Pre-equalized Faster-than-Nyquist Transmission

Mrinmoy Jana, Ahmed Medra, Lutz Lampe, and Jeebak Mitra

***Abstract*—Faster-than-Nyquist (FTN) transmission applies non-orthogonal linear modulation to increase spectral efficiency compared to the well-known orthogonal transmission at Nyquist rate. This comes at a price of inter-symbol interference (ISI), which usually is equalized through receiver processing. In this paper, we investigate the alternative approach of pre-equalization at the transmitter. First, we consider Tomlinson-Harashima precoding (THP) for FTN and propose two novel soft demapping algorithms to generate the soft-input for the error-correction decoder. The developed demappers effectively compensate the modulo-loss associated with conventional THP transmission. Second, we propose a linear pre-filtering strategy to pre-equalize the ISI induced by FTN. We show that the linear pre-equalization approach is equivalent to an orthogonal transmission with a modified pulse shape. It thus yields the optimal error-rate performance while affording higher spectral efficiency. We validate our proposed precoding algorithms through computer simulations of a coded coherent optical communication system as a practical application example for FTN.**

***Index Terms*—Faster-than-Nyquist (FTN) transmission, non-orthogonal signaling, Tomlinson-Harashima precoding (THP), soft demapper, pre-equalization, spectral factorization, spectral leakage.**

I. Introduction

NEXT generation communication systems are evolving towards deploying remarkably improved bandwidth-efficient transmission schemes in order to cope with the growing demand for data rates. One way to accomplish this is by giving up the orthogonality condition in terms of time and frequency spacing of adjacent symbols imposed by the Nyquist criterion. From a communication-theoretic point of view, transmitting at a faster-than-Nyquist (FTN) rate allows us to approach the capacity of a bandlimited channel [2]. From a practical implementation perspective, FTN is advantageous for transmission systems such as coherent optical communication where the application of higher-order modulation formats to increase spectral efficiency renders the system more vulnerable to the non-linear effects of an optical channel [3], [4]. Denser time-frequency packing via FTN is also being considered in the context of new modulation formats for fifth generation (5G) networks [5].

The fact that FTN signaling can be an attractive choice has been discussed extensively in the literature, see [6] and references therein. While the original work by Mazo [7] and other early works (e.g. [8]–[10]) focused on the minimum distance assuming optimal detection to deal with the inter-symbol interference (ISI) introduced by FTN, the development of sub-optimal equalization methods has received significant attention more recently. These include reduced-state versions of maximum a-posteriori probability (MAP) symbol equalization based on the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [11]–[14] and frequency domain equalization (FDE) [15]–[17], often in an iterative fashion together with forward-error-correction (FEC) decoding. However, the complexity of this turbo-equalization is still substantial compared to the absence of FTN equalization in Nyquist transmission. On the other hand, the performance of low-complexity linear equalization methods is usually not sufficient especially when the ISI due to FTN is severe.

We, therefore, turn our attention to pre-equalization techniques which can significantly diminish or completely eliminate the computational burden from equalization at the receiver. To this end, the first key observation is that the FTN introduced ISI is perfectly known at the transmitter. Hence, pre-equalization does not require the feedback of channel state information (CSI) from the receiver to the transmitter. This renders the well-known Tomlinson-Harashima precoding (THP) [18]–[20] an attractive choice for pre-equalization. Indeed, THP for FTN has been considered in several recent publications in the context of 5G mobile wireless communications [21], [22], microwave backhaul links [23] and coherent optical communications [1], [24]–[27]. However, the disadvantages of a coded THP system manifest themselves in the form of the so-called "modulo-loss" and "precoding-loss" [20] and a possible increase in the peak-to-average power ratio (PAPR). While the precoding-loss causes a fixed signal-to-noise ratio (SNR) penalty depending on the modulation format and, as will be shown in this paper, the FTN parameters, the modulo-loss causes an error-rate deterioration by providing inaccurate soft information to the FEC decoder. A few recent works [28]–[30] aim to address the modulo-loss problem by improving the accuracy of the log-likelihood ratio (LLR) computation. However, the presented methods are either computationally prohibitive [30] or performance gains are limited [28], [29].

In this paper, as our first contribution, we propose two computationally efficient demapping algorithms for an FTN-THP system which outperform the existing memoryless demappers from [28], [29] by significant margins. We show that the demappers presented in this work not only compensate for the modulo-loss but also make THP competitive to computationally expensive MAP-based equalization techniques. Having dealt with the modulo-loss, we then investigate the precoding-loss associated with THP. For this, we make the second key observation that FTN-ISI stems entirely from the transmit pulse-shape and the receive matched filter. The transmit pulse-shape thus contributes partially to the ISI and is a part of the

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Huawei Technologies, Canada.

Part of this work was presented at the 42nd European Conf. on Opt. Commun. (ECOC), Germany, Sept. 2016 [1].

Mrinmoy Jana, Ahmed Medra, and Lutz Lampe are with the Department of Electrical and Computer Engineering, University of British Columbia, BC, Canada. Email: mjana@ece.ubc.ca, ahmedm@ece.ubc.ca, lampe@ece.ubc.ca. Jeebak Mitra is with Huawei Technologies, Ottawa, Canada. Email: jeebak.mitra@huawei.com.

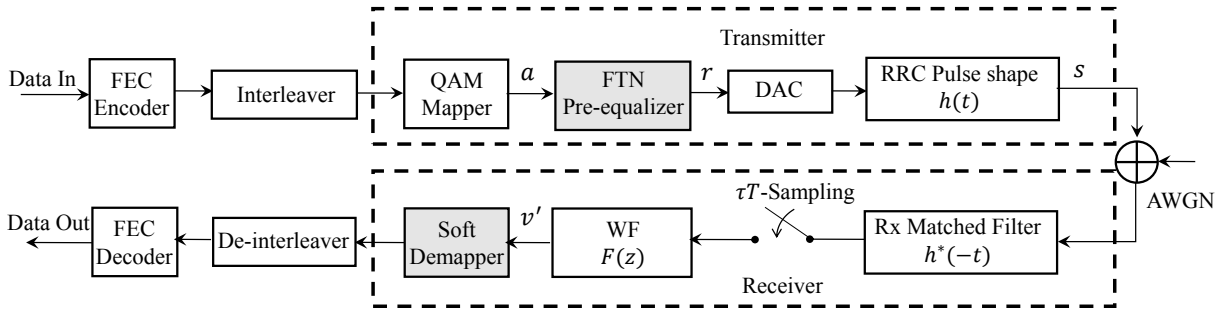


Fig. 1. Baseband system model for a pre-equalized FTN transmission where the shaded blocks at the transmitter and the receiver represent the proposed FTN pre-equalizer and symbol demappers respectively.

transmitter, whereas, a conventional ISI channel in a Nyquist transmission lies outside the transmitter. As a consequence, the precoding-loss for an FTN-THP transmission over an additive white Gaussian noise (AWGN) channel is different from that in a Nyquist-THP transmission over ISI channels. As our second contribution, we derive the analytical expressions for the precoding-loss in an FTN-THP system as a function of the FTN and the pulse-shaping parameters. We show that the precoding-loss of the FTN-THP scheme can be substantial especially when the ISI induced by FTN becomes severe. Motivated by this, we then turn our focus on the linear precoding options. In particular, we propose a linear pre-equalization (LPE) method to pre-compensate for the FTN-ISI. Due to the fact that FTN is different from classical ISI where the channel lies outside the transmitter, linear pre-equalization does not suffer from noise enhancement. It does, however, modify the transmit power spectral density (PSD), and we show that our method converts FTN transmission into orthogonal signaling with an equivalent pulse shape. In doing so, the proposed LPE completely eliminates FTN-ISI. Our method is related to other linear precoding techniques that have been analyzed in the past in conjunction with FTN and partial response signaling (PRS) [31]–[35]. However, these are different in that they are either block-based matrix-precoding techniques or attempt to obtain pre-filter coefficients from optimization problems to maximize distance properties.

The remainder of the paper is organized as follows. The system model is introduced in Section II. In Section III, we propose two novel demappers for FTN-THP and present the analysis for the precoding loss. The new linear pre-filtering method for FTN is proposed in Section IV. In Section V, we validate the proposed methods based on simulations for a coherent optical transmission setup. Finally, Section VI provides concluding remarks.

II. System Model

A. Precoded FTN

We consider the baseband system model for precoded FTN transmission scheme under an AWGN channel shown in Fig. 1. The system model is common for both linear and non-linear pre-equalization methods. As shown in Fig. 1, the data bits are first FEC encoded and then the interleaved and modulated data stream a is precoded with a discrete-time pre-filter to produce the data symbols r. The precoded symbols r are pulse-shaped by a T-orthogonal pulse h and then transmitted with an FTN acceleration factor $\tau < 1$. As in [2], the resulting linearly modulated baseband transmitted signal can be written as

$$s(t) = \sum_k r[k]h(t - k\tau T) \ . \tag{1}$$

For the following, we assume a root-raised-cosine (RRC) pulse-shaping filter h with a roll-off factor β such that $\int_{-\infty}^{\infty} |h(t)|^2 dt = 1$.

At the receiver, the analog received signal, after passing through the matched-filter, is sampled at τT-intervals and then digitally processed by a noise whitening filter (WF) to whiten the colored noise due to FTN. Thereafter, the τT sampled signal v' is sent to a symbol demapper to produce soft information in the form of LLRs for the FEC decoder.

The overall discrete-time channel impulse response between the precoded symbols r and the output of the τT-spaced sampling is given by

$$g[n] = (h * f)(n\tau T) \ , \tag{2}$$

where $f(t) = h^*(-t)$, $\cdot^*$ is complex conjugate, and $*$ denotes the linear convolution. We also introduce $G = \mathcal{Z}(g)$, where $\mathcal{Z}\{\cdot\}$ denotes the z-transform. In a Nyquist-system ($\tau = 1$), T-orthogonality of the pulse-shape h along with the condition $\int_{-\infty}^{\infty} |h(t)|^2 dt = 1$ makes $G(z) = 1$. But for an FTN transmission with $\tau < 1$, $G(z)$ causes ISI across consecutive transmitted symbols.

As THP can be seen as a dual to a decision-feedback equalization (DFE) performed at the receiver [20], we apply a spectral factorization to G (consistent with [20], [36]) as detailed in the following subsection.

B. Spectral Factorization

THP requires the implementation of a feed-forward-filter (FFF) F and a feedback-filter (FBF) B to pre-equalize the ISI due to FTN. As the noise samples after the τT-sampler in an FTN system are colored, the purpose of the FFF is then two-fold: to whiten the received noise samples and to shape the end-to-end channel transfer-function into a causal and minimum-phase response [20]. The FBF is then used as a pre-filter at the transmitter to pre-equalize the overall

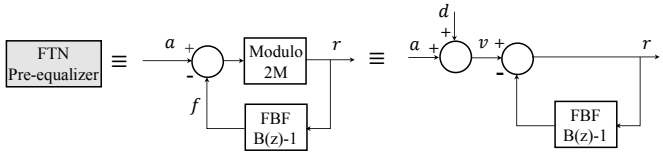


Fig. 2. FTN pre-equalization with THP and the modulo-equivalent linear structure.

effective ISI-channel. Computation of FFF and FBF requires the discrete-time spectral factorization [20]

$$G(z) = \alpha Q(z) Q^* \left(z^{-*}\right) \ , \tag{3}$$

such that $Q(z)$ is casual, monic and minimum-phase and $\alpha > 0$ is a scaling factor used to make $Q(z)$ monic. The necessary and sufficient condition for the realization of the above spectral factorization (see e.g. [20], [37]) can be written in an FTN transmission as

$$\tau T \int_{-\frac{1}{2\tau T}}^{\frac{1}{2\tau T}} |\log\left(G(\mathrm{e}^{\mathrm{j}2\pi f \tau T})\right)| df < \infty \ . \tag{4}$$

Since from (2),

$$G(\mathrm{e}^{\mathrm{j}2\pi f \tau T}) = \frac{1}{\tau T} \sum_{k=-\infty}^{\infty} |\hat{H}(f - k/(\tau T))|^2 \ , \tag{5}$$

where $\hat{H}$ is the Fourier-transform of the pulse-shaping filter h. We note that $G(\mathrm{e}^{\mathrm{j}2\pi f \tau T})$ in (5) is zero in the intervals $\left[-\frac{1}{2\tau T}, -\frac{1+\beta}{2T}\right]$ and $\left[\frac{1+\beta}{2T}, \frac{1}{2\tau T}\right]$ when the FTN acceleration factor $\tau < \frac{1}{1+\beta}$. This causes the condition in (4) to fail, which consequently makes the spectral factorization (3) required for THP unrealizable. Hence, in the following we restrict ourselves to FTN with $\tau \geq \frac{1}{1+\beta}$ for a given β. Once the factorization according to (3) is executed, we obtain the FFF and FBF respectively as

$$F(z) = \frac{1}{\alpha Q^*\left(z^{-*}\right)} \ \text{ and } \ B(z) = Q(z) \ . \tag{6}$$

Using the FFF and FBF computed above, we now proceed to introduce FTN-THP with an improved demapper in the next section, and a new linear pre-equalization method in Section IV.

III. Non-linear Precoding in FTN Systems

In this section, we consider non-linear precoding for FTN in the form of THP.

A. THP-precoded FTN

Since the effective ISI-channel caused by FTN is a-priori known at the transmitter, the filters from Section II-B can be computed and applied for THP without any feedback from the receiver.

Fig. 2 depicts the detailed diagram and the associated linear equivalent structure of the block "FTN Pre-equalizer" from Fig. 1. The modulo operation in a classical THP system as shown in Fig. 2 is used to keep the output stable especially for channels with spectral zeros by bounding it within a well-prescribed range [20]. The input symbols a in Fig. 2 consist of the modulated symbols and the feedback filter B, as given in (6), is a function of the FTN parameter τ and the pulse shape h. As the FTN-ISI is real-valued, without loss of generality, we assume that the symbols a are drawn from an M-ary one-dimensional constellation. In the equivalent linear representation, the modulo operation of THP is replaced by an equivalent addition of a unique sequence d to the data symbols a so that precoded symbols r lie in the interval $[-M, M)$. The combination of a and d produces the intermediate signal v, the elements of which are taken from an extended constellation with more than M signal points. In an ideal noise-free scenario, the signal v' in Fig. 1 is same as v of Fig. 2, and thus, to compensate for THP, conventionally a modulo operation is performed on v' at the receiver. However, for noisy channels and particularly for a relatively low SNR, this modulo operation is sub-optimal which makes the LLR computation by a conventional soft-demapper unreliable. These inaccurate LLRs are then passed on to the FEC decoder as shown in Fig. 1 and thereby causing a performance degradation, especially in an FEC coded transmission, which is commonly known as the "modulo loss".

To overcome this loss, a modified modulo based demapper for a coded THP system was proposed in [28] and its simplified implementation method was also presented recently in [29]. However, the residual modulo-loss of these approaches still causes a significant loss in the bit-error rate (BER). Another near-optimal iterative method was shown in [30]. It is based on a quantized-output THP, and its computational complexity is of the order of MAP equalization. In the following, we present two relatively simpler soft-demapping algorithms which significantly outperform the demapper from [28], which we refer to as Peh-Liang-Demapper (PLD), and are competitive to optimal MAP equalization in terms of BER.

B. Expanded A-priori Demapper (EAD)

In order to counter the modulo-loss, we replace the modulo operation of a conventional THP demapper with our proposed new demapper, referred to as EAD, which now forms the module "Soft Demapper" of Fig. 1. The proposed EAD is based on the linear equivalent model from Fig. 2 and considers the extended constellation of the intermediate data signal v to compute LLRs. Let $\mathcal{A} = \{a^{\kappa}_{\mathrm{PAM}} = \pm 1, \pm 3, \ldots, \pm(M-1)\}$ be the set of M-ary pulse-amplitude modulation (PAM) constellation symbols. Then, the symbols v belong to the extended signal set $\mathcal{V} = \{v[k]\} = \mathcal{A} + 2M\mathbb{Z}$. In particular, we note that the probabilities of the signal points $v \in \mathcal{V}$ are not uniform. Therefore, the EAD computes the LLR value corresponding to the n^{th} bit b_n of the k^{th} data symbol $a[k]$ as

$$\mathrm{LLR}^{\mathrm{EAD}}_{k,n} = \log\left(\frac{\Pr\left(b_n = 1 | v'[k]\right)}{\Pr\left(b_n = 0 | v'[k]\right)}\right) \tag{7}$$

$$= \log \frac{\sum\limits_{c \in \mathcal{C}_{1,n}} \Pr(v'[k] | v[k] = c) \Pr(v[k] = c)}{\sum\limits_{c \in \mathcal{C}_{0,n}} \Pr(v'[k] | v[k] = c) \Pr(v[k] = c)} \ , \tag{8}$$

where $\mathcal{C}_{i,n}$ is the subset of symbols in $\mathcal{V}$ corresponding to the n^{th} bit being equal to $i \in \{0, 1\}$ and $v'[k]$ is the k^{th} received

sample at the demapper input. The relation between v and v' follows as

$$v'[k] = v[k] + \eta[k] \ , \tag{9}$$

where $\eta[k]$ is a zero-mean AWGN with variance σ^2. Note that the colored noise samples after an FTN-sampler at the receiver are whitened by the FFF, as discussed in Section II-B. Introducing the Gaussian probability density function (pdf) of η in (9), we can simplify (8) as

$$\mathrm{LLR}_{k,n}^{\mathrm{EAD}} = \log \frac{\sum\limits_{c\in\mathcal{C}_{1,n}} \Pr(v[k]=c)\, \mathrm{e}^{-\frac{|v'[k]-c|^2}{2\sigma^2}}}{\sum\limits_{c\in\mathcal{C}_{0,n}} \Pr(v[k]=c)\, \mathrm{e}^{-\frac{|v'[k]-c|^2}{2\sigma^2}}} \tag{10}$$

$$\approx \log \frac{\Pr(v[k]=\bar{c}_{1,n})\, \mathrm{e}^{-\frac{|v'[k]-\bar{c}_{1,n}|^2}{2\sigma^2}}}{\Pr(v[k]=\bar{c}_{0,n})\, \mathrm{e}^{-\frac{|v'[k]-\bar{c}_{0,n}|^2}{2\sigma^2}}} \tag{11}$$

$$= \log\left(\frac{\alpha_{1,n}}{\alpha_{0,n}}\right) + \frac{|v'[k]-\bar{c}_{0,n}|^2 - |v'[k]-\bar{c}_{1,n}|^2}{2\sigma^2}, \tag{12}$$

where (11) follows from the nearest neighbor approximation (e.g. [28]), with $\bar{c}_{i,n}$ as the nearest neighbor to the received sample $v'[k]$ representing the n^{th} bit being equal to $i \in \{0,1\}$ and $\alpha_{i,n} = \Pr\left(v[k]=\bar{c}_{i,n}\right)$.

The expressions in (10) and (12) are readily evaluated given the received samples $v'[k]$ and the a-priori probabilities $\Pr\left(v[k]=c\right)$ for the signal points $c \in \mathcal{V}$. To analytically compute these probabilities for a given β and τ, we make use of the following proposition.

Proposition 1. *Expanded constellation symbols $v[k] \in \mathcal{V}$ in Fig. 2 have the following probability mass function (PMF):*

$$\Pr\left(v[k]=a_{M,v}^{(\kappa,i)}\right) = \frac{1}{M}\left[\Phi\left(\frac{M_i^{+}+a_{\mathrm{PAM}}^{\kappa}}{\sigma_f}\right) - \Phi\left(\frac{M_i^{-}+a_{\mathrm{PAM}}^{\kappa}}{\sigma_f}\right)\right], \tag{13}$$

where $a_{\mathrm{PAM}}^{\kappa} \in \mathcal{A}$, $a_{M,v}^{(\kappa,i)} = a_{\mathrm{PAM}}^{\kappa} + 2iM$, $M_i^{+} = (2i+1)M$, $M_i^{-} = (2i-1)M$ for $i \in \mathbb{Z}$, σ_f is the standard deviation of the signal f and $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{\frac{-x^2}{2}}\, dx$.

Proof: See Appendix A. ∎

The standard deviation σ_f in (13) can be computed numerically. Simulation results in Section V show that EAD can offer substantial gains over PLD especially when the FTN-ISI is less severe. The relation between the LLR calculation by EAD and PLD from [28] for severe FTN-ISI is summarized in the following proposition and its corollary.

Proposition 2. *For 2PAM and 4PAM modulations, the LLR expression in (12) becomes equivalent to the approximate LLR expression computed by PLD as given in [28] if the extended constellation symbols of the signal v are assumed to have equal probabilities.*

Proof: See Appendix B. ∎

Corollary 2.1. *For 2PAM and 4PAM modulations, when the FTN-ISI becomes severe (i.e. τ reduces for a given RRC roll-off β), the LLR expressions computed by EAD and PLD become similar.*

Proof: For an M-ary PAM constellation, an upper bound on the maximum number of signal points in $\mathcal{V}$ with non-zero probability is given in [20] as

$$V_{\max} = 2\left\lfloor \frac{M\sum_{k=0}^{P-1}|b[k]|+1}{2} \right\rfloor - 1 \ , \tag{14}$$

where $b = \mathcal{Z}^{-1}(B)$ is the THP feedback filter response, P denotes the length of the ISI channel and the function $\lfloor x \rfloor$ denotes the largest integer contained in x. Therefore, with large P, $\mathcal{V}$ contains more symbols with non-zero probabilities which causes the bell-shaped PMF in (13) to flatten and its shape starts resembling closer to that of a uniform distribution. Then by Proposition 2, LLRs computed by EAD become similar to those computed by PLD. ∎

As evaluated above, the gains offered by EAD reduce for decreasing τ. This can be attributed to the fact that while EAD takes the probabilities of the extended constellation symbols into account, it fails to incorporate the auto-correlation of the intermediate symbol sequence v into the LLR metric in (12). As τ reduces, correlation between successive symbols of v can increase significantly due to severe FTN-ISI. In order to account for this, we present the second demapper design in the following.

C. Sliding-window-EAD (SW-EAD)

The SW-EAD includes L preceding and succeeding observations (corresponding to a sliding window of length $2L+1$) into the computation of LLRs for the current symbol. Depending on the severity of the ISI and the observed auto-correlation of v, a suitable value L is determined. The modified LLR for the n^{th} bit b_n of the k^{th} transmitted symbol $a[k]$ is computed as

$$\mathrm{LLR}_{k,n}^{\text{SW-EAD}} = \log\left(\frac{\Pr(b_n=1|v'[k-L],\ldots,v'[k+L])}{\Pr(b_n=0|v'[k-L],\ldots,v'[k+L])}\right) \tag{15}$$

$$= \log\left(\frac{\sum\limits_{c\in\mathcal{C}_{1,n},v[k-L],\ldots,v[k+L]} \Pr\left(\vec{v'}|\vec{v}_c\right)\Pr\left(\vec{v}_c\right)}{\sum\limits_{c\in\mathcal{C}_{0,n},v[k-L],\ldots,v[k+L]} \Pr\left(\vec{v'}|\vec{v}_c\right)\Pr\left(\vec{v}_c\right)}\right) \tag{16}$$

$$= \log\left(\frac{\sum\limits_{c\in\mathcal{C}_{1,n},v[k-L],\ldots,v[k+L]} \Pr\left(\vec{v}_c\right)\mathrm{e}^{-\frac{\|\vec{v'}-\vec{v}_c\|^2}{2\sigma^2}}}{\sum\limits_{c\in\mathcal{C}_{0,n},v[k-L],\ldots,v[k+L]} \Pr\left(\vec{v}_c\right)\mathrm{e}^{-\frac{\|\vec{v'}-\vec{v}_c\|^2}{2\sigma^2}}}\right), \tag{17}$$

where $\vec{v}_c = [v[k-L],\ldots,v[k]=c,\ldots,v[k+L]]^T$, $\vec{v'} = [v'[k-L],\ldots,v'[k+L]]^T$ and $\|\cdot\|$ denotes the vector norm operator. The computation of (17), using the known extended constellation symbols $v \in \mathcal{V}$ and the received samples v', involves pre-computing and storing the multi-dimensional a-priori probabilities $\Pr(\vec{v}_c)$. For $L=0$, SW-EAD metric (17) reduces to the EAD-computed LLR given in (10). Note that, if $\Delta_v = |\mathcal{V}|$ denotes the cardinality of $\mathcal{V}$, then among the Δ_v^{2L+1} multi-dimensional sequences, only a small fraction, ρ_v^L number of symbol-vectors can have non-zero probabilities, depending

on the values of M, β and τ. It is sufficient to store only these ρ_v^L a-priori probabilities to compute (17).

The SW-EAD can be used recursively in iterations with an FEC decoder. In this case, the extrinsic information provided by the FEC decoder for the coded bits is used to update the a-priori probability $\Pr(\vec{v}_c)$ in (17).

D. Precoding-loss for FTN-THP Systems

In a THP-precoded Nyquist transmission over an ISI-channel, the precoding operation causes an increase in the average transmit power which translates into the precoding loss with respect to an equivalent DFE equalization scheme [20, p. 144]. Moreover, an ideal DFE without error propagation can incur an SNR degradation compared to the matched-filter bound (MFB) depending on the parameter α in (3) [20, p. 67-68]. Therefore, the combined SNR loss of a FTN-THP transmission with respect to ISI-free orthogonal transmission is

$$\mathrm{SNR}_{\mathrm{Loss}}^{\text{FTN-THP}} = \mathrm{P}_{\mathrm{Loss}}^{\text{THP-DFE}} \cdot \mathrm{SNR}_{\mathrm{Loss}}^{\text{DFE-MFB}} , \tag{18}$$

where $\mathrm{SNR}_{\mathrm{Loss}}^{\text{DFE-MFB}} = 1/\alpha$ with α given in (3). While the precoding loss $\mathrm{P}_{\mathrm{Loss}}^{\text{THP-DFE}}$ has been well investigated in the literature (e.g. [20]), the situation is slightly different for FTN-THP systems, where the transmit power and hence the precoding loss depend on the ISI channel through the transmit pulse-shape. In order to quantify the precoding loss, we utilize the results from the following proposition.

Proposition 3. *For an FTN-THP system with the FFF and FBF given in (6), the PSD of the transmitted signal is given by*

$$\Phi_{ss}^{\mathrm{THP}}(f) = \alpha \Phi_{vv}\left(\mathrm{e}^{\mathrm{j}2\pi\tau f T}\right) \frac{\hat{G}(f)}{\sum\limits_k \hat{G}\left(f + \frac{k}{\tau T}\right)} , \tag{19}$$

and the average transmitted power is

$$P_{\mathrm{Avg}}^{\mathrm{THP}} = \frac{\alpha \sigma_v^2}{\tau T} , \tag{20}$$

where σ_v^2 and $\Phi_{vv}\left(\mathrm{e}^{\mathrm{j}2\pi\tau f T}\right)$ are the variance and the PSD of the extended constellation symbols v, respectively, $\hat{G}(f) = |\hat{H}(f)|^2$, and $\alpha > 0$ is the constant given in (3).

Proof: See Appendix C. ∎

Since (i) a THP and a DFE system with the same FFF and FBF perform identically assuming no modulo loss for THP and no error propagation for DFE and (ii) the transmit power for non-precoded FTN is $P_{\mathrm{Avg}} = \frac{\sigma_a^2}{\tau T}$, where σ_a^2 is the variance of the M-ary constellation symbol a (e.g. [2]), we have from (20) that

$$P_{\mathrm{Loss}}^{\text{THP-DFE}} = \alpha \frac{\sigma_v^2}{\sigma_a^2} . \tag{21}$$

The overall SNR-loss of the FTN-THP system as compared to the ISI-free transmission follows from (18) as

$$\mathrm{SNR}_{\mathrm{Loss}}^{\text{FTN-THP}} = \frac{\sigma_v^2}{\sigma_a^2} . \tag{22}$$

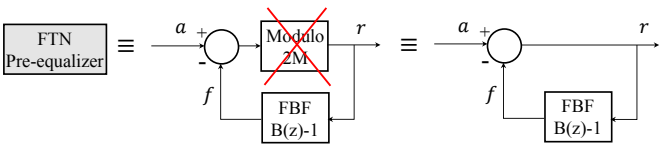


Fig. 3. Linear pre-equalization of FTN ISI.

IV. Linear Pre-equalization for FTN

FTN-THP not only incurs an SNR loss, but it may also complicate pilot-based channel estimation. Since the THP operation results in an expanded signal constellation v' after the WF stage shown in Fig. 1, the receiver lacks the prior knowledge of the exact representation of pilot symbols introduced into the data stream. To alleviate this problem, careful attention to the pilot-symbol design is needed [24] or a coarse detection of the pilot symbol is required before channel estimation [25].

These problems warrant the consideration of linear precoding or pre-equalization methods. We note that linear precoding is done in PRS transmission, albeit the purpose is not pre-equalization but spectral shaping of the transmit signal or the maximization of some performance criteria assuming receiver-side equalization [2], [31]–[33]. Different from this, we propose a linear pre-equalization (LPE) technique to mitigate the ISI introduced through FTN signaling. More specifically, the pre-equalization is achieved through a linear pre-filtering method which is derived from the THP transmitter structure by dropping the modulo operator as shown in Fig. 3. The exclusion of the modulo function renders the overall transmitter of Fig. 3 a linear infinite impulse response (IIR) filtering operation. The minimum-phase property of the feedback filter $B(z)$, as discussed in Section II-B, guarantees the stability of the IIR filter.

In Nyquist transmission over ISI channels, linear pre-equalization is usually not a preferred choice. In particular, linear pre-equalization to eliminate ISI results in an elevation of the average transmitted power, which creates a similar error-rate degradation as the noise-enhancement phenomena encountered in a linear zero-forcing equalization [20]. However, as pointed out earlier, in an FTN transmission, the ISI stems from the transmitter pulse-shape and receiver matched filter. In particular, the feedback filter B in Fig. 3 is a function of the RRC filter h related through (3), (5) and (6). This leads to the following results for the PSD and the average transmit power for FTN-LPE transmission.

Proposition 4. *For an FTN-LPE system in Fig. 3 with the FBF given in (6), the PSD of the transmitted signal is given by*

$$\Phi_{ss}^{\mathrm{LPE}}(f) = \alpha \sigma_a^2 \frac{\hat{G}(f)}{\sum\limits_k \hat{G}\left(f + \frac{k}{\tau T}\right)} , \tag{23}$$

and the average transmit power is

$$P_{\mathrm{Avg}}^{\mathrm{LPE}} = \frac{\alpha \sigma_a^2}{\tau T} , \tag{24}$$

where σ_a^2 is the variance of the input constellation symbols a and $\alpha > 0$ is given in (3).

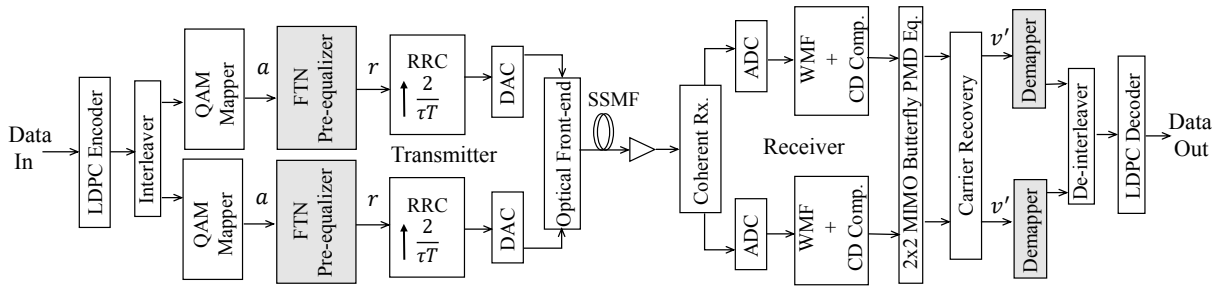


Fig. 4. Block diagram of the precoded FTN dual-polarized coherent optical simulation setup where the shaded blocks at the transmitter and the receiver represent the proposed THP/LPE pre-equalizer and symbol demappers respectively.

Proof: See Appendix C. ∎

Corollary 4.1. *For a Nyquist system, the transmitted PSD becomes* $\Phi_{ss}^{\text{Nyq}}(f) = \frac{\sigma_a^2}{T}\hat{G}(f)$ *with an average transmitted power* $P_{\text{Avg}}^{\text{Nyq}} = \frac{\sigma_a^2}{T}$.

Proof: For a Nyquist system, $\tau = 1$, $\alpha = 1$ from (3) and $\frac{1}{T}\sum_k \hat{G}\left(f + \frac{k}{T}\right) = 1$. Putting these values in (23) and (24) yields the well-known expressions. ∎

Corollary 4.2. *If* $\tau = \frac{1}{1+\beta}$, *the PSD of the transmitted signal becomes rectangular with a bandwidth* $\frac{1+\beta}{T}$.

Proof: The expression $\sum_k \hat{G}\left(f + \frac{k}{\tau T}\right)$ in (23) is the sum of the frequency-shifted replicas of $\hat{G}(f)$, where the frequency shifts are integral multiples of $\frac{1}{\tau T}$. We note that,

$$\hat{G}(f) = 0\,, \text{when } |f| > \frac{1+\beta}{2T}\,. \tag{25}$$

Therefore, when $\tau = \frac{1}{1+\beta}$, there are no overlaps between the replicas of $\hat{G}$ in $\sum_k \hat{G}\left(f + \frac{k}{\tau T}\right)$. Consequently, $\sum_k \hat{G}\left(f + \frac{k}{\tau T}\right) = \hat{G}(f)$ in the frequency range $-\frac{1+\beta}{2T} \le f \le \frac{1+\beta}{2T}$. Thus, from (23) we have

$$\Phi_{ss}^{\text{LPE}}(f) = \begin{cases} \alpha\sigma_a^2, & -\frac{1+\beta}{2T} \le f \le \frac{1+\beta}{2T} \\ 0, & \text{otherwise} \end{cases} \tag{26}$$

∎

To investigate the power-penalty of the FTN-LPE transmission, we use the same procedure which was adopted in Section III-D for the SNR-loss computation of an FTN-THP system. Similar to (18), we can write the combined SNR-loss for the LPE as

$$\text{SNR}_{\text{Loss}}^{\text{FTN-LPE}} = \text{P}_{\text{Loss}}^{\text{LPE-DFE}} \cdot \text{SNR}_{\text{Loss}}^{\text{DFE-MFB}}\,, \tag{27}$$

where $\text{SNR}_{\text{Loss}}^{\text{DFE-MFB}} = 1/\alpha$ as in (18) and following the same reasoning as in Section III-D, the precoding loss $\text{P}_{\text{Loss}}^{\text{LPE-DFE}}$ can be computed from (24) as

$$P_{\text{Loss}}^{\text{LPE-DFE}} = \alpha\,. \tag{28}$$

Hence, the overall SNR-loss of the LPE-THP system as compared to the ISI-free transmission can be written from (27) as

$$\text{SNR}_{\text{Loss}}^{\text{FTN-LPE}} = 1\,. \tag{29}$$

We conclude from (29) that FTN-LPE does not suffer from a power penalty due to channel inversion and achieves the same error-rate performance as an ISI-free transmission. To do so, linear pre-equalization modifies the transmit PSD according to (23) that exhibits τT-orthogonality. In fact, a closer inspection of the FBF and FFF filters for LPE reveals that the combination of the LPE pre-filter and the RRC pulse-shape at the transmitter is equivalent to a new τT-orthogonal square-root Nyquist pulse-shaping filter. Similarly, at the receiver, the RRC filter, combined with the WF, constitutes an equivalent square-root Nyquist matched-filter to the new transmit pulse-shape. Hence, FTN-LPE with whitened matched-filtering and τT sampling is ISI-free.

As an alternative τT-orthogonal signaling scheme, one could directly use a τT-orthogonal RRC filter with roll-off $\hat{\beta} = \tau(1+\beta) - 1$ for transmit pulse-shaping. For instance, an FTN system with T-orthogonal RRC having $\beta = 0.3$ and $\tau = 0.78$ results in an effective $\hat{\beta} = 0.014$ for the direct τT-orthogonal RRC design. We illustrate in Section V that due to the stricter roll-off requirement for this new RRC pulse-shape, the implementation of this filter needs more taps to maintain a given threshold for the out-of-band power leakage compared to that in the proposed LPE-FTN system.

Finally, we remark that the proposed pre-equalization shares similarities with the matrix-based precoded FTN transmission presented in [34], [35]. Similar to FTN-LPE, this method divides the equalization task into pre-filtering at the transmitter and post-filtering at the receiver. However, different from our implementation it is based on block-processing of the transmitted and received symbols and thus it suffers from inter-block ISI [34] or needs guard intervals and thereby reduces spectral efficiency, in addition to introducing a block delay. An alternate way to minimize this additional overhead is to increase the block size which requires more elaborate matrix computations.

V. Numerical Results and Discussion

In this section, we illustrate and validate the proposed pre-equalization techniques by way of numerical results, including error-rate simulations for FTN transmission.

A. Simulation Setup

For the simulations, we consider a coherent optical single-carrier (COSC) transmission system. Optical communication

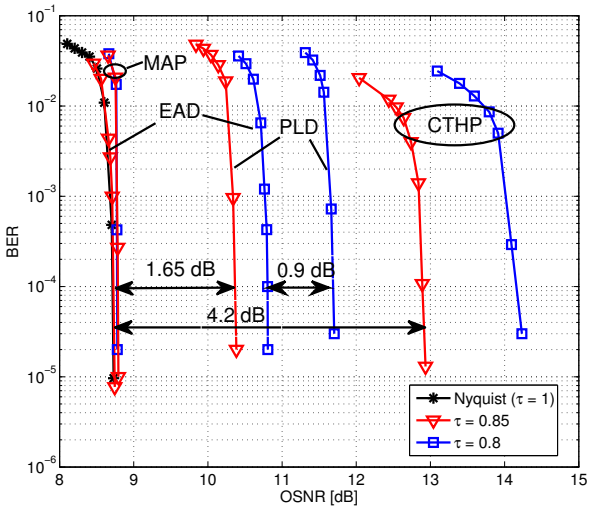


Fig. 5. BER vs. OSNR for FTN-THP with different demappers, illustrating the performance of the proposed EAD. QPSK, $\beta = 0.3$, $\tau = 0.85$ and 0.8.

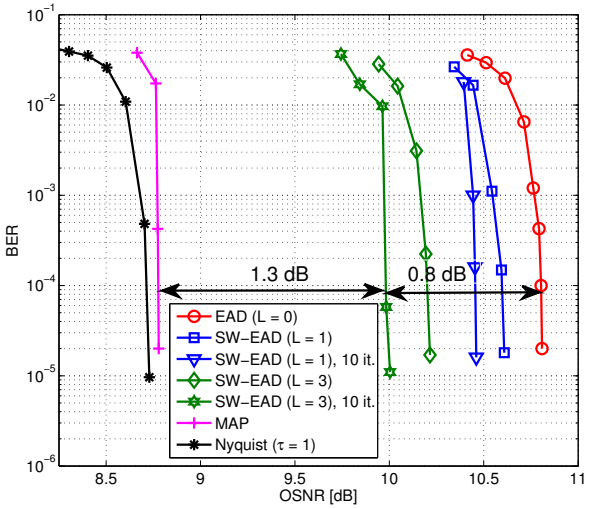


Fig. 7. BER vs. OSNR for FTN-THP with different demappers, illustrating the performance gains with the proposed SW-EAD over EAD. QPSK, $\beta = 0.3$ and $\tau = 0.8$.

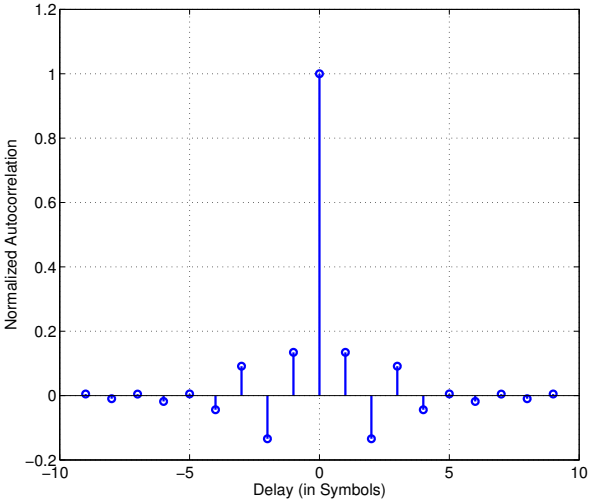


Fig. 6. Auto-correlation of the expanded constellation symbols v for $\beta = 0.3$ and $\tau = 0.8$.

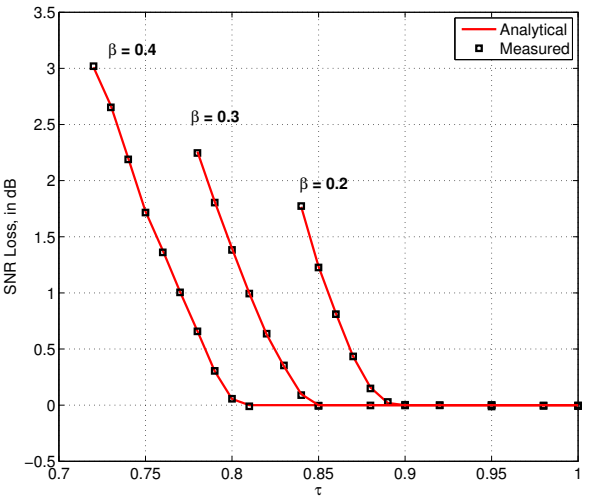


Fig. 8. SNR vs. τ in a QPSK FTN-THP system for varying β.

systems are a prime candidate for the introduction of FTN as the use of higher-order modulation is challenging in such systems [3], [4]. The block diagram of a COSC system with polarization division multiplexing is shown in Fig. 4. The precoding algorithms presented hitherto, considering the AWGN system model in Fig. 1, are directly applicable to linear optical channels as impairments such as chromatic dispersion (CD) and polarization-mode dispersion (PMD) can be compensated through a proper two-dimensional equalizer [38].

In Fig. 4, the transmitter and receiver blocks for the discrete-time baseband modules are same as those in Fig. 1 except that the data processing for each of the two polarizations is performed separately. For our simulations, we use a low-density parity-check (LDPC) code of rate 0.8, a random bit-interleaver, quarternary phase-shift keying (QPSK) and 16-ary quadrature amplitude modulation (16QAM) formats, and a fixed baud rate of 32 Gbaud for all values of τ. The RRC pulse-shaping filter is implemented with 2-times oversampling having 73 time-domain taps with $\beta = 0.3$, and the THP/LPE precoders are designed using 10-taps for the feedback filter. The baseband analog data after the digital-to-analog converter (DAC) is processed by the opto-electronic front-end and transmitted as an optical signal through a 1000 km standard single-mode fiber (SSMF) with CD and mean PMD parameter values of -22.63 ps^2/km and 0.8 ps/$\sqrt{\text{km}}$, respectively, and then is received by the optical coherent receiver. The whitened matched filter (WMF) is combined with the time-invariant frequency domain CD compensator using overlap-and-add method. For PMD compensation, we used a 13-

TABLE I
COMPUTATIONAL COMPLEXITIES OF THE THP-DEMAPPERS FOR EACH BIT AND ITERATION.

Operation	EAD/SW-EAD	PLD	MAP-BCJR
Addition/Subtraction	$\rho_v^L + \Delta_v + 4\rho_v^L\Delta_v - 2$	$2M+2$	$4N_{\text{MAP}} - 2$
Multiplication	$\rho_v^L + \Delta_v$	$M+2$	$6N_{\text{MAP}}$
Division	$\Delta_v + 1$	$M+3$	$2N_{\text{MAP}} + 1$
Non-linear (exp. and log.)	$\rho_v^L + 1$	$M+3$	$4N_{\text{MAP}} + 1$

TABLE II
COMPLEXITY COMPARISON OF THE DEMAPPERS PER BIT PER ITERATION: QPSK, $\beta = 0.3$,$\tau = 0.84$ AND $\tau = 0.8$.

τ	Operations	EAD	PLD	SW-EAD			MAP (6-ISI taps)
				$L=1$	$L=2$	$L=3$	
0.84	ADD.	6	6	52	290	1666	254
	MUL.	8	4	14	36	132	384
	DIV.	5	5	5	5	5	129
	Non-Lin.	5	5	11	33	129	257
	Total	24	20	82	364	1932	1024
0.8	ADD.	6	6	122	1082	6476	254
	MUL.	8	4	28	124	502	384
	DIV.	5	5	5	5	5	129
	Non-Lin.	5	5	25	121	499	257
	Total	24	20	180	1332	7482	1024

tap 2×2 butterfly-type fractionally-spaced adaptive LMS equalizer [38], [39]. Following carrier recovery, the QAM-demapper computes and passes on LLR values to the LDPC decoder.

B. Performance of FTN-THP with Proposed Demappers

We first compare the performance of FTN-THP using the proposed EAD scheme with respect to the conventional THP (CTHP) demapper which employs a modulo operation at the receiver and the modulo-based PLD proposed in [28], [29]. Fig. 5 shows the coded BER performance as a function of the optical SNR (OSNR) for different FTN parameters τ with QPSK modulation. We also include the BER performance with MAP equalization, which considers 6-taps of the ISI channel and performs 10 iterations between the MAP equalizer and LDPC decoder, as a reference. As can be seen from the figure, when ISI is relatively low with $\tau = 0.85$, EAD achieves a performance close to that for the computationally demanding MAP equalization and also to the orthogonal Nyquist-signaling ($\tau = 1$). For this case, EAD outperforms CTHP and PLD by 4.2 dB and 1.65 dB, respectively. When FTN-ISI becomes higher with $\tau = 0.8$, EAD shows a performance gain of 0.9 dB over PLD which is 0.75 dB less compared to the gain with $\tau = 0.85$. The reduction in gap between EAD and PLD with stronger ISI was predicted in Proposition 2.

The loss of performance gain by using EAD with $\tau = 0.8$ can partially be attributed to the correlation between successive symbols of the extended-constellation sequence v of Fig. 2. The auto-correlation sequence of v is plotted in Fig. 6. This correlation is not taken into account while computing the EAD-LLR metric in (12).

Fig. 7 shows the additional performance gains obtained by SW-EAD over EAD. The different curves represent distinct values of L corresponding to the SW-EAD window-length $(2L+1)$ with and without iterations between the demapper and the LDPC decoder. We observe that SW-EAD get improvements of the order of 0.8 dB over EAD which makes it competitive to MAP equalization. With higher values of L, further improvements for SW-EAD are not expected as only up to $3-4$ significant taps are observed in Fig. 6.

The primary reason for the gap in the BER plots between the SW-EAD and MAP equalization can be ascribed to the SNR loss associated with the THP precoding, which was investigated in Section III-D. In Fig. 8, we plot the overall SNR loss (22) of a THP-FTN system compared to an ISI-free transmission as a function of the FTN parameter τ and for different values of β. For each β, we have considered only those values of τ such that $\tau \geq \frac{1}{1+\beta}$, as explained in Section II-B. We observe that for each β, there exists an optimal τ up to which no SNR loss is experienced. For example, with $\beta = 0.3$, this optimal τ is 0.85, which corroborates the BER results in Fig. 5 where FTN-THP transmission yields BER performance close to that of Nyquist-signaling.

C. Computational Complexity Analysis

In this section we present an analysis of the computational cost for the proposed THP-demappers and compare them with the MAP equalization complexity [40]. To reduce the implementation cost of the LLR metric computations for an M^2-ary QAM constellation, we have taken advantage of the fact that the FTN-ISI is real-valued and hence, the in-phase (I) and quadrature (Q) components of the received baseband signals can be individually processed by the demappers and the MAP equalizer. Let P be the number of FTN-ISI taps considered for the MAP-equalization, then the complexity of the BCJR algorithm [40] per bit per iteration is $\mathcal{O}(N_{\text{MAP}})$, where $N_{\text{MAP}} = M^P$. With the quantities L, Δ_v and ρ_v^L as defined in Section III-C, the details of the mathematical operations, required for implementing the EAD and SW-EAD LLR metrics for each bit in each iteration according to (10) and (17) respectively, are furnished in Table I. To illustrate this analysis with further clarity and ease of comparison, two specific examples are provided in Table II with $\tau = 0.84$ and $\tau = 0.8$ for $\beta = 0.3$ and QPSK modulation.

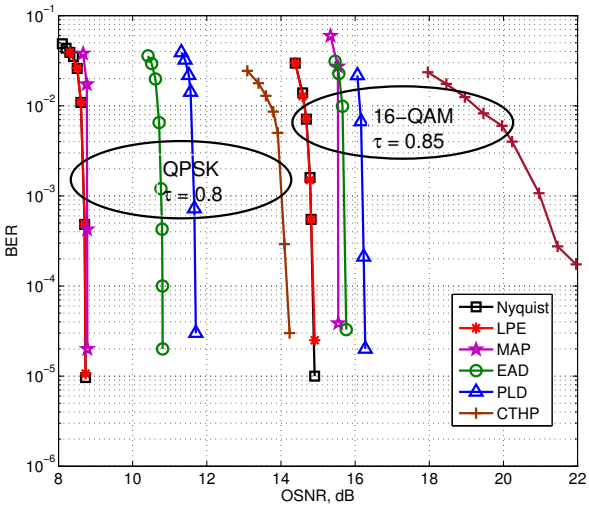


Fig. 9. BER vs. OSNR for FTN with LPE precoding. QPSK with $\tau = 0.8$ and 16QAM with $\tau = 0.85$, $\beta = 0.3$.

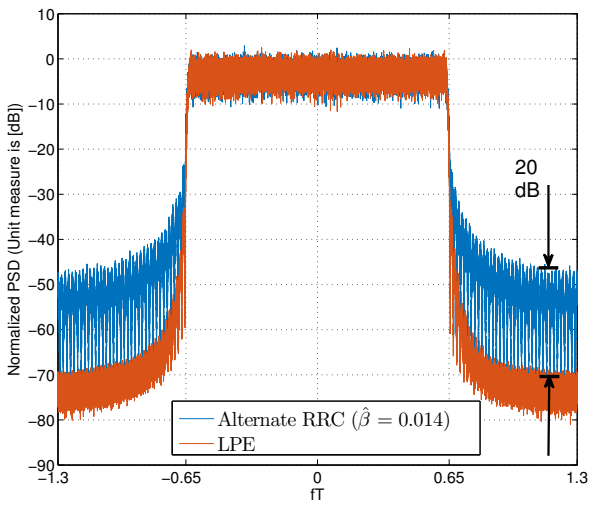


Fig. 11. Normalized PSD of LPE-FTN with $\beta = 0.3$, $\tau = 0.78$ and Nyquist signaling with a τT-orthogonal RRC having $\hat{\beta} = 0.014$ vs. normalized frequency fT using truncated RRC pulses to illustrate spectral leakage.

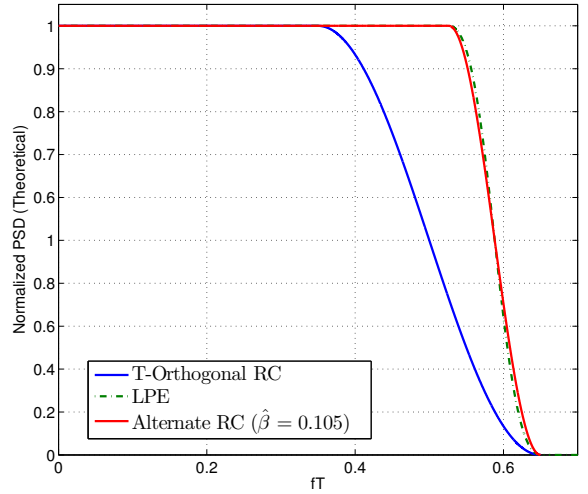


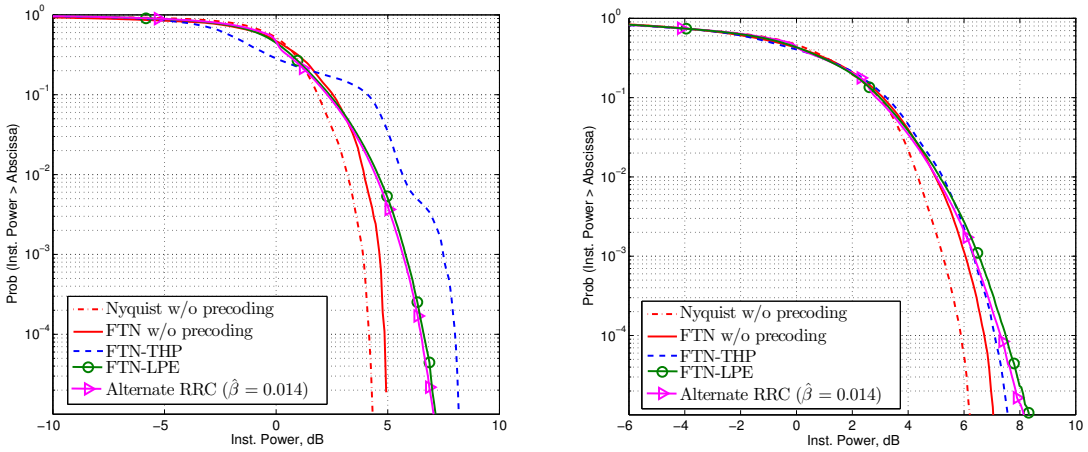
Fig. 10. Normalized PSD of LPE-FTN vs. normalized frequency fT for $\beta = 0.3$, $\tau = 0.85$. Also included are the PSDs for Nyquist signaling with the T-orthogonal RRC with $\beta = 0.3$ and the τT-orthogonal RRC with $\hat{\beta} = 0.105$.

The numbers in Table II reveal that the implementation complexities of the LLR metrics computed by EAD in (10) and PLD are similar for the FTN transmissions studied in this paper, even though EAD demonstrated substantial performance gains over PLD as shown in Section V-B. Moreover, Table II shows that while EAD is significantly more computationally efficient than the MAP equalization, the complexity of the SW-EAD rises with increasing window size, especially for low values of τ. We recall from Section V-B that the SW-EAD performance is always limited even when the window parameter L increases infinitely, as illustrated in Fig. 7. This is because SW-EAD can successfully remove the modulo-loss but fails to improve the power-penalty (SNR-loss) associated with an FTN-THP transmission. Therefore, for a given complexity requirement on the receiver side processing, RRC roll-off β and FTN parameter τ, the window parameter L and the number of iterations between the SW-EAD and the LDPC decoder should be wisely chosen as a desired trade-off between performance and complexity.

D. Performance of Proposed FTN-LPE

As described in Section III-D and shown in Fig. 8, an SNR degradation is inherent to THP precoding for some values of τ and β. Our proposed LPE scheme can overcome this problem. Fig. 9 shows the FTN-LPE BER results for QPSK and 16QAM. We observe that LPE precoding produces an optimal performance, i.e., the BER is identical to that of orthogonal signaling. The figure also includes the BER curves from Fig. 5 with $\tau = 0.8$ to show the gains offered by LPE over THP. Similar observations hold true with higher order modulation, such as 16QAM.

The optimal BER performance of LPE precoded FTN systems comes at the expense of transmitted spectral shape modification, as investigated in Section IV. Fig. 10 plots the normalized analytical transmit PSDs of the LPE precoded FTN system, which was derived in (23). We also include the normalized PSDs for (non-precoded) Nyquist signaling using the underlying T-orthogonal RRC with $\beta = 0.3$ and a τT-orthogonal RRC with $\hat{\beta} = \tau(1+\beta) - 1 = 0.105$ for pulse shaping, respectively. For this comparison, all three systems use the same bandwidth, which implies that the LPE-FTN system and the Nyquist-system with the RRC with $\hat{\beta} = 0.105$ operate at a higher baud rate. We observe that with LPE precoding, the overall PSD behaves as a τT-orthogonal pulse-shape. That is, the PSD has a odd-symmetry about the normalized frequency $fT = \frac{1}{2\tau} = 0.59$ similar to the alternate τT-orthogonal RC.



(a) QPSK (b) 16QAM

Fig. 12. Empirical CCDF of the instantaneous power with average transmit power = 0 dB, $\beta = 0.3$, $\tau = 0.78$.

The advantage of the proposed LPE scheme over a direct τT-orthogonal RRC pulse-shaping is illustrated in Fig. 11 in terms of the out-of-band emission performance. Here, the transmit pulse-shaping filters for both the LPE precoded FTN system with $\beta = 0.3$, $\tau = 0.78$ and the direct τT-orthogonal Nyquist transmission (effective $\hat{\beta} = 0.014$) were implemented by using 73 time domain taps. For both systems, transmit PSDs are computed using the twice-oversampled discrete-time samples before the DAC in Fig. 4. The normalized PSDs are plotted in Fig. 11, as a function of the normalized frequency fT. We observe that LPE transmission results in a significantly lower ($\sim$ 20 dB) spectral leakage in the side-bands. This improved out-of-band emission performance is advantageous for transmission schemes with strict spectral-emission mask requirements to achieve low interference between adjacent channels.

Precoding may cause a possible increase in the PAPR. We demonstrate the PAPR behaviour for the precoded FTN techniques by plotting the empirical complementary cumulative distribution function (CCDF) of the instantaneous power in Fig. 12 for QPSK and 16QAM constellations. The modulation parameters are $\beta = 0.3$ and $\tau = 0.78$, and the different curves correspond to Nyquist signaling with T-orthogonal and τT-orthogonal pulse-shapes ($\hat{\beta} = 0.014$), unprecoded FTN transmission and FTN employing THP and LPE precoding. All transmission schemes are normalized to the same average transmitted power of 0 dB. As can be seen from Fig. 12a, the PAPR of the FTN-THP system with QPSK modulation is relatively higher than that for the LPE precoded FTN system, whereas for 16QAM they perform similarly as presented in Fig. 12b. Furthermore, FTN-LPE transmission yields almost a similar PAPR performance as that of the alternate τT-orthogonal signaling scheme for both QPSK and 16QAM modulation formats.

Finally, we remark that the suitability of the two pre-equalization methods, proposed in this paper, depends on the specific application of FTN. In the current paper, where FTN-LPE has been shown to outperform the FTN-THP scheme, we have restricted the application of FTN to two different channel models, (a) an AWGN channel for the simplicity of the theoretical analysis, and (b) an optical channel as a practical application example to present simulation results. However, the efficiency of the proposed FTN-THP can be more pronounced if we consider FTN transmissions under different channel models. As the functionality of the proposed THP-demappers depends only on the a-priori probabilities of the symbols v and not on the actual channel parameters, they can be directly applied under these circumstances, e.g. (1) in FTN transmissions over multi-path ISI channels, THP would be a suitable choice to pre-equalize the combined ISI due to FTN and the multi-path channel, because LPE may exhibit significant power-loss due to channel inversion, as the multi-path channel lies outside the transmitter; (2) in the previous example of FTN signaling over multi-path channels, a combination of LPE and THP can also be used at the transmitter, where LPE can be employed to pre-mitigate the FTN-ISI, whereas, THP can be applied along with the proposed demappers to pre-equalize the ISI component, arising only from the multi-path channel; (3) in a multi-user, multi-carrier FTN transmission scheme, where frequency-packed sub-channels are allocated to different users, LPE is not good choice for FTN pre-equalization as it requires joint receiver processing in the form of feed-forward filtering, which is not generally a viable option due to the geographical separation of the users. However, THP can be employed in such a scenario with both the feedback and feed-forward filters implemented at the transmitter. While we have not explored the above mentioned FTN applications in detail in the current paper, they can be considered for possible future works as suitable application examples for the proposed FTN pre-equalization methods.

VI. Conclusions

FTN transmission is a non-orthogonal signaling scheme to improve spectral efficiency at the expense of introducing ISI.

As an alternative to computationally demanding equalization at the receiver, in this paper, we have analyzed pre-equalization techniques at the transmitter to mitigate the FTN induced ISI. First, we have considered THP and proposed two new symbol demappers to improve the reliability of the computed LLRs by reducing the modulo-loss. Numerical results for a coherent optical transmission system show that the proposed demappers outperform existing THP demappers by significant margins. Secondly, we have proposed a linear pre-equalization technique which converts the FTN transmission into an orthogonal signaling at a higher baud rate. LPE precoded FTN systems can thus yield optimal ISI-free BER performance. Moreover, the numerical results also suggest that LPE can cause substantially lower out-of-band emission compared to a direct τT-orthogonal RRC pulse-shaping without significant PAPR penalty. In conclusion, we have demonstrated that the proposed FTN pre-equalization techniques are effective means to achieve higher spectral efficiency promised by FTN.

Appendix A
Proof of Proposition 1

To derive the a-priori probabilities, we consider the linear equivalent THP transmitter in Fig. 2 and make use of the assumption that the elements $f[k] = \sum_m b[m]r[k-m]$ of the filter output f are approximately zero-mean Gaussian distributed with variance σ_f^2 [20]. For an M-ary PAM constellation, the expanded signal set $\mathcal{V}$ can be written as $\mathcal{V} = \{\mathcal{A} + 2iM : i \in \mathbb{Z}\}$. Furthermore, from the construction of the equivalent block diagram shown in Fig. 2, one can see that the elements $d[k]$ of the sequence d are from the set $\{2iM : i \in \mathbb{Z}\}$ and $d[k] = 2iM$ if

$$-2iM - M \leq a[k] - f[k] \leq -2iM + M \ . \tag{30}$$

Hence,

$$\begin{aligned} &\Pr\ (v[k] = a_{\text{PAM}}^{\kappa} + 2iM) && (31)\\ &= \Pr\big[\big(-M_i^+ \leq a[k]-f[k] \leq -M_i^-\big) \cap \big(a[k]=a_{\text{PAM}}^{\kappa}\big)\big] && (32)\\ &= \frac{1}{M}\Pr\left[\left(M_i^- + a_{\text{PAM}}^{\kappa} \leq f[k] \leq M_i^+ + a_{\text{PAM}}^{\kappa}\right)\right] && (33)\\ &= \frac{1}{M}\left[\Phi\left(\frac{M_i^+ + a_{\text{PAM}}^{\kappa}}{\sigma_f}\right) - \Phi\left(\frac{M_i^- + a_{\text{PAM}}^{\kappa}}{\sigma_f}\right)\right]. && (34)\end{aligned}$$

Appendix B
Proof of Proposition 2

In an AWGN channel, the nearest-neighbor approximated LLR, computed by EAD for the n^{th} bit of the k^{th} transmitted symbol is given by (12). We can write the nearest neighbors of the received symbol $v'[k]$ for an M-ary PAM as

$$\bar{c}_{0,n} = a_{\text{PAM}}^{\kappa_{0,n}^*} + 2uM \ , \tag{35}$$

$$\bar{c}_{1,n} = a_{\text{PAM}}^{\kappa_{1,n}^*} + 2vM \ , \tag{36}$$

where $a_{\text{PAM}}^{\kappa_{0,n}^*}, a_{\text{PAM}}^{\kappa_{1,n}^*} \in \mathcal{A}$ with $\mathcal{A}$ being the original M-ary PAM signal set and $u, v \in \mathbb{Z}$ such that $|u - v| \leq 1$ for arbitrary bit-mapping.

The nearest neighbors of $v'[k]$ remain invariant after the modulo operation and one additional layer of constellation extension applied in PLD [28], if and only if there exists a $w \in \mathbb{Z}$ such that

$$-M \leq v'[k] - 2wM < M \ , \tag{37}$$

$$-(M+1) \leq a_{\text{PAM}}^{\kappa_{0,n}^*} + 2(u-w)M \leq M+1 \ , \tag{38}$$

$$-(M+1) \leq a_{\text{PAM}}^{\kappa_{1,n}^*} + 2(v-w)M \leq M+1 \ . \tag{39}$$

If $u = v$, then $w = u = v$ satisfies (37)-(39). If $|u - v| = 1$, which means that $\bar{c}_{0,n}$ and $\bar{c}_{1,n}$ lie on different sides of the modulo boundary before the modulo operation, (37)-(39) are satisfied if and only if

$$|v'[k] - \bar{c}_{i^*,n}| \leq 2 \ , \tag{40}$$

where $i^* \in \{0, 1\}$ denotes the index for which $\bar{c}_{i^*,n}$ is located across the modulo boundary from $v'[k]$. It can be easily verified that (40) and hence (37)-(39) are satisfied for any bit-labeling of the PAM constellation when $M = 2$ or 4. However, for $M > 4$, the conditions are not satisfied for some labelings.

Thus, for 2PAM and 4PAM modulations, PLD computes the LLRs as

$$\text{LLR}_{k,n}^{\text{PLD}} = \frac{|v'[k]-\bar{c}_{0,n}|^2 - |v'[k]-\bar{c}_{1,n}|^2}{2\sigma^2} \ . \tag{41}$$

Therefore, under equal probability assumption when $\alpha_{1,n} = \alpha_{0,n}$, comparing (12) and (41) yields

$$\text{LLR}_{k,n}^{\text{EAD}} = \text{LLR}_{k,n}^{\text{PLD}} \ . \tag{42}$$

Appendix C
PSD And Average Transmit Power with Precoding

From (1), the PSD of the FTN transmitted signal s is given by [2], [41]

$$\Phi_{ss}(f) = \frac{1}{\tau T}|\hat{H}(f)|^2 \Phi_{rr}\left(\mathrm{e}^{\mathrm{j}2\pi f \tau T}\right) \ , \tag{43}$$

where Φ_{rr} is the discrete-time Fourier transform of the auto-correlation of the sequence r.

We assume that the constellation symbol sequence a in Fig. 2 and Fig. 3 and the intermediate process v in Fig. 2 can be approximated as a zero-mean wide-sense stationary (WSS) process with auto-correlation sequences $\sigma_a^2\delta$ and ϕ_{vv} respectively, where δ is the Kronecker-delta function. Then, from (3), the z-transforms of the auto-correlation of r for THP and LPE can be written as

$$\Phi_{rr}^{\text{THP}}(z) = \Phi_{vv}(z)\frac{1}{Q(z)Q^*(z^{-*})} = \frac{\alpha\Phi_{vv}(z)}{G(z)} \ , \tag{44}$$

$$\Phi_{rr}^{\text{LPE}}(z) = \sigma_a^2\frac{1}{Q(z)Q^*(z^{-*})} = \frac{\alpha\sigma_a^2}{G(z)} \ , \tag{45}$$

where Φ_{vv} is z-transform of ϕ_{vv}.

Evaluating (44) and (45) on the unit circle and using the relation $\hat{G}(f) = |\hat{H}(f)|^2$, the PSD in (43) becomes

$$\Phi_{ss}^{\text{THP}}(f) = \alpha\Phi_{vv}(e^{\mathrm{j}2\pi f\tau T})\frac{\hat{G}(f)}{\sum\limits_k \hat{G}(f + \frac{k}{\tau T})} \tag{46}$$

for FTN-THP and

$$\Phi_{ss}^{\mathrm{LPE}}(f) = \alpha\sigma_a^2 \frac{\hat{G}(f)}{\sum\limits_k \hat{G}(f + \frac{k}{\tau T})} \tag{47}$$

for FTN-LPE.

In order to compute the average power for the FTN-THP and FTN-LPE systems, we consider an equivalent pulse-shape ψ, such that

$$\Psi(f) = T' \frac{\hat{G}(f)}{\sum\limits_k \hat{G}(f + \frac{k}{T'})} \,, \tag{48}$$

where Ψ is the Fourier transform of ψ and $T' = \tau T$. It can be easily shown that ψ satisfies Nyquist's zero-ISI criterion with respect to the sampling rate T' because

$$\frac{1}{T'} \sum_l \Psi(f + \frac{l}{T'}) = 1 \,. \tag{49}$$

Now, to compute the average power of an FTN-THP system from (46), we can write the autocorrelation of the transmitted signal s, corresponding to a delay $\tilde{\tau}$, as

$$\phi_{ss}^{\mathrm{THP}}(\tilde{\tau}) = \frac{\alpha}{T'} \sum_m \phi_{vv}^m \psi(\tilde{\tau} - mT') \,, \tag{50}$$

where ϕ_{ss}^{THP} is the inverse Fourier-transform of the PSD Φ_{ss}^{THP} and ϕ_{vv}^m is the autocorrelation of v corresponding to a delay m. Therefore, the average power of an FTN-THP system can be written as

$$P_{\mathrm{Avg}}^{\mathrm{THP}} = \int_{\infty}^{\infty} \Phi_{ss}^{\mathrm{THP}}(f) \mathrm{d}f \tag{51}$$

$$= \phi_{ss}^{\mathrm{THP}}(0) \tag{52}$$

$$= \frac{\alpha\sigma_v^2}{\tau T} \,, \tag{53}$$

where the step (52) to (53) follows from (50), using the fact that ψ is a $T'(=\tau T)$-orthogonal Nyquist-pulse as shown in (49) and $\phi_{vv}^0 = \mathbb{E}(|v|^2) = \sigma_v^2$, with $\mathbb{E}(\cdot)$ denoting the expectation operator.

To compute the average power of the FTN-LPE system, we note from the T'-orthogonality of ψ in (49) that

$$\psi(0) = \int_{-\infty}^{\infty} \Psi(f) \mathrm{d}f = 1 \,. \tag{54}$$

Therefore, the average power of the FTN-LPE system follows from (47) and using (54) as

$$P_{\mathrm{Avg}}^{\mathrm{LPE}} = \int_{\infty}^{\infty} \Phi_{ss}^{\mathrm{LPE}}(f) \mathrm{d}f \tag{55}$$

$$= \frac{\alpha\sigma_a^2}{T'} \int_{-\infty}^{\infty} \Psi(f) \mathrm{d}f \tag{56}$$

$$= \frac{\alpha\sigma_a^2}{\tau T} \,. \tag{57}$$

REFERENCES

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

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

[3] G. Colavolpe, T. Foggi, A. Modenini, and A. Piemontese, "Faster-than-Nyquist and beyond: how to improve spectral efficiency by accepting interference," *Optics Express*, vol. 19, no. 27, pp. 26 600–26 609, December 2000.

[4] M. Seimetz, *High-Order Modulation for Optical Fiber Transmission*. Springer Berlin Heidelberg, 2010.

[5] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini, F. Rusek, and A. Ugolini, "Modulation formats and waveforms for 5G networks: Who will be the heir of OFDM?" *IEEE Commun. Mag.*, vol. 31, no. 6, pp. 80–93, Nov. 2014.

[6] J. B. Anderson, F. Rusek, and V. Owall, "Faster-Than-Nyquist Signaling," in *Proc. IEEE*, vol. 101, no. 8, 2013, pp. 1817–1830.

[7] J. E. Mazo, "Faster-than-Nyquist Signaling," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1451–1462, October 1975.

[8] J. E. Mazo and H. J. Landau, "On the Minimum Distance Problem for Faster-than-Nyquist Signaling," *IEEE Trans. Inf. Theory*, vol. 34, no. 6, pp. 1420–1427, November 1988.

[9] D. Hajela, "On Computing the Minimum Distance for Faster than Nyquist Signaling," *IEEE Trans. Inf. Theory*, vol. 36, no. 2, pp. 289–295, March 1990.

[10] A. D. Liveris and C. N. Georghiades, "Exploiting Faster-than-Nyquist Signaling," *IEEE Trans. Commun.*, vol. 51, no. 9, pp. 1502–1511, September 2003.

[11] A. Prlja, J. B. Anderson, and F. Rusek, "Receivers for Faster-than-Nyquist Signaling with and without Turbo Equalization," in *IEEE Int. Symp. on Inf. Theory*, 2008, pp. 464–468.

[12] A. Prlja and J. B. Anderson, "Reduced-Complexity Receivers for Strongly Narrowband Intersymbol Interference Introduced by Faster-than-Nyquist Signaling," *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2591–2601, September 2012.

[13] F. Rusek and J. B. Anderson, "Successive Interference Cancellation in Multistream Faster-than-Nyquist Signaling," in *Int. Conf. on Wireless Commun. and Mobile Computing*, 2006, pp. 1021–1026.

[14] J. Yu, J. Park, F. Rusek, B. Kudryashov, and I. Bocharova, "High Order Modulation in Faster-Than-Nyquist Signaling Communication Systems," in *IEEE Vehicular Tech. Conf. Fall (VTC)*, 2014.

[15] S. Sugiura, "Frequency-Domain Equalization of Faster-than-Nyquist Signaling," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 555–558, October 2013.

[16] S. Sugiura and L. Hanzo, "Frequency-Domain-Equalization-Aided Iterative Detection of Faster-than-Nyquist Signaling," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 2122–2128, May 2015.

[17] T. Ishihara and S. Sugiura, "Frequency-Domain Equalization Aided Iterative Detection of Faster-than-Nyquist Signaling with Noise Whitening," in *IEEE Int. Conf. on Commun. (ICC)*, 2016, pp. 1–6.

[18] M. Tomlinson, "New Automatic Equaliser Employing Modulo Arithmetic," *Electronics Letters*, vol. 7, no. 5, pp. 138–139, March 1971.

[19] H. Harashima and H. Miyakawa, "Matched-Transmission Technique for Channels with Intersymbol Interference," *IEEE Trans. Commun.*, vol. 20, no. 4, pp. 774–780, August 1972.

[20] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*. Wiley-IEEE Press, 2002.

[21] M. E. Hefnawy and H. Taoka, "Overview of Faster-Than-Nyquist for Future Mobile Communication Systems," in *IEEE Vehicular Tech. Conf. Spring (VTC)*, 2013.

[22] F. L. Luo and C. Zhang, *Signal Processing for 5G: Algorithms and Implementations*. Wiley-IEEE Press, 2016.

[23] M. Maso and S. Tomasin, "Pre-equalized Faster Than Nyquist Transmission for 5G Cellular Cicrowave Back-haul," in *IEEE Int. workshop on Signal Proc. advances in Wireless Commun. (SPAWC)*, 2016.

[24] D. Chang, O. Omomukuyo, O. Dobre, R. Venkatesan, P. Gillard, and C. Rumbolt, "Tomlinson-Harashima Precoding with Soft Detection for Faster than Nyquist DP-16QAM Coherent Optical Systems," in *Opt. Fiber Commun. Conf.(OFC)*, 2015, pp. 1–3.

[25] D. Chang, O. Omomukuyo, X. Lin, S. Zhang, O. Dobre, and V. Ramachandran, "Robust Faster-than-Nyquist PDM-mQAM Systems with Tomlinson-Harashima Precoding," *IEEE Photon. Technol. Lett.*, vol. PP, no. 99, May 2016.

[26] R. Rath and W. Rosenkranz, "Tomlinson-Harashima Precoding for Fiber-Optic Communication Systems," in *European Conf. on Opt. Commun. (ECOC)*, 2013, pp. 1–3.
[27] R. Rath, C. Schmidt, and W. Rosenkranz, "Is Tomlinson-Harashima Precoding Suitable for Fiber-Optic Communication Systems?" in *ITG Symp. on Phot. Networks*, 2013, pp. 1–7.
[28] E. C. Y. Peh and Y. C. Liang, "Power and modulo loss tradeoff with expanded soft demapper for LDPC coded GMD-THP MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 714–724, February 2009.
[29] S. Kinjo, "An Efficient Soft Demapper for Tomlinson-Harashima Precoded Systems," *IEICE Comm. Exp.*, vol. 4, no. 3, pp. 89–94, March 2015.
[30] E. Kaminer, D. Raphaeli, and Y. Hayoun, "Iterative Decoding of Coded THP with Quantized Output," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3721–3729, September 2013.
[31] A. Said and J. Anderson, "Bandwidth-Efficient Coded Modulation with Optimized Linear Partial-Response Signals," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 701–713, March 1998.
[32] F. Rusek and J. B. Anderson, "Non Binary and Precoded Faster Than Nyquist Signaling," *IEEE Trans. Commun.*, vol. 56, no. 5, pp. 808–817, May 2008.
[33] Y. J. D. Kim and J. Bajcsy, "On Spectrum Broadening of Pre-Coded Faster-Than-Nyquist Signaling," in *IEEE Vehicular Tech. Conf. Fall (VTC)*, 2010, pp. 1–5.
[34] E. Ringh, "Low complexity algorithms for Faster-than-Nyquist signalling : Using coding to avoid an np-hard problem," Master's thesis, KTH Royal Institute of Technology: School of Engineering Sciences, 2013.
[35] A. Gattami, E. Ringh, and J. Karlsson, "Time Localization and Capacity of Faster-Than-Nyquist Signaling," in *IEEE Global Commun. Conf. (GLOBECOM)*, 2015, pp. 1–7.
[36] R. F. H. Fischer, "Sorted Spectral Factorization of Matrix Polynomials in MIMO Communications," *IEEE Trans. Commun.*, vol. 53, no. 6, pp. 945–951, June 2005.
[37] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. Tata McGraw-Hill, 2002.
[38] 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.
[39] E. Ip and J. M. Kahn, "Digital Equalization of Chromatic Dispersion and Polarization ModeDispersion," *J. Lightw. Technol*, vol. 25, no. 8, pp. 2033–2043, July 2007.
[40] 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.
[41] J. Proakis and M. Salehi, *Digital Communications*. McGraw-Hill Education, 2007.