Design of Time-Frequency Packed WDM Superchannel Transmission Systems

Mrinmoy Jana, Lutz Lampe, and Jeebak Mitra

Abstract—We consider time and frequency-packing (TFP) wavelength division-multiplexing (WDM) superchannel systems for longhaul optical fiber communication. Employing very high baud rates in conjunction with higher-order modulation formats is challenging due to practical constraints. TFP superchannels therefore become an attractive choice to facilitate high data rates with improved spectral efficiency. However, TFP introduces inter-symbol interference (ISI) and/or inter-carrier interference (ICI) that necessitate efficient interference handling techniques. Moreover, phase noise (PN) due to wide laser linewidth (LLW) may cause significant signal distortion in WDM transmission, and the presence of ISI and ICI in TFP systems further complicates carrier phase recovery (CPR). In this paper, we propose a spectrally efficient TFP superchannel design equipped with powerful interference cancellation and PN mitigation techniques, targeting Terabit-per-second (Tbps) data rates. First, we propose a joint ISI and ICI channel estimation algorithm coupled with polarization-recovery and a coarse PN cancellation method. Second, we investigate two iterative CPR algorithms to mitigate the distortion due to the residual PN. Third, a combination of successive and parallel interference cancellation methods for ICI mitigation is investigated in tandem with turbo ISI equalization. The effectiveness of the proposed algorithms is demonstrated through computer simulations of a coded TFP superchannel system in the presence of linear fiber impairments, targeting a throughput of 1.2 Tbps. Numerical results suggest that the proposed TFP design offers more than 2 dB performance gains and as high as 960 km transmission distance improvement over existing competitive super-Nyquist designs. Simulation results also indicate that the proposed design exhibits excellent tolerance to high LLWs and aggressive optical filtering stemming from cascaded reconfigurable optical add-drop multiplexers in the fiber link.

Index Terms—Faster-than-Nyquist (FTN), time-frequency packing (TFP), phase noise (PN), superchannel, wavelength division-multiplexing (WDM), inter-symbol interference (ISI), inter-carrier interference (ICI), factor graph, laser linewidth, wavelength selective switch (WSS).

I. INTRODUCTION

TERABIT-per-second (Tbps) data rates per carrier are being targeted for the next generation optical systems to cope with the increasing traffic in the optical fiber networks [1]–[3]. An attractive choice to enable high data rates over longhaul optical links is to apply time-frequency packing (TFP) super-Nyquist wavelength-division-multiplexing (WDM) superchannel transmission [1], [4]–[9]. As an alternative to employing higher-order modulation

Mrinmoy Jana and Lutz Lampe are with the Department of Electrical and Computer Engineering, University of British Columbia, BC V6T 1Z4, Canada. Email: mjana@ece.ubc.ca, lampe@ece.ubc.ca. Jeebak Mitra is with Huawei Technologies, Ottawa, ON K2K 3J1, Canada. Email: jeebak.mitra@huawei.com.

(HoM) formats and very high baud rates, that are challenging due to nonlinear impairments in fibers and practical limitations of the opto-electronics, TFP superchannels offer spectral efficiency (SE) improvements by allowing controlled overlap of the sub-channels (SCs) in time or frequency or both, and thereby, introducing inter-symbol interference (ISI) and/or inter-carrier interference (ICI) [10]. In the existing literature, time-only [4]–[6], [11] and frequency-only [8], [9], [12], [13] packing have been considered quite extensively, which introduce either ISI or ICI. However, packing the symbols in *both* time and frequency dimensions can theoretically provide higher achievable rates [10], at the price of introducing ISI and ICI simultaneously. Another practical challenge that affects WDM transmission is the signal distortion caused by phase noise (PN) stemming from the spectral linewidth of the transmitter and receiver laser beams [14], [15]. The application of HoM formats makes the communication systems more sensitive to PN. Moreover, the impact of PN is more severe for TFP superchannels, since the presence of ISI and ICI precludes the direct application of off-the-shelf PN mitigation algorithms that are tailored to Nyquist transmission [15]–[17]. Therefore, we need sophisticated signal processing tools to counter TFP interference together with powerful PN cancellation strategies.

A. Background on TFP ISI, ICI and PN Mitigation

For super-Nyquist WDM systems, one approach adopted in the literature [4]–[6], [11] considers suppressing the ICI through aggressive transmit-side filtering of the individual SCs, and the resulting ISI is equalized by Bahl-Cocke-Jelinek-Raviv (BCJR) based ISI cancellation (ISIC) methods. Another body of works [8], [9], [12], [13] allows only spectral overlap, whereby the ICI is mitigated via linear or nonlinear ICI cancellation (ICIC) schemes. However, packing the symbols in only one dimension can be restrictive in terms of achievable rate [10]. While some pioneering works on super-Nyquist systems [10], [18] explored time and frequency packed transmission in additive white Gaussian noise (AWGN) channel scenarios, no proper consideration was given to fiber-optical impairments. Our previous work [1] presents ISIC and ICIC algorithms for optical TFP systems based on the *known* truncated TFP interference channel without performing any channel estimation. As we will show in Section VI, such a strategy cannot achieve the performance of the channel-adaptive interference cancellation proposed in this paper. Moreover, the parallel interference cancellation (PIC) based ICIC approach considered in [1], [10], [18] lacks the benefits of sequential scheduling in a successive interference cancellation (SIC) structure, as shown through our numerical

results. Another interesting method in the existing literature is the application of precoding in optical TFP systems [19]–[23] to pre-equalize the TFP interference at the transmitter. While this is conceptually appealing, the functionality of such precoding is generally limited to a restricted range of time and frequency compression [21]. Furthermore, most of the above works [1], [10], [18], [21] assume perfect carrier phase recovery (CPR). However, ISI/ICI equalization and PN compensation tasks are not always modular, which invites a joint mitigation approach [4], [24]–[26].

PN due to the transmitter and receiver laser linewidth (LLW) causes severe signal distortion in WDM systems [14]–[16], [27], [28]. Conventionally, in Nyquist WDM systems, a feedforward blind phase search (BPS) algorithm is used for CPR [16], which employs hard symbol decisions of the received symbols. However, the presence of both ISI and ICI in TFP WDM systems precludes the feasibility of making error-free hard symbol decisions preceding the forward error correction (FEC) decoder, which renders the BPS algorithm unsuitable for the considered super-Nyquist transmission. More recently, a CPR algorithm based on principal component analysis (PCA) has been presented for Nyquist WDM systems employing square constellations [15]. However, to extract the phase information from the principal components, such a method exploits the geometry of the signal constellation, which gets severely distorted by the ISI and ICI in TFP systems. For the same reason, other sophisticated iterative PN compensation algorithms suitable for Nyquist WDM systems, such as the factor-graph (FG) based CPR [17], can not be directly applied to the super-Nyquist systems without proper consideration of the TFP ISI and ICI. The authors of [4] apply FG-based PN cancellation methods for their time-only packed (TP) systems. However, the CPR method in [4] also needs to be amended before applying to the TFP transmission considered here, because of the absence of ICI and the restriction to quarternary phase-shift keying (QPSK) in [4]. In the following, we briefly highlight the primary contributions of our work.

B. Contributions and Outline

In this paper, we consider both temporal and spectral overlap of the SCs to design a flexible TFP superchannel system under a linear fiber optical channel assumption. To accomplish this, we present signal processing techniques to efficiently mitigate TFP interference, linear fiber impairments, and PN. Following three novel contributions are made.

- A joint ISI and ICI channel estimation method, coupled with polarization mode dispersion (PMD) equalization and a coarse PN estimation (CPNE) is proposed. While the TFP-induced interference channel is a-priori known [1], [29], such an estimation is necessary because (a) estimating the exact number of ISI taps required by the BCJR-ISIC and forcing the remaining ISI-taps to zero can lead to significant performance improvement [4], and (b) interference stemming from additional sources in the optical link, such as the electrical/optical filters, can also be mitigated. In conjunction with PMD equalization and

interference channel estimation, we also jointly perform a coarse CPR, which is beneficial since (a) it facilitates better ISI and ICI channel estimation by minimizing the overall mean squared error (MSE) during the initial pilots transmission phase, and (b) it offers improved boot-strapping for more sophisticated iterative PN mitigation schemes to efficiently compensate for the residual PN.

- A novel, iterative, modulation-format-independent PN estimation method is proposed. Moreover, the FG-based solutions presented in [4] and [17] are also properly adapted to account for the TFP ISI and ICI. While the authors of [4] consider post-FEC hard decisions for the FG metric computations for the case of only TFP-ISI and QPSK transmission, we use soft information to improve robustness against error-propagation in HoM systems in the presence of both ISI and ICI.

- A serial-and-parallel combined interference cancellation (SPCIC) based ICIC scheme is presented. Different from the existing PIC structures [1], [10], [18], [30], the proposed SPCIC encapsulates a conceptual combination of the SIC and the PIC paradigms, to offer performance improvements. The proposed ICIC solution in tandem with BCJR-ISIC is shown to exhibit excellent tolerance to high LLW and aggressive optical filtering due to cascaded reconfigurable optical add-drop multiplexer (ROADM) nodes that may be present in the fiber links.

The remainder of the paper is organized as follows. The system model is introduced in Section II. In Section III, we propose a joint TFP-interference and a CPNE strategy. In Section IV, we investigate two iterative CPR algorithms. Proposed TFP interference mitigation techniques are detailed in Section V, followed by numerical results for optical TFP systems presented in Section VI and a discussion on the results in Section VII. Finally, Section VIII provides concluding remarks.

II. SYSTEM MODEL

In this work, we consider a dual-polarized (DP) TFP WDM transmission system for longhaul optical fiber communication [4], [5]. The schematics of such a system are shown in Fig. 1. For each of the X and Y polarization data streams of the mth SC, $m \in 1, 2, \cdots, N$, with N being the total number of SCs, a low density parity check (LDPC) coded and modulated data stream x_m is upsampled by a factor of 2 and digitally shaped by a root-raised cosine (RRC) pulse p with a roll-off factor β. The digital samples are then transformed into analog signals via digital-to-analog (D/A) converters, followed by conversion to the optical domain using Mach-Zehnder (MZ) modulators. The equivalent baseband transmitted signal for the X-polarization branch can be written as

$$s_{\text{x}}(t) = \sum_l \sum_m x_m[l] p(t - l\tau T) e^{\text{j}\left(2\pi\left(m - \frac{N+1}{2}\right)\Delta f t + \theta_{\text{tx}}^{(m)}(t)\right)}, \quad (1)$$

where $\Delta f = \xi \frac{1+\beta}{T}$ is the frequency-spacing between the adjacent SCs, τ and ξ are the time and frequency compression ratios, respectively, such that $\tau = \xi = 1$ corresponds to the Nyquist WDM system, l is the symbol index, $\frac{1}{\tau T}$ is
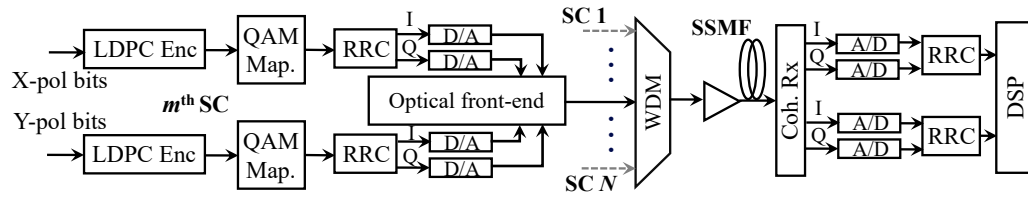
Fig. 1. TFP WDM system model.

the per-SC baud rate, and $\theta_{\mathrm{tx}}^{(m)}$ is the transmitter laser PN corresponding to the mth SC. The transmitted superchannel signal propagates through multiple spans of standard single-mode fibers (SSMFs), whereby the optical signal suffers distortion due to fiber-optical impairments, such as the chromatic dispersion (CD) and the PMD [31]–[34]. At the coherent receiver, the RRC matched-filtered digital samples of the mth SC, distorted by CD, PMD, amplified spontaneous emission (ASE) noise [32], and the receiver laser PN $\theta_{\mathrm{r,x}}^{(m)}$, are fed as inputs to the receiver digital signal processing (DSP) unit as detailed in the next section. The transmitter and the receiver PN for all SCs are modeled as Wiener processes, and the LLW of the lasers determine the standard deviation of these processes [35].

III. INTERFERENCE CHANNEL ESTIMATION AND CPNE

Inspired by the algorithm presented in [4], we propose an adaptive TFP receiver design, where we jointly estimate the 2×2 PMD equalizer tap co-efficients, the TFP-interference channel, and perform a coarse estimation of the laser PN. Different from [4], that restricts itself to ISI channel estimation only, we estimate the ISI and the ICI impulse responses simultaneously, coupled with a coarse CPR. Moreover, we also enforce the real-valued constraint on the ISI-taps adaptation algorithm [10], which offers an additional advantage of employing reduced-complexity BCJR equalization for square quadrature amplitude modulation (QAM) constellations, by separately processing the in-phase (I) and quadrature (Q) components. We will show through the numerical results in Section VI that such an adaptive channel estimation approach facilitates flexible superchannel transmission by offering significant performance advantages over non-adaptive TFP designs, such as [1].

A. DSP Modules

The receiver DSP datapath is shown in Fig. 2 for the mth SC of the TFP superchannel. The received signal for each SC is first downsampled to 2 samples per symbol, and then processed by a time-invariant CD equalizer implemented in the frequency domain through overlap-and-add method [31]–[33]. Thereafter, the CD compensated samples are filtered by a $\frac{T}{2}$-spaced 2×2 PMD equalizer to remove the cross-talk between the two polarization streams. Next, the received samples are further downsampled to produce T-spaced symbols that are processed by a 2-stage PN compensation algorithm for CPR. The estimates of the PMD filter coefficients, TFP-interference channel and a coarse PN are obtained through a pilot symbols-aided least mean square (LMS)-based adaptation algorithm

(see Section III-B). Following the polarization recovery and PN cancellation, the received signal for each polarization stream is processed by the turbo ICIC and the BCJR-ISIC modules. The output of the LDPC decoders in the form of log-likelihood ratios (LLRs) are fed back to mitigate TFP interference and PN, iteratively. The details of the iterative PN cancellation structure, together with ICIC and ISIC operations are presented in Sections IV and V, respectively.

B. LMS Update Equations

We use the column vectors $\boldsymbol{a}_k^{(m)}$ and $\boldsymbol{u}_k^{(m)}$ to denote the constellation symbols and the input samples to the 2×2 PMD equalizer, respectively. These are formulated by stacking the X and Y polarized signals of the mth SC at the kth sample time. The error signal for the LMS adaptation is then computed as the difference between the phase-rotated PMD filter output and the "desired signal" [4], whereby the desired signal computation incorporates the effects of the TFP interference into the clean pilot symbols, as shown below. The combined error signal vector with the X and Y polarization symbols is written as

$$\boldsymbol{\epsilon}_k^{(m)} = [\varepsilon_{k,\mathrm{x}}^{(m)}, \varepsilon_{k,\mathrm{y}}^{(m)}]^{\mathrm{T}} = \underbrace{\boldsymbol{z}_k^{(m)}}_{\text{filtered output}} \odot \underbrace{[\mathrm{e}^{-\mathrm{j}\theta_{\mathrm{x},k}^{(m)}}, \mathrm{e}^{-\mathrm{j}\theta_{\mathrm{y},k}^{(m)}}]^{\mathrm{T}}}_{\text{phase rotation}} - \underbrace{\boldsymbol{d}_k^{(m)}}_{\text{desired signal}} , \quad (2)$$

where

$$\boldsymbol{z}_k^{(m)} = [z_{\mathrm{x},k}^{(m)}, z_{\mathrm{y},k}^{(m)}]^{\mathrm{T}} = \sum_{i=0}^{N_{\mathrm{w}}-1} \boldsymbol{W}_{i,k}^{(m)} \boldsymbol{u}_{k-i}^{(m)} , \quad (3)$$

$$\boldsymbol{d}_k^{(m)} = [\delta_{k,\mathrm{x}}^{(m)}, \delta_{k,\mathrm{y}}^{(m)}]^{\mathrm{T}} = \sum_{j=-L_{\mathrm{s}}}^{L_{\mathrm{s}}} \boldsymbol{h}_{j,k}^{(m)} \odot \boldsymbol{a}_{k-j}^{(m)}$$
$$+ \sum_{n \neq m} \sum_{\nu=-L_{\mathrm{c}}}^{L_{\mathrm{c}}} \boldsymbol{g}_{\nu,k}^{(n,m)} \odot \hat{\boldsymbol{a}}_{k-\nu}^{(n)} . \quad (4)$$

In (2)-(4), $\varepsilon_{k,\mathrm{x/y}}^{(m)}$, $z_{k,\mathrm{x/y}}^{(m)}$ and $\delta_{k,\mathrm{x/y}}^{(m)}$ are respectively the error signal, PMD filtered output and the desired signal corresponding to the X or the Y polarization (the subscript x/y means "X respectively Y") for the mth SC at the kth sample time, $\boldsymbol{W}_{i,k}^{(m)}$ is the ith matrix-tap of the $\frac{T}{2}$-spaced PMD equalizer, N_{w} is the total number of PMD filter taps, $\boldsymbol{h}_{j,k}^{(m)}$ and $\boldsymbol{g}_{\nu,k}^{(n,m)}$ are the jth and the νth symbol-spaced ISI and ICI tap vector, respectively, with L_{s} and L_{c} being the total number of ISI and ICI channel taps, respectively, $[\cdots]^{\mathrm{T}}$ denotes the transpose of a vector, $\odot$ denotes the elementwise vector product, and
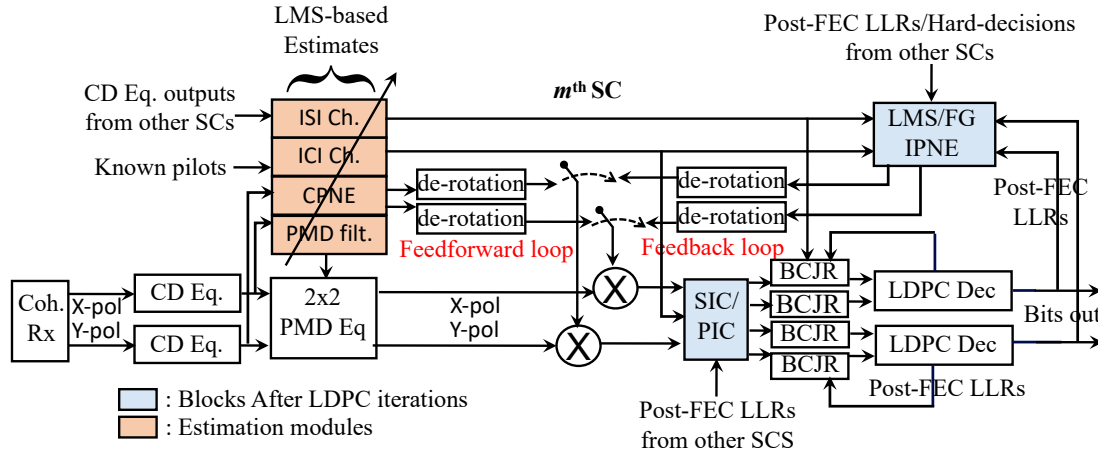
Fig. 2. Jointly estimating the PMD filter coefficients, interference channel, and PN. The estimation and post-LDPC decoding modules are highlighted with different colors.

$\hat{a}_k^{(n)} = a_k^{(n)} e^{\pm j2\pi\Delta fk}$ denotes the rotated constellation symbol for the nth SC, with $\pm$ sign determined from (1) depending on the relative positions of the SCs.

From (2), we define the MSE as

$$\mathcal{M}_{\text{tot}} = \mathbb{E}\left(\sum_{m=1}^{N} \|\boldsymbol{\epsilon}_k^{(m)}\|^2 \right) , \qquad (5)$$

where $\mathbb{E}(\cdot)$ denotes the expectation operator. For ease of formulation of the LMS update equations based on the gradient decent algorithm [25], we exploit the symmetry of the TFP channel by enforcing the real and symmetry assumption on the ISI impulse response [10], and conjugate symmetry on the ICI channel [1], such that

$$\boldsymbol{h}_{j,k}^{(m)} = \left(\boldsymbol{h}_{j,k}^{(m)} \right)^*, \quad m \in \{1, 2, \dots, N\}, -L_{\text{s}} \le j \le L_{\text{s}} ,$$

$$\boldsymbol{h}_{j,k}^{(m)} = \boldsymbol{h}_{-j,k}^{(m)}, \quad m \in \{1, 2, \dots, N\}, -L_{\text{s}} \le j \le L_{\text{s}} ,$$

$$\boldsymbol{g}_{\nu,k}^{(n,m)} = \left(\boldsymbol{g}_{-\nu,k}^{(n,m)} \right)^*, \quad n \in \{1, 2, \dots, N\}, -L_{\text{c}} \le \nu \le L_{\text{c}}.$$

Finally, computing the gradients $\frac{\partial \mathcal{M}_{\text{tot}}}{\partial \boldsymbol{W}_{\alpha,k}^{(m)}}$, $\frac{\partial \mathcal{M}_{\text{tot}}}{\partial \boldsymbol{h}_{\beta,k}^{(m)}}$, $\frac{\partial \mathcal{M}_{\text{tot}}}{\partial \boldsymbol{g}_{\gamma,k}^{(n,m)}}$ and $\frac{\partial \mathcal{M}_{\text{tot}}}{\partial \theta_{\text{x/y},k}^{(m)}}$, where $0 \le \alpha \le N_{\text{w}} - 1$, $0 \le \beta \le L_{\text{s}}$, $0 \le \gamma \le L_{\text{c}}$, the LMS update equations can be written as

$$\boldsymbol{W}_{\alpha,k+1}^{(m)} = \boldsymbol{W}_{\alpha,k}^{(m)} - \mu_{\text{w}} \left([e^{j\theta_{\text{x},k}^{(m)}}, e^{j\theta_{\text{y},k}^{(m)}}]^{\text{T}} \odot \boldsymbol{\epsilon}_k^{(m)} \right) \left(\boldsymbol{u}_{k-\alpha}^{(m)} \right)^{\text{H}}, (6)$$

$$\boldsymbol{h}_{\beta,k+1}^{(m)} = \boldsymbol{h}_{\beta,k}^{(m)} + \mu_{\text{h}} \text{Re}\left[\left(\boldsymbol{\epsilon}_k^{(m)} \odot \left(\boldsymbol{a}_{k-\beta}^{(m)} + \boldsymbol{a}_{k+\beta}^{(m)} \right)^* \right) \right] , \qquad (7)$$

$$\boldsymbol{g}_{\gamma,k+1}^{(n,m)} = \boldsymbol{g}_{\gamma,k}^{(n,m)} + \mu_{\text{g}} \left(\boldsymbol{\epsilon}_k^{(m)} \odot \left(\hat{\boldsymbol{a}}_{k-\gamma}^{(n)} \right)^* + \left(\boldsymbol{\epsilon}_k^{(m)} \right)^* \odot \hat{\boldsymbol{a}}_{k+\gamma}^{(n)} \right), \qquad (8)$$

$$\theta_{\text{x/y},k+1}^{(m)} = \theta_{\text{x/y},k}^{(m)} + \mu_\theta \text{Im}\left[\left(\delta_{\text{x/y},k}^{(m)} \right)^* z_{\text{x/y},k}^{(m)} e^{-j\theta_{\text{x/y},k}^{(m)}} \right] . \quad (9)$$

where $\mu_{\text{w}} > 0$, $\mu_{\text{h}} > 0$, $\mu_{\text{g}} > 0$, and $\mu_\theta > 0$ are the step size parameters, $\text{Re}[\cdot]$, $(\cdot)^*$ and $(\cdot)^{\text{H}}$ denote the real-part, complex conjugation and matrix Hermitian operations, respectively.

C. Data-aided and Decisions-directed Adaptation

To initiate the above adaptive estimation algorithm and accomplish LMS convergence, we transmit a continuous block of pilot symbols at the beginning of data transmission, which is often referred to as the link-setup phase in the literature [4], [9], [11]. Thereafter, we transmit blocks of N_{p} periodic pilot symbols inserted uniformly after every N_{d} data symbols, for the entire transmission duration. During such a periodic pilot transmission, the CPNEs, the PMD equalizer, and the ICI channel taps are slowly adjusted to account for PN tracking, slow rotation of the principal states of polarization (PSP), and slowly time-varying laser drifts, respectively. The pilot symbols density $p = \frac{N_{\text{p}}}{N_{\text{p}} + N_{\text{d}}}$ is chosen to meet a desired trade-off between performance and transmission overhead. Since the ISI channel is not likely to change over the course of the transmission, the ISI impulse response is estimated only once during the link-setup, followed by very slow adjustments based on the post-LDPC symbol-decisions [4].

As the first step towards the 2-stage PN compensation technique presented in this paper, the CPNEs obtained during the periodic pilots transmission are linearly interpolated to account for PN variation over each symbol duration. Thereafter, the interpolated CPNEs are used to bootstrap more powerful FG-based and LMS-based iterative PN mitigation algorithms. Details of such iterative CPR algorithms are presented in the following section.

IV. ITERATIVE PN ESTIMATION (IPNE)

After the LMS-based coarse CPR is accomplished, we employ iterative algorithms to remove the residual PN. At each iteration, the a-posteriori LLRs are fed back from the LDPC decoders for the purposes of (a) iterative PN estimation and compensation, and (b) TFP interference cancellation as detailed in Section V. In this section, we present two IPNE schemes, namely the low-complexity LMS-based IPNE (LIPNE) and the high-performance factor graph based IPNE (FGIPNE) [17]. The LIPNE, which requires low computational cost and buffer-space, offers decent performance for small values of LLW. Additionally, the functionality of the LIPNE does not depend on the modulation format and the explicit knowledge of the PN statistics. On the other hand, the FGIPNE shows excellent tolerance to strong PN and high levels of interference, at the expense of modulation format dependency, relatively higher complexity and storage requirements. Moreover,

the computation of the FGIPNE metrics requir[...]
of the aggregate variance of the combined Wi[...]
of the transmitter and receiver laser PN [17].
the specific TFP application scenario, the most
method can be chosen such that a desired tra[...]
performance and complexity is achieved. An a[...]
trade-off is investigated in Section VI.

A. LIPNE

For this scheme, we refine the LMS based C[...]
in (9) using the LLRs fed back from the LDP[...]
making hard symbol-decisions in each symbo[...]
account for the TFP ISI and ICI, we formulate th[...]
induced "effective" pilots by reconstructing the[...]
$\delta_{x/y,k}^{(m)}$ according to (4) at every iteration, usin[...]
TFP interference channel and the hard-decisio[...]
from all SCs.

While such post-LDPC decision-directed LIPNE is com-
putationally simpler than other state-space based iterative
methods, it may produce sub-optimal performance due to
potential error-propagation when hard-symbol decisions are
erroneous. Moreover, LPINE for TFP systems may also incur
performance degradation when LLW is high. Motivated by
this, we investigate the more powerful FGIPNE approach
presented in [17], whereby we modify the metric computation
to account for the TFP-ISI and ICI.

B. FGIPNE

We take into account the TFP interference by adapting the
state-space based FGIPNE algorithm in [17] to tailor it to
the super-Nyquist transmission. We remark that the authors
of [4] also adopted similar strategies. However, there are the
following two main differences of our method compared to
theirs: (a) as opposed to the post-LDPC hard-decisions, we use
soft values for the FGIPNE metric computations to reduce the
impact of error-propagation, and (b) in addition to the TFP-ISI,
our super-Nyquist systems also introduce ICI, which entails
inter-SC data processing.

Inspired by the technique adopted in [4], we extract the max-
imum a-posteriori probability (MAP) estimates of the individ-
ual PN processes for both polarization streams of all SCs in the
superchannel. For this, we perform Tikhonov-parameterization
based "forward-backward recursive algorithm" [17], whereby
we use the symbol probabilities derived from the LLRs fed
back by the LDPC decoders, iteratively. To account for the
TFP-ISI and ICI, we adequately modify the forward and back-
ward metric computation presented in [17]. Such modifications
together with a relevant review of the FG based algorithm
in [17] are shown in Appendix A.

Following the fine-tuned CPR accomplished through either
the LIPNE or the FGIPNE method, the PN mitigated samples
are processed by the ISIC and ICIC modules, as detailed in
the following.

V. INTERFERENCE CANCELLATION

In this section, we describe the interference cancellation
operations in detail.

Fig. 3. BCJR-ISIC+SPCIC-ICIC, shown for the example of a 3-SC WDM system.

A. Basic Turbo ISIC-ICIC Structure

After compensating for the PN through CPNE and/or IPNE
algorithms, the inputs from all SCs are jointly processed
by the interference canceler. To mitigate the interference
induced by the TFP transmission, the ISI and ICI impulse
responses estimated in Section III are provided as inputs to
the ISIC and ICIC modules, respectively. For ISIC, the BCJR
algorithm based on Ungerboeck's observation model [36], [37]
is employed. For ICIC, we present a new SPCIC method that
involves multiple stages of PIC and SIC scheduling. Such an
ICIC scheme offers significant performance gains over PIC-
only ICIC approaches, such as [1], [10], [18], for the TFP
superchannels considered in this paper.

To describe the basic operational principles of the turbo
ICIC and ISIC, we consider, for example, the X-polarization
of the mth SC in a N-SC WDM superchannel. Following
the CD and PMD compensation, and carrier recovery, the
estimated ICI channel and the LLRs fed back from the LDPC
decoders of the neighboring SCs are used to compute the soft
estimates of the ICI $I_{m-1,m}$ and $I_{m+1,m}$, $m \in \{1, 2, \ldots, N\}$.
For the computational details of such soft estimates, interested
readers are referred to [10], [18], [30]. The soft ICI estimates
thus obtained are then removed from the received samples
$r_{X,m}$ of the mth SC, to produce the signal $r_{X,m}^{\text{clean}}$, as shown
in Fig. 3 for the example of $m = 2$ and $N = 3$. Next, at
every turbo iteration, considering a square-QAM constellation,
BCJR is performed separately on the I and Q branches of both
polarization streams for each SC, using the estimated ISI taps
derived in Section III. At the end of the final iteration count
it_{max}, output bits are generated from the LDPC decoders. The
details of the SIC and PIC scheduling for the ICIC operation
are discussed in the following.

B. ICIC Scheduling: SPCIC

The central SCs in a TFP superchannel suffer from stronger
ICI compared to the edge SCs, with the possibility of less reli-
able LLRs for the first few iterations. Therefore, different from
the PIC approach in [1], [10], we propose to perform ICIC for
the central SCs first, prior to initiating the ICIC operations for
the edge SCs, until an iteration count threshold is reached,
where such threshold is a design parameter depending on the
values of β, τ and ξ. Based on the number of SCs in the

Algorithm 1 SPCIC algorithm in conjunction with BCJR-ISIC, shown for the X-polarization of all SCs.

At the 0th Iteration:
1: **for all** $n \in \{1, 2, \ldots, N\}$ **do**
2:      $r_{\text{X},n}^{\text{clean}} \leftarrow r_{\text{X},n}$
3:      BCJR-ISIC on $r_{\text{X},n}^{\text{clean}}$ (ICI treated as noise)
4: **end for**
ICIC Initiation:
5: **for** it $= 1 : \text{it}_{\max}$ **do**
6:      **if** $N \leq 2$ **then**   # *No SPCIC*
7:          SPCIC not applicable, i.e. schedule PIC-only ICIC.
8:      **else if** $N = 3$ **then**   # *SPCIC, 1 SIC/PIC stage*
9:          $r_{\text{X},2}^{\text{clean}} \leftarrow r_{\text{X},2} - \text{ICI from left SC}$
10:                 $- \text{ICI from right SC}$
11:          **if** it $> \text{it}_1$ **then**
12:             **for all** $n \in \{1, 2, 3\}$ **do**
13:                 $r_{\text{X},n}^{\text{clean}} \leftarrow r_{\text{X},n} - \text{ICI from left SC}$
14:                       $- \text{ICI from right SC}$
15:             **end for**
16:          **end if**
17:      **else if** $N = 4$ **then**   # *SPCIC, 2 SIC/PIC stages*
18:          **if** it $<= \text{it}_1$ **then**
19:             $r_{\text{X},2}^{\text{clean}} \leftarrow r_{\text{X},2} - \text{ICI from left SC}$
20:             $r_{\text{X},3}^{\text{clean}} \leftarrow r_{\text{X},3} - \text{ICI from right SC}$
21:          **else if** it $<= \text{it}_2$ **then**
22:             $r_{\text{X},1}^{\text{clean}} \leftarrow r_{\text{X},1} - \text{ICI from right SC}$
23:             $r_{\text{X},2}^{\text{clean}} \leftarrow r_{\text{X},2} - \text{ICI from left SC}$
24:             $r_{\text{X},3}^{\text{clean}} \leftarrow r_{\text{X},3} - \text{ICI from right SC}$
25:             $r_{\text{X},4}^{\text{clean}} \leftarrow r_{\text{X},4} - \text{ICI from left SC}$
26:          **else**
27:             **for all** $n \in \{1, 2, 3, 4\}$ **do**
28:                 $r_{\text{X},n}^{\text{clean}} \leftarrow r_{\text{X},n} - \text{ICI from left SC}$
29:                       $- \text{ICI from right SC}$

30:             **end for**
31:          **end if**
32:      **else**   # *SPCIC, multple SIC/PIC stages*
33:          $r_{\text{X},2}^{\text{clean}} \leftarrow r_{\text{X},2} - \text{ICI from left SC}$
34:          $r_{\text{X},N-1}^{\text{clean}} \leftarrow r_{\text{X},N-1} - \text{ICI from right SC}$
35:          **if** it $\leq \text{it}_{\lfloor \frac{N+1}{2} \rfloor - 1}$ **then**
36:             **if** it $> \text{it}_1$ **then**
37:                 $r_{\text{X},3}^{\text{clean}} \leftarrow r_{\text{X},3} - \text{ICI from left SC}$
38:                 $r_{\text{X},N-2}^{\text{clean}} \leftarrow r_{\text{X},N-2} - \text{ICI from right SC}$
39:             **end if**
40:             $\vdots$ $\vdots$ $\vdots$
41:             **if** it $> \text{it}_{\lfloor \frac{N+1}{2} \rfloor - 2}$ **then**
42:                 **if** N is odd **then**
43:                     $r_{\text{X},\lfloor \frac{N+1}{2} \rfloor}^{\text{clean}} \leftarrow r_{\text{X},\lfloor \frac{N+1}{2} \rfloor} - \text{ICI from left SC}$
44:                           $- \text{ICI from right SC}$
45:                 **else**
46:                     $r_{\text{X},\lfloor \frac{N+1}{2} \rfloor}^{\text{clean}} \leftarrow r_{\text{X},\lfloor \frac{N+1}{2} \rfloor} - \text{ICI from left SC}$
47:                     $r_{\text{X},\lceil \frac{N+1}{2} \rceil}^{\text{clean}} \leftarrow r_{\text{X},\lceil \frac{N+1}{2} \rceil} - \text{ICI from right SC}$
48:                 **end if**
49:             **end if**
50:          **else**
51:             **for all** $n \in \{1, 2, \ldots, N\}$ **do**
52:                 $r_{\text{X},n}^{\text{clean}} \leftarrow r_{\text{X},n} - \text{ICI from left SC}$
53:                       $- \text{ICI from right SC}$
54:             **end for**
55:          **end if**
56:      **end if**
57:      **for all** $n \in \{1, 2, \ldots, N\}$ **do**
58:          BCJR-ISIC on $r_{\text{X},n}^{\text{clean}}$ (ICI partially/fully removed)
59:      **end for**
60: **end for**

---

superchannel, multiple thresholds corresponding to different stages of SIC scheduling can be designed. The pseudo-code for the SPCIC operation is provided in Algorithm 1.

For the example of a 3-SC superchannel shown in Fig. 3, ICIC only for the 2nd SC is initiated at the beginning of LDPC iterations, up to an iteration count threshold it_1. Thereafter, for every subsequent turbo iteration, ICIC for all 3 SCs is enabled simultaneously via soft-information exchange across them. When the number of SCs is larger than 3, similar SPCIC scheduling can be adopted as detailed in Algorithm 1, where at the first stage of SIC iterations, ICIC is performed for the 2nd SCs from both superchannel edges; for the second stage of iterations, ICIC is conducted on the 3rd SCs from both edges, and so on, until the central SC is reached. After that, ICIC for all SCs are performed simultaneously through a PIC-based scheduling.

VI. NUMERICAL RESULTS

In this section, we present numerical results to show the benefits of the proposed TFP systems. For this, we consider a DP 16-QAM 3-SC WDM superchannel having per SC baud rate 62.5 Gbaud corresponding to a 1.2 Tbps net data rate. We

accomplish suitable BW compression by choosing appropriate values of τ and ξ such that the TFP superchannels fit within an aggregate BW not exceeding 175 GHz. We remark that such data rates and BW constraints serve as a realistic target for the next generation optical networks (see e.g. [38] and references therein). For example, [6], [38] present recent works demonstrating 1000 km fiber transmission of DP 16-QAM dual-carrier Nyquist superchannels achieving 400 Gbps data rate packed within a 75 GHz grid, or equivalently, 1.2 Tbps with an aggregate BW of 225 GHz. By way of our spectrally efficient TFP design in this paper, our proposed superchannels occupy substantially lower BW compared to [6], [38]. We also note that spectrally more efficient Nyquist WDM systems with similar or higher data rates can be realized employing HoM formats, such as [39]. However, such systems have significantly lower transmission reach due to the application of larger signal constellations.

A. Simulation Parameters

The simulation parameters used for our numerical evaluation are listed in Table I. The values of the parameters are chosen in alignment with practical optical fiber systems [14],

TABLE I
SIMULATION PARAMETERS

LDPC parameters	Values
Standard compliance	DVB-S2 [40]
Block length	64800 bits
Code rate	0.8
# Internal iterations	50
Transmission parameters	Values
Modulation	16-QAM
Baud rate per SC	62.5 Gbaud
N	3
β	0.1
τ, ξ	Varying
Launch power	-5 dBm [9]
LLW	$10 - 400$ kHz [27], [41]
Fiber parameters	Values
SSMF span length	80 km [42]
SSMF number of spans	13
CD parameter β_2	-21 ps^2/km [1], [42], [43]
PMD parameter	0.5 ps/$\sqrt{\text{km}}$ [1]
PSP rotation rate	2.5 kHz
Fiber amplifier noise-figure	4.5 dB [42]–[44]
Fiber attenuation	0.2 dB/km [42]–[44]
# WSSs	0 & 4 [1]
Each WSS 3-dB BW	187.5 GHz
Rx DSP parameters	Values
N_{w}	19 $\frac{T}{2}$-spaced
L_{c}	9
L_{s}	3 (64-state BCJR) [1]
it$_1$	3
it$_{\max}$	10 [1], [4], [10]
N_{p}	10
N_{d}	150 & 300

[19], [27], [28], [31]–[33], [41]–[43]. Except for the results in Fig. 7, we simulate a fixed 13 spans of the SSMFs corresponding to a fiber length $L = 1040$ km [20]. For all results except those in Fig. 9 and Fig. 10, a fixed LLW of 75 kHz is considered [28]. We simulate 30000 symbols at the beginning of transmission for link setup [4], [9], followed by $N_{\text{p}} = 10$, $N_{\text{d}} = 150, 300$, corresponding to 6% and 3% pilot-densities, respectively. These values were chosen to minimize the optical signal-to-noise ratio (OSNR) penalty compared to the corresponding zero-PN benchmark systems, without increasing the pilots overhead significantly [45]. $N_{\text{d}} = 150$ is used for all results except those in Fig. 9. A total of 300 codewords are transmitted to evaluate the average performance over both polarizations and all SCs. Moreover, perfect time and frequency synchronization, and no fiber nonlinearities are assumed.

B. Interference Channel Estimation and Cancellation Gains

We first investigate the performance of the proposed joint estimation algorithm by showing the MSE convergence (see (5)) corresponding to the central SC in the 3-SC super-channel, as a function of the (pilot) symbol index in Fig. 4 during the link setup phase, for different operating OSNRs and LLW = 75 kHz. As shown in the figure, for all τ, ξ pairs, the MSEs of both polarization streams settle to stable values ensuring that the equalizer tap-coefficients and the estimated interference channels converge to nearly stationary values.

The estimated TFP-ISI and ICI impulse responses are then used by the ISIC and ICIC modules. The PMD filter co-efficients and the CPNEs after the link setup phase are updated
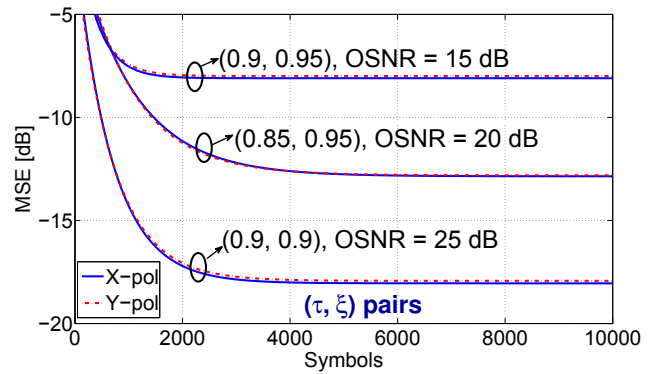
Fig. 4. MSE convergence for 75 kHz LLW, varying τ and ξ. 30000 symbols are transmitted at the beginning for link-setup to ensure MSE convergence, followed by periodic pilots insertion.

only during the periodic pilots as described in Section III, followed by linear interpolation of the data-aided PN estimates to obtain per symbol CPNEs, which are later used to bootstrap the LIPNE or the FGIPNE algorithms.

Next, we show the benefits of our proposed TFP design over the TP super-Nyquist systems, such as [4], [5], [11], [46], in Fig. 5. For this, we plot the bit-error rate (BER) performance averaged over all SCs[1]. For both the TP and the TFP systems, we use 6% pilots for the CPNE, followed by FGIPNE to mitigate the effects of PN. For TFP, SPCIC-ICIC in tandem with BCJR-ISIC is applied to mitigate the TFP-interference, while for TP, only BCJR-ISIC is used and no ICI channel estimation is performed since $\xi = 1$ for such systems. For a fair comparison, we choose the τ, ξ values for the TFP and the equivalent τ value for the TP transmission such that both systems achieve the same SE and hence, same BW, with all systems corresponding to 1.2 Tbps data rate. For example, super-Nyquist WDM systems with $\tau = 0.9$ and $\xi = 0.9$ corresponds to $\tau = 0.84$ and $\xi = 1$ to occupy a BW of 173.25 GHz. For this setting, the proposed TFP system yields a performance improvement of 1.1 dB over the TP transmission, as highlighted in Fig. 5. For higher SE values, such gains due to TFP transmission increase. For example, a TFP (τ, ξ) combination of $(0.85, 0.95)$ offers 2.1 dB OSNR gain over an equivalent time-packed $(0.82167, 1)$ system that occupies the same BW 169.47 GHz. The results show that instead of aggressively packing the SCs in one dimension, relatively benign packing in both the temporal and spectral dimensions can be beneficial when the proposed SPCIC-ICIC together with BCJR-ISIC is employed.

While the average performance over all SCs as shown in Fig. 5 is a crucial metric for a superchannel based transmis-sion, individual performances of the SCs may offer further insight. To show this, in Fig. 6, we have evaluated the individual performance of all SCs in a 3-SC and a 4-SC TFP

[1]No error floor was observed up to a post-FEC BER of 10^{-6} through simulations. Beyond the waterfall region of the post-FEC BER curve, a sharp drop in error rate is observed, and SNRs higher than this threshold are considered to be in the error-free region. This is justified considering that any possible error floor that may occur below this error rate (10^{-6}) can be potentially removed by considering practical concatenated hard-decision outer codes [4], [11].
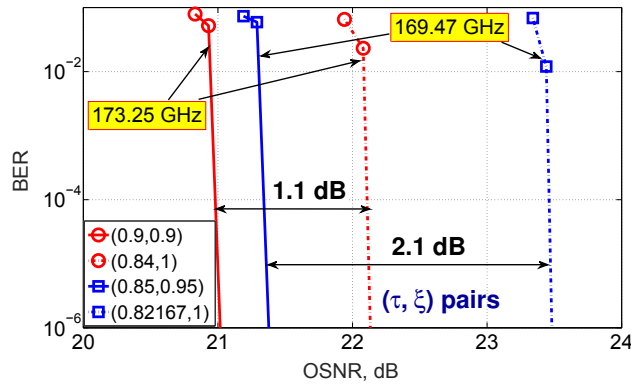
Fig. 5. BER comparison at the same SE, highlighting the benefits of the proposed TFP design over TP. Each color corresponds to a specific SE number, whereas solid and dashed lines represnt TFP and TP schemes, respectively. 1040 km fiber, 75 kHz LLW, CPNE+FGIPNE, 6% pilot density, varying τ and ξ. The aggregate BWs of the superchannels are indicated for each configuration.
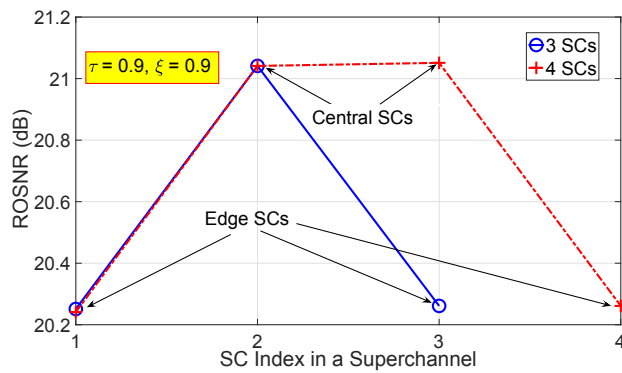

Fig. 6. ROSNR for individual SCs in a 3-SC and 4-SC TFP superchannel. 1040 km fiber, 75 kHz LLW, CPNE+FGIPNE, 6% pilot density, $\tau = 0.9$ and $\xi = 0.9$. The SPCIC parameters, as per Algorithm 1, employed for the TFP systems are as follows: $\mathrm{it}_1 = 3$ for 3 SCs, and $\mathrm{it}_1 = 2, \mathrm{it}_2 = 4$ for 4 SCs, which were chosen to optimize the average performance of the respective superchannels.

superchannel corresponding to $\tau = 0.9, \xi = 0.9$, whereby we plot the required OSNR (ROSNR) for error-free transmission as a function of the SC indices. We observe that the ROSNRs for the central SCs in both these configurations are higher than those for the edges SCs, due to additional ICI from the adjacent neighbors. We also observe that the central and edge SCs in the 3-SC and 4-SC TFP systems perform similarly. This suggests that the average TFP performance may benefit from applying different ξ and/or different coding schemes for different SCs, so as to equalize the ROSNR across SCs, similar to [47].

To further quantify the effectiveness of the proposed TFP design, we show the SE achieved by the proposed TFP systems employing different τ, ξ combinations in Fig. 7, as a function of the transmission distance[2]. For an N-SC DP TFP superchannel with a code rate R_{c}, modulation size M,

---

[2]The SE vs. distance plot in Fig. 7 is evaluated at a fixed launch power of -5 dBm per SC, where the effects of fiber nonlinearity is negligible. In a practical system, a launch power optimization is typically performed [4]. See Section VII-B for a detailed discussion.
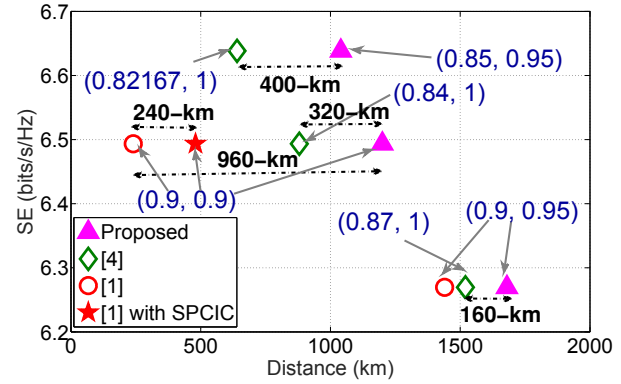
Fig. 7. SE vs. distance evaluated at a fixed launch power of -5 dBm per SC, highlighting the benefits of the proposed TFP design over TP and other TFP designs. 75 kHz LLW, CPNE+FGIPNE, 6% pilot density, varying τ and ξ.
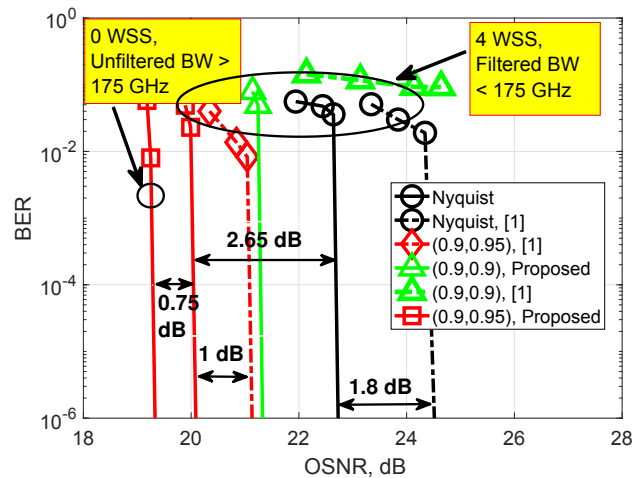
Fig. 8. BER performance of the proposed schemes in a target BW of 175 GHz, for the example of 4-WSSs in the link. This produces an effective 3-dB BW of 166.88 GHz, such that the superchannels fit within the target BW. 1040 km fiber, 75 kHz LLW, CPNE+FGIPNE, 6% pilot density, varying τ and ξ.

and pilot density $p\%$, we compute the SE as

$$\mathrm{SE} = \frac{2NR_{\mathrm{c}} \log_2 M}{\tau(1+\beta)\left[1 + (N-1)\xi\right]} \left(1 - \frac{p}{100}\right) \text{ bits/s/Hz} . \quad (10)$$

In Fig. 7, we also include the SE values of the following two benchmark schemes for reference: (i) the method presented in our previous work [1], employing PIC-ICIC and BCJR-ISIC with the absence of channel estimation, and (ii) time-packed ISI-only systems similar to [4] that achieve the same SE. We observe that a substantial distance improvement of $240 - 960$ km is achieved by the proposed technique over (i) mentioned above. Moreover, to show the benefit of the proposed SPCIC over the PIC-based ICIC scheme, we also evaluate the performance of (i) by replacing the PIC-ICIC with SPCIC, indicated by the legend "[1] with SPCIC" in Fig. 7, which increases the transmission distance by 240 km for $\tau = \xi = 0.9$, as highlighted in the figure. Finally, by comparing the proposed TFP systems with (ii) for the same SE values, we observe a link distance improvement by $2-5$ spans corresponding to $160-400$ km for different τ, ξ pairs.

C. Tolerance to Cascaded ROADMs

The introduction of flexible grid networking [3], [38] in the optical systems facilitates denser frequency allocation, where the reduction of the frequency grid granularity from a conventional 50 GHz to 12.5 GHz is being investigated [46]. This is accomplished by deploying flexible wavelength selective switches (WSSs) in the ROADMs [38]. While propagating through the optical link, a superchannel signal may be subjected to multiple of such WSSs, where the aggregate BW is dictated by the effective 3-dB BW of the cascaded WSSs [3], [38], [46]. In light of some of the previous works that evaluated the performance of 400 Gbps systems in a 75 GHz [6], [38] and a 100 GHz [6] BW, for the following analysis, we set a more ambitious target of packing the 1.2 Tbps superchannels in an aggregate BW of 175 GHz (equivalent to 14×12.5 GHz grids), achieving a higher SE than those works. For this, we have considered an example of 4-WSSs in the optical link such that the effective filtered BW is less than 175 GHz, which enables the superchannels to fit within the target BW. Under such a BW constraint, an optimal τ, ξ pair offers the best performance [1], which we investigate below with our proposed TFP systems. Different from our previous work [1], the proposed design has the added advantage of the ability to estimate the combined interference stemming from the TFP transmission and WSS filtering, both of which can be equalized through the BCJR-ISIC employed in conjunction with the proposed SPCIC-ICIC.

We plot the average BER of all SCs as a function of OSNR in Fig. 8, which shows that significant gains are offered by the proposed design over [1] for all combinations of τ and ξ. For example, OSNR gains of 1 dB and 1.8 dB are highlighted in the figure corresponding to the Nyquist WDM and the $\tau = 0.9, \xi = 0.95$ TFP system, respectively. We also note that the optimal TFP parameter combination of $\tau = 0.9$ and $\xi = 0.95$ produces an OSNR gain of 2.65 dB over the Nyquist WDM system under similar conditions. Additionally, we observe that the performance of the TFP system with such an optimal τ, ξ combination filtered through 4-WSSs is 0.75 dB worse than that achieved by the same system without WSS filters in the optical link having an unconstrained BW larger than the target BW of 175 GHz.

D. Tolerance to Laser Linewidth

In this section, we present results to show the robustness of the proposed TFP design against increasing PN levels. First, we study the effect of varying pilot density and LLW on the proposed CPNE, LIPNE and FGIPNE schemes, for the simplistic case of a Nyquist WDM system. In Fig. 9, we plot the ROSNR penalty of such system over the benchmark transmission that is not impaired with PN, as a function of the LLW of the transmitter and the receiver lasers. Both LIPNE and FGIPNE are bootstrapped with CPNE having 3% and 6% pilot density. In the figure, the label "CPNE-only" refers to the case where no iterative PN mitigation algorithm is applied after the CPNE based coarse CPR. Not surprisingly, all PN mitigation methods produce larger performance degradation with increasing LLW. As shown in the figure, the gains of
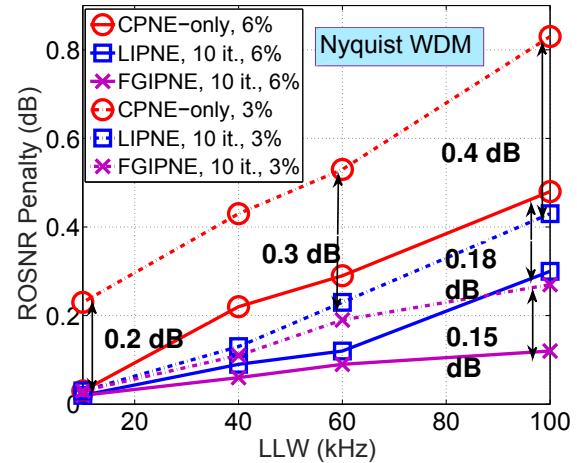

Fig. 9. ROSNR penalty vs. LLW for Nyquist WDM, showing benefits and limitations of CPNE, LIPNE and FGIPNE, having varying pilot densities. Each color represent a specific PN mitigation algorithm, whereas solid and dashed lines correspond to 6% and 3% pilot densities.

LIPNE and FGIPNE over CPNE-only PN mitigation method also increase with larger LLWs for both pilot densities. The plots in the figure also suggest that LIPNE is performing close to FGIPNE for LLW up to 100 kHz with a maximum performance gap of 0.2 dB. Moreover, FGIPNE performs better by 0.15 dB when it is bootstrapped with CPNE using 6% transmission overhead compared to 3% pilot density.

After having investigated the performance of the proposed PN cancellation methods in a Nyquist WDM transmission, we now proceed to evaluate their effectiveness in TFP systems with even higher values of LLWs. In Fig. 10(a)-10(b), we plot the ROSNR penalty over the respective zero-PN systems similar to Fig. 9, for the Nyquist WDM transmission and $\tau = 0.9, \xi = 0.95$ TFP system, respectively. Both IPNE methods are bootstrapped by the CPNE with 6% pilots. In Fig. 10, we also include the "perfect-decision CPNE" scheme corresponding to the LMS-based PN estimation with genie-assisted known transmitted symbols, as a reference. The plots in the figures suggest that while LIPNE is performing decently compared to FGIPNE for the Nyquist WDM transmission, FGIPNE outperforms LIPNE by significant margins in TFP systems, especially when LLW is very high. For example, with 400 kHz LLW, FGIPNE yields 1 dB ROSNR improvement over LIPNE with $\tau = 0.9, \xi = 0.95$ as highlighted in Fig. 10(b), whereas such a gain is restricted to only 0.4 dB for the Nyquist transmission as shown in Fig. 10(a). Moreover, FGIPNE is also able to outperform the perfect-decision CPNE, and it offers 1 dB gain over CPNE-only PN compensation method for both Nyquist WDM and TFP systems.

E. Computational Complexity Analysis

Details of the computational complexity for the proposed systems design are furnished in Table II, where M denotes the modulation order, and the rest of the parameters are already defined in the preceding sections. The numbers in the table correspond to computations required for both polarizations and

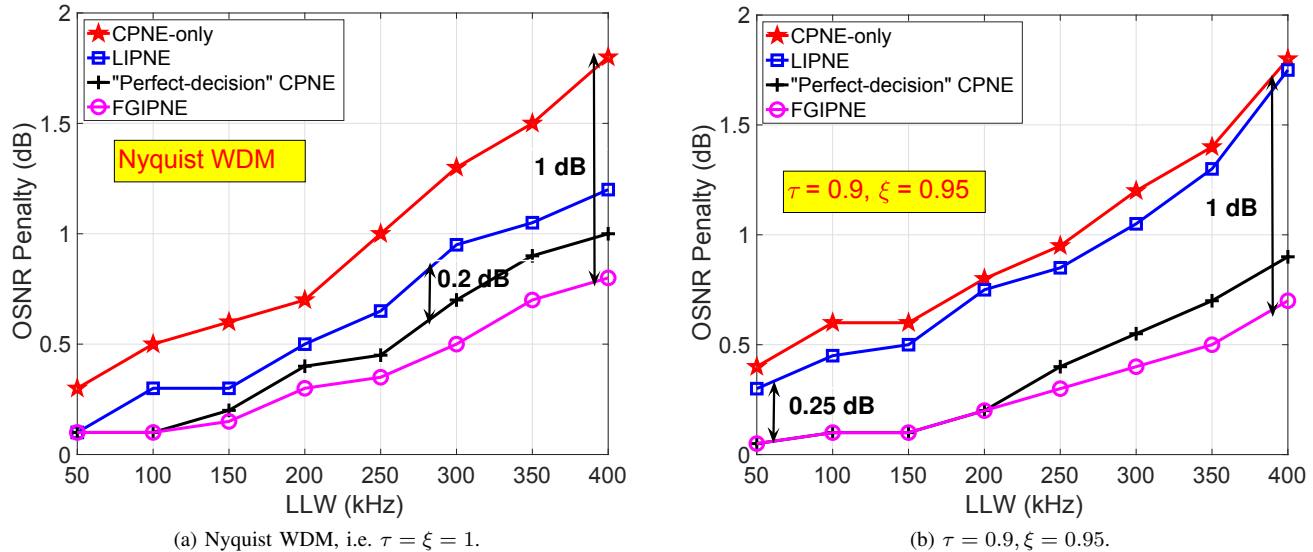(a) Nyquist WDM, i.e. $\tau = \xi = 1$.

(b) $\tau = 0.9, \xi = 0.95$.

Fig. 10. ROSNR penalty vs. LLW, showing benefits and limitations of CPNE, LIPNE and FGIPNE, 6% pilot density. FGIPNE is shown to outperform all other PN mitigation methods for both the Nyquist and TFP systems.

TABLE II
COMPUTATIONAL COMPLEXITY PER CODE SYMBOL.

Task	Item	Add + Sub.	Mul. + Div.
Estimation	PMD	$4NN_{\mathrm{w}}$	$8NN_{\mathrm{w}}$
	ISI	$8N(L_{\mathrm{s}}+1)$	$6N(L_{\mathrm{s}}+1)$
	ICI	$4N(L_{\mathrm{c}}+1)$	$6N(L_{\mathrm{c}}+1)$
	CPNE	$4N$	$10N$
	LIPNE	$4N\mathrm{it}_{\max}$	$10N\mathrm{it}_{\max}$
	FGIPNE	$4N(M+L_{\mathrm{s}}+L_{\mathrm{c}}+5)\mathrm{it}_{\max}$	$4N(M+L_{\mathrm{s}}+L_{\mathrm{c}}+9)\mathrm{it}_{\max}$
Equalization	ISIC	$2N\mathrm{it}_{\max}\left(8M^{\frac{L_{\mathrm{s}}}{2}}-4\right)$	$2N\mathrm{it}_{\max}\left(16M^{\frac{L_{\mathrm{s}}}{2}}+2\right)$
	ICIC	$2N\mathrm{it}_{\max}(M+2L_{\mathrm{c}})$	$2N\mathrm{it}_{\max}(M+2L_{\mathrm{c}}+1)$

all SCs in the superchannel. As shown, the LMS based estimation algorithms for PMD filter coefficients, TFP interference and CPNE scale linearly with the number of SCs. Evidently, BCJR-ISIC constitutes the computationally most challenging module, since the complexity significantly magnifies with the length of the truncated TFP-ISI channel and the constellation size [20], [48]. The computational cost of SPCIC-ICIC, however, increases linearly with the modulation order and TFP-ICI taps length [1]. A comparison between LIPNE and FGIPNE indicates that FGIPNE exhibits a modulation format dependency, and entails slightly more computations compared to LIPNE, with the benefit of substantial performance advantage shown in Section VI. Moreover, FGIPNE also requires additional buffering to store the forward and backward FG metrics [17].

VII. FURTHER DISCUSSIONS

A. Frequency Offset and Laser Drifts

We have not investigated carrier frequency offset (CFO) estimation and mitigation technique in this paper. We remark that any "conventional" carrier frequency recovery algorithm can be applied for the superchannel systems considered here. For example, the authors of [4] employed a data-aided CFO

estimation technique proposed in [49], for their TP superchannels. Moreover, [50] presents a factor-graph based joint CFO and PN estimation algorithm, which can also be integrated into our TFP systems.

Our proposed algorithms are directly applicable to a frequency-comb based superchannel transmission, where the transmitter and receiver lasers are perfectly frequency locked. Although such systems would still require a carrier frequency recovery algorithm to mitigate the common offset that is "shared" by all SCs in the superchannel [51], there will be no relative frequency drifts between them. On the other hand, for individual lasers based systems, any CFO estimation method applied for each SC needs to track the time-varying relative laser drifts between the SCs. Once the estimates for such drifts are obtained, our proposed algorithms are expected be robust to the additional ICI that stems from such drifts, because of the adaptive nature of our ICI channel estimation approach in conjunction with the ICIC scheme. To verify this, we have tested our algorithm for the TFP system having $\tau = 0.9, \xi = 0.9$, with a fixed relative laser drift up to 100 MHz between the SCs, and we observed negligible performance degradation compared to the absence of such drifts.

B. Fiber Nonlinearity

In the performance evaluation, we have ignored the effects of fiber nonlinearity, and to show the transmission distance gains of the proposed TFP design over the TP systems [4], [11] achieving the same SE, we have considered a fixed launch power of -5 dBm, where we assume negligible nonlinear interference and noise (NLIN) effects. In the following, we would like to elaborate on the validity of the latter assumption, and on the meaningfulness of the presented results for systems operating with optimized launch powers and in the nonlinear regime.

To account for NLIN, we have performed additional simulations using the transmission and fiber parameters provided in Table I, and an NLIN coefficient of 1.2 $\text{W}^{-1}\text{km}^{-1}$ [42], with the following two methods: (A) enhanced Gaussian noise (EGN)-model based prediction [52]–[54], and (B) split step Fourier method (SSMF) simulations using VPItransmissionMaker$^{\text{TM}}$ optical systems [55]. For (A), we used the method presented in [54] to compute the SNR correction term for the Nyquist and TFP superchannels having 3 and 9 SCs as

$$\text{SNR}_{\text{corr,dB}} = 10 \log_{10}\left(\sigma^2_{\text{ASE}} + \sigma^2_{\text{NLIN}}\right) - 10 \log_{10}\left(\sigma^2_{\text{ASE}}\right), \quad (11)$$

where σ^2_{ASE} and σ^2_{NLIN} denote the ASE noise variance and the NLIN equivalent noise variance, respectively. For (B), we have investigated the pre-FEC BER performance of the central SC in a 3-SC and a 9-SC Nyquist WDM superchannel, as a function of the launch power.

Based on both (A) and (B), we observed a performance difference less than 0.06 dB and 0.08 dB at -5 dBm launch power between a linear and a nonlinear fiber channel, corresponding to the 3 and 9 SCs, respectively. Such small values prove that the choice of -5 dBm launch power enables us to operate in the linear fiber regime, whereby the effects of NLIN is negligible.

Additionally, to speculate the performance gains offered by the proposed TFP schemes over the TP solutions [4], [11] when a launch power optimization is performed in a practical fiber transmission, we have computed the term $\text{SNR}_{\text{corr,dB}}$ in (11) for a 3-SC system with the TFP configuration $\tau = 0.9, \xi = 0.9$, and the equivalent TP scheme having $\tau = 0.84, \xi = 1$, achieving the same SE, at the optimal launch power obtained from (B). The difference of such correction terms between the two equivalent systems at this optimal launch power turns out to be only 0.035 dB. Similarly, the difference of $\text{SNR}_{\text{corr,dB}}$ between the two equivalent systems $(\tau = 0.85, \xi = 0.95)$ and $(\tau = 0.82167, \xi = 1)$ is observed to be 0.06 dB. Therefore, we reasonably expect that, when a launch power optimization is performed, although the absolute distance values presented in Fig. 7 may change, the relative gains we show through our linear fiber simulations will be translated to similar figures in the presence of NLIN.

Finally, we remark that a part of NLIN manifests itself as nonlinear PN (NLPN) [14], [56]. It would be an interesting future research to investigate if, and to what extent, the PN compensation methods presented in our paper, such as the FGIPNE, are able to mitigate NLPN for the proposed TFP superchannels. Moreover, the combination of the proposed TFP receiver design with other nonlinearity compensation (NLC) techniques seems worth investigating. For example, the NLC algorithm presented in [57] involves reconstruction and removal of the NLIN contributions from the neighboring SCs, which bears a structural resemblance to our proposed ICIC scheme. Therefore, the TFP systems employing the interference mitigation methods as discussed in our paper in conjunction with the mentioned NLC technique may further improve the transmission reach.

VIII. Conclusion

Superchannel data rates of 1 Tbps and more are being targeted in the next generation optical fiber systems to compete with the increasing demands in network traffic. To accomplish such a target, we proposed flexible designs for spectrally efficient TFP superchannel transmission systems achieving Tbps data rates with significantly higher SE values compared to a Nyquist WDM transmission. For this, we have presented sophisticated signal processing tools to efficiently handle TFP interference and accomplish CPR. Our simulation results with a linear fiber channel model suggest that by employing the proposed interference channel and PN estimation methods, together with the BCJR-ISIC and the novel SPCIC-ICIC scheme, the presented TFP WDM systems offer more than 2 dB ROSNR gains and as high as $160 - 960$ km transmission distance improvements over state-of-the-art super-Nyquist superchannel transmission techniques. The proposed TFP design shows outstanding tolerance to PN with LLW up to 400 kHz. Moreover, our system exhibits excellent robustness against additional BW limitation in the form of cascades of ROADM nodes comprised of narrow WSS filters that may be present in the longhaul fiber link.

Appendix A
FGIPNE Metrics Computation

To compute the FGIPNE forward and backward metrics, and thereby, obtain the MAP estimates of the laser PN, we refer to the FG shown in Fig. 2 of [17]. Owing to space limitation, we will not recall the algorithm presented in [17] in detail. Instead, we will briefly revisit only the components that are relevant to our considered TFP system. Using similar notations as in [4], [17], and focusing on the variable node θ_k in Fig. 2 of [17], we note that the product of the three incoming messages $p_d(\theta_k), p_f(\theta_k)$ and $p_b(\theta_k)$ is proportional to the conditional probability density function (PDF) $p(\theta_k|\{r_k\})$, where k is the symbol index and $\{r_k\}$ is the sequence of the received symbols [4], [58] (to familiarize with the concept of FGs and the sum-product algorithm, interested readers are referred to [58]). Thereafter, we employ the iterative forward-backward algorithm as per Section IV.B of [17].

Using similar notations as in [17], we note that the iterative algorithm involves the computation of the parameters α_k and β_k according to (28) of [17], to denote the first and the second-order moments of the transmitted symbols, respectively. To account for the TFP-ISI and ICI in our superchannel transmission, we propose the following modification to the computation of α_k and β_k. For this, we introduce a new variable

$$\gamma_k = \beta_k - \left|\alpha_k\right|^2. \quad (12)$$

Moreover, to indicate the polarization labeling and the SC index, we denote these parameters by $\alpha_{k,\text{x/y}}^{(m)}$, $\beta_{k,\text{x/y}}^{(m)}$ and $\gamma_{k,\text{x/y}}^{(m)}$ corresponding to the mth SC and X or Y polarization, respectively. The modified FGIPNE metrics, stacked as column vectors with X and Y polarization inputs, are formulated as

$$[\alpha_{k,\text{x}}^{(m)}, \alpha_{k,\text{y}}^{(m)}]^{\text{T}} = \sum_{j=-L_\text{s}}^{L_\text{s}} \boldsymbol{h}_{j,k}^{(m)} \odot \mathbb{E}\left(\boldsymbol{a}_{k-j}^{(m)}\right)$$

$$+ \sum_{n \neq m} \sum_{\nu=-L_{\mathrm{c}}}^{L_{\mathrm{c}}} \boldsymbol{g}_{\nu,k}^{(n,m)} \odot \mathbb{E}\left(\hat{\boldsymbol{a}}_{k-\nu}^{(n)}\right) , \qquad (13)$$

$$[\gamma_{k,\mathrm{x}}^{(m)}, \gamma_{k,\mathrm{y}}^{(m)}]^{\mathrm{T}} = \sum_{j=-L_{\mathrm{s}}}^{L_{\mathrm{s}}} \left\langle |\boldsymbol{h}_{j,k}^{(m)}| \right\rangle^2 \odot \mathrm{Var}\left(\boldsymbol{a}_{k-j}^{(m)}\right)$$

$$+ \sum_{n \neq m} \sum_{\nu=-L_{\mathrm{c}}}^{L_{\mathrm{c}}} \left\langle |\boldsymbol{g}_{\nu,k}^{(n,m)}| \right\rangle^2 \odot \mathrm{Var}\left(\hat{\boldsymbol{a}}_{k-\nu}^{(n)}\right) , \quad (14)$$

where $\mathrm{Var}(\cdot)$ and $\left\langle | \cdot | \right\rangle^2$ denote element-wise variance and absolute-square operations, respectively. At each LDPC iteration, the expectations and variances in (13)-(14) are computed for the constellation symbols using the symbol-probabilities obtained from the LLRs fed back by the LDPC decoders. Finally, using the statistical property of the Tikhonov PDF, we obtain the MAP estimate of the PN for the corresponding polarization and SC as [4]

$$\hat{\theta}_{\mathrm{MAP}} = \angle\left(\frac{2 r_k \alpha_k^*}{2\sigma^2 + \gamma_k} + a_{f,k} + a_{b,k}\right), \qquad (15)$$

where $\angle(\cdot)$ denotes the phase angle of a complex scalar, σ^2 is the noise variance per real dimension, and $a_{f,k}$ and $a_{b,k}$ are computed according to (36) and (37) of [17], respectively.

REFERENCES

[1] M. Jana, L. Lampe, and J. Mitra, "Interference Cancellation for Time-Frequency Packed Super-Nyquist WDM Systems," *IEEE Photon. Technol. Lett.*, vol. 30, no. 24, pp. 2099–2102, December 2018.

[2] K. Roberts *et al.*, "High Capacity Transport - 100G and Beyond," *J. Lightw. Technol*, vol. 33, no. 3, pp. 563–578, February 2015.

[3] T. Strasser *et al.*, "Wavelength-Selective Switches for ROADM Applications," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1150–1157, September 2010.

[4] M. Secondini, T. Foggi *et al.*, "Optical Time-Frequency Packing: Principles, Design, Implementation, and Experimental Demonstration," *J. Lightw. Technol*, vol. 33, no. 17, pp. 3558–3570, September 2015.

[5] G. Colavolpe and T. Foggi, "Time-Frequency Packing for High-Capacity Coherent Optical Links," *IEEE Trans. on Commun.*, vol. 62, no. 8, pp. 2986–2995, August 2014.

[6] L. Li, Z. Xiao, L. Liu, and Y. Lu, "Non-Orthogonal WDM Systems with Faster than Nyquist Technology," in *Opt. Fiber Commun. Conf.(OFC)*, 2019, pp. 1–3.

[7] Q. Hu *et al.*, "Inter-Channel Crosstalk Compensation for Time-Frequency Packing Systems," in *European Conf. on Opt. Commun. (ECOC)*, 2017, pp. 1–3.

[8] J. Pan *et al.*, "Inter-Channel Crosstalk Cancellation for Nyquist-WDM Superchannel Applications," *J. Lightw. Technol*, vol. 30, no. 24, pp. 3593–3999, December 2012.

[9] K. Shibahara *et al.*, "Multi-stage Successive Interference Cancellation for Spectrally-Efficient Super-Nyquist Transmission," in *European Conf. on Opt. Commun. (ECOC)*, 2015, pp. 1–3.

[10] F. Rusek, "Partial Response and Faster-than-Nyquist Signaling," Ph.D. dissertation, Lund University, 2007.

[11] G. Meloni, T. Rahman *et al.*, "Experimental Comparison of Transmission Performance for Nyquist WDM and Time-Frequency Packing," *J. Lightw. Technol*, vol. 33, no. 24, pp. 5261–5268, December 2015.

[12] Z. Xiao *et al.*, "ICI Mitigation for Dual-Carrier Superchannel Transmission Based on m-PSK and m-QAM Formats," *J. Lightw. Technol*, vol. 34, no. 23, pp. 5526–5533, December 2016.

[13] S. Chen *et al.*, "Comparison of Advanced Detection Techniques for QPSK Signals in Super-Nyquist WDM Systems," *IEEE Photon. Technol. Lett.*, vol. 27, no. 1, pp. 105–108, January 2015.

[14] M. P. Yankov, T. Fehenberger, L. Barletta, and N. Hanik, "Low-Complexity Tracking of Laser and Nonlinear Phase Noise in WDM Optical Fiber Systems," *J. Lightw. Technol*, vol. 33, no. 23, pp. 4975–4984, December 2015.

[15] J. C. M. Diniz *et al.*, "Low-complexity Carrier Phase Recovery Based on Principal Component Analysis for Square-QAM Modulation Formats," *Optics Express*, vol. 27, no. 11, pp. 15 617–15 626, May 2019.

[16] T. Pfau, S. Hoffmann, and R. Noe, "Hardware-Efficient Coherent Digital Receiver Concept With Feedforward Carrier Recovery for M-QAM Constellations," *J. Lightw. Technol*, vol. 27, no. 8, pp. 989–999, April 2009.

[17] G. Colavolpe, A. Barbieri, and G. Caire, "Algorithms for Iterative Decoding in the Presence of Strong Phase Noise," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1748–1757, September 2005.

[18] D. Dasalukunte, "Multicarrier Faster-than-Nyquist Signaling Transceivers," Ph.D. dissertation, Lund University, 2012.

[19] M. Jana, A. Medra, L. Lampe, and J. Mitra, "Precoded Faster-than-Nyquist Coherent Optical Transmission," in *European Conf. on Opt. Commun. (ECOC)*, 2016.

[20] ——, "Pre-equalized Faster-than-Nyquist Transmission," *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4406–4418, October 2017.

[21] M. Jana, L. Lampe, and J. Mitra, "Precoded Time-Frequency-Packed Multicarrier Faster-than-Nyquist Transmission," in *IEEE Int. workshop on Signal Proc. advances in Wireless Commun. (SPAWC)*, 2019.

[22] M. Jana, J. Mitra, L. Lampe, and A. Medra, "System and Method for Precoded Faster than Nyquist Signaling," U.S. Patent 10 003 390, June, 2018. [Online]. Available: https://patents.justia.com/patent/10003390

[23] M. Jana, J. Mitra, and L. Lampe, "Methods and Systems for Interference Mitigation in a Dual-Polarized Communication System," U.S. Patent 10 425 256, September, 2019. [Online]. Available: https://patents.justia.com/patent/10003390

[24] M. Jana, L. Lampe, and J. Mitra, "Interference and Phase Noise Mitigation in a Dual-Polarized Faster-than-Nyquist Trasnmission," in *IEEE Int. Workshop on Signal Proc. advances in Wireless Commun. (SPAWC)*, 2018.

[25] ——, "Dual-polarized Faster-than-Nyquist Transmission Using Higher-order Modulation Schemes," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5332–5345, November 2018.

[26] X. Qi, N. Wu, H. Wang, and W. Yuan, "A Factor Graph-Based Iterative Detection of Faster-than-Nyquist Signaling in the Presence of Phase Noise and Carrier Frequency Offset," *DSP*, vol. 63, pp. 25–34, April 2017.

[27] J. Hald, L. Nielsen, J. C. Petersen, P. Varming, and J. E. Pedersen, "Fiber Laser Optical Frequency Standard at 1.54 m," *Optics Express*, vol. 19, no. 3, pp. 2052–2063, January 2011.

[28] K. Balskus *et al.*, "Frequency Comb Metrology with an Optical Parametric Oscillator," *Optics Express*, vol. 24, no. 8, pp. 8370–8381, April 2016.

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

[30] A. Barbieri, D. Fertonani, and G. Colavolpe, "Time-frequency Packing for Linear Modulations: Spectral Efficiency and Practical Detection Schemes," *IEEE Trans. on Commun.*, vol. 57, no. 10, pp. 2951–2959, October 2009.

[31] E. Ip and J. M. Kahn, "Digital Equalization of Chromatic Dispersion and Polarization Mode Dispersion," *J. Lightw. Technol*, vol. 25, no. 8, pp. 2033–2043, August 2007.

[32] C. Xie and X. Zhou, *Enabling Technologies for High Spectral-efficiency Coherent Optical Communication Networks*. John Wiley & Sons, Inc, 2016.

[33] G. Colavolpe, T. Foggi, E. Forestieri, and G. Prati, "Robust Multilevel Coherent Optical Systems With Linear Processing at the Receiver," *J. Lightw. Technol*, vol. 27, no. 13, pp. 2357–2369, July 2009.

[34] M. A. Reimer, "Simulation Methods for the Temporal and Frequency Dynamics of Optical Communication Systems," Ph.D. dissertation, The University of Waterloo, 2012.

[35] M. P. Yankov, T. Fehenberger, L. Barletta, and N. Hanik, "Low-Complexity Tracking of Laser and Nonlinear Phase Noise in WDM Optical Fiber Systems," *J. Lightw. Technol*, vol. 33, no. 23, pp. 4975–4984, December 2015.

[36] F. Rusek, M. Loncar, and A. Prlja, "A Comparison of Ungerboeck and Forney Models for Reduced-Complexity ISI Equalization," in *IEEE Global Commun. Conf. (GLOBECOM)*, 2007, pp. 1431–1437.

[37] G. Colavolpe and A. Barbieri, "On MAP symbol detection for ISI channels using the Ungerboeck observation model," *IEEE Commun. Lett.*, vol. 9, no. 8, pp. 720–722, August 2005.

[38] D. Marom *et al.*, "Survey of Photonic Switching Architectures and Technologies in Support of Spatially an Spectrally Flexible Optical Networking [Invited]," *IEEE J. Opt. Commun. Netw.*, vol. 9, no. 1, pp. 1–26, January 2017.

[39] A. Matsushita, M. Nakamura, F. Hamaoka, S. Okamoto, and Y. Kisaka, "High-Spectral-Efficiency 600-Gbps/Carrier Transmission Using PDM-256QAM Format," *J. Lightw. Technol*, vol. 37, no. 2, pp. 470–476, January 2019.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2020.3019257, Journal of Lightwave Technology

13

[40] *Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications (DVB-S2)*, ETSI EN 302 307 V1.2.1 Std. [Online]. Available: www.dvb.org/standards/dvb-s2

[41] G. Bosco, "Advanced Modulation Techniques for Flexible Optical Transceivers: The Rate/Reach Tradeoff," *J. Lightw. Technol*, vol. 37, no. 1, pp. 36–49, January 2019.

[42] S. Makovejs *et al.*, "Characterization of Long-haul 112Gbit/s PDM-QAM-16 Transmission with and without Digital Nonlinearity Compensation," *Optics Express*, vol. 18, no. 12, pp. 12 939–12 947, June 2010.

[43] T. Fehenberger, A. Alvarado, G. Böcherer, and N. Hanik, "On Probabilistic Shaping of Quadrature Amplitude Modulation for the Nonlinear Fiber Channel," *J. Lightw. Technol*, vol. 34, no. 21, pp. 5063–5073, November 2016.

[44] B. Chomycz, *Planning Fiber Optics Networks*.   McGraw Hill Professional, 2009.

[45] C. Yang, F. Yang, and Z. Wang, "Phase Noise Suppression for Coherent Optical Block Transmission Systems: a Unified Framework," *Optics Express*, vol. 19, no. 11, pp. 17 013–17 020, August 2011.

[46] T. Foggi *et al.*, "Spectral Efficiency Optimization in Flexi-Grid Long-Haul Optical Systems," *J. Lightw. Technol*, vol. 33, no. 13, pp. 2735–2742, July 2015.

[47] R. Maher, A. Alvarado, D. Lavery *et al.*, "Increasing the Information Rates of Optical Communications via Coded Modulation: a Study of Transceiver Performance," *Sci. Rep.*, vol. 6, no. 21278, February 2016.

[48] D. Wang and H. Kobayashi, "Matrix Approach for Fast Implementations of Logarithmic MAP Decoding of Turbo Codes," in *IEEE Pacific Rim Conf. on Commun.,Compt. and Signal Proc.*, 2001, pp. 115–118.

[49] U. Mengali and M. Morelli, "Data-aided frequency estimation for burst digital transmission," *IEEE Trans. Commun.*, vol. 45, no. 1, pp. 23 – 25, January 1997.

[50] A. Barbieri, G. Colavolpe, and G. Caire, "Joint Iterative Detection and Decoding in the Presence of Phase Noise and Frequency Offset," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 171 – 179, January 2007.

[51] M. Mazur, J. Schröder, M. Karlsson, and P. A. Andrekson, "Joint Superchannel Digital Signal Processing for Ultimate Bandwidth Utilization," *arXiv e-prints*, p. arXiv:1911.02326, Nov 2019.

[52] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN Model of Non-linear Fiber Propagation," *Optics Express*, vol. 22, no. 13, p. 1633516362, December 2014.

[53] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of Nonlinear Noise in Long, Dispersion-Uncompensated Fiber Links," *Optics Express*, vol. 21, no. 22, pp. 25 685–25 699, November 2013.

[54] ——, "Accumulation of Nonlinear Interference Noise in Fiber-optic Systems," *Optics Express*, vol. 22, no. 12, pp. 14 199–14 211, June 2014.

[55] VPI. Vpitransmissionmaker^TM optical systems. [Online]. Available: https://www.vpiphotonics.com/Tools/OpticalSystems/

[56] M. P. Yankov *et al.*, "Nonlinear Phase Noise Compensation in Experimental WDM Systems With 256QAM," *J. Lightw. Technol*, vol. 35, no. 8, pp. 1438 – 1443, April 2017.

[57] A. Amari, P. Ciblat, and Y. Jaouën, "Inter-Subcarrier Nonlinear Interference Canceler for Long-Haul Nyquist-WDM Transmission," *IEEE Photon. Technol. Lett.*, vol. 28, no. 23, pp. 2760 – 2763, October 2016.

[58] F. R. Kschischang, B. J. Frey, and H. Loeliger, "Factor Graphs and the Sum-Product Algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, February 2001.