Scalable Modulation Technology and the Tradeoff of Reach, Spectral Efficiency and Complexity

Gabriella Bosco*, Dario Pilori, Pierluigi Poggiolini, Andrea Carena, Fernando Guiomar

Politecnico di Torino, DET, C.so Duca Degli Abruzzi 24, 10149 Torino, Italy

ABSTRACT

Bandwidth and capacity demand in metro, regional, and long-haul networks is increasing at several tens of percent per year, driven by video streaming, cloud computing, social media and mobile applications. To sustain this traffic growth, an upgrade of the widely deployed 100-Gbit/s long-haul optical systems, based on polarization-multiplexed quadrature phase-shift keying (PM-QPSK) modulation format associated with coherent detection and digital signal processing (DSP), is mandatory. In fact, optical transport techniques enabling a per-channel bit rate beyond 100 Gbit/s have recently been the object of intensive R&D activities, aimed at both improving the spectral efficiency and lowering the cost per bit in fiber transmission systems.

In this invited contribution, we review the different available options to scale the per-channel bit-rate to 400 Gbit/s and beyond, i.e. symbol-rate increase, use of higher-order quadrature amplitude modulation (QAM) modulation formats and use of super-channels with DSP-enabled spectral shaping and advanced multiplexing technologies. In this analysis, trade-offs of system reach, spectral efficiency and transceiver complexity are addressed.

Besides scalability, next generation optical networks will require a high degree of flexibility in the transponders, which should be able to dynamically adapt the transmission rate and bandwidth occupancy to the light path characteristics. In order to increase the flexibility of these transponders (often referred to as "flexponders"), several advanced modulation techniques have recently been proposed, among which sub-carrier multiplexing, hybrid formats (over time, frequency and polarization), and constellation shaping. We review these techniques, highlighting their limits and potential in terms of performance, complexity and flexibility.

Keywords: Coherent optical systems, advanced modulation formats, flexible transceivers, nonlinear propagation, spectral efficiency.

1. INTRODUCTION

Optical transport networks are currently migrating from a static configuration with little flexibility to the concept of software-defined optical networks.[1] The main reason is that the huge amount of traffic foreseen for the near future cannot be efficiently supported by static increases in network capacity. In fact, the available bandwidth in fibers is limited and an improvement in terms of spectral efficiency always comes at the cost of reduced reach. As a result, a homogeneous increase in data rate or spectral efficiency is not economically viable.[2] In the scenario of next generation optical networks, a key role will be played by flexible and scalable optical transceivers , able to dynamically adapt the modulation format and the transmission rate to the network conditions.[3]

Today, transmission at 100 Gbit/s, enabled by the use of quadrature phase shift keying (QPSK) at 32 Gbaud with polarization multiplexing (PM), has reached maturity and widespread deployment and 200 Gbit/s PM 16-quadrature amplitude modulation (16-QAM) coherent transceivers are already commercially available. Industry efforts are currently devoted to the ASIC development of transceivers for 400 Gbit/s and soon they will move to 1 Tbit/s. Since the transmission speed R is given by the product between the symbol rate R_s and the number of bits per symbol n_{bps} ($R_b = R_s \cdot n_{\mathrm{bps}}$), higher values of R can be obtained by increasing either R_s or n_{bps} (i.e. the order of the modulation format). Note that the first solution would have a strong impact on the hardware complexity, whilst the second one would mainly impact the digital signal processing (DSP) complexity. As an example, a raw transmission rate of 480 Gbit/s can be achieved either with 120-Gbaud PM-QPSK (n_{bps}=4)

*gabriella.bosco@polito.it; www.optcom.polito.it.

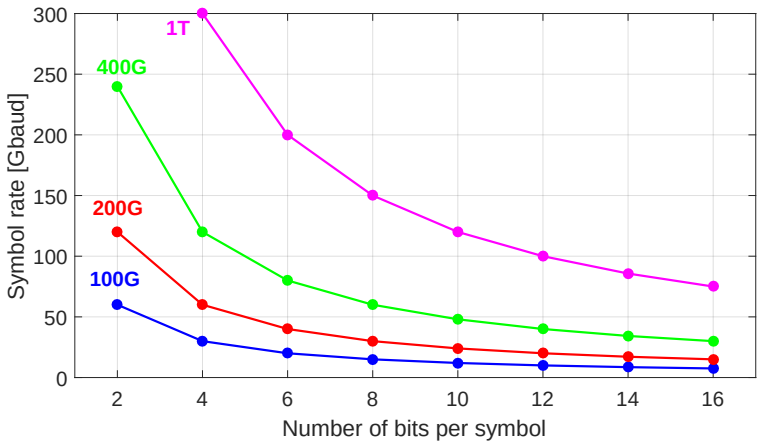

Figure 1. Symbol rate vs. number of bits per symbol needed to achieve a target net transmission speed of 100 Gbit/s, 200 Gbit/s, 400 Gbit/s or 1 Tbit/s (assuming that an FEC with 20% overhead is used).

or 30-Gbaud PM-256QAM (n_{bps}=16). Fig. 1 shows the values of symbol rate and number of bits per symbol needed to achieve a target *net* transmission speed of 100 Gbit/s, 200 Gbit/s, 400 Gbit/s or 1 Tbit/s, assuming that an FEC with 20% overhead is used.

The curves in Fig. 1 highlight the fact that serial interface rates well beyond those provided by state-of-the art components (electro-optic (E/O) converters, O/E converters, DACs, and ADCs) are needed to achieve transmission speeds beyond 200 Gbit/s. Note that this electronic bottleneck can however be efficiently circumvented through the use of *super-channels*, that exploit optical parallelism to provide high-per-channel data rates and better spectral utilization.[4] A *super-channel* is a collection of optical signals that are modulated and multiplexed together with high spectral efficiency (SE) at a common originating site, transmitted and routed over a common optical link as a single entity and received at a common destination site. Typically the super-channels use spectrally efficient advanced modulation formats in combination with advanced multiplexing schemes such as orthogonal frequency-division multiplexing (OFDM)[5] or Nyquist-WDM.[6] In this work, we focus on the use of the Nyquist-WDM technique, briefly describing its generation and detection in Section 2.

The *raw* SE (i.e. without taking into account the FEC overhead) of wavelength-division-multiplexing (WDM) optical systems is defined as the information capacity of a single channel (in bits/s) divided by the frequency spacing Δf (in Hz) between the carriers of the WDM comb:[7]

$$\text{SE} = n_{\text{bps}} \cdot \frac{R_s}{\Delta f} = \frac{n_{\text{bps}}}{\delta f} \tag{1}$$

where δf is the frequency spacing normalized to the symbol rate ($\delta f = \Delta f / R_s$). The total system capacity (defined as the maximum information in bits/s that can be transmitted by the WDM comb) is obtained as the product between the SE and the available bandwidth. The maximization of the SE thus plays an important role in the maximization of the overall system capacity.

In the past years, the SE of optical systems has significantly increased, mainly due to the advent of coherent detection technologies, which enabled the use of polarization-multiplexed high-order modulation formats. However, the use of high-order modulation requires a higher *optical signal-to-noise ratio* (OSNR), which may result in a significantly reduced achievable transmission distance, as shown in Section 3. On the contrary, an increase in R_s has potentially no impact on SE (because the value of Δf needs to be increased accordingly in order to limit the cross-talk between adjacent channels), while it might have a (low) negative impact on the reach, as shown in Section 4.

Besides scalability, a high degree of flexibility will be required in next generation optical networks, i.e. whose transponders should be able to dynamically adapt the transmission rate and bandwidth occupancy to the light

path characteristics. Several advanced modulation techniques have recently been proposed to increase the flexibility of the transponders, among which sub-carrier multiplexing, hybrid formats and constellation shaping. These techniques are reviewed in Sections 5, 6 and 7, highlighting their limits and potential in terms of performance, complexity and flexibility.

2. NYQUIST-WDM

As discussed in the previous section, the SE can be increased by decreasing the normalized frequency spacing between the channels of the WDM comb. To achieve ultimate spectral efficiency, WDM channel spacings are reduced until the optical spectra of neighboring channels start to noticeably overlap. In this limit of ultra-dense WDM systems, linear crosstalk between adjacent WDM channels becomes a main source of degradation.[8] An efficient countermeasure to limit the crosstalk is based on an accurate spectral shaping of each channel of the WDM comb[9–11] . Employing this technique, known in optical communications as *Nyquist-WDM*, the transmission of PM-QPSK WDM signals with channel spacing equal to the symbol rate has been demonstrated over transpacific distances.[12,13]

2.1 Generation of a Nyquist-WDM signal

Nyquist-WDM signal generation and transmission have been demonstrated using either optical filters or DSP-based electrical filters.[14–17] Compared to the optical filter solution, the DSP-based approach (whose schematics is shown in Fig. 2 (left)) is more flexible as the same hardware can be used to generate different filter shapes and modulation formats.[9,18] In practice, the system cost is limited by the complexity, speed, and power consumption. Therefore, it is necessary to optimize the digital Nyquist design for the required signal in order to minimize complexity and speed.[19–21]

For a Nyquist-WDM signal, in order to avoid inter-symbol interference (ISI) among adjacent pulses, a raised-cosine (RC) shape is usually selected. The RC pulse shape in the time and frequency domain for different values of roll-off factor ρ is shown in Fig. 2 (right), where the case $\rho = 0$ corresponds to the rectangular spectrum with bandwidth R_s. Ideally, no cross-talk is present between adjacent WDM channels when $\delta f > (1+\rho)$. In the case in which the channel has a flat frequency response, the optimum receiver filter is matched to the transmit one, thus both assume a square-root raised cosine (SRRC) shape.

The achievable symbol rate and cardinality of the modulation format are limited by the DAC characteristics. In fact, a realistic DAC device is characterized by two main parameters: the number of resolution bits N_{DAC}, which limits the cardinality of the modulation format, and the sampling speed f_{DAC}, which limits the achievable symbol rate $R_s = f_{\mathrm{DAC}}/N_{SpS}$, where N_{SpS} is the number of samples per symbol (also indicated as "oversampling factor"). Typically if f_{DAC} increases, N_{DAC} decreases. The achievable symbol rate can clearly be increased by decreasing the oversampling factor. In doing so, penalties could be incurred due to interference produced by spectral replica of the useful spectrum in the DAC process. In,[22] an oversampling factor as low as 1.15 was used, limiting the penalty due to spectrum replica thanks to the use of ad-hoc anti-alias electrical filters.

Nyquist pulse shaping in the digital domain is typically performed using FIR filters. The required length of the digital filter needed to generate SRRC or RC pulses increases when the roll-off decreases (see Fig. 2). In[23] complexity and performance of digital pulse shaping has been investigated, showing that FIR filters with 17 taps

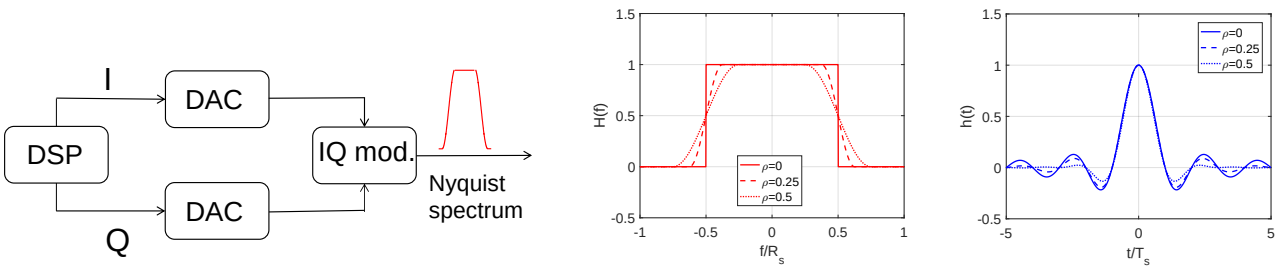

Figure 2. Spectral shaping in the digital domain (left). Raised-cosine pulse in frequency (center) and time domain (right).

allow for a reduction in channel spacing to 1.1 the symbol rate within a 1-dB penalty. A higher number of FIR filter taps are needed for tighter channel spacing. In,[24] a real-time demonstration of generation of Nyquist-like pulses with 14-Gbaud PM-16QAM was reported, showing that the use of a 32-tap FIR filter to shape the signals would allow a channel spacing equal to $1.06 \cdot R_s$ without substantial cross-talk penalty.

2.2 Detection of a Nyquist-WDM signal

In conventional WDM systems, sub-channels are first de-multiplexed in the optical domain and then separately detected. This approach cannot be applied to Nyquist-WDM signals, in which the sub-channels are too closely spaced to be separated through optical filtering without incurring in any penalty. However, sharp filter shapes can be efficiently implemented at the receiver (Rx) in the digital domain, enabling detection of single sub-channels without the need of any tight optical filtering at the Rx. In particular, Nyquist-WDM signals can be detected using a standard coherent receiver,[25] composed of a 90-degree hybrid, followed by four balanced photo-detectors (BPDs), whose functionality is to map the optical field into four electrical signals, corresponding to the in-phase and quadrature field components for the two polarizations. The single sub-channel is selected by tuning the local oscillator (LO) to the center frequency of the sub-channel. The four analog electrical signals are sampled by four analog-to-digital converters (ADCs) with sampling speed f_{ADC} and the signal samples are elaborated by ad-hoc digital signal processing (DSP) algorithms,[25] which perform polarization recovery and compensation of propagation linear (and possibly non-linear) impairments.

Note that anti-aliasing electrical filters can be placed before the ADCs in order to reduce the bandwidth of the signal entering the ADCs, thus relaxing the requirements for the ADC sampling frequency and enabling more efficient DSP for polarization recovery and impairments compensation. Typically, $f_{\mathrm{ADC}} = 2 \cdot R_s$, corresponding to a number of samples per symbol N_{SpS} equal to 2, but lower values can be used without incurring in substantial penalties, as shown in the experiments reported in[12] and,[26] where a value of N_{SpS} equal to 1.67 was used.

In,[20] the matched filter design for SRRC spectrally shaped Nyquist-WDM systems has been addressed. Typically, two equalizers are implemented in the coherent Rx DSP: one CD equalizer and one adaptive butterfly blind equalizer. The CD equalizer is used to compensate for the large amount of accumulated CD in the fiber link and usually is a static frequency domain equalizer (FDE).[27] In[20] it is shown that incorporating a matched filter in the bulk CD equalizer for a SRRC shaped signal can significantly reduce the complexity of the blind equalizer, with no additional complexity added to the CD equalizer. It is also shown that Nyquist-WDM systems with matched filtering are sensitive to the frequency offset between the transmitter (Tx) laser and the local oscillator (LO), and that the induced penalty decreases with increased SRRC roll-off factor.

3. TRADEOFFS BETWEEN SPECTRAL EFFICIENCY AND REACH

The spectral efficiency of WDM optical systems can be significantly increased by increasing the number of constellation points, i.e. using polarization-multiplexed high-order modulation formats such as PM-16QAM or PM-64QAM. However, a reduction of system reach is incurred because higher order modulation formats require an higher SNR to achieve a certain target BER and are typically more sensitive to practical implementation issues or system impairments (e.g. cross-talk, phase noise, ...).

In the following, we report a set of results highlighting the trade-off between SE and reach for different modulation formats (ranging from PM-QPSK to PM-64QAM). In order to analyze a realistic scenario, we assumed to operate with a conservative SNR margin of 3 dB with respect to the ideal back-to-back (BtoB) performance of each format. The effects of nonlinear propagation over a multi-span link with EDFA only amplification (noise figure F=5 dB) are estimated using the EGN model.[28] Four different scenarios are considered: propagation over standard single-mode fiber (SSMF) with either 50 or 100-km span length and propagation over pure silica-core fiber (PSCF) with either 50 or 100-km span length. The parameters of the fibers are shown in Table 1. The symbol rate is R_s=32 Gbaud. The WDM signal is assumed to occupy the entire C-band (B_{WDM}=5 THz) with a spacing among WDM sub-channels equal to $1.05 \cdot R_s$, i.e. the normalized frequency spacing is equal to $\delta f = 1.05$, the lowest value assuring the absence of any linear cross-talk for the employed SRRC spectra (with roll-off 0.05).

The dependence of the SE on total link length is plotted in Fig. 3 for PM-QAM modulation formats with cardinality ranging from 4 to 64, assuming to use soft-decoding FEC schemes.[29] The asymptotic spectral

Table 1. Parameters of the analyzed fiber links.

Link #	Fiber type	Dispersion [ps/nm/km]	Loss [dB/km]	Non-linearity coeff. [$W^{-1}km^{-1}$]	Span length [km]
Link 1	SSMF	16.7	0.2	1.3	100
Link 2	PSCF	20.5	0.165	0.75	100
Link 3	SSMF	16.7	0.2	1.3	50
Link 4	PSCF	20.5	0.165	0.75	50

efficiency achievable by each modulation format at very low distances is equal to $n_{\text{bps}} \cdot \delta f$. The results of Fig. 3 clearly highlight the trade-off between SE and reach, in relation to the different modulation formats. Increasing the cardinality of the constellation, a higher SE can be achieved, but over a shorter transmission distance and/or at a higher required FEC overhead (the required FEC rate is given by the ratio between the achieved SE and the asymptotic SE of the modulation format).

In a terrestrial-like link over SSMF (Link 1) PM-QPSK is the best choice for ultra-long-haul transmissions beyond 5,000 km, while PM-8QAM can be used to achieve a SE around 5-5.5 bit/s/Hz in 3,000 km links. PM-16QAM allows to reach 2,000 km with a SE of 6 bit/s/Hz. The reach of higher-order modulation formats, like PM-32QAM pr PM-64-QAM, is very limited in this kind of systems (unless extremely powerful FEC codes are used), but can be significantly increased by using new generation fibers and shorter span lengths, like in the submarine-like system over PSCF (Link 4), where they can potentially reach 4,000 km and 2,000 km, respectively.

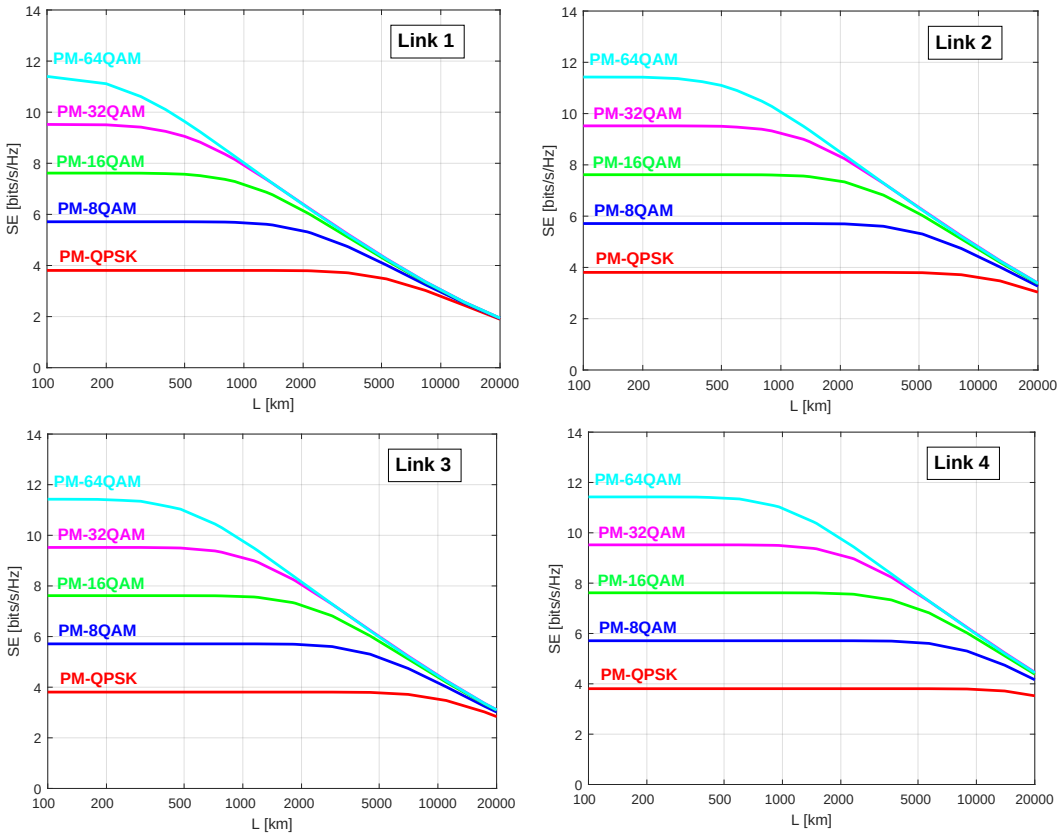

Figure 3. Spectral efficiency versus total link length in the four different transmission scenarios described in Table 1. R_s=32 Gbaud, B_{WDM}=5 THz, $\Delta f = 1.05 \cdot R_s$. Assumption: 3-dB SNR margin from ideal BtoB performance.

4. IMPACT OF SYMBOL RATE ON SYSTEM REACH

Over the last few years, various simulative and theoretical papers[30–35] have presented evidence of a dependence of system performance on the transmission symbol-rate. Exploiting such dependence, maximum-reach (MR) or 'Q-factor' gains were found by performing symbol-rate optimization (SRO). The optimum rates have turned out to be substantially lower than the current 32 Gbaud industry standard, and typically in the 2 to 6 Gbaud range. As for the extent of the MR gains, it has been found to be in the 5%–25% range. Regarding the origin of SRO gains, they were ascribed to peculiar features of fiber non-linear propagation.

Besides theory and simulation, also various *experimental* papers have lately addressed SRO.[36–43] The results have not been unanimous. Note in passing that, given the low value of the optimum rates, these experiments were conducted by breaking up each single optical channel into electrical subcarriers, generated digitally through the transmitter DSP and DACs (see Section 5 for a more detailed description of subcarrier multiplexing). In[37] a 23% MR increase was found when a single 24 Gbaud channel was broken up into 8 subcarriers at 4 Gbaud. In[38] substantial MR gain was found in a WDM experiment. However, the gain could also be attributed to other beneficial effects of using subcarriers in that particular set-up. Other WDM experiments[39–41] found conflicting results, between about 8% MR gain and no gain at all. One ultra-long-haul experiment[36] based on PM-QPSK recently found a 12.5% reach increase over PSCF, from 12,610 km at 32 Gbaud, to 14,180 km at 4 Gbaud. Very recent evidence, from a massive experiment[42] involving a fully-populated C+L band system, has shown a 0.8 dB Q-gain in a trans-oceanic type system with PM-QPSK, potentially leading to about 20% reach increase. However, only 0.2 dB Q-gain or about 5% potential reach increase was found when using PM-16QAM. Overall, at present, experimental evidence still appears somewhat contradictory as to the extent of possible SRO gains.

In this section we look at the topic of SRO from a theoretical/simulative viewpoint, mostly referring to[36] . As shown in[36] , a non-linear interference (NLI) model that appears to work well in predicting SRO performance is the EGN model[28,44,45] . Thanks to the EGN model, it is possible to analytically explore a very wide range of system scenarios. In particular, it is possible to push the total system bandwidth to *C-band*, which is going to be the focus point of this section. We also show split-step simulation results that back-up the analytical EGN predictions, albeit up to 1.5 THz of optical bandwidth because of excessive computational load above this limit.

In order to properly perform the investigation, a suitable test system layout is necessary. The per-channel symbol rate R_s should be the free set-up parameter that is probed. All other system features should be kept fixed. To this purpose, we impose: the modulation format and hence the number of bits per symbol n_{bps}; the *total optical bandwidth* B_{WDM}; the channel spectrum roll-off ρ, assuming an SRRC spectral shape for the transmitted pulses; the *relative* channel spacing $\delta f = \frac{\Delta f}{R_s}$. These parameters determine the system spectral efficiency SE and the total (raw) bit rate $R_{\mathrm{b,tot}}$, which are, respectively: $SE = \frac{n_{\mathrm{bps}}}{\delta f}$ and $R_{\mathrm{b,tot}} = B_{\mathrm{WDM}} \cdot SE$. Changing the value of R_s affects the number of channels that make up the system, and possibly MR performance, but the key parameters B_{WDM}, SE and $R_{\mathrm{b,tot}}$ stay the same.

Regarding the specific link features, we assume same fiber type in every span and a uniform span length L_{span} = 100 km, i.e., a typical terrestrial span length. We choose standard single-mode fiber (SMF) with attenuation 0.22 dB/km, dispersion D=16.7 ps/(nm·km) and non-linearity parameter $\gamma = 1.3$ (W·km)$^{-1}$, respectively. We assume EDFA amplification, with span loss exactly compensated for by amplifier gain. The signal spectrum is raised-cosine with roll-off 0.05, the channel spacing is $1.05 \cdot R_s$, i.e., $\delta f = 1.05$. This test set-up is analyzed using either the EGN model, or simulations. In both cases, at the link output, the total NLI power P_{NLI} falling over the center channel of the WDM comb is assessed. However, the raw P_{NLI} datum is not adequate *per se* to directly compare system performance at different symbol rates. A derived quantity which we called $\tilde{G}_{\mathrm{NLI}}$ provides a more suitable tool:

$$\tilde{G}_{\mathrm{NLI}} = \frac{P_{\mathrm{NLI}}}{R_s\, G_{\mathrm{ch}}^3} \tag{2}$$

This quantity can be viewed as the power spectral density (PSD) of NLI noise impinging on the center WDM channel, normalized versus the transmitted signal PSD, raised to the third power G_{ch}^3. The key features of $\tilde{G}_{\mathrm{NLI}}$, are: it is independent of the power per channel launched into the link; if the *same value* of $\tilde{G}_{\mathrm{NLI}}$ is found among systems using different symbol rates, this means that such systems can potentially achieve *the same maximum reach.* A higher value of $\tilde{G}_{\mathrm{NLI}}$ means smaller MR, a lower value means greater MR. This feature makes $\tilde{G}_{\mathrm{NLI}}$

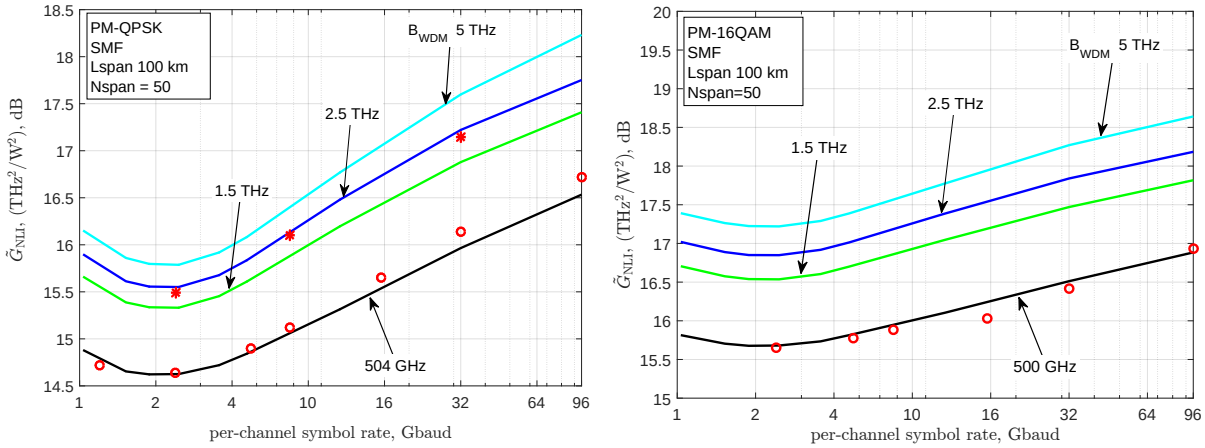

Figure 4. Normalized average NLI noise power spectral density $\tilde{G}_{\rm NLI}$ over the center channel, vs. the symbol rate per channel R_s, for four WDM bandwidth values: 0.5, 1.5, 2.5 and 5 THz (C-band). NLI is measured at 50 spans of SMF. Roll-off is 0.05, channel spacing $1.05{\cdot}R_s$. Left: PM-QPSK modulation. Right: PM-16QAM. Solid lines: calculations using the full EGN model. Markers: dual-polarization split-step simulations: circles for 0.5 THz, '*' markers for 1.5 THz.

very convenient for system comparison across different symbol rates. As for the extent of the potential MR gain or loss, the following approximate, but rather accurate relation holds:[43]

$$(\Delta\mathrm{MR})_{\rm dB} \approx \frac{1}{3}\left(\Delta\tilde{G}_{\rm NLI}\right)_{\rm dB} \tag{3}$$

where Δ represents the *ratio* of two values of the quantity. For example, a 1 dB NLI mitigation, that is $\Delta\tilde{G}_{\rm NLI} = -1$ dB, can be expected to translate into $\Delta\mathrm{MR} = 1/3$ dB, i.e., about 8% MR increase.

The NLI analytical assessment was carried out by numerical integration of the full EGN model according to the formulas reported in[28] . The simulations were carried out based on a full-band split-step method. The NLI noise was measured on the center channel after subtracting a linearly-propagated version of the signal from the non-linearly-propagated one. The Rx compensated statically for polarization rotation and applied an ideal matched filter. No dynamic equalizer was used, to avoid any possible effect of the equalizer adaptivity on NLI estimation. The simulation was completely noiseless: neither ASE noise, nor any other types of noise, such as Rx electrical noise or laser phase/intensity noise, were present. Simulations were run between 2^{16} and 2^{18} symbols, with fewer symbols at lower symbol rates.

Fig. 4-(left) shows the normalized average NLI noise power spectral density $\tilde{G}_{\rm NLI}$ at 50 spans of SMF assuming PM-QPSK transmission. The number of spans was chosen because it was close to the MR of the system at 32 Gbaud, assuming BER 10^{-2}. The total WDM bandwidth $B_{\rm WDM}$ is ramped up from 500 GHz to 1.5, 2.5 and 5 THz (C-band). The abscissa is the symbol rate per channel R_s. The solid curves are obtained using the EGN model, markers are simulations. The general shape of all the curves remains unchanged for the different values of $B_{\rm WDM}$ and, in particular, the optimum symbol rate does not change. However, the NLI mitigation $\Delta\tilde{G}_{\rm NLI}$, between the current industry standard (32 Gbaud) and the optimum rate, increases as $B_{\rm WDM}$ is increased. At C-band it is about 1.8 dB, up from about 1.3 dB at $B_{\rm WDM}$= 500 GHz. We tried to back up this finding by extending the simulations to at least $B_{\rm WDM}$= 1.5 THz. We show three data points (red '*' markers) that, although slightly higher than the analytical curve, confirm the trend of increasing mitigation for larger $B_{\rm WDM}$. Pending further necessary experimental confirmations, Fig. 4-(left) provides the important indication that SRO does not lose effectiveness when $B_{\rm WDM}$ is increased, as opposed for instance to digital back-propagation (DBP). Rather, effectiveness slightly improves. This is quite significant, in view of practical applications.

We then extended the investigation of SRO to PM-16QAM. Fig. 4-(right) shows the results, assuming the same system layout and parameters as for Fig. 4-(left). We chose to measure again NLI at 50 spans, to allow a direct comparison with the PM-QPSK results. Note that 50 spans would still be close to the system MR at 32 Gbaud, if assuming hybrid EDFA-Raman amplification and about 10^{-2} as target BER. The PM-16QAM

simulations for $B_{\mathrm{WDM}} = 500$ GHz (diamond markers) are shown together with the EGN model curves (solid lines). NLI mitigation due to SRO is still clearly present, though its extent is smaller than for PM-QPSK. At C-band it is about 1.15 dB, between 32 Gbaud and the optimum symbol rate, vs. 1.8 dB for PM-QPSK. Interestingly, the optimum symbol rate is the same as that of PM-QPSK, about 2.4 Gbaud.

The results shown in Fig. 4 appear encouraging and show that SRO could be a supporting technique for NLI mitigation in the next generation of optical systems. One key advantage of this technique vs. others, such as DBP, is its non-diminishing return vs. B_{WDM} and the relatively low-complexity implementation. The general indications of Fig. 4 are also in general agreement with the experimental results of[42] , described before, in particular for PM-QPSK. For PM-16QAM, the smaller than expected mitigation found can likely be explained with system impairments hitting that format more seriously than PM-QPSK. Also, the smaller advantage of SRO with PM-16QAM vs. PM-QPSK shown in Fig. 4 has recently been ascribed to the impact of non-linear phase noise (NL-PN). Part of such NL-PN could perhaps be removed by more optimized phase recovery algorithms. This aspect is dealt with in the next section.

Finally, Fig. 4 not only shows mitigation when going from 32 Gbaud down the optimum rate: it also shows an *NLI increase* when moving from 32 Gbaud towards *higher rates*. This strongly suggests that future 64 or 96 Gbaud systems should perhaps be designed with some degree of SRO. If going down all the way to the optimum rate is too complex, a choice for instance of 8 Gbaud (12 subcarriers for a total of 96 Gbaud) would provide most of the mitigation without requiring an excessive number of subcarriers in the channel.

5. SUBCARRIER MULTIPLEXING

Following the predicted benefits of symbol-rate optimization highlighted in the previous section, new optical transmission techniques are required in order to allow for cost-effective operation at relatively low symbol-rates (typically 2–10 Gbaud). Considering that 32 Gbaud optical transceivers are currently commercially available at competitive prices, and if one wants to operate at a nominal symbol-rate of e.g. 2 Gbaud, it is clearly inefficient to require 16× more transceivers (including laser sources) to transmit at the same bit-rate. This has led to the recent interest on electrical subcarrier multiplexing (SCM) systems, in which a high symbol-rate signal is electrically decomposed into a given number of subcarriers, each of which operating at a lower symbol-rate. SCM takes advantage of the high flexibility provided by commercially available high-speed digital-to-analog converters (DACs), allowing to fully design the transmitted signal in the electrical domain. Depending on the optical system under test, several degrees of freedom can be optimized, such as the number of subcarrier components, the inter-subcarrier spacing and the associated modulation format and pulse shaping. The great advantage of DAC-enabled SCM is that all this flexibility does not require any additional hardware, thus allowing to reuse currently available single-carrier transceivers by simply reprogramming the transmitted signal.

The concept of SCM transmission is illustrated in Fig. 5. Note that the only subsystems that have to be adapted from single-carrier transmission are the data assignment to each subcarrier (serial-to-parallel block)

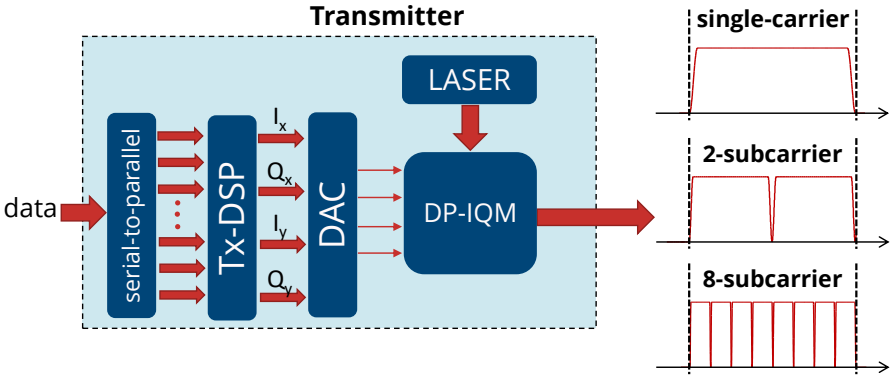

Figure 5. Simplified block diagram illustrating the structure of an SCM transmitter. For simplicity, only one polarization is depicted.

and the corresponding DSP (Tx-DSP block) required for electrical generation of subcarriers, pulse shaping and pre-equalization of the transceiver response. The transition from single-carrier to SCM transmission therefore only requires a software reconfiguration of the transceiver, without implying any additional costs. Also note that, thanks to quasi-Nyquist pulse shaping, the electrical subcarriers can be very tightly spaced, allowing to maintain high SE operation without incurring into inter-symbol interference penalties.

As highlighted in Section 4, the encouraging results of[37,46] have then motivated a multitude of further experimental studies on SCM transmission, with the main aim of identifying and demonstrating the benefits coming from SRO. This investigation, which is still undergoing, has led to somewhat controversial results. While some works have shown maximum reach gains in the range of 10-25%,[37,42,46–49] other works have found little or no performance advantage from SCM transmission.[40,50,51] As highlighted in most of the recently published experimental works on SCM transmission, the increased BtoB penalty experienced at low symbol-rates per subcarrier can in some cases counterbalance the benefits enabled by SRO. These B2B penalties are mostly associated with the reduced tolerance to phase noise at low symbol-rates and with the enhanced detrimental effect of transmitter IQ skew on SCM signals.[52] Improved post-detection DSP algorithms are therefore still needed to unleash the full potential of SCM transmission.

In the following, we review the results originally published in[47] for a PM-QPSK-based SCM system operating at a nominal symbol-rate of 32 Gbaud, including several SCM configurations. The transmitted signal was composed of 19 identical WDM channels with 37.5 GHz separation. Each channel was then transmitted either as 1×32 Gbaud, 8×4 Gbaud or 16×2 Gbaud. The recirculating loop was composed of four spans of pure silica core fiber (PSCF) with an average span length of 108.2 km. More details on the experimental setup conditions can be consulted in.[47] The maximum reach results obtained with each SCM configuration are shown in Fig. 6. The adopted target BER is 10^{-2}. The obtained results show an advantage of ∼12% in terms of extended reach provided by the 8×4 Gbaud SCM configuration over single-carrier transmission. This gain is in good agreement with the analytical predictions provided by the EGN model,[28] represented by the dashed lines in Fig. 6, thus experimentally confirming the SRO effect identified in section 4.

Besides enabling to explore the nonlinear propagation benefits associated with SRO, other advantages can be obtained from SCM transmission. Inherent benefits in terms of system flexibility can be achieved by adjusting the number of subcarriers, modulation formats and spectral occupation to the current load of the network, without the need for any hardware reconfiguration. Another prominent advantage of SCM for long-haul transmission is that chromatic dispersion compensation can be simplified or even completely removed, since the accumulated dispersion perceived at low symbol-rates can be negligible.[53] Also taking advantage of the same properties, several works have been recently published on simplified DBP algorithms optimized for SCM transmission.[54,55]

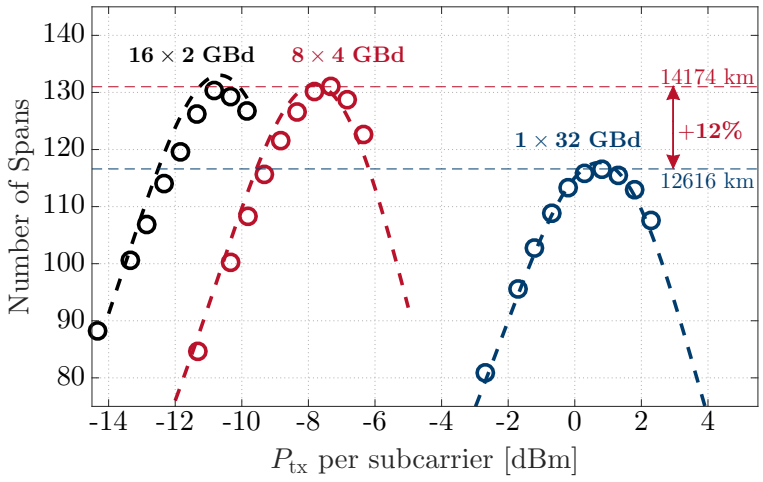

Figure 6. Maximum reach versus launched power per subcarrier at a nominal symbol-rate of 32 Gbaud employing single-carrier and SCM transmission. Markers refer to experimental results. Dashed lines correspond to the analytical predictions provide by the EGN model.

6. HYBRID MODULATION FORMATS

As telecommunications carriers look for a return on capex investments in the installed infrastructures, solutions introducing flexibility in next generation optical networks must avoid the replacement of fibers and link components, like multiplexers, demultiplexers and ROADMs. Under these conditions, a first step towards flexible networks consists in keeping fixed-grid dense WDM transmission on the installed equipment, introducing flexibility only through the replacement of transponders.[56] Future transponders need to become *flexponders* able to adapt their operation to traffic demand and in particular maximizing the throughput for a given reach, trading off the data rate with the lightpath quality of transmission (QoT), i.e., with the generalized optical signal-to-noise ratio (OSNR) taking into account ASE and NLI contribution.

Thanks to the introduction of fast DSP both at the Tx and the Rx, commercial solutions today available on the market can provide some level of flexibility, having the capability to switch among PM-mQAM formats. These solutions though allow a very limited amount of flexibility: as an example a transponder able to switch among PM-QPSK and PM-16QAM, can deliver only 4 or 8 bits per symbol (n_{bps}), i.e. 100G or 200G at 32 Gbaud. With a reduction in maximum reach of about 5 times with respect to PM-QPSK when using PM-16QAM, because the performance scale accordingly to back-to-back (BtoB) where the difference is about 7 dB. Small improvements can be achieved employing non-square PM-mQAM formats, but they add complexity to transponders because there is no more direct mapping of bit tributaries to signal quadratures.

In recent years several solutions for achieving an enhanced flexibility in data-rate have been proposed and demonstrated experimentally: four-dimensional optimized modulation formats,[57,58] coded modulation,[59,60] rate-adaptive modulation[61] and hybrid modulation formats. In this section we focus on this last approach based on the mixing of two standard modulation formats, usually square PM-mQAM, to obtain fractional values of n_{bps}. The hybridization can be done with respect to time, quadrature and frequency. When varying the PM-mQAM modulation with respect to time, a time-division hybrid modulation format (TDHMF) is generated: their performance has been extensively assessed both in simulations and experimental works.[62–67]

A simpler solution called *Flex-PAM*, with a limited amount of flexibility but without any added complexity required to the receiver DSP, is based on the use of a different PAM modulation on each of the four quadratures of a polarization multiplexed channel.[68,69] Recently a new approach based on the sub-carrier modulation has been proposed:[70] it considers to mix PM-mQAM formats on different sub-carriers using the frequency domain as discriminator, thus generating a frequency-division hybrid modulation format (FDHMF).

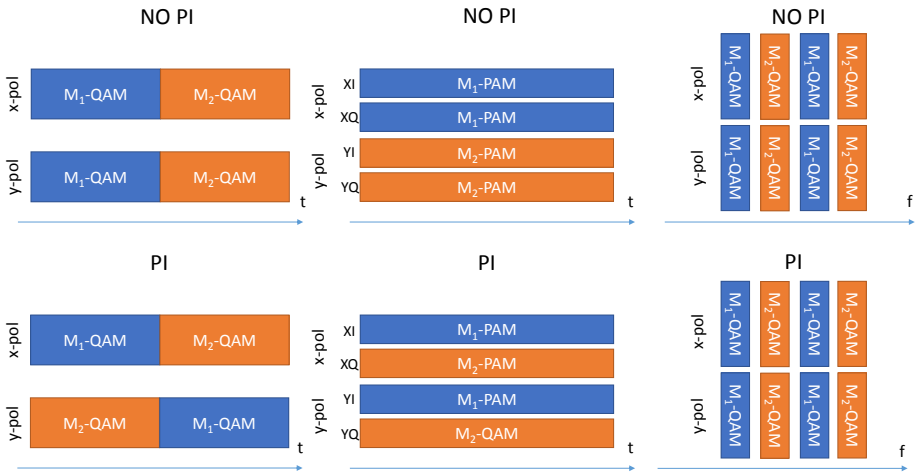

Figure 7. Structure of hybrid modulation formats: TDHMF (left), Flex-PAM (center) and FDHMF (right). Top row: without Polarization Interleaving. Bottom row: with Polarization Interleaving. For simplicity we show only structures for a single frame ratio (K=0.5). Power ratio between formats is not visualized in this figure.

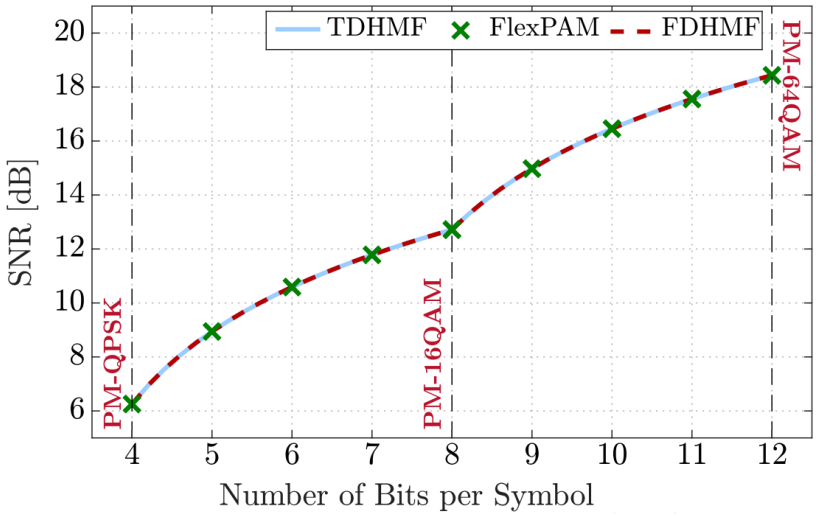

Figure 8. Back-to-back sensitivity of TDHMF, Flex-PAM and FDHMF based on the "same BER" transceiver operation strategy for a number of bits per symbol ranging from 4 (PM-QPSK) up to 12 (PM-64QAM). The target BER is 2×10^{-2}.

6.1 Time-Division Hybrid Modulation Format (TDHMF)

The general structure of a dual-polarization TDHMF frame is illustrated in Fig. 7 top left, In order to reduce the complexity associated with the transmission and detection of TDHMF, in this work each periodic frame is considered to be composed of N symbols distributed among two neighboring square QAM modulation formats with constellation sizes M_1 and M_2. The TDHMF frame is characterized by a given format ratio, $K = N_1/N$, where N_1 represents the number of symbols occupied by the modulation format with the lowest constellation size M_1. The throughput can be fine tuned depending on $n_{\text{bps}} = 2 * (log_2(M_1) * N_1 + log_2(M_2) * N_2)/N$. A key parameter for TDHMF is the power ratio (PR) between the two modulation formats employed, defined as $PR = P_{PM-M_1QAM}/P_{PM-M_2QAM}$. PR depends on the adopted operation strategy for the transponder. In[69] different solutions have been studied, showing that a simple approach based on the principle that both formats must work at same BER, namely the BER threshold for the FEC applied, is very close to the optimal approach where the overall BER of TDHMF is jointly minimized. Fig. 8 reports the BtoB results presented in,[69] based on this "same BER" strategy, showing that a continuous tuning of the required SNR as a function of n_{bps} was achieved. When analyzing the performance of TDHMF under non-linear fiber propagation, a detrimental impact of time-varying optical power was found. In order to counteract it, a strong improvement can be achieved applying polarization interleaving (PI), as illustrated in Fig. 7 bottom left. Note that, from a BtoB perspective, PI has no impact on the performance of TDHMF, but improve the resilience to non-linear propagation. Using the PI solution, in[64] and[69] it was demonstrated that, mixing PM-QPSK and PM16QAM first and then PM-16QAM and PM-64QAM, quasi-continuous trade-off of reach for n_{bps} (i.e. bit-rate) can be achieved. Moreover, a negligible penalty with respect to GN-model predictions has also shown, meaning that TDHMF with PI in non-linear regime behaves similarly to standard modulation formats.

6.2 Flex-PAM

The Flex-PAM frame structure is shown in Fig. 7 top center. Contrary to TDHMF it does not imply a time-varying modulation. Instead, the flexibility is achieved through the hybridization of two M-PAM formats between the four orthogonal quadratures (I and Q of the two polarizations). Consequently, the granularity of Flex-PAM is inherently limited to integer numbers of bits per symbol. Considering that the Flex-PAM frame is composed of the two nearest-sized PAM formats, whose constellations include M_1 and M_2 symbols. Depending on the number of quadratures assigned to each of the two PAM formats employed we can have three different format ratios, K, i.e. 0.25, 0.5 and 0.75, so that $n_{\text{bps}} = 4 * K * log_2(M_1) + 4 * (1 - K) * log_2(M_2)$. BtoB performance

achieved by Flex-PAM is derived in[69] and is based on same principles of PR optimization used for TDHMF for the different operation strategies. In Fig. 8 some results are reported, based on the "same BER" strategy derived in:[69] SNR required at each n_{bps}, limited to integer values, equal those of TDHMF in same conditions. FlexPAM can have a strong power unbalance between quadratures and polarizations that as a strong impact on performances in non-linear fiber propagation. Similarly to TDHMF, to counteract such behavior, we can apply polarization interleaving (PI) as illustrated in Fig. 7 bottom center. PI has not any impact on B2B performance but improve the reach in non-linear propagation.[68] To further improve the resilience to non-linear propagation in[68] it was also proposed to maximize reach by adjusting the PR, moving away from theoretical predictions based on the linear AWGN conditions. Combining PI and PR optimization, it was possible to reduce the penalty shown by FlexPAM with respect to TDHMF and GN model predictions.[69]

6.3 Frequency-Division Hybrid Modulation Format (FDHMF)

As reported in Sec. 4, recently it has been shown that symbol-rate optimization can be applied to reduce non-linear propagation impact. Optimal symbol-rates result to be very low, in the order of few Gbaud, so that SCM has been proposed as a solution to get SRO advantages. Besides the SRO benefits, SCM can also allow an increased flexibility relatively to single-carrier systems: in[70] we proposed the use of frequency domain hybrid modulation formats (FDHMF) as a time-invariant alternative to TDHMF able to achieve similar level of bit-rate granularity. FDHMF structure is shown in Fig. 7 top right: subcarrier carrying different PM-MQAM modulations are mixed over the same optical carrier. FDHMF is exactly the dual of TDHMF in frequency domain. The same theory presented in[69] for TDHMF can be re-used for FDHMF: symbols (in time) must be mapped in subcarriers (in frequency). We can define a format ratio, $K = N_1/N$, where N_1 is the number of subcarriers with a constellations of M_1 symbols and N the total number of subcarriers. Overall throughput in n_{bps} and power ratio PR both have the same expression defined for TDHMF. Same operation strategies can be applied and same result can be obtained: in Fig. 8 FDHMF curves are by definition overlapped to TDHMF. As shown in,[70] FDHMF has a further degree of freedom: applying different permutations of modulation format assignment to each subcarrier a large number of equivalent configurations is possible. For example, three scenarios have been studied: i) frequency interleaving of the low- and high-cardinality formats; ii) allocating the high cardinality format to the edge subcarriers and the low-cardinality format to the center subcarriers; and iii) applying the reverse of ii). Polarization interleaving can be applied also to FDHMF to equalize the power distribution in the channel spectrum and it has some beneficial effects in non-linear propagation. In[70] simulation results are reported for non-linear propagation, showing that FDHMF always achieve the same performance as TDHMF and in some cases can also overcome it.

7. CONSTELLATION SHAPING

Probabilistic (also called statistical) constellation shaping is another method to dynamically adapt the transmit rate of a standard QAM constellation. Using probabilistic shaping, lower-energy (*i.e.* inner) points of a constellation are transmitted with *higher* probability than the other points. This effectively reduces the total bit rate, allowing a potential increase in system reach. In Fig. 9 (left) it is shown the Mutual Information (MI) of probabilistic-shaped 64-QAM using different shaping factors λ, compared with uniform (*i.e.* regular) 64-QAM. The shaping factor λ defines the probability distribution of the points in the constellation,[71] and larger values of λ correspond to higher probabilities of inner points, *i.e.* the constellation is more *shaped*. It is defined such that the probability of transmitting a constellation point a_i is proportional to a Maxwell-Boltzmann distribution

$$\mathcal{P}(a_i) \propto e^{-\lambda |a_i|} \tag{4}$$

Two main characteristics of probabilistic shaping can be seen from Fig. 9 (left). Firstly, shaping reduces the maximum transmit rate (or, equivalently, the number of bits per symbol), represented by the MI *floor* for high values of SNR. This value corresponds to the constellation entropy

$$\mathcal{H}(C) = -\sum_i \mathcal{P}(a_i) \log_2 \mathcal{P}(a_i) \tag{5}$$

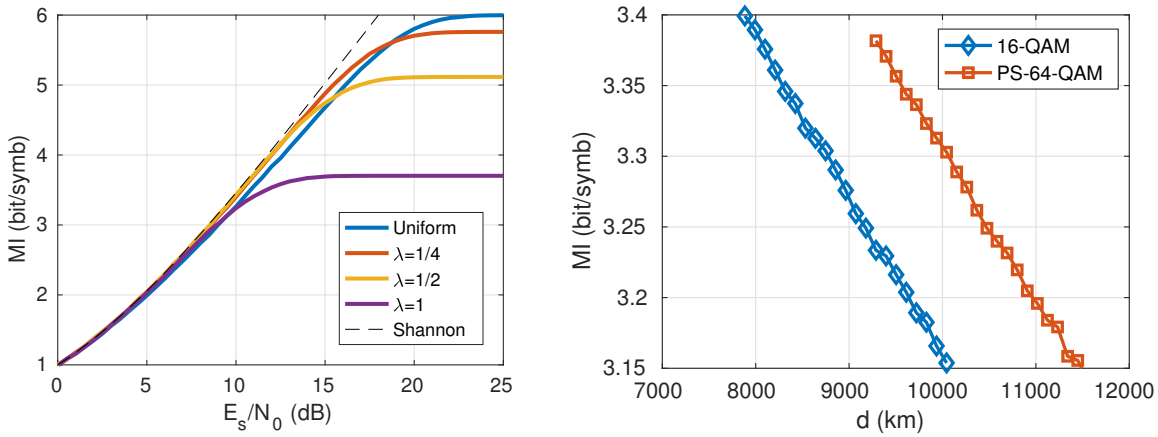

Figure 9. Mutual information vs. SNR for 64-QAM and PS-64-QAM with different shaping factors (*left*). Mutual information vs propagation distance for 16-QAM and PS-64-QAM with $\lambda = 0.879$ (*right*).

This allows to dynamically "tune" the transmit rate, keeping the same constellation (therefore, the same DSP algorithms). The second, more subtle, feature can be seen for low values of SNR, where PS constellations perform slightly better (*i.e.* higher MI at the same SNR) than uniform 64-QAM. The maximum SNR gain is approximately 0.7 dB. This is in accordance with the Shannon's channel capacity theorem, which states that channel capacity is achieved when the transmit constellation is Gaussian-distributed. The application of probabilistic shaping on a QAM constellation effectively makes the constellation shape closer to a Gaussian distribution, increasing its spectral efficiency.

Probabilistic shaping has been recently demonstrated in optical communications[71–74] for different QAM constellations, where there were shown significant improvements in system reach. For instance, in[71] by combining different spatially-coupled LDPC codes and probabilistic shaping on 64-QAM, the authors showed improvements up to 40% in system reach. In[73] the authors optimized the probability distribution of 16-QAM in presence of AWGN and constellation-dependent non-linear interference noise,[28,75] obtaining a 7% increase in system reach.

One of the main issues of probabilistic shaping is that probabilistic shaping algorithms, such as CCDM,[76] require error-free symbols at the receiver. Therefore, probabilistic shaping has to be applied before FEC, and, in order to keep the probability distribution, FEC has to be applied with *ad-hoc* schemes.[71] Unfortunately, these schemes do not allow an easy comparison between PS constellations and smaller uniformly-shaped constellations at the same net data rate, since they require either different FEC rates or constellation entropy.

A possible comparison between uniformly-shaped 16-QAM and PS-64-QAM is shown in Fig. 9 (right). The comparison has been carried out for $\lambda = 0.879$, which gives a constellation entropy (*i.e.* maximum transmit rate) exactly equal to 4 bits/symb, which is the same entropy of a 16-QAM constellation. These two different modulation formats have been transmitted over several spans of simulated PSCF ($\alpha = 0.167$ dB/km, $\beta_2 = -26.75$ ps/(Hz km), $\gamma = 0.75$ 1/(W km), $L_{\text{span}} = 108$ km). Transmit power has been optimized for each distance, and the transmit signal is a WDM comb of 11 channels, each at 32 Gbaud, separated by 50 GHz. As seen in the figure, the maximum reach gain at the same mutual information is 13.75%, which corresponds approximately to the SNR gain in back-to-back.

However, the highest SNR gains of probabilistic shaping are achieved for low values of MI, which corresponds to very high pre-FEC Symbol Error Rates (SERs). High values of SER represent a big challenge for blind DSP algorithms, such as adaptive equalizer and phase recovery. Therefore, at the time of writing, all the published experiments employ some sort of data-aided algorithm, either fully data-aided or pilot-tone based. The development of a fully blind (or with a small initial training sequence) adaptive equalization or carrier recovery algorithm, suitable for probabilistic shaping, is still an open issue, and it will be tackled in future research.

8. CONCLUSIONS

Bandwidth and capacity demand in metro, regional, and long-haul networks is increasing at several tens of percent per year, driven by video streaming, cloud computing, social media and mobile applications. To sustain the huge traffic demand, an upgrade of the widely deployed 100-Gbit/s long-haul optical systems is mandatory. In this paper, different available options to scale the per-channel bit-rate to 400 Gbit/s and beyond have been reviewed, analyzing the effects of symbol-rate increase and use of higher-order quadrature amplitude modulation (QAM) modulation formats, together with the use of DSP-enabled spectral shaping and advanced multiplexing technologies. In this analysis, tradeoffs between spectral efficiency and reach have also been highlighted.

Beside an increased capacity to keep up with traffic growth, next generation optical networks will require higher degrees of flexibility. A first step towards flexible networks that avoids the replacement of fibers and link components, like multiplexers, demultiplexers and ROADMs, consists in keeping fixed-grid dense wavelength-division multiplexing transmission on the installed equipment and introducing flexibility only through the replacement of transponders. In this paper, three advanced modulation techniques that have been recently proposed to increase the flexibility of the transponders have been reviewed, i.e. sub-carrier multiplexing, hybrid formats and constellation shaping. The limits and potential of these techniques have been highlighted, in terms of performance, complexity and flexibility.

ACKNOWLEDGMENTS

This work was partially supported by the European Commission through a Marie Sklodowska-Curie individual fellowship, project Flex-ON (653412), by the Cisco University Research Program Fund, a corporate advised fund of Silicon Valley Community Foundation and by Cisco Systems within a SRA contract. The authors gratefully acknowledge the support of NVIDIA Corporation, with the donation of the Tesla K40 GPU used for this research, and thank Andreas Bisplinghoff, Chris Fludger and Fabrizio Forghieri for useful discussions.

REFERENCES

[1] O. Gerstel et al., "Elastic optical networking: A new dawn for the optical layer?," *IEEE Signal Process. Mag.*, vol. 50, no. 2, pp. s12–s20, Feb. 2012.

[2] D.A. Morero, M.A. Castrillòn, A. Aguirre, M.R. Hueda, O.E. Agazzi, "Design Tradeoffs and Challenges in Practical Coherent Optical Transceiver Implementations," *J. of Lightw. Technol.*, vol. 34, no. 1, pp. 121–136, Jan. 1, 2016.

[3] M. Jinno et al., "Spectrum-efficient and scalable elastic optical path network: Architecture, benefits, and enabling technologies," *IEEE Signal Process. Mag.*, vol. 47, no. 11, pp. 66–72, Nov. 2009.

[4] X. Liu, S. Chandrasekhar, P. Winzer, "Digital signal processing techniques enabling multi-Tb/s superchannel transmission," *IEEE Signal Process. Mag.*, vol. 31, no. 2, pp. 16–24, Mar. 2014.

[5] W. Shieh, I. Djordjevic, *OFDM for Optical Communications*, Elsevier Academic Press, 2010.

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

[7] G. Bosco, "Spectrally Efficient Multiplexing: NYQUIST-WDM," book chapter in *Enabling Technologies for High Spectral-efficiency Coherent Optical Communication Networks*, John Wiley & Sons, Inc., New Jersey, pp. 123–156, 2016.

[8] P.J. Winzer, M. Pfennigbauer, R.-J. Essiambre, "Coherent Crosstalk in Ultradense WDM Systems," *J. Lightw. Technol.*, vol. 23, no. 4, pp. 1734–1744, Apr. 2005

[9] G. Bosco, "Spectral Shaping in Ultra-Dense WDM Systems: Optical vs. Electrical Approaches," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, USA, Paper OM3H.1 (2012).

[10] R. Schmogrow, P.C. Schindler, W. Freude, J. Leuthold, "Digital Pulse-Shaping for Spectrally Efficient and Flexible Coherent Optical Networks," in *Proc. Advanced Photonics for Communications*, San Diego, USA, Paper NM4D.3 (2012).

[11] M. Mazurczyk, "Spectral Shaping in Long Haul Optical Coherent Systems With High Spectral Efficiency," *J. Lightw. Technol.*, vol. 32, no. 16, pp. 2915–2924, Aug. 2014.

[12] E. Torrengo, et al., "Transoceanic PM-QPSK Terabit superchannel transmission experiments at Baud-rate subcarrier spacing," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Torino, Italy, paper We.7.C.2 (2010).

[13] J.-X. Cai, "100G Transmission Over Transoceanic Distance With High Spectral Efficiency and Large Capacity," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3845–3856, Dec. 2012.

[14] J. Fickers et al., "Design Rules for Pulse Shaping in PDM-QPSK and PDM-16QAM Nyquist-WDM Coherent Optical Transmission Systems," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Amsterdam, The Netherlands, paper We.1.C.2 (2012).

[15] C. Xie, G. Raybon, P.J. Winzer, "Hybrid 224-Gb/s and 112-Gb/s PDM-QPSK transmission at 50-GHz channel spacing over 1200-km dispersion-managed LEAF spans and 3 ROADMs," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, USA, paper PDPD2 (2011).

[16] X. Zhou, L. Nelson, R. Issac, P. Magill, B. Zhu, D. Peckham, "1200km Transmission of 50GHz spaced, 5504-Gb/s PDM-32- 64 hybrid QAM using Electrical and Optical Spectral Shaping," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, USA, paper OM2A.2 (2012).

[17] J.-X. Cai et al,"Transmission of 96100G pre-filtered PDM-RZ-QPSK channels with 300 % spectral efficiency over 10,608 km and 400 % spectral efficiency over 4368 km," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, San Diego, USA, paper PDPB10 (2010).

[18] M. Mazurczyk, "Optical spectral shaping and high spectral efficiency in long haul systems," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, San Francisco, USA, paper Tu3J.4 (2014).

[19] T. Duthel et al., "DAC enabled spectrally efficient CPQPSK at 28Gbaud," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Amsterdam, The Netherlands, paper Tu.4.A.4 (2012).

[20] J. Wang, C. Xie, Z. Pan, "Optimization of DSP to generate spectrally efficient 16QAM Nyquist-WDM signals," *Photon. Technol. Lett.*, vol. 25, no. 8 , pp. 772–775, Aug. 2013.

[21] S.O. Arik, K.-P. Ho, J.M. Kahn, "Optical network scaling: roles of spectral and spatial aggregation," *Opt. Exp.*, vol. 22, no. 24, pp. 29868–29887 (2014).

[22] A. Nespola et al., "1306-km 20x124.8-Gb/s PM-64QAM Transmission over PSCF with Net SEDP 11,300 (b·km)/s/Hz using 1.15 samp/symb DAC," *Opt. Exp.*, vol. 22, no. 1, pp. 1796–1805 (2014).

[23] J. Wang, C. Xie, Z. Pan, "Generation of Spectrally Efficient Nyquist-WDM QPSK Signals Using Digital FIR or FDE Filters at Transmitters," *J. Lightw. Technol.*, vol. 30, no. 23, pp. 3679–3686, Dec. 2012.

[24] R. Schmogrow et al., "Real-Time Nyquist Pulse Modulation Transmitter Generating Rectangular Shaped Spectra of 112 Gbit/s 16QAM Signals," in *Proc. of 2011 OSA Summer Topical Meeting on Signal Processing in Photonics Communications (SPPCom)*, Toronto, Canada, paper SPMA5 (2011).

[25] S.J. Savory, "Digital filters for coherent optical receivers," *Opt. Exp.*, vol. 16, no. 2, pp. 804–817 (2008).

[26] M. Salsi et al., "38.75 Tb/s Transmission Experiment over Transoceanic Distance," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, London, UK, Paper PD3.E.2 (2013).

[27] R. Kudo, T. Kobayashi, K. Ishihara, Y. Takatori, A. Sano, Y. Miyamoto, "Coherent optical single carrier transmission using overlap frequency domain equalization for long-haul optical systems," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3721–3728, Aug. 2009.

[28] A. Carena et al., "EGN model of non-linear fiber propagation," *Opt. Exp.*, vol. 22, no. 13, pp. 16335–16362 (2014).

[29] G. Bosco, P. Poggiolini, A. Carena, V. Curri, F. Forghieri, "Analytical results on channel capacity in uncompensated optical links with coherent detection," *Opt. Exp.*, vol. 19, no. 26, pp. B440–B451 (2011).

[30] W. Shieh, Y. Tang, "Ultrahigh-speed signal transmission over nonlinear and dispersive fiber optic channel: the multicarrier advantage," *IEEE Photonics J.*, vol. 2, no. 3, pp. 276–283, June 2010.

[31] C. Behrens, R.I. Killey, S.J. Savory, Ming Chen, P. Bayvel, "Nonlinear transmission performance of higher-order modulation formats," *IEEE Photon. Technol. Lett.*, vol. 23, no. 6, pp. 377–379, Mar. 2011.

[32] L.B. Du, A.J. Lowery, "Optimizing the subcarrier granularity of coherent optical communications systems," *Opt. Exp.*, vol. 19, no. 9, pp. 8079–8084 (2011).

[33] Q. Zhuge, B. Châtelain, D.V. Plant, "Comparison of intra-channel nonlinearity tolerance between reduced-guard-interval CO-OFDM systems and Nyquist single carrier systems," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, USA, paper OTh1B.3 (2012).

[34] A. Bononi, N. Rossi, P. Serena, "Performance dependence on channel baud-rate of coherent single-carrier WDM systems," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, London, UK, paper Th.1.D.5 (2013).

[35] N. Rossi, P. Serena, A. Bononi, "Symbol-rate dependence of dominant nonlinearity and reach in coherent WDM links," *J. Lightw. Technol.*, vol. 33, no. 14, pp. 3132–3143, Jul. 2015.

[36] P. Poggiolini et al., "Analytical and experimental results on system maximum reach increase through symbol rate optimization," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1872–1885, Apr. 2016.

[37] M. Qiu, Q. Zhuge, X. Xu, M. Chagnon, M. Morsy-Osman, D.V. Plant "Subcarrier multiplexing using DACs for fiber nonlinearity mitigation in coherent optical communication systems," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, San Francisco, USA, paper Tu3J.2 (2014).

[38] F. Yaman et al., "First quasi-single-mode transmission over transoceanic distance using few-mode fibers," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, USA, post-deadline paper Th5C.7 (2015).

[39] H. Nakashima et al., "Experimental investigation of nonlinear tolerance of subcarrier multiplexed signals with spectrum optimization," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Valencia, Spain, paper Mo.3.6.4 (2015).

[40] J. Fickers, A. Ghazisaeidi, M. Salsi, G. Charlet, P. Emplit, F. Horlinet, "Multicarrier offset-QAM for long-haul coherent optical communications," *J. of Ligthw. Technol.*, vol. 32, no. 4, pp. 4671–4678, Dec. 2014.

[41] A. Carbo, J. Renaudier, R. Rios-Mller, P. Tran, G. Charlet, "Experimental analysis of non linear tolerance dependency of multicarrier modulations versus bandwidth efficiency," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Valencia, Spain, paper Th.2.6.6, Sep. 2015.

[42] J.-X. Cai, M. Mazurczyk, O.V. Sinkin, M. Bolshtyansky, D.G. Foursa, A. Pilipetskii, "Experimental Study of Subcarrier Multiplexing Benefit in 74 nm Bandwidth Transmission up to 20,450 km," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Dusseldorf, Germany, paper W.3.D.4 (2016).

[43] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, F. Forghieri, "The GN model of fiber non-linear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 2014.

[44] R. Dar, M. Feder, A. Mecozzi, M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Exp.*, vol. 21, no. 22, pp. 25685–25699 (2013).

[45] P. Serena, A. Bononi, "A time-domain extended gaussian noise model," *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1459–1472, Apr. 2015.

[46] M. Qiu et al., "Digital subcarrier multiplexing for fiber nonlinearity mitigation in coherent optical communication systems," *Opt. Exp.*, vol. 22, no. 15, pp. 18770–18777 (2014).

[47] A. Nespola et al., "Experimental demonstration of fiber nonlinearity mitigation in a WDM multi-subcarrier coherent optical system," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Valencia, Spain (2015).

[48] A. Nespola et al., "Effectiveness of digital back-propagation and symbol-rate optimization in coherent WDM optical systems," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, USA (2016).

[49] F. Buchali, W. Idler, K. Schuh, L. Schmalen, T. Eriksson, G. Böcherer, P. Schulte, F. Steiner, "Study of electrical subband multiplexing at 54 GHz modulation bandwidth for 16QAM and probabilistically shaped 64QAM," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Valencia, Spain (2016).

[50] A.C. Meseguer, J. Renaudier, R. Rios-Muller, P. Tran, G. Charlet, "Impact of bandwidth efficiency in non-linear tolerance of multicarrier modulations," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1787–1792, Apr. 2016.

[51] A. Carbó, J. Renaudier, P. Tran, G. Charlet, "Experimental analysis of non linear tolerance dependency of multicarrier modulations versus number of WDM channels," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Anaheim, USA, paper Tu3A.6 (2016).

[52] G. Bosco, S.M. Bilal, A. Nespola, P. Poggiolini, F. Forghieri, "Impact of the transmitter IQ-skew in multi-subcarrier coherent optical systems," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, USA (2016).

[53] M. Malekiha, I. Tselniker, D.V. Plant, " Chromatic dispersion mitigation in long-haul fiber-optic communication networks by sub-band partitioning," *Opt. Exp.*, vol. 23, no. 25, pp. 32654–32663 (2015).

[54] F. Zhang, Q. Zhuge, M. Qiu, X. Xu, W. Wang, Y. Gao, M. Chagnon, D. Plant, "Advanced and low-complexity digital backpropagation for subcarrier-multiplexing systems," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, USA, paper. Th3D.4 (2015).

[55] F. Zhang, Q. Zhuge, M. Qiu, W. Wang, M. Chagnon, D.V. Plant, " XPM model-based digital backpropagation for subcarrier-multiplexing systems," *J. Lightw. Technol.*, vol. 33, no. 24, pp. 5140–5150, Dec. 2015.

[56] G. Wellbrock, "How Will Optical Transport Deal With Future Network Traffic Growth?", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Cannes, France, paper Th.1.2.1 (2014).

[57] J. Fischer, S. Alreesh, R. Elschner, F. Frey, M. Nolle, C. Schmidt-Langhorst, C. Schubert, "Bandwidth-variable transceivers based on four-dimensional modulation formats," *J. Lightw. Technol.*, vol. 32, no. 16, pp. 2886–2895, Aug. 2014.

[58] J. Renaudier et al., "Experimental transmission of Nyquist pulse shaped 4-D coded modulation using dual polarization 16QAM set-partitioning schemes at 28 Gbaud," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Anaheim, USA, paper OTu3B.1 (2013).

[59] L. Beygi, E. Agrell, J. Kahn, M. Karlsson, "Rate-adaptive coded modulation for fiber-optic communications," J. Lightw. Technol., vol. 32, no. 2, pp. 333–343, Jan. 2014.

[60] E.L.T. de Gabory, T. Nakamura, H. Noguchi, W. Maeda, S. Fujita, J. Abe, K. Fukuchi, "Experimental demonstration of the improvement of system sensitivity using multiple state trellis coded optical modulation with QPSK and 16QAM constellations," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, USA, paper W3K.3 (2015).

[61] G.H. Gho, J.M. Kahn, "Rate-adaptive modulation and low-density parity-check coding for optical fiber transmission systems," *J. Opt. Commun. Netw.*, vol. 4, no. 10, pp. 760–768, Oct. 2012.

[62] W.-R. Peng, I. Morita, H. Tanaka, "Hybrid QAM transmission techniques for single-carrier ultra-dense WDM systems," in *Proc. 16th Opto-Electronics and Communications Conf. (OECC)*, Taiwan, pp. 824–825 (2011).

[63] D. van den Borne, S.L. Jansen, "Dynamic capacity optimization using flexi-rate transceiver technology," in *Proc. 17th Opto-Electronics and Communications Conf. (OECC)*, Busan (Korea), pp. 769–770 (2012).

[64] V. Curri, A. Carena, P. Poggiolini, R. Cigliutti, F. Forghieri, C.R. Fludger, T. Kupfer, "Time-division hybrid modulation formats: Tx operation strategies and countermeasures to nonlinear propagation," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, San Francisco, USA, paper Tu3A.2 (2014).

[65] X. Zhou, L.E. Nelson, "400G WDM transmission on the 50 GHz grid for future optical networks," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3779–3792, Dec. 2012.

[66] Q. Zhuge, M. Morsy-Osman, X. Xu, M. Chagnon, M. Qiu, D. Plant, "Spectral efficiency-adaptive optical transmission using time domain hybrid QAM for agile optical networks," *J. Lightw. Technol.*, vol. 31, no. 15, pp. 2621–2628, Aug. 2013.

[67] R. Rios-Muller et al., "Experimental comparison between hybrid-QPSK/8QAM and 4D-32SP-16QAM formats at 31.2 GBaud using Nyquist pulse shaping," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, London, UK, paper Th.2.D.2 (2013).

[68] R. Li, V. Curri, A. Carena, "Bit-rate maximization for elastic transponders operating in WDM uncompensated amplified links," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Anaheim, USA, paper Th2A.45 (2016).

[69] F.P. Guiomar, R. Li, C.R.S. Fludger, A. Carena, V. Curri, "Hybrid Modulation Formats Enabling Elastic Fixed-Grid Optical Networks," *J. Opt. Commun. Netw.*, vol. 8, no. 7, pp. A92–A100, Jul. 2016.

[70] F.P. Guiomar, A. Carena, "Achieving Fine Bit-Rate Granularity with Hybrid Subcarrier Modulation," in *Proc. Signal Processing in Photonic Communications (SPPCom)*, paper Sp.W3F.2 (2016).

[71] F. Buchali et al., "Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration," *J. Lightw. Technol.*, vol. 34, no. 7, pp. 1599–1609, Apr. 2016.

[72] M.P. Yankov et al., "Constellation shaping for fiber-optic channels with QAM and high spectral efficiency," *IEEE Photon. Technol. Lett.*, vol. 26, no. 23, pp. 2407–2410, Dec. 2014.

[73] F.R. Kschischang et al., "Probabilistic 16-QAM Shaping in WDM Systems," *J. Lightw. Technol.*, vol. 34, no. 18, pp. 4285–4292, Sep. 2016.

[74] S. Chandrasekhar et al., "High-spectral-efficiency transmission of PDM 256-QAM with Parallel Probabilistic Shaping at Record Rate-Reach Trade-offs", in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Dusseldorf, Germany, paper. Th.3.C.1 (2016).

[75] R. Dar et al., "On shaping gain in the nonlinear fiber-optic channel," *IEEE International Symposium on Information Theory*, Hawaii, USA, pp. 2794–2798 (2014).

[76] P. Schulte et al., "Constant composition distribution matching," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 430–434, Jan. 2016.