









800G DSP ASIC Design Using Probabilistic Shaping and Digital Sub-Carrier Multiplexing

Han Sun, Mehdi Torbatian , Mehdi Karimi, Robert Maher, Sandy Thomson, Mohsen Tehrani , Yuliang Gao, Ales Kumpera, George Soliman , Aditya Kakkar, Mohammad Osman, Ziad A. El-Sahn, Clayton Doggart , Weikun Hou, Shailesh Sutarwala, Yuejian Wu, Mohammad Reza Chitgarha, Vikrant Lal , Huan-Shang Tsai, Scott Corzine, Jiaming Zhang , John Osenbach, Sanketh Buggaveeti, Zulfikar Morbi, Miguel Iglesias Olmedo , Irene Leung, Xian Xu, Parmijit Samra, Vince Dominic, Steve Sanders, Mehrdad Ziari, Antonio Napoli , Bernhard Spinnler, Kuang-Tsan Wu, and Parthiban Kandappan

Abstract—The design of application-specific integrated circuits (ASIC) is at the core of modern ultra-high-speed transponders employing advanced digital signal processing (DSP) algorithms. This manuscript discusses the motivations for jointly utilizing transmission techniques such as probabilistic shaping and digital sub-carrier multiplexing in digital coherent optical transmission systems. First, we describe the key-building blocks of modern high-speed DSP-based transponders working at up to 800G per wave. Second, we show the benefits of these transmission methods in terms of system level performance. Finally, we report, to the best of our knowledge, the first long-haul experimental transmission – e.g., over 1000 km – with a real-time 7 nm DSP ASIC and digital coherent optics (DCO) capable of data rates up to 1.6 Tb/s using two waves ($2 \times 800G$).

Index Terms—ASIC design, digital sub-carrier multiplexing, DSP, high-capacity systems, high-order modulation formats, probabilistic shaping.

I. INTRODUCTION

OPTICAL telecommunications have experienced tremendous progress over the last decades. Just one decade ago,

Manuscript received January 21, 2020; revised March 29, 2020 and May 12, 2020; accepted May 15, 2020. Date of publication May 21, 2020; date of current version September 1, 2020. (Corresponding author: Antonio Napoli.)

Han Sun, Mehdi Torbatian, Mehdi Karimi, Sandy Thomson, Mohsen Tehrani, Yuliang Gao, Ales Kumpera, George Soliman, Aditya Kakkar, Mohammad Osman, Ziad A. El-Sahn, Clayton Doggart, Weikun Hou, Shailesh Sutarwala, Yuejian Wu, and Kuang-Tsan Wu are with Infinera Canada Inc., Ottawa, ON K2K 2X3, Canada (e-mail: hsun@infinera.com; mtorbatian@infinera.com; kKarimi@infinera.com; sthompson@infinera.com; mTehrani@infinera.com; YGao@infinera.com; akumpera@infinera.com; gSoliman@infinera.com; aKakkar@infinera.com; mOsman@infinera.com; zElsahn@infinera.com; cDoggart@infinera.com; WHou@infinera.com; sSutarwala@infinera.com; yWu@infinera.com; kwu@infinera.com).

Robert Maher, Mohammad Reza Chitgarha, Vikrant Lal, Huan-Shang Tsai, Scott Corzine, Jiaming Zhang, Zulfikar Morbi, Miguel Iglesias Olmedo, Irene Leung, Xian Xu, Parmijit Samra, Vince Dominic, Steve Sanders, Mehrdad Ziari, and Parthiban Kandappan are with Infinera Corp., Sunnyvale, CA 94089 USA (e-mail: rmaher@infinera.com; MChitgarha@infinera.com; vlal@infinera.com; ctsai@infinera.com; SCorzine@infinera.com; JiZhang@infinera.com; zMorbi@infinera.com; molmedo@infinera.com; lLeung@infinera.com; XXu@infinera.com; PSamra@infinera.com; VDominic@infinera.com; ssanders@infinera.com; MZiari@infinera.com; PKandappan@infinera.com).

John Osenbach and Sanketh Buggaveeti are with Infinera Corp., Allentown, PA 18106 USA (e-mail: JOsenbach@infinera.com; SBUGGAVEETI@infinera.com).

Antonio Napoli and Bernhard Spinnler are with Infinera Corp., 81377, Munich, Germany (e-mail: anapoli@infinera.com; BSpinnler@infinera.com).

Color versions of one or more of the figures in this article are available online at <https://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JLT.2020.2996188

10G was the leading transponder card with direct-detection systems capable of transmitting up to ~ 1 Tb/s. The explosion of Internet pushed for a rapid deployment of 100G. This was possible by jointly utilizing coherent detection technology, digital signal processing (DSP) algorithms and high-speed digital-to-analog, and analog-to-digital converters (DAC / ADC) [1]–[4]. Thanks to coherent detection, phase and amplitude information is preserved. This enables DSP algorithms to compensate for component limitations and linear fiber propagation impairments, e.g., accumulated dispersion (D_{acc}) and polarization mode dispersion (PMD). The capacity was finally doubled by introducing polarization-multiplexing [3], [5].

Upcoming innovation (e.g., 5G, high-speed optical access networks, etc.) will further increase the compound annual growth rate (CAGR) of Internet traffic, which is now $\sim 26\%$, with peaks of $\sim 46\%$ in metro [6]. To cope with the current rapid traffic evolution, several options are available, among them: (I) a more efficient management of the resources based on elastic and cognitive optical networks [7]; (II) usage of high-order modulation formats and high symbol rate [22], [23]; (III) implementation of advanced DSP and forward error correction (FEC) algorithms [8]; (IV) transmission with advanced digital communication techniques such as probabilistic shaping (PS) and digital sub-carrier multiplexing (DSCM) [9], [10]; (V) enabling beyond C-band transmission [11]; and finally (VI) development of new fiber types such as multi-mode and –core [12], [13]. Choices (I) – (V) are currently being exploited and the selection criteria mainly depends on system requirements such as reach, cost, footprint, performance, and power consumption.

An efficient management of the resources is fundamental (I), because it can postpone costly deployment of new fibers by efficiently utilizing the existing optical fiber infrastructures [7], [14]. With (II), the network architecture can be simplified by utilizing high-order modulation formats – beyond 16 quadrature amplitude modulation (QAM) – at symbol rates ≥ 90 GBaud, thus enabling single-wavelength transmission up to 800G.¹ Consequently, more complex DSP algorithms, to compensate the numerous impairments affecting the signal within the transceiver

¹Note that the chip is actually dual-wave, i.e., 1.6 Tb/s using two waves ($2 \times 800G$).

and the channel – such as the compensation of narrow bandwidth [15], [16], device nonlinearities [17], [18] and effects such as IQ skew [19], [20] – will be required with the joint utilization of advanced FEC codes/schemes (III). At the moment, first 600G commercial single-wavelength transponders [21] are available, delivering a total capacity – in C-band – up to 38.4 Tb/s/fiber over legacy medium and long distances when using 400G 64QAM in 50 GHz grid [21].

Furthermore, the introduction of advanced digital communication techniques such as PS and DSCM would enable fine data rate adaptivity and provide robustness against fiber dispersive effects, respectively. Next, the simplification and flexibility of the network management layer is also achieved thanks to water-filling techniques (i.e., availability of several operational modes), which allows an efficient utilization of the network resources based on its instantaneous status. In modern transponders [21], [22], the number of operational modes increases enormously, and in recently announced 800G line-cards [22], [23] – employing PS and DSCM [23] – the number of modes becomes even larger. This paves the way to the realization of true elastic optical networks so that dynamic traffic and time-frequency varying characteristics of the channel can be instantaneously addressed [7]. Finally, we can further boost fiber capacity, (V), by transmitting beyond C-band with C+L-band.

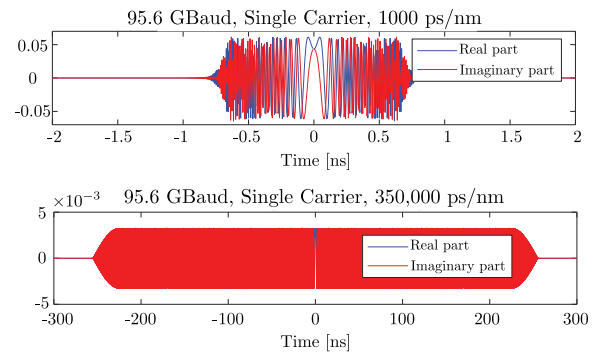
This paper focuses on the application-specific integrated circuits (ASIC) design for DSP for high-end transponders employing DSCM and PS. Section II provides the motivations for implementing them, while Section III presents their theoretical background. Section IV describes key-modules within the ASIC discussing their implementation complexity, and for the FEC, the benefit of gain sharing. Section V shows the potentialities of the joint usage of PS and DSCM to enhance transmission performance and system flexibility via water-filling. Section VI reports, to the best of our knowledge, the first real-time transmission over 1000 km of fiber with our 2×800 G DSP ASIC transponder operating at one wave of 95.6 GBaud 800 Gb/s. Section VII discusses the conclusions.

II. HIGH-SPEED OPTICAL COMMUNICATION SYSTEMS

For nearly ten years, the architecture for equalization of D_{acc} and PMD have remain largely unchanged [1]–[4], [24], [25]. In fact, electronic dispersion compensation² has been carried out in the frequency domain using fast Fourier transform (FFT), inverse-FFT (I-FFT) and overlap-and-save method as primary modules; while the residual D_{acc} and PMD compensation have been realized in time domain using a set of 2×2 tap delay line filters organized in a multiple-input multiple output (MIMO) structure. This architecture had been successful as a first-generation DSP, however lurking under the surface is the issue of equalization enhanced phase noise (EPPN) [26] that stems from non-zero laser linewidth (LW) and digital dispersion compensation. EPPN is worse if the symbol rate of the system is increased as, for example, in current transponders

²From now on, we will use the term digital dispersion compensation, or simply dispersion compensation, to refer to electronic dispersion compensation.

a) Dispersion impulse response



b) Phase noise process

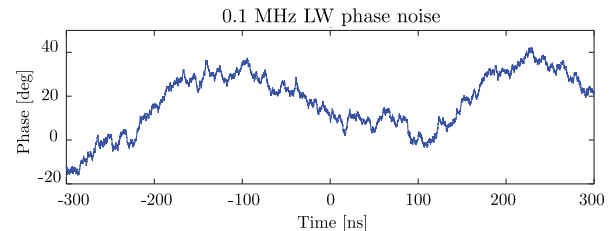


Fig. 1. (a) Dispersion impulse response at 1000 ps/nm and 350,000 ps/nm, single carrier 95.6 GBaud and (b) Phase noise process of 0.1 MHz linewidth.

[22], [23] and beyond. Recent hero experiments achieved symbol rates up to 190 GBaud [27] as an effort to further increase capacity and reduce the cost of transmission. Digital D_{acc} compensation capabilities have also increased due to the transpacific distances offered by the latest, low-loss, large effective area fiber, and the introduction of L-band systems. For modern submarine optical systems, dispersion compensation range now needs to be as high as 350,000 ps/nm. On the other hand, due to physical limitations, commercial external cavity laser LW have not reduced much below ~ 100 KHz. Ultra-low LW lasers are possible, but they are not practical for commercial use for many reasons including costs and geometry. EPPN was ignored in early generations of coherent systems but it is becoming one of the dominant effects, and it must be properly addressed, either by mitigating it or by properly designing the transponder.

Fig. 1(a) illustrates the impulse response of a digital dispersion compensating filter assuming a 95.6 GBaud single carrier signal with 1000 ps/nm and 350,000 ps/nm of dispersion. The time span of the dispersion compensating filter occupies ± 0.7 ns at 1000 ps/nm, but it spans ± 250 ns at 350,000 ps/nm. The span of the filter scales linearly with dispersion. The assumption in digital dispersion compensation is that the laser phase (in particular the receiver laser) must be static within the time span of the dispersion compensating filter in the receiver. However, laser phase noise breaks this assumption. Fig. 1(b) illustrates a time capture of a laser phase noise, which is a random walk process that can vary significantly within the time span of the filter especially when dispersion is large. In the figure, the laser phase moves between -20 deg to $+40$ deg in 200 ns, well within the time span of the dispersion compensating filter at 350,000 ps/nm. Thus, when the input dispersed optical signal is

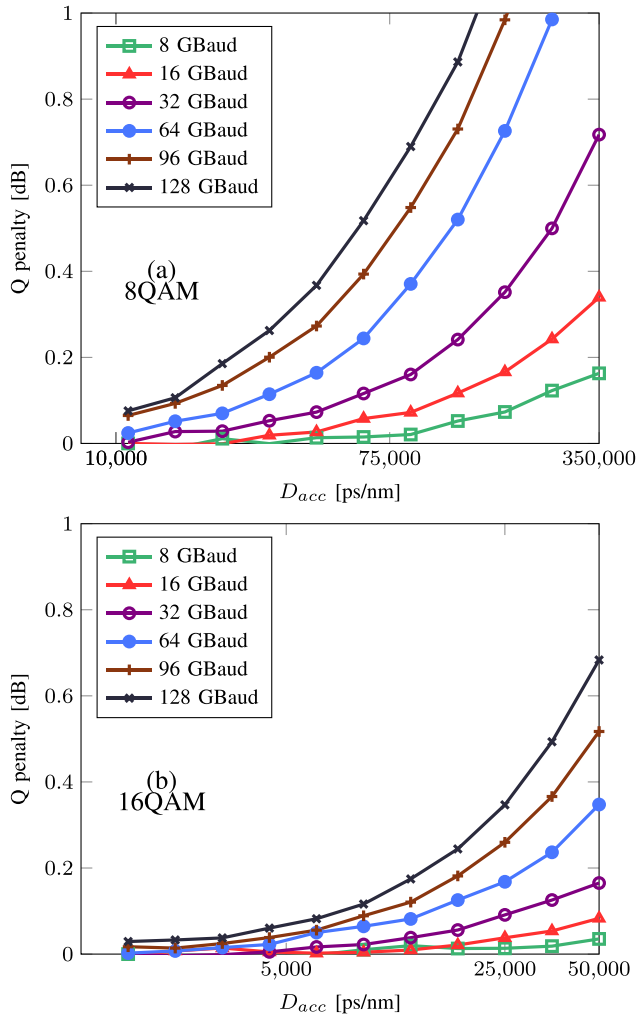


Fig. 2. Numerically evaluated system penalty from equalization enhanced phase noise for (a) 8QAM and (b) 16QAM with laser linewidth 100 KHz.

modulated by the laser phase noise and then convoluted through the dispersion compensating filter, noise can be generated in the process [28]. Tx and Rx laser phase noise can contribute to the EEPN noise process based on the amount of dispersion compensation occurring at their respective sides. The effect of EEPN noise grows linearly as a function of D_{acc} , laser LW, and symbol rate. A more detailed analytic and experimental verification of the EEPN process can be found in [29].

In Fig. 2(a) we report the system Q-penalty in [dB] normalized with respect to 7 dB versus D_{acc} – on a log scale – at different symbol rates using a laser with 100 kHz LW, and digital dispersion compensation at the receiver side. This numerical analysis assumes a DSP engine that can compensate well beyond 350,000 ps/nm of accumulated dispersion. The penalty displayed is entirely caused by EEPN. We report the results for two different modulation formats: 8QAM in Fig. 2(a) and 16QAM in Fig. 2(b). With respect to Fig. 2(a), at 0.2 dB penalty, an 8 GBaud signal can tolerate $16 \times$ more D_{acc} than a 128 GBaud signal, consistent with the symbol rate ratio. At 350,000 ps/nm, symbol rates of 32 GBaud, or higher, will encounter considerably larger Q penalties and will be impractical.

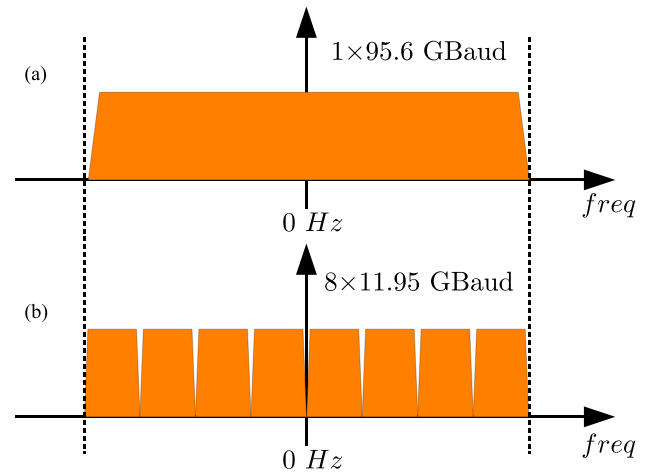


Fig. 3. Visualization of a single carrier (a) and a digital sub-carrier channel employing 8×11.95 GBaud (b).

For long haul terrestrial transmission, the accumulated fiber dispersion is less than in the previous example but can, for instance, be as high as 50,000 ps/nm. In this scenario – though can be dependent on other factors – the dominant modulation format could be 16QAM. Under these transmission conditions, Fig. 2(b) highlights that there is still a significant benefit in terms of reduced Q-penalty when transmitting at lower symbol rates, and that EEPN still represents a significant problem at 64 GBaud and above. There are existing techniques in the literature that attempt to mitigate the EEPN, among them we would like to mention very fast equalizer [30], and an approach involving additional photonic circuitry that detects the laser frequency employing an extra set of A/D converters [31], [32]. Both approaches require additional circuitry in either digital and/or analog domain, which inevitably complicates the overall system design.

III. DIGITAL COMMUNICATION AND TRANSMISSION

Key-aspects related to DSCM and PS are discussed within the next subsections. Both parts serve as motivation for the deployment of DSCM and PS.

A. Digital Sub-Carrier Multiplexing

The main conclusion from Fig. 2(a-b) is that operating at low symbol rates significantly reduces the impact of EEPN. However, if the symbol rate is reduced by a factor N , the number of lasers and parallel modulators will be increased by the same factor which is undesirable from the point of view of device complexity. Although dense photonic integration has been commercially successful for a long time [33], the ability to fully utilize the bandwidth (BW) of the electrical channel is also important to maximize the system capacity. It is in light of these considerations that we evaluate the possibilities and advantages offered by digitally multiplexing lower symbol rate sub-carriers to form a high BW signal. As an example, we consider the implementation of a 95.6 GBaud symbol rate system using 8×11.95 GBaud sub-carriers, multiplexed at near Nyquist sub-carrier spacing as visualized in Fig. 3, which shows

the spectrum of a single carrier 95.6 GBaud signal (Fig. 3(a)) and 8 sub-carriers (Fig. 3(b)) – each at 11.95 GBaud. In both cases, the spectrum is shaped to near Nyquist with roll-off factor equal to 0.0625. The frequency spacing is chosen to pack the 8 sub-carriers together without any overlap in frequency, thus making them mutually independent via frequency separation. The overall spectral occupancy is the same as that of a single carrier signal with the same roll-off factor. Other configurations are possible, however, for ease of implementation, the number of sub-carriers should ideally be powers of 2. Hence, 4 and 16 sub-carriers systems are obvious possibilities. Very small sub-carrier symbol rates are considered difficult to implement from a carrier phase tracking point of view, and very large sub-carrier symbol rate may not get the full benefit of the EEPN noise reduction. An 8 sub-carriers configuration was chosen for our 800G system. For further description of DSCM system please refer to [34].

B. Probabilistic Shaped Modulations for Maximizing Capacity

A universal metric to measure the optimality of a modulation technique over a communication channel is the distance in dB to the capacity of the channel. For additive white Gaussian noise (AWGN) channels, it has been proven that the capacity achieving distribution of the transmitted symbols is Gaussian. The distance to capacity for this channel is calculated with respect to the Shannon capacity formula for Gaussian channels [35]. One must note that the fiber optic channel is not entirely AWGN. For long distance multi-span amplified fiber transmission, amplifier noise is a dominant factor, but there is a non-negligible portion of signal distortion caused by fiber nonlinearity. This fiber nonlinearity aspect is only partially addressed with the Super-Gaussian distribution scheme to be described in later section of this manuscript. This leaves room for improvements in future systems.

Recent advances in DSP and FEC design for coherent optical modems have enabled the use of higher order modulation formats that increase spectral efficiency (SE) and reduce the distance to capacity. However, a gap to capacity still exists for a wide range of SE. A fraction of that gap is caused by overhead that is being used for pilots, training sequences, and control signals. Another fraction is due to use of FEC schemes with fixed-length codewords. A significant portion of this gap is recoverable by a more sophisticated design and optimization of the FEC and modulation formats. For example, traditional modulation formats – such as the widely used MQAM – are subject to an inherent loss in the performance because, with uniform discrete signal distribution, the required signal power to noise power ratio (SNR) for error free communication is away from the Shannon limit no matter how strong the FEC technique used here is. This performance gap (also known as shaping gain), increases at higher SE values and reaches up to 1.53 dB for an *hyper*-cube constellation with large cardinality, large square constellations expanded over n complex dimensions [36]. Thus, one big step to reduce the distance to capacity is to feed the channel with a capacity achieving distribution which is Gaussian for AWGN channels, assuming that the fiber optic channel is dominated by amplifier noise. Constellation shaping

is a technique to tailor the statistical distribution of the input symbols according to the capacity achieving distribution of the channel. Probabilistic shaping (PS) [37] and geometric shaping (GS) [38] are possible approaches to constellation shaping. In GS, the physical location of constellation points on a complex plane is adjusted to approximate a desired distribution. For a constellation with fixed cardinality, more points are located close to the origin, and less points are located as the distance from the origin is increased. As this in fact implements a quantized version of the desired distribution, it is more effective at high SE values where there are sufficient constellation points to finely quantize the desired distribution. The encoding and decoding of GS is relatively simple as no dimension expansion is involved and a symbol-by-symbol processing is enough to achieve the shaping gain. However, other aspects such as binary labeling, bit accurate implementation, and effective equalization are also challenging. For example, part of the shaping gain achieved by GS is normally optimized with respect to bit-interleaving coded modulation (BICM) systems due to lack of ideal binary gray labeling.

In PS, the input information data bits are encoded such that when they are mapped to a specific 2-D constellation such as 64QAM, the probability of occurrence of each of the constellation points follows a desired discrete probability mass function, which is designed to approximate the Gaussian distribution. The objective is to shape the amplitude distribution of the constellation, symbols with the same amplitude levels are equiprobable. This simplifies the problem by only applying the distribution over the amplitude levels of the constellation. For square shaped constellations with points located on a Cartesian grid such as 64QAM, this becomes even easier by applying the corresponding 1-D distribution on I- and Q-dimensions separately [37]. As a result, using the 64QAM base constellation set, a PS scheme can carry all SE values between 2 to 6 bits per symbol per polarization by having different distributions on the constellation points. This is a considerable advantage over GS, which requires the design of a bespoke constellation for each SE. The binary gray labeling is still in place for square QAM constellations causing no loss.

The desired distribution in PS is constructed by dimension expansion over multiple symbols of the base constellation. To that end, a specific encoder module is required to generate the desired distribution from uniform input information bits. Accordingly, a PS decoder module is required at the receiver side to recover information bits from PS encoded data. The main complexity of PS is the implementation of the encoder and decoder modules. Apart from that, the significant flexibility that it provides to finely tune the SE for each specific link, plus the SNR gain advantage as compared to GS and uniform modulation formats, makes PS a considerable powerful tool for coherent optical modems. In the following, it is shown – for a specific scenario – that the encoder and the decoder modules can be constructed with a reasonable implementation complexity and limited power consumption.

Different algorithms exist for implementation of the PS encoder and decoder modules [39]–[41]. The mainstream approaches follow the principles of lossless source entropy coding such as enumerative coding and arithmetic coding [41]. While

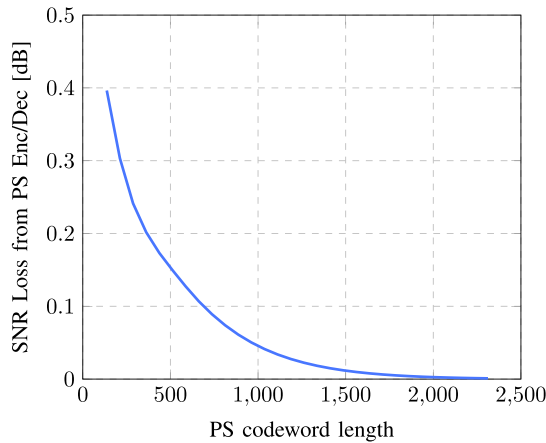


Fig. 4. Example of PS Enc/Dec SNR performance loss versus codeword length.

these approaches can directly be used for distribution matching purpose, the main challenge is to guarantee a fixed-to-fixed mapping in which a fixed length vector of bits of length k is always mapped to a fixed length vector of symbols of length n . Constant composition distribution matcher (CCDM) is a variant of the arithmetic coding approach which has been specifically designed for constellation shaping purposes [39]. This algorithm guarantees a fixed-to-fixed mapping between input bit sequences and output codewords. For a generic PS encoder/decoder design, the size of the base constellation, e.g., 16/32/64/128, and the PS codeword length n need to be chosen. The base constellation determines the range of the SE values that can be covered. The codeword length determines the precision at which a desired distribution can be implemented. A quantized version of the desired probability distribution is implemented within a length n of symbols from the base constellation. Thus, longer codeword length n results in a more accurate implementation of the desired distribution. At the same time the implementation complexity increases with increase of n .

For a specific base constellation, it can be demonstrated that almost all shaping gain can be obtained with PS codewords with length ~ 1000 symbols. For this range of n , ASIC implementation of PS encoder and decoder modules is quite feasible with reasonable implementation complexity and power consumption. Fig. 4 shows the loss in noise tolerance (SNR loss) versus PS codeword length for a net spectral efficiency corresponding to 16QAM modulation format. The SNR loss is relative to ideal implementation. Shorter codeword length results in a higher rate loss through the PS encode/decode engine, which then translates to a higher loss in noise tolerance of the system. The results obtained here are specific to the particular encoding and decoding algorithm used. Fig. 5 shows some example constellation diagrams for different spectral efficiencies.

Because the PS encode / decode engine has been designed in an ASIC with long enough codeword length, it is also easy to provide very fine tuning of SE. The implemented PS engine itself can achieve a fine SE step size of the order of a percent of a bit per symbol per polarization. This fine granularity allows for the maximization of the channel capacity by adjusting the symbol distribution of the transmitted signal between each

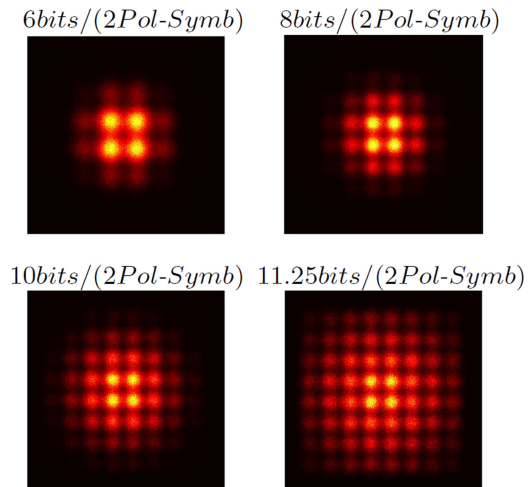


Fig. 5. Example PS constellations at different SE.

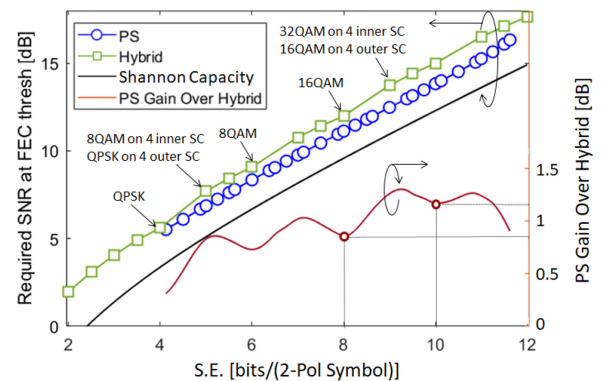


Fig. 6. SNR gain of PS modulation over hybrid-QAM.

modem pair. This capacity-maximizing capability is at the heart of [23]. Another important aspect is the raw performance of the PS scheme, and this is theoretically calculated by the SNR tolerance of the modulation. The achieved required SNR (calculated) versus spectral efficiency for PS (blue) and hybrid (green) modulations are shown in Fig. 6. The unconstrained Shannon capacity (black line) is also plotted on top for comparison. The x-axis is normalized such that QPSK is at 4 bits/(2Pol-Symbol) and uniform 64QAM is at 12 bits/(2Pol-Symbol). The FEC overhead is assumed to be 20%. The net spectral efficiency of the entire system including FEC and shaping is the value on the x-axis divided by 1.2 to account for the FEC overhead. As an example, a conventional QPSK format has spectral efficiency of 4 bits per symbol for 2 polarizations, but the net spectral efficiency including FEC is 4 bits/(2Pol-Symbol)/1.2 = 3.33 bits/(2Pol-Symbol). The conventional formats such as QPSK, 8QAM, 16QAM are labelled with an arrow on the figure for the green curve. The left y-axis visualizes the required SNR to reach FEC threshold. The FEC is a 20% overhead LDPC whose performance curves, as measured, are shown in Fig. 9. A superior modulation will generate a lower required SNR and hence higher noise tolerance and longer system reach. The PS and hybrid modulations are compared against each other and their differences in SNR tolerance is plotted in red with a second

y-axis on the right hand side of Fig. 6. Hybrid modulations are constructed using standard uniform modulation such as QPSK and 8/16/32/64QAM. For example, one can use different modulation on each sub-carrier and thereby achieving SE half way between QPSK and 8QAM, or half way between 16QAM and 32QAM, and so on. The relative power of QPSK and 8QAM symbols are optimized to maximize performance of the hybrid modulations. As labeled on the red curve, a SNR gain of 1.17 dB and 0.87 dB – right y-axis in red – is achieved in favor of PS modulation when compared against standard 32QAM and 16QAM respectively. As the curve shows, the shaping gain is smaller at 6bit/8bit/10bit spectral efficiencies, and higher in between them. This creates a wavy feature for the red curve. This is due to the fact that hybrid modulations, that are constructed using a mixture of 8QAM and 16QAM symbols, will not be as optimum as the standard modulation themselves, and hence the PS will have a slightly higher gain over hybrid schemes.

IV. ALGORITHM DESIGN

A. Frequency Domain Engine for DSCM and Benefits for Dispersion Compensation Complexity

At the receiver side of a coherent transponder a frequency domain equalizer (FDE) compensates for the bulk accumulated dispersion (a second MIMO equalizer compensates for the residual dispersion and other linear effects such as PMD). The FDE structure can be utilized for both sub-carrier de-multiplexing and dispersion compensation. Thus, within the receiver side, the complexity does not increase, and in fact, the dispersion compensation complexity is somewhat reduced due to the use of lower symbol rate sub-carrier signals. The following section compares the complexity of an FDE designed for compensating a single carrier 95.6 GBaud signal with respect to the one utilized to compensate the same amount of dispersion for an 8 digital sub-carrier DSP engine.

It is well known that the number of taps of the filter, required for dispersion compensation, increases (or decreases) proportionally to the square of the symbol rate [42]. Therefore, if the symbol rate is reduced, for example, by a factor $\times 8$, the length of the impulse response is reduced by $\times 64$. This can translate into some direct savings in dispersion compensation complexity. It is worth noting that the complexity does not scale as $\times 64$. The dispersion compensation of a DSCM system requires the compensation of the group delay distortion on each sub-carrier individually plus a delay element on each sub-carrier. The equation below illustrates that when compensating for dispersion (β_2) on an 11.95 GBaud signal at frequency offset ω_0 , one must also include a delay element (2^{nd} term: $2 \cdot \beta_2 \cdot \omega_0$). The constant phase in the 3^{rd} term will be compensated by carrier recovery, placed after the MIMO.

$$e^{j\beta_2 \cdot (\omega + \omega_0)^2} = e^{j\beta_2 \cdot \omega^2} \cdot e^{j2\beta_2 \cdot \omega \cdot \omega_0} \cdot e^{j\beta_2 \cdot \omega_0^2}$$

In time domain, one can show in Fig. 7 that the real part (blue) and the imaginary one (red) of the impulse response for a single carrier 95.6 GBaud signal can be broken down into 8 smaller impulse responses (each $1/8^{th}$ of the span), but with delay between them. The bulk delay between the sub-carriers is compensated by

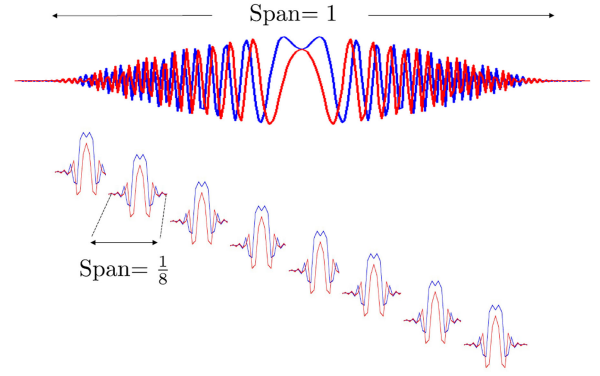


Fig. 7. Dispersion impulse response for single carrier versus 8 sub-carriers.

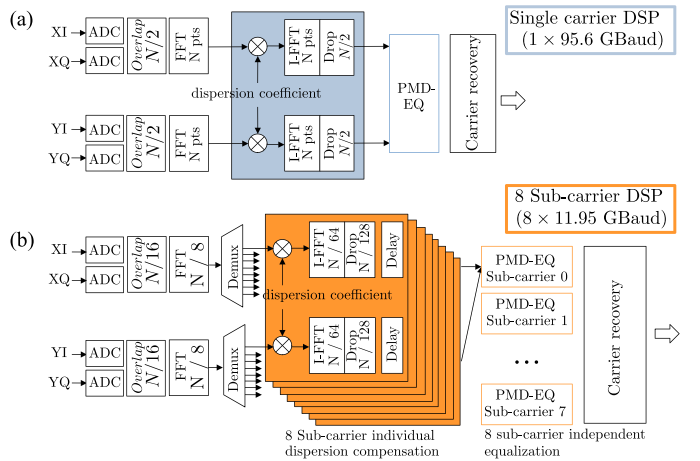


Fig. 8. FDE block diagram: (a) for single carrier; (b) for digital sub-carrier multiplexing.

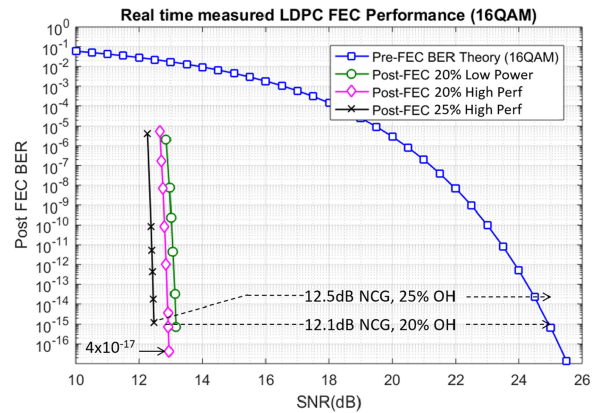


Fig. 9. LDPC SD-FEC experiment using 28 nm ASIC.

a single barrel shifter on each sub-carrier independently, which is a more computationally efficient solution when compared to using multiplications to implement a long filter. The high-level block diagram for the receiver side DSP for single carrier and 8 sub-carrier signals are illustrated in Fig. 8. The sampled signals after ADCs are first used to compensate for the bulk accumulated dispersion. For sub-carrier DSP, the dispersion compensation and sub-carrier de-multiplexing are executed all in one step in

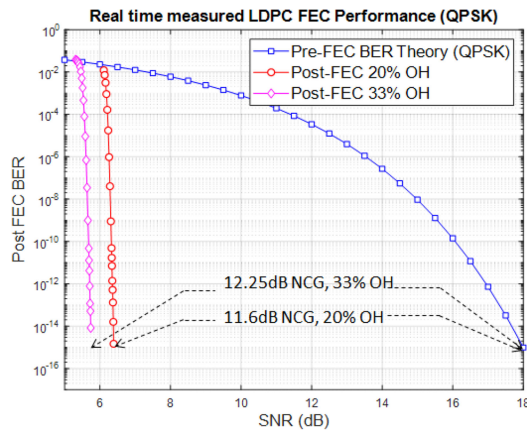


Fig. 10. LDPC SD-FEC experiment using 7 nm 800G ASIC.

the frequency domain, by using different size of FFTs and IFFTs and overlap and save method. A delay function is performed on each sub-carrier to de-skew the data due to the dispersion in the fiber. Next, a conventional 2×2 MIMO adaptive equalizer is used to perform polarization demultiplexing and compensate for residual dispersion and PMD (applied to each sub-carrier). Since sub-carriers are separated from each other in frequency, there is no interference between them, and hence no requirement for equalization that spans adjacent sub-carriers. Finally, carrier recovery and cycle slip correction are performed on the MIMO outputs, i.e., the equalized symbols. The rest of the DSP includes framing, de-mapping and FEC decoding, which are not shown in the diagram.

Example of implementation: Assuming that the FFT size required to compensate for large dispersion for a 95.6 GBaud DSP is N . For the same D_{acc} , an 8×11.95 GBaud DSP uses an FFT size of $N/8$, which is $\times 8$ smaller. The IFFT size for 95.6 GBaud DSP is also N , but for 64×11.95 GBaud DSP, it would be only $N/64$ ($64 \times$ smaller). The complexity of an FFT scales according to $\log_2(\cdot)$ (FFT size) and thus the combined FFT and IFFT would bring a complexity saving of $0.68 \times$ for 8 sub-carriers versus single carrier, which leads to power saving. We found that frequency domain dispersion coefficient multipliers for the two DSP architectures have identical complexity. PMD MIMO equalizer and as well carrier recovery functions would also be nearly identical for both DSP architectures. Other DSP functions after carrier recovery such as symbol de-mapper and FEC are also identical. For the receiver side, by combining sub-carrier de-multiplexing together with dispersion compensation, the DSP complexity is in fact reduced and not increased. On the transmitter side, sub-carrier multiplexing is a necessary additional function. As at the receiver, it can be combined with dispersion compensation at the transmitter for added tolerance to nonlinear effects caused by propagation over fiber [10]. Finally, we found there is no complexity issue when multiplexing or de-multiplexing sub-carriers.

B. SD-FEC Performance and FEC Gain Sharing

Soft decision FEC (SD-FEC) is an integral part of the overall DSP system. SD-FEC based on low density parity check codes

(LDPC) are known to be capacity approaching codes [43]. Allowing the parity check equations to span over multiple frames through spatial coupling of frames [44] or convolutional type [45], the parity check matrix can be made sparse, and message passing algorithm can truly approach maximum likelihood decoding. Message passing algorithms can be modified and simplified to reduce complexity for example by employing layered decoding algorithm [46] and min-sum algorithm [47]. LDPC codes have parity check matrices that are generated using random choices of edges between information bits and parity check bits [48]. This random connectivity defines the edges of the underlying Tanner graph and it must follow a certain pre-defined degree distribution. In LDPC theory, given an appropriately chosen variable node or check node degree distribution, it has been proven that this alone is enough to guarantee performance close to Shannon capacity. However, there are almost an infinite number of ways to generate the parity check matrix for exactly the same degree distribution. The ensemble of all those possible matrices will have average performance close to Shannon. However, the error floor of a randomly generated graph is difficult to analyze and the final error floor performance must be validated through real-time measurement or field-programmable gate array (FPGA) emulation. Note that input pre-FEC BER threshold alone is not a correct way of assessing FEC performance. For example, consider a FEC design that tolerates input BER of 2×10^{-2} , but using 20% OH versus another FEC design that tolerates the same BER of 2×10^{-2} , but instead it uses 30% OH. In this case, the 20% OH FEC is a much better design because it would require less signal BW and achieves higher SE. The proper metric for evaluation of the FEC performance is the net coding gain (NCG), whose definition normalizes out the necessary signal BW differences due to differences in FEC OH or redundancy. A higher OH code should still have higher NCG than a lower OH code even after the BW normalization. As an example, an uncoded 16QAM system can achieve an output BER = 10^{-15} at a channel SNR of 24.95 dB. With a 20% OH FEC coded system, the 16QAM symbols can be loaded with a channel SNR of 12.04 dB to reach output BER = 10^{-15} . The NCG can be calculated as

$$24.95 - 12.04 - 10 \times \log_{10}(1.2) = 12.12 \text{ dB.}$$

The NCG must factor-in the extra noise the system may experience when system BW is increased by 20%. Fig. 9 shows a real-time measured LDPC FEC performance internal to a 28 nm CMOS ASIC that was manufactured and integrated into Infinera's ICE4 system. Binary bits are first LDPC encoded and then gray mapped into 16QAM symbols. The symbols are noise loaded by an internal noise generator circuit. Demapper is used to generate the log-likelihood ratios (LLR) values for the binary LDPC decoding in a BICM scheme. The uncoded 16QAM system is shown as the blue curve, and the measurement is shown for 20% and 25% overhead codes. The error floor of the code is verified down past 4×10^{-17} . The modulation assumed is 16QAM, and 12.1 dB NCG is achieved at 20% OH and 12.5 dB NCG at 25% OH. The third curve (with low power in the legend) is a measurement result using half the number of

iterations resulting in 0.2 dB less NCG. Fig. 10 shows a real-time measurement of the LDPC FEC used in the 7 nm 2×800 G ASIC. Measured using QPSK modulation format (hence a binary NCG), it is capable of delivering 11.6 dB NCG at 20% OH, and 12.25 dB NCG at 33% OH. The 33% OH curve was measured down to 10^{-14} BER, and is extrapolated from 10^{-14} to 10^{-15} . The encoding for the 20% OH code is exactly the same as the code in Fig. 9. When used with 16QAM format, this 20% code is expected to reach 12.2 dB NCG, a slight increase of 0.1 dB NCG over the previous generation design shown in Fig. 9. This is a result of a slight increase in decoding complexity.

The 7 nm ASIC is designed to accommodate two waves worth of DSP, each at 800 Gb/s for a total of 1.6 Tb/s. This dense integration allows for SD-FEC gain sharing to be utilized as a form of nonlinear or polarization-dependent loss (PDL) mitigation [49], [50]. At the transmitter, the two FEC encoding engines produce two codewords that are interleaved (or mixed) and sent over to $2 \times$ DSP engines, which then modulate onto two waves of 8 sub-carriers each. At the receiver side, data and symbols from 16 sub-carriers are de-interleaved reversing the interleaving pattern in the transmitter, and then sent to two separate FEC decoding engines. In this arrangement, both FEC decoders experience the exact same average BER even though the two physical channels may be well separated in frequency, and experience different BER due to erbium doped fiber amplifier (EDFA) gain tilt or nonlinear impairments. Since each wave is widely tunable across the whole C-band, FEC gain sharing can be applied over any two pairs of waves widely separated in frequency as, for example, in dispersion managed submarine fiber systems where BER in one end of the spectrum may be significantly different than another part. The worst channel in a WDM system can be arranged to have its FEC codeword shared with the best channel, and as such, both channels experience the same averaged BER, and more Q margin is obtained for the worst case channel. In an extreme case, the worst channel may have an error rate higher than FEC threshold, but after FEC gain sharing, its error rate may be brought under the FEC threshold. The capacity improvements and benefits in this regime have been demonstrated in real-time commercial systems [49]. The other application is PDL mitigation. PDL penalty in the coherent system is strongly dependent on incident polarization angle onto the principal PDL axis in the fiber, which varies significantly over frequency due to the birefringent nature of the fiber. Therefore the BER varies significantly for different wavelength division multiplexing (WDM) channels even in absence of nonlinear effects or EDFA gain tilt. In another real-time commercial system, the Q penalties from PDL effects of the channel can be mitigated by 40–50% using the FEC gain sharing technique, resulting in 0.5 dB gain in Q margin to FEC threshold for a fiber system with 3 dB average PDL [50].

Fig. 11 illustrates a layout diagram of a 7 nm DSP ASIC capable of two waves of TX and RX DSP, with each wave operating at maximum capacity of 800 Gb/s. The modulation format employed on each of the 8 sub-carriers is a probabilistically-shaped modulation using 64QAM as the base constellation. The aggregate symbol rate of the 8 sub-carriers exceeds 90 GBaud in order to achieve 800 Gb/s per wave.

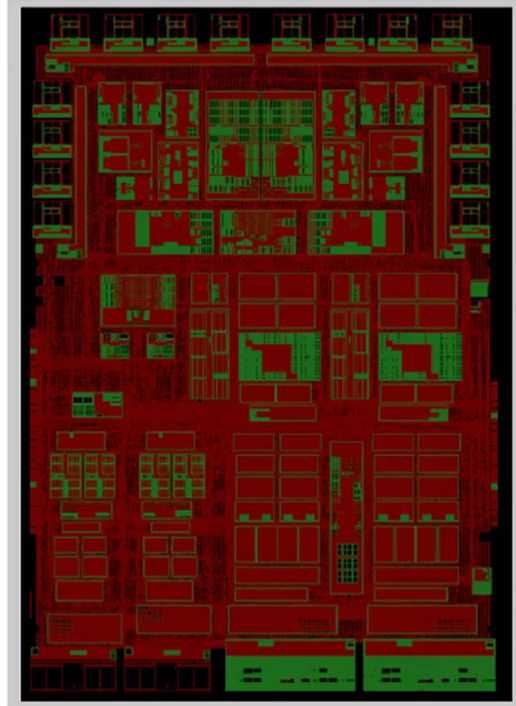


Fig. 11. 2×800 G 7 nm ASIC layout diagram.

V. BENEFIT OF PS AND DSCM

A. Fiber Kerr Nonlinearity Tolerant Modulations

Although the PS modulation outperforms regular equiprobable QAM in the linear regime, when signals are propagated over long distances of fiber, other opportunities exist to further optimize system performance. The main reason is that the Maxwell-Boltzmann distribution of the constellation points is optimized for an AWGN channel and, for a long-haul fiber transmission, this assumption is not entirely valid so the system, at optimum launch power, can have a significant nonlinear component. In this case, we should use other considerations such as the nonlinear parameters of fiber, to optimize the distribution of the constellation points. A generalized form of Maxwell-Boltzmann distribution has been proposed in [51], and implemented in the 7 nm real-time ASIC. This Super-Gaussian distribution significantly reduces nonlinear interference (NLI) variance of received signal by decreasing the standardized moments of the signal. Thereby, the overall received SNR of the signal would be improved compared to Maxwell-Boltzmann distribution. The equation below illustrates the probability mass function of the transmitted symbols for a Super-Gaussian distribution.

$$p(x_i) = e^{-\lambda x_i^P} / I, i = 0, 1, \dots, m - 1,$$

where m is the cardinality of 1-D amplitude levels of the constellation (e.g., $m = 4$ for 64QAM), and $1/I$ is the scaling factor to normalize the sum of the probabilities to one. The parameter “ P ” in the equation is equal to 2 in the Maxwell-Boltzmann distribution, but equal to a value larger than 2 in the Super-Gaussian distribution. It can be seen in Fig. 12 that the Super-Gaussian

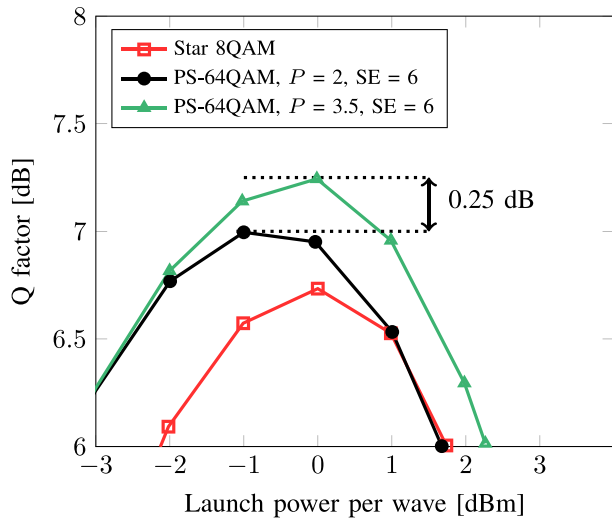


Fig. 12. System Q benefit of Super-Gaussian PS signals.

distribution (with $P = 3.5$) provides 0.25 dB higher Q at optimum launch power compared to Maxwell-Boltzmann distribution (with $P = 2$). The simulation environment is 9360 km (151 spans) of TeraWave fiber. There are $16 \times$ WDM channels, each consists of 8 sub-carriers at 8.25 GBaud. Each sub-carrier is modulated with either Star-8QAM at 6 bits per dual-pol symbol, or shaped modulation based on 64QAM symbol set achieving the same SE. The optimum launch power is also improved by slightly less than 1 dB. All this translates to longer system reach for the Super-Gaussian shaped signal.

B. Capacity optimization of Digital Sub-Carrier Multiplexed Transmission by Means of the Water-Filling Approach

In practical high symbol rate systems at 90 GBaud and above, the signal quality at high frequencies presents a critical issue that needs to be addressed to reduce the implementation penalty for high order modulation such as 64QAM or high SE shaped signals. The high frequency signal quality degrades due to various factors such as: (a) signal loss at high frequencies, a.k.a. BW limitation; (b) uncontrolled reflections in the RF chain at high frequencies; (c) higher nonlinearities generated by high frequency contents of the signal through nonlinear devices such as a Mach-Zehnder (MZ) driver electronics; (d) I- and Q-channel transfer function mis-match at high frequencies that can be device dependent and time-varying from thermal effects and aging. There are several effects in real, practical systems that cannot be effectively compensated by DSP means, and hence degrade achievable SNR. It is well known that if the channel SNR is a function of frequency, then it is advantageous to modulate different data rates in different parts of the spectrum in order to maximize the net throughput. This technique is known as water-filling and is a classic technique deployed in OFDM systems [52]. The combination of DSCM and PS is a natural and cost effective way to implement water-filling for the real-time ASIC. In Fig. 13, the inner-most sub-carriers {6, 7, 0, 1} may be modulated with high spectral efficiency shaped signal while sub-carrier {2, 5} might be modulated at slightly lower spectral

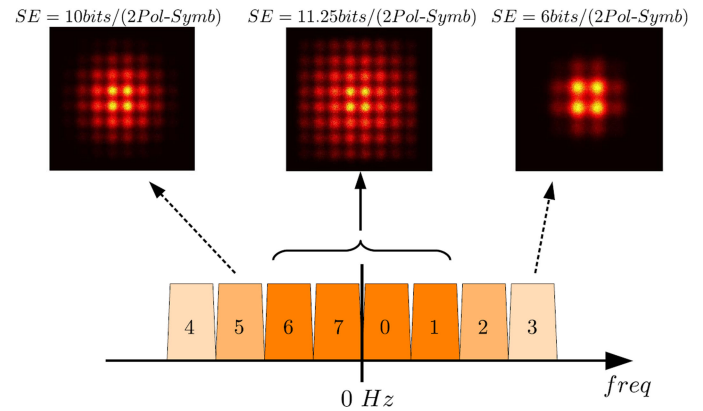


Fig. 13. Sub-Carrier water-filling using PS modulation.

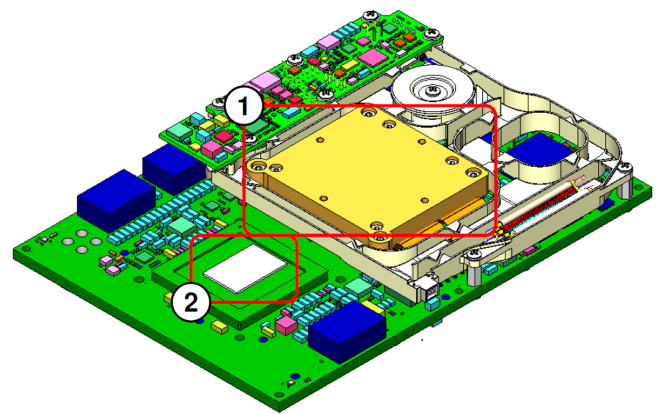


Fig. 14. 1.6 Tb/s DCO. Bounding box ① highlights the hybrid integrated 2-wavelength electro-optical module, while bounding box ② highlights the $2 \times$ 800G DSP ASIC.

efficiency and thus able to tolerate a higher amount of noise and distortions, and the two outer-most sub-carriers {4, 3} are modulated with the lowest spectral efficiency consistent with the reduced channel SNR on the outer-most sub-carriers. With the channel SNR function of frequency, the modulation should also be chosen appropriately for the achievable SNR in that frequency band. This water-filling technique is only possible, with it and represents a unique capability of digital sub-carrier modulation.

VI. EXPERIMENTAL DEMONSTRATION OF A $2 \times$ 800G REAL-TIME TRANSPONDER

In this section, we report the first, to the best of our knowledge, experimental demonstration of a real-time 2-wave 1.6 Tb/s digital coherent optics (DCO) module, which is shown in Fig. 14. The DCO is comprised of a $2 \times$ 800G DSP ASIC and hybrid integrated electro-optical module, which consists of a 2-wave photonic integrated circuit (PIC), flip-chip bonded with a corresponding 2-wave analog driver ASIC [53]. The integrated DCO is capable of generating and receiving two optical wavelengths with data rates ranging from 100 Gb/s to 800 Gb/s per-wavelength, using symbol rates that can be tuned from 32 GBaud to 100 GBaud.

A. Experimental Results

1) *Optical Back-to-Back*: The DCO was initially operated in optical loop-back and the modulation was set to 95.6 GBaud PS-64QAM, which provided a net data rate of 800 Gb/s at an information rate of 10.5 bit/sym. The roll-off factor of the digital pulse shaping filter in the TX DSP was set to 6.25%, therefore the optical BW of a single wavelength channel was 101.62 GHz. Fig. 15(a) shows the transmitted optical spectrum at the output of the DCO, where the TX laser frequency was set to 193.48125 THz. The optical output of the DCO was attenuated before being passed directly into the receive port of the DCO. Fig. 15(b) shows the corresponding received spectrum, which was read directly from the DSP ASIC. The power uniformity of the sub-carriers demonstrates the very high electro-optical BW response of the DCO. The sub-carriers are numbered from 0 to 3 for the upper side band SCs and 4 through 7 for the lower side band SCs.

Fig. 15(c) shows the experimentally recorded Q-factor (averaged over both polarizations) for each of the 8 digital sub-carriers that comprise the 800 Gb/s 95.6 GBaud PS-64QAM channel. Water-filling, as described in Section V-B, was employed to ensure consistent performance across all sub-carriers. While the received Q-factor ranged from 8.3 dB to 8.6 dB across all 8 sub-carriers, it is the mean Q-factor (8.45 dB) of the entire 800 Gb/s signal that is important for the performance of the FEC engine. Fig. 16 shows the experimentally captured probabilistically shaped constellations of all 8 sub-carriers and both polarization. Water-filling can be clearly observed from the constellation data. The inner sub-carriers were set to the highest information rate and therefore the densest modulation format. As the frequency of the SC increases, the information rate is reduced, while still maintaining a net bit rate of 800 Gb/s.

2) *Transmission*: The setup for the full C-band transmission experiment is illustrated in Fig. 17(a). A single 800 Gb/s 95.6 GBaud PS-64QAM channel, from the real-time 2-wavelength ICE6 DCO,³ was transmitted at the center of the C-band (193.48125 THz). The remainder of the C-band was loaded using spectrally shaped amplified spontaneous emission (SS-ASE) noise [54]. The ASE aggressor was multiplexed onto the line system using a commercially available reconfigurable optical add-drop multiplexer (ROADM), which consisted of a wavelength selective switch (WSS), EDFA and an optical control channel. The 800 Gb/s 95.6 GBaud PS-64QAM channel under test (CUT) was also multiplexed using the same ROADM, which amplified the entire WDM signal to provide a total launch power of 19 dBm. The launch power of the CUT was set to 4 dBm.

The transmission link consisted of ten 100 km fiber spans, each followed by a commercially available hybrid Raman EDFA booster amplifier. The fiber used in this experiment was CorningTXF fiber, which has a combination of ultra-low attenuation (0.168 dB/km typical at 1550 nm), large effective area (125 μm^2 typical at 1550 nm) and high dispersion

³ICE6 is Infinera's 6th generation coherent optical engine.

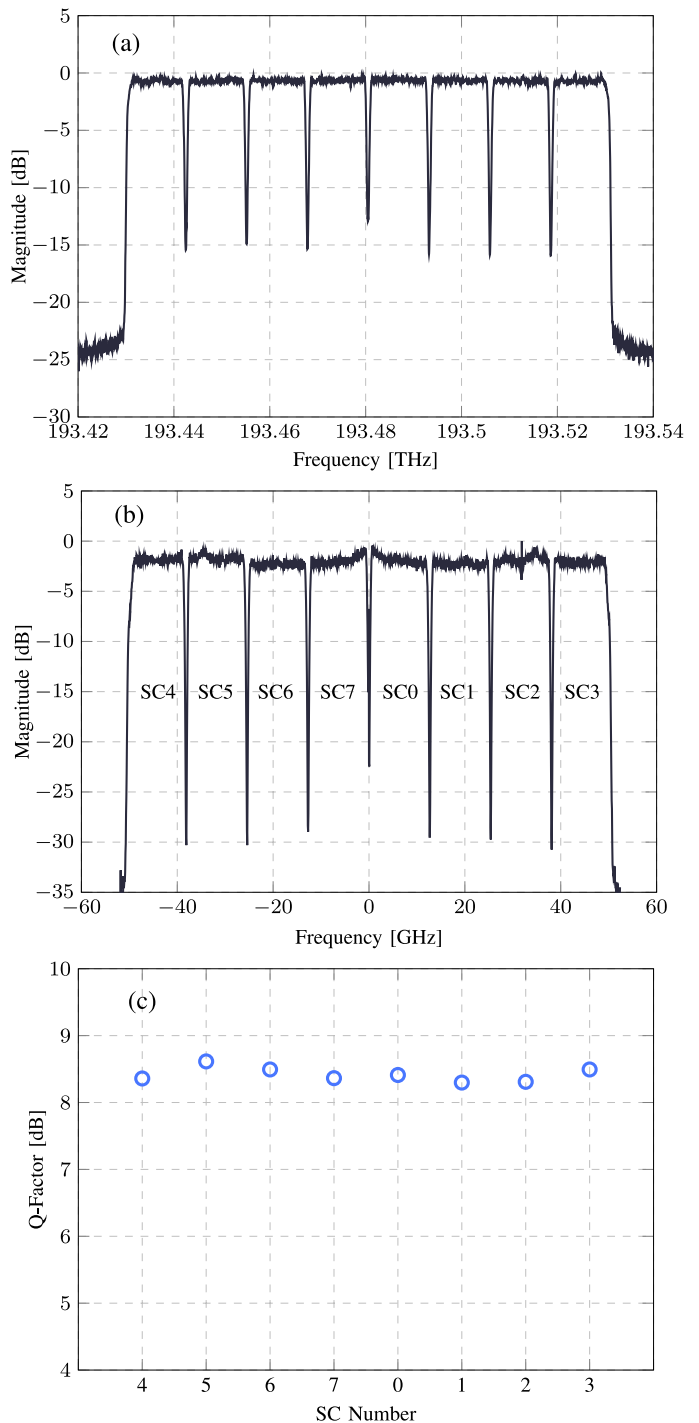


Fig. 15. (a) Experimentally measured optical spectrum at the output of the DCO. (b) Experimentally measured PSD of the received signal as reported by the DSP ASIC. (c) Experimentally measured B2B Q-factor for a single optical channel operating at 800 Gb/s 95.6 GBaud PS-64QAM.

(~ 21 ps/nm.km at 1550 nm). TXF fiber is based on a silica-core design, which allows for a lower nonlinear refractive index ($\sim 2.2 \times 10^{-20}$ m²/W) compared to Germanium-doped fibers. The combination of ultra-low loss and improved tolerance towards nonlinear effects enables longer reach, higher transmission capacity, and improved system margin. As an ITU-T Recommendation G.654.E compliant fiber with a maximum cable cut-off

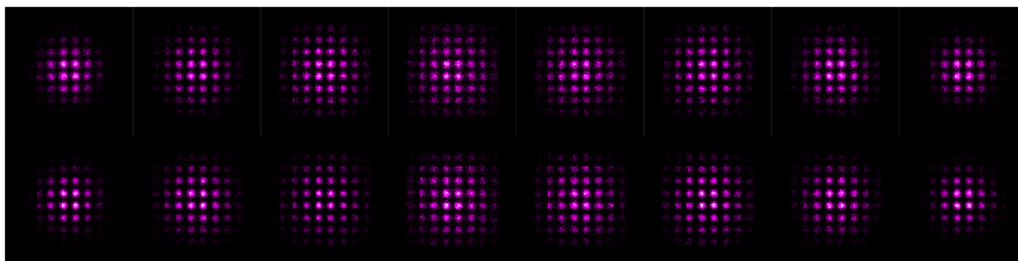


Fig. 16. Experimentally measured 800 Gb/s 95.6 GBaud PS-64QAM constellations for 8 sub-carriers per polarization when the DCO is operated in full optical loop-back. The top row is the XPOL, while the bottom row is the YPOL. The constellations are ordered as shown in Fig. 15(b).

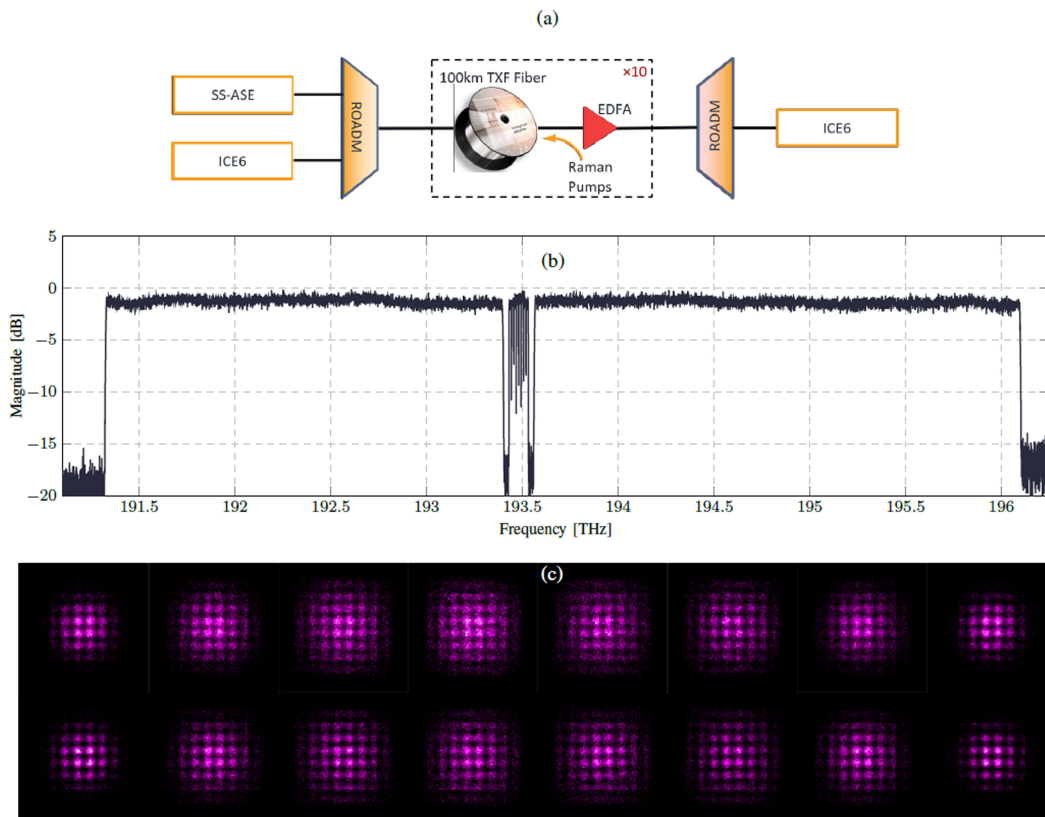


Fig. 17. (a) Full C-band transmission test bed. (b) WDM optical spectrum with a single real-time 95.6 GBaud PS-64QAM channel located at the center of the C-band, with ASE aggressors. (c) Corresponding X- and Y-POL constellations after transmission over 1000 km of CorningTXFfiber.

wavelength of 1520 nm, TXF fiber is specifically designed to operate for long-haul and regional terrestrial networks that use the C or C+L transmission bands.

Fig. 17(b) illustrates the transmitted spectrum at the output of the ROADM, with the real-time 800 Gb/s PS-64QAM signal placed at the center of the C-band. After transmission over the 1000 km line system, the ROADM at the receive port, which was set to a pass-band BW of 150 GHz, dropped the CUT and fed it directly to the receive port of a second ICE6 DCO. The receive DCO subsequently attenuated the optical signal to -6 dBm in order to achieve optimal performance.

The average Q-factor (over 2 polarizations and 8 sub-carriers) for the 800 Gb/s 95.6 GBaud PS-64QAM, which was experimentally measured over approximately 10 hours is shown in Fig. 18. The mean Q-factor achieved over the 10 hour measurement window was 6.1 dB, with the 3σ Q-factor bounded between 6 dB

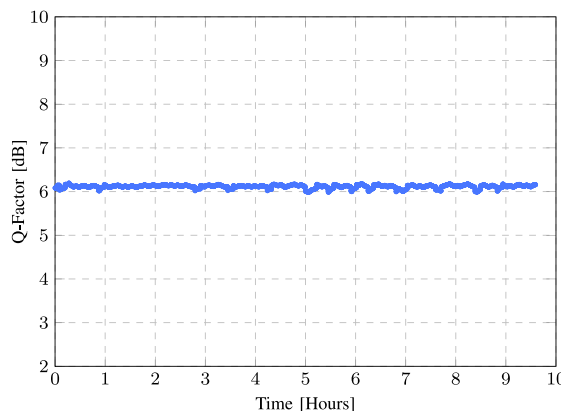


Fig. 18. Average Q-factor of the 800 Gb/s 95.6 GBaud PS-64QAM channel after transmission over 1000 km of CorningTXFfiber. Q-values were experimentally measured with a time interval of 5 seconds over nearly 10 hours.

and 6.2 dB. This provided a Q-factor margin of 0.4 to 0.6 dB to the FEC threshold, thus providing zero post-FEC errors. The corresponding received constellations for all sub-carriers are displayed in Fig. 17(c).

VII. CONCLUSIONS

We report on the challenges and realization of a $2 \times 800\text{G}$ transponder for next generation optical networks. We motivated and discussed the main features of digital sub-carrier multiplexing, probabilistic shaping and the need for implementing them in a real-time transponder. We also described the main required algorithms and benefits arising by the usage of these methods. Last section reports on the first, to the best of our knowledge, real-time experiment using a latest 7 nm ASIC. Each ASIC and PIC is designed capable of two waves of 800G each.

ACKNOWLEDGMENT

The authors would like to thank Corning for the loan of the Corning TFX fiber used in this work. The authors would also like to thank Dr. Ulrich Gaubatz and Paul Doolan for their valuable comments and the anonymous reviewers who helped to improve our article with their constructive comments.

REFERENCES

- [1] H. Sun, K.-T. Wu, and K. Roberts, "Real-time measurements of a 40 Gb/s coherent system," *Opt. Express*, vol. 16, no. 2, pp. 873–879, 2008.
- [2] M. Kuschnerov *et al.*, "DSP for coherent single-carrier receivers," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3614–3622, Aug. 2009.
- [3] K. Roberts *et al.*, "Performance of dual-polarization QPSK for optical transport systems," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3546–3559, Aug. 2009.
- [4] Y. Han and G. Li, "Coherent optical communication using polarization multiple-input-multiple-output," *Opt. Express*, vol. 13, no. 19, pp. 7527–7534, 2005.
- [5] M. S. Alfiad *et al.*, "111-Gb/s transmission over 1040-km field-deployed fiber with 10G/40G neighbors," *IEEE Photon. Technol. Lett.*, vol. 21, no. 10, pp. 615–617, May 2009.
- [6] "Cisco Visual Networking Index: Forecast and Methodology," Jun. 2017. [Online]. Available: <https://www.cisco.com/c/en/us/solutions/service-provider/visual-networking-index-vni/index.html>
- [7] J. Slovak *et al.*, "Aware optical networks: Leaving the lab," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 11, no. 2, pp. A134–A143, Feb. 2019.
- [8] K. Sugihara, K. Ishii, K. Dohi, K. Kubo, T. Sugihara, and W. Matsumoto, "Scalable SD-FEC for efficient next-generation optical networks," in *Proc. 42nd Eur. Conf. Opt. Commun.*, 2016, pp. 1–3.
- [9] M. Torbatian, D. Chan, H. H. Sun, S. Thomson, and K.-T. Wu, "Distribution matching for probabilistic constellation shaping with an arbitrary input/output alphabet," U.S. Patent App. 16/152,353, May 16, 2019.
- [10] D. Pileri, L. Bertignono, A. Nespola, F. Forghieri, and G. Bosco, "Comparison of probabilistically shaped 64QAM with lower cardinality uniform constellations in long-haul optical systems," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 501–509, Jan. 2018.
- [11] A. Ferrari *et al.*, "Assessment on the achievable throughput of multi-band ITU-T G.652.D fiber transmission systems," *J. Lightw. Technol.*, to be published, doi: [10.1109/JLT.2020.2989620](https://doi.org/10.1109/JLT.2020.2989620).
- [12] R. Van Uden *et al.*, "Ultra-high-density spatial division multiplexing with a few-mode multicore fibre," *Nature Photon.*, vol. 8, pp. 865–870, 2014. [Online]. Available: <https://doi.org/10.1038/nphoton.2014.243>
- [13] T. Kobayashi *et al.*, "1-Pb/s (32 SDM/46 WDM/768 Gb/s) C-band dense SDM transmission over 205.6-km of single-mode heterogeneous multicore fiber using 96-Gbaud PDM-16QAM channels," in *Proc. Opt. Fiber Commun. Conf.*, 2017, Paper Th5B–1.
- [14] Y. Pointurier, "Design of low-margin optical networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 9, no. 1, pp. A9–A17, Jan. 2017.
- [15] A. Napoli *et al.*, "Novel DAC digital pre-emphasis algorithm for next-generation flexible optical transponders," in *Proc. Opt. Fiber Commun. Conf.*, 2015, Paper Th3G–6.
- [16] Y. Yoffe *et al.*, "Low-resolution digital pre-compensation enabled by digital resolution enhancer," *J. Lightw. Technol.*, vol. 37, no. 6, pp. 1543–1551, Mar. 2019.
- [17] P. W. Berenguer *et al.*, "Nonlinear digital pre-distortion of transmitter components," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1739–1745, Apr. 2016.
- [18] A. Napoli, M. M. Mezghanni, S. Calabro, R. Palmer, G. Saathoff, and B. Spinnler, "Digital predistortion techniques for finite extinction ratio IQ Mach-Zehnder modulators," *J. Lightw. Technol.*, vol. 35, no. 19, pp. 4289–4296, Oct. 2017.
- [19] J. C. M. Diniz, F. Da Ros, E. P. da Silva, R. T. Jones, and D. Zibar, "Optimization of DP-M-QAM transmitter using cooperative coevolutionary genetic algorithm," *J. Lightw. Technol.*, vol. 36, no. 12, pp. 2450–2462, Jun. 2018.
- [20] G. Khanna, S. Calabrò, B. Spinnler, E. De Man, and N. Hanik, "Joint adaptive pre-compensation of transmitter I/Q skew and frequency response for high order modulation formats and high baud rates," in *Proc. Opt. Fiber Commun. Conf.*, 2015, Paper M2G–4.
- [21] "Infinera 600G," 2019. [Online]. Available: <https://www.infinera.com/press-release/infinera-sets-highest-performance-600gtransmission-record>
- [22] "Ciena 800 G," 2019. [Online]. Available: <https://www.ciena.com/insights/articles/Ciena-unveils-WaveLogic-5-800G-and-somuch-more.html>
- [23] "Infinera 800 G," 2019. [Online]. Available: <https://www.infinera.com/press-release/infinera-announces-ice6-the-next-generationinfinite-capacity-engine-featuring-800g-waves>
- [24] S. Savory *et al.*, "Digital equalisation of 40Gbit/s per wavelength transmission over 2480 km of standard fibre without optical dispersion compensation," in *Proc. IEEE Eur. Conf. Opt. Commun.*, 2006, pp. 1–2, doi: [10.1109/ECOC.2006.4800978](https://doi.org/10.1109/ECOC.2006.4800978).
- [25] J. C. Geyer, C. R. Fludger, T. Duthel, C. Schulien, and B. Schmauss, "Efficient frequency domain chromatic dispersion compensation in a coherent polmux QPSK-receiver," in *Proc. Opt. Fiber Commun. Conf.*, 2010, Paper OWV5.
- [26] W. Shieh and K.-P. Ho, "Equalization-enhanced phase noise for coherent-detection systems using electronic digital signal processing," *Opt. Express*, vol. 16, no. 20, pp. 15 718–15 727, 2008.
- [27] X. Chen *et al.*, "All-electronic 100-GHz bandwidth digital-to-analog converter generating PAM signals up to 190 gbaud," *J. Lightw. Technol.*, vol. 35, no. 3, pp. 411–417, Feb. 2017.
- [28] S. Yamamoto, N. Edagawa, H. Taga, Y. Yoshida, and H. Wakabayashi, "Analysis of laser phase noise to intensity noise conversion by chromatic dispersion in intensity modulation and direct detection optical-fiber transmission," *IEEE J. Lightw. Technol.*, vol. 8, no. 11, pp. 1716–1722, Nov. 1990.
- [29] A. Kakkar *et al.*, "Comprehensive study of equalization-enhanced phase noise in coherent optical systems," *J. Lightw. Technol.*, vol. 33, no. 23, pp. 4834–4841, Dec. 2015.
- [30] X. Zhou and L. Nelson, "Advanced DSP for 400 gb/s and beyond optical networks," *J. Lightw. Technol.*, vol. 32, no. 16, pp. 2716–2725, Aug. 2014.
- [31] A. Kakkar *et al.*, "Overcoming EEPN in coherent transmission systems," in *Proc. IEEE Conf. Lasers Electro-Opt.*, 2016, pp. 1–2.
- [32] G. Goldfarb, "Apparatus for compensating optical signal impairments," U.S. Patent 8,923,707, Dec. 30, 2014.
- [33] R. Going *et al.*, "Multi-channel InP-based coherent PICs with hybrid integrated SiGe electronics operating up to 100 GBd, 32QAM," in *Proc. IEEE Eur. Conf. Opt. Commun.*, 2017, pp. 1–3, doi: [10.1109/ECOC.2017.8346096](https://doi.org/10.1109/ECOC.2017.8346096).
- [34] D. Krause, A. Awadalla, A. S. Karar, H. Sun, and K.-T. Wu, "Design considerations for a digital subcarrier coherent optical modem," in *Proc. Opt. Fiber Commun. Conf.*, 2017, Paper Th1D–1.
- [35] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2012.
- [36] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 913–929, May 1993.
- [37] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idrer, "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.

- [38] N. S. Loghin, J. Zöllner, B. Mouhouche, D. Anzorregui, J. Kim, and S.-I. Park, "Non-uniform constellations for ATSC 3.0," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 197–203, Mar. 2016.
- [39] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 430–434, Jan. 2016.
- [40] C. Pan and F. R. Kschischang, "Probabilistic 16-QAM shaping in WDM systems," *J. Lightw. Technol.*, vol. 34, no. 18, pp. 4285–4292, Sep. 2016.
- [41] T. Cover, "Enumerative source encoding," *IEEE Trans. Inf. Theory*, vol. 19, no. 1, pp. 73–77, Jan. 1973.
- [42] B. Spinnler, "Equalizer design and complexity for digital coherent receivers," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1180–1192, Sep./Oct. 2010.
- [43] A. Alvarado, E. Agrell, D. Lavery, and P. Bayvel, "LDPC codes for optical channels: Is the "FEC limit" a good predictor of post-FEC BER?" in *Proc. Opt. Fiber Commun. Conf.*, 2015, Paper Th3E–5.
- [44] L. Schmalen, V. Aref, J. Cho, D. Suikat, D. Rösener, and A. Leven, "Spatially coupled soft-decision error correction for future lightwave systems," *J. Lightw. Technol.*, vol. 33, no. 5, pp. 1109–1116, Mar. 2015.
- [45] L. Schmalen, D. Suikat, D. Rösener, and A. Leven, "Evaluation of left-terminated spatially coupled LDPC codes for optical communications," in *Proc. IEEE Eur. Conf. Opt. Commun.*, 2014, pp. 1–3, doi: [10.1109/ECOC.2014.6964006](https://doi.org/10.1109/ECOC.2014.6964006).
- [46] D. Chang *et al.*, "LDPC convolutional codes using layered decoding algorithm for high speed coherent optical transmission," in *Proc. IEEE Opt. Fiber Commun. Conf. Exhib./Nat. Fiber Optic Engineers Conf.*, 2012, pp. 1–3.
- [47] I. B. Djordjevic, M. Arabaci, and L. L. Minkov, "Next generation FEC for high-capacity communication in optical transport networks," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3518–3530, Aug. 2009.
- [48] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.
- [49] E. Abbess *et al.*, "Capacity improvement using dual-carrier FEC gain sharing in submarine optical communications," in *Proc. Opt. Fiber Commun. Conf.*, 2016, Paper Th2A–51.
- [50] J. Rahn *et al.*, "Transmission improvement through dual-carrier FEC gain sharing," in *Proc. IEEE Opt. Fiber Commun. Conf. Expo. Nat. Fiber Opt. Engineers Conf.*, 2013, pp. 1–3, doi: [10.1364/OFC.2013.OW1E.5](https://doi.org/10.1364/OFC.2013.OW1E.5).
- [51] M. N. Tehrani, M. Torbatian, H. Sun, P. Mertz, and K.-T. Wu, "A novel non-linearity tolerant super-gaussian distribution for probabilistically shaped modulation," in *Proc. IEEE Eur. Conf. Opt. Commun.*, 2018, pp. 1–3, doi: [10.1109/ECOC.2018.8535379](https://doi.org/10.1109/ECOC.2018.8535379).
- [52] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 171–178, Feb. 2003.
- [53] V. Lal *et al.*, "1.6 Tbps coherent 2-channel transceiver using a monolithic Tx/Rx InP PIC and single SiGe ASIC," in *Proc. IEEE Opt. Fiber Commun. Conf. Expo. Nat. Fiber Opt. Engineers Conf.*, 2020, p. M3A-2.
- [54] D. J. Elson *et al.*, "Investigation of bandwidth loading in optical fibre transmission using amplified spontaneous emission noise," *Opt. Express*, vol. 25, no. 16, pp. 19 529–19 537, 2017.