed at contaminated connector end faces, microcracks, or tight fiber bends. The fiber fuse threshold depends on the fiber design and is estimated to be between 1 and 2 W (30–33 dBm) for the safe operation of deployed communication fibers [107], [108]. In other words, a fiber capacity limit can certainly not be represented by a single number, such as the frequently quoted "100 Tb/s," even if that number captures the rough order of magnitude of the limit of practical long-haul fiber-optic communication system designs operating over today's commercial system bandwidths.

To illustrate the interplay of the parameters B, L, and $\mathrm{SNR}_{\mathrm{TRX}}$, Fig. 3 depicts a range of estimated capacities and optimum total launch powers P for a lightwave system as a function of transmission distance, calculated with the GN model (for details on the model, see the Appendix) for three transmission bands: the C-band (1530–1565 nm or 4.4 THz), the $C + L$-band (1530–1625 nm or 11.5 THz), and the $S + C + L$-band (1460–1625 nm or 20.9 THz). Some realistic assumptions for a system with lumped optical amplification are given in Table 1 where n_{sp} is the amplifier's spontaneous emission factor corresponding to the given noise figure if the amplifier is operated with sufficient gain, i.e., $G \gg 1$, as discussed in the Appendix. Assumptions about the fiber parameters are also stated in the Appendix. As exemplified in Fig. 3(a), if a system of this kind is operated only in the C-band, it could never reach the 100-Tb/s mark; if operated in the $C + L$-band, it could transmit 100 Tb/s up to ~3 000 km; and if operated in the $S + C + L$ band it could potentially transmit 100 Tb/s even beyond 10 000 km. The markers in Fig. 3(a) show recent experimental records. The main reason why short-reach transmission experiments for the C-, $C + L$-, and $S + C + L$-bands experience difficulties in reaching capacities above ~60, ~140, and ~260 Tb/s is due to a maximum transponder SNR below 25 dB, and why particularly ultrawideband systems have difficulties in reaching their Shannon limit estimates (cf. Fig. 11) is also due to higher amplifier noise figures and higher implementation penalties resulting from higher band-splitter/combiner losses and stimulated Raman scattering [113]. As shown in

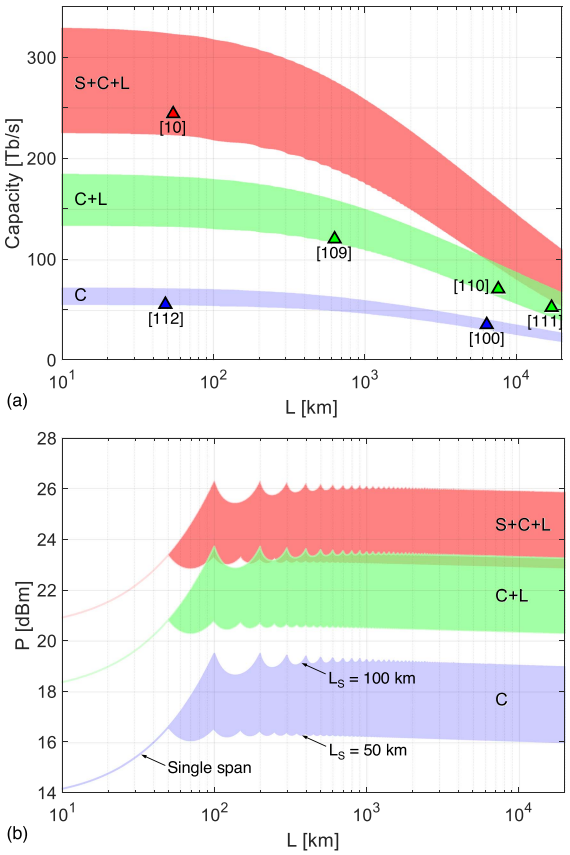

Fig. 3. *(a) Estimated range of the SMF capacity for the C-band (blue), the C + L-band (green), and the S + C + L band (red) assuming lumped amplification. The markers represent published data from hero experiments operating in the C-, C + L-, and S + C + L-bands. The upper bound of each transmission band is given by an SNR_{TRX} of 25 dB, η of 1, and a span length L_S of 50 km. (b) Estimated range of optimal total launch power P as a function of transmission distance assuming lumped amplification. The upper bound of each transmission band is given by a span length L_S of 100 km.*

Fig. 3(b), the estimated optimum signal launch power becomes nearly independent of distance [7] for transmission distances exceeding a few span lengths. For short transmission distances below the respective span lengths of 50 and 100 km, the system consists of only a single amplified span. At intermediate transmission distances of less than ∼10 span lengths, the effect of a shorter final span, following an integer number of spans, manifests itself in a wavy curve, as explained in more detail in Section IV-A3 (in conjunction with Fig. 8). We note that, despite the idealization assumed in our model based on a homogeneous WDM system, the estimated ranges of power are in good agreement with the total launch powers used in the experiments. From Fig. 3(b), it is also evident that, even for a fully loaded $S + C + L$-band system, the total system launch power is still about a factor of 3–5 (i.e., ∼5–7 dB) lower than the power threshold of the fiber fuse. However, when used in conjunction with distributed Raman amplification instead of or in addition to lumped optical amplification, optical power levels within a fiber

Table 1 Calculation Model Parameters

	C	C+L	S+C+L
B [THz]	4.4	11.5	20.9
Bandwidth fill factor	1	0.97	0.95
SNR_{TRX} range [dB]	20 - 25	20 - 25	20 - 25
n_{sp} (Noise figure [dB] if G >>1)	1.3 (4)	1.6 (5)	2 (6)
η range	0.8 - 1	0.63 - 1	0.5 - 1
Amplifier span length range [km]	50 - 100	50 - 100	50 - 100

may get substantially higher. This may leave only little margin to further increase the transmission bandwidth due to the need for several Raman pumps per transmission band with typically >100-mW optical output power per pump [114]. In addition, once system bandwidths exceed the Raman gain bandwidth of ∼13 THz, Raman pumps and signal wavelengths start to spectrally overlap, which reduces fiber capacity due to the need to depopulate signals in lower wavelength regions of the system bandwidth, where Raman pumps are situated to amplify the longer wavelength regions of the system bandwidth [115].

IV. METHODS TO INCREASE CAPACITY

This section briefly summarizes methods to increase system capacity and discusses their effectiveness to handle exponential traffic growth. In this context, it is important to look for ways to scale capacity by large factors. A mere factor of 2 in increased system capacity corresponds to only two years of traffic growth at an annual traffic growth rate of 40% (1.5 years at a 60% growth rate). At a 40% traffic growth rate, a capacity increase by a factor of 30 is needed to ensure network scalability for another ten years.

A. Capacity Gain by Improving the SNR

1) Reducing Fiber Loss and Nonlinearity: Both fiber loss and fiber nonlinearity are responsible for noise accumulation during transmission: loss needs to be compensated by OAs that add noise, and nonlinearity both limits the maximum signal launch power (which makes amplification necessary in the first place) and also causes signal distortions that manifest, to a good approximation, as GN in modern coherent systems without in-line optical dispersion compensation, which is the basis for an equivalent NLIN term in the SNR [cf. (9)]. Since the announcement of the first 0.2-dB/km fiber in 1979 [21], much research effort has been devoted to further reduce the impact of fiber loss and nonlinearity, resulting in today's record-low loss and nonlinear coefficients of ∼0.14 dB/km and ∼0.6 $(W \cdot km)^{-1}$ obtained with pure silica-core fibers [24], [25].

On a parallel research path, hollow-core fibers (HCFs) have been investigated for over two decades. These fibers promise substantially reduced propagation loss and nonlinearity, far beyond the fundamental limits of solid-core fibers [116]. In recent years, impressive progress in fabricating a special variant of low-loss HCF known as

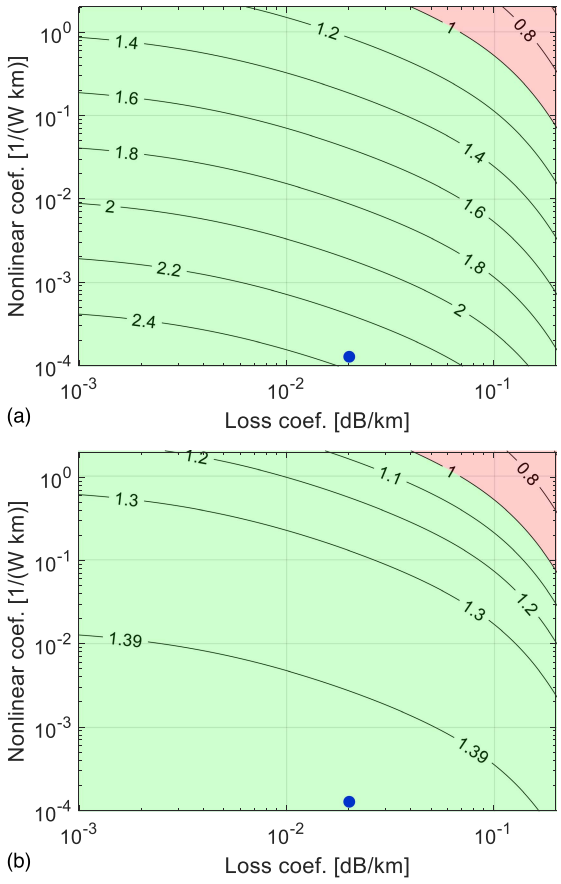

Fig. 4. *SE gain (contour lines) of a lumped-amplified 1000-km NANF system over a distributed-amplified SMF reference system (a) without and (b) with a transponder SNR limit of 30 dB. The green area depicts all NANF parameter combinations that result in an SE gain > 1 relative to the SMF reference system. The blue marker indicates the SE gain for the best NANF parameters predicted to date [117].*

the nested antiresonant nodeless fiber (NANF) [117] has heightened expectations that this type of HCF may be able to replace solid-core SMFs in the future. Simulations predict loss coefficients to be as low as ∼0.02 dB/km and nonlinear coefficients that are three to four orders of magnitude lower compared to those of their solid-core counterparts [117]. However, the capacity gains of such fibers on a systems' level are fairly small, as both a reduced fiber loss and nonlinearity coefficient "only" improve the SNR, which merely results in a logarithmic impact on capacity [cf. (6)]. To quantify the potential capacity benefits of a NANF system relative to an SMF system, Fig. 4 shows the SE gain of a lumped-amplified NANF system over an ideal distributed-amplified (reference) SMF system: 1) without and 2) with a transponder SNR limit of 30 dB. Lumped amplification for the NANF system has to be assumed because distributed Raman amplification relies on fiber nonlinearities and due to their very low nonlinearities being inherently difficult to achieve in HCFs. The SMF parameters are the same as those used for Fig. 2 (for fiber parameter details, see the Appendix). Both systems have a total link length of 1000 km and operate across the $C+L$-band. Note that the choice of using the same system bandwidth for both fiber types in this comparison was not dictated by limitations in the bandwidth of either fiber but by limitations in the bandwidth of practically available optical amplifiers.

The NANF system is equipped with ideal lumped amplifiers ($n_{\mathrm{sp}} = 1$) at uniform span lengths of 100 km. The NANF's dispersion parameter is assumed as 2 ps/nm/km [118]. Its loss coefficient α is varied between 0.001 and 0.2 dB/km, and its nonlinear coefficient γ is varied between 0.0001 and 1.3 $(\mathrm{W}\cdot\mathrm{km})^{-1}$. The green area depicts all $\{\alpha, \gamma\}$ combinations that result in a NANF SE gain > 1 relative to the SMF reference system. The blue marker near the bottom of the figures represents an "ideal NANF" with the best loss and nonlinear coefficients numerically predicted to-date ($\alpha = 0.02$ dB/km and $\gamma = 0.00013\ \mathrm{W}^{-1}\mathrm{km}^{-1}$) [117]. The total optimal launch power (across all WDM channels) was calculated to be ∼40 dBm for the ideal NANF and ∼18 dBm for the SMF reference system. Disregarding the practical difficulties of reliably generating such power levels with uniform spectral density across the $C+L$-band, SE gains of ∼2.4 could, in principle, be achieved for the ideal NANF and ideal transponders. An optimistic practical transponder noise floor of 30 dB reduces this value to ∼1.4. A noise floor of 20 dB, reflecting the current state of the art in high-speed commercial transponders, reduces the benefit of NANFs to only ∼1.1.

To better understand why a transponder noise floor has such a dramatic impact on the SE gain potential of an ideal NANF system, Fig. 5 depicts the SE as a function of transmission distance L for the cases of the ideal NANF system, the ideal distributed-amplified SMF reference system, and a more realistic lumped-amplified low-loss SMF

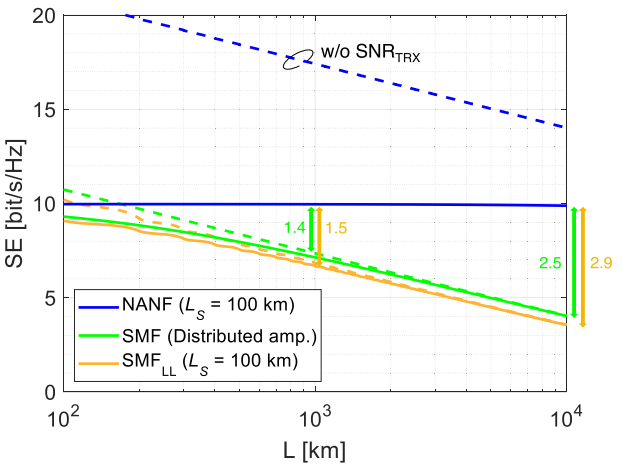

Fig. 5. *SE as a function of transmission distance L for two different fiber types and amplification scenarios without (dashed lines) and with (solid lines) a transponder SNR limit of 30 dB. Arrows with numbers indicate the SE gain for the relevant fiber types and distances. (SMF: standard SMF; $\mathrm{SMF}_{\mathrm{LL}}$: low-loss SMF; and L_S: amplifier span length).*

system (SMF_{LL}, see the Appendix for parameter details). Due to the transponder SNR limit, the SE of the NANF is clamped to ∼10 bit/s/Hz. In other words, it is the noise floor of practical transponders that prevent NANFs from reaching their full potential. The SE gain with respect to the SMF systems increases with transmission distance but does not exceed a factor of 3 (a factor of 2 in the case of $SNR_{TRX} = 20$ dB) even for trans-Pacific distances up to 10 000 km [84]. But, even without a transponder noise floor, Fig. 5 shows that the SE gain of NANFs is limited to between 2× and 4× (ratio of dashed lines) across relevant transmission distances, satisfying traffic demand increases for no more than four years, even in the most optimistic scenario.

Finally, we also note that these SE gains were obtained *without* considering intermodal interference (IMI)[8] within the NANF. We estimate that IMI would have to be less than −60 dB/km across the whole transmission band in order to not significantly impact transmission performance at trans-Pacific distances. In this respect, we also note that the effect on the SE due to crosstalk in form of IMI at a certain distance is identical to that of a transponder noise floor at the same value since both noise terms constitute SNR_0 [cf. (9b)]. For example, the individual noise contribution of IMI with −60 dB/km at $L = 1000$ km is identical to a transponder noise floor at −30 dB, and thus, IMI by itself reduces the SE to 10 bit/s/Hz (cf. blue solid line in Fig. 5) at $L = 1000$ km. Furthermore, the combined effect of IMI and an actual transponder noise floor would bring down the SE to 9 bit/s/Hz as the $1/SNR_0$ term dominates all other 1/SNR terms in this particular example with an ultralow-loss and ultralow-nonlinearity NANF at $L = 1000$ km. The relationship of these noise terms is further discussed in Section V-A1. The latest NANF designs have demonstrated an IMI suppression down to −56 dB/km on average, with a variation of ±2 dB/km across the C-band, uncabled, and under controlled laboratory conditions [119]. Although these accomplishments are very impressive from a technological perspective, from a systems perspective, we conclude that even the most optimistic IMI suppression, and loss and nonlinear characteristics of NANFs are not able to offer a sufficient SE gain benefit with respect to SMFs that would be needed to effectively deal with exponential traffic growth in the long term.

2) Reducing the Effect of Nonlinearities in the Digital Domain: The SNR of (9) can also be improved by reducing P_{NLIN} through digital NLC [104], [105], which can be performed on only one channel (intrachannel NLC) or across multiple WDM channels (interchannel NLC), as shown in Fig. 6. Interchannel NLC is limited by three key aspects: first, in order to perform interchannel NLC, one needs to get digitally sampled versions of multiple

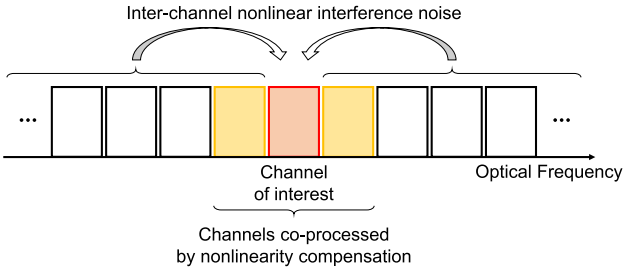

Fig. 6. ***NLIN is induced by all WDM channels that at least partially copropagate with the channel of interest, but only those WDM channels that share the entire transmission path and are digitized into a common ASIC can be used for effective digital NLC (after [105]).***

WDM channels into a common ASIC for joint processing. However, typical coherent DSP ASICs process only a single channel (with some exceptions having two processor cores) due to the large involved algorithmic complexities that need to be performed in real time at high speeds and make multichannel DSPs quickly untenable.[9] Even if such NLC coprocessing were done, multichannel NLC has been shown to quickly lead to diminishing returns [105]. Second, to effectively compensate for interchannel nonlinearities, WDM channels should propagate alongside each other from the same transmitter to the same receiver. This scenario applies to point-to-point systems (such as is the case for most submarine links) but does not apply to terrestrial optical mesh networks, where individual wavelengths are routinely routed to different fibers throughout the network at reconfigurable optical add/drop MUX (ROADM) nodes, and neighboring channels at any given receiver will generally only have shared a fraction of their nonlinear propagation path. Third, the nonlinear mixing of signals and ASE sets a fundamental limit to NLC [7]. Finally, just as with the reduction of fiber loss and nonlinearity coefficients, even a heroic 3-dB increase in SNR due to NLC only amounts to a 1-bit/s/Hz increase in the per-polarization SE, which is negligible in the bigger picture of an annual traffic growth rate of 40% or more.

3) Reducing the Noise Generated by Optical Amplification: The SNR of a transmission link also depends on the type and amount of optical amplification. Owing to the loss of SMF, transmission distances significantly exceeding 100 km make periodic in-line amplification mandatory to obtain sufficient SNR at the receiver. In-line amplification is carried out with either lumped or distributed amplification, as illustrated in Fig. 7.

The by-far most widely deployed *lumped* OA is the EDFA. Submarine cables typically use regular amplification spans between 50 and 100 km, while terrestrial systems place amplifiers approximately every 80–120 km, depending on

[8]By design, NANFs cannot be made purely single-moded. IMI represents a length- and wavelength-dependent crosstalk to the data-carrying fundamental mode from undesired copropagating and constantly cross-coupling higher order modes.

[9]At 100 GBaud, 1.5× oversampling, and 8 bits of ADC resolution, the aggregate bit rate of the digitized samples of only a single WDM channel (with two quadratures and two polarizations) amounts to 4.8 Tb/s that need to be real-time processed by the coherent DSP.

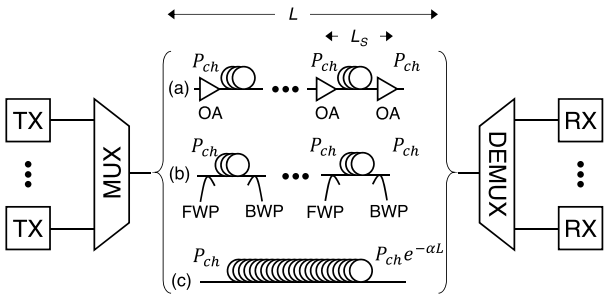

Fig. 7. *Schematic of a transmission link with (a) lumped amplification, (b) distributed amplification, and (c) no amplification. (TX: transmitter; RX: receiver; MUX: multiplexer; DEMUX: demultiplexer; OA: optical amplifier; FWP: forward pump; BWP: backward pump; P_{ch}: per-channel signal power; L: link distance; L_S: span length; and α: fiber loss coefficient).*

the available locations of repeater huts. Each amplifier nominally compensates for the loss of the preceding span, i.e., the amplifier has a gain of $G = e^{\alpha L_S}$, where L_S is the length of the preceding span. The ASE power accumulated per WDM channel over the total transmission distance can be expressed as $P_{\text{ASE}} = 2n_{\text{sp}}N_s hfR_S(e^{\alpha L_S} - 1)$, where n_{sp} is the amplifier's spontaneous emission factor (whose relation to the amplifier's noise figure is discussed in the Appendix), N_S is the number of spans, h *is* Planck's constant, f is the optical carrier frequency, and R_S is the symbol rate.

The most commonly used *distributed* amplification scheme is Raman amplification [114], where several pump lasers operating at wavelengths about 100 nm (∼13 THz) shorter than the signal wavelengths amplify the signals via stimulated Raman scattering by directly pumping the transmission fiber, usually with an optical power on the order of a few 100 mW. Raman pumping schemes comprise both forward pumping (FWP) and backward pumping (BWP). Higher order bidirectional pumping (where secondary and sometimes tertiary Raman pumps are used to pump the primary Raman pumps) generates the least variation in the gain experienced by the signal as it propagates down the transmission fiber, thus approximating the ideal case of maintaining constant signal power along the link. This, however, comes at the expense of MPI that must be taken into account in practical system designs [120]. The ASE noise power depends on the Raman gain coefficient, the fiber loss, and the pump power distribution along the fiber. Under the assumption of ideal distributed amplification, where the signal power and gain stay constant across the entire fiber, the ASE power is given by $P_{\text{ASE}} = 2hfR_S K_T \alpha L$, which represents the minimum ASE noise power generated by an amplified system [7]. K_T is proportional to the phonon occupancy factor and amounts to ∼1.14 for realistic Raman amplification [114]. Assuming ideal distributed amplification with $K_T = 1$ in the GN model results in an optimistic estimate for the capacity of a system that uses distributed amplification. As indicated by the black double arrows in Fig. 8, for regional to long-haul lightwave systems (i.e., $N_S \gg 1$), the SE degradation (relative to ideal distributed amplification) due to lumped amplification amounts to ∼0.6 bit/s/Hz per polarization for $L_S = 50$ km and ∼1.6 bit/s/Hz for $L_S = 100$ km. The amplifiers' spontaneous emission factor is assumed to be 1 (corresponding to a noise figure of 3 dB if the amplifier is operated with sufficient gain, i.e., $G \gg 1$). All other parameter assumptions are the same as those used in Fig. 2 (for fiber parameter details, see the Appendix).

Note that, for systems with short transmission distances below the respective span lengths of 50 and 100 km, the system consists of only a single amplified span with a span length equal to the transmission distance. This lets the curves representing lumped amplification in Fig. 8 converge toward the curve representing distributed amplification with decreasing transmission distance. In a similar vein, for cases where the transmission distance is not an integer multiple of the span length, the final span consists of an amplified span with a reduced length. The presence of a shorter final span manifests in small ripples that are only visible for a few spans. Also note that this graph already indicates the potential of a substantial capacity gain in power-limited systems when taking out every other amplifier (i.e., increasing the span length by a factor of 2) and using the freed-up amplifier power to set up a parallel spatial path to double the reduced capacity. For example, at 1000 km, a 50-km span system has an SE of 6.6 bit/s/Hz. Taking out every other amplifier reduces the SE to 5.5 bit/s/Hz, which adds up to an aggregate SE of 11 bit/s/Hz when using the freed-up amplifiers in a second spatial path. This strategy, which will be discussed in more detail in Section IV-C, works provided that: 1) doubling the total number of transponders and the total

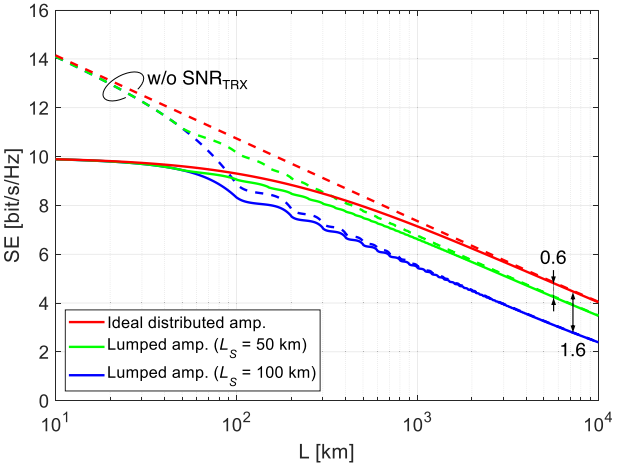

Fig. 8. *SE as a function of transmission distance L for distributed amplification (red), lumped amplification with 50-km spans (green), and lumped amplification with 100-km spans (blue). Solid lines assume SNR_{TRX} of 30 dB, while dashed lines assume no limitations due to SNR_{TRX}.*

power consumed by those transponders is tolerable and 2) the amplifiers do not consume significantly more power when providing twice as much gain for the same output power.

Raman amplification achieves gain bandwidths of ∼13 THz (100 nm) if pump wavelength and signal wavelength are kept in disjoint wavelength bands [114], [121]. Larger amplification bandwidths can only be achieved by sharing the system bandwidth between the signal and Raman pumps [115]. On the other hand, doped amplifiers typically achieve only ∼5 THz of gain bandwidth. Semiconductor OAs (SOAs) [122], [123] can achieve much wider lumped amplification bandwidths than doped amplifiers but generally attain worse noise performance than EDFAs.

Parametric amplifiers [124] have attracted considerable attention in the research community as they promise a reduced noise figure at the expense of more complicated to manage phase-sensitive amplification. A reach improvement of up to 5.6× has been demonstrated experimentally [125]. However, these benefits of parametric amplification come at the cost of: 1) using twice the bandwidth due to the need to copropagating an idler (i.e., a phase-conjugated copy of the signal) and the pump signal; and 2) increased system complexity due to the need for pump recovery and dispersion control at each amplifier stage. Sacrificing half the available system bandwidth for a potential 3-dB SNR advantage will in most scenarios reduce system capacity, as bandwidth is a prelog factor to capacity, while the SNR is inside the logarithm [cf. (6)]. It is, thus, unlikely that parametric amplification can effectively deal with exponential traffic growth in operational networks in the future.

B. Capacity Gain Through Multiple Transmission Bands

The costs associated with installing new regional and long-haul fiber are very high, on the order of $20 000 per kilometer, provided that ducts already exist. These costs are by far dominated by cable installation costs, as standard SMF strands cost less than $10 per kilometer in volume [48]. Therefore, and in contrast to short-reach applications, where cable deployment costs are much lower, regional and long-haul systems must exploit the vast installed fiber base to scale capacity as much as possible. This is done through multiband transmission systems, e.g., by deploying an L-band system in addition to an existing C-band system on the same fiber to double capacity. When considering multiband systems to boost the capacity of the installed fiber base, the loss distribution of already deployed fibers across the intended operating bands becomes a key consideration. Fig. 9 shows the loss distribution of various fiber types as a function of optical frequency. Note that, in contrast to typically published representations, the figure has the *optical frequency* as its lower x-axis and, only for reference purposes, shows the *optical wavelength* at its upper x-axis. Plotting wideband systems on a linear wavelength axis can be misleading because the inverse relationship between frequency and wavelength makes "bandwidths" on a wavelength axis appear larger at longer wavelengths than they are when expressed in terms of frequency. For example, 150 nm centered in the L-band and 100 nm centered in the O-band both correspond to a frequency bandwidth of ∼17.5 THz. Yet, it is the bandwidth measured in units of frequency (and not in units of wavelength) that matters in terms of system capacity [cf. (1)]. As can be seen from the blue double arrows in Fig. 9, the $C+L$-band has a bandwidth of 11.5 THz at a ∼0.2-dB/km loss level. When tolerating a loss coefficient of 0.23 dB/km, the SMF bandwidth extends to 23 THz, and 65 THz when tolerating a loss coefficient of 0.4 dB/km, which increases the system bandwidth by a factor of 2 and ∼6 relative to the $C+L$-band, respectively. The need to compensate for twice as much dB-loss has negative implications on the system design and energy budget, but the effect only impacts capacity logarithmically (through the SNR), as opposed to the linear capacity scaling when expanding system bandwidth. Shortening amplification spans through new repeater huts (if operationally viable) can help to mitigate higher losses at the edges of the system operating bands.

Apart from the SMF loss spectrum, Fig. 9 also shows experimentally measured loss curves of three recently fabricated NANFs (NANF-A/B/C) [119], [126] together with the results of an optimistic simulation of a NANF [117] discussed in Section IV-A1, which predicts an unrivaled 0.2-dB/km bandwidth of 110 THz, about 10× that of deployed SMF. Note, however, that the motivation to go to UWB systems is primarily to reuse *existing* fiber infrastructure. Therefore, the deployment of NANFs has to compete against the new installation of multiple parallel SMF strands or other value propositions. For example, the ultralow loss of these fibers, provided that cabling and splices do not add substantial loss once a NANF cable is actually deployed over long distances in the field, may enable long-haul transmission systems *without any optical amplification* [see also Fig. 7(c)] [127]: Assuming a uniform loss coefficient of 0.02 dB/km across the $C+L$-band (11.5 THz), an analysis similar to that carried out in Section IV-A1 reveals a maximum SE gain of ∼1.5 for an *unamplified* NANF relative to a distributed-amplified SMF system for a transmission distance of 2000 km, as depicted in Fig. 10. The results apply for an $\mathrm{SNR}_{\mathrm{TRX}}$ of 30 dB. An $\mathrm{SNR}_{\mathrm{TRX}}$ of 20 dB reduces the benefit of an unamplified ultralow-loss NANF system to ∼1.2, as shown in Fig. 10(b), showing the SE gain of an unamplified ideal NANF system relative to an SMF system with distributed amplification as a function of $\mathrm{SNR}_{\mathrm{TRX}}$ (or SNR_0) for different transmission distances. Note, however, that such low losses will be difficult to maintain in a deployed systems context as splices and cabling will unavoidably induce additional losses. As evident from Fig. 10(a), an additional loss of only ∼0.015 dB/km suffices to place the ideal NANF system on the "SE gain = 1" contour line, where the

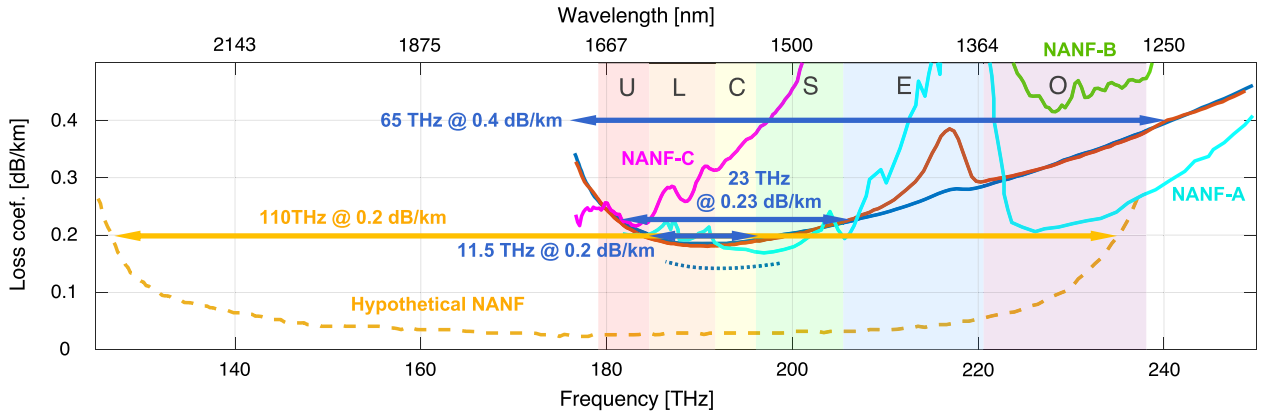

Fig. 9. *Fiber loss coefficient as a function of optical frequency (lower x-axis) and wavelength (upper x-axis) for widely deployed SMF with (red solid) and without (blue solid) water peak in the E-band; an ultralow-loss SMF (blue dotted) [25]; three experimentally fabricated NANFs (NANF-A/B/C, cyan [119], and green and purple [126]); and a hypothetical UWB NANF (orange dashed) predicted by optimistic simulations [117]. (O-band: 1260–1360 nm; E-band: 1360–1460 nm; S-band: 1460–1530 nm; C-band: 1530–1565 nm; L-band: 1565–1625 nm, and U-band: 1625–1675 nm.)*

NANF system shows no SE gain relative to the SMF system anymore. Also note in this context, that, in *unamplified* systems, *any* additional losses between transmitter and receiver, such as optical component insertion losses or nonideal photodetector responsivities, impact the SNR dB for dB, while amplified systems allow for the compensation of such lumped loss elements through dedicated lumped amplifiers with negligible impact on the SNR.

Consequently, the practical systems' benefit of NANFs in terms of scaling system capacity is mainly its increased *bandwidth expansion factor* (if it can be realized as predicted), and whatever bandwidth expansion factor may be achievable has to compete against an equivalent number of parallel SMFs. An additional, capacity-unrelated benefit of HCFs in general (and NANFs in particular) is the 50% increased propagation speed of light due to signals traveling mostly in air instead of glass. This can be a benefit in latency-sensitive systems, for applications such as high-frequency trading, memory disaggregation, distributed computing, and mobile fronthaul and backhaul. To assess whether or not HCFs might be beneficial in such situations, one has to assess whether the propagation delay is a limiting factor relative to an application's other critical time constants, including latencies from signal processing and communication protocols. Furthermore, the estimated (gain flattened!) optimum total optical launch power for an unamplified UWB system based on NANFs would range from many 10 to more than 100 Watts depending on the transmission distance, which comes with huge challenges in handling these high levels of optical launch powers at MUXs, splitters, connectors, or splices at the transmitter.

Aside from the fiber type and the problems with optical amplification and power handling, developing system components, such as lasers, high-speed photodetectors, modulators, and filters that operate equally well over *multiple optical frequency bands,* is a huge challenge in practice. To illustrate this challenge in experimental practice, Fig. 11 shows the actually achieved SE in recently reported transmission experiments as a function of system bandwidth. To account for the impact of different transmission distances across experiments, the measured SEs are *normalized* to the estimated SE at the respective transmission distance as per (6), and SEs are reported as a percentage of those estimated for each point. Since the SE estimate is made for an $\mathrm{SNR}_{\mathrm{TRX}}$ of 25 dB [cf. the orange solid line in Fig. 2(c)], the measured SE above 100% [96] represents a short-reach small-bandwidth experiment that was conducted with a higher $\mathrm{SNR}_{\mathrm{TRX}}$ and, hence, is above the capacity limit estimate for $\mathrm{SNR}_{\mathrm{TRX}} = 25$ dB. Although these measurements were taken under very different experimental conditions, such as different amplification schemes and different transmission distances, it can be clearly seen that wider system bandwidths make it much more difficult to stay close to the theoretical estimate. In other words, the capacity gain from operating a system across a wider bandwidth is partially lost in practice by a reduced SE due to difficulties in building wideband system components. This is a general truth throughout engineering applications: systems with a small *relative* bandwidth B_{rel} (i.e., a small system bandwidth divided by the system's average carrier frequency f_c, $B_{\mathrm{rel}} = B/f_c$) are much easier to build than systems with a large relative bandwidth [3], [9]. The C-band has a relative bandwidth of 2.3%, a recent SOA-based system [122] has a relative bandwidth of 6.6%, and the hypothetical NANF of Fig. 9 has a nearly octave-spanning relative bandwidth of 61%.

The need to develop different system component technologies for each band implies that "multiband" parallelism is not the same type of parallelism that has so successfully been used for WDM or spectral superchannels [131], [132], where multiple carrier frequencies can be simultaneously passed through the same devices within

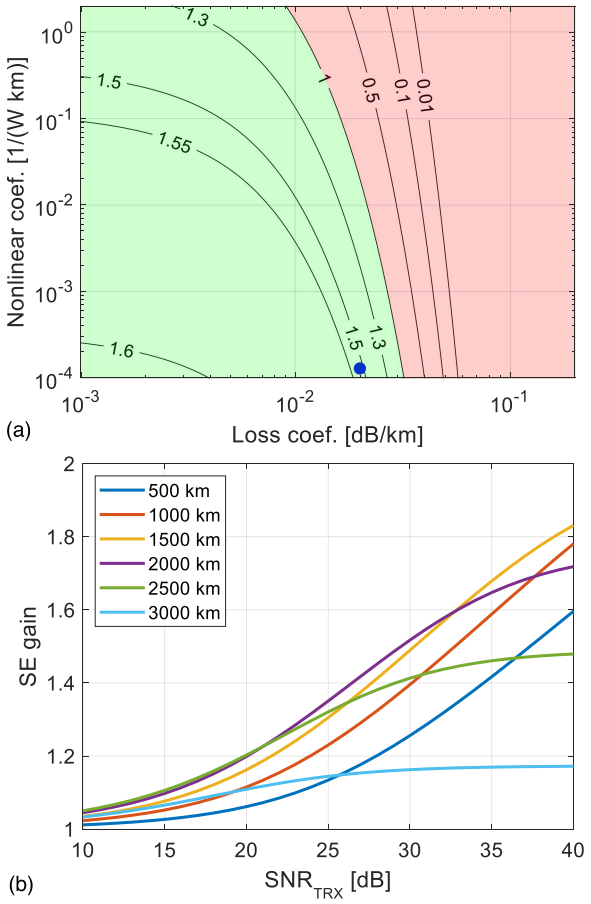

Fig. 10. *(a) SE gain (contour lines) for a 2000-km unamplified NANF system over an ideally amplified SMF reference system with a 30-dB noise floor. The green area depicts all parameter combinations that result in an SE gain > 1 relative to the SMF reference system. The blue marker indicates the SE gain for an ideal NANF [117]. (b) SE gain for an unamplified ideal NANF system over an ideally amplified SMF reference system as a function of transponder SNR at different transmission distances.*

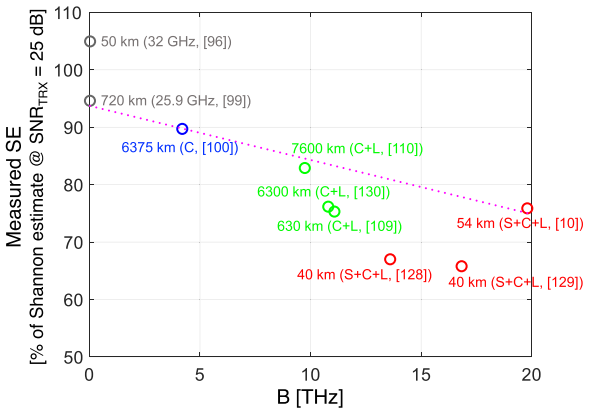

Fig. 11. *Experimentally measured SE of recent transmission record experiments as a function of system bandwidth B. SEs are shown normalized, as a percentage of the estimated SE as per (6). assuming $SNR_{TRX} = 25$ dB. The numbers indicate transmission distances. The marker above 100% represents a short-reach small-bandwidth experiment that was operated with a slightly higher SNR_{TRX} and, hence, falls above the Shannon limit estimate for $SNR_{TRX} = 25$ dB. The colors represent different classes of system bandwidths (gray: small system bandwidth; blue: C-band; green: C + L-band; and red: S + C + L-band). The magenta dotted line depicts a trend line, as a guide to the eye only.*

the same wavelength band. The latter entails an immense benefit in terms of energy and cost reduction by means of integration. It is, thus, important to keep in mind that *parallelism* is most effective if the system components that are used in parallel are actually based on a *common technology platform*, i.e., use a reasonably small relative bandwidth. Using different technologies to accommodate multiple bands (e.g., through the use of widely different dopants in doped OAs) does not leverage parallelism most efficiently. As noted earlier, UWB transmission also interferes with distributed amplification as Raman pumps share spectrum with signals and raise concerns with regard to damage caused by a potential fiber fuse and eye safety for the system's maintenance personnel.

C. Capacity Gain Through Multiple Spatial Channels

The concept of exploiting spatial parallelism to increase data throughput is highly effective as it simply replicates the *same technology within the same optical operating band* M times in parallel. Spatial parallelism is successfully deployed in many other areas, such as in microchips, where the same transistors are used billions of times, in computers using parallel data buses, in cellular wireless systems reusing frequencies in spatially disjoint cells from macrocells to femtocells, and in multifiber cables for short-reach optical interfaces. How many parallel spatial channels should be used as opposed to how much information should be packed onto a single high-capacity spatial path is a tradeoff between the cost of the *transmission medium* (including its cabling, deployment, and connectorization) and the cost of the *transponders* (including their PDM and WDM multiplexing mechanisms). For short transmission distances, where the transponder cost is much higher than the cost of the fiber link, it is generally cheaper to refrain from excessive multiplexing, while multiplexing is essential in long-haul transmission to amortize the cable cost across as many transponders as possible. Table 2 quantifies this tradeoff, showing the approximate cost of a spatial fiber path in terms of 100-Gb/s transponder

Table 2 Cost per Fiber Path in Units of Transponder Cost/100G

	Distance	Cost
Trans-Pacific	11,000 km	1,200
Trans-Atlantic	6,000 km	700
Terrestrial Long-Haul	1,000 km	65
DCI	100 km	1
Intra-DC	1 km	0.3
HPC	100 m	0.1

units [47], [48]. At trans-Pacific distances, the cost per spatial path (including fiber, cable, and amplifiers) amounts to about 1200 times the unit transponder cost per 100 Gb/s. Hence, it is mandatory to multiplex a large number of WDM channels onto each fiber to balance transponder costs with cable costs. On the other hand, in high-performance compute (HPC) systems with transmission reaches of 100 m and less, the cost per fiber path can be ∼10% (or less) of the cost of a 100-Gb/s transponder, which currently makes multiplexing of any form an unnecessary endeavor. It is much cheaper in such short-reach systems to deploy massively parallel fiber bundles than to multiplex many wavelengths onto a single fiber.

As today's long-haul optical communication systems are being operated close to their nonlinear Shannon limit estimates, with essentially all practical capacity-increasing methods exhausted, it is natural to migrate to spatially parallel systems. However, the important question is how this migration should take place to meet the growing demand for capacity most economically. Simply deploying M parallel systems leads to an M-fold capacity increase at M times the cost. This lets the cost per bit remain constant, which is not what society expects of digital communication services. SDM systems must, therefore, be built to share system components among spatial channels, as shown in Fig. 12, thereby reducing the cost and energy overhead that are dominant in each individual component [133]. Sharing becomes most effective by *array integration* across all system components, including transponders, fibers, splices and connectors, OAs, MUXs, filters, and switches. SDM research over the past decade has been mostly concentrated on novel fiber designs and demonstrating transmission records on those new media. Pursuing massive array integration of other system components is at least as important, though.

In addition to massive array integration, SDM offers other system-level benefits. For example, *power-limited* systems can significantly benefit from SDM even without any array integration, as has recently been shown [46], [47], [134] and has been implemented in the latest generation of submarine cables [14]. To illustrate this point, Fig. 13(a) shows the achievable *aggregate* SE (dubbed $\mathrm{SE_{SDM}}$) of a power-limited system

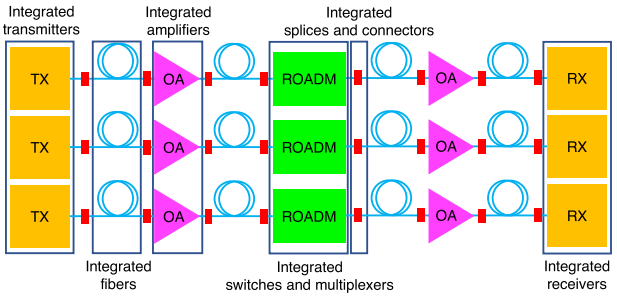

Fig. 12. ***Component sharing in an SDM system by means of integration.***

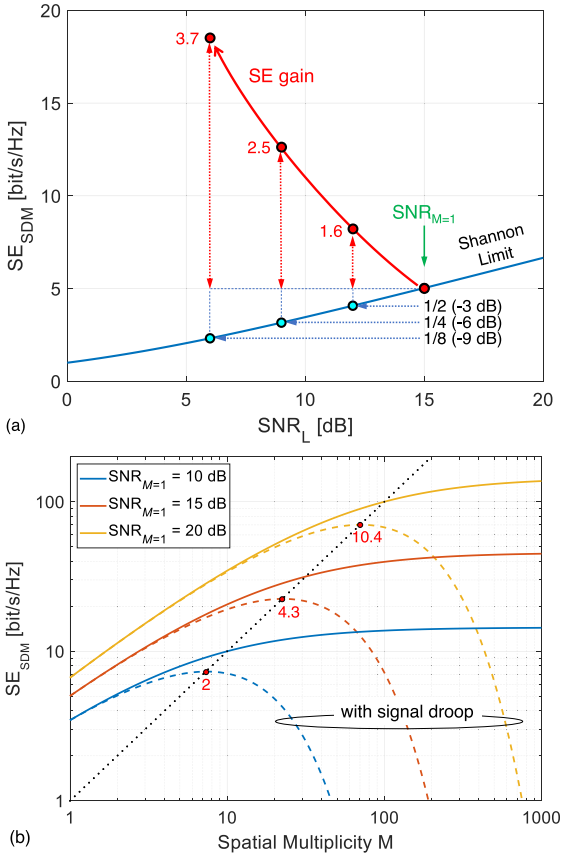

Fig. 13. ***(a) Aggregate SE ($\mathrm{SE_{SDM}}$) obtained by distributing the signal power across multiple spatial paths. The available power is assumed to provide an $\mathrm{SNR_L}$ of $\mathrm{SNR}_{M=1} = 15$ dB for a single spatial path. (b) $\mathrm{SE_{SDM}}$ as a function of spatial multiplicity for different single-path $\mathrm{SNR_L}$ values at $M = 1$ ($\mathrm{SNR}_{M=1}$) for a power-limited system. The solid lines show the case when the SNR is calculated assuming no signal droop (constant-gain OAs), and the dashed lines show the case when the SNR is calculated according to the generalized droop model (constant-output-power OAs) with $N_S = 100$ spans. The numbers in red give the aggregate SE gain relative to the SE at $M = 1$. The black dotted line depicts the optimum $\mathrm{SE_{SDM}}$ for all M in the presence of signal droop.***

as a function of $\mathrm{SNR_L}$. $\mathrm{SE_{SDM}}$ is the total transmitted information per polarization and per unit bandwidth across all parallel spatial paths. The blue curve shows the maximum achievable SE of a single spatial path. For the following analysis, we assume that our reference system ($\mathrm{SNR}_{M=1}$) operates at a sufficiently high SNR but not above the nonlinear optimum[10] [cf. Fig. 2(b)]. As shown in Fig. 13(a), when considering a system that consists of a single spatial path to which a total $\mathrm{SNR_L}$ of 15 dB is available due to limited electrical system power supplying

[10]To be more precise, we assume that the reference system operates in a regime where the $1/\mathrm{SNR}_1$ term dominates all other 1/SNR terms, i.e., the ASE power dominates all other noise terms. If $1/\mathrm{SNR}_{-2}$ (i.e., $1/\mathrm{SNR_{NLIN}}$) cannot be neglected, i.e., if the reference system operates very close to or even slightly beyond the nonlinear optimum, $\mathrm{SE_{SDM}}$ will be slightly lower than described here. On the other hand, if the $1/\mathrm{SNR}_0$ term is dominant, i.e., if the reference system is limited by $1/\mathrm{SNR_{TRX}}$ or $1/\mathrm{SNR_{XT}}$, $\mathrm{SE_{SDM}}$ will be higher than described here. This could be the case for a NANF system, as discussed in conjunction with Fig. 5.

the in-line OAs, we get an SE of $\log_2(1+32) \approx 5$ bit/s/Hz per polarization. Splitting the available power among two parallel spatial paths yields $\mathrm{SE_{SDM}} = 2\log_2(1+16) \approx 8$ bit/s/Hz (equivalent to an *aggregate* SE gain of 1.6×, although the SE *per spatial path* drops by 19%). Using four parallel spatial paths yields $\mathrm{SE_{SDM}} = 4\log_2(1+8) \approx 13$ bit/s/Hz (equivalent to an aggregate SE gain of 2.5× although the SE per spatial path drops by 37%). Increasing spatial parallelism at the cost of a smaller SNR (and, hence, a smaller SE) per spatial path brings about an overall capacity gain in a power-limited system since the prelog multiplier outside the logarithm is much more valuable than the associated SNR dilution. This concept has become particularly attractive for power-limited submarine cables, where system designers and operators realized that, by diluting the fixed amount of total power over a larger number of fiber pairs, the overall system cost per bit can be significantly reduced [46], [47], [134], [135]. This also prompts the question of what the limits in capacity gain due to parallelism and power dilution may be.

Under the assumption of a uniform lossless distribution of the signal power across all spatial paths, ideal Nyquist pulses, and no implementation penalty (i.e., $\eta = 1$), $\mathrm{SE_{SDM}}$ as a function of spatial multiplicity M is shown in Fig. 13(b), with the Shannon-limited SNR as a parameter that reflects the total power available to the system [$\mathrm{SNR}_{M=1}$; cf. Fig 13(a)]. Since $\mathrm{SNR}_{M=1}$ is fixed, both the span length and total launch power are also assumed not to change with increasing M as these parameters affect $\mathrm{SNR}_{M=1}$. Two cases are shown:

Case 1 (solid lines) assumes *constant-gain amplifiers* resulting in the same *signal* power at the amplifier output for all spans. This case is dubbed the "no droop" case with $\mathrm{SE_{SDM}}$ given by

$$\mathrm{SE_{SDM}} = M\log_2\left(1+\frac{\mathrm{SNR}_{M=1}}{M}\right). \tag{10}$$

Case 2 (dashed lines) assumes *constant-output-power amplifiers*. While constant-gain and constant-power cases converge for high SNRs, they significantly differ for low SNRs as the amplifier output power consists of both signal and ASE power. This effect has become known as the "signal droop" effect [135], [136], [137] since the signal power decreases relative to the amplifier output power further down a long transmission line, and the accumulated ASE takes on an increasingly larger portion of the amplifier's total output power [138]. From the generalized droop model [137], $\mathrm{SE_{SDM}}$ follows as

$$\mathrm{SE_{SDM}} = M\log_2\left(1+\frac{1}{\left(1+\frac{M}{\mathrm{SNR}_{N_S=1,M=1}}\right)^{N_S}-1}\right) \tag{11}$$

with N_S being the number of amplifier spans and $\mathrm{SNR}_{N_S=1,M=1}$ representing the SNR after the first span.

In case 1, $\mathrm{SE_{SDM}}$ saturates at the value $\mathrm{SNR}_{M=1}\log_2(\mathrm{e})$ for $M \to \infty$, and for typical $\mathrm{SNR}_{M=1}$ values ranging from 10 to 20 dB, the $\mathrm{SE_{SDM}}$ obtained by ideally distributing the total power over many parallel paths ranges from about 14 to 144 bit/s/Hz, which corresponds to an SE gain of about 4× to 21× relative to the SE at $M = 1$.

In contrast, for case 2, which represents the way today's power-limited submarine systems are operated, a maximum $\mathrm{SE_{SDM}}$ is reached at a much lower spatial multiplicity. Note that the achievable $\mathrm{SE_{SDM}}$ in the signal droop case also depends on N_S. For the hypothetical case of $N_S = 1$, there is no signal droop, and the SEs are identical to case 1 (solid lines). The dashed curves represent the case $N_S = 100$, which is virtually identical to all cases where $N_S > 100$. In this case, the maximum achievable $\mathrm{SE_{SDM}}$ is given by[11] $\mathrm{SE_{SDM}} = M \approx \mathrm{SNR}_{M=1}/\log_2(\mathrm{e})$, which corresponds to an optimum SE per spatial path of 1 bit/s/Hz per polarization and an optimum SNR per spatial path [represented by the compound fraction in (11)] of 0 dB, consistent with [139]. Consequently, the maximum achievable aggregate SE gain (relative to the SE at $M = 1$) is 2×, 4.3×, and 10.4× for an $\mathrm{SNR}_{M=1}$ of 10, 15, and 20 dB, respectively.

It is important to note that, in the above analysis, the spatial multiplicity M does not necessarily equal the actual number of fibers or fiber pairs contained in a submarine cable but rather represents the number of parallel spatial paths that share a limited power supply that would result in $\mathrm{SNR}_{M=1}$ if it were used on a single spatial path. For example, a submarine cable that was designed to use eight fiber pairs operating near the nonlinear optimum at $\mathrm{SNR}_{M=1} = 12$ dB per fiber [e.g., as represented by the cyan curve in Fig. 2(b)], reflecting a certain amount of available electrical supply power, could be upgraded to 88 fiber pairs (i.e., $M = 11$) to boost the $\mathrm{SE_{SDM}}$ of the cable by a factor of 2.7 with the *same* amount of electrical supply power. In practical systems, the optimum SE and SNR per spatial path depend on the total link length and various other link parameters. Hence, optimization with regard to span length, power efficiency, and cost [47], [140], as well as operational overheads and some unavoidable loss due to power distribution among the spatial paths, may lead to a slightly higher optimum SE (and, consequently, also a higher optimum SNR) per spatial path, which, however, reduces $\mathrm{SE_{SDM}}$. In this respect, we also remark that the main purpose of this analysis is to show that in a power-limited system signal droop leads to a finite M when attempting to maximize capacity solely by means of power dilution. Increasing capacity beyond that point will inevitably require adding more electrical power to the system.

An interesting side benefit of diluting power over multiple parallel systems is that each fiber is pushed to operate in the *linear* regime, which makes fiber nonlinearities

[11] This value is approached for a sufficiently large number of spans ($N_S \gg 1$) and $M \gg 1$. For $M = 1$, the optimum is found at $\mathrm{SE_{SDM}} = \mathrm{SNR}_{M=1} = 1$.

a nonissue for cost-optimized high-capacity submarine cables in the future. This is particularly interesting as these systems have traditionally been the most nonlinearity-plagued fiber-optic communication systems. Further discussions on applications and implications of SDM in submarine systems can be found, e.g., in [140], [141], and [142].

V. PARALLEL SYSTEM ARCHITECTURES AND TRADEOFFS

The integration of system components is key to the overall value proposition of spatially parallel systems, as, without integration benefits, M parallel systems have M times the cost of a single system, which does not reduce the cost per bit. This section briefly introduces the main concepts of SDM integration that have been proposed and discusses various tradeoffs. Further details are provided in other papers contained within this Special Issue, including [60], [67], [143], [144], [145], and [146].

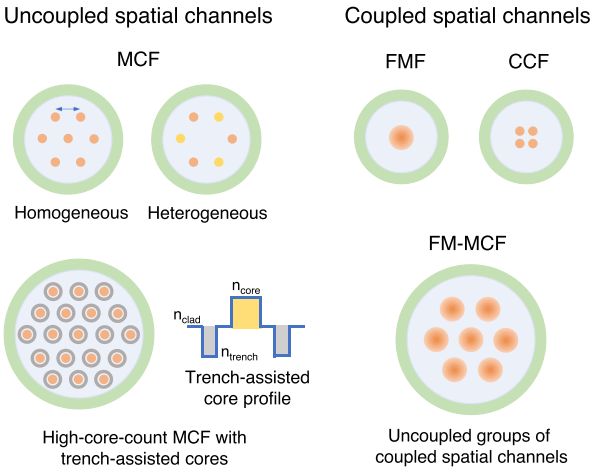

Fig. 14. *SDM fiber designs (cores: yellow/orange; cladding: light blue; buffer coating: light green; n: refractive index; and Λ: core pitch).*

A. Transmission Medium

The most straightforward spatially parallel system architecture uses a high-density fiber bundle, i.e., a cable with multiple individual fiber strands that connect two network end points in parallel. To reduce the overall cable diameter and weight, with implications discussed in Section V-A5, a reduced cladding and coating diameter may be used for the fiber strands [143]. Cables with an outer cable diameter of 23-mm housing 3000 parallel optical fiber strands [147] and a diameter of 29-mm housing 6912 parallel optical fiber strands [148] with a coating diameter of 200 μm are already deployed.

In order to further reduce cable diameter and weight, dedicated SDM fibers have been intensively explored in research labs over the past decade. These can be broadly classified into fibers with *uncoupled spatial paths* and fibers with *coupled spatial paths*, as illustrated in Fig. 14. Whether a system is considered to have uncoupled spatial paths or coupled spatial paths depends on whether or not MIMO DSP across multiple parallel spatial paths is needed at the receiver to undo the crosstalk (XT) between spatial paths in the process of recovering the signals.[12] If the XT between parallel spatial paths is *below* this system-specific "MIMO threshold," the exact amount of crosstalk is irrelevant as XT is absorbed as a noise term (P_{XT}) into the achievable SNR, and systems operate within the allocated crosstalk penalty budget. If the crosstalk is *above* the MIMO threshold and MIMO DSP is used to separate the signals traveling on parallel spatial paths, the exact amount of XT is again irrelevant, as MIMO DSP can cope with *arbitrary* amounts of XT, i.e., a signal launched into any of the coupled-path SDM fiber's M parallel spatial paths at the transmitter may be equally distributed among all paths during transmission. The definition of a "MIMO threshold" results in a clean (but transponder- and system-specific) separation of fibers into uncoupled and coupled variants.

1) Uncoupled Spatial Parallelism (Core Multiplexing): The concept of having parallel spatial channels (i.e., cores) physically separated while embedding them in the same cladding structure has led to the design of MCFs. Individual cores can be nominally identical (homogenous MCFs) or deliberately different (heterogeneous MCFs) in terms of their size and refractive index [149]. The trench-assisted core profile [150], [151] uses a refractive index trench around each core to suppress intercore XT and has been popular owing to its effectiveness and low fabrication complexity, which is similar to that of bend-insensitive SMFs [152]. The heterogeneous core layout [153] can additionally suppress intercore XT but comes at an increased fabrication complexity.

In MCFs (sometimes also referred to as weakly coupled MCFs), intercore XT arises due to random perturbations in the refractive index caused by structural fluctuations, manufacturing imperfections, bends, or twists [154], [155], in addition to a small but nonnegligible overlap of the field from the interfering core, i.e., the core where the XT signal originates from, with the surrounding (interfered-upon) cores, i.e., the cores where the XT signal is coupled to. Recent studies have shown that intercore XT is influenced by the strength of intra-core random polarization mode coupling [156], [157]. We further note that although XT is a strictly stochastic quantity [158], [159] and its variance (i.e., the amount of random deviations from a mean value over time) can depend on the modulation format as well as the symbol rate of the transmitted signal [160], for modulation schemes and symbol rates typically used in conjunction with coherent detection, XT can be considered virtually stationary and thus is sufficiently described by its mean value. Intercore XT is a function of the propagation

[12]Coherent systems on standard SMF always use a 2 × 2 MIMO DSP to separate the two polarization components at the receiver, which are always fully coupled in realistic transmission scenarios.

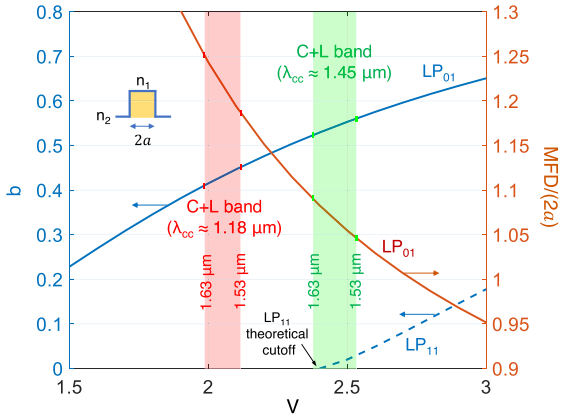

Fig. 15. *Normalized propagation constant b and normalized mode field diameter MFD/(2a) as a function of normalized frequency V. The solid lines represent the fundamental (LP_{01}) mode, and the dashed line represents the next-higher order (LP_{11}) mode whose propagation is generally undesired in an SMF core. The colored areas depict the frequency range of the $C+L$-band for two different cable cutoff wavelengths. The inset shows the step-index refractive index profile of a standard SMF considered here.*

distance L and is defined as $\mathrm{XT}(L) = P_2(L)/P_1(L)$, where P_1 is the power in the core under consideration measured at distance L when no interferers are present, and P_2 is the power that has accumulated due to XT from all interferers at distance L in the core under consideration when no power is coupled into the core under consideration at distance 0. The unit "dB/km," commonly used to describe intercore XT, refers to the amount of XT accumulating in 1 km of fiber. For example, in the case of an MCF with a per-km XT of -60 dB/km, the accumulated XT after 1000 km of propagation amounts to -30 dB.

As the amount of intercore XT depends on how far the optical field reaches into the cladding, XT can be suppressed by operating the fiber closer to the theoretical cutoff wavelength of the next-higher order mode. Similar to SMFs, a typical MCF is designed with a cable cutoff wavelength of $\lambda_{\mathrm{cc}} < 1260$ nm to guarantee single-mode operation including the O-band (cf. Fig. 9) [161]. Operating a single-mode core designed for $\lambda_{\mathrm{cc}} < 1260$ nm at 1550 nm results in a mode with a broader cladding field (and consequently higher XT) than what would be obtained if the core were designed for, e.g., $\lambda_{\mathrm{cc}} < 1460$ nm. To illustrate this effect, Fig. 15 shows the normalized propagation constant b and normalized mode field diameter (MFD) as a function of normalized frequency V for a standard SMF with a step-index profile. These fiber parameters are defined as [69], [162]

$$b = \frac{n_{\mathrm{eff}}^2 - n_2^2}{n_1^2 - n_2^2}, \quad V = \frac{2\pi a}{\lambda} n_1 \sqrt{2\Delta}, \quad \Delta = \frac{n_1^2 - n_2^2}{2n_1^2}, \text{ and}$$

$$\mathrm{MFD} = 2\sqrt{2\frac{\int_0^\infty \psi^2 r dr}{\int_0^\infty \left(\frac{d\psi}{dr}\right)^2 r dr}}$$

where $2a$ represents the core diameter, n_{eff} is the mode's effective refractive index (or modal index), which is the ratio of the mode's propagation constant (or wavenumber) to that of a plane wave in free space, Δ is the relative index difference between core (n_1) and cladding (n_2), and ψ stands for the mode's field distribution. The red and green areas in Fig. 15 depict the frequency range of the $C+L$-band for the cases of $\lambda_{\mathrm{cc}} \approx 1180$ nm (obtained here with $2a = 8.5$ μm and $\Delta = 0.35\%$) and $\lambda_{\mathrm{cc}} \approx 1450$ nm (obtained here with $2a = 8.5$ μm and $\Delta = 0.5\%$), respectively. For an MCF designed with a cable cutoff of $\lambda_{\mathrm{cc}} \approx 1180$ nm, the fundamental (LP_{01}) mode at 1550 nm has an MFD of $\sim 1.2\times$ the core diameter, which induces XT onto nearby cores much more strongly compared to an MCF with a cable cutoff of $\lambda_{\mathrm{cc}} \approx 1450$ nm, where the MFD is only $\sim 1.05\times$ the core diameter. While the latter case is preferred for intercore XT suppression, it also increases the probability of exciting the LP_{11} mode, which may lead to MPI problems for such designs.

The concept of the cable cutoff wavelength is used to obtain a realistic estimate of the attenuation of the first higher order mode (LP_{11} mode in the case of SMFs) in cabled fibers subject to perturbations such as bends and twists [163]. Note that, for a standard SMF, the cable cutoff wavelength is about 150 nm shorter than the theoretical mode cutoff wavelength. The latter is defined as the wavelength where $b = 0$ for the LP_{11} mode, i.e., where the modal index becomes equal to the refractive index of the cladding. Above this wavelength, the relevant waveguide mode turns into a cladding mode and, thus, loses its light-guiding property. How much the cable cutoff wavelength differs from the theoretical cutoff wavelength depends on the core profile. For example, if the core is surrounded by a deep index trench, it is common to find the theoretical cutoff wavelength much closer to (or even slightly below) the cable cutoff wavelength as the propagation loss in such more bend-resistant fibers is predominately governed by leakage loss in a straight fiber rather than by bend loss in a bent fiber [164]. Nevertheless, the cable cutoff wavelength remains the relevant parameter to quantify attenuation (i.e., suppression) of the first higher order mode.

Fig. 16 shows the MFD and the intercore XT as a function of wavelength across the $C+L$-band for the two cable cutoff wavelengths discussed above. We first note that intercore XT increases by $\sim$10 dB across the $C+L$-band. In general, intercore XT increases with wavelength at a rate of $\sim$0.1–0.15 dB per nm (depending on the core profile and layout) [165]. Signals with the longest wavelength in a WDM system, therefore, experience the largest XT. Hence, it is this XT value that needs to be considered when evaluating the maximum impact of XT on a system's transmission quality (see Figs. 17 and 18). As expected from Fig. 15, a significant reduction in XT (in this example, from ~ -30 to ~ -65 dB/km) can be obtained if the cable cutoff wavelength is moved closer to the lower edge of the C-band [166], provided that the operation at shorter wavelength bands than the C-band is not required. XT is

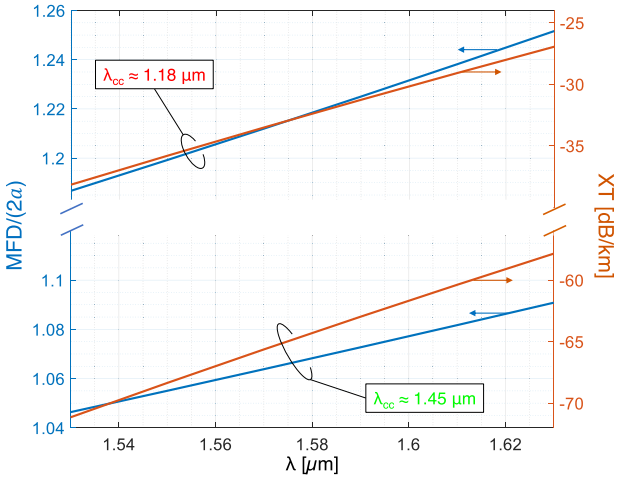

Fig. 16. *Normalized mode field diameter MFD/(2a) and intercore XT of the LP_{01} mode as a function of wavelength for the cases of $\lambda_{cc} \approx 1180$ nm (top) and $\lambda_{cc} \approx 1450$ nm (bottom) corresponding to the red and green areas of Fig. 15.*

also inversely proportional to the core pitch Λ (see Fig. 14). In terms of XT suppression, it is, therefore, desirable to place the cores as far apart as possible, which stands against the desire to pack as many cores as possible into a given MCF cladding diameter. Furthermore, cores placed too close to the cladding boundary exhibit excess loss due to the optical field leaking into the lossy coating. An outer cladding thickness (i.e., the distance from the outermost core's center to the cladding-coating boundary) of $\sim$35 μm is typically needed for single-mode MCFs with a trench-assisted core profile to keep the excess loss below 0.01 dB/km at $\lambda = 1.625$ μm [166], [167]. Like intercore XT, this value depends on the core profile and bend condition of the fiber. For MCFs that deliberately support multiple propagation modes per core in their few-moded cores, the outer cladding thickness and the core pitch have to be increased with an increasing number of modes propagating in the cores of an FM-MCF, as higher order mode fields tend to spread farther into the cladding [149], [155]. A minimum required core pitch and outer cladding thickness limit the number of cores that can be accommodated within a given cladding diameter [60]. Increasing the cladding diameter is problematic because the cladding size critically affects the fiber's mechanical reliability [168]. Typical SMFs have a cladding diameter of 125 μm, and a cladding diameter of 230–250 μm is believed to be the practical upper limit for silica fibers to be reliably used in optical networks. While MCFs with a cladding diameter of 242 μm [169] enclosing as many as 32 uncoupled cores have been demonstrated in conjunction with high-capacity transmission record experiments [170], keeping the typical SMF cladding diameter of 125 μm is currently preferable in terms of mass production, as well as cabling and connectorization. This currently limits the maximum number of cores to 4 [60] for operation in the $C + L$-band if XT and its impact on SE are to be kept sufficiently low across a wide range of transmission distances (see Figs. 17 and 18). One way to increase the core density at a fixed MCF cladding diameter is to only allow for opposite propagation directions in adjacent cores [171]. Since XT only occurs between copropagating signals, directionally interleaved propagation results in an effective increase of the XT-generating core-to-core spacings at the highest possible core density. The resulting MCFs in a deployed scenario then have a "polarity" though, with clearly defined "input ports" and "output ports."

Mathematically, the noise contribution due to linear crosstalk is captured in (9) through $P_{\mathrm{XT}} = \xi P_{\mathrm{ch}} = (\xi_1 + \Upsilon(\xi_2 L))P_{\mathrm{ch}}$ with ξ_1 describing length-independent XT, such as the XT from neighboring WDM channels, and ξ_2 representing length-dependent XT, such as intercore XT in MCFs. In a well-engineered transmission system, ξ_1 is negligibly small. The function Υ depends on the core layout and ensures that the noise contribution due to XT cannot become stronger than the signal from which the XT originates (i.e., saturates for large L). It is beyond the scope of this article to describe the function Υ and its effect on SE for all kinds of core layouts and crosstalk strengths. Suffice it to say that, if XT is sufficiently small, i.e., $\xi_2 L \ll 1$, the approximation $\Upsilon(\xi_2 L) \approx \xi_2 L$ holds, and XT scales linearly with fiber length L [172], [173], i.e., increases by 3 dB for every doubling in the transmission distance. This motivates the specification of XT in terms of "dB/km" and also represents the regime where it is generally not necessary to resort to MIMO DSP to compensate for intercore XT.

Fig. 17 shows the impact of intercore XT on the SE as a function of transmission distance for an aggregate XT

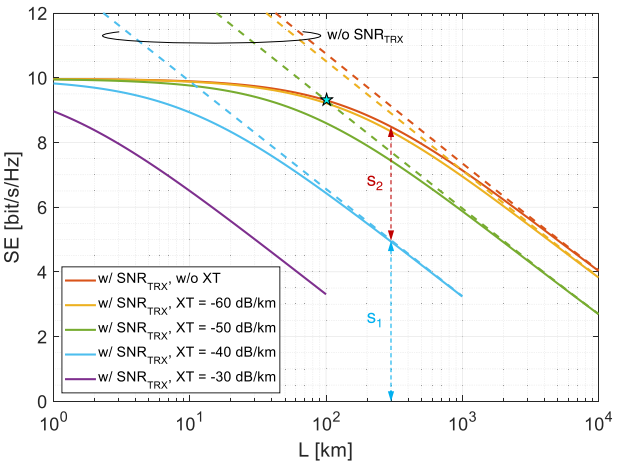

Fig. 17. *SE as a function of transmission distance L with and without (aggregate) intercore XT, and with (solid lines) and without (dashed lines) an SNR_{TRX} of 30 dB. All lines are plotted up to an accumulated maximum XT of -10 dB. The marker indicates that the effect of intercore XT of -50 dB/km at $L = 100$ km on the SE is the same as that of a transponder floor at -30 dB.*

ranging from −60 to −30 dB/km. The calculations assume distributed amplification, no implementation penalty, and a homogenous WDM system across the $C + L$-band operating over a standard SMF. Parameter assumptions are the same as those for Fig. 2. The *aggregate* XT is the sum of the XT arising from all parallel spatial channels contributing to XT within the core of interest. The curves are plotted up to the point where the noise contribution due to aggregate XT stays below −10 dB, and thus, the approximation $\Upsilon \approx \xi_2 L$ holds. In this regime, signal loss due to XT and interactions between XT and fiber nonlinearities are negligibly small and are, therefore, not modeled. Importantly, though, it can be shown that the optimal signal launch power is mainly a tradeoff between ASE and NLIN, and is independent of XT [cf. (A12)]. Furthermore, note that, at a fixed transmission distance, the effect of XT on the SE is identical to that of $\mathrm{SNR}_{\mathrm{TRX}}$, as they both make up the signal-independent SNR contribution SNR_0 [cf. (9b)]. This is depicted in Fig. 17 by the cyan star marker where the individual contributions of an XT of −50 dB/km and a transponder noise floor at −30 dB ($1/\mathrm{SNR}_{\mathrm{TRX}}$) result in the same SE at $L = 100$ km. The combined effect is depicted by the green solid curve. At distances $L > 100$ km where the accumulated XT becomes gradually larger than $1/\mathrm{SNR}_{\mathrm{TRX}}$, the transponder noise becomes irrelevant. Conversely, at distances $L < 100$ km where XT is smaller than $1/\mathrm{SNR}_{\mathrm{TRX}}$, the effect of a transponder noise floor is noticeable. The presence of a transponder noise floor, therefore, makes short-reach systems up to about 100 km more resistant to XT, as $\mathrm{SNR}_{\mathrm{XT}} \gg \mathrm{SNR}_{\mathrm{TRX}}$ for such systems. On the other hand, for transmission distances where the system is not strongly influenced by $\mathrm{SNR}_{\mathrm{TRX}}$, the SE reduction (in bit/s/Hz) due to XT depends only on ξ_2 and is nearly independent of the transmission distance, reflected by XT curves being parallel to the XT-free curve at long transmission distances in Fig. 17 [174], [175].

Fig. 18(a) shows the percentage SE penalty due to XT as a function of transmission distance, with the length-dependent XT factor ξ_2 as a parameter. The SE penalty due to a certain XT (e.g., −40 dB/km) at a certain distance (e.g., 300 km) is given by $s_2/(s_1 + s_2)$ (with s_1 and s_2 as defined in Fig. 17 for an $\mathrm{SNR}_{\mathrm{TRX}}$ of 30 dB). For an SE penalty of less than 5% up to 10 000 km, an aggregate XT of about −60 dB/km must be maintained. Consequently, the optimum number of cores of an MCF (for any fixed cladding diameter up to ∼250 μm) is nearly independent of the transmission distance (i.e., independent of the system application), which opens the door to manufacturing universal MCFs for deployment in various system contexts [174]. Looking at the same situation from a MIMO perspective, Fig. 18(b) addresses the question of how much capacity one could gain back if one were to implement MIMO at the receiver to undo the XT by showing the MIMO gain. The MIMO gain is defined by $(s_1 + s_2)/s_1$. A MIMO gain of 3 dB implies that MIMO processing of parallel XT-inducing propagating paths doubles system capacity at the expense of significant DSP and

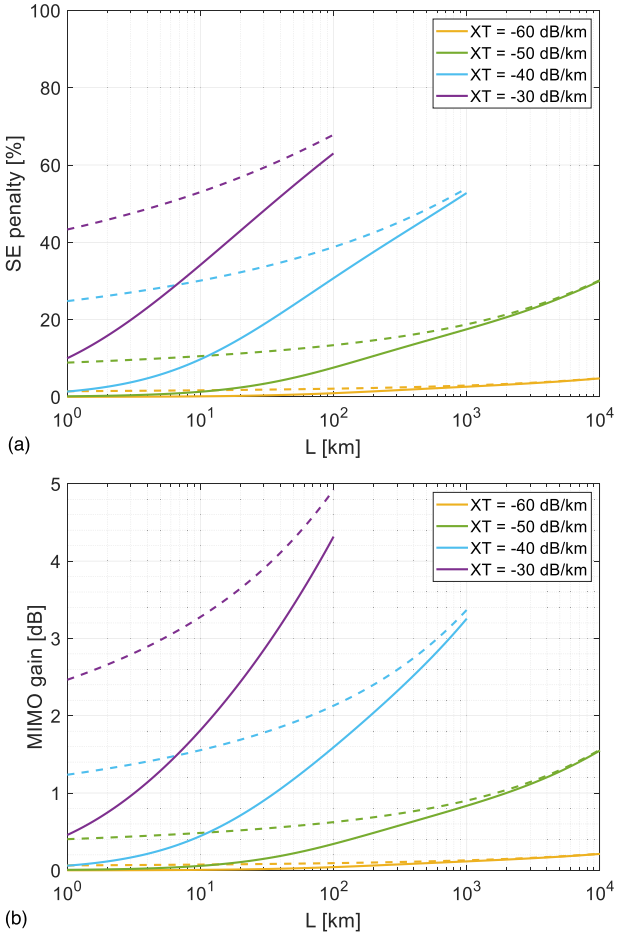

Fig. 18. *(a) Percentage SE penalty and (b) MIMO gain for various (aggregate) intercore XT values as a function of transmission distance L with (solid lines) and without (dashed lines) $\mathrm{SNR}_{\mathrm{TRX}}$ of 30 dB. All lines are plotted up to an accumulated maximum XT of −10 dB.*

hardware complexity. Note that, due to the dependence of XT on the wavelength (cf. Fig. 16), the SE reduction and MIMO gains due to XT pertain to the longest-wavelength WDM channels and the overall WDM system capacity loss due to XT will, in general, be smaller than the SE loss at the longest operating wavelength. Refined system designs may use this fact to transmit more information in the shorter wavelength portion of the spectrum than in the longer wavelength portion (e.g., by using different PCS factors across the band).

2) Coupled Spatial Parallelism (Mode Multiplexing): Once XT between spatial paths (whether between adjacent cores in an MCF or between different propagation modes in a single core) becomes too high to be tolerated in terms of an SE loss or, for fixed system settings (including modulation and FEC), becomes too high to allow for error-free operation at the allocated XT margin, the system exceeds its "MIMO threshold." In this case, MIMO DSP is needed to recover the individual signals at the receiver, and we speak of a coupled-mode or coupled-core transmission

system. Note that, in an unperturbed waveguide (i.e., a waveguide that does not change its properties in the longitudinal direction), true waveguide modes are by definition orthogonal and propagate independently and without any coupling. LP fiber "modes" are not true waveguide modes but represent a linear combination of true fiber modes, resulting in *deterministic* coupling at a given beat length between them, e.g., between the $LP_{11a,x}$ and $LP_{11b,y}$ modes [176]. This beating phenomenon is also known under the term "modal birefringence" [177]. The same is true when launching a linear combination of supermodes into an array of parallel waveguides by exciting only a single waveguide of that array, as it is done when sending light through, e.g., a directional coupler or an ideal (i.e., straight and unperturbed) CCF. The coupling process of deterministic mode coupling does not vary along the transmission path.

More important for coupled-mode fiber transmission systems, however, is *random* mode coupling along the transmission path due to random field perturbations induced by fiber imperfections, by temperature variations and external perturbations, such as fiber twists and bending affecting the deterministic mode beating, by splices and connectors with random offsets, and by nonideal spatial multiplexers/demulitplexers (MUXs/DEMUXs) (see Sections V-B and V-C). Random mode coupling, which is always present in some form in a realistic transmission scenario using polarization and/or spatial modes, is also the main reason why a coupled-mode transmission system beyond the MIMO threshold needs to resort to MIMO DSP.

In an FMF [145], a special kind of MMF, spatial channels are typically represented by spatial (pseudo)modes, such as the LP modes [69]. An MMF can easily support more than a hundred spatial modes without the need to increase the cladding diameter. An FMF for MIMO-assisted SDM transmission is designed to support typically a small amount of low-order spatial modes to keep the MIMO complexity at the receiver low. A CCF is typically made from a few (e.g., 4 or 7; see Fig. 20) homogeneous single-mode cores that are located just at the right distance to each other to enable strong random coupling [144]. Such a fiber supports as many spatial supermodes as it has cores. Due to the strong random field coupling between the spatially separated cores, uniform CCFs exhibit more uniform mode coupling characteristics compared to FMFs. Irrespective of the coupled-mode fiber type, strong random coupling between modes is preferable to suppress transmission impairments, such as differential mode delay (DMD) and mode-dependent loss (MDL) [178], [179], both of which affect the performance of the MIMO DSP in terms of system outage probabilities [180]. Note, in this context, that purely unitary (i.e., lossless) mode coupling does not affect MIMO system capacity. For systems in which the transmitter does not use channel state information (CSI), which is the typical assumption in optical MIMO-SDM systems, nonunitary mode coupling (i.e., MDL) due to components closer to the transmitter (i.e., before channel noise loading) has a stronger impact on MIMO capacity than MDL of components closer to the receiver [180]. In systems using CSI at the transmitter, MDL occurring prior to random channel perturbations (e.g., in the mode MUX) could be perfectly precompensated at the transmitter and would have no impact on the system capacity.

DMD (also referred to as differential group delay (DGD), group delay spread, or modal dispersion) is caused by the differences in group delay between the spatial modes. Modal dispersion lets a single impulse launched into an FMF spread over a time window whose width is proportional to the DMD. To compensate for this spread, MIMO DSP requires enough filter taps (i.e., a sufficiently long memory or filter length) to accommodate the DMD [181]. For strong random coupling (as is found, e.g., in optimized CCFs), the impulse response takes on a Gaussian shape, and DMD (proportional to the width of this Gaussian distribution) scales moderately, with the square root of transmission distance [182]. This scaling is also reflected in the unit $ps/km^{0.5}$ typically used to describe modal dispersion of strongly coupled spatial modes, in analogy to its traditional use in describing the DGD due to PMD in SMFs. It was shown both theoretically [183], [184] and experimentally [185] that strong coupling among spatial modes also reduces nonlinear impairments and, thus, enhances transmission performance.

MDL is a measure of the power variations between modes due to individual modal loss or gain. MDL accumulates during propagation and, when specified as the ratio of the strongest to the weakest mode, can range from 1 (0 dB) for no power variations between modes to infinity if the power of one mode drops to zero, i.e., a mode is completely filtered out. MDL extends the definition of polarization-dependent loss (PDL) to coupled-mode systems. MDL is caused by nonideal optical network elements, such as OAs or MUXs, as well as by perturbations along the fiber such as bends and imperfect splices. MDL affects the quality of the MIMO channel and, in conjunction with amplifier noise, sets a fundamental limit to the performance of a mode multiplexed transmission system. The specification of MDL as the ratio of strongest to weakest modal powers (peak-to-peak MDL) lends itself best to lumped system components, such as mode MUXs, optical filters, and amplifiers. In distributed systems, the combination of random mode coupling and MDL leads to a statistical modal loss distribution that requires knowledge of statistical MDL parameters, such as the root-mean-square value (rms MDL) [186], and, ideally, the probability distribution of MDL [180] for conclusive statements on the impact of MDL on system capacity and outage. Also, note that the modes determining MDL, in general, do not correspond to the modes that are launched into the fiber (e.g., LP modes) but are found as the eigenmodes of the end-to-end MIMO channel, including all its random and deterministic perturbations.

While CCFs are usually designed to form only a single group of strongly coupled modes (with their effective

refractive indices or propagation constants very close to each other), FMFs and MMFs typically form multiple groups of strongly coupled modes, with relatively weak random coupling between mode groups (MGs) [187]. This results in a less favorable linear growth of DMD with distance, which increases the MIMO DSP complexity if the crosstalk between MGs needs to be removed as well. This linear growth is reflected in the unit ps/km, which is typically used to describe the modal dispersion of weakly coupled modes. The formation of MGs in FMFs can suppress coupling between MGs to a negligible level over modest distances such that MIMO only needs to be performed *within* MGs but not across MGs, i.e., MIMO is only used to compensate for intra-MG coupling but not for inter-MG coupling [188], [189]. Successful transmission of up to 1800 km of two MGs was recently demonstrated in conjunction with coherent detection and MIMO processing only within each MG [190]. Optimizing the separation of the effective refractive indices of the various MGs allows for fiber designs suitable for such systems. If all modes within an MG carry the same information, each MG acts as an individually isolated spatial channel, and MIMO DSP can be avoided altogether, which is referred to as MG multiplexing. MG multiplexing without MIMO processing at the receiver in combination with OOK and direct detection has been demonstrated for various transmission scenarios [191], [192], [193], [194] and transmission distances up to ∼150 km [195].

3) Combinations of Coupled and Uncoupled Spatial Parallelisms: Combinations of coupled and uncoupled SDM fiber types to achieve both core multiplexing and mode multiplexing have also been demonstrated. In *few-mode* MCFs (FM-MCFs), each core of an uncoupled-core MCF supports multiple coupled spatial modes. The number of spatial channels provided by the MCF can, thus, be more easily multiplied by tolerating a modest increase in MIMO complexity per core. Each core effectively forms a local MG that is independently maintained over long distances and does not couple much with other (spatially separated) MGs (i.e., cores). The MIMO threshold between nominally uncoupled groups of coupled spatial channels, e.g., the cores in an FM-MCF, was found to be the same as the MIMO threshold between individual single-mode channels, e.g., the cores in a single-mode MCF [196]. However, the fact that higher order mode fields spread further into the cladding [149], [155] requires a larger core pitch for such fibers, a larger outer cladding thickness, and, often, also a heterogeneous trench-assisted core profile in the fiber design to suppress intercore XT, which results in cladding diameters much larger than 125 μm of SMFs, FMFs, or MMFs. Nevertheless, FM-MCFs have so far achieved the highest spatial channel count in a single strand of fiber. Fibers supporting up to 120 spatial channels (12 cores × ten spatial modes per core, with a cladding diameter of 217 μm [197]) have been reported for fabricated fibers and up to 114 spatial channels (19 cores × six spatial

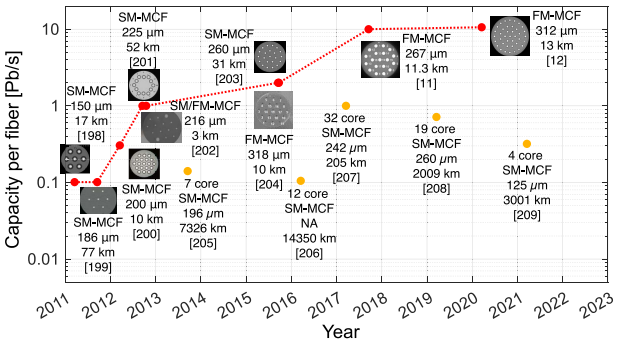

Fig. 19. ***Per-fiber capacity of record single-span MCF transmission experiments (red markers) and a few representative multispan MCF transmission experiments (orange markers). Additional information is given on fiber type, cladding diameter, and transmission distance. (SM-MCF: single-mode MCF; FM-MCF: few-mode MCF.)***

modes with a cladding diameter of 267 μm [11] and 38 cores × three spatial modes with a cladding diameter of 312 μm [12]) have been used in conjunction with high capacity transmission experiments achieving transmission capacities of ∼10 Pb/s (see Fig. 19). Note that, in accordance with (2), polarization is treated separately from spatial modes and is, therefore, not included in the above count of "spatial channels."

4) Transmission Experiments Using SDM Fibers: Figs. 19 and 20 summarize the evolution of record single-span experiments together with a few representative multispan experiments that have used SDM-specific fiber types. FM-MCFs with 114 spatial channels have demonstrated the so-far highest data throughput of more than 10 Pb/s in a single strand of fiber [11], [12]. CCFs with up to seven spatial channels (seven coupled single-mode cores) have demonstrated SDM transmission over trans-Pacific distances over more than 9000 km, outperforming SMFs in terms of their nonlinear transmission characteristics [185]. As shown in Fig. 19, all MCFs used in experiments with transmission rates above 2 Pb/s have employed larger cladding diameters than 250 μm. In view of this constraint and the experimental results presented so far at top conferences, we conclude that it will be difficult to boost the number of spatial channels incorporated in a single strand of silica MCF much beyond a factor of a few tens compared to SMF in practical long-distance transmission scenarios. On the other hand, all fibers underlying the results summarized in Fig. 20 have a cladding diameter of 125 μm, with the tolerable MIMO DSP complexity at the transponder dictating the maximum number of spatial channels.

Apart from these impressive transmission results in the laboratory, a practically important and highly relevant question when assessing the scalability of MCFs, FM-MCFs, FMFs and CCFs concerns the related savings on a systems' level (including fiber fabrication). So far, bundles of SMFs are still more cost-effective compared to SDM-specific

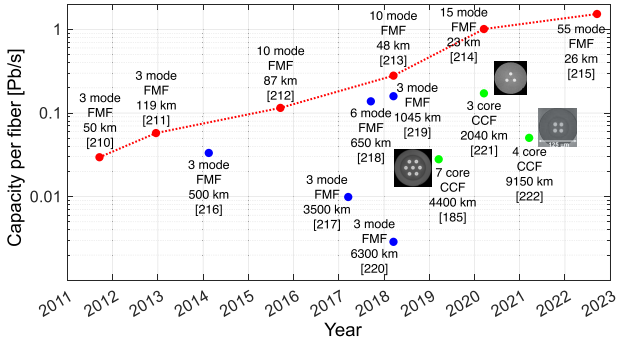

Fig. 20. ***Per-fiber capacity of record single-span FMF transmission experiments (red markers), multispan FMF transmission experiments (blue markers), and CCF transmission experiments (green markers). Additional information is given on spatial mode count, fiber type, and transmission distance. In all cases, the cladding diameter is 125 μm.***

fiber types and avoid higher order MIMO-DSP altogether. Furthermore, given the wide range of possible cladding diameters and the plethora of conceivable core layouts and mode multiplicities, it is clear that standardization efforts [60] are crucial for the further development of SDM components and their adoption in commercial systems.

5) Cabling: The number of optical fibers that can be packed into an optical cable depends mainly on the allowable cable diameter and weight. *Submarine* cables are strongly limited in their allowable diameter to ensure the highest level of reliability under deep sea conditions. In *terrestrial long-haul, regional,* and *metro systems*, limited available duct space and difficulties in pulling long cables through existing ducts are of consideration, as is the ability to transport heavy cable drums to locations accessible by construction crews for deployment. The shorter practically deployable individual cable segments have to be, the more splice points are needed, which adversely affects loss and deployment costs. While DCI systems often use newly deployed ducts, it is preferable in most other terrestrial network segments to utilize existing ducts as efficiently as possible, as the deployment of new duct infrastructure outside a datacenter campus environment can be prohibitively expensive. Fig. 21 depicts an example of how the fiber count can be upgraded in an existing duct. Fig. 21(a) shows a duct (e.g., 60 mm in diameter) housing a single high-density optical cable (A1) that can accommodate more than 100 optical fibers with a cable diameter of ~10 mm, as is typically used today. Such high-density optical cables are made from partially bonded optical fiber ribbons that, similar to traditional edge-bonded or encapsulated fiber ribbons, have the benefit of enabling very dense packing in the cable while allowing for simultaneous fusion splicing of multiple optical fibers using a ribbon splicer. In one upgrade scenario [see Fig. 21(b)], the duct is gradually filled with additional high-density optical cables of the same type (A2, A3, and A4). In another upgrade scenario [see Fig. 21(c)], the remaining duct space is filled with a single ultrahigh-density optical cable with a larger cable diameter (A2'). Such ultrahigh fiber count cables can be fabricated by leveraging the technology used for high-density cable and optical fibers with a thinner coating diameter.

Although cables with a larger diameter provide more instantaneous capacity, their use increases the installation workload, i.e., a thicker and heavier cable restricts the installation length, which also increases the overall number of splice points. Moreover, the installation of an ultrahigh-density cable makes it more difficult to add further cables at a later point in time. Manufacturing of these cables also entails a longer total fiber length (TFL) per cable, where TFL refers to the product of the cable length times the number of fibers in the cable. With requirements toward carbon neutrality of optical networks getting more stringent in the coming years, the CO_2 emission generated during the life cycle of a fiber cable cannot be neglected anymore, with fiber drawing being the dominant factor, as recent studies have revealed [223]. The TFL per cable will, therefore, have a significant impact on choosing cable types. Against this backdrop, a sensible choice could be to first use multiple high-density cables [A1–A4 in Fig. 21(b)] whose installation workload and TFL per cable remain the same. Using cables with smaller diameters also facilitates the adjustment of the remaining duct space relative to the cable space. Eventually, however, this upgrade scenario will result in the need for another duct once the original duct is filled up with the maximum number of cables. Before this point is reached, the smaller diameter optical cables have the benefit of facilitating a gradual replacement of the cables containing conventional SMFs [A1–A4 in Fig. 21(b)] with cables containing SDM fibers, e.g., a replacement with B–D in Fig. 21(d) may accommodate SDM fiber cables with an increasingly higher spatial multiplicity compared to their SMF counterparts A1–A4. SDM technology may, thus, enable a gradual increase in capacity while keeping the installation workload and TFL per cable constant. Assuming that four cables are the maximum number of cables that the duct of Fig. 21 can

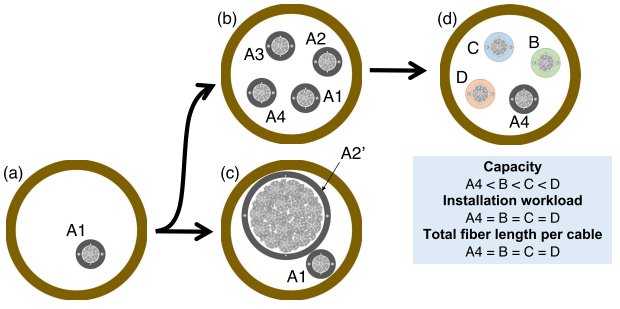

Fig. 21. ***Example of the sustainable use of duct resources.***

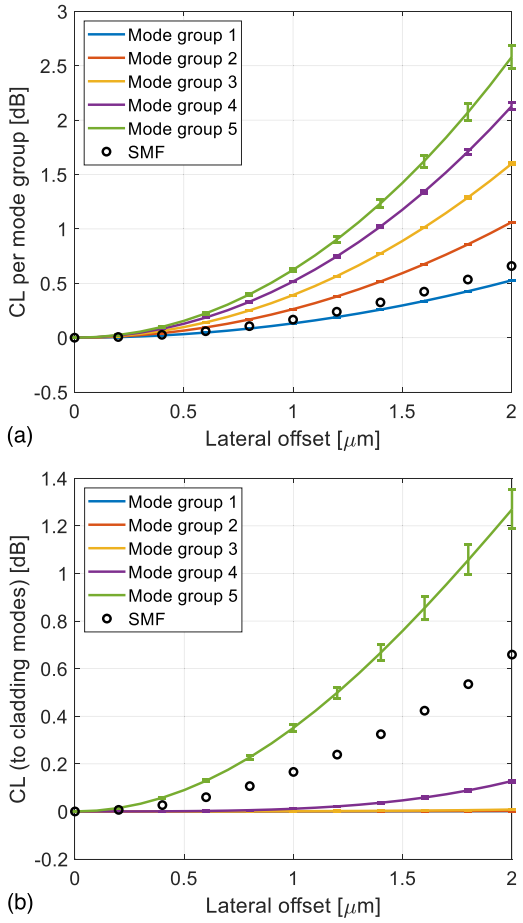

Fig. 22. *Calculated average power coupling loss (CL): (a) between the same MGs and (b) between MGs and nonpropagating cladding modes due to a splice as a function of a lateral offset at the splice point for an FMF supporting five MGs or an SMF. The vertical bars indicate loss variations within each MG.*

support while still allowing for cable replacements, duct sustainability could be more easily achieved. This would be done by first adding cable B, next moving services from cable A1 to B, and, finally, taking out cable A1. The same process would be repeated until the duct is filled up with SDM cables of maximum spatial multiplicity. Switching from the SMF cable to an SDM cable with higher spatial multiplicity may also mitigate the risk of link loss during cable replacement as the newly added SDM cable with higher spatial multiplicity provides a higher capacity than the replaced cable right away.

Although the above considerations indicate that physical limitations and sustainability of ducts and cables serve as key motivations to use MCFs in long-haul networks, we note that there is currently no commercial technology available that allows for *simultaneous* fusion splicing of MCF arrays or ribbons in a field-deployable setting. This is one of the remaining technical challenges that need to be resolved to realize commercially viable MCF-based networks.

B. Splices and Connectors

Splices and connectors are essential elements of optical transmission systems and networks. As most cores of an MCF are not located in the fiber center, the alignment of MCFs upon splicing or connectorization requires precise control of the core position in both radial and angular directions with respect to the fiber center. Both lead to a lateral offset between coupling core pairs at splices and connectors. The open markers in Fig. 22 show the coupling loss (CL) of a single-mode core as a function of a lateral core-to-core offset and apply to both SMFs and MCFs with single-mode cores. For MCFs, the lateral core-to-core offset can be due to a lateral offset between the two MCFs but can also be due to an angular misalignment of the two MCFs, which results in a lateral offset between MCF cores that increases with the distance of the core from the fiber center, making the cores closer to the cladding-coating boundary more prone to CLs. This is especially critical for high-core-count MCFs with large cladding diameters. For example, an angular misalignment of 1° causes a lateral offset of ∼0.87 μm at a distance of 50 μm from the fiber center and a lateral offset of ∼1.75 μm at a distance of 100 μm from the fiber center. As a point of reference, typical angular alignment tolerances for polarization-maintaining SMFs today are ±3° (for standard grade) and ±1.5° (for high grade). Similar to polarization-maintaining fibers, to mitigate issues associated with angular misalignment of MCFs, fusion splicers equipped with both end- and side-view functions can be used to inspect the core arrangement on the fiber facet while rotating the fiber before splicing [224], [225], [226]. Also, new types of MU-, SC-, and LC-type MCF fiber connectors have been developed, which avoids unintentional rotation of the fiber upon connection [227], [228]. The latest reported LC-type MCF connectors have demonstrated less than 0.07-dB insertion loss on average for a 125-μm-cladding 8-core fiber with a fiber-center-to-core-center distance of 40.5 μm and are ready for mass production [229]. Multi-MCF connectors have also been demonstrated with up to 256 cores (32 × 8-core MCFs) [230]. In addition, due to the 2-D core layout, there are also some network applications where MCF cables need to distinguish between “input ports” and “output ports” (“MCF polarity”) to guarantee the connection of cores with the same core ID [231]. An example is the directionally interleaved cores discussed in Section V-A1.

FMFs, on the other hand, employ only a single core in the fiber center. Angular misalignment is, therefore, irrelevant for cores with a circularly symmetric refractive index profile. However, the tolerances to the lateral misalignment are more severe for higher order modes, as illustrated in Fig. 22, with respect to an FMF designed to support 5 MGs (i.e., 30 spatial and polarization modes) at a wavelength of 1.55 μm with a graded-index profile (Δ = 1%) and a core diameter of 28 μm [232]. Fig. 22(a) shows the MG CL, i.e., the power loss of an MG when coupled

to the same MG after the splice. The curves represent the average loss within the group, while the vertical bars (only visible in the higher order MGs) indicate the loss variations within each MG. MG 1 containing the two fundamental (polarization) modes of LP_{01} exhibits a slightly better offset tolerance than that of the standard SMF (core diameter of 8.5 μm and a step-index profile with $\Delta = 0.35\%$). Note, though, that the coupling of any particular mode to the *same* mode includes both true power loss and power coupling to *other* modes. In a MIMO-system context, coupling to other propagating modes is irrelevant, which makes the power loss at the splice point the much more relevant quantity as it directly leads to an increase in MDL [233], [234]. Fig. 22(b), therefore, shows the CL of various fiber modes to nonpropagating cladding modes after a lateral offset splice. In the case of an SMF, the MG CL is equivalent to the loss due to coupling to the cladding modes. Being closest to the cladding index in terms of modal indices, the highest-order MG naturally is affected most strongly, and this MG (i.e., MG 5 in our example) will also experience the strongest degradations in transmission quality, unless MGs strongly couple to each other. If they strongly couple, MG coupling in the fiber following the offset splice randomizes the effect. On the other hand, the low-order MGs are more robust against lateral offsets as they mainly exchange power among themselves, which could be compensated for by MIMO DSP across MGs.

We note that all these calculations are based on the fiber's circular-symmetric true mode field profiles to avoid numerical artifacts that may lead to a misleading dependence of the results on noncircular symmetric LP mode profiles (of modes within the same MG, e.g., LP_{11a} and LP_{11b}). LP modes are often used to describe mode propagation in optical fibers in a simplified approximation, but their individual effects average out over longer transmission distances due to strong mode coupling of degenerate or nearly degenerate modes within the same MG.

C. Spatial Multiplexers and Demultiplexers

The role of a spatial MUX and DEMUX is to efficiently and as unitarily as possible couple light from a bundle of SMFs onto the cores of an MCF or the modes of an FMF, MMF, or CCF, and vice versa. In the case of a MUX for a single-mode MCF or CCF, the coupling mechanism is straightforward and maps an array of single-mode input beams (e.g., originating from an SMF bundle) to an array of single-mode output beams whose spatial distribution is matched to the core layout of the MCF or CCF. This has been achieved with free-space optics [235], [236], 3-D waveguides [237], [238], and reduced-cladding or pitch-reduced fiber bundles [239], [240], including pluggable types [241]. The MUXs and DEMUXs for an MCF or CCF are also often referred to as fan-in–fan-out (FIFO) devices. FIFOs for MCFs are not XT free, but the level of XT is usually small (typically below -50 dB on average) compared to the intercore XT accumulating along the fiber upon transmission.

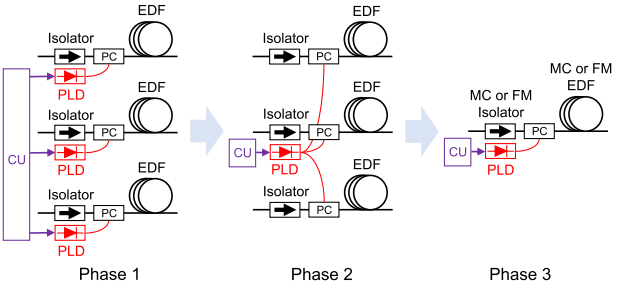

Fig. 23. *Integration phases of OAs for parallel systems (CU: control unit; PLD: pump laser diode; PC: pump combiner; EDF: erbium-doped fiber; MC: multicore; and FM: few-mode).*

In the case of an FMF, a MUX couples spatially separated single-mode input beams onto the modes of the FMF. This coupling does not need to happen to any particular mode basis, as long as the orthogonality of the spatial channels is sufficiently preserved, i.e., as long as the MUX implements a (nearly) unitary transformation between its input fundamental modes of SMFs and its output guided FMF modes. Numerous devices have been developed over recent years that can achieve this transformation. Among these are phase masks [242], [243], directional couplers implemented either in form of 3-D waveguides [244] or planar waveguides [245], photonic lanterns that can be directly fused to both the SMF array and FMF [246], and multiplane light converters [247] that were even demonstrated to work in conjunction with >1000 modes [248]. Modal XT due to an FMF MUX is somewhat higher (typically below -25 dB on average) than the typical XT of FIFOs but again negligible compared to the modal XT accumulating along the fiber. However, care must be taken that the MUX's nonideal (i.e., nonunitary) coupling behavior does not significantly contribute to the overall system MDL.

D. Optical Amplifiers

One of the most widely used active network elements, the OA, is expected to greatly benefit from array integration. Wall-plug efficiencies (or energy conversion efficiencies) of amplifier line cards used in today's networks range from slightly less than 1% to a few percent [48], which is much below the fundamental OA efficiency limit. The pump-to-signal photon conversion efficiency is ~60% assuming a 980-nm pump and a 1550-nm signal, and typical pump lasers have efficiencies of up to ~50%. This indicates that there is still ample headroom for improvement in the efficiency of amplifying optical signals. The development toward a fully integrated SDM amplifier is likely to happen in several phases, as illustrated in Fig. 23. In the first phase, each spatial channel has its own EDF and pump laser unit, but common housekeeping functions, such as laser diode (LD) control, monitoring, and thermoelectric cooling, are shared among all units. In the second phase, the output of a single or a few high-power LDs is distributed among multiple spatial channels

to pump their individual EDFs. This is done both for energy efficiency reasons and to provide redundancy in the case of LD failure. The latest submarine cable systems employing pumping schemes known under the term "pump farming" [249] are already pertinent to this phase of SDM amplifier development. Individual EDFs, isolators and WDM couplers for pump combining, however, are still used for each spatial channel. This phase may also comprise methods that efficiently reuse the unabsorbed pump light after passing through the EDF [250]. Finally, in the third phase, the input and output fibers, as well as the entire EDFA, are built from SDM components, such as MCFs or FMFs and multicore or few-mode isolators [251] with one or a few high-power LDs pumping the cores of the MCF or modes in the FMF. In this respect, two pumping concepts have been devised in recent years: EDF cores are pumped either directly, or they can be pumped indirectly via the cladding [252].

So far, SDM-OAs have not yet achieved efficiencies exceeding those of conventional single-mode EDFAs. Core pumping of MCFs requires either fan-out devices to couple the pump light to each core individually [253] or uses WDM couplers built with free-space optics [254]. Cladding pumping can make use of side-pumping schemes and high-power multimode LDs, and has been demonstrated for both MC- and FM-EDFAs [255]. FM-EDFAs can be a significant source of MDL (due to gain variations among the modes), and various amplifier designs have been devised with the goal to minimize MDL [256], [257], [258]. As it will be difficult to completely eliminate MDL in amplifiers solely by optimizing the refractive index profile of the active fiber, the use of gain flattening filters in the spatial domain at the amplifier output, analogous to the way they are being used to compensate for gain variations in the spectral domain in conjunction with single-mode EDFAs, has been investigated to deal with residual MDL [259].

E. Network Nodes and Switching

Optical network nodes are used to aggregate, disaggregate, and switch optical signals between their various inputs and outputs, transparently on a wavelength granularity, thereby enabling the ability to perform optical mesh networking. Today's network nodes are implemented as colorless, directionless, and contentionless ROADMs, enabling optical signal routing of any wavelength, in any nodal direction, and without blocking other wavelength channels, provided that the respective wavelength slot is still available at the targeted output port.

In future networks, ROADMs will have to increasingly complement wavelength switching with spatial switching to accommodate high degrees of spatial parallelism [260], [261]. In addition, as the bit rate of logical router interfaces in the core network approaches the capacity of a full WDM system, wavelength switching will be phased out in favor of space switching only [48]. This will not only substantially simplify the network architecture

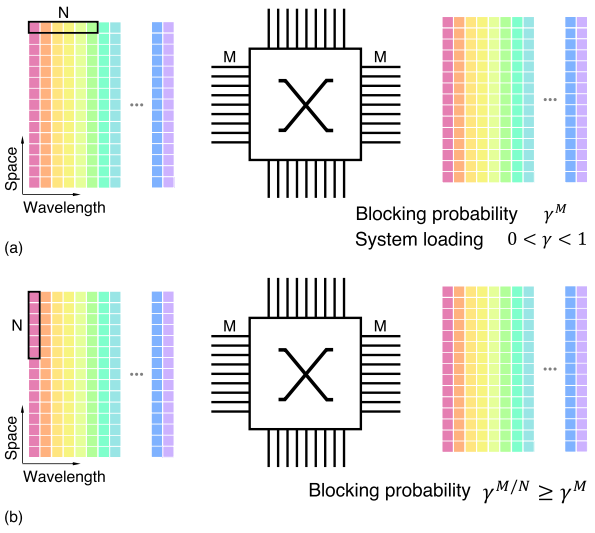

Fig. 24. ***Comparison of the blocking probability of a ROADM node that uses either (a) spectral superchannels or (b) spatial superchannels (after [3]).***

itself but will also make wavelength routing algorithms and amplifier transient control systems no longer needed. Transient control is needed today to avoid the transient impact of OAs due to the sudden wavelength add/drop actions onto other channels in a WDM system. Wavelength routing algorithms are needed because transparent optical network nodes currently cannot perform wavelength conversion [262]. A routing request needs to occupy the same wavelength range from source to destination, which can lead to stranded bandwidth and defragmentation issues in WDM networks.

It is important to note that spatial switching and spectral switching are fundamentally different in terms of their ability to perform contention resolution, as wavelength conversion is a nonlinear operation and, hence, not easily done all-optically, while spatial path conversion is a linear operation that is easily implemented. To illustrate this point and the associated networking implications, Fig. 24 shows two optical switching scenarios for a node with four directions and M parallel spatial paths connected to each direction, each spatial path carrying a WDM system [3]. In the upper scenario [see Fig. 24(a)], logical ports are implemented as *spectral superchannels*, i.e., a logical data stream consists of multiple wavelength signals on a single spatial path. In the lower scenario [see Fig. 24(b)], *spatial superchannels* are being used, i.e., a logical data stream consists of same-wavelength signals on multiple spatial paths. Hybrid cases are also possible. To illustrate network operational differences between spectral and spatial superchannels, assume that a superchannel occupying N contiguous spatial or wavelength slots is to be routed from west to east in the nodes of Fig. 24. Further assume that the system loading at the output port (East) is given by $0 < \gamma < 1$, i.e., the probability that a superchannel occupying N contiguous spatial or wavelengths slots is already occupied

at the output is γ. In the case of a *spectral superchannel*, routing is impossible if all M spatial paths already carry a spectral superchannel in the needed wavelength range, which leads to a blocking probability of γ^M. In contrast, routing for the *spatial superchannel* is impossible if there are not enough wavelength slots available at the output port to fit a spatial superchannel of N spatial slots, leading to a blocking probability of $\gamma^{M/N} > \gamma^M$. This example shows the asymmetry in space and wavelength when it comes to switching applications, which is fundamentally rooted in the fact that spatial switching is a linear process, while wavelength switching is inherently nonlinear and therefore cannot be easily implemented [3]. Note, though, that the choice between spatial and spectral superchannels is not always optional. For example, if MIMO-SDM is being used, all coupled modes belonging to a single transmitter–receiver pair must *stay together* upon transitioning through an optical network node. Despite research reports being published now and then stating otherwise, the constituents of a coupled-path spatial superchannel *cannot* be independently routed without destroying the integrity of the signal. This is analogous to polarization multiplexing in an SMF system: a network node in an SDM system must never route the two polarizations of a single wavelength channel to different output ports.

Whether spatial or spectral superchannels are the better option depends on the network scenario: the use of MIMO forces spatial superchannels, and spatial superchannels also have benefits in terms of transceiver hardware simplicity and power consumption. On the other hand, optically routed mesh networks favor spectral superchannels as a logical entity in terms of blocking probability and in terms of switch node architectures, especially once the logical port capacity of an IP router becomes comparable to the capacity of an entire fiber [48]. For a detailed discussion on the node architecture and networking of SDM networks, the interested reader may refer to [263] in this Special Issue.

F. Transponder Hardware

In the context of transponder hardware, SDM integration asks for massive array integration in order to make the transponders addressing M spatial paths more cost-efficient than M individual transponders. While this general desire also applies to the integration of spectral superchannel transponders (i.e., transponders addressing N wavelengths), spatial instead of wavelength integration offers distinct advantages, including the use of a single laser source, no need for spectrally gain-flattened optical amplification within a transponder, and no need for skew compensation of different-wavelength signals sent in parallel. In a spectral superchannel, CD causes significant temporal walk-off between wavelength constituents of a spectral superchannel. For example, the two outermost wavelength constituents of a spectral superchannel with ten wavelengths on a 100 GHz grid in the C-band experience a walk-off of ~120 ns after 1000 km of SMF propagation. This skew needs to be compensated at the receiver in order to interpret the data as a single logical stream.

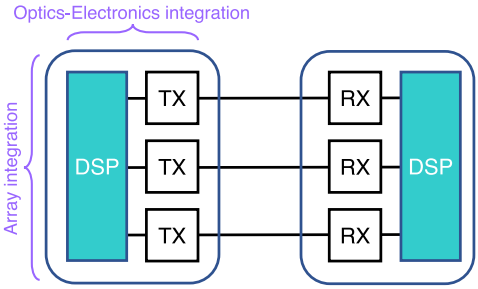

Fig. 25. *Transponder array integration (after [264]).*

As shown in Fig. 25, and as further discussed in [264], there are three aspects to transponder array integration: 1) *optoelectronic array integration*, i.e., the placement of multiple elements of the same kind onto a commonly integrated chip, such as modulator arrays, electrical amplifiers arrays, or multichannel DSP-ASICs; 2) *optics–electronics integration*, i.e., the close integration (hybrid or monolithic) of the optoelectronic array and the CMOS DSP ASIC; and 3) *holistic DSP integration*, i.e., the codesign of DSP functions to compensate for performance shortcomings due to the high integration density of the optoelectronic hardware components. An example of the latter is the possibility to digitally compensate for crosstalk in densely integrated modulator arrays [265]. Note that multichannel ASICs are limited both by their digital (co)processing abilities and by their ability to interface with multiple high-speed data signals (cf. footnote 9 for an example of the required data rates). The DSP complexity to perform MIMO processing seems manageable at a constant interface rate of a spatial superchannel [266].

G. System Considerations and Tradeoffs

Table 3 summarizes the main characteristics of the above-discussed three types of SDM fibers (MCF, FMF, and CCF) in terms of the link, OA, transponder, and node design. A green circle indicates similar performance (or complexity) to the equivalent network element in SMF systems, and an orange triangle indicates higher complexity.

MCF and CCF employ traditional refractive index profiles, which allows the fabrication of pure silica cores with similar loss characteristics as SMFs. On the other hand, the graded-index profile of FMFs requires doping in the core, leading to intrinsically higher propagation losses. In terms of fabrication complexity, the presence of multiple cores not located in the fiber center of MCFs and CCFs requires not only precise control of the core pitch but also the development of new fabrication processes of the fiber preforms. In addition, the drawing of MCFs with cladding diameters exceeding that of the SMF reduces production yield for a fixed preform volume. A preform with a larger volume than the one currently used for SMFs would

Table 3 Comparison of SDM Fiber Characteristics

Network element	Characteristics	MCF	FMF	CCF
Link	Propagation loss	●	▲	●
	Fabrication complexity	▲	●	▲
	Cabling	●	●	●
	Splicing	▲	●	▲
OA	PSM	●	●	●
	Native solution	▲	▲	▲
TRX	Receiver complexity	●	▲	▲
Node	Spatial switching	●	▲	▲

OA: optical amplifier, PSM: parallel single-mode fibers, TRX: transponder

require significant modifications in both manufacturing facilities and manufacturing processes.

The FMF requires a precise control of the graded-index profile shape to suppress DMD but can be manufactured with traditional fabrication processes. All SDM fiber types are compatible with existing cable designs. All fiber types also fulfill the standard requirements regarding bending loss. Compared to SMFs, MCFs need to manage intercore XT, and FMFs and CCFs need to manage modal properties, such as DMD and MDL. All these properties are sensitive to cabling and installation whose effects need to be considered in the fiber design phase. In terms of connectivity, the need for angular alignment in the case of MCF and CCF makes the connection and splicing process more complex. Ribbon splicing (or mass-fusion splicing) of MCFs is currently still an unresolved technical challenge. Note, however, that also the lateral alignment during the splicing of FMFs can be critical to suppress MDL, as discussed in Section V-B. Regarding the connectivity to amplifiers, all SDM fibers can be readily connected via FIFO devices or spatial mode MUX/DEMUXs (in conjunction with an array of parallel SMFs) to existing single-mode OAs. With improved technology, more energy-efficient SDM OAs (i.e., native solutions) are expected to replace single-mode amplifiers in the future. MCFs can be also designed to be directly connected to integrated transmitter and receiver arrays, while FMFs and CCFs typically need mode MUXs and MIMO DSP to jointly detect and process multiple spatial channels. When applying MCFs to network nodes, all spatial and spectral channels can be switched individually. On the other hand, in a mode-multiplexed link using FMFs or CCFs, individual coupled modes cannot be separately switched. This limits the switching granularity of FMFs and CCFs to an MG or even the whole optical fiber at a given wavelength as a switching unit.

In summary, the MCF offers better compatibility with existing transmission technologies, but the nature of the 2-D core arrangement requires specific fabrication processes and specific network operation if MCF polarity is of concern. Also, the number of cores that can be accommodated in MCFs with a standard cladding diameter of 125 μm is strongly limited. Expanding the cladding diameter will require extensive further investment in manufacturing facilities and processes. On the other hand, FMFs may require MIMO DSP (for spatial demultiplexing) but are compatible with traditional single-core technologies and network operation. CCFs also require MIMO DSP, but the effect of DMD can be minimized, which makes them suitable for long-distance point-to-point links.

VI. CONCLUSION

We have examined a variety of methods to increase capacity in optical networks, confirming that massive spatial parallelism (SDM) is the *only* viable option to effectively deal with an unabated growth in network traffic. Spatial parallelism may be supported by UWB systems. For terrestrial networks, UWB systems will be used as long as they make economic sense due to the high costs of new fiber deployments. On the other hand, the power constraints of submarine systems, which comprise newly deployed cables for each installation anyway, as well as the massive amounts of data transfer and associated energy challenges of data centers, are already making SDM the preferred choice. In the long term, efficiently integrated spatial parallelism will eventually become an unavoidable requirement and will dominate all network segments. ■

APPENDIX

A. Model for the Nonlinear Fiber Channel

Throughout the analyses presented in this article, we assume a WDM system loaded with ideal Nyquist (sinc) pulses, i.e., WDM channels with rectangular-shaped spectra ($B_{\text{ch}} = R_S$), and ideal Gaussian modulation constellations as approximated, e.g., by PCS with a sufficiently large QAM template. The system bandwidth of a WDM system loaded with N spectral channels is then $B = NB_{\text{ch}}$. Furthermore, we assume that all spectral channels within the system bandwidth B have uniform launch power P_{ch}. Except for Fig. 3, the transmission medium is assumed to be a standard SMF with a uniform loss coefficient $\alpha = 0.2$ dB/km, dispersion parameter $D = 17$ ps/nm/km, and a nonlinear coefficient $\gamma = 1.3\ (\text{W}\cdot\text{km})^{-1}$. In the case of Fig. 3, we assume a low-loss and low-nonlinearity fiber (SMF_{LL}), with a uniform loss coefficient $\alpha = 0.16$ dB/km, dispersion parameter $D = 21$ ps/nm/km, and a nonlinear coefficient $\gamma = 0.8\ (\text{W}\cdot\text{km})^{-1}$ corresponding to typical fibers used in ultralong-haul submarine transmission systems.

The SNR of the linear channel without implementation penalty and without linear crosstalk, i.e., $\eta = 1$ and $P_{\text{XT}} = 0$, is given by

$$\text{SNR}_{\text{L}} = \frac{P_{\text{ch}}}{P_{\text{SH}} + P_{\text{ASE}} + P_{\text{TRX}}}. \tag{A1}$$

The SNR of the nonlinear channel is given by

$$\text{SNR}_{\text{NL}} = \frac{P_{\text{ch}}}{P_{\text{SH}} + P_{\text{ASE}} + P_{\text{TRX}} + P_{\text{NLIN}}}. \tag{A2a}$$

Note that, in the main text, we omit the subscript "NL" when we refer to $\mathrm{SNR}_{\mathrm{NL}}$. In a dual-polarization transmission system, such as exclusively considered in this article, P_{ch} represents the sum of the channel's launch power in both polarizations. In an optically amplified (multispan) system, the received signal power equals the signal launch power, as the signal power is restored at each amplifier output (*including* an amplifier following the last span, i.e., directly at the receiver). In an unamplified single-span link, the received power is $P_{\mathrm{ch}}e^{-\alpha L}$ [see Fig. 7(c)] with α expressed in[13] Neper/m or Neper/km. In this case, the $\mathrm{SNR}_{\mathrm{NL}}$ can be written as

$$\mathrm{SNR}_{\mathrm{NL,\,single}} = \frac{P_{\mathrm{ch}}e^{-\alpha L}}{P_{\mathrm{SH}} + P_{\mathrm{TRX}}e^{-\alpha L} + P_{\mathrm{NLIN}}e^{-\alpha L}}. \quad \text{(A2b)}$$

The individual noise terms are given as follows (all normalized to the LO power P_{LO}, which appears in the numerator and in the denominator of the SNR and consequently cancels out).

1) LO Shot Noise:

$$P_{\mathrm{SH}} = 2hfR_S \quad \text{(A3)}$$

where h is Planck's constant and f is the optical frequency. We note that other shot noise terms exist, but these are negligible in a well-designed coherent receiver with $P_{\mathrm{LO}} \gg P_{\mathrm{ch}}$ considered for the analyses presented in this article.

2) LO-ASE Beat Noise:

1) For ideal *distributed* amplification, this noise term is

$$P_{\mathrm{ASE}} = 2hfR_S\alpha L \quad \text{(A4)}$$

with R_S being the symbol rate (cf. Section III–A) and where we set the factor K_T equal to 1 (cf. Section IV-3). We note that, also, other beat noise terms exist, but these, too, are negligible in a well-designed coherent receiver with balanced detection considered for the analyses presented in this article.

2) For *lumped* amplification, this noise term is [267]

$$P_{\mathrm{ASE}} = 2n_{\mathrm{sp}}N_s hfR_S\,(G-1) \quad \text{(A5)}$$

with n_{sp} being the spontaneous emission factor, $N_s = L/L_S$ being the span number, $G = e^{\alpha L_s}$ being the amplifier gain compensating for the propagation loss of the span preceding the amplifier, and L_S being the span length. The amplifier's noise figure F can be obtained from the spontaneous emission factor as

$$F = 2n_{\mathrm{sp}}\left(1 - \frac{1}{G}\right) + \frac{1}{G}. \quad \text{(A6)}$$

[13]The conversion from the more typically used fiber loss unit of dB/km is $\alpha_{\mathrm{Neper/km}} = (\ln(10)/10)\alpha_{\mathrm{dB/km}}$.

Under the assumption that the amplifier is operated with sufficient gain, i.e., $G \gg 1$, $n_{\mathrm{sp}} = 1$ corresponds to a noise figure of ∼3 dB, and $n_{\mathrm{sp}} = 1.6$ corresponds to a noise figure of ∼5 dB.

3) Transponder Noise:

$$P_{\mathrm{TRX}} = \kappa P_{\mathrm{ch}} \quad \text{(A7)}$$

with κ being a constant determining the transponder noise floor.

4) NLI Noise:

$$P_{\mathrm{NLIN}} = \chi {P_{\mathrm{ch}}}^3. \quad \text{(A8)}$$

While several models for NLIN exist in the literature [61], [62], [63], [64], [65], [66], we employ the GN model [87], [88], which provides analytical formulas for the NLI coefficient χ as follows.

1) For ideal distributed amplification [91, eq. (44)]

$$\chi = \frac{16}{27\pi}\frac{\gamma^2 L}{|\beta_2|\,{R_S}^2}\mathrm{asinh}\left(\frac{\pi^2}{3}\,|\beta_2|\,LB^2\right) \quad \text{(A9)}$$

with the group velocity dispersion parameter $\beta_2 = -D\lambda^2/(2\pi c)$, λ being the wavelength, and c being the speed of light in a vacuum.

2) For lumped amplification [91, eq. (16)]:

$$\chi = \frac{8}{27\pi}\frac{\gamma^2}{|\beta_2|\,{R_S}^2}N_s^{1+\epsilon}\frac{\left(1-e^{-\alpha L_s}\right)^2}{\alpha} \times \mathrm{asinh}\left(\frac{\pi^2}{2}\frac{1}{\alpha}\,|\beta_2|\,B^2\right) \quad \text{(A10)}$$

where ϵ is an adjustment factor that accounts for the degree of coherence between subsequent amplification spans. ϵ depends on the system bandwidth B and is typically on the order of or smaller than 0.1 for $B >$ 0.2 THz [87], [268]. Since (A10) is not reliable in conjunction with scenarios where the span loss (αL_s) drops below ∼10 dB (which may be the case for either short span lengths or ultralow propagation losses as predicted for HCFs), (A10) was recently extended in [269] and heuristically simplified in [84] to

$$\chi = \varepsilon_1 \mathrm{asinh}\,(\varepsilon_2) \quad \text{(A11)}$$

with

$$\varepsilon_1 = \frac{8}{27\pi}\frac{\gamma^2}{|\beta_2|\,{R_S}^2}\frac{N_s}{\alpha}\,\frac{\left(1-e^{-\alpha L_s}\right)^3}{1-e^{-\alpha L_s}-\alpha L_s e^{-\alpha L_s}}$$

and

$$\varepsilon_2 = \frac{2\pi^2}{3}\,|\beta_2|\,\frac{N_s}{\alpha}\,\frac{1-e^{-\alpha L_s}-\alpha L_s e^{-\alpha L_s}}{1-e^{-\alpha L_s}}B^2.$$

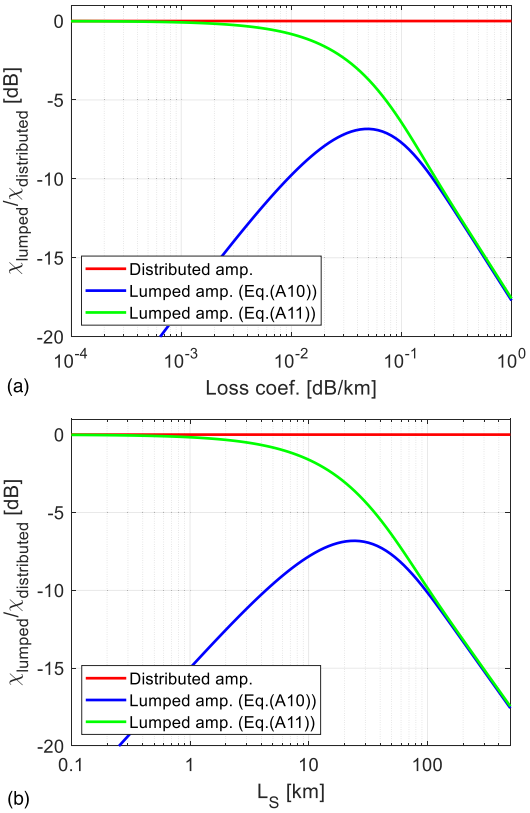

Fig. 26. *Normalized NLI coefficient for $N_S = 1$ as a function of (a) loss coefficient α at fixed $L_S = 100$ km and (b) span length L_S at fixed $\alpha = 0.2$ dB/km.*

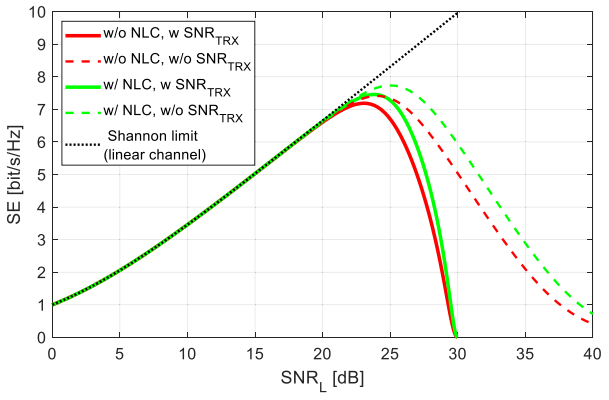

Fig. 27. *SE as a function of SNR_L before (in red) and after (in green) NLC for $L = 1000$ km with (solid lines) and without (dashed lines) an SNR_{TRX} of 30 dB.*

In Fig. 26, we compare (A11) with (A10) for $N_s = 1$ and $B = 4.4$ THz (C-band) as a function of loss and span length. The NLI coefficient χ given by (A11) converges to that of (A9) when the span loss (i.e., either loss coefficient α or span length L_S) approaches zero. It also converges to that of (A10) when the span loss exceeds ∼10 dB. For $N_s > 1$, nearly identical curves are obtained if ϵ in (A10) is set to ∼0.08. Note that (A11) does not require the adjustment factor ϵ. We conclude, therefore, that (A11) reliably predicts χ for a wide range of parameter choices.

Next, by taking the first derivative of (A2a) with respect to P_{ch}, one finds the channel's optimum launch power and optimum SNR_{NL} as

$$P_{\text{ch}}^{\text{opt}} = \sqrt[3]{\frac{P_{\text{SH}} + P_{\text{ASE}}}{2\chi}} \tag{A12}$$

and

$$\text{SNR}_{\text{NL}}^{\text{opt}} = \frac{1}{\kappa + \frac{3}{2}\sqrt[3]{2\chi\left(P_{\text{SH}} + P_{\text{ASE}}\right)^2}}. \tag{A13}$$

Note that the optimum launch power is independent of any transponder noise floor or linear crosstalk (i.e., independent of SNR_0). The curves shown in Fig. 2(b) are obtained by plotting the SE given by the SNR of the nonlinear channel (SNR_{NL}) as a function of the SNR of the linear channel (SNR_{L}) for a range of launch power values.

The above equations do not consider NLC at the receiver. Modern coherent receivers, however, often perform digital single-channel NLC on the spectral channel of interest. NLC can be taken into account analytically [88] by subtracting P_{NLIN} of a single spectral channel ($P_{\text{NLIN}}^{\text{SC}}$), obtained by setting $B = B_{\text{ch}}$ in the above equations, from that of the total P_{NLIN} as follows:

$$P_{\text{NLIN}}^{\text{NLC}} = P_{\text{NLIN}} - P_{\text{NLIN}}^{\text{SC}} \tag{A14}$$

and substituting P_{NLIN} in (A2) with $P_{\text{NLIN}}^{\text{NLC}}$. To give an example, an improvement in SE of ∼0.3 bit/s/Hz can be achieved with single-channel NLC for the case of distributed amplification over 1000 km of standard SMF with $R_S = 100$ GHz, $B = 4.4$ THz, and $\kappa = 10^{-3}$, as shown in Fig. 27. Details on the system benefits of NLC are summarized, e.g., in [104] and [105].

Acknowledgment

The authors gratefully acknowledge insightful discussions with respect to the Gaussian noise (GN) model with P. Poggiolini, G. Bosco, and M. Ranjbar Zefreh (Politecnico di Torino) and comments on the manuscript from C. Antonelli (University of L'Aquila).

REFERENCES

[1] E. B. Desurvire, "Capacity demand and technology challenges for lightwave systems in the next two decades," *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4697–4709, Dec. 2006.
[2] R. W. Tkach, "Scaling optical communications for the next decade and beyond," *Bell Labs Tech. J.*, vol. 14, no. 4, pp. 3–9, 2010.
[3] 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.
[4] *CRU: The Global Optical Fiber and Cable Market*. Accessed: Oct. 2022. [Online]. Available: https://www.tticables.com/news/cru-the-global-optical-fiber-and-cable-market-is-in-a-recovery-period-and-is-expected-to-grow-by-7-in-2021
[5] R.-J. Essiambre, G. J. Foschini, G. Kramer, and P. J. Winzer, "Capacity limits of information transport in fiber-optic networks," *Phys. Rev. Lett.*, vol. 101, no. 16, Oct. 2008, Art. no. 163901.
[6] A. Chraplyvy, "Plenary paper: The coming capacity crunch," in *Proc. 35th Eur. Conf. Opt. Commun. (ECOC)*, 2009, p. 1.
[7] R. 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, 2010.
[8] *GEN90P Coherent Trials and Deployments*. Accessed: Oct. 2022. [Online]. Available: https://cignal.ai/2021/06/5th-gen-coherent-trials-and-deployments/
[9] P. J. Winzer, "Would scaling to extreme ultraviolet or soft X-ray communications resolve the capacity crunch?" *J. Lightw. Technol.*, vol. 36, no. 24, pp. 5786–5793, Dec. 15, 2018.
[10] B. J. Puttnam, R. S. Luís, G. Rademacher, M. Mendez-Astudillio, Y. Awaji, and H. Furukawa, "S, C- and L-band transmission over a 157 nm bandwidth using doped fiber and distributed Raman amplification," *Opt. Exp.*, vol. 39, no. 6, pp. 10011–10018, 2022.
[11] D. Soma et al., "10.16 peta-bit/s dense SDM/WDM transmission over low-DMD 6-mode 19-core fibre across C+L band," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2017, Art. no. Th.PDP.A1.
[12] G. Rademacher et al., "10.66 peta-bit/s transmission over a 38-core-three-mode fiber," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2020, Art. no. Th3H.1.
[13] *Cisco/Acacia Announces 1.2Tbps DSP and Coherent Module*. Accessed: Oct. 2022. [Online]. Available: https://cignal.ai/2021/09/cisco-acacia-announces-1-2tbps-dsp-and-coherent-module/
[14] *Dunant Submarine Cable System*. Accessed: Oct. 2022. [Online]. Available: https://www.subcom.com/documents/2021/Dunant_RFS_Final_3FEBRUARY2021.pdf
[15] S. Varughese, S. Edirisinghe, M. Stephens, B. Boyanov, and P. Mertz, "SDM enabled record field trial achieving 300+Tbps trans-atlantic transmission capacity," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2022, Art. no. M1F.2.
[16] *Grace Hopper*. Accessed: Oct. 2022. [Online]. Available: https://www.submarinenetworks.com/en/systems/trans-atlantic/grace-hopper
[17] *NEC to Build a Transatlantic Cable*. Accessed: Oct. 2022. [Online]. Available: https://www.nec.com/en/press/202110/global_20211008_01.html
[18] *HMN Tech Launches 32FP Petabit-Level Repeater Prototype*. Accessed: Oct. 2022. [Online]. Available: https://www.submarinenetworks.com/en/vendors/hmn-tech/hmn-tech-launches-32fp-petabit-level-repeater-prototype
[19] T. Izawa and N. Inagaki, "Materials and processes for fiber preform fabrication—Vapor-phase axial deposition," *Proc. IEEE*, vol. 68, no. 10, pp. 1184–1187, 1980.
[20] *Milestones: Invention and Development of Vapor-Phase Axial Deposition (VAD) Method for Mass Production of High-Quality Optical Fiber for Telecommunication*. Accessed: Oct. 2022. [Online]. Available: http://ieeemilestones.ethw.org/Milestones: Invention_and_development_of_vapor-phase_axial_deposition_(VAD)_method_for_mass_production_of_high-quality_optical_fiber_for_telecommunication
[21] T. Miya, Y. Terunuma, T. Hosaka, and T. Miyashita, "Ultimate low-loss single-mode fibre at 1.55 μm," *Electron. Lett.*, vol. 4, no. 15, pp. 106–108, 1979.
[22] K. C. Kao and G. A. Hockham, "Dielectric-fibre surface waveguides for optical frequencies," *Proc. Inst. Elect. Eng.*, vol. 113, no. 7, pp. 1151–1158, Jul. 1966.
[23] (1970). *Milestones: World's First Low-Loss Optical Fiber for Telecommunications*. Accessed: Oct. 2022. [Online]. Available: https://ethw.org/Milestones:World%27s_First_Low-Loss_Optical_Fiber_for_Telecommunications,_1970
[24] Y. Yamamoto, Y. Kawaguchi, and M. Hirano, "Low-loss and low-nonlinearity pure-silica-core fiber for C- and L-band broadband transmission," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 321–326, Jan. 15, 2016.
[25] Y. Tamura et al., "The first 0.14-dB/km loss optical fiber and its impact on submarine transmission," *J. Lightw. Technol.*, vol. 36, no. 1, pp. 44–49, Jan. 1, 2018.
[26] R. J. Mears, L. Reekie, I. M. Jauncey, and D. N. Payne, "Low-noise erbium-doped fibre amplifier operating at 1.54 μm," *Electron. Lett.*, vol. 23, no. 19, pp. 1026–1028, 1987.
[27] 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.
[28] T. Li, "The impact of optical amplifiers on long-distance lightwave telecommunications," *Proc. IEEE*, vol. 81, no. 11, pp. 1568–1579, Nov. 1993.
[29] A. R. Chraplyvy, A. H. Gnauck, R. W. Tkach, and R. M. Derosier, "8 $\times$ 10 Gb/s transmission through 280 km of dispersion-managed fiber," *IEEE Photon. Technol. Lett.*, vol. 5, no. 10, pp. 1233–1235, Oct. 1993.
[30] G. P. Agrawal, *Nonlinear Fiber Optics*, 6th ed. New York, NY, USA: Academic, 2019.
[31] D. A. Fishman, W. A. Thompson, and L. Vallone, "LambdaXtreme transport system: R&D of a high capacity system for low cost, ultra long haul DWDM transport," *Bell Labs Tech. J.*, vol. 11, no. 2, pp. 27–53, Aug. 2006.
[32] H. Kogelnik, R. M. Jopson, and L. E. Nelson, "Polarization-mode dispersion," in *Optical Fiber Telecommunications IV B*. New York, NY, USA: Academic, 2002, ch. 15.
[33] P. J. Winzer et al., "100-Gb/s DQPSK transmission: From laboratory experiments to field trials," *J. Lightw. Technol.*, vol. 26, no. 20, pp. 3388–3402, Oct. 15, 2008.
[34] A. Farbert, S. Langenbach, N. Stojanovic, C. Dorschky, T. Kupfer, and C. Schulien, "Performance of a 10.7-Gb/s receiver with digital equalizer using maximum likelihood sequence estimation," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2004, Art. no. Th4.1.5.
[35] D. McGhan, C. Laperle, A. Savehenko, C. Li, G. Mak, and M. O'Sullivan, "5120 km RZ-DPSK transmission over G652 fiber at 10 Gb/s with no optical dispersion compensation," in *Proc. OFC/NFOEC Tech. Opt. Fiber Commun. Conf.*, 2005, Art. no. PDP27.
[36] 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.
[37] R. Noe, "PLL-free synchronous QPSK polarization multiplex/diversity receiver concept with digital I&Q baseband processing," *IEEE Photon. Technol. Lett.*, vol. 17, no. 4, pp. 887–889, Apr. 2005.
[38] D. Ly-Gagnon, S. Tsukamoto, K. Katoh, and K. Kikuchi, "Coherent detection of optical quadrature phase-shift keying signals with carrier phase estimation," *J. Lightw. Technol.*, vol. 24, no. 1, pp. 12–21, Jan. 2006.
[39] S. J. Savory et al., "Digital equalization of 40 Gbit/s per wavelength transmission over 2480 km of standard fibre without optical dispersion compensation," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2006, Art. no. Th2.5.5.
[40] A. Leven, N. Kaneda, A. Klein, U.-V. Koc, and Y. Chen, "Real-time implementation of 4.4 Gbit/s QPSK intradyne receiver using field programmable gate array," *Electron. Lett.*, vol. 42, no. 24, pp. 1421–1422, Nov. 2006.
[41] C. R. S. Fludger, T. Duthel, T. Wuth, and C. Schulien, "Uncompensated transmission of 86 Gbit/s polarization multiplexed RZ-QPSK over 100 km of NDSF employing coherent equalization," in *Proc. Eur. Conf. Optical Commun. (ECOC)*, 2006, Art. no. Th4.3.3.
[42] G. Charlet, N. Maaref, J. Renaudier, H. Mardoyan, P. Tran, and S. Bigo, "Transmission of 40 Gb/s QPSK with coherent detection over ultra-long distance improved by nonlinearity mitigation," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2006, Art. no. Th4.3.4.
[43] H. Sun, K.-T. Wu, and K. Roberts, "Real-time measurements of a 40 Gb/s coherent system," *Opt. Exp.*, vol. 16, no. 2, pp. 873–879, Jan. 2008.
[44] F. Derr, "Coherent optical QPSK intradyne system: Concept and digital receiver realization," *J. Lightw. Technol.*, vol. 10, no. 9, pp. 1290–1296, Sep. 1992.
[45] R. A. Linke and A. H. Gnauck, "High-capacity coherent lightwave systems," *J. Lightw. Technol.*, vol. 6, no. 11, pp. 1750–1769, Nov. 1988.
[46] J. D. Downie, "Maximum capacities in submarine cables with fixed power constraints for C-band, C+L-band, and multicore fiber systems," *J. Lightw. Technol.*, vol. 36, no. 18, pp. 4025–4032, Sep. 15, 2018.
[47] 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.
[48] P. J. Winzer, "Transmission system capacity scaling through space-division multiplexing: A techno-economic perspective," in *Optical Fiber Telecommunications VII*. New York, NY, USA: Academic, 2020, ch. 8.
[49] B. G. Lee et al., "120-Gb/s 100-m transmission in a single multicore multimode fiber containing six cores interfaced with a matching VCSEL array," in *Proc. IEEE Photon. Soc. Summer Topicals*, Jul. 2010, Art. no. TuD4.4.
[50] B. Zhu et al., "Seven-core multicore fiber transmissions for passive optical network," *Opt. Exp.*, vol. 18, no. 11, pp. 11117–11122, 2010.
[51] J. Sakaguchi et al., "109-Tb/s (7$\times$97$\times$172-Gb/s SDM/WDM/PDM) QPSK transmission through 16.8-km homogeneous multi-core fiber," in *Proc. Opt. Fiber Commun. Conf., Nat. Fiber Optic Eng. Conf.*, 2011, Art. no. PDPB6.
[52] B. Zhu et al., "Space-, wavelength-, polarization-division multiplexed transmission of 56-Tb/s over a 76.8-km seven-core fiber," in *Proc. Opt. Fiber Commun. Conf., Nat. Fiber Optic Eng. Conf.*, 2011, Art. no. PDPB7.
[53] S. Berdagué and P. Facq, "Mode division multiplexing in optical fibers," *Appl. Opt.*, vol. 21, no. 11, pp. 1950–1955, 1982.
[54] H. R. Stuart, "Dispersive multiplexing in multimode optical fiber," *Science*, vol. 289, pp. 281–283, Jul. 2000.
[55] R. Ryf et al., "Space-division multiplexing over 10 km of three-mode fiber using coherent 6$\times$6 MIMO processing," in *Proc. Opt. Fiber Commun. Conf., Nat. Fiber Optic Eng. Conf.*, 2011, Art. no. PDPB10.
[56] R. Ryf et al., "Record transmission performance over 4-core coupled-core fiber," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2017.
[57] C. Antonelli et al., "Towards space-division multiplexed transmission: The world-wide first and unique testbed with deployed special fibers for the photonics community," *IEEE Photon. Soc.*, vol. 34, no. 2, pp. 5–10, Jun. 2020.
[58] R. S. Luís et al., "Evaluation of dynamic skew on spooled and deployed multicore fibers using O-band signals," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2020, Art. no. T4J.4.
[59] M. Mazur et al., "Real-time MIMO transmission over field-deployed coupled-core multi-core fibers," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2022, Art. no. Th4B.8.
[60] T. Matsui, P. L. Pondillo, and K. Nakajima, "Weakly coupled multicore fiber technology, deployment, and systems," *Proc. IEEE*, vol. 110, no. 11, pp. 1772–1785, Nov. 2022, doi: 10.1109/JPROC.2022.3202812.
[61] A. Splett, C. Kurzke, and K. Petermann, "Ultimate transmission capacity of amplified optical fiber communication systems taking into account fiber nonlinearities," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, vol. 2, 1993, pp. 41–44.
[62] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fibre communications," *Nature*, vol. 411, pp. 1027–1030, Jun. 2001.
[63] 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, pp. 1095–1101, Jul. 2002.
[64] 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.
[65] R. Dar, M. Shtaif, 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.
[66] 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.
[67] M. Shtaif, C. Antonelli, A. Mecozzi, and X. Chen, "Challenges in estimating the information capacity of the fiber-optic channel," *Proc. IEEE*, vol. 110, no. 11, pp. 1655–1678, Nov. 2022, doi: 10.1109/JPROC.2022.3197188.
[68] J. Cho and P. J. Winzer, "Probabilistic constellation shaping for optical fiber communications," *J. Lightw. Technol.*, vol. 37, no. 6, pp. 1590–1607, Mar. 15, 2019.
[69] D. Gloge, "Weakly guiding fibers," *Appl. Opt.*, vol. 10, no. 10, pp. 2252–2258, 1971.
[70] G. Böcherer, P. Schulte, and F. Steiner, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4651–4665, Dec. 2015.
[71] T. Fehenberger, G. Böcherer, A. Alvarado, and N. Hanik, "LDPC coded modulation with probabilistic shaping for optical fiber systems," in *Proc. Opt. Fiber Commun. Conf.*, 2015, Art. no. Th.2A.23.
[72] F. Buchali, G. Bocherer, W. Idler, L. Schmalen, P. Schulte, and F. Steiner, "Experimental demonstration of capacity increase and rate-adaptation by probabilistically shaped 64-QAM," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2015, Art. no. PDP3.4.
[73] 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.
[74] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.
[75] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.
[76] D. Forney, *Principles of Digital Communication II*. Cambridge, MA, USA: MIT, 2005.
[77] P. J. Winzer, "High-spectral-efficiency optical modulation formats," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3824–3825, Aug. 23, 2012.
[78] X. Chen et al., "All-electronic 100-GHz bandwidth digital-to-analog converter generating PAM signals up to 190 GBaud," *J. Lightw. Technol.*, vol. 35, no. 3, pp. 411–417, Feb. 1, 2017.
[79] H. Mardoyan et al., "222-GBaud on-off keying transmitter using ultra-high-speed 2:1-selector and plasmonic modulator on silicon photonics," in *Proc. 45th Eur. Conf. Opt. Commun. (ECOC)*, 2019, Art. no. PD2.3.
[80] H. Mardoyan et al., "First 260-GBd single-carrier coherent transmission over 100 km distance based on novel arbitrary waveform generator and thin-film lithium niobate I/Q modulator," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2022, Art. no. Th3C.2.
[81] L. Nelson and S. L. Woodward, "Demultiplexing via a dual-polarization coherent receiver," in *Proc. IEEE Summer Topicals Meeting*, 2010, Art. no. WA1.2.
[82] A. Ghazisaeidi et al., "Submarine transmission systems using digital nonlinear compensation and adaptive rate forward error correction," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1886–1895, Apr. 15, 2016.
[83] L. Galdino et al., "On the limits of digital back-propagation in the presence of transceiver noise," *Opt. Exp.*, vol. 25, no. 4, pp. 4564–4578, 2017.
[84] W. Klaus and P. J. Winzer, "Hollow-core fiber capacities with receiver noise limitations," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2022, Art. no. M2C.4.
[85] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Exp.*, vol. 21, no. 22, pp. 25685–25699, 2013.
[86] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Accumulation of nonlinear interference noise in fiber-optic systems," *Opt. Exp.*, vol. 22, no. 12, pp. 14199–14211, Jun. 2014.
[87] 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.
[88] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 15, 2014.
[89] M. Nakazawa, M. Yoshida, M. Terayama, S. Okamoto, K. Kasai, and T. Hirooka, "Observation of guided acoustic-wave Brillouin scattering noise and its compensation in digital coherent optical fiber transmission systems," *Opt. Exp.*, vol. 26, no. 7, pp. 9165–9181, 2018.
[90] M. A. Bolshtyansky et al., "Impact of spontaneous guided acoustic-wave Brillouin scattering on long-haul transmission," in *Proc. Opt. Fiber Commun. Conf.*, 2018, Art. no. M4B.3.
[91] E. R. Hartling et al., "Design, acceptance and capacity of subsea open cables," *J. Lightw. Technol.*, vol. 39, no. 3, pp. 742–756, Feb. 1, 2021.
[92] M. S. Faruk and K. Kikuchi, "Compensation for in-phase/quadrature imbalance in coherent-receiver front end for optical quadrature amplitude modulation," *IEEE Photon. J.*, vol. 5, no. 2, Apr. 2013, Art. no. 7800110.
[93] X. Chen, J. Cho, A. Adamiecki, and P. Winzer, "16384-QAM transmission at 10 BBd over 25-km SSMF using polarization-multiplexed probabilistic constellation shaping," in *Proc. 45th Eur. Conf. Opt. Commun. (ECOC)*, 2019, Art. no. PD3.3.
[94] T. Zami et al., "Simple self-optimization of WDM networks based on probabilistic constellation shaping," *J. Opt. Commun. Netw.*, vol. 12, no. 1, p. A82, 2020.
[95] F. Buchali et al., "128 GSa/s SiGe DAC implementation enabling 1.52 Tb/s single carrier transmission," *J. Lightw. Technol.*, vol. 39, no. 3, pp. 763–770, Feb. 1, 2021.
[96] 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.*, 2018, Art. no. Th4C.5.
[97] Y. Wakayama et al., "2048-QAM transmission at 15 GBd over 100 km using geometric constellation shaping," *Opt. Exp.*, vol. 28, no. 12, pp. 18743–18759, 2021.
[98] S. Okamoto, M. Terayama, M. Yoshida, K. Kasai, T. Hirooka, and M. Nakazawa, "Experimental and numerical comparison of probabilistically shaped 4096 QAM and a uniformly shaped 1024 QAM in all-Raman amplified 160 km transmission," *Opt. Exp.*, vol. 26, no. 3, pp. 3535–3543, Feb. 2018.
[99] T. Omiya, M. Yoshida, and M. Nakazawa, "400 Gbit/s 256 QAM-OFDM transmission over 720 km with a 14 bit/s/Hz spectral efficiency by using high-resolution FDE," *Opt. Exp.*, vol. 21, no. 3, pp. 2632–2641, 2013.
[100] 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, Art. no. Th5C.2.
[101] W. Shieh and K. P. Ho, "Equalization-enhanced phase noise for coherent detection systems using electronic digital signal processing," *Opt. Exp.*, vol. 16, no. 20, pp. 15718–15727, 2008.
[102] C. Xie, "Local oscillator phase noise induced penalties in optical coherent detection systems using electronic chromatic dispersion compensation," in *Proc. Opt. Fiber Commun. Conf. Nat. Fiber Optic Engineers Conf.*, 2009, Art. no. OMT4.
[103] P. J. Winzer, "Energy-efficient optical transport capacity scaling through spatial multiplexing," *IEEE Photon. Technol. Lett.*, vol. 23, no. 13, pp. 851–853, Jul. 1, 2011.
[104] J. C. Cartledge, F. P. Guiomar, F. R. Kschischang, G. Liga, and M. P. Yankov, "Digital signal processing for fiber nonlinearities," *Opt. Exp.*, vol. 25, no. 3, pp. 1916–1936, 2017.
[105] R. Dar and P. J. Winzer, "Nonlinear interference mitigation: Methods and potential gain," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 903–930, Feb. 15, 2017.
[106] R. Kashyap, "The fiber fuse—From a curious effect to a critical issue: A 25th year retrospective," *Opt. Exp.*, vol. 21, no. 5, pp. 6422–6441, 2013.
[107] T. Morioka, "New generation optical infrastructure technologies: 'EXAT initiative' towards 2020 and beyond," in *Proc. 14th Optoelectron. Commun. Conf.*, 2009, pp. 1–2, doi: 10.1109/OECC.2009.5213198.
[108] A. M. Rocha, F. Domingues, M. Facao, and P. S. Andre, "Threshold power of fiber fuse effect for different types of optical fiber," in *Proc. 13th Int. Conf. Transparent Opt. Netw.*, Jun. 2011, p. 13.
[109] M. Ionescu et al., "91 nm C+L hybrid distributed Raman–erbium-doped fibre amplifier for high capacity subsea transmission," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2018, Art. no. Mo4G.2.
[110] 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," *J. Lightw. Technol.*, vol. 36, no. 1, pp. 114–121, Mar. 19, 2018.
[111] 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.
[112] V. Bajaj, F. Buchali, M. Chagnon, S. Wahls, and V. Aref, "54.5 Tb/s WDM transmission over field deployed fiber enabled by neural network-based digital pre-distortion," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2021, Art. no. M5F.2.
[113] D. Semrau, R. Killey, and P. Bayvel, "Achievable rate degradation of ultra-wideband coherent fiber communication systems due to stimulated Raman scattering," *Opt. Exp.* , vol. 25, no. 12, pp. 13024–13034, 2017.
[114] J. Bromage, "Raman amplification for fiber communication systems," *J. Lightw. Technol.*, vol. 22, no. 1, pp. 79–93, 2004.
[115] A. H. Gnauck, R. M. Jopson, and P. J. Winzer, "Demonstration of counter-propagating Raman pump placed near signal-channel wavelengths," *IEEE Photon. Technol. Lett.*, vol. 29, no. 1, pp. 154–157, Jan. 1, 2017.
[116] T. D. Bradley et al., "Towards low loss hollow core optical fibers," *Proc. SPIE*, vol. 11713, pp. 117130A-1–117130A-8, Mar. 2021.
[117] F. Poletti, "Nested antiresonant nodeless hollow core fiber," *Opt. Exp.*, vol. 22, no. 20, pp. 23807–23828, 2014.
[118] Y. Hong et al., "Hollow-core NANF for high-speed short-reach transmission in the S+C+L-bands," *J. Lightw. Technol.*, vol. 39, no. 19, pp. 6167–6174, Oct. 2021.
[119] G. T. Jasion et al., "0.174 dB/km hollow core double nested antiresonant nodeless fiber (DNANF)," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2022, Art. no. Th4C.7.
[120] J. Bromage, P. J. Winzer, and R.-J. Essiambre, "Multiple path interference and its impact on system design," in *Raman Amplifiers for Telecommunications*. New York, NY, USA: Springer-Verlag, 2004, ch.15.
[121] S. Namiki and Y. Emori, "Ultrabroad-band Raman amplifiers pumped and gain-equalized by wavelength-division-multiplexed high-power laser diodes," *IEEE J. Sel. Topics Quantum Electron.*, vol. 7, no. 1, pp. 3–16, Jan. 2001.
[122] 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.
[123] A. Arnould et al., "Experimental characterization of nonlinear distortions of semiconductor optical amplifiers in the WDM regime," *J. Lightw. Technol.*, vol. 38, no. 2, pp. 509–513, Jan. 15, 2020.
[124] P. A. Andrekson and M. Karlsson, "Fiber-based phase-sensitive optical amplifiers and their applications," *Adv. Opt. Photon.*, vol. 12, no. 2,

pp. 367–428, 2020.
[125] S. L. I. Olsson, H. Eliasson, E. Astra, M. Karlsson, and P. A. Andrekson, "Long-haul optical transmission link using low-noise phase-sensitive amplifiers," *Nature Commun.*, vol. 9, no. 1, p. 2513, Dec. 2018.
[126] H. Sakr et al., "Hollow core NANFs with five nested tubes and record low loss at 850, 1060, 1300 and 1625 nm," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2021, Art. no. F3A.4.
[127] F. Poletti and P. Poggiolini, "Potential system impact of low-loss antiresonant hollow core fibers," in *Proc. 45th Eur. Conf. Opt. Commun. (ECOC)*, 2019, Art. no. P51.
[128] F. Hamaoka et al., "150.3-Tb/s ultra-wideband (S, C, and L bands) single-mode fibre transmission over 40-km using >519 Gb/s/A PDM-128QAM signals," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2018, pp. 1–3.
[129] 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.
[130] 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.
[131] S. Chandrasekhar, X. Liu, B. Zhu, and D. W. Peckham, "Transmission of a 1.2-Tb/s 24-carrier no-guard-interval CO-OFDM superchannel over 7200-km of ultra-large-area fiber," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2009, Art. no. PD2.6.
[132] G. Bosco, V. Curri, A. Carena, P. Poggiolini, and F. Forghieri, "On the performance of Nyquist-WDM terabit superchannels based on PM-BPSK, PM-QPSK, PM-8 QAM or PM-16 QAM subcarriers," *J. Lightw. Technol.*, vol. 29, no. 1, pp. 53–61, Jan. 1, 2011.
[133] R. S. Tucker, "Green optical communications—Part I: Energy limitations in transport," *IEEE J. Sel. Topics Quantum Electron.*, vol. 17, no. 2, pp. 245–260, Mar. 2011.
[134] A. Pilipetskii, "High capacity submarine transmission systems," in *Proc. Opt. Fiber Commun. Conf.*, 2015, Art. no. W3G.5.
[135] O. V. Sinkin et al., "SDM for power-efficient undersea transmission," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 361–371, Jan. 15, 2018.
[136] J.-C. Antona, A. C. Meseguer, and V. Letellier, "Transmission systems with constant output power amplifiers at low SNR values: A generalized droop model," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2019, Art. no. M1J.6.
[137] A. Bononi, J.-C. Antona, M. Lonardi, A. Carbo-Meseguer, and P. Serena, "The generalized droop formula for low signal to noise ratio optical links," *J. Lightw. Technol.*, vol. 38, no. 8, pp. 2201–2213, Apr. 15, 2020.
[138] C. R. Giles and E. Desurvire, "Propagation of signal and noise in concatenated erbium-doped fiber optical amplifiers," *J. Lightw. Technol.*, vol. 9, no. 2, pp. 147–154, Feb. 1991.
[139] J. D. Downie, X. Liang, V. Ivanov, P. Sterlingov, and N. Kaliteevskiy, "SNR model for generalized droop with constant output power amplifier systems and experimental measurements," *J. Lightw. Technol.*, vol. 38, no. 12, pp. 3214–3220, Jun. 15, 2020.
[140] J.-X. Cai et al., "9 Tb/s transmission using 29 mW optical pump power per EDFA with 1.24 Tb/s/W optical power efficiency over 15,050 km," *J. Lightw. Technol.*, vol. 40, no. 6, pp. 1650–1657, Mar. 15, 2022.
[141] A. N. Pilipetskii and G. Mohs, "Technology evolution and capacity growth in undersea cables," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2020, Art. no. W4E.2.
[142] M. Spalding, M. Bolshtyansky, and O. Sinkin, "Vision for next generation undersea optical fibers and cable designs," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Dec. 2020, Art. no. Tu2E-1.
[143] T. Sasaki, F. Sato, P. Pondillo, A. Bertaina, and P. Weimann, "Ultrahigh fiber count and high-density cables, deployments, and systems," *Proc. IEEE*, vol. 110, no. 11, pp. 1760–1771, Nov. 2022, doi: 10.1109/JPROC.2022.3199645.
[144] T. Hayashi et al., "Randomly-coupled multi-core fiber technology," *Proc. IEEE*, vol. 110, no. 11, pp. 1786–1803, Nov. 2022, doi: 10.1109/JPROC.2022.3182049.
[145] P. Sillard et al., "Few-mode fiber technology, deployments, and systems," *Proc. IEEE*, vol. 110, no. 11, pp. 1804–1820, Nov. 2022, doi: 10.1109/JPROC.2022.3207012.
[146] N. Fontaine, J. Carpenter, S. Gross, R. A. Correa, S. L. Saval, and Y. Jung, "Photonic lanterns, 3D waveguides, multiplane light conversion and other components that enable space division multiplexing," *Proc. IEEE*, vol. 110, no. 11, pp. 1821–1834, Nov. 2022, doi: 10.1109/JPROC.2022.3207046.
[147] Y. Tsujimoto, N. Yamashita, A. Namazue, and K. Osato, "3000-fiber optical cable using 200 μm fiber," *Fujikura Techn. Rev.*, no. 50, pp. 1–4, Dec. 2020. [Online]. Available: https://www.fujikura.co.jp/eng/rd/gihou/2063698_11754.html and https://www.fujikura.co.jp/eng/rd/gihou/backnumber/pages/__icsFiles/afieldfile/2021/05/12/50e_01.pdf
[148] *Ultra-High Fiber Count (UHFC) 6912-Fiber Rollable Ribbon Cable and Ribbon Fiber Mass Fusion Splicer S124M12 and Related Tools*. Accessed: Oct. 2022. [Online]. Available: https://www.furukawa.co.jp/en/rd/review/fr051/fr51_10.pdf
[149] K. Saitoh and S. Matsuo, "Multicore fiber technology," *J. Lightw. Technol.*, vol. 34, no. 1, pp. 55–66, Jan. 1, 2015.
[150] K. Takenaga et al., "Reduction of crosstalk by trench-assisted multi-core fiber," in *Proc. Opt. Fiber Commun. Conf., Nat. Fiber Optic Eng. Conf.*, 2011, p. OWJ4.
[151] T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, and E. Sasaoka, "Ultra-low-crosstalk multi-core fiber feasible to ultra-long-haul transmission," in *Proc. Opt. Fiber Commun. Conf., Nat. Fiber Optic Eng. Conf.*, 2011, Art. no. PDPC2.
[152] L.-A. de Montmorillon et al., "Next generation SMF with reduced bend sensitivity for FTTH networks," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2006, Art. no. Mo3.3.2.
[153] M. Koshiba, K. Saitoh, and Y. Kokubun, "Heterogeneous multi-core fibers: Proposal and design principle," *IEICE Electron. Exp.*, vol. 6, no. 2, pp. 98–103, Jan. 2009.
[154] T. Hayashi, "Multi-core fibers for space division multiplexing," in *Handbook of Optical Fibers*. Cham, Switzerland: Springer, 2019, ch.3.
[155] K. Saitoh, "Multi-core fiber technology for SDM: Coupling mechanisms and design," *J. Lightw. Technol.*, vol. 40, no. 5, pp. 1527–1543, Mar. 1, 2022.
[156] C. Antonelli, G. Riccardi, T. Hayashi, and A. Mecozzi, "Role of polarization-mode coupling in the crosstalk between cores of weakly coupled multi-core fibers," *Opt. Exp.*, vol. 28, no. 9, pp. 12847–12861, 2020.
[157] C. Antonelli, T. Hayashi, and A. Mecozzi, "Random polarization-mode coupling explains inter-core crosstalk in uncoupled multi-core fibers," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Dec. 2020, Art. no. Th1A-1.
[158] T. Hayashi, T. Sasaki, and E. Sasaoka, "Behavior of inter-core crosstalk as a noise and its effect on Q-factor in multi-core fiber," *IEICE Trans. Commun.*, vol. 97, no. 5, pp. 936–944, 2014.
[159] R. S. Luís et al., "Time and modulation frequency dependence of crosstalk in homogeneous multi-core fibers," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 441–447, 2016.
[160] G. Rademacher, R. S. Luís, B. J. Puttnam, Y. Awaji, and N. Wada, "Crosstalk dynamics in multi-core fibers," *Opt. Exp.*, vol. 25, no. 10, pp. 12021–12028, 2017.
[161] *Characteristics of a Single-Mode Optical Fibre and Cable*, document ITU-T Recommendation G.652, ed. 9, 2016.
[162] K. Petermann, "Constraints for fundamental-mode spot size for broadband dispersion-compensated single-mode fibers," *Electron. Lett.*, vol. 19, pp. 712–714, Dec. 1983.
[163] *Characteristics of a Single-Mode Optical Fibre and Cable*, document ITU-T Recommendation G.650.1, ed. 8, 2020.
[164] L.-A. de Montmorillon, M. Bigot-Astruc, and P. Sillard, "Cutoff mechanisms in bend-insensitive single-mode fibers," in *Proc. Opt. Fiber Commun. Conf., Nat. Fiber Optic Eng. Conf.*, 2011, Art. no. OTuA1.
[165] F. Ye et al., "Wavelength-dependence of inter-core crosstalk in homogeneous multi-core fibers," *IEEE Photon. Technol. Lett.*, vol. 28, no. 1, pp. 27–30, Jan. 1, 2016.
[166] T. Matsui, Y. Yamada, Y. Sagae, and K. Nakajima, "Standard cladding diameter multi-core fiber technology," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2021, Art. no. Tu6B4.
[167] K. Takenaga et al., "A large effective area multi-core fiber with an optimized cladding thickness," *Opt. Exp.*, vol. 19, no. 26, p. B543, 2011.
[168] S. Matsuo et al., "Large-effective-area ten-core fiber with cladding diameter of about 200 μm," *Opt. Lett.*, vol. 36, no. 23, pp. 4626–4628, 2011.
[169] Y. Sasaki et al., "Crosstalk-managed heterogeneous single-mode 32-core fibre," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2016, Art. no. W2B2.
[170] 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, Art. no. Th5B.1.
[171] T. Ito, E. L. T. de Gabory, M. Arikawa, Y. Hashimoto, and K. Fukuchi, "Reduction of influence of inter-core cross-talk in MCF with bidirectional assignment between neighboring cores," in *Proc. Opt. Fiber Commun. Conf., Nat. Fiber Optic Eng. Conf.*, 2013, Art. no. OTh3K.2.
[172] K. Takenaga et al., "An investigation on crosstalk in multi-core fibers by introducing random fluctuation along longitudinal direction," *IEICE Trans. Commun.*, vol. E94-B, no. 2, pp. 409–416, 2011.
[173] M. Koshiba, K. Saitoh, K. Takenaga, and S. Matsuo, "Analytical expression of average power-coupling coefficients for estimating intercore crosstalk in multicore fibers," *IEEE Photon. J.*, vol. 4, no. 5, pp. 1987–1995, Oct. 2012.
[174] J. M. Gené and P. J. Winzer, "A universal specification for multicore fiber crosstalk," *IEEE Photon. Technol. Lett.*, vol. 31, no. 9, pp. 673–676, May 1, 2019.
[175] J. M. Gené, P. J. Winzer, H. Chen, R. Ryf, T. Hayashi, and T. Sasaki, "Towards broadly optimum multi-core fiber designs," in *Proc. 45th Eur. Conf. Opt. Commun. (ECOC)*, 2019, pp. 1–4.
[176] J. von Hoyningen-Huene, R. Ryf, and P. Winzer, "LCoS-based mode shaper for few-mode fiber," *Opt. Exp.*, vol. 21, no. 15, pp. 18097–18110, Jul. 2013.
[177] H. Kogelnik and P. J. Winzer, "Modal birefringence in weakly guiding fibers," *J. Lightw. Technol.*, vol. 30, no. 14, pp. 2240–2245, Jul. 2012.
[178] 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.
[179] A. Andrusier, M. Shtaif, 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. 2014.
[180] 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.
[181] S. O. Arik, K.-P. Ho, and J. M. Kahn, "Group delay management and multiinput multioutput signal processing in mode-division multiplexing systems," *J. Lightw. Technol.*, vol. 34, no. 11, pp. 2867–2880, Jun. 1, 2016.
[182] K.-P. Ho and J. M. Kahn, "Statistics of group delays in multimode fiber with strong mode coupling," *J. Lightw. Technol.*, vol. 29, no. 21, pp. 3119–3128,

Nov. 1, 2011.
[183] 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.
[184] C. Antonelli, M. Shtaif, 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.
[185] R. Ryf et al., "Coupled-core transmission over 7-core fiber," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2019, Art. no. Th4B.3.
[186] K. Choutagunta, S. O. Arik, K. P. Ho, and J. M. Kahn, "Characterizing mode-dependent loss and gain in multimode components," *J. Lightw. Technol.*, vol. 36, no. 18, pp. 3815–3823, Sep. 15, 2018.
[187] C. Antonelli, A. Mecozzi, M. Shtaif, and P. J. Winzer, "Random coupling between groups of degenerate fiber modes in mode multiplexed transmission," *Opt. Exp.*, vol. 21, no. 8, pp. 9484–9490, 2013.
[188] C. Koebele et al., "40km transmission of five mode division multiplexed data streams at 100 Gb/s with low MIMO-DSP complexity," in *Proc. 37th Eur. Conf. Exhib. Opt. Commun.*, 2011, pp. 1–3.
[189] M. Salsi, C. Koebele, G. Charlet, and S. Bigo, "Mode division multiplexed transmission with a weakly-coupled few-mode fiber," in *Proc. Opt. Fiber Commun. Conf.*, 2012, pp. 1–3.
[190] M. Zuo et al., "Long-haul intermodal-MIMO-free MDM transmission based on a weakly coupled multiple-ring-core few-mode fiber," *Opt. Exp.*, vol. 30, no. 4, pp. 5868–5878, 2022.
[191] G. Labroille et al., "30 Gbit/s transmission over 1 km of conventional multi-mode fiber using mode group multiplexing with OOK modulation and direct detection," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2015, pp. 1–3.
[192] K. Benyahya et al., "Multiterabit transmission over OM2 multimode fiber with wavelength and mode group multiplexing and direct detection," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 355–360, Jan. 15, 2018.
[193] K. Benyahya et al., "High-speed bi-directional transmission over multimode fiber link in IM/DD systems," *J. Lightw. Technol.*, vol. 36, no. 18, pp. 4174–4180, Sep. 15, 2018.
[194] K. Benyahya et al., "200 Gb/s transmission over 20km of FMF fiber using mode group multiplexing and direct detection," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2018, pp. 1–3.
[195] L. Shen et al., "MIMO-free 20-Gb/s×4×2 WDM-MDM transmission over 151.5-km single-span ultra low-crosstalk FMFs," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2018, Art. no. We1.E.6.
[196] S. Randel, A. Sierra, R. Ryf, and P. J. Winzer, "Crosstalk tolerance of spatially multiplexed MIMO systems," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2012, Art. no. P4.08.
[197] T. Sakamoto et al., "120 spatial channel few-mode multi-core fibre with relative core multiplicity factor exceeding 100," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2018, pp. 1–3.
[198] J. Sakaguchi et al., "109-Tb/s (7×97×172-Gb/s SDM/WDM/PDM) QPSK transmission through 16.8-km homogeneous multi-core fiber," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2011, Art. no. PDPB6.
[199] B. Zhu et al., "112-Tb/s space-division multiplexed DWDM transmission with 14-b/s/Hz aggregate spectral efficiency over a 76.8-km seven-core fiber," *Opt. Exp.*, vol. 19, no. 17, pp. 16665–16671, 2011.
[200] J. Sakaguchi et al., "19-core fiber transmission of 19×100×172-Gb/s SDM-WDM-PDM-QPSK signals at 305Tb/s," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2012, Art. no. PDP5C.1.
[201] H. Takara et al., "1.01-Pb/s (12 SDM/222 WDM/456 Gb/s) crosstalk-managed transmission with 91.4-b/s/Hz aggregate spectral efficiency," in *Proc. Eur. Conf. Exhib. Opt. Commun.*, 2012, Art. no. Th.3C.1.
[202] D. Qian et al., "1.05Pb/s transmission with 109b/s/Hz spectral efficiency using hybrid single- and few-mode cores," in *Proc. Frontiers in Optics (FiO)*, 2012, Art. no. FW6C.3.
[203] B. J. Puttnam et al., "2.15 Pb/s transmission using a 22 core homogeneous single-mode multi-core fiber and wideband optical comb," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2015, Art. no. PD3.1.
[204] D. Soma et al., "2.05 peta-bit/s super-nyquist-WDM SDM transmission using 9.8-km 6-mode 19-core fiber in full C band," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2015, Art. no. PD3.2.
[205] K. Igarashi et al., "1.03-Exabit/s km super-Nyquist-WDM transmission over 7,326-km seven-core fiber," in *Proc. 39th Eur. Conf. Exhib. Opt. Commun. (ECOC)*, 2013, pp. 1–3.
[206] A. Turukhin et al., "105.1 Tb/s power-efficient transmission over 14,350 km using a 12-core fiber," in *Proc. Opt. Fiber Commun. Conf.*, 2016, Art. no. Th4C.1.
[207] 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, Art. no. Th5B.1.
[208] 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, Art. no. Th4B.1.
[209] B. J. Puttnam, R. S. Luís, 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, Art. no. F3B.3.
[210] E. Ip et al., "88×3×112-Gb/s WDM transmission over 50 km of three-mode fiber with inline few-mode fiber amplifier," in *Proc. 37th Eur. Conf. Exhib. Opt. Commun.*, 2011, pp. 1–3.
[211] V. Sleiffer et al., "73.7 Tb/s (96×3×256 Gb/s) mode-division-multiplexed DP-16 QAM transmission with inline MM-EDFA," *Opt. Exp.*, vol. 20, no. 26, pp. B428–B438, 2012.
[212] R. Ryf et al., "10-mode mode-multiplexed transmission over 125-km single-span multimode fiber," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2015, pp. 1–3.
[213] S. Beppu et al., "402.7-Tb/s MDM-WDM transmission over weakly coupled 10-mode fiber using rate-adaptive PS-16 QAM signals," *J. Lightw. Technol.*, vol. 38, no. 10, pp. 2835–2841, May 15, 2020.
[214] G. Rademacher et al., "1.01 Peta-bit/s C+L-band transmission over a 15-mode fiber," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2020, pp. 1–4.
[215] G. Rademacher et al., "1.53 Peta-bit/s C-band transmission in a 55-mode fiber," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2022, Art. no. Th3C.3.
[216] E. Ip et al., "146×6×19 Gbaud wavelength-and mode-division multiplexed transmission over 10×50 km spans of few-mode fiber with a gain-equalized few-mode EDFA," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 790–797, 2014.
[217] G. Rademacher et al., "3500-km mode-multiplexed transmission through a three-mode graded-index few-mode fiber link," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2017, pp. 1–3.
[218] J. van Weerdenburg et al., "138 Tbit/s transmission over 650 km graded-index 6-mode fiber," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2017, pp. 1–3.
[219] G. Rademacher et al., "159 Tbit/s C+L band transmission over 1045 km 3-mode graded-index few-mode fiber," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2018, Art. no. Th4C.4.
[220] 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. Postdeadline Papers*, 2018, Art. no. Th4C.6.
[221] G. Rademacher et al., "High capacity transmission in a coupled-core three-core multi-core fiber," *J. Lightw. Technol.*, vol. 39, no. 3, pp. 757–762, Feb. 1, 2021.
[222] D. Soma et al., "50.47-Tbit/s standard cladding ultra-low-loss coupled 4-Core fiber transmission over 9,150 km," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2021, Art. no. W7D.3.
[223] [Online]. Available: https://www.jectec.or.jp/03cable/data/LCA333.PDF
[224] K. Saito, T. Sakamoto, T. Matsui, K. Nakajima, and T. Kurashima, "Side-view based angle alignment technique for multi-core fiber," in *Proc. Opt. Fiber Commun. Conf.*, 2016, Art. no. M3F.3.
[225] M. Takahashi, T. Fujii, R. Sugizaki, M. Tsukamoto, and Y. Arashitani, "Field usable fusion splicing technique for multicore fiber," in *Proc. IEEE Photon. Soc. Summer Topicals Meeting Ser. (SUM)*, 2020, pp. 1–2.
[226] K. Yoshida, A. Takahashi, T. Konuma, K. Yoshida, and K. Sasaki, "Fusion splicer for specialty optical fiber with advanced functions," *Fujikura Techn. Rev.*, vol. 41, pp. 10–13, Sep. 2021.
[227] R. Nagase, "Optical connectivities for multicore fiber," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2020, Art. no. Th3I.1.
[228] Y. Abe, K. Shikama, and S. Asakawa, "Multi-core fiber connector technology for low-loss physical-contact connection," *NTT Techn. Rev.*, vol. 15, no. 6, pp. 1–6, 2017.
[229] T. Morishima, K. Manabe, S. Toyokawa, T. Nakanishi, T. Sano, and T. Hayashi, "Simple-structure LC-type multi-core fiber connector with low insertion loss," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2020, Art. no. Th3I.2.
[230] T. Morishima et al., "MCF-enabled ultra-high-density 256-core MT connector and 96-core physical-contact MPO connector," in *Proc. Opt. Fiber Commun. Conf. Postdeadline Papers*, 2017, Art. no. Th5D.4.
[231] T. Hayashi, T. Nagashima, A. Inoue, H. Sakuma, T. Suganuma, and T. Hasegawa, "Uncoupled multi-core fiber design for practical bidirectional optical communications," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2022, Art. no. M1E.1.
[232] P. Sillard et al., "Low-differential-mode-group-delay 9-LP-mode fiber," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 425–430, Jan. 15, 2016.
[233] S. Warm and K. Petermann, "Splice loss requirements in multi-mode fiber mode-division-multiplex transmission links," *Opt. Exp.*, vol. 21, no. 1, pp. 519–532, Jan. 2013.
[234] T. Sakamoto et al., "Characteristic of splicing misalignment induced mode dependent loss for coupled multi-core fibre," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2017, pp. 1–3.
[235] W. Klaus et al., "Free-space coupling optics for multicore fibers," *IEEE Photon. Technol. Lett.*, vol. 24, no. 21, pp. 1902–1905, Nov. 1, 2012.
[236] Y. Tottori, H. Tsuboya, T. Kobayashi, and M. Watanabe, "Integrated optical connection module for 7-core multi-core fiber and 7 single mode fibers," in *Proc. IEEE Photon. Soc. Summer Topical Meeting Ser.*, Jul. 2013, Art. no. MC3.2.
[237] R. R. Thomson et al., "Ultrafast-laser inscription of a three dimensional fanout device for multicore fiber coupling applications," *Opt. Exp.*, vol. 15, no. 18, pp. 11691–11697, 2007.
[238] Y.-C. Ling, S. Yuan, and S. J. Ben Yoo, "Low-loss three-dimensional fan-in/fan-out devices for multi-core fiber integration," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2021, Art. no. Th1A.37.
[239] K. Watanabe, T. Saito, K. Imamura, and M. Shiino, "Development of fiber bundle type fan-out for multicore fiber," in *Proc. 17th Opto-Electronics Commun. Conf.*, Jul. 2012, pp. 475–476.
[240] V. I. Kopp, J. Park, M. Wlodawski, J. Singer, D. Neugroschl, and A. Z. Genack, "Pitch reducing optical fiber array and multicore fiber for space-division multiplexing," in *Proc. IEEE Photon. Soc. Summer Top. Meeting Ser.*, 2013, pp. 99–100.
[241] O. Shimakawa, M. Shiozaki, T. Sano, and A. Inoue, "Pluggable fan-out realizing physical-contact and low coupling loss for multi-core fiber," in *Proc. Opt. Fiber Commun. Conf., Nat. Fiber Optic Eng. Conf.*, 2013, Art. no. OM3I.2.

[242] W. Mohammed, M. Pitchumani, A. Mehta, and E. G. Johnson, "Selective excitation of the LP mode in step index fiber using a phase mask," *Opt. Eng.*, vol. 45, no. 7, 2006, Art. no. 074602.
[243] R. Ryf et al., "Mode-division multiplexing over 96 km of few-mode fiber using coherent 6×6 MIMO processing," *J. Lightw. Technol.*, vol. 30, no. 4, pp. 521–531, Feb. 2012.
[244] S. Gross, N. Riesen, J. D. Love, and M. J. Withford, "Mode-division multiplexing using femtosecond laser written 3-dimensional tapered couplers," in *Proc. Optoelectron. Commun. Conf. Austral. Conf. Opt. Fibre Technol.*, 2014, pp. 156–158.
[245] N. Hanzawa et al., "Demonstration of PLC-based six-mode multiplexer for mode division multiplexing transmission," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2015, pp. 1–3.
[246] S. G. Leon-Saval, N. K. Fontaine, J. R. Salazar-Gil, B. Ercan, R. Ryf, and J. Bland-Hawthorn, "Mode-selective photonic lanterns for space-division multiplexing," *Opt. Exp.*, vol. 22, no. 1, pp. 1–9, 2014.
[247] G. Labroille, B. Denolle, P. Jian, P. Genevaux, N. Treps, and J.-F. Morizur, "Efficient and mode selective spatial mode multiplexer based on multi-plane light conversion," *Opt. Exp.*, vol. 22, pp. 15599–15607, Dec. 2014.
[248] N. K. Fontaine et al., "Hermite-Gaussian mode multiplexer supporting 1035 modes," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2021, p. M3D.4.
[249] P. Pecci et al., "Pump farming as enabling factor to increase subsea cable capacity," in *Proc. SubOptic*, 2019, Art. no. OP14-4.
[250] A. Gnauck, P. J. Winzer, R. Jopson, and E. C. Burrows, "Efficient pumping scheme for amplifier arrays with shared pump laser," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2016, Art. no. M2.A.1.
[251] Y. Jung, A. Wood, S. Jain, Y. Sasaki, S.-U. Alam, and D. J. Richardson, "Fully integrated optical isolators for space division multiplexed (SDM) transmission," *APL Photon.*, vol. 4, no. 2, Feb. 2019, Art. no. 022801.
[252] P. M. Krummrich and S. Akhtari, "Selection of energy optimized pump concepts for multi core and multi mode erbium doped fiber amplifiers," *Opt. Exp.*, vol. 22, no. 24, pp. 30267–30280, 2014.
[253] Y. Tsuchida et al., "Simultaneous 7-core pumped amplification in multicore EDF through fibre based fan in/out," in *Proc. Eur. Conf. Exhib. Opt. Commun.*, 2012, Art. no. Tu.4.F.2.
[254] J. Sakaguchi et al., "19-core MCF transmission system using EDFA with shared core pumping coupled via free-space optics," *Opt. Exp.*, vol. 22, no. 1, pp. 90–95, 2014.
[255] K. S. Abedin et al., "Seven-core erbium-doped double-clad fiber amplifier pumped simultaneously by side-coupled multimode fiber," *Opt. Lett.*, vol. 39, no. 4, pp. 993–996, 2014.
[256] Q. Kang et al., "Minimizing differential modal gain in cladding-pumped EDFAs supporting four and six mode groups," *Opt. Exp.*, vol. 22, no. 18, pp. 21499–21507, 2014.
[257] L. Bigot, G. L. Cocq, and Y. Quiquempois, "Few-mode erbium-doped fiber amplifiers: A review," *J. Lightw. Technol.*, vol. 33, no. 3, pp. 588–596, Feb. 2015.
[258] Z. S. Eznaveh et al., "Ultra-low DMG multimode EDFA," in *Proc. Opt. Fiber Commun. Conf.*, 2017, Art. no. Th4A.4.
[259] J. Zhu et al., "Few-mode gain-flattening filter using LPFG in weakly-coupled double-cladding FMF," *J. Lightw. Technol.*, vol. 39, no. 13, pp. 4439–4446, Jul. 1. 2021.
[260] M. Jinno, "Evolution scenarios for spatial channel networks toward 1-Pb/s optical transport and beyond," in *Proc. Opto-Electron. Commun. Conf. (OECC)*, 2020, pp. 1–3.
[261] R. S. Luís et al., "Experimental demonstration of a petabit per second SDM network node," *J. Lightw. Technol.*, vol. 38, no. 11, pp. 2886–2896, Apr. 20, 2020.
[262] E. Ciaramella, "Wavelength conversion and all-optical regeneration: Achievements and open issues," *J. Lightw. Technol.*, vol. 30, no. 4, pp. 572–582, Feb. 15, 2012.
[263] D. M. Marom, Y. Miyamoto, D. T. Neilson, and I. Tomkos, "Optical switching in future fiber-optic networks utilizing spectral and spatial degrees of freedom," *Proc. IEEE*, vol. 110, no. 11, pp. 1835–1852, Nov. 2022, 2022.
[264] P. J. Winzer and D. T. Neilson, "From scaling disparities to integrated parallelism: A decathlon for a decade," *J. Lightw. Technol.*, vol. 35, no. 5, pp. 1099–1115, Mar. 2017.
[265] X. Chen et al., "Characterization and digital pre-compensation of electro-optic crosstalk in silicon photonics I/Q modulators," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, 2016, Art. no. Tu.3A.5.
[266] S. Randel, P. J. Winzer, M. Montoliu, and R. Ryf, "Complexity analysis of adaptive frequency-domain equalization for MIMO-SDM transmission," in *Proc. 39th Eur. Conf. Exhib. Opt. Commun. (ECOC)*, 2013, pp. 1–3.
[267] G. P. Agrawal, *Fiber-Optic Communication Systems*, 5th ed. Hoboken, NJ, USA: Wiley, 2021.
[268] G. Bosco et al., "Experimental investigation of nonlinear interference accumulation in uncompensated links," *IEEE Photon. Technol. Lett.*, vol. 24, no. 14, pp. 1230–1232, Jul. 2012.
[269] M. Ranjbar Zefreh and P. Poggiolini, "A GN-model closed-form formula supporting ultra-low fiber loss and short fiber spans," 2021, *arXiv:2111.04584*.

ABOUT THE AUTHORS

Werner Klaus received the M.Sc. degree (Hons.) in electrical engineering from the Technical University of Vienna (TU Wien), Vienna, Austria, in 1991, and the Ph.D. degree in electronics engineering from The University of Tokyo, Tokyo, Japan, in 1995.

Since 1995, he has been working at the National Institute of Information and Communications Technology (NICT), Tokyo, on various aspects of optical communications ranging from device to coherent optical system design both in free-space and modern fiber optics, including quantum key distribution (QKD) and space-division multiplexing (SDM). He is currently a Chief Senior Researcher with NICT.

Dr. Klaus has served as a Technical Committee Member of the Optical Fiber Communication Conference (OFC) from 2019 to 2021. He also serves as a member of the Technical Committee on Extremely Advanced Optical Transmission Technologies (EXAT) organizing the EXAT International Symposium that has been held in Japan biyearly since 2011 and has become an important platform for information exchange on the latest trends and developments of space-division multiplexing (SDM) technology between Japanese and foreign researchers. He is a member of Optica. He has served as the Program Chair of Optica's Advanced Photonics Congress from 2017 to 2018. Since 2018, he has been serving as Associate Editor for the JOURNAL OF LIGHTWAVE TECHNOLOGY.

Peter J. Winzer (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the Technical University of Vienna (TU Wien), Vienna, Austria, in 1998.

From 2000 to 2019, he worked at Bell Labs, Holmdel, NJ, USA, where he performed research on fiber-optic communication systems and networks, set multiple high-speed optical transmission records, and contributed to optical communications product developments. Following his involvement in estimating the optical fiber Shannon capacity, he investigated space-division multiplexing (SDM) to scale optical transport systems. In 2020, he founded the startup Nubis Communications, New Providence, NJ, USA. He has widely published and patented. He is actively involved with the IEEE and Optica.

Dr. Winzer is a Highly Cited Researcher, a Bell Labs Fellow, a Fellow of Optica, and an Elected Member of the U.S. National Academy of Engineering. He received multiple recognitions for his work, including the John Tyndall Award and an Honorary Doctorate from the Technical University of Eindhoven. He was the Program Chair of European Conference on Optical Communication (ECOC) in 2009 and the Program/General Chair of the Optical Fiber Communication Conference (OFC) from 2015 to 2017. He has served as the Editor-in-Chief of the JOURNAL OF LIGHTWAVE TECHNOLOGY from 2013 to 2018.

Kazuhide Nakajima (Member, IEEE) received the Ph.D. degree in electrical engineering from Nihon University, Chiba, Japan, in 2005.

In 1994, he joined Nippon Telegraph and Telephone (NTT) Corporation, Ibaraki, Japan, where he was engaged in research on optical fiber design and related measurement techniques. He is currently a Senior Distinguished Researcher with NTT.

Dr. Nakajima is a member of the Optica, the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan, and the Japan Society of Applied Physics (JSAP). He received the Best Paper Award in OptoElectronics and Communications Conference (OECC) in 1996 and 2019, the Best Paper Award in IEICE in 2011, the Achievement Award in IEICE in 2012, and the Maejima Hisoka Award in 2016. He has been a Rapporteur of Q5/SG15 of International Telecommunication Union–Telecommunication Sector (ITU-T) since 2009.