Turbo DFE Assisted Time-Frequency Packing for Probabilistically Shaped Terabit Superchannels

Mrinmoy Jana, Lutz Lampe, and Jeebak Mitra

***Abstract*—We present a probabilistically shaped (PS) time-frequency-packing (TFP) wavelength-division multiplexing superchannel system employing higher order modulation (HoM) formats, with an objective to improve the spectral efficiency (SE). However, TFP systems suffer from inter-symbol interference (ISI) and/or inter-carrier interference (ICI). Additionally, the bandwidth limitations of the wavelength selective switches in the fiber link may cause severe ISI for the edge sub-channels (SCs) in a superchannel. Mitigation of such interferences is particularly challenging for HoM systems, since the implementation of the well-known turbo equalization schemes, such as the Bahl-Cocke-Jelinek-Raviv equalizer, is computationally challenging for larger constellations. In this paper, we investigate an expectation propagation based, computationally efficient, turbo decision feedback equalizer for ISI cancellation, in tandem with a parallel interference cancellation based ICI mitigation. By optimizing the parameters in the shaping and the packing dimensions, we show through our numerical results that the proposed PS-TFP superchannels enabling 1.8 Tbps data rates offer up to 1.05 dB packing gain over an unpacked system using the same modulation format, and a 1.15 dB shaping gain over an unshaped system that uses a lower modulation order to achieve the same SE.**

***Index Terms*—Faster-than-Nyquist (FTN), time-frequency packing (TFP), probabilistic shaping (PS), decision feedback equalization (DFE), expectation propagation (EP), superchannel.**

I. Introduction

TO ENABLE high spectral efficiency (SE) in the next generation optical fiber networks, probabilistic shaping (PS) using higher order modulation (HoM) formats has received a considerable attention more recently [1]–[8]. In addition to providing an ultimate shaping gain up to 1.53 dB, PS can also offer rate adaptivity for systems employing a fixed forward-error-correction (FEC) code. In the recent past, several computationally efficient PS architectures, such as the symbol-level constant composition distribution matching (SL-CCDM) based probabilistic amplitude shaping (PAS) [1]–[3] and the bit-level product distribution matching (PDM) [4], [5] based PS scheme, have grown popularity for efficiently realizing FEC coded PS systems.

Another interesting avenue to improve the SE is to enable time-frequency packing (TFP) wavelength-division-multiplexing (WDM) superchannel transmission [9], [10]. Due to practical limitations of the opto-electronics to implement high baud-rate single-carrier systems, such a superchannel transmission is an attractive choice to achieve high data rates [11]. However, TFP introduces inter-symbol interference (ISI) and/or inter-carrier interference (ICI). Furthermore, the bandwidth (BW) limitations of the reconfigurable optical add-drop multiplexers (ROADMs) present in the fiber link may further degrade the signal quality of the edge sub-channels (SCs) in a superchannel. We have shown in our previous work [6] that combining TFP with PS can be doubly beneficial to counter such narrow filtering effects, since TFP can reduce the aggregate signal BW by packing the symbols in time and/or frequency, while still retaining the shaping gains provided by PS. Reaping these benefits of PS-TFP superchannels entails successful mitigation of the interference. The application of HoM formats may further complicate interference cancellation because the implementation of the well-known turbo equalization schemes, such as the Bahl-Cocke-Jelinek-Raviv (BCJR) equalizer, is computationally expensive for signal constellations larger than 16-ary quadrature amplitude modulation (16-QAM) [12]. On the other hand, low-complexity linear equalization is usually restrictive in performance [6]. This warrants the investigation of computationally efficient, high-performance, alternative equalization methods.

In this work, we present a spectrally efficient PS-TFP Terabit superchannel transmission employing 64-QAM modulation format, in the presence of severe BW limitations due to the wavelength selective switches (WSSs) of the ROADMs. We apply an expectation propagation (EP) based decision feedback equalization (DFE) [13] for the ISI cancellation (ISIC), in conjunction with a turbo parallel interference cancellation (PIC) based ICI cancellation (ICIC) [14]. Since the computational complexity of EP-DFE is independent of the modulation order [13], it is suitable for HoM formats. Our numerical results show that, by employing the EP-DFE ISIC together with turbo-PIC ICIC, the proposed PS-TFP superchannels enabling 1.8 Tbps data rate and occupying 300 GHz BW offer up to 1.05 dB packing gain over an unpacked Nyquist WDM system using the same modulation format. Moreover, the proposed superchannels yield a 1.15 dB shaping gain over an unshaped TFP transmission

Manuscript received August 10, 2021; revised October 14, 2021; accepted October 25, 2021. Date of publication October 29, 2021; date of current version November 11, 2021. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Huawei Tech., Canada. *(Corresponding author: Lutz Lampe.)*

Mrinmoy Jana and Lutz Lampe are with the Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: mjana@ece.ubc.ca; lampe@ece.ubc.ca).

Jeebak Mitra is with the Huawei Tech., Ottawa, ON L3R 5A4, Canada (e-mail: jeebak.mitra@huawei.com).

Digital Object Identifier 10.1109/JPHOT.2021.3123810

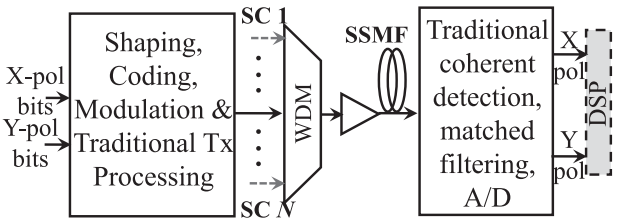


Fig. 1. Transmitter and receiver block diagram.

that employs a lower modulation order to achieve the same SE.

II. PS-TFP System Model

A simplified block diagram of a dual-polarized (DP) PS-TFP WDM superchannel transmission is shown in Fig. 1. Uniformly distributed bits for each polarization (pol.) stream and SC undergo traditional transmitter-side digital signal processing (DSP) of a typical coded PS WDM system [3], [6], comprised of distribution matching (DM), low-density parity-check (LDPC) encoding, QAM mapping, and a root raised cosine (RRC) pulse-shaping with a roll-off factor α. For DM, we either use a SL-CCDM [1]–[3] for high performance or a PDM [4], [5] for low complexity. For the TFP transmission, we use a time-packing factor of τ and a frequency-packing factor of ξ, such that $\tau = \xi = 1$ corresponds to a Nyquist WDM system [6]. After the transmitter DSP, the WDM signal is generated using electrical digital-to-analog converters and an optical frontend employing a driver and a Mach-Zehnder (MZ) modulator. Thereafter, the optical signal is transmitted over multiple spans of a standard single-mode fiber (SSMF), whereby the signal suffers from various fiber impairments and optical filtering due to WSSs. At the receiver, each SC is coherently detected, and after analog-to-digital (A/D) conversion and matched filtering, the received samples are fed as inputs to the receiver-side DSP unit, as detailed in the next section.

III. Interference Estimation & Mitigation

Schematics of the receiver DSP are shown in Fig. 2. After the chromatic dispersion (CD) equalization, the received samples of the X and Y pol. are filtered together by a 2×2 adaptive polarization mode dispersion (PMD) equalizer. We modify the least mean square (LMS) based adaptation algorithm of the PMD equalizer to estimate additional parameters as described below.

A. Interference Channel & Noise Covariance Estimation

To design the interference mitigation module for our system, our first key observation is that a channel estimation algorithm must be in place to estimate the impulse response of the interference channel. This is motivated by the fact that (a) adaptively estimating the ISI and ICI taps, rather than truncating the known TFP channel, minimizes the residual interference [6], and (b) interference stemming from additional filtering in the optical link, such as the BW limitations caused by the electrical/optical filters and the WSSs, can also be mitigated. To this end, our second key observation is that an EP-DFE ISIC (see Section III-C) requires the statistics of the equivalent additive noise samples [13], which are colored for TFP systems [6], [12]. In light of the above, we modify the traditional PMD filter adaptation algorithm to also estimate the ISI channel $h_{\rm s}$, the ICI channel $h_{\rm c}$, and the noise covariance matrix $\boldsymbol{R_{ww}}$. To accomplish this, we transmit a continuous block of pilot symbols at the beginning of data transmission, followed by periodic pilots inserted uniformly across the subsequent data frames. We apply the LMS update rules provided in [6, Lemma 1] to estimate $h_{\rm s}$ and $h_{\rm c}$. For the estimation of $\boldsymbol{R_{ww}}$, we accumulate the LMS error signal over multiple pilot-blocks to extract the noise statistics.

Algorithm 1: EP-DFE Algorithm Accounting for the Colored Noise Covariance Matrix $\boldsymbol{R_{ww}}$.

Input: $\boldsymbol{y}, \boldsymbol{H}, \boldsymbol{R_{ww}}$
Output:
Initialize with $L_{\rm a}^{(0)}(\boldsymbol{d}_k) = 0,$ for all symbol index k.
for GI index $G = 0 : \mathcal{G} - 1$**do**
  $\forall k = 0, 1, \ldots, K-1$, use $L_{\rm a}^{(G)}(\boldsymbol{d})$ to compute $\mathcal{P}_k^{(G)}$ as in [13, (17)].
  Set $(x_k^{{\rm d}(0)}, v_k^{{\rm d}(0)}) \leftarrow (x_k^{\rm p}, v_k^{\rm p})$ using [13, (21)].
  for SI index $S = 0 : \mathcal{S} - 1$**do**
    for Symbol index $k = 0 : K - 1$**do**
      Obtain $(x_k^{{\rm e}(S)}, v_k^{{\rm e}(S)})$ using $\boldsymbol{\Sigma}_k^{\rm d}$ from (1), followed by [13, (27)].
      Use [13, (19)-(20)] to update $\mathcal{D}_k^{(S+1)}$, and generate $(x_k^{{\rm d}(S+1)}, v_k^{{\rm d}(S+1)})$ with [13, (22)-(23)].
    end for
  end for
  Compute $L_{\rm e}^{(G)}(\boldsymbol{d}_k)$ using $\mathcal{D}_k^{(G,\mathcal{S}-1)}$ with [13, (29)], $\forall k$, and provide them to the decoder, to obtain $L_{\rm a}^{(G+1)}(\boldsymbol{d}_k), \forall k$.
end for

B. Turbo-PIC ICIC

Once the ICI channel $h_{\rm c}$ is estimated, we employ a turbo PIC ICIC algorithm [6] for each SC to mitigate the effects of ICI due to the neighboring SCs. For this, we use $h_{\rm c}$ and the a-posteriori log-likelihood ratios (LLRs) fed back from the LDPC decoders of the adjacent SCs to iteratively subtract the soft estimates of the ICI from the received samples [14]. Thereafter, the received symbols are fed as inputs to the EP-DFE ISIC as detailed next.

C. EP-DFE ISIC

DFE is a well-known nonlinear equalization technique. However, traditional DFE involves hard-decisions of symbols, which introduce error propagation [15]. To reduce its impact, soft-decisions assisted turbo DFE, such as the EP-DFE [13], has been investigated for FEC coded systems. The block-digram of our EP-DFE implementation is shown in Fig. 2, whereby the equalizer exchanges soft information with the demapper.

Details of such message passing mechanism are provided in Algorithm 1. For this, the method presented in [13, Algorithm 1]

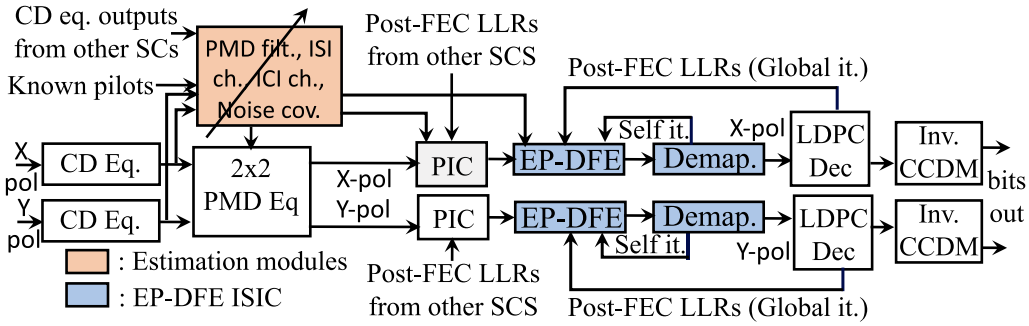


Fig. 2. Receiver-side DSP processing for each SC of the proposed PS-TFP superchannel transmission.

is properly adjusted to account for the colored noise samples of a TFP transmission, and applied to a practical WDM system, in tandem with the channel estimation method described in Section III-A. In Algorithm 1, k, $\boldsymbol{y}$, $\boldsymbol{H}$, $\mathcal{G}$, $\mathcal{S}$, $L_{\rm a}$, and $L_{\rm e}$ denote the symbol index, the K received symbols vector as defined in [13, (2)], the Toeplitz matrix representation of the estimated ISI channel $h_{\rm s}$ as defined in [13, (2)], the maximum number of "global iterations (GI)" between the LDPC decoder and the equalizer, the maximum number of "self iterations (SI)" between the equalizer and the demapper, apriori LLRs, and extrinsic LLRs, respectively. By accounting for the noise covariance matrix $\boldsymbol{R_{ww}}$ estimated as per Section III-A, we calculate the following parameter to be used for the EP-DFE metrics computation

$$\boldsymbol{\Sigma}_k^{\rm d} = \boldsymbol{R_{ww}} + \boldsymbol{H}_k \boldsymbol{V}_k^{\rm d} \boldsymbol{H}_k^{\rm H} \,, \tag{1}$$

where the superscript d corresponds to the metric-computations by the EP-DFE "demapper" [13], $\boldsymbol{H}_k$ is defined as in [13, (3)], and $\boldsymbol{V}_k^{\rm d}$ is defined as in [13, (25)]. We note that this is in contrast with [13], which assumes white noise samples at the EP-DFE input, and thereby, considers $\boldsymbol{R_{ww}}$ to be a scaled identity matrix.

As shown in Algorithm 1, in every global iteration for each pol., the LDPC decoder feeds back the extrinsic LLRs to the equalizer, and the equalizer, in turn, performs $\mathcal{S}$ self iterations with the demapper to exchange messages in the form of soft means and variances. Finally, at the end of $\mathcal{G}$ global iterations, hard-decoded bits are passed on to the inverse-DM module. The values of $\mathcal{G}$ and $\mathcal{S}$ are design parameters subject to a given receiver complexity budget.

IV. Numerical Results

To present our numerical results, we consider DP 64-QAM PS-TFP superchannels having 3 SCs, with a per-SC data rate of 600 Gbps, such that the 1.8 Tbps superchannels are packed within an aggregate BW of 300 GHz. We simulate a 800 km SSMF with 17 ps/nm/km CD, 0.1 ps/$\sqrt{\rm km}$ PMD, and 1 WSS in the optical link with a 3-dB BW of 300 GHz. We adopt $\alpha = 0.1$, LDPC codes with a rate $r = 0.8$ and codeword-length 64800 bits[1], and 6% pilots density. Moreover, we consider varying amount of TFP compression by changing the values of τ and ξ, and varying Maxwell-Boltzman (MB) parameter that determines the amount of shaping. Baud rates per SC are varied as per Table I. The parameter η in Table I denotes the uncoded bits per symbol (BPS) per pol. Furthermore, 9 taps for the estimated ICI channel, maximum GI count $\mathcal{G}$ of 10, and different values for the maximum SI count $\mathcal{S}$, estimated ISI channel taps length $L_{\rm ce}$, and DFE taps length $N_{\rm dfe}$ are considered. Perfect frequency synchronization and phase-lock are assumed for our simulations.

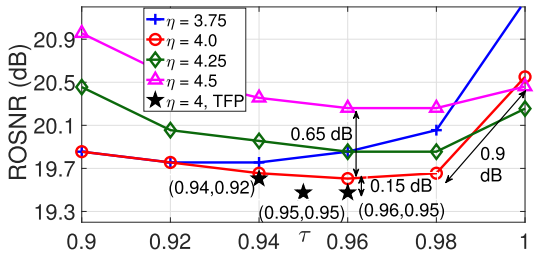
TABLE I
Simulation Configurations

Baud rate per SC (Gbaud)	Superchannel BW without TFP (GHz)	WSS 3-dB BW (GHz)	MB parameter	$\eta=$ Uncoded BPS per pol. bits/symb./pol.
88.33	291.49	300	0.075	4.50
93.53	308.65		0.087125	4.25
99.38	327.95		0.10065	4.00
106	349.80		0.1159	3.75

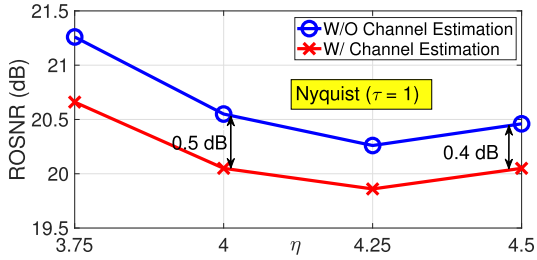
In Fig. 3(a), we show the required optical signal-to-noise ratio (ROSNR) as a function of τ for the *time-only-packed* systems to achieve an error-free transmission, with different η values. A more shaped system requires a higher baud rate to achieve the same data rate, and consequently, suffers more from the BW limitations of the WSSs. Lower values of τ can reduce such effects by decreasing the superchannel BW. Therefore, an optimal combination of η and τ is required to produce the best ROSNR performance [6]. For example, $\eta = 4$ bits/symbol/pol. together with $\tau = 0.96$ is shown to yield the lowest ROSNR in Fig. 3(a), which corresponds to a 0.9 dB packing gain compared to the unpacked Nyquist WDM superchannel with $\tau = 1$ for the same η. Moreover, by compressing the signal in *both time and frequency*, an optimal TFP configuration, corresponding to $\tau = 0.96, \xi = 0.95$, achieves 0.15 dB additional ROSNR reduction, and thereby, offering an aggregate packing gain of 1.05 dB.

Next, we highlight the ability of the EP-DFE to mitigate the WSS induced ISI in a Nyquist WDM system in the absence of TFP. For this, we show the ROSNR as a function of η for $\tau = 1$ in Fig. 3(b). As shown in the figure, EP-DFE in tandem with

[1]The LDPC codes used in this paper are compliant with the second generation digital video broadcasting standard for satellite (DVB-S2) applications [16]. They are encoded as irregular repeat accumulate codes. LDPC decoding is performed by iterative standard message passing algorithm [16], with the maximum number of LDPC internal iterations set to 50.



(a) ROSNR vs. τ.



(b) ROSNR vs. η for the Nyquist WDM system.

Fig. 3. ROSNR performance. $L_{\rm ce}=7$, $N_{\rm dfe}=9$, $\mathcal{S}=3$ unless otherwise mentioned.

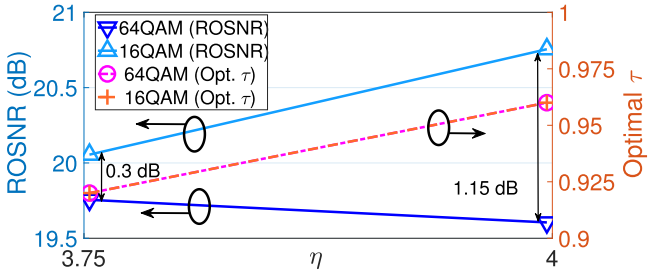


Fig. 4. Lowest ROSNR and the optimal τ vs. η.

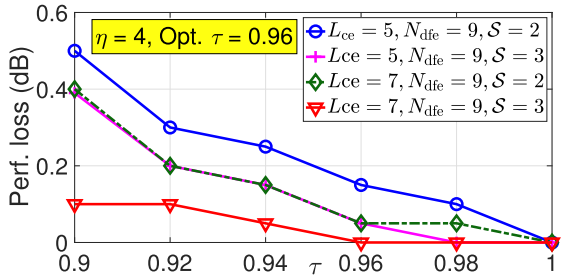


Fig. 5. Performance degradation compared to the case with $L_{\rm ce}=13$, $N_{\rm dfe}=15$, $\mathcal{S}=3$.

the proposed channel estimation algorithm achieves 0.4 0.5 dB performance improvement over a traditional system that does not employ any ISIC algorithm.

In Fig. 4, we show the following two sets of plots: (i) we compare the ROSNR performance at the optimal τ between a 16-QAM and a 64-QAM based PS-TFP superchannel systems to achieve the same SE and data rate, and (ii) the optimal τ values for the said 16-QAM and the 64-QAM based systems. The plots corresponding to (i) above shows that, at the optimal time packing, a 64-QAM system outperforms a 16-QAM based transmission by 0.3 dB and 1.15 dB for $\eta=3.75$ bits/symbol/pol. and 4 bits/symbol/pol., respectively. For $\eta=4$ bits/symbol/pol., the 16-QAM based system corresponds to an unshaped transmission, while the 64-QAM system employs a shaped constellation. This results in an improved ROSNR for the 64-QAM transmission, which can be primarily attributed to the shaping gain. However, with $\eta=3.75$ bits/symbol/pol., both systems transmit shaped constellations, and thereby, reducing the ROSNR margin between the two modulation formats. Furthermore, we also observe from Fig. 4 that the optimal τ values, corresponding to (ii) above, are same for both 16-QAM and 64-QAM systems. For details on packing gains and how ROSNR changes with varying τ for a 16-QAM based system, the interested reader is referred to [6].

The ROSNR performance degradation with different EP-DFE parameters compared to a benchmark configuration with $L_{\rm ce}=13$, $N_{\rm dfe}=15$, $\mathcal{S}=3$ is shown in Fig. 5. The plots indicate that the lower values of $L_{\rm ce}$ and/or $\mathcal{S}$ incur higher performance loss, which becomes more significant as τ reduces. Such an observation suggests that the mitigation of a stronger TFP induced interference requires longer estimated channel taps and EP-DFE taps, and more SI, as expected.

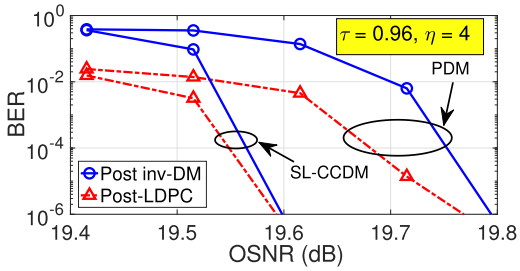
Finally, in Fig. 6(a) and Fig. 6(b), we investigate the post-LDPC and post-inverse DM bit-error-rate (BER) performance of the proposed superchannels, as a function of the optical signal-to-noise ratio (OSNR), to compare (a) SL-CCDM vs. PDM based PS schemes, and (b) EP-DFE vs. BCJR based ISIC methods, respectively, at $\tau=0.96$ and $\eta=4$ bits/symbol/pol. For (a), an SL-CCDM is shown to yield a 0.18 dB gain over PDM, at the price of a higher serialism requirement for the DM implementation [3]–[5]. For (b), EP-DFE is shown to outperform a 64-state BCJR equalizer, corresponding to $L_{\rm ce}=5$, by 0.2 dB. To explain such an observation, we argue that the application of an estimated 5-tap channel for the ISIC algorithm, instead of the actual impulse response, introduces sub-optimality even for the BCJR ISIC, otherwise known as an optimal equalizer[2].

V. Further Discussions

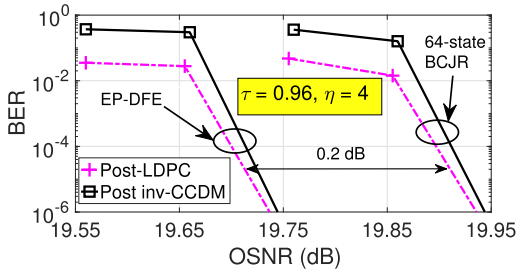
A. LDPC Code Rate

For the numerical results presented in Section IV, we have used a fixed LDPC code rate $r=0.8$, and different values of SE were achieved by varying the amount of shaping via the parameter η as per Table I. To accomplish different target SE values, an alternative approach is to vary r. However, for

[2]Due to the prohibitive complexity of BCJR, evaluating the performance with the actual (longer) channel was challenging for this case. However, with a lower modulation order and a mild packing, EP-DFE and BCJR performances were identical when the actual ISI impulse response was considered by both ISIC algorithms.



(a) BER vs. OSNR for SL-CCDM and PDM. $L_{\rm ce} = 7,\ N_{\rm dfe} = 9,\ \mathcal{S} = 3$



(b) BER vs. OSNR for EP-DFE ($L_{\rm ce} = 5,\ N_{\rm dfe} = 9,\ \mathcal{S} = 2$) and a 64-state BCJR.

Fig. 6. Post-LDPC and post-inverse-DM BER performance vs. OSNR.

TABLE II
INVESTIGATING DIFFERENT VALUES FOR THE LDPC CODE RATE r.

η bits/symb./pol.	r	τ	Baud rate per SC (Gbaud)	Superchannel TP BW (GHz)	ROSNR (dB)
4.25	0.75	0.96	99.77	316.07	19.96
4.25	0.8		93.53	296.30	19.86
4.00	0.8		99.38	314.84	19.66

TABLE III
INVESTIGATING DIFFERENT VALUES FOR THE RRC ROLL-OFF α.

η bits/symb./pol.	α	τ	Baud rate per SC (Gbaud)	Superchannel TP BW (GHz)	ROSNR (dB)
4.00	0.05	0.96	99.38	300.53	19.66
4.00	0.1		99.38	314.84	19.66

TABLE IV
COMPUTATIONAL COMPLEXITY

Operations	BCJR	EP-DFE (Gauss. Elimn.)
ADD/SUB	$8M^{\frac{L_{\rm ce}-1}{4}}-4$	$\mathcal{S}\left(\frac{N_{\rm dfe}(N_{\rm dfe}-1)(2N_{\rm dfe}+5)}{6}+30\right)$
MUL/DIV	$16M^{\frac{L_{\rm ce}-1}{4}}+2$	$\mathcal{S}\left(\frac{N_{\rm dfe}\left(N_{\rm dfe}^2+3N_{\rm dfe}-1\right)}{3}+42\right)$
Nonlin.	$8M^{\frac{L_{\rm ce}-1}{4}}+2$	0

(a) Parametric expression.

Operations	BCJR		EP-DFE (Gauss. Elimn.)	
	$L_{\rm ce} = 5$	$L_{\rm ce} = 7$	$\mathcal{S} = 2$	$\mathcal{S} = 3$
ADD/SUB	508	4092	612	918
MUL/DIV	1026	8194	726	1089
Nonlin.	514	4098	0	0

(b) Specific example for $M = 64$, $N_{\rm dfe} = 9$.

this goal, varying η is more beneficial than varying r due to the following two reasons: (a) PS facilitates more flexible and fine-resolution rate-adaptivity [1]–[3], and (b) PS offers shaping gains. To investigate this in more detail, we have performed additional simulations with $r = 0.75$ to achieve the same target data rate of 600 Gbps per SC in the presence of similar BW restrictions imposed by the WSS filters.

In Table II, we show the ROSNR values for the following three configurations: (i) $\eta = 4.25$ bits/symbol/pol., $r = 0.75$, (ii) $\eta = 4.25$ bits/symbol/pol., $r = 0.8$, and (iii) $\eta = 4$ bits/symbol/pol., $r = 0.8$. For this investigation, we have used $\tau = 0.96$, and all other simulation parameters are kept same as in Fig. 3(a). By comparing (i) with (ii) from the first two rows of Table II, we observe that, for the same η, (ii) with $r = 0.8$ achieves 0.1 dB better ROSNR compared to (i) with $r = 0.75$. This can be attributed to the larger TP superchannel BW of (i), which suffers from more WSS induced ISI. Finally, a comparison between (i) and (iii) from the first and third rows, respectively, of Table II reveals that, for the same data rate and a similar TP superchannel BW (and hence, a similar net SE value), (iii) with $r = 0.8$ and $\eta = 4$ bits/symbol/pol. outperforms (i) with $r = 0.75$ and $\eta = 4.25$ bits/symbol/pol. by a margin of 0.3 dB. Such a performance improvement at the same SE can be attributed to the additional shaping gain achieved by (iii) by employing a lower η.

B. RRC Roll-off

To analyze the effects of reduced roll-off factors on the PS-TFP system performance, we have performed additional simulations with $\alpha = 0.05$, as opposed to $\alpha = 0.1$ used in Section IV. A fixed value of $\tau = 0.96$ and $\eta = 4$ bits/symbol/pol. are considered for this investigation. All other simulation parameters are kept same as in Fig. 3(a). The summary of the ROSNR results is shown in Table III. The table entries suggest that the two systems having different α values perform similarly. While the superchannel with $\alpha = 0.05$ occupies lower aggregate BW, and hence, suffers from less WSS-ISI compared to that with $\alpha = 0.1$, the system with $\alpha = 0.05$ suffers from more TP induced ISI for a given time packing factor τ [17]. Therefore, the overall system performance is governed by the combined impact of these two sources of interferences. When α reduces further, careful optimization of the packing, shaping and the receiver DSP parameters is required to meet the desired performance-complexity target.

VI. COMPUTATIONAL COMPLEXITY

The computational complexity numbers for the EP-DFE and the BCJR ISIC are furnished in Table IV, assuming a Gaussian elimination method to implement the matrix inversion of $\mathbf{\Sigma}_k^{\rm d}$ as required in [13, (26)]. These values indicate that the complexities of both ISIC methods are comparable for the 64-QAM based systems when $L_{\rm ce} = 5$. However, for larger values of $L_{\rm ce}$, which were shown to offer performance improvements for the

EP-DFE in Section IV, the complexity of BCJR is substantially higher. We also remark that the EP-DFE complexity can be further reduced by employing computationally more efficient matrix inversion techniques, such as the Strassen method and the Coppersmith-Winograd algorithm [13], which we have not explored in this work.

VII. Conclusion

To improve the SE of the optical fiber systems, we have considered PS-TFP WDM superchannels transmission using HoM formats. However, TFP introduces ISI and/or ICI. Moreover, the WSSs present in the fiber link may impose severe BW limitations for the edge SCs. To counter such interferences, we investigated an EP-DFE ISIC in tandem with a turbo-PIC ICIC. Our numerical results showed that, by optimizing of the shaping and packing parameters, the proposed PS-TFP superchannels enabling 1.8 Tbps data rates achieved a 1.05 dB packing gain over an unpacked Nyquist WDM system using the same modulation format, and a 1.15 dB shaping gain over an unshaped TFP system that uses a lower modulation order to achieve the same SE.

References

[1] G. Böcherer, P. Schulte, and F. Steiner, "Probabilistic shaping and forward error correction for fiber-optic communication systems," *J. Lightw. Technol.*, vol. 37, no. 2, pp. 230–244, Jan. 2019.

[2] G. Bosco, "Advanced modulation techniques for flexible optical transceivers: The rate/reach tradeoff," *J. Lightw. Technol.*, vol. 37, no. 1, pp. 36–49, 2019.

[3] Y. Gültekin, T. Fehenberger, A. Alvarado, and F. M. J. Willems, "Probabilistic shaping for finite blocklengths: Distribution matching and sphere shaping," *Entropy*, vol. 22, no. 5, May 2020, Art. no. 581.

[4] G. Böcherer, P. Schulte, and F. Steiner, "High throughput probabilistic shaping with product distribution matching," Feb. 2017, *arXiv: 1702.07510*.

[5] M. Pikus and W. Xu, "Bit-level probabilistically shaped coded modulation," *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 1929–1932, May 2017.

[6] M. Jana, L. Lampe, and J. Mitra, "Probabilistic shaping in time-frequency-packed terabit superchannel transmission," *IEEE Photon. Technol. Lett.*, vol. 32, no. 17, pp. 1065–1068, Jul. 2020.

[7] M. Jana, L. Lampe, J. Mitra, and C. Li, "Jointly shaped dual polarization systems," in *Proc. Opt. Fiber Commun. Conf.*, 2021, pp. 1–3.

[8] M. Jana, L. Lampe, J. Mitra, and C. Li, "Multi-dimensional probabilistic shaping for optical superchannels," in *Proc. Eur. Conf. Opt. Commun.*, 2021, pp. 1–4.

[9] L. Li, Z. Xiao, L. Liu, and Y. Lu, "Non-orthogonal WDM systems with faster than Nyquist technology," in *Proc. Opt. Fiber Commun. Conf.*, 2019, pp. 1–3.

[10] M. Jana, L. Lampe, and J. Mitra, "Design of time-frequency packed WDM superchannel transmission systems," *J. Lightw. Technol.*, vol. 38, no. 24, pp. 6719–6731, Aug. 2020.

[11] M. Mazur, J. Schröder, M. Karlsson, and P. A. Andrekson, "Joint superchannel digital signal processing for effective inter-channel interference cancellation," *J. Lightw. Technol.*, vol. 38, no. 20, pp. 5676–5684, Oct. 2020.

[12] M. Matar, M. Jana, J. Mitra, L. Lampe, and M. Lis, "A turbo maximum-a-posteriori equalizer for faster-than-Nyquist applications," in *Proc. IEEE Int. Symp. Field-Programm. Custom Comput. Machines*, 2020, pp. 1–5.

[13] S. Şahin, A. M. Cipriano, C. Poulliat, and M. L. Boucheret, "Iterative equalization with decision feedback based on expectation propagation," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4473–4487, Oct. 2018.

[14] F. Rusek and J. B. Anderson, "Multistream faster than Nyquist signaling," *IEEE Trans. Commun.*, vol. 57, no. 5, pp. 1329–1340, May 2009.

[15] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*. Hoboken, NJ, USA: Wiley, 2002.

[16] H. Ge, "Investigation of LDPC code in DVB-S2," Master's thesis, Dept. Elect. Eng., Linköping Univ., 2012.

[17] F. Rusek, "Partial response and faster-than-Nyquist signaling," Ph.D. dissertation, Dept. Elect. Inf. Technol., Lund Univ., 2007.