Next-generation long-haul optical links: Higher spectral efficiency through time-frequency packing

Next-Generation Long-Haul Optical Links: Higher Spectral Efficiency through Time-Frequency Packing

Giulio Colavolpe and Tommaso Foggi
University of Parma, Dipartimento di Ingegneria dell'Informazione, and CNIT Research Unit, I-43124 Parma, Italy

***Abstract*—We consider realistic long-haul optical links, where nonlinear effects represent the main impairment, and investigate the application of time-frequency packing with low-order constellations as a the most viable solution to increase the spectral efficiency. We will see that this technique allows to overcome the relevant theoretical and technological issues related to this spectral efficiency increase and is more effective than the simple adoption of high-order modulation formats which are more sensitive to nonlinear effects.**

I. Introduction

The ever-growing bandwidth demand on internet and data networks has pushed the research in the field of optical communications towards more sophisticated transmission techniques [1]–[3]. Although this field has been characterized for decades by simple and poorly-performing modulation formats and transmission systems, since bandwidth requirements were met with both cost-effective and technological affordable solutions, nowadays optical communication channels have rapidly passed from 10 Gbps to 100 Gbps, whereas the next challenging step will lead to 1 Tbps. Among the more severe limitations involved in such a system upgrade, the technological and practical issues of processing high data rates on a single channel, and the optical channel impairments related to the required transmit power (i.e., the nonlinear effects) are the most prominent. Since a higher capacity will hardly be reached with single-channel transmissions, many different solutions and related implementation techniques have being devised and proposed, in order to exploit at best the optical channel capacity under current technological costraints. Basically, all the proposed solutions, irrespectively of the particular transmission technique, are based on multi-carrier transmissions or so called "superchannels" [4], which means that the goal capacity is reached binding up as many single-channels together as necessary, in an efficient way. The technique investigated in this paper belongs to this class and allows an effective and complexity-limited "packing" of the channels with the aim of greatly improving the spectral efficiency of the transmission, also considering how optical medium impairments affect the transmitted signals.

As in most digital communication systems, orthogonal signaling is the paradigm traditionally employed also in the design of long-haul optical systems. This paradigm consists of ensuring the absence of intersymbol interference (ISI) and, in multi-carrier scenarios, also the absence of inter-carrier interference (ICI), i.e., the absence of interference between adjacent channels. As an example, in single-carrier coherent optical systems, possibly employing polarization multiplexing (PM), given a conventional transmitter, with Mach-Zehnder (MZ) modulators and return to zero (RZ) or non-return to zero (NRZ) shaping pulses, when group velocity dispersion (GVD) and polarization mode dispersion (PMD) are effectively compensated for and nonlinear effects are limited, proper filtering and sampling at the receiver ensure that even a symbol-by-symbol detector enables an almost-optimal performance since ISI and ICI are very limited [5]. Example of multi-carrier transmission systems based on this paradigm are represented by orthogonal frequency-division multiplexing (OFDM) [6] (see also [7] and references therein) and Nyquist wavelength-division multiplexing (WDM) [8] (and similar bandwidth narrowing techniques such as [9]), where the use of proper shaping pulses allows to remove, at least theoretically, both ISI and ICI without using guard bands and thus without wasting resources. On the other hand, when practical transmit or receive filters are considered and nonlinear effects come into play, orthogonality is no more guaranteed and an unwanted interference appears. In addition, in these orthogonal signaling systems, the spectral efficiency (SE) can be improved only by increasing the constellation cardinality, thus employing modulation formats more sensitive to nonlinear effects and crosstalk.

An alternative paradigm to increase the SE is represented by time-frequency packing (TFP) [10]–[12]. In this case, low-order modulations, such as quaternary phase shift keying (QPSK), are employed but the spacing between two adjacent pulses in the time domain (i.e., the symbol interval) is reduced well below that corresponding to the Nyquist rate. Similarly, the frequency separation between two adjacent channels can be also reduced, with the aim of maximizing the *achievable spectral efficiency*, which is thus used as a performance measure instead of the minimum Euclidean distance (as in more classical *faster-than-Nyquist* signaling schemes, see [11] and references therein). In addition, rather than the optimal receiver, a complexity-constrained detector is considered. In other words, controlled ISI and ICI are introduced and partially coped with at the receiver through a single-channel maximum a posteriori (MAP) symbol detector, which is designed to take into account only a limited amount of ISI or, through a suboptimal multiuser detector (MUD), which enables the joint processing of multiple sub-channels and thus coping with a limited amount of ICI (and possibly ISI).

Advanced signal processing plays a key role in TFP systems since, besides the fundamental working principles which entails decoding and soft detection techniques, significant performance improvements derive from a skilful combination of pulse shaping, coding, the adoption of a proper linear filtering [13], of a proper MAP symbol detection strategy, and, possibly, of a proper multiuser processing. Improving the SE without increasing the constellation cardinality can be considerably convenient since it is well known that low-order constellations are more robust to channel impairments such as nonlinearities, whose effects are already increased by the higher transmitted

978-1-4799-1353-4/13/$31.00 ©2013 IEEE

power needed to obtain higher SE values, and phase noise. In this paper, the computation of the SE is performed by resorting to the simulation-based method described in [14], which allows to also take into account nonlinear effects.

The remainder of this paper is organized as follows. The system model is described in Section II. The framework that we use to evaluate the SE is then described in Section III, whereas the adopted detectors are described in Section IV. Numerical results are reported in Section V with a detailed comparison with alternative approaches in the literature, possibly based on the paradigm of orthogonal signaling. Finally, some conclusions are drawn in Section VI.

II. Preliminaries and System Model

Let us consider a multi-carrier system over an additive white Gaussian noise (AWGN) channel, where equally-spaced adjacent carriers are associated to the same linear modulation format and shaping pulse $p(t)$. The complex envelope of the transmitted signal can be expressed as[1]

$$s(t) = \sum_k \sum_\ell x_k^{(\ell)} p(t - kT - \tau^{(\ell)}) e^{j(2\pi \ell F t + \theta^{(\ell)})} \quad (1)$$

where T is the symbol interval, $x_k^{(\ell)}$ the symbol transmitted over the ℓth carrier during the nth symbol interval, $\tau^{(\ell)}$ and $\theta^{(\ell)}$ the delay and the initial phase of the ℓth carrier, respectively, and F the frequency spacing between two adjacent carriers.

In such a scenario, the capacity-achieving distribution of the transmitted symbols is Gaussian. Hence, in order to maximize the system spectral efficiency, independent and uniformly distributed (i.u.d.) Gaussian symbols $\{x_k^{(\ell)}\}$ must be employed along with a shaping pulse $p(t)$ having a rectangular spectrum (sinc pulse) with bandwidth $B = 1/2T$ and a frequency spacing $F = 1/T$, i.e., no guard band is employed between two adjacent carriers. No ISI or ICI occur since orthogonal signaling is employed.

Practical systems necessarily deviate from this paradigm. First of all, the transmitted symbols $\{x_k^{(\ell)}\}$ are not Gaussian but usually belong to a properly normalized zero-mean M-ary complex constellation χ. Under these conditions, instead of trying to approach as close as possible the impractical condition of having a shaping pulse with rectangular spectrum and to reduce as much as possible the guard band, as in Nyquist-WDM [8] systems, TFP technique intentionally introduces both ISI and ICI to improve the spectral efficiency [10],[11]. In other words, for a given shaping pulse the symbol time T and the frequency spacing F are properly optimized, as explained in Section III, to maximize the spectral efficiency intentionally violating the orthogonal signaling paradigm.

Let us now consider a realistic optical system. In this case, the possibility to generate a transmitted signal with expression (1) is strictly related to the availability of a linear modulator. In other words, let us consider the transmitted signal associated to the carrier for $\ell = 0$. If pulse $p(t)$ has support larger than T, this signal cannot be directly generated through a MZ modulator unless it is properly linearized. This is due to the nonlinear transfer function of the MZ modulator between the electrical signal at its input and the optical signal at its output. We could, however, use a MZ modulator to generate a linearly modulated signal with shaping pulse having support at most T and then "stretch" the transmitted pulses through an optical filter, thus obtaining an effect similar to that obtained with time packing. Hence, in this case, time-packing is not an available option but we have a viable surrogate. The degrees of freedom are thus the frequency spacing F and the bandwidth B of the optical filter used at the MZ output [11], which in the present analysis is always a 4th-order Gaussian filter. Fig. 1 shows a schematic of the considered transmission system, irrespectively of the constellation size. Blocks related to the receiver will be explained later.

We consider optical channels impaired by GVD and PMD, and in particular uncompensated links where chromatic dispersion compensation is only performed with a fixed-tap equalizer in the electrical domain at the receive side. We also account for nonlinear effects, as it will be explained in Section V, where a description of the simulated optical links will be provided. We will compare systems based on TFP and employing quaternary constellations with other known systems, based on higher-order modulations, which show good results in terms of spectral efficiency, namely Nyquist-WDM [8] and receiver-side duobinary shaping [9]. In the latter system, an electrical two-tap filter is used to force a duobinary shape to the received signal, thus limiting the ICI, at the expense of employing a MAP sequence (or symbol) detector to cope with the introduced ISI. No comparison will be performed with OFDM systems since they present a few drawbacks, from an implementation point of view, for optical links and are also less efficient than the systems considered here [7].

III. Spectral Efficiency Optimization

We describe the framework used to evaluate the performance limits of the optical systems considered in this paper and to perform the optimization of the optical filter bandwidth and frequency spacing. We perform this investigation by constraining the complexity of the employed receiver. In particular, we assume that parallel independent detectors are employed, one for each carrier and each polarization. For this reason, without loss of generality we only consider the detection of symbols $\mathbf{x}^{(0)} = \{x_k^{(0)}\}$ of carrier with $\ell = 0$.[2] In addition, we also consider low-complexity receivers taking into account only a portion of the actual channel memory. Under these constraints, we compute the information rate (IR), i.e., the average mutual information when the channel inputs are independent and uniformly distributed (i.u.d.) random variables belonging to the considered constellation. Provided that a proper *auxiliary channel* can be defined for which the adopted low-complexity receiver is optimal, the computed IR represents an achievable lower bound of the IR of the actual channel, according to mismatched detection [14], [15].

Denoting by $\mathbf{y}^{(0)}$ a proper discrete-time received sequence used for detection of $\mathbf{x}^{(0)}$, the achievable IR, measured in bit

[1] In the following, we will consider the adoption of polarization multiplexing. In this case, $s(t)$ is the signal transmitted on one state of polarization.

[2] Assuming a system with an infinite number of carriers, the results do not depend on the specific considered carrier.

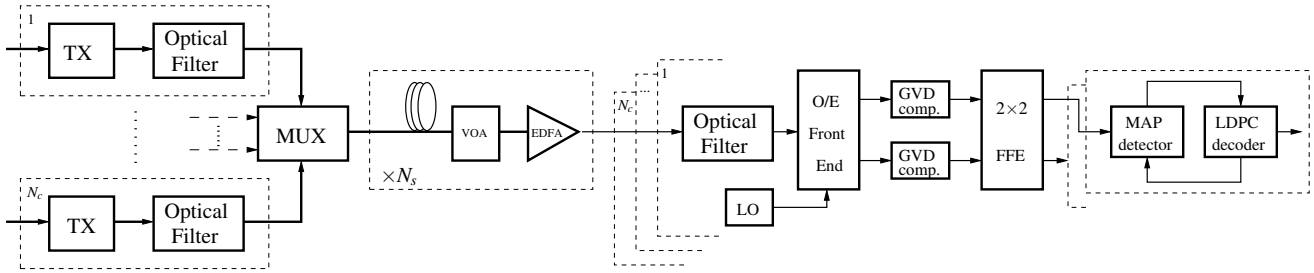

Figure 1. Schematic of the transmission system, where N_c is the number of channels, N_s the number of link spans, VOA a *variable optical attenuator*, EDFA an *erbium-doped fiber amplifier*, LO a *local oscillator*, O/E Front End the *opto-electronic front end* [5], GVD comp. a *linear equalizer* aimed at compensating for the chromatic dispersion, the 2×2 DD-FFE is a two-dimensional *decision-directed feed-forward equalizer*, and the iterative detection/decoding between a proper detector and the low-density parity-check (LDPC) decoder is performed on blocks of symbols whose length n depends, as we will see later, on the constellation size.

per channel use, can be obtained as

$$I(\mathbf{x}^{(0)};\mathbf{y}^{(0)}) = \lim_{K\to\infty} \frac{1}{K} E\left\{ \log \frac{p(\mathbf{y}^{(0)}|\mathbf{x}^{(0)})}{p(\mathbf{y}^{(0)})} \right\}, \qquad (2)$$

where K is the number of transmitted symbols. The probability density functions $p(\mathbf{y}^{(0)}|\mathbf{x}^{(0)})$ and $p(\mathbf{y}^{(0)})$ are computed by using the optimal MAP symbol detector for the auxiliary channel, while the expectation in (2) is with respect to the input and output sequences generated *according to the actual channel model* [14]. We use F as a measure of the bandwidth of each single carrier. The achievable SE is thus

$$\eta = \frac{1}{FT} I(\mathbf{x}^{(0)};\mathbf{y}^{(0)}) \quad \text{[b/s/Hz]}. \qquad (3)$$

The aim of the proposed technique is to find the values of F and the bandwidth B of the optical filter after the MZ modulator providing, for each value of the signal-to-noise ratio (SNR), the maximum value of SE achievable by that particular receiver, optimal for the considered specific auxiliary channel. Namely, we compute

$$\eta_{\mathrm{M}} = \max_{F,B>0} \eta(F,B)\,. \qquad (4)$$

Typically, the dependence on the SNR value is not critical, in the sense that we can identify two or at most three SNR regions for which the optimal spacings practically have the same value.

For fair comparisons in terms of SE, we need a proper definition of the SNR. We define the SNR as the ratio P/N between the signal power and the noise power (in the considered bandwidth). Denoting by K the number of carriers, P/N can as written as, under the assumption of a large number of carriers to avoid boundary effect,

$$\frac{P}{N} = \lim_{K\to\infty} \frac{KP_C}{2N_0((K-1)F+B)} = \frac{P_C}{2N_0F}\,, \qquad (5)$$

where P_C is the power for each carrier and $N_0/2$ is the two-sided power spectral density of the amplified spontaneous emission (ASE) noise per polarization as if the channel were linear. P_C is independent of the bandwidth B. The SNR definition as given in (5) is independent of the transmit waveform and its parameters. This provides a common measure to compare the performance of different solutions in a fair manner.

IV. Auxiliary Channel Models

The system model described in Section III is representative of the considered scenario and has been employed in the information-theoretic analysis and in the simulations results. In this section, we describe the employed auxiliary channel models and the corresponding optimal MAP symbol detectors. As explained in Section III, they are used to compute lower bounds on the SE for the considered channel [14]. Since these lower bounds are achievable by those receivers, we will say that the computed lower bounds are the SE values of the considered channel when those receivers are employed. As described below, the considered receivers are those often used in practical systems.

First of all, we assume that no compensation of the nonlinear effects is foreseen at the receiver. Several works propose the adoption of digital backpropagation [16] as the most effective nonlinearities compensation method, but on the other hand the complexity entailed in the considered multichannel systems becomes impractical. In addition, the presence or absence of digital backpropagation only affects the maximum values of achievable SE but not our conclusions. Hence, the considered receivers are optimal for an auxiliary channel where nonlinearities are absent. As far as GVD and PMD are concerned, it is assumed that they are perfectly compensated (i.e., they are absent in the corresponding auxiliary channel). As known, in the absence of nonlinear effects perfect compensation is possible through a proper two-dimensional equalizer [5].[3] Under these assumptions, the above mentioned independent receivers, one for each carrier and each polarization, have to also take into account the ISI intentionally or accidentally introduced in the system. As an example, in the case of TFP or the adoption of the receiver-side duobinary shaping [9], the receiver takes into account (a portion of) the ISI intentionally introduced. In the case of Nyquist-WDM systems, a receiver coping with the unwanted ISI deriving from the adopted practical filters and shaping pulses is considered instead.

[3] In Fig. 1, this equalizer is represented as the cascade of two fixed one-dimensional equalizers, one for each polarization, aimed at compensating for the GVD, and a short adaptive two-dimensional equalizer to cope with the PMD.

As mentioned, for the proposed TFP technique, the carrier spacing and the bandwidth of the transmit optical filter are optimized to maximize the spectral efficiency. For a fair comparison, in the case of Nyquist-WDM systems we also optimize the transmit optical filter bandwidth and the channel spacing, with the constraint that the spacing is not lower than the Nyquist bandwidth, i.e., $1/T$ (otherwise we fall in the domain of the TFP technique). Under these conditions the memory at the receiver is usually limited to at most $L = 1$ interfering symbol. On the other hand, the memory at the receiver in the case of the technique described in [9] is, by definition, $L = 1$. The memory introduced by the TFP technique is, instead, potentially very large. To limit the receiver complexity with a limited performance degradation, we apply a channel shortening (CS) technique [17]. In fact, when the memory of the channel is too large to be taken into account by a full complexity detector, an excellent performance can be achieved by properly filtering the received signal before adopting a reduced-state detector [17]. A very effective CS technique for general linear channels is described in [13].

In the case of adoption of CS, we are looking for an auxiliary channel and the corresponding optimal MAP symbol detector. As mentioned, in the auxiliary channel model we do not consider nonlinearities whereas GVD and PMD are assumed perfectly compensated for. In addition, independent receivers, one for each carrier and each polarization are considered here. Hence, the receiver assumes that, apart from AWGN, only one carrier is present. Without loss of generality we cansider the receiver for carrier with $\ell = 0$. A set of sufficient statistics $\mathbf{y}^{(0)}$ can thus be obtained by sampling the output of a matched filter (MF). The kth element of $\mathbf{y}^{(0)}$, under the above mentioned assumption that only carrier with $\ell = 0$ is present, reads

$$y_k^{(0)} = \sum_i x_{k-i}^{(0)} g_i + n_k \,,$$

where

$$g_i = \int h(t) h^*(t - iT) \mathrm{d}t$$

being $h(t)$ the convolution of the shaping pulse $p(t)$ after the MZ modularor and the optical transmit filter impulse response, and n_k a Gaussian process with $\mathrm{E}\{n_{k+i} n_k^*\} = 2N_0 g_i$. Vector $\mathbf{y}^{(0)}$ can be written as

$$\mathbf{y}^{(0)} = \mathbf{G}\mathbf{x}^{(0)} + \mathbf{n} \tag{6}$$

where $\mathbf{G}$ is a Toeplitz matrix obtained from the sequence $\{g_i\}$ whereas $\mathbf{n}$ is a vector collecting the noise samples. According to the CS approach, the considered auxiliary channel is based on the following channel law [13]

$$p(\mathbf{y}^{(0)}|\mathbf{x}^{(0)}) \propto \exp\left(2\Re(\mathbf{y}^{(0)H}\mathbf{H}^r\mathbf{x}^{(0)}) - \mathbf{x}^{(0)H}\mathbf{G}^r\mathbf{x}^{(0)}\right) , \tag{7}$$

where $(\cdot)^H$ denotes transpose conjugate, whereas $\mathbf{H}^r$ and $\mathbf{G}^r$ are Toeplitz matrices obtained from proper sequences $\{h_i^r\}$ and $\{g_i^r\}$, and are known as *channel shortener* and *target response*, respectively [13]. Matrix $\mathbf{H}^r$ represents a linear filtering of the sufficient statistics (6), and $\mathbf{G}^r$ is the ISI to be set at the detector (different from the actual ISI) [13]. In (7), the noise variance has been absorbed into the two matrices. In order to reduce the complexity, we constrain the target response used at the receiver to

$$g_i^r = 0 \quad |i| > L_r \,, \tag{8}$$

which implies that the memory of the detector is L_r instead of the true memory L of the channel. The CS technique finds a closed form of the optimal $\{h_i^r\}$ and $\{g_i^r\}$ which maximize the achievable IR (2). If the memory L_r is larger than or equal to the actual channel memory the trivial solution is $\mathbf{H}^r = \mathbf{I}/2N_0$ and $\mathbf{G}^r = \mathbf{G}/2N_0$, where $\mathbf{I}$ is the identity matrix. Interestingly, when $L_r = 0$ the optimal channel shortener becomes a minimum mean square error (MMSE) feedforward equalizer. With the constraint (8), the optimal MAP symbol detector for the auxiliary channel with law (7) is described in [18].

V. Simulation Results

We here report the optimal spectral efficiency η_{M} as a function of P/N, estimated as if the channel were linear, for different polarization-multiplexed systems. We assume perfect synchronization as we are interested in the evaluation of achievable bounds for the spectral efficiency. The employed shaping pulses are those resulting from the use of RZ pulses with duty cycle 50% (or NRZ in case of Nyquist-WDM systems with non-optimized bandwidths, see later), a MZ modulator, and a 4th-order Gaussian optical transmit filter with 3-dB optical filter bandwidth B (specified later). The considered modulation formats are QPSK and 16/64/256-ary quadrature amplitude modulations (QAMs). In all numerical results, we considered systems with 8 carrier (subchannels) at 140 Gbps each, irrespectively of the employed modulation format (thus, the baud rate changes for each format). It can be noticed that, since the bandwidth of each subchannel is highly reduced by filtering and multilevel modulations are considered, the required sampling rate is always within the state-of-the-art technology, i.e, well below 80 Gsample/s. At the receiver side, after the optical receive filter, whose 3-dB optical filter bandwidth B_R will be specified later, two non-adaptive one-dimensional equalizers and a two-dimensional (2-D) decision-directed (DD) adaptive feedforward equalizer (FFE) with 25 taps process the signals received over two orthogonal states of polarization to compensate for GVD and PMD and to complete (along with the optical filter) the implementation of the MF [5].[4] In the case of systems employing receiver-side duobinary shaping, a further digital filter is present to perform the required shaping [9]. The output is provided to the MAP symbol detectors (on for each carrier and each polarization) which iteratively exchange information with the decoders for a maximum of 50 iterations.

All the considered constellations can be viewed, with a proper rotation, as two independent signals transmitted over the in-phase and quadrature components, respectively. Hence, at the receiver side, we may use two identical and independent detectors, one working on the in-phase and the other one on the quadrature component. This is beneficial in case of adoption of a MAP detector. In fact, when L interfering symbols are

[4] It other words, the FFE taps are updated using the MF output as target response, so that the equalizer does not remove the ISI induced by narrow filtering.

Table I. SMF LENGTHS FOR EACH SPAN IN THE SIMULATED OPTICAL LINK.

span #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SMF (km)	70.8	75.5	55.1	52.1	40.1	67.	53.2	50	80.3	79.1	53.6	75.1	90.3	54.2	99.4

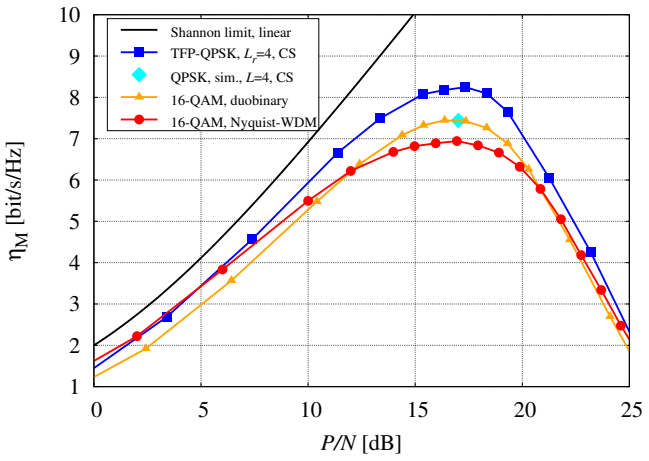

Figure 2. Maximum achievable spectral efficiency on the considered uncompensated link, for TFP-QPSK, 16-QAM with Nyquist-WDM spacing, and 16-QAM with receiver-side duobinary shaping. The Shannon limit in the linear regime is also reported as a reference, along with the simulated performance of a TFP-QPSK system employing a rate-4/5 LDPC code.

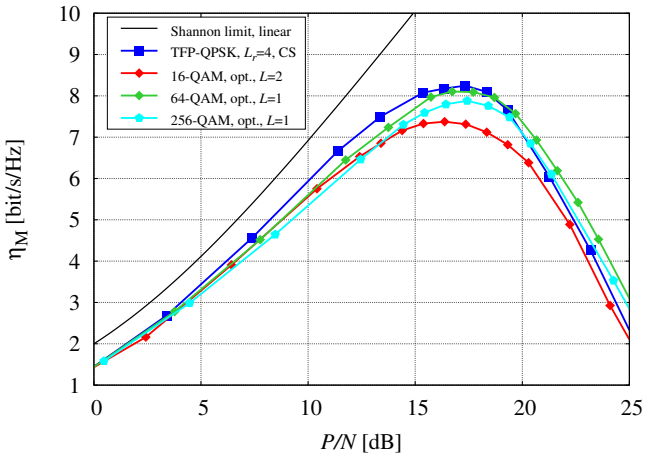

Figure 3. Maximum achievable spectral efficiency on the considered uncompensated link for TFP-QPSK, and 16/64/256-QAM with optimized bandwidths.

taken into account, we have two detectors working on a trellis with $(\sqrt{M})^L$ states instead of a single detector working on a trellis with M^L states. Hence, for a given complexity, a larger memory can be taken into account.

We consider an existing link of standard single-mode fiber (SMF), whose spans are described in Table I. The fiber dispersion is 16.63 $\mathrm{ps/nm/km}$, the PMD parameter is 0.1 $\mathrm{ps}/\sqrt{\mathrm{km}}$, the attenuation is 0.23 $\mathrm{dB/km}$, the nonlinear index γ is equal to 1.3 $\mathrm{W^{-1}km^{-1}}$, and the noise figure of all amplifiers is equal to 6 dB. For this link, Fig. 2 shows the achievable spectral efficiency η_M for a TFP system employing QPSK (TFP-QPSK), a Nyquist-WDM system using a 16-QAM, and a system employing receiver-side duobinary shaping, still using 16-QAM. In the case of TFP, the frequency spacing F and the optical 3-dB filter bandwidth B have been optimized for each value of P/N, and it is $B_R = B$. For the Nyquist-WDM system we used $B = 1.1/T$ and $B_R = 1/T$ as suggested in [8] whereas, for receiver-side duobinary shaping we used $B = B_R = 1/T$ as in [9]. The Shannon Limit [19] in the absence of nonlinearities is also shown for comparison. It can be observed that the TFP-QPSK outperforms the other two systems in spite of their use of a higher-order modulation.

These information-theoretic results can be approached by using proper coding schemes. As an example, we simulated the bit-error rate (BER) a TFP-QPSK system using $B = 0.325/T$, $F = 0.43/T$, and employing a rate-4/5 low-density parity-check (LDPC) code having codewords of 64800 bits. Assuming a reference for the BER of 10^{-7}, the performance of this system has been reported in Fig. 2 in the Shannon plane. It may be observed that, despite the lack of an optimization in the code design, we have a loss of less than 1 b/s/Hz from the theoretical results, obtaining a SE of 7.5 b/s/Hz. The observed loss is mainly due to the presence of nonlinear effects which require a careful redesign of the codes (we used a good code for the linear channel). We would like to mention that on this link an experimental field trial demonstration has been also conducted, reaching a spectral efficiency of more than 5 b/s/Hz despite the constraint to use (poorly performing) 1st-order Gaussian filters, not flexible and penalizing frequency grids, neighbouring data channels with commercial traffic, and no time left to perform the necessary system optimizations. The relevant results will be reported in a later paper.

Fig. 3 reports the same TFP-QPSK curve in Fig. 2 and a comparison with Nyquist-WDM systems employing high-order QAM formats, namely, 16-QAM, 64-QAM, and 256-QAM. For a fair comparison, this time we used RZ pulses with duty cycle 50% for Nyquist-WDM systems also. For these systems we also optimized, from a spectral efficiency point of view, the bandwidths B and B_R of the optical filters and used $F = B$ (as mentioned, we imposed the constraint $F \geq 1/T$). The optimized values corresponding to the peak of the η_M curves are: $B = 1/T$ and $B_R = 0.7/T$ for 16-QAM, $B = 1.1/T$ and $B_R = 0.85/T$ for 64-QAM, and $B = 1.25/T$ and $B_R = 0.9/T$ for 256-QAM.[5] At the receiver, we used a MAP symbol detector taking into account a memory $L = 2$ for 16-QAM and, for complexity reasons, a memory $L = 1$ for 64- and 256-QAM, although a value $L = 1$ would be sufficient in all cases, i.e., there is no significant difference for $L > 1$. Again, TFP-QPSK outperforms all considered M-QAM formats. Although, in principle, 16-QAM, 64-QAM, and 256-QAM with polarization multiplexing could achieve

[5]The fact that we found $B_R < B$ and not $B = B_R$ is related to the fact that we have $F = B$ and the contraint $F \geq 1/T$. Hence, in this case is more convenient to reduce B_R with respect to B.

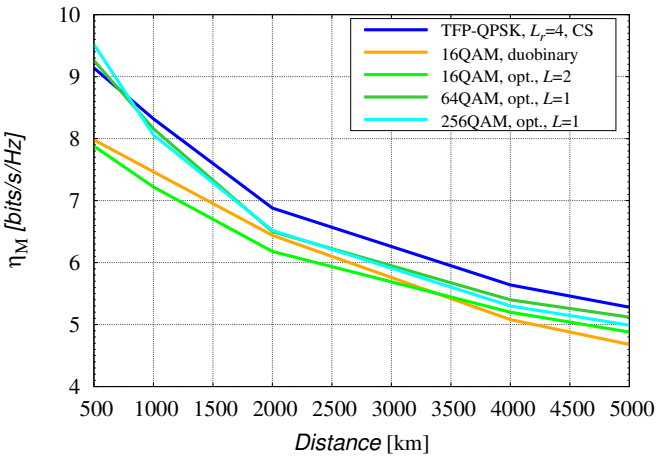

Figure 4. Maximum achievable spectral efficiency of of the considered systems, as a function of distance.

spectral efficiency values (with sinc pulses) of 8, 12, and 16 b/s/Hz, respectively, they would be reached, in the linear regime, only for higher values of P/N, whereas the optimum launch power corresponds to an SNR value which privileges TFP-QPSK.

Finally, in Fig. 4 we show the maximum spectral efficiency η_M at the optimal launch power, as a function of distance on a simplified (which means identical spans of length 100 km, the same amplifier noise figure as before, no polarization mode dispersion) uncompensated link, for all so far considered systems. TFP-QPSK shows to reach the highest spectral efficiency in the considered range 1000-5000 km.

VI. Conclusions

We compared different techniques to improve the spectral efficiency of realistic long-haul optical systems. The most promising solution results to be that based on time-frequency packing which is related to the use of narrow optical filtering and a tight packing of the carriers in frequency, giving up the signal orthogonality in the time and frequency domains, and on the adoption of detectors able to cope with the interference intentionally introduced in the system. This solution outperforms other solutions proposed in the literature and based on higher-order modulations, showing that, when nonlinear effects are present, the spectral efficiency cannot be trivially increased by increasing the modulation order. We also reported simulation results on a modulation and coding format which, on a realistic optical link, reaches a spectral efficiency of 7.5 b/s/Hz with a polarization-multiplexed time-frequency-packed QPSK, with a loss of less than 1 b/s/Hz from the information-theoretic results.

Acknowledgment

This work was supported in part by Ericsson and by the Italian Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) under the FIRB project Coherent Terabit Optical Networks (COTONE).

References

[1] S. Bigo, "Coherent optical long-haul system design," in *Proc. Optical Fiber Commun. Conf.*, (Los Angeles, CA), 2012. paper OTh3A.1.

[2] R. Essiambre and R. W. Tkach, "Capacity trends and limits of optical communication networks," *Proceedings of the IEEE*, vol. 100, pp. 1035–1055, May 2012.

[3] S. Chandrasekhar and X. Liu, "Enabling components for future high-speed coherent communication systems," in *Proc. Optical Fiber Commun. Conf. (OFC'09)*, (Los Angeles, CA, USA), Mar. 2011. paper OMU5.

[4] S. L. Jansen, "Multi-carrier approaches for next-generation transmission: Why, where and how?," in *Proc. Optical Fiber Commun. Conf.*, (Los Angeles, CA), 2012. paper OTh1B.1.

[5] G. Colavolpe, T. Foggi, E. Forestieri, and G. Prati, "Robust multilevel coherent optical systems with linear processing at the receiver," *J. Lightwave Tech.*, vol. 27, pp. 2357–2369, July 1 2009.

[6] J. Zhao and A. Ellis, "Electronic impairment mitigation in optically multiplexed multicarrier systems," *J. Lightwave Tech.*, vol. 29, no. 3, pp. 278–290, 2011.

[7] A. Barbieri, G. Colavolpe, T. Foggi, E. Forestieri, and G. Prati, "OFDM vs. single-carrier transmission for 100 Gbps optical communication," *J. Lightwave Tech.*, vol. 28, pp. 2537–2551, September 1 2010.

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

[9] J. Li, E. Tipsuwannakul, T. Eriksson, M. Karlsson, and P. A. Andrekson, "Approaching Nyquist limit in WDM systems by low-complexity receiver-side duobinary shaping," *J. Lightwave Tech.*, vol. 30, pp. 1664–1676, June 1 2012.

[10] A. Barbieri, D. Fertonani, and G. Colavolpe, "Time-frequency packing for linear modulations: spectral efficiency and practical detection schemes," *IEEE Trans. Commun.*, vol. 57, pp. 2951–2959, Oct. 2009.

[11] G. Colavolpe, T. Foggi, A. Modenini, and A. Piemontese, "Faster-than-Nyquist and beyond: how to improve spectral efficiency by accepting interference," *Opt. Express*, vol. 19, pp. 26600–26609, Dec 2011.

[12] J.-X. Cai, C. Davidson, A. Lucero, H. Zhang, D. Foursa, O. Sinkin, W. Patterson, A. Pilipetskii, G. Mohs, and N. Bergano, "20 Tbit/s transmission over 6860 km with sub-Nyquist channel spacing," *J. Lightwave Tech.*, vol. 30, pp. 651 –657, feb.15, 2012.

[13] F. Rusek and A. Prlja, "Optimal channel shortening for MIMO and ISI channels," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 810–818, Feb. 2012.

[14] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavčić, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inform. Theory*, vol. 52, pp. 3498–3508, Aug. 2006.

[15] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai, "On information rates for mismatched decoders," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1953–1967, Nov. 1994.

[16] R. Essiambre, G. Kramer, P. J. Winzer, F. G. J., and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightwave Tech.*, vol. 28, pp. 662–701, Feb. 2010.

[17] D. D. Falconer and F. Magee, "Adaptive channel memory truncation for maximum likelihood sequence estimation," *Bell System Tech. J.*, vol. 52, pp. 1541–1562, Nov. 1973.

[18] G. Colavolpe and A. Barbieri, "On MAP symbol detection for ISI channels using the Ungerboeck observation model," *IEEE Commun. Letters*, vol. 9, pp. 720–722, Aug. 2005.

[19] C. Shannon, "A mathematical theory of communication," *Bell System Tech. J.*, pp. 379–423, July 1948.